# DHII

## A BOOK OF RESEARCH ON
## COMPUTER SCIENCE AND APPLICATIONS


**Jyoti Nivas College Autonomous**

DEPARTMENT OF COMPUTER SCIENCE (PG)

**Our motto:** Let your light shine

**Our Vision:** Communion, Excellence, Service and Relevance

**Our Mission:** To Turn Out Intellectually Enlightened

# Editorial Board

# DHII

Published by:
Research Publication and Development Cell
Jyoti Nivas College Autonomous, Bengaluru– 560095, India

Edited, owned, printed, and published by Dr Sr Lalitha Thomas from Research Publication and Development Cell, Jyoti Nivas College Autonomous, Bengaluru-560 095, India and printed at National Printing Press, 580, K.R. Garden, Koramangala, Bengaluru-560 095.

ISBN : 978-93-5768-303-6

# Message

**"Sooner or later, those who win are those who think they can." — Paul Tournier**

It gives me immense pleasure to present to you the twenty fourth issue of the Research journal - '*Dhii*' June 2023. It is a journal by the students of MCA that has both internal and external expert reviewer teams. The journal publishes articles on current research in the areas of Information Technology. This issue has thirteen articles. The time spent at the Master's degree must make a difference. In a way it is a stepping-stone to a different life altogether. Since, this is the ideal time for learning and teaching and learning thrives in an ambience wherein new knowledge is created, the course has been planned carefully to enable the students for further academic pursuits.

The purpose of this journal is to encourage students at the Master's degree level to develop an interest in research and the journal is the first step to create research culture, to make research a part of their curriculum, to generate the interest among the young academicians at an early stage. In their final year, the students have the opportunity to work on a research project in an area of their choice, under the guidance of one of the faculty members. This is to offer students an opportunity to engage with research and prepare them for higher levels of academia. Their endeavors at contributing articles, compiling and editing teaches them valuable lessons on research.

While *'Dhii'* hones the research skills in students, it challenges the staff to keep abreast with the latest areas of research which in turn helps to improve the quality of their teaching.

My sincere appreciation goes to our student contributors, who worked with ease, enthusiasm and joy. I thank my dear staff, who journeyed with our students and worked towards successful release of the twenty fourth issue. Congratulations to the staff and students of MCA. In our constant striving for excellence, we continue to march ahead keeping in mind the college motto 'Let Your Light Shine'.

**God Bless You**
**Dr. Sr. Lalitha Thomas**
**Principal, Jyoti Nivas College Autonomous**

# *Names of the Reviewers*

## MCA

### External Reviewer

- Dr. YETHIRAJ N G**,** Assistant Professor and P G Coordinator, Department of Computer Science, Maharani‘s Science college for women, Bangalore

### Internal Reviewer

- Dr**.** IRENE GETZI**,** Associate Professor, Department of MCA, Jyoti Nivas College**,** Bangalore.
- Dr. SHILPA ABHANG, Associate Professor, Department of MCA, Jyoti Nivas College**,** Bangalore.
- Dr. RAJESH DHARMARAJ, Associate Professor, Department of MCA, Jyoti Nivas College**,** Bangalore.
- Dr. SWARNAMUGI M, Assistant Professor, Department of MCA, Jyoti Nivas College**,** Bangalore.
- Dr. SENTHIL VADIVU, Assistant Professor, Department of MCA, Jyoti Nivas College**,** Bangalore.

# DHII
# JOURNAL OF THE MCA DEPARTMENT,
# JYOTI NIVAS COLLEGE AUTONOMOUS

## Volume XXIV June 2023

# *Foreword*

Definiteness of purpose is the starting point of all achievement.

–W. Clement Stone

The purpose of this journal is to foster and reward the scholarly efforts of our students as well as to provide a valuable learning experience. June 2023 of 'DHII' showcases 13 outstanding research papers from the field of computer science focusing the active areas of research namely Digital Image Processing, Internet Of Things, Data Visualization and Machine Learning.

The first paper of Dhii is based on "Object Detection in Color Images".
Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, fruits, or cars) in digital images and videos. Well-researched domains of object detection include face detection and pedestrian detection. Object detection has applications in many areas of computer vision, including image retrieval and video surveillance. It is widely used in computer vision tasks such as face detection, face recognition, video object co-segmentation. It is also used in tracking objects, for example tracking a ball during a football match, tracking movement of a cricket bat, or tracking a person in a video.

The second paper is based on "Credit Card fraud detection using Logistic Regression".
Credit card fraud detection is presently the most frequently occurring problem in the present world. This is due to online transactions and e-commerce platforms. The fraud happens when the card was misused any stolen by the fraudster. In this world, we humans are facing a lot of credit card problems in some or other way. To detect the fraud this system was introduced. This project aims to focus mainly on machine learning algorithms. The algorithms used are Logistic Regression. The results are based on accuracy. The logistic regression model algorithm that gives the more accuracy considered as the best algorithm to detect the fraud.

The third paper is based on "Text Based Plagiarism Checker"
Student assessment is a vital part of the teaching and learning process. The idea of the research project is to propose a plagiarism based system for online paper evaluation. The input response sheet from the students must pass both a syntactic and a semantic similarity check with the evaluation method. Plagiarism in student work is a pervasive and aggravating issue in the academic process. Traditional manual plagiarism detection by humans is a challenging, imprecise, and time-consuming method.

The fourth paper is based on "The comparative study of noise removal techniques in compound images**".**
In the advancing world many new techniques are being discovered to reflect data. Image is used to represent data, compound images are frequently in use.

Noise is present many images captured, this causes visual disturbance and degrades the quality by corrupting the properties of the image. In order to maintain good quality of images, noise should be removed, this helps to regain quality of the image. Noise removal in an image can be done by applying the filters such as Median, Wiener, Gaussian filter, etc. This paper deals with Speckle, Salt & pepper, Gaussian noise and filters such as Mean, Median, Gaussian filters.

The fifth paper is based on "Drowsiness Detection System"
Drowsiness is a state of close sleep due to fatigue which reduces mental alertness and reduces person's ability to function. The Sleep Detection System helps detect if the person in front of the camera is working, drowsy or asleep. This helps to reduce the risk of accidents while driving, to wake up in the classroom whena student is present and to be alert at meetings and presentations when an employee is present. The main goal of the Sleep Scheme is to develop an automated system for diagnosing whether a person is active or sleepy using dlib and face recognition technology. When the subject is drowsy or drowsy it is characterized by marking features of the face and eyelids and sounding the alarm to wake them up.

The sixth paper is based on "Currency recognition system using Image Processing"
All currencies around the world look totally different from each other. For instance the size of the paper is different, the same as the color and pattern. The staffs who work at places like money exchange offices have to distinguish between different types of currencies and that is not an easy job. They have to remember the symbol of each currency. This may result into wrong recognition, so they need an efficient and foolproof system to aid in their work.

The seventh paper is based on "Face Recognition Attendance System".
The Face Recognition System Project is designed to identify a person's face at school / college to mark the presence of their students. The Face Recognition System therefore uses real-time face recognition as a real- world solution that assists with day-to-day student management activities.
The task is extremely difficult as removing a real-time background is still a challenge. Real-time facial recognition is also used, and a quick Analysis of the Main Part is used to detect faces obtained with a high degree of accuracy. Faces are used to mark student presence. This program automatically records student presence and marks student presence. The program is therefore designed to be an effective module that incorporates facial recognition to manage student attendance records.

The eighth paper is based on "Wine quality prediction"
Wine manufacturing is a difficult task since taste is the least predicted of the human senses. A good wine quality can be predicted and is very useful in the certification phase, since currently the taste analysis is performed by human tasters, being clear it will be a subjective approach. An automated predictive system can be integrated into a decision support system, helping the speed and quality of the performance. Furthermore, a feature selection process can help to analyze the impact of the analytical tests. If it is concluded that several input variables are highly relevant to predict the wine quality, since in the production process some variables can be controlled, this information can be used to improve the wine quality. Classification models used here are 1) Random Forest 2) Stochastic Gradient Descent 3) SVC 4) Logistic Regression.

The ninth paper is based on “Face Mask Detection using YOLOV5”
Since the infectious coronavirus disease in the new world (2020), it has become a Public health problem in India and even around the world. This pandemic is having devastating effects on societies and economies around the world. The increase in the number of COVID-19 tests gives more information about the epidemic spread, which may lead to the possibility of surrounding it to prevent further infections. Multidisciplinary efforts have been organized to slow the spread of the pandemic, wearing a face mask that prevents the transmission of droplets in the air and maintaining an appropriate physical distance between people, and reducing close contact with each other can still be beneficial in combating this pandemic. I would like to build a system that can detect faces in images and identify if the detected faces are wearing masks or not. I will firstly explore mask and no mask classification in images classification and also the count the number of faces detected in the image. I have used CNN, YOLOV5 and Tensor-flow to know the accuracy in the mask detection.

The tenth paper is based on “Credit Card Fraud Detection using Random Forest”
Credit card fraud detection is presently the most frequently occurring problem in the present world. This is causes to the rise in both online transactions and e-commerce platforms. Credit card fraud happens when the card was stolen for any of the unauthorized purposes or even when the fraudster uses the credit card information for his use. In the present world, we are facing a lot of credit card problems. To detect the fraudulent activities the credit card fraud detection system was introduced. This project aims to focus mainly on machine learning algorithms. The algorithms used are Random forest. The results of the algorithms are based on accuracy. The Random forest model algorithm that has the greatest accuracy considered as the best algorithm that is used to detect the fraud.

The eleventh paper is based on “Sign language detection”
A real-time sign language detection that can scan photos and identify indicators quickly at the rate of streaming images is absolutely necessary. This design proposes leveraging neural networks for sign language translation and text tracking and identification. For those learning sign language for the first time, a system that utilizes hand detection has been created. This methodology is based on the explicit skin-color space thresholding method of skin-color modelling. The Convolutional Neural Network (CNN) model was used to classify the images after being fed the photos. For image training, Keras was utilized. Of this, 90.04% was attributable to ASL alphabet recognition, 93.44% to number recognition, and 97.52% to static word identification. The method is real-time and utilized for quick computing.

The twelfth paper is based on “Text to speech Synthesis”
A Text-to-Speech (TTS) system converts normal language text into speech. In this system, the input text is taken, and output is given as speech. The quality of a speech synthesizer is judged by its naturalness and intelligibility.
TTS systems are based on a complex pipeline. TTS is one of the major NLP applications. Text-to-speech discussion involves three key stages of text analysis, text processing and waveform production i.e., speech construction. TTS is an application that converts text into speech, the user enters text and receives output as audio.

The thirteenth paper is based on “Object detection using deep learning algorithms”
The goal of detection is to output an image's labels and borders. Object detection is a critical subject in the science of computer vision. It offers significant research and application value in the areas of monitoring and autonomous driving. In recent years, deep learning has made a breakthrough in the investigation of picture categorization and led to the rapid growth of object vision identification.
In this paper, we use the YOLO deep learning algorithm. The YOLO algorithm generates a bounding box and prints the object's confidence level. For the production of blobs, we used a Deep Neural Network, which aids in the analysis of shape features of objects such as area, length, position, and direction of lumps, and we smoothed the image using Gaussian blur.

***Dr. Shilpa Abhang***
***Associate Professor***
***Department of MCA, Jyoti Nivas College Autonomous***

***DHII***
***Journal Dept. of MCA***
***Jyoti Nivas College Autonomous***

# OBJECT DETECTION IN COLOR IMAGES

*Vidya Rathod*

***Abstract***

*Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, fruits, or cars) in digital images and videos. Well-researched domains of object detection include face detection and pedestrian detection. Object detection has applications in many areas of computer vision, including image retrieval and video surveillance. It is widely used in computer vision tasks such as face detection, face recognition, video object co-segmentation. It is also used in tracking objects, for example tracking a ball during a football match, tracking movement of a cricket bat, or tracking a person in a video.*

***Keywords::*** *object_recognition, face_ recognition.*

.

## I. INTRODUCTION

Face Object detection is a well-known computer technology connected with computer vision and image processing that focuses on detecting objects or its instances of a certain class (such as humans, buildings, flowers, animals, fruits) in digital images and videos. There are various applications of object detection like face detection, character recognition, and vehicle calculator, pedestrian detection. Object detection can be used for various purposes including retrieval and surveillance. In these various basic concepts used in object detection while making use of OpenCV library of python, improving the efficiency and accuracy of object detection are presented**.**

Object detection methods generally fall into either machine learning based approaches or deep learning-based approaches. In machine learning approaches, it becomes necessary to define features using technique such as Support Vector Machine (SVM) to do classification whereas in deep learning techniques are able to do end-to-end object detection without specifically defining features and are typically based on Convolutional Neural Networks (CNN).

## II. LITERATURE REVIEW

Object detection is widely used in computer vision tasks such as image annotation, activity recognition, face detection, vehicle recognition and tracking objects.

**Paper 1:** In paper 1 authors have proposed to analyze and review the previous approach towards object tracking and detection using video sequences through different phases. Also, identify the gap and suggest a new approach to improve the tracking of object over video frame.

**Paper 2:** In paper 2 authors have proposed an Easynet model looks at the whole image at test time so its predictions are informed by global context. In prediction time easy model generates scores for the presence of the object in a particular category with single network evaluation.

**Paper 3:** In paper 3 authors have proposed a deep learning technique used to detect a live object, localize an object. Categories an object, extract features, appearance information and so on. To build deep learning based real time object detector with OpenCV need access webcam in an efficient manner and to apply object detection to each of the frame.

**Paper 4**: In paper 4 authors have proposed an efficient approach to localize visible object or region of object from an image, using less memory and computing power. And also implemented CNN based classification which has proved to provide better performance than baseline works. Fruit Detection and Recognition (FDR) is very proficient regardless of different limitations such as high and poor image quality, complex background, different fruits of shape and color, multiple overlapped fruits, existence of non-fruit object in the image.

**Paper 5**: In paper 5 authors have proposed a work explains the development of a framework to automatically detect the smoking scenes using neural network model and display the required warning message. The challenge to detect the smoking scenes in video clips is that only small portion of the smoking event and displayed for fraction of second. Object detection is used to detect based on shapes, colors and sizes.

**Table:** Existing methods used

| Papers | Techniques used |
|---|---|
| A Review of Detection and Tracking of Object from Image and Video Sequences | Tracking algorithm to smoothen the video sequence |
| Object Detection and Recognition in Images | Feature Extraction pattern recognition algorithm Object recognition algorithms rely on matching |
| Real-Time Object Detection System using Caffe Model | Single Size Multi Box Detector<br>R-CNN<br>Caffe Model |
| Multi Class Fruit Classification Using Efficient Object Detection and Recognition Techniques | Fruit Detection or Localization Process<br>CNN recognition model |
| Development of framework for detecting smoking scene in video clips | Region-Based Convolutional Neural Networks<br>Fast R-CNN |

## III. PROPOSED METHODOLOGY

The aim is to build a system that can detect the object from the given image. In our proposed work we train the system by giving some train data and then we try to detect the object. According to features, shape and size of the object we train data.

**HAAR cascade classifier:** It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images. The classifier is an xml file and has lot of definitions/patterns inside. When some object matches with those defined patterns then code will identify and categorizes that object.

**1)** First step is to collect the **Haar Features**. A Haar feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums.

**2) Integral Images** are used to make image super-fast.

For example, consider the image below. Top row shows two good features. The first feature selected seems to focus on the property that the region of the eyes is often darker than the region of the nose and cheeks. The second feature selected relies on the property that the eyes are darker than the bridge of the nose. But the same windows applying on cheeks or any other place is irrelevant.

**3) Adaboost** which both selects the best features and trains the classifiers that use them. During the detection phase, a window of the target size is moved over the input image, and for each subsection of the image and Haar features are calculated.

**4) Cascade Classifier**

- The cascade classifier consists of a collection of stages, where each stage is an ensemble of weak learners. The weak learners are simple classifiers called decision stumps.
- Each stage is trained using a technique called boosting. Boosting provides the ability to train a highly accurate classifier by taking a weighted average of the decisions made by the weak learners.
- Each stage of the classifier labels the region defined by the current location of the sliding window as either positive or negative. Positive indicates that an object was found and negative indicates no objects were found.

## IV. IMPLEMENTATION

This research is implemented in MATLAB, OPENCV, and PYTHON environment It works on Windows OS and it is user friendly.

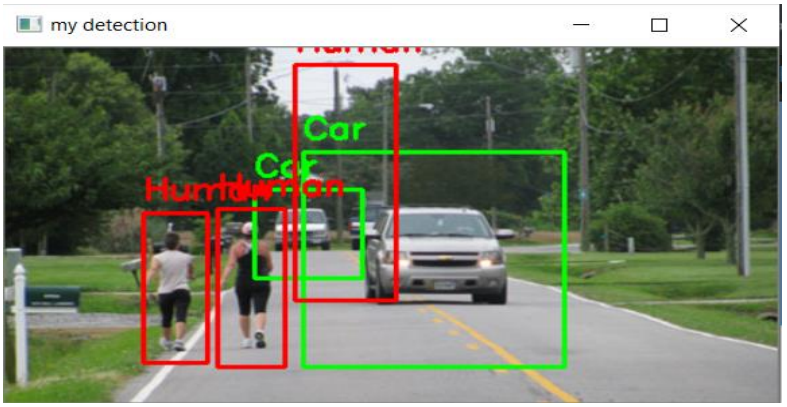
Dataset and Environment used: All the images are imported in .jpg and .png format. The data set includes images with car and human focused in image.

Environment: Operating System-Windows 10, Programming Language-Python 3.8, Editor-Visual code editor, Python libraries- Open CV, Pandas, NumPy, Math, Matplotlib, NumPy. Dual core 1.7GHz processor,8GB RAM, 500GB Storage space.

**Image1**

**Input Image**

**Output Image**




**Explanation:**

1. Read the file using OpenCV and create an instance.
2. Define the classifiers. Convert the color image into grey image for faster processing, as most of the cases color is not an important factor to identify the objects rather the patterns. Now create trackers for individual entities (car/human/etc.) by passing the classifiers via OpenCV Cascade Classifier method.
3. Apply the trackers on the grey image to identify the position of the objects (car/human/etc.) multi-dimensional array values appear. The array contains the location of the objects detected by the program.
4. Iterate through the above multi-dimensional array and draw rectangle around each object. Finally required output is detected.

**Image 2**

**Input Image**

**Output Image**


**Explanation:**

1. Read the file using OpenCV and create an instance.
2. Define the classifiers. Convert the color image into grey image for faster processing, as most of the cases color is not an important factor to identify the objects rather the patterns. Now create trackers for individual entities (car/human/etc.) by passing the classifiers via OpenCV Cascade Classifier method.
3. Apply the trackers on the grey image to identify the position of the objects (car/human/etc.) multi-dimensional array values appear. The array contains the location of the objects detected by the program.
4. Iterate through the above multi-dimensional array and draw rectangle around each object. Finally required output is detected.

**Applications of Object Detection:**

1. **Self-Driving Cars**-Self Driving Cars may use Object detection and recognition system to identify pedestrians and cars on the roads and then make the suitable decision in accordance.
2. **Face Detection**-Another application of Object detection and recognition is Face Detection. examples- Facebook recognizes people before they are tagged in images.
3. **Medical Science**-Object Detection and recognition system may help Medical science to detect diseases. For e.g.-Detecting Tumors and various cancers.
4. **Text Recognition**-Text recognition deals with recognizing letters/symbols, individual words and series of words. Ex-Recognizing handwriting of a person.
5. **Hand Gesture Recognition**- Hand Gesture Recognition deals with recognition of hand poses, and sign languages.

## V. CONCLUSION AND FUTURE SCOPE

**Conclusion:**

Object recognition is one of the fundamental tasks in computer vision. It is the process of finding or identifying instances of objects in digital images, stored videos and real time videos. Object detection refers to the capability of computer and software systems to locate objects in an image/scene and identify each object. There are many ways object detection can be used as well in many fields of practice. Like every other computer technology, a wide range of creative and amazing uses of object detection will definitely come from the efforts of computer programmers and software developers.

**Future Scope:**

Design and simulation of complex video sequence and test them using same tracking algorithm. In the potential scenario, occlusion is used for an object with the same color for the moving objects or else using bigger occlusion with longer occlusion time. Increasing the number of the object help to identify the efficiency and functionality of the tracking algorithm.

## VI. REFERENCES

**1**. Mukesh Tiwari Dr. RakeshSinghai, A Review of Detection and Tracking of Object from Image and Video Sequences , International Journal of Computational Intelligence Research ISSN 0973-1873 Volume 13, Number 5 (2017), pp. 745-765 © Research India Publications.

**2**. Sandeep Kumar, Aman Balyan, Manvi Chawla, Object Detection and Recognition in Images, Computer Science &Engineering Department, Maharaja Surajmal Institute of Technology, New Delhi, India. © 2017 IJEDR | Volume 5, Issue 4 | ISSN: 2321-9939.

**3**. Vaishali, Shilpi Singh,Real-Time Object Detection System using Caffe Model ,International Research Journal of Engineering and Technology (IRJET) ,Volume: 06 Issue: 05 | May 2019.

**4**. Rafflesia Khan and Rameswar Debnath ,Multi Class Fruit Classification Using Efficient Object Detection and Recognition Techniques ,I.J. Image, Graphics and Signal Processing, 2019, 8, 1-18 Published Online August 2019 in MECS.

**5**. G, Shashank B. N, Athri G Rao ,Development of framework for detecting smoking scene in video clips, Indonesian Journal of Electrical Engineering and Computer Science Vol. 13, No. 1, January2019,

# CREDIT CARD FRAUD DETECTION USING LOGISTIC REGRESSION

*Swetha M*

***Abstract***

*Credit card fraud detection is presently the most frequently occurring problem in the present world. This is due to online transactions and e-commerce platforms. The fraud happens when the card was misused any stolen by the fraudster. In this world, we humans are facing a lot of credit card problems in some or other way. To detect the fraud this system was introduced. This project aims to focus mainly on machine learning algorithms. The algorithms used are Logistic Regression. The results are based on accuracy. The logistic regression model algorithm that gives the more accuracy considered as the best algorithm to detect the fraud.*

***Keywords::*** Logistic Regression, ANN Algorithm.

## I. INTRODUCTION

Whenever we hear the word Credit Card the first thing that pops in our mind is the frauds that are associated with these cards. Although a credit card has many advantage when they are used in a proper manner but it damages can be caused to it by many fraud activities too. But in today's fast growing world these frauds can be detected with a huge knowledge of machine learning algorithms. After the implementation of the model, we can use it further to identify the fraud, a new transaction that is occurring as fraud or not. Usually, our focus here is to detect 100% fraud transactions that is being occur. In today's world this is being the major concern, which demands the attention of the fields such as Machine Learning, Deep Learning Artificial Intelligence, etc. where the solution of this issue can be real-time. Our aim is to predict the accuracy of the fraud detection with the help of different algorithms. Further this analysis can be used to implement detecting the fraud.

➢ Major problem is that online payment does not require physical card. Anyone who knows the details of the card can make fraud transactions. Card holder comes to know only after the fraud transaction is carried out.

## II. METHODOLOGY

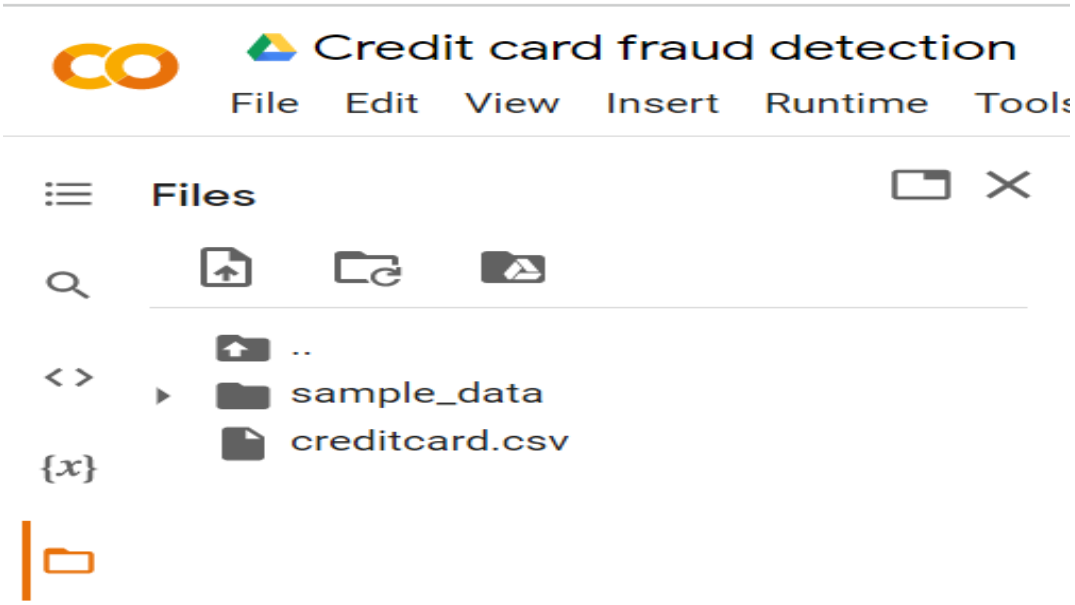
Fraud detection using artificial intelligence is proposed. The method uses logistic regression which helps to build the classifier that prevent frauds in credit card transactions.

To handle dirty data and to ensure a major degree of detecting accuracy, a pre-processing step is used.

