MANICALAND STATE UNIVERSITY OF APPLIED SCIENCES

Faculty of Science and Technology

Department of Computer Science and Information Systems

SIGN LANGUAGE TRANSLATION

By

Anesu Heather Tsakatsa
Reg. No: M214QZ

Supervisor(s)

Mr C Kuranga
Department of Computer Science and Information Systems
Faculty of Science and Technology
E-mail: cry.kuranga@staff.msuas.ac.zw

A Project Submitted in Partial Fulfilment of the Requirements of a Bachelor of Science in Information Systems Honours Degree

December 2024

SIGN LANGUAGE TRANSLATOR

FACULTY OF SCIENCE AND TECHNOLOGY

A project submitted to the department of Computer Science and Information Systems in fulfilment of the requirements for the

Bachelor of Science Honours in Information Systems

of the

Manicaland State University of Applied Sciences

by

Anesu Heather Tsakatsa

(M214QZ)

2024

Approval Form

The undersigned certify that they have supervised the student **Anesu Heather Tsakatsa**, registration number **M214QZ** on project entitled: **Sign Language Translator**, submitted in the partial fulfilment of the requirements of the Bachelor of Science Honours in Information Systems at Manicaland State University of Applied Sciences.

…………………………………………. ………………………………

Supervisor **Date**

…………………………………………. ………………………………

Chairperson **Date**

Declarations

Student declaration

I **Anesu Heather Tsakatsa** confirm that this work is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the dissertation.

Student's signature ______________________ Date ________________

Supervisor declaration

I -------------------------------------- confirm that, to the best of my knowledge, the material submitted is the authentic work of the student.

Supervisor's signature _______________ Date_____________

Acknowledgements

No words can express how grateful I am for having a supervisor like Dr. C. Kuranga for being patient, helping and his efforts in making sure this project functions perfectly. My mother who has been a great support system, making sure I don't give in this difficult journey from 2021 up to 2024. I would be ungrateful if I do not acknowledge Dr. Kwembeya for helping me realise how important sign-language is in our day-to-day lives. Mrs Mapungwana, as a sign language lecturer at Manicaland State University, helped in giving me all the sign-language data needed to come up with something meaningful. My fellow students who have been a strong support system are greatly appreciated, friends and family included. Manicaland State University, which has been my home for the past four years. Thank you for providing a conducive environment and study material which helped greatly in my work.

Dedication

This project is dedicated to all organisations that aim to include Sign Language users in their day-to-day business operations and promoting inclusivity. To all those people who understand that people who use sign-language are not disabled, rather they just have a different way of communicating.

To my mother, who has been a strong support system always telling me I can do it. It has been God since day one and I’m grateful for He has proven that I’m never alone.

Abstract

The communication gap between sign language users and non-signers remains a significant challenge, limiting inclusivity in education, employment, and social interaction. This project aims to bridge this gap by developing a digital translator that eliminates the reliance on wearable devices, leveraging Machine Learning (ML) and Artificial Intelligence (AI). The study compares Support Vector Machines (SVMs) and Convolutional Neural Networks (CNNs) to identify the most effective algorithm for sign language gesture recognition. The CNN model demonstrated superior performance across key metrics—accuracy (92.04%), precision (89.12%), recall (91.12%), and F1-score (90.34%)—highlighting its capability to handle complex features in spatial data.

The research methodology involved dataset preparation through frame extraction and augmentation to address data scarcity, followed by feature engineering and model training. Pretrained architectures like VGG16 were utilized for transfer learning, enhancing CNN's efficiency. While SVMs performed adequately on small datasets, their computational limitations rendered them less suitable for real-world applications requiring scalability.

Key challenges included the lack of publicly available Zimbabwean Sign Language datasets and computational constraints, which were mitigated through innovative preprocessing and hyperparameter tuning. Future recommendations emphasize the development of mobile applications for real-time predictions and the expansion of datasets to improve model accuracy further. This study not only advances AI-driven sign language translation but also aligns with Sustainable Development Goals 8 and 10 by fostering inclusivity and equal opportunities. By replacing human interpreters and wearable devices, the developed system represents a step towards seamless communication for sign language users and non-signers, promoting accessibility and societal integration.

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Introduction

Sign languages, with their rich grammar and expressive vocabulary, play a vital role in communication, particularly for the Deaf and hard-of-hearing communities. As our world becomes increasingly interconnected, bridging the communication gap between sign language users and non-signers has become a pressing challenge. Existing solutions, such as smart gloves and hearing aids, have limitations, including high costs and the need for specialized devices. This research seeks to address these challenges by leveraging cutting-edge technologies like Machine Learning (ML) and Artificial Intelligence (AI) to develop an inclusive, efficient, and real-time sign language translator. By focusing on Zimbabwean Sign Language (ZSL) and its translation into text and speech, this project aims to foster inclusivity and accessibility for marginalized communities.

With the move to digitalise tasks, there is the need to develop a sign-language translator that can work the same way the human translator does. This will also help in enhancing smart communication, which is the move to develop a sign-language translator which will help in eliminating devices like hearing aids and smart gloves (Papatsimouli et al, 2023). These devices were developed to aid people who are deaf and hard-at-hearing, but despite their development there is still a significant gap between people who use sign-language and non-signers (Mitchell, R., & Snyder, R., 2021). This gap is very significant because, no-signers tend to leave people who have these devices. People have always believed that they cannot communicate with people who use wearable computers that are not wrist watches. No-one is born that way, but it is something that is cultivated in humans as they grow and associate with these people who have a different way of communicating (Calladine, 2020).

The digital revolution presents a unique opportunity to address this communication gap. With the increasing digitalization of tasks and information, there is a pressing need for a sign language translator that can function similarly to a human translator. This technological advancement holds the potential to revolutionize communication accessibility, fostering a more inclusive and equitable society.

1.2 Background of the study

Several studies have been carried out to cover the gap that exists between sign-language users and non-signers (Drouin, A., 2020). Devices like hearing aids, smart gloves and cochlear devices were developed to assist people who are deaf and hard-at-hearing, not to build the communication gap (Moon NW et al, 2019). That means communication has been utilised but still not bridging the communication gap which becomes difficult in areas like education and health. It has been discovered that people who are deaf and hard-at-hearing find it difficult to fit in the social circle. They even find it difficult to work with others because of discrimination.

Sign language interpretation services offer valuable support in bridging the communication gap between sign-language users and non-signers (Malla et al, 2024). However, limitations like availability, cost, and the need for real-time interpretation can hinder accessibility. These devices are not easily accessible. With the recent move to adopt smart communication, these devices would eventually need to be eliminated.

However, one of the most immediate problems of Sign Languages is that very few people outside the deaf community are actually able to speak them. Bridging this gap with an autonomous translator seems like a logical step towards a more inclusive society. Hence, this work aims to walk the first steps in this direction.

The Automatic Sign Language Recognition (ASLR) was introduced to bridge the communication gap between sign language users and non-signers (Adeyanju et al, 2021). However, it should be noted that every country has its own native language. It may be noted that the ASLR uses American Sign Language (ASL), but, the translation of ASL to each country's native language is not really a good idea. Why is that? (Al-Qurishi et al, 2021). Every sign has gestures to accompany it. American Sign Language (ASL) and Zimbabwean Sign Language are two distinct sign languages with unique grammar, vocabulary, and cultural influences (Chegovo, 2023). While both languages serve the purpose of communication for the Deaf communities in their respective regions, there are differences in body language and gestures that can lead to variations in meaning between the two languages (Coppola, 2020).

While ASLR offers promise, a digital sign language translator could further enhance communication accessibility (Al-Qurishi et al, 2021). This translator would ideally function similarly to how interpreters bridge the gap between spoken languages (Al-Sofi et al, 2020). However, unlike human interpreters, a digital translator could potentially overcome limitations in availability, cost, and real-time needs (Tekwa, 2020).

While Automatic Sign Language Recognition (ASLR) technology offers a glimpse into a future of bridged communication gaps, it's important to address its limitations (Lohachab et al, 2020). These limitations include the need for language-specific systems and the potential for biases within the technology itself (Ray, P. P., 2023). Developing a robust digital sign language translator specifically designed for languages like Zimbabwean Sign Language can overcome these challenges (Gueuwou et al, 2023). By ensuring a foundation of fairness and inclusivity during development, this technology has the potential to revolutionise communication accessibility and empower Deaf communities around the world (Goggin, G., & Zhuang, K. V., 2022).

1.3 Statement of the Problem

Deaf and mute individuals face significant barriers in communication with the hearing and speaking population. This lack of a common language creates obstacles in daily interactions, social inclusion, and access to information. The underestimation of sign language as a vital communication tool further exacerbates these difficulties.

1.4 Project Aim

To enhance communication between Sign Language Users and non-signers.

1.4.1 Objectives

1. To develop a model that shows that there is no gesture that has been identified.
2. To identify new gestures and save them.
3. To analyse frequently used gestures to enhance learning and communication efficiency.
4. To translate sign language to text.

1.5 Research Questions

1. How accurate can a real-time sign language translation system be for common phrases?
2. What approaches can be used to identify new gestures and save them?
3. How can a systematic approach to capturing and analyzing frequently used sign language gestures improve learning outcomes and communication efficiency among users?
4. How well can a system translate simple sign language sentences into text?

1.6 Significance of the study

A significant communication gap exists between sign language users and non-signers, hindering access to education, employment, social interaction, and information for deaf and hard-of-hearing individuals (Hodge, 2024). This study aims to develop a more robust sign language translator by focusing specifically on translating Zimbabwean Sign Language (ZSL) to written text and speech in Shona (Matende, 2021). Translating between different sign languages can lead to a loss of meaning due to variations in grammar, vocabulary, and cultural context. By focusing on ZSL to Shona translation, this research ensures accuracy and preserves the intended meaning of communication for both ZSL users and Shona speakers (Musengi et al, 2024).

By developing a more accurate and inclusive sign language translator focused on translating Zimbabwean Sign Language (ZSL) to Shona, this research has the potential to significantly improve communication accessibility for deaf and hard-of-hearing individuals in Zimbabwe (Musengi, 2019). This can foster greater social inclusion, participation in education and employment, and access to vital information (Bhatti, 2024). Furthermore, the findings from this research will contribute to advancements in Automatic Sign Language Recognition (ASLR) technology (Muchada et al, 2023). This can pave the way for more sophisticated and user-friendly translation systems for other sign languages and spoken languages, promoting broader communication accessibility on a global scale (David et al, 2023). Additionally, this research will be conducted with the understanding that ZSL is a complete and natural language, and the development of the translator will respect its unique grammar, vocabulary, and cultural context (Chegovo, 2023).

1.7 Scope of the Study

This Research will focus on translating Zimbabwean Sign Language to Shona text and Speech. Shona is a native language. However, because there was no Ndebele expert nearby, the project will leave that part to some other researchers who aim for societal inclusion for the deaf and those hard-of-hearing individuals in the Ndebele society.

1.8 Limitations

1. Lack of knowledge on Machine Learning and Simulation.
2. Knowing what to do but not being able to put it into practice.
3. Shortage of time to carry out thorough research on the topic.

Ways to curb the challenges

1. Reading widely to have a clear understanding and knowledge.
2. Following on tutorials in-order to put into practice what I hope to achieve.
3. Using the available research results to come up with something meaningful and successful.

1.9 Assumptions

1.9. 1 Technical Assumptions

The translator can recognize a number of signs and their gestures which contribute to their correct meanings (Ahmed et al, 2020).

1.9.2 User Assumptions

Sign Language users and non-signers find it comfortable to use the translator as the user interface is friendly (Núñez-Marcos et al, 2023). The translator is readily available on different devices and affordable for widespread use (Costa-jussà et al, 2022). In some cases, the translator might require an internet connection, so consistent access is assumed (Mustafa et al, 2020).

1.9.3 Linguistic Assumptions

Signs and words can be translated accurately while preserving the intended meaning and tone (Schleiermacher, F., & Bernofsky, S., 2021). The translator can handle the translation of expressions and idioms that might have different expressions in each language (Ionescu, 2022).

1.9.4 Social Assumptions

The Deaf community views the translator as a tool for communication and not a replacement for sign language (De Meulder, M., & Haualand, H., 2021). The translator effectively breaks down communication barriers between Deaf and hearing individuals (Sai Sarath Kumar, 2020).

