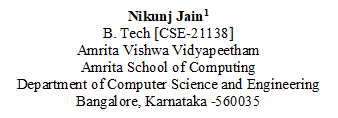
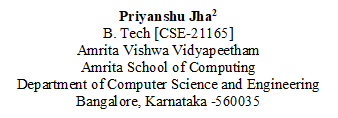
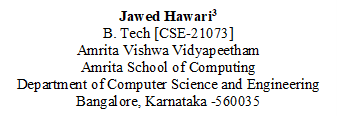
**Detection of Face and Eye using Machine learning Techniques**

*Corresponding author:Nikunj Jain*



**Abstract**—**Facial recognition has been nowadays a key role in computer field. The rapid advancement of Machine learning techniques has been revolutionizing the field of computer vision. This paper presents a comprehensive investigation into the application of Machine learning techniques. The study comprehensively reviews the various state-of-the-art facial and ocular algorithms. It depicts the modern learning techniques. With this we try to concatenate various models or techniques prevailed in the technology. A detailed examination of dataset and evaluation metrics are used for assessing detection performance is also presented to provide a foundation for comparative analysis. The core of this paper is used for designing and implementing of novel models for face and eye detection. The technique can be used in various places as applications such as Face detection in Airports for quick security, iris detection for authentication, implementing with CCTV for detecting for any type of fraudulence. In conclusion, this paper contributes to the ongoing discourse on machine learning approaches to face and eye recognition by providing new examples, performance reviews, and ethical considerations. The results of this study have important implications for the development of more accurate, more efficient, and more responsible facial and visual recognition systems, widely used in various industries**

**Keywords: Machine learning , Face Detection , Multi-Scale Training, Concatenate models**

**I.Introduction**

Face detection is a fundamental and important problem in computer vision and pattern recognition, which has been widely studied over the past few decades. Face detection is one of the important key steps towards many subsequent face-related applications. Face detection is a fundamental and important problem in computer vision and pattern recognition, which has been widely studied over the past few decades. Face detection is one of the important key steps towards many subsequent face-related applications.

Face recognition is a unique challenge in computer vision, and it is often considered an important aspect of object recognition. Researchers are trying to tackle face recognition by drawing inspiration from successful deep learning techniques applied to all object recognition tasks. An influential framework in this area is an approach based on convolutional neural networks (CNNs). This approach extends the capability of CNNs to efficiently solve object original method and its improved versions.

The recognition of human faces in images and videos regardless of factors such as context, lighting, obstruction, expression, and so on has become a major area of ​​interest in biometric analytics. Due to its relevance to this application, more attention has been given to this work recently Face recognition faces many challenges such as position, lighting conditions, obstacles, and facial expressions. Furthermore, the computational requirements of this system can severely limit real-time processing. To address these challenges, researchers have explored various approaches, including approaches based on shapes, materials, shapes and facial features.

**II.Literature Survey**

**1. Multiple face detection based on machine learning**

Hajar Filali1, Jamal Riffi2, Adnane Mohamed Mahraz3, Hamid Tairi4

This paper examines the rise of facial recognition, spurred by its widespread use. The research examines various approaches, including machine learning-based approaches, to overcome the challenges associated with traditional 2D-based approaches, such as feedback, illumination manipulation, etc. This four approaches a discrete (Haar-AdaBoost, LBP-AdaBoost, GF-SVM, GF-NN) analysis and comparison analysis based on processing time, face recognition accuracy, and false positives are helpful for computing the field of vision and human-computer interaction is enhanced by illuminating the severe weaknesses of this face recognition strategies

In summary, the research paper addresses the significant and challenging task of facial detection and explores various methods and techniques, with a focus on machine learning approaches to overcome the limitations of traditional methods. The comparative analysis provides valuable insights into the performance of different face detection methods, contributing to the ongoing advancements in this important field of computer vision and human-computer interaction.

**2. Face Recognition System Using Machine Learning Algorithm**

Sudha Sharma1, Mayank Bhatt2, Pratyush Sharma3

Facial recognition, the cornerstone of artificial intelligence, has found widespread applications in our daily lives from protecting smartphones to instantly recognizing Facebook users in photos Despite the adoption, the real world in face recognition is a formidable challenge due to factors such as different lighting conditions, partially face blocking and changes in settings. It also introduces a new method of integrating principal component analysis (PCA). By experimenting with various machine learning techniques such as linear discriminant analysis, multilevel perceptrons, naïve Bayes, support vector machines, the analysis achieves impressive descriptive accuracies of up to 97% and 100% using PCA and linear discriminant analysis is used, respectively

