VISVESVARAYA TECHNOLOGICAL UNIVERSITY

JNANA SANGAMA, BELAGAVI-590018, KARNATAKA

A Project Report
on

Deep neural framework for continuous sign language recognition by iterative training

Submitted in partial fulfillment of the requirements for the VII Semester of degree of ***Bachelor of Engineering in Information Science and Engineering*** *of Visvesvaraya Technological University, Belagavi*

by

SUPRITH SATISH(1RN19IS161)
VIJAY PAYYAVULA(1RN19IS174)
VISHAL VEERARAJ SHETTY(1RN19IS178)
YASH S SINDHE(1RN19IS186)

Under the Guidance of
Mrs. Shyla N
Assistant Professor
Department of ISE

Department of Information Science and Engineering

RNS Institute of Technology

Dr. Vishnuvaradhana Road, Rajarajeshwari Nagar post,
Channasandra, Bengaluru-560098

2022-2023

RNS INSTITUTE OF TECHNOLOGY

Dr. Vishnuvaradhana Road, RajarajeshwariNagar post,
Channasandra,Bengaluru - 560098

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

CERTIFICATE

Certified that the project work entitled ***Deep neural framework for continuous sign language recognition by iterative training*** has been successfully completed by **Vishal Veeraraj Shetty (1RN19IS178), Yash S Sindhe (1RN19IS186), Suprith Satish (1RN19IS161) and Vijay Payyavula (1RN19IS174),** bonafide students of **RNS Institute of Technology, Bengaluru** in partial fulfillment of the requirements for the award of degree in **Bachelor of Engineering in Information Science and Engineering** of **Visvesvaraya Technological University, Belgavi** during academic year **2022-2023**. The project report has been approved as it satisfies the academic requirements in respect of project work for the said degree.

Mrs. Shyla N	**Dr. Suresh L**	**Dr. M K Venkatesha**
Project Guide	Professor and HOD	Principal
Assistant Professor	Department of ISE	RNSIT
Department of ISE	RNSIT	

External Viva

Name of the Examiners	**Signature with Date**
1. ______________________	1. ______________________
2. ______________________	2. ______________________

DECLARATION

We, **Suprith Satish(1RN19IS161), Vijay Payyavula(1RN19IS174), Vishal Veeraraj Shetty(1RN19IS178), Yash S Sindhe(1RN19IS186),** students of VII Semester BE, in Information Science and Engineering, RNS Institute of Technology hereby declare that the Project entitled ***Deep neural framework for continuous sign language recognition by iterative training*** has been carried out by us and submitted in partial fulfillment of the requirements for the *VII Semester of degree of* ***Bachelor of Engineering in Information Science and Engineering*** *of Visvesvaraya Technological University, Belgavi* during academic year 2022-2023.

Place : Bangalore

Date : 29-12-2022

SUPRITH SATISH (1RN19IS161)

VIJAY PAYYAVULA(1RN19IS174)

VISHAL VEERARAJ SHETTY(1RN19IS178)

YASH S SINDHE(1RN19IS186)

Abstract

Sign Language plays an indispensable role in the lives of people having speaking and hearing disabilities. Recognition of American Sign Language using Computer Vision is very challenging due to its increasing complexity and high interclass variations. In this paper, convolutional neural networks (CNNs) are used to recognize the Alphabets. We are detecting the sign language using CNN algorithm.

More than 100 different sign languages are utilized across the world for understanding the deaf and speech impaired groups, they are Indian Sign Language, American Sign Language (ASL), Italian Sign Language, and many others. More than thousands and millions of people and about millions of people in India are using Sign Language as their primary mode of communication. American Sign Language is the widely used sign language and it ranks fourth most used language in North America. It is a surprise to know that not only in the USA, ASL is also used in more than 30 other nations, where English is considered as the primary mode of communication. Around a million people in the USA and various parts of the globe use ASL as their primary mode of communication.

ACKNOWLEDGEMENT

At the very onset we would like to place our gratefulness to all those people who helped us in making this project a successful one.

Coming up, this project to be a success was not easy. Apart from the shear effort, the enlightenment of our very experienced teachers also plays a paramount role because it is they who guide us in the right direction.

First of all, we would like to thank Management of RNS Institute of Technology for providing such a healthy environment for the successful completion of project work.

In this regard, we express our sincere gratitude to the the principal Dr. M K Venkatesha, for providing us all the facilities in this college.

We are extremely grateful to our own and beloved Professor and Head of Department of Information science and Engineering, Dr. Suresh L, for having accepted to patronize us in the right direction with all his wisdom.

We place our heartfelt thanks to Ms. Shyla N, Assistant Professor, Department of information science and Engineering for having guided for project and all the staff members of our department for helping us out at all times.

We thank Dr. Prakasha S Project coordinator, Department of Information Science and Engineering for supporting and guiding us all through.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CNN	CONVOLUTION NEURAL NETWORK
FTTC	FAULT-TOLERANT COOPERATIVE CONTROL
HLD	HIGH-LEVEL DESIGN
HOG	HISTOGRAM ORIENTATION GRADIENT
LoRa	LONG RANGE WIRELESS DATA TELEMETRY
ML	MACHINE LEARNING
RGB	RED GREEN BLUE
SMC	SLIDING MODEL CONTROL
UAV	UNMANNED AERIAL VEHICLES
VGG	VISUAL GEOMETRY GROUP
WSN	WIRELESS SENSOR NETWORK

Chapter 1

INTRODUCTION

1.1 Overview

The Sign Language Alphabets are created by using hand and facial gestures in order to express the thoughts of the speaking and hearing disabled population to standard people, who may not purely understand sign language. This develops a vast communication gap between normal people and the people with disabilities. There are certain technological developments to help the speaking and hearing disabled people as it is impossible for them to seek help all the time and assist them in their daily activities as a translator.

A novel approach can clarify sign gestures into decipherable text with the help of technology so that it becomes understandable for the normal public. This approach will surely reduce the communication gap and initiate the communication between impaired and standard people in a more feasible way.

