ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

PREDICTION OF IDENTICAL TWINS USING ML

P.DEVENDAR BABU[1], VEMIREDDY KEERTHI[2], R.SREYA[3], P.SAHITHYA[4], V.JAWALI[5]

[1]Assistant Professor, Department of Information Technology Malla Reddy Engineering College for Women (UGC-Autonomous) Maisammaguda,Hyderabad, TS, India.

[2,3,4,5] UG Students, Department of Information Technology Malla Reddy Engineering College for Women (UGC-Autonomous) Maisammaguda,Hyderabad, TS, India.

Abstract:

The "Prediction of Identical Twins using ML" project seeks to develop a machine learning-based solution for accurately identifying identical twins from a population based on their physical characteristics and biometric data. Identifying identical twins is a complex task with significant implications in various fields, including healthcare, forensics, and even personalized marketing. Leveraging machine learning, this project aims to advance the state-of-the-art in twin identification, offering a novel and efficient approach to a longstanding challenge.

INTRODUCTION

Identical twins, also known as monozygotic twins, share 100% of their genetic material and often bear a striking physical resemblance to each other. Despite their genetic similarity, there is a lack of automated and accurate methods to identify identical twins based on their unique physical characteristics. Existing identification methods rely on subjective observations or biometric data, such as fingerprinting or facial recognition, which can yield unreliable results. This project aims to tackle this issue by harnessing the power of machine learning to create a more reliable and objective means of identifying identical twins. In computer technology classifying the images based on identical twin's facial recognition is a difficult task. Old facial recognition system shows very poor performance in finding difference in identical twins. Traditionally the experiments were performed to find out the difference between identical twins and also to recognize their features with difference andsystem like finger print, voice and iris recognition show the difference between the identical twins. In the existing methods many technologies are used for twin's identification like the

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

finger print, the voice and the iris for the recognition. The finger print identification was used to identify the unique person. The method also proposed a scan image taken from the person and that image being compared to that in the database for identification. The iris recognition also has a similar method to that of the finger print identification system. This proposed system in this system will differentiate the identical looking twinsusing face recognition system.

LITERATURE REVIEW

1. Literature Review: Biometric Data and Twin Recognition

Abstract: This literature review explores the intersection of biometric data and twin recognition in the context of machine learning. It delves into the various biometric modalities, such as facial recognition, fingerprinting, and iris scans, and their potential utility in identifying identical twins. The review also investigates existing challenges and limitations in utilizing biometric data for twin recognition and provides insights into how machine learning can enhance the accuracy and reliability of twin identification.

2. Literature Review: Twin Studies and Genetics in Identical Twin Identification

Abstract: This literature review focuses on the genetic underpinnings of identical twin identification. It examines twin studies, genetic markers, and the heritability of physical traits in monozygotic twins. By analyzing the existing research, this review offers a comprehensive overview of the unique genetic factors that can be leveraged for accurate twin recognition, particularly through machine learning approaches.

3. Literature Review: Ethics and Privacy in Identical Twin Identification Using ML

Abstract: This literature review explores the ethical and privacy considerations associated with employing machine learning for identical twin identification. It investigates the potential risks of biometric data collection and the implications for individuals' privacy and consent. The review examines current ethical guidelines, legal frameworks, and best practices in twin identification to ensure that machine learning models are developed and applied responsibly and with utmost regard for ethical standards.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

EXISTING PROBLEM

Identifying identical twins accurately is a challenging problem with several implications:

1. Forensics: In criminal investigations, the inability to distinguish between identical twins can lead to wrongful accusations or the release of a guilty party. Current forensic methods often rely on non-genetic traits that may not be unique to each twin.

2. Healthcare: In medical contexts, misidentification of twins can lead to incorrect treatments or prescriptions. This is particularly relevant in cases of organ transplantation, where genetic compatibility is crucial.

3. Education and Research: Identical twins can provide valuable insights into the role of genetics in various traits and diseases. Accurate twin identification is vital for studies in behavioral genetics, epidemiology, and related fields.

