Some Machine Learning Approaches to Enhance Electromyography-based Hand Gesture Recognition

The thesis submitted in partial fulfilment

for the Award of Degree

DOCTOR OF PHILOSOPHY

by

SHASHANK KUMAR SINGH

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY

(BANARAS HINDU UNIVERSITY)

VARANASI -221005

INDIA

Roll No.: 17071508 **Year: 2024**

CERTIFICATE

It is certified that the work contained in this thesis entitled "Some Machine Learning Approaches to Enhance Electromyography-based Hand Gesture Recognition" submitted by "SHASHANK KUMAR SINGH" (Roll No.: 17071508) has been carried out under my supervision and that it has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all the requirements of Comprehensive Examination, Candidacy and State-of-the-Art (SOTA) for the award of Ph.D. Degree.

Amrita

Signature of Supervisor

Dr. AMRITA CHATURVEDI

Assistant Professor

Department of Computer Science and Engineering

Indian Institute of Technology

(Banaras Hindu University)

Varanasi-221005 , India

पर्यवेक्षक/Supervisor
संगणक विज्ञान एवं अभियांत्रिकी विभाग
Department of Computer Sc. & Engg
भारतीय प्रौद्योगिकी संस्थान
Indian Institute of Technology
(काशी हिन्दू विश्वविद्यालय)
(Banaras Hindu University)
वाराणसी/Varanasi-221005

DECLARATION BY THE CANDIDATE

I, SHASHANK KUMAR SINGH , certify that the work embodied in this thesis is my own bonafide work and carried out by me under the supervision of Dr. AMRITA CHATURVEDI from DECEMBER-2017 to AUGUST-2024, at DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, Indian Institute of Technology (BHU) Varanasi. The matter embodied in this thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credits to the research workers wherever their works have been cited in my work in this thesis. I further declare that I have not willfully lifted up any other's work, paragraphs, text, data, results, etc., reported in journals, books, magazines, reports, dissertations, theses, etc., or available on websites and have not included them in this thesis and have not cited as my own work.

Shashank

Date : 20/08/2024

Place : Varanasi

Signature of the Student

(SHASHANK KUMAR SINGH)

CERTIFICATE BY THE SUPERVISOR

It is certified that the above statement made by the student is correct to the best of my/our knowledge.

Amrita

Signature of Supervisor

Date : 20/8/2024

Place : Varanasi

(Dr. AMRITA CHATURVEDI)
Department of Computer Science and Engineering
Indian Institute of Technology
(Banaras Hindu University)
Varanasi-221005

Signature of Head of Department

विभागाध्यक्ष/HEAD
संगणक विज्ञान एवं अभियांत्रिकी विभाग
Department of Computer Sc. & Engg
भारतीय प्रौद्योगिकी संस्थान
Indian Institute of Technology
(काशी हिन्दू विश्वविद्यालय)
(Banaras Hindu University)
वाराणसी/Varanasi-221005

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: Some Machine Learning Approaches to Enhance Electromyography-based Hand Gesture Recognition

Name of the Student: SHASHANK KUMAR SINGH

Copyright Transfer

Date : 20/8/2024

Place : Varanasi

Shashank

(SHASHANK KUMAR SINGH)

ACKNOWLEDGEMENTS

"We must find time to stop and thank the people who make a difference in our lives." – John F. Kennedy

I am immensely thankful to my thesis advisor, Dr. Amrita Chaturvedi, for her constant support, guidance, and encouragement throughout my research journey. Your profound insights and advice were crucial, and this thesis could not have been accomplished without your expert guidance and mentorship. I also extend my gratitude to my thesis committee member, Dr. Lakshmanan Kailasam, whose insightful feedback and suggestions significantly improved my work.

I am grateful to Prof. Sanjay Kumar Singh, Head of the Computer Science and Engineering Department, for his support and kindness in facilitating my research. My thanks also go to other faculty members of the Department, Prof. A.K. Tripathi, Prof. Rajeev Srivastava, and Dr. Ruchir Gupta, for their motivational and extensive guidance throughout my academic tenure.

