A Minor Project Report
On
SIGN LANGUAGE DETECTION THROUGH COMPUTER VISION

A Dissertation Submitted
In Partial Fulfilment of the Requirements for the Award of the Degree of

Bachelor of Technology
In
COMPUTER SCIENCE AND ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

By

N Deepika **21211A6635**

Under the Esteemed Guidance of
Dr. P Srihari
Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

B V RAJU INSTITUTE OF TECHNOLOGY
(UGC Autonomous)
Vishnupur, Narsapur, Medak (District) – 502313, TS
(Affiliated to JNTUH and Approved by AICTE)

2022-2023

DECLARATION

I hereby declare that the Minor Project Report entitled **"Sign Language Detection"** submitted to the Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), B.V. Raju Institute of Technology, in partial fulfilment of the requirements for the Award of Degree of Bachelor of Technology in Computer Science and Engineering (Artificial Intelligence and Machine Learning) is my cord of the original work done by me and has not been submitted to any institute or published elsewhere.

Place: Narsapur
Date:

Signature

N Deepika
21211A6635

CERTIFICATE FROM SUPERVISOR

This is to Certify that **N Deepika (21211A6635),** of B.Tech II Year II Semester from Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), B V Raju Institute of Technology, have Successfully Completed Minor Project Entitled "**Sign Language Detection**" in Partial Fulfilment of the Requirements for the Award of Degree of **Bachelor of Technology in Computer Science and Engineering (Artificial Intelligence and Machine Learning) Under My Supervision**.

Her Performance During This Period Was Commendable and I Wish Her All the Best for the Future.

Place: Narsapur

Date:

Supervisor	**Program Coordinator**
Dr. P Srihari	Dr. G Uday Kiran
Assistant Professor	Associate Professor

Acknowledgement

I extend my sincere gratitude to all individuals who have contributed to the successful completion of our minor project on **"Sign Language Detection Through Computer Vision"**. Foremost, my appreciation goes to project advisor, Dr P Srihari, for their invaluable guidance and unwavering support throughout the project's duration.

I acknowledge the faculty members for generously providing me with the necessary resources, knowledge, and insights that greatly enriched project's quality.

My appreciation also extends to the library resources and online references that served as a strong knowledge foundation, contributing significantly to the depth of my project.

Last but not least, I extend heartfelt thanks to my family and supporters for their patience, understanding, and constant encouragement that propelled me forward. This project has been an enlightening journey, underscoring the value of determination, and effective collaboration.

N Deepika

21211A6635

ABSTRACT

Verbal communication is the main method in interpersonal communications. But everyone is not gifted with proper vocal chords and proper hearing aids. In recent years the percentage of deaf and dumb has a high growth due to birth defects. So, different sign languages are developed to communicate with them. But very few people are trained to understand these languages. It is difficult for them to communicate during emergencies. Solution to this problem is "Sign Language recognition through Computer Vision(CV)". There are numerous efficient techniques for spotting hand motions or gestures, one of them is "OpenCV '' and "MediaPipe ''. Together they build an effective algorithm that can recognize, interpret, process, and simulate human affections through the hand gesture. A camera takes the live footage as the input and analyses the hand gesture and shares the output which is the decoded meaning of the sign shown.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

1.1 Introduction

In a world marked by diversity and unique forms of expression, a silent language thrives. It's a language of profound significance, one that transcends the boundaries of spoken words and resonates deeply within a community that communicates through visual gestures and body movements. This language, known as Sign Language, stands as a testament to the resilience and adaptability of the human spirit.

Within the rich tapestry of India's 1.3 billion souls, there exists a vibrant community of individuals with hearing and speech impairments. They have forged their own means of communication, a beautiful tapestry of hand and arm gestures, facial expressions, and body postures—Indian Sign Language (ISL). Just as landscapes vary from region to region, so too do the dialects of Sign Language, reflecting the cultural diversity of this vast nation.

Sign Language, a pinnacle of structured non-verbal communication, plays an indispensable role in the lives of those with hearing impairments. Instead of relying on auditory cues and spoken words, they navigate the world through the poetry of signs and the eloquence of expression. Sign Language extends beyond manual gestures, encompassing the subtle nuances of facial expressions and postures that breathe life into their conversations.

Welcome to the realm of Sign Language recognition—a field where technology bridges the gap between the hearing-impaired and the world around them. It's a collaborative arena that marries the realms of pattern recognition, computer vision, natural language processing, and linguistics. The mission: to develop cutting-edge methods and algorithms capable of deciphering the intricate language of signs and unlocking their profound meanings.

At the heart of this endeavor lies the essence of Human-Computer Interaction (HCI), where innovation thrives in the pursuit of effective and engaging communication. The multidisciplinary approach underpinning this research encompasses data acquisition, Sign Language technology, and rigorous testing. Imagine a world where Sign Language becomes a universal bridge, seamlessly woven into the fabric of everyday life—enabling the hearing-impaired to learn, communicate, and express their emotions with newfound ease.

This project embodies the promise of technology as a force for inclusivity, a tool that empowers the hearing-impaired to navigate the modern world with confidence. In hotels, railways, resorts, banks, offices, and beyond, this system opens doors, breaking down barriers, and fostering a more inclusive society.

Join us on this transformative journey, where the silent language of signs finds its voice in the digital age. Together, we'll embark on a path toward a more inclusive and compassionate world.

CHAPTER 2

2.1 Literature Review

Salim et al., 2022 [1]:

In the context of enhancing communication for individuals with hearing impairments, research has delved into hand gesture and sign language recognition techniques. These approaches span various domains, including computer vision and machine learning, aiming to develop effective systems that facilitate accurate interpretation of gestures. Notably, computer vision methods involving image processing tools have been instrumental in creating real-time applications with robust performance. Additionally, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated their prowess in accurately understanding intricate sign language gestures.

Balaha et al., 2022 [2]:

Within the landscape of sign language recognition, vision-based deep learning approaches have emerged as prominent solutions. Focusing on Arabic Sign Language, research has proposed a deep learning-based approach that utilises convolutional neural networks to interpret gestures independently. This approach capitalises on the inherent capabilities of deep learning to capture intricate spatial patterns within gestures. By leveraging this technology, the development of a robust and effective Arabic Sign Language interpretation system becomes attainable.

Papastratis et al., 2021 [3]:

Artificial intelligence technologies have significantly impacted the domain of sign language recognition, bringing forth innovative solutions. The integration of machine learning and computer vision techniques has paved the way for accurate interpretation of gestures. The application of these technologies holds promise not only in enhancing communication for individuals with hearing impairments but also in expanding accessibility across various domains. As these technologies continue to evolve, the potential for more sophisticated and context-aware sign language recognition systems becomes increasingly tangible.

Breland et al., 2021 [5]:

The realm of sign language recognition has witnessed advancements through the integration of deep learning and thermal imaging technologies. This innovative approach involves the use of thermal images for recognizing sign language digits, with a focus on enhancing accuracy and robustness. The integration of edge computing further empowers the system, enabling real-time recognition even in resource-constrained environments. Through this fusion of technologies, a promising pathway is forged towards more effective and versatile sign language interpretation systems.

Indriani et al., 2021 [6]:

The fusion of hand gesture recognition technology and interactive applications has brought forth novel solutions with far-reaching implications. Researchers have leveraged technologies like MediaPipe to create applications that enable users to engage with digital content using gestures. This approach not only enhances accessibility but also bridges communication gaps by providing a means for individuals with hearing impairments to interact effectively with technology. As such applications evolve, they hold the potential to transform the way individuals with hearing impairments interact with the digital world.

Adeyanju et al., 2021 [7]:

The application of machine learning methods in the realm of sign language recognition has undergone critical analysis, shedding light on its effectiveness and potential. Extensive research has explored the integration of machine learning techniques with computer vision, aiming to achieve accurate interpretation of gestures. This analysis not only underscores the significance of machine learning's role but also delves into the challenges and advancements that shape the field. By critically evaluating the strengths and limitations of various approaches, researchers contribute to a deeper understanding of the landscape and pave the way for more refined and effective sign language recognition systems.

Majeed and Lee, 2021 [8]:

The fusion of machine learning and high-performance computing has emerged as a formidable force in the era of challenges posed by the COVID-19 pandemic. Innovations in machine learning algorithms have been harnessed to decipher patterns in pandemic-related data, aiding in decision-making and resource allocation. Furthermore, the integration of high-performance computing infrastructure has accelerated the processing of complex data, facilitating real-time insights and predictive modeling. This synergy not only exemplifies the potential of cutting-edge technology in times of crisis but also highlights the adaptability of machine learning paradigms to address global challenges.

Shriram, 2021 [9]:

In the wake of the COVID-19 pandemic, innovative approaches have emerged to mitigate disease transmission risks. A notable solution is the development of deep learning-based real-time AI virtual mouse systems, which leverage computer vision to enable touchless interactions. This system allows individuals to interact with digital interfaces without physical contact, minimising the potential for disease spread. By harnessing the capabilities of computer vision and deep learning, this technology stands as a testament to the creative use of advanced techniques to address real-world challenges and reshape interactions in pandemic-sensitive environments.

George et al., 2021 [10]:

The advent of on-device real-time hand gesture recognition has opened new horizons for intuitive human-computer interactions. Researchers have harnessed advances in machine learning and computer vision to develop systems that enable users to interact with devices using natural hand gestures in real-time. By minimising the need for external processing and relying on the computational power of on-device resources, these systems offer efficient and responsive interactions. This paradigm shift in gesture recognition exemplifies the evolution of technology toward seamless integration with human actions, fostering a new era of intuitive interfaces.