There is a need to have sample Excel sheet dataset which we will use in testing our model. The csv files need to be in the same directory with your script with an extension of .csv.

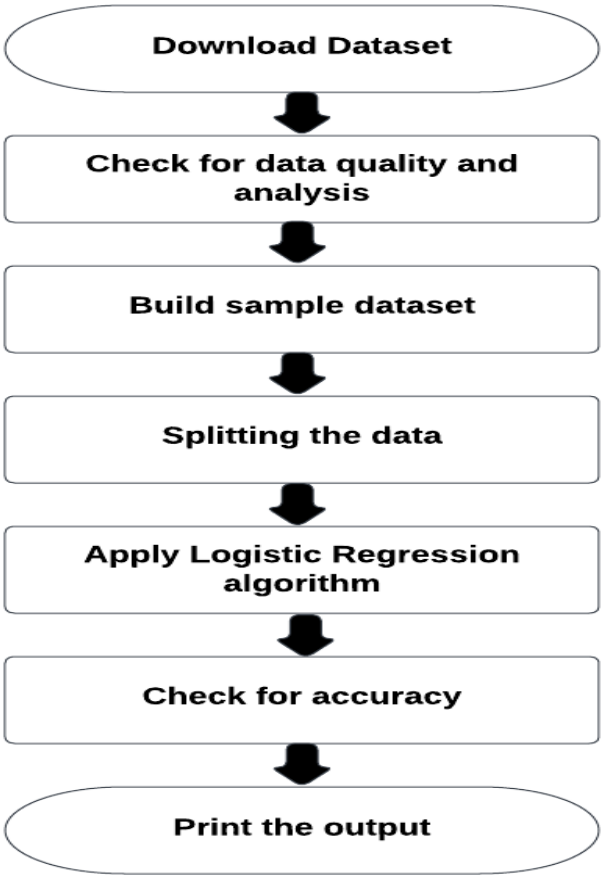
**Logistic Regression:** Logistic regression is one of the popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is also used for predicting the categorical dependent variable by using a given set of independent variables. Logistic regression helps to predict the output of a categorical dependent variable. Therefore, the outcome will be a categorical value. It is denoted as either Yes or No, 0 or 1, true or False, etc. but instead of giving the perfect value as 0 and 1, it gives the binary values which lie between 0 and 1.

To solve this sample Excel sheet dataset required which will be used in testing. CSV files which are being tested should be in the directory with the extension .CSV




Dataset file Image A

## 2.1 WorkFlow




## III. EXPERIMENTAL RESULT

Accuracy table

| Method Used | Accuracy On Training data | Accuracy on Testing data |
|---|---|---|
| **Accuracy on Testing data** | 94.28 | 91.37 |
| **Accuracy on Testing data** | 84.60 | 81.00 |

We have used balanced dataset to predict the fraudulent transaction, which is downloaded from kaagle.com. We have to import the require packages and check the dataset quality followed by splitting the data into train data and test data to get the accuracy.

Using ANN algorithm, logistic regression, decision tress comparatively we are getting the accuracy of loss 84% in ANN, and in logistic regression accuracy rate of Test data and training data is 94.28%, 91.37%.

**Correlation with class**

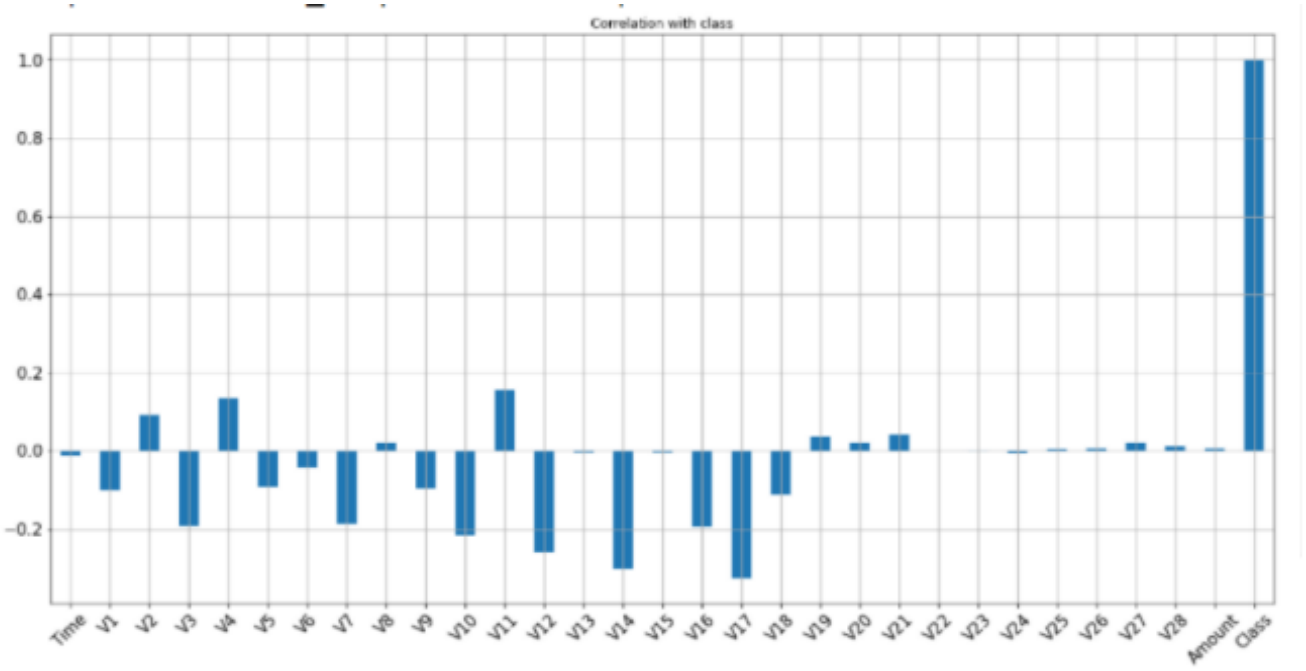



Image b

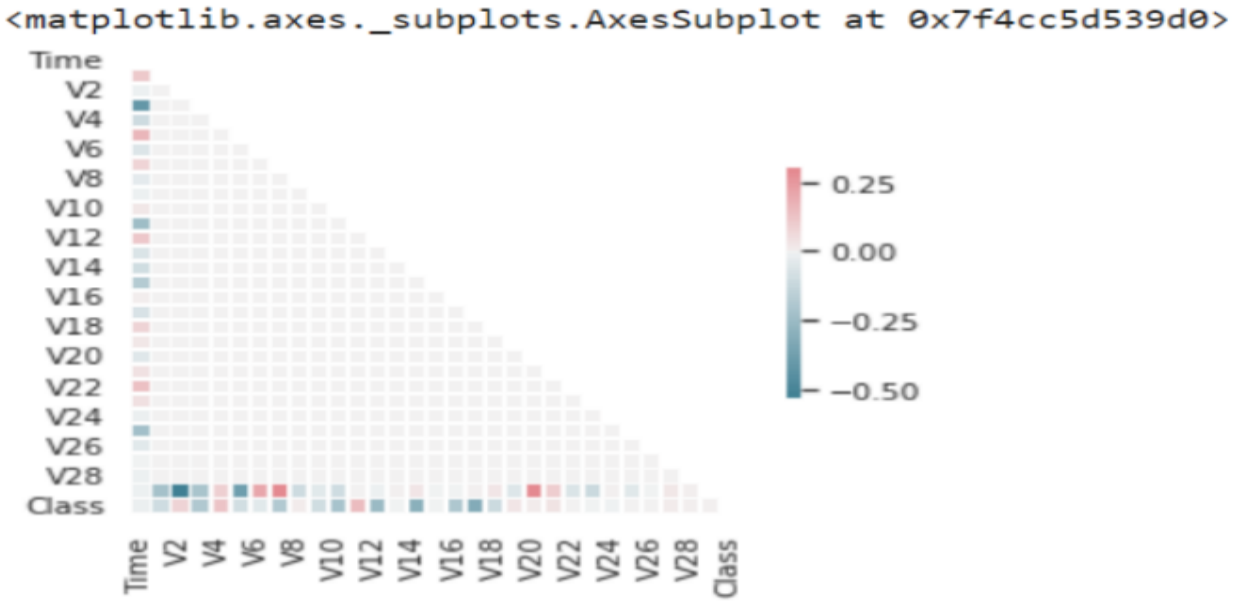



Image c

## IV.CONCLUSION

In credit card fraud detection paper, Machine learning technique of logistic regression, artificial neural network were used to detect the fraud that happens in credit card system. Sensitivity, specificity, accuracy is used to assess the performance for the proposed system.

The accuracy for logistic regression for training data is 94.28 and testing data is 91.37and ANN accuracy loss of 84% respectively. By comparing both the method, found that logistic regression is better than the ANN.

## V. REFERENCES

**1.** Sanobar khan, Sanovar, Suneel Kumar, Mr Hitesh Kumar, "Credit card fraud detection using machine learning" Published year: June 2021.

**2**. Vaishnavi Nath, "Credit card fraud detection using machine learning algorithms" Published year: 2019.

**3**. Dr Gaurav Gupta. "Machine learning approach for credit card fraud detection" Published year: September 2019.

**4.** Prof Amol Jadav, "Credit card fraud detection using machine learning" Published year: September 2020.

**5**. S P Mani raj, "Credit card fraud detection using machine learning and data Science" Published year: September 2019.

**6**. Dr Jitendra Kumar, Ashutosh Kumar Singh, "Credit Card Fraud Detection using Machine Learning: A Study" Published year: August 2021.

**7**. Harshita Anand, Richa Gautam Raman Chaudhry, "Credit Card Fraud Detection using Machine Learning" Published year: May 202.

# TEXT BASED PLAGIARISM CHECKER

*Misbah Rafia Khanum, Neha Kouser*

***Abstract***

*Student assessment is a vital part of the teaching and learning process. The idea of the research project is to propose a plagiarism based system for online paper evaluation. The input response sheet from the students must pass both a syntactic and a semantic similarity check with the evaluation method. Plagiarism in student work is a pervasive and aggravating issue in the academic process. Traditional manual plagiarism detection by humans is a challenging, imprecise, and time-consuming method.*

***Keywords::*** Cosine *similarity, Vectorization, Word embedding, Plagiarism*

## I. INTRODUCTION

Plagiarism is the sense of stealing ideas continues to be fought against as with other forms of research misconduct (e.g., Fraud and Duplication)
Plagiarism in student work is a pervasive and aggravating issue in the academic process. Traditional manual plagiarism detection by humans is a challenging, imprecise, and time-consuming method.
Many students engage in some form of academic fraud and plagiarism. Teachers typically find it challenging to manually detect plagiarism in student work. If detection is carried out automatically, it becomes simpler, faster, and more effective (i.e. via a computerised system). The practise of finding instances of plagiarism in a document's text or code is known as plagiarism detection. Manual or computer-aided detection methods are both possible. Manual detection is labor - intensive and impractical when there are numerous papers to compare.

## II. LITERATURE REVIEW

**Paper 1**: One of the major problems in the research world is plagiarism, which is a bigger concern in the educational system. This research compares unigram, bigram, and trigram of vector space model with cosine similarity measure to identify efficient plagiarism detection tool on selecting optimal intra-corpal plagiarism detection for text-based assignments.

**Paper 2:** Students benefit from the automatic verification of online tests. Because of this, instruction can be the main focus of formal practical sessions. Nevertheless, there is a chance that students would abuse such systems, discrediting the evaluation process. This essay focuses on instances of plagiarism found in a SQL learning environment**.**

**Paper 3:** Any document is basically a set of words, keywords, and some other terms. One can analyse a document on any of the above mentioned basis. It makes the process easy to owe someone's ideas and work or take the credit of other's work. In academics it is very common to copy others assignments from online sites. So there is a need to identify students from plagiarism.

**Paper 4:** Computers have changed the way education works. Computers offer resources that supported entire process of learning with the help of advanced technologies that enable easier and faster understanding of complex concepts. This is called e-learning type of education that uses a variety of technologies to deliver skills and knowledge. Education can be provided inside classrooms, or can be distributed through online classes, using content such as internet, CD-ROM, etc.

**Paper 5:** In the modern era, the usage of freely accessible internet resources is encouraging a culture of easy copy and paste, which leads to plagiarism in a variety of research papers and academic reports. Many strategies and methodologies have been established for the plagiarism detection to assess the originality in the research documents in this situation of the expanding research and development publications. This essay outlines numerous strategies for spotting and avoiding plagiarism in books, journals, and academic works.

**Paper 6:** The rapid development of technology has made it easier to get knowledge via a variety of channels, which has made plagiarism possible. Effective software systems are required to monitor academic integrity. This paper conducts a study on plagiarism with an emphasis on the quickly developing field of extrinsic text plagiarism detection. Based on the state of the art at the time, the various extrinsic detection approaches are discussed.

**Paper 7:** The preprocessing module adds some additional screening while also converting the incoming documents to a standard format. Plain text is the most frequent format used by Plag Zap. It transforms files between the various formats it can handle. Linux sub-processes and temporary files were employed for the integration. Custom code was used to implement further filters. Although everything worked as intended, sporadic conversion errors and high CPU utilisation prompted us to try out alternative solutions.

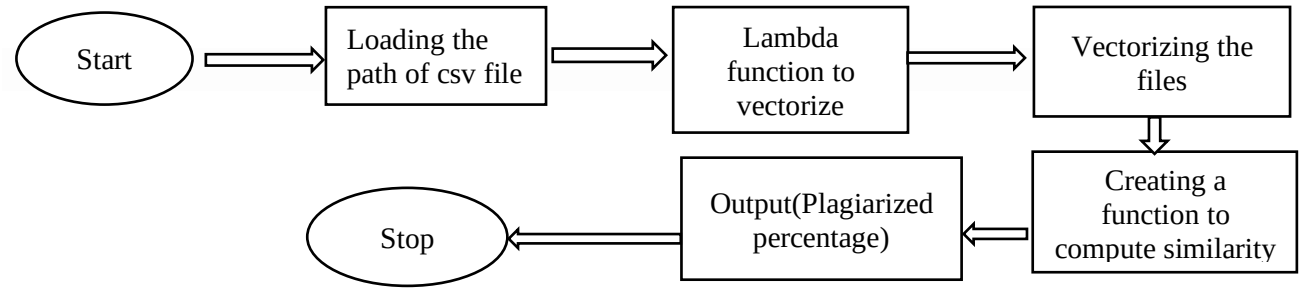
## III. PROPOSED METHODOLOGY

By determining if the text has been duplicated using cosine similarity, this method attempts to create a plagiarism checker.

### 3.1 Similarity in cosine

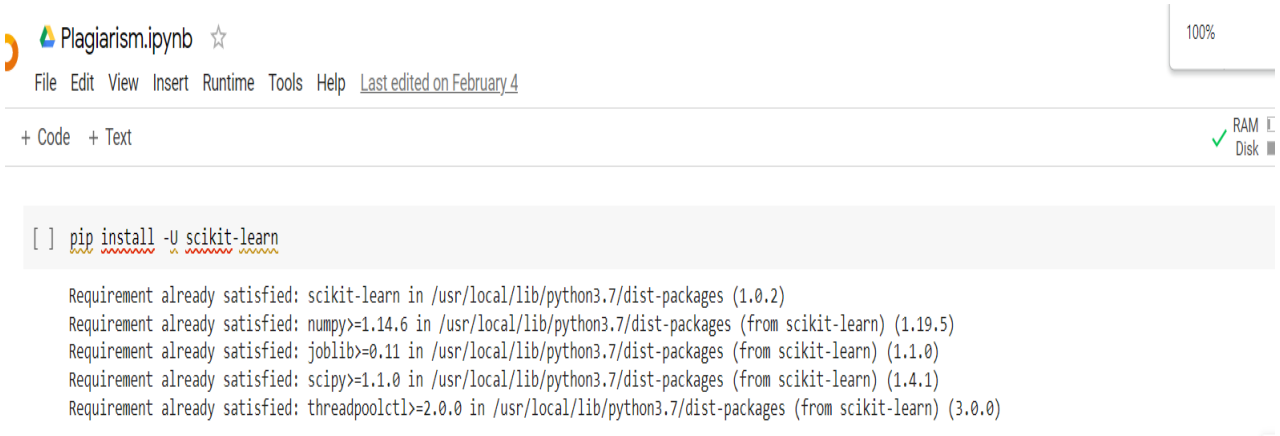
By estimating the value of cosine similarity between vector representations of students' text assignments, we will apply the fundamental notion of vector, dot product to ascertain how closely two texts are similar. The student's homework must include sample text documents that we will utilise to evaluate our model. The csv files must have the.csv extension and be located in the same directory as your script.

**FLOW DIAGRAM:**




## IV. IMPLEMENTATION AND RESULTS

The screenshots of the completed implementation are as follows: - Scikit-Learn needs to be set up on your computer



Let's import all required modules first.
Using the OS Module to import text files, TfidfVectorizer to do word embedding on our textual data, and cosine similarity to calculate plagiarism are the next steps.

```
import os
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity
```

Using List Comprehension to read all text files
To load all of the path text files in our project directory, we will use the ideas of a list comprehension.

```
[ ] student_files = [doc for doc in os.listdir() if doc.endswith('.csv')]
    print(student_files)
```

```
['Sample Survey - Data.csv', 'Description - Sheet1.csv']
```

Lambda function for vectorization and similarity calculation

Two lambda functions must be written, one to translate text into arrays of integers and another to determine how similar the two are.

Textual data is first converted into an array of numbers using a lambda function.

To calculate the similarity between the two csv files, use a second lambda function.

To vectorize the loaded student files, add the two lines below to the text data.

```
[ ] s_vectors = list(zip(student_files, vectors))
    print(s_vectors)
```

```
[('Sample Survey - Data.csv', array([0.37997836, 0.53404633, 0.        , 0.53404633, 0.        ,
       0.53404633])), ('Description - Sheet1.csv', array([0.44943642, 0.        , 0.6316672 , 0.        , 0.6316672 ,
       0.        ]))]
```

```
[ ] plagiarism_results = set()
    print(set())
```

```
set()
```

Making a Similarity Calculation Function

The primary part of our script that controls the entire similarity computation procedure is shown below.

```
[ ]  for data in check_plagiarism():
         print(data)
```

```
('Description - Sheet1.csv', 'Sample Survey - Data.csv', 17.077611319011652)
```

It determines the content that is plagiarised in percentage format by comparing the similarity between the two CSV files, and the percentage of plagiarised content is then displayed.

The lowest is anticipated to be 0, while the maximum is dependent on the amount of content that has been copied or reproduced from another website.

## V. CONCLUSION

The similarities between the two csv files are checked for plagiarism using a plagiarism checker, which computes them using cosine similarity. The test csv files ought to be located in the same directory. By comparing it with internet repositories, a plagiarism checker aids in the removal of poorly rephrased terms.

If for some reason your professor still has questions about certain elements of your paper, you can obtain a printed copy of the plagiarism check to show that you've taken the appropriate safeguards. By comparing it with internet repositories, a plagiarism checker aids in the removal of poorly rephrased terms. You can show proof that you've taken the essential safeguards by showing a printed copy of the plagiarism check.

## VI. REFERENCES

**1.** Plagiarism Detection on Electronic Text based Assignments Using Vector Space Model. by Roshan Ragel, Mohamed Abdul Cader, Jiffriya, MAC, and Akmal Jahan, University of Peradeniya 24.12.2014, University of Peradeniya, Peradeniya, Sri Lanka.

**2**. Online assessment and checking of sql: identifying and preventing plagiarism, by Gordon Russell and Andrew Cumming, 11 August 2015.

**3.** Plagiarism Checker in Text Mining, Ankita Bahl, Anu Saini, and Mitali Singh, Maharaja Surajmal Institute of Technology, January 2016.

**4.** Plagiarism detection in homework assignments and term papers by Darko Puflovi and Leonid Stoimenov Electronic Engineering Department, University of Ni, 24-25 September 2015, Belgrade, Serbia.

**5.** Detecting The Plagiarism For Text Documents On The World Wide Web, by Durga Bhavani, Dasaril, and Dr. Venu Gopala Rao, was published on October 10, 2014, at Jawaharlal Nehru Technological University in Hyderabad.

**6**. Study on Extrinsic Text Plagiarism Detection Techniques and Tools, Vani K1 and Deepa Gupta2, Amrita School of Engineering, Amrita University, Amrita Vishwa Vidyapeetham, September 10, 2016.

**7.** University of Constanta, Romania and Naval Academy, 10 September 2016. Elena Bautu and Andrei Bautu, PlagZap - A Textual Plagiarism Detection System for Student Assignments Built with Open-Source Software."

# THE COMPARATIVE STUDY OF NOISE REMOVAL TECHNIQUES IN COMPOUND IMAGES

*G Harshitha*

***Abstract***

In the advancing world many new techniques are being discovered to reflect data. Image is used to represent data, compound images are frequently in use. Noise is present many images captured, this causes visual disturbance and degrades the quality by corrupting the properties of the image. In order to maintain good quality of images, noise should be removed, this helps to regain quality of the image. Noise removal in an image can be done by applying the filters such as Median, Wiener, Gaussian filter, etc. This paper deals with Speckle, Salt & pepper, Gaussian noise and filters such as Mean, Median, Gaussian filters.

***Keywords::*** *Mean filter, Median filter, Gaussian filter, Speckle noise, Salt & pepper noise, Gaussian noise*

## I .INTRODUCTION

Image processing is the technique to perform operations to enhance the image, this helps us to achieve the information about the image [1]. The software used in image processing is digital image processing, it enhances both analog and digital signals, image and voice signals, etc. Image sharpening, image restoration, remote sensing, noise removal, machine vision are few of the tasks that digital image process helps the users to perform. Fast and high-quality image is distributed, but in some cases the process is time consuming and only qualified user will be able to work on it with comfort.

Many visual contests are available, it may be of any kind such as text, picture and graphics. These are very complicated and when downloaded, the contents gets corrupted by the noise[1]. The quality is regained by doing some operations on the downloaded content like preprocessing and if it is an image the image filtering is done in order to remove the noise that occurred during image acquisition. This process helps in providing good visual enhancement and the filtering techniques depends on the type of image that is being chosen [3].

Compound images are images that contains graphical representation, picture of a person or photo, text. Compound images involves all the different types of visual representation [1].

## II. LITERATURE REVIEW

**Paper 1**: Hlaing Htake Khaung Tin Proposed preprocessing is done to recognize overall noise present in an image, then the sourse of the noise is understood. Certain filters such as Linear, Median, Adaptive filters are used to remove the noise in an image and histogram equalization is done.

**Paper 2**: Noise is applied to all the images that is taken as input and filters like Median, Mean, Weiner, Center Weighted median, Averaging filter are the filters that is being worked on [1]. Performance evaluation is done in order to find which filter is best suited for specific filter was proposed by Zinat Afrose.

**Paper 3**:Sukhjinder Kaur proposed that before applying image processing tool to the image noise removal is necessary [6]. Comparision of the performance of each filter is measured through parameters like PSNR, MSE, RMSE, SSIM, PSNR.

**Paper 4**:Abdalla Mohamed Hambal, Dr. Zhijun Pei, Faustini Libent Ishabailu worked on coloured images. Denoising of the images are done by applying noise removal techniques that involues all kinds of filters [5]. It was stated that denoising capability can corrupt colour component of the image.

**Paper 5**:Bhawna Dhurv, Neetu Mittal, Megha Modi proposed that comparative analysis is done by determining the parameters – time of execution and entropy.Noise removal serves as one of the crucial steps for improving the quality of the image[5].

**Paper 6**: Muna Khalid Jasim, RehanHamdullah Najm, Emran Hassn Kanan, Hamza Esam Alfaar, Mohammed Otair proposed Mean, Median and Weiner filter can be used for comparative analysis of noise removal techniques [4]. MSE, MAE, PSNR values of each noise with respect to the filter used are compared.

**Paper 7**:Roman Garnett, Timothy Huegerich, Charles Chui proposed that noise removal algorithm with Impulse Detector can be used to remove the noise, that involves Bilateral filtering, image restoration for Gaussian noise, Implusive noise, Non-linear noise and Mixed noise.

**Paper 8**:Suresh kumar, papendra kumar, manoj gupta, ashok kumar nagawat [7]
worked on a comparative analysis of the filters used for denoising the images. The filters such as Median filter and Weiner filter was compared using different performance analysis parameters.
III. PROPOSED METHODOLOGY

## 1. Mean filter

Mean filter woks pixel by pixel replacing the center value of the pixel by average or mean of all the pixel value that is present in the window. The window in which the replacement of the pixels occur are called window or kernel [1]. It is used for image smoothening by reducing the intensity or brightness variation between the pixels surround each pixel. Mean filter is a linear filter in which the output pixels value is the linear combination of all the pixel surrounding each pixel [6].

**(a)** **(b)**

**Fig.1: (a)Original image, (b)Mean filter**

## 2. Median filter

The filter is replaces pixel by pixel by calculating median value of the neighbouring pixel [3]. The median is calculated by arranging all the values of the pixel from the neighbouring pixel into numerical order and the pixel will be replaced by middle pixel value. This happened over the entire image [7].

