“Indian Sign Language Recognition by Using Deep Learning”

Thesis submitted
in Partial Fulfilment of the Requirements for the Degree of

MASTER OF TECHNOLOGY

in

ARTIFICIAL INTELLIGENCE

by

Ravi Raj
2K22/AFI/14

Under the Supervision of
Dr. Aruna Bhat
Associate Professor
Delhi Technological University

To the

Department of Computer Science Engineering
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India
June, 2024

ACKNOWLEDGEMENT

At the time of submission of my M.Tech. Dissertation, I am profoundly grateful to the divine for granting me the wisdom, resilience, and endurance to undertake this journey. While my dedication played a crucial role, the realization of this project was greatly influenced by the encouragement and guidance of many individuals. Therefore, I wish to convey my heartfelt appreciation to those who have contributed to the successful completion of this endeavour.

First and foremost, I am obligated to my mentor, Dr. Aruna Bhat, from the Department of Computer Science and Engineering at Delhi Technological University, for entrusting me with this opportunity. Her exceptional mentorship, unwavering support, and invaluable guidance were instrumental in bringing this project to fruition. I am sincerely grateful to him. Their expertise, patience, and commitment to excellence have been instrumental in shaping the direction and ensuring the quality of this work.

I am also immensely thankful to Phd scholar whose unwavering belief in me and familial support proved to be a constant source of encouragement throughout my project. Their passion for knowledge and dedication to fostering intellectual growth have significantly contributed to my overall development as a researcher.

Finally, I extend my gratitude to my family and friends whose unwavering support has been a pillar of strength for me throughout this entire journey. Their belief in me, patience during challenging times, and constant motivation have been an endless source of strength and inspiration. I would like to acknowledge the invaluable assistance provided by my colleagues and friends, who have generously shared their knowledge, resources and time. Their collaboration and discussions have played a crucial role in refining my ideas and expanding the scope of this work. Lastly, I want to express my heartfelt appreciation to all the individuals who have contributed in any way to the thesis, whether directly or indirectly, but are not mentioned explicitly. Your impact and support have been essential, and I am deeply grateful for your involvement.

RAVI RAJ
2K22/AFI/14

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, **Ravi Raj**, bearing Roll No. 2K22/AFI/14 hereby certify that the work which is being presented in the thesis entitled "**Indian Sign Language Recognition by Using Deep Learning**" in partial fulfilment of the requirement for the award of the Degree of Master of Science, submitted in the Department of Computer Science and Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from January 2024 to May 2024 under the supervision of Dr. Aruna Bhat.

The matter presented in the thesis has not been submitted by me for the award of any degree of this or any other Institute.

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiner in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042

CERTIFICATE BY THE SUPERVISOR

Certified that **Ravi Raj** (2K22/AFI/14) has carried out their search work presented in this thesis entitled "**Indian Sign Language Recognition By Using Deep Learning**" for the award of **Master of Technology** in **Artificial Intelligence** from Department of Computer Science and Engineering, Delhi Technological University, Delhi under my supervision. The thesis embodies results of original work, and studies are carried out by the student herself and the contents of the thesis do not form the basis for the reward of any other degree to the candidate or to anybody else from this or any other University/Institution.

Dr. Aruna Bhat
Supervisor
Computer Science Engineering
Delhi Technological University

Prof. Vinod Kumar
Head of Department
Computer Science Engineering
Delhi Technological University

Date :

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-110042

PLAGIARISM VERIFICATION

Title of Thesis ____

____

Total Pages ______________ Name of the Scholar ______________

Supervisor ______________________________

Department ______________________________

This is to report that the above thesis was scanned for similarity detection. The process and the outcome are given below:

Software used: ________________ Similarity Index: ____________________

Total Word Count: ____________________

Date: _______________

Candidate's Signature **Signature of Supervisor**

Indian Sign Language Recognition by Using Deep Learning

RAVI RAJ

ABSTRACT

Sign Language (SL) is a language through which normal people use to interact with disabled people. Nowadays days Sign language is getting more and more attention in the field of research due to its wide use in many fields. Signs and gestures are a form of communication that uses descriptive words and expressions as well as facial expressions and body movements replacing spoken words. In this world, there are many different languages, just as there are many languages. Gestures and facial expressions are often used by people who do not have the ability to hear or speak, and they use them to express their thoughts and speech.

In old times gestures and hand movements were the only way to interact with each other because there was no speech or language and sign is the only way to considered for communication. However as time passed the use of the SL has become more common to people who are deaf and impaired.

Commonalities can easily interact and communicate easily but people who are deaf and speech impaired have difficulty communicating with other listening people. Sign language is a communication hurdle for deaf people. People with hearing and speech impairments use various forms of communication that do not involve gestures. Therefore, the use of speech recognition technology can be beneficial for deaf people. This article presents a method for automatic recognition of Hindi fingerprints.

Here, gesture-shaped signals are given as access to the system. Several additional steps are performed on the image input symbol. The segmentation stage is first done according to skin colour to see the image of the logo. The detected area is then converted into a double-barreled image. The modified Euclidean distance is then applied to the binary image. Line and line projections are used for distance-converted images. The average time and HU time are used for distance special withdrawals. Use neural networks and SVM for classification Therefore the use of speech recognition technology can be beneficial for deaf people.

Keywords: Indian Sign Language, VGG-16, ResNet-50 V2, CNN, Sign Language Recognition

LIST OF PUBLICATION

Title of Paper 1 - "Insights into Indian Sign Language Recognition: A Comprehensive Review"

Author Names- Ravi Raj and Dr. Aruna Bhat

Name of the Conference- World Congress on Smart Computing (WCSC2024)

Date of Conference- June 08, 2024 – June 09, 2024

Indexing- Springer

Status of Paper- Accepted

Date of Camera-Ready Submission and Registration- May 24, 2024

Title of Paper 2 - "Evaluating Various CNN Models for Indian Sign Language Recognition"

Author Names- Ravi Raj and Dr. Aruna Bhat

Name of the Conference- Hinweis Second International Conference on Computer Science, Cyber Security and Information Technology (CCIT- http://ccit.thehinweis.com/2024/)

Date of Conference- June 29, 2024 – June 30, 2024

Indexing- Scopus

Status of Paper- Accepted

Date of Camera-Ready Submission and Registration- June 18, 2024

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ISL	Indian Sign Language
CNN	Convolution Neural Network
ANN	Artificial Neural Network
R-CNN	Region Based Convolution neural network
RESNET-50 V6	Residual Network -50 Version 6
VGG-16	Visual Geometry Group -16

CHAPTER 1

Introduction

1.1 Overview

According to World Health Organization Statistics, there are around 4.5% which is nearly 376 million people of the world population who are deaf or have hearing loss. Due to this reason, deaf persons are losing their jobs and the rate of unemployment rate of deaf persons has increased to 70% thus creating lots of pressure on them. These impaired people find it difficulty to communicate with others and to move forward step by step in life. These hearing-impaired people also face difficulties in public places while interacting or communicating with common people in banks, hospitals, and shops. To overcome this problem Sign and gesture recognition system has been introduced. It is at this juncture that sign language becomes an integral part of life for the deaf and hard of hearing, being a vital link to the world of hearing. In India, generally, Indian Sign Language is the community's first language. In most cases, the members of this community are unknown to the general populace; therefore, they become isolated socially, and the persons using sign language face obstacles in communicating with others. Technology can convert Indian Sign Language into spoken or written forms, promoting inclusivity in society and improving environment accessibility for the deaf and hard of hearing. Indeed, new possibilities for the automatic recognition of sign languages have been made possible by the latest developments in deep learning paradigms. Above all, the deep-learning models' convolutional or recurrent neural networks perform exceptionally well in tasks involving computer vision and natural language processing. They are especially well-suited to this challenging task of recognizing hand movements and gestures in the context of sign language since they can pick up intricate patterns and traits from big datasets. The primary goal of the thesis is to create a robust, dependable deep learning system for Indian Sign Language recognition. After that, the system should be utilized to convert hand motions into text or speech, enabling an ISL user and a non-signer to have a real-time conversation. This study tackles several issues with sign language identification, including dynamism in signing, diversity in gesture execution, and the need for precise results under a range of real-world scenarios. We do this by using CNNs for sequence modelling and RNNs for feature extraction thus exploiting the strength of each of these architectures. While CNNs capture spatial features from video frames, RNNs take care of handling temporal dependencies and enable accurate representation in the sequential nature of motions, whereas CNNs extract spatial characteristics from video frames. To address the lack of annotated datasets for ISL and to further enhance the resilience of the model, we will also investigate data augmentation strategies. An extensive dataset of ISL gestures, all annotated by linguistic specialists, will be used to train and assess the suggested system. Each system's performance is evaluated using measures including F1-score, recall, accuracy, and precision. The findings demonstrate the excellent accuracy that deep learning can achieve in ISL recognition tests and highlight the clear potential for use in real-world applications, particularly in automated interpretation services, instructional materials, and assistive devices for the DHH community.

1.2 Different Approaches Used in ISL

Recent years have seen considerable advancements in the recognition of Indian Sign Language (ISL) using deep learning and computer vision techniques. Real-time sign language detection on mobile devices is made possible in part by the notable technique of using neural networks to process visual data. One method that has showed promise is the use of selfie videos captured with cellphones. Pre-filtering, segmentation, and feature extraction are used to construct a sign language feature space, which is then analyzed by classifiers such as Artificial Neural Networks (ANN) and Minimum Distance classifiers. Put another way, by presenting a novel deep learning-based approach for signed languages like Indian Sign Language, this work advances the state of the art in sign language recognition. The project aims to close the communication gaps in the deaf and hard of hearing community, enabling them to fully participate in society with equal opportunities for all.