Facial recognition has tremendous utility in human-computer interaction applications and contributes to various aspects of modern life, including identity verification, security, and personal device authentication This review explores key areas of face recognition, and emphasizes the important extraction and classification steps . Principal Component Analysis (PCA) is an important technique for dimensionality reduction during feature extraction, while designing machine learning algorithms for efficient classification While addressing the challenges of real-world scenarios, this study highlights the importance emphasizing the need for continued improvement of facial recognition systems

**3. Masked Face Recognition Using Deep Learning: A Review**

Ahmad Alzu’bi1, Firas Albalas2, Tawfik AL-Hadhrami3, Lojin Bani Younis4, Amjad Bashayreh5

The paper addresses the increasingly important and complex area of ​​face masks (MFR) in an era of epidemic mask mandates and increasing safety concerns MFR has gained prominence as it plays an important role in implementation diversity, confronting to ensure security and secure reliability From modifications to facial recognition systems down to the rise of threats such as pandemics and fraud, systems and solutions developed in this area has accelerated, thereby introducing new complexities and requirements Typically, facial recognition systems focus on non-obstructive faces, but real-world events often require the identification of masked individuals, a need that has become especially important in the COVID-19 pandemic in

While prioritizing public safety, it also raises the challenge of authentication without the need for individuals to unmask, particularly in situations such as access roads or immigration This has led to a paradigm shift in facial recognition in 2010, with many organizations adapting existing algorithms and datasets to masked facial recognition. Deep learning, the powerful microscopic artificial intelligence plays an important role in improving MFR. Research efforts have focused on improving the accuracy of system detection in the face of masks or occlusions using various deep learning techniques, including sparse representations, auto coders, and bidirectional deep networks. Despite these advances, challenges such as computer complexity and Roe remain

**4. Face Recognition and Identification using Deep Learning Approach**

KH Teoh1 , RC Ismail2, SZM Naziri3 , R Hussin4, MNM Isa5 and MSSM Basir6

In recent years, the rapid development of artificial intelligence has led to innovations such as self-driving cars and automated supermarkets. One crucial aspect of artificial intelligence is computer vision, which aims to replicate human vision electronically, allowing machines to perceive and interpret images. Computer vision not only enables machines to "see" but also to react by detecting, identifying, and processing images, much like the human visual system. However, this presents a significant challenge as the real world is three-dimensional, while visual sensors typically provide only two-dimensional images. This paper explores the intersection of computer vision, face recognition, and deep learning to tackle these challenges, emphasizing the importance of developing efficient methods for analyzing three-dimensional objects in two-dimensional images.

Face recognition, a subset of computer vision, plays a vital role in biometric identification based on facial images. While humans can easily recognize individuals by their faces, computerized methods are required for scalable and accurate face recognition. Despite extensive research, face recognition still faces challenges such as misalignment, pose variation, illumination changes, and expression variability. This paper highlights the need to address these issues and explores the potential of deep learning, a machine learning technique that utilizes neural networks with multiple hidden layers, to improve face recognition accuracy. Additionally, the paper discusses the efficiency and performance differences between OpenCV and MATLAB in the context of computer vision tasks, emphasizing OpenCV's advantages in terms of speed and efficiency for real-time image processing.

**5. A Review on Using Machine Learning to Conduct Facial Analysis in Real Time for Real-Time Profiling**

Shynu T1, S. Suman Rajest2, R. Regin3, Steffi. R4

This paper describes a lie detection system that uses the Facial Landmark Detection System provided by the OpenCV Toolkit to analyse individual facial expressions and subtle eye movements and track suspects’ facial muscle movements and eye blinks loudly as the system answers a series of questions. It uses eye opening (EAR) to monitor eye opening in each photo, with the aim of detecting deceptive behaviour during interviews. The combination of machine learning and facial mark recognition allows for deeper analysis of fraudulent behaviour, and helps interviewers check the accuracy of answers, also considering possible future improvements such as heat tracking to lie the sight is accurate.

The literature review delves into the historical contemporary landscape of lie detection methods, emphasizing the importance of accurate deception detection in a variety of contexts including criminal investigations, intelligence interrogations, and court proceedings tree The study explores how lie detection evolves through development The ongoing pursuit of more accurate and reliable authentication methods, including EEG-based methods highlights the difficulties and challenges associated with fraud examining it in real-world situations, it emphasizes.

**6. Real-Time Face Detection and Face Recognition: Study of Approaches**

Siddhartha Singh Bhadauriya1, Sachin Kushwaha2, Shweta Meena3

This study delves into various face detection and recognition techniques, providing a comprehensive overview of their strengths and weaknesses. It aims to make the complex field of face detection accessible to individuals with varying levels of technological expertise. The study covers fundamental concepts, methodologies, and key terms, ensuring that even those less familiar with technology can grasp the basics and explore this domain further.