My colleagues and friends at the department have provided a dynamic and supportive environment, for which I am very grateful. Special thanks are due to Mr. Anurag Tiwari, Mr. Rakesh Kumar, Miss. Shruti Bajpai, Mr. Ankit Jaiswal, Mr. Nirbhay Tagore, Mr. Supriya Chanda, and Mr. Ramakant Kumar, whose assistance, encouragement, and feedback were invaluable. This thesis benefitted immensely from their thoughtful critiques and ongoing support.

I appreciate the department's staff, including Mr. Ravi Kumar Bharti, Mr. Shubham Pandey, Mr. Prakhar Kumar, Mr. Manoj Kumar Rai, Mr. Viplav Biswas, and Mr. Akhilesh Kumar Pandey, for their help in providing the necessary resources and creating a conducive research environment. Special thanks to the non-teaching staff for their unwavering support. This journey has been both challenging and fulfilling, and I am grateful to everyone who has been a part of it.

I would also like to acknowledge the technical staff of the UG II Laboratory for their indispensable support in providing devices that were crucial for my experiments. Their expertise and willingness to help at every step were fundamental to the success of my research. Additionally, I extend my gratitude to the staff at the Center for Computational and Information Sciences (CCIS) for their invaluable assistance and support. Their dedication and technical expertise greatly contributed to the smooth execution of my research activities.

Shashank Kumar Singh

Contents

List of Figures

List of Tables

Abbreviations

1D-CNN	**O**ne-**D**imensional **C**onvolutional **N**eural **N**etwork
3D	**T**hree **D**imensional
AD	**A**lzheimer's **D**isease
AI	**A**rtificial **I**ntelligence
ANN	**A**rtificial **N**eural **N**etwork
ASL	**A**merican **S**ign **L**anguage
AUC-ROC	**A**rea **U**nder the **C**urve - **R**eceiver **O**perating **C**haracteristic
AUC	**A**rea **U**nder the **C**urve
BGRU	**B**idirectional **G**ated **R**ecurrent **U**nit
BiLSTM	**Bi**directional **L**ong **S**hort-**T**erm **M**emory
BLSTM	**B**idirectional **L**ong **S**hort-**T**erm **M**emory
CSV	**C**omma-**S**eparated **V**alues
CNN	**C**onvolutional **N**eural **N**etwork
DAE	**D**enoising **A**uto**E**ncoder
DARWIN	**D**ataset for **A**lzheimer's **R**esearch on **W**riting k**I**Nematics
DCNN	**D**eep **C**onvolutional **N**eural **N**etwork
DTW	**D**ynamic **T**ime **W**arping
EMG	**E**lectro**M**yo**G**raphy
FFT	**F**ast **F**ourier **T**ransform
FPR	**F**alse **P**ositive **R**ate
GRU	**G**ated **R**ecurrent **U**nit
HCI	**H**uman-**C**omputer **I**nteraction

HCR	**H**andwritten **C**haracter **R**ecognition
HGR	**H**and **G**rasp **R**ecognition
HMM	**H**idden **M**arkov **M**odel
IMU	**I**nertial **M**easurement **U**nit
IoT	**I**nternet of **T**hings
KNN	**K-N**earest **N**eighbors
KRE	**K**raskov **E**ntropy
LDA	**L**inear **D**iscriminant **A**nalysis
LSTM	**L**ong **S**hort-**T**erm **M**emory
MCC	**M**atthews **C**orrelation **C**oefficient
MDAE	**M**ultimodal **D**eep **A**uto**E**ncoder
MDC	**M**odel of **D**ynamic **C**ontraction
ML	**M**achine **L**earning
MSE	**M**ean **S**quared **E**rror
MSC	**M**odel of **S**tatic **C**ontraction
MU	**M**otor **U**nit
MUAP	**M**otor **U**nit **A**ction **P**otential
MSSA	**M**odified **S**alp **S**warm **A**lgorithm
PCA	**P**rincipal **C**omponent **A**nalysis
PSO	**P**article **S**warm **O**ptimization
ReLU	**Re**ctified **L**inear **U**nit
RF	**R**andom **F**orest
RNN	**R**ecurrent **N**eural **N**etwork
ROC-AUC	**R**eceiver **O**perating **C**haracteristic - **A**rea **U**nder the **C**urve
ROC	**R**eceiver **O**perating **C**haracteristic
SHAP	**Sh**apley **A**dditive ex**P**lanations
SLRS	**S**ign **L**anguage **R**ecognition **S**ystem
SSA	**S**alp **S**warm **A**lgorithm
SSAE	**S**tacked **S**parse **A**uto**E**ncoder