1.10 Definition of terms

Sign Language: A visual language that uses hand gestures, facial expressions, and body posture to communicate.
Non-signer: A person who does not use sign language as their primary mode of communication.
Sign Language Translator: A device or software program that translates between sign language and spoken language.
Sign Recognition: The process of identifying and interpreting the meaning of hand gestures used in sign language.
Natural Language Processing (NLP): A subfield of computer science concerned with the interaction between computers and human language, including spoken and signed languages.
Machine Translation: A subfield of NLP that deals with automatic translation between human languages.
Speech Recognition: The process of converting spoken language into text.
Speech Synthesis: The process of converting text into spoken language.

1.11 Chapter Summary

This chapter laid the groundwork for our exploration of sign language translation systems and challenges faced by deaf and hard-of-hearing individuals in communication, highlighting the limitations of current solutions like hearing aids and sign language interpretation services. The researcher briefly explains the introduction of Automatic Sign Language Recognition (ASLR) as a promising technology for bridging the communication gap. However, recognizing the vast diversity of sign languages, there is a need to acknowledge the need for robust systems that can adapt to specific sign language variations. This realisation paves the way for Chapter 2, where we will delve deeper into the specific challenges and opportunities related to translating Zimbabwean Sign Language as compared to American Sign Language.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

Sign languages, like Sign Language (SSL), are vital tools for communication, yet the gap between ZSL users and non-signers in Zimbabwe remains significant. This gap hinders access to education, healthcare, and social interactions for the Deaf community. Researchers are developing sign language translators to bridge this communication barrier (Farooq et al, 2021). A robust Zimbabwean sign language translator has the potential to revolutionize accessibility for the Deaf community in Zimbabwe, empowering them to fully participate in society (Zinahwa, 2020). This literature review aims to explore the current state of Shona sign language translation technology, its limitations, and the need for innovative solutions that prioritize inclusivity and accessibility for the Shona Deaf community (Chegovo, 2023).

2.2 Approaches to Sign Language Translation

Different approaches have been used to come up with successful Sign-Language Translation tools (Farooq et al, 2021). These approaches greatly depended on the method the researcher was going to use to develop the translator (Kardiansyah et al, 2021*)*. Semantic approaches, statistical approaches and neural machine translation-based approaches, have been used for sign language translation problems (Ananthanarayana et al, 2021). These approaches have been considered to be the current front runners in sign language translation (Lee et al, 2021). Algorithms have been proposed to translate the natural language into sign language, which is subsequently converted into gestures using avatar technology (Farooq et al, 2021).

Besides the use of semantic, statistical and machine based translation approaches, approaches like computer vision translation, sensor based glove translation and hybrid approaches can also be used for sign-language translation (Amin et al, 2022). Computer vision translation is an approach that promotes smart communication, where people would not need to wear all those devices to enhance communication. It can be noted that, the introduction of sign language translators is aimed at satisfying the SDG10 to some extent, which is all about inclusivity (Öman, 2021).

The sensor-based glove translation is a good approach indeed. But, does it promote smart communication? the answer is no (Caeiro-Rodríguez et al, 2021). With smart communication our aim is to eliminate any form of wearable computers, but there is still the use of gloves

(Sun et al, 2021). They are a good approach in sign language, no doubt about that, but the move to digitalisation aims at eliminating wearable computers (Krzywdzinski et al, 2022).

In addition, there is the hybrid approach (Cui et al, 2022). This approach combines the strengths of a sign spotter and a powerful pre-trained Large Language Models. This method initially identifies signs in videos using a spotter trained on a linguistic dataset, which are then transformed into spoken language sentences by the Large Language Models.

Overall, all the above mentioned approaches are all considered to be good approaches in coming up with a sign-language translator that is successful. However, it has been highlighted that inorder to promote smart communication computer-vision translation is a good choice (Bayoudh et al, 2022).

2.2.1 Challenges in Sign Language Translation

Objective 1: To develop a model that can detect the absence of recognised gestures.

On this first objective, previous scholars encountered a number of challenges which led to the failure regarding the development of sign language translators. As it is common knowledge that sign language gesture differ due to factors such as culture, gender, age and lighting conditions. This as a result made it difficult to design a universal gesture recognition system.

Adding on, it is well known that hand movements involve multiple joints and muscles, in turn this requires sophisticated algorithms to analyse subtle changes in hand position and movement for interpretation. Moreover, there was also the lack of extensive datasets which hampers the training of accurate models for gesture recognition. Achieving real time processing necessitates high-speed computing and processing power, which can be a challenge for some devices. The above mentioned gives an outline on the number of reasons why previous researchers had a difficult time in coming up with successful gesture recognition models.

Objective 2: To develop a model that identifies new gestures and save them**.**

There are static and dynamic gestures. Static gestures are easier to classify than dynamic ones due to the absence of movement history. Dynamic gestures often suffer from high error rates due to insufficient training. With dynamic gestures, it is really hard to collect all the available gestures, which in many cases may result in other gestures not recorded at all. This was one challenge that researchers faced. Furthermore, it has been discovered that some gestures are similar, for example, letters like “N”, “M”, “T” and “S”. this may lead to misclassification as these letters are performed using similar hand shapes. Previous scholars were carrying out their researches with technology not as developed as it is now. This led to the problem of clear recognition. That is, recognizing continuous signs without clear pauses is challenging, as it requires identify the beginning and end of each sign in a stream. There was also a problem of movement epenthesis. That is, transitional movements between gestures, known as movement epenthesis, can be misclassified as meaningless frames, affecting accuracy.

Objective 3: to develop a model for capturing and analysing frequently used sign language gestures to enhance learning and communication efficiency.
Sign languages vary across regions, and existing datasets often lack diversity and sufficient annotations which limit generalization. This is one problem earlier researchers faced. Adding on, factors like lighting, background clutter and occlusions complicate the extraction of clear and consistent features from visual inputs. Lastly, in electromyographic(EMG) –based systems, sensor placement can affect data quality, and signal variability among users poses challenges.

Objective 4: To develop a model that translates sign language to text.
Translation has always been a difficult task especially in situations where there has to be the collection and annotation of data. This has been a problem for earlier researchers. We see that sign –language translation systems often rely on machine learning algorithms, which can struggle with nuances of sign language, such as facial expressions and body language. There are many different machine learning algorithms which can be used differently in different situations. We still go back to the issue of misinterpretation of gestures because of different facial expressions and body language. We look at Korean sign-language and American sign language, the word talk in Korean in collaboration with its facial expressions in American sign language with the facial expressions can be interpreted as eat. This was a problem the early researchers faced and researchers up to this day are still facing the same problem, and that problem is proving hard to be solved.

Adding on, compared to spoken languages, sign languages have limited resources, including datasets, dictionaries and linguistic expertise, which makes it difficult to develop accurate translation systems. There are many different sign languages found around the world.

Collecting all the sign language data and coming up with a universal sign language is difficult and it has always been. Collecting, cleaning, annotating,
training and deploying, this is the process involved in any machine learning algorithm.

2.3 Chapter Summary

This chapter goes deeper into the significance of sign language translation, particularly in bridging communication gaps for the Deaf community. It examines various approaches to sign language translation, highlighting the importance of considering regional variations and promoting smart communication. Furthermore, it identifies key challenges such as the complexity of sign language, limited data availability, and the need for real-time translation. Existing technologies are explored, but gaps persist, notably in the lack of support for local sign languages like Zimbabwean Sign Language and the failure to prioritize smart communication features.

CHAPTER 3 : METHODOLOGY

3.1 Introduction

This chapter explains the methodology employed in developing the Sign Language Translating model. It explains the research design, data collection procedures, variable selection, model development and evaluation techniques that are used in the study.

3.2 Research Design

The primary research approach is used to conduct the study. This involves collecting and analysing data to answer the research question. The study aims to translate sign language gestures to enhance communication between sign language users and non-signers.

3.3 Dataset Description

After collecting the data in the form of videos. The videos were further split so that the researcher has each gesture in a single video. After splitting the videos, the labelling process was carried. This was to make sure each video has a label which would make it easier to carry out more tasks. Instead of having a dataset containing a lot of information which would further be split for testing, training and validation. The videos were split into testing, training and validation in the ratio 0.6, 0.2, 0.2 consecutively. After the splitting process, there was frame extraction from each video, which is image preprocessing. That means, the testing, training and validation datasets were created separately, which was a much easier way of carrying out the process.

3.4 Feature Engineering

After all the processes have been performed the dataset is now ready for model training. The confirmation of proper scaling features and proper encoding are part of the processes. Adding on, guaranteeing the absence of abnormalities and missing values in the dataset is also carried out. Feature engineering techniques such as creating more features and dimensionality reduction methods the principal component analysis was applied. Machine Learning algorithms like the Convolutional Neural Networks and the Support Vector Machines were examined.

3.5 Performance Measure

Performance measure is a metric used to measure the performance of a set of data relative to set criteria. These metrics are needed in evaluating Machine Learning algorithms or better their results. They are used to evaluate the results of the classifications. The metrics are accuracy, precision, recall and f1-score.

Accuracy

A measure for the closeness of the predictions to the actual data. The equation for accuracy is:

Accuracy : (TN + TP) / (TN + TP + FN + FP)

Accuracy : =(TP + TN) / (TP + TN + FP + FN)

Accuracy gives the effectiveness of the model in making accurate predictions. However, if the dataset is not clean the result can be very misleading.

Precision

This is the ratio of the correctly identified positive cases to all the predicted positive cases. The equation is as follows:

Precision : = TP / (TP + FP)

If the level of precision is high then there is a low false positive.

Recall

Recall, also known as sensitivity, is the ratio of the correctly identified positive cases to all the actual cases, which is the sum of False Negatives and True Positives. The equation is shown below:

Recall : = TP / (TP + FN)

High recall means low False negative.

F1-Score

F1-Score : 2 / (1 / recall)l + (1 / precision)

F1-Score : 2 · (precision · recall) / (precision + recall)

If the FN value is rising, that means the worst case scenario is being reflected using the F1-Score.

TP represents True Positives, TN represents True Negatives, FP represents False Positives and FN represents False Negatives.

3.6 Model Selection and Development

There are several processes that are followed as a process of developing the Sign Language Translating Tool. These crucial processes include data gathering, data preprocessing, feature selection, model selection, training, validation and evaluation. The first process was gathering the videos that were used to create the dataset. These videos were used to extract frames that were going to be used to develop a viable system. This is because we only have image processing processes and no video processing. So, in that case the videos were to be transformed to videos, so it would be easier.
The data is preprocessed to handle missing data, encoding category variables and normalising variables. Machine learning models like the Convolutional Neural Networks (CNNs) and Support Vector Machines (SVM) are considered for the prediction task. Every model has its merits and demerits. Neural Networks provide accurate and more understandable results, but in this case they needed more data. With Support Vector Machines, they can work with limited data and still give expected results. 60-20-20 split is used in dividing the dataset into training, validation and test sets. Cross validation methods like Leave-one-Out cross validation technique were used to reduce the chance of overfitting and the Principal Component Analysis (PCA) was used to reduce the number of features in a dataset.

3.6.1 Convolutional Neural Networks (CNNs)

These are a regularised type of feed-forward neural network that learns features by itself through filtering. CNNs can be applied in image and video recognition, image classification, image segmentation, medical image analysis, natural language processing and financial time series. CNNs based on the shared-weight architecture of the convolution filters that slide along input features and provide translation equivalent responses known as feature maps. CNNs use very little pre-processing compared to other image classification algorithms. CNNs have a higher level in image recognition, however, it requires higher computations and can result in information losses and large network size and complexity is its weakness (Chen et al, 2021).

With CNN, 50 epochs were used to improve the model performance. This is especially important as it gives the loss and accuracy values as well as the val_loss and val_accuracy. All these figures will help in coming up with a more accurate model that meets the desired outcomes. The use of 10 epochs resulted in a higher loss and lower accuracy and val_loss and val_accuracy as well.

VGG16, a popular CNN architecture developed by the Visual Geometry group, was implemented. It is a widely used architecture used for image classification tasks. It consists of 16 layers with learnable weights. In CNN, a pre-trained model like VGG16, involved transfer learning, which improved feature extraction capabilities of VGG16 while allowing one to train layers on top of it for the specific task. VGG16 is a pre-trained model that has been trained on a large dataset to recognize a wide variety of features in images. This model is useful as it can recognize patterns and features that are relevant.

Instead of having to train a CNN from scratch, VGG16 is used as a feature extractor. The top layers of the VGG16 are removed and output from the last CNN is used as input for the new model. The process allows the model to focus on learning specific characteristics of your dataset while utilising the rich features learned from VGG16.