The paper discusses several face detection techniques, including geometric-based, feature-based, and Haar-like feature-based approaches, with a particular focus on the latter. In the realm of face recognition, the study explores lazy learner's, neural networks, and holistic face recognition approaches. The primary objective is to showcase the development of a model capable of real-time face detection, a crucial component in modern security systems. The proposed model utilizes the K-nearest neighbors algorithm, Haar cascade classifier for object detection, and OpenCV, an open-source Python library renowned for computer vision, machine learning, and image processing. The research underscores the significance of automated face detection in enhancing security systems and its applicability in real-time scenarios.

**III. Methodology**

*Entropy Calculation and Feature Selection:* The entropy for each characteristic or feature at the root node was determined in the first phase. The metric used to evaluate impurity was Information Gain. The goal was to determine which feature, when employed as the root node of the decision tree, maximizes Information Gain.

*Decision Tree Construction and* Depth: A Decision Tree was constructed using DecisionTreeClassifier. The model was trained with the data provided. The accuracy of this model was evaluated on the training set, and the depth of the decision tree was determined using the model's get\_depth() function.

*Visualization of Decision Tree*: The decision tree graph was plotted using the plot\_tree() function in conjunction with Matplotlib and scikit-learn's tree module. The resulting visualization helped in comprehending the structure of the decision tree.

*Decision Tree Construction and Accuracy*: A Decision Tree classifier was created and then fitted with the training data. The model's accuracy on the training set was determined, and subsequently, the accuracy on the test set was also evaluated. Also, a graphical representation of the decision tree was plotted.

*Decision Tree with Max Depth Constraint*: In this, we implemented a maximum depth constraint when constructing the Decision Tree classifier. This constraint helped to limit the tree's depth also preventing it from becoming overly complex. After constructing the model with this constraint, we proceed to evaluate its performance by measuring its accuracies on both the training as well as test datasets. We also visualized the resulting Decision tree to see how the imposed constraint influenced the tree's structure and complexity.

*Criterion Comparison*: In this, the criterion for the Decision Tree was studied. Initially, the model was built using the default criterion. Later, the criterion was changed to "Entropy". The differences in the models and graphical representation between the default and entropy criterion were investigated.

*Random Forest Classifier*: In this, we constructed the Random Forest classifier on the data set. A comparison was made between the performance metrics of the Decision Tree and Random Forest classifiers. This evaluation helped to highlight the differences in accuracy, precision, recall, and other relevant metrics.

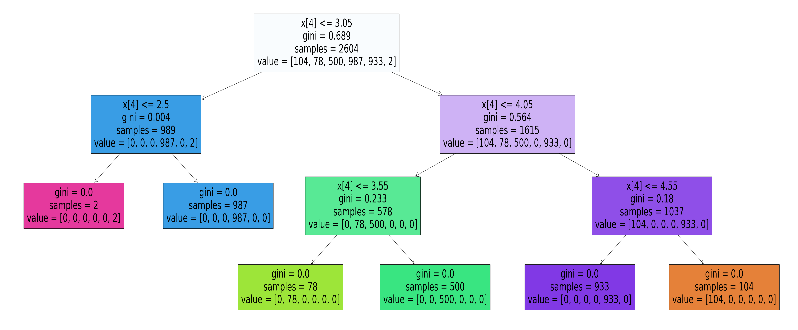
*Random Forest Parameters and Attributes*: Finally, a thorough study of the various parameters and model attributes of the Random Forest classifier was conducted. The importance of these parameters and attributes in shaping the model's behavior was explored. Key parameters, such as the number of trees, maximum depth, and minimum samples per leaf, were noted, along with attributes like feature importance and out-of-bag score. The analysis provided valuable insights into how different settings impact the Random Forest model.