More than 100 different sign languages are utilized across the world for understanding the deaf and speech impaired groups, they are Indian Sign Language, American Sign Language (ASL), Italian Sign Language, and many others. More than thousands and millions of people and about millions of people in India are using Sign Language as their primary mode of communication. American Sign Language is the widely used sign language and it ranks fourth most used language in North America. It is a surprise to know that not only in the USA, ASL is also used in more than 30 other nations, where English is considered as the primary mode of communication. Around a million people in the USA and various parts of the globe use ASL as their primary mode of communication.

ASL is quite a vast and complex language, which is created by using hand, finger actions, and facial gestures to express the thoughts of speech-impaired people. ASL Language is recognized as an appropriate and original language, where it contains more variations like any other language. For example: Italian and German. Sign language is a magnificent piece of interaction art that is favorable to a wide community of speaking and hearing disabled people.

Chapter 2

LITERATURE SURVEY

A literature survey or a literature review in a project report shows the various analyses and research made in the field of interest and the results already published, considering the various parameters of the project and the extent of the project. Literature survey is mainly carried out in order to analyze the background of the current project which helps to find out flaws in the existing system & guides on which unsolved problems we can work out. So, the following topics not only illustrate the background of the project but also uncover the problems and flaws which motivated to propose solutions and work on this project.

A literature survey is a text of a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. Literature reviews use secondary sources, and do not report new or original experimental work. Most often associated with academic-oriented literature, such as a thesis, dissertation or a peer-reviewed journal article, a literature review usually precedes the methodology and results sectional though this is not always the case. Literature reviews are also common in are search proposal or prospectus (the document that is approved before a student formally begins a dissertation or thesis). Its main goals are to situate the current study within the body of literature and to provide context for the particular reader. Literature reviews are a basis for researching nearly every academic field. demic field. A literature survey includes the following:

- Existing theories about the topic which are accepted universally.
- Books written on the topic, both generic and specific.
- Research done in the field usually in the order of oldest to latest.
- Challenges being faced and on-going work, if available

Literature survey describes about the existing work on the given project. It deals with the problem associated with the existing system and also gives user a clear knowledge on how to deal with the existing problems and how to provide solution to the existing problems.

Real-Time Hand Gesture Recognition Using Finger Segmentation(2020):

Hand gesture recognition is very significant for human-computer interaction. In this work, we present a novel real-time method for hand gesture recognition. In our framework, the hand region is extracted from the background with the background subtraction method. Then, the palm and fingers are segmented so as to detect and recognize the fingers. Finally, a rule classifier is applied to predict the labels of hand gestures.

The experiments on the data set of 1300 images show that our method performs well and is highly efficient. Moreover, our method shows better performance than a state-of-art method on another data set of hand gestures.

Dog Breed Classifier for Facial Recognition using Convolutional Neural Networks(2021):

This paper dealt with the breed classification of dogs. To classify dog breed is a challenging part under a deep convolutional neural network. A set of sample images of a breed of dogs and humans are used to classify and learn the features of the breed. The images are converted to a single label of dimension with image processing.

The images of human beings and dogs are considered for breed classification to find the existing percentage of features in humans of dogs and dogs of humans. This research work has used principal component analysis to shorten the most similar features into one group to make an easy study of the features into the deep neural networks. And, the facial features are stored in a vector form.

This prepared vector will be compared with each feature of the dog into the database and will give the most efficient result. In the proposed experiment, 13233 human images and 8351 dog images are taken into consideration. The images under test are classified as a breed with the minimum weight between test and train images. This paper is based on research work that classifies different dog breeds using CNN. If the image of a dog is supplied then the algorithm will work for finding the breed of dog and features similarity in the breed and if the human image is supplied it determines the facial features existing in a dog of human and vice-versa.

User-Independent American Sign Language Alphabet Recognition Based on Depth Image and PCANet Features(2019) :

Sign language is the most natural and effective way for communication among deaf and normal people. American Sign Language (ASL) alphabet recognition (i.e. fingerspelling) using marker-less vision sensors is a challenging task due to the difficulties in hand segmentation and appearance variations among signers.

Existing color-based sign language recognition systems suffer from many challenges such as complex background, hand segmentation, large inter-class and intra-class variations. In this paper, we propose a new user independent recognition system for the American sign language alphabet using depth images captured from the low-cost Microsoft Kinect depth sensor. Exploiting depth information instead of color images overcomes many problems due to their robustness against illumination and background variations.

Hand region can be segmented by applying a simple preprocessing algorithm over depth image. Feature learning using convolutional neural network architectures is applied instead of the classical hand-crafted feature extraction methods. Local features extracted from the segmented hand are effectively learned using a simple unsupervised Principal Component Analysis Network (PCANet) deep learning architecture.

Two strategies of learning the PCANet model are proposed, namely to train a single PCANet model from samples of all users and to train a separate PCANet model for each user, respectively. The extracted features are then recognized using linear Support Vector Machine (SVM) classifiers. The performance of the proposed method is evaluated using a public dataset of real depth images captured from various users. Experimental results show that the performance of the proposed method outperforms state-of-the-art recognition accuracy using leave-one-out evaluation strategy.

User-Independent American Sign Language Alphabet Recognition Based on Depth Image and PCANet Features(2017) :

Sign language is the most natural and effective way for communication among deaf and normal people. American Sign Language (ASL) alphabet recognition (i.e. fingerspelling) using marker-less vision sensors is a challenging task due to the difficulties in hand segmentation and appearance variations among signers. Existing color-based sign language recognition systems suffer from many challenges such as complex background, hand segmentation, large inter-class and intra-class variations. In this paper, we propose a new user independent recognition system for the American sign language alphabet using depth images captured from the low-cost Microsoft Kinect depth sensor. Exploiting depth information instead of color images overcomes many problems due to their robustness against illumination and background variations.

Hand region can be segmented by applying a simple preprocessing algorithm over depth image. Feature learning using convolutional neural network architectures is applied instead of the classical hand-crafted feature extraction methods. Local features extracted from the segmented hand are effectively learned using a simple unsupervised Principal Component Analysis Network (PCANet) deep learning architecture. Two strategies of learning the PCANet model are proposed, namely to train a single PCANet model from samples of all users and to train a separate PCANet model for each user, respectively.

