HAND GESTURE RECOGNITION AND HAND TRACKING FOR MEDICAL
APPLICATIONS

BY

HAJAR SHARIF

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mechanical Engineering
in the Graduate College of the
University of Illinois Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Professor Srinivasa M. Salapaka, Chair
Professor Thenkurussi K. Kesavadas, Director of Research
Professor Ramavarapu S. Sreenivas
Associate Professor Heidi Phillips

Abstract

Hand gestures are a mean of communication and a prevalent type of body language that conveys messages through different shapes constructed by palm and fingers. Hand gesture recognition (HGR) has been of interest in many research fields such as sign language translation, musical creation, and virtual environment control. There are also several studies on HGR for robotics, prosthetic, and rehabilitation applications. In this dissertation, the application of HGR for addressing two challenges in the medical field is presented. The first challenge is to develop a quantitative metric to improve rehabilitation of neurological conditions, with a focus on improvement in performing activities of daily living (ADL), while the second challenge is to develop ATTENTIVE, an automated and quantitative assessment system, to enhance a better evaluation of surgical skills proficiency.

Many neurological conditions lead to motor impairment of upper extremity that includes muscle weakness, altered muscle tone, joint laxity, and impaired motor control. As a result, common activities such as reaching, picking up objects, and holding onto them are compromised. Therefore, such patients will experience disability in performing ADL such as eating, writing, performing housework, and so on. Several evaluation methods are commonly used to assess problems in performing ADL. Despite the wide application of these methods, all of them are subjective techniques, i.e. they are either questionnaires or qualitative scores assigned by a medical professional. We hypothesize that providing a more quantitative metric can enhance evaluation of the rehabilitation progress, and lead to a more efficient rehabilitation regimen tailored to the specific needs of each individual patient.

Since the first step of developing a metric is to distinguish different ADL activities using hand gesture data, in this dissertation the focus is on classification of ADL tasks using hand gestures. Data analysis pipelines were developed to take in data, collected by the leap motion controller as well as the electromyography and inertial measurement unit sensors, from the lower arm during completion of certain ADL tasks. These pipelines output classification accuracies to distinguish the ADL tasks. Different preprocessing, feature extraction, and classification methods were tested on the data from healthy adults to detect the best structure and parameters for the proposed pipelines. The developed pipelines can be trained and their parameters can be tuned based on data from an intact-adult population. Then, The tuned pipelines can be set as the references. Subsequently, hand motion data from a neurological patient completing the same tasks in the same data collection setup can be fed into the reference pipelines to obtain the classification accuracies. The achieved accuracies indicate how close a patient's hand motions and muscle activation are to the hand motions and muscle activation of the healthy population. This method enhances assessment of the overall performance of a patient in a quantitative fashion. In addition, the acquired confusion matrices provide insight into the patient's performance in completing each individual task.

The second section of this dissertation includes the design of ATTENTIVE; an evidence-supported, automated, robust, real-time, comprehensive, quantitative assessment system for evaluating proficiency in

basic surgical skills. Since ATTENTIVE provides quantitative feedback, it can have a variety of applications in teaching surgical skills either in traditional settings or within incorporation of the augmented reality systems. As of now, the presence of an automated and quantitative assessment system to provide feedback on surgical tasks performance is lacking, and expert surgeon's involvement is necessary to provide feedback to the surgical trainees. As a result, a trainee's opportunities to receive feedback on one's performance is restricted to the availability of an expert surgeon, which is limited due to pre-existing high workload of the expert surgeons. ATTENTIVE can eliminate such restriction that in turn may result in surgical trainees' performance improvement and superior surgical outcomes over the long run.

In this work, the idea and pipeline for developing ATTENTIVE are presented. Next, the apparatus and experimental setup and protocol to investigate the feasibility of ATTENTIVE were designed and built. Afterwards, data was collected from 65 participants completing four basic surgical tasks. The participants were students, residents, and expert surgeons in the fields of veterinary and human medicine. To benefit from both sensor-based and vision-based HGR methods for solving the problem in hand, Azure Kinect DK, Leap Motion Controller, and Myo armband were used to collect data from the lower arm of the participants. ATTENTIVE's workflow consists of three major steps including separating the main task part from the preparation and cleanup after the task completion, classifying the input surgical task, and assigning a performance score to the input task. In this dissertation, many parts of these three steps are completed, and algorithms to complete the rest are determined and implemented to a vast extend. The details of the data analysis steps are beyond the scope of the abstract and are presented in the second section's chapters.

The last chapter of the current dissertation contains preliminary work on design and fabrication of a wearable device, named iBand, to collect biosignals and kinematic data from the lower arm. Different components of iBand have been selected and calibrated to read synchronized data from the lower arm, transfer them to a computer via Bluetooth, and save them as separate files. An easy-to-work user interface has been developed for iBand to enable user to save the data in the desired folder and with the desired file name. In addition, the user interface enhances a real-time data observation in which the user can choose the sensor from which the collected signal is displayed. Upon completion, the iBand can replace the discontinued Myo armband for research and daily life applications.

To Mother and Father, and parents-in-law.
To my better half, Hossein.
To my lovely children, Fatemeh, Ali, and Mohammad, without whom, this would have been completed two
years earlier!
And to all the hearts that seek truth, peace, and love.

Acknowledgments

First and foremost, I would like to acknowledge and give my warmest thanks to my supervisor, Professor Thenkurussi K. Kesavadas, for being the best mentor I could have asked for. It was truly a great opportunity and pleasure to work under his supervision and benefit from his immense knowledge and plentiful experience. Professor Kesavadas supported me through all my research ventures and gave me complete liberty to shape my dissertation. He patiently gave me opportunities to fall and rise while he was always available to support me through my PhD journey and to nudge me in the right direction to ensure that I met my professional goals. One of the most important skills that I learnt from him is how to think wildly and crazily about a research topic and then, shape it within the boundaries of real life. The lessons I have learnt from him are invaluable and I will cherish them for a lifetime.

I would like to express my sincere gratitude to the respected members of my dissertation committee, Professor Heidi Phillips, Professor Srinivasa M. Salapaka, and Professor Ramavarapu S. Sreenivas for taking the time out of their busy schedules to carefully review my dissertation and provide me with insightful questions and suggestions.

I would like to extend my special thanks to Professor Phillips for being a part of my graduate journey and an invaluable source of knowledge and support from the first day we started our collaboration on writing a proposal about "Skill Assessment in surgery" to obtain a research grant up to today. She and her student, Dr. Jennifer Kuzminsky, shared their knowledge and experience with me generously and patiently. Collaboration with Professor Phillips and her wonderful team at the College of Veterinary Medicine and Small Animal Clinic at the University of Illinois Urbana-Champaign (UIUC) was a valuable opportunity for me to expand my understanding and knowledge about the challenges in the clinical world. I sincerely thank her and her team.

I sincerely appreciate the support of Jump Trading Simulation and Education Center for providing us with their facility for data collection and the residents and surgeons in OSF Saint Francis Medical Center, located in Peoria, Illinois, who devoted their precious time to participate in our study. Particularly, I would like to thank Dr. Paul M. Jeziorczak, Safura Sultana, and Dr. William F. Bond. In addition, I am deeply grateful to Professor Meenakshy Aiyer, Professor Jessica R Hanks, and Professor Richard C Anderson from University of Illinois College of Medicine Peoria, for assisting us in recruiting participants and Professor Minh N. Do for sharing equipment from his lab to complete the "Hand Gesture Recognition in Surgical Skill Assessment" study.

I was fortunate to collaborate with some of the smartest people on the planet, who were incredible sources of knowledge and encouragement. I would like to express my sincere gratitude to my mentors and colleagues Dr. Pramod Chembrammel and Dr. Yao Li whose experience, deep knowledge, and attention to detail have instilled in me the need to deliver high quality work at every iteration of a project. In addition, I offer my special thanks to my wonderful colleagues Ahmadreza Eslaminia, Seung Byum Seo, and Reihaneh Jahedan

for their dedication to our collaborative research. I also acknowledge the efforts of the undergraduate students who helped with developing the iBand, specially Sumedh Vemuganti who was my first mentee in the iBand project and his dedication to work improved my confidence as a successful mentor.

I offer my regards and blessings to all members of the Health Care Engineering Systems Center at UIUC, particularly Dr. Antonios Michalos, Michelle L. Osborne, Anusha Muralidharan, Lydia Lee, Harris Junaid Nisar, George Heintz, and Dr. Inki Kim who supported me in all respects during the completion of my research. Thank you All, and I look forward to a continued association with each and everyone of you.

Further, I am grateful to my friends, who made my graduate journey fun and bearable. Particularly, I thank Minoo Roghani, Zahra Dastgheib, Dr. Sara Bahramian, Maryam Badiee, and Fatemeh Ameri for their support and friendship.

I would also like to express my gratitude to my family members, my parents, my parents-in-law, my siblings, my siblings-in-law, and all my close relatives for their unwavering love and support, without which I would not have succeeded. I specially thank my mother, Dr. Tahereh Zandieh, who has been my role model in being persistent and strong. She has been an endless source of courage, support, and love throughout my life.

My special thanks and sincere appreciation are extended to the love of my life,my husband - Dr. Hossein Bagheri, who was the best unintended outcome of attending the university and who selflessly supported me through each and every step to ensure that I can achieve my professional and personal goals. I also appreciate my lovely children, Fatemeh, Ali, and Mohammad, who were sweeter and stronger than any coffee to keep me alert! They surely did a great job to ensure that I keep a balance between personal and professional life.

Last but not least, I want to thank the participants in my studies who generously devoted their precious time to make this research happen.

Table of contents

Preface

Hand gestures are a mean of communication and a prevalent type of body language that conveys messages through different shapes constructed by palm and fingers. The stable hand shapes are usually referred to as "static gestures" whereas a series of hand movements form "dynamic gestures". Another terminology with an interchangeable meaning is "hand posture" which focuses on the hand shape whereas "hand gesture" focuses on the movement of hand [1], [2].

Hand gesture recognition (HGR) has been of interest in many research fields such as sign language translation [3]–[5], musical creation [6], and virtual environment control [7]–[11]. There are also several studies on HGR for robotics, prosthetics, and rehabilitation applications [12]–[21]. Another field of interest is using HGR to evaluate proficiency in medical skills [22].

HGR methods can be divided into two categories based on the applied hardware to detect the hand gestures [1], [23]. The first category is the sensor-based methods and includes equipment that come in contact with person's body, e.g. sensor-embedded-bracelets, electromyography (EMG) sensors, or sensor-embedded-gloves [24], [25]. These methods are superior in providing information about gesture without background distraction; however, they are less natural, bulkier, and usually more expensive [1]. In contrast, the second category contains contactless, vision-based methods by means of different types of cameras such as Red-Green-Blue (RGB), Time-Of-Flight (TOF), and depth cameras [2], [26]–[28]. These methods do not interfere with the natural way of forming hand gestures; however, several factors such as the number and positioning of cameras, the hand visibility, and algorithms applied on the captured videos can affect the performance of these techniques [1].

In this dissertation, the application of HGR for addressing two challenges in the medical field has been studied using a selection of hardware from both sensor-based and vision-based categories. The first section of this manuscript, i.e. Chapters 1 to 4, contains application of HGR methods in assessment of activities of daily living. The second section, i.e. Chapters 5 to 10, includes the HGR methodologies used in developing a quantitative method to evaluate surgical proficiency. Chapter 12 includes preliminary work on design and fabrication of a wearable device, named iBand, to collect biosignals and motion data from the lower arm. This device may have future potential in both the research topics presented in this dissertation.

Part I
HAND GESTURE RECOGNITION FOR EVALUATING DEXTERITY IN PERFORMING DAILY LIFE ACTIVITIES

Chapter 1

Introduction

1.1 Motivation

Cutoff of blood flow to an area of brain leads to oxygen deprivation and cell death. This phenomenon is known as stroke that is one of the leading mortality and motor and cognitive disability causes worldwide [29], [30]. In United States, every 40 seconds one person has a stroke and in every 3.5 minutes one person loses life due to it [31]. Based on the rehabilitation intervention, as illustrated in Figure 1.1, the phases of stroke are defined as acute, within the first month, subacute, between first and sixth month, and chronic, after the sixth month post stroke [30].

The most common deficit after stroke is hemiparesis of the contralateral upper limb, which is experienced by 80% and 40% of stroke patients in acute and chronic phases of the stroke, respectively [32], [33]. Motor impairment of upper extremity includes muscle weakness, altered muscle tone, joint laxity, and impaired motor control. As a result, common activities such as reaching, picking up objects, and holding onto them are compromised [33]. Therefore, the stroke patients will experience disability in performing activities of daily living (ADL) such as eating, writing, performing housework, and so on.

Several evaluation methods are commonly used to assess the effect of stroke in performing ADL [30], [34], [35]. The most famous assessment methods for evaluating upper limb functionality include Motricity index of arm and leg, Fugl-Meyer motor assessment, Motor assessment scale, Fatigue severity scale, Barthel index, Functional independence measure, Frenchay activities index, Rankin scale, Frenchay arm test, Action research arm test, and Wolf motor function test [35]–[37]. More details about these assessment methods are provided in Appendix A. Despite the wide application of these methods, all of them are subjective techniques, i.e., they are either questionnaires or qualitative scores assigned by a medical professional [30], [35].

We hypothesize that providing a more quantitative metric can enhance evaluation of the rehabilitation progress, and lead to a more efficient rehabilitation regimen tailored to the specific needs of each individual patient. For instance, such a quantitative methodology can help to defer the plateau in the patient's recovery. 'Plateau' is a term that is used to explain a stage of stroke recovery at which functional improvement is not observed (see Figure 1.1), and is determined by clinical observations, empirical research, and patient reports. In spite of the importance of plateau time as an indication of the time to discharge a patient form physiotherapy post stroke, researchers have questioned the reliability of current methods for determining plateau [38], [39]. Demain et.al [38] implemented a standard critical appraisal methodology, and found that the definition of recovery is ambiguous. For instance, there is a 12.5 -26 weeks variability in plateau time for

ADL. A few parameters have been pinpointed to cause this inconsistency. For example, they found that only 50% out of 1022 of physiotherapists use the validation tools to determine plateau, the duration and frequency of measurements are not consistent, and the measurement tools are qualitative [38]–[40]. An early and unnecessary discharge from physiotherapy can leave the patient with a permanent yet potentially preventable disability. Whereas a more reliable technique to indicate the start of plateau can help to determine the time to adjust the rehabilitation regimen and minimize the neuromuscular adaptation that in turn can delay the plateau [40].
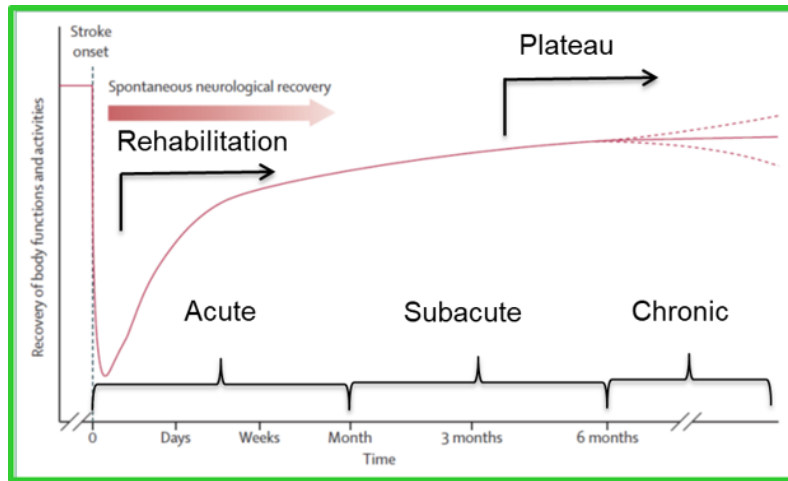
Figure 1.1: Stroke timeline [30]

The term "Activities of Daily Living" have been used in many fields such as rehabilitation, occupational therapy, and gerontology to describe patient's ability to perform daily tasks to maintain an unassisted living [41]. Since this term is very qualitative, researchers have proposed many sub-categories of ADL such as physical self-maintenance, activities of daily living, and instrumental activities of daily living [42] to assist physicians or occupational therapists to evaluate patient's ability in performing ADL in a more justifiable fashion [41], [43], [44].

A fundamental step toward developing a quantitative ADL assessment methodology is to distinguish different ADL tasks. Different taxonomies, such as the occurrence rate of a certain grasp in performing ADL or involvement of different fingers and hand palm, have been utilized for grasp classification [35], [45]. In this work, two types of sensors, i.e. EMG and IMU sensors, and an LMC infrared camera were used to record data from the lower arm of healthy-adult participants performing some common daily activities. Then, the data is processed using statistical and machine learning techniques to pave the way for development of a quantitative ADL assessment metric. The data collection procedure and the outcome of this research are presented in chapters 1.2-3, respectively.

1.2 Materials and Methods

1.2.1 Participants and tasks

Nine healthy adults including three females and six males, with intact hands were recruited to participate in this study, and informed consent was obtained from all the participants. The age range of the participants

was 25-62 years with an average of 37 years. This study was approved by Institutional Review Board office of University of Illinois Urbana-Champaign, and there was no limitation in terms of occupation, gender or ethnicity for recruiting the participants. The IRB approval letter is presented in Appendix B.

Each subject attended one session of data collection, and six of the participants completed two sets of tasks while two of them only completed one due to time limitations. Each set of tasks contained eight randomly distributed tasks, and the order of the tasks in the two sets was different. The subjects were asked to rest for 45 seconds between tasks to avoid muscle fatigue. During the task, the subjects were seated on a regular office chair with back support. Each task was performed by the participant's dominant hand, and composed of static and dynamic phases. In the static phase, the participants were instructed to rest their forearms on a regular office desk to avoid tremor and hold an object, as listed in Table 1.1, for around 10 seconds similar to how they would hold it in daily life. In the dynamic phase of the task, they were instructed to utilize the object over the entire range of motion that is usually performed in daily living at their own pace. Each dynamic task was repeated continuously for 5 times without any rest intervals in between. Table 1 and Figure 1.2 demonstrate the ADL tasks.

Table 1.1: Dynamic tasks of the ADL dataset [46].

Object	Dynamic Task
Cup	Grabbing a cup from the table top and bringing it to mouth to pretend drinking from the cup and put it back on the table
Fork	Bringing pretended food from a paper plate on the table to the person's mouth
Key	Locking/unlocking a pretended door lock while holding a car key
Knife	Cutting a pretended stake by moving the knife back and forth
Nail Clipper	Holding a nail clipper and pressing/releasing its handles
Pen	Tracing one line of uppercase letter "A"s, with 4 randomly distributed font sizes
Spherical Doorknob*	Rotating a doorknob clockwise and counter clockwise
Spoon	Bringing pretended food from a paper plate on the table to the person's mouth
* A cup was used instead of a spherical doorknob and the participants were instructed to mimic the hand posture of holding a spherical doorknob	
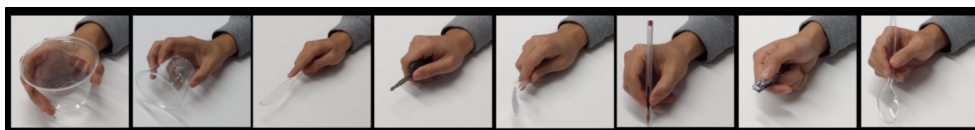
Figure 1.2: ADL tasks

The selected tasks in this study include a variety of ADLs from physical self-maintenance; e.g. utilizing spoon, fork, knife; and activities of daily living, e.g. writing. In addition, based on Cutkosky grasp taxonomy, the tasks in this study include precision grasps such as holding pen, spoon, and spherical doorknob as well as power grasps like holding glass, knife, and nail clipper [35], [47], [48]. These tasks include a diverse palm/finger involvement and facilitate the analysis of hand and fingers grasp over the entire range of motion that is typically used in ADL.

1.2.2 Data collection

A leap motion controller (LMC) [49] and a Myo armband [50] were used to collect data from the dominant lower arm of each participant. This way, both vision-based and sensor-based data collection methods for HGR were applied in this study.

LMC is a low-cost TOF camera for hand motion tracking that contains a pair of stereo infrared cameras and three infrared LEDs. Using the infrared light data, the device creates a grayscale stereo image of the hands. As shown in Figure 1.3, LMC camera is designed to either be placed on a surface, e.g. on an office desk, facing upward or be mounted on a virtual reality headset. LMC reads the sensor data and performs any necessary resolution adjustments in its local memory. Then streams the data to Ultraleap's hand tracking software on computer via USB, compatible with both USB 2.0 and USB 3.0 connections. LMC is a marker-free visual-based sensor, and its interaction zone is between 10 cm to 80 cm, from the device, and in a 140°×120° field of view, as shown in Figure 1.4 [49], [51], [52].
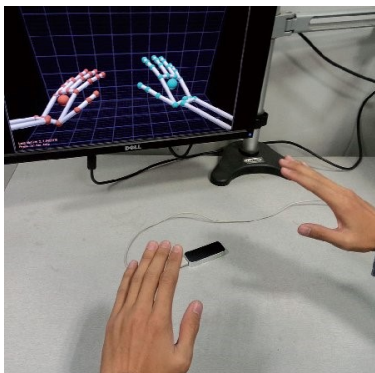
Figure 1.3: Leap Motion Controller connected to the Leap Motion Visualizer software showing the hands on top of the LMC camera [52]
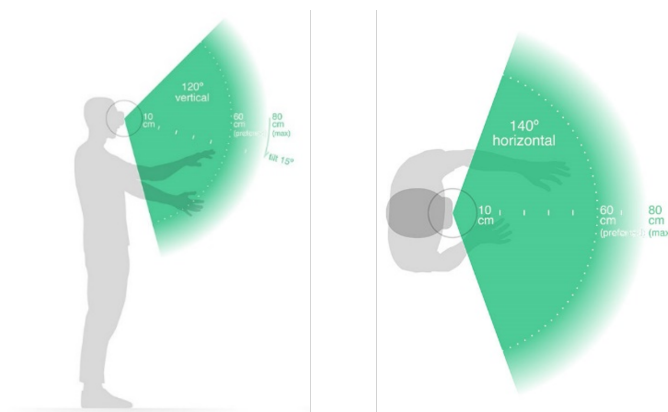
Figure 1.4: LMC's interaction zone [53]

Myo armband is a low-cost, wireless, consumer-grade, and easy-to-use bracelet that contains eight dry surface-EMG sensors and one nine-axis (3-axis acceleration, 3-axis gyroscope, and 3-axis geomagnetic) IMU sensor. The sampling rate for the EMG and IMU sensors are 200 Hz and 50 Hz, respectively [54], [55]. This armband was originally designed for remote computer control as it transfers signals from sensors to the computer via Bluetooth. However, Myo armband found its place as a reliable device for collecting EMG

and IMU data for research in the field of HGR [56]–[62]. Figure 1.5 shows, Myo armband along with its placement on the lower arm, eight sample EMg signals from the eight sensors, and five hand gestures that Myo software uses to control the computer.
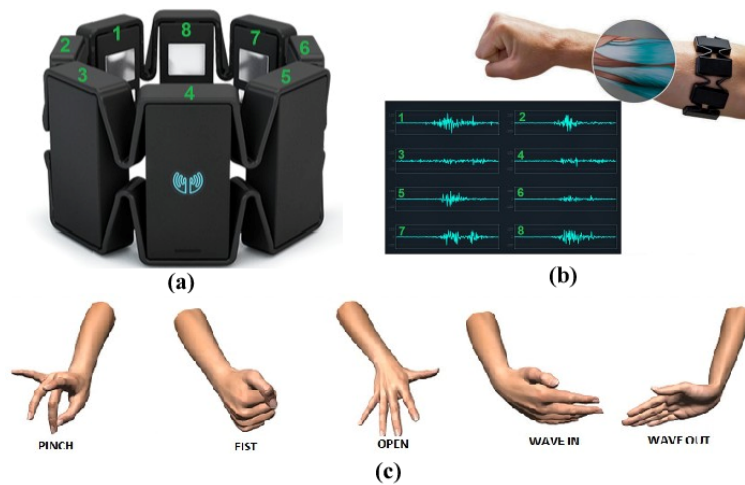
Figure 1.5: (a) Myo armband with eight EMG sensors, (b) armband placement on lower arm with eight sample signals from the EMG sensors, and (c) five hand gestures for controlling computer [63]

More information about the data collected each of these sensors, data analysis procedure, and the results are presented in chapters 2, 3 and 4.

Chapter 2

Hand Gesture Recognition Using Surface Electromyography

Abstract

Surface electromyography has become one of the popular methods for recognizing hand gestures. In this paper, the performance of four classification methods on sEMG signals have been investigated. These methods are developed by combinations of two feature extraction methods, including Mean Absolute Value and Short-Time Fourier Transform, and two classifiers, including Support Vector Machine and Convolutional Neural Network. These classification methods achieved an accuracy over 97% on the NinaPro dataset 1. In addition, a new dataset, which includes the Activities of Daily Living, was proposed and an accuracy over 98% was obtained by applying the presented classification methods[1].

Clinical Relevance— This methodology can provide the basis for a robust quantitative technique to evaluate hand grasps of stroke patients in performing activities of daily living that in turn can lead to a more efficient rehabilitation regimen.

2.1   Introduction

Stroke is one of the leading mortality and disability causes worldwide. Several evaluation methods have been used to assess the effects of stroke in performing activities of daily living (ADL). However, these methods are subjective [30]–[35]. Presence of a quantitative ADL realization assessment method, which enhances a better evaluation of hand's function and therefore a more efficient rehabilitation regimen, is lacking. A fundamental step toward such a quantitative method is to distinguish different hand gestures while performing ADL tasks.

A recent advancement of gesture recognition using surface electromyography (sEMG) has demonstrated a promising outlook for sign language, prosthetics, robotics, virtual reality, and clinical assessment [4], [10], [64]–[67]. We hypothesize such sEMG-based gesture recognition methods can be applied to ADL tasks, as well. The established gesture recognition approach is to extract features from sEMG signals and feed them to classifiers. Traditional features such as spectrograms, and mean absolute values (MAV), continue to be used as the state of the art, while the choice of classifiers has converged to Convolutional Neural Network (CNN) in recent literature as it significantly outperforms other classifiers [55], [60], [68]–[70].

---

[1]Includes previously published material [46]

In this paper, a dataset that includes sEMG data from 8 intact subjects while performing 8 different ADL tasks, referred to as "ADL dataset" hereafter, is presented. Compared to other available datasets for hand posture, data collection method in ADL dataset is relatively simple and inexpensive while this dataset includes more complicated tasks that better mimic the daily living activities. Particularly, the ADL dataset includes dynamic tasks to capture hand posture in performing ADL tasks over the entire range of motion being used in daily life. The second scope of this paper is to evaluate the performance of four classification systems in classifying hand gestures in ADL dataset. These systems are composed of the different combinations of MAV and Short-Time Fourier Transform (STFT) feature extraction methods with Support Vector Machine (SVM) and CNN classifiers. Then, the classification results are compared to those of the NinaPro dataset 1 to verify robustness of the systems. The NinaPro datasets are commonly used as benchmark for sEMG studies because of their reliable protocols and standard movements. NinaPro dataset 1 includes the kinematic and sEMG (10 electrodes) data from the upper limb of 27 intact subjects performing 52 hand postures with 10 repetitions of each posture and hand rest posture between the repetitions [71].

The remaining sections are organized as follows: Section 2.2 introduces ADL dataset, along with the description of the feature extraction methods and the classifiers. The classification results are presented and discussed in Section 2.3. Finally, Section 2.4 contains a summary of the current chapter.

2.2   Materials and methods

2.2.1   Subjects and Data Acquisition

Eight healthy adults, two females and six males, with intact hands were recruited to participate in this study and informed consent was obtained from all the participants. The age range of the participants was 25-62 years with an average of 37 years. This study was approved by Institutional Review Board office of University of Illinois at Urbana-Champaign and there was no limitation in terms of occupation, gender or ethnicity for recruiting the participants. Each subject attended one session of data collection, and six of the participants completed two sets of tasks while two of them only completed one due to time limitations. Each set of tasks contained eight randomly distributed tasks and the order of the tasks in the two sets was different. There was around 45 s rest between tasks to avoid muscle fatigue. During the task completion, the subjects were seating on a regular office chair with back support. Each task was performed by the participant's dominant hand and composed of static and dynamic phases. In the static phase, the participants were instructed to rest their forearms on a regular office desk to avoid tremor and hold a particular object for around 10 seconds in a way that they would hold it in daily life. In the dynamic phase, they were instructed to utilize the object over the entire range of motion that is usually performed in daily living at their own pace. Each dynamic task was repeated continuously for 5 times without any rest intervals in between. Table 1.1 provides a summary of the dynamic tasks. Figure 2.1 demonstrates the set of ADL tasks as well as their most similar tasks from the NinaPro dataset 1.
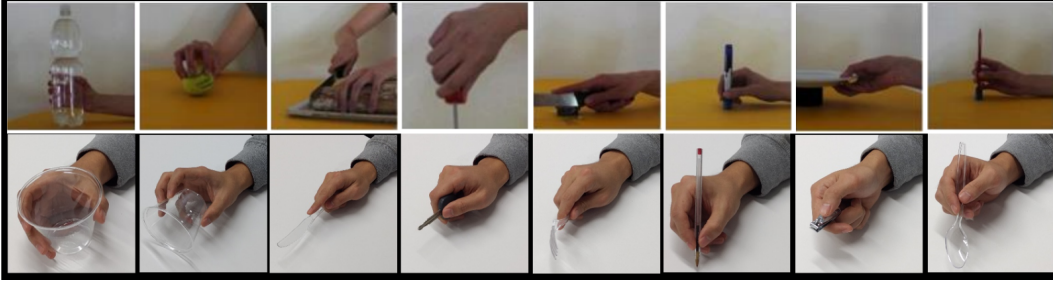
Figure 2.1: Row 1: Tasks from NinaPro dataset 1 that are the most similar to tasks of the ADL dataset, Row 2: Tasks from the ADL dataset

To record sEMG signals, a Myo armband [50] was placed on the participant's dominant forearm, around two inches from the elbow joint. Myo armband is a low-cost consumer-grade armband that contains an array of eight dry-electrodes with a sampling rate of 200 Hz. It is an easy-to-use armband as a user can simply slip the bracelet on and no preparation is required [55]. Figure 2.2 shows the armband placement on a participant's forearm during the static and dynamic phases of the experiment.
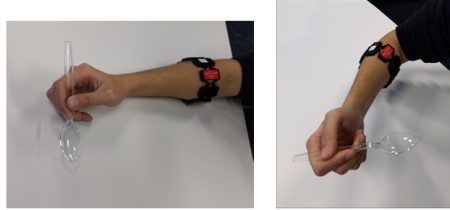
Figure 2.2: A participant performing static task (left) and dynamic task (right) while wearing Myo armband

2.2.2 Features and Classifiers

A first order Butterworth low-pass filter with a 1 Hz cutoff frequency was applied on rectified ADL dataset and NinaPro dataset 1. Next, both datasets were standardized per channel, i.e. each channel of sEMG data was standardized to have a mean of 0 and a standard deviation of 1. Then, different combinations of two feature extraction methods and two classification algorithms were tested on both datasets. MAV and STFT were implemented to extract time domain and time-frequency domain features, respectively. MAV was calculated over a 0.15 Second sliding window on each sEMG channel. Therefore, at each sampling point, the number of extracted features with MAV is the same as the number of sEMG sensors in each dataset. A Hamming window of size L with an overlap of $L/2$ was applied to calculate STFT. The window size was designed to provide a frequency resolution of 20 Hz. Due to spatial and temporal correlation between the neighboring sensors, the obtained matrices from the feature extraction algorithms are treated as 2-dimensional feature maps with the dimensions of number-of-samples $\times n$ as opposed to 1-dimensional data streams on the independent channels, where n is the number of extracted features. Afterwards, the extracted features were fed into SVM with radial basis kernel and CNN algorithms.

The proposed architecture of CNN is illustrated in Figure 2.3. PyTorch was used for the framework to build CNN [72]. CNN architecture is composed of 3 convolution layers and 1 linear layer. Three convolution layers have output channels of 16, 32 and 32 in sequential order, and each convolution layer consists of 2 x 2 filters with a stride of 1 and zero padding of 1. Rectified Linear Unit (ReLU) activation function and batch

normalization function are applied at the end of each convolution layer and maximum pooling function is applied at the end of the first and second layers. Fifty percent dropout is implemented at the end of the fully connected layer, i.e. after the linear function in Figure 2.3. A learning rate for training the CNN algorithm was set to 0.004.
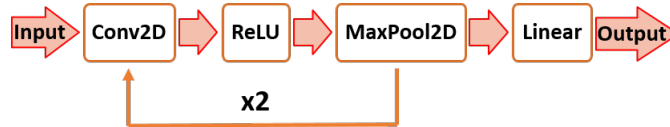
Figure 2.3: Proposed CNN architecture

2.3   Results and Discussion

The reported accuracies in this section are based on 5-fold cross validation method. Despite the simplicity of the presented CNN architecture, which leads to a lower computational cost, the achieved classification accuracy by applying MAV/CNN system on the entire NinaPro dataset 1 is over 97% that is comparable to the previous literature [70], [71], [73]. As a result, this architecture is selected for classifying the ADL dataset.

Table 2.1 demonstrates the achieved classification accuracies by applying different classification systems on static subset, dynamic subset, and the combination of both subsets from ADL dataset. For all the subsets, regardless of the choice of feature extraction method, SVM algorithm provides better classification accuracies. This may be correlated to the fact that CNN algorithm requires larger dataset to achieve a high accuracy while SVM performs superbly on smaller datasets as well.

Table 2.1: Evaluation of different classification systems on ADL dataset (numbers are accuracy in percent)

	SVM		CNN	
	MAV	*STFT*	*MAV*	*STFT*
Static	99	87	90	81
Dynamic	92	80	79	65
Static & Dynamic	95	79	79	65

In an assessment of the performance of feature extraction methods, it can be inferred that MAV features result in higher accuracies for each classification algorithm. Particularly, the combination of MAV/SVM outperforms other classification systems. However, the combination of STFT/SVM obtains accuracies that are comparable with the accuracies achieved by MAV/CNN for all data subsets, specially for dynamic subset and the combination of static and dynamic tasks. This observation suggests that a robust classification algorithm may compensate for a less impeccable feature extraction method to some extent. Similarly, a high performance feature extraction technique can partially alleviate the adverse effect of an imperfect classification method.

In comparing between different subsets, a higher accuracy was achieved on the static subset. This result can be correlated to the larger intra class variability of the dynamic tasks, which lead to lower accuracies in dynamic subset and the combination of both dynamic and static subsets. Despite the mutable nature of the dynamic tasks that makes it challenging to classify them, the MAV/SVM algorithm could achieve 92%

classification accuracy, which is only 7% lower than the achieved accuracy for classifying the static subset with the same classification system. To the best of our knowledge, this paper is the first to classify dynamic ADL tasks.

The confusion matrix for ADL dynamic subset classified by STFT/SVM system is presented in Figure 2.4. The row numbers represent the true labels, and the column numbers correspond to the predicted labels. Uniform distribution of the off-diagonal elements in the confusion matrix confirms that the SVM model has not been overfitted to any of the classes. The majority of misclassification cases are between the classes of utilizing spoon, fork, and cup. The correlation between hand gestures in holding fork, spoon, and pen is in-line with the fact that some of the participants utilized extremely similar hand grasps in holding these items, as presented in Figure 2.1, to the extent that even for a human eye, it was barely possible to distinguish between their hand postures while holding these objects. The correlation between these three tasks is observed in the confusion matrix of static dataset, Figure 2.5, as well. In contrast, in the dynamic dataset, the main source of classification error is misclassifying the three classes with the similar gross arm motions. Since these gross arm motions are generated by the muscles in the upper arm and shoulder while the Myo armband is placed on the lower arm, these results indicate that the detected signals by Myo armband during the gross forearm motion are very noisy. In other words, the signals from other sources such as skin movement are high compared to the signals from the muscle groups that are responsible for generating the fine hand grasps.
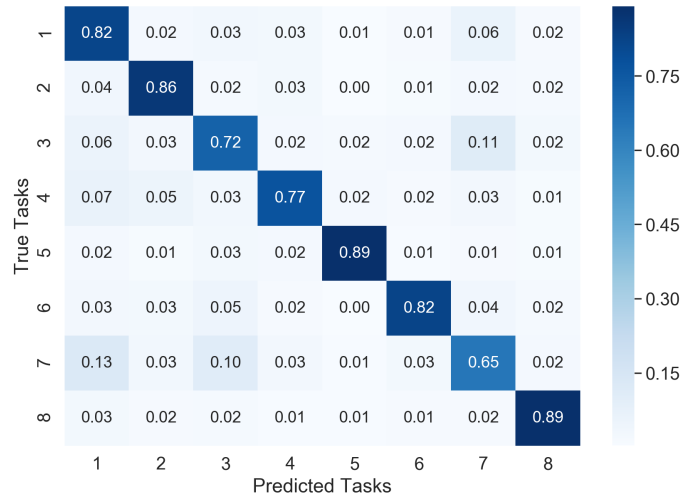
Figure 2.4: Confusion matrix for dynamic ADL dataset classified by STFT/SVM system
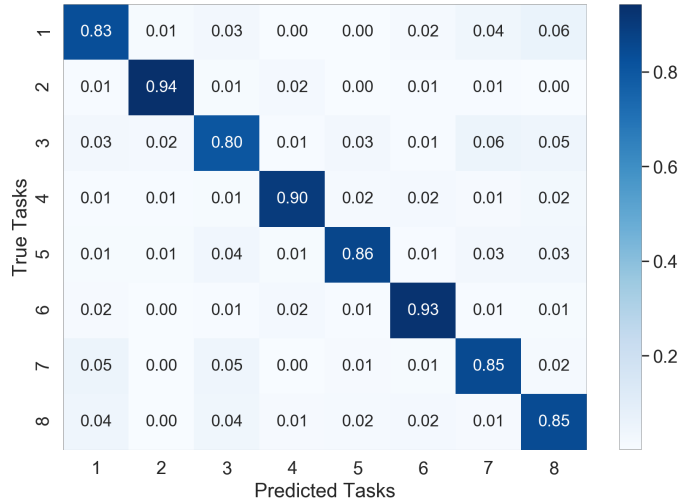
Figure 2.5: Confusion matrix for static ADL dataset classified by STFT/SVM system

Table 2.2 presents a comparison between static subset of ADL dataset and NinaPro dataset 1. In order to facilitate a meaningful comparison between the two datasets, only 8 tasks from NinaPro dataset 1 that include the most similar hand postures to the tasks of ADL dataset are considered in this analysis. These tasks are presented in Figure 2.1.

Table 2.2: Comparison between static subset of the ADL dataset and Ninapro dataset 1 (numbers are accuracy in percent)

	SVM		CNN	
	MAV	*STFT*	*MAV*	*STFT*
Static ADL subset	99	87	90	81
NinaPro dataset 1	99.5	97	98.9	97

Similar to the results that are observed in evaluating the performance of different classification systems on different subsets of the ADL dataset, MAV/SVM system outperforms other algorithms on NinaPro dataset 1 too. In contrast to the ADL dataset, regardless of the choice of features, the CNN algorithm provides high classification rates similar to those of SVM. These results prove the proposed CNN algorithm can achieve high classification accuracies if the size of the dataset is large enough.

All the classification systems, demonstrated higher performance in classifying the tasks of NinaPro dataset 1. This observation can be anticipated since the ADL dataset includes several hand gestures that appear the same, as shown in Figure 2.1. In particular, a person holding a spoon and a person holding a pen share similar grasps. Therefore, the ADL dataset contains more overlapping classes that can adversely affect the classification accuracy, whereas the NinaPro dataset 1 includes more distinct classes. Nonetheless, NinaPro dataset 1 contains data from a higher number of participants that can increase the intraclass variability and therefore, reduce the classification rate. As a result, by applying a robust classification system such as MAV/SVM, the classification accuracies for the two datasets are similar. Although the overlapping classes in the ADL dataset makes it more challenging to classify the tasks in the ADL dataset than the NinaPro dataset 1, they are also more suitable for the real world applications as they better represent the real ADL tasks. Therefore, expanding the ADL dataset by including data from more participants can provide the community

with a benchmark for both static and dynamic ADL tasks. More advanced models, e.g. multi-view CNN [70] or auto-encoder model [74], can improve the classification accuracy in such a benchmark.

2.4 Conclusion

In this paper, four systems of classification have been presented. These systems are made from the combinations of MAV and STFT as feature extraction methods, followed by SVM and CNN as classifiers. The classification performance of the proposed systems have been tested on the ADL dataset and the NinaPro dataset 1. Furthermore, to address a lack of dynamic ADL tasks in the currently available datasets, the ADL dataset has been presented and the quantitative experiments have been conducted to probe the properties of this new dataset. A higher classification performance in static tasks compared to dynamic tasks was observed that indicates the higher noise in dynamic signals. For the ADL dataset, SVM algorithm provides higher classification rate compared to CNN, regardless of the choice of features. In contrast, both classifiers result in similar classification accuracies on NinaPro dataset 1. From these observations, it can be deduced that the presented CNN algorithm may achieve higher accuracies if the size of the ADL dataset will be increased. Generally, STFT features result to a lower accuracy compared to the MAV features.