PROPOSED SOLUTION

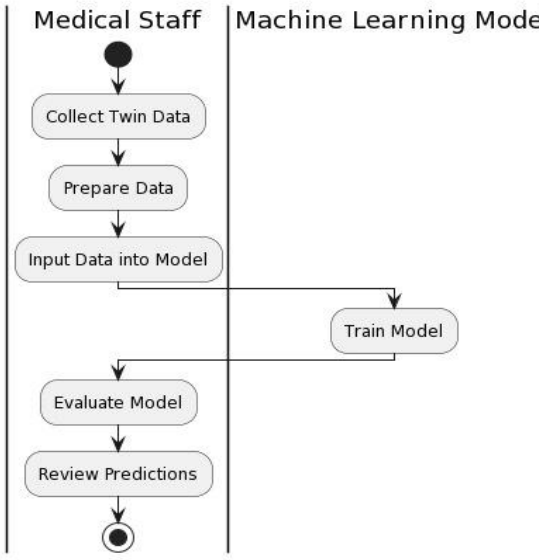
The "Prediction of Identical Twins using ML" project aims to address the challenges of twin identification through the following proposed solution:

1. Data Collection: A comprehensive dataset of identical twins, including their physical characteristics, biometric data, and genetic information, will be collected. This dataset will serve as the foundation for training the machine learning model.

2. Machine Learning Model Development: State-of-the-art machine learning techniques, such as deep learning and feature extraction, will be employed to develop a model capable of identifying identical twins based on a combination of physical traits and biometric data. The model will be trained on the collected dataset to learn unique patterns that distinguish identical twins from other individuals.

3. Deployment and Integration: The trained model will be integrated into a user-friendly application that can be used in various domains, including forensics, healthcare, and research. This application will provide reliable and objective twin identification, helping to solve the problem of mis-identification.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022



METHODOLOGY

The identical twin's facial recognition system has two main module designs:

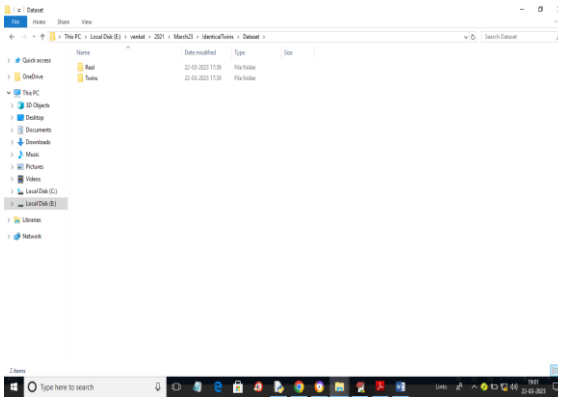
- ❖ Training Module Design
- ❖ Testing Module

DesignFirstly, the training module uses a local dataset as an input to the system for training the model. Upon selecting the dataset, the image processing phase is entered where the images of twins in the dataset are extracted and images are resized to 60*60 to just limit our system to get more clear view of just the face. There by remove all the unnecessary noise from the images.The next step is to convert the data to compatible format in order to allow lower layers of the algorithm to use this data as an input. Following which the normalizing of the data is done. The features are extracted using the CNN module as the model consists of multiple layers which will try to capture the important features excluding the noise and the negative values. Once the final layer gives the features the AWS Cloud is connected and the features of images and labels are stored in the database.

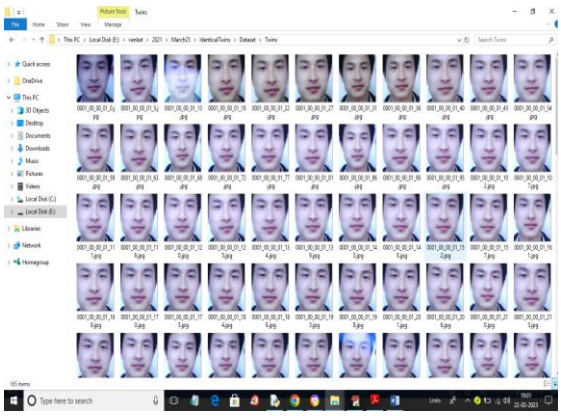
In real words twins faces are exists and this twins can utilize advantages to dupe peoples in examination or any other organizations. To detect such twins we are applying machine learning algorithms such as Naïve Bayes and Random Forest which may get trained on possible Real and Twins faces. Once after training we can input face to this trained model to identify weather face is Real or Twin. Before training we are applying various image processing techniques such as applying Bilateral Filters to enhance image quality and then convert image to Black & White format and then apply Object detection technique to detect face from image. This processed image will be input to Machine learning algorithm to train a model.

For training we are using below images dataset

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022



In above screen we have two folders called Twins and Real and just go inside any folder to view images like below screen



So by using above images we will evaluate performance of both Random Forest and Naïve Bayes Algorithm.