The extracted features are then recognized using linear Support Vector Machine (SVM) classifiers. The performance of the proposed method is evaluated using a public dataset of real depth images captured from various users. Experimental results show that the performance of the proposed method outperforms state-of-the-art recognition accuracy using leave-one-out evaluation strategy

Content-based Retrieval and Real Time Detection from Video Sequences Acquired by Surveillance Systems(2020) :

In this paper, a surveillance system devoted to detect abandoned objects in unattended environments is presented to which image processing content based retrieval capabilities have been added for making inspection tasks from operators.

Video-based surveillance systems generally employ one or more cameras connected to a set of monitors. This kind of system needs the presence of a human operator, who interprets the acquired information and controls the evolution of the events in a surveyed environment. During the last years efforts have been performed to develop systems supporting human operators in their surveillance task, in order to focus the attention of operators when unusual situations are detected.

Image sequence databases are also managed by the proposed surveillance system in order to provide operators with the possibility of retrieving in a second time the interesting sequences that may contain useful information for discovering causes of an alarm. Experimental results are shown in terms of the probability of correct detection of abandoned objects and examples about the retrieval sequences.

Robust Real-Time Periodic Motion Detection(2017) :

We describe new techniques to detect and analyze periodic motion as seen from both a static and a moving camera. By tracking objects of interest, we compute an object's self-similarity as it evolves in time. For periodic motion, the self-similarity measure is also periodic and we apply TimeFrequency analysis to detect and characterize the periodic motion. The periodicity is also analyzed robustly using the 2D lattice structures inherent in similarity matrices. A real-time system has been implemented to track and classify objects using periodicity. Examples of object classification (people, running dogs, and vehicles), person counting, and nonstationary periodicity are provided.

A learning-based prediction-and-verification segmentation scheme for hand sign image sequence(2019) :

We present a prediction-and-verification segmentation scheme using attention images from multiple fixations. A major advantage of this scheme is that it can handle a large number of different deformable objects presented in complex backgrounds. The scheme is also relatively efficient.

The system was tested to segment hands in sequences of intensity images, where each sequence represents a hand sign in American Sign Language. The experimental result showed a 95 percent correct segmentation rate with a 3 percent false rejection rate.

Design of Sign Language Recognition Using E-CNN(2021) :

Sign language reduces the barrier for communicating with the humans having impaired of speech and hearing, on the other hand Sign language cannot be easily understood by common people. Therefore, a platform is necessary that is built using an algorithm to recognize various signs; it is called Sign Language Recognition (SLR).

It is a technique that simplifies the communication between speech and hearing impaired people with normal people, the main aim of SLR is to overcome the aforementioned drawback. In this manuscript it is aimed to review various techniques that have been employed in the recent past for SLR that are employed at various stages of recognition.

Adding to the above, various images based with or without the glove employed for detection, their advantages and difficulties encountered during this process. Also, segmentation, feature extraction, methods used for feature vector quantization and reduction techniques are discussed in detail. Techniques like CNN are used as a classification tool.

Chapter 3

PROBLEM IDENTIFICATION

Sign language is the most natural and effective way for communication among deaf and normal people. American Sign Language (ASL) alphabet recognition (i.e. fingerspelling) using marker-less vision sensors is a challenging task due to the difficulties in hand segmentation and appearance variations among signers.

Existing color-based sign language recognition systems suffer from many challenges such as complex background, hand segmentation, large inter-class and intra-class variations, and Principal Component Analysis Network (PCANet) deep learning architecture. Two strategies of learning the PCANet model are proposed on this detection we can use.

The existing systems are time-consuming and can be detrimental in real time scenarios. Hence we overcome this shortcoming/limitation by integrating Machine Learning technology to speed up the detection and classification process.

Using CNN, we achieve:

1. High Accuracy
2. Time Saving
3. Faster Process

Chapter 4

OBJECTIVES

The features that we are implementing in this project are as follows:

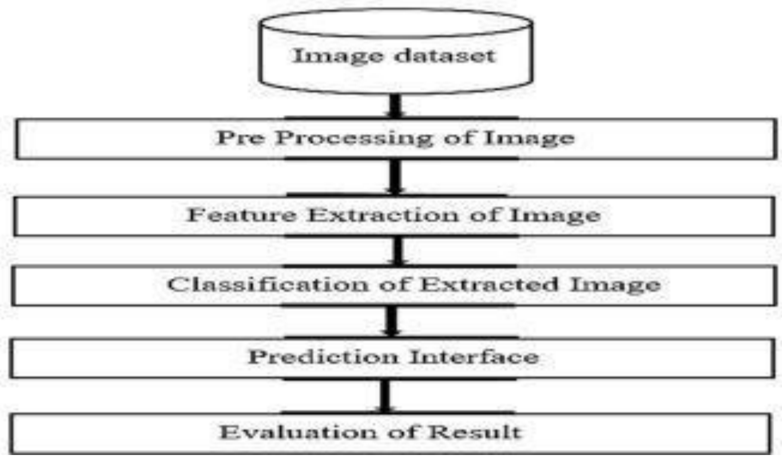
- In this project, convolutional neural networks (CNNs) are used to recognize the English Alphabets. We are detecting sign language using CNN algorithm. Hence we are using Machine Learning to detect sign language.
- Sign language is the most natural and effective way for communication among deaf and normal people. American Sign Language (ASL) alphabet recognition (i.e. fingerspelling) using marker-less vision sensors is a challenging task due to the difficulties in hand segmentation and appearance variations among signers.
- Existing color-based sign language recognition systems suffer from many challenges such as complex background, hand segmentation, large inter-class and intra-class variations.
- We are developing a ML project to predict the sign language.We are detecting the sign language. Using CNN algorithm we are predicting the sign language.