Future works will investigate three directions. CNN architecture will be improved implementing more advanced architectures such as multi-view CNN or auto-encoder model. Furthermore, more feature types and classifiers will be investigated. Finally, the ADL dataset will be expanded by recruiting more participants and including more tasks, so the dataset can better capture the distribution of the sEMG signals in various ADL tasks.

Chapter 3

Feasibility Study: Hand Gesture Recognition Using Leap Motion Controller

3.1 Introduction

Analysis of collected data from Leap Motion Controller (LMC) [49] for HGR is composed of two sections. The first section , i.e. the current chapter[1], includes a preliminary study while in the second section, i.e. chapter 4, the main dataset has been presented and analyzed. LMC is a TOF camera that contains a pair of stereo infrared cameras and three illumination LEDs. As shown in Figure 1.3, LMC camera is designed to either be placed on a surface, e.g. on an office desk, facing upward or be mounted on a virtual reality headset. In this study, a 7-degrees-of-freedom robotic arm, i.e. Cyton Gamma 300 [75], is used to hold the LMC at an optimum position to minimize occlusion. The experimental set-up, and hand model in LMC with the global coordinate system (GCS) are provided in Figures 3.1a and 3.1b, respectively.

In this work, the performance of different methodologies to classify a variety of hand grasps commonly being used by human adults in performing ADL is presented. Nine sets of preliminary data, each containing 3 trials of 3 different tasks, are collected. These tasks are selected to include a diverse palm/finger involvement while addressing some of the most basic ADL tasks that are affected by stroke. These tasks facilitate the kinematic analysis of hand and fingers over the entire range of motion that is typically used in ADL. The tasks include (1) tracing two lines of letter "A" with randomly distributed font sizes on an A4 paper while holding a pen, (2) holding a spoon and moving it from a plate on a desk to mouth while the participant seats on a chair, and (3) holding and twisting a spherical doorknob clockwise and counterclockwise. Based on Cutkosky's grasps taxonomy [48], all these tasks are categorized as precision, opposed to power grasps. The first two tasks are prismatic grasps while the last one is a circular grasp.

---

[1]Includes previously presented material:

- Hajar Sharif, Pramod Chembrammel, and Thenkurussi Kesavadas , "A Methodology for Classification of Activities of Daily Living Based on Anatomical Data Obtained Using an Infrared Camera", poster presented as EMBC 2018.

- Hajar Sharif, Pramod Chembrammel, Thenkurussi Kesavadas, "Classification of Activities of Daily Living Based on Hand Kinematics Obtained Using an IR Depth Sensor", poster presented at IMECE 2018.
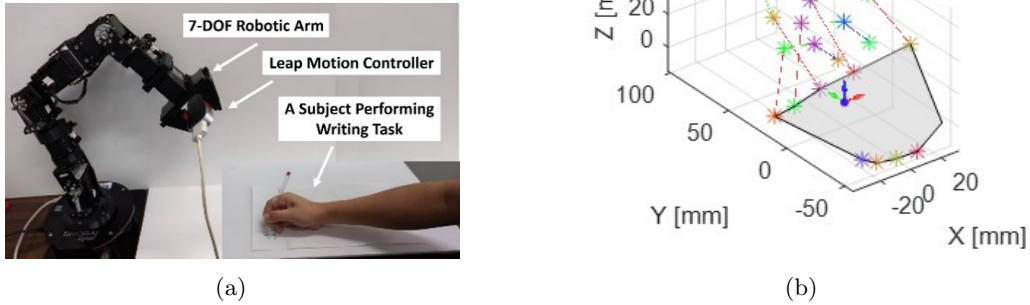
Figure 3.1: **(a)** Experimental setup, and **(b)** hand model in global coordinate system

3.2 Data analysis

In this study, two different methods were investigated for classifying hand grasps.

3.2.1 Method 1: Features obtained from the distal phalanges vectors

While a participant performs a particular task, referred to as a trial hereafter, in each sample i.e. each frame of the LMC camera, 17 features are defined for each finger based on the orientation of the vector of distal phalanges, with respect to a hand coordinate system (HCS) [76]. Therefore, the data of each trial can be presented as an $n \times 85$ matrix where n is the number of frames in the trial and 85 is the number of features for all 5 fingers. In this work, we propose to investigate the singular values (SVs) of these feature matrices in order to distinguish different tasks.

Histogram of dominant SVs of all the trials is presented in Figure 3.2a. This pattern is observed in the majority (6 out of 9) of the data sets, therefore those 6 data sets are further analyzed. Figure 3.2b demonstrates the 5 largest SVs, located on dotted, dashed, and solid lines for tasks 1 to 3, respectively. Figure 3.2b also shows the fitted normal distribution to the largest two SVs of the six data sets shown in blue, green, and red for tasks 1 to 3, respectively. Means and variances of these distributions are significantly different. Moreover, T-test (p-value=0.05) on empirical cumulative distribution function of SVs between task-1 and other tasks confirmed that the differences in SVs' distribution are statistically significant.
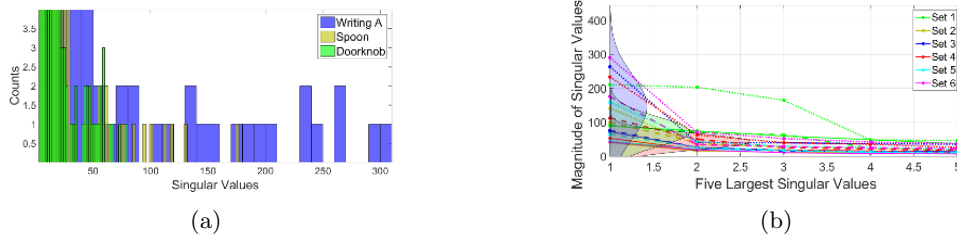
Figure 3.2: **(a)** Histogram of SVs of 9 Data Sets, and **(b)** the 5 largest SVs and distributions fitted to the largest 2 SVs of 6 data sets for all 3 tasks shown in blue, green, and red for Writing, Spoon, and Doorknob tasks, respectively

16

3.2.2 Method 2: Features obtained from the interphalangeal angles

In each trial of an ADL task, the coordinates of hand joints are detected from which the 15 interphalangeal angles are calculated at every frame of the depth sensor. Therefore, an $n \times 15$ dimensional data matrix is acquired per trial, where n is the number of frames and 15 is the number of interphalangeal angles. The singular values (SVs) of these angle matrices are used to distinguish different tasks. T-test (p-value=0.05) confirms that the differences in means and variances of fitted Normal distributions to the dominant SVs of the three ADL tasks are statistically significant.

3.3 Conclusion

Results of the feasibility study prove that the proposed data collection setup and analysis methods are promising methodologies to classify different ADL tasks. To investigate their effectiveness further, a larger dataset is presented in chapter 4.

Chapter 4

Main Study: Hand Gesture Recognition Using Leap Motion Controller

Abstract: Stroke is one of the leading mortality and disability causes worldwide. Several evaluation methods have been used to assess the effect of stroke in performing activities of daily living (ADL). However, these methods are qualitative. One first step toward developing a quantitative evaluation method is to classify different ADL tasks based on the hand grasps. In this work, a dataset has been presented that includes the data collected by a leap motion controller from the hand grasps of healthy adults performing eight common ADL tasks. Then a set of features, in time and frequency domains, were combined with two well-known classifiers, i.e. support vector machine and convolutional neural network, to classify the tasks, and over 99% classification accuracy was achieved[1].

keyword: Leap Motion Controller; activities of daily living; hand grasps classification

4.1 Introduction

Many neurological conditions lead to motor impairment of upper extremity that includes muscle weakness, altered muscle tone, joint laxity, and impaired motor control [32], [33]. As a result, common activities such as reaching, picking up objects, and holding onto them are compromised. Therefore, such patients will experience disability in performing activities of daily living (ADL) such as eating, writing, performing housework, and so on [33].

Several evaluation methods are commonly used to assess problems in performing ADL [30], [34], [35]. Despite the wide application of these methods, all of them are subjective techniques, i.e. they are either questionnaires or qualitative scores assigned by a medical professional [30], [35]. We hypothesize that providing a more quantitative metric can enhance evaluation of the rehabilitation progress, and lead to a more efficient rehabilitation regimen tailored to the specific needs of each individual patient.

For instance, such a quantitative methodology can help to defer the plateau in the patient's recovery. 'Plateau' is a term that is used to explain a stage of stroke recovery at which functional improvement is

---

[1] Includes preprinted material at [77]

not observed (see Figure 1.1), and is determined by clinical observations, empirical research, and patient reports. In spite of the importance of plateau time as an indication of the time to discharge a patient form physiotherapy post stroke, researchers have questioned the reliability of current methods for determining plateau [38], [39]. Demain et.al [38] implemented a standard critical appraisal methodology, and found that the definition of recovery is ambiguous. For instance, there is a 12.5 -26 weeks variability in plateau time for ADL. A few parameters have been attributed with causing such inconsistency, among which, the qualitative nature of the assessment metrics can be mentioned [38]–[40]. An early and unnecessary discharge from physiotherapy can leave the patient with a permanent yet potentially preventable disability. Whereas a more reliable technique to indicate the start of plateau can help to determine the time to adjust the rehabilitation regimen and minimize the neuromuscular adaptation that in turn can delay the plateau [40].

The term "Activities of Daily Living" have been used in many fields such as rehabilitation, occupational therapy, and gerontology to describe patient's ability to perform daily tasks to maintain an unassisted living [41]. Since this term is very qualitative, researchers have proposed many sub-categories of ADL such as physical self-maintenance, activities of daily living, and instrumental activities of daily living [42] to assist physicians or occupational therapists to evaluate patient's ability in performing ADL in a more justifiable fashion [41], [43], [44].

A fundamental step towards developing a quantitative ADL assessment methodology is to distinguish different ADL tasks based on the hand gesture data. Based on the applied hardware to detect the hand gestures, hand gesture recognition (HGR) methods can be divided into sensor-based and vision-based categories [1]. In the sensor-based methods, the equipment for data collection is exposed to user's body whereas in the vision-based techniques different types of cameras are used for data acquisition [24], [25]. The vision-based methods do not interfere with the natural way of forming hand gestures; however, several factors such as the number and positioning of cameras, the hand visibility, and algorithms applied on the captured videos can affect the performance of these techniques [1].

Leap Motion Controller (LMC) is a marker-free vision-based hand-tracking sensor that has shown to be a promising tool for HGR applications [78], [79]. Several researchers used LMC to detect signs using hand gestures for American [80], [81], Arabic[82]–[85], Indian [86]–[88], and other sign languages [51], [89]–[93]. LMC has applications in education [94] and navigating robotic arms [95], [96]. Researchers have investigated LMC applications in medical fields [97], [98] including, but not limited to, upper extremity rehabilitation [99]–[102], wheelchair maneuvering [103], and surgery [104], [105]. Bachmann et al. [106] reviewed the application of LMC for 3D human-computer interface, and some studies focused on using LMC for real-time HGR [107], [108].

In this dissertation we used collected data from healthy adults to develop the first stage of quantitative techniques that have a wide range of applications in improving the outcome of assessments of many common neurological conditions. We have demonstrated two classification schemes based on SVM and CNN that can efficiently classify ADL tasks. These classifiers use the features extracted by existing feature engineering methods from the collected data. In addition, we have generated a dataset containing hand motion data collected using LMC while the participants performed a variety of common ADL tasks. We have tested the performance of the proposed classification schemes using this dataset.

The selected tasks in this dataset include a variety of ADLs from physical self-maintenance; e.g. utilizing spoon, fork, knife; and activities of daily living, e.g. writing. In addition, based on Cutkosky grasp taxonomy, the tasks in this study include precision grasps such as holding pen, spoon, and spherical doorknob as well as power grasps like holding glass, knife, and nail clipper [35], [47], [48]. These tasks include a diverse

palm/finger involvement and facilitate the analysis of hand grasp over the entire range of motion that is typically used in ADL.

4.2 Materials and Methods

4.2.1 Subjects and Data Acquisition

In this study, an LMC is employed to collect data from the dominant arm of the participants while performing tasks. LMC is a low-cost marker-free visual-based sensor that works based on TOF concept for hand motion tracking. It contains a pair of stereo infrared cameras and three infrared LEDs. Using the infrared light data, the device creates a grayscale stereo image of the hands. As shown in Figure 1.3, LMC is designed to either be placed on a surface, e.g. on an office desk, facing upward or be mounted on a virtual reality headset. To collect the ADL data, a 7-degrees-of-freedom robotic arm, i.e. Cyton Gamma 300 [75], was used to hold the LMC at an optimum position to minimize occlusion. The experimental setup, and hand model in LMC with the global coordinate system (GCS) are provided in Figures 3.1a and 3.1b, respectively. LMC reads the sensor data and performs any necessary resolution adjustments in its local memory. Then streams the data to Ultraleap's hand tracking software on computer via USB, compatible to both USB 2.0 and USB 3.0 connections. LMC's interaction zone is between 10 cm to 80 cm, from the device, and in a 140°×120° typical field of view, as shown in Figure 1.4 [49], [51], [52].

Nine healthy adults including three females and six males, with intact hands were recruited to participate in this study, and informed consent was obtained from all the participants. The age range of the participants was 25-62 years with an average of 37 years. This study was approved by Institutional Review Board office of University of Illinois at Urbana-Champaign, and there was no limitation in terms of occupation, gender or ethnicity for recruiting the participants.

Each subject attended one session of data collection, and six of the participants completed two sets of tasks while two of them only completed one due to time limitations. Each set of tasks contained eight randomly distributed tasks, and the order of the tasks in the two sets was different. The subjects were asked to rest for 45 seconds between tasks to avoid muscle fatigue. During the task, the subjects were seated on a regular office chair with back support. Each task was performed by the participant's dominant hand, and composed of static and dynamic phases. In the static phase, the participants were instructed to rest their forearms on a regular office desk to avoid tremor and hold an object, as listed in Table 1.1, for around 10 seconds similar to how they would hold it in daily life. In the dynamic phase of the task, they were instructed to utilize the object over the entire range of motion that is usually performed in daily living at their own pace. Each dynamic task was repeated continuously for 5 times without any rest intervals in between. Table 1.1 and Figure 1.2 demonstrate the ADL tasks.

4.2.2 Preprocessing

LMC provides the coordinates of hand joints and palm center, demonstrated in Figure 4.1, in 3-dimensional space. It also provides the coordinates of three orthonormal vectors at the palm center, which form the hand coordinate system (HCS), as shown in Figure 4.2. These coordinates are in units of millimeter with respect to the LMC frame of reference. The origin of LMC's frame of reference is located at the top, center of the hardware, as presented in Figure 4.3. Therefore, while a participant performs a particular task, referred to as a trial hereafter, in each sample i.e. each frame of the depth sensor, 84 coordinate values are recorded.

The output of the LMC for each trial is a matrix of $n \times 84$, where n is the number of samples, i.e. number of frames.
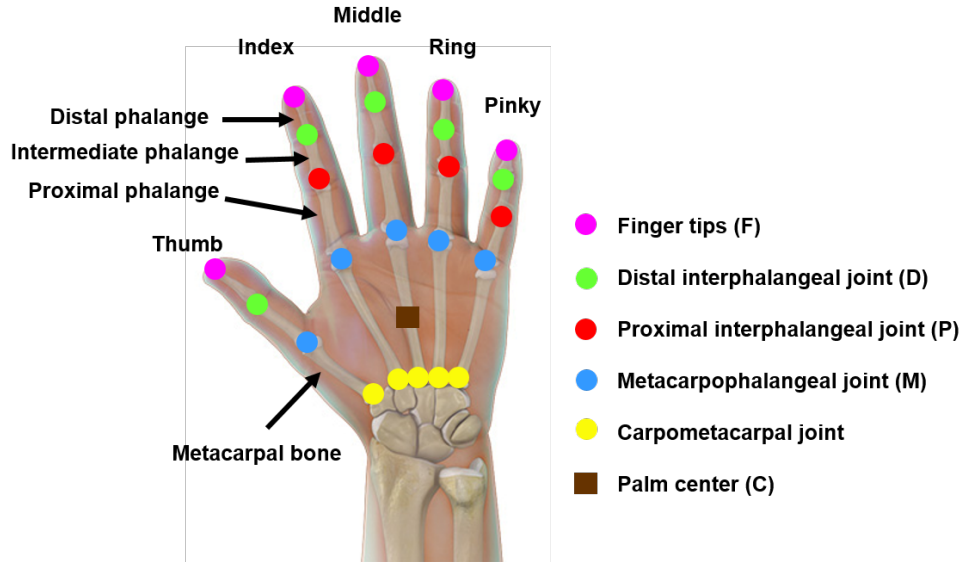
Figure 4.1: Hand joints and palm center [109]
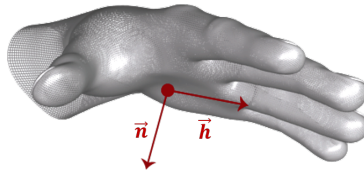
Figure 4.2: Hand coordinate system [110]
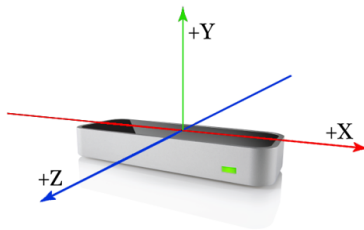
Figure 4.3: Leap Motion Controller frame of reference [111]

Change of basis

The first preprocessing step is transforming LMC data from the LMC coordinate system to GCS using Denavit-Hartenberg parameters [112] of the Cyton robot since the LMC is rigidly attached to the end-effector of Cyton.

Once the LMC data is transformed to GCS, the data is linearly translated into hand palm center. Afterwards, by using a change of basis matrix at each frame, data is transferred from GCS to HCS based on Equation 4.1. In this equation, A is the change-of-basis matrix, or transition matrix, and its columns are the

coordinates of the basis vectors of HCS in the GCS at each frame [113]. $XHCS$ and $XGCS$ are the data matrices in HCS and GCS, respectively.

$$XHCS = inverse(A) \times XGCS, \tag{4.1}$$

During the trials, the hand grasps, i.e. the relative position and orientation of the fingers and palm, do not change. In this work, the hand grasps are used for classifying different ADL tasks. Therefore, upper limb trajectories during the dynamic phase of the tasks, e.g. the entire-hand motions from plate to mouth while performing the "spoon" task, that are captured in the GCS should be removed. Transforming data from GCS to HCS eliminates gross hand motions and leaves the hand grasp information.

Filtering

At the next step, the transformed data is filtered using a median filter on a window size of 5 sampling points, i.e. 1/6 Second.

4.2.3 Features and Classifiers

Feature extraction

Choice of features to represent the raw data can significantly affect the performance of the classification algorithms [114]. In this work, three groups of features, as presented in Table 4.1, were calculated for each trial and later combined for classification. Followings, the features are explained in details.

Table 4.1: Features categories

Time-domain	Geometrical	AFA, ATD, DPUV, FHA, FTE, JA, NPTD
	Non-geometrical	MAV, RMS, VAR, WL
Frequency-domain	DFT	
Description of acronyms: Adjacent Fingertips Angle(AFA), Adjacent Tips Distance (ATD), Distal Phalanges Unit Vectors (DPUV), Fingertip-$\overrightarrow{h}$ Angle (FHA), Fingertip Elevation (FTE), Joint Angle (JA), Normalized Palm-Tip Distance (NPTD), Mean Absolute Value (MAV), Root Mean Square (RMS), Variance (VAR), Waveform Length (WL), Discrete Fourier Transform (DFT)		

<u>**Geometrical features in time-domain**</u>

In order to compensate for different hand sizes, the features should be normalized. The geometrical features that represent the angles are divided by π whereas the distance features are normalized to M. M is the accumulative Euclidean distance between the palm center and tip of the middle finger. At each sampling point, M is calculated by summation over distance between the palm center and metacarpophalangeal joint and lengths of all three bones of the middle finger, as presented in Equation 4.2. Since there were fewer variations between participants' hand grasps while performing the "cup" task, the coordinates in this task were used for M calculation. The final length for normalization was calculated by averaging M over the first 30 sampling points, i.e. the first second, of the first trial of the "cup" task.

$$M = \left|\overrightarrow{CM}\right| + \left|\overrightarrow{MP}\right| + \left|\overrightarrow{PD}\right| + |\overrightarrow{DF}| \tag{4.2}$$

1. Adjacent Fingertips Angle (AFA): This feature demonstrates the angle between every two adjacent fingertip vectors, which is the angle between vectors from the palm center to fingertips. AFA is calculated by Equation 4.3, where F_i represents Fingertip location. This feature is normalized to the interval of [0,1] by dividing the angles by π. Lu et.al [115] achieved 74.9% classification accuracy using the combination of this feature and hidden conditional neural field (HCNF) as the classifier.

$$AFA = \frac{\angle(F_i,\ F_{i+1})}{\pi} \quad i = 1,\ 2,\ \ldots,\ 4 \tag{4.3}$$

2. Adjacent Tips Distance (ATD): This feature represents the Euclidean distance between every two adjacent fingertips and is calculated by Equation 4.4 in which F_i represents Fingertip's location. There are four spaces between the five fingers of each hand, so there are four ATD in each hand. This feature is normalized to the interval of [0, 1] by dividing the calculated distances by M. Lu et.al [115] achieved 81.2% accuracy using the combination of this feature and HCNF.

$$ATD = \frac{|F_i - F_{i+1}|}{M} \quad i = 1, 2, \ldots,\ 4 \tag{4.4}$$

3. Distal Phalanges Unit Vectors (DPUV) [116]: For each finger, distal phalanges vector is defined as the vector from distal interphalangeal joint to the fingertip as presented in Figure 4.1. This feature is normalized by dividing to its norm.

4. Normalized Palm-Tip Distance (NPTD): This feature represents the Euclidean distance between Palm Center and each fingertip. NPTD is calculated by Equation 4.5 where F_i represents Fingertip location, and C is the location of the palm center. This feature is normalized to the interval [0,1] by dividing the distance by M. Lu et.al [115] achieved 81.9% accuracy using the combination of this feature and HCNF, while Marin et.al [117] achieved 76.1% accuracy using the combination of Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel and Random Forest (RF) algorithms.

$$NPTD = \frac{|F_i - C|}{M} \quad i = 1,\ 2,\ \ldots,\ 5 \tag{4.5}$$

5. Joint Angle (JA) [118], [119]: This feature represents the angle between every two adjacent bones at interphalangeal and metacarpophalangeal joints. For example, for distal interphalangeal joint, θ is derived by Equation 4.6.

$$\theta = \arccos\left(\frac{\overrightarrow{DF} \cdot \overrightarrow{PD}}{\left|\overrightarrow{DF}\right|\left|\overrightarrow{PD}\right|}\right) \tag{4.6}$$

6. Fingertip-$\overrightarrow{h}$ Angle (FHA): This feature determines the angle between the vector from the palm center to the projection of every fingertip on the palm plane and $\overrightarrow{h}$, which is the finger direction of the hand coordinate system as presented in Figure 4.3. FHA is calculated by Equation 4.7 in which the F_i^p is the projection of the F_i on the palm plane. Palm plane is a plane which is orthogonal to vector $\overrightarrow{n}$ and contains $\overrightarrow{h}$. By dividing the angles by π this feature is normalized to the interval of [0,1]. Lu et.al [115]and Marin et.al [117] achieved 80.3% and 74.2% accuracies classifying FHA features by HCNF and the combination of RBF-SVM with RF as classifiers, respectively.

$$FHA = \frac{\angle(F_i^p - C, h)}{\pi} \quad i = 1, 2, \ldots, 5 \tag{4.7}$$

7. Fingertip Elevation (FTE): Another geometrical feature is Fingertip elevation that defines the fingertip distance from palm plane. FTE is calculated by Equation 4.8 in which "sgn" is the sign function, and $\overrightarrow{n}$ is the normal vector to the palm plane. Like previous features, the F_i^p is the projection of the F_i on the palm plane. Lu et.al [115] achieved 78.7% accuracy using the combination of this feature and HCNF, while Marin et.al [117] achieved 73.1% accuracy classifying FTE features by the combination of SVM with RBF kernel and RF.

$$FTE = \frac{sgn\left((F_i - F_i^p) . \overrightarrow{n}\right) \|F_i - F_i^p\|}{M} \quad i = 1, 2, \ldots, 5 \tag{4.8}$$

Non-geometrical features in time-domain

In order to compensate for the variations imposed by different participants' hand sizes, the filtered data is normalized to M that is described in "geometrical features in time-domain" section. All the non-geometrical time-domain features were calculated over a sliding window with the size of 15 samples, which equals to 0.5 Second, with no overlap between the windows.

1. Mean Absolute Value (MAV): MAV is calculated by taking an average over the absolute value of the signal's amplitude, using Equation 4.9. MAV has shown promising results for classifying hand gestures [10], [46], [62], [114].

$$MAV = \frac{1}{N} \sum_{n=1}^{N} |X_n| \tag{4.9}$$

2. Root Mean Square (RMS): similar to MAV, RMS feature represents the signal in an average sense. RMS feature is calculated using Equation 4.10, where X_n is sampling point and N is the number of samples in the moving window [114], [120].

$$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^{N} X_n^2} \tag{4.10}$$

3. Variance (VAR): Variance of a signal is related to the deviation of the sampling points from their average, $\bar{x}$, and is calculated by Equation 4.11. Variance is the mean value of the square of these deviations [114].

$$VAR = \frac{1}{N-1} \sum_{n=1}^{N} (x_n - \bar{x})^2 \tag{4.11}$$

4. Waveform length (WL): Waveform length is derived by summation over the numerical derivative of the samples, and is given by Equation 4.12 [114], [120], [121].

$$WL = \sum_{n=1}^{N-1} |X_{n+1} - X_n| \tag{4.12}$$

Frequency-domain features

Discrete Fourier Transform (DFT): Since the coordinates are transferred to HCS, it is a valid assumption to assume the grasps, and therefore the joints coordinates, are constant through the entire task. Therefore, DFT is used to transfer signals from time domain to the frequency domain. numpy.fft.fft was used to extract DFT feature, based on Equation 4.13 where $W_N = e^{-j2\pi/N}$ [122].

$$
\begin{cases}
X[k] = \sum_{n=0}^{N-1} x[n]W_N^{nk} & k = 0, 1, \ldots, \text{N-1} \\
x[n] = \frac{1}{N}\sum_{n=0}^{N-1} X[k]W_N^{-nk} & n = 0, 1, \ldots, \text{N-1}
\end{cases}
\tag{4.13}
$$

Classification

The data matrix for each feature was formed by concatenating the features from all the trials of all the tasks. Size of the obtained matrix was $n \times m$, where n is the number of sampling points from all the trials of all the tasks, and m is the number of feature components. Data matrices were standardized, to have zero mean and unit variance, per column before being fed to the machine learning algorithms.

SVM is well-known to be a strong classifier for hand gestures [85], [105], [123]–[128]. It is a robust algorithm for high dimensional datasets with smaller number of sampling points. SVM maps data into a higher dimensional space and separates classes using an optimal hyperplane. In this study, scikit-learn library [129] was used to implement SVM with Radial Basis Function (RBF), and the parameters were determined heuristically [130].

Moreover, a Convolutional Neural Network (CNN) was implemented in PyTorch [72], [131] for classifying the tasks. CNN and its variations have shown to be efficient algorithms for hand gesture classification [132]–[135]. The proposed architecture of CNN is illustrated in Figure 2.3. CNN architecture is composed of three convolution layers and one linear layer. Three convolution layers have output channels of 16, 32 and 32 in sequential order, and each convolution layer consists of 2×2 filters with a stride of 1 and zero padding of 1. Rectified Linear Unit (ReLU) activation function and batch normalization function are applied at the end of each convolution layer and maximum pooling function is applied at the end of the first and second layers. Fifty percent dropout is implemented at the end of the fully connected layer, i.e. after the linear function in Figure 2.3. The learning rate, epoch, and batch size for training the CNN algorithm were set to 0.01, 20, and 40, respectively. The hyperparameters were determined experimentally.

4.3   Results and Discussion

PCA dimensionality reduction, adaptive learning rate for training the CNN algorithm, and different data filtering schemes were tested and got rejected as they showed to be detrimental to the classification accuracy. The 5-fold cross validation performance metrics of CNN and SVM algorithms in classifying ADL tasks on the pure data, i.e. filtered data in HCS, as well as different combinations of features have been presented in Table 4.2 and Table 4.3, respectively. Precision, recall, and F1-score are calculated using sklearn.metrics.precision_recall_fscore_support function, by setting average = 'macro', which calculates these metrics for each class and reports their average value.

Both algorithms do better at classifying some of the time-domain features when compared to their performance when classifying pure data. Among the time-domain, non-geometrical features, VAR and WL represent the data poorly as they are calculated based on the variations in the signal over time (Equation 4.11 and Equation 4.12). Since the data have been transformed to HCS, the grasps, and consequently the

coordinates of the joints, can be assumed constant over time. Therefore, VAR and WL are very similar in different tasks, and cannot be used to discriminate tasks from each other. Similarly, DFT feature can be assumed as the frequency decomposition of DC signals with different amplitudes. As a result, the inter-class variability in this feature is not high enough to achieve high classification accuracies.

Based on Table 4.2 and Table 4.3, SVM and CNN have comparable accuracies in classifying geometrical features. However, SVM outperforms CNN when features are combined. This can be correlated to the capability of SVM in classifying high dimensional datasets even when the number of samples are not proportionally high.

The achieved classification accuracies using AFA and FTE feature is lower than the achieved accuracies by a similar study [115]; however, the tasks classified in the two studies are very different. The ADL dataset includes tasks in many of them the fingers are flexed while the hand holds an object. This minimizes the variation in AFA and FTE among the tasks. In addition, to have a meaningful comparison between the results of different studies, the inclusion or exclusion of the gross hand motions in the classification should be taken into account. In the current analysis, information of the gross hand motions is removed from the data.

As demonstrated in Table 4.2 and Table 4.3, ATD and JA are the best features for classifying the tasks using both algorithms. ATD-CNN combination achieved over 99% classification accuracy and over 97% for both precision and recall. JA showed a better performance when combined with the SVM algorithm. JA-SVM combination achieved over 90% for both accuracy and precision and a recall over 89%. Moreover, combining two or more time-domain features can improve the classification performance using the same classifiers. Confusion matrices for both classifiers and sample geometrical features, which achieved over 70% accuracy, are presented in Table 4.4. The uniform distribution of off-diagonal elements in these matrices shows that the algorithms were not over fitted to any of the classes using these features.

Table 4.2: Performance metrics for different combinations of features and CNN as classifier using 5-fold cross validation. All the numbers are percentage values (%). Different sets of features, based on Table4.1, are determined by different colors

Feature	CNN			
	Accuracy	Precision	Recall	F-Score
Pure data	63.5	50.5	40.2	41.2
MAV	85.1	80.5	80.1	80.2
RMS	84.1	78.1	77.8	77.9
VAR	34.8	32.9	23.5	23.3
WL	36.7	31.4	29.2	29.6
AFA	57.3	54.4	52.3	52.6
ATD	99.88	97.5	97.3	97.4
DPUV	72	68.8	68	68.3
FHA	70.2	66.1	65.3	65.5
FTE	41.5	29	25.4	24.5
JA	77.4	74.4	73.9	74.2
NPTD	71.5	68.4	67.6	67.9
DFT	58.4	53.4	50.4	51.4
JA + DPUV	80.4	77.1	76.6	76.8
JA+NPTD	78.8	74.3	73.7	74
MAV+RMS	84	79.5	78.9	79.2
MAV+JA+NPTD	88.4	83.8	83.6	83.7
MAV+JA+NPTD+DPUV	87.59	82.9	82.5	82.7

Table 4.3: Performance metrics for different combinations of features and SVM as classifier using 5-fold cross validation. All the numbers are percentage values (%). Different sets of features, based on Table 4.1, are determined by different colors

Feature	SVM			
	Accuracy	Precision	Recall	F-Score
Pure data	68.9	70.5	67.3	68.2
MAV	79.5	81.1	77.9	78.99
RMS	76.3	78.4	74.7	75.8
VAR	24.6	61.1	21.8	20.3
WL	29.4	48.7	26.7	25.6
AFA	49.6	57.1	46.9	47.5
ATD	75.1	80.3	74.2	76.1
DPUV	79.3	79.2	78.3	78.7
FHA	64.2	66.2	62.5	63.3
FTE	30.8	50.8	29	30.4
JA	90.3	90.2	89.9	90
NPTD	79.3	79.2	78.3	78.6
DFT	52.4	77.6	50	54.3
JA + DPUV	94.7	94.4	94.4	94.4
JA+NPTD	92.3	92.2	91.9	92
MAV+RMS	79	80.5	77.5	78.4
MAV+JA+NPTD	92.5	92.3	91.9	92
MAV+JA+NPTD+DPUV	95.1	94.8	94.7	94.8

Table 4.4: Confusion matrices for sample combinations of features and classifiers. All the values obtained through 5-fold cross validation and are in percentage (%)

DPUV-SVM

True \ Predicted	Cup	Doorknob	Knife	Key	Fork	Writing	Nail Clipper	Spoon
Cup	75.8	4.4	4.0	6.1	3.6	2.5	2.3	1.4
Doorknob	3.9	64.6	8.8	7.6	3.7	3.3	5.6	2.3
Knife	1.6	4.3	78.8	4.8	1.7	3.6	4.0	1.1
Key	1.9	4.0	6.9	77.9	2.1	3.3	2.4	1.6
Fork	1.4	2.4	2.6	3.8	82.8	3.0	2.2	1.9
Writing	1.4	3.7	3.9	4.0	1.8	81.6	2.5	1.2
Nail Clipper	0.9	1.6	2.6	1.5	0.6	1.4	91.0	0.4
Spoon	1.9	2.4	3.7	5.2	8.1	3.0	1.6	74.0

DPUV-CNN

True \ Predicted	Cup	Doorknob	Knife	Key	Fork	Writing	Nail Clipper	Spoon
Cup	64.9	4.5	5.6	8.3	4.4	5.4	4.2	2.6
Doorknob	4.5	45.9	14.2	10.3	4.9	8.4	8.4	3.3
Knife	1.8	5.4	66.7	7.9	2.9	6.2	6.8	2.2
Key	3.1	4.8	8.9	66.3	3.9	6.1	4.1	2.6
Fork	2.1	2.9	3.6	4.5	75.2	5.3	3.5	3.0
Writing	2.4	3.5	5.5	5.2	3.2	73.4	4.2	2.6
Nail Clipper	1.1	2.2	5.1	2.9	1.4	3.4	83.0	1.0
Spoon	2.1	2.8	4.7	6.6	6.2	5.5	3.3	68.7

JA-SVM

True \ Predicted	Cup	Doorknob	Knife	Key	Fork	Writing	Nail Clipper	Spoon
Cup	87.5	3.3	2.1	3.3	1.0	1.1	1.1	0.7
Doorknob	2.9	80.3	4.7	2.8	1.0	1.4	6.4	0.5
Knife	1.0	3.5	88.7	1.6	0.5	1.4	3.0	0.2
Key	1.4	2.9	3.4	88.9	0.8	1.3	0.9	0.3
Fork	0.9	1.1	1.3	1.5	92.6	1.4	0.5	0.6
Writing	1.0	1.3	1.8	1.1	0.9	92.5	0.8	0.6
Nail Clipper	0.3	0.7	1.0	0.4	0.1	0.2	97.3	0.1
Spoon	1.3	1.7	1.4	1.2	1.3	1.3	0.7	91.2

JA-CNN

True \ Predicted	Cup	Doorknob	Knife	Key	Fork	Writing	Nail Clipper	Spoon
Cup	69.5	4.6	5.2	7.9	3.8	3.6	2.0	3.3
Doorknob	4.6	56.8	11.8	7.7	4.1	5.5	6.7	2.7
Knife	1.9	6.1	73.3	5.5	2.2	4.8	4.6	1.6
Key	3.3	4.7	8.4	71.1	3.7	4.4	2.1	2.3
Fork	2.5	2.6	3.6	3.8	79.5	3.9	2.0	2.3
Writing	1.8	3.1	6.1	3.8	2.4	78.0	2.6	2.3
Nail Clipper	0.9	2.2	3.8	1.4	0.9	2.0	87.8	1.0
Spoon	2.3	2.3	4.2	5.8	3.3	4.0	2.4	75.8

NPTD-SVM

True \ Predicted	Cup	Doorknob	Knife	Key	Fork	Writing	Nail Clipper	Spoon
Cup	75.6	4.4	4.0	6.1	3.6	2.5	2.3	1.4
Doorknob	3.8	64.8	8.8	7.6	3.7	3.4	5.6	2.3
Knife	1.6	4.4	78.8	4.8	1.6	3.6	4.0	1.1
Key	1.9	3.9	6.8	78.0	2.1	3.3	2.3	1.6
Fork	1.3	2.4	2.6	3.8	82.8	3.0	2.2	1.9
Writing	1.4	3.7	3.8	3.9	1.8	81.7	2.5	1.1
Nail Clipper	0.8	1.7	2.6	1.5	0.6	1.4	90.9	0.4
Spoon	1.9	2.4	3.7	5.2	8.1	3.1	1.6	74.0

NPTD-CNN

True \ Predicted	Cup	Doorknob	Knife	Key	Fork	Writing	Nail Clipper	Spoon
Cup	64.8	5.0	6.1	8.2	4.3	5.0	4.3	2.3
Doorknob	4.6	47.5	14.9	9.7	4.9	7.5	8.0	3.0
Knife	1.9	6.1	67.3	7.5	2.9	5.8	6.4	2.0
Key	3.2	5.6	9.6	65.4	3.9	5.9	4.0	2.5
Fork	2.3	3.4	3.7	4.5	75.2	4.9	3.4	2.7
Writing	2.5	4.1	5.9	5.2	3.4	72.1	4.2	2.5
Nail Clipper	1.3	2.7	6.0	2.9	1.7	3.3	81.2	1.0
Spoon	2.4	3.3	5.0	6.4	6.9	5.6	3.3	67.0

4.4 Conclusions and Future work

In this work, several classification systems have been presented. These systems are made from the combination of a variety of time-domain and frequency-domain features, followed by SVM and CNN as classifiers. The

classification performance of the systems has been tested on a proposed ADL dataset. The ADL dataset includes the leap motion controller data collected from the upper limb of healthy adults performing eight common ADL tasks. To the best of authors' knowledge, this is the first ADL dataset collected by LMC that includes both static hand grasps and dynamic hand motion of participants using real daily-life objects.

In this work, the data were transformed to the HCS, so only the grasp information, and not the gross hand motions, were used for classification. Over 99% classification accuracy and over 97% precision and recall were achieved by applying CNN on the "adjacent fingertips distance" feature. Eleven classification systems achieved over 80%, six over 90%, classification accuracy with high precision and recall. Although for the individual features CNN and SVM had comparable performance, for the combination of features, SVM outperformed the CNN algorithm. From these observations, it can be deduced that the presented CNN algorithm may achieve higher accuracies if the size of the ADL dataset is increased.

The findings of this study pave the way for developing an ADL-assessment-metric in two folds. The first fold is the immediate application of these findings in evaluating patient's performance, while the second fold is concerned with the long-term application.

In the current study, a data analysis pipeline was developed that takes in the LMC data from hand motion and outputs a classification accuracy to distinguish different ADL tasks. Different preprocessing, feature extraction, and classification methods were tested on the data from healthy adults to detect the best structure and parameters for the proposed pipeline. The developed pipeline can be set as a reference. Then, hand motion data from a neurological patient completing the same tasks in the same data collection setup can be fed into the reference pipeline to obtain the classification accuracy. The achieved accuracy indicates how close a patient's hand motions are to the hand motions of the healthy population. This method enhances assessment of the overall performance of a patient in a quantitative fashion. In addition, the acquired confusion matrix provides insight into the patient's performance in completing each individual task.

As for the long-term application, the features that achieved a higher classification rate can be used for further analysis and for developing other metrics as they represent different classes in a more distinguishable way. For instance, the distribution of these features, in each ADL task, among the healthy adults can be set as a reference metric. In this scenario, the location of patient's hand data in the reference distribution can be used to evaluate the patient's performance and the rehabilitation progress. More analysis of the data from healthy adults as well as collecting the same data from the neurological patients are required to complete this metric.

In conclusion, the future work can investigate three directions. Firstly, other classifiers can be investigated to increase the algorithm's speed. Furthermore, the LMC data can be transformed back to the global coordinate system to include gross hand motions and implement time series algorithms for classification. Finally, the ADL dataset can be expanded by recruiting more healthy and neurological-patient participants, so the proposed methodology will be advanced further toward development of a quantitative assessment method. Particularly, data from the neurological patients are crucial to generalize the findings of the current study for clinical applications.

