ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 4, 2024

https://doi.org/10.5281/zenodo.13945262

USING MACHINE LEARNING TECHNIQUES TO IDENTIFY THE GENDER OF CYBER ATTACKERS

[1]MUDDHAM NIRMALA, [2]ALIGETI SOUMYA

[1]Assistant Professor,[2]Student

Department of CSE

Sree Chaitanya College of Engineering, Karimnagar

ABSTRACT

Although they may take many different forms, cyberattacks basically include the installation of a virus on a target computer system with the intention of gathering data, monitoring the user, or gaining remote control of the device. It is illegal behaviour that, if done extensively, has the potential to destroy whole countries' economy. To successfully counter cyberattacks, one must get a thorough knowledge of the motivations behind the assaults as well as their regular behaviour. As a result, the present research examines the gender of people who conduct cyberattacks using a social science viewpoint. To be more precise, this means using machine learning techniques like artificial neural networks (ANN), K-nearest neighbour (KNN), and support vector machines (SVM). The results show that SVM provides a trustworthy method for identifying the gender of cyberattackers, which aids profilers in looking into such crimes.

1.INTRODUCTION

Face recognition is a hybrid technology that combines biometic and machine learning methods, offering excellent accuracy and dependability. This technology may be used to automatically recognise a person's face from databases. Open computer vision has been more popular in recent years, finding usage in a variety of applications such as robots and security cameras. This technology is used in identity, authorisation, validation, and authentication processes. In affluent nations, the government generates facial recognition datasets by comparing suspected facial expressions with pre-trained datasets and database information. The three processes of face identification are as follows: (1) face detection; (2) feature extraction; and (3) face recognition. For accurate identification and tracking of moving objects, camera setup is crucial.

Important information regarding face shape is encoded in facial feature points. Accurate positioning and tracking of face feature points are crucial. Typically, a local search is used to identify and track each feature point by finding the best matching location. Research on facial recognition using edge-based detection is quite limited. The edges are easy to process in addition to conveying important face-related info. The Viola Jones technique uses AdaBoost to choose a small number of important characteristics before building a classifier. The Viola Jones technique, which concentrates on the positive aspects of the face, effectively combines more composite classifiers into a

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.13945262
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

cascade structure, increasing the detector's speed significantly.

The most crucial and time-consuming activity for police searching for criminals is criminal identification, which is also the most challenging since it must be done everywhere. Cities and other public areas with a high population density will provide more challenges. Manual identification methods may provide additional information about offenders. Thus, by identifying offenders' faces, this article suggests an automated criminal identification method. This will assist law enforcement in locating and apprehending offenders in public areas.

Identity theft may be detected in two ways. When using the Manual Identification System (MIS), police officers conduct searches on people in public areas to identify them. It takes a long time to provide the right care, and there's a danger that criminals won't receive the treatment they need because they'll be aware that police officers may flee the scene with ease. Since the MIS is taking longer than expected, we won't be able to adequately attend to everyone. However, using an automatic identification system (AIS) in a public setting eliminates the requirement for surveillance. Here, every procedure used in the system is automated.

2. LITERATURE SURVEY

A Study on various state of the art of the Art Face Recognition System Using Deep Learning Techniques

Sketch recognition is one of the most important areas that have evolved as an integral component adopted by the agencies of law administration in current trends of forensic science. Matching of derived sketches to photo images of face is also a difficult assignment as the considered sketches are produced upon the verbal explanation depicted by the eye witness of the crime scene and may have scarcity of sensitive elements that exist in the photograph as one can accurately depict due to the natural human error. Substantial amount of the novel research work carried out in this area up late used recognition system through traditional extraction and classification models. But very recently, few researches work focused on using deep learning techniques to take an advantage of learning models for the feature extraction and classification to rule out potential domain challenges.