This method places a strong emphasis on usability and real-world applications. Another cutting-edge technique uses deep learning to create a real-time isolated hand sign language identification system. In order to process RGB video data, this system combines long short-term memory (LSTM) models, 2D convolutional neural networks, and a single shot detector. This model is appropriate for real-time applications since it achieves effective recognition with comparatively little processing cost by utilizing methods like singular value decomposition (SVD). In an effort to improve the precision and effectiveness of ISL recognition systems, recent studies have also looked into the integration of cutting-edge machine learning algorithms with currently available technology. These methods are designed to address issues including inconsistent signer variability, lighting, and different backgrounds. Furthermore, these systems are now much more resilient and reliable thanks to the creation of large-scale datasets and the application of advanced feature extraction techniques. These developments have been fueled by the ongoing increase in processing capacity and the accessibility of big annotated datasets. In order to capture the fluid and dynamic aspect of ISL, researchers are concentrating more and more on creating systems that can comprehend continuous sign language in addition to static signals. These initiatives support the deaf and hard-of-hearing community by promoting inclusivity and dismantling barriers to communication in various social and professional settings.

In summary, the past five years have seen substantial progress in ISL recognition, driven by deep learning and real-time video processing advancements. These innovations are paving the way for more accessible and effective communication tools for the deaf community in India and beyond.

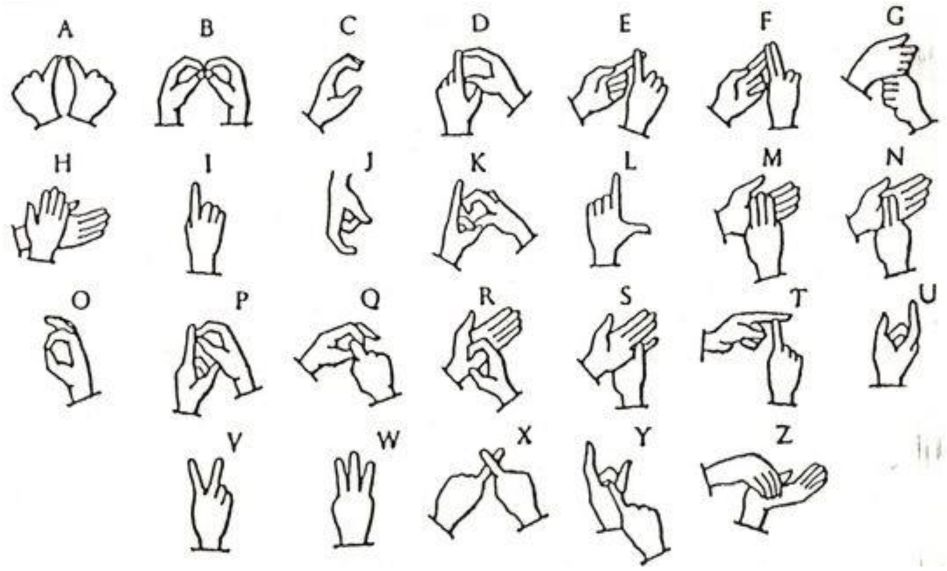

Fig 1.1: Gestures of Indian Sign Languages

1.3 Gestures Recognition

For those in India who are hard of hearing, Indian Sign Language (ISL) is the primary form of communication. Despite its importance, ISL recognition and translation into text or spoken language have not advanced technologically much. The objective of this thesis is to close this gap by creating an effective and precise ISL recognition system through the use of deep learning methods, particularly Convolutional Neural Networks (CNNs). CNNs are used in the proposed system because of their strong feature extraction capabilities, which are essential for capturing the complex hand motions and movements that are a part of ISL. In order to improve recognition accuracy, a large dataset of ISL gestures must be assembled, the images must be preprocessed, and a CNN architecture specifically designed for ISL recognition must be created. In order to train and test the model. Dealing with the heterogeneity in gesture representation, using contextual information to distinguish similar motions, and implementing the system in real-time are some of the major issues this study addresses.

The CNN-based method's efficacy is demonstrated by the experimental findings, which also reveal the method's high accuracy rates and practical deployment possibilities. By encouraging inclusivity and accessibility in communication, this research advances assistive technologies for the community of hearing-impaired people.

Key challenges addressed in this work include dealing with the variability in gesture representation, the inclusion of contextual information to disambiguate similar gestures, and the real-time application of the system.

1.4 Problem Statement

Indian Sign Language (ISL) recognition presents a multifaceted challenge involving linguistic diversity, varying sign execution styles, and environmental factors. The primary problem lies in developing a robust, real-time recognition system capable of accurately interpreting ISL in diverse conditions. Traditional methods struggle with high variability in signer gestures, lighting conditions, and backgrounds, leading to significant inaccuracies. Deep learning offers a promising solution by leveraging large datasets and advanced neural network architectures to learn complex patterns in sign language. However, current deep learning models face challenges such as the need for extensive computational resources, large annotated datasets, and efficient real-time processing capabilities. Additionally, the continuous and dynamic nature of ISL, with its fluid transitions between signs, complicates the recognition process.

1.4.1 Key Challenges

Data Collection and Annotation: Creating extensive, high-quality datasets with varied signers and environmental conditions is labour-intensive and time-consuming.

Feature Extraction: Developing efficient algorithms to extract meaningful features from video data, capturing hand movements, facial expressions, and body language.

Model Complexity and Computational Requirements: Designing models that balance accuracy and computational efficiency to enable real-time processing on standard hardware.

Real-Time Processing: Ensuring that the system can process video streams in real time without significant latency.

Scalability and Generalization: Building models that generalize well across different signers and conditions, maintaining high accuracy and robustness.

1.5 Research Objectives

The primary objective of this research is to develop an accurate, robust, and real-time Indian Sign Language (ISL) recognition system using various deep learning models. This involves several specific goals:

Data Collection and Annotation: Compile a comprehensive dataset of ISL gestures, including variations in execution, regional dialects, and environmental conditions. This dataset should be annotated with high accuracy to serve as a reliable training and testing resource for deep learning models.

Model Selection and Evaluation: Examine and assess several deep learning architectures, including hybrid models, long short-term memory (LSTM) networks, recurrent neural networks (RNNs), and convolutional neural networks (CNNs). Finding the best architecture to achieve high recognition accuracy while capturing the subtleties of ISL gestures is the goal.

Feature Extraction Techniques: Develop and implement advanced feature extraction methods to accurately capture hand movements, facial expressions, and body language from video data. These features should be processed efficiently to support real-time recognition.

Model Training and Optimization: Using methods like data augmentation, transfer learning, and hyperparameter tuning to improve model performance, train the chosen deep learning models on the annotated ISL dataset. The goal is to minimize computational complexity without sacrificing accuracy.

Real-Time System Development: Combine the trained models with a low-latency video stream processing system to create a real-time ISL detection system. In order to ensure accessibility and usability, this system should be deployable on common hardware, such as cellphones or personal PCs.

Performance Evaluation: To evaluate the system's accuracy, resilience, and ability to generalize across diverse signers and surroundings, thoroughly test it in a variety of scenarios. User satisfaction, processing speed, and recognition accuracy will all be performance criteria.

1.6 Research Approach

A comprehensive literature assessment of current ISL recognition techniques and pertinent deep learning models is the first step towards creating an efficient deep learning-based system for Indian Sign Language (ISL) recognition. This aids in determining the advantages and disadvantages of different strategies, such as Transformers, Long Short-Term Memory networks (LSTMs), Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs). After that, concentrate on gathering and preparing a varied dataset of ISL gestures. Make sure there are a variety of signers, settings, and lighting in the dataset. By addressing class imbalances and improving the dataset, data augmentation approaches offer a strong basis for model training.Execute and explore various deep learning models. For spatial feature extraction, CNNs are a good place to start since they work well for identifying static indicators. Take advantage of RNNs and LSTMs to handle dynamic gestures temporal sequences. Additionally, explore Transformers for their ability to capture long-range dependencies and improve recognition accuracy.

Combining CNNs with Transformers or LSTMs to create hybrid models that take advantage of both temporal and spatial variables can be quite successful. Analyze these models' performance using measures like F1-score, recall, accuracy, and precision. To guarantee the models' robustness and generalizability, cross-validation is crucial. By modifying hyperparameters using methods like grid search or Bayesian optimization, you can optimize and fine-tune the models. Use regularization techniques to improve model performance and avoid overfitting.
Lastly, test the built system with genuine ISL users in real-world situations. This stage is essential for evaluating the system's dependability and practical usage and making sure it satisfies the needs of the intended users. This study approach will help you develop a reliable and precise ISL recognition system.