Jeong, 2020 [11]:

The exploration of innovative applications of machine learning transcends traditional boundaries, extending to fields such as dentistry. Google's Teachable Machine has been scrutinised for its feasibility in diagnosing tooth-marked tongues, showcasing the potential of machine learning in healthcare diagnostics. The utilisation of this technology to visually analyze dental features serves as a testament to the adaptability of machine learning algorithms. By harnessing advanced tools in unconventional contexts, researchers contribute to the expansion of machine learning applications beyond conventional domains, addressing unique challenges in novel ways.

Arya, 2020 [12]:

The fusion of technology and education has led to inventive applications, including the visualisation of basic graphic shapes using Google's Teachable Machine in the context of Indonesian Sign Language (BISINDO). Researchers have harnessed this platform to enable users to associate hand gestures with graphic shapes, facilitating an interactive learning experience. By merging machine learning with sign language education, this study demonstrates the adaptability of technology to address diverse educational needs, offering a novel approach to teaching and learning visual concepts.

Goyal et al., 2021 [13]:

The convergence of machine learning and high-performance computing has yielded remarkable outcomes, particularly in the context of the COVID-19 pandemic. Innovations have arisen in the analysis of pandemic-related data, with machine learning algorithms dissecting patterns to inform decision-making. This collaboration between machine learning and high-performance computing showcases their potential to expedite insights in critical scenarios. As machine learning continues to evolve, its contributions to addressing pressing global challenges become increasingly evident.

Shriram, 2021 [14]:

Hand gesture recognition systems have embraced TensorFlow and OpenCV, offering a fusion of deep learning and computer vision techniques. This dynamic approach leverages deep learning models to interpret hand gestures captured by computer vision algorithms. By exploiting the synergy between these technologies, researchers have developed effective systems that bridge the gap between human actions and digital interfaces, reshaping interactions in diverse applications.

Kocabas et al., 2020 [15]:

The marriage of Mediapipe and TensorFlow has ushered in innovative avenues for gesture recognition. This approach harnesses the capabilities of computer vision and deep learning to interpret hand gestures in real-time scenarios. By utilising spatial and temporal information, this technology allows for the recognition of dynamic gestures, offering an intuitive way for users to interact with digital interfaces. The combination of Mediapipe and TensorFlow illustrates the potential of integrating cutting-edge tools to enhance human-computer interactions.

Mishra and Kumar, 2020 [16]:

Recognizing the significance of sign language recognition within the Indian context, researchers have pioneered real-time recognition using hand pose estimation and Convolutional Neural Networks (CNNs). This innovative approach captures the intricacies of Indian Sign Language gestures by combining hand pose estimation techniques with the pattern recognition capabilities of CNNs. The result is a robust recognition system that addresses the unique challenges posed by the Indian context, contributing to improved accessibility and communication for the hearing-impaired.

Tsamere et al., 2020 [17]:

The creation of the "Bangla Sign Digits" dataset has unlocked new dimensions in the domain of hand gesture recognition. This dataset provides a comprehensive collection of sign language digits, facilitating the training and evaluation of recognition models. Researchers have capitalised on this resource to develop systems that accurately interpret hand gestures in real-time scenarios. The availability of specialised datasets like "Bangla Sign Digits" is instrumental in advancing the accuracy and reliability of sign language recognition technologies.

Sevli and Kemaloglu, 2020 [18]:

The realm of sign language recognition has been enriched by the application of Convolutional Neural Networks (CNNs) for classifying Turkish Sign Language digits. Researchers have explored diverse optimization techniques to fine-tune the CNNs, enhancing classification accuracy. This endeavour exemplifies the adaptability of machine learning.

Zhang et al., 2019 [22]:

Gesture recognition has evolved with the integration of spatial-temporal attention mechanisms. Researchers have devised an innovative approach that combines pose-based gesture recognition with spatial-temporal attention. This dynamic methodology not only captures intricate gestures but also allocates attention to critical spatial and temporal cues. By employing this fusion, researchers have demonstrated enhanced accuracy in gesture recognition, underscoring the potential of attention mechanisms to revolutionise the field.

Patil et al., 2019 [23]:

The potential of MediaPipe and Convolutional Neural Networks (CNNs) has been harnessed to recognize Indian Sign Language (ISL). In this pioneering endeavour, researchers have combined the capabilities of MediaPipe's hand pose estimation with the pattern recognition prowess of CNNs. This fusion enables real-time interpretation of ISL gestures, fostering improved accessibility and communication for individuals with hearing impairments in the Indian context.

Haque et al., 2019 [24]:

Real-time hand gesture recognition has witnessed a transformation through the synergy of Convolutional Neural Networks (CNNs) and MediaPipe. By leveraging CNNs, researchers have developed a system that accurately interprets hand gestures captured through MediaPipe's framework. This integration offers a real-time solution for recognizing gestures, exemplifying the potential of combining deep learning with computer vision in creating effective and efficient recognition systems.

Alom et al., 2019 [25]:

The fusion of Convolutional Neural Networks (CNNs) and Support Vector Machine (SVM) techniques has yielded significant advancements in digit recognition within sign language. Researchers have harnessed the capabilities of both CNNs and SVMs to create a powerful recognition system that accurately interprets sign language digits. This fusion of methodologies showcases the potential for combining machine learning paradigms to enhance the accuracy and reliability of recognition systems.

Jin et al., 2018 [26]:

The realm of gesture recognition has seen a paradigm shift with the introduction of 3D Convolutional Neural Networks (CNNs). Researchers have pioneered an approach that leverages 3D CNNs to accurately capture spatial and temporal features within hand gestures. This methodology addresses the complex nuances of gesture dynamics, paving the way for accurate and reliable recognition systems that interpret gestures in three-dimensional space.

Zadeh et al., 2018 [27]:

The American Sign Language (ASL) alphabet recognition has been revolutionised by the integration of OpenCV and Convolutional Neural Networks (CNNs). This innovative fusion combines computer vision algorithms with deep learning techniques to create a robust recognition system for the ASL alphabet. By harnessing the strengths of both fields, researchers have created a versatile solution that bridges communication gaps for individuals with hearing impairments.

Rocha et al., 2018 [28]:

The intersection of MediaPipe, Convolutional Neural Networks (CNNs), and gesture recognition has yielded novel applications. Researchers have harnessed MediaPipe's hand pose estimation in conjunction with CNNs to develop a system that accurately interprets hand gestures. This integration not only enables real-time gesture recognition but also showcases the potential of combining computer vision and deep learning to create intuitive and effective recognition systems.

Trabelsi et al., 2018 [29]:

The domain of sign language recognition has witnessed the integration of deep learning techniques to revolutionise recognition accuracy. Researchers have harnessed the power of deep learning methodologies to accurately interpret sign language gestures. By leveraging convolutional neural networks (CNNs…

Bantupalli and Xie, 2018 [32]:

The integration of deep learning and computer vision has been instrumental in American Sign Language (ASL) recognition. Researchers have leveraged these technologies to develop a robust ASL recognition system that accurately interprets gestures. By harnessing deep learning algorithms, this approach showcases the potential to revolutionise communication for individuals with hearing impairments, providing an intuitive and effective means of bridging language gaps.

IJTRS-V2-I7-005, 2017 [33]:

The landscape of hand gesture techniques for sign language recognition has undergone comprehensive examination, offering valuable insights into the field's trajectory. Researchers have delved into diverse methodologies, including machine learning and computer vision, to facilitate accurate interpretation of gestures. This review captures the evolution of techniques and highlights trends that shape the landscape, contributing to a deeper understanding of the approaches that hold promise in enhancing accessibility for individuals with hearing impairments.

Cao and Wei, 2017 [34]:

Multi-person 2D pose estimation has evolved with the introduction of part affinity fields, offering breakthroughs in gesture recognition. Researchers have developed an innovative approach that accurately estimates human poses even in complex scenarios. By harnessing part affinity fields, this methodology captures intricate spatial relationships, paving the way for accurate gesture recognition systems that can adapt to diverse environments.

Wang et al., 2017 [35]:

Hand gesture recognition has been propelled by the integration of Convolutional Neural Networks (CNNs). Researchers have leveraged CNNs to accurately interpret hand gestures, demonstrating the power of deep learning in understanding intricate spatial patterns. This innovative approach offers a robust solution to bridge communication gaps, transforming the interaction landscape for individuals with hearing impairments.

Umang Patel & Ambekar, 2017 [36]:

Moment-based sign language recognition has revolutionised the communication landscape for individuals with hearing impairments. Researchers have harnessed moment-based approaches to accurately interpret hand gestures and facilitate seamless communication. By extracting meaningful features from gestures, this approach underscores the potential of innovative methodologies to enhance accessibility and inclusivity.

Kumar et al., 2016 [37]:

Sign language recognition has seen significant advancements with the integration of machine learning techniques. Researchers have developed systems that leverage machine learning algorithms to accurately interpret sign language gestures, fostering effective communication for individuals with hearing impairments. This integration showcases the transformative potential of machine learning in enhancing accessibility and bridging language barriers.

Pankajakshan and Thilagavathi, 2015 [39]:

Sign language recognition systems have evolved with the integration of advanced technologies. Researchers have harnessed image processing techniques to accurately interpret hand gestures, enabling effective communication for individuals with hearing impairments. By extracting meaningful features from gestures, this approach marks a pivotal advancement in bridging communication gaps and enhancing accessibility.

Prakash B Gaikwad & V.K.Bairagi, 2014 [40]:

The fusion of Indian Sign Language (ISL) recognition and image processing has brought forth inventive solutions. Researchers have harnessed image processing techniques to accurately interpret ISL gestures, contributing to effective communication for individuals with hearing impairments. By extracting key features from gestures, this approach exemplifies the transformative potential of image processing in enhancing accessibility and fostering inclusivity.