To implement this project we have designed following modules

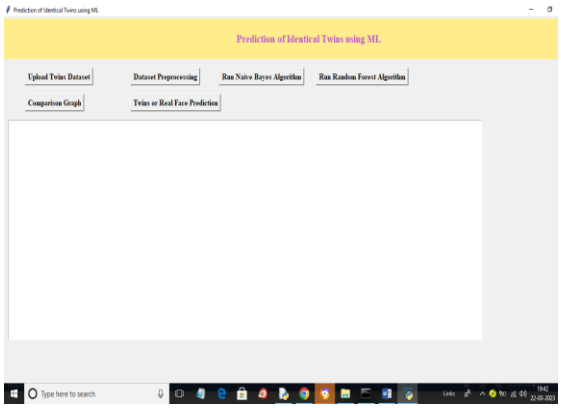
1) Upload Twins Dataset: using this module we will upload dataset to application and then apply filtration and object detection techniques
2) Dataset Preprocessing: using this module we will normalized and then shuffle and split dataset into train and test where application using 80% dataset for training and 20% for testing
3) Run Naive Bayes Algorithm: 80% processed train images will be input to Naive Bayes Algorithm to train a model and this model will be applied on 20% test images to calculate prediction accuracy
4) Run Random Forest Algorithm: 80% processed train images will be input to Random Forest Algorithm to train a model and this model will be applied on 20% test images to calculate prediction accuracy
5) Comparison Graph: using this module we will plot comparison graph between both algorithms
6) Twins or Real Face Prediction: using this module we will upload test images and then algorithm will predict weather image is real or belongs to twins.

SCREEN SHOTS

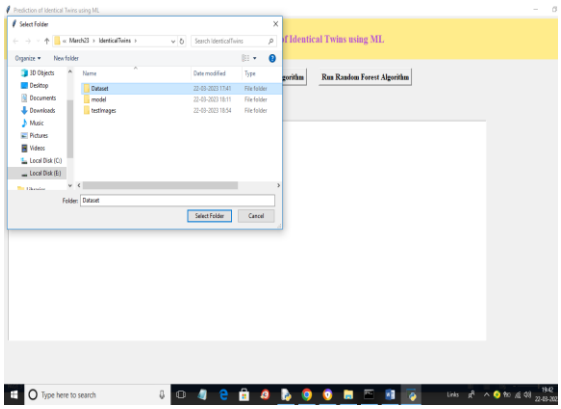
To run project double click on 'run.bat' file to get below screen

ISSN PRINT 2319 1775 Online 2320 7876

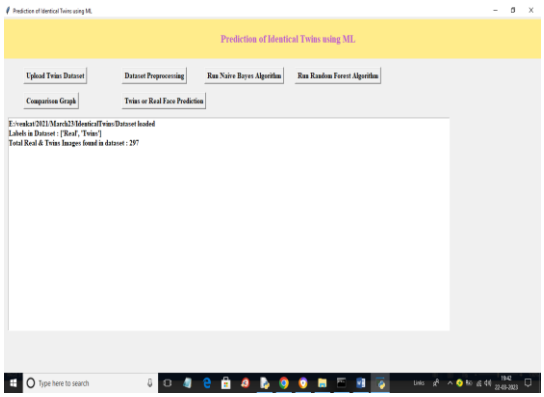
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022



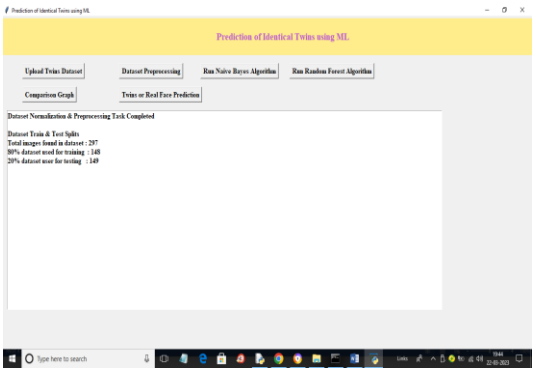
In above screen click on 'Upload Twins Dataset' button to upload dataset and get below output