Part II
HAND GESTURE RECOGNITION FOR EVALUATING SURGICAL SKILLS PROFICIENCY

Chapter 5

Introduction

5.1 Motivation

Acquiring excellent surgical and microsurgical skills are critical to avoid high stakes and morbidity to the patients. Gaining surgical skills requires significant practice and mentorship, yet opportunities to gain on-the-job surgical training are scarce. Administrative, legal, and ethical pressures often understandably preclude exposure of surgical patients to novice surgeons on the steepest part of the learning curve, which as a result, negatively affects many surgical training programs. In addition, fewer and fewer medical training programs use animals in the delivery of medical education due to societal pressures and perceptions.

As a result, development of non-living models and simulations using augmented and virtual reality technology is critical to advancement of surgical training in medical education. Evidence suggests that many highly technical skills may be acquired and refined outside of the operating arena through the use of surgical simulations [136]–[143]. Figures 5.1a and 5.1b demonstrate a microsurgery lab and a virtual reality simulator for intraocular surgery training, respectively.

<div align="center">(a)   (b)</div>

Figure 5.1: **(a)** Microsurgery Training Lab [144] , and **(b)** A virtual reality simulator for surgery training [145]

Many models and some simulation-based exercises for surgical, microsurgical, and minimally invasive surgical training and assessment have been described, but few have been validated as teaching and learning tools [136]–[138]. Validation of any surgical simulation system is needed to ensure skills transfer to real-life

surgical situations. Robust and quantitative skill assessment methods are required to perform such system validation. In an effort to standardize evaluation techniques, the use of checklists, validated global rating scores, time to complete a specified task, and assessment of independent completion of a specified task without the help of a supervising surgeon have been advocated. These standard evaluation techniques are rarely objective and consider the overall performance rather than focusing on the correct performance of the specific steps of the procedure [142].

Moreover, with traditional surgical training, as well as with the surgical simulation, evaluation of a trainee's performance involves the direct observation, supervision, and input of a surgical expert. Direct observation by an expert surgeon often takes one of two forms: The surgeon directly observes the trainee during a procedure in real time or watches recorded videos of the trainee performing a procedure after completion of the procedure to assess the trainee's skill level [146], [147]. The direct observational evaluation of a novice by an expert surgeon is very time-consuming and subjective [146]–[148].

Standard evaluation techniques are mainly concerned with one of the following three categories [149]–[153]:

1. Surgeon's hands: correct instrument grasp, correct hand position, number of hand motions, hand-motion trajectory, hand tremor, magnitude of force applied to instruments and tissue, etc.

2. Tissue: accuracy and precision of tissue handling, use of irrigation to prevent tissue desiccation, dissection technique, degree of trauma inflicted on tissue, etc.

3. Overall process performance: suture placement, use of appropriate number of sutures, accuracy of knot-tying technique, completion time, flow of operation, etc.

There have been a number of studies to quantitatively evaluate hand motion for assessment of microsurgical competency [154]–[158]. These studies chose task completion time, number of hand movements, and motion path length as the measurable outcomes for assessment. There have been fewer studies that focused on the hand/needle path [159] or path/velocity of both hands [160], [161] through task completion.

Incorporation of surgical simulation into longitudinal training courses and surgical curricula is sorely lacking and requires robust assessment methods [162]. In addition, provision of real-time quantitative feedback has been lacking with the use of many surgical simulation techniques [145], [163], [164]. Therefore, development of an evidence-supported, **A**utomated, robus**T**, real-**T**ime, comprehensiv**E**, quaN**tita**TIVE assessment system (ATTENTIVE) is crucial to achieve higher training performance and pave the way for more widespread inclusion of advanced simulators such as virtual and augmented reality into the surgical curricula.

In the second section of this dissertation, i.e. current section, it has been proposed to use a combination of different sensors to evaluate kinetic, kinematic, and video data to devise ATTENTIVE by including a variety of assessment parameters to understand any correlation among them. Consequently, the ATTENTIVE system is applied to evaluate surgical performance of novice and expert surgeons, from human and veterinary medicine fields, in performing basic surgical tasks. The outcome of ATTENTIVE will be compared against the outcome of the gold standard methods such as Global Rating Scale (GRS), Objective Structured Assessment of Technical Skill (OSATS), and Objective Structured Clinical Examination (OSCE) to evaluate its concurrent validity and construct validity.

Specifically, to improve performance as a part of a comprehensive training and assessment tool, ATTEN-TIVE should provide feedback to the surgeon on independent parameters, allowing the surgeon to refine and improve her/his techniques. For example, by providing feedback that allows a surgeon to adjust to a more

correct hand posture and hand pressure while manipulating instruments and tissue, hand tremor, muscle fatigue, and tissue trauma can be reduced, and consequently a higher overall performance can be achieved [165].

The chapters of this section are organized as follows. Section 5.2 describes the material and methods. Consequently, Chapter 6 presents clustering of the Kinect DK RGB data.The next section, i.e. chapters 7 to 9 includes the classification of surgical tasks based on the data from Kinect DK, leap motion controller, and Myo armband sensors, respectively. Chapters 10 and 11 contains development of metrics for evaluating surgical skill proficiency based on the data from Kinect DK and Myo armband, respectively.

5.2   Materials and methods

Hand motion analysis has been a longstanding research interest among computer vision scientists due to its important and global applications that can be extended to include assessment of surgical proficiency [166]–[170].

In this dissertation, it is hypothesized that ATTENTIVE can be developed by applying machine learning techniques to data acquired from surgeons' lower arms while performing specific surgical tasks. The pipeline for ATTENTIVE is presented in Figure 5.2.
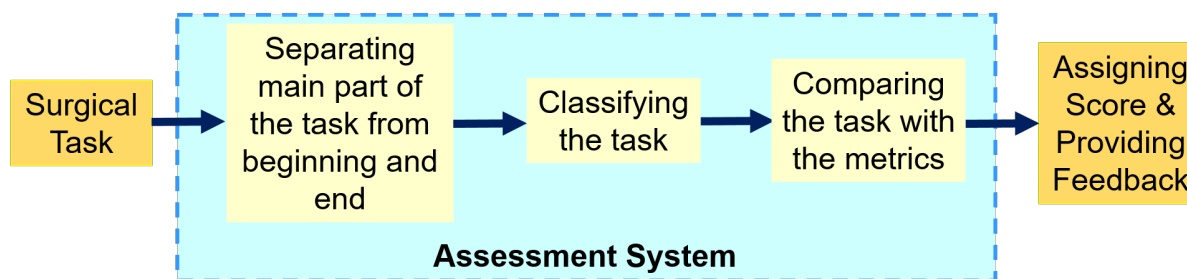
Figure 5.2: Pipeline for ATTENTIVE

Input of the assessment system are biosignal and motion data from the lower arm of a user performing a particular surgical task. The first step of the assessment system is to separate the main part of the task from the beginning and end, i.e. separating the main task from preparing and loading the suture, etc. Then, the second step is to determine the type of the input task. For instance, in this work, four basic surgical tasks were included, and classification algorithms were trained to discriminate between them. Subsequently, the last step is to compare the task against the developed metrics to assign scores and provide feedback accordingly.

5.2.1   Experimental setup

To develop such an assessment system, an experimental setup was designed and implemented to monitor lower arm's motions and muscle activation during completion of basic surgical tasks. The experimental setup, as presented in Figure 5.3, includes a surgical benchtop with standard height to mimic surgical room benchtop, surgical rotating stool with adjustable height, and an apparatus. The apparatus is made of hollow T-Slotted framing, with 4-slot rail and 1.5" × 1.5" cross-section [171] and holds the vision-based sensors as well as the lights. The apparatus is built as an enclosure with only one open side to provide a controlled lighting condition to mimic the lighting of the surgical room.

Both vision-based and sensor-based HGR methods are applied in this study. A Kinect camera, a leap motion controller, a Canon camcorder, and a PTZ Optic camera track the lower arm motions. Each of these cameras provide a different view angle and zooming level to capture different information from the lower arm motion. Sample frames from the Kinect DK and PTZ Optic cameras are presented in Figure 5.3. The participants wear a Myo armband, on the dominant arm, that includes EMG and IMU sensors to provide information on muscle activation and lower arm position and orientation, respectively. Then they wore blue gown and put disposable medical-grade orange and green gloves on their dominant and non-dominant hands, respectively. These colors are designed to facilitate distinguishing the hands in RGB videos of the Kinect DK and other cameras. The participants had the freedom to identify their dominant hands, and also determine if they were ambidextrous in surgery.

The surgical tools and a high-fidelity synthesized suturing pad, with a straight incision line are placed on the surgical desk. The participants were provided by written and video instructions of the surgical tasks that were placed on a side table. Figures 5.4a and 5.4b shows a participant wearing a Myo armband and performs the surgical tasks.
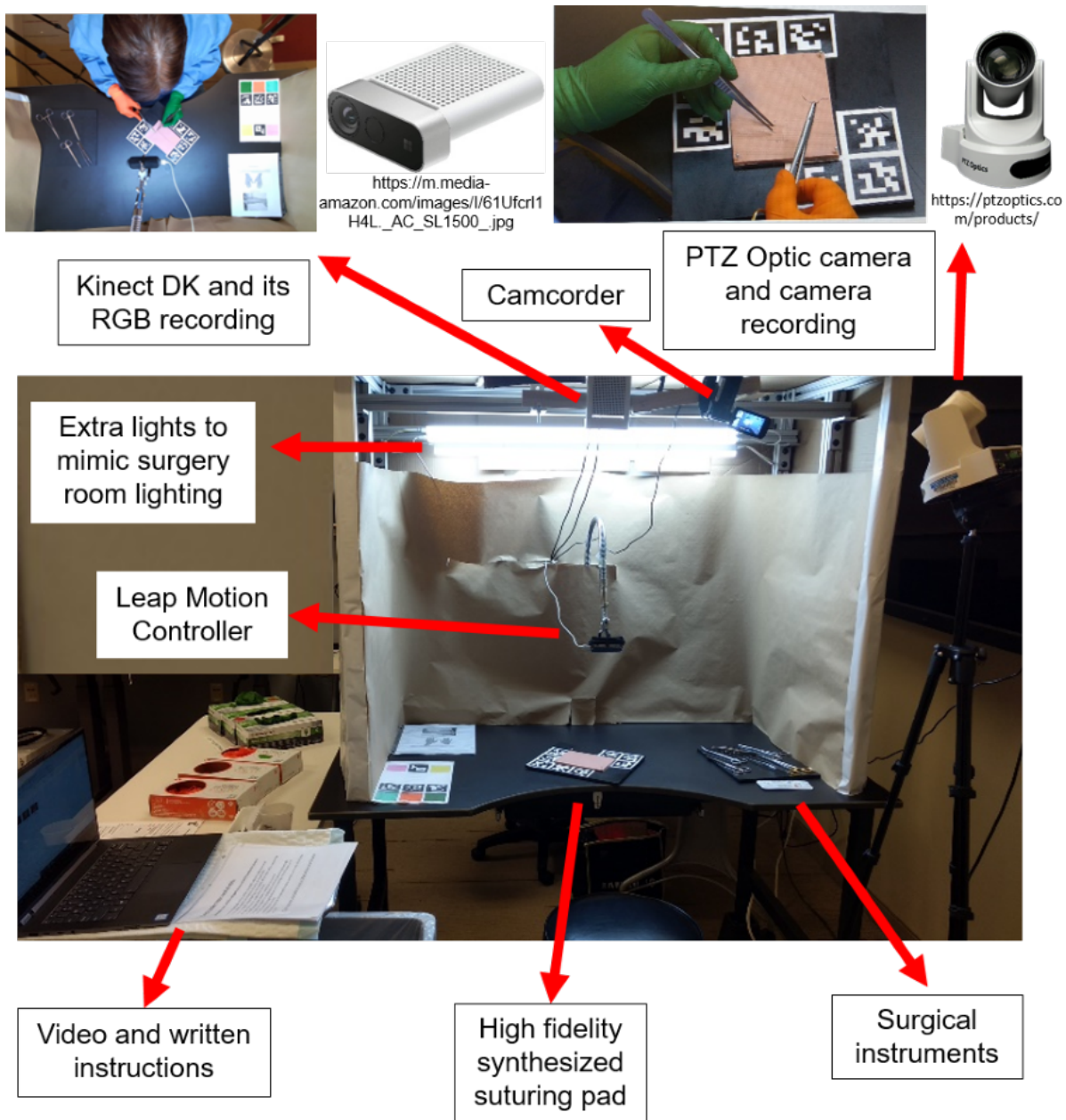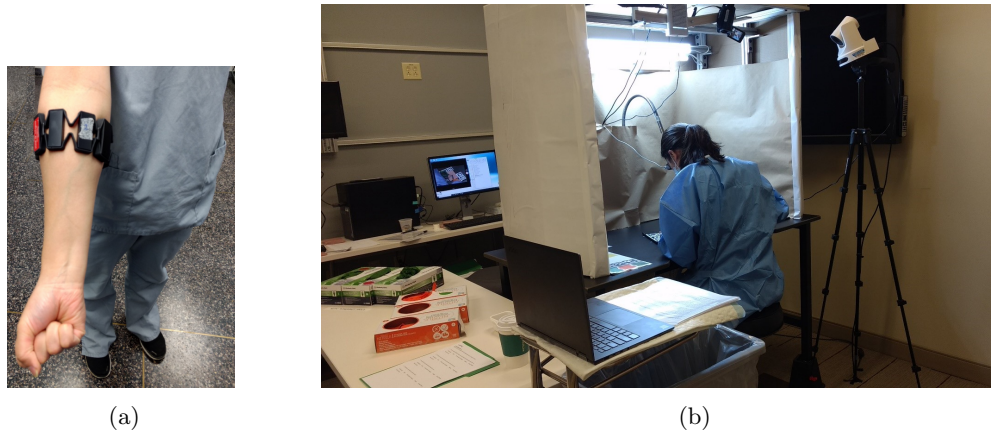
Figure 5.3: Experimental setup

Figure 5.4: A participant **(a)** wears a Myo armband and **(b)** performs the surgical tasks
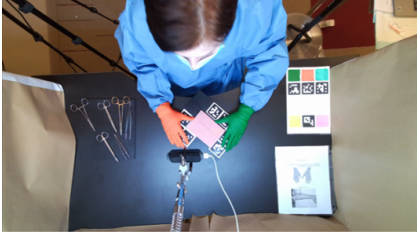
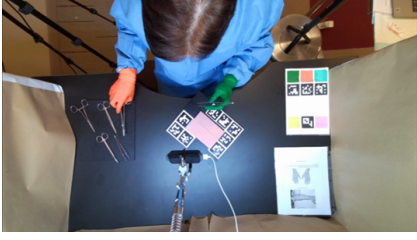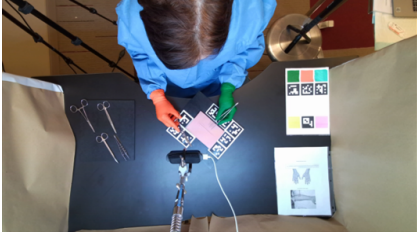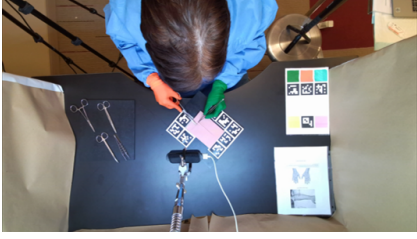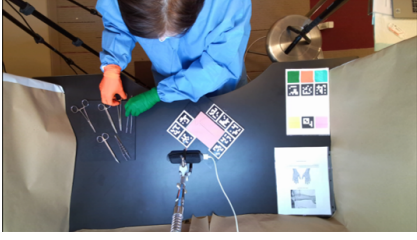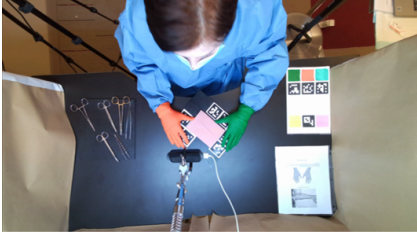## 5.2.2 Participants and Data Acquisition Protocol

The goal of the current study is to find out how the surgical performance can be evaluated through utilizing collected data without human interference. In order to achieve this goal, two sets of data collection, namely "Longitudinal Study" and "Proficiency Level Study", were executed. For both experiments, the participants completed four basic surgical tasks including suturing, knot tying, continuous suturing, and buried suture and knot. The tasks were performed on a high-fidelity synthesized tissue with a straight incision line. These synthesized suturing pads were fabricated by the collaborators from the Small Animal Clinic at the University of Illinois Urbana-Champaign (SAC-UIUC).

The surgical tasks were described through written step-by-step instructions as well as video recordings of an expert performing these tasks. Appendix C includes the written instructions for the surgical tasks as provided to the participants. For each surgical task, the participants could watch the video once, but could refer to the written instructions as often as they needed.

In order to collect baseline data and keep consistency in data acquisition, the participants were instructed to follow a certain protocol as demonstrated in Table 5.1. For each surgical task, after watching the video and reading the written instructions, the participants placed their hands in a neutral position on the surgical bench and announced that they are ready to start the task (Neutral Position in Table 5.1). At this point the Canon camcorder, Azure Kinect DK, and the leap motion controller started collecting data. Then using human voice cue, the participants were instructed to pick up the instruments they intend to use, load the suture, and hold the instruments with the correct technique (Instrument Preparation and Grasp in Table 5.1). Once their hands stabilized while holding the loaded instruments and they announced that they are ready, the Myo armband started collecting data. After 5 seconds of collecting Myo armband baseline data, the participants were instructed via a human voice cue to start the task (Task step in able 5.1)). Once they completed the assigned surgical task, they returned the instruments to their original location (Returning Instrument in Table 5.1) and placed their hands on the surgical desk in the neutral position (second Neutral Position in Table 5.1). at this point, all the sensors stopped data collection.

Table 5.1: Task completion protocol

Neutral Position	
Preparing Instrument	
Grasp	
Task	
Returning Instruments	
Neutral Position	

Animal Clinic at Urbana-Champaign and JUMP Simulation Center at Peoria are provided in Appendices D to F. In addition, the participants filled out questionnaires before and after their participation to evaluate how much the experimental setup, especially the Myo armband, interfere with their performance. The questionnaires completed by the participants in the longitudinal study and the rest of the participants are provided in appendices G and H, respectively.

Longitudinal Study:

The main goal of this study is to propose metrics that can capture the potential performance changes between different weeks in a longitudinal study. A cohort of sixteen second-year veterinary students from SAC-UIUC participated in a longitudinal study, during which data were collected on completion of the first, third, and fifth weeks of a suturing course. In each data collection session, the participants completed the four tasks in the chronological order. Task four was introduced only at the last two sessions.

Proficiency Level Study:

The main goal of this study is to propose metrics that can detect the differences in performance of experts, intermediates, and novice participants. Pre-clinical students, residents, and expert surgeons from two cohorts including veterinary and human medicine professionals participated in this study. There were 28 veterinary participants from SAC, and 21 human-medicine participants from JUMP Simulation center and OSF hospital at Peoria, Illinois, USA. In each data collection session, the participants completed two sets of tasks back-to-back. The tasks of the first set were performed in the chronological order while the task order of the second set was randomized.

Table 5.2 provides more details about the number of completed tasks by the participants of each group.

Table 5.2: Number of completed tasks by each group of the participants

Groups (The labels are based on prior experience level)		Number of Participants	Number of Completed Tasks			
			Task 1	Task 2	Task 3	Task 4
Veterinary Medicine	Experts	10	19	19	19	19
	Longitudinal Studies	16	48	48	48	32
	Novice & Intermediates	18	35	36	35	36
Human Medicine		21	40	40	37	40
Total		65	142	143	139	127

5.2.3   Data Acquisition Devices

Azure Kinect DK

An Azure Kinect DK [172] includes one 1MP (mega pixel) depth sensor and one 12MP RGB video camera. The RGB and depth videos are aligned and can be synchronized in time. The depth camera works based on TOF concept and provides wide and narrow field-of-view (FOV) options. The Kinect DK is equipped with an array of seven microphones and an IMU sensor for tracking orientation of the camera in three-dimensional (3D) space [173]. For this study, "k4arecorder.exe" was used to record the Kinect DK data using "k4arecorder.exe -c 1080p –d NFOV_UNBINNED –depth-delay 0 -r 30 –imu ON File_Name.mkv" command in Windows 10 Command Prompt application. The RGB video with 1080p resolution (1080×1920 pixel progressive-

scan/non-interlaced video mode) was recorded at a sampling rate of 30 frames per second (FPS). The depth data was set to be at a narrow field of view with no delays with respect to the RGB video. Figure 5.5 demonstrates the components of Azure Kinect DK.

1. 1-MP depth sensor with wide and narrow field-of-view
2. 7-microphone array for far-field speech and sound capture
3. 12-MP RGB video camera for an additional color stream that's aligned to the depth stream
4. Ac celerometer and gyroscope (IMU) for sensor orientation and spatial tracking
5. External sync pins to easily synchronize sensor streams from multiple Kinect devices

Figure 5.5: Azure Kinect DK components [173]

Canon R700 Camcorder

A Canon R700 camcorder was used to record the whole procedure. The data from the camcorder can be combined with the RGB data from Kincet DK to serve as a stereo camera to reconstruct the 3D lower arm and track it through the frames. This goal is beyond the scope of teh current dissertation.

Leap Motion Controller

LMC's components and function are described in section 4.2. In this study, LMC's data were collected at a sampling rate of 30 fps. As presented in section 4.2, a Cyton Gamma 300 [75] was used to hold the LMC at an optimum angle to minimize occlusion. However, after collecting data from some of the participants, the Cyton robotic arm crashed, and was replaced with an ordinary webcam holder to keep the LMC at the same position and orientation as much as possible. After collecting data from a few of the participants, it

was observed that while some participants shake their legs vigorously during task completion, the webcam holder and consequently the LMC vibrated, which in turn makes the data noisy. Therefore, a sturdier holder composed of a three-finger-clamp [174] mounted on a 24" heavy-duty flexible gooseneck tube arm [175], a 12" rigid arm [176], and a table clamp [177] replaced the webcam holder to hold the LMC at the same position and orientation as the Cyton robotic arm.

Myo Armband

In the resting position of a muscle, i.e. when the muscle is not contracted, an ionic equilibrium exists between the inner and outer spaces of a muscle cell [178]. This generates surface EMG signals with an amplitude in the microvolts to millivolts range [179]. When the Alpha motor neurons cause contraction in skeletal muscles, an endplate potential, which is a voltage, is generated at the muscle fibers/cells [180]. This endplate potential leads to ion exchange between inner and outer spaces of the muscle cells followed by backward ion exchange to return to the previous ion distribution. Such ion exchanges are known as muscle fibers' depolarization and repolarization, respectively. Motor Unit Action Potential (MUAP) is the combination of action potentials from all the muscle cells that are controlled by a single motor unit [178], [181], [182]. An EMG signal is composed of several MUAP.

EMG can be detected using invasive, by inserting needle/wire electrode into muscle tissue, or non-invasive, by placing electromyography electrodes on the skin surface, methods [178]. The later is known as surface electromyography or sEMG.

The intensity and firing rate of MUAPs affect the EMG signal. Therefore, the EMG signals contain valuable information about the muscle activities, and have wide application in diagnoses of the neuromuscular disorders. Since the occurrence of action potentials are at random intervals, the EMG signal can get positive or negative values at any moment. [182].

There are several noise sources that affect an EMG signal. These include ambient electromagnetic radiation, noises from the electronic equipment, motion artifacts caused by relative movement of skin, electrodes, and underlying muscles, and lastly the cross talk as the EMG sensors collect data from several muscle groups underlying the skin [178], [182].

Detection of hand force and posture using electromyography (EMG) has been studied widely in robotics and in controlling hand prosthetics [66], [183]–[198]. Moreover, researchers have studied EMG data for assessing tremor in afflicted patients and evaluating fatigue in laparoscopic surgeons [65], [199], [200]. EMG signals, have been widely used in hand gesture classification [201], [202] including surgical applications [203].

In this work, a Myo Armband [50] was used to detect muscle activity and motion and rotation of the hand and forearm. An array of eight sEMG, referred to as EMG hereafter, sensors and a 9-axis Inertial Measurement Unit (IMU) comprise the Myo Armband, shown in Figure 1.5. Myo armband is a low-cost, wireless, and easy-to-use device. The output EMG data from the Myo armband is scaled between -128 and 127 [204].

PTZ Optics Camera

A PTZ Optics Camera provided highly focused over-the-shoulder view from the participants hands and the suturing pad. A sample frame of this camera is presented in Figure 5.3. The data of this camera was used by expert surgeons to evaluate participants' performance in completing the tasks.

5.2.4 Metrics Overview

A number of metrics to evalute surgical skills' performance were extracted from OSCE, GRS, and OSAT checklists as well as the literature review. These parameters where categorized in groups including time and motion, respect for tissue, instrument handling, flow of operation and forward planning, and miscellaneous. The parameters of these groups are presented in Tables 5.3 to 5.7.

Table 5.3: Time and Motion metrics for evaluating surgical performance

Parameter	Definition
Necessary/unnecessary moves [205], [206]	How many tries to adjust the needle appropriately. How many time stabbing needle to pass through the suture
Duration [205]	Time for completing each task
Motion density [161]	Presence of hand at each region, e.g. pixel of a recorded video over time
Motion economy [161]	Efficiency of movement, or conservation of energy in any trajectory
Movement rates (speed) [207], [208]	Second derivative of position with respect to time
Jerk [161]	Third derivative of position with respect to time
Path length per cycle [161]	Sum of change in position to complete a cycle
Smoothness [209]	Jerk index and curvature of the path
Total path length [161], [210]	Sum of change in position to complete a task

Table 5.4: Respect for tissue metrics for evaluating surgical performance

Parameter
Unnecessary force on tissue [205], [206]
Damage to tissue [205], [206]
Surgeon's force and tension when manipulating the tissue [161]
Single attempt at passage through benchtop model tissue [206]
Adequate bites taken >2 mm and <8 mm from edge of incision [206]
Minimal damage with forceps [206]

Table 5.5: Instrument handling metrics for evaluating surgical performance

Parameter
Fluid/awkward moves with instrument [205], [206]
Selecting appropriate instruments [206]
Selecting appropriate suture [206]
Correct needle holding technique [206]
Needle driver stabilized with good hand position [206]
Follow-through on curve of needle at entrance on >90% bites [206]
Follow-through on curve of needle at exit on >90% bites [206]
Use of forceps to handle the needle [206]

Table 5.6: Flow of operation and forward planning metrics for evaluating surgical performance

Parameter	Definition
Effortless flow from one move to the next vs. frequent stops [205], [206]	Zero velocity in the consecutive frames more than a threshold, e.g. zero velocity in 4 out of 5 consecutive frames
Fluidity of motion - Periods of idle time [161], [211]	Fluidity of motion is a measure of hesitancy, pauses, changes in direction, and resets
Cycle frequency [161]	How many cycles are completed in a certain period of time

Table 5.7: Miscellaneous metrics for evaluating surgical performance

Parameter
Six throws per knot [206]
Quality of final product [206]
Bites taken perpendicular to incision line [206]
Square knots [206]
Correct tension for knot [206]
Continue suturing along the incision line [206]
Equal bites each side (continuous line suturing) [206]

Table 5.8 contains some new metrics proposed in this dissertation along with the selected metrics from the previous tables to be investigated for evaluation of proficiency in performing basic surgical tasks.

Table 5.8: List of selected metrics

Metric	Notes
Duration	
Motion density	For a certain percentage of time, how many pixels are spanned by the hands? The less the better
	Ratio of hand presence in certain cells
Median speed	
Maximum acceleration	An indication of sudden movements
Jerk	A notion of smoothness
Total path length	Can be calculated by sum(diff(L2norm(x,y,z position) formula and should be minimized
Number of frames in which hand is blocked by head observing from an overhead camera	May not be a very sensitive indicator
Potential difference in muscle activation's magnitude between novice and expert participants [200]	

Part II.A
CLUSTERING BASIC
SURGICAL TASKS

Chapter 6

Clustering RGB Videos of Basic Surgical Tasks

6.1   Introduction

Unsupervised learning is a category of machine learning in which the algorithms do not need to know the data labels. Unsupervised learning algorithms extract information/features from data to detect the underlying pattern of data that can represent the data in a meaningful way, in contrast to pure random noise [212]. Clustering is the unsupervised classification of the underlying patterns that divides the data in different groups based on the similarity of features within each group. It has a numerous application in different fields of science and engineering from medicine to human-machine interaction and wireless sensor networks [213]–[216]. As a result, several algorithms have been developed for clustering data [217], [218]. One common application is clustering the videos [219]–[222].

Soomro et al. [222] proposed an unsupervised learning method to label the action classes and localize them in the videos spatiotemporally using undirected graphs. Murali et al. [223] presented Transition State Clustering with Deep Learning (TSC-DL) to segment surgical tasks using videos along with the kinematic data obtained from the surgical robots. [224]–[226] utilized the combination of the kinematic data from the surgical robot and the RGB videos as well. Primus et al. [227] segmented the endoscopic videos based on the motion pattern changes into (1) no movement, (2) motion caused by camera movement, and (3) movements of the instrument. Twinanda et al. [228] applied clustering methods on RGB-D clips recorded in an operating room for surgical phase, e.g., adjusting bed, moving patient to bed, mixing cement, and so on, recognition. Nara et al. [229] employed video imaging and an ultrasonic location aware system to capture intraoperative movement. Jamal [230] et al. applied clustering methods to analyze the surgical operating room workflow. When the RGB videos are utilized as the input to the clustering algorithms, either when they are used alone or in combination with other data modalities, there are two methods to provide the algorithm's input. In the first method, the clustering algorithm takes in a whole video as the input while in the second approach the video is divided into shorter clips each containing one specific action/cluster [228], [231]–[235].

Kukleva et al. [232] developed an unsupervised model for temporal action determination in untrimmed videos. This model takes advantage of the sequential essence of actions to detect an embedding of features based on their relative temporal position. Then, the clustering is performed on the embedded features and the mean temporal appearance for each cluster is calculated. Afterwards, based on the overall proximity of

each frame to each cluster, the videos are segmented. The proposed method was evaluated on three datasets and outperformed the fully supervised method of RNN+HMM [236] and the weakly supervised method of NN-vit [237].

As presented in Figure 5.2, the first step in ATTENTIVE workflow is to extract the Task section from the entire data. This can be formulated as a clustering problem, in which the data should be divided into pre-Task, surgical-Task, and post-Task clusters (see Table 5.1). In this chapter, the methodology for clustering the Task section of the RGB videos obtained from the Kinect DK camera (see chapter 5.2.3).

6.2 Material and Methods

Since ATTENTIVE should be fully automated, the method used by [232] is selected for clustering the Kinect data. Their presented algorithm works with the untrimmed videos; therefore, the clustering can be performed on the raw RGB videos.

In order to develop a ground truth to evaluate the performance of the clustering algorithm, the Kinect DK video frames should be annotated. Based on Table 5.1, six clusters were defined for each video. As of now, the Kinect videos collected from "Novice and Intermediate" participants are fully annotated for all six clusters. In addition, the first three clusters, i.e. Neutral Position, Instrument Preparation, and Grasp; are annotated in the collected videos from all 65 participants. at last, the fourth cluster, i.e. Task, is annotated for the "Expert", and half of the "Longitudinal Study" participants.

6.3 Conclusion and Future Work

The future work of this chapter includes completing the annotation of the entire dataset and utilizing the clustering model developed by [232] to segment the Task portion from each video. Based on the obtained results, it should be decided weather other algorithms are needed to be used.

6.4 Acknowledgment

Part II.B
CLASSIFYING BASIC
SURGICAL TASKS

Chapter 7

Classification of Basic Surgical Tasks Using Kinect Dk Data

7.1 Introduction

As presented in Figure 5.2, after separating the Task section, the first step of developing ATTENTIVE is classifying different surgical tasks. Classifying the surgical tasks based on the video, and other data modalities, has been studied over the past couple of decades.

Zappella et al. [238] segmented the videos from robotic surgery into short clips each containing a single gesture such as grabbing a needle, passing a needle, and so on. They applied three algorithms on their dataset and achieved over 93% classification accuracy. First, they classified the clips when each video clip was modeled as the output of a linear dynamical system (LDS). In the second method, they applied a bag-of-features (BOF) approach on the extracted spatio-temporal features from the video clips. In the third method, they combined BOF and LDS approaches using multiple kernel learning. The microscopic video frames from twelve surgical phases of cataract surgery were defined as classes in [239]. Dynamic time warping and hidden Markov model achieved 94% classification accuracy. Petscharnig et al. [240] used two CNN architectures, i.e. AlexNet and GoogLeNet, to classify video shots of gynecological surgical actions such as dissection, injection, etc. Zhang et al. [241] leveraged RetinaNet to detect bounding boxes around the hand and track the hands in videos from open surgery with 75% mean average precision. Azari et al. [22] applied decision trees, random forests, and hidden Markov models for classification of surgical maneuvers, i.e. transition, suturing, and knot tying, in 2-second video clips with 79% prediction accuracy. Twinanda [242] extracted visual features including color information, salient point, and image gradients to classify eight different laparoscopic surgeries using Multiple Kernel Learning as the classifier and achieved 91% accuracy. An F1-score of about 68% was achieved by Primus et al. [243] in classifying video frames into eleven phases of the cataract surgery utilizing CNN.

RGB and depth videos have been widely used for hand detection and hand tracking for different applications [17], [168], [244]–[251]. In this chapter, the methodology and results for classifying surgical task by utilizing the RGB-D data obtained from the Azure Kinect DK camera (see chapter 5.2.3) are presented.

7.2 Material and Methods

7.2.1 Subjects and Data Acquisition

The experimental setup and data collection protocol are explained in Chapter 5. In this work, orange and green color gloves were used to detect the dominant and non-dominant hands, respectively. The participants were instructed to put on a blue gown to facilitate a more accurate color extraction from the RGB videos. For this study, so far the data from 'Novice and Intermediate' and 'Expert' groups are analyzed.

7.2.2 Preprocessing

Data Wrangling

The output of the Kinect DK camera were synchronized streams of RGB and depth (D) videos, which were saved in .mkv file format. The first analysis step is to extract RGB and depth frames from the mkv files using a free and open-source software, FFmpeg [252].

Figure 7.1 demonstrates the pipeline for extracting the hand coordinates in three dimensional (3D) space from the Azure Kinect DK data. At each sampling point, the X and Y coordinates are extracted from the RGB frame while the Z coordinate is obtained from the depth frame. The Z axis is perpendicular to the plane of the RGB frame. Since the RGB and depth streams are synchronized, the overall goal of the preprocessing step can be summarized as (1) overlaying the RGB and depth frames (registration), (2) finding the pixels that contain hands from the RGB frame using the color of gloves, and (3) extracting the (X,Y,Z) coordinates of those pixels from the registered RGB and depth frames.
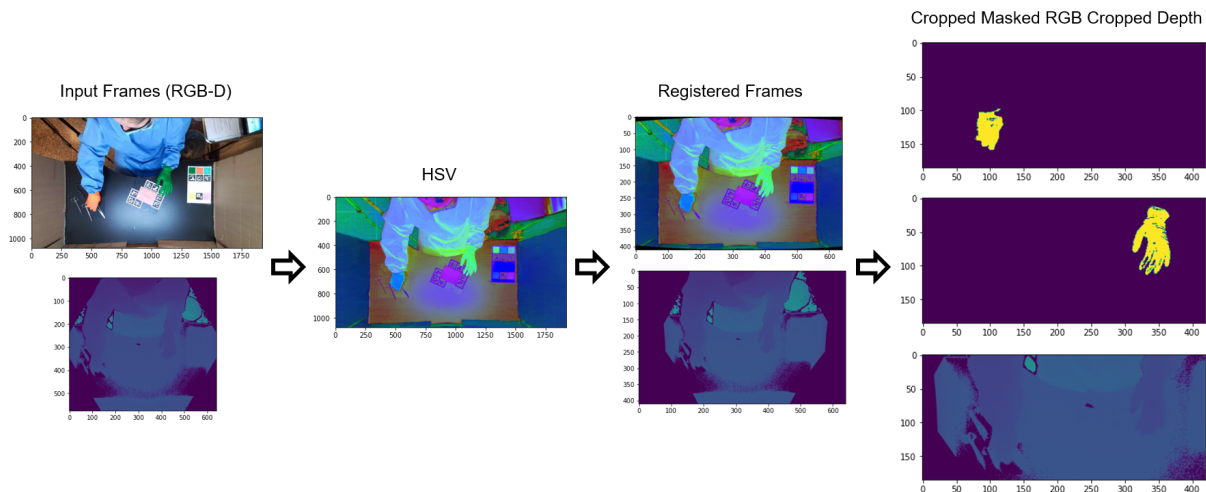
Figure 7.1: Pipeline for extracting hand coordinates in 3D space from Azure Kinect DK data

The first step for registering the RGB and depth frames is to transfer the RGB frames to hue, saturation, value (HSV) color space. As the participants move their hands within the camera view during the task completion, different shades of the orange and green colors for the gloves are recorded in the RGB stream. Although in the RGB color space many of these shades that are present in different frames are recognized as different colors, in HSV color space all of them are detected as the same color, and subsequently are are assigned to the same object that changes its position and orientation between the frames.

The next analysis step is to register the RGB and depth frames. In order to achieve this goal, first both RGB and depth cameras should be calibrated. Camera calibration for the depth and RGB cameras were performed in Matlab R2016a and OpenCV library [253] with Python programming language [254], respectively. Camera calibration, for both cameras, includes finding the corners of squares in a checkerboard plane with known dimensions and deriving the camera intrinsic and extrinsic matrices using the coordinates of these corners in world units and in image units, i.e. pixels.

The corners were found in several frames in which the rigid checkerboard plane had a variety of orientations and positions with respect to the cameras.

The checkerboard pattern for calibrating the depth camera was created in a way that the black squares were hollows, see Figure 7.2, so there will be a contrast in depth at the edges and corners.

Once the camera parameters were obtained, each frame was undistorted using these parameters. Then, the RGB and depth frames were cropped and resized to ensure the salient features in the scene, e.g. table edges and hands, in both frames were mapped to each other as accurately as possible.

Figure 7.2: Hollow checkerboard for calibrating the depth camera of Azure Kinect DK

Afterwards, the pixel coordinates of the hands should be obtained from the RGB frames. For each video, the glove colors, i.e. the (h, s, v) values in HSV color space, are extracted from the first frame and used to generate masks for the dominant and non-dominant hands for all the video frames. As shown in Figure 7.2, there is a sheet of paper on the table that includes Aruco markers [255] and square cuts from the gloves. Using OpenCV in Python, the Aruco markers were detected and the HSV color of the gloves cuts, knowing their distance from the Markers centers, were extracted. The last step in Figure 7.1 demonstrates the generated masks for each hand and the corresponding depth frame.

Once the (X,Y,Z) coordinates (in pixel) for all the points of both hands are extracted at each frame, the average of each coordinate at each frame is computed and used for the rest of the analyses. As a result, each video will be represented by two matrices, one for each hand, of $n \times 3$ in which n is the number of the video frames and three columns are the averages of (X,Y,Z) coordinates at each frame.

Filtering

The beginning and end of the Task section, see Table 5.1, of each video were manually determined and used to extract the Task section of the coordinate matrices, also referred to as trials hereafter. Afterwards, a moving average filter with the window size of 5, i.e. 1/6 of second, was applied on each column of the truncated data matrices.

7.2.3   Features and Classifiers

Features

In this chapter, two types of features including the time domain (TD) and time-frequency domain features are taken into account. The TD features are categorized as geometrical and non geometrical features, and are listed in Table 7.1. TD features, except of the motion density maps (MDM) , were extracted from each column, i.e. each coordinate, of the filtered coordinate matrices. The non geometrical TDs were calculated using a sliding window of size 10 samples, which equals to 1/3 of a second.

   Once the TD features were extracted for all the trials from the four classes, for each feature, the feature matrices were concatenated in one matrix to compute the mean and standard deviation of the entire dataset. These parameters were calculated for each column of the concatenated matrix. Afterwards, the obtained mean and standard deviation values were utilized to standardize each column of the feature matrices, i.e. elementwise subtraction of the mean and division by the standard deviation obtained from their corresponding columns of the concatenated matrix.