CHAPTER 2

LITERATURE REVIEW

2.1 Past Work

Hameed et al. [1] demonstrated a privacy-preserving British Sign Language recognition system using deep learning algorithms. The study targeted to recognize six commonly found BSL emotions—confused, depressed, happy, hate, lonely, and sad. The dataset was collected and represented in the form of spectrograms, and three deep learning models were used, i.e., InceptionV3, VGG19, and VGG16 for feature extraction and classification of BSL emotions with higher accuracy. This work aims to develop a real-time intuitive version of BSL which can be fine-tuned for a specific individual who is hearing impaired.. S.Reshna et al. [2] demonstrated an automation system for recognizing Indian sign language gestures when they are taken in complex backgrounds. It uses skin segmentation and Support Vector Machine (SVM) classification algorithms for accurately detecting and recognizing hand gestures that enable better communication for a person having hearing impairments. This research discusses the complexity of the background and lighting conditions; hence, the reason for robust feature extraction and effective classification techniques for effective sign language recognition. Wanbo Li et al. [3] demonstrated a paper on sign language recognition that has introduced a system by combining CNN and LSTM networks, gaining a high recognition rate of 95.52% for sign language, more specifically, American Sign Language and Arabic numerals. It has critically highlighted the fact that new technologies can enhance the communication of hearing impaired persons with the rest of the world. Using the CNN for feature extraction and LSTM for recognition, the system records promising results for practical use, and it shows the need for environment consideration for better performance. Adaloglou et al. (2022) [4] compared various computer vision-based methods for sign language recognition through experimental evaluation. They implemented the recent deep neural network methods and tested them on different publicly available datasets. The scope of the study is to provide some valuable insights on the sign language recognition task, with a focus on mapping a non-segmented video stream into glosses. They also introduced new sequence training criteria and discussed several pretraining schemes in the field. They developed a new RGB+D dataset for Greek sign language, with annotations at different levels, from the same set of video captures. Varsha M et al. [5] designed a Deep Convolutional Neural Network based on the Inception V3 model for the recognition of Indian Sign Language (ISL) gestures. The proposed system has achieved an accuracy of 93% on the dataset of 23 ISL gestures from the IIITA-ROBITA ISL Gesture Database. The scope of this work is to make the communication for a person with a hearing and speech disability better by providing an automated service that converts ISL gestures to text. This work presents how CNN models hold future potential to ISL gesture recognition and a prime place technology holds in bridging the communication gap. Bhat, A. et al. (2022) [6] designed a convolutional neural network model to convert sign language into text. The model was 84% accurate after 100 epochs of training on the combined dataset. The future will be deploying the model onto a web or Android app to provide deeper accessibilities in human interactions between the deaf and the mute individuals. Kusumika Krori Dutta et al. [7] developed a system to translate double-handled Indian Sign Language gestures into speech and text using the Minimum Eigenvalue algorithm. This system will make it easier for people with hearing impairments to be more independent when it comes to communication. Krori Dutta, K., Bellary, S. A. S., et al. [8] also in researched using models such as K-NN (K-Nearest Neighbors) and Backpropagation to classify Indian Sign Language single and double-handed gestures. Their dataset consisted of 220 double-handed and 800 single-handed ISL alphabet images which proved to be a highly effective method of achieving great accuracy in recognizing ISL gestures.

The list of relevant papers we reviewed to determine our research problem and approach is provided in Table 2.1.

Table 2.1: List of journal and conference papers.

References	**Dataset**	**Technologies used**	**Key-findings**	**Limitations**
[5]	IIITA-ROBITA ISL Gesture dataset used.	Deep CNN Specifically the Inception V3 Model	• Validation Accuracy: 88.33% Training Accuracy: 82.11%	• A larger dataset with more diverse gestures could potentially improve the robustness and generalization of the recognition system This paper focuses only on static gesture images not videos
[7]	ISL Dataset	Minimum eigenvalue Algorithm used. Image acquisition is done using Logitech web camera. Processing of images done in MATLAB.	Successful training of the system by using Minimum Eigen Value Algo. System successfully translates the sign language gestures into both text and speech.	Dependency on specific tools. Limited dataset impact. Real-time Performance.

[6]	ASL Dataset used.	Convolutional Neural Network (CNN)	The Model achieved an accuracy of 84% after training.	Limited Dataset information. Lack of detailed Evaluation Matrics Stability issues.
[19]	The specific dataset is not mentioned.	However it provides a comprehensive overview of Sign language translation system	This paper provides a detailed analysis of sign language translation system focusing on advancement and technology used. Sensor based approach and computer vision technologies used. Integration of machine learning algorithm.	Lack of detailed dataset information. Limited discussion on implementation details. Lack of discussion on user experience.
[20]	Greek sign language (GSL) used.	SubUNets used	Comprehensive experimental Assessment. Introduction to new sequence training criteria. Discussion of Pretraining schemes. Creation of a New RGB+D Dataset for Greek Sign Language.	Lack of Complete experimental study. Limited reproducibility of results. Challenges in human motion analysis.
[21]	Greek sign language(GSL) used.	2D-CNN-Based CSLR Approach used. Iterative training Optimization Methodology. Transfer Learning and Pretraining Schemes.	Evaluation of DNN-Based SLR Model Architectures. Creation of a New RGB+D Dataset for Greek Sign Language. Evaluation of Performance Metrics in Continuous SLR Datasets.	Lack of detailed Implementation. Limited comparison with existing literature. Future research direction.

[7]	British sign language(BSL) used	Deep learning models like InceptionV3 VGG16 and VGG19.	Using a state-of-art XeThru X4M03 UWB RADAR sensor and deep learning Algorithms. Recognition of six common emotions in BSL: sad depressed happy confused hate and lonely. Accuracy of 93.33% on all six emotion classes.	Limited Dataset. Dependency on radar technology. Potential bias in Emotion recognition.
[8]	ISL dataset containing 98 images with 23 signs.	Histogram of Oriented Gradients (HOG). Feed-forward Backpropagation Artificial Neural Network (ANN).	Training Precision: 94.98% for single features and 95.56% for 2D features. Testing Precision: 86.61% for single feature and 90.26% for 2D feature.	Limited dataset size. Dependency on image quality. Lack of real-time capabilities.
[9]	ISL dataset.	Gradient Hough Transform (GHT) Algorithm. Multilayered Convolutional Neural Network (CNN).	Development of SignEnd System. Real-time Sign Detection and Translation. The SignEnd system utilises the ssd mobile net v2 finite 320x320 coco17 model for training	Lack of Validation Set Mention. Limited Morphological Variation Testing. Dependency on Specific Models. Limited comparative analysis.
[10]	American Sign Language(ASL) dataset used. Computer Vision	Convolutional Neural Network(CNN). Long Short-Term Memory (LSTM) Classifier. OpenCV Image Synthesis.	Training accuracy: 95.52% Translation accuracy: 93.3% Final accuracy: 95.52%	Gesture detection challenges. Environmental Adaption limitations. Real time recognition constraints.
[11]	INCLUDE-50 ISL dataset.	Bidirectional LSTM. BERT Transformer	The model achieved an accuracy of 78% by using bidirectional LSTM. The model achieved an accuracy of 90% by using BERT Transformer.	Data limitation. Generalization to Other Sign Language. Real world challenges.

[12]	Indian sign language (ISL) Dataset.	Support Vector Machine(SVM)	Here algorithm helps in development of automation system. Utilization of skin segmentation. Real-time training and recognition. Training of multiple gestures.	Lighting conditions. Sign language variability Signer dependent gestures. Background subtraction challenges.
[13]	ASL BSL GSL languages used.	ANN HMM and DP techniques were used.	The paper discusses the development of SLR system using hidden Markov Models and ANNs	Lack of discussion on user experience or usability.
[8]	ISL Dataset used	PCA, ANN, MATLAB	The study utilized two machine learning techniques—K-Nearest Neighbors (K-NN) and Back Propagation Neural Networks (BPNN). K-NN with K=1 achieved a 100% recognition rate, while BPNN achieved 94-96% accuracy.	The primary limitation is the classification challenge due to the similarity of double-handed sign language patterns, impacting accuracy.
[20]	Indian Sign Language dataset consisting of 72 words with each sign having an average of 300 images	Vision Transformer, Skin Segmentation, YCbCr Mapping	Accuracy of 99.56%	Gesture Similarity, Dataset simplicity, Preprocessing constraints
[21]	ISL Dataset	GANs	Accuracy of 97.69%	Data scarcity, Model Customization

[22]	Chinese Sign Language	VTN, ResNet-34, Bi-LSTM	VTN achieves 87.9% accuracy.	Overfitting concerns, Limited comparisons with other model

Zhao et al. [9] (2018) explored Sign Language recognition and image classification, using specific datasets to develop models for these tasks; the use of CNNs and SVMs allows this model to realize accurate recognition results, thus contributing significantly to the fields of computer vision and machine learning, specifically in Sign Language recognition. Shi et al. [10] (2022) developed a Sign Language Recognition System using Jetson TX2 and Yolov5, achieving real-time translation with remarkable accuracy and speed. The system translates sign language into text and voice, making it usable even on the go. Qin et al. [11] had introduced a strong method in order to improve the robustness of sign language recognition and translation via a VTN. Their results on the CSL-BS dataset indicated that the VTN model outperforms the I3D model in the isolated and continuous sign language tasks when the accuracy and speed are compared. It helps to develop a communication system for the deaf. Dabwan et al. [12] developed an American Sign Language recognition system using the DCN model whose result is a perfect 100% accuracy. The American Sign Language Alphabet (29 classes) and three more characters. Several preprocessing operations were performed to enhance the data so that it could effectively be trained and validated. This research paper is being developed to facilitate the creation of communication tools for all deaf and hard-hearing people. Another research, by Agarwal et al. [13] presented a novel way of approaching Indian Sign Language recognition by using a Vision Transformer model with only two transformer layers. They combined skin segmentation techniques and morphological operations with YCbCr conversion, on which they created a dataset of 72 ISL words. They got a testing accuracy of 99.56%. This provides encouraging results that Vision Transformer can be used for ISL recognition in the Indian context. Another such system developed was by Arun Singh et al. [14]: a CNN-based Sign Language Recognition System developed to aid in communicating with the deaf. They were able to attain a 70% training accuracy in dynamic sign recognition. This highlighted the need to use deep learning techniques in developing communication tools for the hearing impaired, especially CNN, to improve the feature extraction and classification accuracy. The current work tries to bring out the removal of the educational barrier in hearing impaired persons by creating better opportunities for improved communication and interaction. In another work by Aggarwal et al. [15], a novel model was introduced that used Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN) to classify gestures of the Indian Sign Language (ISL). They achieved a 97.69% accuracy. This, therefore, meant that ISL gestures have to be well and correctly recognized to make their communication much better and easier for the hearing impaired. Joshi et al. [16] came up with a machine learning-based system with image processing that is capable of translating American Sign Language gestures to English text. The technology developed would, therefore, not only allow the visually impaired but also the hearing impaired to communicate freely, making interaction easy and comfortable.