**(a)** **(b)**

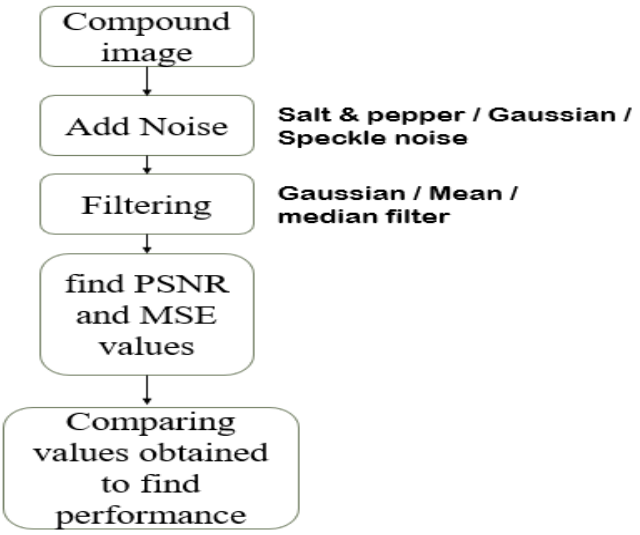
**Fig.2: (a)Original image, (b)Median filter**

### 3. Gaussian filter

Gaussian filter usually is used to smoothen and blur the image, this degree of smoothening which is done is studied by standard deviations of the Gaussian [6]. Gaussian filter is called as low pass filter using high frequency component and blurring specified place in the image. Symmetric kernel is always considered while performing all the operations on the image.

(a) (b)

**Fig.3: (a)Original,(b)Gaussian filter**




Add noise algorithm for all the images in the data set. Filter/remove the noise from the images using mean, median and gaussian at a time for each image. To find which filter gives better performance PSNR values is found by comparing noise image and filtered image. Then MSE values is found by comparing noise image and filtered image.

## IV. EXPERIMENTAL RESULTS

Compound images are be taken as the input for processing and various noise such as salt & pepper noise, gaussian noise, speckle noise are added to the compound image, then the noise added is filtered out using filters such as gaussian filter, mean filter, median filter. The comparisons of the filters best suited for the particular noise will be studied using the values such as PSNR(peak signal to noise values), MSE(mean square error).

**(a)** **(b)** **(c)** **(d)** **(e)**

**Fig 4: (a)Original image, (b)Salt and pepper noise, (c)Median filter, (d)Mean filter, (e)Gaussian filter**

**1. Performance evaluation**

Performance evaluation of the filters such as Gaussian, Mean, Median filter are done through performance parameters like PSNR (Peak signal to noise ratio) and MSE (Mean square error). The PSNR values obtained must range between 30 to 50 db and the lower the MSE values lower will be the error of the output operated on. Average of all the values that are performed on 5 images are stated in the in the tables below.

**Table 1: PSNR values of the compound images with Salt & pepper, Gaussian, Speckle noise.**

| **Noise** | **Median (db)** | **Mean (db)** | **Gaussian (db)** |
|---|---|---|---|
| Salt & pepper | 32.67 | 28.02 | 27.93 |
| Gaussian | 32.74 | 28.51 | 28.52 |
| Speckle | 31.68 | 27.71 | 27.60 |

**Table 2: MSE values of the compound images with Salt & pepper, Gaussian, Speckle noise.**

| **Noise** | **Median filter** | **Mean filter** | **Gaussian filter** |
|---|---|---|---|
| Salt & pepper | 34.97 | 98.29 | 104.62 |
| Gaussian | 34.88 | 91.64 | 91.34 |
| Speckle | 44.36 | 110.30 | 112.81 |

## V. CONCLUSION

The research depicts to remove the noise i.e., salt & pepper, gaussian and speckle noise in the compound image using the filters – median, mean and gaussian. The performance of the filters is measured and compared using PSNR and MSE values of each image and each filter. From the performance analysis it is concluded that median filtering gives better results.

## VI .REFERENCES


**1**. Zinat Afrose, "A Comparative Study on Noise Removal of Compound Images using Different Types of Filters", International Journal of Computer Applications (0975-888) Volume 47-No.14, June 2012.

**2**. Hlaing Htake Khaung Tin, "Removal of Noise by Median Filtering in Image Processing", The 6th Parallel and Soft Computing (PSC 2011).

**3**. Bhawa Dhruv, Neetu Mittal, Megha Modi, "Analysis of Different Filters for Noise Reduction in Images", 2017 recent Developments in Control, Automation & Power Engineering (RDCAPE).

**4**. Muna Khalid Jasim, RehanHamdullah Najm, Essam Alfaar. Mohammed Otair, "Image Noise Removal Techniques: A Comparative Analysis", Volume 8, No.6, November -December 2019.

**5**. Abdalla Mohamed Hambal, Dr. Zhijun Pei, Faustini Libent Ishabailu, "Image Noise Reduction and filtering Techniques", International Journal of Science and Research (IJSR) ISSN (online): 2319-7064.

**6.** Sukhjinder Kaur, "Noise Type and Various Removal Techniques", International Journal of Advanced Research in Electrics and Communication Engineering (IJARECE), Volume 4 Issue 2, February 2015.

**7**. Suresh kumar, papendra kumar, manoj gupta, ashok kumar nagawat, "Performance comparison of median and wiener filter in image de-noising",international journal of computer applications (0975 – 8887) volume 12– no.4, november 2010.

**8.** Amir Said and Alexander I. Drukarev, "Simplified segmentation for compound image compression," Proc. IEEE Int. Conf. Image Processing, vol. 1, pp. 24-28, Oct. 1999.

**9**. Piotr S. Windyga, "Fast Impulsive Noise Removal", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 10, NO. 1, JANUARY 2001.

**10**. Roman Garnett, timothly Huegerich, Charles Chui, Fellow, IEEE, and Wenjie He, Member, IEEE, "A Universal Noise Removal Algorithm With an Impulse Detector"IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 14, NO. 11, NOVEMBER 2005

# DROWSINESS DETECTION SYSTEM

*Neha SreeshKumar*

***Abstract***

Drowsiness is a state of close sleep due to fatigue which reduces mental alertness and reduces a person's ability to function. The Sleep Detection System helps detect if the person in front of the camera is working, drowsy or asleep. This helps to reduce the risk of accidents while driving, to wake up in the classroom whena student is present and to be alert at meetings and presentations when an employee is present.

The main goal of the Sleep Scheme is to develop an automated system for diagnosing whether a person is active or sleepy using dlib and face recognition technology. When the subject is drowsy or drowsy it is characterized by marking features of the face and eyelids and sounding the alarm to wake them up.

***Keywords::*** drowsiness_detection.

## I. INTRODUCTION

People have been inventing machines and devising strategies to simplify and protect their lives, to perform routine tasks such as going to work, or for interesting purposes such as air travel.

This document is a report on a review of research conducted and a project undertaken in the field of computer programs to improve the system of drowsiness detection to prevent driver accidents, student absenteeism and staff meetings. In addition, this paper cites an overview of the authors' recognition to help improve the field in which it is used in order to obtain assistance in the most efficient way.

In particular, in this project the camera is used to capture video of the title and the facial features are identified and marked. We are mainly focused on the student mark and eyeglasses. Based on changes in the length and width of the eyelids the system indicates whether a person is asleep, drowsy or active. If a person continues to sleep for more than 15 seconds the alarm starts to sound and a message is sent to the number stored in the code stating that the owner of this number is asleep.

When a person opens his eyes again the alarm stops and continues when he falls asleep again.

The main goal of the Drowsiness Detection system is to develop an automated system for diagnosing whether a person is active or sleepy using dlib and face recognition technology.

When the interviewee is drowsy or asleep it is marked by marking the facial features and eyelashes and then an alarm sounds to wake up as well as a message sent to a number stored in a code saying that someone is asleep.

But sometimes, some points and observations made by the system are not true enough. Therefore, to provide data and other perspectives on the existing problem, in order to improve its use and increase the solution, thisproject has been carried out.

## II. LITERATURE REVIEW

**Paper 1:** In paper 1, they proposed a new way to compare driver drowsiness with multiple eyelid movements based on cognitive-compulsive technique (PLSR), to address the problem of tight collinear relationships between the eyelid movement features and, thus, predict sleep patterns. The predicted accuracy and durability of the model developed thus are guaranteed, indicating that it offers a new way of combining multiple factors to improve our ability to detect and predict drowsiness.

**Paper 2:** In paper 2, they suggested that the driver's eye movements could detect sleepiness under the hood or test conditions. The latest eye tracking performance based on measures of car fatigue is being tested. These steps are calculated mathematically and categorically based on a large 90-hour database of real road driving. The results show that sleep detection works well for some drivers as long as blink vision is effective. Despite the proposed improvements, however, there are still problems with poor lighting conditions and with people wearing glasses. In short, camera-based sleep measurements provide an important contribution to the sleep indicator, but they are not reliable enough to be the only indicator.

**Paper 3:** In paper 3, they suggested that in order to reduce the number of such accidents, a modified driver assistanceprogram module should be introduced, which addresses automatic driver sleep deprivation and driver disturbance. Artificial intelligence algorithms are used to process visual information to detect, track and analyze both the driver's face and eyes to calculate indications of drowsiness and disturbance. This real-time system works during the night conditions due to the proven infrared lighting system. Finally, examples of different driver images taken from a real night car are displayed to validate the proposed algorithms.

**Paper 4**: In paper 4, they proposed a new method of analyzing the facial expression of the driver through Hidden Markov Model (HMM) based dynamic modeling to detect drowsiness. They have implemented the algorithm using a simulated driving setup. Experimental results verified the effectiveness of the proposed method.

**Paper 5:** In paper 5, introduced visual analysis of eye state and head pose (HP) to monitor vehicle driver monitoring.Many of the available methods of visual detection of careless driving patterns depend on blindfolds or head angles to determine the driver's drowsiness or level of disturbance. The proposed scheme utilizes visual features such as optical indicator (EI), student performance (PA), and HP to extract important information from the driver's negligence. The vector support machine (SVM) divides the sequence of video segments into warning or non-warning driving events. Test results show that the proposed scheme provides high-level accuracy with acceptably low errors and false alarms for people of different races and genders in real-world driving situations.

***Table: Existing papers have used the following techniques***

| PAPERS | TECHNIQUES USED |
|---|---|
| A Survey on Drowsiness DetectionTechniques | • SVM<br>• CNN<br>• ANN |
| Survey on Vehicle Drowsiness DetectionSystem | • Drowsiness<br>• Haar Cascade Classifier<br>• Hough Circle<br>• Image Processing<br>• Real Time Drowsiness Detection<br>• OpenCV |
| Drowsy Driver Detection and Alert System | • Open CV<br>• Drowsy<br>• Machine learning<br>• Image processing<br>• Eye landmarks |
| Drowsiness Detection using ComputerVision Technology | • Background Elimination<br>• Face Detection<br>• Eye Detection<br>• Detection of Lip Movement |
| IoT-Based Smart Alert System for DrowsyDriver Detection | • Drowsy eyes and face recognition<br>• Collision detection<br>• Prediction of drowsiness |

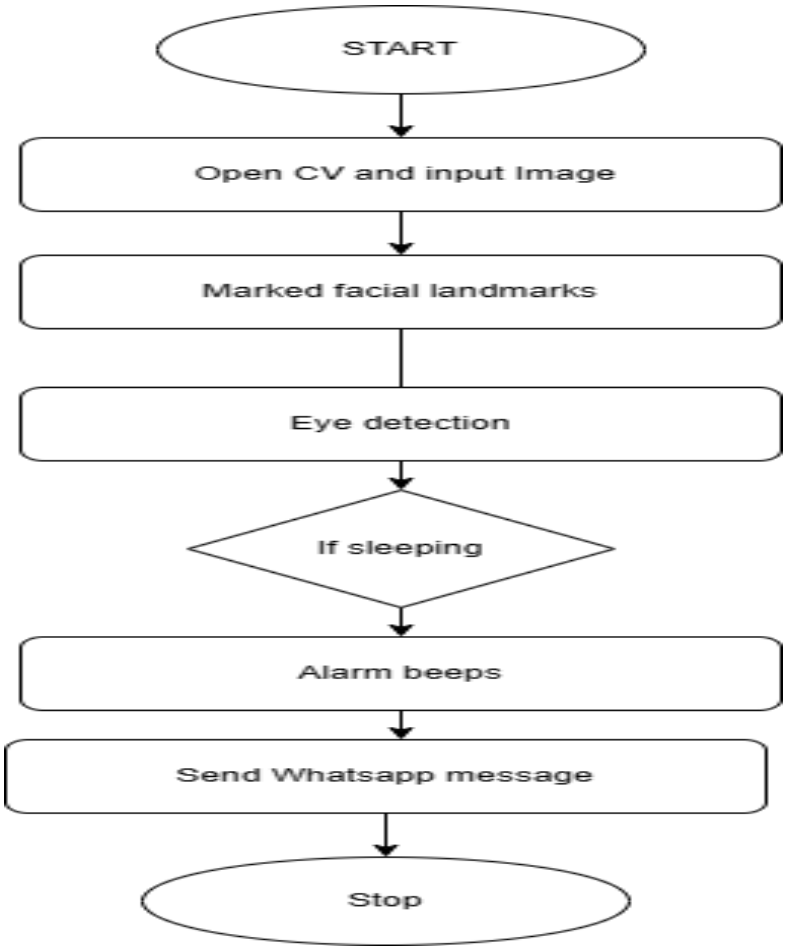
## III. PROPOSED WORK

In this program we create an app that detects if the person in front of the camera is working, drowsy or asleep. When a person is drowsy or drowsy it is manifested by marking facial features and eyelashes and then an alarm goes off to wake him up as well as a message sent to a number stored in a code saying that someone is asleep. When a person opens his eyes again the alarm stops and continues when he falls asleep again. The proposed system overcomes problems in the existing system.

The objectives of proposed system are :

- ☐ To identify whether the subject is active, drowsy or sleeping.
- ☐ Beeps alarm if the subject is sleeping.
- ☐ Send a whatsapp message saying the person is sleeping.
- ☐ Alarm stops as the subject becomes active

***Flowchart:***



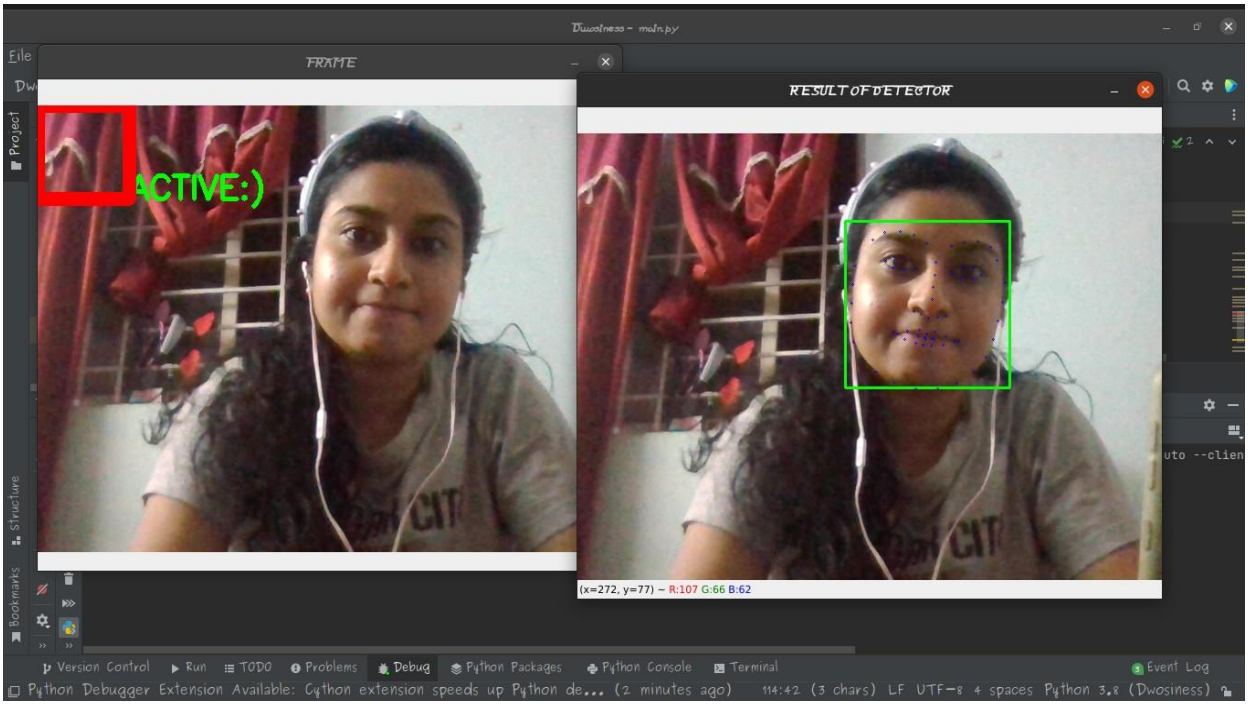
## IV. METHODOLOGY

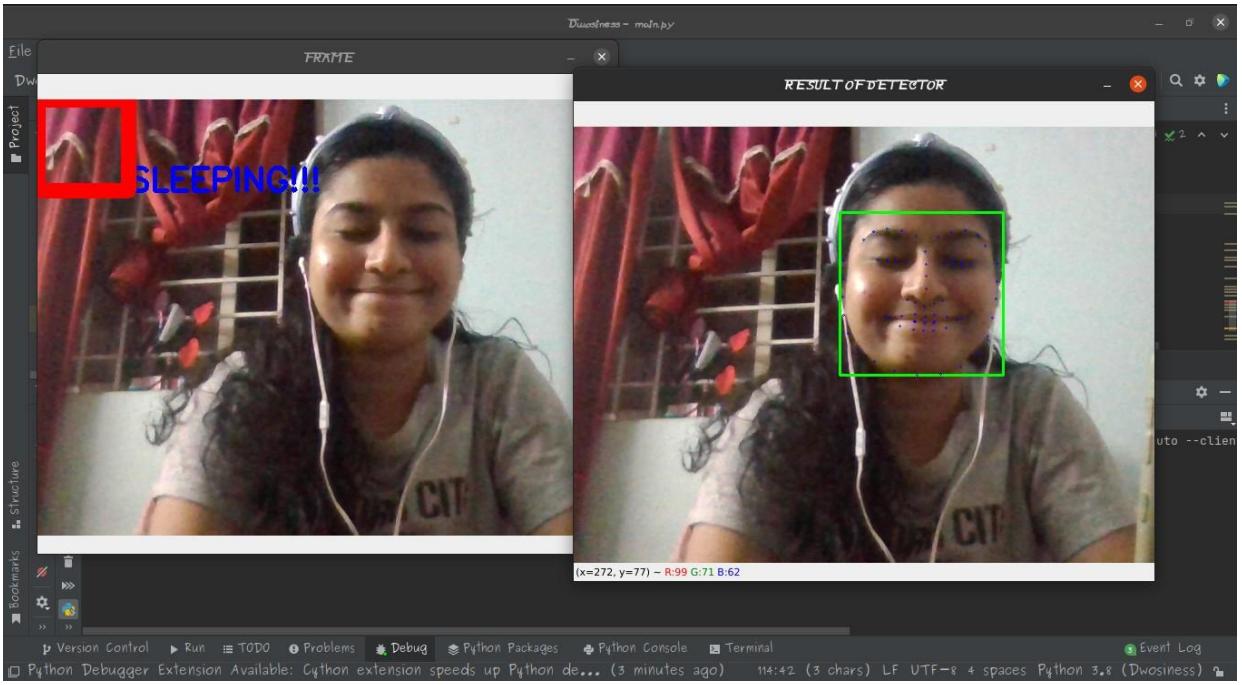
### *Convolutional Neural Network (CNN)*

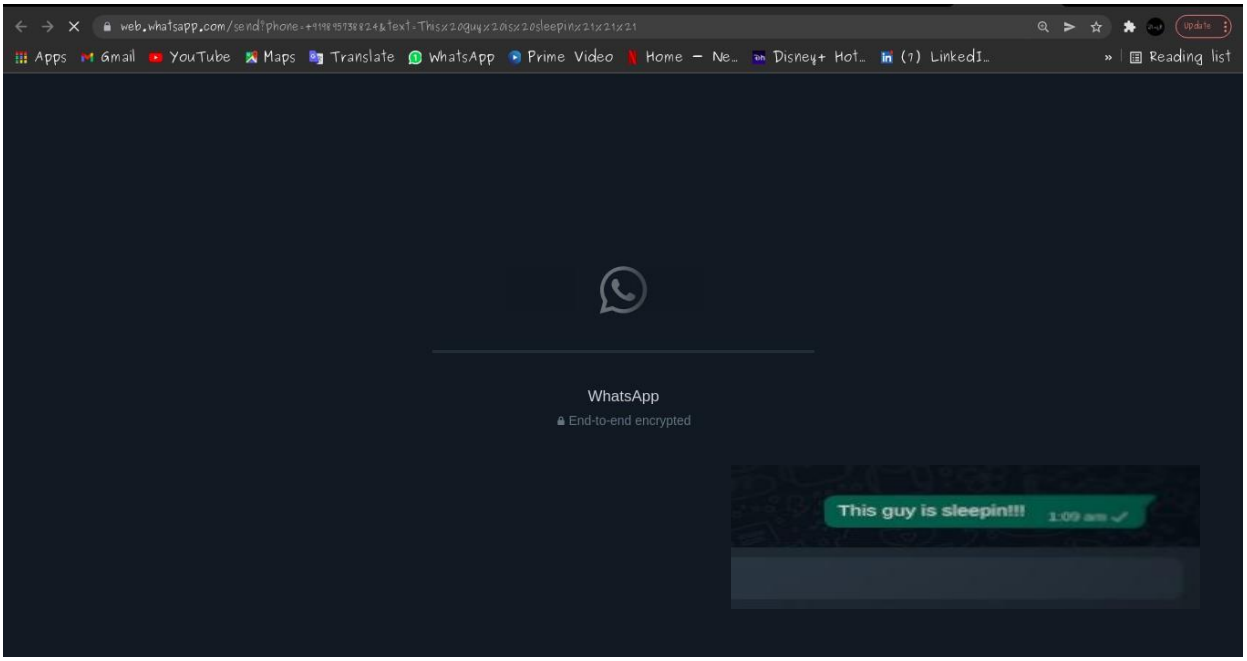
CNN is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet. It usesa special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.

The Convolutional neural network (CNN), a class of artificial intelligence networks that is well-known for its various forms of computer vision, is attracting interest in all different fields, including radiology. CNN is designed to automatically and automatically study the layout of features by distributing backwards using multiple building blocks, such as convolution layers, integration layers, and fully integrated layers. This review article provides an overview of CNN's core concepts and its application to various radiological applications, and discusses its challenges and future directions in the field of radiology. Two challenges to using CNN in radiological operations, a small data set and overload, will also be included in this article, as well as mitigation strategies. Familiarity with the ideas and benefits, as well as the limitations, of CNN is essential to harnessing its power in diagnostic radiology, with the aim of increasing the work of radiologists and improving patient care.

## V. IMPLEMENTATION




FRAME
SLEEPING!!!
RESULT OF DETECTOR
(x=272, y=77) – R:99 G:71 B:62


WhatsApp
End-to-end encrypted
This guy is sleepin!!!

## VI. CONCLUSION

The program aims to build an efficient sleep detection system using face recognition and the formation of 68 parameters. The proposed system will be able to detect the subject's state of sleep with a web camera and sound an alarm when the subject is temporarily drowsy and a message is sent via the whats app to the number already coded. This may reduce the number of accidents, students may actively participate in class and employees may be better at their job.

## VII. FUTURE ENHANCEMENT

In ongoing work, we will try to add additional features to the system such as the ECG and EEG tester, verbal attention  to see if the subject is yawning or not, Instant messaging and calls are no longer available live when the subject is in driving mode. . In the future our app overcomes mistakes when it happens and discovers new features offered to employees for easy and convenient use.

## VIII. REFERENCES

**1**. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=621668

**2**. https://ieeexplore.ieee.org/document/6518125

**3.** https://ieeexplore.ieee.org/document/6907440?arnumber=6907440

**4.** https://ieeexplore.ieee.org/document/6856149

**5.** http://www.ijirset.com/upload/december/21-A%20Dedicated%20System.pdf

# CURRENCY RECOGNITION SYSTEM USING IMAGE PROCESSING

*Harshitha M*

***Abstract***

*All currencies around the world look totally different from each other. For instance the size of the paper is different, the same as the color and pattern. The staffs who work at places like money exchange offices have to distinguish between different types of currencies and that is not an easy job. They have to remember the symbol of each currency. This may result into wrong recognition, so they need an efficient and foolproof system to aid in their work.*

***Keywords*:: *currency_recognition.***

## I. INTRODUCTION

The aim of our system is to help people who need to recognize different currencies, and work with convenience and efficiency. With development of modern banking services, automatic methods for paper currency recognition become important in many applications such as vending machines. It is verydifficult to count different denomination notes in a bunch. This project proposes an image processing technique for paper currency recognition and conversion. The extracted region of interest can be used with Pattern Recognition .The image processing approach is discussed to detect the features of paper currency. Image Processing involves changing the nature of an image in order to improve its pictorial informationfor human interpretation. There are various techniques for currency recognition that involve texture, pattern or colour based. We use digital image processing techniques to find region of interest, The proposed system will work on two images, one is original image of the paper currency and other is the test image on which verification is to be performed Template matching is often used as a simple method to classify currency.

## II. LITERATURE REVIEW

**Paper 1**: In paper 1, here objective is to propose a system for automated currency recognition using image processing techniques.

The proposed method can be used for recognizing both the country or origin as well as the denomination or value of a given banknote. This method works by first identifying the country of origin using certain predefined areas of interest, and then extracting the denomination value using characteristics such as size, color, or text on the note, depending on how much the notes within the same country differ.