A facial expression recognition system using robust face features from depth videos and deep learning

This work proposes a depth camera-based robust facial expression recognition (FER) system that can be adopted for better human machine interaction. Although video-based facial expression analysis has been focused on by many researchers, there are still various problems to be solved in this regard such as noise due to illumination variations over time. Depth video data in the helps to make an FER system person-independent as pixel values in depth images are distributed based on distances from a depth camera. Besides, depth images should resolve some

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.13945262
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

privacy issues as real identity of a user can be hidden. The accuracy of an FER system is much dependent on the extraction of robust features. Here, we propose a novel method to extract salient features from depth faces that are further combined with deep learning for efficient training and recognition. Eight directional strengths are obtained for each pixel in a depth image where signs of some top strengths are arranged to represent unique as well as robust face features, which can be denoted as Modified Local Directional Patterns (MLDP).

Deep Learning Face Representation by Joint Identification-Verification

The key challenge of face recognition is to develop effective feature representations for reducing intra-personal variations while enlarging inter-personal differences. In this paper, we show that it can be well solved with deep learning and using both face identification and verification signals as supervision. The Deep Identification-verification features (DeepID2) are learned with carefully designed deep convolutional networks. The face identification task increases the inter-personal variations by drawing DeepID2 features extracted from different identities apart, while the face verification task reduces the intra-personal variations by pulling DeepID2 features extracted from the same identity together, both of which are essential to face recognition. The learned DeepID2 features can be well generalized to new identities unseen in the training data. On the challenging LFW dataset, 99.15% face verification accuracy is achieved. Compared with the best previous deep learning result on LFW, the error rate has been significantly reduced by 67%.

3. EXISTING SYSTEM:

In existing system, we propose an automatic criminal identification system for Police Department to enhance and upgrade the criminal distinguishing into a more effective and efficient approach. Using technology, this idea will add plus point in the current system while bringing criminals spotting to a whole new level by automating tasks. Technology working behind it will be face recognition, from the footage captured by the CCTV cameras; our system will detect the face and recognize the criminal who is coming to that public place. The captured images of the person coming to that public place get compared with the criminal data we have in our database. If any person's face from public place matches, the system will display their image on the system screen and will give the message with their name that the criminal is found and present in this public place. This system matching more than 80% of the captured images with database images.

DISADVANTAGES:

- It doesn't efficient for large volume of data's.
- Training time is more.
- The process is implemented without removing the noise.
- Prediction is not accurate.

4. PROPOSED SYSTEM:

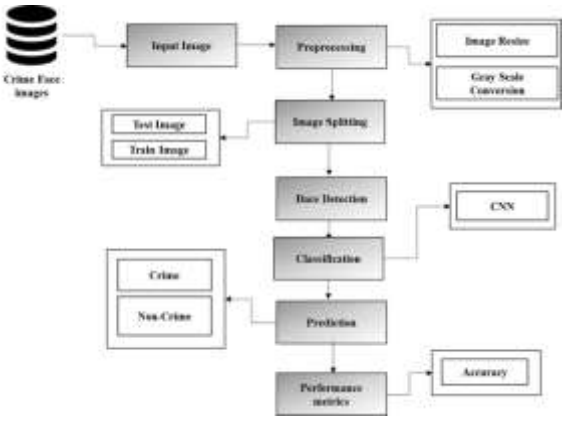
INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.13945262
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

In this system, the Crime face images dataset is collected from dataset repository. Then, we have to implement the image pre-processing step. Here, we can implement the image resize and grayscale conversion. Then, we can split the images into test images and train images. The test image is used for prediction and train image is used for evaluation. Then, we have to implement the deep and machine learning algorithm such as Convolutional Neural Network (CNN) and random forest. The experimental results shows that the accuracy. Finally, we can identify the input image is either crime or not. If the person is crime, the system can display the details of crime. Finally, the experimental results shows that accuracy and error rate.

ADVANTAGES:

- It is efficient for large number of datasets.
- Time consumption is low.
- The process is implemented with removing noise.
- It display the crime details.
- Prediction is accurate.