   The motion density feature was calculated for each trial. To calculate motion density feature, first each grayscale RGB frame was divided into a 32×32 grid. Then, the number of frames in which the mean of the hand was presented in each block of the grid was detected to generate motion density maps as demonstrated in Figures 7.3a and 7.3b for the dominant and non-dominant hands, respectively. The brighter cells show the regions with more hand presence in them through the trial.

| (a) | (b) |

Figure 7.3: Motion density maps calculated for an Expert participant performing a continues suturing task: **(a)** dominant hand, and **(b)** non-dominant hand

Classifiers

The motion density maps were saved as images and were classified by a CNN algorithm with three two-dimensional convolutional layers. Three convolution layers have 32, 64 and 64 output filters in sequential order, and each convolution layer consists of 3×3 kernels with a stride of 0 and no zero padding. Rectified Linear Unit (ReLU) activation function and maximum pooling function were applied at the end of each convolution layer. The maximum pooling function was applied on 2×2 window after the first two layers and 4×4 window after the last convolution layer. Fifty % dropout was implemented at the end of the last two layers. A fully connected layer with Softmax activation function was implemented as the very last layer of the CNN model. The categorical cross entropy as the loss function and an Adam optimizer, with adaptive learning rate, were used for training the CNN. The number of epochs and batch size were set to 100 and 16, respectively. The CNN model was implemented with Keras library [256] with Tensorflow [257]. Five-fold cross validation was applied to evaluate the performance of the proposed CNN model in classifying the motion density maps into four classes, i.e. four surgical tasks. Precision, recall, and F1-score were calculated using sklearn.metrics.precision_recall_fscore_support function, by setting average = 'macro', which calculates these metrics for each class and reports their average value.

The TD features are treated as time series as their changes over the time defines the characteristics of each surgical task. Depending on how long it took for a participant to complete each surgical task, each trial has a different length. In order to prepare data to be fed into the classification algorithm, the trials need to have the same dimensions. Therefore, the length of the longest trial among all the classes, i.e. the four surgical tasks, was set as the reference length for each TD feature. Then, the rest of the feature matrices were made the same length by repeating each feature matrix at the end of itself as many time as needed to make it the same length as the reference length.

A combination of CNN and Long short-term memory (LSTM) was implemented in Keras library with Tensorflow. The model was composed of two identical one-dimensional convolutional layers followed by three identical LSTM layers. The convolution layers had 50 output filters and kernel of size 6 with a stride of 1 and padding="same" to ensure the output size is the same as the input size. Rectified Linear Unit (ReLU) activation function and maximum pooling function, with pool size of 30, followed by batch normalization and 20% dropout were applied at the end of each convolution layer. The output dimension of the LSTM layer, which is the same as the number of the hidden nodes, was set to 50 for all three LSTM layers. A Dropout of 20% was implemented after the first two LSTM layers. A fully connected layer with Softmax activation function was implemented as the very last layer of the model. The categorical Cross entropy as the loss function and an Adam optimizer, with adaptive learning rate, were used for training the model. The number of epochs and batch size were set to 40 and 3, respectively. Five-fold cross validation was applied to evaluate the performance of the proposed model in classifying the TD features into four classes. Precision, recall, and F1-score were calculated using sklearn.metrics.precision_recall_fscore_support function, by setting average = 'macro'.

7.3   Results and Discussion

The analysis results are presented in Tables 7.1 to 7.4 for dominant and non-dominant hands of 'Expert' and 'Novice and Intermediate' groups, respectively. The classification accuracies of both hands of both groups are aggregated in Table 7.5. The majority of features could provide over 85%, some over 90%, classification accuracy for both hands in both groups.

The best accuracies are achieved using MAV and RMS features, which provide an average sense of the position in a sliding window over the entire task completion time. In all cases, variance was the least efficient feature for classifying the tasks. Variance is a notion of the deviation of sampling points from their average. In this work, the variance was calculated over a sliding window on the position along X, Y, and Z axes. Although the change of position was a promising feature to distinguish different surgical tasks, since the tasks are performed within a confined area in the three-dimensional space, the averaged squared deviations from the mean are not sufficiently large and variant over the time to discriminate the tasks. Therefore, although the achieved accuracy using the variance feature is comparable with some studies from the literature, variance performed poorly for task classification compared to the features employed in the current study.

Better classification results can be obtained using X and Y position compared to the Z position. This can be associated with the larger changes in the hand position in (X,Y) plane compared to the Z direction. Nevertheless, it is interesting to observe that the dominant hand's Z coordinate of the 'Novice and Intermediate' group provides a significantly higher accuracy, compared to the 'Expert' group. This may suggest that the non-Expert participants exert redundant hand motion along the Z axis, and therefore, the Z coordinate of the hand position can be used as a distinctive metric to evaluate the performance of non-Experts in completing the surgical tasks. In all cases, i.e. for both hands and both groups, the L2 norm of the (X,Y,Z) coordinates, which provides information about the 3D motion, outperformed the individual coordinates.

In all cases, speed was a better feature for classification compared to its derivatives with respect to time, i.e. acceleration and jerk. This observation can be correlated to the fairly slow and smooth hand motions while performing surgical tasks. All of these three parameters as well as the position along X-axis and Y-axis achieved higher classification rates for the 'Expert' group, which may suggest a potentially more task specific hand motion and hand control in the 'Expert' group.

In comparing the two groups (see Figure 7.5), 6 out of 11 features provided a higher classification performance in the 'Expert' group and the rest performed better for the 'Novice and Intermediate' participants. In addition, the difference of each feature's accuracies between the two groups is less than 5% except for the Z-axis position, which was discussed earlier in this chapter. In the 'Expert' group, higher accuracies for different features were obtained equally from both hands whereas the dominant hand achieved better results for most features in the 'Novice and Intermediate' group. These observations suggest that the 'Expert' group used their non-dominant hand in a more task specific way that can result in a higher efficiency in completing the tasks.

Table 7.1: Performance metrics for time domain features obtained from the dominant hand of the 'Expert' group participants using 5-fold cross validation. All the numbers are percentage values (%). Different sets of features are determined by different colors.

Feature		Expert - Dominant Hand			
		Accuracy	Precision	Recall	F1-Score
Geometrical Features in Time Domain	Position (X-axis)	88	88.5	88.2	88.1
	Position (Y-axis)	84	84.7	84.2	83.3
	Position (Z-axis)	64	66.7	64.3	63.9
	L2norm of Position Coordinates	90.7	90.8	90.8	90.7
	Speed	92	92.1	92.1	92.1
	Acceleration	89.3	89.5	89.5	89.4
	Jerk	89.3	89.7	89.5	89.3
	Motion Density Map	84	85.3	84.2	83.5
Non Geometrical Features in Time Domain	MAV	96	96.3	96	96
	RMS	93.3	93.7	93.3	93.4
	VAR	76	76.8	76.2	75.1
	WL	78.7	80.8	78.7	78.7
Time-Frequency Domain	STFT	89.3	91.2	89.5	89

Table 7.2: Performance metrics for time domain features obtained from the non-dominant hand of the 'Expert' group participants using 5-fold cross validation. All the numbers are percentage values (%). Different sets of features are determined by different colors.

Feature		Expert - Non-Dominant Hand			
		Accuracy	Precision	Recall	F1-Score
Geometrical Features in Time Domain	Position (X-axis)	92	92.5	92	92
	Position (Y-axis)	88	89.3	88.1	87.8
	Position (Z-axis)	65.3	69.6	65.8	64.7
	L2norm of Position Coordinates	92	92	92.1	92
	Speed	90.7	91.8	90.8	90.9
	Acceleration	89.3	89.5	89.5	89.5
	Jerk	85.3	85.3	85.4	85.1
	Motion Density Map	81.3	82.9	81.6	81
Non Geometrical Features in Time Domain	MAV	97.3	97.5	97.4	97.3
	RMS	98.7	98.8	98.7	98.7
	VAR	70.1	72.2	71	69.7
	WL	86.7	87	87	86.6
Time-Frequency Domain	STFT	89.3	90.7	90.7	90.7

Table 7.3: Performance metrics for time domain features obtained from the dominant hand of the 'Novice and Intermediate' group participants using 5-fold cross validation. All the numbers are percentage values (%). Different sets of features are determined by different colors.

Feature		Novice and Intermediate - Dominant Hand			
		Accuracy	Precision	Recall	F1-Score
Geometrical Features in Time Domain	Position (X-axis)	89.4	89.4	89.5	89.4
	Position (Y-axis)	87.3	87.3	87.4	87.2
	Position (Z-axis)	75.3	76.1	75.5	75.3
	L2norm of Position Coordinates	92.9	93.1	93.1	92.8
	Speed	90.9	90.9	90.9	90.9
	Acceleration	87.4	87.3	87.4	87.3
	Jerk	88.7	89.3	88.8	88.9
	Motion Density Map	87.4	87.2	87.4	87.1
Non Geometrical Features in Time Domain	MAV	95.1	95.3	95.1	95.1
	RMS	93.8	94	93.7	93.6
	VAR	77.6	77.8	77.5	77.3
	WL	90.2	90.8	90.2	90.3
Time-Frequency Domain	STFT	86.6	86.7	86.7	86.6

Table 7.4: Performance metrics for time domain features obtained from the non-dominant hand of the "Novice and Intermediate" group participants using 5-fold cross validation. All the numbers are percentage values (%). Different sets of features are determined by different colors.

Feature		Novice & Intermediate - Non-Dominant Hand			
		Accuracy	Precision	Recall	F1-Score
Geometrical Features in Time Domain	Position (X-axis)	86.6	86.6	86.7	86.6
	Position (Y-axis)	85.9	87.4	85.9	85.8
	Position (Z-axis)	69.0	69.6	69.2	68.3
	L2norm of Position Coordinates	89.5	89.6	89.5	89.4
	Speed	88.1	88.1	88.1	88.1
	Acceleration	87.3	87.7	87.3	87.4
	Jerk	81.6	82.9	81.8	82.0
	Motion Density Map	88.1	88.7	88	88.2
Non Geometrical Features in Time Domain	MAV	95.8	95.8	95.8	95.8
	RMS	95.8	95.9	95.8	95.8
	VAR	74.6	74.4	74.8	74.3
	WL	89.4	89.8	89.5	89.6
Time-Frequency Domain	STFT	92.3	92.4	92.3	92.2

Table 7.5: Aggregated classification accuracies. All the numbers are percentage values (%) and obtained by applying 5-fold cross validation. Different sets of features are determined by different colors.

Group:	Novice and Intermediate		Expert	
Hand:	Dominant	Non-Dominant	Dominant	Non-Dominant
Position (X-axis)	89.4	86.6	88	**92**
Position (Y-axis)	87.3	85.9	84	**88**
Position (Z-axis)	**75.3**	69	64	65.3
L2norm of (X,Y,Z) coordinates	**92.9**	89.5	90.7	92
Speed	90.9	88.1	**92**	90.7
Acceleration	87.4	87.3	**89.3**	**89.3**
Jerk	88.7	81.6	**89.3**	85.3
Motion Density Map	87.4	**88.1**	84	81.3
MAV	95.1	95.8	96	**97.3**
RMS	93.8	95.8	93.3	**98.7**
VAR	**77.6**	74.6	76	70.1
WL	**90.2**	89.4	78.7	86.7
STFT	86.6	**92.3**	89.3	89.3

7.4 Conclusion and Future Work

In this work, by utilizing a Kinect DK camera, the position of dominant and non-dominant hands was tracked during the completion of the basic surgical tasks. Colored gloves were used to increase the accuracy in separating the hands in RGB frames. Then the depth and RGB frames were registered to extract the 3D position of the hands through the videos. Knowing the 3D position, several time domain features were extracted and used with LSTM and CNN to classify four basic surgical tasks. Over 98% and 92% task classification accuracies were obtained for the 'Expert' and 'Novice and Intermediate' groups, respectively. The future work investigates four directions. First the frequency and time-frequency domain features such as discrete wavelet transform and STFT as well as the combination of different features should be tested for task classification. In addition, other classification algorithms can be implemented to enhance real-time classification. Moreover, in the current work, the average of the hand pixels in each frame were calculated and utilized for classification. Instead, in each frame other features such as the hand borders or the point cloud can be used for task classification to provide a more complete picture. At last, the data from other participants can be included in the analysis.

Chapter 8

Classification of Basic Surgical Tasks Using Leap Motion Controller Data

8.1 Introduction

As presented in Figure 5.2, after separating the Task section, the first step of developing ATTENTIVE is classifying different surgical tasks. In this chapter, the methodology and results for classifying data obtained from the Leap Motion Controller (see chapter 5.2.3) is presented.

8.2 Material and Methods

The experimental setup and data acquisition method are explained in Chapters 4 and 5 in details. In this chapter the analysis workflow and results for all the participants are presented.

8.2.1 Preprocessing

The data collected from each participant completing each surgical task each time is saved in three matrices for right hand, left hand, and both hands. The X, Y, and Z coordinates of the hand joints and palm center, presented in Figure 4.1, as well as the basis vectors of the hand coordinate system, shown in Figure 4.2, are saved in the data matrix columns for each frame, i.e. each row of the data matrices. In some frames, LMC fails to detect both hands correctly, and it misses one or both. Therefore, the lengths of the data matrices for different hands are not equal.

As presented in section 4.2, a Cyton Gamma 300 [75] was used to hold the LMC at an optimum angle to minimize occlusion. However, after collecting data from some of the participants, the Cyton robotic arm crashed, and was replaced with an ordinary webcam holder to keep the LMC at the same position and orientation as much as possible. After collecting data from a few of the participants, it was observed that while some participants shake their legs vigorously during task completion, the webcam holder and consequently the LMC vibrated, which in turn makes the data noisy. Therefore, a sturdier holder composed of a three-finger-clamp [174] mounted on a 24" heavy-duty flexible gooseneck tube arm [175], a 12" rigid arm [176], and a table clamp [177] replaced the webcam holder to hold the LMC at the same position and orientation as the Cyton robotic arm.

To compensate for the discrepancy in the data collection setup, the data was transformed to a global coordinates system. In particular, first, all the data points were translated to the LMC center of coordinate system. Then, a rotation from LMC coordinate system to a global coordinate system was applied based on the formula 8.1. A is the change-of-basis matrix, or transition matrix, and its columns are the coordinate vectors of the basis vectors of hand coordinate system in the LMC coordinate system at each frame. Since at the beginning of all the tasks, the hands are in the neutral position, the global coordinate system is defined as the hand coordinate system in the first 15 frames, i.e. the first 0.5 Seconds of each task. The LMC center of coordinate system for translation and the matrix A are obtained from the average of the first 15 frames and applied to the entire task.

$$XGCS = inverse(A) \times XLMCS \tag{8.1}$$

At the next step, the transformed data is filtered using a median filter on a window size of 5 sampling points, i.e. 1/6 Second.

As presented in Table 5.1 the participants were instructed to follow a certain protocol to complete the surgical tasks. In this work, in an attempt to remove the common steps among the four surgical tasks, the "Neutral Position" was removed from the beginning of the LMC data matrices. The end point of the Neutral Position section was defined based on changes in the acceleration along the Z-axis, when the acceleration magnitude was larger than a set threshold, which was determined heuristically. The Neutral Position section of the data was removed and the rest of it was used for further processing to classify the surgical tasks.

8.2.2 Features and Classifiers

Two sets of features, including time domain and time-frequency domain features, were extracted from the LMC data.

Time domain features:

The time domain features listed in Table 4.1 were extracted from the four classes. Next, for each feature, the feature matrices were concatenated in one matrix to compute the mean and standard deviation of the entire dataset. These parameters were calculated for each column of the concatenated matrix. Subsequently, the obtained mean and standard deviation values were utilized to standardize each column of the feature matrices, i.e. elementwise subtraction of the mean and division by the standard deviation obtained from their corresponding columns of the concatenated matrix.

Time-frequency domain features:

Three features were extracted from discrete wavelet transform coefficients [258], [259] including:

1. The sum of wavelet coefficients

2. The sum of squared wavelet coefficients

3. The energy value of each signal that is defined as the mean of squares of the wavelet coefficients.

The discrete wavelet transform was implemented in Python using pywt package [260]. Daubechies 7 was picked as the mother wavelet [261]. Sharp changes in the signal were highlighted by Discrete wavelet coefficients using three levels of decomposition [262].

Classifiers

The features were treated as time series as their changes over time defines the characteristics of each surgical task. Depending on how long it took for a participant to complete a surgical task, each trial has a different length. In order to prepare data to be fed into the classification algorithms, the sampling points, i.e. trials, need to have the same dimensions. In this work, the length of the longest trial among all the classes was selected as the reference length. Then, the rest of the trials were made the same length by repeating each trial at the end of itself as many time as needed to make it the same length as the reference length.

A combination of CNN and Long short-term memory (LSTM) was implemented in Keras library with Tensorflow. The model was composed of two identical one-dimensional convolutional layers followed by three identical LSTM layers. Several combinations of hyperparameters were tested to achieve the highest classification performance.

For all the features, the convolution layers had kernel of size 6 with a stride of 1 and padding="same" to ensure the output size is the same as the input size. Rectified Linear Unit (ReLU) activation function and maximum pooling function, with pool size of 30, followed by batch normalization and 20% dropout were applied at the end of each convolution layer. A Dropout of 20% was also implemented after the first two LSTM layers. A fully connected layer with Softmax activation function was implemented as the very last layer of the model. The categorical Cross entropy as the loss function and an Adam optimizer, with adaptive learning rate, were used for training the model. Five-fold cross validation was applied to evaluate the performance of the proposed model in classifying the features into four classes. Precision, recall, and F1-score were calculated using sklearn.metrics.precision_recall_fscore_support function, by setting average = 'macro'.

8.3   Results and Discussion

Five-fold cross validation classification accuracies obtained by applying the CNN-LSTM model using different combinations of the hyper parameters are presented in Table 8.1. These results were achieved by using MAV and DPUV as the input features. Regardless of the choice of the hyperparameters, the model is not capable to achieve an accuracy over 50%. In Chapter 4, it was observed that MAV and DPUV are very suitable features to represent the LMC data. Similarly, the CNN-LSTM model showed high performance in classifying surgical tasks when applied on the Kinect DK and EMG data as presented in Chapters 7 and 9, respectively. From these observations, it can be deduced that the data preparation procedure that was applied on the raw LMC data in this study requires improvement. For instance, although the "Neutral Position" step was removed from the beginning of each data matrix, there is still a huge overlap between the classes due to other steps from Table 5.1 that are common between the classes. In other words, several steps such as "Instrument Preparation", "Grasp", etc still present in all the classes that in turn result in huge inter class similarity.

Table 8.1: Five-fold cross validation classification accuracies obtained by applying the CNN-LSTM model using different combinations of the hyper parameters

Batch Size	# of Epochs	CNN Output Filters	LSTM Output Filters	Classification Accuracy (%)
10	40	50	50	46
10	40	100	100	31
10	40	100	50	48
5	40	200	50	43
3	40	200	200	39
10	40	200	200	45
20	40	20	20	43
20	100	20	20	50
20	40	50	50	38
20	100	50	50	31

8.4 Conclusion and Future Work

Despite the high performance of the MAV and DPUV features in classifying the LMC data in a previous study, i.e. Chapter 4, the obtained classification rates in the current study were below 50%. The CNM-LSTM model used for classifying the surgical tasks demonstrated high performance in classifying the surgical tasks classes based on the Kinect DK and the Myo armband data. As a result, the low classification performance may be correlated to the low inter class variability due to the actions that are common between the tasks. Therefore, the first step for the future work is to apply clustering algorithms to segment the "Task" section from the other actions in the video. In addition, the classification should be performed on each participant group separately, particularly the data from participants with human and veterinary medicine background should be tested separately to detect the potential intraclass variabilities. Moreover, the randomized grid search cross validation method should be applied for selecting the hyperparameters to achieve the highest performance. Furthermore, other features in time and frequency domains should be utilized for classification. Finally, other classification algorithms can be implemented to enhance real-time classification.

Chapter 9

Classification of Basic Surgical Tasks Based on EMG-IMU Sensor Fusion

9.1 Introduction

As presented in Figure 5.2 after separating the Task section, the first step of developing ATTENTIVE is classifying different surgical tasks. In this chapter, the methodology and results for classifying data obtained from the Myo armband is presented. The details about the Myo armband and its application i classifying hand gestures are presented in Chapters 5.2.3 and 2 in details. [1]

9.2 Material and Methods

9.2.1 Subjects and Data Acquisition

To record sEMG signals, a Myo armband [50] was used in this study. Myo armband is a low-cost consumer-grade armband that contains an array of eight dry-electrode surface EMG sensors with the sampling rate of 200 Hz and an Inertial Measurement Unit (IMU) sensor with 50 Hz sampling rate. It is an easy-to-use armband as a user can simply slip the bracelet on and no preparation is required [55]. Figure 5.4a shows the armband placement on a participant's forearm during data collection. To maintain consistency between participants in collecting EMG data, the Myo armband was worn on the participant's dominant arm in a way that a particular sensor was always placed on the flexor carpi radialis muscle, as presented in [263], and the rest of the sensors are spaced out evenly around the lower arm.

In this work, the data analysis workflow obtained from the participants in the "Longitudinal Study" group is presented.

---

[1]Partial findings of this research was presented at: Sharif H., Phillips H., Kuzminsky J. Seo S. B., McNeil L., Kesavadas T., "Automated Surgical Tasks Classification Using Surface Electromyography", presented at American College of Surgeons meeting on March 2, 2022

9.2.2 Preprocessing

Data Wrangling

After rectifying the EMG signals, EMG and IMU signals were synchronized by down sampling EMG data from 200 Hz to 50 Hz. Synchronizing EMG and IMU data facilitates separation of particular parts of the EMG signal based on the IMU data. In this work, the signal was separated into two sections. The first section, referred to as tool-holding or Grasp section, is when a participant holds the surgical tools with correct technique after loading the suture (See Table 5.1). During the data collection session, the participants were instructed to hold the instruments with correct technique for 5 seconds and start the task when they were verbally instructed. The second part of the EMG signal is when the participants performed the surgical tasks, referred to as the Task section.

To find the Grasp section of the signal and separating it from the Task section, a moving window of the size of 25 samples, i.e. 0.5 Seconds, was applied. The start point of the Task section was defined based on changes in the acceleration along y-axis, when the average acceleration magnitude was larger than a set threshold. This threshold worked for the entire longitudinal study dataset except of one trial, for which the separation point was detected manually. The access and threshold were determined heuristically. The Task section of the signals were used for further processing to classify the surgical tasks.

The EMG signals needed to be normalized to compensate for variations in muscle activation between the participants and for each participant in different data collection days. To achieve this goal, first the mean of signal at the very last second was calculated. In this last second the participants were instructed to place their hands on the surgical table in a neutral position. The EMG signal from each sensor was normalized by dividing the signal to the signal's mean during the last second obtained from the same sensor.

Subsequently, a first order Butterworth low-pass filter with a 1 Hz cutoff frequency [71] was applied on each channel of EMG data. The eight EMG signals are concatenated in a data matrix with eight columns, one column for each EMG sensor, and n rows, for n sampling points.

9.2.3 Features and Classifiers

Features

In this chapter, four time domain (TD) features are taken into account [264]. These features are listed in Table 9.3. TD features were extracted from each column of the filtered EMG matrices. They were calculated using a sliding window of size 8 samples, which equals to 0.16 S for downsampled EMG signal to 50 HZ.

Once the TD features were extracted for all the trials from the four classes, for each feature, the feature matrices were concatenated in one matrix with eight columns to compute the mean and standard deviation of the entire dataset. These parameters were calculated for each column of the concatenated matrix. Afterwards, the obtained mean and standard deviation values were utilized to standardize each column of the feature matrices, i.e. elementwise subtraction of the mean and division by the standard deviation obtained from their corresponding columns of the concatenated matrix.

Classifiers

The TD features were treated as time series as their changes over the time defines the characteristics of each surgical task. Depending on how long it took for a participant to complete each surgical task, each trial has a different length. In order to prepare data to be fed into the classification algorithm, the trials need to have

the same dimensions. Therefore, the length of the longest trial among all the classes, i.e. the four surgical tasks, was set as the reference length for each feature. In this work, two techniques, i.e. zero-padding and repetition, were implemented to make the trials equalled size. To make the lengths of the rest of the trials equal to the reference length, zeros were added at the end of the trials in zero-padding technique whereas in the repetition method, the feature matrices were made the same length by repeating each feature matrix at the end of itself as many time as needed.

A combination of CNN and Long short-term memory (LSTM) was implemented in Keras library with Tensorflow. The model was composed of two identical one-dimensional convolutional layers followed by three identical LSTM layers. The hyperparameters of the model were customized for each feature to enhance achieving the highest classification performances. These parameters are presented in 9.1.

For all the features, the convolution layers had kernel of size 6 with a stride of 1 and padding="same" to ensure the output size is the same as the input size. Rectified Linear Unit (ReLU) activation function and maximum pooling function, with pool size of 30, followed by batch normalization and 20% dropout were applied at the end of each convolution layer. A Dropout of 20% was also implemented after the first two LSTM layers. A fully connected layer with Softmax activation function was implemented as the very last layer of the model. The categorical Cross entropy as the loss function and an Adam optimizer, with adaptive learning rate, were used for training the model. Five-fold cross validation was applied to evaluate the performance of the proposed model in classifying the TD features into four classes. Precision, recall, and F1-score were calculated using sklearn.metrics.precision_recall_fscore_support function, by setting average = 'macro'.

Table 9.1: Hyper parameters for the CNN-LSTM model

	Batch Size	# of Epochs	CNN Output Filters	LSTM Output Filters
MAV	3	80	20	20
RMS, VAR, WL	3	100	40	40

9.3 Results and Discussion

The analysis results are presented in Tables 9.2 and 9.3 for the repetition and zero-padding methods, respectively. For all the features, the zero-padding method outperformed the repetition technique. When applying the zero-padding method, all the features achieved over 88% accuracy with F1-scores over 87%. In addition, the difference between the lowest and highest accuracy values for different features is less than 2%, which suggests that all the features are equally efficient in classifying the surgical tasks. While the repetition method is utilized, the MAV feature outperforms other features.

Although the zero-padding method achieved higher accuracies, this achievement can be interpreted as the algorithm learns to classify based on the number of padded zeros that represents the task length rather than using other features purely. However, since different features provided different accuracies despite their similar lengths, it can be concluded that the algorithm learned based on a combination of each feature and the task completion time.

Table 9.2: Performance metrics for time domain features using 5-fold cross validation. All the numbers are percentage values (%).

Feature	CNN-LSTM – Repetition			
	Accuracy	Precision	Recall	F1-Score
MAV	82.4	81.2	80.5	80.6
RMS	69.5	69.9	68.5	68.3
VAR	77.3	76.1	75.8	75.6
WL	79	77.9	77.6	77.7

Table 9.3: Performance metrics for time domain features using 5-fold cross validation. All the numbers are percentage values (%).

Feature	CNN-LSTM – zero-padding			
	Accuracy	Precision	Recall	F1-Score
MAV	89.2	90.3	88	88.7
RMS	89.2	90.5	88.5	89.2
VAR	88.6	90.5	86.5	87.4
WL	90.3	90.8	89.3	89.9

The confusion matrices obtained by applying the zero paddling technique are demonstrated in Figures 9.1a to 9.1d for all the TD features. Uniform distribution of the off-diagonal elements in the confusion matrices confirm that the model has not been overfitted to any of the classes. The majority of misclassification cases are when task 4 was falsely classified as task 2. This correlation is in-line with the similarity in task completion time between these two tasks. In addition, since the number of trials from task 4 is less than the number of trials of task 2, the model might learn the characteristics of task 2 better than the characteristics of task 4, and as a result detects task 4 as task 2 and not the other way around.
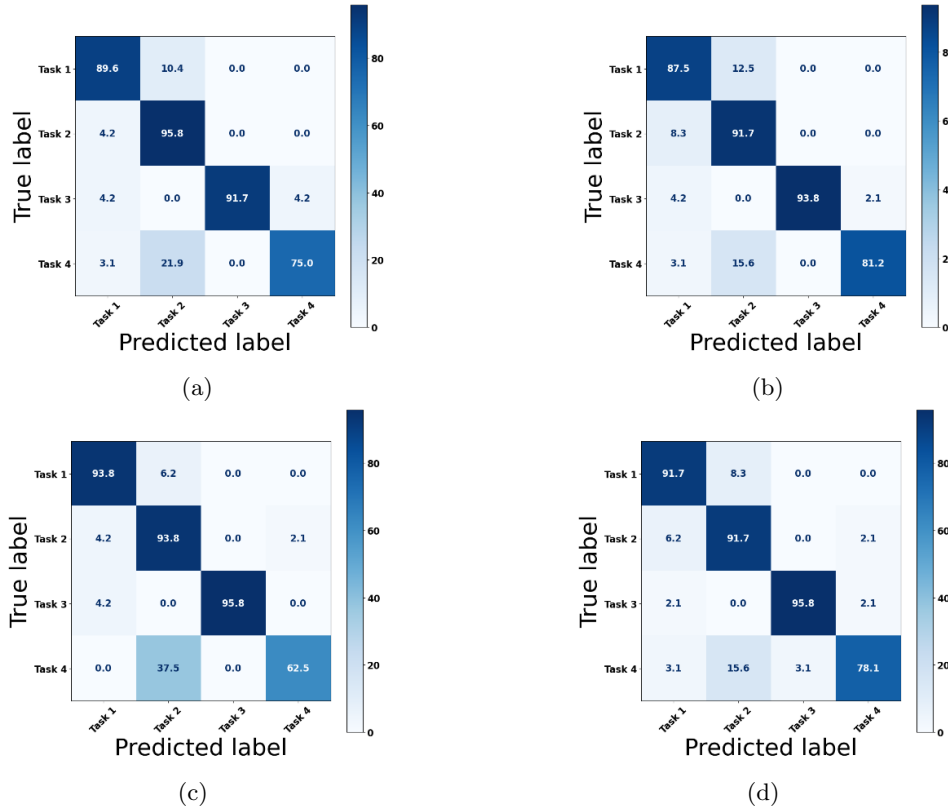
Figure 9.1: Confusion matrices for the zero-padding method using 5-fold cross validation. All the numbers are percentage values (%). Features include **(a)** MAV, **(b)** RMS, **(c)** VAR, and **(d)** WL.

9.4 Conclusion and Future Work

Four time domain features and two data wrangling methods were applied in this work to prepare the EMG data collected from the lower arm of the "Longitudinal Study" participants. The zero-padding method outperformed the repetition technique, which can be correlated to the number of zeros at the end of each data matrices. The number of padded zeros is a notion of the task completion time. Therefore, it can be deduced that the zero-padding technique applied a combination of task completion time and each time domain feature. Classification using just the task completion time, i.e. without the features, can provide more insight on this topic.

As the next step of this work, the frequency and time-frequency domain features such as discrete wavelet transform and STFT as well as the combination of different features should be tested for task classification. In addition, other classification algorithms can be implemented to enhance real-time classification. Moreover, the data from other participants can be included in the analysis. At last, currently the surgical task completion part was isolated from the rest of each trial using the parameters that were determined heuristically. In order for ATTENTIVE to be fully automated, the task separation needs to be automated too. One approach toward achieving this goal is utilizing clustering algorithms to segment the task section part from each EMG data matrix.

Part II.C
EVALUATING
SURGICAL PROFICIENCY

Chapter 10

Quantitative Evaluation of Surgical Skills Proficiency Using Kinect DK Data

10.1 Introduction

Due to the importance of developing a quantitative assessment system, many researchers have attempted to propose metrics and methodologies to achieve this goal. Tables 5.3 to 5.8 provide a comprehensive list of metrics. Watson et al. [265] used IMU data to classify the expert and novice participants performing venous anastomoses on a benchtop model. crowdsourcing and deep neural networks were applied on video data to track instruments and divide the skill level into low level and high level while performing invasive minimally surgery [266]. Aggarwal et al. [210] utilized ROVIMAS video-based motion tracking device to evaluate dexterity of two groups of experienced and inexperienced surgeons in performing laparoscopic surgery. The surgical tool's motion data along with the location of surgeon's eye gaze on the screen were utilized to distinguish expert from novice participants performing a cadaveric eye laparoscopic surgery [267]. Barnhill et al. developed a standard protocol for editing videos of surgical performances to obtain a reproducible assessment of surgical proficiency [268]. Tanin [269] developed and experimentally evaluated a deep learning model to assess skill level in performing cataract surgery using raw surgery videos. Such evaluation can supplement human review.

As presented in Figure 5.2, the last step in ATTENTIVE workflow is to assign performance scores to each input datum. In this chapter the workflow for metric development using the Kinect DK data is presented[1].

10.2 Material and Methods

10.2.1 Subjects and Data Acquisition

The experimental setup and data acquisition method are explained in Chapters 5 and 7 in details.

---

[1]This research was presented at American College of Surgeons meeting on March 2, 2022 as: Sharif H., Phillips H., Kuzminsky J., Li Y., McNeil L., Kesavadas T., "Automated Surgical Skills Evaluation Using RGB-D data in a Longitudinal Study"

In order to validate the developed metrics, they should be compared against the medically validated assessment methods, including OSCE checklists, OSATS, and GRS. In order to provide such clinically approved scores a PTZ Optics camera, presented in 5.2.3, was utilized to provide a highly focused over-the-shoulder view from the participants hands and the suturing pad. A sample frame of this camera is presented in Figure 10.1. Three veterinary and two human medicine expert surgeons used validated assessment methods to evaluate the performance of the participants based on these video recordings, which were provided to them as blinded data.
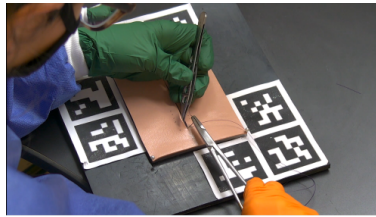
Figure 10.1: A sample frame of PTZ Optics camera

Since the evaluation process by the expert surgeons has not been completed at the time of writing this dissertation, the data from eight participants from the "Longitudinal Study" group was utilized to develop assessment metrics and detect the potential improvement in surgical performance over time. It should be emphasized that the "Expert" and "Novice and Intermediate" groups that are mentioned in Table 5.2, are based on the experience level. Therefore, it does not guarantee that any participant's performance is necessarily correlated to the assigned label based on the past experience. For instance, an expert surgeon who is adept at performing surgery on real tissue may not demonstrate a high performance in completing the surgical tasks on a synthesized model.

10.2.2 Preprocessing

Data Preprocessing approach is the same as the approach presented in Chapter 7.2.2. For some of the first participants, the orange glove was not available due to shortage of supply incurred by Covid-19 pandemic. Therefore, as shown in Figure 10.2a, the participants were putting on purple gloves on their dominant hands. This purple color was very similar to the gown color in both RGB and HSV domains that resulted in a high noise level. Therefore, a different strategy, i.e. YOLOv5 algorithm [270], was applied to extract hand pixels from the RGB frame. YOLOv5 was trained to find a bounding box around each hand at each frame as presented in 10.2.
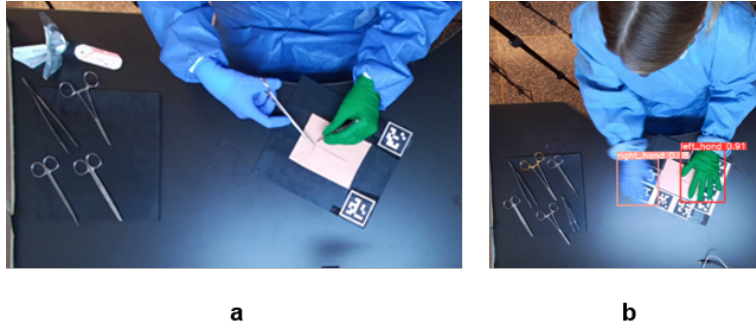
Figure 10.2: **(a)** Purple glove for the dominant hand **(b)** bounding boxes generated by YOLOv5 in a sample frame

Metric

Task completion time and motion density map were selected as the first set of metrics in this study.

10.3   Results and Discussion

The motion density maps of both hands for four participants performing Task 1 are demonstrated in Figure 10.3. An ideal motion density map would contain a few brighter regions rather than a lot of regions with less brighter color. Such a motion density map conveys a higher efficiency in hand motion, i.e. less redundant hand motions as the hands are expected to stay in the suture pad area during the task completion. An improvement in motion density map from week 1 to week 5 was observed in all the four participants presented in Figure 10.3.
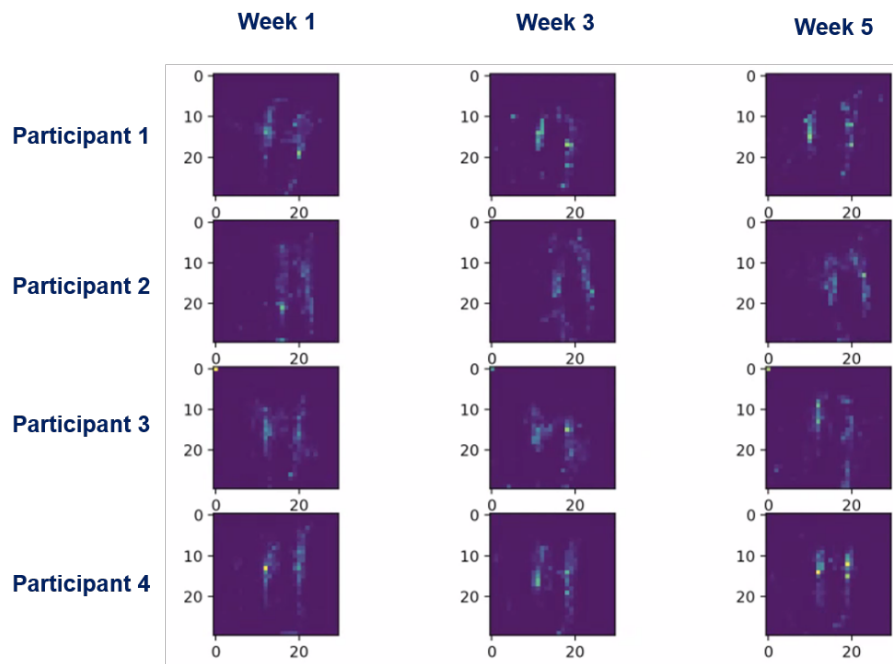
Figure 10.3: Motion density maps of both hands for four participants

The difference in task completion time between different weeks and the OSCE scores assigned by the expert surgeons are provided in Tables 10.1 and 10.2, respectively. The changes in task completion time versus changes in the OSCE score are demonstrated in Figure 10.4. For some of the participants, a reduction in task completion time, i.e. negative values in Table 10.1, and improvement in the OSCE score from week one to weeks three or/and five were observed. However, the correlation between improvement in task completion time and OSCE score is not apparent.

Table 10.1: The difference in task completion time (for needle passing task) between different weeks

	C1	C2	C3	C4	D1	D2	D3	D4
W3-W1	286	-447	-549	749	18	-409	-67	-545
W5-W1	-175	-544	-508	511	501	134	617	-436
W5-W3	-461	-97	41	-238	483	543	684	109

Table 10.2: OSCE scores

	C1	C2	C3	C4	D1	D2	D3	D4
W1	23	21.5	23.5	21.5	24	24	24	24
W3	23.5	23	23.5	23	24	23.5	24	24
W5	24	22.5	23	22	23.5	23	24	24
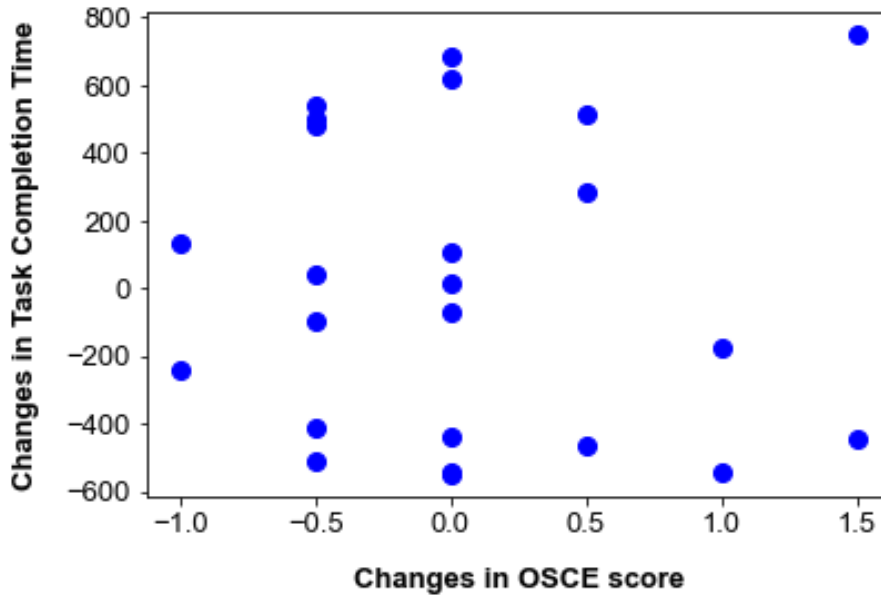
Figure 10.4: Changes in task completion time versus changes in the OSCE score for needle passing task

10.4 Conclusion and Future Work

Upon receiving the OSCE, GRS, and OSAT scores from the expert surgeon evaluators for all the participants, the following steps can be taken to develop a robust assessment system:

1. The statistical analyses should be performed on data from more participants and using more features.

2. Classification algorithms can be applied to classify participants into two or more groups based on the validated scores as the class labels.