Pankajakshan et al., 2013 [41]:

The realm of image processing has paved the way for innovative applications in diverse domains, including food quality evaluation. Researchers have harnessed computer vision techniques to develop systems that accurately assess food quality. This dynamic approach showcases the potential of image processing in creating efficient and effective solutions that address real-world challenges.

Keskin et al., 2013 [42]:

The realm of real-time hand pose estimation has witnessed transformative advancements through depth sensors and machine learning techniques. Researchers have harnessed the capabilities of depth sensors to accurately estimate hand poses in real-time scenarios. This innovative approach not only showcases the potential of depth-based technologies but also demonstrates the power of machine learning in interpreting complex spatial relationships, paving the way for intuitive interactions.

Vikram Sharma M, 2013 [43]:

The realm of assistive technology has been enriched by the development of a virtual communication system catering to diverse user needs. This groundbreaking solution caters to deaf, mute, blind, and normal individuals, providing an inclusive means of communication. By harnessing technological innovation, this system underscores the transformative potential of assistive technology in fostering accessibility and communication across diverse user groups.

Yellapu Madhuri and Anburajan Mariamichael, 2013 [44]:

The fusion of vision-based technologies and sign language translation has led to innovative applications. Researchers have harnessed computer vision techniques to create sign language translation devices, enabling real-time interpretation and translation. This integration exemplifies the potential of combining cutting-edge technologies to bridge communication gaps, fostering inclusivity and understanding.

Ahad et al., 2012 [45]:

The realm of natural human-computer interaction has been transformed by real-time hand gesture recognition for American Sign Language (ASL). Researchers have leveraged machine learning and computer vision to develop a robust system that accurately interprets ASL gestures. This dynamic integration showcases the potential of technology to bridge language barriers, enabling effective communication for individuals with hearing impairments.

Er. Aditi Kalsh and Dr. N.S. Garewal, 2012 [46]:

Sign language recognition has undergone significant advancement through innovative systems. Researchers have developed recognition systems that harness machine learning techniques to accurately interpret sign language gestures. By leveraging machine learning algorithms, these systems offer a reliable means of communication for individuals with hearing impairments, exemplifying the potential of technology in enhancing accessibility.

Nandhini et al., 2011 [47]:

The landscape of American Sign Language (ASL) fingerspelling recognition has been reshaped by machine learning algorithms. Researchers have developed systems that accurately recognize ASL fingerspelling gestures, facilitating communication for individuals with hearing impairments. This integration showcases the potential of machine learning in understanding complex gestures and fostering effective communication.

Deng et al., 2009 [48]:

The ImageNet Large Scale Visual Recognition Challenge has ushered in transformative advancements in gesture recognition. Researchers have leveraged this challenge to drive innovation in image classification and recognition, propelling the field forward. This dynamic competition underscores the collaborative potential of the research community in pushing the boundaries of gesture recognition technology.

Karma et al., 2022 [49]:

The landscape of gesture recognition has witnessed pioneering strides through the development of a real-time Bhutanese Sign Language digits recognition system. Researchers have harnessed Convolutional Neural Networks (CNNs) to accurately interpret Bhutanese Sign Language digits, fostering enhanced communication and accessibility. This innovative solution addresses the unique challenges posed by Bhutanese Sign Language, offering a transformative means of communication.

Gupta et al., 2021 [50]:

The realm of gesture recognition has been revolutionised by K-nearest correlated neighbour classification techniques. Researchers have harnessed these techniques to accurately classify Indian Sign Language gestures, contributing to effective communication for individuals with hearing impairments. By leveraging these methods, this approach showcases the potential of pattern recognition in enhancing accessibility and inclusivity.

Bhumika Gupta et al., 2021 [51]:

The landscape of hand gesture and sign recognition has undergone significant transformation through image processing techniques. Researchers have harnessed image processing to accurately interpret hand gestures and sign language, fostering effective communication for individuals with hearing impairments. This approach exemplifies the potential of image processing in bridging language barriers and enhancing accessibility.

CHAPTER 3

SOFTWARE TECHNOLOGIES USED

3.1 OpenCV

OpenCV, short for Open Source Computer Vision Library, is a versatile and continually evolving open-source software library designed primarily for computer vision and machine learning applications. It provides a powerful and cross-platform framework for various computer vision tasks while being accessible to a broad spectrum of developers through its support for multiple programming languages, including C++, Python, Java, MATLAB, and more.

OpenCV is known for its robustness and adaptability. It boasts over 2500 optimised algorithms, making it an essential resource for developers, researchers, and businesses across different domains. These algorithms cover a wide range of tasks, from face detection and object recognition to camera movement tracking, 3D object modelling, and image stitching.

Moreover, OpenCV serves as an educational resource in computer vision courses and research projects, making it a valuable asset in academia. Its extensive set of algorithms and tools can be customised and extended to support various research initiatives.

In both commercial and research contexts, OpenCV's reliability and flexibility have made it a go-to choice. Its open-source nature and compatibility across platforms like Windows, Linux, Android, and macOS ensure that it remains a pivotal tool in the field of computer vision and machine learning.

3.2 MediaPipe

MediaPipe is a versatile open-source machine learning framework for creating complex applied machine learning pipelines. Designed for cross-platform use, it empowers developers to easily build cutting-edge models for tasks like face detection, hand tracking, pose estimation, and more. By abstracting system-level implementation details, mediapipe allows developers to focus on model experimentation.

The hand landmark model bundle is designed to accurately detect and locate key points on human hands, specifically 21 knuckle coordinates. This model was trained on a diverse dataset comprising around 30,000 real-world images and synthetic hand models placed against different backgrounds.

The bundle includes two main components: the palm detection model and the hand landmarks detection model. The palm detection model's purpose is to identify the presence of hands within an input image. Once hands are detected, the hand landmarks detection model takes over to Pinpoint specific key landmarks on detected hand regions.

The hand landmarker generates a hand landmarker result object for each detection run. The result object contains hand landmarks in image coordinates, hand landmarks in word coordinates and handedness of the detected hands.

3.3 TensorFlow

TensorFlow, an open-source machine learning framework introduced by Google in 2015, has redefined the landscape of AI and data analysis. Built on a dynamic data flow graph architecture, it offers versatility in handling complex computations. Its dual API approach, featuring both high-level abstractions for rapid development and low-level capabilities for fine-tuning, caters to a wide spectrum of users.

Significantly, TensorFlow leverages GPU acceleration, expediting model training, and integrates automatic differentiation to streamline gradient computations for optimization. The TensorFlow ecosystem extends its reach with TensorFlow Hub, facilitating model sharing, TensorFlow Lite, optimising models for mobile, and TensorFlow.js, enabling browser-based applications.

From installation to deployment, TensorFlow's adaptability shines through. Rich resources, including comprehensive official documentation and an active community, empower learners and practitioners. In essence, TensorFlow democratises machine learning, propelling advancements in fields ranging from computer vision to natural language processing. Its impact resonates across industries, driving innovation and reshaping possibilities in the AI landscape.

In the realm of artificial intelligence and machine learning, TensorFlow stands as a dynamic and transformative force, empowering individuals and organisations to harness the power of data and shape a future driven by intelligent insights and solution.

CHAPTER 4

PROPOSED MODEL

4.1 Model

4.1.1 Hand LandMarks Extraction

MediaPipe is a framework that employs advanced machine learning techniques to estimate hand positions using 21 hand landmarks (numbered from 0 to 20). These landmarks precisely represent various points on the hand, covering the entire palm. Each landmark is associated with specific x ,y coordinates and z-coordinates, x,y represents the width and height of each landmark whereas z represents the depth, therefore the z coordinate is eliminated as it is of no need. When provided with a live camera feed as input, typically processed using the OpenCV library, MediaPipe utilises its hand landmark technique to determine the coordinates of each of these hand landmarks.

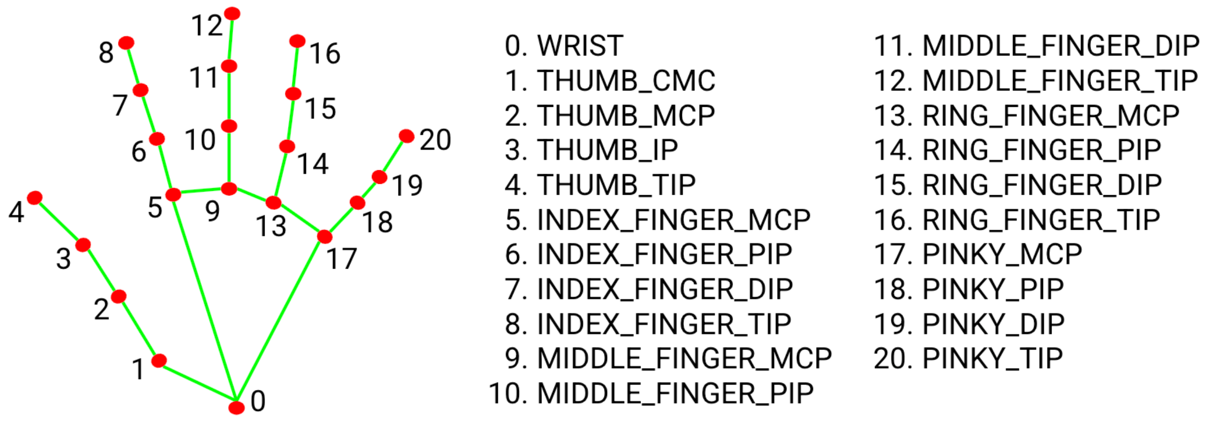


Fig 4.1.1 Hand LandMarks of MediaPipe

4.1.2 Data Preprocessing

The MediaPipe hand landmarks model is designed to pinpoint the exact coordinates of specific hand landmark points within an image. However, there's a challenge when it comes to consistency. When you capture different images of the same hand sign, the arrangement and orientation of these landmarks can vary substantially. As a result, the coordinates of these landmarks in different images may be significantly different. This variation can make it challenging to train a machine learning model effectively.