5. SYSTEM ARCHITECTURE



6. IMPLEMENATION

Modules Description

- Input image
- Preprocessing
- Image splitting
- Classification
- Performance Estimation

MODULES DESCRIPTION:

IMAGE SELECTION:

- The dataset, crime face image dataset is implemented as input. The dataset is taken from dataset repository.
- The input dataset is in the format '.png, '.jpg.
- In this step, we have to read or load the input image by using the imread () function.
- In our process, we are used the tkinter file dialogue box for selecting the input image.

IMAGE PREPROCESSING:

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024
https://doi.org/10.5281/zenodo.13945262

- In our process, we have to resize the image and convert the image into gray scale.
- To **resize an image**, you call the resize () method on it, passing in a two-integer tuple argument representing the width and height of the resized image.
- The function doesn't modify the used image; it instead returns another Image with the new dimensions.
- Convert an Image to **Grayscale** in Python Using the Conversion Formula and the matplotlib Library.
- We can also convert an image to grayscale using the standard RGB to grayscale conversion formula that is imgGray = 0.2989 * R + 0.5870 * G + 0.1140 * B.

IMAGE SPLITTING:

- During the machine learning process, data are needed so that learning can take place.
- In addition to the data required for training, test data are needed to evaluate the performance of the algorithm in order to see how well it works.
- In our process, we considered 70% of the input dataset to be the training data and the remaining 30% to be the testing data.
- Image splitting is the act of partitioning available data into two portions, usually for cross-validator purposes.
- One Portion of the data is used to develop a predictive model and the other to evaluate the model's performance.
- Separating data into training and testing sets is an important part of evaluating data mining models.
- Typically, when you separate a data set into a training set and testing set, most of the data is used for training, and a smaller portion of the data is used for testing.

CLASSIFICATION:

- In our process, we have to implement the deep and machine learning algorithm such as Convolutional Neural Network (CNN) and RF.
- **CNN** In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery.
- They have applications in image and video recognition, recommender systems, image classification, medical image analysis, natural language processing, brain-computer interfaces, and financial time series.
- CNNs are regularized versions of multilayer perceptron's. Multilayer perceptron's usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer.
- **Random forest** is a commonly-used machine learning algorithm trademarked by Leo Breiman and

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024
https://doi.org/10.5281/zenodo.13945262

Adele Cutler, which combines the output of multiple decision trees to reach a single result.

- Its ease of use and flexibility have fueled its adoption, as it handles both classification and regression problems.

RESULT GENERATION:

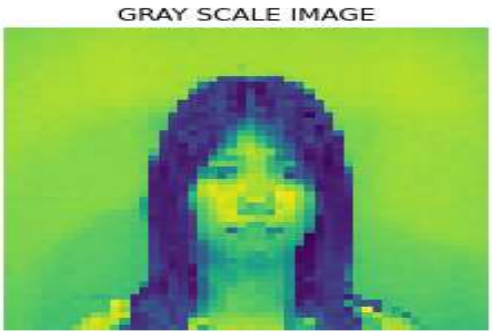
The Final Result will get generated based on the overall classification and prediction. The performance of this proposed approach is evaluated using some measures like