IV. Result

1. Results obtained from A4 question
2. Training set accuracy: 1.0

Tree depth : 3

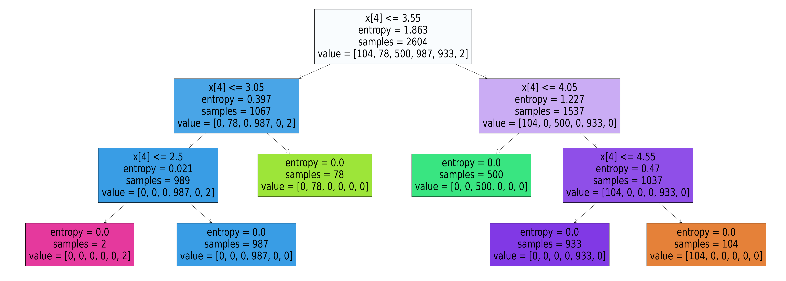
1. Decision Tree diagram



1. Results obtained from A6
2. Training set accuracy: 1.0

Tree depth : 3

1. Decision Tree diagram



1. Results from problem A7
2. Accuracy of Random Forest Classifier: 0.999012508229098
3. Accuracy of Decision Tree Classifier : 0.999835418038183
4. Classification report matrices of Random Forest Classification

Precision recall f1-score support

0 1.00 1.00 1.00 225

1 1.00 1.00 1.00 203

2 1.00 1.00 1.00 1089

3 1.00 1.00 1.00 2286

4 1.00 1.00 1.00 2267

5 1.00 0.00 0.00 6

accuracy 1.00 6076

macro avg 1.00 0.83 0.83 6076

weighted avg 1.00 1.00 1.00 6076

1. Confusion matrix of Random Forest Classifier

[[ 225 0 0 0 0 0]

[ 0 203 0 0 0 0]

[ 0 0 1089 0 0 0]

[ 0 0 0 2286 0 0]

[ 0 0 0 0 2267 0]

[ 0 0 0 6 0 0]]

1. Classification report of matrices of Decision Tree classifier

precision recall f1-score support

0 1.00 1.00 1.00 220

1 1.00 1.00 1.00 199

2 1.00 1.00 1.00 1083

3 1.00 1.00 1.00 2346

4 1.00 1.00 1.00 2224

5 0.80 1.00 0.89 4

...

accuracy 1.00 6076

macro avg 0.97 1.00 0.98 6076

weighted avg 1.00 1.00 1.00 6076

1. Confusion matrix Decision Tree classifier

[[220 0 0 0 0 0]

[ 0 199 0 0 0 0]

[ 0 0 1083 0 0 0]

[ 0 0 0 2345 0 1]

[ 0 0 0 0 2224 0]

[ 0 0 0 0 0 4]]

***Model Performance*:** The performance of model on general basis is same in both methods (Random Forest Classifier and Decision Tree Classifier). The model got the same accuracy in both methods. Further performance of model based on performance parameters are described below.

*Precision*: Based on precision Random Forest classification model perform better than Decision Tree model. In the precision of RFC, the accuracy to measure 6th label is 1.0 whether in Decision Tree classification it is 0.8. That meant there is 20% that Decision Tree classification will fail to make true positive decision.

*Recall*: Decision Tree Classification outperformed Random Forest Classification in terms of recall. For label 6th Decision tree has got 1.0 value, which means the model is able identify the true positive cases. However, in case of Random Forest classification recall value for 6th label is 0.0, which means that model is not able to identify true positive instances. In other words, the system failed to identify the instances for 6th label.

*F1-Score*: Decision Tree classification performs better compare to Random Forest Classification. Since F1-Score is combination of both precision and recall, it gives an overall idea about the model’s ability to identify and capture all the true positive instances for particular label. Random Forest Classifier has f1-score as 0.00 and Decision Tree Classification has the f1-score as 0.89.

*Confusion Matrix*: We use confusion matrix to check how well the model performs to on identifying the correct labels out of given samples. On observing confusion matrix of both classifiers, Random Forest classifier has six false positive values on the other hand Decision Tree classifier has one false negative value. Confusion matrix of Random Forest Classifier completely fails to identify the instances for 6th class, which is not good for model. Therefore, compare to Random Forest classifier Decision Tree Classifier performs better as from confusion matrix we can see it is successfully identifying all the classes with slight error.

*Overfit and Underfit* : Random Forest Classifier performs overfit because it has accuracy 0.999 yet it is unable to predict the last label. Decision Tree Classifier is able to predict the instances corresponding to all the classes. Thus, Decision Tree Classifier has underfit condition.

**The usage of pruning to avoid overfitting in decision trees**

Pruning is a technique that is used when there a high level of complexity found in model. It can be used with any classifier techniques such as Decision Tree classification, Random Forest Classification, etc. It avoids overfitting by cancelling all the branches that are not significant in tree model. As per our dataset, pruning is required to be applied on Decision Tree classifier. In problem statement A5, the tree classifier used criterion as ‘max\_depth=5’ and produces a tree of height 3. In that tree every branch is leading to a valid class label, which means all of them are significant for the model. So, in this case we are avoid using pruning in the classifier. However, if we use then the performance of the model would increase for sure.