Chapter 5

METHODOLOGY

5.1 System Implementation

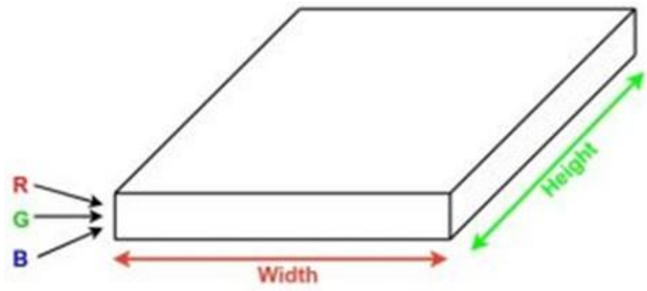
Structural Design



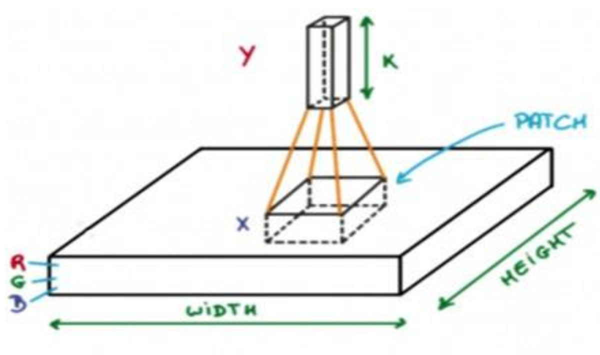
Convolution Neural Network:

Convolutional neural network is the special type of feed forward artificial neural network in which the connectivity between the layers is inspired by the visual cortex. Convolutional Neural Network (CNN) is a class of deep neural networks which is applied for analyzing visual imagery. They have applications in image and video recognition, image classification, natural language processing etc.

Convolutional Neural Networks or covnets are neural networks that share their parameters. Imagine you have an image. It can be represented as a cuboid having its length, width (dimension of the image) and height (as images generally have red, green, and blue channels).



Now imagine taking a small patch of this image and running a small neural network on it, with say, k outputs and represent them vertically. Now slide that neural network across the whole image, as a result, we will get another image with different width, height, and depth. Instead of just R, G and B channels now we have more channels but lesser width and height. His operation is called Convolution. If patch size is same as that of the image it will be a regular neural network. Because of this small patch, we have fewer weights.

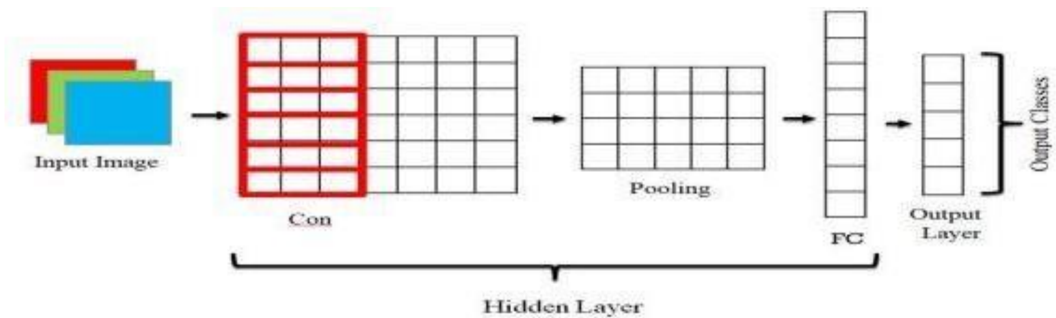


Now let's talk about a bit of mathematics which is involved in the whole convolution process.

- Convolution layers consist of a set of learnable filters (patch in the above image). Every filter has small width and height and the same depth as that of input volume (3 if the input layer is image input)
- For example, if we have to run convolution on an image with dimension 34x34x3. Possible size of filters can be axax3, where 'a' can be 3, 5, 7, etc but small as compared to image dimension.
- During forward pass, we slide each filter across the whole input volume step by step where each step is called stride (which can have value 2 or 3 or even 4 for high

dimensional images) and compute the dot product between the weights of filters and patch from input volume.

- As we slide our filters we'll get a 2-D output for each filter and we'll stack them together and as a result, we'll get output volume having a depth equal to the number of filters. The network will learn all the filters.



Layers used to build ConvNets

A covnets is a sequence of layers, and every layer transforms one volume to another through a differentiable function.

Types of layers:

Let's take an example by running a covnets on an image of dimension 32 x 32 x 3.

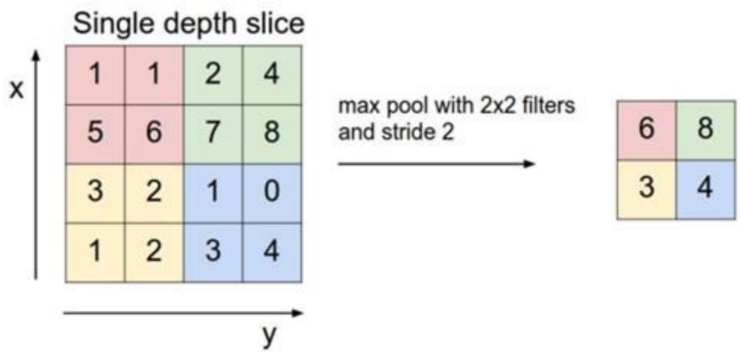
1. **Input Layer:** This layer holds the raw input of an image with width 32, height 32 and depth 3.

2. **Convolution Layer:** This layer computes the output volume by computing dot product between all filters and image patch. Suppose we use a total 12 filters for this layer we'll get output volume of dimension 32 x 32 x 12.

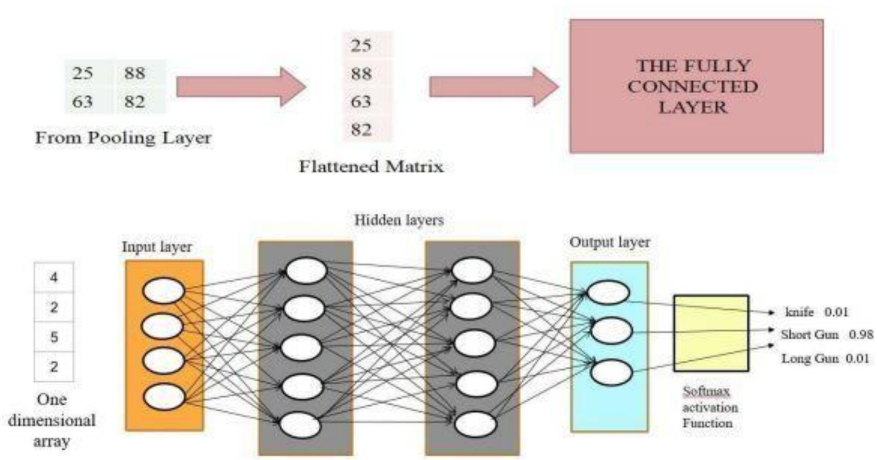
3. **Activation Function Layer:** This layer will apply element wise activation function to the output of the convolution layer. Some common activation functions are RELU: max(0,

x), Sigmoid: 1/(1+e^-x), Tanh, Leaky RELU, etc. The volume remains unchanged hence output volume will have dimension 32 x 32 x 12.