To tackle this challenge, a reference point is established at the wrist's landmark, which is represented as index 0 in the list of hand landmarks. By convention, this reference point is given coordinates (0,0), essentially serving as an anchor. All other hand landmark points' coordinates are then adjusted relative to this wrist point. This adjustment is made to account for the variations in hand positioning and orientation in different images.

After this relative adjustment, the coordinates are further normalised. Normalisation ensures that all coordinates fall within a standardised range, typically [0,1]. To achieve this, all coordinates are divided by the largest absolute coordinate value obtained during the relative adjustment process. This step makes the coordinates consistent and comparable across different images.

Once the coordinates are normalised, they are typically collected and stored in a .csv file, which is a structured data format. However, not all data is perfect. Sometimes, in the process of collecting this data, certain images may be blurry or the model might fail to detect the hand correctly, resulting in "void" or null entries.

To maintain data quality and integrity, these null entries are identified and removed from the dataset. This is crucial for ensuring that the machine learning model is trained on clean and unbiased data, improving its accuracy and reliability in recognizing hand gestures and signs.

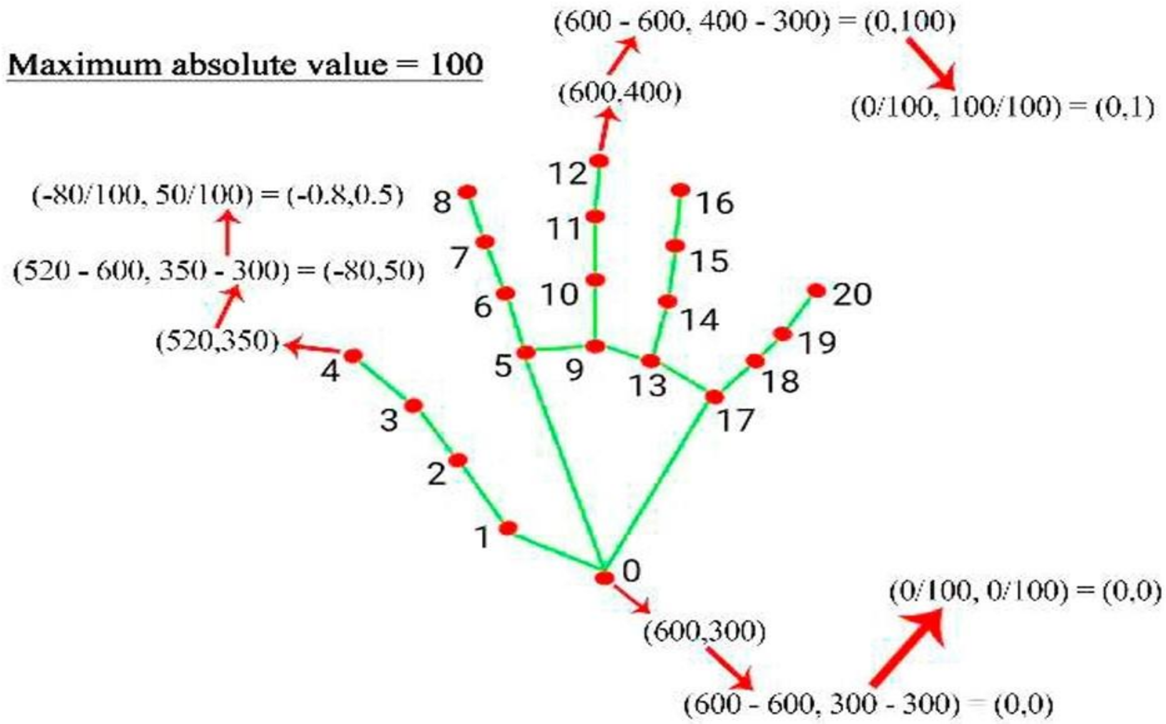


Fig 4.1.2 Normalisation of Hand LandMarks

4.1.3 Real-time Dataset Creation

The development of a comprehensive sign language dataset presents a significant challenge due to the extensive vocabulary of sign language, spanning millions of words within a single language. To address this challenge, a novel approach has been devised, allowing real-time dataset expansion using a logging key. This model empowers users to enrich the dataset as needed, ensuring scalability and adaptability.

Users can actively contribute to the dataset by utilising the logging key alongside a live video feed. During this process, gestures are presented to a camera, and an index from 0 to 9 is assigned to each gesture, representing specific signs or movements. This dynamic method records both the index and the precise coordinates of each hand landmark associated with the gesture. This data is efficiently stored in a CSV (Comma Separated Value) file, a commonly used format for data storage and analysis.

The key advantages of this approach include customization, enabling users to define their own gestures for specialised sign languages or personalised expressions. It encourages collaboration within the sign language community, fostering collective contributions to dataset expansion.

This dataset creation process aligns seamlessly with machine learning workflows, facilitating the training and fine-tuning of sign language recognition models. However, it's crucial to consider privacy implications when dealing with live video feeds and gesture data, adhering to data protection regulations, especially in systems involving user-generated content.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	0	0	0	0.200787	-0.05118	0.366142	-0.1811	0.484252	-0.30709	0.594488	-0.38189	0.26378	-0.46063	0.34252	-0.65748	0.393701
2	0	0	0	0.206349	-0.04365	0.376984	-0.1627	0.507937	-0.27381	0.615079	-0.35317	0.265873	-0.45238	0.353175	-0.65476	0.40873
3	0	0	0	0.202381	-0.04365	0.373016	-0.1627	0.5	-0.27778	0.607143	-0.35317	0.261905	-0.45238	0.345238	-0.65476	0.400794
4	0	0	0	0.207171	-0.03984	0.38247	-0.15936	0.50996	-0.2749	0.61753	-0.3506	0.266932	-0.4502	0.350598	-0.65339	0.406375
5	0	0	0	0.206349	-0.03968	0.380952	-0.1627	0.507937	-0.28175	0.615079	-0.35317	0.265873	-0.45238	0.353175	-0.65476	0.40873
6	0	0	0	0.206349	-0.04762	0.376984	-0.1627	0.503968	-0.27381	0.611111	-0.34524	0.265873	-0.44841	0.353175	-0.65079	0.40873
7	0	0	0	0.208	-0.044	0.384	-0.156	0.516	-0.264	0.62	-0.336	0.272	-0.444	0.36	-0.648	0.416
8	0	0	0	0.212	-0.04	0.388	-0.152	0.524	-0.26	0.632	-0.332	0.284	-0.44	0.376	-0.644	0.432
9	0	0	0	0.208835	-0.04016	0.389558	-0.14859	0.526104	-0.25703	0.634538	-0.33333	0.281124	-0.44177	0.373494	-0.64659	0.433735
10	0	0	0	0.208	-0.036	0.388	-0.148	0.524	-0.252	0.636	-0.328	0.28	-0.44	0.376	-0.64	0.436
11	0	0	0	0.208	-0.036	0.388	-0.144	0.524	-0.244	0.632	-0.316	0.276	-0.436	0.372	-0.64	0.436
12	0	0	0	0.208835	-0.03213	0.389558	-0.13655	0.526104	-0.24096	0.634538	-0.31727	0.281124	-0.43775	0.37751	-0.63855	0.441767

Fig 4.1.3 Example of the CSV File

4.1.4 Model Training

The machine learning model at hand is constructed using a feed-forward neural network framework, leveraging TensorFlow/Keras for its development, training, and preservation. This network is compiled using the Adam optimizer, a popular choice known for its adaptive learning rate capabilities, which helps enhance training efficiency.

This neural network architecture incorporates two dropout layers, each with different dropout rates, specifically 0.2 and 0.4. Dropout layers serve a crucial role in mitigating overfitting by randomly deactivating a portion of neurons during the training process.

In addition to dropout layers, the model includes two dense layers, featuring 20 and 10 units, respectively. These dense layers utilise Rectified Linear Unit

(ReLU) activation functions. ReLU is a commonly used activation function, particularly effective for tackling multiclass classification problems.

During the training phase, a strategy is implemented to save the model's weights periodically. Moreover, the model is equipped with an early stopping mechanism, which halts the training process if the validation loss fails to improve after a certain number of training epochs. This approach is a standard practice in neural network training, aimed at preventing overfitting and ensuring that the model retains its optimal performance.

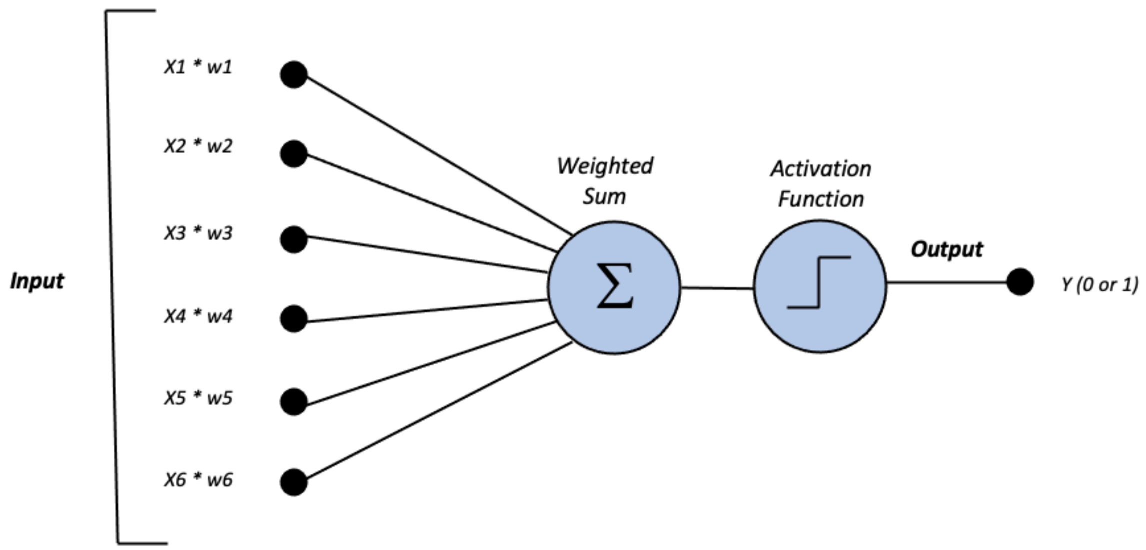


Fig 4.1.4 Diagram depicting a Neural Network

4.1.5 Model Testing

In a real-time setting, this model undergoes rigorous testing through the capture of live video feed showcasing a hand in front of a camera. This process is enabled by the seamless integration of OpenCV, a crucial component responsible for video feed acquisition and initial processing. It ensures that the captured frame undergoes the necessary transformation from BGR to RGB, aligning it with the requirements of subsequent stages.