3. Clustering participants into two or more groups utilizing different features.

4. The regression methods can be applied to assign score to an input data in a more continuous and numeric fashion, rather than just assigning a discrete Novice, intermediate, or expert label to the data.

5. Data from different sensors can be fused together to enhance a multi-modal score assignment.

10.5   Acknowledgment

Chapter 11

Quantitative Evaluation of Surgical Skills Proficiency Using EMG-IMU Sensor Fusion

11.1 Introduction

Due to the importance of developing a quantitative assessment system, many researchers have attempted to propose metrics and methodologies to achieve this goal. Tables 5.3 to 5.8 provide a comprehensive list of metrics. Watson et al. [265] used IMU data to classify the expert and novice participants performing venous anastomoses on a benchtop model. crowdsourcing and deep neural networks were applied on video data to track instruments and divide the skill level into low level and high level while performing invasive minimally surgery [266]. Aggarwal et al. [210] utilized ROVIMAS video-based motion tracking device to evaluate dexterity of two groups of experienced and inexperienced surgeons in performing laparoscopic surgery. The surgical tool's motion data along with the location of surgeon's eye gaze on the screen were utilized to distinguish expert from novice participants performing a cadaveric eye laparoscopic surgery [267].

Pe´rez-Duarte [200] found statistically significant differences in the mean EMG signals from the forearm muscles between novice and expert groups while performing suturing. Based on their observations, Novices had higher muscle activation compared to experts performing the same suturing task. Increased muscle activation can be correlated with increased physical effort, which can lead to fatigue, pain, and sometimes more serious lesions [200].

As presented in Figure 5.2, the last step in ATTENTIVE workflow is to assign performance scores to each input datum. In this chapter the workflow for metric development using the Myo armband data is presented.

11.2 Material and Methods

11.2.1 Subjects and Data Acquisition

To record sEMG signals, a Myo armband [50] was used in this study. Myo armband is a low-cost consumer-grade armband that contains an array of eight dry-electrode surface EMG sensors with the sampling rate

of 200 Hz and an Inertial Measurement Unit (IMU) sensor with 50 Hz sampling rate. It is an easy-to-use armband as a user can simply slip the bracelet on and no preparation is required [55]. Figure 5.4a shows the armband placement on a participant's forearm during data collection. To maintain consistency between participants in collecting EMG data, the Myo armband was worn on the participant's dominant arm in a way that a particular sensor was always placed on the Flexor carpi radialis muscle, as presented in [263], and the rest of the sensors are spaced out evenly around the lower arm. The experimental setup and data acquisition method are explained in Chapters 5 and 9 in details.

In order to validate the developed metrics, they should be compared against the medically validated assessment methods, including OSCE checklists, OSATS, and GRS. In order to provide such clinically approved scores a PTZ Optics camera, presented in 5.2.3, was utilized to provide a highly focused over-the-shoulder view from the participants hands and the suturing pad. A sample frame of this camera is presented in Figure 10.1. Three veterinary and two human medicine expert surgeons used validated assessment methods to evaluate the performance of the participants based on these video recordings, which were provided to them as blinded data.

Since the evaluation process by the expert surgeons has not been completed at the time of writing this dissertation, the data from eight participants from the "Longitudinal Study" group was utilized to develop assessment metrics and detect the potential improvement in surgical performance over time. It should be emphasized that the "Expert" and "Novice and Intermediate" groups that are mentioned in Table 5.2, are based on the experience level. Therefore, it does not guarantee that any participant's performance is necessarily correlated to the assigned label based on the past experience. For instance, an expert surgeon who is adept at performing surgery on real tissue may not demonstrate a high performance in completing the surgical tasks on a synthesized model.

11.2.2 Preprocessing

After rectifying the EMG signals, EMG and IMU signals were synchronized by down sampling EMG data from 200 Hz to 50 Hz. Synchronizing EMG and IMU data facilitates separation of particular parts of the EMG signal based on the IMU data. In this work, the signal was separated into two sections. The first section, referred to as tool-holding or Grasp section, is when a participant holds the surgical tools with correct technique after loading the suture. During the data collection session, the participants were instructed to hold the instruments with correct technique for 5 seconds and start the task when they were verbally instructed. The second part of the EMG signal is when the participants performed the surgical tasks, referred to as the Task section. These sections are presented in 5.1.

To find the Grasp section of the signal and separating it from the Task section, a moving window of the size of 25 samples, i.e. 0.5 Seconds, was applied. The start point of the Task section was defined based on changes in the acceleration along y-axis, when the average acceleration magnitude was larger than a set threshold. This threshold worked for the entire longitudinal study dataset except of one trial, for which the separation point was detected manually. The access and threshold were determined heuristically. The Task section of the signals were used for further processing to develop evaluation metrics.

In order to capture the changes in muscle activation between the sessions and correlate it to surgical proficiency level, the EMG signal was normalized to compensate for variations in muscle activation between the participants and for each participant in different data-collection sessions. To achieve this goal, first the mean of signal at the very last second was calculated. In this last second the participants were instructed to place their hands on the surgical table in a neutral position. The EMG signal from each sensor was

normalized by dividing the signal to the signal-mean during the last second obtained from the same sensor.

Subsequently, a moving average filter with a window size of 5 samples, 0.1 second, was used for smoothing the signal of each sensor. In an attempt to evaluate the effect of different filtering schemes on the final results, a first order Butterworth low-pass filter with a 1 Hz cutoff frequency [71] was applied on each channel of EMG data. It was observed that the choice of filtering method doesn't affect the final results, i.e. teh distribution of data over different weeks.

11.3   Results and Discussion

The mean of EMG signal's magnitude for each sensor and each task before standardization of the dataset was calculated. Standardizing data changes the range of the values; however, it does not affect the distribution of data among the weeks. As an example, data from sensors 2 and 7 as well as the average of signal magnitude's mean over the eight sensors are plotted in Figure 11.1. Sensors 2 and 7 were placed on the larger muscle groups in the lower arm, sensor 2 on the Flexor carpi radialis muscle. The Box plots of means of signal's magnitude, aggregated for all the participants, are presented in Figure 11.2. These plots do not show a significant trend in improvement of signal's magnitude.
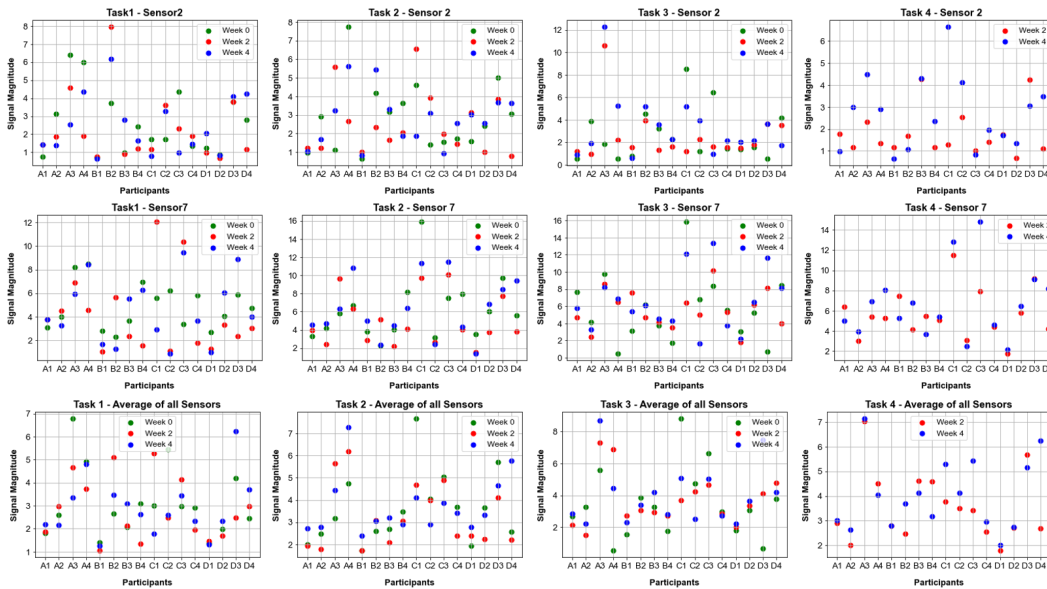
Figure 11.1: Sample graphs of mean of signal magnitude
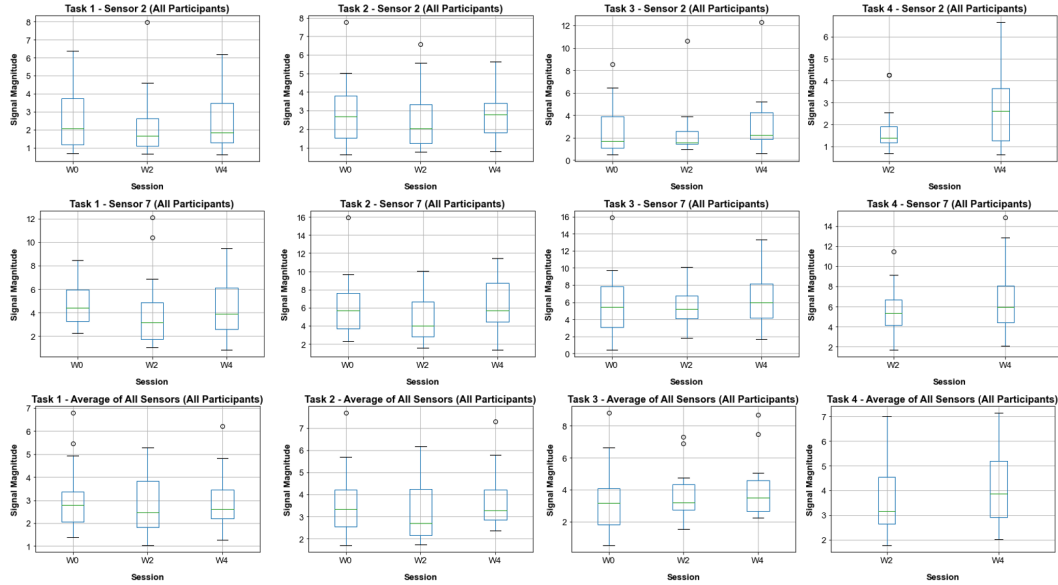
Figure 11.2: Sample Box plots of mean of signal magnitude for all participants

Two expert surgeons, who had not participated in the study, evaluated the participants' performance using three grading scales, i.e. GRS, OSAT, and OSCE. The Box plots of their evaluation scores are presented in Figures 11.3 to 11.5. No difference is observed between evaluations for OSCE score; however, evaluator B assigned higher scores using OSAT and Global metrics. According to both evaluators, the accumulated OSCE score was improved from the first week to the second session and did not significantly change between the second and the last sessions. A similar trend is observed in the mean of signal's magnitude (see figure 11.2) for tasks 1 and 2 between the first and second sessions; however, the trend is different for the second and the last sessions.

To evaluate the changes in the mean of EMG signal's magnitude between different sessions in a more quantitative way, a repeated samples T-test was run by applying the Python command scipy.stats.ttest_rel. The ratio of the larger sample variance, where sample refers to data from all the participants at each week, to the smaller sample variance is 1.586, which is less than 2. As a result, it is legitimate to assume that the population variances are equal, and the t-test results are reliable. The T-test results are presented in 11.1. All the p-values are greater than 0.05 that means there is no statistically significant differences in the average of signal magnitude between the sessions. These findings were not inline with OSCE and OSAT checklists, that demonstrate an improvement in performance from the first to the second session. However, such findings are in agreement with the GRS scores that suggest no significant changes between different weeks. For many of the comparisons, data from sensor 7, which had collected data from Extensor Carpi Radialis muscle group and Brachioradialis, provided the smallest p-value among the sensors. This may suggest that the data from this sensor is more reliable for metric development and less affected by different types of noises.
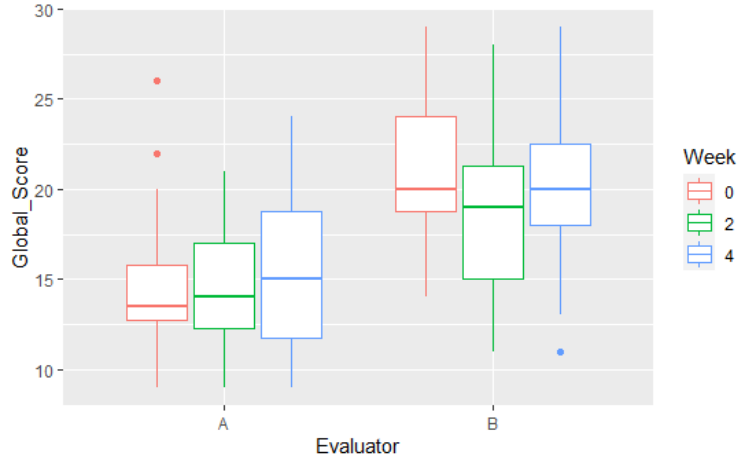
Figure 11.3: Box plot of GRS scores provided by the expert evaluators
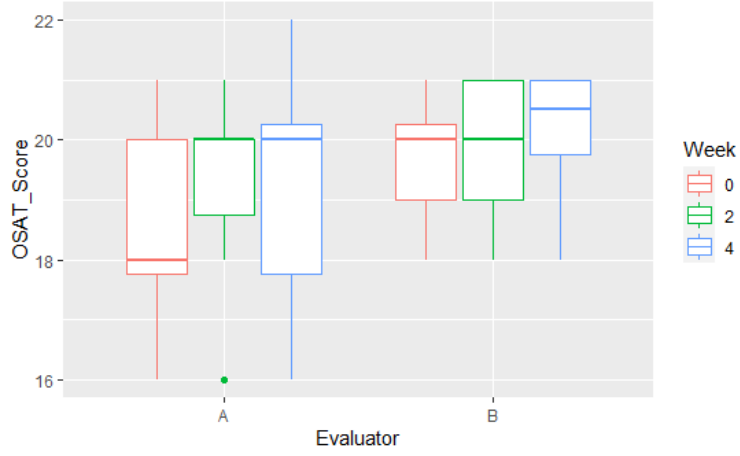
Figure 11.4: Box plot of OSAT scores provided by the expert evaluators
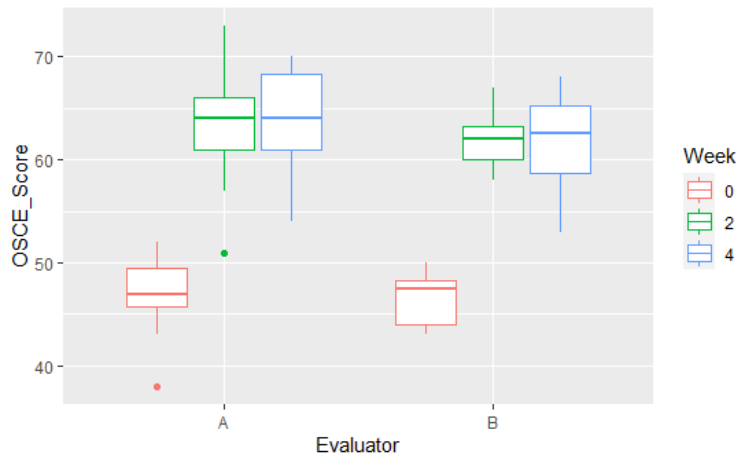
Figure 11.5: Box plot of OSCE scores provided by the expert evaluators

Table 11.1: p-value obtained from running the T-test on the EMG signal's mean from different weeks. Each column contains the data for one sensor, Shown as S.

Task	Groups	S 1	S 2	S 3	S 4	S 5	S 6	S 7	S 8
Needle Passing	W0-W2	0.82	0.88	0.69	0.53	0.29	0.79	0.09	0.2
	W0-W4	0.72	0.79	0.74	0.84	0.95	0.95	0.19	0.57
	W2-W4	0.96	0.94	0.94	0.66	0.16	0.81	0.56	0.4
Knot Tying	W0-W2	0.95	0.6	0.98	0.93	0.88	0.81	0.32	0.35
	W0-W4	-0.3	-0.9	-0.4	-1.4	0.69	-1.4	-0.06	-0.2
	W2-W4	0.83	0.66	0.36	0.53	0.18	0.67	0.24	0.25
Continuous Suturing	W0-W2	0.42	0.86	0.73	0.56	0.37	0.99	0.93	0.53
	W0-W4	-1.36	-0.75	-0.5	-0.76	-0.82	0.15	0.62	-0.67
	W2-W4	0.7	0.37	0.36	0.84	0.96	0.85	0.38	0.94
Buried Suturing	W2-W4	0.37	0.08	0.8	0.7	0.55	0.85	0.36	0.39

11.4 Conclusions and Future work

Average of EMG signal's magnitude was utilized to investigate the changes in participants dexterity in performing basic surgical tasks over the course of a longitudinal study. No significant changes in the average of EMG signal's magnitude was observed, which is inline with GRS evaluation scheme.

Upon receiving the OSCE, GRS, and OSAT scores from the expert surgeon evaluators for all the participants, the following steps can be taken to develop a robust assessment system:

1. The statistical analyses should be performed on data from more participants and using more features.

2. Classification algorithms can be applied to classify participants into two or more groups based on the validated scores as the class labels.

3. Clustering participants into two or more groups utilizing different features.

4. The regression methods can be applied to assign score to an input data in a more continuous and numeric fashion, rather than just assigning a discrete Novice, intermediate, or expert label to the data.

5. Data from different sensors can be fused together to enhance a multi-modal score assignment.

Part III
iBAND Design and Fabrication

Chapter 12

iBAND Design and Fabrication

12.1   Introduction

In the previous chapters a Myo armband was used to collect EMG and IMU data from the lower arm. However, the Myo armband became discontinued. In order to facilitate collecting EMG and IMU data from the lower arm, we proposed to build a wearable technology, to replace Myo armband, as a part of ATTENTIVE. To achieve this goal, we recruited five undergraduate students from Mechanical Science and Engineering as well as Electrical and Computer Engineering departments of the University of Illinois at Urbana-Champaign to help with building iBand. The first two students worked on iBand for 2-months (one student full time and the other one as an hourly position). Three students worked on iBand as their senior design project. Hajar Sharif, the author of this dissertation, supervised and worked closely with the undergraduate students to design and build the iBand.

The original idea was to equip iBand with a Galvanic Skin Response (GSR) and a Pulse sensor on top of an array of EMG sensors and one IMU sensor. The second goal was to make iBand wireless. Specially, In addition, the iBand should have a rechargeable battery that can be charged both on the iBand and on an external station. The next design criteria was to keep the iBand price lower, or comparable if better features are offered, than the alternative EMG sensors on the market. A list of alternative systems is presented in Table 12.1. Moreover, given the iBand is supposed to be part of ATTENTIVE, it is necessary for it to be light and not bulky, so wearing it does not interfere with a user's performance while performing surgical tasks. furthermore, the iBand should come with a user-friendly graphical user interface (GUI) for recording, saving, and visualizing the data.

Table 12.1: Commercially available products

Name of Device	Price (in 2020)
Delsys Systems Trigno Avanti [271]	$20,000
Biometrics Datalite Semg [272]	$17,000
Noraxon Ultium EMG [273]	$20,000
Oymotion Gforce-Pro [274]	$1,250
Thalmic Lab Myo Armband [50]	$200

12.2   Materials and Methods

Based on the desired design criteria, the following components were purchased off the shelf and their performance was validated:

1. Arduino Micro [275]

2. HiLetgo HC-05 Wireless Bluetooth RF Transceiver Master Slave Integrated Bluetooth Module [276]

3. SparkFun 9 axes IMU sensor [277]

Figures 12.1a and 12.1b demonstrate the translational and rotational validations of IMU sensor, respectively..

(a)   (b)
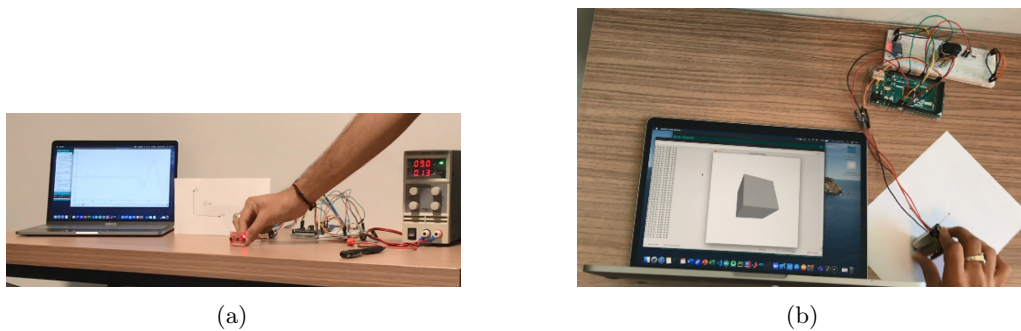
Figure 12.1: Validating (a) translational and (b) rotational functionalities of IMU sensor

The pin connections of Arduino Micro processor and the first iBand prototype, before adding the EMG sensors, are presented in Figures 12.2a and 12.2b, respectively.

(a)   (b)
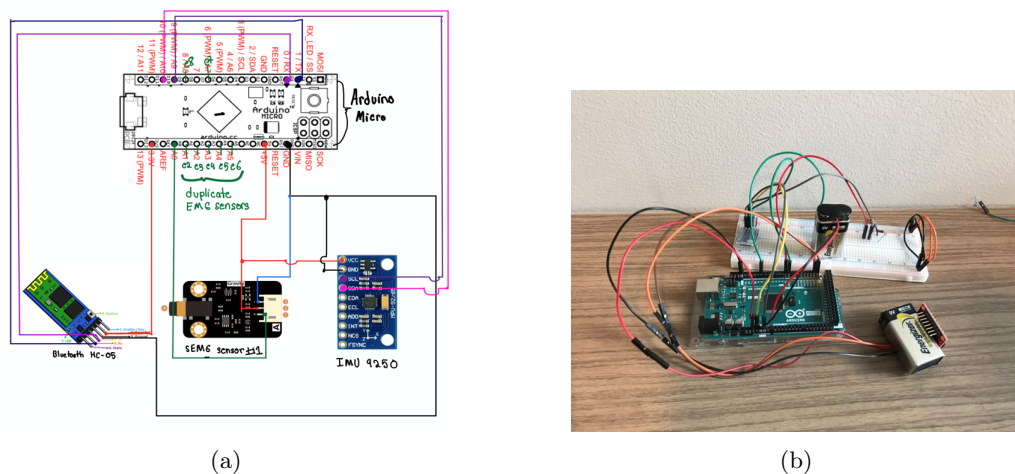
Figure 12.2: (a) Pin connections of Arduino Micro, and (b) first iBand prototype

The first phase of the iBand project was performed at the begging months of the Covid-19 pandemic. As a result, the choice of sensors and equipment was more limited than usual. Based on the specifications of the available EMG sensors in the market, Gravity EMG sensor by OYMotion [278] was purchased and tested as the first option. Usually, two types of EMG electrodes, i.e. needle and surface electrodes, are used for

clinical and research applications. The surface electrodes may be sticky electrode, dry electrode with gel, or dry electrode without gel [279]–[281]. Although the dry electrode EMG sensors were required for iBand, sticky electrodes were used as a reference to compare the dry electrode sensor's performance. OYMotion EMG sensor failed to collect EMG data robustly. Therefore, some additional EMG sensors were purchased. Each EMG package is composed of two main pieces, i.e. EMG electrodes and EMG circuitry to filter, rectify, and amplify the signal. Different combinations of EMG electrodes and EMG sensor circuitries were tested to determine the best combination. The combination of MFI medical gold-plated bar EMG electrode [282] and OYMotion circuitry demonstrated the best performance, so they were selected for the rest of the project. The second iBand prototype, which includes two EMG sensors, is presented in Figure 12.3. In this prototype, the Arduino Micro was replaced with Arduino Nano to reduce iBand's size. Figures 12.4a and 12.4b show the EMG signals from the EMG electrodes on back and front of the lower arm when the fingers are respectively in neutral position and when a couple of fingers are flexed or extended.
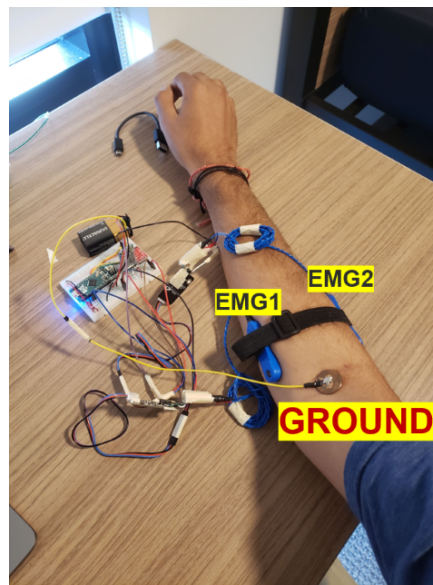
Figure 12.3: Second prototype of iBand with two EMG sensors

(a)   (b)
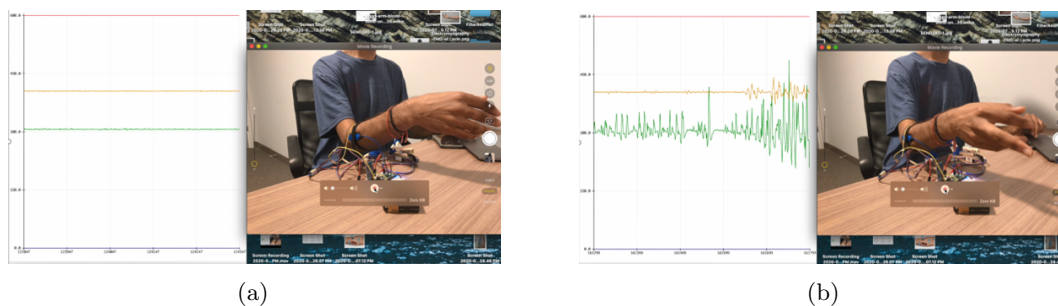
Figure 12.4: **(a)** Collected signals by the EMG sensors while the hand is in the neutral position, and **(b)** when two of the fingers are extended or flexed

Through the experiment, it was observed that using one ground for each MFI EMG electrode reduces the noise significantly. In the next iBand prototype, the enclosure was built to keep the EMG sensors in place

83

and pennies were used in between each pair of metal surfaces of each MFI electrode to function as the ground electrode. The black capsules in the enclosure hold the EMG circuitries and are connected to each other with rubber bands to enhance adjustable size for the iBand. The iBand prototype and the penny electrodes are presented in Figures 12.5a and 12.5b, respectively. In the third prototype, the Arduino and the breadboard were replaced with a printed circuit board (PCB) to reduce the size.

The pulse and GSR sensors were used to collect data from different areas of the lower arm. These sensors are designed to collect data from the fingers, and they function properly while are placed on the fingers. However, they were unable to collect reliable signals from other areas of the lower arm, except for the pulse sensor that can be used inside the elbow joint or on the wrist. As a result, there is a compromise between collecting reliable signals utilizing these two sensors and the convenience of using iBand. The reason is placing sensors inside the elbow joint can make discomfort for the user and interfere with bending the elbow through its whole range of motion. Similarly, wearing iBand on the wrist can interfere with users' performance for delivering delicate tasks such performing surgery. Moreover, It should be confirmed through research whether the wrist is an optimum place to collect the EMG signals to monitor fingers' flexion/extension. One idea to resolve this issue is to build iBand into two sections including the main iBand's body, which holds the electronics as well as the EMG and IMU sensors, and an extension that holds the GSR and pulse sensors.
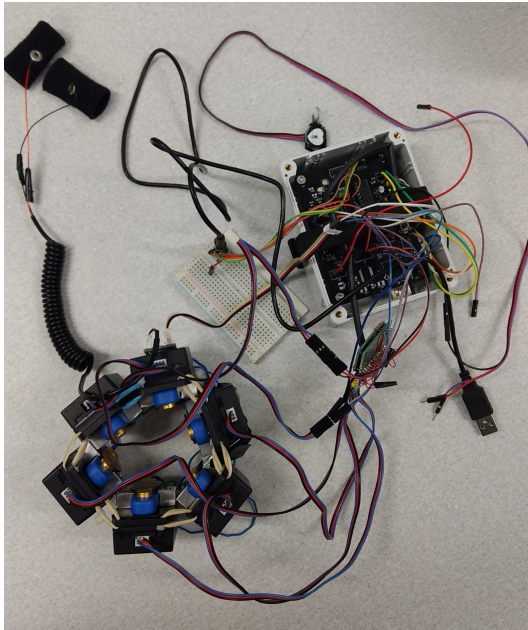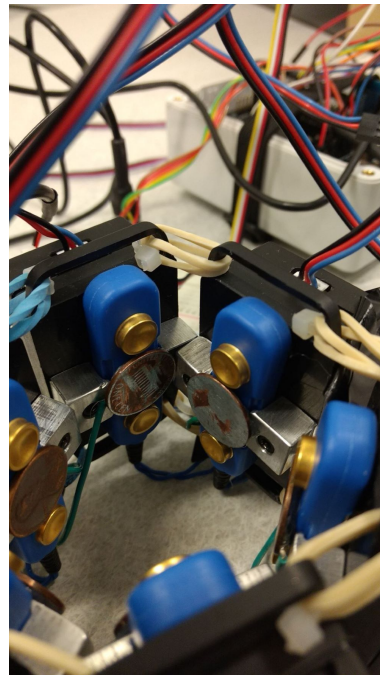
(a)   (b)

Figure 12.5: **(a)** Third prototype of iBand with enclosure and PCB, and **(b)** pennies as the ground electrode for the EMG sensors

Python scripts were developed to receive packs of data from IMU and EMG sensors. In addition, a GUI was developed to save the data from different sensors as synchronized data in separate files in the folder and with the file names specified by the user. Arduino IDE was used to develop the firmware code for the microcontroller to facilitate utilizing publicly available libraries for the Arduino. A multi-threaded approach was implemented using Arduino Threads library in Python [283] to collect data from multiple sensors in a

parallel manner. Then, the collected data was packaged in the form of a JavaScript Object Notation (JSON). JSON is a lightweight and compact framework that is widely used to send data, and many libraries have been developed on top of it. Specifically, the ArduinoJSON library [284] that was used in iBand to package sensor data into a JSON.

12.3   Conclusion and Future Work

This chapter includes a brief description of the design and fabrication of a wearable technology, named iBand, to collect biosignals and motion data from the lower arm. Different components of iBand have been designed, purchased, assembled, and calibrated to read reliable synchronized data from the lower arm, transfer data to a computer via Bluetooth, and save the data on computer in separate files. An easy-to-work user interface has been developed for iBand to enable a user to save the data in the desired folder and determine the file names. In addition, the user interface enhances a real-time data observation in which the user can choose the sensor from which the collected signal will be displayed as a line graph.

As the next step of developing iBand , the codes should be modified to remove a lag that currently exists between the collected and displayed data. In addition, a new enclosure should be designed and fabricated to enhance wearing iBand as easy as wearing a bracelet, i.e. just sliding up the iBand on the lower arm. Moreover, the decision on whether or not to include the pulse and GSR sensors in the iBand should be finalized. Inclusion of these sensors can be facilitated through either wearing iBand on the wrist, rather than closer to the elbow, or adding a second section to iBand. The second section can be either an attached extended arm or a completely detached part from the main iBand's body. In this way, the pulse and GSR sensors can be placed on the areas of the lower arm that are more ideal for collecting their corresponding signals. Furthermore, the iBand should be downsized. This goal can be achieved through some changes in iBand. First of all, the EMG circuits may be replaced by Python scripts to filter the noise. Additionally, the extra wires and the MFI blue enclosures to hold the golden plate electrodes can be removed or be replaced by more compact alternatives. at last, the PCB can be downsized by utilizing both sides of it.

Chapter 13

Conclusion and Future Directions

In this dissertation, the application of HGR for addressing two challenges in the medical field is presented. The first challenge is to develop a metric to improve rehabilitation of neurological conditions, with a focus on improvement of ADL performance, while the second challenge is to develop a metric to enhance a better evaluation method for training the basic surgical skills.

13.1 Hand gesture recognition for evaluating dexterity in performing the activities of daily living

Metric development process for ADL rehabilitation assessment is presented in Chapters 1 to 4. Since the first step of developing a metric is to distinguish different ADL activities using hand gesture data, in this dissertation the focus is on classification of ADL tasks using hand gestures. Muscle activation signals (EMG) and hand kinematics data were collected from the lower arm of healthy adults while performing eight common daily tasks. A variety of neurological conditions affect the hand grasp in some upper limb positions more than other postures [285]. Therefore, the data from the healthy participants were collected when the tasks were performed over the entire range of motion used in daily life to ensure capturing of motion details over the entire required range for performing ADL decently. Then, different features in time and frequency domains were extracted from the data and were fed into two classification algorithms to detect the best data analysis pipeline for classification of the ADL tasks. Over 99% classification accuracy and over 97% precision and recall were achieved in the presented studies.

These findings pave the way for developing an ADL-assessment-metric in two folds. The first fold is the immediate application of these findings in evaluating neurological patient's performance, while the second fold is concerned with the long-term applications.

The immediate application benefits from the data analysis pipelines that were developed in the current dissertation. The data analysis pipelines take in raw data, collected by using the leap motion controller as well as the EMG and IMU sensors, from the lower arm during completion of certain ADL tasks. These pipelines output classification accuracies to distinguish the ADL tasks. Different preprocessing, feature extraction, and classification methods were tested on the data from healthy adults to detect the best structure and parameters for the proposed pipelines. The developed frameworks can be set as the references for evaluating patients' data obtained from leap motion controller and Myo armband, or similar devices. Then, hand motion and muscle activation data of a neurological patient completing the same tasks in the same data

collection setup can be fed into the reference pipelines to obtain the classification accuracies. The achieved accuracies indicate how close a patient's hand motion and muscle activation are to the hand motion and muscle activation of the healthy population. This method enhances assessment of the overall performance of a patient in a quantitative fashion. In addition, the acquired confusion matrices provide insight into the patient's performance in completing each individual task.

As for the long-term application, the features that achieved a higher classification rate can be used for further analysis and for developing other metrics as they represent different classes, i.e. ADL tasks, in a more distinct way. For instance, the distribution of these features for each ADL task obtained from data of the healthy adults can be set as a reference metric. In this scenario, the location of each feature obtained from patient's hand data in the reference distribution can be used to evaluate the patient's performance and the rehabilitation progress. More analysis of the data from healthy adults as well as collecting the same data from the neurological patients are required to complete the long-term-application metric.

In conclusion, the future work of the first section of this dissertation can investigate three directions:

- Other classifiers can be investigated to increase the algorithm's speed and provide the capability for real-time applications.

  Currently, the rehabilitation routines are mainly hospital-centric that is costly and requires patient's travel to the therapy centers. The restricted mobility of some neurological patients, such as stroke patients, aggravates the situation. As a result, after a while, the patients continue therapy at home on their own accord [286]. Nevertheless, while still in the hospital setting, the therapy programs should be complemented by prescribed exercises to be performed at home [286]. The therapy being practiced at home is assessed based on the self-reports by the patient and/or their family members that are qualitative measures of patient's performance [287].

  In the hospital setting, a therapist monitors the patient's performance while the absence of a human chaperone at home hinders the therapy progress and negatively affects the patient's motivation towards therapy [288], [289]. Thus, home-based therapy paradigms such as tele-robotic-rehabilitation systems have been presented as a viable alternative so a therapist can remotely evaluate and adjust the rehabilitation progress [290]–[292]. These systems are still limited as they call for a therapist's remote involvement that is not always possible due to high workload of the therapists [293]. Such limitation can be removed by development of intelligent robotic-rehabilitation systems that are capable to assess the patient's performance autonomously in real-time and tailor the rehabilitation regimen accordingly.

- The data from different sensors can be fused together and be used as the input for metric development. There are two methods to develop evaluation metrics using the features from different sensors:

  1. Sensor fusion: in this method, the best features, which represent the extracted data from each sensor in the most distinctive way, are fused together to form a larger feature map. Consequently, the constructed feature map will be used as the input of the assessment system. This method may provide a higher accuracy as the feature map includes a wider range of information from the lower arm obtained from a variety of sensors. For instance, it contains both kinematic and muscle activation data. However, this method lacks the ability to provide feedback on the individual parameters that affect the performance. For example, it cannot determine how much the performance is affected by muscle activation and how much by the lower arm speed of motion.

2. Sensor specific metrics: in this method, the scores are calculated using features obtained from each individual sensor's data. As a result, feedback on individual aspects of the performance can be provided. Although this approach requires development of a methodology to combine the scores from different sensors to output an overall performance score.

It is noteworthy to mention these methods are applicable to metric development for evaluating performance in executing ADL tasks and proficiency in performing surgical skills.

- The ADL dataset can be expanded by recruiting more healthy and neurological-patient participants, so the proposed methodologies will be advanced further toward development of a quantitative assessment system.

The findings of the first section of this dissertation provides the backbone for the data collection setup and the data analysis methodology to develop a quantitative ADL assessment metric. Nevertheless, data from a larger group of healthy adults should be included to increase robustness of the metrics and eliminate the unfairness that can be imposed into the results due to diverse nature of hand grasps. For the sake of clarification, different healthy adults may use different hand grasps to perform the same ADL task. Despite the diversity in these hand grasps, all of them are considered as normal/healthy grasps and should be included in development of the reference for grasp evaluation. Data paucity can result in skewed metrics that fail to capture the manifold normal grasps and consequently, a higher error in assessing the patient's performance in conducting the ADL tasks. A larger number of healthy-adult participants can increase the probability of inclusion of the diversified hand grasps that in turn makes the metric more robust and reliable.

In addition, the study in the first section of this dissertation was conducted with healthy participants. Therefore, although the findings of this study proves the applicability of the proposed methodology to distinguish different ADL tasks and potentially to develop a quantitative ADL assessment metric, the inclusion of data from the neurological patients are crucial to generalize the outcome of the study for clinical applications. Particularly, the experimental setup should be tested in the clinical setting with neurological patients to investigate its feasibility and efficacy as a clinical assessment tool. For instance, a neurologically impaired participant may have difficulty to form a particular hand grasp with the affected limb or apply sufficient force to form the required grip for holding onto the objects. In such cases the grasp in the affected limb is formed with the help of the intact limb to open/close the impaired fingers. This may impose some challenges during the data collection or data analysis. Moreover, the outcome of the developed metrics should be compared against the assessment results obtained from the widespread and clinically accepted assessment tools to ensure their functionality in evaluating the neurological patient's dexterity in performing ADL tasks.

In addition, a larger sample size, for both intact and neurologically impaired participants, provides a better estimation of the distribution of each feature for each ADL task over the entire population. Once such distributions are obtained for both healthy and neurologically impaired participants, a judicial judgment can be made to select the feature-task combinations with the maximum difference among the intact and neurologically impaired participants. If for a particular feature-task combination the distributions of both groups overlap a lot, that feature-task combination may fail to serve as an effective metric to monitor the rehabilitation progress.

13.2 Hand gesture recognition for surgical skill assessment

The second part of this dissertation includes the design of ATTENTIVE; an evidence-supported, automated, robust, real-time, comprehensive, quantitative assessment system for evaluating proficiency in performing basic surgical tasks.

Acquiring excellent surgical and microsurgical skills are critical to avoid high stakes and morbidity to the patients. Gaining surgical skills requires significant practice and mentorship, yet opportunities to gain on-the-job surgical training are scarce. Administrative, legal, and ethical pressures often understandably preclude exposure of surgical patients to novice surgeons on the steepest part of the learning curve, which as a result, negatively affects many surgical training programs. In addition, fewer and fewer medical training programs use animals in the delivery of medical education due to societal pressures and perceptions. As a result, development of non-living models and simulations using artificial intelligence technologies is critical to advancement of surgical training in medical education. Evidence suggests that many highly technical skills may be acquired and refined outside of the operating arena through the use of surgical simulations [136]–[143].

Many models and some simulation-based exercises for surgical, microsurgical, and minimally invasive surgical training and assessment have been described, but few have been validated as teaching and learning tools [136]–[138]. Validation of any surgical simulation system is needed to ensure skills transfer to real-life surgical situations. Robust and quantitative skill assessment methods are required to perform such system validation.

Moreover, with traditional surgical training, as well as with the surgical simulation, evaluation of a trainee's performance involves the direct observation, supervision, and input of a surgical expert, which is very costly, time-consuming and subjective. It also limits the trainees' opportunities for receiving feedback on their performance to the availability of an expert surgeon. Incorporation of surgical simulation into longitudinal training courses and surgical curricula is sorely lacking and requires robust assessment methods [162]. In addition, provision of real-time quantitative feedback has been lacking with the use of many surgical simulation techniques [145], [163], [164]. Therefore, development of an ATTENTIVE assessment system is crucial to achieve higher training performance and pave the way for more widespread inclusion of advanced simulators such as virtual/augmented reality into the surgical curriculum.