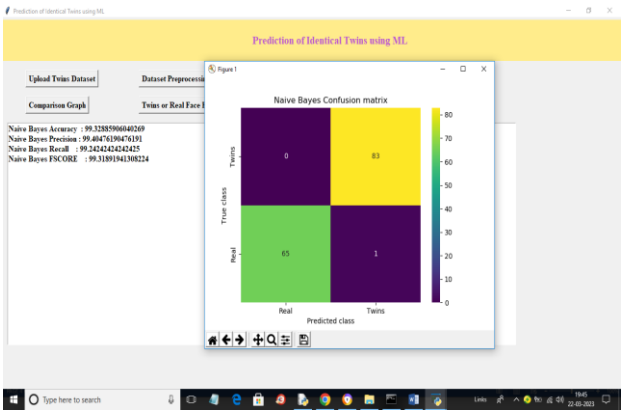
In above screen selecting and uploading 'Dataset' folder and then click on 'Select Folder' button to load dataset and get below output



In above screen we can see dataset loaded and we can see available labels and images in the dataset and now click on 'Dataset Preprocessing' button to normalize, shuffle and split dataset into train and test and will get below output



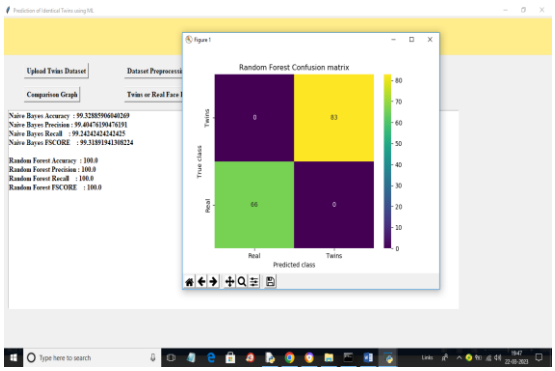
In above screen we can see dataset processed and we can see total images used for train and test and now click on 'Run Naïve Bayes Algorithm' button to train Naïve Bayes and get below output



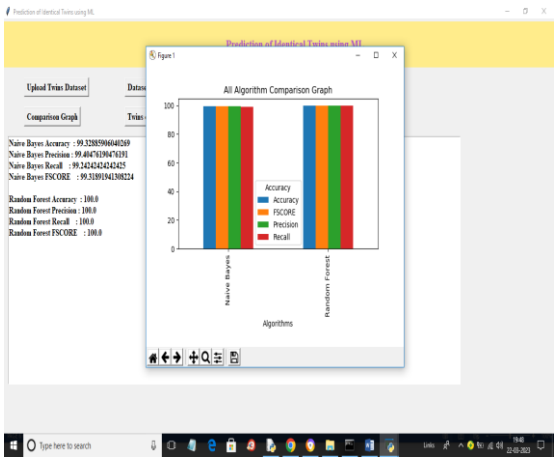
In above screen with Naïve Bayes we got accuracy as 99% and we can see other metrics also and in confusion matrix graph x-axis represents Predicted Labels and y-axis represents True Labels and green and yellow boxes contains Correct Prediction count and blue boxes represents incorrect prediction count

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

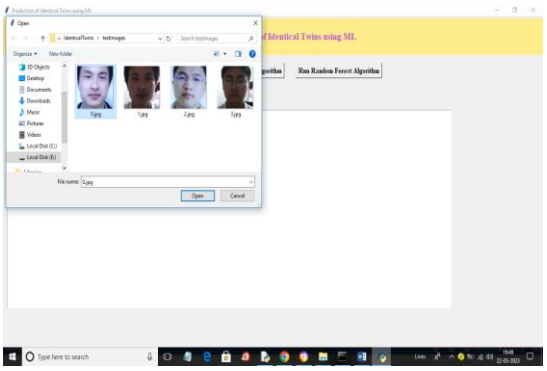
which is 1 only and now close above window and then click on 'Run Random Forest' button to train Random Forest and get below output



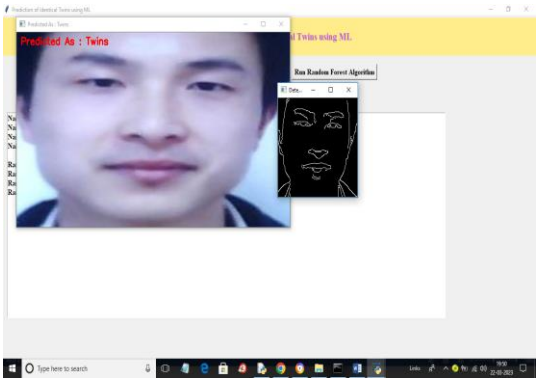
In above screen with Random Forest we got 100% accuracy and we can see confusion graph also and now click on 'Comparison Graph' button to get below graph