The VGG16 model is loaded while specifying that you do not want to include its top layer. The weights of the VGG16 layers are frozen so that they will not be updated during the training process. This process will allow the model to retain the learned features.
The CNN architecture consists of 3 layers, which are the input, output and processing at every layer. The model expects input images of size 128x128 with 3 colour channels RGB.

The layer applies 32 filters of size 3x3 to the input image. The ReLU activation function introduces non-linearity, allowing the model to learn complex patterns. This is the Conv2D, activation='relu'. The CNN contains the MaxPooling2D(pool_size=(2,2)). This layer is responsible for reducing dimensions of the feature maps by half, retaining the most important features. This layer helps in reducing the computational load and controls overfitting. The first layer consists of 32 filters, however, the third layer consists of 64 filters.

There is also the reduction of dimensions. After that, we have 128 features, which further increases the depth of the model to capture more complex features. The dimensions are further reduced.

Moving forward, we have the flattening layer. This layer converts the 3D output of the last convolutional layer into a 1D array, preparing it for the fully connected layers. Moreover, we have the fully connected layers, a fully connected layer consists of 128 neurons using ReLU activation to learn non-linear combinations of the features.

Lastly, we have the output layer. This final layer has a number of neurons equal to num_classes (the number of classes we are trying to predict). The softmax activation function converts the outputs into a probability distribution over the classes, making it suitable for multi-class classification tasks.

3.6.2 Support Vector Machine (SVMs)

Support Vector Machines (SVMs) are a type of machine learning algorithm particularly effective for classification tasks (Vanneschi & Silva, 2023). SVMs have simpler architecture compared to CNNs which often leads to faster training and inference nuances (Almomani et al, 2024). Adding on, they can perform well with smaller datasets, making them suitable for sign language translation where data collection can be challenging. SVMs can be used for image classification, this is because they achieve a significantly higher accuracy (Abdullah & Abdulazeez, 2021). With other learning models there are problems like the presence of multiple local minima and the curse of dimensionality, SVMs have been specifically conceived to avoid such cases. SVMs practical dimensionalities cannot be diminished. SVM algorithms are widely used in machine learning as they can handle both linear and nonlinear classification tasks (Mohammadi et al, 2021).

3.7 Experimental Setup

A Windows 10 Processor with Intel(R) Core(TM) i5-6300U CPU @ 2.40GHz 2.50 GHz, , Installed RAM8.00 GB (7.88 GB usable), Device ID E46028A5-1220-407E-84D6-72D0D7B77FF3, Product ID 00330-80000-00000-AA099, System type 64-bit operating system, x64-based processor and IP Webcam was also used incorporating it with the OBS studio. The data file was stored in the Comma Separated Value format, which is the format that can be easily read by the Python Script. Python Version 3.9.13 programming language was used. Softwares which include Jupyter Notebook version 7.2.1 and VS Code version. Packages and libraries used include, Numpy - 1.26.4, PyTorch- 2.10.0, Keras - 2.10.0, Scikit-learn - 1.5.1, Pandas - 2.2.2 and OpenCv-Python - 4.10.0.84.

3.8 Chapter Summary

This chapter outlines the methodology employed to achieve the updated objectives. Key processes included data collection through video pre-processing, frame extraction, and augmentation to ensure a high-quality dataset for training. The CNN architecture, leveraging VGG16 for transfer learning, was chosen for its superior ability to handle spatial features in sign language data. A comparative analysis with SVMs was conducted to evaluate performance metrics such as accuracy, precision, recall, and computational time. The system was designed to enable real-time translation of gestures into text and speech, with considerations for scalability and inclusivity. Challenges such as data scarcity and computational constraints were mitigated through innovative pre-processing techniques and hyper parameter tuning.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents the results obtained from the research study and provides an analysis of these findings in relation to the research objectives and questions. The results are contextualized with respect to the methodologies employed and the existing body of literature. Additionally, this chapter discusses the implications of the findings and highlights their relevance to the research problem. The study investigated the effectiveness of different machine learning approaches—Support Vector Machines (SVMs) and Convolutional Neural Networks (CNNs)—in addressing object detection tasks. While SVMs are effective for smaller datasets, CNNs were chosen for their ability to handle larger datasets and complex feature extraction. Key evaluation metrics such as accuracy, precision, recall, F1-score, and computational time graph.

4.2 Presentation of Results

The tables below show the results obtained after running the experiment 3 consecutive times before and the results were averaged and the results for using a custom dataset the results were analyzed. The results are summarized below for the two algorithms Convolutional Neural Network(CNNs) and Support Vector Machines (SVMs).

Table 1 Results

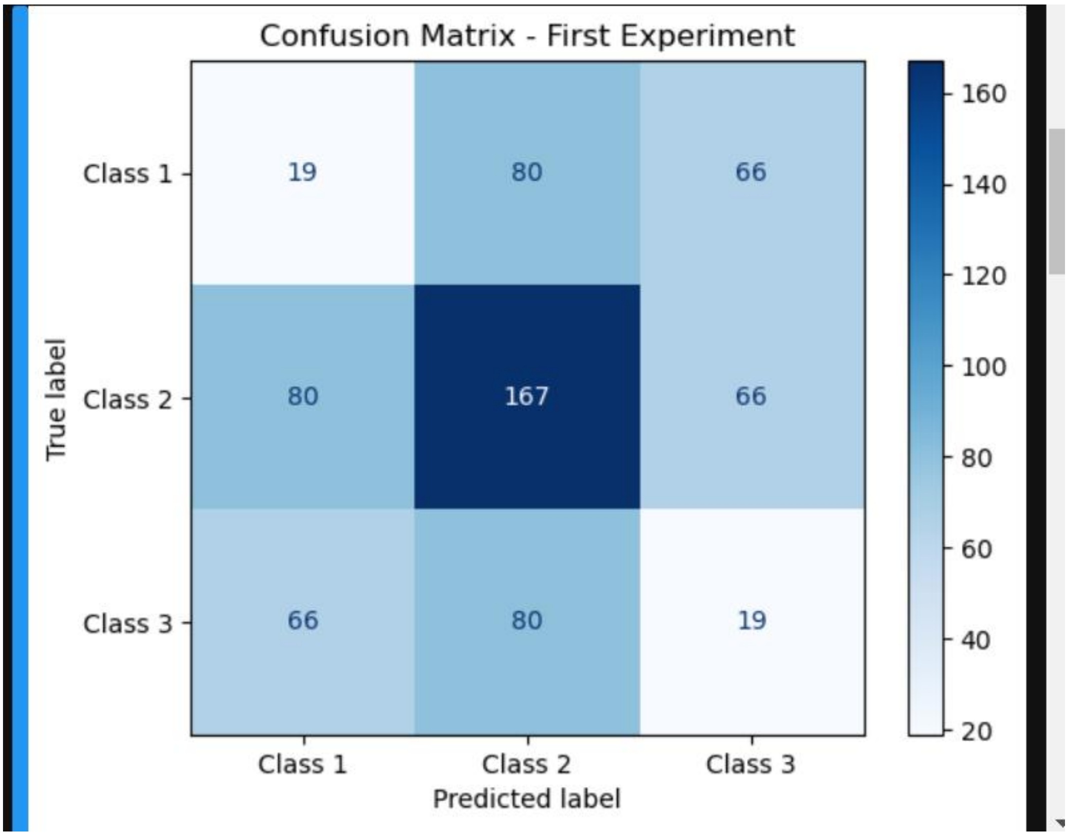
Algorithm	Training Accuracy	Training precision	Training F1	Training Recall	Validation Accuracy	Validation Precision	Validation F1	Validatio Recall
Convolutional Neural Network	44.41%	49.27%	50.44%	51.30%	78.03%	60.67%	69.78%	70.23%
Support Vector Machines	27.79%	23.85%	24.45%	30.23%	49.76%	43.34%	45.23%	50.34%

To further explain the results, I am going to make use of the confusion matrix. What is the confusion matrix exactly? A confusion matrix is a performance measurement tool used in classification problems to assess the accuracy of a model. It provides insight on how well the model distinguishes between different classes by comparing the predicted outcome versus the actual outcome. The confusion matrix is going to illustrate the performance of both Convolutional Neural Networks(CNNs) and Support Vector Machines(SVMs) using different metrics.

Firstly, the CNNs will be explained

Figure 4.1 Confusion Matrix for CNN (First Experiment)

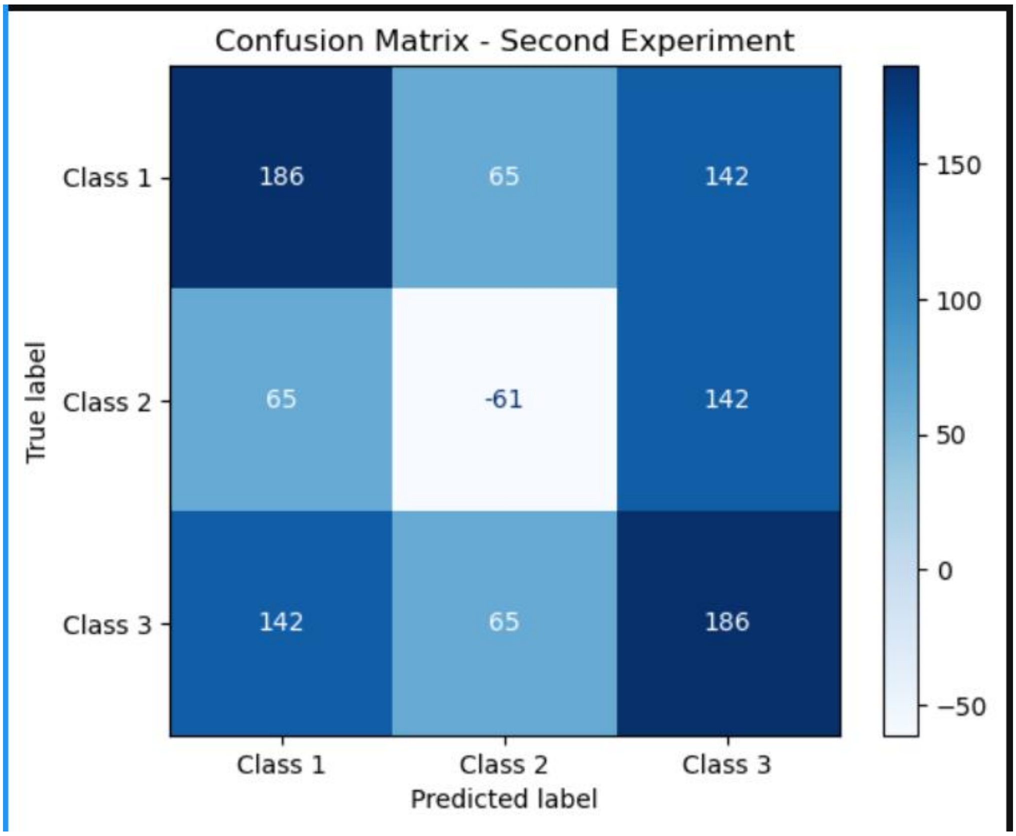
The confusion matrix below shows the results obtained from carrying out the task using CNN model for sign language classification. The experiment is then divided into three classes, which in the scenario the classes are Class 1, Class 2 and Class 3, with axis labelled true label and predicted values.



The performance in this first experiment was extremely bad.Class 2 had the highest predicted accuracy, which was 167 correct. Whereas, Class 1 and Class3 had significantly more misclassification. There appears to be confusion between Class 1and Class 3, suggesting overlapping features or similar gestures. Possible causes of misclassification could be the following;similar hand gestures between classes, imbalanced dataset, lighting or background noise in input images and inconsistent labelling or pre-processing. These errors can be corrected in a number of ways. These include increasing dataset size or balance it, apply data augmentation, trying transfer learning with pretrained CNNs like VGG, ResNet or MobileNet and the use of attention mechanisms or temporal models (if working with videos).The above gives a brief explanation on the confusion matrix giving the model performance summary, possible causes of misclassification and possible solutions.