**Paper 2**: In paper 2, The proposed system gives an approach to verify the Indian currency notes. Verification of currency note is done by the concepts of image processing. This article describes extraction of various features of Indian currency notes. MATLAB software is used to extract the features of the note. The proposed system has got advantages like simplicity and high performance speed. The result will predict whether the currency note is fake or not.

**Paper 3**: In paper 3, The proposed system is based on image processing and makes the process robust and automatic. This system is based on our knowledge about computer science technologies like Digital Image Processing, python and also a small step to implement in a system that is most important for industrial development. The system extracts the information from image based on features such as colour and texture. With the help of these features the system determines currency name and denomination of the currency note. After this the exchange rate of currency will be extracted from internet with the help ofonline exchange rate api url. At the end output displayed on the screen.

**Paper 4**: In paper 4, There are 50 currencies all over the world, with each of them looking totally different. For e.g. the size of the paper may vary, and also color and pattern may different. People have to remember the features of each currency. This may cause some problems (e.g.wrong recognition),so people need an efficient and exact system to help their work. The aim of our proposed system is to help people who need to recognize their currency, and work with convenience and efficiency. For bank staffs, there is a Currency Sorting Machine which is an electronic device which helps them to recognize currencies. The main working processes of Currency Sorting Machine are image acquisition and recognitions.

**Paper 5**: In paper 5, All currencies around the world look totally different from each other. For instance the size of the paper is different, the same as the colour and pattern. The staffs who work at places like money exchange offices have to distinguish between different types of currencies and that is not an easy job. They have to remember the symbol of each currency. This may result into wrong recognition, so they need an efficient and foolproof system to aid in their work. . The aim of our system is to help people who need to recognize different currencies, and work with convenience and efficiency. With development of modern banking services, automatic methods for paper currency recognition become important in many applications such as vending machines.

It is very difficult to count different denomination notes in a bunch. This project proposes an image processing technique for paper currency recognition and conversion. The extracted region of interest (ROI) can be used with Pattern Recognition and Neural Networks matching technique.

**Table: Existing papers have used the following techniques.**

| PAPERS | TECHNIQUES USED |
|---|---|
| Currency Recognition System Using image processing | • K-means clustering<br>• Template Matching |
| Currency Recognition System Using image processing | • Template Matching |
| Currency Recognition System Using image processing | • Template Matching |
| Currency Recognition System Using image processing | • Pattern Recognition Technique |
| Currency Recognition System Using image processing | • Template Matching |

## III. PROPOSED WORK

The proposed system will work on two images, one is original image of the paper currency and other is the test image on which verification is to be performed. In this project we have built a system that can detect currency. The system uses, ORB algorithm. First it takes the input of the given image and preprocess the given image and converts the RGB image into the gray scale image. The extracted features can be used for recognition, classification and retrieval of currency notes.

## Workflow

**Step 1:** Image of paper currency will be acquired by simple scanner or digital camera.

**Step 2:** The image acquired is RGB image and then it will be converted into gray scale.

**Step 3:** Edge detection of the whole gray scale image will be performed

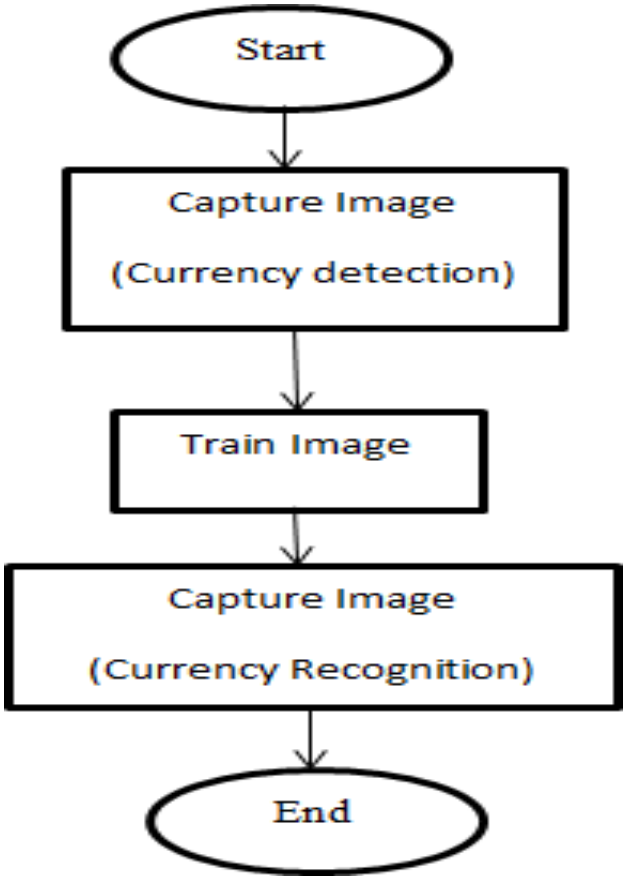
**Step 4:** After detecting edges, the characteristics of the paper currency will besegmented.

**Step 5:** After segmentation, the characteristics of the paper currency will be extracted.

**Step 6:** The characteristics of test image are compared with the original pre-storedimage in the system.

**Step 7:** If it matches then the currency is genuine otherwise counterfeit.

**Flowchart**




## IV. METHODOLOGY

**Feature matching using ORB algorithm in Python-OpenCV**

ORB is a fusion of FAST keypoint detector and BRIEF descriptor with some added features to improve theperformance. FAST is Features from Accelerated Segment Test used to detect features from the provided image.

It also uses a pyramid to produce multiscale-features. Now it doesn't compute the orientation and descriptors for the features, so this is where BRIEF comes in the role.

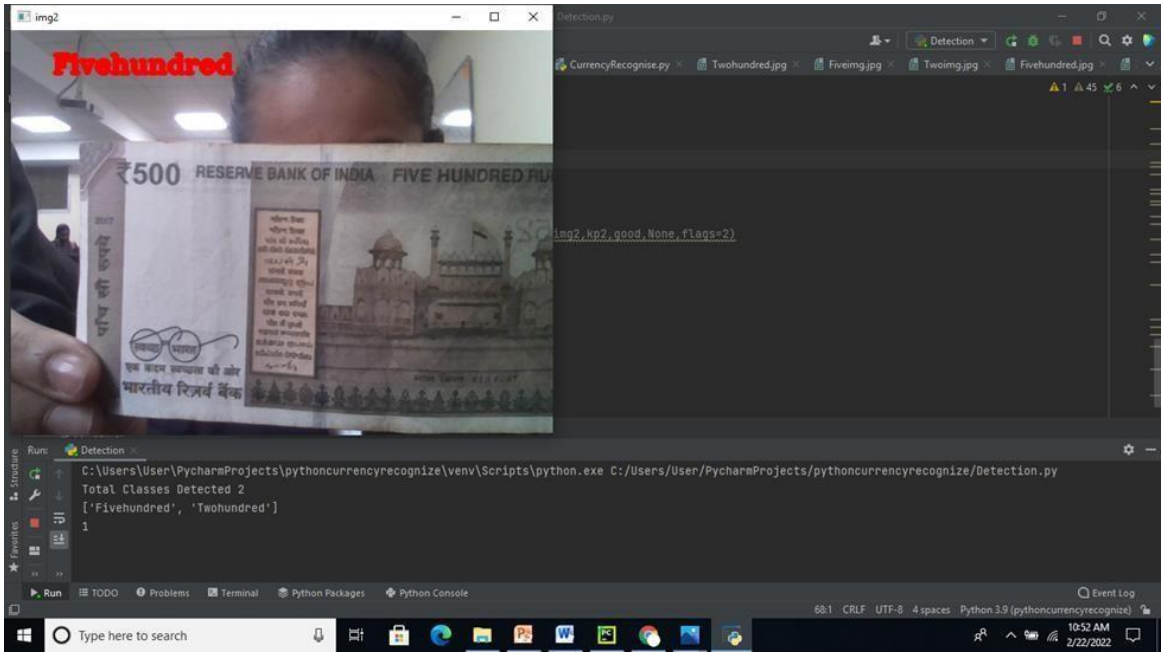
ORB uses BRIEF descriptors but as the BRIEF performs poorly with rotation. So what ORB does is to rotate the BRIEF according to the orientation of key points. Using the orientation of the patch, its rotation matrix is found and rotates the BRIEF to get the rotated version.

**Brute Force Matcher**

Brute Force Matcher is used for matching the features of the first image with another image. It takes one descriptor of first image and matches to all the descriptors of the second image and then it goes to the second descriptor of first image and matches to all the descriptor of the second image and so on.

## V. IMPLEMENTATION






## VI. CONCLUSION

By using digital image processing, analysis of Currency image is more accurate as well as this method is efficient in terms of cost and time consuming compared to existing techniques. Day by day research work is increasing in this field and various image processing techniques are implemented in order to get more accurate result. The proposed system is worked effectively for extracting feature ofIndian currency images.

## VII. REFERENCES

**1.** https://ieeexplore.ieee.org/document/8284300

**2.** https://www.researchgate.net/publication/323056402_Currency_recognition_system_using_i mage-processing

**3.** https://www.ijser.org/researchpaper/Currency-Recognition-System-Using-Image- Processing.pdf

**4.** https://www.irjet.net/archives/V4/i3/IRJET-V4I3658.pdf

**5.** https://www.rroij.com/open-access/currency-recognition-using-image-processing.pdf

**6.** https://www.internationaljournalssrg.org/IJCSE/2020/Volume7-Issue6/IJCSE-V7I6P101.pdf

# FACE RECOGNITION ATTENDANCE SYSTEM

*Harsha P C*

***Abstract***

The Face Recognition System Project is designed to identify a person's face at school / college to mark the presence of their students. The Face Recognition System therefore uses real-time face recognition as a real- world solution that assists with day-to-day student management activities. The task is extremely difficult as removing a real-time background is still a challenge. Real-time facial recognition is also used and a quick Analysis of the Main Part is used to detect faces obtained with a high degree of accuracy. Faces are used to mark student presence. This program automatically records student presence and marks student presence. The program is therefore designed to be an effective module that incorporates facial recognition to manage student attendance records.

***Keywords::*** *face_regonition, attendance_system*

## I. INTRODUCTION

Face Recognition System is a student identification system using face recognition algorithm based on high definition vision and other computer technologies. The development of this program is aimed at achieving digital integration into the traditional system of capturing people and keeping paper records. Current styles are boring and time-consuming. Existing records can be easily used for personal recording. The normal process of existence and current biometric systems are at stake in proxies. This paper is therefore proposed to address all of these issues. The proposed system allows the use of the HOG (Histogram Of Oriented Grades) algorithm. After face recognition, presence reports will be created and stored in excel format. The system is tested under various conditions such as light, head movement, distance difference between the student and the cameras. The proposed System proved to be efficient and robust for classroom entry without the use of time and manual labor. The built-in system is economical and requires minimal installation.

## II. LITERATURE REVIEW

**Paper 1**: In paper 1,This approach creates a student-centered approach based on facial recognition.

This method,like the others, begins by inserting an image or uploaded to a memory or camera. Both LBP and PCA feature releases the methods are studied in detail and computerized in this way for comparison. LBPis was developed in this way to reduce the illumination effect. The advanced integration algorithm for LBP andPCA is also designed for independent selection to increase accuracy.

**Paper 2:** In paper 2, In this automated project design project the project uses the LabVIEW face recognitionalgorithm software. The program requires a video recording device and a valid LabVIEW algorithm used successfully. It sees the face and marks its presence correctly. This program will do avoid unnecessary wastage of classroom time that is often wasted on pronunciation.

**Paper 3**: In paper 3, The project uses a machine learning algorithm commonly used in computer visualization. In thisproject you are working on a Haar Cascade classifier used to extract facial features. This is basically the discovery of an object. An algorithm used to identify the face in a photo or real-time video. The model created for this training is available on OpenCV .These models are eye-catching, face detection etc. The mainfeatures of the image make it easy to find edges or lines in the image, or to select areas where there is a sudden change. thickness of pixels.

**Paper 4:** In paper 4, In this paper, the concept of dual technology is Student mobility and the response process is initiated by machine learning. The system automatically detects student performance and keeps student records such as school attendance and their response to subjects such as Science, English, etc. The presence of a student can therefore be made by observation. At first glance, attendance details and details aboutstudent marks are available as a response.

**Paper 5:** In paper 5, The system will be presented with a photo by camera or memory and should automatically detect the face number on it. After identifying the face, the system should cut the face from the image and save it in memory to see the image to be made in the second step. The system should be able to automaticallycalculate the number of faces found in the image. The second step will be part of the alert where the system will be able to identify faces from the database andcompare the input data from the first step. The software will be used on this face filtering system automatically. The software will work seamlessly to facilitate interoperability between multiple tasks as needed.Because the system has two steps, the second phase of the program will involve image training in a database that will be used for recognition. Because the system has two steps, the second phase of the program will involve image training in a database that will be used for recognition.

**Table: Existing papers have used the following techniques**

| PAPERS | TECHNIQUES USED |
|---|---|
| Face Recognition based Automated Student Attendance System | • SVM<br>• LBP<br>• PCA |
| Face Detection And Recognition Student Attendance System | • LabVIEW<br>• Vision Acquisition<br>• Pattern Extraction<br>• Feature extraction<br>• OpenCV |
| A Counterpart Approach to Attendancesystem using Machine Learning Techniques | • OpenCV<br>• Haar Cascade |
| Automatic Attendance By Face Recognition System | • Face Detection<br>• SVM<br>• Image Acquisition |
| Face Recognition System | • LBP<br>• Image Extraction<br>• Open CV |

## III. PROPOSED WORK

In the proposed system, student photos are stored in. Then, on a case-by-case basis, a sample of two real-life images and one false image was observed. After comparing the accompanying images they are said to be realand the real value is printed. If the comparative images are not the same then a false value is displayed.

In the Project Navigation section the student photos are already stored on the website, and during school hours the student's face is filmed with a webcam and a similar view of the website. If the same is said to be true it means that the student's name is displayed on the screen.

During detection if the same is true then attendees are automatically marked on the excel sheet.

**Workflow:**

Step 1: Save the image to a website.

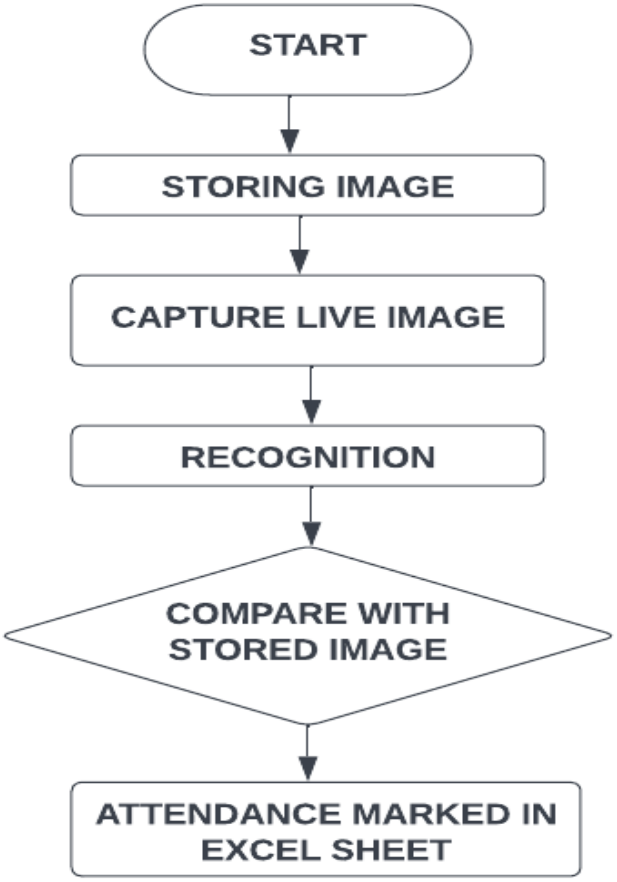
Step 2: Captures the face with a webcam.

Step 3: The awareness process begins

Step 4: Compare the captured image with the saved website image.

Step 5: Add Attendees to Excel Sheet.

**Flowchart:**




## IV. METHODOLOGY

**Object Gradient Histogram (HOD):**

HOG, or Histogram of Oriented Gradients, is a descriptive feature often used to extract objects from image data. It is often used in computer vision tasks to find an object.

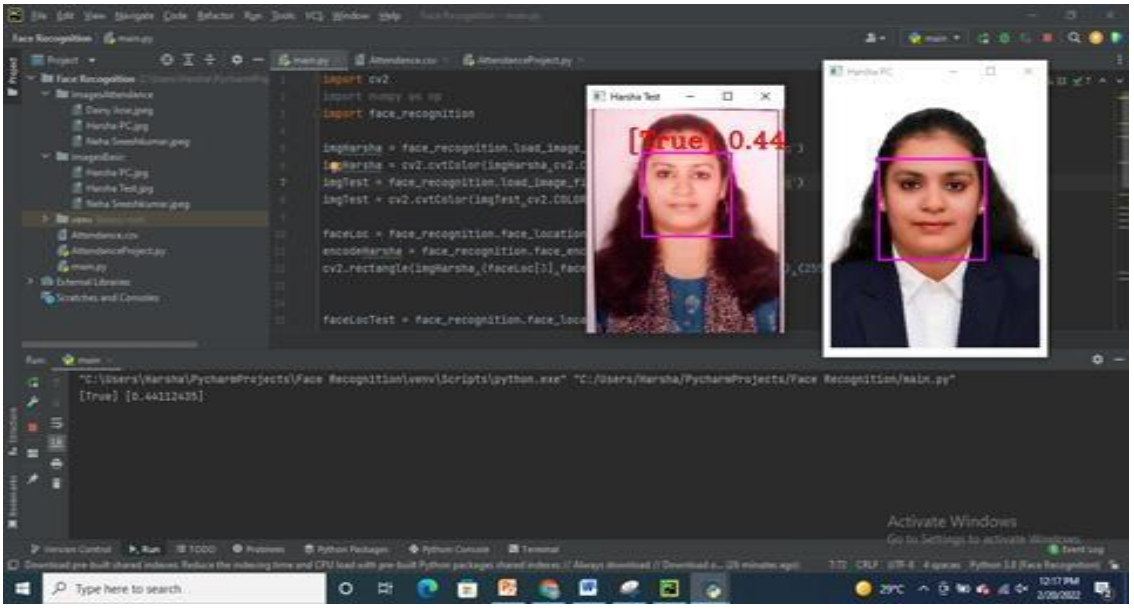
The HOG adjective focuses on the structure orstate of an object.
In the case of edge features, we only show whether the pixel is horizontal or not. HOG can also give negativeresults. This is done by removing the gradient and setting (or not specifying the size and direction) of the edges.

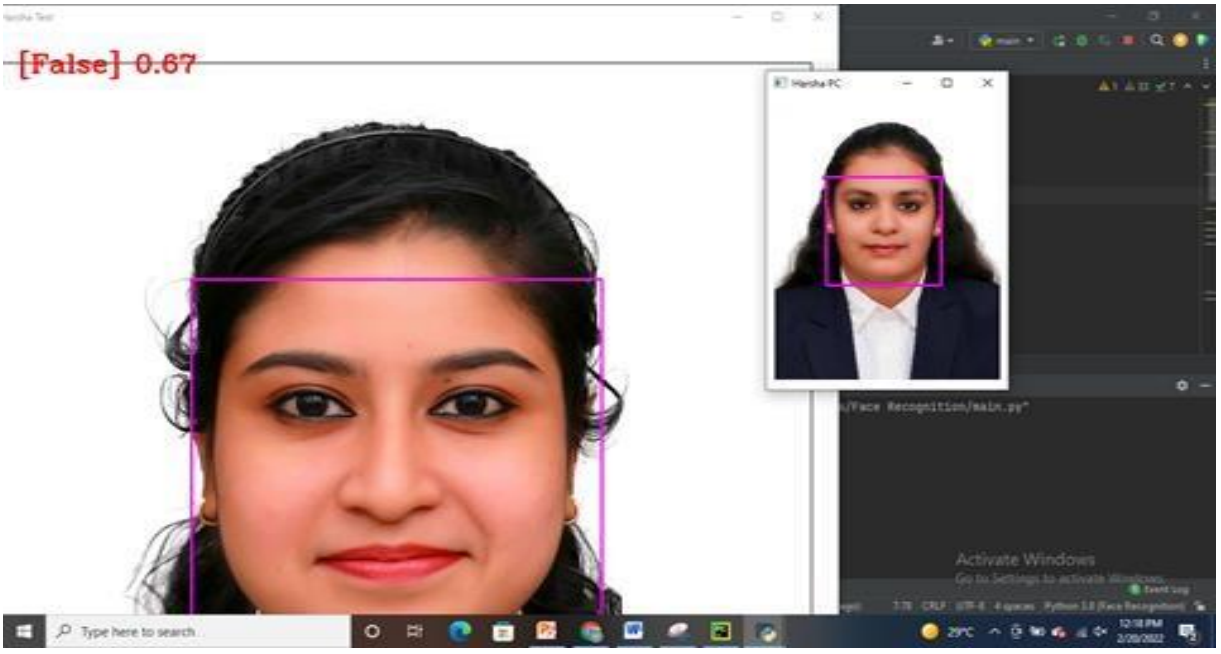
Additionally, these guidelines are categorized into 'local' categories. This means that the whole image isdivided into smaller regions and in each region, gradients and shapes are calculated.

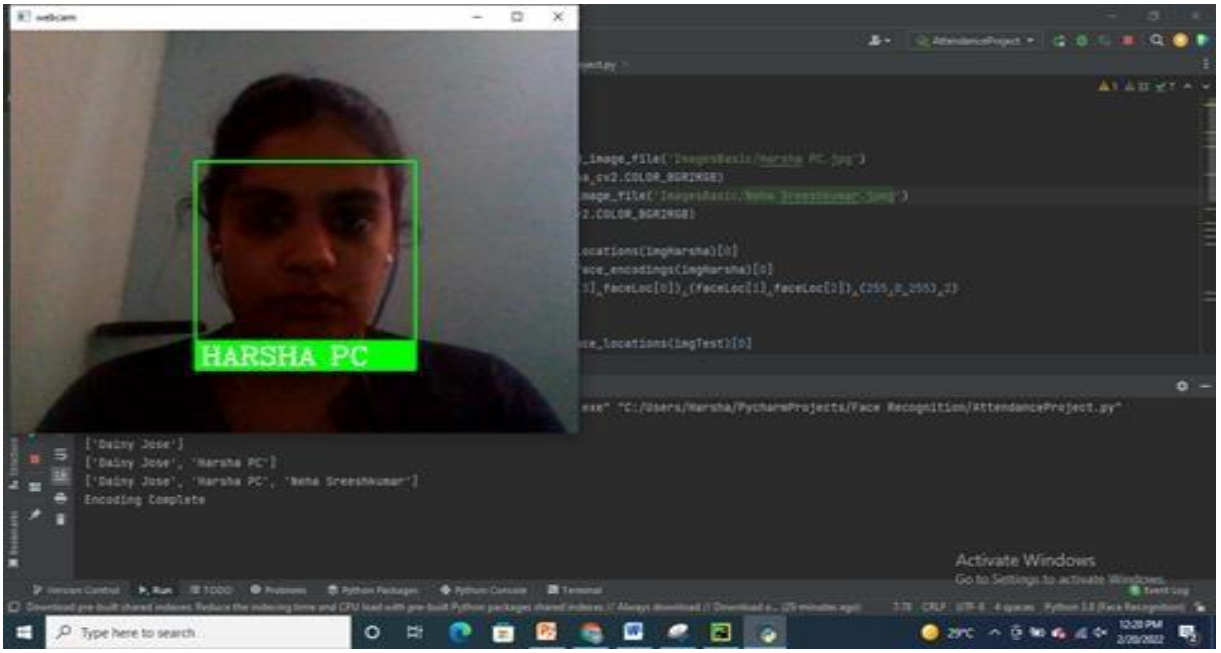
Eventually the HOG will produce a Histogram for each of these regions separately. Histograms are createdusing gradients and pixel values, hence the name 'Histogram of Oriented Gradients'. The HOG feature description calculates the appearance of gradient shapes in local parts of the image.

