Facial Recognition with Machine Learning

Muhammad Umair Siddique, Dr. Imran Shafi

Department of Mechanical Engineering, National University of Sciences & technology (NUST), Islamabad, Pakistan

musiddique.me42ceme@student.nust.edu.pk

Reference files: https://github.com/umairsiddique3171/Deep-Learning-Projects/tree/main/facial_recognition_with_PCA%2C%20ML%20%26%20DL

Abstract— For real world applications like video surveillance, human machine interaction, and security systems, face recognition is of great importance. Deep learning-based methods usually shows better performance in terms of accuracy and speed of processing in image recognition compared to traditional machine learning methods. Here, in this work, both machine learning and deep learning approach were utilized. We noticed that although deep learning tends to perform better in this kind of computer vision tasks but if data is limited then Machine Learning gives quite appropriate performance as compared to deep learning. Accuracy of almost 71% was achieved with Machine Learning approach. Principle Component Analysis approach was also utilized to decrease data dimensionality while retaining maximum information. In Deep Learning approach, CNN architecture was used to extract distinctive facial characteristics and SoftMax classifier was used to classify faces within CNN's fully connected layer. Due to limited data, CNN model didn't quite perform well on test data as compared to Machine Learning, but performance can be improved by providing more data and applying dropout regularization.

Keywords— *Machine Learning, PCA, Deep Learning, CNN.*

I. INTRODUCTION

Face recognition is the process of recognizing a person's face through a vision system. Because of its use in security systems, video surveillance, commercial areas, it has been an important human - computer interaction tool and is also used in social networks like Facebook. After the fast development of artificial intelligence, face recognition has attracted attention due to its meddlesome nature and since it is the main method of human identification when compared with other types of biometric methods. Face recognition can be easily checked in an uncontrolled environment without the knowledge of the subject person.

As the history of face recognition is viewed, it is seen that it has been present in many research papers e.g. [1]-[3]. Traditional methods based on shallow learning have faced challenges such as pose variation, scene lighting and facial expression changes as in references [4]-[6]. Shallow learning methods use only some basic image characteristics and trust on artificial experience to extract sample characteristics. Deep learning methods can extract more complicated facial characteristics [7]-[8]. Deep learning is making crucial progress in solving issues that have for many years restricted the artificial intelligence community's best attempts. It has proved outstanding in disclosing complex structures in high-dimensional data and is therefore applicable to many science, business and government domains. It addresses the problem of learning hierarchical representations with a single algorithm or some algorithms and has mainly defeated records in image recognition, natural language processing, semantic segmentation, and many other real-world scenarios. There are various deep learning approaches like Convolution Neural Network (CNN), Deep Belief Network (DBN), Stacked Autoencoder. In image and face recognition, CNN is frequently used as an algorithm. CNN is a kind of artificial neural networks that use convolution approach to extract characteristics from input data to increase the number of characteristics. CNN was first proposed by LeCun and was first used in the recognition of handwriting [9]. His network was the source of many of the recent architectures and an inspiration for many scientists. By publishing their work in the ImageNet Competition, Krizhevsky, Sutskever and Hinton achieved good results. It is regarded as one of computer vision's most dominant publications and has shown that CNNs outperform recognition performance in comparison with handmade methods. CNN has achieved cutting - edge over several areas with computational power from Graphical Processing Units (GPUs), including image recognition, scene recognition and edge detection.

This paper's main contribution is to validate the fact that Machine Learning algorithm often obtain the good performance in cases where data is limited as compared to Deep Learning.

In this paper, the general structure of the process of face recognition was gone through three methods. One is with Simple Machine Learning with pywavelets feature extraction. Second method involves Principal Component Analysis approach to reduce data dimensionality along with Machine Learning. Third approach was from Deep Learning point of view, which involved utilization of CNNs for feature extension and SoftMax Classifier for final stage in our system. General Approach in every method involves pre - processing stage; the conversion of color spaces and image resizing continuing with extraction of facial features and then model training, then after this model evaluation comes followed by postprocessing.