SSDAE	**S**tacked **S**parse **D**enoising **A**uto**E**ncoder
sEMG	**s**urface **E**lectro**M**yo**G**raphy
SVD	**S**ingular **V**alue **D**ecomposition
SVM	**S**upport **V**ector **M**achine
TQWT	**T**unable-**Q** **W**avelet **T**ransform
TPR	**T**rue **P**ositive **R**ate
t-SNE	**t**-**D**istributed **S**tochastic **N**eighbor **E**mbedding
TP	**T**rue **P**ositive
UWB	**U**ltra-**W**ide**B**and
XAI	**E**xplainable **A**rtificial **I**ntelligence
XGB	**X**G**B**oost

Symbols

D	Dataset consisting of sEMG time series recordings
N	Number of sEMG time series recordings
$TS_i(u)$	Time series recording i with data points $u_1, u_2, \ldots, u_N$
$f_{i,m}$	Feature extracted for a window size where m is the total number of features
Ins_i	i^{th} instance in the dataset
l_k	Class label of hand grasp movement corresponding to the instance Ins_i
$v(S)$	Real-valued profit/payoff for coalition S
$\delta(S, j)$	Marginal contribution of player j to coalition S
$\varphi_j(X, v)$	Shapley value for player j in the coalition game
$G(X_d(t_n), t_n)$	Deterministic part of the Langevin equation
$H(X_d(t_n), t_n)$	Stochastic part of the Langevin equation
$\tau(t_n)$	White Gaussian noise in the Langevin equation
X	Raw sEMG data matrix
T	Number of time samples
$\Phi(X)$	Feature vector generated from raw data
M	Number of features
$\Phi_{S_i}(X)$	Set of relevant features selected by algorithm i
$\Phi_S(X)$	Combined feature vector from multiple methods
h	Classifier function
θ	Parameters of the classifier
$\hat{Y}$	Predicted gesture label

Y	True gesture label
$L(\theta, S)$	Classification loss
ℓ	Loss function (e.g., cross-entropy loss)
A	Classification accuracy
$A_{\min}$	Specified accuracy threshold
$T_s(u)$	Time series corresponding to each sEMG sensor
m	Data points in time series
$W_s(p, s)$	Window segment of length p
z	Total number of samples
l_i	Annotation (label)

PREFACE

In recent years, the field of Human-Computer Interaction (HCI) has seen a marked increase in efforts to create more user-friendly and interactive interfaces through a variety of inputs. These inputs range from traditional methods like keyboards and mouse to advanced technologies such as voice commands, eye tracking, Brain-Computer Interfaces (BCI), augmented and virtual reality, haptic feedback systems, and hand gestures.

Hand Gesture Recognition (HGR) is a key component of Human-Computer Interaction (HCI), aimed at creating technologies that that interpret human commands through gestures. This involves computer systems recognizing and understanding human hand movements to enable device interaction. A promising area in HGR is the use of biosignals generated during hand gestures to establish correspondence with related inputs. These systems play a vital role in improving human-machine and human-robot interactions, enhancing human abilities by utilizing the electrical characteristics of the human nervous system. They make use of signals from tissues, muscles, organs, or the central nervous system to develop interfaces for computers.

One such notable input method for HGR models is surface Electromyography (sEMG) signals, a biosignal that can track skeletal muscles and measure electrical activity generated during static and dynamic gestures. This biosignal can also approximate the intensity and specific type of muscle engagement, providing a clear correlation with the intended motion.

sEMGs are non-invasive and hold significant value in areas like medical rehabilitation, prosthetics, sports science, and beyond. However, sEMG processing faces challenges like signal interference, individual variability, motion artifacts, complex data processing needs, muscle fatigue effects, and sensitivity to electrode placement. Despite these challenges, advancements in machine learning and signal processing are enhancing the effectiveness and applicability of sEMG-based hand gesture recognition systems.