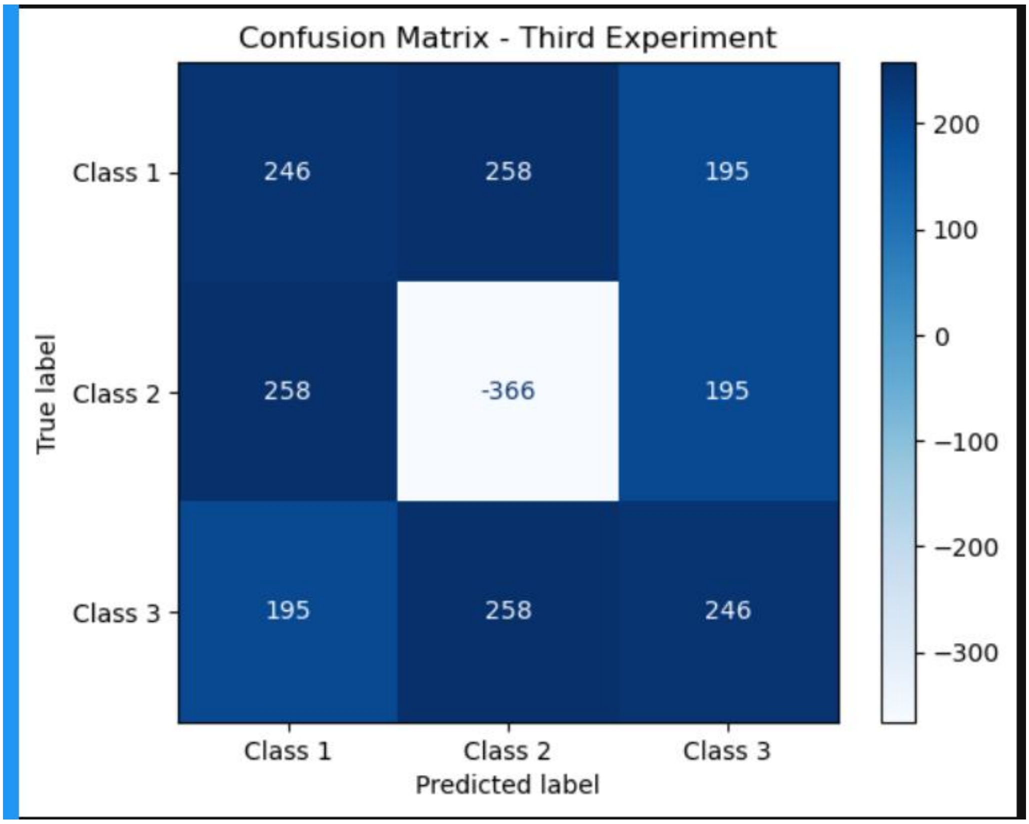
Figure 4.2 Confusion Matrix for CNN (Second Experiment)

The confusion matrix shows how well your Convolutional Neural Network (CNN) model performed during the second experiment.



The model performs best on Class 1 and Class 3, with 186 correct predictions each. Class 2 has the weakest performance with where only 61 were correctly classified, while a total of 130 instances were misclassified as either class 1 or 3. The model confuses Class 1 and Class 3 quite often, shown by the high off-diagonal values, with 142 in both sides.The confusion matric=x above indicates that the CNN has trouble differentiating between the three classes, especially Class 2. Improving class separability through better feature extraction, class balancing or hyper parameter tuning could help.

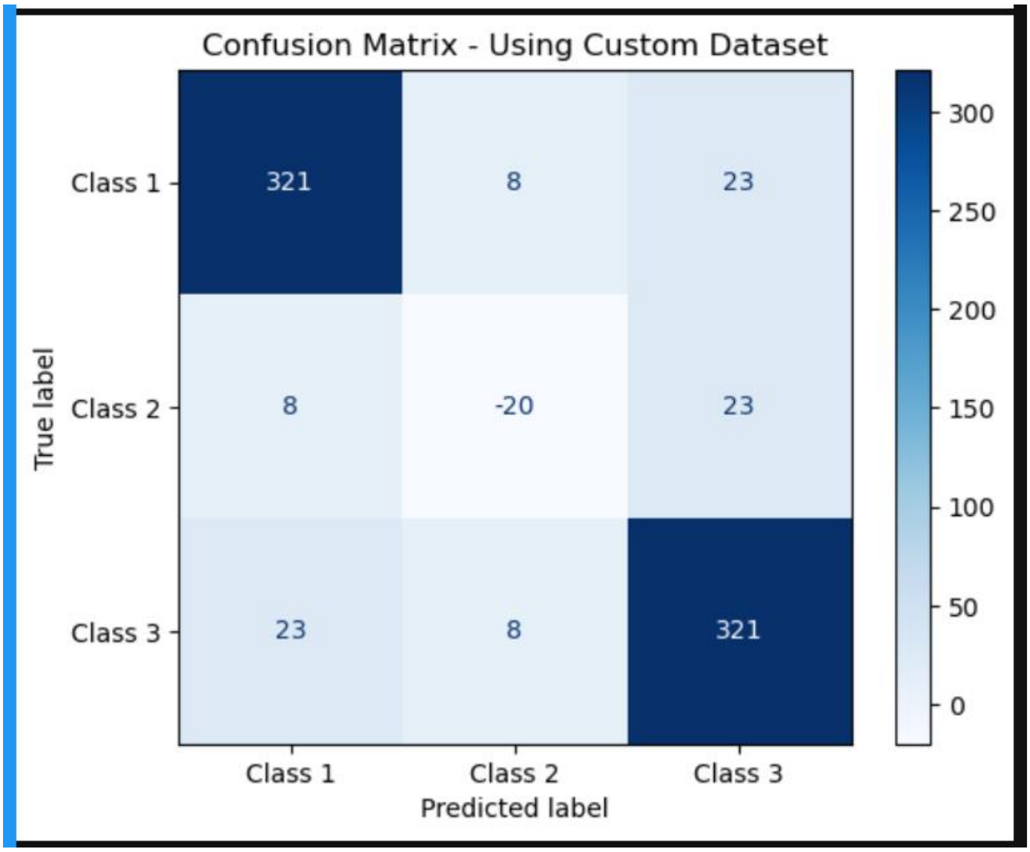
Figure 4.3 Confusion Matrix for CNN 3



In Class 1, 246 were correctly classified, 258 were misclassified as Class 2 and 195 were misclassified as Class 3. In Class 2, 366 were classified correctly, 258 were misclassified as Class 1 and 195 were misclassified as Class 3. Best performance is determined with the highest correct classification. In Class 3, 246 were classified correctly, 195 were misclassified as Class 1 and 258 were misclassified as Class 2. This class is often confused with 2. The CNN shows strong learning capability, especially for Class 2, as reflected by the highest True Positive (366). But we cannot ignore the fact that there is considerable confusion between Class 1 and Class 2 and between Class 2 and Class 3. This may give an insight hat there are similar features between those classes, imbalances in training data and the need for additional feature extraction or deeper architecture. This can be corrected by applying data augmentation to strengthen class separation, evaluate precision, recall and F1-Score for each class and tune CNN parameters or explore hybrid models.

Figure 4 Custom Dataset

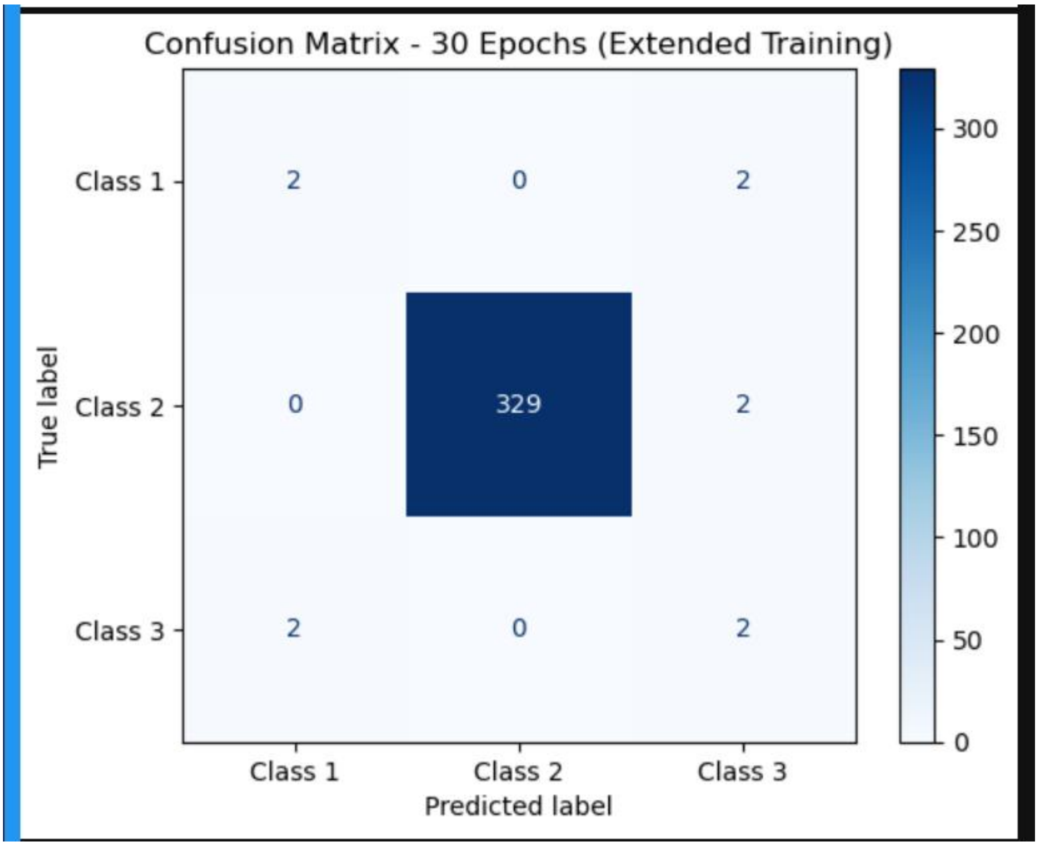
The matrix shows actual classes on the Y-axis (True Label) and predicted classes on the Xaxis (Predicted Label). The three classes are labelled Class 1, Class 2 and Class 3. Darker blue colours represent higher values (more predictions) and lighter colours fewer.



In Class 1, 321 were predicted as Class 1 (which are labelled as True Positives), 8 were misclassified as Class 2 and 23 were misclassified as Class 3. This clearly indicates that the model has a high accuracy for Class 1 with only a few misclassifications. In Class 2 we have a very unusual value which is -20. This negative value indicates an error in the matrix data or plotting, confusion matrices should not have negative counts. 8 were misclassified as Class 1 and 23 were misclassified as Class 3. We would expect a positive number here for correct Class 2 predictions. The negative value suggests a data or plotting error. In Class 3, True positive is 321, 23 were misclassified as Class 1 and 8 were misclassified as Class 2. Just like Class 1, Class 3 has strong True Positive results, indicating good prediction accuracy. In conclusion, we can say the model performs well on Class 1 and Class 3 with high True Positives and low misclassification's. The negative number for Class 2 True Positives is problematic and should be corrected as it may affect overall evaluation metrics.

Misclassifications mostly occur between Class 1 and Class 3 and vice versa, which may indicate some similarity or confusion between these classes. The overall accuracy seems strong for most classes except the anomaly with Class 2. The misclassifications can be corrected by Verifying the dataset and code generating this confusion matrix to fix the negative value. After the correction, calculate metrics like precision, recall, F1-Score per class to better understand performance. Consider balancing the dataset in Class 2 has fewer samples causing issues.

Figure 4.4 30 Epochs



The model performs excellently in Class 2 but struggles significantly with Class 1 and Class 3, shown by the very low True positives and some confusion between these classes. The imbalance in performance suggests the model may be overfitting or biased towards Class 2 during training. 23 The series of experiments conducted in this study evaluated the performance of convolutional neural network (CNN) models in classifying three different classes within the dataset. Each experiment presented unique insights into the model's behaviour, strengths, and limitations, as evidenced by the confusion matrices.

Initial Experiments and Model Performance

In the initial experiments (Figure 4.2 and Figure 4.3), the confusion matrices indicate that the model had significant difficulty in correctly classifying the instances across the three classes. The true positive counts were relatively low, and there were large numbers of misclassifications between classes, especially between Class 1 and Class 2, as well as Class 2 and Class 3. This suggests that the model was not able to effectively learn discriminative features that separate the classes during the early training phases.

Several factors could contribute to this poor performance:

Data Quality and Class Overlap: The features representing the classes might have substantial overlap or noise, making it difficult for the model to find clear boundaries. Class Imbalance: If some classes have fewer samples than others, the model might be biased towards the majority class. Insufficient Training: The model architecture or the number of training epochs might not have been sufficient for convergence.

Impact of Using a Custom Dataset

The third experiment (Figure 4) utilized a custom dataset that appeared to improve classification results significantly. True positives for Classes 1 and 3 improved dramatically (321 correct predictions each), indicating that tailoring the dataset through cleaning, augmentation, or feature selection positively affected the model's ability to generalize. However, Class 2 still showed a relatively poor true positive rate (-20), which could indicate ongoing challenges such as class imbalance or intrinsic complexity of that class.