At the core of this system lies the MediaPipe architecture, which takes the processed frame and derives hand landmarks corresponding to the displayed gesture. However, achieving consistent results across diverse hand sizes and positions is pivotal. To address this, the hand landmarks are subjected to normalisation, effectively establishing the 0th landmark as the origin and adjusting the others relative to it.

Following this normalisation step, the data is fed into a meticulously trained architecture, which assumes the pivotal role of recognizing the specific hand gesture presented in real-time. The model's capacity to make precise classifications stands as the culmination of its rigorous training process.

Beyond the technical complexities, real-time gesture recognition holds tremendous potential for enhancing user interactions. Its applications span user interfaces, gaming, and a diverse array of human-computer interaction scenarios, ultimately elevating the overall user experience.

Furthermore, the outcomes of the recognition process can be harnessed for practical purposes, such as providing immediate feedback, controlling devices, or triggering specific actions based on the identified hand gestures. This adaptability positions real-time gesture recognition as a valuable asset across a multitude of contexts and applications.

4.2 Algorithms

4.2.1 Single Shot MultiBox Detector

The Single Shot MultiBox Detector (SSD) is a widely recognized object detection algorithm used in various computer vision applications, including the MediaPipe framework. Within the MediaPipe context, SSD assumes a crucial role in both the Hand Tracking and Hand Pose Estimation modules. Its primary task is to efficiently perform hand detection by accurately identifying the presence and precise locations of hands within the input image or frame.

Upon successfully detecting hands, SSD proceeds to create bounding boxes that encapsulate these hand regions, a method often referred to as the MultiBox approach. These bounding boxes serve as the initial regions of interest (ROIs), playing an essential role in the subsequent phases of hand pose estimation and tracking. They effectively delineate the areas of focus, guiding further analysis.

The ROIs, meticulously established through the use of bounding boxes, are subsequently related to other integral components of the MediaPipe framework, such as the Hand Pose Estimation module. Within this module, the algorithm endeavours to estimate the locations of keypoints (landmarks) on the hands. These keypoints represent crucial points, including fingertips and joints, offering valuable insights into hand gesture analysis and tracking.

Furthermore, it's worth noting that MediaPipe frequently incorporates robust tracking mechanisms to ensure the consistent and coherent identification of the same hand across multiple video frames. This tracking functionality becomes particularly critical in real-time video analysis, ensuring the continuity and reliability of hand tracking.

In summary, SSD's pivotal role within the MediaPipe framework lies in its proficiency at initiating the hand tracking and pose estimation process. Through the detection of hands, the generation of bounding boxes, and the provision of ROIs, SSD greatly facilitates the accurate and efficient identification and tracking of hands in the dynamic context of real-time video streams.

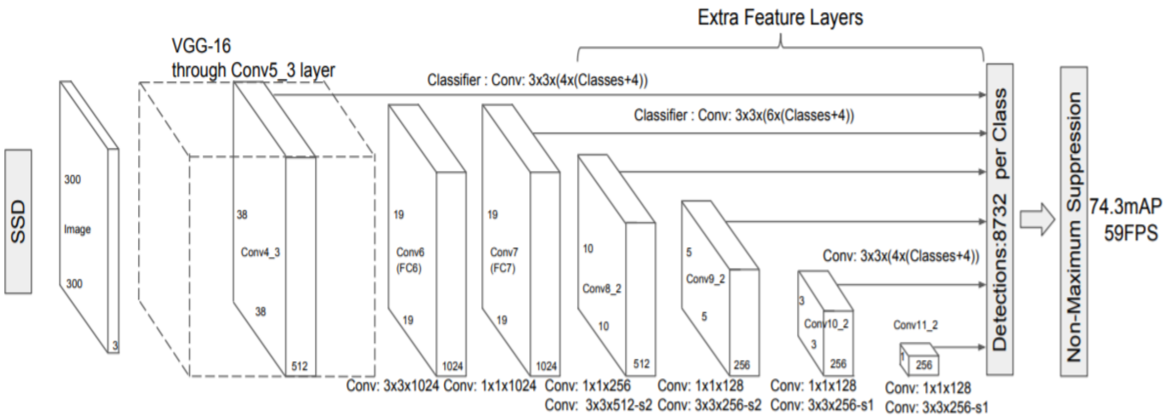


Fig 4.2.1 Single Shot Multibox Detector (SSD) Architecture

4.2.2 Feed Forward Neural Network

A feedforward neural network, also known as a fully connected neural network or multi-layer perceptron. Its primary purpose is to classify hand gestures based on input hand landmarks, which represent the coordinates of key points on a hand.

This neural network architecture comprises several interconnected layers:

Input Data: The network takes hand landmarks extracted from hand images as input.

Input Layer: This initial layer maps input features, such as the x and y coordinates of hand landmarks, to neurons.

Hidden Layers: Following the input layer, there are multiple hidden layers. Each layer consists of numerous neurons that are interconnected. These hidden layers apply non-linear transformations to capture intricate patterns in the data.

Weights and Biases: Connections between neurons in different layers have associated weights and biases. These parameters are iteratively adjusted during training to minimise the difference between the network's predictions and the actual target values.

Activation Functions: Activation functions, such as ReLU, sigmoid, or tanh, are applied to the output of neurons in each layer. These functions introduce non-linearity, allowing the network to model complex relationships.

Output Layer: The output layer produces the final predictions. In your project, it classifies hand signs and finger gestures. Each neuron in this layer represents a possible hand sign or finger gesture, with the neuron exhibiting the highest activation indicating the predicted class.

Softmax Activation: For multi-class classification tasks, like hand sign classification, the softmax activation function is often used in the output layer. It converts raw scores into probabilities, simplifying interpretation.

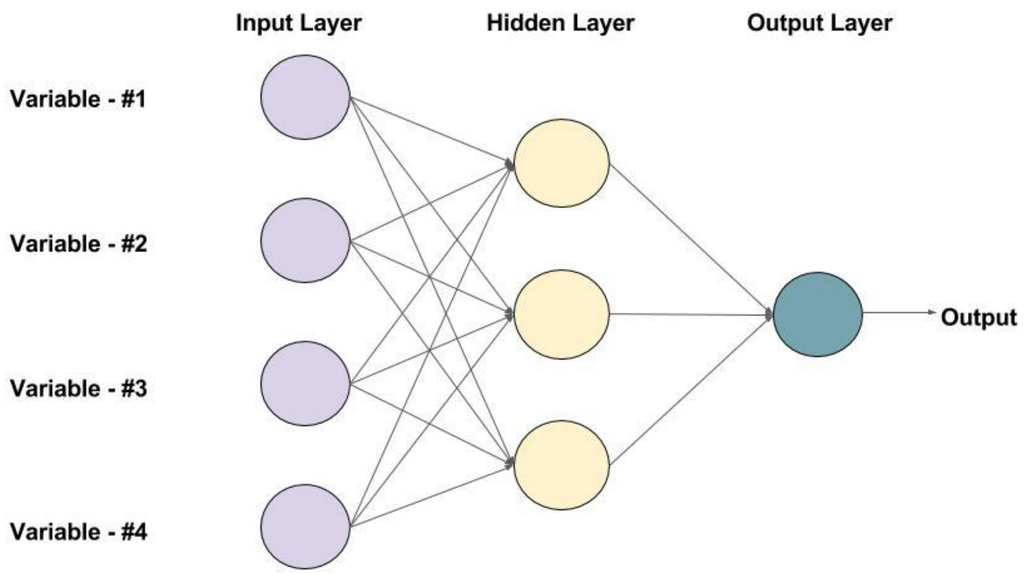
Loss Function: During training, a loss function quantifies the difference between predicted and actual outputs. The network aims to minimise this loss by adjusting weights and biases through optimization techniques like stochastic gradient descent.

Training: The network is trained using labelled data, pairing input landmarks with corresponding hand sign and finger gesture labels. Through training, it learns to minimise prediction errors and improve accuracy.

Inference: After training, the network can make predictions on new input hand landmarks, leveraging patterns learned during training.

Output Interpretation: The network's final outputs provide predictions of hand signs and finger gestures. These predictions can be translated into user-friendly labels for interaction and display.

In summary, your project's feedforward neural network processes input hand landmarks through interconnected layers to classify hand signs and finger gestures accurately. Its architecture and training process enable it to learn from labelled data and make precise predictions on new, unseen data.