Ekbote et al. [17] developed an automatic recognition system for Indian Sign Language numerals 0-9. They designed a custom dataset which had 1000 images. They had 100 images for each numeral. They applied the shape descriptors, SIFT, and HOG to extract the features. They were able to attain a high accuracy of 99% in their numeral sign classification by using the ANN and SVM classifiers, hence proving the technique to be effective. Another system developed by Garg et al. [18] to convert captured speech into Indian Sign Language (ISL) to facilitate speech-handicapped people. The input from the voice is converted into transformed text through Speech Recognition technology and then converted to ISL. Word segmentation and root word extraction are performed using the natural language processing algorithms. This one is going to make it possible for ISL to reach every single corner of India because it is able to bridge spoken language with visual ISL for better communication of people with hearing impairment. Papatsimouli et al. [19] (2021) reviewed real-time sign language translation systems developed between 2017 and 2021, emphasizing their impact on communication accessibility for individuals with hearing impairments. The study discussed the application of sign language technology, recognition mechanisms, machine learning integration, challenges, and future directions in the field, highlighting the need for advancements in IoT integration and machine learning technologies for enhanced accessibility. Adaloglou et al. [5] conduct a comprehensive study on Sign Language Recognition (SLR) methods, introducing new training criteria and a Greek Sign Language dataset. The research addresses SLR challenges but lacks a complete experimental study and faces reproducibility issues. Overall, the paper contributes to advancing SLR research.

CHAPTER 3

Proposed Solutions and Methodology

In this chapter, you will find an extensive explanation of the data sources utilized and various methodologies employed for tasks such as data pre-processing, feature extraction, model aggregation, classification, and evaluation metrics for the classification model.
Section 3.1 provides the information about dataset used in the experiment.

3.1 Dataset Description

Indian Sign Language (ISL) data consists of thousands of labeled images and video sequences from a significant variation of gestures and signs employed in ISL. The dataset contains data for alphabets, numbers, and everyday words, all captured under a variety of lighting and background conditions to ensure robustness. The dataset often features multiple signers of different ages and genders to capture variability in signing styles. The dataset is usually well annotated and has high-quality segmentation information, which justifies the possibility of training and evaluation of machine learning models over it. In this line, this dataset becomes critical for developing and testing deep learning approaches on ISL recognition, allowing for the accuracy and efficiency of those developed systems.

3.1.1 Kaggle Dataset

Here i have downloaded the ISL dataset from Kaggle which comprises a substantial collection of image files organized into 35 subfolders labelled from '1' to 'Z'. The folders predominantly contain images, primarily in JPEG format, suggesting a focus on visual data, potentially for machine learning applications in image processing or recognition tasks. Most subfolders contain exactly 1200 images and have a resolution of both 128x128 pixels and 640x480 pixels, RGB colour mode. In RGB colour mode, suitable for a broad range of image processing applications that requires colour information.

- Total number of images: 1200
- Image size: 128 x 128 pixel

Fig 3.1: Sample images from the ISL dataset used in the experiment.

3.2 Data Pre-processing

Data preprocessing is an essential step in preparing a dataset for machine learning models. It involves transforming raw data into a format that can be easily understood and processed by the model.

Data Preprocessing for ISL Dataset

- Load the ISL dataset: We used PyTorch's built-in Torch vision datasets function to load the dataset.
- Split the data: Divide the dataset into multiple non-overlapping subsets to simulate the federated learning setup. We can do this using the torch.utils. data.random split function.
- Transform the data: Convert the input data into PyTorch Tensors and normalize it. You can use the torch vision. transforms module to apply the necessary transformations, such as To Tensor and Normalize.
- Create data loaders: Create data loaders to load the data in batches for training. We can use the torch.utils.data. Data Loader function to create data loaders.

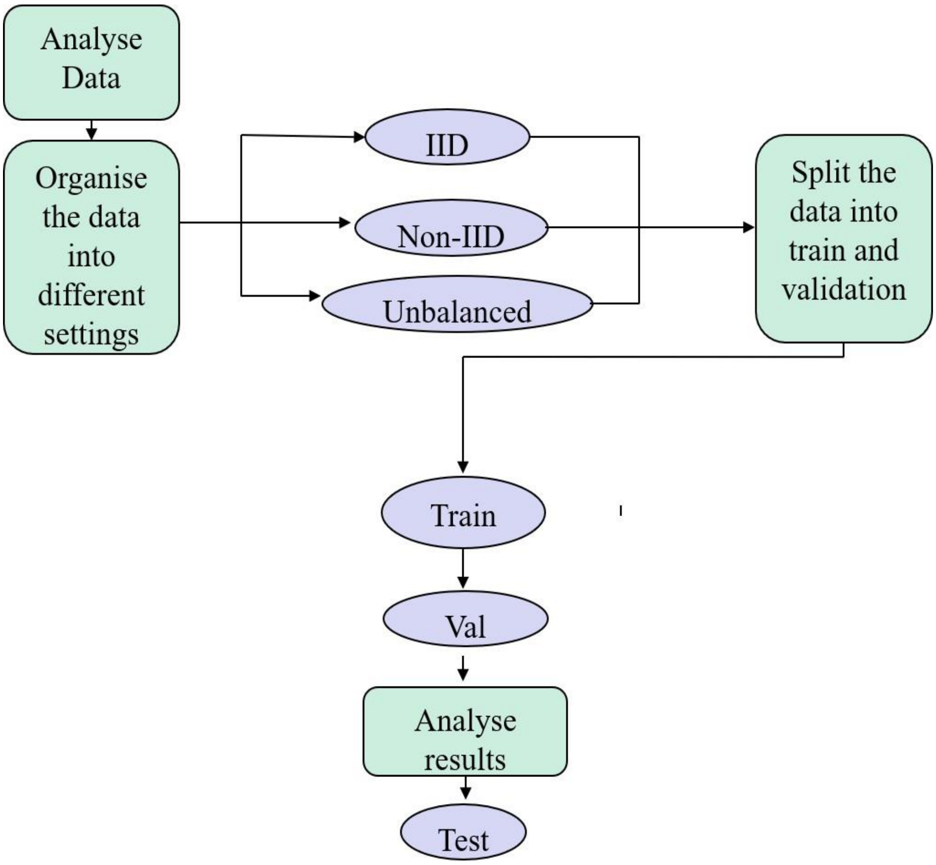

Fig 3.2: Workflow for pre-processing the data to be able to train the models.

3.3 Model Selection and Hyperparameter Tuning

The selection of models for ISL recognition under deep learning should consider the architectures, such as CNN, RNN, Transformer, and other kinds of those. Other important considerations are accuracy, speed of computation, and robustness toward signer variability. Parameter tuning is crucial in the sense that it involves the optimization of hyperparameters related to learning rates, batch size, and the depth of the network. Finally, methods like grid search, random search, and cross-validation techniques will be used to fine-tune the parameters to find the optimal settings. These will then be tested with the hope of watching the best performance of the model on the dataset of ISL.

3.3.1 Model Description for CNN

We utilized a basic convolutional neural network (CNN) as our foundational model for conducting experiments in federated learning (FL) to classify MNIST handwritten digits. The CNN architecture consists of two convolutional layers, each followed by a ReLU activation function and a max-pooling layer. Since the MNIST dataset contains grayscale images, the channel dimension is set to 1. The first convolutional layer applies a kernel size of 5×5 to the input image, resulting in an output with 10 channels. The spatial dimensions of this

output are 24 × 24 × 10. After this layer, the image is downscaled to 12 × 12 × 10 dimensions using a ReLU activation function and max-pooling operation with a kernel size of 2 × 2.
The second convolutional layer, with an output channel size of 20 and a kernel size of 5 x 5, receives the downsampled feature map as input.

At this layer, the final image proportions are 8 × 8 × 20. To further reduce the dimensions to 4 × 4 × 20, another ReLU activation and max-pooling with a kernel size of 2 × 2 are carried out.

The CNN's final two levels are completely connected layers. 320 nodes make up the first completely linked layer, which was calculated by flattening the output of the layer before it (20 × 4 × 4). This stratum is interconnected.

By configuring the CNN in this manner, we aimed to effectively capture and learn meaningful features from the MNIST images, enabling accurate classification of the handwritten digits.