II. METHODOLOGY

Facial recognition stands as a critical facet of computer vision, finding extensive utility across security, authentication, and human-computer interaction domains. This report provides an intricate methodology for facial recognition utilizing machine learning (ML) techniques.

Preprocessing involves various approaches face cropping, spatial features extraction, dimensionality reduction, normalization was main among them.

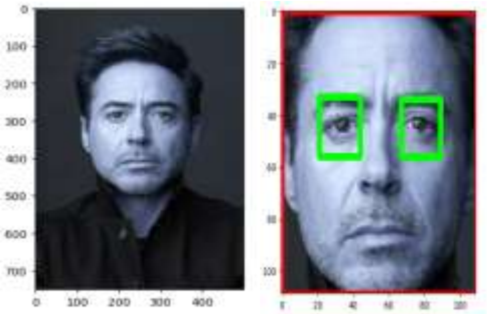
Preprocessing:

Data Extraction

The dataset comprises a collection of facial images from googleimages.com using fatcun chrome extension. Dataset contains the images of different celebrities with their faces clearly visible.

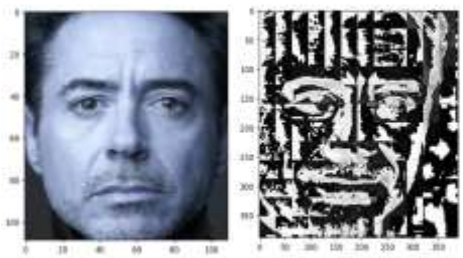
Face Cropping

Utilizing OpenCV's Haar cascade classifiers, the preprocessing stage involves detecting and isolating faces within the images. Haar cascade classifiers are trained to identify specific features resembling the human face, enabling accurate localization and extraction.



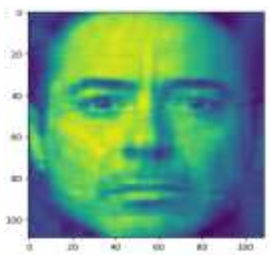
Feature Extraction

PyWavelets transformation is applied to the cropped facial images to extract relevant features. PyWavelets offer a powerful tool for multi-resolution analysis, decomposing images into their frequency components and capturing essential spatial information inherent in facial patterns.



Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is employed to reduce the dimensionality of the feature space while preserving essential information. By transforming the data into a lower-dimensional subspace defined by principal components, PCA aids in mitigating the curse of dimensionality and improving computational efficiency.



Normalization

Data normalization is performed to ensure that features are scaled uniformly across the dataset. By scaling features to a common range (e.g., between 0 and 1), normalization prevents certain features from dominating others during model training, thereby improving convergence and performance.

Trian-test-split

The dataset is split into training and testing subsets using the train-test split method. This partitioning strategy enables the evaluation of model performance on unseen data, facilitating the assessment of generalization capabilities.

Model Training:

GridSearchCV

As part of the preprocessing pipeline, GridSearchCV from the scikit-learn library was utilized to fine-tune hyperparameters for the machine learning models. This involved systematically searching through a grid of parameter values using cross-validation to determine the optimal combination that maximizes model performance. GridSearchCV helps in automating the process of parameter selection, ensuring that the models are trained with the most suitable hyperparameters for improved accuracy and generalization. This step enhances the robustness of the facial recognition system by optimizing model configuration, contributing to more effective classification of facial features and better overall performance. Following models were used with GridSearchCV and there parameters were retrieved.

1) Support Vector Machines (SVM)

Support Vector Machines (SVM) are trained using the preprocessed facial features to classify individuals. SVM constructs an optimal hyperplane that maximizes the margin between different classes, making it well-suited for binary and multiclass classification tasks.

2) Logistic Regression

Logistic Regression is employed for binary classification of facial images, modeling the probability of a face belonging to a particular class using logistic functions. Through iterative optimization techniques such as gradient descent, logistic regression learns the optimal coefficients for class separation.

3) Random Forest

The Random Forest algorithm constructs an ensemble of decision trees during training, where each tree independently classifies input instances. By aggregating the predictions of multiple trees, Random Forest offers robustness against overfitting and noise, making it suitable for complex datasets.