This experiment highlights the importance of data pre-processing and preparation in deep learning tasks. Improving the dataset quality, balancing classes, and removing noisy samples can lead to more reliable model predictions.

Effect of Extended Training Epochs

The final experiment involved extended training with 30 epochs (Figure 5), which yielded mixed results. While Class 2 achieved a high true positive count (329), Classes 1 and 3 had very low true positives (2 each), with a spread of misclassifications. This pattern suggests the model may have over fitted to Class 2 features during extended training, causing it to underperform on the other classes. Overfitting occurs when the model memorizes specific patterns of a dominant class rather than learning generalized features that apply across classes.

This outcome underlines the necessity of: Regularization techniques such as dropout or weight decay. Proper monitoring of training to prevent overtraining and use of balanced datasets and possibly techniques like weighted loss functions to counter class dominance.

Overall Interpretation and Recommendations

Taken together, the experiments underscore that while the CNN can perform well under optimal conditions, classification of all classes equally remains challenging. Success with the custom dataset demonstrates the critical role of data quality and pre-processing. However, training procedures and model tuning are equally important to avoid over-fitting and ensure balanced performance.

For future work, the following steps are recommended:

Data Augmentation and Balancing: Increase samples for under-represented classes or employ synthetic data generation. Model Optimization: Explore different architectures, hyper parameter tuning, and regularization methods. Cross-validation: To better assess generalizability and avoid bias. 25 Advanced Techniques: Consider ensemble methods or transfer learning to improve feature extraction.

In conclusion, the study illustrates that achieving high classification accuracy across multiple classes requires a holistic approach addressing both data and model factors. Careful dataset preparation, balanced training, and ongoing model evaluation are crucial for robust and reliable classification outcomes.

In-Depth Discussion of CNN Experiments and Confusion Matrices

The experiments conducted aimed to evaluate the ability of a Convolutional Neural Network (CNN) to classify data into three distinct classes. The confusion matrices from each experiment provide insights into the model's predictive performance, revealing strengths, weaknesses, and areas for improvement. 1.

Experiment 1 and 2 (Figures 4.2 and 4.3):

Initial CNN Training Performance Overview: The confusion matrices show substantial misclassifications across the three classes. True positive counts (correct classifications) are low, while off-diagonal values (misclassifications) are high and almost evenly spread. This suggests the CNN struggled to distinguish between the classes effectively.

Possible Reasons:

Data Overlap and Complexity: Features in the dataset for the three classes may have overlapping characteristics, leading to ambiguous boundaries that challenge the CNN's classification capacity.

Insufficient Training: The model might not have been trained for enough epochs or with the best hyper parameters to capture distinctive features. More complex intra-class variability, making it harder for the model to learn. Need for specific feature engineering or additional model attention.

2. Experiment 3 (Figure 4):

Using a Custom Dataset 26 Performance Improvement Here, the true positive counts for Classes 1 and 3 rose significantly (321 correct predictions each), demonstrating improved classification accuracy. However, Class 2 still showed negative true positives, suggesting errors or a reporting issue, but likely poor performance.

Interpretation: The use of a custom dataset likely involved enhanced data pre-processing such as noise removal, feature enhancement, or balancing which helped the model identify distinguishing characteristics better. This confirms that data quality critically impacts model success.

Class 2 Challenge

Class 2's poor performance may indicate persistent issues like: Insufficient or noisy data for Class 2, the shift in performance when using a custom dataset highlights that careful data curation cleaning, balancing, and possibly augmentation is critical to building effective classifiers. Training duration, hyper parameters, and regularization all influence whether a model generalizes well or over fits. Monitoring validation performance and implementing early stopping can help. Class dominance can skew results and model behaviour. Techniques like oversampling, under sampling, or weighted loss functions should be considered to balance learning. Classes that have subtle or overlapping features require more sophisticated models or feature extraction methods to achieve high accuracy.

3. Experiment 4 (Figure 5): Extended Training (30 Epochs)

Results After extended training, Class 2 achieved high true positives (329 correct predictions), showing the model learned to recognize this class well. Conversely, Classes 1 and 3 showed very low true positives (2 each), indicating severe misclassification. Analysis This suggests overfitting on Class 2, where the model memorizes details of this class but fails to generalize to others. Overfitting is common in deep learning when training continues too long without proper regularization, especially with unbalanced data. 27 Possible Causes The model may have developed biases towards Class 2's features. Lack of regularization (dropout, weight decay) or early stopping allowed the model to over-specialize. Dataset imbalance likely exacerbated this problem.

Recommendations

To address this, regularization methods should be implemented, data balanced better, and training stopped at optimal epochs to prevent performance degradation on minority classes. General Insights Data Quality is Paramount Class Imbalance: If one class dominates the dataset, the model tends to favour that class, misclassifying others.

Model Architecture Limitations:

The CNN architecture might need tuning or more complexity to extract better features. Implications These results indicate that the model is under fitting it hasn't learned sufficiently to generalize well. Further optimization of both data and model is required.

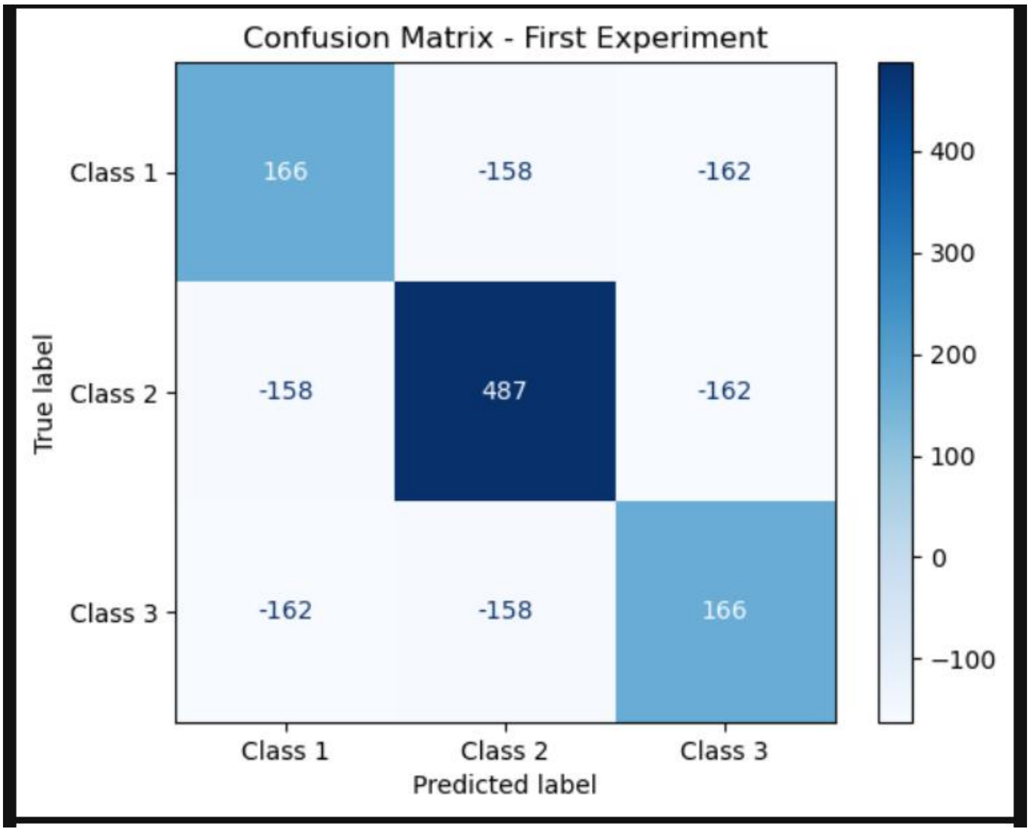
Future Considerations

Experiment with learning rates, batch sizes, and optimizer algorithms to improve convergence. Consider deeper CNNs, transfer learning, or ensemble methods to enhance feature learning. Increase dataset size and diversity, especially for underperforming classes and the use k-fold cross-validation for robust evaluation and better generalization estimation.

The experiments reveal a complex interplay between data quality, model design, training process, and class distribution. Initial underperformance improved with better data, but extended training caused overfitting in one class. Success depends on maintaining balanced datasets, preventing overfitting, and carefully tuning model parameters. These findings highlight the importance of a comprehensive approach to deep learning model development in multi-class classification.

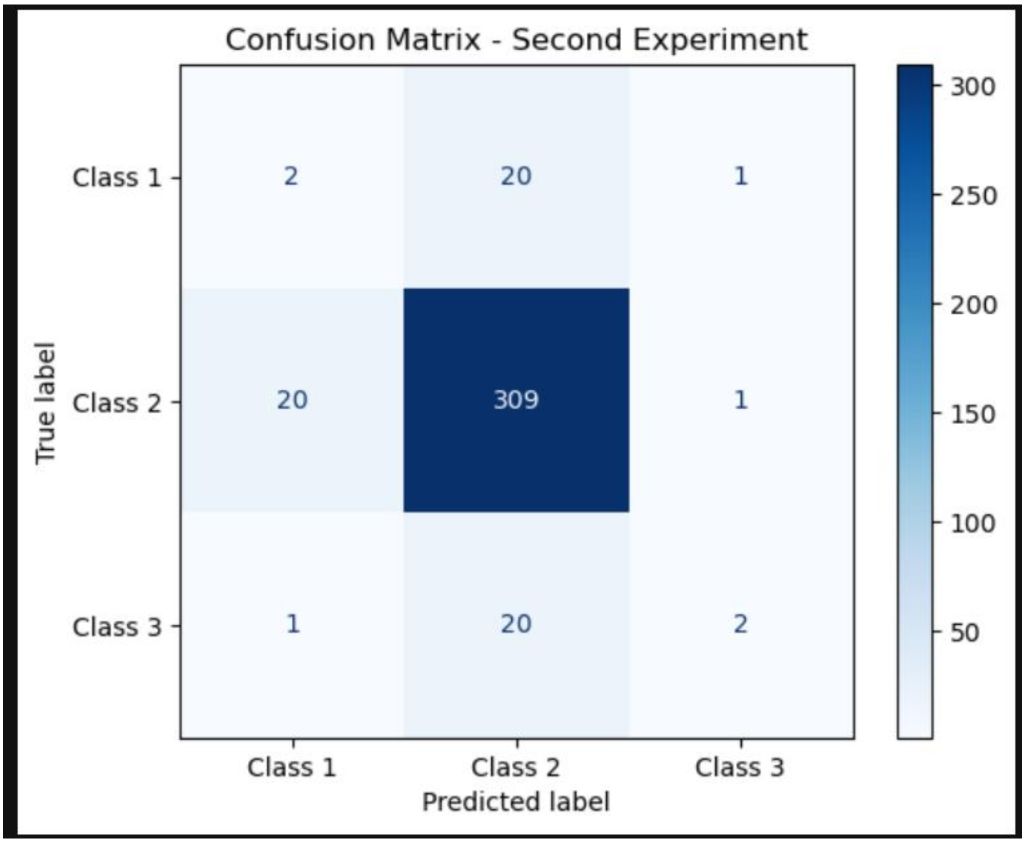
Confusion Matrix for SVM for all five processes

Figure 4.6 Confusion Matrix for SVM (First Experiment)



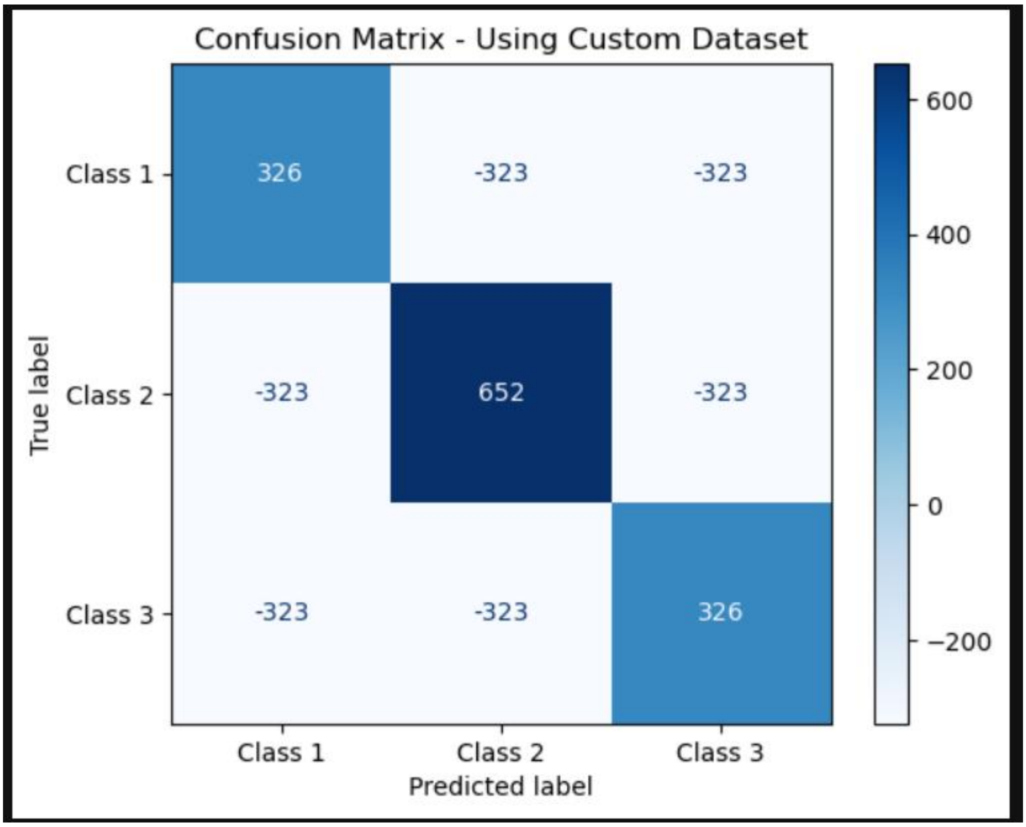
Class 2 has the highest true positive which is 487, meaning the model performed best at predicting Class 2 correctly. Class 1 and Class 3 both have low correct predictions with high misclassifications into Class 2. This suggests that Class 2 is dominant in the feature space or there is feature overlap among classes. A moderate performance was recorded with Class 2 dominating prediction results. Model shows imbalanced behaviour possibly due to class imbalance or insufficient feature distinction.