4. **Pool Layer:** This layer is periodically inserted in the covnets and its main function is to reduce the size of volume which makes the computation fast, reduces memory and also prevents overfitting. Two common types of pooling layers are **max pooling** and **average pooling**. If we use a max pool with 2 x 2 filters and stride 2, the resultant volume will be of dimension 16x16x12.



5. **Fully-Connected Layer:** This layer is a regular neural network layer which takes input from the previous layer and computes the class scores and outputs the 1-D array of size equal to the number of classes.



Chapter 6

SYSTEM DESIGN

6.1 System requirements

System requirement specifications gathered by extracting the appropriate information to implement the system. It is the elaborative conditions which the system needs to attain. Moreover, the SRS delivers a complete knowledge of the system to understand what this project is going to achieve without any constraints on how to achieve this goal. This SRS does not provide the information to outside characters but it hides the plan and gives little implementation details.

6.1.1 Specific Requirement

- Require access to a client session of Python and Keras toolbox for job submission.
- A shared file system between user desktops and clusters.
- Maximum of Python workers per physical CPU core.

6.1.2 Hardware Requirement

- Processor: Intel core
- Processor Speed: 1.86 GHz.
- RAM: 4GB+
- Hard Disk Space: 500 GB+
- Monitor: 15 VGA Color

6.1.3 Software Requirement

- Operating system: Windows 10
- Coding Language: Python
- Software Tool: Keras

6.2 Functional and Non-functional Requirements

6.2.1 Functional Requirements:

- System should do minimal computations on its own.
- System should capture images.
- System should automatically detect fire on its own.
- System should automatically be intimate.

6.2.2 Non-Functional Requirements:

- The Camera is used take Video
- Requirement data will be stored in the python database
- System should be reliable
- System should be flexible for future enhancements
- System should be Easily Implementable
- System should be Easy to Implement.
- Cost of Implementation should be low.

6.3 System Design

- The proposed system includes five modules. The initial stage is the image acquisition stage through which the real world sample is recorded in its digital form using a digital camera.
- In the next stage of the research, the image was subjected to a pre-processing stage. Making use of it the size and complexity of the image is reduced.
- The precise digital information is subjected to a segmentation process which separates the rotten portion of the Animal samples.
- The feature extraction aspect of an image analysis focuses on identifying inherent features of the objects present within an image,
- Classification maps the data into specific groups or classes.

High-level design (HLD) explains the architecture that would be used for developing a software product. The architecture diagram provides an overview of an entire system, identifying the main components that would be developed for the product and their interfaces. The HLD uses possibly nontechnical to mildly technical terms that should be understandable to the administrators of the system. In contrast low level design further exposes the logical detailed design of each of these elements for programmers.

High level design is the design which is used to design the software related requirements. In this chapter complete system design is generated and shows how the modules, sub modules and the flow of the data between them are done and are integrated. It consists of very simple phases and shows the implementation process.

Design Consideration:

- **Input:**

  The input to the system is a video or sequence of images of a person making a sign.

- **Preprocessing:**

  The input video or images are preprocessed to extract relevant features, such as hand shape, position, and motion. This may involve applying image processing techniques, such as cropping, resizing, and normalizing the images.. Images are taken as the input and output for image processing techniques.

  It is the analysis of image to image transformation which is used for the enhancement of image. Firstly, we convert the RGB image to a gray scale image. It helps to reduce the complexity in the image and also make the work easy. Then by using min-max scalar method converts the gray scale values into binary values.

  The obtained binary values are taken as the input for the further process. In the obtained binary matrix consider one value region as white and zero value region black. By using these values, the region of interest can be identified. So that the values are useful for feature extraction and identification of regions of interest.

- **Feature extraction:**

  The extracted features are then used to extract more meaningful and relevant features that can be used as input to the machine learning model. This may involve applying dimensionality reduction techniques, such as principal component analysis (PCA) or linear discriminant analysis (LDA). In this stage extract the required feature from the identified region which is obtained from the previous step. That region is compressed by converting a reduced size matrix to control over fitting. The reduction of the matrix size helps in reducing the memory size of the images. Then the flattening process is applied to the reduced matrix, in which the reduced matrix is converted to a one-dimension array, which is used for final detection.

- **Machine learning model:**

  The extracted features are fed into a machine learning model, such as a support vector machine (SVM), decision tree, or neural network. The model is trained on a labeled dataset of sign language gestures, allowing it to learn and recognize different signs.

- **Output:**

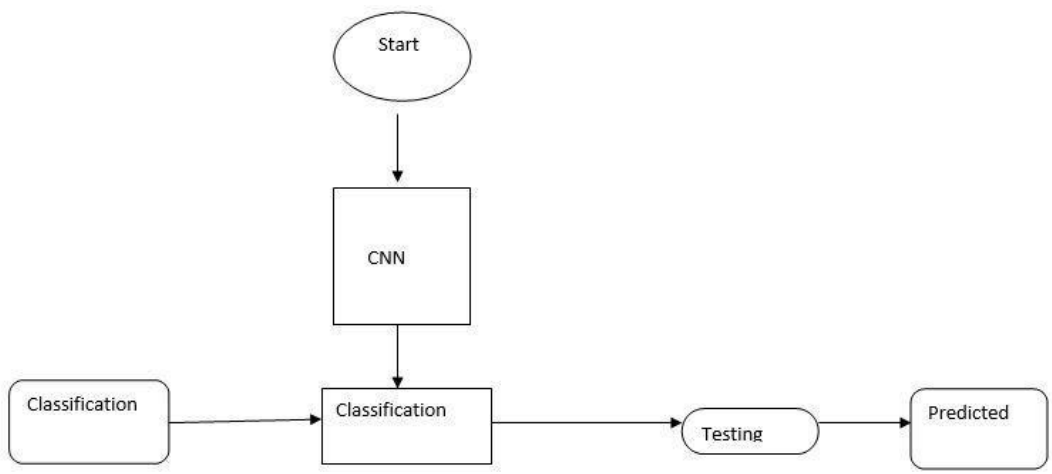
  The machine learning model outputs a prediction of the sign being made in the input video or images. This prediction can be used to generate text or audio output, allowing individuals who do not understand sign language to access the information being communicated.