The structure of ATTENTIVE is presented in Figure 5.2. In this work, the idea and framework for developing ATTENTIVE are presented. Next, the apparatus and experimental setup and protocol to investigate the feasibility of ATTENTIVE were designed and built. Afterwards, data was collected from 65 participants completing four basic surgical tasks. The participants were students, residents, and experienced surgeons in the fields of veterinary and human medicine. To benefit from both sensor-based and vision-based HGR methods for solving the problem in hand, Azure Kinect DK, Leap Motion Controller, and Myo armband were used to collect data from the lower arm of the participants. The details of analyzing data from these sensors and the obtained results are presented on Chapters 6 to 11. These chapters are divided into three parts based on ATTENTIVE pipeline (see Figure 5.2).

The first part includes automated clustering of the video frames to facilitate extraction of the main task portion of the videos from the rest of the sections. The literature review on this topic and an implemented algorithm to perform clustering were presented in Chapter 6. However, the annotation of the dataset, i.e. to assign a label to each frame of each video, is half-way done. Therefore, the future work of this chapter includes finishing dataset annotation and testing the performance of the clustering algorithm on the annotated dataset.

The next part includes Chapters 7 to 9 that are concerned with classification of the four surgical tasks using data collected from each device. Over 95% classification accuracy was achieved from Kinect DK data. The preprocessing of Myo armband and LMC data as well as the implementation of feature extraction and classification algorithms are completed. The future work includes extracting more frequency domain features from the data and testing the classification performance of the implemented algorithms.

The last part contains development of assessment metrics based on the collected data from each sensor. This part includes the Chapters 10 and 11. Metrics were extracted from the preprocessed data and were used to evaluate the changes in performance of the participants in the longitudinal study over the course of four weeks. The outcomes were tested against the gold standard methods that are being wildly used in clinical settings. In order to use the gold standard methods, expert surgeons, who had not participated in the study, evaluated the performance of the participants by watching the recorded videos from the participants' lower arms. These expert evaluators were not provided with any information about the participants' experience level or expertise background, so they assigned scores only based on the participants' performance in completing the assigned surgical tasks. Some of the presented metrics in this dissertation showed similar results to the gold standard methods.

The future work of this part includes using clustering algorithms to separate participants based on their performance levels. In addition, the proposed metrics should be investigated on the entire dataset, i.e. including data from all the participants, to develop a metric system for evaluating surgical tasks. These goals can be achieved once the scores based on the clinical standards are received from the expert evaluators.

In the first two sections of this dissertation, data were collected from the lower arm of the human participants to develop assessment metrics. Since experimentally collecting data, specially from human participants, is usually costly and time consuming, a valid question that may arise is whether the hand grasps can be artificially generated using kinematic and mechanical models to complement the experimental data from the human hand. Besides, adding model-based data to empirical data for metric development facilitates incorporation of more variability into the reference metric by including individualistic parameters, such as knuckle size, that vary from one person to the other. Therefore, the combination of the augmented and experimentally collected data may represent a wider range of demographics and consequently is more generalizable and reliable for clinical applications. System identification methods and algorithms such as autoencoder can be applied on the experimental data to extract a hand model and confirm its reliability.

13.3   iBAND design and fabrication

The last chapter of the current dissertation, i.e. Chapter 12, includes preliminary work on design and fabrication of a wearable device, named iBand, to collect biosignals and motion data from the lower arm. Different components of iBand have been designed, purchased, assembled, and calibrated to read reliable synchronized data from the lower arm, transfer data to a computer via Bluetooth, and save the data on computer in separate files. An easy-to-work user interface has been developed for iBand to enable a user to save the data in the desired folder and determine the file names. In addition, the user interface enhances a real-time data observation in which the user can choose the sensor from which the collected signal will be displayed as a line graph.

As the next step of developing iBand , the codes should be modified to remove a lag that currently exists between the collected and displayed data. In addition, a new enclosure should be designed and fabricated to enhance wearing iBand as easy as wearing a bracelet, i.e. just sliding up the iBand on the lower arm.

Moreover, the decision on whether or not to include the pulse and GSR sensors in the iBand should be finalized. Inclusion of these sensors can be facilitated through either wearing iBand on the wrist, rather than closer to the elbow, or adding a second section to iBand. The second section can be either an attached extended arm or a completely detached part from the main iBand's body. In this way, the pulse and GSR sensors can be placed on the areas of the lower arm that are more ideal for collecting their corresponding signals. Furthermore, the iBand should be downsized. This goal can be achieved through some changes in iBand. First of all, the EMG circuits may be replaced by Python scripts to filter the noise. Additionally, the extra wires and the MFI blue enclosures to hold the golden plate electrodes can be removed or be replaced by more compact alternatives. at last, the PCB can be downsized by utilizing both sides of it.

Final note: In this dissertation, artificial intelligence (AI) methods are applied to address two challenges in the medical and healthcare field to answer one question; "Can AI model evaluate human performance for sensitive medical applications?".

We recognize that AI has limitations. Obtaining data from a narrow set of human subjects may lead to invalid evaluation and, accordingly, erroneous adjustments in the methodologies that, in turn, adversely affect rehabilitation or surgical performance. Thus, like any other such systems, much more diverse data must be used to develop models before they can be deployed for real-world use.

We proposed data collection protocols and setups, developed AI models, investigated their performance on empirical data, and achieved promising results that proved the feasibility of developing reliable AI models for evaluating human dexterity in performing ADL and surgical tasks. Despite such good results, we must realize that the long-term performance of AI systems and their reliability and applicability in various situations depending on the robustness of the developed algorithms and the data used for training the algorithms along with hyperparameter selections. Hence, care must be taken to obtain data from a more diverse end-user demographic during the development stage to increase the robustness of the final AI-dependent products for healthcare applications. A robust AI-driven product can be developed by including data from participants in different geographic regions, with participants selected from a variety of races, ethnicity, physical abilities, and experience levels. This can only be achieved through broader collaboration across the globe and multi-center efforts for collecting data leading to a much more diverse and sizeable dataset. This dissertation is a first step towards developing methodologies that can set the stage for a much more significant demonstration of the role of AI in the future of healthcare performance evaluation.

References

[1] H. I. Mohammed, J. Waleed, and S. Albawi, "An inclusive survey of machine learning based hand gestures recognition systems in recent applications," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 1076, 2021, p. 012 047.

[2] M. Oudah, A. Al-Naji, and J. Chahl, "Hand gesture recognition based on computer vision: A review of techniques," *journal of Imaging*, vol. 6, no. 8, p. 73, 2020. DOI: 10.3390/jimaging6080073.

[3] S. S. Fels and G. E. Hinton, "Glove-talk: A neural network interface between a data-glove and a speech synthesizer," *IEEE transactions on Neural Networks*, vol. 4, no. 1, pp. 2–8, 1993. DOI: 10.1109/72.182690.

[4] S. Kim, J. Kim, S. Ahn, and Y. Kim, "Finger language recognition based on ensemble artificial neural network learning using armband emg sensors," *Technology and Health Care*, vol. 26, no. S1, pp. 249–258, 2018. DOI: 10.3233/THC-174602.

[5] F. R. Khan, H. F. Ong, and N. Bahar, "A sign language to text converter using leap motion," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 6, no. 6, pp. 1089–1095, 2016.

[6] D. Arfib and L. Kessous, "Gestural control of sound synthesis and processing algorithms," in *International Gesture Workshop*, Springer, 2001, pp. 285–295. DOI: 10.1007/3-540-47873-6_30.

[7] N. L. Hakim, T. K. Shih, S. P. Kasthuri Arachchi, W. Aditya, Y.-C. Chen, and C.-Y. Lin, "Dynamic hand gesture recognition using 3dcnn and lstm with fsm context-aware model," *Sensors*, vol. 19, no. 24, p. 5429, 2019. DOI: 10.3390/s19245429.

[8] C. Lee, S. Ghyme, C. Park, and K. Wohn, "The control of avatar motion using hand gesture," in *Proceedings of the ACM symposium on Virtual reality software and technology*, 1998, pp. 59–65.

[9] R. A. Bolt, ""put-that-there" voice and gesture at the graphics interface," in *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, 1980, pp. 262–270.

[10] X. Li, Z. Zhou, W. Liu, and M. Ji, "Wireless semg-based identification in a virtual reality environment," *Microelectronics Reliability*, vol. 98, pp. 78–85, 2019. DOI: 10.1016/j.microrel.2019.04.007.

[11] T.-H. Tsai, C.-C. Huang, and K.-L. Zhang, "Design of hand gesture recognition system for human-computer interaction," *Multimedia tools and applications*, vol. 79, no. 9, pp. 5989–6007, 2020. DOI: 10.1109/TSMCC.2011.2161077.

[12] Q. Gao, J. Liu, and Z. Ju, "Hand gesture recognition using multimodal data fusion and multiscale parallel convolutional neural network for human–robot interaction," *Expert Systems*, vol. 38, no. 5, e12490, 2021. DOI: 10.1111/exsy.12490.

[13] D. Dai, W. Zhuang, Y. Shen, L. Li, and H. Wang, "Design of intelligent mobile robot control system based on gesture recognition," in *International Conference on Artificial Intelligence and Security*, Springer, 2020, pp. 101–111. DOI: 10.1007/978-3-319-22879-2_53.

[14] X. Li, "Human–robot interaction based on gesture and movement recognition," *Signal Processing: Image Communication*, vol. 81, p. 115 686, 2020. DOI: 10.1016/j.image.2019.115686.

[15] M. Meghana, C. U. Kumari, J. S. Priya, *et al.*, "Hand gesture recognition and voice controlled robot," *Materials Today: Proceedings*, vol. 33, pp. 4121–4123, 2020. DOI: 10.1016/j.matpr.2020.06.553.

[16] G. Airò Farulla, D. Pianu, M. Cempini, *et al.*, "Vision-based pose estimation for robot-mediated hand telerehabilitation," *Sensors*, vol. 16, no. 2, p. 208, 2016. DOI: 10.3390/s16020208.

[17] M. Yeasin and S. Chaudhuri, "Visual understanding of dynamic hand gestures," *Pattern Recognition*, vol. 33, no. 11, pp. 1805–1817, 2000. DOI: 10.1016/S0031-3203(99)00175-2.

[18] S. M. Pool, J. M. Hoyle, L. A. Malone, *et al.*, "Navigation of a virtual exercise environment with microsoft kinect by people post-stroke or with cerebral palsy," *Assistive Technology*, vol. 28, no. 4, pp. 225–232, 2016. DOI: 10.1080/10400435.2016.1167789.

[19] F. V. Tenore, A. Ramos, A. Fahmy, S. Acharya, R. Etienne-Cummings, and N. V. Thakor, "Decoding of individuated finger movements using surface electromyography," *IEEE transactions on biomedical engineering*, vol. 56, no. 5, pp. 1427–1434, 2008. DOI: 10.1109/TBME.2008.2005485.

[20] T. Lorrain, N. Jiang, and D. Farina, "Surface emg classification during dynamic contractions for multifunction transradial prostheses," in *2010 annual international conference of the IEEE engineering in medicine and biology*, IEEE, 2010, pp. 2766–2769. DOI: 10.1109/IEMBS.2010.5626587.

[21] M. Atzori, A. Gijsberts, I. Kuzborskij, *et al.*, "Characterization of a benchmark database for myoelectric movement classification," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 23, no. 1, pp. 73–83, 2014. DOI: 10.1109/TNSRE.2014.2328495.

[22] D. P. Azari, Y. H. Hu, B. L. Miller, B. V. Le, and R. G. Radwin, "Using surgeon hand motions to predict surgical maneuvers," *Human factors*, vol. 61, no. 8, pp. 1326–1339, 2019. DOI: 10.1177/0018720819838901.

[23] M. J. Cheok, Z. Omar, and M. H. Jaward, "A review of hand gesture and sign language recognition techniques," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 1, pp. 131–153, 2019. DOI: 10.1007/s13042-017-0705-5.

[24] T. Allevard, E. Benoit, and L. Foulloy, "Hand posture recognition with the fuzzy glove," in *Modern Information Processing*, Elsevier, 2006, pp. 417–427. DOI: 10.1016/B978-044452075-3/50035-2.

[25] P. Garg, N. Aggarwal, and S. Sofat, "Vision based hand gesture recognition," *International Journal of Computer and Information Engineering*, vol. 3, no. 1, pp. 186–191, 2009.

[26] H.-Y. Lai, H.-Y. Ke, and Y.-C. Hsu, "Real-time hand gesture recognition system and application," *Sensors and Materials*, vol. 30, no. 4, pp. 869–884, 2018. DOI: 10.18494/SAM.2018.1790.

[27] I. U. Rehman, S. Ullah, D. Khan, *et al.*, "Fingertip gestures recognition using leap motion and camera for interaction with virtual environment," *Electronics*, vol. 9, no. 12, p. 1986, 2020. DOI: 10.3390/electronics9121986.

[28] Z. Ren, J. Meng, and J. Yuan, "Depth camera based hand gesture recognition and its applications in human-computer-interaction," in *2011 8th international conference on information, communications & signal processing*, IEEE, 2011, pp. 1–5. DOI: 10.1109/ICICS.2011.6173545.

[29] http://www.stroke.org/understand-stroke/what-stroke, [Accessed: July.12, 2017].

[30] P. Langhorne, J. Bernhardt, and G. Kwakkel, "Stroke rehabilitation," *The Lancet*, vol. 377, no. 9778, pp. 1693–1702, 2011. DOI: 10.1016/S0140-6736(11)60325-5.

[31] https://www.cdc.gov/stroke/facts.htm, [Accessed: Sept.4, 2022] 2022/9/4.

[32] S. C. Cramer, G. Nelles, R. R. Benson, *et al.*, "A functional mri study of subjects recovered from hemiparetic stroke," *Stroke*, vol. 28, no. 12, pp. 2518–2527, 1997. DOI: 10.1161/01.STR.28.12.2518.

[33] S. M. Hatem, G. Saussez, M. Della Faille, *et al.*, "Rehabilitation of motor function after stroke: A multiple systematic review focused on techniques to stimulate upper extremity recovery," *Frontiers in human neuroscience*, vol. 10, p. 442, 2016. DOI: 10.3389/fnhum.2016.00442.

[34] http://www.strokeassociation.org/STROKEORG/AboutStroke/Impact-of-Stroke-Stroke-statistics/ _UCM/_310728/_Article.jsp#.WNPkhvnytAh, [Accessed: July.12,2017].

[35] M. T. Duruoz, *Hand function*. Springer, 2016. DOI: 10.1007/978-3-030-17000-4_3.

[36] https://www.neurorehabdirectory.com/treatment-interventions/arm-and-hand-recovery/, [Accessed: July.4,2017].

[37] O. H. Gündüz and C. Ş. Toprak, "Hand function in stroke," in *Hand Function*, Springer, 2019, pp. 125–135. DOI: 10.1007/978-1-4614-9449-2_8.

[38] S. Demain, R. Wiles, L. Roberts, and K. McPherson, "Recovery plateau following stroke: Fact or fiction?" *Disability and rehabilitation*, vol. 28, no. 13-14, pp. 815–821, 2006. DOI: 10.1080/09638280500534796.

[39] S. Lennon, "Physiotherapy practice in stroke rehabilitation: A survey," *Disability and rehabilitation*, vol. 25, no. 9, pp. 455–461, 2003. DOI: 10.1080/0963828031000069744.

[40] S. J. Page, D. R. Gater, and P. Bach-y-Rita, "Reconsidering the motor recovery plateau in stroke rehabilitation," *Archives of physical medicine and rehabilitation*, vol. 85, no. 8, pp. 1377–1381, 2004. DOI: 10.1016/j.apmr.2003.12.031.

[41] K. Matheus and A. M. Dollar, "Benchmarking grasping and manipulation: Properties of the objects of daily living," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2010, pp. 5020–5027. DOI: 10.1109/IROS.2010.5649517.

[42] S. Katz, "Assessing self-maintenance: Activities of daily living, mobility, and instrumental activities of daily living.," *Journal of the American Geriatrics Society*, 1983. DOI: 10.1111/j.1532-5415.1983.tb03391.x.

[43] A. M. Dollar, "Classifying human hand use and the activities of daily living," in *The Human Hand as an Inspiration for Robot Hand Development*, Springer, 2014, pp. 201–216. DOI: 10.1007/978-3-319-03017-3_10.

[44] M. P. Lawton and E. M. Brody, "Assessment of older people: Self-maintaining and instrumental activities of daily living," *The gerontologist*, vol. 9, no. 3_Part_1, pp. 179–186, 1969.

[45] M. Vergara, J. L. Sancho-Bru, V. Gracia-Ibáñez, and A. Pérez-González, "An introductory study of common grasps used by adults during performance of activities of daily living," *Journal of Hand Therapy*, vol. 27, no. 3, pp. 225–234, 2014. DOI: `10.1016/j.jht.2014.04.002`.

[46] H. Sharif, S. B. Seo, and T. K. Kesavadas, "Hand gesture recognition using surface electromyography," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 2020, pp. 682–685. DOI: `10.1109/EMBC44109.2020.9175770`.

[47] J. Rowson and A. Yoxall, "Hold, grasp, clutch or grab: Consumer grip choices during food container opening," *Applied Ergonomics*, vol. 42, no. 5, pp. 627–633, 2011. DOI: `10.1016/j.apergo.2010.12.001`.

[48] M. R. Cutkosky *et al.*, "On grasp choice, grasp models, and the design of hands for manufacturing tasks.," *IEEE Transactions on robotics and automation*, vol. 5, no. 3, pp. 269–279, 1989.

[49] https://www.ultraleap.com/product/leap-motion-controller/, [Accessed: July.12,2022].

[50] https://developerblog.myo.com/, [Accessed: July.4,2017].

[51] T. Guzsvinecz, V. Szucs, and C. Sik-Lanyi, "Suitability of the kinect sensor and leap motion controller—a literature review," *Sensors*, vol. 19, no. 5, p. 1072, 2019.

[52] N. Yu, C. Xu, K. Wang, Z. Yang, and J. Liu, "Gesture-based telemanipulation of a humanoid robot for home service tasks," in *2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, IEEE, 2015, pp. 1923–1927. DOI: `10.1109/CYBER.2015.7288241`.

[53] https://www.ultraleap.com/company/news/blog/how-hand-tracking-works/, [Accessed: July.12,2022].

[54] https://developerblog.myo.com/myocraft-logging-imu-and-raw-emg-data/, [Accessed: July.4,2017].

[55] U. Côté-Allard, C. L. Fall, A. Drouin, *et al.*, "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 27, no. 4, pp. 760–771, 2019. DOI: `10.1109/TNSRE.2019.2896269`.

[56] K. Nymoen, M. R. Haugen, and A. R. Jensenius, "Mumyo–evaluating and exploring the myo armband for musical interaction," 2015.

[57] P. Visconti, F. Gaetani, G. A. Zappatore, and P. Primiceri, "Technical features and functionalities of myo armband: An overview on related literature and advanced applications of myoelectric armbands mainly focused on arm prostheses," *International Journal on Smart Sensing and Intelligent Systems*, vol. 11, no. 1, pp. 1–25, 2018. DOI: `10.21307/ijssis-2018-005`.

[58] J. L. Betthauser, J. T. Krall, R. R. Kaliki, M. S. Fifer, and N. V. Thakor, "Stable electromyographic sequence prediction during movement transitions using temporal convolutional networks," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, IEEE, 2019, pp. 1046–1049. DOI: `10.1109/NER.2019.8717169`.

[59] W. Caesarendra, T. Tjahjowidodo, Y. Nico, S. Wahyudati, and L. Nurhasanah, "Emg finger movement classification based on anfis," in *Journal of Physics: conference series*, IOP Publishing, vol. 1007, 2018, p. 012005.

[60] Y. Yamanoi, Y. Ogiri, and R. Kato, "Emg-based posture classification using a convolutional neural network for a myoelectric hand," *Biomedical Signal Processing and Control*, vol. 55, p. 101574, 2020. DOI: `10.1016/j.bspc.2019.101574`.

[61] B. Tveit, "Analyzing behavioral biometrics of handwriting using myo gesture control armband," M.S. thesis, UiT Norges arktiske universitet, 2018.

[62] Z. Zhang, K. Yang, J. Qian, and L. Zhang, "Real-time surface emg pattern recognition for hand gestures based on an artificial neural network," *Sensors*, vol. 19, no. 14, p. 3170, 2019. DOI: 10.3390/s19143170.

[63] M. E. Benalcázar, A. G. Jaramillo, J. A. Zea, A. Paez, and V. H. Andaluz, "Hand gesture recognition using machine learning and the myo armband," *2017 25th European Signal Processing Conference (EUSIPCO)*, pp. 1040–1044, 2017. DOI: 10.23919/EUSIPCO.2017.8081366.

[64] L. Resnik, H. H. Huang, A. Winslow, D. L. Crouch, F. Zhang, and N. Wolk, "Evaluation of emg pattern recognition for upper limb prosthesis control: A case study in comparison with direct myoelectric control," *Journal of neuroengineering and rehabilitation*, vol. 15, no. 1, pp. 1–13, 2018. DOI: 10.1186/s12984-018-0361-3.

[65] F. Vial, P. Kassavetis, S. Merchant, D. Haubenberger, and M. Hallett, "How to do an electrophysiological study of tremor," *Clinical Neurophysiology Practice*, vol. 4, pp. 134–142, 2019. DOI: 10.1016/j.cnp.2019.06.002.

[66] M. Gandolla, S. Ferrante, G. Ferrigno, *et al.*, "Artificial neural network emg classifier for functional hand grasp movements prediction," *Journal of International Medical Research*, vol. 45, no. 6, pp. 1831–1847, 2017. DOI: 10.1177/0300060516656689.

[67] S. Ferguson and G. R. Dunlop, "Grasp recognition from myoelectric signals," in *Proceedings of the Australasian Conference on Robotics and Automation, Auckland, New Zealand*, vol. 1, 2002.

[68] A. Ameri, M. A. Akhaee, E. Scheme, and K. Englehart, "Regression convolutional neural network for improved simultaneous emg control," *Journal of neural engineering*, vol. 16, no. 3, p. 036 015, 2019. DOI: /10.1088/1741-2552/ab0e2e.

[69] W. Yang, D. Yang, Y. Liu, and H. Liu, "Decoding simultaneous multi-dof wrist movements from raw emg signals using a convolutional neural network," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 5, pp. 411–420, 2019. DOI: 10.1109/THMS.2019.2925191.

[70] W. Wei, Q. Dai, Y. Wong, Y. Hu, M. Kankanhalli, and W. Geng, "Surface-electromyography-based gesture recognition by multi-view deep learning," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 10, pp. 2964–2973, 2019. DOI: 10.1109/TBME.2019.2899222.

[71] M. Atzori, A. Gijsberts, C. Castellini, *et al.*, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Scientific data*, vol. 1, no. 1, pp. 1–13, 2014. DOI: 10.1038/sdata.2014.53.

[72] A. Paszke, S. Gross, S. Chintala, *et al.*, "Automatic differentiation in pytorch," 2017.

[73] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, and A. Skodras, "A hilbert curve based representation of semg signals for gesture recognition," in *2019 International Conference on Systems, Signals and Image Processing (IWSSIP)*, IEEE, 2019, pp. 201–206. DOI: 10.1109/IWSSIP.2019.8787290.

[74] M. Zia ur Rehman, A. Waris, S. O. Gilani, *et al.*, "Multiday emg-based classification of hand motions with deep learning techniques," *Sensors*, vol. 18, no. 8, p. 2497, 2018. DOI: 10.3390/s18082497.

[75] http://new.robai.com/assets/Cyton-Gamma-300-Arm-Specifications_2014.pdf, [Accessed: July.12,2022].

[76] *https://www.researchgate.net/publication/338020205_Grasp_Modelling_Classification_From _Demonstrations_Using_Depth_Measurement_of_Fingertip_Pose/citations*, [Accessed: Dec.12,2021].

[77] H. Sharif, A. Eslaminia, P. Chembrammel, and T. Kesavadas, "Classification of activities of daily living based on grasp dynamics obtained from a leap motion controller," *Sensors*, vol. 22, no. 21, p. 8273, 2022.

[78] D. G. Alonso, A. Teyseyre, A. Soria, and L. Berdun, "Hand gesture recognition in real world scenarios using approximate string matching," *Multimedia Tools and Applications*, vol. 79, no. 29, pp. 20773–20794, 2020. DOI: `10.1007/s11042-020-08913-7`.

[79] I. A. Stinghen Filho, B. B. Gatto, J. Pio, E. N. Chen, J. M. Junior, and R. Barboza, "Gesture recognition using leap motion: A machine learning-based controller interface," in *Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2016 7th International Conference. IEEE*, 2016.

[80] C.-H. Chuan, E. Regina, and C. Guardino, "American sign language recognition using leap motion sensor," in *2014 13th International Conference on Machine Learning and Applications*, IEEE, 2014, pp. 541–544.

[81] T.-W. Chong and B.-G. Lee, "American sign language recognition using leap motion controller with machine learning approach," *Sensors*, vol. 18, no. 10, p. 3554, 2018. DOI: `10.3390/s18103554`.

[82] M. Mohandes, S. Aliyu, and M. Deriche, "Arabic sign language recognition using the leap motion controller," in *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, IEEE, 2014, pp. 960–965. DOI: `10.1109/ISIE.2014.6864742`.

[83] B. Hisham and A. Hamouda, "Arabic static and dynamic gestures recognition using leap motion.," *J. Comput. Sci.*, vol. 13, no. 8, pp. 337–354, 2017. DOI: `10.3844/jcssp.2017.337.354`.

[84] A. Elons, M. Ahmed, H. Shedid, and M. Tolba, "Arabic sign language recognition using leap motion sensor," in *2014 9th International Conference on Computer Engineering & Systems (ICCES)*, IEEE, 2014, pp. 368–373. DOI: `10.1109/ICCES.2014.7030987`.

[85] B. Hisham and A. Hamouda, "Arabic sign language recognition using ada-boosting based on a leap motion controller," *International Journal of Information Technology*, vol. 13, no. 3, pp. 1221–1234, 2021. DOI: `10.1007/s41870-020-00518-5`.

[86] P. Karthick, N. Prathiba, V. Rekha, and S. Thanalaxmi, "Transforming indian sign language into text using leap motion," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 4, p. 5, 2014.

[87] P. Kumar, H. Gauba, P. P. Roy, and D. P. Dogra, "A multimodal framework for sensor based sign language recognition," *Neurocomputing*, vol. 259, pp. 21–38, 2017. DOI: `10.1016/j.neucom.2016.08.132`.

[88] P. Kumar, R. Saini, S. K. Behera, D. P. Dogra, and P. P. Roy, "Real-time recognition of sign language gestures and air-writing using leap motion," in *2017 Fifteenth IAPR international conference on machine vision applications (MVA)*, IEEE, 2017, pp. 157–160. DOI: `10.23919/MVA.2017.7986825`.

[89] D. Zhi, T. E. A. de Oliveira, V. P. da Fonseca, and E. M. Petriu, "Teaching a robot sign language using vision-based hand gesture recognition," in *2018 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, IEEE, 2018, pp. 1–6. DOI: `10.1109/CIVEMSA.2018.8439952`.

[90] A. Anwar, A. Basuki, R. Sigit, A. Rahagiyanto, and M. Zikky, "Feature extraction for indonesian sign language (sibi) using leap motion controller," in *2017 21st International Computer Science and Engineering Conference (ICSEC)*, IEEE, 2017, pp. 1–5. DOI: `10.1109/ICSEC.2017.8443926`.

[91] L. O. R. Nájera, M. L. Sánchez, J. G. G. Serna, R. P. Tapia, and J. Y. A. Llanes, "Recognition of mexican sign language through the leap motion controller," in *Proceedings of the International Conference on Scientific Computing (CSC)*, The Steering Committee of The World Congress in Computer Science, Computer . . ., 2016, p. 147.

[92] M. Simos and N. Nikolaidis, "Greek sign language alphabet recognition using the leap motion device," in *Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, 2016, pp. 1–4.

[93] L. E. Potter, J. Araullo, and L. Carter, "The leap motion controller: A view on sign language," in *Proceedings of the 25th Australian computer-human interaction conference: augmentation, application, innovation, collaboration*, 2013, pp. 175–178.

[94] M. A. Castañeda, A. M. Guerra, and R. Ferro, "Analysis on the gamification and implementation of leap motion controller in the ied técnico industrial de tocancipá," *Interactive Technology and Smart Education*, 2018. DOI: `10.1108/ITSE-12-2017-0069`.

[95] D. Bassily, C. Georgoulas, J. Guettler, T. Linner, and T. Bock, "Intuitive and adaptive robotic arm manipulation using the leap motion controller," in *ISR/Robotik 2014; 41st International Symposium on Robotics*, VDE, 2014, pp. 1–7.

[96] S. Chen, H. Ma, C. Yang, and M. Fu, "Hand gesture based robot control system using leap motion," in *International Conference on Intelligent Robotics and Applications*, Springer, 2015, pp. 581–591. DOI: `10.1007/978-3-319-22879-2_53`.

[97] U. A. Siddiqui, F. Ullah, A. Iqbal, *et al.*, "Wearable-sensors-based platform for gesture recognition of autism spectrum disorder children using machine learning algorithms," *Sensors*, vol. 21, no. 10, p. 3319, 2021. DOI: `10.3390/s21103319`.

[98] S. Ameur, A. B. Khalifa, and M. S. Bouhlel, "Hand-gesture-based touchless exploration of medical images with leap motion controller," in *2020 17th International multi-conference on systems, signals & devices (SSD)*, IEEE, 2020, pp. 6–11. DOI: `10.1109/SSD49366.2020.9364244`.

[99] A. Karashanov, A. Manolova, and N. Neshov, "Application for hand rehabilitation using leap motion sensor based on a gamification approach," *Int. J. Adv. Res. Sci. Eng*, vol. 5, no. 2, pp. 61–69, 2016.

[100] M. Alimanova, S. Borambayeva, D. Kozhamzharova, *et al.*, "Gamification of hand rehabilitation process using virtual reality tools: Using leap motion for hand rehabilitation," in *2017 First IEEE International Conference on Robotic Computing (IRC)*, IEEE, 2017, pp. 336–339. DOI: `10.1109/IRC.2017.76`.

[101] Z.-r. Wang, P. Wang, L. Xing, L.-p. Mei, J. Zhao, and T. Zhang, "Leap motion-based virtual reality training for improving motor functional recovery of upper limbs and neural reorganization in subacute stroke patients," *Neural regeneration research*, vol. 12, no. 11, p. 1823, 2017. DOI: `10.4103/1673-5374.219043`.

[102] W.-J. Li, C.-Y. Hsieh, L.-F. Lin, and W.-C. Chu, "Hand gesture recognition for post-stroke rehabilitation using leap motion," in *2017 International Conference on Applied System Innovation (ICASI)*, IEEE, 2017, pp. 386–388. DOI: `10.1109/ICASI.2017.7988433`.

[103] A. Škraba, A. Koložvari, D. Kofjač, and R. Stojanović, "Wheelchair maneuvering using leap motion controller and cloud based speech control: Prototype realization," in *2015 4th Mediterranean Conference on Embedded Computing (MECO)*, IEEE, 2015, pp. 391–394. DOI: `10.1109/MECO.2015.7181952`.

[104] T. Travaglini, P. Swaney, K. D. Weaver, and R. Webster III, "Initial experiments with the leap motion as a user interface in robotic endonasal surgery," in *Robotics and mechatronics*, Springer, 2016, pp. 171–179. DOI: `10.1007/978-3-319-22368-1_17`.

[105] W. Qi, S. E. Ovur, Z. Li, A. Marzullo, and R. Song, "Multi-sensor guided hand gesture recognition for a teleoperated robot using a recurrent neural network," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 6039–6045, 2021. DOI: `10.1109/LRA.2021.3089999`.

[106] D. Bachmann, F. Weichert, and G. Rinkenauer, "Review of three-dimensional human-computer interaction with focus on the leap motion controller," *Sensors*, vol. 18, no. 7, p. 2194, 2018. DOI: `10.3390/s18072194`.

[107] R. Nogales and M. Benalcázar, "Real-time hand gesture recognition using the leap motion controller and machine learning," in *2019 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, IEEE, 2019, pp. 1–7. DOI: `10.1109/LA-CCI47412.2019.9037037`.

[108] J. Rekha, J. Bhattacharya, and S. Majumder, "Hand gesture recognition for sign language: A new hybrid approach," in *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV)*, The Steering Committee of The World Congress in Computer Science, Computer . . ., 2011, p. 1.

[109] https://www.upperlimbclinics.co.uk/images/hand-anatomy-pic.jpg, [Accessed: July.12,2022].

[110] https://developer-archive.leapmotion.com/documentation/python/devguide/Leap_Overview.html , [Accessed: July.12,2022].

[111] *https://developer-archive.leapmotion.com/documentation/csharp/devguide/Leap_Coordinate_Mapping. html #:~text=Leap%20Motion%20Coordinates,10cm%2C%20z%20%3D%20%2D10cm , [Accessed: July.12,2022].*

[112] J. J. Craig, *Introduction to robotics: mechanics and control*. Pearson Educacion, 2005.

[113] *Change of basis*, https://math.hmc.edu/calculus/hmc-mathematics-calculus-online-tutorials/linear-algebra/change-of-basis.

[114] K. Patel, "A review on feature extraction methods," *Int. J. Adv. Res. Electr. Electron. Instrum. Eng*, vol. 5, pp. 823–827, 2016. DOI: `10.15662/IJAREEIE.2016.0502034`.

[115] W. Lu, Z. Tong, and J. Chu, "Dynamic hand gesture recognition with leap motion controller," *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1188–1192, 2016. DOI: `10.1109/LSP.2016.2590470`.

[116] Q. Yang, W. Ding, X. Zhou, D. Zhao, and S. Yan, "Leap motion hand gesture recognition based on deep neural network," in *2020 Chinese Control And Decision Conference (CCDC)*, IEEE, 2020, pp. 2089–2093. DOI: `10.1109/CCDC49329.2020.9164723`.

[117] G. Marin, F. Dominio, and P. Zanuttigh, "Hand gesture recognition with jointly calibrated leap motion and depth sensor," *Multimedia Tools and Applications*, vol. 75, no. 22, pp. 14 991–15 015, 2016. DOI: `10.1007/s11042-015-2451-6`.

[118] D. Avola, M. Bernardi, L. Cinque, G. L. Foresti, and C. Massaroni, "Exploiting recurrent neural networks and leap motion controller for the recognition of sign language and semaphoric hand gestures," *IEEE Transactions on Multimedia*, vol. 21, no. 1, pp. 234–245, 2018. DOI: `10.1109TMM.2018.2856094`.

[119] R. Fonk, S. Schneeweiss, U. Simon, and L. Engelhardt, "Hand motion capture from a 3d leap motion controller for a musculoskeletal dynamic simulation," *Sensors*, vol. 21, no. 4, p. 1199, 2021. DOI: `10.3390/s21041199`.

[120] I. M. Khairuddin, S. N. Sidek, A. P. A. Majeed, M. A. M. Razman, A. A. Puzi, and H. M. Yusof, "The classification of movement intention through machine learning models: The identification of significant time-domain emg features," *PeerJ Computer Science*, vol. 7, e379, 2021. DOI: `10.7717/peerj-cs.379`.

[121] S. Abbaspour, M. Lindén, H. Gholamhosseini, A. Naber, and M. Ortiz-Catalan, "Evaluation of surface emg-based recognition algorithms for decoding hand movements," *Medical & biological engineering & computing*, vol. 58, no. 1, pp. 83–100, 2020. DOI: `10.1682/JRRD.2010.09.0177`.

[122] Kehtarnavaz, N.; Mahotra, S. Digital Signal Processing Laboratory: LabVIEW-Based FPGA Implementation; Universal-Publishers, 465 2010.

[123] B. Kumar and M. Manjunatha, "Performance analysis of knn, svm and ann techniques for gesture recognition system," *Indian Journal of Science and Technology*, vol. 9, no. 1, pp. 1–8, 2016. DOI: `10.17485/ijst/2017/v9iS1/111145`.

[124] J. Huo, K. L. Keung, C. K. Lee, and H. Y. Ng, "Hand gesture recognition with augmented reality and leap motion controller," in *2021 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, IEEE, 2021, pp. 1015–1019. DOI: `10.1109/IEEM50564.2021.9672611`.

[125] F. Li, Y. Li, B. Du, H. Xu, H. Xiong, and M. Chen, "A gesture interaction system based on improved finger feature and we-knn," in *Proceedings of the 2019 4th International Conference on Mathematics and Artificial Intelligence*, 2019, pp. 39–43. DOI: `10.1145/3325730.3325759`.

[126] S. Sumpeno, I. G. A. Dharmayasa, S. M. S. Nugroho, and D. Purwitasari, "Immersive hand gesture for virtual museum using leap motion sensor based on k-nearest neighbor," in *2019 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, IEEE, 2019, pp. 1–6. DOI: `10.1109/CENIM48368.2019.8973273`.

[127] I. Ding Jr and M.-C. Hsieh, "A hand gesture action-based emotion recognition system by 3d image sensor information derived from leap motion sensors for the specific group with restlessness emotion problems," *Microsystem Technologies*, pp. 1–13, 2020. DOI: `10.1007/s00542-019-04503-2`.

[128] R. Nogales and M. Benalcázar, "Real-time hand gesture recognition using knn-dtw and leap motion controller," in *Conference on Information and Communication Technologies of Ecuador*, Springer, 2020, pp. 91–103. DOI: `10.1145/3357160.3357670`.

[129] https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html, [Accessed: June.30,2022].

[130] M. Sha'Abani, N. Fuad, N. Jamal, and M. Ismail, "Knn and svm classification for eeg: A review," *InECCE2019*, pp. 555–565, 2020. DOI: `10.1007/978-981-15-2317-5_47`.

[131] https://pytorch.org/, [Accessed: June.30,2022].

[132]   K. Kritsis, M. Kaliakatsos-Papakostas, V. Katsouros, and A. Pikrakis, "Deep convolutional and lstm neural network architectures on leap motion hand tracking data sequences," in *2019 27th European Signal Processing Conference (EUSIPCO)*, IEEE, 2019, pp. 1–5. DOI: `10.23919/EUSIPCO.2019.8902973`.

[133]   C. R. Naguri and R. C. Bunescu, "Recognition of dynamic hand gestures from 3d motion data using lstm and cnn architectures," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, IEEE, 2017, pp. 1130–1133. DOI: `10.1109/ICMLA.2017.00013`.

[134]   K. Lupinetti, A. Ranieri, F. Giannini, and M. Monti, "3d dynamic hand gestures recognition using the leap motion sensor and convolutional neural networks," in *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, Springer, 2020, pp. 420–439. DOI: `10.1101/740548`.

[135]   A. Ikram and Y. Liu, "Skeleton based dynamic hand gesture recognition using lstm and cnn," in *2020 2nd International Conference on Image Processing and Machine Vision*, 2020, pp. 63–68. DOI: `10.1145/3421558.3421568`.

[136]   A. of Surgical Research, "Guidelines for training in surgical research with animals," *Journal of Investigative Surgery*, vol. 22, no. 3, pp. 218–225, 2009. DOI: `10.1080/08941930902904542`.

[137]   S. W. Baran, E. J. Johnson, and J. Kehler, "An introduction to electronic learning and its use to address challenges in surgical training," *Lab animal*, vol. 38, no. 6, pp. 202–210, 2009. DOI: `10.1038/laban0609-202`.

[138]   S. W. Baran, E. J. Johnson, M. A. Stephens, and J. Kehler, "Development of electronic learning courses for surgical training of animal research personnel," *Lab animal*, vol. 38, no. 9, pp. 295–304, 2009. DOI: `10.1038/laban0909-295`.

[139]   T. Kesavadas, A. Kamerkar, and A. Anand, *Touch-Based Interactive NURBS Modeler Using a Force/Position Input Glove, Virtual Engineering, Chap 2*. Momentum Press, 2010.

[140]   T. Kesavadas and K. Guru, "Role of simulators in robotic surgery," *Pediatric Robotic and Reconstructive Urology: A Comprehensive Guide*, pp. 19–28, 2011. DOI: `10.1002/9781444345292.ch3`.

[141]   S. S. Y. Tan and S. K. Sarker, "Simulation in surgery: A review," *Scottish medical journal*, vol. 56, no. 2, pp. 104–109, 2011. DOI: `10.1258/smj.2011.011098`.

[142]   S. R. Dawe, J. A. Windsor, J. A. Broeders, P. C. Cregan, P. J. Hewett, and G. J. Maddern, "A systematic review of surgical skills transfer after simulation-based training: Laparoscopic cholecystectomy and endoscopy," *Annals of surgery*, vol. 259, no. 2, pp. 236–248, 2014. DOI: `10.1097/SLA.0000000000000245`.

[143]   W. R. Foshay and P. T. Tinkey, "Evaluating the effectiveness of training strategies: Performance goals and testing," *ILAR journal*, vol. 48, no. 2, pp. 156–162, 2007. DOI: `10.1093/ilar.48.2.156`.