This thesis presents the design and development of innovative sEMG-based hand gesture recognition techniques, focusing on static and complex dynamic hand gestures. This research encompasses the recognition of Sign Language gestures, hand grasp gestures (static gestures), and handwritten characters (dynamic gestures) using data-centric methods, primarily within a machine learning framework. Additionally, explainable AI methods are predominantly used throughout this thesis to enhance the interpretability of the models developed.

The first contribution is the development of an efficient sEMG-based static hand gesture recognition pipeline for the Sign Language. This includes a novel ensemble feature selection method combining four filter-based techniques—ANOVA, Chi-square, Mutual Info, and ReliefF. A new feature combiner exploits feature-feature and feature-class correlation thresholds, resulting in reduced and representative feature subsets for classifying sign language gestures, aiming to improve recognition accuracy.

As a second contribution, we have introduced an sEMG-based hand grasp recognition (HGR) pipeline that can effectively distinguish between six different static hand grasps using a minimal number of sEMG sensors. To achieve this, we employed a game theory-based feature selection approach to determine the representative feature subset by treating the feature selection as a cooperative game with a transferable utility function. The proposed pipeline was validated using standard performance metrics, and it demonstrated competitive recognition accuracy on a publicly available dataset using single-channel sEMG. This high accuracy with minimal sensors supports the development of cost-effective devices.

The third contribution focuses on sEMG-based dynamic hand gesture recognition. We leveraged deep representation learning to build an efficient and robust

Handwritten Character Recognition (HCR) pipeline. A Stacked Sparse Denoising Autoencoder(SSDAE) network is designed to obtain an effective deep feature representation. The resultant learned feature representation is further introduced into traditional machine learning classifiers, achieving potentially superior results in HCR tasks. Our proposed approach can potentially contribute to real-time Human-Computer Interaction (HCI) systems, promoting the digitization of handwritten notes and enhancing clinical applications that analyze handwriting tasks. For our experiments, we created a new dataset, denoted as $S-HCR$, which includes EMG signals collected corresponding to lowercase English alphabets from participants while writing on paper with a pen.

For our fourth contribution, we developed an efficient multi-modal handwritten character recognition (HCR) pipeline by integrating electromyography (EMG) and inertial measurement units (IMU), building upon our second contribution, which focused solely on sEMG. The system employs a feature fusion approach, incorporating a modified deep autoencoder to achieve an effective combined data representation for both sEMG and IMU signals. For our experiments, we collected a new dataset comprising sEMG and IMU data corresponding to 26 isolated handwritten English alphabets on a whiteboard, referred to as $MM-HCR$. The feature fusion technique enhances the system's performance in terms of recognition accuracy. With its higher recognition accuracy, the proposed method shows potential for applications in digitizing handwritten notes from whiteboards in a smart classroom and in clinical handwriting analysis for conditions such as Alzheimer's.

The fifth contribution addresses high-dimensional feature vectors in the $MM-HCR$ dataset by employing a cost-effective meta-heuristic-based feature selection approach. The Enhanced Salp Swarm Algorithm, a metaheuristic optimization technique, selects relevant information from EMG and IMU features, yielding a robust representation of handwriting dynamics. This method results in a smaller feature subset vector and improved accuracy, showcasing the potential of the integrated feature selection and fusion methodology.

Another key contribution is the extensive use of explainable AI methods to enhance the interpretability of sEMG-based hand gesture recognition models. Techniques like SHAP (Shapley Additive exPlanations) are predominantly implemented where applicable to provide insights into the decision-making process of the machine learning models, ensuring transparency and trustworthiness in practical applications. However, the focus was more functional than foundational and the thesis does not attempt to build formal causal or theoretical models based on explainability.

The thesis concludes with major findings in sEMG-based static and dynamic hand gesture recognition and emphasizes the use of explainable AI methods to improve interpretability. These findings hold significant promise in biomedical signal processing, electromyography, and HCI, with practical applications in Sign Language interpretation, smart classrooms, and clinical analysis.

Keywords: Biomedical signal processing, Deep learning, Human-computer interaction, Sign language, Feature selection, Explainable AI.