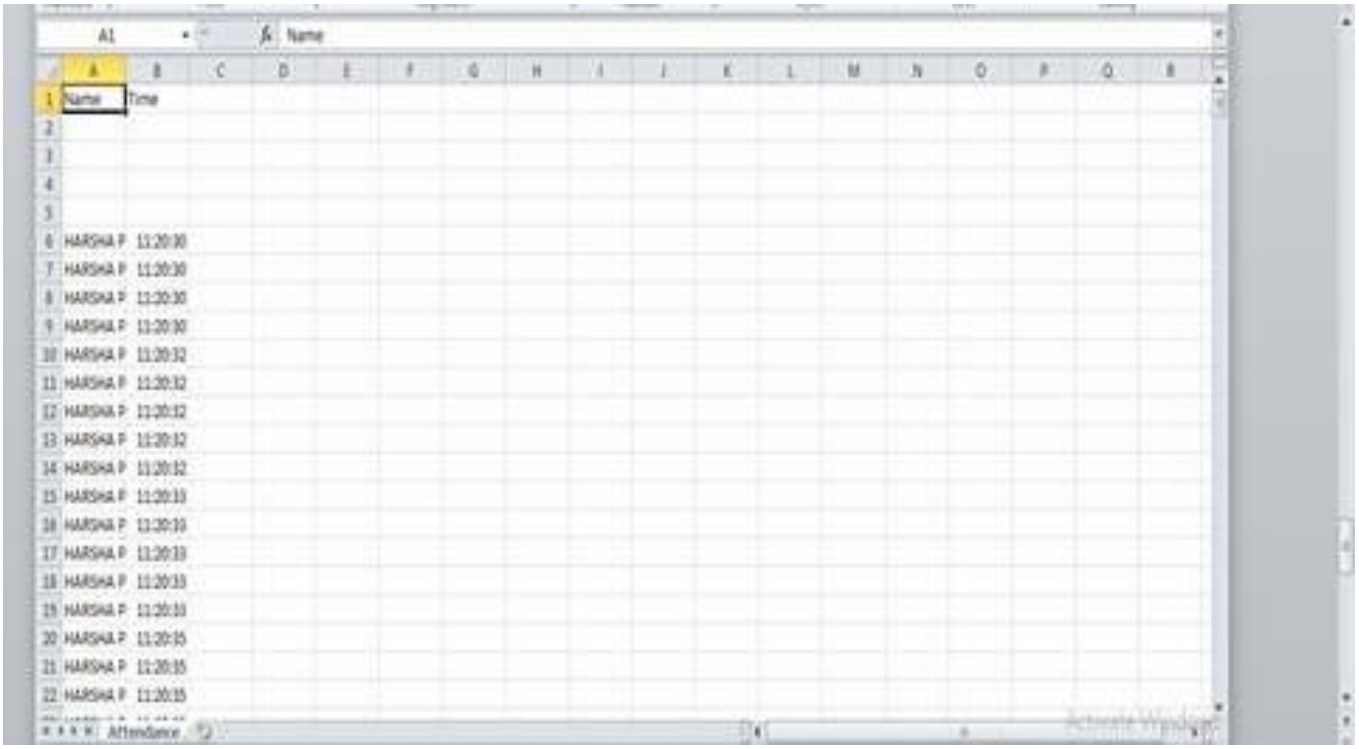
The HOG adjective has a few important advantages over other adjectives. As it works on local cells, it doesnot change geometry and image transformation, except for the shape of an object. Such changes will only occur in regions of large spaces.

## V. IMPLEMENTATION




[False] 0.67
Harsha PC
Activate Windows


HARSHA PC
['Dainy Jose']
['Dainy Jose', 'Harsha PC']
['Dainy Jose', 'Harsha PC', 'Neha Sreeshkumar']
Encoding Complete
Activate Windows




## VI. CONCLUSION

The presence of a face recognition system is very efficient and it is easier to keep records in the system than in papers or books. Reduce teacher time and keep classes of different travel pages at different times and days.

This program can be used to have better results regarding presence management. The system will save time, reduce the amount of work that managers have to do and will replace stationery with electrical tools and reduce the amount of staff needed for this purpose. So the system with the expected results has been developed but thereis still room for improvement

## VII. FUTURE ENHANCEMENT

In the ongoing work, we aim to improve the effectiveness of facial recognition through communication between our system, students and teachers. On the other hand, our system can be improved by integrating video and videostreaming services, providing in-depth applications in the field of higher education, curriculum management (CMS) and skills development support (FD). and allowed them to change their positions.

## VIII. REFERENCES

**1.** https://www.pmu.edu.sa/attachments/academics/pdf/udp/coe/dept/ee/face_detection_system_ report.pdf

**2**. https://nevonprojects.com/face-recognition-attendance-system

**3.** https://www.slideshare.net/NiteshDubey31/project-synopsis-on-face-recognition-in-e-attendance

**4.** https://www.slideshare.net/ShreyaDandavate/face-recognition-attendance-system-96913577

**5.** https://www.questjournals.org/jses/papers/Vol5-issue-2/D05021829.pdf

# WINE QUALITY PREDICTION

*Sudagani Sai Sarika*

***Abstract***

*Wine manufacturing is a difficult task since taste is the least predicted of the human senses. A good wine quality can be predicted and is very useful in the certification phase, since currently the taste analysis is performed by human tasters, being clear it will be a subjective approach. An automated predictive system can be integrated into a decision support system, helping the speed and quality of the performance. Furthermore, a feature selection process can help to analyze the impact of the analytical tests. If it is concluded that several input variables are highly relevant to predict the wine quality, since in the production process some variables can be controlled, this information can be used to improve the wine quality. Classification models used here are 1) Random Forest 2) Stochastic Gradient Descent 3) SVC 4) Logistic Regression.*

***Keywords ::*** *Regression, Classification, Support Vector Classifier, Ensemble methods, Random Forest, KNN classifier, Naïve Bayes.*

## Problem Statement

To predict the wine quality all the features are not necessary. Therefore, this research aims to predict what all features of the wine are important to get the results by implementing some of the machine learning classification algorithms and ensemble methods.

## I. INTRODUCTION

In recent years, there is a gradual increase in the wine consumption as it has been proven that wine has a good correlation with the heart rate variability. With the increase in the consumption wine manufacturing industries are looking forward for alternatives to produce good quality wine at low cost. Different wines have different importance. Although most of the chemicals are same for different type of wine based on the chemical tests, the quantity of each chemical have different level of concentration for different type of wine.

These days it is really important to classify different wine for quality assurance. In the past due to lack of technological resources it become difficult for most of the industries to classify the wines based on the chemical analyses as it takes lot of time and also need more money.

These days with the help of the machine learning techniques it is possible to classify the wines as well as it is possible to figure out the importance of each chemical analyses parameters in the wine and which one to ignore for reduction of cost. The performance comparison with different feature sets will also help to classify it in a more different way. In this paper machine learning approach is proposed by considering based feature selection considering the classifiers, linear classifiers and probabilistic classifiers to predict the quality in wine.

## 2. MATERIAL AND METHODOLOGIES

We discuss our suggested framework in the following sections, followed by descriptions of algorithms, datasets, and performance evaluation criteria.

### 2.1. Proposed framework

In a proposed system, an intelligent approach is projected by considering algorithms based on feature selection yet as simulated annealing based on feature selection considering the nonlinear classifiers, linear classifiers, and probabilistic classifiers to predict the quality in red wine yet as the white wine. Proposed with Random forest and SVM Classifiers gives the best result compared to the existing model.

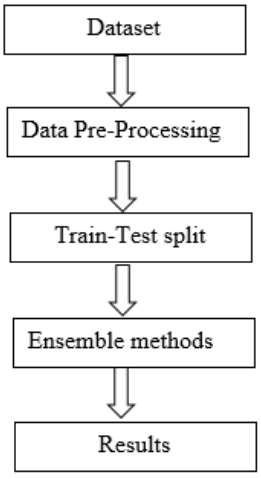



Figure 1.Workflow of the proposed system

### 2.2. Pre-processing

Before being utilized as an input for training the models, the corpus gathered from the Internet is preprocessed. Unwanted factors such as authors, date posted, URL, and category are removed from the articles. Articles with no body text or a body text of less than 20 words are also eliminated. For consistency of format and structure, multicolumn articles are converted to single column articles. To achieve consistency of format and structure, these actions are applied to all datasets.

Different hyper parameters are used to train the learning algorithms to obtain maximum accuracy for a particular dataset while maintaining an ideal balance of variance and bias. Each model is optimized for the optimum result by training it numerous times with a set of distinct parameters using a grid search.

### 2.3. Ensemble Methods

We proposed that existing ensemble techniques be combined with textual features as feature input to improve overall accuracy when determining whether an item is true or untrue. Because more than one model is trained using a particular technique to minimize the total error rate and increase the model's performance, ensemble learners have higher accuracies. The rationale underpinning ensemble modelling is similar to that which we are already used to in our daily lives, such as obtaining numerous expert opinions before making a decision in order to reduce the risk of making a bad judgement or having an unfavorable outcome.

A classification algorithm, for example, can be trained on a specific dataset with a specific set of parameters to produce a decision boundary that roughly fits the data. The output of that algorithm is influenced not only by the parameters used to train the model, but also by the type of training data used. The model may over fit and generate biased findings over unseen data if the training data contains less variance or homogenous data. As a result, methods like as cross validation are utilized to reduce the danger of overfitting. To establish numerous decision limits using randomly picked data points as training data, a variety of models can be trained on different sets of parameters. As a result, these issues can be handled and mitigated utilizing ensemble learning approaches by training many algorithms and combining their results for a near-optimal output.

#### 2.3.1. Bagging Classifier

Bagging classifier, or bootstrap aggregating, is an early ensemble method for reducing variance (overfitting) over a training set. One of the most popular bagging classifier variants is the random forest model. To reduce overall variance, the bagging model selects the class for a classification problem based on main votes measured by number of trees, but the data for each tree is selected using random sampling with replacement from the entire dataset. The bagging model, on the other hand, averages numerous estimates for regression issues.

#### 2.3.2. Boosting Classifier

Another commonly used ensemble strategy for training weak models to become strong learners is boosting. A forest of randomized trees is trained for this purpose, with the final prediction depending on the majority vote outcome from each tree.

This method uses an incremental approach to help poor learners to accurately categorize data points that are frequently misclassified. To classify a problem, identical weighted coefficients are utilized at first for all data points. In subsequent rounds, the weighted coefficients are reduced for correctly classified data points and increased for incorrectly classified data points.

Each round's succeeding tree learns to reduce the previous round's faults and improve overall accuracy by properly identifying data points that were misclassified in earlier rounds. One important drawback of boosting ensemble is that it may over fit to the training data, resulting in inaccurate predictions for unknown cases. There are a variety of boosting techniques available that can be utilized for classification and regression. Gradient Boosting methods were employed in our experiments for classification.

### 2.4. Classification Methods

#### 2.4.1. Random Forest Classifier

Random forest (RF) is a supervised learning model that is an enhanced version of decision trees (DT). RF is made up of a huge number of decision trees that work together to predict the outcome of a class, with the final prediction based on the class that obtained the most votes. Due to little correlation across trees, the error rate in random forest is low when compared to other models. Our random forest model was trained with various parameters, such as varying numbers of estimators, in a grid search to find the optimal model that can accurately predict the outcome.

#### 2.4.2. Support Vector Classifier

The SVM model is used for both non-linear and linear data. It uses a nonlinear mapping to convert the main preparing information into a higher measurement. The model searches for the linear optimum splitting hyperplane in this new measurement. A hyperplane can split the data into two classes with an appropriate nonlinear mapping to suitably high measurements and for the finding, this hyperplane SVM uses the support vectors and edges. The SVM model is a representation of the models as a point in space, the different classes are isolated by the gap to map with the aim that instances are wide as would be careful. The model can perform out a nonlinear form of classification.

#### 2.4.4. K-NN Classifier

A k-nearest-neighbor algorithm, often abbreviated k-nn, is an approach to data classification that estimates how likely a data point is to be a member of one group or the other depending on what group the data points nearest to it are in.The k-nearest-neighbor is an example of a "lazy learner" algorithm, meaning that it does not build a model using the training set until a query of the data set is performed.

## 3. DATASET

The dataset used in this research is open source and publicly available on the internet. The dataset has 12 features. Among twelve features 11 features are independent variables and one dependent variable. Two categories of wine analyzed red wine and white one.

In our analysis we have considered nearly 1000 and above samples, among that 1599 samples were red wine and 4898 samples were white wine. The dataset, which comprises the required number of features for training and testing, is available on Kaggle.

## 4. PERFORMANCE MATRIX

To evaluate algorithm performance, we used a confusion matrix. Examining the confusion matrix is a far more effective way to evaluate a classifier's performance. A confusion matrix is a table that displays how well a classification model (or "classifier") performs on a set of test data with known true values. The confusion matrix is simple in itself, but the related nomenclature might be confusing. The confusion matrix is a four-parameter tabular representation of a classification model's performance on the test set: true positive, false positive, true negative, and false negative (See Table 1).

| | | True Labels | |
|---|---|---|---|
| | | Positive (1) | Negative (0) |
| Predicted Labels | Positive (1) | True positive (TP) | False Positive (FP) |
| | Negative (0) | False negative (FN) | True negative (TN) |

Table 1

### i) Accuracy

Accuracy is a commonly used metric that represents the percentage of accurately predicted true or erroneous observations. The following equation can be used to calculate the accuracy of a model's performance:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}.$$

In most cases, a high accuracy value indicates a good model; however, because we are training a classification model in this case, an article predicted as true when it was actually false (false positive) can have negative consequences; similarly, if an article was predicted as false when it contained factual data, trust issues can arise. As a result, we employed three other metrics, precision, recall, and F1-score, to account for the wrongly classified observation.

ii) **Recall**

The total number of positive classifications out of true class is referred to as recall. It reflects the proportion of values anticipated to be true out of the total number of true values in our example.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

**iii) Precision**

Precision score, on the other hand, is the proportion of true positives to all real occurrences predicted. Precision, in this example, refers to the number of values identified as true out of all the positively predicted (true) values:

iv) **F1-Score**

The F1-score is a measure of the precision-to-recall trade-off. The harmonic mean between the two is calculated. As a result, it accounts for both false positive and false negative results. The following formula can be used to calculate F1-score

$$\text{F1} - \text{score} = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

### 4.1. Results and Discussion

Figure-2 and Figure-3 are the data visualization techniques used in the research study for better approaches. On the datasets under consideration, Table 2 highlights the accuracy of each approach. The bagging classifier clearly attain the maximum accuracy on the Data set, which is 0.93. The accuracy of the Boosting Classifier - Adaptive Boosting Classifier and Random Forest Classifier was 0.925, the K-NN Classifier and SVM gave an accuracy of 0.88 and the Naïve Bayes was 0.85.

| Classifier | Training Accuracy | Testing Accuracy |
|---|---|---|
| Random forest Classifier | 0.99 | 0.925 |
| K Neighbouring Classifier | 0.92 | 0.88 |
| Support Vector Classifier | 0.86 | 0.88 |
| AdaBoost Classifier | 0.90 | 0.925 |
| Bagging Classifier | 0.99 | 0.93 |
| Naive Bayes | 0.84 | 0.85 |

Table 2

The average accuracy of all methods over the dataset is likewise shown in Table 2. Bagging classifier is the best-performing algorithms overall (accuracy 1.0). However, accuracy score alone isn't a fair

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}.$$

way to judge a model's performance, therefore we also include recall, precision, and F1-score when evaluating learning models. Bagging classifier achieved

$$\text{F1} - \text{score} = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

the highest results in terms of average precision of 0.95. Random Forest Classifier had a precision of 0.99.

Bagging classifier have the best recall performance. Random Forest Classifier, which has a recall of 0.99, is right behind it. The algorithms performed similarly in terms of F1-score as they did in terms of precision and recall. The bagging classifier had the best F1-score of all the approaches. Even all the other approaches too.

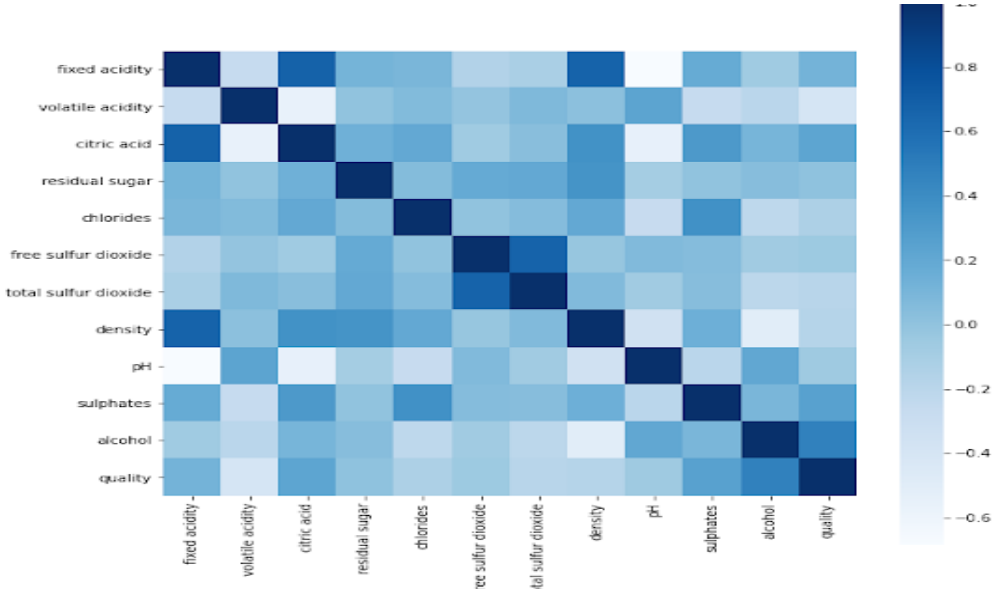



Figure 2

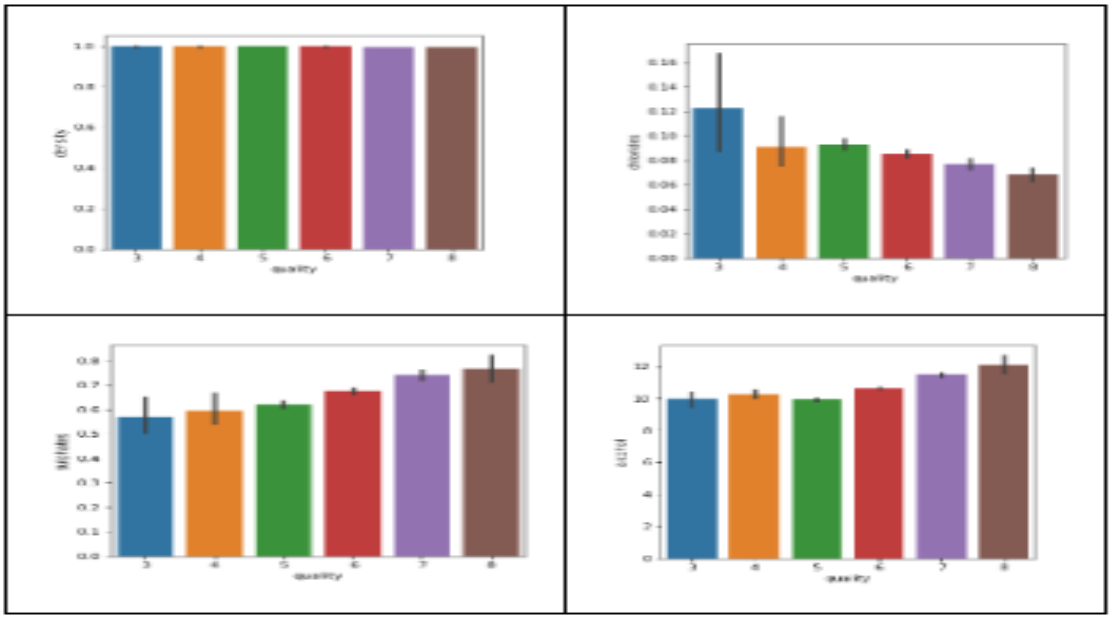


Figure 3

## 5. CONCLUSION

The work of manually classifying necessitates a thorough understanding of the domain as well as the ability to spot unimportant features for the study. The topic of classifying the suitable and important features and unimportant features using machine learning models and ensemble approaches was tackled in this study. The information we used in our research came from the Internet and consisted of data about the wine from various domains that covered most of its features.

The study's main goal is to find best quality by classifying the important and selective features of the wine.

To achieve optimal accuracy, the learning models were trained and parameter-tuned. Some models have been shown to be more accurate than others. To compare the results of each method, performance measures were employed. When compared to individual learners, ensemble learners have a higher overall score on all performance indicators.

There are numerous outstanding difficulties in the wine quality prediction that require research. Identification of essential aspects involved in the study.

In the future, broad data set may be used for experiments and other machine learning techniques may be explored for prediction of wine quality, and will expand this analysis to include feature development methods to test whether or not the model's predictive power may be increased.

## 6. REFERENCES

**1 .**https://www.kaggle.com/c/fake-news

**2.** Gupta, Y., 2018. Selection of important features and predicting wine quality using machine learning techniques. Procedia Computer Science, 125, pp.305-312.

**3.** Kumar, S., Agrawal, K. and Mandan, N., 2020, January. Red wine quality prediction using machine learning techniques. In 2020 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE.

**4.** Dahal, K.R., Dahal, J.N., Banjade, H. and Gaire, S., 2021. Prediction of wine quality using machine learning algorithms. Open Journal of Statistics, 11(2), pp.278-289.

**5.** Gupta, U., Patidar, Y., Agarwal, A. and Singh, K.P., 2020. Wine quality analysis using machine learning algorithms. In Micro-Electronics and Telecommunication Engineering (pp. 11-18). Springer, Singapore.

**6.** Frank, I.E. and Kowalski, B.R., 1984. Prediction of wine quality and geographic origin from chemical measurements by parital least-squares regression modeling. Analytica Chimica Acta, 162, pp.241-251.

**7.** Gambetta, J.M., Cozzolino, D., Bastian, S.E. and Jeffery, D.W., 2016. Towards the creation of a wine quality prediction index: Correlation of Chardonnay juice and wine compositions from different regions and quality levels. Food analytical methods, 9(10), pp.2842-2855.

**8.** Ough, C.S. and Singleton, V.L., 1968. Wine Quality Prediction from Juice Brix/Acid Ratio and Associated Compositional Changes forWhite Riesling'andCabernet Sauvignon'. American Journal of Enology and Viticulture, 19(3), pp.129-138.

**9.** Chiu, T.H.Y., Wu, C. and Chen, C.H., 2021. A Generalized Wine Quality Prediction Framework by Evolutionary Algorithms. International Journal of Interactive Multimedia & Artificial Intelligence, 6(7).

***DHII***
***Journal Dept. of MCA***
***Jyoti Nivas College Autonomous***

# FACE MASK DETECTION USING YOLOV5

*Sandhya M*

***Abstract***

*Since the infectious coronavirus disease in the new world (2020), it has become a Public health problem in India and even around the world. This pandemic is having devastating effects on societies and economies around the world. The increase in the number of COVID-19 tests gives more information about the epidemic spread, which may lead to the possibility of surrounding it to prevent further infections. Multidisciplinary efforts have been organized to slow the spread of the pandemic, wearing a face mask that prevents the transmission of droplets in the air and maintaining an appropriate physical distance between people, and reducing close contact with each other can still be beneficial in combating this pandemic. I would like to build a system that can detect faces in images and identify if the detected faces are wearing masks or not. I will firstly explore mask and no mask classification in images classification and also the count the number of faces detected in the image. I have used CNN, YOLOV5 and Tensor-flow to know the accuracy in the mask detection.*

***Keywords*:: ***masked and non-masked faces, CNN, Bounding Boxes, machine learning.*

## I. INTRODUCTION

The novel coronavirus covid-19 had brought a new normal life. India is struggling to get out of this virus attack and the government implemented lockdown for the long way. Lockdown placed a pressure on the global economy. So the government gave relaxations in lockdown. Declared by the WHO that a potential speech by maintaining distance and wearing a mask is necessary. The biggest support that the government needs after relaxation is social distancing and wearing of masks by the people. But many people are getting out without a face mask this may increase the spread of covid-19. Economic Times India has stated that the survey clearly points that people are aware but they are not wearing the mask due to some discomfort in wearing and carelessness. This may result in the easy spreading of covid-19 in public places.

The world health organization has clearly stated that until vaccines are found the wearing of masks and social distancing are key tools to reduce in spread of viruses. So it is important to make people wear masks in public places. In densely populated regions it is difficult to find the persons not wearing the face mask and warn them.

Hence we are using image processing techniques for identification of persons wearing and not wearing facemasks.

Not wearing a mask is done. The trained dataset is made by using machine learning technique which is the deciding factor of the results. The algorithm created by means of using a trained dataset will find the persons with and without wearing face mask. The yolov5 algorithm is an object detection and the annotated images can used easily and this algorithm works faster and using tensor-flow the training and accuracy can see in the graph.

## II. PROBLEM STATEMENT

Identifying the identity of masked faces is a challenging problem for the multiple image sets since the features required to accurately predict the identity of an individual is reduced from the whole face to just the eye and sometimes the forehead this task should me included in multiple face images. This study is built on existing yolov5 model architecture trained on human faces to solve the problem of identifying a person's identity when wearing a face mask and giving the count of it.

## III. RELATED WORK

Since generally, most of the papers talk about the face recognizing of wearing mask and not wearing mask. In this research paper we are focusing on detecting the wearing and not wearing face mask using YOLOv5 algorithm to help in spreading the corona virus decease. In many papers the authors have been discussed about the algorithms of PCA, SVM, Res-Net, RCNN, MobileNet and so on, but in our research work we have chosen YOLOv5 algorithm because it gets faster and easy way of finding the accuracy level. We got a good accuracy of 0.87 which is a good value of what we got.

## IV. DATASTS CHARACTERISTICS

This research of dataset is collected from humans in the loop website. The dataset named as Medical mask dataset. This dataset consists of total size of images are 6024 and the number of classes which are,

1. Face_with_mask
2. Face_with_mask_incorrect

3. Face_no_mask
4. Face_other_covering

And 16 other classes for masks, face shields, hats, scarves and other accessories. But from all these classes only two are used which is Face_with_mask and Face_no_mask for the research work.

## V. THE PROPOSED MODEL

The introduced model includes Keras, MobileNet V2, OpenCV, CNN, YOLOv5 and TensorFlow to implement a model. Keras is more powerful and user-friendly for deep learning frame work with TesorFlow because it is a backend engine it is easy to use and deploy, lightweight, and also it can be implemented for almost any kind of task in the machine learning industries. To make the work simple and easy the best to use the keras.