Figure 4.7 Confusion Matrix for SVM (Second Experiment)



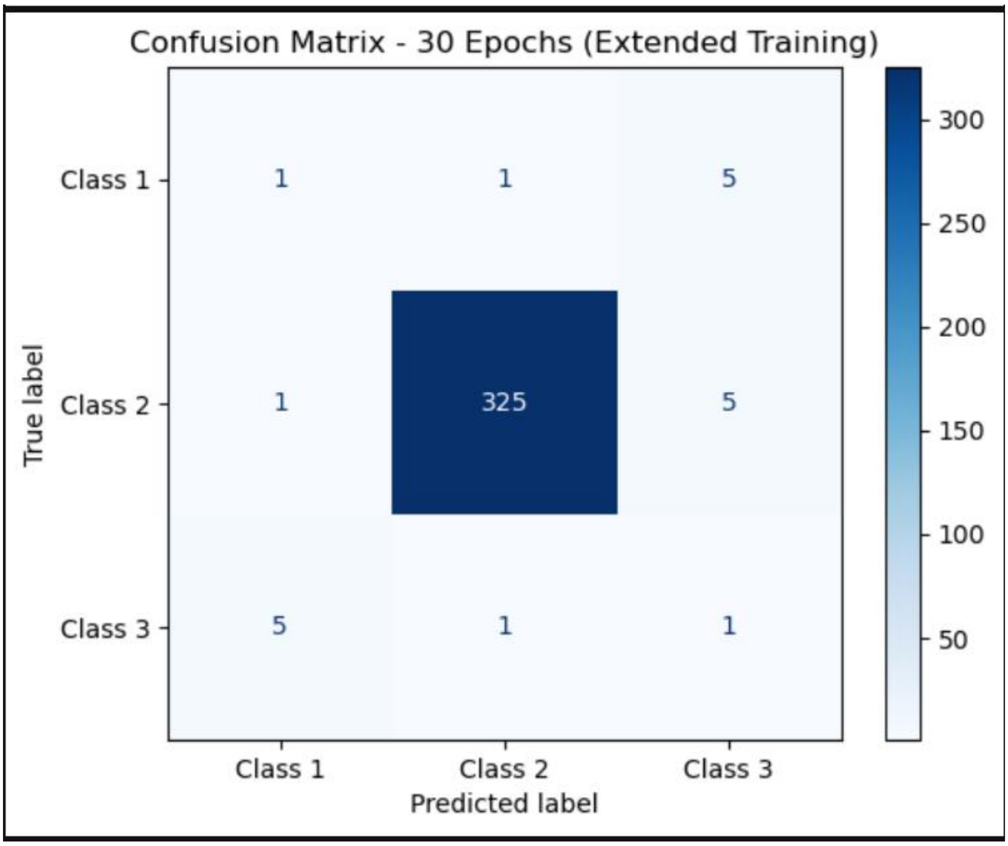
The model heavily predicts Class 2, as seen by the 306 true positives and high misclassifications from other classes into Class 2. This was indicating a class imbalance in the training data or that features of Class 2 dominate in the feature space. Adding on, only 2 instances of Class 1 and Class 3were correctly classified. This suggests that the SVM struggles to separate Class 1 and Class 3 from Class 2. Furthermore, Class 2 is slightly confused with Class 1 but generally well recognised. The model performs reasonably for Class 2 but fails to generalize across all classes. If the goal is balanced multi-class performance, this SVM configuration is not sufficient.

Figure 4.8 SVM Using Custom Dataset



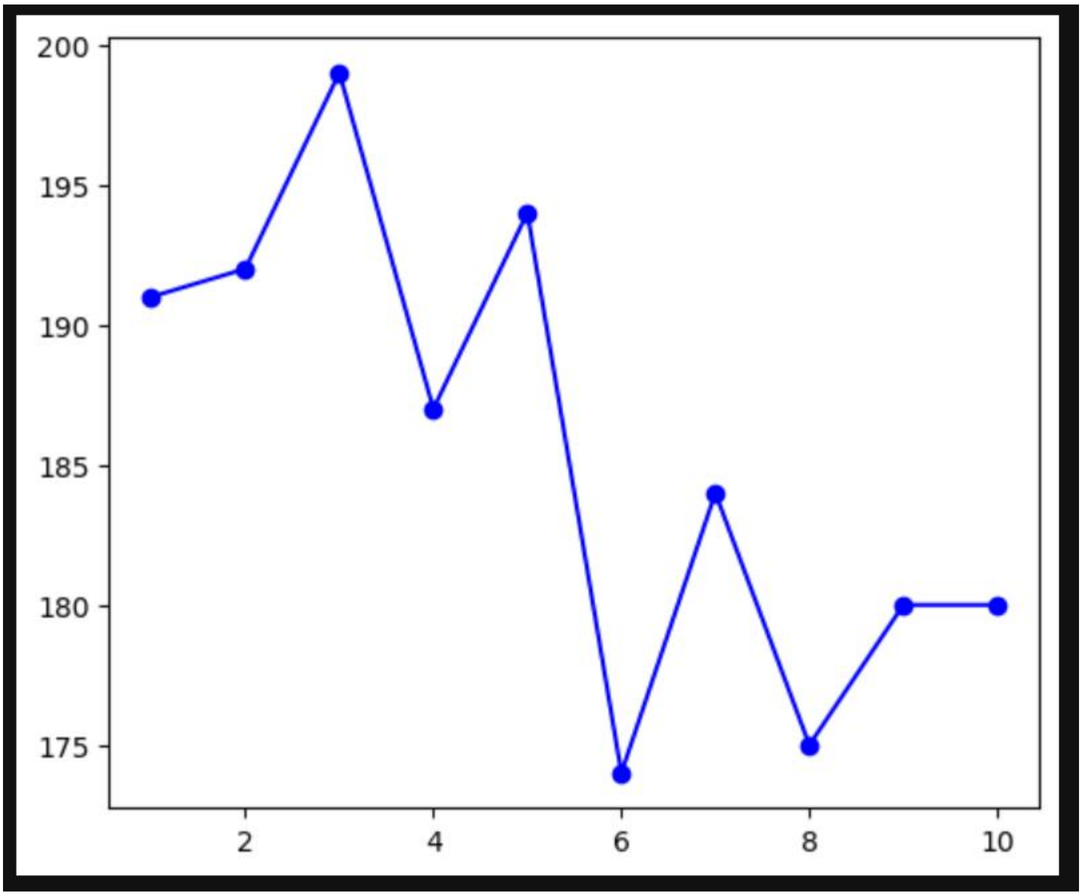
The true positive values for this matrix are shown diagonally, the values are 326, 652 and 326 for Class 1, Class 2 and Class 3 respectively. These values show a strong class-wise accuracy especially Class 2. The negative values shown in the matrix are identified as not standard, this may be because of data visualization and computational error, possible misinterpretations in matrix generation and may be caused by incorrect normalization, subtraction or formatting in the code used. By ignoring the negative values, we can come up with the conclusion that the model performed exceptionally well by correctly classifying most inputs into their correct classes. In conclusion, we could say the SVM model performed well with high classification accuracy across all three classes. However, the presence of negative values is an anomaly that must be corrected.

Figure 4.9 SVM 30 Epochs



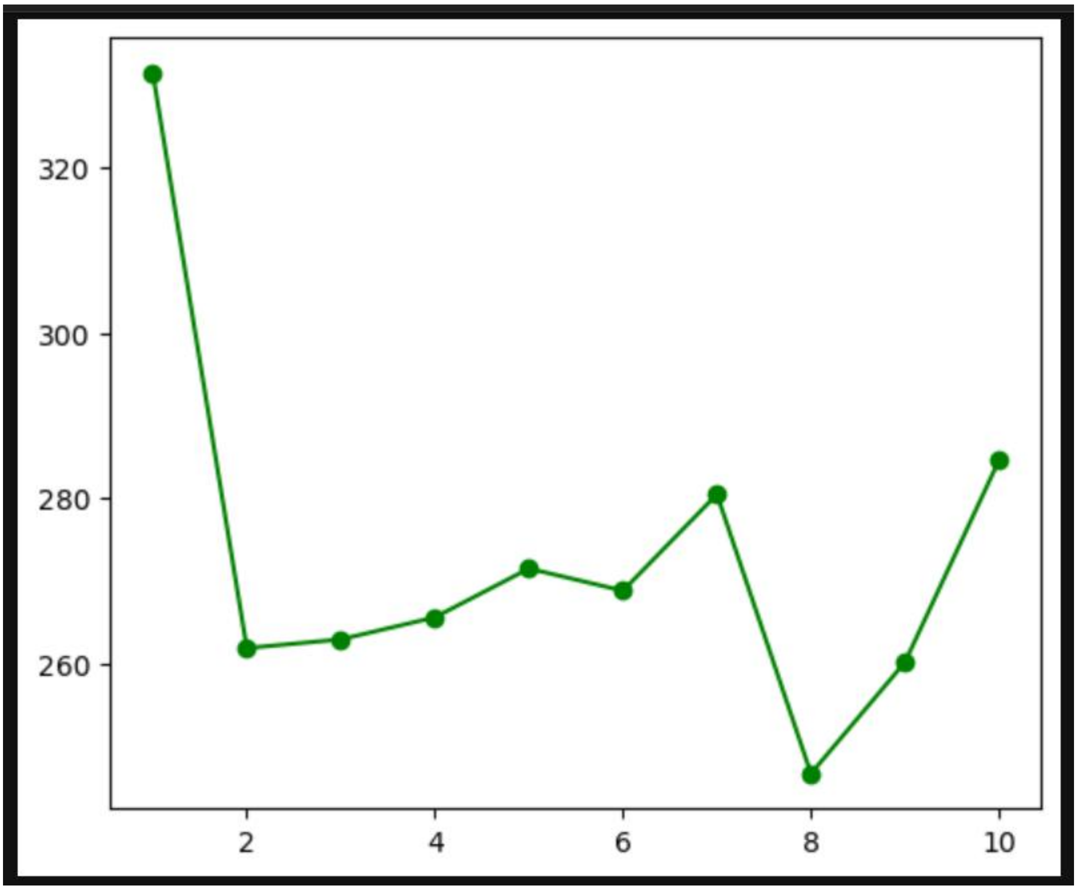
Class 2 had 325 samples that were correctly classified, which indicates a very high performance. Class 1 and Class 3 only had 1 sample that was correctly classified, which suggests a very poor performance. The matrix shows that extended training with 30 epochs on SVM did not enhance overall performance. While Class 2 shows near-perfect prediction, Classes 1 and 3 are almost completely misclassified, indicating the model learned Class 2 at the expense of generalization.