- **Evaluation:**

  The performance of the system is evaluated using metrics such as accuracy, precision, and recall. The system can be fine-tuned and improved based on these results.

System Architecture:

A system architecture is the conceptual model that defines the structure, behavior, and more views of a system. A system architecture can consist of system components and the sub-systems developed that will work together to implement the overall system.

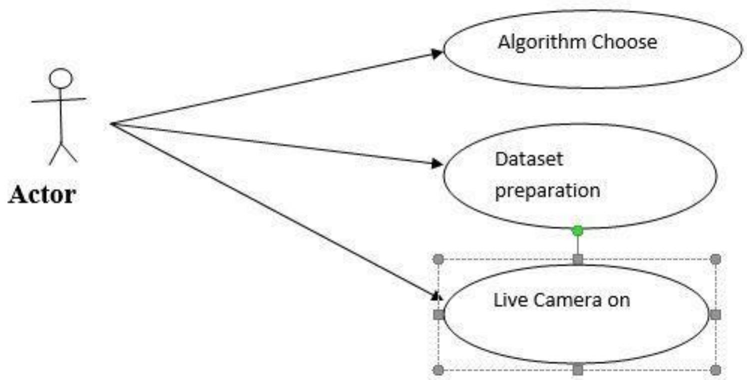


The figure shows the system architecture for the proposed system. The input image is preprocessed and converted to gray scale image to get the clear vision of the image. Then it will be converted into binary values. In the next step identify the part which needs to proceed further. Then required features are extracted by In the CNN convolution layer. By passing those features into different layers of CNN we get a compressed image, that feature is used for detection of sign language.

Specifications using use case diagrams:

Unified Modeling Language (UML) is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created, by the Object Management Group.

A use case diagram at its simplest is a graphical representation of a user's interaction with the system and depicts the specifications of a use case. A use case diagram can portray the different types of users of a system and the various ways that they interact with the system.



6.4 Data flow diagram

A data flow diagram (DFD) is a graphic representation of the "flow" of data through an information system. A data flow diagram can also be used for the visualization of data processing (structured design). It is common practice for a designer to draw a context level DFD first which shows the interaction between the system and outside entities.

Data flow diagrams show the flow of data from external entities into the system, how the data moves from one process to another, as well as its logical storage. There are only four symbols:

1. Squares representing external entities, which are sources and destinations of information entering and leaving the system.

2. Rounded rectangles representing processes, in other methodologies, may be called 'Activities', 'Actions', 'Procedures', 'Subsystems' etc. which take data as input, do processing to it, and output it.

3. Arrows representing the data flows, which can either be electronic data or physical items. It is impossible for data to flow from data store to data store except via a process, and external entities are not allowed to access data stores directly.

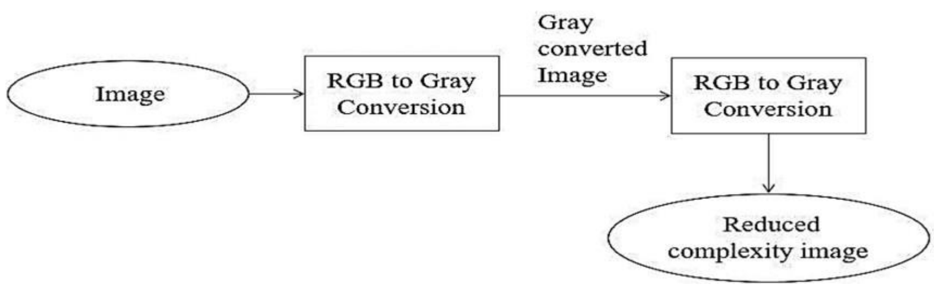
4. The flat three-sided rectangle representing data stores should both receive information for storing and provide it for further processing.

5. It is also used to analyze a particular problem and the solution for it in steps.

6. A user loads the data and the system reads the data provided by the user.

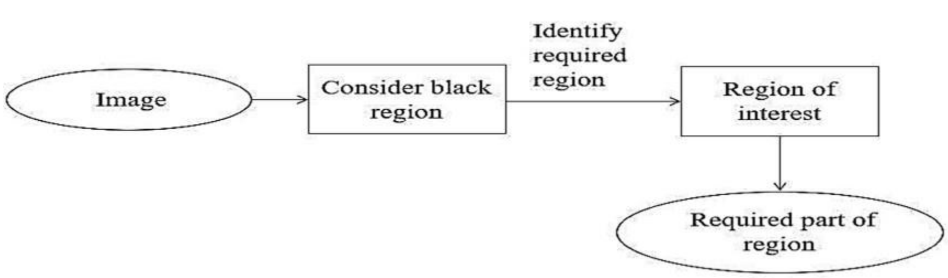
7. Based on feature extraction and classifier the model will be trained and tested.

- **Data Flow Diagram for Pre-processing**



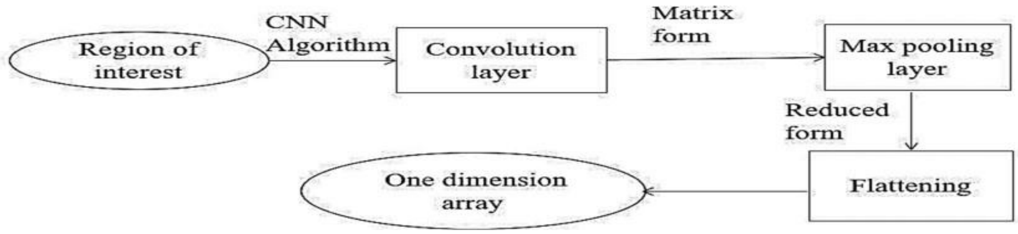
The figure shows that the image is given as input. As we give the color image so that RGB image is converted into gray scale values to reduce complexity in the image. For efficient feature extraction gray scale values are converted into binary values. Then the image with reduced complexity is sent to the next process.

- **Data Flow Diagram for Identification**



The figure shows that the image with reduced complexity is considered as input. Here the region with the value of one is considered as black that region is considered for the next process.