In above graph x-axis represents algorithm names and y-axis represents accuracy and other metrics in different colour bars and in both algorithms Random Forest got high performance and now click on 'Twins or Real Face Prediction' button to upload test image and get below output

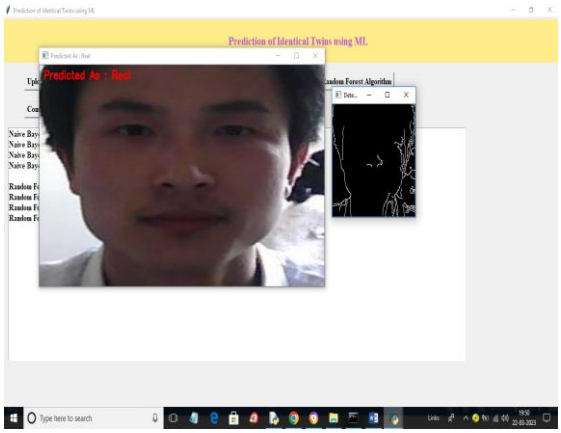


In above screen selecting and uploading '0.jpg' image and then click on 'Open' button to load image and get below output



In above screen in red colour text we can see image predicted as Twins and we can see detected object in face in black and white colour and similarly you can upload and test other images

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022



In above screen image predicted as real.

CONCLUSION

The current research work mainly concentrates on identifying the identical looking twins on basis of the features extracted. The proposed system uses Python as the programming language for development of the project. The proposed system uses various numberof frame works like Open Computer Vision Library (OpenCV), TesnorFlow, NumPy, Keras module and Convolutional Neural Networks (CNN) algorithm. Flask frame work is used to build a Graphical User Interface for the system. MySQLYog is used to connect to cloudand store in the database using the queries. Amazon Public Cloud is used to store the features in the cloud. Service used is Database as a service. CNN algorithm is designed to process data through multiple layers as shown in figure 3. This type of neuralnetworks is used in applications like image recognition or face recognition systems.The practical performance of this project is analyzed and it shows that the system works perfectly for the limited dataset. This proposed work recognizes the identical twins and displays the result on the bases of features extracted, which may be any one among the three possible outputs of the system.

REFERANCES

[1.]Face recognition techniques for differentiate similar faces and twin faces R. Prema ; P. Shanmuga priya IEEE 2018.

[2.]Face recognition using wavelet transforms-based feature extraction & spatial differentiation-based pre-processing Shridhar S Shanbhag; Sourabh B; K. Manikantan; S.Ramachandran IEEE 2018

[3.]Distinguishing Identical Twins by Face Recognition P. Jonathon Phillips, Patrick J. Flynn, Kevin W. Bowyer, Richard W. Vorder Bruegge, Patrick J. Grother, George W. Quinn, Matthew Pruitt IEEE 2018.

[4.]Differentiating Identical Twins by Using Conditional Face Recognition Algorithms Swati Y. Dhote, A. D. Gotmare , M. S.Nimbarte IJSR 2015.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

[5.]Face Recognition: Issues, Methods and Alternative Applications Waldemar Wójcik, Konrad Gromaszek and Muhtar Junisbekov INTEC 2016.

[6.]Face Recognition: From Traditional to Deep Learning Methods Daniel Saez Trigueros, Li Meng, Margaret Hartnett IJCT 2018.

[7.]Performance Analysis of Human Face Recognition Techniques Sharmila; Raman Sharma; Dhanajay Kumar; VaishaliPuranik; KritikaGautham2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU) Year: 2019 | Conference Paper | Publisher: IEEE.

[8.]Blur and Motion Blur Influence on Face Recognition Performance Katarina Knežević; EmilijaMandić; RankoPetrović; BrankaStojanovi2018 14th Symposium on Neural Networks and Applications (NEUREL) Year: 2018 | Conference Paper | Publisher: IEEE.

[9.]AlisterCordiner, Philip Ogunbona, and WanqingLi, "Face Detection Using Generalised Integral Image Features,"IEEE 2009.

[10.]JanarthanyNagendrarajah, M.U.S. Perera, "Recognition of Expression Variant Faces –APrinciple Component Analysis Based Approach for Access Control,"IEEE 2010.