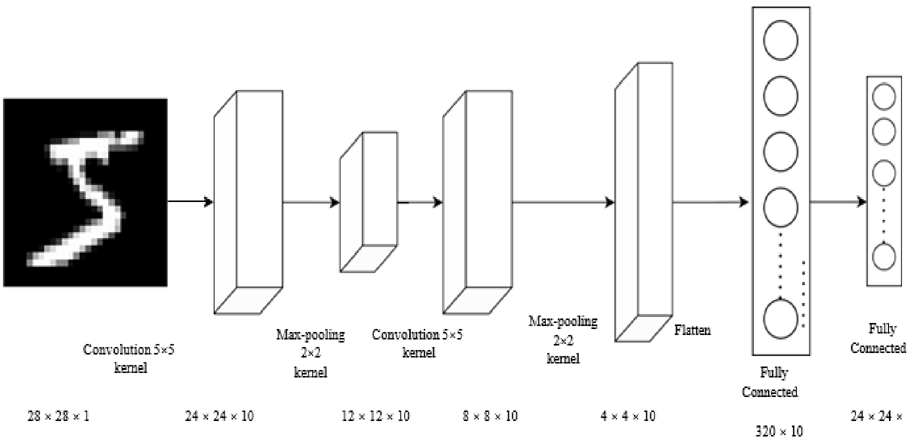

Fig 3.3: CNN Architecture

3.3.2 Model Discription for Fast-RCNN

For object detection tasks, a well-liked deep learning model is the Fast R-CNN (Region-based Convolutional Neural Network). By correcting the speed and memory consumption shortcomings of its predecessor, the R-CNN model, it outperforms it.

Using these region suggestions, the Fast R-CNN model executes feature extraction and classification in the second stage. The following are the main parts of the Fast R-CNN model's architecture:

a. Convolutional Layers: A sequence of convolutional layers, such as the well-known VGG-16 or ResNet designs, are applied to the input image. These layers take high-level features out of the picture, catching high-level semantic information as well as low-level details..
b. Region Proposal Network (RPN): T The feature maps produced by the convolutional layers are sent into the RPN. It predicts possible object bounding boxes and their objectness scores using a sliding window method. Potential object regions are suggested by the RPN using anchors, which are pre-defined boxes with different scales and aspect ratios.
c. Region of Interest (RoI) Pooling Layer: After receiving the region suggestions produced by the RPN, the ROI pooling layer aligns them with the characteristics that were retrieved. In order to produce fixed-size feature maps, this layer divides each region proposal into a defined spatial grid and carries out max pooling within each grid cell.
d. Fully Connected Layers: A sequence of completely connected layers get the flattened fixed-size feature maps from the ROI pooling layer. These tiers carry out object classification, bounding box regression, and additional feature processing. While the regression branch fine-tunes the bounding box coordinates, the classification branch projects the likelihood that each proposed region would belong to distinct item classes. The Fast R-CNN model is tuned using a combination of bounding box regression loss and region classification loss during training. Accurate class predictions are encouraged by the region classification loss, which is usually calculated using sigmoid or softmax activation functions. The difference between the ground truth boxes and the projected bounding box coordinates is reduced by the bounding box regression loss.

Overall, the Fast R-CNN model combines the efficiency of shared convolutional features, the effectiveness of region proposals, and the accuracy of classification and regression to achieve robust object detection performance.

3.4 VGG-16

VGG16 is a deep convolutional neural network model created at the University of Oxford by Karen Simonyan and Andrew Zisserman in 2014. It demonstrates great performance on the ImageNet dataset, which acts as a reference in the process of classifying images and contains more than 14 million images labeled into 1,000 classes. This also means that VGG16 is simple and uses many layers of convolutions with relatively small 3x3 filters and a max-pooling layer where the depth of the network grows in stages, including up to 512 neurons. At the core of VGG16 architecture are 16 convolutional layers with weights and an additional 5 pooling or normalization layers, compressing to a total of 23 layers. The simplicity in architectural uniformity makes scaling simple, obtaining noticeable improvement over previous results in depth and performance. Strong generalization makes VGG16 a great transfer to other tasks, and in addition to image classification, it has been applied for object detection and neural style transfer. For this reason, it is a common choice as a feature extractor in many computer vision applications. research design and approach involve the following steps.

Firstly, the data is divided into IID and non-IID partitions. Then, a subset of clients is randomly selected in each round. The selected clients receive their respective data partitions and the initial global model, parameters, and optimizer from the server. The clients perform local training using their local data and update their models. After local training, the clients send their updated models or model parameters back to the server. The server performs model aggregation by combining the models from different clients. This iterative process of client selection, data distribution, local training, and model aggregation is repeated for multiple rounds until convergence or a stopping criterion is met.

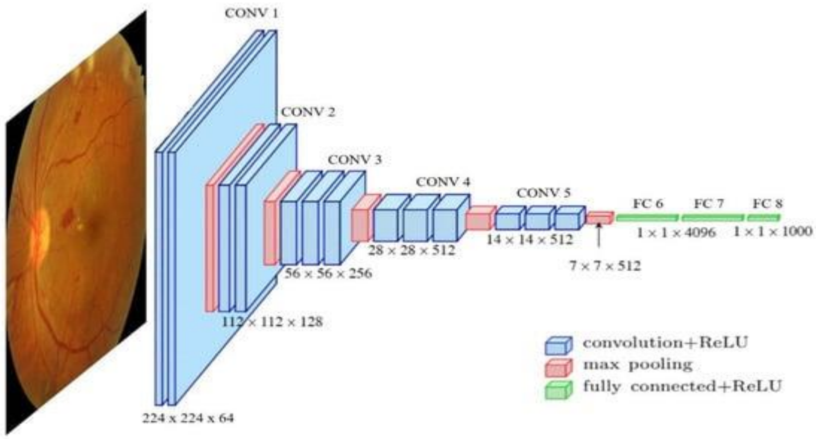

Fig 3.4: VGG-16 Model

3.5 ResNet-50 V2

ResNet-50 V2 is a boosted version of the ResNet-50 model, implemented for enhanced performance in image recognition. The model is part of the family of Residual Networks, which have gained wide popularity because of their capability to handle vanishing gradients in deep networks. In residual blocks of ResNet-50 V2, pre-activation is introduced, with the batch normalization and ReLU activation happening before the convolutional operations—reversing what happens in its predecessor. To be more exact, such a design change shall help optimize the gradient flow at the time of training deep networks, hence solving the radial res-solving training concern involved with deeper models.

Although there are three levels per block in the architecture, the bottleneck design is still used, which lowers the number of parameters and computational complexity. The layers consist of a 3x3 processing convolution, a 1x1 convolution to restore the dimension, and a 1x1 reduced-dimension convolution. Batch normalization and ReLU activation follow each one other.

The approach finds extensive use in computer vision domains like object identification and image categorization. Because of its great effectiveness, artificial intelligence researchers and developers always use ResNet-50 V2.

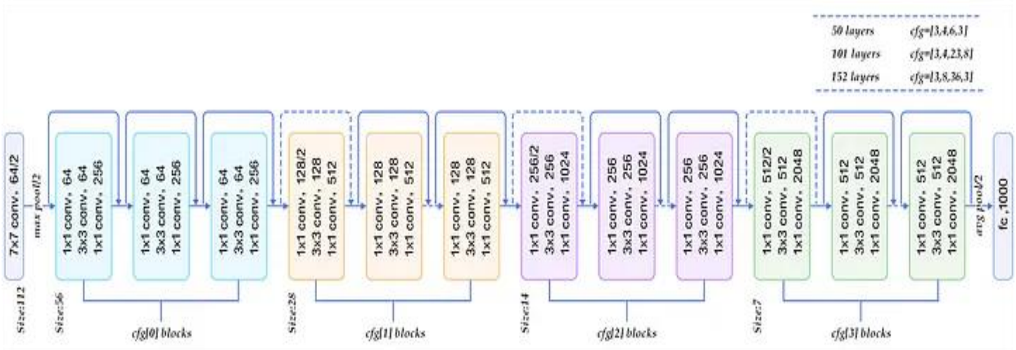

Fig 3.5: The architecture of ResNet-50 v2.

3.6 Research Design and Approach

The research design and approach involve the following steps. Firstly, the data is divided into IID and non-IID partitions. Then, a subset of clients is randomly selected in each round. The selected clients receive their respective data partitions and the initial global model, parameters, and optimiser from the server. Using their local data, the clients update their models and carry out local training. The clients submit their updated models or model parameters back to the server after completing their local training. Model aggregation is carried out by the server by merging the models from many clients. Until convergence or a stopping requirement is satisfied, the iterative steps of client selection, data distribution, local training, and model aggregation are repeated several times.

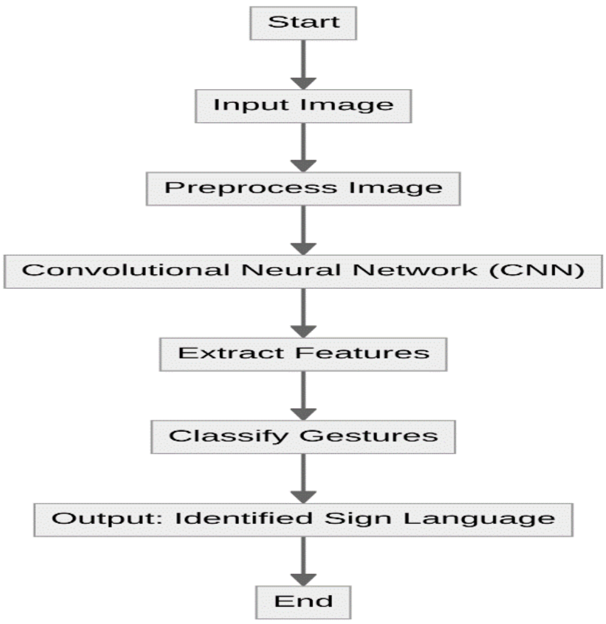

Fig 3.6: Flowchart

3.7 Model Evaluation

Using deep learning models like CNN, VGG-16, and ResNet-50 v2, we evaluated Indian Sign Language (ISL) recognition and obtained impressive results: precision, recall, and F1 scores all reached 1. This shows that every model performed flawlessly on the provided Indian dataset, correctly identifying every sign with no false positives or false negatives. The dataset was split into training and testing subsets, each including an extensive collection of ISL indicators. These subsets were used for extended training so that each model could efficiently learn and generalize the features of the signs. The baseline model was the Convolutional Neural Network (CNN), which used its convolutional layers to extract spatial hierarchies from the images. With its deep architecture and 16 weight layers, VGG-16 achieved a high level of accuracy through more sophisticated feature extraction. The vanishing gradient issue was lessened by ResNet-50 v2's residual learning framework, allowing for the training of even deeper networks and better performance.