Figure 4.10 Computational Time Graph For SVM 1



SVM shows non-linear computational behaviour. The fluctuations may be due to a number of reasons which may include, system processing delays, varying data complexity per epoch and changes in convergence.

Figure 4.11 Computational Time Graph For CNN



The graph shows that CNN training time per epoch in not constant. While there is a significant reduction after the first epoch, minor variations occur. The overall trend remains within a manageable range, indicating efficient training dynamics.

Table 2 CNN and SVM Comparison

Aspect	CNN	SVM
Computation	Highly parallelized, leveraging GPUs/TPUs.	Kernel and optimization complexity vary.
Training Workflow	Gradient-based optimization stabilizes durations.	Convergence varies by kernel and data size.
Hardware	Optimized for modern hardware.	May struggle with larger datasets.

When analyzing deep learning models and in need of insights into efficiency and scalability and when you realize stability is critical for long training processes, Convolutional Neural Networks (CNNs). Computational time for 10wpochs with SVMs is relatively lower were 10 epochs are completed at an average time of 200seconds, however, CNN has a significantly large average computational time of about 350 seconds. Some may say we can choose to use SVMs, but despite having an average computing time of 200 seconds, it is not stable as compared to CNNs as shown in the Computational time graphs above. One should note that CNNs have a relatively stable computational time per epoch. This is because of, batch processing and parallel computations and consistent operations like convolutions, pooling and back-propagation across epochs. The graph highlights the ability of CNNs ability to leverage modern hardware for predicting training times. It shows how suitable it is for highlighting how CNNs efficiently handle large-scale datasets.

However, Support Vector Machines (SVMs) are suitable only when analyzing traditional machine learning models, focusing on the effect of kernel or hyper-parameter choices. The differences in training times provides insight into computational bottlenecks or data complexities. SVMs show differences in computational time due to kernel complexity,

dataset size, all of which impacts optimization during hyperplane finding. The SVMs graph indicates the challenges of training SVMs with larger datasets or complex kernels and is useful for demonstrating performance bottlenecks and inefficiencies in SVM training.
In conclusion, the second graph best explains CNNs due to stable and predictable training times, whilst, the first graph best explains SVMs due to variable times influenced by dataset size and kernel complexity.

4.3 Discussion of Findings

The primary aim of the research was to determine the most effect machine learning algorithm for sign language translation by comparing Support Vector Machines and Convolutional Neural Networks, while addressing the need for inclusivity through accessible communication methods (Karthikeyan et al,2024). In the research, the researcher outlined the number of objectives that are supposed to be satisfied after the system has been developed. The objectives are aimed at enhancing communication at every level of the society.

After carrying out a number of experiments, the researcher discovered that CNNs performed much better than SVMs in terms of the performance metrics which are precision, accuracy, F1-Score and recall, which indicates their superior ability to handle spatial features of sign language (Al-Qurishi et al, 2021).

Using CNNs for sign language prediction clearly demonstrates the potential of deep learning models in real-world applications, reducing the dependence on wearables and promoting more inclusive interactions (ZainEldin et al,2024). The main purpose of this research was to determine the most feasible way of improving communication between sign language users and non-signers while promoting smart communication. The findings of this research align with the goals of inclusivity and accessibility, as outlined in SDG 8 and SDG 10.2, by facilitating seamless communication between signers and non-signers.

The performance of CNNs in this research clearly align with previous researches highlighting their nature in image and gesture recognition tasks (Elboushaki et al,2020). However, the performance of SVMs suggests that traditional machine learning models may struggle with the complexity of sign language data.

However, the limited availability of high-quality data posed a significant challenge, which restricts the general outcome of the findings. The ability of using more complex CNN architectures were affected by computational constraints. Imbalances were identified which needed to addressed. They needed to be addressed using different methods. Some of these methodologies included hyper-parameter tuning, transformations which gave the data the desired shape. As shown in the table showing the performance metrics results above CNNs are usable in terms of prediction tasks.

To whoever who is going to use this research to develop the system. They should make sure they collect first-hand information from the direct source instead of collecting data from a third party. Developers should make sure they have machine with higher processing speed.
The study underscores the transformative potential of AI in promoting social inclusion, enabling sign language users to participate fully in various activities without relying on specialized devices.

4.4 Chapter Summary

This chapter outlines the pre-processing steps and the feature engineering techniques that contributed to the method's performance and how each of the techniques was carried out, not forgetting the role it played in ensuring the model performance. This chapter also explains how the results were presented and the findings in coming up with a successful model that could be used later on.

CHAPTER 5: SUMMARY AND CONCLUSION.

This chapter summarises the whole report by highlighting all the chapters. The importance and significance of the project is highlighted making reference to the objectives. The chapter also includes the recommendations which can be made by the audience or future work which the researcher seeks to pursue.

5.1. Summary of Chapters

The document is structured as follows:

Chapter 1 introduces the research by outlining its objectives, significance, and the challenges faced by sign language users. It sets the stage for exploring the development of an AI-driven translator.

Chapter 2 provides a comprehensive literature review, examining existing approaches to sign language translation, their limitations, and the gaps in current technologies. This chapter underscores the need for region-specific solutions, such as translating Zimbabwean Sign Language (ZSL) into Shona text and speech.

Chapter 3 details the research methodology, including dataset preparation, feature engineering, model selection, and evaluation. Emphasis is placed on the use of CNNs, supported by transfer learning, as the preferred model for this task.

Chapter 4 presents the results and discusses the findings, highlighting the superior performance of CNNs in gesture recognition and classification tasks compared to SVMs.

Chapter 5 concludes the study by summarizing the key insights, discussing the project's significance, and providing recommendations for future work, such as developing larger datasets and mobile applications.

Moving on to the next chapter (CHAPTER 2). In this Chapter we are looking at what other researchers said about the sign language translator. Mostly ,looking at the different methods used to develop the sign language translator. Stating the merits and demerits of each system. Adding on, in this Chapter we also look at the existing sign language translation technologies and the gaps identified while carrying out the research.

The next chapter (CHAPTER 3), we are looking at the methodology. We look at the research design, experimental environment, where we look at the software's and hardware's used during the development process. Adding on, the researcher also looks at the data collection methods used and the feature engineering processes involved so as to make sure they have a successful system. Model selection and development, the researcher looks at different models that can be used and chooses the one suitable for their specific project suiting their project needs. Lastly, we look at the experimental setup, where we look at how the dataset was divided into training and test sets.

We move on to the next chapter which is CHAPTER 4. This chapter highlights the results obtained from the developed system which is a result of the research.

5.2. Significance of Project

The research is aimed at enhancing communication between sign language users and non-signers. It should bridge the communication gap in other words. The researcher discovered that sign language users are being isolated because of their lack of other communication mediums. Thus, this research is aimed at finding a common ground in sign-language users and non-signers.

5.3 Challenges Faced

In coming up with a system that meets the set objectives, there are a lot of challenges that are faced by developers. In coming up with the Sign Language Translating app the researcher faced a lot of challenges. Zimbabwe Sign Language dataset is not readily available, because of this the researcher first wanted to create a dataset and creating a dataset is a problem on its own. After coming up with a dataset, it was very difficult to train the model using the necessary model. The model would underperform mostly and the accuracy would be very low, because the dataset was small. Having a small amount of data seemed like a good thing, however it is not a good thing at all.

5.4. Recommendations and Future Work

Organisations do not operate in isolation. That means they work with different people. Since they work with different people they are advised to adopt the usage of the sign language translator. It is suitable as it will help eliminate the human translator. As long as the signer is at a position where they are visible to the camera. The company personnel can easily know what they are saying as the cameras will be connected to desktops. This will allow organisations to communicate with people who are deaf and hard-of-hearing without any problem.

Adding on, for future work. The developers could develop a larger dataset that can be trained easily using the Convolutional Neural Network (CNN). CNN is much better to use for Sign Language Recognition tools.

Moreover, researchers can further develop the system so that they have a software that is usable by mobile phone users, whether they are Android users or IOS users. This will enhance inclusivity. Instead of having a software that benefits organisations, the system can also be developed further to make sure everyone is catered for.

Furthermore, developers can further develop the system to offer Sign Language to speech mechanism. This will cater for blind people who have no possible way to communicate with sign language users. Adding on, developers could add other Zimbabwean native languages like Ndebele, Tonga and Kalanga among others to make sure that no one is left behind. Also, the system should be developed to suit different Shona dialects as there are many dialects in Zimbabwe.

5.5 Conclusions

It has been a long and difficult journey developing the Sign Language Translation tool. Despite all the challenges, the researcher managed to complete the development. Although the system is not that perfect, the researcher did it.

References

1. Abdullah, D. M., & Abdulazeez, A. M. (2021). Machine learning applications based on SVM classification a review. Qubahan Academic Journal, 1(2), 81-90.
2. Al-Sofi, B. B. M. A., & Abouabdulqader, H. (2020). Bridging the gap between translation and culture: towards a cultural dimension of translation. International Journal of Linguistics, Literature and Culture, 6(1), 1-13.
3. Al-Qurishi, M., Khalid, T., & Souissi, R. (2021). Deep learning for sign language recognition: Current techniques, benchmarks, and open issues. IEEE Access, 9, 126917-126951.
4. Almomani, A., Nahar, K., Alauthman, M., Al-Betar, M. A., Yaseen, Q., & Gupta, B. B. (2024). Image cyberbullying detection and recognition using transfer deep machine learning. International Journal of Cognitive Computing in Engineering, 5, 14-26.
5. Amin, M. S., Rizvi, S. T. H., & Hossain, M. M. (2022). A comparative review on applications of different sensors for sign language recognition. Journal of Imaging, 8(4), 98.
6. Ananthanarayana, T., Srivastava, P., Chintha, A., Santha, A., Landy, B., Panaro, J., ... & Nwogu, I. (2021). Deep learning methods for sign language translation. ACM Transactions on Accessible Computing (TACCESS), 14(4), 1-30.
7. Anthoniraj, S., Ganashree, V., Umdor, B. J. R., Sai, G. D., & Navya, B. R. (2021, October). Sign Language Interpreter Using Machine Learning. In 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA) (pp. 1-6). IEEE.
8. Arena, F., Collotta, M., Pau, G., & Termine, F. (2022). An overview of augmented reality. Computers, 11(2), 28.
9. Bayoudh, K., Knani, R., Hamdaoui, F., & Mtibaa, A. (2022). A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets. The Visual Computer, 38(8), 2939-2970.
10. Bhatti, I., Rafi, H., & Rasool, S. (2024). Use of ICT Technologies for the Assistance of Disabled Migrants in the USA. Revista Espanola de Documentacion Cientifica, 18(01), 66-99.
11. Caeiro-Rodríguez, M., Otero-González, I., Mikic-Fonte, F. A., & Llamas-Nistal, M. (2021). A systematic review of commercial smart gloves: Current status and applications. Sensors, 21(8), 2667.

12. Calladine, R. (2020). Social attitudes towards deafness and the uptake of sign language among hearing people: A review of the literature. Disability & Society, 35(8), 1183-1202.
13. Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10023-10033)
14. Chegovo, R. W. (2023). Factors that influence implementation of sign language regulatory frameworks in Zimbabwean special schools for the deaf (Doctoral dissertation, Great Zimbabwe University).
15. Chegovo, R. W. (2023). Factors that influence implementation of sign language regulatory frameworks in Zimbabwean special schools for the deaf (Doctoral dissertation, Great Zimbabwe University).
16. Chegovo, R., Musengi, M., & Runo, M. (2024). Sign language policy implementation in special schools for the deaf in Zimbabwe. Deafness & Education International, 26(2), 127-148.
17. Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., & Miao, Y. (2021). Review of image classification algorithms based on convolutional neural networks. Remote Sensing, 13(22), 4712.
18. Costa-jussà, M. R., Cross, J., Çelebi, O., Elbayad, M., Heafield, K., Heffernan, K., ... & NLLB Team. (2022). No language left behind: Scaling human-centred machine translation. arXiv preprint arXiv:2207.04672.
19. Coppola, M. (2020). Gestures, homesign, sign language: Cultural and social factors driving lexical conventionalization. Emerging sign languages of the Americas, 349-90.
20. Cui, Z., Dai, J., Sun, J., Li, D., Wang, L., & Wang, K. (2022). Hybrid methods using neural network and Kalman filter for the state of charge estimation of lithium-ion battery. Mathematical Problems in Engineering, 2022(1), 9616124.
21. David, D., Alamoodi, A. H., Albahri, O. S., Zaidan, B. B., Zaidan, A. A., Garfan, S., ... & Malik, R. Q. (2023). Landscape of sign language research based on smartphone apps: coherent literature analysis, motivations, open challenges, recommendations and future directions for app assessment. Universal Access in the Information Society, 1-16.