The MobileNet is an model developed by Google and it was released in 2017 and contributed a lot to the evolution of computer vision. V2 is still used in several applications, and works better than MobileNet V1 and the classic ConvNets. This is because it uses computationally cheaper convolutions that later versions of MobileNet, but still contains important architecture improvements over older networks.

When implementing this architecture the Preprocessing steps is similar to classic image classification pipeline. It gathers the images which each of it is contain only one object. It makes all the images in same size. The augmented dataset with Keras ImageDataGenerator. MobileNet converts all the images into NumPy arrays labels in binary notation.

A CNN have four components they are convolutional layer, the max-pooling layer, densely connected layer and output layer. The densely connected one performs the actual classification task and activates the perceptron in the output layer that corresponds to the predicted value. The conv2D and MaxPooling2D layers these two combined together for simple image classification tasks. For a real time detection the convolution operation cost more and makes slow, so Conv2D layer is faster and which colloid with MobileNet. The disadvantages is it shows less accurate. OpenCV helps in dealing with images, video and camera. It gives a great performance in small devices.

A YOLOv5 is powerful and easy to use object detection models available. We can also call YOLO as You Only Look Once and the latest version is 5. It is a novel convolutional neural network that detects objects in real-time with great accuracy.

This approaches uses a single neural network to process the entire picture, the separates it into parts and predicts bounding boxes and probabilities for each components.

The bounding boxes are weighted by expected probability. The method just looks once at the image in the sense that it makes predictions after only one forward propagation run through the neural network. It then delivers detected items after non-max suppression. Fig.1 is an architecture of YOLov5 where three mainly considered are Backbone, Neck and Head.

Backbone is the model mostly used for an extraction key features from an image. Neck is used to create feature pyramids and generalizing when it comes an object scaling. It aids in the identification of the same object in various sizes and scales. Head is responsible for the final detection step. It uses anchor boxes to construct final output vectors with class probabilities, object-ness scores and bounding boxes.

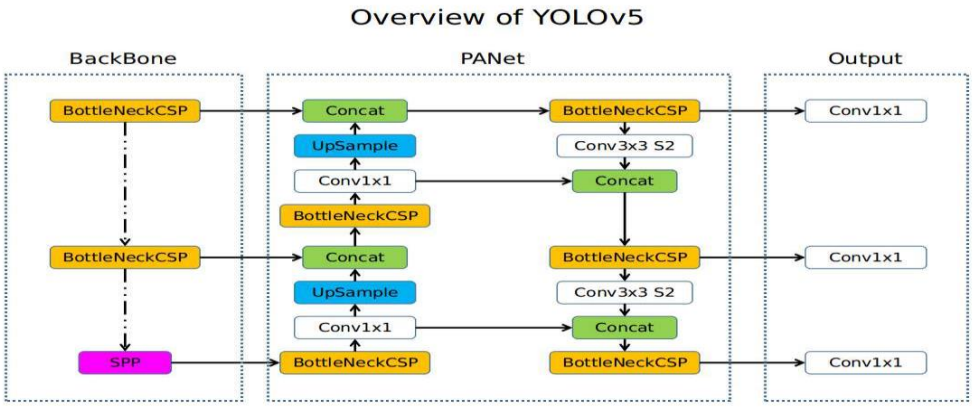



Fig. 1 An architecture of YOLOv5.

## VI. PROCESS FLOW

Data Acquisition is composed of two words namely Data and Acquisition, where data is the raw data and figures, which could be structured and unstructured and acquisition means acquiring data for the given task at hand. Data acquisition meaning is to collect data from known sources before it can be stored, cleaned, preprocessed and used for further mechanisms. It is the process of retrieving relevant business information, transforming the data into the required business form, and loading it into the designated system.Data annotation is the categorical and labelling of data for AI applications. The training data must be properly categorized and annotated for specific use case. With high-quality, human-powered data annotation.

YOLOV5 configuration is a complete set of details for the dataset that are used in training the model which is defined by YAML data file. The parameters used in the configurations are train, test and val all these locations of images are trained, tested and validated and also number of classes in the dataset are mentioned that we are using in the research. Naming of the classes in the dataset, the index of the classes in the list would be used as an identifier for class names in the code.

To train the model, since we are using small dataset and we don't have objects per images we can start with the smallest of pertained models of yolov5s to solve the problem simple and to prevent the overfitting.

The detection of the model will give the result of the image is without mask or with mask of wearing in an image.

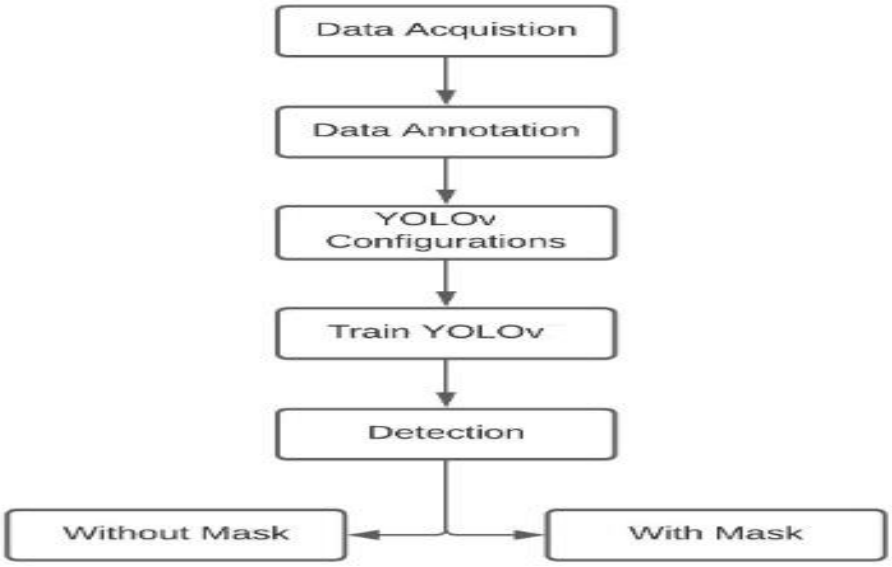



Fig. 2 Process flow of YOLOV5

## VII. METRICES

The metrics that is used in the work is mAP which is mean Average Precision for object Detection, it is one of the popular metric in measuring the accuracy of object detectors like SSD, Faster R-CNN, etc. It computes the average precision value for recall value over o to 1.

Precision: This measures how accurate is the predictions and the percentage of the predictions are correct.

Recall: This measures how good that the findings of all positives.

Mathematical definitions:

Precision = TP / ( TP + FP )

RECALL = TP / ( TP + FN )

TP = Ture positive

TN = Ture negative

FP = False positive

FN = False negative

F1 = 2 * ( Precision * recall ) / ( Precision + recall )

```
Validating runs/train/exp/weights/best.pt...
Fusing layers...
Model Summary: 213 layers, 7015519 parameters, 0 gradients, 15.8 GFLOPs
               Class     Images     Labels          P          R     mAP@.5 mAP@.5:.95: 100% 19/19 [00:06<00:00,  2.83it/s]
                 all        290       1079      0.892      0.752      0.826      0.412
                mask        290        822      0.922      0.851      0.905      0.513
              nomask        290        257      0.861      0.654      0.746       0.31
Results saved to runs/train/exp
```

Fig. 2 mAP values

## VIII. RESULT

The training dataset can plot a validation batch obtained during training and inspect the confidence score of each label. Although 0.8 we got the result which is good for the model were it supports the YOLOv5 sacrifices accuracy for good detection speed. We have tested the new unseen images manually create a new directory to upload some images and run the image to get good result.

```
Fusing layers...
Model Summary: 213 layers, 7015519 parameters, 0 gradients, 15.8 GFLOPs
image 1/15 /content/drive/MyDrive/face_data/test/173.jpg: 416x416 3 nomasks, Done. (0.028s)
image 2/15 /content/drive/MyDrive/face_data/test/27.jpg: 416x416 1 nomask, Done. (0.029s)
image 3/15 /content/drive/MyDrive/face_data/test/crowd_mask181.jpg: 256x416 12 masks, Done. (0.024s)
image 4/15 /content/drive/MyDrive/face_data/test/crowd_mask23.jpg: 224x416 1 mask, Done. (0.025s)
image 5/15 /content/drive/MyDrive/face_data/test/crowd_mask38.jpg: 288x416 1 mask, Done. (0.025s)
image 6/15 /content/drive/MyDrive/face_data/test/crowd_mask62.jpg: 320x416 7 masks, 3 nomasks, Done. (0.026s)
image 7/15 /content/drive/MyDrive/face_data/test/crowd_mask65.jpg: 256x416 12 masks, 3 nomasks, Done. (0.024s)
image 8/15 /content/drive/MyDrive/face_data/test/image_117.jpg: 288x416 1 mask, Done. (0.025s)
image 9/15 /content/drive/MyDrive/face_data/test/image_5.jpg: 320x416 1 mask, Done. (0.026s)
image 10/15 /content/drive/MyDrive/face_data/test/image_503.jpg: 288x416 1 mask, Done. (0.026s)
image 11/15 /content/drive/MyDrive/face_data/test/image_577.jpeg: 416x288 5 masks, Done. (0.023s)
image 12/15 /content/drive/MyDrive/face_data/test/image_602.jpg: 416x288 1 nomask, Done. (0.023s)
image 13/15 /content/drive/MyDrive/face_data/test/image_605.jpg: 416x288 1 mask, Done. (0.023s)
image 14/15 /content/drive/MyDrive/face_data/test/image_609.jpg: 416x288 1 nomask, Done. (0.023s)
image 15/15 /content/drive/MyDrive/face_data/test/new_116.jpg: 416x416 1 nomask, Done. (0.028s)
Speed: 0.5ms pre-process, 25.2ms inference, 1.8ms NMS per image at shape (1, 3, 416, 416)
Results saved to runs/detect/exp
15 labels saved to runs/detect/exp/labels
```

## IX. CONCLUSION

The increase in the number of COVID-19 teste gives more information about the spread, which may lead to the possibility of surroundings it to prevent further infections. Multidisciplinary efforts have been organized to slow the spread of the pandemic, wearing a face mask that prevents the transmission of spreading in the air and maintaining an appropriate physical distance between people and reducing close contact with each other can still be beneficial in combating this pandemic. For this we have come up with a result by using YOLOv5 model the best object detection and the dataset that is used are experimented on and different training on a specific dataset while testing over other datasets to prove the efficiency of the proposed model. The presented works concluded that the YOLOv5 model gives the good accuracy of a good speed.

## X. REFERENCES

**1**. Qin, B. and Li, D., 2020. Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19. Sensors, 20(18), p.5236.

**2.** Loey, M., Manogaran, G., Taha, M.H.N. and Khalifa, N.E.M., 2021. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement, 167, p.108288.

**3.** Sethi, S., Kathuria, M. and Kaushik, T., 2021. Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread. Journal of biomedical informatics, 120, p.103848.

**4.** Wu, P., Li, H., Zeng, N. and Li, F., 2022. FMD-Yolo: An efficient face mask detection method for COVID-19 prevention and control in public. Image and Vision Computing, 117, p.104341.

**5.** Mbunge, E., Simelane, S., Fashoto, S.G., Akinnuwesi, B. and Metfula, A.S., 2021. Application of deep learning and machine learning models to detect COVID-19 face masks-A review. Sustainable Operations and Computers, 2, pp.235-245. (2002) The IEEE website. [Online]. Available: http://www.ieee.org/

**6.** Ieamsaard, J., Charoensook, S.N. and Yammen, S., 2021, March. Deep learning-based face mask detection using yoloV5. In 2021 9th International Electrical Engineering Congress (iEECON) (pp. 428-431). IEEE.FLEXChip Signal Processor (MC68175/D), Motorola, 1996.

**7.** Yang, G., Feng, W., Jin, J., Lei, Q., Li, X., Gui, G. and Wang, W., 2020, December. Face mask recognition system with YOLOV5 based on image recognition. In 2020 IEEE 6th International Conference on Computer and Communications (ICCC) (pp. 1398-1404). IEEE.

**8.** Sharma, V., 2020. Face mask detection using YOLOv5 for COVID-19 (Doctoral dissertation, California State University San Marcos).Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, IEEE Std. 802.11, 19

# CREDIT CARD FRAUD DETECTION USING RANDOM FOREST

*Joice Rani*

***Abstract***

*Credit card fraud detection is presently the most frequently occurring problem in the present world. This is causes to the rise in both online transactions and e-commerce platforms. Credit card fraud happens when the card was stolen for any of the unauthorized purposes or even when the fraudster uses the credit card information for his use. In the present world, we are facing a lot of credit card problems. To detect the fraudulent activities the credit card fraud detection system was introduced. This project aims to focus mainly on machine learning algorithms. The algorithms used are Random forest. The results of the algorithms are based on accuracy. The Random forest model algorithm that has the greatest accuracy considered as the best algorithm that is used to detect the fraud.*

***Keywords::*** *credit_card,random_forest*

## I. INTRODUCTION

Whenever we hear the word Credit Card the first thing that pops in our mind is the frauds that are associated with these cards. Credit card has become an indispensable part of our lives. Although a credit card has many advantages when used in a proper manner but damages can be caused to it by many fraudulent activities as well. But in today's advanced world these frauds can be detected with a vast knowledge of machine learning algorithms. The Credit Card Anomaly Detection Problem includes modeling past credit card transactions with the ones that turned out to be fraud. After the implementation of this model, we can use it further to identify, a new transaction that is occurring as fraudulent or not. Basically, our focus here is to detect 100% fraud transactions that is being occur by minimizing the incorrect fraud classification In today's world this is the major concern, which demands the attention of the fields such as Machine Learning, Artificial Intelligence, Deep Learning etc. where the solution of this issue can be automated. Our aim is to predict the accuracy/precision of the fraud detection through different algorithms. Further this analysis can be used to implement the fraud detection model.

**Problem statement**

Credit card fraud detection is to predict fraudulent credit card transactions with the help of machine learning models. As we are moving towards the digital world cybersecurity is becoming a crucial part of our life. Credit card frauds are increasing heavily , financial loss is increasing drastically. Every year due to fraud Billions of amounts are lost. Many machine learning algorithms are implemented to detect real world credit card fraud.

## II. RELATED WORK

1. Credit card fraud detection using machine learning in this paper, As observed this whole data set have just 2 days transactions results, their only a portion of data which could have made accessible if the protude was to be utilized in the mercantile measure. It depends on artificial intelligence techniques; this code shall just extend their productivity of tempo as most data are vent onto this. Methodology that they have used for prediction is new machine/soft learning/soft learning, random forest algorithms, KNN Algorithm.
2. Credit card fraud detection using machine learning algorithms. In this paper, customers are grouped based on their transactions and extract behavioural patterns to develop a profile for every cardholder. Then different classifiers are applied on three different groups later rating scores are generated for every type of classifier. This dynamic changes in parameters lead the system to adapt to new cardholder's transaction behaviour's timely Methodology that they have used for prediction is Clustering method, Sliding-window method.
3. Machine learning approach for credit card fraud detection. in this paper, The whole data is divided into test and training sets, by applying the SVM classifier in this work. Here, input is the test and training set which provides the future values. Experimental results, evaluated the performance of proposed modal as compared to existing technique in terms of accuracy. Methodology that they have used for prediction is SVM training algorithm, design of proposed security solution.
4. Credit card fraud detection using machine learning**. i**n this paper, The work based on dataset and conclude :There are total 4000 points in their dataset, out of which 3988 points belongs to class label '0' and 12 belongs to class label '1'. This means that their model has performed well despite having very imbalanced dataset. Methodology that they have used for prediction is KNN technique, T-SNE technique, removing outliers, fraudulent.
5. Credit card fraud detection using machine learning and data Science In this paper, While the algorithm does reach over 99.6% accuracy, its precision remains only at 28% when a tenth of the data set is taken into consideration.

However, when the entire dataset is fed into the algorithm, the precision rises to 33%. This high percentage of accuracy is to be expected due to the huge imbalance between the number of valid and number of genuine transactions Methodology that they have used for prediction is applications of machine learning, data science, isolation forest algorithm, local outlier factor, automated fraud detection.

6. Credit Card Fraud Detection using Machine Learning: A Study In this paper, Among all the techniques discussed, it is found that Neural Networks detects frauds with high precision and performs best. But as they are expensive to train and can also be over trained in case of fewer nodes as in case of LSTM . So, in order to minimize this cost, we can pair neural network with some augmentation techniques like Genetic Algorithms or Artificial Immune Systems by selecting optimized weight edges and eliminating those weights that causes error. Methodology that they have used for prediction is Hidden Markov Model, Decision Trees, Logistic Regression ,Support Vector Machines (SVM), Genetic Algorithm, Artificial Neural Networks Random Forests , Bayesian Belief Network.

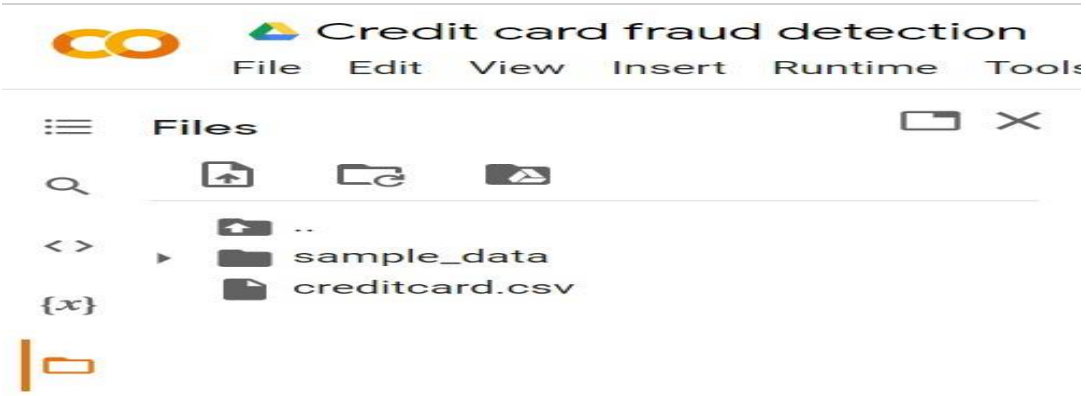
7. Credit Card Fraud Detection using Machine Learning In this paper, As the entire dataset consists of only two days transaction records, it's only a fraction of data that can be made.

## III. METHODOLOGY

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. As the name suggests, Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset. Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. the random forest combines multiple trees to predict the class of the dataset, it is possible that some decision trees may predict the correct output, while others may not. But together, all the trees predict the correct output. Therefore, below are two assumptions for a better Random forest classifier: There should be some actual values in the feature variable of the dataset so that the classifier can predict accurate results rather than a guessed result. The predictions from each tree must have very low correlations. Random Forest predicts the output of a categorical dependent variable.

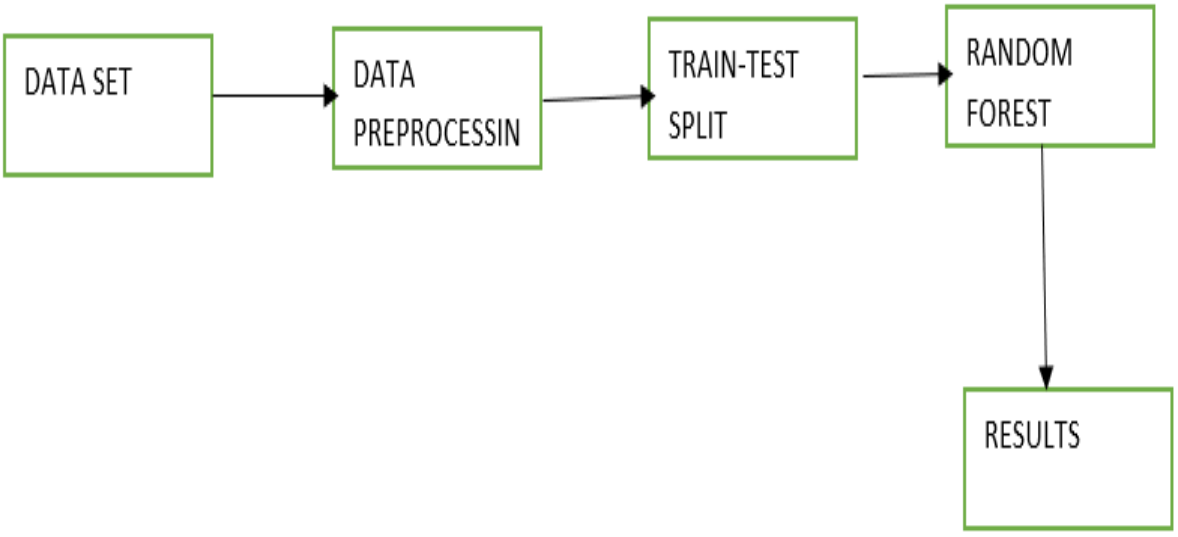
Therefore, the outcome must be a categorical or discrete value.It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.

To solve this sample Excel sheet dataset required which will be used in testing. CSV files which are being tested should be in the same directory with the extension .CSV




Dataset file Image a

**Work flow**

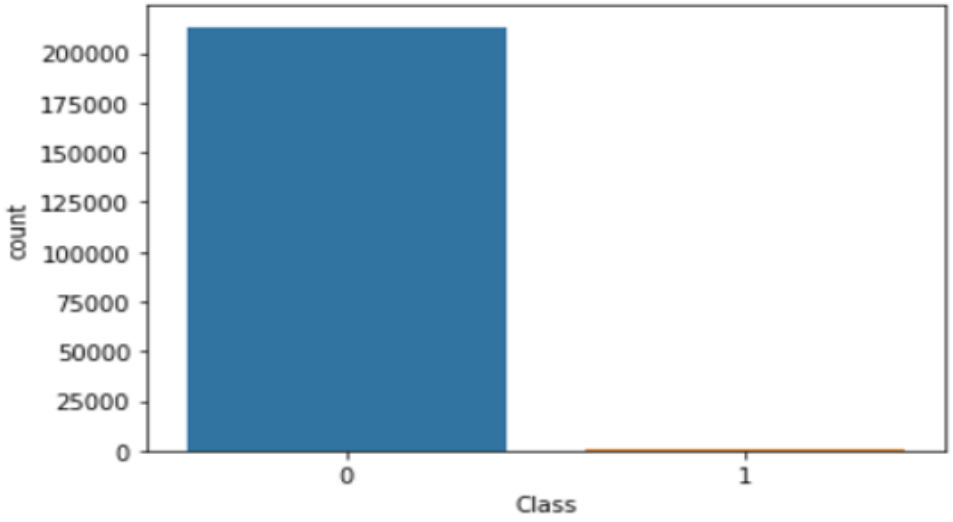


## IV. EXPERIMENTAL RESULTS

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 1.00 | 1.00 | 1.00 | 71079 |
| 1 | 0.87 | 0.80 | 0.83 | 123 |
| accuracy | | | 1.00 | 71202 |
| macro avg | 0.93 | 0.90 | 0.92 | 71202 |
| weighted avg | 1.00 | 1.00 | 1.00 | 71202 |

F1 Score: 0.8305084745762712

<matplotlib.axes._subplots.AxesSubplot at 0x7fc92cd11610>




we have used balanced dataset to predict the fraudulent transaction, which is downloaded from kaagle.com. we will import the require packages and check the dataset quality followed by splitting the data into train data and test data to get the accuracy . Using ANN algorithm , Random forest, decision tress comparatively we are getting the accuracy of loss 84% in ANN, Where as in logistic regression accuracy rate of Test data and training data is 94.28% , 91.37%.

## V. CONCLUSION

In this paper, Machine learning technique of logistic regression, artificial neural network were used to detect the fraud in credit card system. Sensitivity, specificity, accuracy is used to evaluate the performance for the proposed system. The accuracy for logistic regression for training data is 94.28 and testing data is 91.37and ANN accuracy loss of 84% respectively. By comparing both the method, found that logistic regression is better than the ANN.

## VI. REFERENCES

**1.** Sanobar khan, Sanovar, Suneel Kumar, Mr Hitesh Kumar, "Credit card fraud detection using machine learning" Published year: June 2021.

**2.** Vaishnavi Nath, "Credit card fraud detection using machine learning algorithms" Published year: 2019. Dr Gaurav Gupta. "Machine learning approach for credit card fraud detection" Published year: September 2019.

**3.** Prof Amol Jadav, "Credit card fraud detection using machine learning" Published year: September 2020.

**4.** S P Mani raj, "Credit card fraud detection using machine learning and data Science" Published year: September 2019.

# SIGN LAGUAGE DETECTION

*Baby*

***Abstract***

*A real-time sign language detection that can scan photos and identify indicators quickly at the rate of streaming images is absolutely necessary. This design proposes leveraging neural networks for sign language translation and text tracking and identification.*

*For those learning sign language for the first time, a system that utilizes hand detection has been created. This methodology is based on the explicit skin-color space thresholding method of skin-color modelling. The predetermined skin-color range will separate pixels (hand) from non-pixels (background). The Convolutional Neural Network (CNN) model was used to classify the images after being fed the photos. For image training, Keras was utilized. With the right illumination and a consistent background, the system achieved an average testing accuracy of 93.67%, outperforming earlier related studies. Of this, 90.04% was attributable to ASL alphabet recognition, 93.44% to number recognition, and 97.52% to static word identification. The method is real-time and utilized for quick computing.*

***Keywords::*** *sign_language*

## I. INTRODUCTION

Building a nation requires effective communication. All community members, including the deaf, benefit from effective communication because it improves understanding. A total of 1.23% of Filipinos are either deaf, mute, or have hearing impairments.