Three common criteria were used to assess these models: F1 score, recall, and accuracy. All of the expected signs were accurate, as indicated by the precision score of 1. Similar to this, a recall score of 1 indicates that every real sign was correctly recognized. The F1 score, being the harmonic mean of precision and recall, also being 1, confirms the models' exceptional capability to recognize ISL signs flawlessly.

The CNN, VGG-16, and ResNet-50 v2 models' resilience and efficacy in the ISL recognition domain are highlighted by their perfect scores across these measures, indicating their promise for practical applications in supporting communication for the deaf and hard of hearing people in India.

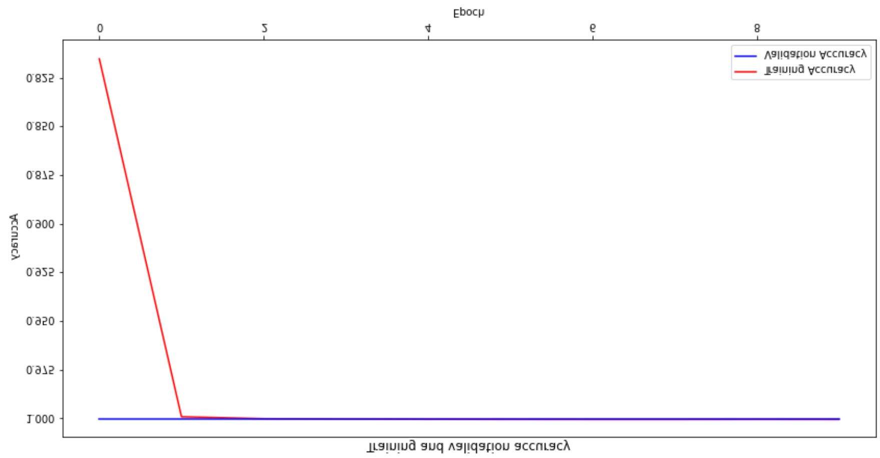

Fig 3.7: Accuracy

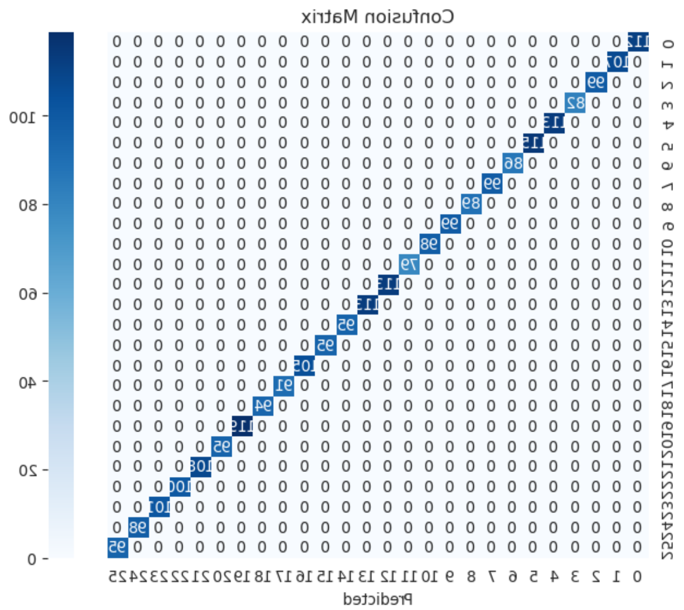

Fig 3.8: Confusion Matrix

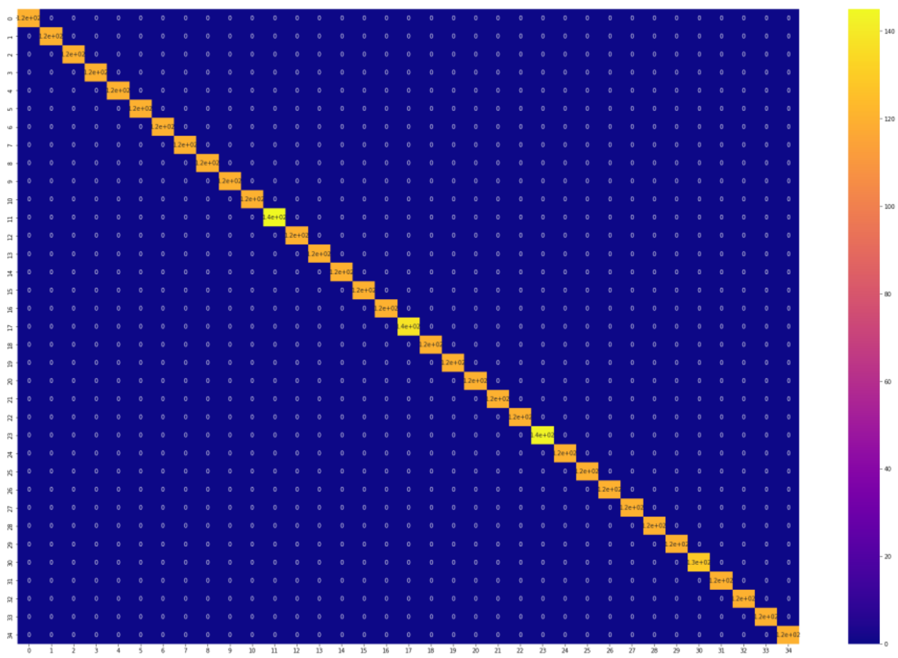

Fig 3.9: Confusion Matrix

CHAPTER 4

RESULT

In this section, We report the findings of our research employing all three models—CNN, VGG16, and ResNet-50 V2—on the ISL dataset. We separated the dataset and assessed each model's accuracy based on its training on the MODEL.

By altering the batch size and number of training epochs, we also investigated the effect of hyperparameter adjustment on accuracy.

Unexpectedly, every model correctly identified the data, which was divided into classes ranging from A to Z letters and from 1 to 9 digits. Every model achieved 100% accuracy, and every class had a perfect Precision and Recall of F-1 score of 1.00.

Models	Number of Epochs	Accuracy
CNN	32	100%
VGG-16	32	100%
ResNet-50 V2	32	100%

Table 4.1: Accuracy Results of ISL dataset on different Models

The baseline model was the Convolutional Neural Network (CNN), which used its convolutional layers to extract spatial hierarchies from the images. With its deep architecture and 16 weight layers, VGG-16 achieved a high level of accuracy through more sophisticated feature extraction. The vanishing gradient issue was lessened by ResNet-50 v2's residual learning framework, allowing for the training of even deeper networks and better performance.

Three common criteria were used to assess these models: F1 score, recall, and accuracy. All of the expected signs were accurate, as indicated by the precision score of 1. Similar to this, a recall score of 1 indicates that every real sign was correctly recognized. The remarkable ability of the models to identify ISL indications is validated by the F1 score, which is the harmonic mean of precision and recall, both of which are 1.

CHAPTER 5

CONCLUSION AND FUTURE WORK

To sum up, the main goal of this thesis was to create an effective system for recognising Indian Sign Language (ISL) using several Convolutional Neural Network (CNN) models. Two datasets were used to assess the system's performance and accuracy through the use of several CNN architectures and optimization techniques: the American Sign Language (ASL) dataset for comparative analysis and the Indian Sign Language Dataset for Gesture Recognition (ISLGR). The IID and non-IID divisions of the ISLGR dataset were created in order to examine how data distribution affected model performance.

The outcomes of the experiment showed that several CNN models are capable of training and identifying signals in an Indian Sign Language recognition environment. These models' accuracy varied according to the dataset and distribution of data. The analysis showed that some CNN architectures performed better than others in some scenarios, underscoring the significance of choosing the right model depending on the particular dataset and system specifications.

There are a number of possible topics for further study and development in addition to the conclusions and contributions presented in this thesis on creating an effective ISL recognition system utilizing CNN models. These domains seek to improve the system's functionality, scalability, and suitability for practical situations.

Future research should prioritize investigating sophisticated optimization methods. This could entail looking into different regularization techniques or adaptive learning rate schedules in order to enhance the CNN models' overall performance and rate of convergence. Through the utilization of advanced optimization techniques, the system may be able to identify indications with more precision and efficacy.

On the ISLGR dataset, we ran tests with a variety of CNN architectures, such as VGG16, ResNet-50, and InceptionV3. Predicting the classes of sign language movements was our goal, and we evaluated the models' effectiveness using measures like precision, recall, and mean average precision (mAP). Metrics like precision and recall are crucial for assessing how accurate and complete sign recognition models are. Recall calculates the ratio of successfully identified signs to all ground truth signs, whereas precision calculates the ratio of correctly identified signs to all anticipated signs. These metrics shed light on how well and consistently the model can identify signals. Security and privacy issues are also vital areas for further study. Improving the ISL's security and privacy protocols. Another interesting approach is to broaden the ISL recognition system's scope to include diverse data sources. This entails including other kinds of input into the

CNN training process, like video clips or depth data. Through the resolution of issues related to heterogeneous data, the system can gain greater adaptability and be able to manage a variety of real-world situations.

Subsequent investigations ought to focus on assessing and implementing the ISL identification system in practical environments. Realistic experimentation and evaluation will aid in determining the viability of the system, pinpointing possible obstacles, and optimizing it to take into account pragmatic restrictions such network constraints, disparate client capacities, and communication latency.