[144]   *https://www.surgery.wisc.edu/wp-content/uploads/2018/12/20180908_133248-1200×900.jpg*, [Accessed: Nov.2, 2019].

[145]   https://www.vrmagic.com/medical-simulators/eyesi-surgical, [Accessed: Nov.2, 2019].

[146]   S. Takazawa, T. Ishimaru, K. Harada, *et al.*, "Pediatric thoracoscopic surgical simulation using a rapid-prototyped chest model and motion sensors can better identify skilled surgeons than a conventional box trainer," *Journal of Laparoendoscopic & Advanced Surgical Techniques*, vol. 26, no. 9, pp. 740–747, 2016. DOI: `10.1089/lap.2016.0131`.

[147] S. Yokoyama, K. Mizunuma, Y. Kurashima, *et al.*, "Evaluation methods and impact of simulation-based training in pediatric surgery: A systematic review," *Pediatric Surgery International*, vol. 35, no. 10, pp. 1085–1094, 2019. DOI: 10.1007/s00383-019-04539-5.

[148] T. S. Kashikar, T. F. Kerwin, A. C. Moberly, and G. J. Wiet, "A review of simulation applications in temporal bone surgery," *Laryngoscope investigative otolaryngology*, vol. 4, no. 4, pp. 420–424, 2019. DOI: 10.1002/lio2.277.

[149] C. L. Temple and D. C. Ross, "A new, validated instrument to evaluate competency in microsurgery: The university of western ontario microsurgical skills acquisition/assessment instrument [outcomes article]," *Plastic and reconstructive surgery*, vol. 127, no. 1, pp. 215–222, 2011. DOI: 10.1097/PRS.0b013e3181f95adb.

[150] W. Chan, N. Niranjan, and V. Ramakrishnan, "Structured assessment of microsurgery skills in the clinical setting," *Journal of Plastic, Reconstructive & Aesthetic Surgery*, vol. 63, no. 8, pp. 1329–1334, 2010. DOI: 10.1016/j.bjps.2009.06.024.

[151] J. Martin, G. Regehr, R. Reznick, *et al.*, "Objective structured assessment of technical skill (osats) for surgical residents," *British journal of surgery*, vol. 84, no. 2, pp. 273–278, 1997.

[152] P. U. Kalu, J. Atkins, D. Baker, C. J. Green, and P. E. Butler, "How do we assess microsurgical skill?" *Microsurgery: Official Journal of the International Microsurgical Society and the European Federation of Societies for Microsurgery*, vol. 25, no. 1, pp. 25–29, 2005. DOI: 10.1002/micr.20078.

[153] A. Ghanem, A. E. Kim, M. Singh, and S. Myers, "Trauma'12 winning abstract–end product assessment in microsurgery simulation training–a new insider view," *Annals of Medicine and Surgery*, vol. 2, no. 1, p. 36, 2013. DOI: 10.1016/S2049-0801(13)70027-X.

[154] E. D. Grober, S. J. Hamstra, K. R. Wanzel, *et al.*, "Validation of novel and objective measures of microsurgical skill: Hand-motion analysis and stereoscopic visual acuity," *Microsurgery*, vol. 23, no. 4, pp. 317–322, 2003. DOI: 10.1002/micr.10152.

[155] C.-A. E. Moulton, A. Dubrowski, H. MacRae, B. Graham, E. Grober, and R. Reznick, "Teaching surgical skills: What kind of practice makes perfect?: A randomized, controlled trial," *Annals of surgery*, vol. 244, no. 3, p. 400, 2006. DOI: 10.1097/01.sla.0000234808.85789.6a.

[156] G. M. Saleh, V. Gauba, D. Sim, D. Lindfield, M. Borhani, and S. Ghoussayni, "Motion analysis as a tool for the evaluation of oculoplastic surgical skill: Evaluation of oculoplastic surgical skill," *Archives of Ophthalmology*, vol. 126, no. 2, pp. 213–216, 2008. DOI: 10.1001/archophthalmol.2007.62.

[157] D. G. Ezra, R. Aggarwal, M. Michaelides, *et al.*, "Skills acquisition and assessment after a microsurgical skills course for ophthalmology residents," *Ophthalmology*, vol. 116, no. 2, pp. 257–262, 2009. DOI: 10.1016/j.ophtha.2008.09.038.

[158] E. D. Grober, M. Roberts, E.-J. Shin, M. Mahdi, and V. Bacal, "Intraoperative assessment of technical skills on live patients using economy of hand motion: Establishing learning curves of surgical competence," *The American journal of surgery*, vol. 199, no. 1, pp. 81–85, 2010. DOI: 10.1016/j.amjsurg.2009.07.033.

[159] G. Forestier, F. Petitjean, P. Senin, *et al.*, "Surgical motion analysis using discriminative interpretable patterns," *Artificial intelligence in medicine*, vol. 91, pp. 3–11, 2018. DOI: 10.1016/j.artmed.2018.08.002.

[160] M. Uemura, M. Tomikawa, T. Miao, *et al.*, "Feasibility of an ai-based measure of the hand motions of expert and novice surgeons," *Computational and mathematical methods in medicine*, vol. 2018, 2018. DOI: `10.1155/2018/9873273`.

[161] D. P. Azari, *Quantifying surgical skill*. The University of Wisconsin-Madison, 2018.

[162] I. Badash, K. Burtt, C. A. Solorzano, and J. N. Carey, "Innovations in surgery simulation: A review of past, current and future techniques," *Annals of translational medicine*, vol. 4, no. 23, 2016. DOI: `10.21037/atm.2016.12.24`.

[163] N. E. Seymour, A. G. Gallagher, S. A. Roman, *et al.*, "Virtual reality training improves operating room performance: Results of a randomized, double-blinded study," *Annals of surgery*, vol. 236, no. 4, p. 458, 2002. DOI: `10.1097/00000658-200210000-00008`.

[164] M. Wilson, A. Middlebrook, C. Sutton, R. Stone, and R. McCloy, "Mist vr: A virtual reality trainer for laparoscopic surgery assesses performance.," *Annals of the Royal College of Surgeons of England*, vol. 79, no. 6, p. 403, 1997.

[165] M. A. Roa and R. Suárez, "Grasp quality measures: Review and performance," *Autonomous robots*, vol. 38, no. 1, pp. 65–88, 2015. DOI: `10.1007/s10514-014-9402-3`.

[166] A. Gardner, C. A. Duncan, J. Kanno, and R. Selmic, "3d hand posture recognition from small unlabeled point sets," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, 2014, pp. 164–169. DOI: `10.1109/SMC.2014.6973901`.

[167] S. Li, X. Ma, H. Liang, *et al.*, "Vision-based teleoperation of shadow dexterous hand using end-to-end deep neural network," in *2019 International Conference on Robotics and Automation (ICRA)*, IEEE, 2019, pp. 416–422. DOI: `10.1109/ICRA.2019.8794277`.

[168] B. K. Chakraborty, D. Sarma, M. K. Bhuyan, and K. F. MacDorman, "Review of constraints on vision-based gesture recognition for human–computer interaction," *IET Computer Vision*, vol. 12, no. 1, pp. 3–15, 2018. DOI: `10.1049/iet-cvi.2017.0052`.

[169] S. Sridhar, F. Mueller, M. Zollhöfer, D. Casas, A. Oulasvirta, and C. Theobalt, "Real-time joint tracking of a hand manipulating an object from rgb-d input," in *European Conference on Computer Vision*, Springer, 2016, pp. 294–310. DOI: `10.1007/978-3-319-46475-6_19`.

[170] H. Liang, J. Yuan, D. Thalmann, and Z. Zhang, "Model-based hand pose estimation via spatial-temporal hand parsing and 3d fingertip localization," *The Visual Computer*, vol. 29, no. 6, pp. 837–848, 2013. DOI: `10.1007/s00371-013-0822-4`.

[171] https://www.mcmaster.com/47065T102/ [Accessed: July.10, 2022].

[172] https://azure.microsoft.com/en-us/services/kinect-dk/#overview, [Accessed: July.7, 2022].

[173] https://azure.microsoft.com/en-us/services/kinect-dk/#industries, [Accessed: July.7,2022].

[174] *https://snakeclamp.com/collections/attachments/products/three-finger-clamp-2, [Accessed: July.5,2021].*

[175] *https://snakeclamp.com/collections/flexible-gooseneck-and-rigid-arms/products/24-heavy-duty -flexible-gooseneck-arm-chrome, [Accessed: July.5,2021].*

[176] *https://snakeclamp.com/collections/flexible-gooseneck-and-rigid-arms/products/12-rigid-arm-chrome, [Accessed: July.5,2021].*

[177]   https://snakeclamp.com/collections/mounting-bases/products/table-clamp, [Accessed: July.5,2021].

[178]   N. Nazmi, M. A. Abdul Rahman, S.-I. Yamamoto, S. A. Ahmad, H. Zamzuri, and S. A. Mazlan, "A review of classification techniques of emg signals during isotonic and isometric contractions," *Sensors*, vol. 16, no. 8, p. 1304, 2016. DOI: 10.3390/s16081304.

[179]   W. M. B. W. Daud, A. B. Yahya, C. S. Horng, M. F. Sulaima, and R. Sudirman, "Features extraction of electromyography signals in time domain on biceps brachii muscle," *International Journal of Modeling and Optimization*, vol. 3, no. 6, p. 515, 2013. DOI: 10.7763/IJMO.2013.V3.332.

[180]   https://training.seer.cancer.gov/anatomy/muscular/structure.html, [Accessed: Sept.15, 2022].

[181]   H. A. Javaid, M. I. Tiwana, A. Alsanad, *et al.*, "Classification of hand movements using myo armband on an embedded platform," *Electronics*, vol. 10, no. 11, p. 1322, 2021. DOI: 10.3390/electronics10111322.

[182]   M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of emg signal analysis: Detection, processing, classification and applications," *Biological procedures online*, vol. 8, no. 1, pp. 11–35, 2006. DOI: 10.1251/bpo115.

[183]   D. Yang, J. Zhao, Y. Gu, L. Jiang, and H. Liu, "Estimation of hand grasp force based on forearm surface emg," in *2009 International Conference on Mechatronics and Automation*, IEEE, 2009, pp. 1795–1799. DOI: 10.1109/ICMA.2009.5246102.

[184]   M. DiCicco, L. Lucas, and Y. Matsuoka, "Comparison of control strategies for an emg controlled orthotic exoskeleton for the hand," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, ieee, vol. 2, 2004, pp. 1622–1627. DOI: 10.1109/ROBOT.2004.1308056.

[185]   I. B. Abdallah, Y. Bouteraa, and C. Rekik, "Design and development of 3d printed myoelectric robotic exoskeleton for hand rehabilitation.," *International Journal on Smart Sensing & Intelligent Systems*, vol. 10, no. 2, 2017.

[186]   C. Sapsanis, G. Georgoulas, and A. Tzes, "Emg based classification of basic hand movements based on time-frequency features," in *21st Mediterranean conference on control and automation*, IEEE, 2013, pp. 716–722. DOI: 10.1109/MED.2013.6608802.

[187]   J. Vogel, C. Castellini, and P. van der Smagt, "Emg-based teleoperation and manipulation with the dlr lwr-iii," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2011, pp. 672–678. DOI: 10.1109/IROS.2011.6094739.

[188]   Y. Huang, K. Low, A. H. McGregor, and K. Kong, "Clinical-based engineering assessment and data interpretation of hand strength for task-oriented robotic rehabilitation," *Advanced Robotics*, vol. 25, no. 15, pp. 1991–2018, 2011. DOI: 10.1163/016918611X588871.

[189]   S. Bao and B. Silverstein, "Estimation of hand force in ergonomic job evaluations," *Ergonomics*, vol. 48, no. 3, pp. 288–301, 2005. DOI: 10.1080/0014013042000327724.

[190]   R. Gurram, S. Rakheja, and G. Gouw, "A study of hand grip pressure distribution and emg of finger flexor muscles under dynamic loads," *Ergonomics*, vol. 38, no. 4, pp. 684–699, 1995. DOI: 10.1080/00140139508925140.

[191]   S.-W. Shin, S.-H. Jeong, and S.-T. Chung, "Evaluation of hand grip strength and emg signal on visual reaction," *Korean Journal of Sport Biomechanics*, vol. 24, no. 2, pp. 161–166, 2014. DOI: 10.5103/KJSB.2014.24.2.161.

[192] R. Barański and A. Kozupa, "Hand grip-emg muscle response.," *Acta Physica Polonica, A.*, vol. 125, 2014.

[193] A. B. Ajiboye and R. Weir, "Muscle synergies as a predictive framework for the emg patterns of new hand postures," *Journal of neural engineering*, vol. 6, no. 3, p. 036 004, 2009. DOI: `10.1088/1741-2560/6/3/036004`.

[194] G. Lange, C. Y. Low, K. Johar, F. A. Hanapiah, and F. Kamaruzaman, "Classification of electroencephalogram data from hand grasp and release movements for bci controlled prosthesis," *Procedia Technology*, vol. 26, pp. 374–381, 2016. DOI: `10.1016/j.protcy.2016.08.048`.

[195] A. H. Oskouei, M. G. Paulin, and A. B. Carman, "Intra-session and inter-day reliability of forearm surface emg during varying hand grip forces," *Journal of Electromyography and Kinesiology*, vol. 23, no. 1, pp. 216–222, 2013. DOI: `10.1016/j.jelekin.2012.08.011`.

[196] S. B. Akben, "Classification of hand movements related to grasp by using emg signals," in *2015 19th National Biomedical Engineering Meeting (BIYOMUT)*, IEEE, 2015, pp. 1–4. DOI: `10.1109/BIYOMUT.2015.7369445`.

[197] G. Ouyang, X. Zhu, Z. Ju, and H. Liu, "Dynamical characteristics of surface emg signals of hand grasps via recurrence plot," *IEEE journal of biomedical and health informatics*, vol. 18, no. 1, pp. 257–265, 2013. DOI: `10.1109/JBHI.2013.2261311`.

[198] N. Ho, K. Tong, X. Hu, *et al.*, "An emg-driven exoskeleton hand robotic training device on chronic stroke subjects: Task training system for stroke rehabilitation," in *2011 IEEE international conference on rehabilitation robotics*, IEEE, 2011, pp. 1–5. DOI: `10.1109/ICORR.2011.5975340`.

[199] C. W. Hess and S. L. Pullman, "Tremor: Clinical phenomenology and assessment techniques," *Tremor and other hyperkinetic movements*, vol. 2, 2012. DOI: `10.7916/D8WM1C41`.

[200] F. J. Pérez-Duarte, F. M. Sánchez-Margallo, I. D.-G. Martıén-Portugués, *et al.*, "Ergonomic analysis of muscle activity in the forearm and back muscles during laparoscopic surgery: Influence of previous experience and performed task," *Surgical Laparoscopy Endoscopy & Percutaneous Techniques*, vol. 23, no. 2, pp. 203–207, 2013. DOI: `10.1097/SLE.0b013e3182827f30`.

[201] A. K. Mukhopadhyay and S. Samui, "An experimental study on upper limb position invariant emg signal classification based on deep neural network," *Biomedical signal processing and control*, vol. 55, 2020. DOI: `10.1016/j.bspc.2019.101669`.

[202] M. Uhrich, R. Underwood, J. Standeven, N. Soper, and J. Engsberg, "Assessment of fatigue, monitor placement, and surgical experience during simulated laparoscopic surgery," *Surgical endoscopy*, vol. 16, no. 4, pp. 635–639, 2002. DOI: `10.1007/s00464-001-8151-5`.

[203] R. Bieck, R. Fuchs, and T. Neumuth, "Surface emg-based surgical instrument classification for dynamic activity recognition in surgical workflows," *Current Directions in Biomedical Engineering*, vol. 5, no. 1, pp. 37–40, 2019. DOI: `10.1515/cdbme-2019-0010`.

[204] N. Nasri, S. Orts-Escolano, F. Gomez-Donoso, and M. Cazorla, "Inferring static hand poses from a low-cost non-intrusive semg sensor," *Sensors*, vol. 19, no. 2, p. 371, 2019. DOI: `10.3390/s19020371`.

[205] https://www.csats.com/osats [Accessed: July.5,2022].

[206]   Kuzminsky, J., Phillips, H., Sarif H,. Moran, C., Gleason, H,. McCoy, AM., Lanzo S., McNeil L., Kesavadas T. "Longitudinal evaluation of student progression in suturing through use of computer models and standardized grading rubrics on a novel skin pad model", Under preparation.

[207]   C. E. Glarner, Y.-Y. Hu, C.-H. Chen, *et al.*, "Quantifying technical skills during open operations using video-based motion analysis," *Surgery*, vol. 156, no. 3, pp. 729–734, 2014. DOI: `10.1016/j.surg.2014.04.054`.

[208]   L. L. Frasier, D. P. Azari, Y. Ma, *et al.*, "A marker-less technique for measuring kinematics in the operating room," *Surgery*, vol. 160, no. 5, pp. 1400–1413, 2016. DOI: `10.1016/j.surg.2016.05.004`.

[209]   A. Ghasemloonia, Y. Maddahi, K. Zareinia, S. Lama, J. C. Dort, and G. R. Sutherland, "Surgical skill assessment using motion quality and smoothness," *Journal of surgical education*, vol. 74, no. 2, pp. 295–305, 2017. DOI: `10.1016/j.jsurg.2016.10.006`.

[210]   R. Aggarwal, T. Grantcharov, K. Moorthy, *et al.*, "An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room," *Annals of surgery*, vol. 245, no. 6, p. 992, 2007.

[211]   A.-L. D. D'Angelo, D. N. Rutherford, R. D. Ray, *et al.*, "Idle time: An underdeveloped performance metric for assessing surgical skill," *The American journal of surgery*, vol. 209, no. 4, pp. 645–651, 2015.

[212]   Z. Ghahramani, "Unsupervised learning," in *Summer school on machine learning*, Springer, 2003, pp. 72–112.

[213]   A. E. Ezugwu, A. M. Ikotun, O. O. Oyelade, *et al.*, "A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects," *Engineering Applications of Artificial Intelligence*, vol. 110, p. 104 743, 2022.

[214]   S. H. Lee, S.-M. Park, S. S. Yeo, *et al.*, "Parkinson's disease subtyping using clinical features and biomarkers: Literature review and preliminary study of subtype clustering," *Diagnostics*, vol. 12, no. 1, p. 112, 2022.

[215]   C. Magoo and M. Singh, "An implemented review for intent creation using different clustering techniques," in *2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT)*, 2022, pp. 83–91. DOI: `10.1109/CCiCT56684.2022.00027`.

[216]   H. C. Yadav and D. Patel, "A review of different routing and clustering protocol for wireless sensor network,"

[217]   A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[218]   M. Ghesmoune, M. Lebbah, and H. Azzag, "State-of-the-art on clustering data streams," *Big Data Analytics*, vol. 1, no. 1, pp. 1–27, 2016.

[219]   M.-U. Asad, R. Mustafa, and M. S. Hossain, "An efficient strategy for face clustering use in video surveillance system," in *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, 2019, pp. 12–17. DOI: `10.1109/ICIEV.2019.8858532`.

[220]   U. Damnjanovic, V. Fernandez, E. Izquierdo, and J. M. Martinez, "Event detection and clustering for surveillance video summarization," in *2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, 2008, pp. 63–66. DOI: `10.1109/WIAMIS.2008.53`.

[221] J. Lin, S.-h. Zhong, and A. Fares, "Deep hierarchical lstm networks with attention for video summarization," *Computers & Electrical Engineering*, vol. 97, p. 107 618, 2022.

[222] K. Soomro and M. Shah, "Unsupervised action discovery and localization in videos," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 696–705. DOI: 10.1109/ICCV.2017.82.

[223] A. Murali, A. Garg, S. Krishnan, *et al.*, "Tsc-dl: Unsupervised trajectory segmentation of multi-modal surgical demonstrations with deep learning," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 4150–4157. DOI: 10.1109/ICRA.2016.7487607.

[224] A. Zia, C. Zhang, X. Xiong, and A. M. Jarc, "Temporal clustering of surgical activities in robot-assisted surgery," *International journal of computer assisted radiology and surgery*, vol. 12, no. 7, pp. 1171–1178, 2017.

[225] A. Zia, Y. Sharma, V. Bettadapura, E. L. Sarin, and I. Essa, "Video and accelerometer-based motion analysis for automated surgical skills assessment," *International journal of computer assisted radiology and surgery*, vol. 13, no. 3, pp. 443–455, 2018.

[226] H. Zhao, J. Xie, Z. Shao, Y. Qu, Y. Guan, and J. Tan, "A fast unsupervised approach for multi-modality surgical trajectory segmentation," *IEEE Access*, vol. 6, pp. 56 411–56 422, 2018.

[227] M. J. Primus, K. Schoeffmann, and L. Böszörmenyi, "Segmentation of recorded endoscopic videos by detecting significant motion changes," in *2013 11th International Workshop on Content-Based Multimedia Indexing (CBMI)*, 2013, pp. 223–228. DOI: 10.1109/CBMI.2013.6576587.

[228] A. P. Twinanda, E. O. Alkan, A. Gangi, M. de Mathelin, and N. Padoy, "Data-driven spatio-temporal rgbd feature encoding for action recognition in operating rooms," *International journal of computer assisted radiology and surgery*, vol. 10, no. 6, pp. 737–747, 2015.

[229] A. Nara, C. Allen, and K. Izumi, "Surgical phase recognition using movement data from video imagery and location sensor data," in *Advances in Geocomputation*, Springer, 2017, pp. 229–237.

[230] M. A. Jamal and O. Mohareri, "Multi-modal unsupervised pre-training for surgical operating room workflow analysis," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2022, pp. 453–463.

[231] B. van Amsterdam, M. J. Clarkson, and D. Stoyanov, "Gesture recognition in robotic surgery: A review," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 6, pp. 2021–2035, 2021. DOI: 10.1109/TBME.2021.3054828.

[232] A. Kukleva, H. Kuehne, F. Sener, and J. Gall, "Unsupervised learning of action classes with continuous temporal embedding," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 12 066–12 074.

[233] B. Peng, J. Lei, H. Fu, Y. Jia, Z. Zhang, and Y. Li, "Deep video action clustering via spatio-temporal feature learning," *Neurocomputing*, vol. 456, pp. 519–527, 2021.

[234] L. Zelnik-Manor and M. Irani, "Statistical analysis of dynamic actions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1530–1535, 2006. DOI: 10.1109/TPAMI.2006.194.

[235] S. H. Khorasgani, Y. Chen, and F. Shkurti, "Slic: Self-supervised learning with iterative clustering for human action videos," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 16 091–16 101.

[236] L. Ding and C. Xu, "Weakly-supervised action segmentation with iterative soft boundary assignment," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6508–6516.

[237] A. Richard, H. Kuehne, A. Iqbal, and J. Gall, "Neuralnetwork-viterbi: A framework for weakly supervised video learning," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2018, pp. 7386–7395.

[238] L. Zappella, B. Béjar, G. Hager, and R. Vidal, "Surgical gesture classification from video and kinematic data," *Medical image analysis*, vol. 17, no. 7, pp. 732–745, 2013.

[239] F. Lalys, L. Riffaud, D. Bouget, and P. Jannin, "A framework for the recognition of high-level surgical tasks from video images for cataract surgeries," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 966–976, 2012. DOI: 10.1109/TBME.2011.2181168.

[240] S. Petscharnig and K. Schöffmann, "Learning laparoscopic video shot classification for gynecological surgery," *Multimedia Tools and Applications*, vol. 77, no. 7, pp. 8061–8079, 2018.

[241] M. Zhang, X. Cheng, D. Copeland, *et al.*, "Using computer vision to automate hand detection and tracking of surgeon movements in videos of open surgery," in *AMIA Annual symposium proceedings*, American Medical Informatics Association, vol. 2020, 2020, p. 1373.

[242] A. P. Twinanda, J. Marescaux, M. D. Mathelin, and N. Padoy, "Towards better laparoscopic video database organization by automatic surgery classification," in *International Conference on Information Processing in Computer-Assisted Interventions*, Springer, 2014, pp. 186–195.

[243] M. J. Primus, D. Putzgruber-Adamitsch, M. Taschwer, *et al.*, "Frame-based classification of operation phases in cataract surgery videos," in *International Conference on Multimedia Modeling*, Springer, 2018, pp. 241–253.

[244] M. Karbasi, Z. Bhatti, R. Aghababaeyan, *et al.*, "Real-time hand detection by depth images: A survey," *Jurnal Teknologi*, vol. 78, no. 2, 2016.

[245] R. Lun, "Human activity tracking and recognition using kinect sensor," 2018.

[246] M. F. Kıraç, "Real-time human hand pose estimation and tracking using depth sensors," Ph.D. dissertation, Bogaziçi University, 2013.

[247] C. Qian, X. Sun, Y. Wei, X. Tang, and J. Sun, "Realtime and robust hand tracking from depth," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1106–1113.

[248] S. Qin, X. Zhu, H. Yu, S. Ge, Y. Yang, and Y. Jiang, "Real-time markerless hand gesture recognition with depth camera," in *Pacific-Rim conference on multimedia*, Springer, 2012, pp. 186–197.

[249] G. Airò Farulla, L. O. Russo, C. Pintor, *et al.*, "Real-time single camera hand gesture recognition system for remote deaf-blind communication," in *International Conference on Augmented and Virtual Reality*, Springer, 2014, pp. 35–52.

[250] B. Verma and A. Choudhary, "Dynamic hand gesture recognition using convolutional neural network with rgb-d fusion," in *Proceedings of the 11th Indian Conference on Computer Vision, Graphics and Image Processing*, 2018, pp. 1–8.

[251] M. Yuan, F. Farbiz, C. M. Manders, and K. Y. Tang, "Robust hand tracking using a simple color classification technique," in *Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, 2008, pp. 1–5.

[252] https://github.com/FFmpeg/FFmpeg/commit/d08bcbffff [Accessed: Oct.12,2022].

[253] https://opencv.org/ [Accessed: Oct.12,2022].

[254] https://www.python.org/ [Accessed: Oct.12,2022].

[255] https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html [Accessed: Oct.12,2022].

[256] https://keras.io/, [Accessed: July.12, 2022].

[257] https://www.tensorflow.org/, [Accessed: July.12, 2022].

[258] C. E. Heil and D. F. Walnut, "Continuous and discrete wavelet transforms," *SIAM review*, vol. 31, no. 4, pp. 628–666, 1989.

[259] A. M. Nasrabadi, A. R. Eslaminia, P. R. Bakhshayesh, M. Ejtehadi, L. Alibiglou, and S. Behzadipour, "A new scheme for the development of imu-based activity recognition systems for telerehabilitation," *Medical Engineering & Physics*, vol. 108, p. 103 876, 2022.

[260] G. Lee, R. Gommers, F. Waselewski, K. Wohlfahrt, and A. O'Leary, "Pywavelets: A python package for wavelet analysis," *Journal of Open Source Software*, vol. 4, no. 36, p. 1237, 2019.

[261] I. Daubechies, *Ten lectures on wavelets*. SIAM, 1992.

[262] M. G. Tsipouras, A. T. Tzallas, G. Rigas, S. Tsouli, D. I. Fotiadis, and S. Konitsiotis, "An automated methodology for levodopa-induced dyskinesia: Assessment based on gyroscope and accelerometer signals," *Artificial intelligence in medicine*, vol. 55, no. 2, pp. 127–135, 2012.

[263] G. S. Kasman and S. L. Wolf, "Surface emg made easy: A beginner's guide for rehabilitation clinicians," 2002.

[264] P. Kaczmarek, T. Mańkowski, and J. Tomczyński, "Putemg—a surface electromyography hand gesture recognition dataset," *Sensors*, vol. 19, no. 16, p. 3548, 2019.

[265] R. A. Watson, "Use of a machine learning algorithm to classify expertise: Analysis of hand motion patterns during a simulated surgical task," *Academic Medicine*, vol. 89, no. 8, pp. 1163–1167, 2014.

[266] H. Law, K. Ghani, and J. Deng, "Surgeon technical skill assessment using computer vision based analysis," in *Machine learning for healthcare conference*, PMLR, 2017, pp. 88–99.

[267] N. Ahmidi, G. D. Hager, L. Ishii, G. Fichtinger, G. L. Gallia, and M. Ishii, "Surgical task and skill classification from eye tracking and tool motion in minimally invasive surgery," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2010, pp. 295–302.

[268] C. W. Barnhill, S. J. Kaplan, A. A. Alseidi, and S. B. Deal, "A comparative analysis of video-based surgical assessment: Is evaluation of the entire critical portion of the operation necessary?" *Surgical Endoscopy*, pp. 1–5, 2022.

[269] U. H. Tanin, "Deep video analysis methods for surgical skills assessment in cataract surgery," Ph.D. dissertation, Carleton University, 2022.

[270] https://github.com/ultralytics/yolov5 [Accessed: Oct.12,2022].

[271] https://delsys.com/trigno-avanti/ [Accessed: Oct.20,2022].

[272] https://www.biometricsltd.com/datalite.htm [Accessed: Oct.20,2022].

[273] *https://www.noraxon.com/our-products/ultium-emg/?gclid=Cj0KCQjw48OaBhDWARIsAM d966C2AUzXQvJtE7uANz47QbjaQqddc5CXrr0WdrQhewLxXegqnddzCCgaAm4-EALw_wcB [Accessed: Oct.20,2022].*

[274] http://www.oymotion.com/en/product32/149 [Accessed: Oct.20,2022].

[275] *https://www.amazon.com/Arduino-Micro-Headers-A000053-Controller/dp/B00AFY2S56/ref =sr_1_3?dchild=1&keywords=arduino%2Bmicro&qid=1591218072&sr=8- 3&th=1 [Accessed: Oct.20,2022].*

[276] *https://www.amazon.com/HiLetgo-Wireless-Bluetooth-Transceiver-Arduino/dp/B071YJG8DR /ref=pd_lpo_147_t_0/146-3308786-5424240?_encoding=UTF8&pd_rd_i=B071YJG8DR&pd_rd_r =6e04f799-9831-4084-b554-6fd70ee3628b&pd_rd_w=Yu7y0&pd_rd_wg=WyHQf&pf_rd_p=7b36d 496-f366-4631-94d3-61b87b52511b&pf_rd_r=XG4BWQ9FX5ZSCESX752B&refRID=XG4 BWQ9FX5ZSCESX752B&th=1 [Accessed: Oct.20,2022].*

[277] *https://www.amazon.com/SparkFun-PID-13762-IMU-Breakout/dp/B01JQ79FZS/ref=sr_1_2?d child=1&keywords=mpu+9250+breakout&qid=1591211984&sr=8-2 [Accessed: Oct.20,2022].*

[278] https://www.dfrobot.com/product-1661.html [Accessed: Oct.20,2022].

[279] *https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electromyography-emg #vm_a_cb2f0571 [accessed: oct.20,2022].*

[280] *https://bio-medical.com/neurotrode-dry-eeg-emg-ecg-sensors-50-pack.html [Accessed: Oct.20,2022].*

[281] *https://www.medline.com/jump/product/x/Z05-PF75170 [Accessed: Oct.20,2022].*

[282] *https://mfimedical.com/products/gold-plated-bar-emg-electrode [Accessed: Oct.20,2022].*

[283] *https://www.arduino.cc/reference/en/libraries/arduinothread/ [Accessed: May.14,2022].*

[284] *https://arduinojson.org [Accessed: May.14,2022].*

[285] K. Sathian, L. J. Buxbaum, L. G. Cohen, *et al.*, "Neurological principles and rehabilitation of action disorders: Common clinical deficits," *Neurorehabilitation and neural repair*, vol. 25, no. 5_suppl, 21S–32S, 2011.

[286] S. Rosewilliam, C. A. Roskell, and A. Pandyan, "A systematic review and synthesis of the quantitative and qualitative evidence behind patient-centred goal setting in stroke rehabilitation," *Clinical rehabilitation*, vol. 25, no. 6, pp. 501–514, 2011.

[287] H. Zheng, N. D. Black, and N. D. Harris, "Position-sensing technologies for movement analysis in stroke rehabilitation," *Medical and biological engineering and computing*, vol. 43, no. 4, pp. 413–420, 2005.

[288] E. T. Esfahani, S. Pareek, P. Chembrammel, M. Ghobadi, and T. Kesavadas, "Adaptation of rehabilitation system based on user's mental engagement," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers, vol. 57106, 2015, V003T14A017.

[289] S. Pareek, "Iart: An intelligent assistive robotic therapy system for home-based stroke rehabilitation," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2020.

[290] D. J. Reinkensmeyer, C. T. Pang, J. A. Nessler, and C. C. Painter, "Web-based telerehabilitation for the upper extremity after stroke," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 10, no. 2, pp. 102–108, 2002.

[291] A. Alamri, M. Eid, R. Iglesias, S. Shirmohammadi, and A. El Saddik, "Haptic virtual rehabilitation exercises for poststroke diagnosis," *IEEE transactions on instrumentation and measurement*, vol. 57, no. 9, pp. 1876–1884, 2008.

[292] R. Huq, E. Lu, R. Wang, and A. Mihailidis, "Development of a portable robot and graphical user interface for haptic rehabilitation exercise," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, IEEE, 2012, pp. 1451–1457.

[293] A. C. Lo, P. D. Guarino, L. G. Richards, *et al.*, "Robot-assisted therapy for long-term upper-limb impairment after stroke," *New England Journal of Medicine*, vol. 362, no. 19, pp. 1772–1783, 2010.

[294] C. Collin and D. Wade, "Assessing motor impairment after stroke: A pilot reliability study.," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 53, no. 7, pp. 576–579, 1990.

Appendices

Appendix A

Assessment Methods to Evaluate Functionality of Upper Extremity

This Appendix includes samples of assessment metrics that are utilized clinically.

A.1   Action Research Arm Test

Shirley Ryan Abilitylab

Action Research Arm Test

PURPOSE

The ARAT assesses upper limb functioning using observational methods.

LINK TO INSTRUMENT

INSTRUMENT DETAILS

ACRONYM

ARAT

AREA OF ASSESSMENT

Activities of Daily Living
Coordination
Dexterity
Upper Extremity Function

ASSESSMENT TYPE

Observer

ADMINISTRATION MODE

Paper & Pencil

COST

Free

DIAGNOSIS/CONDITIONS

Brain Injury.   Multiple Sclerosis.   Stroke Recovery

POPULATIONS

Stroke   Mixed Populations

KEY DESCRIPTIONS

- The ARAT's is a 19 item measure divided into 4 sub-tests (grasp, grip, pinch, and gross arm movement). Performance on each item is rated on a 4-point ordinal scale ranging from:
  3) Performs test normally
  2) Completes test, but takes abnormally long or has great difficulty
  1) Performs test partially
  0) Can perform no part of test

- Lyle's decision rules:
  1) Patients who achieve a maximum score on the first (most difficult) item are credited with having scored 3 on all subsequent items on that scale.
  2) If the patient scores less than 3 on the first item, then the second item is assessed.
  3) This is the easiest item, and if patients score 0 then they are unlikely to achieve a score above 0 for the remainder of the items and are credited with a zero for the other items.
  4) The maximum score on the ARTS is 57 points (possible range 0 to 57).

- Items can also be summed (van der Lee et al., 2002).

- A standardized scoring protocol has been published by Yozbatiran, 2008.

NUMBER OF ITEMS

19

EQUIPMENT REQUIRED

- Various sized wood blocks
- Cricket ball
- Stone
- Jug and glass
- Tube
- Washer and bolt
- Ball bearing
- A marble

TIME TO ADMINISTER

10 minutes

DEPENDENT ON THE NUMBER OF ITEMS PERFORMED

A.2 The Barthel Index

THE BARTHEL INDEX

Patient Name: _____

Rater Name: _____

Date: _____

Activity	Score

FEEDING
0 = unable
5 = needs help cutting, spreading butter, etc., or requires modified diet
10 = independent _____

BATHING
0 = dependent
5 = independent (or in shower) _____

GROOMING
0 = needs to help with personal care
5 = independent face/hair/teeth/shaving (implements provided) _____

DRESSING
0 = dependent
5 = needs help but can do about half unaided
10 = independent (including buttons, zips, laces, etc.) _____

BOWELS
0 = incontinent (or needs to be given enemas)
5 = occasional accident
10 = continent _____

BLADDER
0 = incontinent, or catheterized and unable to manage alone
5 = occasional accident
10 = continent _____

TOILET USE
0 = dependent
5 = needs some help, but can do something alone
10 = independent (on and off, dressing, wiping) _____

TRANSFERS (BED TO CHAIR AND BACK)
0 = unable, no sitting balance
5 = major help (one or two people, physical), can sit
10 = minor help (verbal or physical)
15 = independent _____

MOBILITY (ON LEVEL SURFACES)
0 = immobile or < 50 yards
5 = wheelchair independent, including corners, > 50 yards
10 = walks with help of one person (verbal or physical) > 50 yards
15 = independent (but may use any aid; for example, stick) > 50 yards _____

STAIRS
0 = unable
5 = needs help (verbal, physical, carrying aid)
10 = independent _____

TOTAL (0–100): _____

The Barthel ADL Index: Guidelines

1. The index should be used as a record of what a patient does, not as a record of what a patient could do.
2. The main aim is to establish degree of independence from any help, physical or verbal, however minor and for whatever reason.
3. The need for supervision renders the patient not independent.
4. A patient's performance should be established using the best available evidence. Asking the patient, friends/relatives and nurses are the usual sources, but direct observation and common sense are also important. However direct testing is not needed.
5. Usually the patient's performance over the preceding 24-48 hours is important, but occasionally longer periods will be relevant.
6. Middle categories imply that the patient supplies over 50 per cent of the effort.
7. Use of aids to be independent is allowed.

References

Mahoney FI, Barthel D. "Functional evaluation: the Barthel Index."
Maryland State Medical Journal 1965;14:56-61. Used with permission.

Loewen SC, Anderson BA. "Predictors of stroke outcome using objective measurement scales."
Stroke. 1990;21:78-81.

Gresham GE, Phillips TF, Labi ML. "ADL status in stroke: relative merits of three standard indexes."
Arch Phys Med Rehabil. 1980;61:355-358.

Collin C, Wade DT, Davies S, Horne V. "The Barthel ADL Index: a reliability study."
*Int Disability Study.*1988;10:61-63.

Copyright Information

A.3 Fatigue Severity Scale

Shirley Ryan
Abilitylab

Fatigue Severity Scale

PURPOSE

The 9-item scale which measures the severity of fatigue and its effect on a person's activities and lifestyle in patients with a variety of disorders.

LINK TO INSTRUMENT

INSTRUMENT DETAILS

ACRONYM

FSS

AREA OF ASSESSMENT

Activities of Daily Living
Life Participation
Sleep

ASSESSMENT TYPE

Patient Reported Outcomes

ADMINISTRATION MODE

Paper & Pencil

COST

Free

DIAGNOSIS/CONDITIONS

Cancer Rehabilitation, Multiple Sclerosis,

Parkinson's Disease + Neurologic Conditions, Stroke Recovery

POPULATIONS

Stroke Spinal Injuries Parkinson's Disease Non-Specific Patient Population

Musculoskeletal Conditions Multiple Sclerosis Mixed Populations Chronic Pain

KEY DESCRIPTIONS

— A 9-item questionnaire with questions related to how fatigue interferes with certain activities and rates its severity according to a self-report scale.

— The items are scored on a 7 point scale with 1 = strongly disagree and 7= strongly agree.

— The minimum score = 9 and maximum score possible = 63. Higher the score = greater fatigue severity.

— Another way of scoring: mean of all the scores with minimum score being 1 and maximum score being 7.

— Lerdal (2011) found that items 1 and 2 in the FSS-9 should not be used in a mean score because the FSS-7 shows better validity and reliability and is likely more sensitive for measuring change in fatigue.

NUMBER OF ITEMS

9

EQUIPMENT REQUIRED

— Pen

TIME TO ADMINISTER

Less than 5 minutes

REQUIRED TRAINING

No Training

INSTRUMENT REVIEWERS

Initially reviewed by Avisha Shah in 10/2012; Updated by Terry Ellis, PT, PhD, NCS and Laura Savella SPT in 2013; Updated by Kent Bubel, OTS, Eva Jarek, OTS and Camara Singleton, OTS in 2016.

ICF DOMAIN

Activity
Participation

MEASUREMENT DOMAIN

General Health

PROFESSIONAL ASSOCIATION RECOMMENDATION

A.4   Frenchay Activities Index

Shirley Ryan Abilitylab

Frenchay Activities Index

PURPOSE

Assesses a broad range of activities of daily living in domains of domestic chores, leisure/work, and outdoor activities.

LINK TO INSTRUMENT

INSTRUMENT DETAILS

ACRONYM

FAI

AREA OF ASSESSMENT

Activities of Daily Living

ASSESSMENT TYPE

Patient Reported Outcomes

ADMINISTRATION MODE

Paper & Pencil

COST

Free

DIAGNOSIS/CONDITIONS

Stroke Recovery

POPULATIONS

Stroke   Older Adults and Geriatric Care   Non-Specific Patient Population

Limb Loss and Amputation

KEY DESCRIPTIONS

— The items included move beyond the scope of ADL scales, which tend to focus on issues related to self care and mobility.