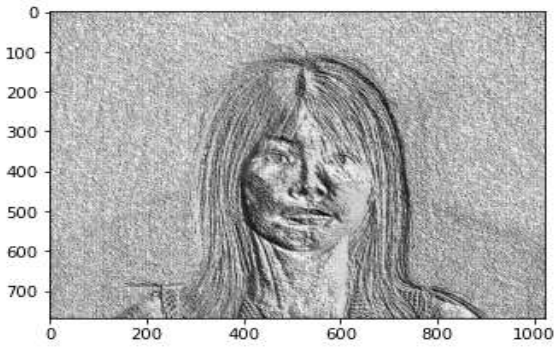
- **Accuracy**

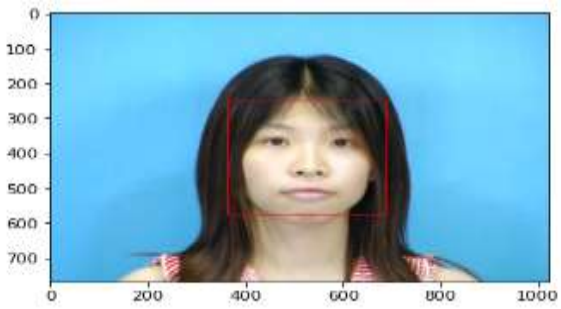
Accuracy of classifier refers to the ability of classifier. It predicts the class label correctly and the accuracy of the predictor refers to how well a given predictor can guess the value of predicted attribute for a new data.

AC= (TP+TN)/ (TP+TN+FP+FN)

7. SCREEN SHOTS







INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024
https://doi.org/10.5281/zenodo.13945262

8. CONCLUSION AND FEATURE ENHANCEMENT

Finally, we deduce that the facial photos came from a dataset store. Here, picture pre-processing methods like greyscale conversion and image resizing are put into practice. We are using machine learning and deep learning algorithms like CNN and RF. Subsequently, the experimental findings verify that correctness. This classification method helps us determine if the individual in the input picture is a criminal or not.

FUTURE ENHANCEMENT

For improved performance or efficiency, we will hybridise transfer learning in future work, combine two distinct machine learning algorithms, or mix two distinct deep learning algorithms.

REFERENCES

[1] "The Goode Intelligence Biometric Survey 2021." Goode Intelligence. Apr. 2021. [Online]. Available: https://www.goodeintelligence.com/report/the-goode-intelligence-biometric-survey-2021/

[2] S. Bhattacharjee, A. Mohammadi, A. Anjos, and S. Marcel, "Recent advances in face presentation attack detection," in Handbook of Biometric Anti-Spoofing. Cham, Switzerland: Springer, 2019, pp. 207–228.

[3] P. Bontrager, W. Lin, J. Togelius, and S. Risi, "Deep interactive evolution," in Proc. Int. Conf. Comput. Intell. Music Sound Art Des., 2018, pp. 267–282.

[4] H. H. Nguyen, J. Yamagishi, I. Echizen, and S. Marcel, "Generating master faces for use in performing wolf attacks on face recognition systems," in Proc. IJCB, 2020, pp. 1–10.

[5] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch, "Face recognition systems under morphing attacks: A survey," IEEE Access, vol. 7, pp. 23012–23026, 2019.

[6] P. Bontrager, A. Roy, J. Togelius, N. Memon, and A. Ross, "DeepMasterPrints: Generating MasterPrints for dictionary attacks via latent variable evolution," in Proc. BTAS, 2018, pp. 1–9.

[7] M. Une, A. Otsuka, and H. Imai, "Wolf attack probability: A new security measure in biometric authentication systems," in Proc. ICB, 2007, pp. 396–406.

[8] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in Proc. CVPR, 2019, pp. 4690–4699.

[9] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," in Proc. ICLR, 2014.

[10] I. Goodfellow et al., "Generative adversarial nets," in Proc. NIPS, 2014, pp. 2672–2680.

[11] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in Proc. ICML, 2017, pp. 214–223.

[12] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in Proc. NIPS, 2017, pp. 5769–5779.

[13] H. Huang, Z. Li, R. He, Z. Sun, and T. Tan, "IntroVAE: Introspective variational

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.13945262
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

autoencoders for photographic image synthesis," in Proc. NIPS, 2018, pp. 52–63.
[14] A. Razavi, A. van den Oord, and O. Vinyals, "Generating diverse highfidelity images with VQ-VAE-2," in Advances in Neural Information Processing Systems, 2019, pp. 14866–14876.
[15] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," in Proc. ICLR, 2018.