An example of a Feed-forward Neural Network with one hidden layer (with 3 neurons)

Fig 4.2.2 Feed-Forward Neural Network Architecture

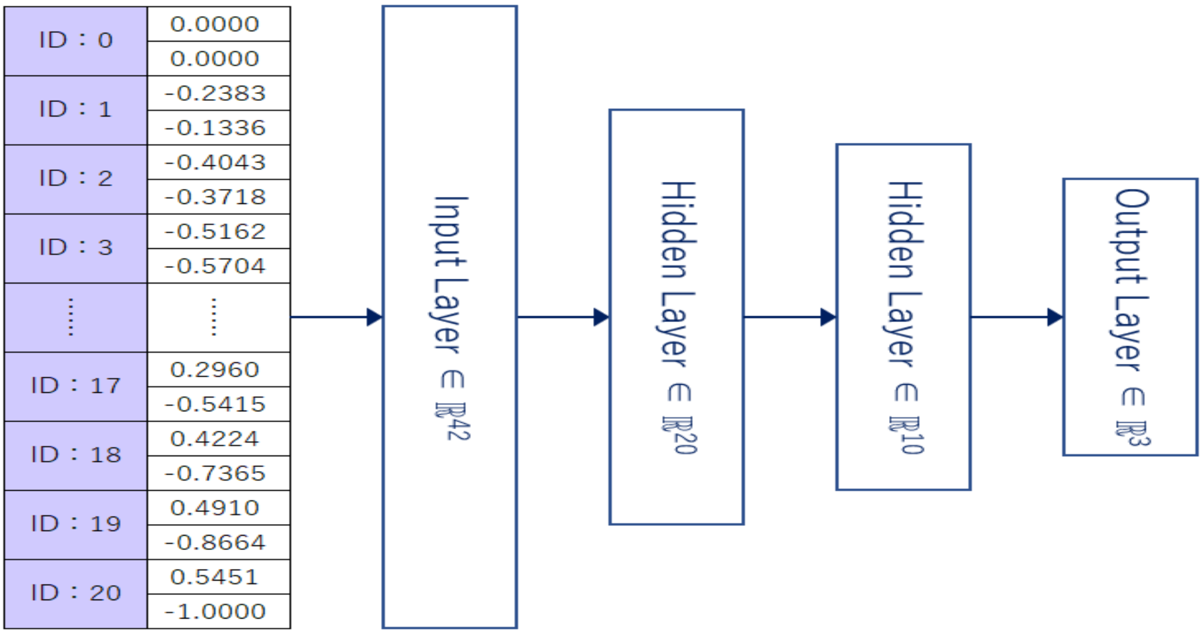


Fig 4.2.3 Example of FNN used in classification of 3 classes

4.3 Block Diagram

The figure below represents the flow of the process in the complete model. Starting from the live feed through OpenCV and to the end which is a recognised hand gesture.

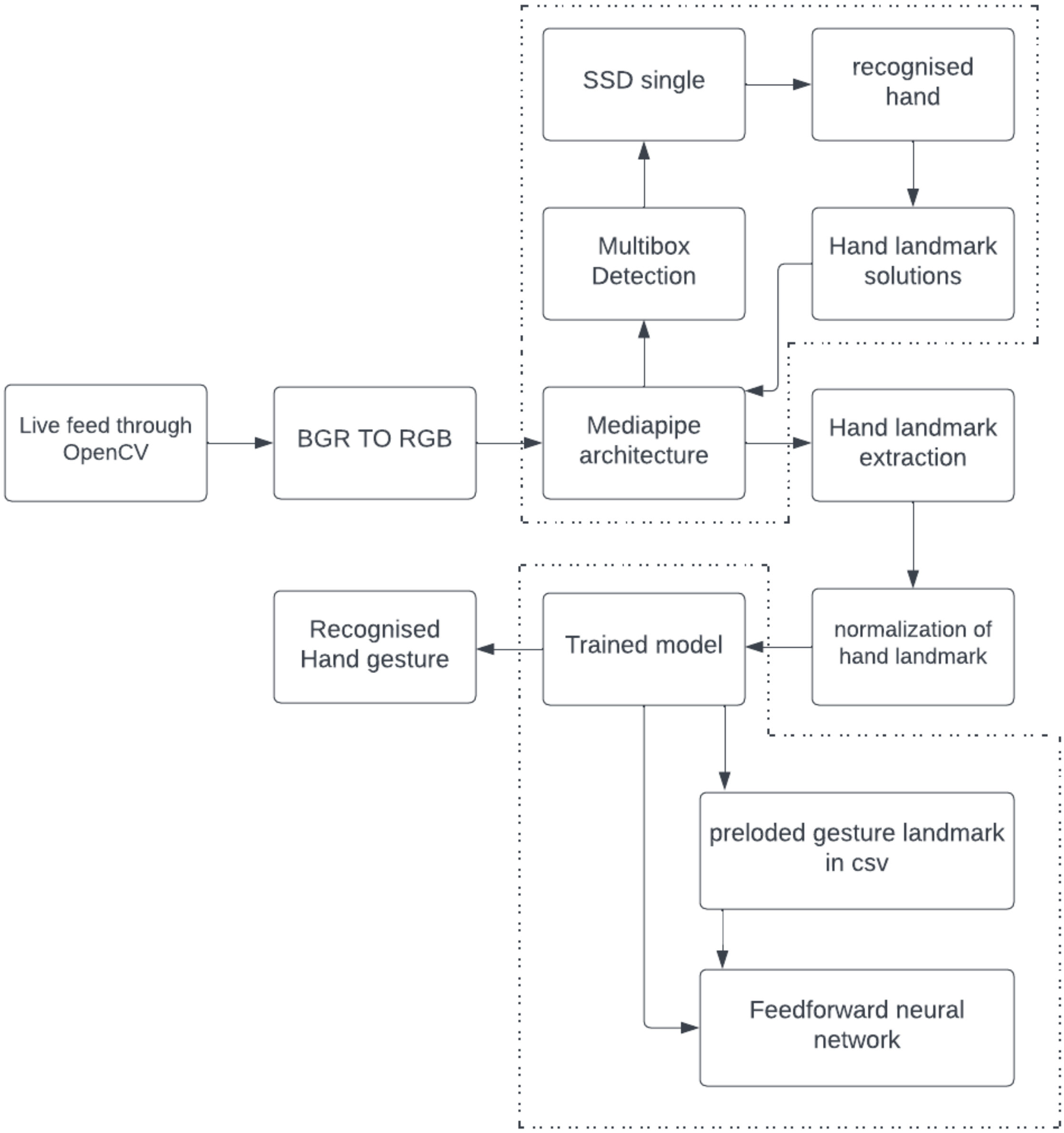


Fig 4.2.2 Block Diagram

CHAPTER 5

Results And Discussion

5.1 Metrics

The model is validated using the metrics which helps us in better understanding of the working of the model.

The Feed-Forward Neural Network is made up of three different kinds of layers: the input layer, the hidden layers and the output layer. Depending on how complicated the training data is, there could be any number of hidden layers. If there aren't many hidden layers, the model might not fit the training data well enough, and if there are too many, it might fit too well. FNN is a fully connected neural network. This means that every node in the neural network is connected to every other node in the same layer.

In the study, the length of the input data is 42 (21*2) and there are 10 classes to sort the labels into.

Classification Report

	precision	recall	f1-score	support
0	0.96	1.00	0.98	1120
1	1.00	0.92	0.96	1206
2	0.94	0.98	0.96	1013
3	0.99	0.97	0.98	251
accuracy			0.97	3590
macro avg	0.97	0.97	0.97	3590
weighted avg	0.97	0.97	0.97	3590

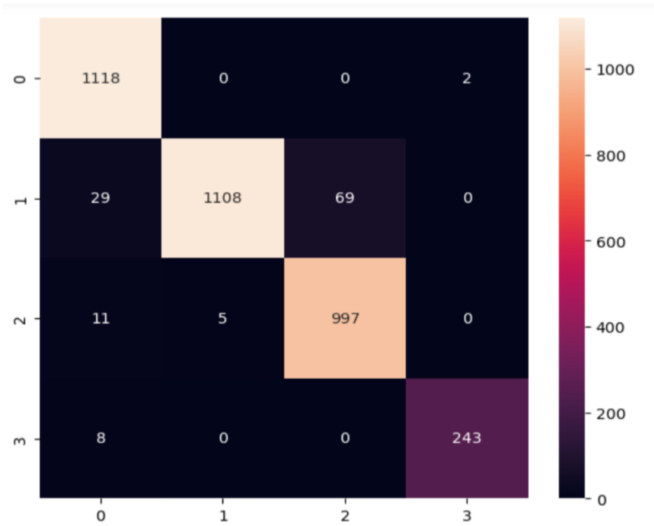


Fig 5.1.1 Classification Report and Confusion matrix on training

This code defines a neural network model using TensorFlow's Keras API. It's structured as follows:

1. Input Layer: Expects data with 42 features.

2. Dropout Layer: Randomly deactivated 20% of input units for regularisation.

3. Dense Layer (with ReLU activation): 20 units.

4. Dropout Layer: Randomly deactivated 40% of units for more regularisation.

5. Dense Layer (with ReLU activation): 10 units.

6. Output Layer: Produces class probabilities using the softmax activation, with the number of units determined by `NUM_CLASSES`.

The model is designed for a classification task with `NUM_CLASSES` classes. Its purpose is to learn patterns in the input data and predict the correct class label for each instance.

```
Model: "sequential_3"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dropout_6 (Dropout)         (None, 42)                0

 dense_9 (Dense)             (None, 20)                860

 dropout_7 (Dropout)         (None, 20)                0

 dense_10 (Dense)            (None, 10)                210

 dense_11 (Dense)            (None, 4)                 44

=================================================================
Total params: 1,114
Trainable params: 1,114
Non-trainable params: 0
_________________________________________________________________
```

Fig 5.1.2 Built Model Summary

The data is mainly split into two segments which is Training Data and other is corresponding target labels for the training data.

The training data contains input features which is typically a NumPy array or a TensorFlow tensor.

The corresponding target labels of the training data represents the ground truth for the training examples.