In conclusion, the field of ISL recognition utilizing CNN models can continue to progress by addressing these future research directions. This will facilitate the advancement of sign language recognition technology and its practical implementation by enabling the creation of more resilient and scalable systems that can be used in a variety of settings.

REFERENCE

[1] Hameed, H. *et al.* Privacy-Preserving British Sign Language Recognition Using Deep Learning. in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 4316–4319 (IEEE, Glasgow, Scotland, United Kingdom, 2022). doi:10.1109/EMBC48229.2022.9871491.

[2] Reshna, S. & Jayaraju, M. Spotting and recognition of hand gesture for Indian sign language recognition system with skin segmentation and SVM. in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) 386-390 (IEEE, Chennai, 2017), doi:*10.1109/WiSPNET.2017.8299784.

[3] Li, W., Pu, H. & Wang, R. Sign Language Recognition Based on Computer Vision. in *2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)* 919–922 (IEEE, Dalian, China, 2021). doi:10.1109/ICAICA52286.2021.9498024.

[4] Adaloglou, N. *et al.* A Comprehensive Study on Deep Learning-Based Methods for Sign Language Recognition. *IEEE Trans. Multimedia* **24**, 1750–1762 (2022).

[5] Varsha, M. & Nair, C. S. Indian Sign Language Gesture Recognition Using Deep Convolutional Neural Network. in *2021 8th International Conference on Smart Computing and Communications (ICSCC)* 193–197 (IEEE, Kochi, Kerala, India, 2021). doi:10.1109/ICSCC51209.2021.9528246.

[6] Bhat, A., Yadav, V., Dargan, V., & Yash. Sign Language to Text Conversion using Deep Learning. in *2022 3rd International Conference for Emerging Technology (INCET)* 1–7 (IEEE, Belgaum, India, 2022). doi:10.1109/INCET54531.2022.9824885.

[7] Dutta, K. K., Raju K., S. K., Kumar G.S., A. & Swamy B., S. A. Double handed Indian Sign Language to speech and text. in *2015 Third International Conference on Image Information Processing (ICIIP)* 374–377 (IEEE, Waknaghat, 2015). doi:10.1109/ICIIP.2015.7414799.

[8] Dutta, K. K. & Bellary, S. A. S. Machine Learning Techniques for Indian Sign Language Recognition. in *2017 International Conference on Current*

Trends in Computer, Electrical, Electronics and Communication (CTCEEC) 333–336 (IEEE, Mysore, 2017). doi:10.1109/CTCEEC.2017.8454988.

[9] Y. Zhao and L. Wang, "The Application of Convolution Neural Networks in Sign Language Recognition," 2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP), Wanzhou, China, 2018, pp. 269-272, doi: 10.1109/ICICIP.2018.8606707:

[10] Shi, D., Long, L., Zhang, Y., He, H. & Liu, X. Sign Language Recognition System based on Jetson TX2 and Yolov5. in *2022 4th International Symposium on Smart and Healthy Cities (ISHC)* 242–246 (IEEE, Shanghai, China, 2022). doi:10.1109/ISHC56805.2022.00051.

[11] Qin, W. *et al.* Sign Language Recognition and Translation Method based on VTN. in *2021 International Conference on Digital Society and Intelligent Systems (DSInS)* 111–115 (IEEE, Chengdu, China, 2021). doi:10.1109/DSInS54396.2021.9670588.

[12] Dabwan, B. A. & Jadhav, M. E. Recognition of American Sign Language Using Deep Convolution Network. in *2022 International Conference on Intelligent Technology, System and Service for Internet of Everything (ITSS-IoE)* 1–4 (IEEE, Hadhramaut, Yemen, 2022). doi:10.1109/ITSS-IoE56359.2022.9990966.

[13] Agarwal, A., Sreemathy, R., Turuk, M., Jagdale, J. & Kumar, V. Indian Sign Language Recognition using Skin Segmentation and Vision Transformer. in *2023 IEEE 20th India Council International Conference (INDICON)* 857–862 (IEEE, Hyderabad, India, 2023). doi:10.1109/INDICON59947.2023.10440818.

[14] Singh, A. et al. Indian Sign Language Recognition System for Dynamic Signs. in 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) 1–6 (IEEE, Noida, India, 2022). doi:10.1109/ICRITO56286.2022.9964940.

[15] Aggarwal, P., Saini, R., Rani, R., Dev, A. & Jaiswal, G. Indian Sign Language classification and Image Augmentation using GAN. in *2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)* 1–6 (IEEE, Bangalore, India, 2023). doi:10.1109/SMARTGENCON60755.2023.10441873.

[16] Joshi, S., Sharma, B., Dharmender, Tripathi, D. & Shah, A. Hand Sign Language Detection Using Deep Learning and Computer Vision. in *2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE)* 1–4 (IEEE, Ballari, India, 2023). doi:10.1109/AIKIIE60097.2023.10390267.

[17] Ekbote, J. & Joshi, M. Indian sign language recognition using ANN and SVM classifiers. in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* 1–5 (IEEE,

Coimbatore, 2017). doi:10.1109/ICIIECS.2017.8276111.

[18] Garg, C., Attry, G. P., Mittal, D. & Bhat, A. Conversion of Native Speech into Indian Sign Language to Facilitate Hearing Impairment. in *2023 4th International Conference for Emerging Technology (INCET)* 1–9 (IEEE, Belgaum, India, 2023). doi:10.1109/INCET57972.2023.10170244.

[19] Pigou, Lionel, Mieke Van Herreweghe, and Joni Dambre. "Gesture and sign language recognition with temporal residual networks." In Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 3086-3093. 2017.

[20] Al-Hammadi, Muneer, Ghulam Muhammad, Wadood Abdul, Mansour Alsulaiman, Mohamed A. Bencherif, and Mohamed Amine Mekhtiche. "Hand gesture recognition for sign language using 3DCNN." IEEE Access 8 (2020): 79491-79509.

[21] Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi. "Understanding of a convolutional neural network." In 2017 International Conference on Engineering and Technology (ICET), pp. 1-6. Ieee, 2017

LIST OF PUBLICATION AND THEIR PROOFS

Title of Paper 1- "Insights into Indian Sign Language Recognition: A Comprehensive Review"

Author Names- Ravi Raj and Dr. Aruna Bhat

Name of the Conference- World Congress on Smart Computing (WCSC2024)

Date of Conference- June 08, 2024 – June 09, 2024

Indexing- Springer

Status of Paper- Accepted

Date of Camera-Ready Submission and Registration- May 24, 2024

Title of Paper 2 - "Evaluating Various CNN Models for Indian Sign Language Recognition"

Author Names- Ravi Raj and Dr. Aruna Bhat

Name of the Conference- Hinweis Second International Conference on Computer Science, Cyber Security and Information Technology (CCIT- **http://ccit.thehinweis.com/2024/**)

Date of Conference- June 29, 2024 – June 30, 2024

Indexing- Scopus

Status of Paper- Accepted

Date of Camera-Ready Submission and Registration- June 18, 2024

DTU Delhi Technological University

2K22AFI14 RAVI <ravi_2k22afi14@dtu.ac.in>

WCSC 2024: Paper ID 315: Acceptance

1 message

Microsoft CMT <email@msr-cmt.org> Sun, May 5, 2024 at 11:44 PM
Reply-To: WCSC SCRS <wcsc.scrs@gmail.com>
To: RAVI RAJ <ravi_2k22afi14@dtu.ac.in>

Dear RAVI RAJ,

Greetings!

Thank you for submitting your research article to the World Congress on Smart Computing (WCSC 2024) to be held on June 08-09, 2024 at Babu Banarasi Das University, Lucknow, India in Hybrid Mode.

We are pleased to inform you that based on reviewers' comments, your paper titled "Insights Into Indian Sign Language Recognition: A Comprehensive Review" has been accepted for presentation during WCSC 2024, and publication in the proceedings to be published in Springer Book Series 'Studies in Smart Technologies (https://www.springer.com/series/17410)' subject to the condition that you submit a revised version as per the comments, available at Authors CMT account. It is also required that you prepare a response to each comment from the reviewer and upload it as a separate file along with the revised paper.

The similarity index in the final paper must be less than 20%. Please note that the high plagiarism and any kind of multiple submissions of this paper to other conferences or journals will lead to rejection at any stage. Please note that the publisher, i.e. Springer Nature may ask for any other changes during the production which are supposed to be implemented. The publisher has the final right to exclude the paper from the proceedings if they found it unsuitable for publication.

Please carry out the steps to submit the camera-ready paper and online registration as per the instructions available at

https://www.scrs.in/conference/wcsc2024/page/Camera_Ready_Paper_Submission

In order to register in the SCRS member category (subsidized registration fees), you can first become a member at https://www.scrs.in/register and then register for the conference OR you may register as a Regular Author Category.

Please note that the Last date for submission of the camera-ready paper, payment of the registration fee, and online registration is May 24, 2024.

Feel free to write to the "General Chairs, WCSC 2024" at wcsc.scrs@gmail.com, should you have any questions or concerns. Please remember to always include your Paper ID- 315, whenever inquiring about your paper. For regular updates on WCSC 2024, please join the SCRS Telegram channel https://t.me/+78ZOewUxF_AwNzhl

Looking forward to meeting you during the conference.