It is possible to communicate with others through sign language. However, sign language is difficult to master and is generally not understood by hearing individuals. As a result, there is still a definite divide between the hearing majority and the hearing impaired.

A lot of work has gone into developing a sign language recognition (SLR) system during the last few decades.SLR is divided into two basic categories: continuous sign classification and recognition of individual sign languages. While Starner et al. and Vogler et al. pay attention to continuous SLR,Zhang et al. and Wang et al. concentrate on isolated SLR.

The hidden Markov model (HMM) utilizes continuous SLR because it implicitly allows for the segmentation of the data stream into its continuous signals, completely avoiding the challenging segmentation problem.

Based on its input, the SLR architecture can be divided into two basic categories: data gloves-based and vision-based. Smart gloves are used by Chouhan et al. [6] to collect data on hand locations, joint alignment, and velocity utilising microcontrollers and specialised sensors like accelerometers and flex sensors. Other motion sensor-based methods for gathering signs exist, including electromyography (EMG) sensors, RGB cameras, Kinect sensors, jump motion controllers, or combinations of these. Higher accuracy is one of this strategy's benefits, while limited movement is one of its disadvantages.

The use of vision-based approaches, which include camera input, has grown in popularity in recent years (web camera, stereo camera, or 3D camera). Color-coded gloves were employed by Sandjaja and Marcos to facilitate hand detection. It is also feasible to combine the two architectures; this is known as a hybrid architecture.

These vision-based systems typically have two primary components to their architecture. The first step is feature extraction, which involves taking the desired features out of a movie using computer vision or image processing methods. The second component, the recognizer, should be able to correctly identify testing data on which machine methods were used by learning patterns from training data using the extracted and defined characteristics.

The goal of the proposed study is to create a system that can translate static sign movements into words and recognise them. To collect the data from the signer and allow for offline use, a vision-based approach utilising a web camera is developed. The system's goal is to act as a learning aid for people interested in learning more about the fundamentals of sign language, such as the alphabet, numerals, and typical static signals. The system's accuracy was increased by the proponents' use of a convolutional neural network (CNN) as the recognizer, a white background, and a specific area for image processing of the hand. The study's scope covers fundamental static signs, numbers, and ASL alphabets (A–Z). The system's capacity to generate words by fingerspelling without the use of sensors or other external technologies is one of the study's key characteristics.

## II. RELEATED WORK

A Real-Time American Sign Language Recognition System using Convolutional Neural Network for Real Datasets, using real colouring photos from a PC camera, a ConvNet algorithm was used to create a real-time ASL detection system.

The model, which was tested on fresh datasets, is the first ASL recognition model to classify a total of 26 letters, including (J & Z), with two new classes for space and delete. It was designed to include a wide range of qualities, such as various lightings, skin tones, backgrounds, and circumstances. The experimental findings had great accuracy, with training and validation accuracy of 98.53% and 98.84%, respectively.

P Subha Rajan and Dr. G. Balakrishnan created a technique for understanding gestures in Indian Sign Language, proposing that each gesture be recognised using 7 bit orientation and generation through RIGHT and LEFT scan. The next procedure was a laborious way to identify indicators that took around six modules.

T. Shanableh had created a technique for identifying individual Arabic sign language motions in a user-independent fashion. To make the procedure of colour segmenting the signer's hands easier, the signers in this approach wore gloves. Two alternative classification algorithms, K-NN and polynomial networks, were used to evaluate the performance of the proposed user-independent feature extraction scheme. Many scientists used specialised equipment to decipher sign language.

As a result, the suggested technology cannot be integrated into a typical PC camera. The ASL Finger Spelling Data sets from the University of Surrey's Centre for Vision were used to train CNNs for real-time ASL identification. The data set contains 24 static signs without the letters (J and Z) that were photographed over the course of five different sessions against a background with a similar level of illumination. A real-time hand posture identification model is created by Tang, Lu, Wang, Huang, and Li using deep learning-based CNNs. They made use of public datasets (MSRGesture3D). 12 dynamic ASL gestures were gathered from 10 participants for the dataset. A 94.17% accuracy rate was attained.

The main form of communication for the deaf community is sign language. Interacting with hearing individuals can be challenging for the millions of people who suffer from hearing loss worldwide.

The creation of automatic sign language recognition (SLR) systems is the primary goal of sign language recognition (SLR), which aims to improve communication with the deaf community. Prior to recent years, Arabic SLR (ArSLR) specifically did not garner much attention.

This paper gives a thorough evaluation of two distinct recognition methods for continuous ArSLR, namely a sequential data-suitable modified k-nearest neighbour method and Hidden Markov Models (HMMs) methods based on two separate toolkits.

Real-Time Face Tracking and Recognition System Using Kanade-Lucas-Tomasi and Two-DimensionalPrincipal Component Analysis The KLT (Kanade-Lucas-Tomasi) tracker and 2DPCA are used to create a proposed system for face tracking and recognition in a video series (Two-Dimensional principal Component Analysis). The Viola-Jones face identification technique is used to identify every face in the image or video sequence before the KLT algorithm is used to track faces. KLT maintains long-term tracking when faces enter or exit the frame after being identified in the preceding frames.

Humans can communicate effectively using hand gestures, but sign language is the most expressive and natural form of communication for the deaf and hard of hearing. This paper suggests a real-time hand gesture system. The system's experimental design uses a fixed-position, inexpensive web camera with a 10 mega pixel resolution positioned on top of the computer monitor to take pictures utilising the Red, Green, and Blue [RGB] colour space from a fixed distance. Feature matching, region extraction, feature extraction, and image preprocessing are the four processes that make up this task.

Additionally, academics' approaches to creating Sign Language Recognition differ. Each method has a particular advantage over the others, and researchers are continually experimenting with various techniques to improve their own sign language recognition. Each technique also has restrictions when compared to other techniques. The goal of this work is to review the techniques for recognising sign language and identify the most effective methodology that has been employed by researchers. As a result, other researchers can learn more about the techniques employed and possibly create future Sign Language Application Systems that are superior.

A Mobile Application of American Sign Language Translation via Image Processing Algorithms This research proposes a novel framework to recognise photographs of various sign language movements using well-known image processing techniques. To separate the hand motion from its background, we first use Canny edge detection and seeded region growth. Then, using the Speeded Up Robust Features (SURF) technique, feature points are retrieved from the features that were obtained using the Bag of Features method (BoF).

The trained dataset is then utilised to identify future inputs of sign language gestures. Support Vector Machine (SVM) is then employed to classify our dataset of gesture images. Experimental results demonstrate that the proposed framework can recognise and translate 16 different American Sign Language movements with an overall accuracy of 97.13%. It has been successfully implemented on smartphone platforms.

## III. METHODOLOGY

Figure (1) shows a general work flow for a real time sign language detection. There are five basic processes in real time sign language detection, each with its own critical significance.

a. Data Collection

b. Preprocessing

c. CNN Configuration

d. Pass dataset to CNN and build model

### A. Data Collection

The systematic process of obtaining observations or measurements is known as data collection.

### B. Preprocessing

The term "background subtraction" refers to the process of removing the backdrop from a photograph. This method includes separating the background from the foreground. A mask that is developed in accordance with the user's preferences is used to accomplish this. This method is used to detect through stationary cameras. Tracking an item requires background removal, which can be accomplished in a variety of methods.

We are surrounded by a lot of information, which our eyes pick up selectively based on our preferences and which varies for each person. On the other hand, machines observe everything and absorb every image before converting the information into 1s and 0s that a computer can comprehend. How is this conversion carried out Pixel. The smallest component of a digital image that may be projected or displayed on a display device is this. For our photos, we have values consisting of only one value (grayscale), which are automatically assigned by the computer based on the strength of the darkness or level.

The image has multiple intensity levels at various points, and these levels are represented by numbers. Grayscale or RGB are the two most often used methods for distinguishing images. Think of a black-and-white photo when describing a grayscale image; these photos have just two hues. White is the colour with the highest intensity, while black is thought to be or is used as a measurement of the lowest intensity. According to the levels of darkness, the values are assigned by the computer. Red, green, and blue (abbreviated as RGG) make up RGB. A real colour is created by the combination of all these hues.

Only by employing these three hues can any colour that exists on the earth be described. Every pixel is examined by the computer, which then determines its value and allocates it to an array.

**C. CNN Configuration**

Without the idea of convolutional neural networks, our entire model cannot be achieved. One of its classes of deep neural networks includes CNN. CNN is used in a variety of contexts, most of which include visual elements. The weights of the neurons that make up CNN can be changed to achieve the desired results. These biases and weights can be learned.

The neurons perform dot products, which causes nonlinearity. The assumption made by CNN that all the inputs are explicit images is the primary distinction between convolutional neural networks and regular neural networks. The entire network expresses a single function. Given that our project involves photographs, this will therefore be the optimal way to train the model. As a result, CNN will be able to benefit from the architecture's ability to be sensiblely limited with images as input; the configuration of neurons is done in three dimensions: width, depth, and height. The volume needed for activation is referred to as depth.

**D. Pass dataset to CNN and build model**

We are aware that a convolutional network has several layers and that a function is used to convert the volume of one form to another (differentiable). The pooling layer, convolutional layer, and fully-connected layer are the layers that make up the CNN architecture. A convolutional network architecture will be produced by stacking these correctly. The loss function must be decreased as these layers now extract characteristics from the image that are specific to a property. utilising this formula. It is possible to denote the sample's positive classes by M. The CNN score for each positive class is represented by Sp, and the scaling factor is represented by 1/M.

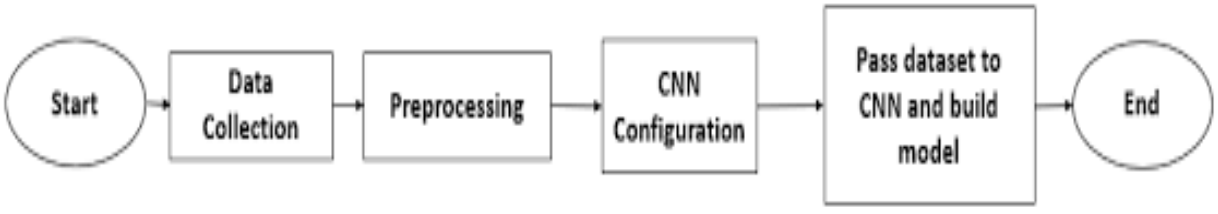



Figure (1). Work flow of the proposed system

## IV. EXPERIMENTAL RESULT

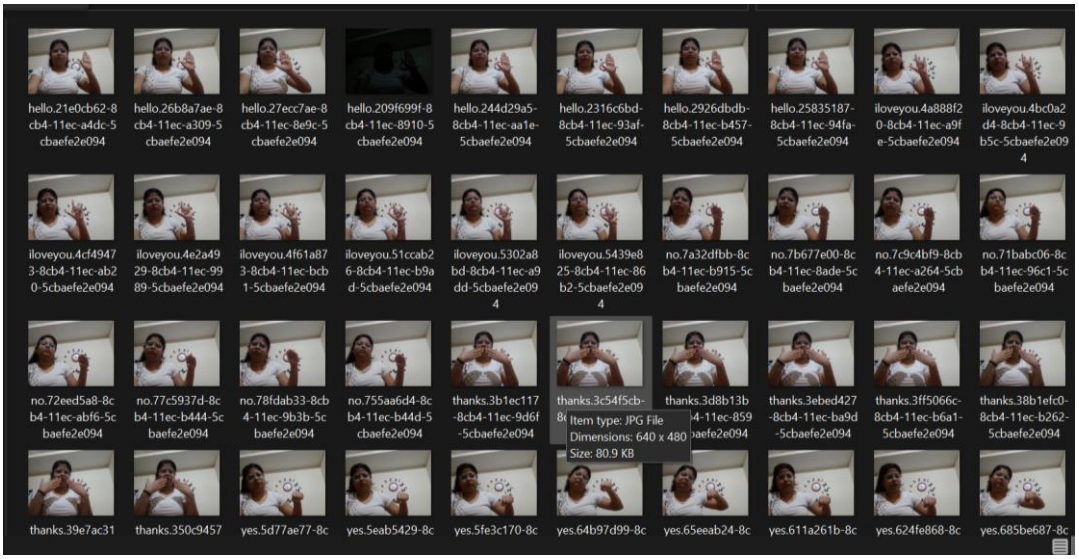



Figure (2). Input Images

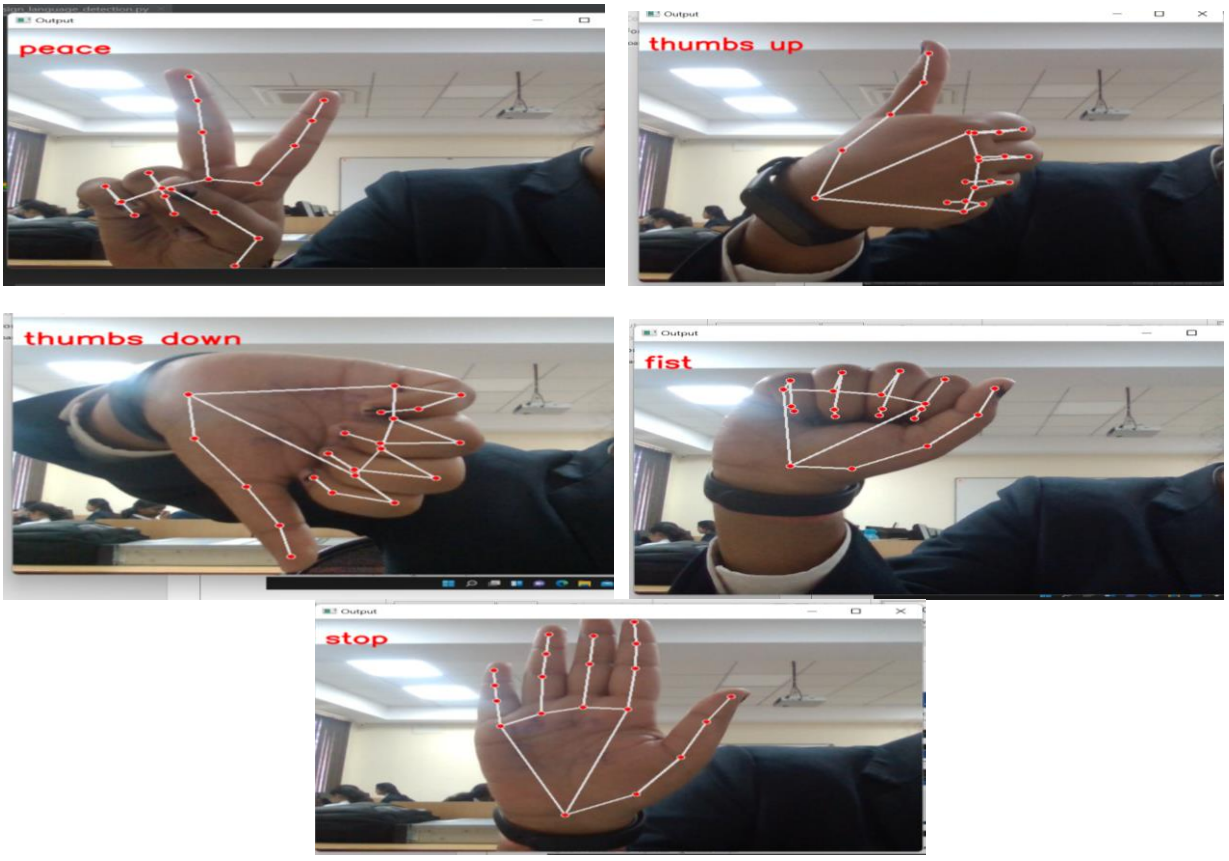



Figure (3) Output Image

## V. CONCLUSION

Using OpenCV and Python, we created a sign language detection for the proposed system. For the purposes of gesture identification and detection, I used the MediaPipe and Tensorflow frameworks, respectively.

Additionally, a traditional neural network was utilised to predict the sign, along with a widely known image processing method. Future recognition of sign languages will benefit further from the suggested method, allowing people to understand sign language.

## VI. REFERENCES


1. Kadhim, R.A. and Khamees, M., 2020. A real-time american sign language recognition system using convolutional neural network for real datasets. Tem Journal, 9(3), p.937.

**2**. Nikam, A.S. and Ambekar, A.G., 2016, November. Sign language recognition using image based hand gesture recognition techniques. In 2016 online international conference on green engineering and technologies (IC-GET) (pp. 1-5). IEEE.

**3**. Nyaga, C.N. and Wario, R.D., 2018, May. Sign language gesture recognition through computer vision. In 2018 IST-Africa Week Conference (IST-Africa) (pp. Page-1). IEEE.

**4**. Tolentino, L.K.S., Juan, R.S., Thio-ac, A.C., Pamahoy, M.A.B., Forteza, J.R.R. and Garcia, X.J.O., 2019. Static sign language recognition using deep learning. Int. J. Mach. Learn. Comput, 9(6), pp.821-827.

**5**. Jin, C.M., Omar, Z. and Jaward, M.H., 2016, May. A mobile application of American sign language translation via image processing algorithms. In 2016 IEEE Region 10 Symposium (TENSYMP) (pp. 104-109). IEEE.

**6**. Anderson, R., Wiryana, F., Ariesta, M.C. and Kusuma, G.P., 2017. Sign language recognition application systems for deaf-mute people: a review based on input-process-output. Procedia computer science, 116, pp.441-448.

**7**. Dawod, A.Y., 2021. Hand Gesture Recognition Based Sign Language Interpretation in Real-Time. LINGUISTICA ANTVERPIENSIA, pp.2987-3007.

**8**. Barnouti, N.H., Al-Mayyahi, M.H.N. and Al-Dabbagh, S.S.M., 2018, October. Real-Time Face Tracking and Recognition System Using Kanade-Lucas-Tomasi and Two-Dimensional Principal Component Analysis. In 2018 International Conference on Advanced Science and Engineering (ICOASE) (pp. 24-29). IEEE.

**9**. Hassan, M., Assaleh, K. and Shanableh, T., 2019. Multiple proposals for continuous arabic sign language recognition. Sensing and Imaging, 20(1), pp.1-23.

**10**. Kadhim, R.A. and Khamees, M., 2020. A real-time american sign language recognition system using convolutional neural network for real datasets. Tem Journal, 9(3), p.937.

# TEXT TO SPEECH SYNTHESIS

*Dainy Jose*

*Abstract*

*A Text-to-Speech (TTS) system converts normal language text into speech. In this system, the input text is taken and output is given as speech. The quality of a speech synthesizer is judged by its naturalness and intelligibility.TTS systems are based on a complex pipeline. TTS is one of the major NLP applications. Text-to-speech discussion involves three key stages of text analysis, text processing and waveform production i.e., speech construction. TTS is an application that converts text into speech, the user enters text and receives output as audio.In this research paper, we have implemented English text to speech system based on Tacatron-2 model. In this method we trained the model based on a dataset with all words in English language and their pronunciations. The system is composed of a recurrent sequenceto-sequence feature prediction network that maps character embeddings to mel-scale spectrograms, followed by a modified WaveNet model acting as a vocoder to synthesize time-domain waveforms from those spectrograms.The benefits of speech synthesis have been many, including computers that can read books to people, better hearing aids, more simultaneous telephone conversations on the same cable, talking machines for vocally impaired or deaf people and better aids for speech therapy.*

*Keywords:: text,speech.*

## I. INTRODUCTION

A Text-to-Speech (TTS) system converts normal language text into speech. In this system, the input text is taken and output is given as speech. The quality of a speech synthesizer is judged by its naturalness and intelligibility.

The Tacotron 2 and WaveGlow model form a text-to-speech system that enables user to synthesise a natural sounding speech from raw transcripts without any additional prosody information.

The Tacotron 2 model produces mel spectrograms from input text using encoder-decoder architecture.

WaveGlow is a flow-based model that consumes the mel spectrograms to generate speech. Tacotron, a sequential-to-sequential architecture for producing large spectrograms from letter sequences, simplifies the normal pipeline of speech by altering the production of these language and acoustic elements with a single network of sensors trained from a single data source. To quote the effects of the magnitude spectrogram, Tacotron uses the Griffin-Lim algorithm to scale the stage, followed by a short Fourier transformation.

This was simply a reflection of the future trends of neural vocoder, as Griffin-Lim produces artistic elements with lower sound quality than WaveNet. This implementation of Tacotron 2 model differs from the model described in the paper. Our implementation uses Dropout instead of Zoneout to regularize the LSTM layers.

The benefits of speech synthesis have been many, including computers that can read books to people, better hearing aids, more simultaneous telephone conversations on the same cable, talking machines for vocally impaired or deaf people and better aids for speech therapy.

## II. LITERATURE REVIEW

**Paper 1**: In paper 1, they describe a mixed approach to speech production in which powerful waveform generating fragments of a set of format-based compound rules are combined with preprogrammed natural speech fragments to produce a synthetic tone. This multidisciplinary approach integrates speech that combines legal-based integration with natural language waveform integration, and suggests ways in which a hybrid-based text-to-speech system can harness the power of each approach. Specifically, such a system may have a much higher sound quality than current systems based on rules, while other speech features can be improved by using the ability to carefully control parameters in integrated components. In addition, since a large percentage of segments could be performed by flight, with a smaller set of rules, system size may be reduced. Although this hybrid was originally used as a tool for improved voice quality research in format-based integration, it has produced so much good results that we now see it as a potentially powerful and effective method of converting text into speech. Potential advantages of the method include a small waveform interface for the waveform assembly, the development of specific speech signals for less sensitive space, and improved unit selection or productivity and efficiency. Additionally, this method has already proven its use as a research and development tool in both format and integration based on format.

**Paper 2**: In paper 2, the mathematical language integration system based on Markov hidden models (HMMs) has grown in popularity over the past few years. This simultaneous system simulates spectrum, excitement, and duration of speech using contextual HMMs and generates speech wave forms from HMMs themselves.

With integrated integration we need very different expression on voice characteristics, speech styles, and emotions, we need to record large or large data on this variety in order to achieve the integration we desire without compromising on quality. However, recording such a large database is very difficult and expensive.

In the system, contextual HMMs are trained from the native speech website, and we can generate speech waves from HMMs themselves. This program offers the ability to model different styles without the need for recording very large database.

**Paper 3**: In paper 3, Text-to-speech (TTS) synthesis translates input text carelessly into clear and audible speech. The TTS system consists mainly of two components: native language processing and digital signal processing. It is necessary to remove the speech delay when the speech waves are made to coalesce. Therefore, direct domain integration can produce speech smoothly. For two or more characters, the system method is most suitable. Audio quality is understandable. Therefore, a straightforward combination of domain and phonetic support is very easy and smooth to use unlike other methods involving many complex algorithms. But in unit selection, implementation is not as easy as these two methods. The outgoing sentence of the sentence has a slight error. However, speech output is better than phoneme-based synthesis.

**Paper 4:** In paper 4, they propose a novel framework that enables us to manipulate and control forms in HMM-based speech integration. In this framework, the dependence between formats and spectral features is modelled by partial line conversion; Successfully mapped format parameters are in the Gaussian distribution methods over spectral synthesis parameters. The spectral envelope features produced under the influence of formats in this way may be transferred to high quality voice generators to produce a speech wave. This offers two major advantages over conventional frameworks. First, we can achieve spectral conversions by changing the forms only in those parts where we want to control, while the user must specify all forms by hand in the synthesis of a standard format (e.g., Klatt). Second, this can produce high quality speech. This function demonstrates the proposed method of controlling vowels in integrated speech by deceiving F1 and F2 without compromising compound quality. Taken together, this synthesizer will be very useful in other areas, such as speech recognition and phonetic research, where the Klatt model is currently the main tool. The experimental results also showed that the proposed method provides control of vowels in integrated speech with F1 and F2 manipulation, without compromising compound quality.

**Paper 5**: In paper 5, they explore various ways of compiling Text-to-speech developed and used by a few keyword research researchers and research groups around the world over the past decades. Ongoing research continues the discourse smoothly with the help of the main methods namely the mathematical method of integrating speech. This paper provides a brief overview of language integration in Indian languages, summarizing and comparing the features of the various techniques used. Concatenative synthesis is the best form of text to speech that provides better audio output compared to other composite methods. Noise level is better when using Di-phone synthesis.

**Table: Existing papers have used the following techniques**

| PAPERS | TECHNIQUES USED |
|---|---|
| Integration of Rule-Based Formant Synthesis and Waveform Concatenation: A Hybrid Approach to Text-To-Speech Synthesis | • Speech synthesis<br>• Natural languages<br>• Synthesizers<br>• Speech enhancement<br>• Concatenated codes |
| The HMM-based Speech Synthesis System (HTS) Version 2.0 | • Markov hidden models (HMM) |
| Text To Speech Conversion Using Different Speech Synthesis | • Phoneme based text to speech<br>• Unit selection method |
| Formant-controlled HMM-based Speech Synthesis | • Markov hidden models (HMM) |
| Review on Text-To-Speech Synthesizer. | • Di-phone synthesis<br>• Concatenative synthesis |

## III. PROPOSED WORK

The Tacotron 2 and WaveGlow model form a text-to-speech system that enables user to synthesise a natural sounding speech from raw transcripts without any additional prosody information.
This is an end-to-end speech synthesis. In this we just feed the character sequence in the input and directly predict our mel spectrogram, and later we can generate the audio in the form of wave format using vocoder.