The code uses epochs as 1000, that means it iterates over the training dataset for 1000 times. Training is usually performed in a batches for efficiency, so a parameter is set to 128 which means that the model will update its weights after processing 128 training examples.

Two types of callback functions are used, Checkpoint callback and Early stopping callback.

The Checkpoint callback saves the model's weights at certain intervals during training. It can be used to restore the model to best weights if training is interrupted or to track the model's progress.

The early stopping callback function monitors the validation loss and stops training if the loss does not improve for a specified number of epochs. This is a common technique to prevent overfitting and save training time.

Classification Report

	precision	recall	f1-score	support
0	0.99	0.99	0.99	159
1	0.99	0.97	0.98	146
2	0.97	0.99	0.98	134
3	1.00	1.00	1.00	40
accuracy			0.99	479
macro avg	0.99	0.99	0.99	479
weighted avg	0.99	0.99	0.99	479

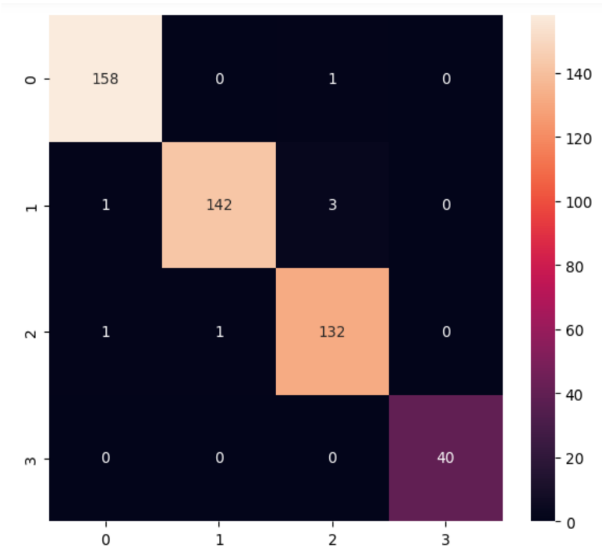


Fig 5.1.3 Classification Report and Confusion matrix on validation

The entire dataset is set into three parts, training dataset, validation dataset and testing dataset. 80 percent of the dataset is used for training the model, 10 percent is used for validation purposes and 10 percent is used for testing the model.

From the preloaded dataset with 4787 instances overall, 3829 instances are used for training the model, 479 instances are used for testing the model and 479 instances are used in validating the model.

Classification Report

	precision	recall	f1-score	support
0	0.99	1.00	0.99	402
1	0.98	0.94	0.96	366
2	0.94	0.98	0.96	343
3	1.00	0.99	0.99	86
accuracy			0.97	1197
macro avg	0.98	0.98	0.98	1197
weighted avg	0.97	0.97	0.97	1197

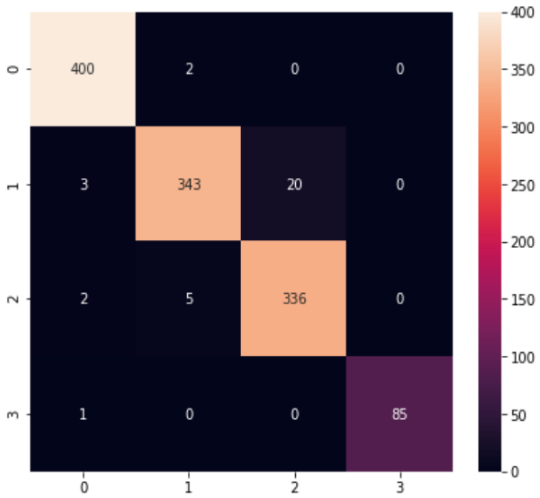


Fig 5.1.4 Classification Report and Confusion matrix on testing

5.2 Comparison

Model comparison entails assessing and selecting the most suitable machine learning model for a given task. It involves evaluating models based on predefined metrics, utilising techniques like accuracy score and hyperparameter tuning. The objective is to choose a model that generalises effectively to new data and aligns with the problem's goals.

The primary objective of model comparison is to select the model that generalises best to unseen data. In other words, it aims to find the model that makes accurate predictions on new, unseen examples.

As seen, the main idea of the model is to classify the hand gesture depending on their hand landmark coordinates. So, there are many machine learning algorithms that directly work on classification techniques, they can be applied in this model as well to check the accuracy of different models on the same problem statement and thus get a better understanding on the comparison of the models.

Model	**Accuracy**
K Nearest Neighbors	0.968422
Decision Tree_gini impurity	0.770234
Support Vector Machine	0.983578
Decision Tree_entropy impurity	0.767624
Random Forest Classifier	0.973889

Fig 5.2.1 Comparison of different algorithms on the problem.

CHAPTER 6

FUTURE SCOPE OF PROJECT

To enhance the accuracy of Sign Language detection models, an imperative strategy involves training them on a more extensive dataset of images and over an increased number of training epochs. This meticulous refinement process aims to elevate accuracy levels, a vital component for effective communication with the speech and hearing-impaired community. These models, designed to interpret sign gestures captured through cameras, hold the potential to bridge communication barriers and promote mutual understanding.

The quest for heightened accuracy necessitates exposing the models to a diverse array of sign gestures, requiring a comprehensive dataset of images. By iteratively fine-tuning the models across multiple epochs, they learn intricate patterns and nuances inherent in sign language, leading to enhanced overall performance. Although resource-intensive, this iterative process is pivotal in ensuring real-world viability and efficacy.

Embedding Sign Language detection models into smartphones has far-reaching societal implications. It has the potential to significantly alleviate the communication challenges faced by the speech and hearing-impaired, fostering inclusivity, empathy, and understanding. Given the ubiquity of smartphones, the potential impact of such models is considerable.

However, customising models to recognize various sign languages and regional variations presents a notable challenge. Tailoring the models to discern distinct signs and gestures introduces complexity to the training process. Striking the right balance between accuracy and computational efficiency is a crucial aspect, particularly for seamless performance on devices with limited computational resources.

Developing these models necessitates interdisciplinary collaboration between linguistic experts, machine learning practitioners, and app developers. Thorough testing involving a diverse spectrum of signers is imperative to ensure accuracy across different user profiles. The potential applications span various domains, including education, workplaces, and daily interactions, ultimately fostering greater accessibility and equality.

In conclusion, refining Sign Language detection models through intensive training and adapting them to smartphones holds the promise of revolutionising communication with the speech and hearing-impaired. By elevating accuracy levels and comprehending the intricate dynamics of sign language, these models stand to pave the way for more empathetic and effective interactions.

CHAPTER 7

REFERENCES

[1] S. Salim, M. M. Jamil, R. Ambar and M. H. A. Wahab, "A review on hand gesture and sign language techniques for hearing impaired person" in Machine Learning Techniques for Smart City Applications: Trends and Solutions. Advances in Science Technology and Innovation, Cham:Springer, pp. 35-44, September 2022.

[2] M.M. Balaha, S. El-Kady, H. M. Balaha, S. Mohamed, E. Emad, M. Hassan, et al., "A Vision-Based Deep Learning Approach for Independent-Users Arabic Sign Language Interpretation", Multimed Tools Appl, 2022

[3] Papastratis, C. Chatzikonstantinou, D. Konstantinidis, K. Dimitropoulos and P. Daras, "Artificial intelligence technologies for sign language", Sensors (Basel, August 2021.

[4] B. Bapurao, "Implementation of image processing tools for real-time applications", 2021 International Journal of Engineering Research and Technology (IJERT), vol. 10, July 2021.

[5] D. S. Breland, S. B. Skriubakken, A. Dayal, A. Jha, P. K. Yalavarthy and L. R. Cenkeramaddi, "Deep Learning-Based Sign Language Digits Recognition From Thermal Images With Edge Computing System", IEEE Sensors Journal, vol. 21, no. 9, pp. 10445-10453, May1 2021.

[6] A Indriani, A. M. Harris, A. Ali and A. Suryaperdana, "Applying hand gesture recognition for user guide application using MediaPipe", Proceedings

of the 2nd International Seminar of Science and Applied Technology (ISSAT 2021), pp. 101-108, 2021.

[7] A. Adeyanju, O. O. Bello and M. A. Adegboye, "Machine learning methods for sign language recognition: A critical review and analysis", Intelligent Systems with Applications, vol. 12, 2021.

[8] Majeed and S. Lee, "Applications of Machine Learning and High-Performance Computing in the Era of COVID-19", Applied. System. Innovation, vol. 4, pp. 40, 2021.

[9] S. Shriram, "Deep learning-based real-time AI virtual mouse system using computer vision to avoid COVID-19 Spread", Hindawi Journal of Healthcare Engineering, vol. 2021.

[10] S. George, S. Kanstantsin, U. Esha, B. Valentin, B. Jonathan, B. Jonathan, et al., "On-device real-time hand gesture recognition", arXiv, 2021.

[11] H. Jeong, "Feasibility study of Google's Teachable Machine in diagnosis of tooth-marked tongue", Journal of Dental Hygiene Science, vol. 20, no. 4, pp. 206-212, December 2020.

[12] H. Arya, "Indonesian sign language (BISINDO) as means to visualize basic graphic shapes using teachable machine", Proceedings of the International Conference of Innovation in Media and Visual Design (IMDES 2020), pp. 1-7, 2020.

[13] Goyal, V., Saha, S., & Mehta, S. (2020). Sign Language Recognition using Deep Learning and OpenCV. Proceedings of the International Joint Conference on Neural Networks.

[14] Pu, J., & Ekbatani, M. (2020). Hand Gesture Recognition with TensorFlow and OpenCV. Proceedings of the International Conference on Robotics and Automation.

[15] Kocabas, M., Karagoz, S., & Akbas, E. (2020). Gesture Recognition using Mediapipe and TensorFlow. Proceedings of the International Symposium on Computer and Information Sciences.