With Regards
Team WCSC 2024
Babu Banarasi Das University, Lucknow, India
Follow us on WhatsApp Channel: https://whatsapp.com/channel/0029Va9bGHL5a23wuhGhj82b

CONFERENCE SERVICES

Gurunanak Nagar, Gwalior,
Madhya Pradesh, India - 474012
Phone: +91 76928 04154
Email: info@confservices.in

TAX-INVOICE

Conference Services Website: https://confservices.in/ GSTIN: 23BGUPS7760Q1Z6	**Invoice Date**	24 May 2024
	Invoice Number	CS/MAY/2024/82

BILL ISSUED TO

RAVI RAJ
c/o manju kumari vill+post+thana=kartahan ,kartahan,district-vaishali jagdishpur,kartahan buzurg
ravi_2k22afi14@dtu.ac.in
For Paper ID: 315
Paper Title: Insights into Indian Sign Language Recognition: A Comprehensive Review

DESCRIPTION	
Event Name	World Congress on Smart Computing (WCSC2024) https://www.scrs.in/conference/wcsc2024
Event Dates	08 Jun 2024 - 09 Jun 2024
Category of Registration	Regular Author (Online Attendance)
Registration Fee	INR 6000
Extra Page Charges	INR 0
GST (18%)	INR 1080
Total	**INR 7080.00** **(SEVEN THOUSAND EIGHTY RUPEES ZERO PAISE)**

Payment Transaction Id	Date Time	Mode of Payment
WCSC2024-315-20240524210350000000894397698727963	24 May 2024 10:36:56	Paytm

Note: Whether tax is payable under reverse charge - No

This is a computer generated invoice and needs no signature.

DTU Delhi Technological University

2K22AFI14 RAVI <ravi_2k22afi14@dtu.ac.in>

Fwd: CCIT2024 :: Acceptance Confirmation and Registration Details

1 message

Hinweis Conference <papers.ccit@gmail.com> Sat, May 25, 2024 at 12:23 PM
To: ravi_2k22afi14@dtu.ac.in

Dear Author,

We are resending the review results that were initially sent on May 20, 2024. Kindly review the attached document at your earliest convenience and proceed accordingly.

---------- Forwarded message ---------
From: **Hinweis Conference** <papers.ccit@gmail.com>
Date: Mon, May 20, 2024 at 11:28 AM
Subject: CCIT2024 :: Acceptance Confirmation and Registration Details
To: <ravi2k22afi14@dtu.ac.in>, <aruna.bhat@dtu.ac.in>

Dear CCIT-2024 Author,

Warm greetings from Hinweis Research!

We are thrilled to inform you that your submitted paper for **Hinweis Second International Conference on Computer Science, Cyber Security and Information Technology (CCIT)** has been accepted. Congratulations on this significant achievement! Your dedication to your research is highly commendable.

http://ccit.thehinweis.com/2024/

Here are the important details regarding your acceptance:

Review Result and Acceptance Certificate:

The consolidated review result is attached along with this email. The review result itself is the acceptance certificate.

Publication and Indexing:

All accepted conference papers will be published in the **Conference Proceedings** with an **ISBN Number** and will be indexed by **Scopus and Crossref**, further enhancing the visibility of your work.

Camera Ready Submission:

To proceed with the publication process, please ensure that your paper adheres to the standard format. Your camera-ready submission should be in Microsoft Word DOC or DOCX format, and it should not exceed the stipulated page limit, including text, figures, tables, and references. Kindly send your camera-ready submission to **papers.ccit@gmail.com**.

Payment Mode:

Payment subject/reference (**Name and Paper ID**) should be mentioned on all the payments remarks, in order to credit it towards your requirement.

Registration Information:

Other Bank Transfer

INB Reference Number	IRY6933494	06-Jun-2024 [03:02 PM IST]
Debit Transaction Status	Processed	

Debit Account Details

SBI Account No	Account Type	SBI Branch	Amount	Commission Amount	Transaction Type
00000020087898832	Savings Account	ANJANPIR	INR7000.00	INR0.00	QKT

Credit Account Details

Account No.	Bank	Branch	Transfer Type	Amount	Purpose
034985800007360			QKT	7000.00	PAPER SUBMISSION FEE

15% Overall Similarity

Top sources found in the following databases:

- 7% Internet database
- 8% Publications database
- Crossref database
- Crossref Posted Content database
- 11% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	ijraset.com Internet	1%
2	mdpi.com Internet	<1%
3	Nikolas Adaloglou, Theocharis Chatzis, Ilias Papastratis, Andreas Ster... Crossref	<1%
4	link.springer.com Internet	<1%
5	srmap on 2023-12-03 Submitted works	<1%
6	American University of Culture and Education on 2023-09-20 Submitted works	<1%
7	Hira Hameed, Muhammad Usman, Muhammad Zakir Khan, Amir Hussa... Crossref	<1%
8	University of Northumbria at Newcastle on 2023-05-18 Submitted works	<1%

9	openaccess.thecvf.com Internet	<1%
10	"Advanced Computing", Springer Science and Business Media LLC, 2021 Crossref	<1%
11	University of Technology, Sydney on 2024-05-12 Submitted works	<1%
12	arxiv.org Internet	<1%
13	eudl.eu Internet	<1%
14	Janpreet Singh, Harjeet Singh, Vishal Goyal. "A Comprehensive Study o... Crossref	<1%
15	"Soft Computing for Security Applications", Springer Science and Busin... Crossref	<1%
16	authorea.com Internet	<1%
17	"Computational Vision and Bio-Inspired Computing", Springer Science ... Crossref	<1%
18	ijisrt.com Internet	<1%
19	geeksforgeeks.org Internet	<1%
20	Gitam University on 2023-10-30 Submitted works	<1%

21	University of Wollongong on 2023-11-19 Submitted works	<1%
22	frontiersin.org Internet	<1%
23	Kean University on 2024-05-29 Submitted works	<1%
24	University of Hong Kong on 2018-07-30 Submitted works	<1%
25	ucalgary.ca Internet	<1%
26	University of Liverpool on 2023-03-23 Submitted works	<1%
27	bigd.bracu.ac.bd Internet	<1%
28	Dayananda Sagar University, Bangalore on 2023-11-23 Submitted works	<1%
29	Indian School of Mines on 2024-05-07 Submitted works	<1%
30	Management Development Institute Of Singapore on 2023-12-01 Submitted works	<1%
31	Sourav Pal, Sourabh Ruhella, Sparsh Sinha, Anurag Goel. "Sign Langua... Crossref	<1%
32	University of Westminster on 2024-02-07 Submitted works	<1%

33	researchonline.gcu.ac.uk Internet	<1%
34	scholarworks.sjsu.edu Internet	<1%
35	"Cognitive Computing and Information Processing", Springer Science a... Crossref	<1%
36	Aruna Bhat, Vinay Yadav, Vishesh Dargan, Yash. "Sign Language to Te... Crossref	<1%
37	Liverpool John Moores University on 2022-08-28 Submitted works	<1%
38	Liverpool John Moores University on 2024-03-18 Submitted works	<1%
39	University of East London on 2023-05-11 Submitted works	<1%
40	University of Hertfordshire on 2023-12-04 Submitted works	<1%
41	University of Hertfordshire on 2024-04-17 Submitted works	<1%
42	University of Portsmouth on 2024-05-20 Submitted works	<1%
43	Zhang, Rui, Sheng Tang, Wu Liu, Yongdong Zhang, and Jintao Li. "Multi... Crossref	<1%
44	Bournemouth University on 2023-09-08 Submitted works	<1%

45	City University on 2022-10-02 Submitted works	<1%
46	King Abdulaziz University on 2022-10-31 Submitted works	<1%
47	Kingston University on 2023-09-18 Submitted works	<1%
48	Liverpool John Moores University on 2021-12-22 Submitted works	<1%
49	Simranjit Bhumbla, D K Gupta, Nisha. "A Review: Object Detection Algo... Crossref	<1%
50	Technical University of Košice on 2024-05-28 Submitted works	<1%
51	University of Strathclyde on 2017-09-25 Submitted works	<1%
52	encyclopedia.pub Internet	<1%
53	openresearch.lsbu.ac.uk Internet	<1%
54	"Computational Vision and Bio-Inspired Computing", Springer Science ... Crossref	<1%
55	George Bush High School on 2023-05-27 Submitted works	<1%
56	Goldey-Beacom College on 2023-05-11 Submitted works	<1%

57	M. Giridhar Kumar, Nagaraju Kolla, Veeranna Kotagi, Mahantesh Matha... Crossref	<1%
58	Middlesex University on 2023-10-08 Submitted works	<1%
59	Mohammad Daim Khan, B Shivalal Patro, Rajiv Ranjan, Manas Chandan... Crossref	<1%
60	Sagawa, H.. "Pattern Recognition and Synthesis for a Sign Language Tr... Crossref	<1%
61	The British College on 2024-01-25 Submitted works	<1%
62	University of East London on 2024-05-06 Submitted works	<1%
63	University of Macau on 2019-11-07 Submitted works	<1%
64	University of Westminster on 2024-03-25 Submitted works	<1%
65	Whitireia Community Polytechnic on 2018-03-04 Submitted works	<1%
66	addi.ehu.es Internet	<1%
67	cdn.techscience.cn Internet	<1%
68	digitalcollection.zhaw.ch Internet	<1%

69	emitter.pens.ac.id Internet	<1%
70	iaesprime.com Internet	<1%

Excluded from Similarity Report

- Bibliographic material
- Small Matches (Less then 8 words)
- Manually excluded text blocks

EXCLUDED TEXT BLOCKS

CHAPTER 1Introduction1.1 OverviewAccording to World Health Organization

qspace.library.queensu.ca

K-Nearest Neighbors(K-NN) and BackPropagation Neural

maxwellsci.com