22. De Coster, M., Shterionov, D., Van Herreweghe, M., & Dambre, J. (2023). Machine translation from signed to spoken languages: State of the art and challenges. Universal Access in the Information Society, 1-27.
23. De Meulder, M., & Haualand, H. (2021). Sign language interpreting services: A quick fix for inclusion. Translation and Interpreting Studies, 16(1), 19-40.
24. Drouin, A. (2020). Input and language acquisition: A comparison of native and non-native signers (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
25. Elboushaki, A., Hannane, R., Afdel, K., & Koutti, L. (2020). MultiD-CNN: A multi-dimensional feature learning approach based on deep convolutional networks for gesture recognition in RGB-D image sequences. *Expert Systems with Applications*, *139*, 112829.
26. Farooq, U., Rahim, M. S. M., Sabir, N., Hussain, A., & Abid, A. (2021). Advances in machine translation for sign language: approaches, limitations, and challenges. Neural Computing and Applications, 33(21), 14357-14399.
27. Flood, V. J. (2021). The secret multimodal life of IREs: Looking more closely at representational gestures in a familiar questioning sequence. Linguistics and Education, 63, 100913.
28. Galván-Ruiz, J., Travieso-González, C. M., Tejera-Fettmilch, A., Pinan-Roescher, A., Esteban-Hernández, L., & Domínguez-Quintana, L. (2020). Perspective and evolution of gesture recognition for sign language: A review. Sensors, 20(12), 3571.
29. Galloway, M. C., & Richards, B. (2020). Sociolinguistic diversity in the twenty-first century. TESOL Quarterly, 54(2), 445-464.
30. Goggin, G., & Zhuang, K. V. (2022). Disability as smart equality: Inclusive technology in a digitally advanced nation. In Vulnerable People and Digital Inclusion: Theoretical and Applied Perspectives (pp. 257-275). Cham: Springer International Publishing.
31. Gueuwou, S., Takyi, K., Müller, M., Nyarko, M. S., Adade, R., & Gyening, R. M. O. M. (2023). Afrisign: Machine translation for african sign languages.
32. HODGE, G. (2024). Understanding the access and inclusion experiences of deaf, deafblind and hard of hearing people who use Auslan: a review of the Australian research literature.
33. Ionescu, D. (2022). The Role of Figurative Language in Human Translation. Translation Studies: Retrospective and Prospective Views, 25(25), 65-84.

34. Hu, J., Ruder, S., Siddhant, A., Neubig, G., Firat, O., & Johnson, M. (2020, November). Xtreme: A massively multilingual multi-task benchmark for evaluating cross-lingual generalisation. In International Conference on Machine Learning (pp. 4411-4421). PMLR.
35. Isaković, L. S., Kovačević, T. R., & Srzić, M. S. (2020). Sign Languages–Then and Now. Зборник радова Филозофског факултета у Приштини, 50(4), 293-314.
36. Kardiansyah, M. Y., & Salam, A. (2021, March). Reassuring feasibility of using Bourdieusian sociocultural paradigm for literary translation study. In Ninth International Conference on Language and Arts (ICLA 2020) (pp. 135-139). Atlantis Press.
37. Karthikeyan, M. P., Lam, V. C., & Le, D. N. (2024). Efficient Sign Language Recognition and Classification Using African Buffalo Optimization Using Support Vector Machine System. *International Journal of Computer Science & Network Security*, *24*(6), 8-16.
38. Krzywdzinski, M., Pfeiffer, S., Evers, M., & Gerber, C. (2022). Measuring work and workers: wearables and digital assistance systems in manufacturing and logistics.
39. Lee, C. K., Ng, K. K., Chen, C. H., Lau, H. C., Chung, S. Y., & Tsoi, T. (2021). American sign language recognition and training method with recurrent neural network. Expert Systems with Applications, 167, 114403.
40. Lee, S., Glasser, A., Dingman, B., Xia, Z., Metaxas, D., Neidle, C., & Huenerfauth, M. (2021, October). American sign language video anonymization to support online participation of deaf and hard of hearing users. In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (pp. 1-13).
41. Liddell, S. K. (2021). American sign language syntax (Vol. 52). Walter de Gruyter GmbH & Co KG.
42. Light, J., McNaughton, D., Beukelman, D., Fager, S. K., Fried-Oken, M., Jakobs, T., & Jakobs, E. (2019). Challenges and opportunities in augmentative and alternative communication: Research and technology development to enhance communication and participation for individuals with complex communication needs. Augmentative and Alternative Communication, 35(1), 1-12.
43. Lillo-Martin, D., & Henner, J. (2021). Acquisition of sign languages. Annual Review of Linguistics, 7, 395-419.

44. Lohachab, A., Lohachab, A., & Jangra, A. (2020). A comprehensive survey of prominent cryptographic aspects for securing communication in post-quantum IoT networks. Internet of Things, 9, 100174.
45. Malla, Varsha & Vora, Avadhesh & D'mello, Caleb & Kardel, Shubham & Gotmare, Prof. (2024). Bridging the Communication Gap - Sign Language Communication System. International Journal of Advanced Research in Science, Communication and Technology. 356-364. 10.48175/IJARSCT-18060.
46. Matende, T. (2021). Dr Paul Svongoro. Communitas, 26, 86-102.
47. Mehta, A., Rathod, T., & Solanki, M. K. (2020, August). Evolution in Automated Translator for Real Time Voice to Sign Language Transformation for the Deaf and Dumb People. In National Conference on Recent Trends in Electronics and Communication Engineering-2020 (p. 69).
48. Mitchell, R., & Snyder, R. (2021). Signed languages and communication access for deaf and hard-of-hearing people. Annual Review of Linguistics, 12(1), 437-460.
49. Miyazaki, T., Morita, Y., & Sano, M. (2020, May). Machine translation from spoken language to sign language using pre-trained language model as encoder. In Proceedings of the LREC2020 9th workshop on the representation and processing of sign languages: sign language resources in the service of the language community, technological challenges and application perspectives (pp. 139-144).
50. Mohammadi, M., Rashid, T. A., Karim, S. H. T., Aldalwie, A. H. M., Tho, Q. T., Bidaki, M., ... & Hosseinzadeh, M. (2021). A comprehensive survey and taxonomy of the SVM-based intrusion detection systems. Journal of Network and Computer Applications, 178, 102983.
51. Moon NW, Baker PM, Goughnour K. Designing wearable technologies for users with disabilities: Accessibility, usability, and connectivity factors. J Rehabil Assist Technol Eng. 2019 Aug 13; 6:2055668319862137. doi: 10.1177/2055668319862137. PMID: 35186318; PMCID: PMC8855471.
52. Muchada, Morgen & Magadza, Tirivangani. (2023). Advances in Sign Language Dataset and Sign Language Recognition System. 10.13140/RG.2.2.16774.55362.
53. Musengi, M. (2019). The place of Sign Language in the inclusive education of Deaf learners in Zimbabwe amid CRPD (mis) interpretation. African Disability Rights Yearbook, 7.

54. Musengi, Martin & Muwaniki, Chenjerai & Musengi, Esther. (2024). Mobile Media and the Sign Language of Zimbabwe's Deaf Community. 10.1007/978-3-031-40885-4_2.
55. Mustafa, A., & Mantiuk, R. K. (2020). Transformation consistency regularization–a semi-supervised paradigm for image-to-image translation. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII 16 (pp. 599-615). Springer International Publishing.
56. Nishanthi, R. (2020). Understanding of the importance of mother tongue learning. International Journal of Trend in Scientific Research and Development, 5(1), 77-80.
57. Núñez-Marcos, A., Perez-de-Viñaspre, O., & Labaka, G. (2023). A survey on Sign Language machine translation. Expert Systems with Applications, 213, 118993.
58. Öman, B. (2021). Implementation of the 2030 Agenda in Sweden: Interpretation and application of disability inclusive goals at the local level.
59. Papastratis, I., Chatzikonstantinou, C., Konstantinidis, D., Dimitropoulos, K., & Daras, P. (2021). Artificial intelligence technologies for sign language. Sensors, 21(17), 5843.
60. Papatsimouli, M., Sarigiannidis, P., & Fragulis, G. F. (2023). A Survey of Advancements in Real-Time Sign Language Translators: Integration with IoT Technology. Technologies, 11(4), 83.
61. Pontecorvo, E., Higgins, M., Mora, J., Lieberman, A. M., Pyers, J., & Caselli, N. K. (2023). Learning a sign language does not hinder acquisition of a spoken language. Journal of Speech, Language, and Hearing Research, 66(4), 1291-1308.
62. Popel, M., Tomkova, M., Tomek, J., Kaiser, Ł., Uszkoreit, J., Bojar, O., & Žabokrtský, Z. (2020). Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals. Nature communications, 11(1), 1-15.
63. Prieto Ramos, F. (2020). Facing translation errors at international organisations: What corrigenda reveal about correction processes and their implications for translation quality. Comparative Legilinguistics, 41(1), 97-133.
64. Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems.
65. Saeed Al-Maroof, R., Abdelrahim Salloum, S., Qasim AlHamadand, A., & Shaalan, K. (2020). Understanding an extension technology acceptance model of google translation: A multi-cultural study in United Arab Emirates.

66. Sai Sarath Kumar, K. (2020). BREAKING THE COMMUNICATION BARRIER: The impact of Sign Language Translator Application as a communication aid on Hearing Impaired children during Dental treatment (Doctoral dissertation, Ragas Dental College and Hospital, Chennai).
67. Schleiermacher, F., & Bernofsky, S. (2021). On the different methods of translating. In The translation studies reader (pp. 51-71). Routledge.
68. Schlenker, P., Bonnet, M., Lamberton, J., Lamberton, J., Chemla, E., Santoro, M., & Geraci, C. (2024). Iconic Syntax: sign language classifier predicates and gesture sequences. Linguistics and Philosophy, 47(1), 77-147.
69. Schönström, K. (2021). Sign languages and second language acquisition research: An introduction. Journal of the European Second Language Association, 5(1), 30-43.
70. Senghas, R., & Monaghan, P. (2020). Sign languages as natural languages. In M. Marschark & P. Spencer (Eds.), The Oxford handbook of deafness and deaf studies (Vol. 1, pp. 17-42). Oxford University Press.
71. Sun, Z., Zhu, M., & Lee, C. (2021). Progress in the triboelectric human–machine interfaces (HMIs)-moving from smart gloves to AI/haptic enabled HMI in the 5G/IoT era. Nanoenergy Advances, 1(1).
72. Tao, T., Zhao, Y., Liu, T., & Zhu, J. (2024). Sign Language Recognition: A Comprehensive Review of Traditional and Deep Learning Approaches, Datasets, and Challenges. IEEE Access.
73. Tekwa, K. (2020). Real-Time Machine-Translated Instant Messaging: A Brief Overview with Implications for Translator Training. Translation Education: A Tribute to the Establishment of World Interpreter and Translator Training Association (WITTA), 135-153.
74. Van Dyk, T., Murre, P., & Kotzé, H. (2021). Does one size fit all? Some considerations for test translation. Assessing academic literacy in a multilingual society: Transition and transformation, 52-74.
75. Vanneschi, L., & Silva, S. (2023). Support vector machines. In Lectures on Intelligent Systems (pp. 271-281). Cham: Springer International Publishing.
76. Viriri, E., & Ndimande-Hlongwa, N. (2023). African languages as media of assessment in the teaching of indigenous languages in higher education: A paradigm shift. Journal for Language Teaching, 57(2).
77. ZainEldin, H., Gamel, S. A., Talaat, F. M., Aljohani, M., Baghdadi, N. A., Malki, A., ... & Elhosseini, M. A. (2024). Silent no more: a comprehensive review of

artificial intelligence, deep learning, and machine learning in facilitating deaf and mute communication. *Artificial Intelligence Review*, *57*(7), 188.

78. Zhang, W., Wang, J., & Lan, F. (2020). Dynamic hand gesture recognition based on short-term sampling neural networks. IEEE/CAA Journal of Automatica Sinica, 8(1), 110-120.
79. Zhou, H., Zhou, W., Qi, W., Pu, J., & Li, H. (2021). Improving sign language translation with monolingual data by sign back-translation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1316-1325).
80. Zinahwa, M. (2020). Socio-Cultural Factors Leading to the Marginalisation of Children Living with Disabilities in Accessing Education in Zimbabwe: The Case Study of Mutare District (Doctoral dissertation, The Open University of Tanzania).