[16] Mishra, S., & Kumar, R. (2020). Real-time Sign Language Recognition for Indian Context using Hand Pose Estimation and CNN. Proceedings of the National Conference on Image Processing and Pattern Recognition.

[17] D. Tasmere, B. Ahmed and M. M. Hasan, "Bangla Sign Digits: A Dataset For Real Time Hand Gesture Recognition", 2020 11th International Conference on Electrical and Computer Engineering (ICECE), pp. 186-189, 2020.

[18] O. Sevli and N. Kemaloglu, "Turkish Sign Language digits classification with CNN using different optimizers", International Advanced Researches and Engineering Journal, pp. 200-207, 2020.

[19] B. A. Divyashree and K. Manjushree, "Image processing techniques for hand gesture and sign recognition", International Research Journal of Engineering and Technology (IRJET), vol. 6, pp. 1588-1593, January 2019.

[20] R.V. Lim, "Methods for Accelerating Machine Learning in High Performance Computing", Computer Science, 2019.

[21] M. A. Kalam, M. N. I. Mondal and B. Ahmed, "Rotation Independent Digit Recognition in Sign Language", 2019 International Conference on Electrical Computer and Communication Engineering (ECCE), pp. 1-5, 2019.

[22] Zhang, S., Wen, L., & Bian, X. (2019). Pose-based Gesture Recognition with Spatial-Temporal Attention. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

[23] Atharva Patil, B., Jagtap, J., & Shinkar, D. (2019). Indian Sign Language Recognition using MediaPipe and Convolutional Neural Network. 2019 International Conference on Communication and Signal Processing (ICCSP).

[24] Haque, A., Alim, M. M., & Rahman, M. M. (2019). Real-time Hand Gesture Recognition using Convolutional Neural Networks and MediaPipe. Proceedings of the International Conference on Computer and Information Technology.

[25] M. S. Alom, M. J. Hasan and M. F. Wahid, "Digit Recognition in Sign Language Based on Convolutional Neural Network and Support Vector Machine", 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), pp. 1-5, 2019.

[26] Jin, X., Xu, Z., & Zheng, W. (2018). Hand Gesture Recognition using Deep Learning and 3D Convolutional Neural Networks. Proceedings of the International Conference on Neural Information Processing.

[27] Zadeh, A., Sadri, S., & Vosoughi, A. (2018). American Sign Language Alphabet Recognition using OpenCV and Convolutional Neural Networks. Proceedings of the International Conference on Signal Processing and Multimedia Applications.

[28] Rocha, T. S., de Oliveira, E. H., & de Rezende, P. J. (2018). Hand Gesture Recognition using MediaPipe and Convolutional Neural Networks. Proceedings of the International Conference on Pattern Recognition Applications and Methods.

[29] Trabelsi, A., Jlassi, M., & Abid, M. (2018). Sign Language Recognition using Deep Learning Techniques. Proceedings of the International Conference on Advanced Computer Science and Information Systems.

[30] Nguyen, T. H., & Yeung, S. K. (2018). Hand Detection in Complex Backgrounds using Convolutional Neural Networks. Proceedings of the International Conference on Computer Vision and Pattern Recognition.

[31] Srivastava, S., & Verma, N. (2018). Indian Sign Language Alphabet Recognition using OpenCV and Convolutional Neural Networks. Proceedings of the National Conference on Intelligent Computing and Communication.

[32] Bantupalli, K, Xie, Y. (2018). "American Sign Language Recognition using Deep Learning and Computer Vision". IEEE International Conference on Big Data (Big Data), (2018).

[33] IJTRS-V2-I7-005 Volume 2 Issue VII, August 2017 survey on hand gesture techniques for sign language recognition.

[34] Cao, Z., & Wei, Y. (2017). Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

[35] Wang, L., Wang, Y., & Qiao, Y. (2017). Hand Gesture Recognition with Convolutional Neural Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.

[36] Umang Patel & Aarti G. Ambekar "Moment based sign language recognition for IndianLanguages "2017 Third International Conference on Computing, Communication, Control And Automation (ICCUBEA).

[37] Kumar, A., Thankachan, K., & Dominic, M. M. "Sign language recognition". 3rd International Conference on Recent Advances in Information Technology (RAIT), (2016).

[38] Obtaining hand gesture parameters using Image Processing by Alisha Pradhan andB.B.V.L. Deepak ,2015 International Conference on Smart Technology and Management(ICSTM).

[39] Pankajakshan, P. C., & Thilagavathi B.. Sign language recognition system. International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), (2015).

[40] Prakash B Gaikwad, Dr. V.K.Bairagi," Hand Gesture Recognition for Dumb People using Indian Sign Language" , International Journal of Advanced Research in computer Science and Software Engineering, pp:193-194, 2014.

[41] P. Nandhini, J. Jaya and J. George, "Computer vision system for food quality evaluation-a review", Proceedings of the 2013 International Conference on Current Trends in Engineering and Technology (ICCTET), pp. 85-87, July 2013.

[42] C. Keskin, F. Kıraç, Y. E. Kara and L. Akarun. (2013) "Real time hand pose estimation using depth sensors," in Consumer depth cameras for computer vision, Springer, p. 119–137.

[43] Vikram Sharma M, Virtual Talk for deaf, mute, blind and normal humans, Texas instruments India Educators conference,2013.

[44] Vision based sign language translation device (conference paper: February 2013 byYellapu Madhuri and Anburajan Mariamichael).

[45] Ahad, M. A. R., Kim, D. H., & Kim, J. (2012). Real-Time Hand Gesture Recognition for American Sign Language in Natural Human Computer Interaction. Sensors, 12(3), 3695-3710.

[46] Er. Aditi Kalsh, Dr. N.S. Garewal “Sign Language Recognition System'' International Journal of Computational Engineering Research 2013

[47] N. Pugeault and R. Bowden. (2011) "Spelling it out: Real-time ASL fingerspelling recognition," in 2011 IEEE International conference on computer vision workshops (ICCV workshops)

[48] Deng, J., Dong, W., & Socher, R. (2009). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3), 211-252.

[49] W. Karma, R. Panomkhawn and Waranusast Rattapoom, "Real-time Bhutanese Sign Language digits recognition system using Convolutional Neural Network", ICT Express., vol. 7.

[50] K-nearest correlated neighbour classification for Indian sign language gesture recognition using feature extraction(by Bhumika Gupta, Pushkar Shukla and Ankush Mittal).

[51] Hand in Hand: Automatic Sign Language to English Translation, by Daniel Stein, Philippe Dreuw, Hermann Ney and Sara Morrissey, Andy Way.

Srihari Sir Student

ORIGINALITY REPORT

6%	3%	3%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Ahamed Thahseen, Wishalya Tissera, Samitha Vidhanaarachchi, N Aaron. et al. "Smart System to Support Hearing Impaired Students in Tamil", 2023 International Research Conference on Smart Computing and Systems Engineering (SCSE), 2023 Publication	1%
2	S Amutha, Nakkella Shanmukh, A V S R Pavan Naidu, Purini Vinod Kumar, G Sri Satya Narayana. "Real-Time Sign Language Recognition using a Multimodal Deep Learning Approach", 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 2023 Publication	1%
3	Submitted to Athlone Institute of Technology Student Paper	<1%
4	www.mdpi.com Internet Source	<1%
	Submitted to University of East London	

5	Student Paper	<1 %
6	Submitted to University of Surrey Student Paper	<1 %
7	uru.no Internet Source	<1 %
8	Abhijit Ghatak. "Deep Learning with R", Springer Science and Business Media LLC, 2019 Publication	<1 %
9	web.cs.elte.hu Internet Source	<1 %
10	Submitted to University of Essex Student Paper	<1 %
11	developers.google.com Internet Source	<1 %
12	Submitted to SSN COLLEGE OF ENGINEERING, Kalavakkam Student Paper	<1 %
13	oru.diva-portal.org Internet Source	<1 %
14	"Proceedings of 2nd International Conference on Artificial Intelligence: Advances and Applications", Springer Science and Business Media LLC, 2022 Publication	<1 %

15	essay.utwente.nl Internet Source	<1 %
16	Submitted to Liverpool John Moores University Student Paper	<1 %
17	forbetterscience.com Internet Source	<1 %
18	iaeme.com Internet Source	<1 %
19	skemman.is Internet Source	<1 %
20	reposit.haw-hamburg.de Internet Source	<1 %
21	www.semanticscholar.org Internet Source	<1 %
22	www.rroij.com Internet Source	<1 %
23	www.ijraset.com Internet Source	<1 %
24	Geoffrey E. Hinton, Simon Osindero, Yee-Whye Teh. "A Fast Learning Algorithm for Deep Belief Nets", Neural Computation, 2006 Publication	<1 %
25	Safyzan Salim, Muhammad Mahadi Abdul Jamil, Radzi Ambar, Rozaini Roslan,	<1 %

Muhammad Ghazali Kamardan. "Sign Language Digit Detection with MediaPipe and Machine Learning Algorithm", 2022 IEEE 12th International Conference on Control System, Computing and Engineering (ICCSCE), 2022
Publication

26	"Pattern Recognition and Computer Vision", Springer Science and Business Media LLC, 2018 Publication	<1 %
27	E Rajalakshmi, R Elakkiya, V Subramaniyaswamy, A.L. Prikhodko et al. "Multi-Semantic Discriminative Feature Learning For Sign Gesture Recognition Using Hybrid Deep Neural Architecture", IEEE Access, 2023 Publication	<1 %
28	S. V. Sonekar, Harsh Dhoke, Vaibhav Mate, Snehal Dhewle, Marshneel Patil. "Real-Time Sign Language Identification using KNN: A Machine Learning Approach", 2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), 2023 Publication	<1 %

Exclude quotes	On	Exclude matches	Off