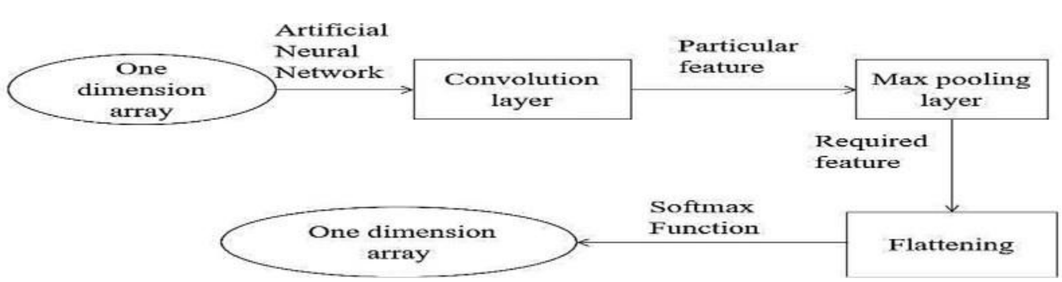
❖ Data Flow Diagram for Feature Extraction



The figure shows that the region of interest from the identification step is considered as input. The region of interest is obtained from converting RGB color image to the gray scale image by using minimax scalar method. For that region CNN algorithm is applied. A CNN consists of an input layer and an output layer, as well as multiple hidden layers between them. The hidden layer basically consists of the convolution layer, pooling layer, relu layer and fully connected layers.

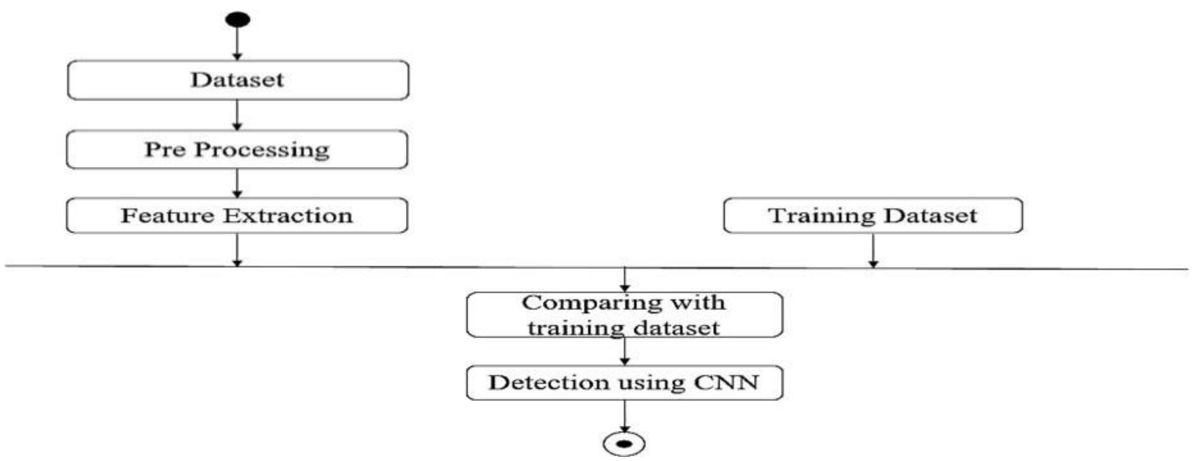
In this the RGB color image is converted into grayscale image by using minimax scalar method. The binary valued image is given as input to the convolution layer. In the convolution layer binary matrix is multiplied with a filter to extract features from the region.

❖ Data Flow Diagram for Classification and Detection



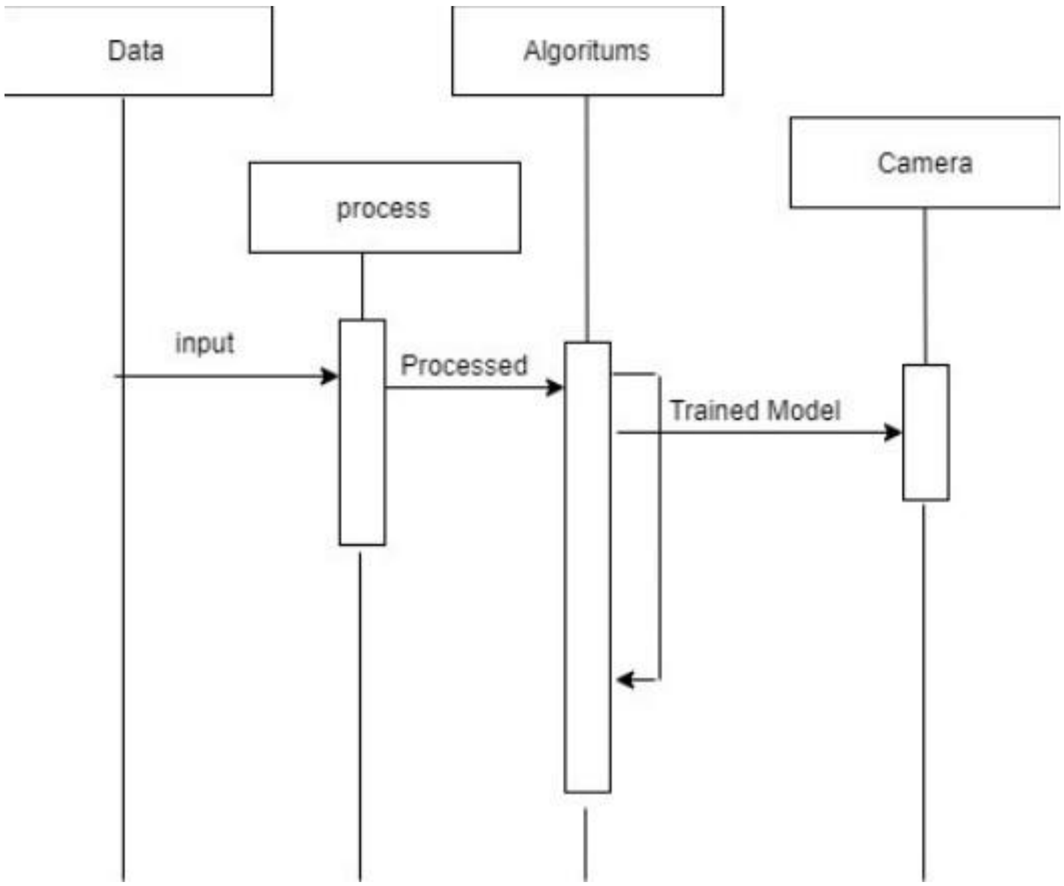
The figure shows that the one-dimension array is sent to a fully connected layer of CNN. Artificial neural network method is applied to this layer. Firstly, a one-dimension array is sent to the input layer. Some particular feature which is required for the detection is identified by the hidden layer of ANN. The continued connection from hidden layer to output layer will help to identify accurate results. By considering all the features, the output layer gives the result with some predictive value. These values are calculated by using SoftMax activation function. SoftMax activation function provides predictive values.