The system basically. Sort of encoder-decoder sort of system or sequence -to sequence model where we have the text processing encoder which basically consumes the character embedding. Then it generates mel spectrogram, from there we can generate audio using wavenet vocoder.

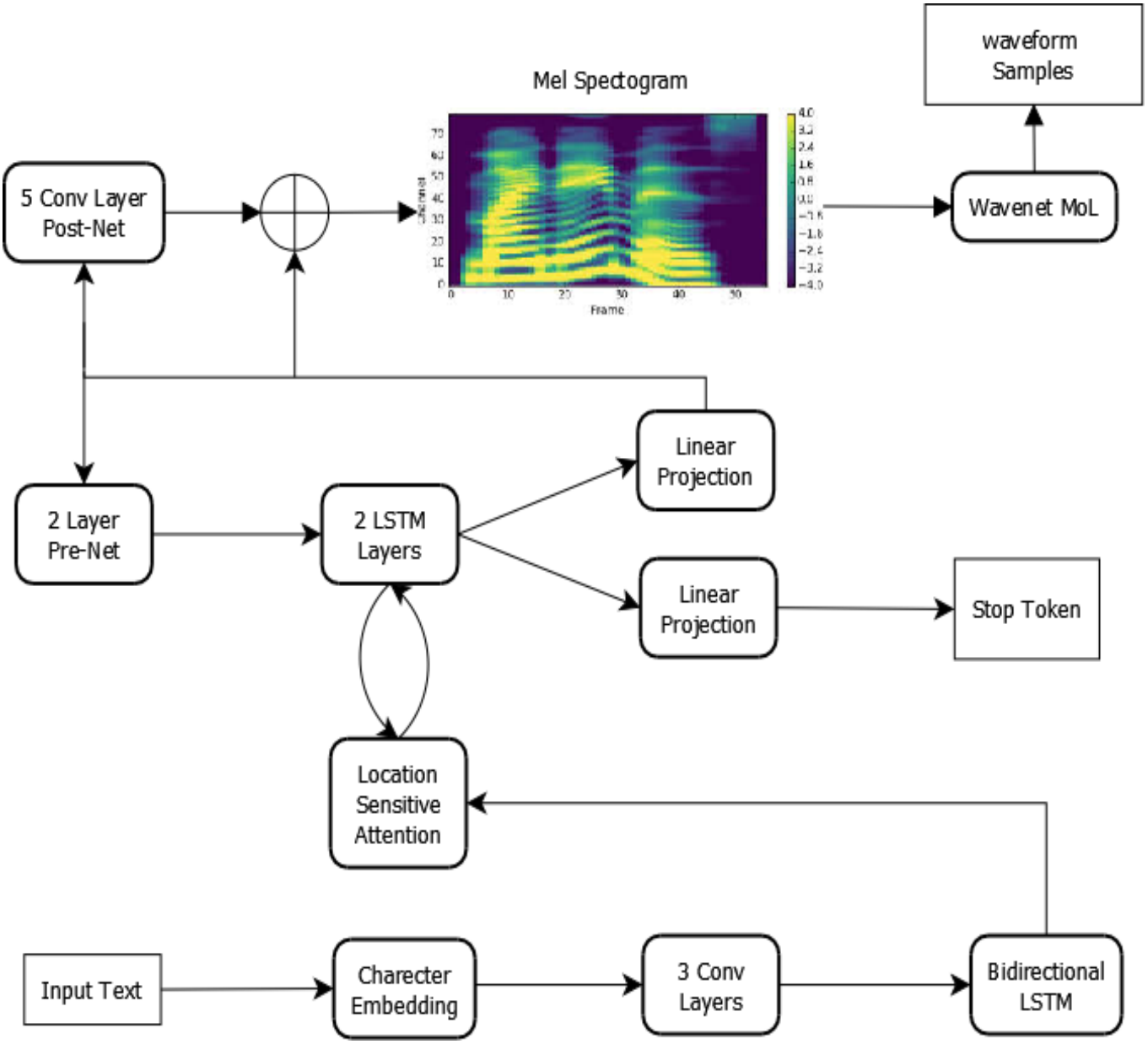
**Workflow:**

- ❖ Pretrained Tacotron2 and Waveglow models are loaded.
- ❖ Tacotron2 generates mel spectrogram given tensor represantation of an input text or pdf file.
- ❖ Waveglow generates sound given the mel spectrogram
- ❖ The output sound is saved in an 'audio.wav' file

**Algorithm**

- The Griffin-Lim Algorithm (GLA) is a phase reconstruction method based on the redundancy of the short-time Fourier transform. It promotes the consistency of a spectrogram by iterating two projections, where a spectrogram is said to be consistent when its inter-bin dependency owing to the redundancy of STFT is retained.
- Short-time Fourier transform (STFT) is a sequence of Fourier transforms of a windowed signal. STFT provides the time-localized frequency information for situations in which frequency components of a signal vary over time, whereas the standard Fourier transform provides the frequency information averaged over the entire signal time interval.
- Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series.

**Architectural Diagram**



## IV. METHODOLOGY

### Tacotron 2 and WaveNet model

The Tacotron 2 and WaveNet model form a text-to-speech system that enables user to synthesize a natural sounding speech from raw transcripts without any additional prosody information. The Tacotron 2 model produces mel spectrograms from input text using encoder-decoder architecture. WaveNet model that consumes the mel spectrograms to generate speech.

### Intermediate Feature Representation

The mel-frequency spectrogram is related to the linear-frequency spectrogram, that is, the size of the Fourier transform (STFT) short-term. It is obtained by using an indirect change in frequency The axis of the STFT, motivated by rated responses from the individual a hearing system, and summarizes content that is usually a few ratings. Using such a scale of hearing has the effect of emphasizing details on lower frequencies, which are important to speech comprehension, while de-emphasizing high frequency details, which they are dominated by fricatives and other explosions and usually do it is not necessary to model with high fidelity. Because of these buildings, features taken from the mel scale were used as the basis representation of speech recognition for decades.

### Spectrogram Prediction Network

In line with the prediction of the spectrogram framework, the convergence of the output LSTM decoder and the context of attention are set down in scalar and passed on sigmoid performance to predict Chances are that the output sequence is complete. The "stop token" prediction is used during speculation to allow the model to operate more efficiently decide when a generation should be cut off instead of always productive fixed term.

### WaveNet Vocoder

We use a modified version of the WaveNet architecture from to invert the mel spectrogram feature representation into time-domain waveform samples. The Tacotron 2 and WaveGlow model form a text-to-speech system that enables user to synthesise a natural sounding speech from raw transcripts without any additional prosody information. The Tacotron 2 model produces mel spectrograms from input text using encoderdecoder architecture. WaveGlow is a flow-based model that consumes the mel spectrograms to generate speech.

## V. CONCLUSION

This paper describes Tacotron 2, a fully neural TTS system that combines a sequence-tosequence recurrent network with attention to predicts mel spectrograms with a modified WaveNet vocoder. The resulting system synthesizes speech with Tacotron-level prosody and WaveNetlevel audio quality. This system can be trained directly from data without relying on complex feature engineering, and achieves state-of-the-art sound quality close to that of natural human speech.
Also, this method is used in android TTS, which is used for mobile applications. It consists of a large database for all the words in English. But there are few abbreviations and formats missed out in this database like reading passwords, security number, etc., i.e., Ph., Dr., etc. In this research we have proposed a work to overcome these challenges.

## VI. FUTURE ENHANCEMENT

Future enhancement for our research work is to implement text to speech with our own database. With consists of new words and better pronunciations. Also, we will include more symbols and corresponding pronunciation in the database. So, it will cover almost all the possible combinations of text. We will also include reading of given input text with the emotion. Also include the speech synthesis from large input as text or pdf file.

## VIII. REFERENCES

**1**. Design and Development of a Text-To-Speech Synthesizer System. - A.Chauhan, Vineet Chauhan, Gagandeep Singh, ChhaviChoudhary, Priyanshu Arya

**2**. Integration Of Rule-Based Formant Synthesis And Waveform Concatenation: A Hybrid Approach To Text-To-Speech Synthesis. - Susan R. Hertz

**3**. Natural Language Processing techniques in Text-To-Speech synthesis and Automatic Speech Recognition. - Alexandre Trilla

4. The HMM-based Speech Synthesis System (HTS) Version 2.0 - Heiga Zen, Takashi Nose, Junichi Yamagishi, Shinji Sako, Takashi Masuko, AlanW. Black

**5**. Text To Speech Conversion Using Different Speech Synthesis - Hay Mar Htun, TheingiZin, HlaMyoTun.

# Object Detection Using Deep Learning Algorithms

*Aishwarya K V*

*Abstract*

*One of the most popular and significant branches of computer vision is object detection. Deep learning algorithms on detecting task and performance on object detectors are highly developed, and have been utilized more in people's lives for security monitoring, autonomous driving, and so on. Object detection is the process of identifying, classifying, and localizing objects in a picture. The goal of detection is to output an image's labels and borders. Object detection is a critical subject in the science of computer vision. It offers significant research and application value in the areas of monitoring and autonomous driving. In recent years, deep learning has made a breakthrough in the investigation of picture categorization and led to the rapid growth of object vision identification.*
*In this paper, we use the YOLO deep learning algorithm. The YOLO algorithm generates a bounding box and prints the object's confidence level. For the production of blobs, we used a Deep Neural Network, which aids in the analysis of shape features of objects such as area, length, position, and direction of lumps, and we smoothed the image using Gaussian blur.*

*Keywords::* Bounding box, Confidence, shape features, length, position, smoothened

## I. INTRODUCTION

Humans can quickly detect and recognize objects in their environment regardless of their conditions, regardless of their position or whether they are alone. Partially occluded, upside down, different in color or texture etc. As a result, humans make object detection appear simple. Using a computer to discover and recognize the same objects. To extract some information on, a lot of processing is required. The shapes and objects depicted in a photograph. Surveillance, cancer detection, vehicle detection, and underwater object detection are all applications where object detection is crucial. For various purposes, several strategies have been utilized to precisely and effectively detect the item.
Deep learning models are now widely used across the whole field of computer vision, including both generic and domain-specific object detection.

Deep learning networks are used as the backbone and detection network in most state-of-the-art object detectors to extract features from input photos, classification, and localization. Object detection is a computer technology that deals with finding instances of semantic items of a specific class (such as individuals, buildings, or cars) in digital photos and videos.
It is related to computer vision and image processing. Multi category detection, edge detection, salient object detection, posture detection, scene text detection, face detection, and pedestrian detection are some of the well-researched object detection domains. Object detection has been widely used in many domains of modern life, including security, military, transportation, medical, and life fields, as a key aspect of scene interpretation.

This is where the YOLO approach fits in. We will not select the portions of the image that are of interest in this case. Instead, we forecast the classes and bounding boxes of the entire image in a single run of the algorithm and use a single neural network to recognize multiple objects. When compared to other classification methods, the YOLO approach is quick. Our algorithm processes 45 frames per second in real time. In the background, the YOLO algorithm makes localization errors but predicts fewer false positives. The efficiency of detecting families such as CNN and R-CNN was examined, and the YOLO method was introduced to improve the efficiency.
The main steps in object detection are as follows: Image conversion, Gaussian Filter, Resizing the image, bounding the image using YOLO algorithm.

## II. RELATED WORK

The methods proposed by Heba Hakim and Ali Fadhil are for training the dataset they have used back-propagation pooling layers, back-propagation fully connected layers.

## III. METHODOLOGY

The proposed model is described in depth in this paper. The main work flow of object detection is depicted in the diagram below. There are five main phases in detecting objects in total.

A. Pre-processing
    i. Converting the image from BGR to RGB
    ii. Blurring the image
    iii. Resizing the image
B. Creation of blobs
C. Bounding boxes

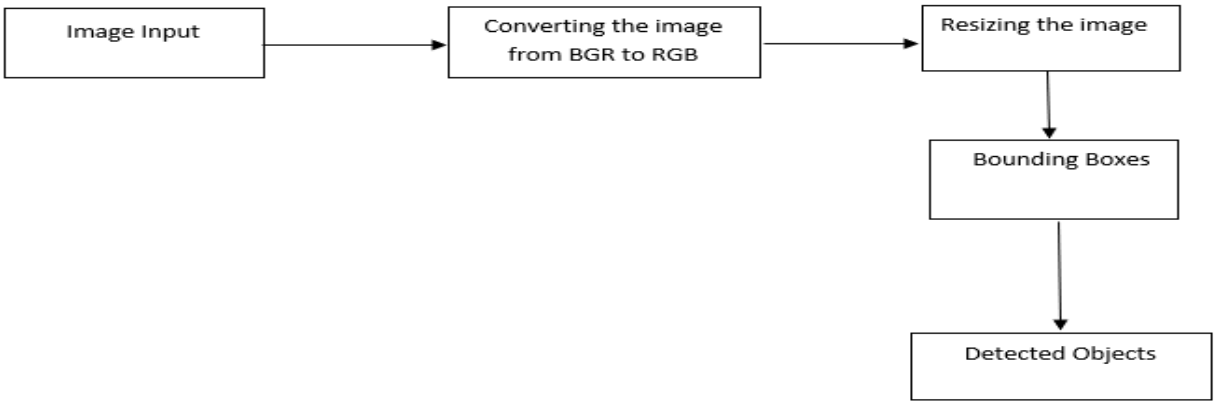



Figure (1). Work flow of proposed system

### A. Pre – Processing

The goal of pre-processing is to improve the picture quality in the data by removing superfluous distortion. Image functions are critical for subsequent processing. The image pixel values are transformed to 420 * 420 images, which are then fed into the neural network. This is done to prevent the neural network from becoming too dense. The following procedures are followed in the pre-processing of images.

We read the picture, image shape, and image size before pre-processing the images because images will vary in shape and size.

### Converting image from BGR to RGB

The photos must be converted from BGR (blue, green, red) to RGB (red, blue, green) because the dataset contains BGR images. In general, OpenCV reads images in BGR format, whereas Matplotlib reads images in RGB format. The main reason for converting the image from BGR to RGB is different pixel orderings exist in various processing image libraries. Because we're using Matplotlib to read the image, we'll need to convert it from BGR to RGB. Since our model works well with RGB format images, the conversion will be done appropriately. Converting the images from BGR to RGB makes object detection easier.

**Blurring the image**

We use the Gaussian blur to improve the textures of the image in different compartments because images often contain noise. Gaussian blur is used to reduce noise and remove speckles from images. It's also critical to remove high frequency components that aren't associated with the gradient filter being used, as this could result in false edges being detected.

Smoothening an image by blurring it implies removing the outliner, which might cause noise in the image. The Gaussian blur was used to blur or smooth the image.

## Resizing the image

The images will be converted to a smaller size using the Open CV library. Reduced size may result in a reduction in vector features and computational complexity for neural networks.

Because the images in the dataset are of varying sizes, resizing the image is necessary. Since the neural network receives inputs only of the same size therefore the image has to be reduced to a predetermined size before applying them to neural network. Larger the fixed size less shrinking required. Less shrinkage means less deformation of features.

**A. Creation of blobs**

In this step, we have created the blobs by using function from Open CV. Blob Stands for Binary Large Object and refers to the binary image's connected pixels. The adjective "large" refers to an entity of a definite size, whereas "little" binary objects are typically noise. When it comes to BLOB analysis, there are three steps.

**BLOB Extraction:**

Blob extraction is the process of extracting BLOBs (objects) from a binary image. A BLOB is a collection of connected pixels. The connectivity, or which pixels are neighbors of another pixel, can be used to identify whether two pixels are related or not. There are two kinds of interconnection. There are two types of connectivity: 8-connectivity and 4-connectivity. The 8-connectivity is significantly superior to the 4-connectivity.

**BLOB Representation:**

The term "BLOB representation" simply refers to the process of converting a large integer into a few smaller numbers. The next step is to categories the various BLOBs after they have been extracted.

In the BLOB representation procedure, there are two steps.

Each BLOB is given a set of characteristics in the first stage, and then some matching methods are used to compare the features of each BLOB in the second step.

**BLOB Classification:**

We determine the type of BLOB here, such as whether the BLOB is a circle or not. The difficulty here is how to tell which BLOBs are circle and which aren't based on the characteristics we discussed earlier. In most cases, we'll need to create a prototype model of the object we're looking for this purpose.

**B. Bounding Boxes**

We used our You Only Look Once (YOLO) technique to construct bounding boxes in this stage, which aid in the detection of items in the image. YOLO detects objects in real time with the help of a neural network. Because of its precision and speed, this is the most common option. According to my research, YOLO is the best algorithm for detecting objects in real time when compared to SSD, CNN, R- CNN, D-SSD, Fast R- CNN, and Faster R- CNN. In a single run of the Algorithm, YOLO predicts the bounding boxes and classes of the entire image. To deal with the problem, YOLO employs Non-Maximal Suppression.

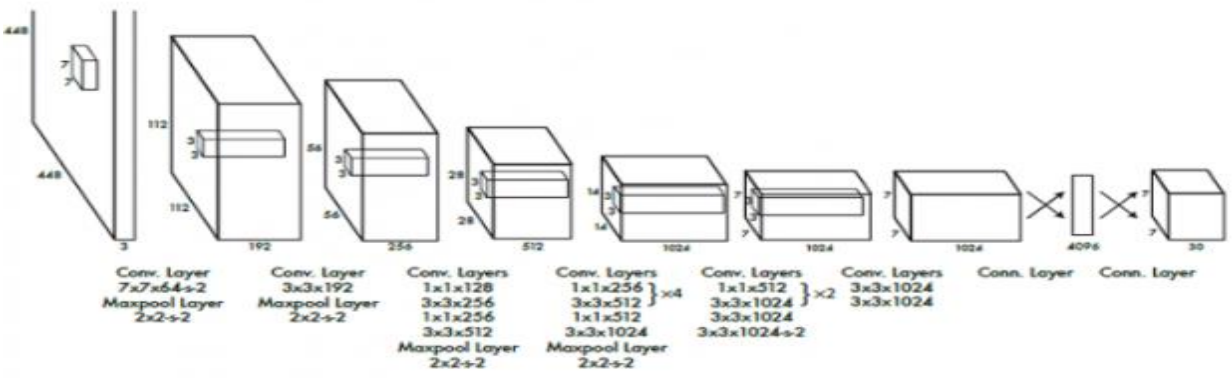



Figure (2). Architecture of YOLO algorithm

Basically, YOLO algorithm works by using the three techniques:

a) Residual blocks
b) Bounding Boxes
c) Intersection over union

a) **Residual blocks:** To begin, the image is divided into a number of grids, each of which has a dimension of S * S. There will be a lot of grids with the same dimensions. Each grid cell must recognise the objects that appear within it. If an object appears in the centre of a grid, it is this that is responsible for object detection.

b) **Bounding Boxes:** Bounding boxes are outliners that draw attention to a certain object in the image. The image's bounding boxes are defined by their width, height, class, and bounding box centre. The likelihood of an object appearing in the bounding box is represented by this number.

c) **Intersection Over union:** Intersection over union is used by YOLO to create an output box that surrounds the objects. The prediction of bounding boxes and confidence scores is the responsibility of each grid cell. This removes bounding boxes that aren't the same size as the genuine one. The transfer learning framework from DARKNET is used to build YOLO.

## IV. DATA SET

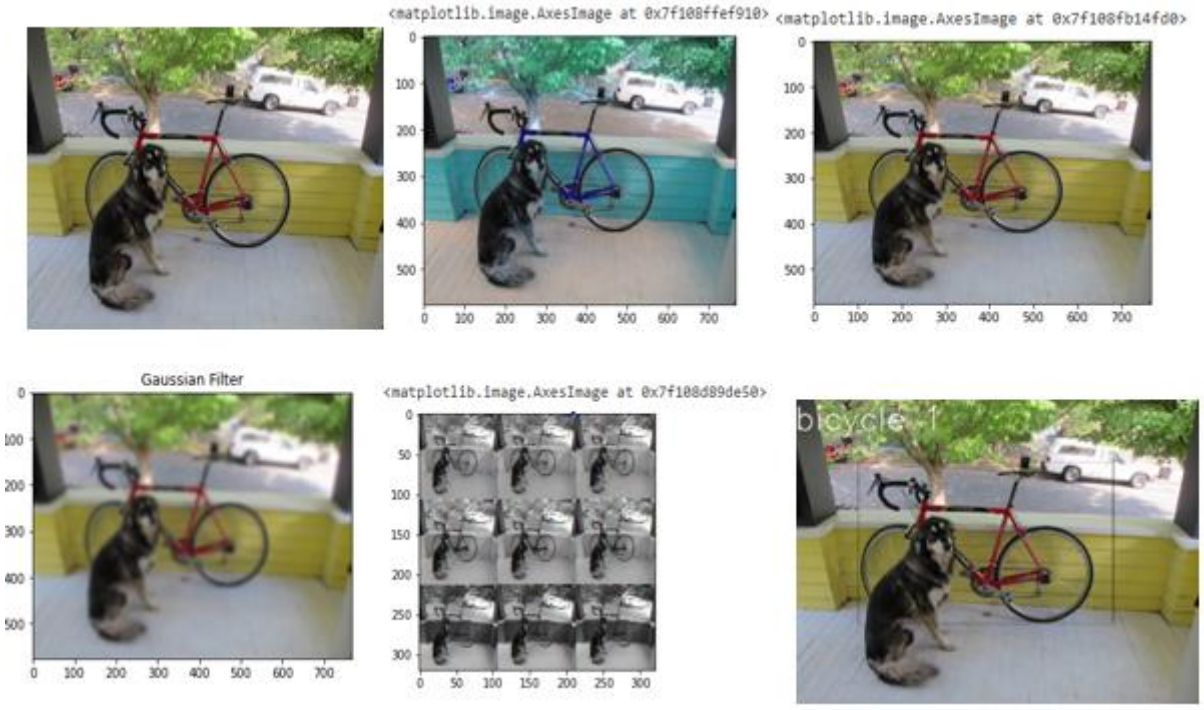
I have taken coco dataset from kaggle site. For object detection, key point detection, stuff segmentation, panoptic segmentation, dense pose, and image captioning, COCO provides numerous annotation types. JSON is used to store the annotations. Please note that all annotations can be accessed and modified using the COCO API specified on the download page. The core data structure of all annotations is the same. I used an MS coco dataset that contains two photos, a cfg file, a weights file (which is a pre-trained, speed-optimized file), and a list of COCO names that includes a list of items that the model can recognise.

## V. EXPERIMENTAL RESULT

Google collab is responsible for the system algorithm. Figure (2) shows an image from the dataset. We transformed the image from BGR to RGB, then used Gaussian blur to eliminate the image's noise.

We then used the YOLO algorithm. After all of the approaches have been used Figure(3) illustrates the result.

Figure (2). Input image




## VI. CONCLUSION

We proposed the YOLO technique in this paper for the aim of detecting objects with a single neural network. This algorithm is broader, and it outperforms others. When generalising from natural images, numerous methodologies are used to various domains. The algorithm is easy to create and maintain can be trained on a whole image immediately proposal for a region.

The classifier is limited to a specific region via strategies. YOLO When anticipating borders, it has access to the full image and it also forecasts that there will be fewer false positives in the background areas. When compared to other classification algorithms, this algorithm is superior is a significantly more efficient and quicker algorithm to utilise in real-world situations time.

## VII. REFERENCES


**1.** Hakim, H. and Fadhil, A., 2021, February. Survey: Convolution Neural networks in Object Detection. In *Journal of Physics: Conference Series* (Vol. 1804, No. 1, p. 012095). IOP Publishing.

**2.** Chahal, K. and Dey, K., A survey of modern object detection literature using deep learning. arXiv 2018. *arXiv preprint arXiv:1808.07256*.

**3**. Sommer, L.W., Teutsch, M., Schuchert, T. and Beyerer, J., 2016, March. A survey on moving object detection for wide area motion imagery. In *2016 IEEE winter conference on applications of computer vision (WACV)* (pp. 1-9). IEEE.

**4**. Sualeh, M. and Kim, G.W., 2019. Dynamic multi-lidar based multiple object detection and tracking. *Sensors*, *19*(6), p.1474.

**5**. Deshpande, H., Singh, A. and Herunde, H., 2020. Comparative analysis on YOLO object detection with OpenCV. *International Journal of Research in Industrial Engineering*, *9*(1), pp.46-64.

**6**. Bai, Q., Li, S., Yang, J., Song, Q., Li, Z. and Zhang, X., 2020. Object detection recognition and robot grasping based on machine learning: A survey. *IEEE Access*, *8*, pp.181855-181879.

**7**. Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, *28*.

**8**. Hechun, W. and Xiaohong, Z., 2019, August. Survey of deep learning based object detection. In *Proceedings of the 2nd International Conference on Big Data Technologies* (pp. 149-153).

**9**. Sumit, S.S., Watada, J., Roy, A. and Rambli, D.R.A., 2020, April. In object detection deep learning methods, YOLO shows supremum to Mask R-CNN. In *Journal of Physics: Conference Series* (Vol. 1529, No. 4, p. 042086). IOP Publishing.

**10**. Ahmed, M., Hashmi, K.A., Pagani, A., Liwicki, M., Stricker, D. and Afzal, M.Z., 2021. Survey and Performance Analysis of Deep Learning Based Object Detection in Challenging Environments. Sensors, 21(15),p.5116


Department of Computer Science (PG)
Jyoti Nivas College Autonomous
Hosur Road, Bangalore 560095