- Can be separated into 3 domains:
  1) Domestic chores
  2) Leisure/work
  3) Outdoor activities

- The frequency with which each item or activity is undertaken over the past 3 or 6 months (depending on the nature of the activity) is assigned a score of 1 – 4 where a score of 1 is indicative of the lowest level of activity.

- The scale provides a summed score from 15 – 60.

- A modified 0–3 scoring system introduced by Wade et al. (1985) yields a score of 0 – 45.

NUMBER OF ITEMS

15

TIME TO ADMINISTER

5 minutes

REQUIRED TRAINING

No Training

AGE RANGES

Adult	Elderly Adult
18 - 64	65 +
YEARS	YEARS

INSTRUMENT REVIEWERS

Initially reviewed by Jason Raad, MS, in 2010; Updated with references for the amputation population by Nick Lefere, SPT and Matt Morris, SPT in 2011; Updated with references for stroke, geriatric, and amputee populations by Erika Gentry, SPT and Ashley Kanelos, SPT in 11/2012. Updated with references for chronic lower back pain, high utilizers of healthcare, traumatic limb injury, and mild cognitive impairment by Susan Felson, OTS, Christine Leung, OTS, Monica O'Connor, OTS in 3/2016.

ICF DOMAIN	MEASUREMENT DOMAIN
Activity	Activities of Daily Living
Participation	

A.5 FUGL-MEYER ASSESSMENT - UPPER EXTREMITY

This section includes a sample of "FUGL-MEYER ASSESSMENT for UPPER EXTREMITY".

Rehabilitation Medicine, University of Gothenburg

FUGL-MEYER ASSESSMENT
UPPER EXTREMITY (FMA-UE)
Assessment of sensorimotor function

ID:
Date:
Examiner:

Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S: The post-stroke hemiplegic patient. A method for evaluation of physical performance. Scand J Rehabil Med 1975, 7:13-31.

A. UPPER EXTREMITY, sitting position

I. Reflex activity		none	can be elicited
Flexors: biceps and finger flexors (at least one)		0	2
Extensors: triceps		0	2
Subtotal I (max 4)			

II. Volitional movement within synergies, without gravitational help			none	partial	full
Flexor synergy: Hand from contralateral knee to ipsilateral ear. From extensor synergy (shoulder adduction/ internal rotation, elbow extension, forearm pronation) to flexor synergy (shoulder abduction/ external rotation, elbow flexion, forearm supination). Extensor synergy: Hand from ipsilateral ear to the contralateral knee	Shoulder	retraction	0	1	2
		elevation	0	1	2
		abduction (90°)	0	1	2
		external rotation	0	1	2
	Elbow	flexion	0	1	2
	Forearm	supination	0	1	2
	Shoulder	adduction/internal rotation	0	1	2
	Elbow	extension	0	1	2
	Forearm	pronation	0	1	2
Subtotal II (max 18)					

III. Volitional movement mixing synergies, without compensation		none	partial	full
Hand to lumbar spine hand on lap	cannot perform or hand in front of ant-sup iliac spine	0		
	hand behind ant-sup iliac spine (without compensation)		1	
	hand to lumbar spine (without compensation)			2
Shoulder flexion 0°- 90° elbow at 0° pronation-supination 0°	immediate abduction or elbow flexion	0		
	abduction or elbow flexion during movement		1	
	flexion 90°, no shoulder abduction or elbow flexion			2
Pronation-supination elbow at 90° shoulder at 0°	no pronation/supination, starting position impossible	0		
	limited pronation/supination, maintains starting position		1	
	full pronation/supination, maintains starting position			2
Subtotal III (max 6)				

IV. Volitional movement with little or no synergy		none	partial	full
Shoulder abduction 0 - 90° elbow at 0° forearm pronated	immediate supination or elbow flexion	0		
	supination or elbow flexion during movement		1	
	abduction 90°, maintains extension and pronation			2
Shoulder flexion 90° - 180° elbow at 0° pronation-supination 0°	immediate abduction or elbow flexion	0		
	abduction or elbow flexion during movement		1	
	flexion 180°, no shoulder abduction or elbow flexion			2
Pronation/supination elbow at 0° shoulder at 30°- 90° flexion	no pronation/supination, starting position impossible	0		
	limited pronation/supination, maintains start position		1	
	full pronation/supination, maintains starting position			2
Subtotal IV (max 6)				

V. Normal reflex activity assessed only if full score of 6 points is achieved in part IV; compare with the unaffected side		0 (IV), hyper	lively	normal
biceps, triceps, finger flexors	2 of 3 reflexes markedly hyperactive or 0 points in part IV	0		
	1 reflex markedly hyperactive or at least 2 reflexes lively		1	
	maximum of 1 reflex lively, none hyperactive			2
Subtotal V (max 2)				
Total A (max 36)				

Approved by Fugl-Meyer AR 2010
1
Updated 2015-03-11

B. WRIST support may be provided at the elbow to take or hold the starting position, no support at wrist, check the passive range of motion prior testing		none	partial	full
Stability at 15° dorsiflexion elbow at 90°, forearm pronated shoulder at 0°	less than 15° active dorsiflexion dorsiflexion 15°, no resistance tolerated maintains dorsiflexion against resistance	0	1	2
Repeated dorsifexion / volar flexion elbow at 90°, forearm pronated shoulder at 0°, slight finger flexion	cannot perform volitionally limited active range of motion full active range of motion, smoothly	0	1	2
Stability at 15° dorsiflexion elbow at 0°, forearm pronated slight shoulder flexion/abduction	less than 15° active dorsiflexion dorsiflexion 15°, no resistance tolerated maintains dorsiflexion against resistance	0	1	2
Repeated dorsifexion / volar flexion elbow at 0°, forearm pronated slight shoulder flexion/abduction	cannot perform volitionally limited active range of motion full active range of motion, smoothly	0	1	2
Circumduction elbow at 90°, forearm pronated shoulder at 0°	cannot perform volitionally jerky movement or incomplete complete and smooth circumduction	0	1	2
Total B (max 10)				

C. HAND support may be provided at the elbow to keep 90° flexion, no support at the wrist, compare with unaffected hand, the objects are interposed, active grasp		none	partial	full
Mass flexion from full active or passive extension		0	1	2
Mass extension from full active or passive flexion		0	1	2
GRASP				
a. Hook grasp flexion in PIP and DIP (digits II-V), extension in MCP II-V	cannot be performed can hold position but weak maintains position against resistance	0	1	2
b. Thumb adduction 1-st CMC, MCP, IP at 0°, scrap of paper between thumb and 2-nd MCP joint	cannot be performed can hold paper but not against tug can hold paper against a tug	0	1	2
c. Pincer grasp, opposition pulpa of the thumb against the pulpa of 2-nd finger, pencil, tug upward	cannot be performed can hold pencil but not against tug can hold pencil against a tug	0	1	2
d. Cylinder grasp cylinder shaped object (small can) tug upward, opposition of thumb and fingers	cannot be performed can hold cylinder but not against tug can hold cylinder against a tug	0	1	2
e. Spherical grasp fingers in abduction/flexion, thumb opposed, tennis ball, tug away	cannot be performed can hold ball but not against tug can hold ball against a tug	0	1	2
Total C (max 14)				

D. COORDINATION/SPEED, sitting, after one trial with both arms, eyes closed, tip of the index finger from knee to nose, 5 times as fast as possible		marked	slight	none
Tremor	at least 1 completed movement	0	1	2
Dysmetria at least 1 completed movement	pronounced or unsystematic slight and systematic no dysmetria	0	1	2
		≥ 6s	2 - 5s	< 2s
Time start and end with the hand on the knee	at least 6 seconds slower than unaffected side 2-5 seconds slower than unaffected side less than 2 seconds difference	0	1	2
Total D (max 6)				

TOTAL A-D (max 66)			
H. SENSATION, upper extremity eyes closed, compared with the unaffected side		anesthesia	hypoesthesia or dysesthesia	normal
Light touch	upper arm, forearm palmary surface of the hand	0 0	1 1	2 2
		less than 3/4 correct or absence	3/4 correct or considerable difference	correct 100%, little or no difference
Position small alterations in the position	shoulder elbow wrist thumb (IP-joint)	0 0 0 0	1 1 1 1	2 2 2 2
				Total H (max12)

J. PASSIVE JOINT MOTION, upper extremity, sitting position, compare with the unaffected side				**J. JOINT PAIN** during passive motion, upper extremity		
	only few degrees (less than 10° in shoulder)	decreased	normal	pronounced pain during movement or very marked pain at the end of the movement	some pain	no pain
Shoulder Flexion (0° - 180°) Abduction (0°-90°) External rotation Internal rotation	0 0 0 0	1 1 1 1	2 2 2 2	0 0 0 0	1 1 1 1	2 2 2 2
Elbow Flexion Extension	0 0	1 1	2 2	0 0	1 1	2 2
Forearm Pronation Supination	0 0	1 1	2 2	0 0	1 1	2 2
Wrist Flexion Extension	0 0	1 1	2 2	0 0	1 1	2 2
Fingers Flexion Extension	0 0	1 1	2 2	0 0	1 1	2 2
Total (max 24)				**Total** (max 24)		

A. UPPER EXTREMITY	/36
B. WRIST	/10
C. HAND	/14
D. COORDINATION / SPEED	/ 6
TOTAL A-D (motor function)	/66

H. SENSATION	/12
J. PASSIVE JOINT MOTION	/24
J. JOINT PAIN	/24

A.6 Functional Independence Measure

This section includes a sample of "Functional Independence Measure".

Functional Independence Measure

Identifying and definitional attributes

Metadata item type:	Glossary Item
Synonymous names:	FIM
METeOR identifier:	495857
Registration status:	Health, Standard 11/04/2014
	Tasmanian Health, Endorsed 20/12/2016
	Independent Hospital Pricing Authority, Standard 31/10/2012
Definition:	An assessment of the severity of patient disability.
Context:	The Functional Independence Measure (FIM™) instrument is a basic indicator of patient disability. FIM™ is used to track the changes in the functional ability of a patient during an episode of hospital rehabilitation care.

Collection and usage attributes

Guide for use:	Patient function is assessed using the FIM™ instrument at the start of a rehabilitation episode of care and at the end of a rehabilitation episode of care. Admission assessment is collected within 72 hours of the start of a rehabilitation episode. Discharge assessment is collected within 72 hours prior to the end of a rehabilitation episode.
Comments:	FIM™ is comprised of 18 items, grouped into 2 subscales - motor and cognition.

The motor subscale includes:

- Eating
- Grooming
- Bathing
- Dressing, upper body
- Dressing, lower body
- Toileting
- Bladder management
- Bowel management
- Transfers - bed/chair/wheelchair
- Transfers - toilet
- Transfers - bath/shower
- Walk/wheelchair
- Stairs

The cognition subscale includes:

- Comprehension
- Expression
- Social interaction
- Problem solving
- Memory

Each item is scored on a 7 point ordinal scale, ranging from a score of 1 to a score of 7. The higher the score, the more independent the patient is in performing the task associated with that item.

1 - Total assistance with helper

2 - Maximal assistance with helper

3 - Moderate assistance with helper

4 - Minimal assistance with helper

5 - Supervision or setup with helper

6 - Modified independence with no helper

7 - Complete independence with no helper

The total score for the FIM motor subscale (the sum of the individual motor subscale items) will be a value between 13 and 91.

The total score for the FIM cognition subscale (the sum of the individual cognition subscale items) will be a value between 5 and 35.

The total score for the FIM instrument (the sum of the motor and cognition subscale scores) will be a value between 18 and 126.

Source and reference attributes

Submitting organisation:	Independent Hospital Pricing Authority
Origin:	FIM™ is a trademark of the Uniform Data System for Medical Rehabilitation, a division of UB Foundation Activities Incorporated.
	Australasian Rehabilitation Outcomes Centre holds the territory license for the use of the FIM™ instrument in Australia.
Reference documents:	Uniform Data System for Medical Rehabilitation 2009. The FIM System® Clinical Guide, Version 5.2. Buffalo: UDSMR.
	Australasian Rehabilitation Outcomes Centre, University of Wollongong 2012. What is the FIM™ Instrument? Viewed 19 September 2012, http://ahsri.uow.edu.au/aroc/whatisfim/index.html

Relational attributes

Metadata items which use this glossary item:	Activity based funding: Admitted sub-acute and non-acute hospital care DSS 2013-2014 Independent Hospital Pricing Authority, Standard 11/10/2012
	Admitted subacute and non-acute hospital care DSS 2014-15 Health, Superseded 13/11/2014
	Admitted subacute and non-acute hospital care DSS 2015-16 Health, Superseded 19/11/2015 Independent Hospital Pricing Authority, Proposed 05/09/2014
	Admitted subacute and non-acute hospital care NBEDS 2016-17 Health, Superseded 03/11/2016
	Admitted subacute and non-acute hospital care NBEDS 2017-18 Health, Standard 03/11/2016 Independent Hospital Pricing Authority, Recorded 04/08/2016
	Admitted subacute and non-acute hospital care NBEDS 2018-19 Health, Standardisation pending 27/09/2017
	Episode of admitted patient care—clinical assessment score, code NN Independent Hospital Pricing Authority, Standard 30/10/2012

125

A.7 Motor Assessment Scale

Shirley Ryan
AbilityLab

Motor Assessment Scale

PURPOSE

Assesses everyday motor function in stroke patients.

LINK TO INSTRUMENT

INSTRUMENT DETAILS

ACRONYM

MAS

AREA OF ASSESSMENT

Activities of Daily Living
Functional Mobility

ASSESSMENT TYPE

Observer

ADMINISTRATION MODE

Paper & Pencil

COST

Free

DIAGNOSIS/CONDITIONS

Stroke Recovery

POPULATIONS

Stroke

KEY
DESCRIPTIONS

— 8 items that assess 8 areas of motor function

Patients perform each task 3 times, only the best performance is

recorded

Items (with the exception of the general tonus item*) are assessed using a 7-point scale (0 to 6)

A score of 6 indicates optimal motor behavior Item scores (with the exception of the general tonus item) are summed to provide an overall score (out of 48 points)

Completing a higher-level item suggests successful performance on lower-level items and thus lower-items can be skipped.

*For the general tonus item, the score is based on continuous observations throughout the assessment. A score of 4 on this item indicates a consistently normal response, a score > 4 indicates persistent hypertonus, and a score < 4 indicates various degrees of hypotonus (Carr et al. 1985).

NUMBER OF ITEMS

8

EQUIPMENT REQUIRED

— Stopwatch
— 8 Jellybeans
— Polystyrene cup
— Rubber ball
— Stool
— Comb
— Spoon
— Pen
— 2 Tea cups
— Water
— Prepared sheet for drawing lines
— Cylindrical shaped object like a jar
— Table

TIME TO ADMINISTER

15 minutes

REQUIRED TRAINING

No Training

AGE RANGES

Adult

18 - 64

YEARS

Elderly Adult

65 +

YEARS

ICF DOMAIN

Activity

MEASUREMENT DOMAIN

Activities of Daily Living
Motor

CONSIDERATIONS

— The general tonus item may be difficult to assess because no guidelines regarding where it should be tested or how the item should be scored exist (e.g. when there is variance between leg, arm, and trunk).

— This item is often omitted for these reasons (Poole & Whitney, 1988; Malouin et al, 1994; Loewen & Anderson, 1990).

— Reliability has only been established in stable patients

— Research suggests the Fugl-Meyer Assessment better discriminates between levels of motor recovery in the early recovery and among more disabled patients (Malouin et al, 1994).

Do you see an error or have a suggestion for this instrument summary? Please e-mail us!

Stroke

NORMATIVE DATA

Mixed Stroke Sample: (English et al. 2006; $n = 61$ chronic and acute patients; mean age = 65.2 (13.1) years; mean time between admission and discharge assessments = 56.4 (38.1) days)

MAS scores at admission:

Item #	Dimension	Mean (SD)
1	Rolling	3.3 (2.0)

A.8 Motricity Index

The Motricity Index for Motor Impairment After Stroke [294]

Overview: The Motricity Index can be used to assess the motor impairment in a patient who has had a stroke.

Tests for Each Arm:

1. pinch grip: using a 2.5 cm cube between the thumb and forefinger

    - 19 points are given if able to grip cube but not hold it against gravity
    - 22 points are given if able to hold cube against gravity but not against a weak pull
    - 26 points are given if able to hold the cube againt a weak pull but strength is weaker than normal

2. elbow flexion from 90° so that the arm touches the shoulder

    - 14 points are given if movement is seen with the elbow out and the arm horizontal

3. shoulder abduction moving the flexed elbow from off the chest

    - 19 points are given when the shoulder is abducted to more than 90° beyond the horizontal against gravity but not against resistance

Tests for Each Leg:

1. ankle dorsiflexion with foot in a plantar flexed position

    - 14 points are given if there is less than a full range of dorsiflexion

2. knee extension with the foot unsupported and the knee at 90°

    - 14 points are given for less than 50
    - 19 points are given for full extension yet it can be easily pushed down

3. - 14 points are given if there is less than a full range of passive motion
    - 19 points are given if the hip is fully flexed yet it can be easily pushed down

MRC Grade	MRC Score	Points for Pinch Grip	Points for Other Tests
no movement	0	0	0
palpable flicker but no movement	1	11	9
movement but not against gravity	2	19	14
movement against gravity	3	22	19
movement against resistance	4	26	25
normal	5	33	33

arm score for each side = SUM(points for the 3 arm tests) + 1

leg score for each side = SUM(points for the 3 leg tests) + 1

side score for each side = ((arm score for side) + (leg score for side)) / 2

Interpretation:

- minimum score: 0

- maximum score: 100

A.9 Wolf Motor Function Test

This section includes a sample of "Wolf Motor Function Test".

Shirley Ryan
Λbilitylab

Wolf Motor Function Test

PURPOSE

The WMFT is a quantitative measure of upper extremity motor ability through timed and functional tasks.

LINK TO INSTRUMENT

INSTRUMENT DETAILS

ACRONYM

WMFT

AREA OF ASSESSMENT

Dexterity
Strength
Upper Extremity Function

ASSESSMENT TYPE

Performance Measure

ADMINISTRATION MODE

Computer

COST

Free

DIAGNOSIS/CONDITIONS

Brain Injury, Stroke Recovery

POPULATIONS

Stroke   Brain Injury

**KEY
DESCRIPTIONS**

- The original version consisted of 21 item; the widely used version of the WMFT consists of 17 items.

- Composed of 3 parts:
  1) Time
  2) Functional ability
  3) Strength

- Includes 15 function-based tasks and 2 strength based tasks
  1) Performance time is referred to as WMFT-TIME
  2) Functional ability is referred to as WMFT-FAS

- Items 1-6 involve timed functional tasks, items 7-14 are measures of strength, and the remaining 9 items consist of analyzing movement quality when completing various tasks.

- Examiner should test the less affected upper extremity followed by the most affected side.

- Uses a 6-point ordinal scale:
  -"0" = "does not attempt with the involved arm" to
  -"5" = "arm does participate; movement appears to be normal."

NUMBER OF ITEMS

21
or 17

EQUIPMENT REQUIRED

- Standardized table (54 inches long, 30 inches wide, and 29 inches high) and chair

- Standardized test item template

- Height-adjustable bedside table

- Box (one that does not require patient to flex or abduct shoulder more than 90 degrees)

- Individual wrist weights, 1-20 pounds

- 12-oz beverage can, unopened

- 7" pencil with 6 flat sides

- 2" paper clip

- 3 checkers

- Three 3" x5" note cards

- Standardized lock and key board at 45 degree angle

- Standardized face towel

- Standardized basket

- — Dynamometer
- — Talcum powder to reduce friction as needed
- — Stopwatch
- — Video camera (optional)

TIME TO ADMINISTER

35 minutes

REQUIRED TRAINING

No Training

AGE RANGES

Adult	Elderly Adult
18 - 64	65 +
YEARS	YEARS

INSTRUMENT REVIEWERS

Initially reviewed by Jason Raad MS and the Rehabilitation Measures Team; Updated by Irene Ward, PT, DPT, NCS and the TBI EDGE task force of the Neurology Section of the APTA in 2012.

BODY PART

Upper Extremity

ICF DOMAIN

Activity

MEASUREMENT DOMAIN

Motor

PROFESSIONAL ASSOCIATION RECOMMENDATION

Recommendations for use of the instrument from the Neurology Section of the American Physical Therapy Association's Multiple Sclerosis Taskforce (MSEDGE), Parkinson's Taskforce (PD EDGE), Spinal Cord Injury Taskforce (PD EDGE), Stroke Taskforce (StrokEDGE), Traumatic Brain Injury Taskforce (TBI EDGE), and Vestibular Taskforce (Vestibular EDGE) are listed below. These recommendations were developed by a panel of research and clinical experts using a modified Delphi process.

For detailed information about how recommendations were made, please visit: http://www.neuropt.org/go/healthcare-professionals/neurology-section-outcome-measures-recommendations

Appendix B

IRB Approval for ADL Studies

This appendix includes the IRB approval letter for the presented study in the Part I of this dissertation.

ILLINOIS

OFFICE OF THE VICE CHANCELLOR FOR RESEARCH

Office for the Protection of Research Subjects
805 W. Pennsylvania Ave., MC-095
Urbana, IL 61801-4822

Notice of Approval: New Submission

July 17, 2018

Principal Investigator	T Kesavadas
CC	Hajar Sharif
Protocol Title	*Understanding of grasps in activities of daily living*
Protocol Number	18529
Funding Source	National Science Foundation
Review Type	Expedited 4, 6
Status	Active
Risk Determination	no more than minimal risk
Approval Date	July 17, 2018
Expiration Date	July 16, 2019

This letter authorizes the use of human subjects in the above protocol. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) has reviewed and approved the research study as described.

The Principal Investigator of this study is reponsible for:
- Conducting research in a manner consistent with the requirements of the University and federal regulations found at 45 CFR 46.
- Requesting approval from the IRB prior to implementing modifications.
- Notifying OPRS of any problems involving human subjects, including unanticipated events, participant complaints, or protocol deviations.
- Notifying OPRS of the completion of the study.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
IORG0000014 • FWA #00008584
217.333.2670 • irb@illinois.edu • oprs.research.illinois.edu

Appendix C

Written Instructions of the Surgical Tasks

Task 1. Needle-passing (using half-bites):

1. Picks up the needle with the correct instruments and correctly grasp the needle

2. Proceeds the needle to the incision in the bench top model

3. Passes the needle through the tissue from one side of the incision and out the middle of the incision

4. Grasp the needle appropriately and then reposition the needle with your instruments to finish passing the needle through the other side of the incision at the corresponding dot marked

5. Correctly re-grasp the needle in preparation for a new bite

6. Proceeds the needle to the middle of the incision in the bench top model

7. Passes the needle through the tissue on the other side of the incision

8. Grasp the needle appropriately and then reposition the needle with your instruments to finish passing the needle through the other side of the incision at the corresponding dot marked

9. Place instruments on the table and then place your hands on the table until otherwise instructed

***Once instructed, please pull the suture out of the model and set everything down.

Task2. Knot-tying: the participant:

1. Picks up the needle with the correct instruments

2. Proceeds the needle to the incision in the bench top model

3. Passes the needle through the tissue from one side of the incision and out the middle of the incision

4. Grasp the needle appropriately and then reposition the needle with your instruments to finish passing the needle through the other side of the incision at the corresponding dot marked

5. Correctly re-grasp the needle in preparation for a new bite

6. Proceeds the needle to the middle of the incision in the bench top model

7. Passes the needle through the tissue on the other side of the incision

8. Grasp the needle appropriately and then reposition the needle with your instruments to finish passing the needle through the other side of the incision at the corresponding dot marked

9. Tie a single surgeons knot

10. Follow the surgeons knot with 4 extra throws for a total of 6 throws

11. Appropriately cut the suture to the correct length using the correct instrument

12. Place instruments on the table and then place your hands on the table until otherwise instructed

***Once task 2 is completed, please set everything down–DO NOT CUT OR REMOVE THE SUTURE.

Task 3. Continuous suturing:

1. Picks up the needle with the correct instruments

2. Proceeds the needle to the incision in the bench top model

3. Passes the needle through the tissue from one side of the incision and out the middle of the incision

4. Grasp the needle appropriately and then reposition the needle with your instruments to finish passing the needle through the other side of the incision at the corresponding dot marked

5. Correctly re-grasp the needle in preparation for a new bite

6. Proceeds the needle to the middle of the incision in the bench top model

7. Passes the needle through the tissue on the other side of the incision

8. Grasp the needle appropriately and then reposition the needle with your instruments to finish passing the needle through the other side of the incision at the corresponding dot marked

9. Tie a single surgeons knot

10. Follow the surgeons knot with 4 extra throws for a total of 6 throws

11. Appropriately cut the free end of the suture to the correct length using the correct instrument

12. Passes the needle through the tissue from one side of the incision at the dot marked to the other side at the corresponding dot marked for a total of 3 times, then pass 6 more times without marked locations

13. Extracts the needle out of the tissue and re-grasp the needle after each pass

14. Repeats 9 more times for total of 10 passes

15. Tie a square knot applying appropriate tension on both the suture and suture loop

16. Follow the surgeons knot with 4 extra throws for a total of 6 throws

17. Appropriately cut the suture to the correct length using the correct instrument

18. Place instruments on thetable and then place your hands on the table until otherwise instructed

***Once task 3 is completed, please set everything down –DO NOT CUT OR REMOVE THE SUTURE.

Task 4. Buried suture:

1. Pick up the needle with the correct instruments

2. Proceed the needle to the incision in the bench top model

3. Pass the needle through the tissue from deep to superficial and out the middle of the incision. Your needle and suture should not breach the tissue face or engage the subcutaneous tissue

4. Correctly re-grasp the needle in preparation for a new bite

5. Proceed the needle to the middle of the incision in the bench top tissue model

6. Pass the needle through the tissue on the other side of the incision from superficial to deep. Your needle and suture should not breach the tissue face or engage the subcutaneous tissue

7. Tie a single surgeon's throw so that it buries in the tissue pulling parallel with the incision

8. Follow the surgeon's throw with 3 extra throws for a total of 5throws

9. Appropriately cut the suture to the correct length using the correct instrument

10. Place instruments on the table and then place your hands on the table until otherwise instructed

***Once Task 4is completed, please set everything down –DO NOT CUT OR REMOVE THE SUTURE.

Appendix D

IRB Approval for Surgical Skills Assessment Studies

This appendix includes the IRB approval letter for the presented study in the part II of this dissertation.

I ILLINOIS

OFFICE OF THE VICE CHANCELLOR
FOR RESEARCH & INNOVATION
Office for the Protection of Research Subjects
805 W. Pennsylvania Ave., MC-095
Urbana, IL 61801-4822

Notice of Approval: New Submission

December 11, 2020

Principal Investigator	Heidi Phillips
CC	Thenkurussi Kesavadas, Hajar Sharif, Annette McCoy, Jennifer Kuzminksy, Leslie McNeil
Protocol Title	*Skill Assessment in Surgery and Microsurgery*
Protocol Number	21177
Funding Source	Jump ARCHES Endowment
Review Type	Expedited 4, 6, 7
Status	Active
Risk Determination	No more than minimal risk
Approval Date	December 11, 2020
Closure Date	December 10, 2025

This letter authorizes the use of human subjects in the above protocol. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) has reviewed and approved the research study as described.

The Principal Investigator of this study is responsible for:
- Conducting research in a manner consistent with the requirements of the University and federal regulations found at 45 CFR 46.
- Using the approved consent documents, with the footer, from this approved package.
- Requesting approval from the IRB prior to implementing modifications.
- Notifying OPRS of any problems involving human subjects, including unanticipated events, participant complaints, or protocol deviations.
- Notifying OPRS of the completion of the study.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
IORG0000014 • FWA #00008584
217.333.2670 • irb@illinois.edu • oprs.research.illinois.edu

Appendix E

First IRB Amendment Approval for Surgical Skills Assessment Studies

This appendix includes the IRB approval letter for the first amendment for the presented study in the part II of this dissertation.

UNIVERSITY OF ILLINOIS
URBANA-CHAMPAIGN

Office of the Vice Chancellor for Research & Innovation
Office for the Protection of Research Subjects
805 W. Pennsylvania Ave., MC-095
Urbana, IL 61801-4822

Notice of Approval: Amendment #01

September 24, 2021

Principal Investigator	Heidi Phillips
CC	Thenkurussi Kesavadas; Paul Jeziorczak; Hajar Sharif; Leslie Klis McNeil; Jennifer Kuzminsky; Annette McCoy
Protocol Title	*Skill Assessment in Surgery and Microsurgery*
Protocol Number	21177
Funding Source	Jump ARCHES Endowment through the Health Care Engineering Systems Center
Review Type	Expedited 4, 6, 7
Amendment Requested	• Updating the research team • Adding phase 2 of study at new research site (Jump Simulation Center in Peoria)
Status	Active
Risk Determination	No more than minimal risk
Approval Date	September 24, 2021 (amendment approval date)
Closure Date	December 10, 2025

This letter authorizes the use of human subjects in the above protocol. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) has reviewed and approved the research study as described.

The Principal Investigator of this study is responsible for:
- Conducting research in a manner consistent with the requirements of the University and federal regulations found at 45 CFR 46.
- Using the approved consent documents, with the footer, from this approved package.
- Requesting approval from the IRB prior to implementing modifications.
- Notifying OPRS of any problems involving human subjects, including unanticipated events, participant complaints, or protocol deviations.
- Notifying OPRS of the completion of the study.

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN
IORG0000014 • FWA #00008584
217.333.2670 • irb@illinois.edu • oprs.research.illinois.edu

140

Appendix F

Second IRB Amendment Approval for Surgical Skills Assessment Studies

This appendix includes the IRB approval letter for the second amendment for the presented study in the part II of this dissertation.

UNIVERSITY OF ILLINOIS
URBANA-CHAMPAIGN

Office of the Vice Chancellor for Research & Innovation
Office for the Protection of Research Subjects
805 W. Pennsylvania Ave., MC-095
Urbana, IL 61801-4822

Notice of Approval: Amendment #02

October 25, 2021

Principal Investigator	Heidi Phillips
CC	T Kesavadas; Hajar Sharif; Annette McCoy; Paul Jeziorczak; Leslie McNeil; Jennifer Kuzminsky
Protocol Title	*Skill Assessment in Surgery and Microsurgery*
Protocol Number	21177
Funding Source	Jump ARCHES Endowment through the Health Care Engineering Systems Center
Review Type	Expedited 4, 6, 7
Amendment Requested	Adding a recruitment flyer to distribute among residents
Status	Active
Risk Determination	No more than minimal risk
Approval Date	October 25, 2021 (amendment approval date)
Closure Date	December 10, 2025

This letter authorizes the use of human subjects in the above protocol. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) has reviewed and approved the research study as described.

The Principal Investigator of this study is responsible for:
- Conducting research in a manner consistent with the requirements of the University and federal regulations found at 45 CFR 46.
- Using the approved consent documents, with the footer, from this approved package.
- Requesting approval from the IRB prior to implementing modifications.
- Notifying OPRS of any problems involving human subjects, including unanticipated events, participant complaints, or protocol deviations.
- Notifying OPRS of the completion of the study.

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN
IORG0000014 • FWA #00008584
217.333.2670 • irb@illinois.edu • oprs.research.illinois.edu

Appendix G

Questionnaires for Longitudinal Study (Surgical Skill Assessment)

ID#

Pre-participation Survey:

1. Please describe your current position within the Veterinary Teaching Hospital:

   1st year student   2nd year student   3rd year student   Resident   Faculty

2. Please identify your dominant hand:

   RIGHT   **LEFT**

3. Have you previously worn an electomyographic (EMG) armband?

   YES   **NO**

4. Have you had any previous experience with suturing or using surgical instruments? Circle your answer.

   1 .... 2 .... 3 .... 4 5

   1 = No experience   3 = Moderate experience   5 = Significant experience

   Please describe:

5. Have you practiced suturing on a synthetic skin model previously?

   YES   **NO**

6. Have you had formal training under guidance of a board-certified surgeon? Circle your answer.

   1 .... 2 .... 3 .... 4 5

   1 = No experience   3 = Moderate experience   5 = Significant experience

   Please describe:

7. Have you had any practical surgical *experience* (i.e. assisting a primary vet or surgeon, working in a shelter environment, doing spays/neuters/laceration repairs)? Circle your answer.

1 .... 2 .... 3 .... 4 5

1 = No experience   3 = Moderate experience   5 = Significant experience

Please describe:

8. Have you had any other training that could be related to this project?

Please describe:

9. Estimate how many cadavers , if any, you have practiced suturing previously. Circle your answer.

1 .... 2 .... 3 .... 4 5
1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Please describe:

2. Estimate how many live animals, if any, you have sutured previously. Circle your answer.

1 .... 2 .... 3 .... 4 5
1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Please describe:

3. Estimate how many hours you have practiced suturing on tissue models. Circle your answer.

1 .... 2 .... 3 .... 4 5
1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Post- participation Survey:

Please rate your level of agreement with the following statements by circling your answer:

1. The electromyography (EMG) armband was uncomfortable and interfered with my performance.

   Strongly agree   Agree   Neither Agree nor Disagree   Disagree   Strongly Disagree

   Please describe:

2. The camera canopy and experimental set-up interfered with my performance.

   Strongly agree   Agree   Neither Agree nor Disagree   Disagree   Strongly Disagree

   Please describe:

3. The synthetic tissue model received and held suture well.

   Strongly agree   Agree   Neither Agree nor Disagree   Disagree   Strongly Disagree

   Please describe:

4. This experience was stressful or caused me anxiety.

   Strongly agree   Agree   Neither Agree nor Disagree   Disagree   Strongly Disagree

   Please describe:

Are you willing to continue in this study and return for more data collection sessions? If your answer is "YES", we will keep your information until all the data collection sessions are completed.

YES   **NO**

Practice

Estimate how many hours you practiced suturing on the tissue model. Circle your answer.

Week 1:

1 .... 2 .... 3 .... 4 5

1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Week 2:

1 .... 2 .... 3 .... 4 5

1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Week 3:

1 .... 2 .... 3 .... 4 5

1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Week 4:

1 .... 2 .... 3 .... 4 5

1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Please circle the 3 skills you practiced most:

Week 1:

Instrument handling	Respect for tissue, gentle tissue handling
Using correct needle holding technique	Apposing tissue with correct amount of tension
Passage of needle through the tissue	Knot tying
Spacing of suture bites from incision edge	Suture handling
Orientation of suture bites relative to incision	Time and motion
Use of forceps to handle the needle	Flow of task

145

Week 2:

Instrument handling	Respect for tissue, gentle tissue handling
Using correct needle holding technique	Apposing tissue with correct amount of tension
Passage of needle through the tissue	Knot tying
Spacing of suture bites from incision edge	Suture handling
Orientation of suture bites relative to incision	Time and motion
Use of forceps to handle the needle	Flow of task

Week 3:

Instrument handling	Respect for tissue, gentle tissue handling
Using correct needle holding technique	Apposing tissue with correct amount of tension
Passage of needle through the tissue	Knot tying
Spacing of suture bites from incision edge	Suture handling
Orientation of suture bites relative to incision	Time and motion
Use of forceps to handle the needle	Flow of task

Week 4:

Instrument handling	Respect for tissue, gentle tissue handling
Using correct needle holding technique	Apposing tissue with correct amount of tension
Passage of needle through the tissue	Knot tying
Spacing of suture bites from incision edge	Suture handling
Orientation of suture bites relative to incision	Time and motion
Use of forceps to handle the needle	Flow of task

Recommendations for Improvement

Below please find a list of tasks on which

✓	Steps – Preparing and Handling Supplies
	Selects Mayo-Hegar needle holder, Brown-Adson forceps, and suture scissors
	Holds needle drivers with ring finger, not index finger, in handle ring
	Holds thumb forceps from below, like a pencil, not with palm over top of forceps
✓	Steps – Simple Interrupted Pattern
	Grasps needle in tip of needle holder 1/3 to 1/2 distance from swedged end to point (at least 1/3 needle away from suture end)
	Needle point faces student when driver held palm down and pointed in the direction of progress in closing incision.
	Controls suture length in hand holding needle driver
	Uses thumb forceps to lift "tissue" of near side onto needle as it penetrates through from external face of "tissue"
	Uses thumb forceps to stabilize "tissue" of far side of incision as needle rotates through from internal to external face of "tissue"
	Pulls needle out of tissue with instrument, , preferably needle drive, NOT fingers
	Uses fingers to draw suture through tissue, controlling it within the zone of sterility
	Leaves tail of suture appropriate length for tying surgeon's knot
	Positions needle driver parallel with incision, between suture tail and length with needle
	Wraps suture length with needle around driver toward suture tail (this direction alternates between away or toward surgeon with each throw)
	Breaks throw symmetrically and tightens by pulling suture lengths perpendicular to incision with tension in the plane of the table (not lifting upward)
	Avoids over-tightening suture
	Completes knot with recommended 5 total throws (surgeon's plus 3)
	Trims suture to leave 2 mm tag ends
	Knots are square

✓	Steps – Simple Continuous Pattern
	Begins as for Simple Interrupted, taking half "bites" from tissue face to incision, incision to tissue face)
	Knot anchoring suture length has 6 total throws (surgeon's plus 4)

	Advances along incision 5 mm (3 to 8) then takes a second "bite" transversely through incision using thumb forceps correctly to control tissue
	Pulls needle out of tissue with instrument, NOT fingers
	Resets needle driver for next bite after pulling needle through tissue
	Uses fingers to draw suture through tissue, controlling it within the zone of sterility
	Tissue edges are evenly apposed
	Suture bites are parallel to each other and evenly spaced
	Leaves loop of last bite for tail of appropriate length for tying surgeon's knot
	Positions needle driver parallel with incision, between suture tail and suture length with needle
	Wraps length with needle around driver toward suture tail (this direction alternates between away or toward surgeon with each throw)
	Breaks throw symmetrically and tightens by pulling suture lengths perpendicular to incision with tension in the plane of the table (not lifting upward)
	Opens driver with jaw still hooked in loop so that loop collapses and acts like a single length while tightening knot
	Avoids over-tightening suture
	Completes knot with 7 or 8 total throws
	Suture trimmed to leave 2 mm tag ends
	Knots are square

Comments:

Appendix H

Questionnaires for Surgical Skill Assessment Study

ID#

Pre-participation Survey:

1. Please describe your current position within the Hospital:

Student	Intern	Resident	Attending

2. Please identify your dominant hand:

RIGHT   LEFT

3. Have you previously worn an electomyographic (EMG) armband?

YES   NO

4. Have you had any previous experience with suturing or using surgical instruments? Circle your answer.

1 .... 2 .... 3 .... 4 5

1 = No experience   3 = Moderate experience   5 = Significant experience

Please describe:

5. Have you practiced suturing on a synthetic skin model previously?

YES   NO

6. Have you had formal training under guidance of a board-certified surgeon? Circle your answer.

1 .... 2 .... 3 .... 4 5

1 = No experience   3 = Moderate experience   5 = Significant experience

Please describe:

7. Have you had any practical surgical *experience* (i.e. assisting a surgeon, performing surgery, doing laceration repairs, etc)? Circle your answer.

1 .... 2 .... 3 .... 4 5

1 = No experience   3 = Moderate experience   5 = Significant experience

Please describe:

8. Have you had any other training that could be related to this project?

Please describe:

9. Estimate how many cadavers, if any, you have practiced suturing previously. Circle your answer.

1 .... 2 .... 3 .... 4 5

1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Please describe:

2. Estimate how many people, if any, you have sutured previously. Circle your answer.

1 .... 2 .... 3 .... 4 5

1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Please describe:

3. Estimate how many hours you have practiced suturing on tissue models. Circle your answer.

1 .... 2 .... 3 .... 4 5

1 = 0   2= 1-10   3 = 11-25   4= 26-50   5 = 50+

Post- participation Survey:

Please rate your level of agreement with the following statements by circling your answer:

1. The electromyography (EMG) armband was uncomfortable and interfered with my performance.

   Strongly agree   Agree   Neither Agree nor Disagree   Disagree   Strongly Disagree
   Please describe:

2. The camera canopy and experimental set-up interfered with my performance.

   Strongly agree   Agree   Neither Agree nor Disagree   Disagree   Strongly Disagree
   Please describe:

3. The synthetic tissue model received and held suture well.

   Strongly agree   Agree   Neither Agree nor Disagree   Disagree   Strongly Disagree
   Please describe:

4. This experience was stressful or caused me anxiety.

   Strongly agree   Agree   Neither Agree nor Disagree   Disagree   Strongly Disagree
   Please describe:

Are you willing to continue in this study and return for more data collection sessions? If your answer is "YES", we will keep your information until all the data collection sessions are completed.

                YES   **NO**