Activity Diagram:



Here the preprocessing of the image by converting the RGB to grayscale image and feature extraction is done by the first layer that is the convolution layer of the neural network and detection done by using fully connected layers of the convolutional neural network.

Sequence diagram



Here the image dataset is given and pre-processing of the image is done. The processed data is given to feature extraction and here comparing and testing of image is done and by applying CNN algorithm the detection is done.

Chapter 7

CONCLUSION

As a result of our literature review, we are able to determine that it is possible for us to build a powerful system to detect sign language with good precision and accuracy. This system can be used in public spaces to allow people with hearing impairment to communicate effectively.

In this project the training data is taken as a huge number of images so that the system learns to classify them depending on which sign is used to symbolize which alphabet.

Here the system processes the image containing hand signs based on RGB to grayscale conversion process.We used CNN technique to develop the sign language predictor, was trained and tested on the model above and it was observed that the predictor produced a better accuracy while using CNN than Non-CNN based techniques.

In conclusion, the use of machine learning techniques in the development of a sign language recognition system has proven to be an effective and efficient approach. The system was able to accurately classify and recognize different signs with a high degree of accuracy, demonstrating its potential for practical application in real-world situations.

The use of machine learning algorithms allowed the system to learn and adapt to different signs over time, improving its overall performance and accuracy. In the future, it is likely that the use of machine learning in sign language recognition will continue to evolve and improve, opening up new opportunities for individuals who use sign language as their primary mode of communication.

Chapter 8

REFERENCES

[1] S. Y. Kim, H. G. Han, J. W. Kim, S. Lee, and T. W. Kim, “A hand gesture recognition sensor using Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021) IEEE Xplore Part Number

[2] Y. Cui and J. Weng, “A learning-based prediction-and-verification segmentation scheme for hand sign image sequence,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 8, pp. 798–804, 1999, doi: 10.1109/34.784311.

[3] A. Kumar and A. Kumar, “Dog Breed Classifier for Facial Recognition using Convolutional Neural Networks,” pp. 508–513, 2020.

[4] K. L. Bouman, G. Abdollahian, M. Boutin, and E. J. Delp, “A low complexity sign detection and text localization method for mobile applications,” IEEE Trans. Multimed., vol. 13, no. 5, pp. 922–934, 2011, doi: 10.1109/TMM.2011.2154317.

[5] J. Wu, L. Sun, and R. Jafari, “A Wearable System for Recognizing American Sign Language in RealTime Using IMU and Surface EMG Sensors,” IEEE J. Biomed. Heal. Informatics, vol. 20, no. 5, pp. 1281–1290, 2016, doi: 10.1109/JBHI.2016.2598302.

[6] C. Zhang, W. Ding, G. Peng, F. Fu, and W. Wang, “Street View Text Recognition With Deep Learning for Urban Scene Understanding in Intelligent Transportation Systems,” IEEE Trans. Intell. Transp. Syst., pp. 1–17, 2020, doi: 10.1109/tits.2020.3017632.

[7] M. Safeel, T. Sukumar, K. S. Shashank, M. D. Arman, R. Shashidhar, and S. B. Puneeth, “Sign Language Recognition Techniques- A Review,” 2020 IEEE Int. Conf. Innov. Technol. INOCON 2020, pp. 1–9, 2020, doi: 10.1109/INOCON50539.2020.9298376.

[8] W. Aly, S. Aly, and S. Almotairi, “Userindependent american sign language alphabet recognition based on depth image and PCANet features,” IEEE Access, vol. 7, pp. 123138–123150, 2019, doi: 10.1109/ACCESS.2019.2938829.

[9] S. Adhikari, S. Thapa, and B. K. Shah, “Oversampling based Classifiers for Categorization of Radar Returns from the Ionosphere,” Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020, no. Icesc, pp. 975–978, 2020, doi: 10.1109/ICESC48915.2020.9155833.

[10] S. Hayani, M. Benaddy, O. El Meslouhi, and M. Kardouchi, “Arab Sign language Recognition with Convolutional Neural Networks,” Proc. 2019 Int. Conf. Comput. Sci. Renew. Energies, ICCSRE 2019, pp. 2019–2022, 2019, doi:

10.1109/ICCSRE.2019.8807586. [11] Y. Li, X. Wang, W. Liu, and B. Feng, “Pose Anchor: A Single-Stage Hand Keypoint Detection Network,” IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 7, pp. 2104–2113, 2020, doi: 10.1109/TCSVT.2019.2912620.

[12] X. Shen and F. L. Chung, “Deep network embedding for graph representation learning in signed networks,” arXiv, vol. 50, no. 4, pp. 1556– 1568, 2019.

[13] Y. Zhu, M. Liao, M. Yang, and W. Liu, “Cascaded Segmentation-Detection Networks for Text-Based Traffic Sign Detection,” IEEE Trans. Intell. Transp. Syst., vol. 19, no. 1, pp. 209–219, 2018, doi: 10.1109/TITS.2017.2768827.

[14] B. K. Shah, V. Kedia, R. Raut, S. Ansari, and A. Shroff, “Evaluation and Comparative Study of Edge Detection Techniques,” vol. 22, no. 5, pp. 6–15, 2020, doi: 10.9790/0661-2205030615.

[15] B. K. Shah, V. Kedia and R. K. Jha, "Integrated Vendor-Managed Time Efficient Application to Production of Inventory Systems," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 275-280, doi: 10.1109/ICICT50816.2021.9358504.