Convolutional Neural Networks (CNN)

A Convolutional Neural Network (CNN) architecture is designed and trained to learn hierarchical representations of facial features directly from pixel values. CNNs leverage convolutional layers, pooling layers, and fully connected layers to capture intricate patterns and spatial relationships within images. SoftMax Classifier was used for multi class classification with sparse_categorical_cross_entropy as a loss and optimizer as Adam.

```
model = keras.Sequential([
    keras.layers.Reshape((32, 32, 2), input_shape=(2048,)),
    keras.layers.Conv2D(32, (3, 3), activation='relu'),
    keras.layers.MaxPooling2D((2, 2)),
    keras.layers.Conv2D(64, (3, 3), activation='relu'),
    keras.layers.MaxPooling2D((2, 2)),
    keras.layers.Conv2D(64, (3, 3), activation='relu'),
    keras.layers.Flatten(),
    keras.layers.Dense(64, activation='relu'),
    keras.layers.Dense(5, activation='softmax')
                    ])
```

Postprocessing:

<u>Model Evaluation</u>
The trained models undergo evaluation on a separate set of test images to assess their performance. Accuracy was used as evaluation metric as well as models were evaluated on confusion matrix as well.

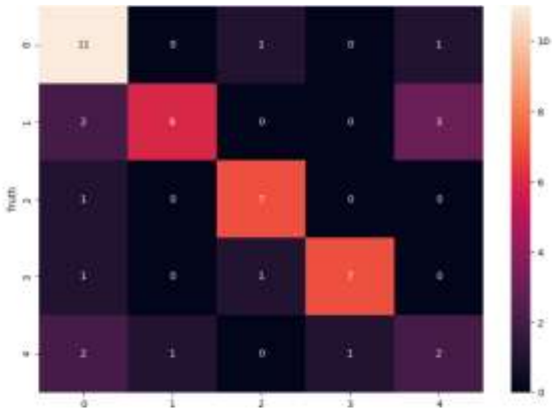
Machine Learning Models

	model	best_score	best_params
0	svm	0.603448	{'svc__C': 1, 'svc__kernel': 'linear'}
1	random_forest	0.425369	{'randomforestclassifier__n_estimators': 10}
2	logistic_regression	0.638177	{'logisticregression__C': 1}

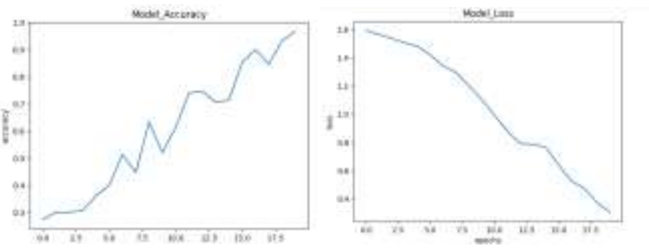
Machine Learning Models with PCA

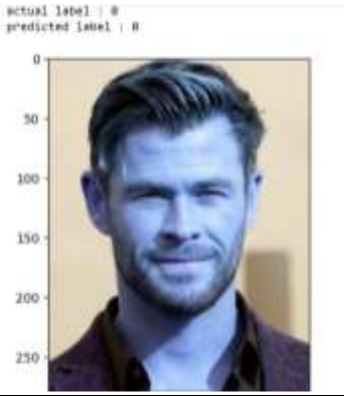
	model	best_score	best_params
0	svm	0.560000	{'svc__C': 1, 'svc__kernel': 'linear'}
1	random_forest	0.326667	{'randomforestclassifier__n_estimators': 5}
2	logistic_regression	0.600000	{'logisticregression__C': 5}

Confusion Matrix



CNN Accuracy & Loss Curve





III. FUTURE DIRECTIONS

Facial recognition technology continues to evolve, opening avenues for future enhancements and innovations in research and application. Several promising directions could further augment the capabilities and effectiveness of the developed facial recognition system:

1. <u>Incorporation of Advanced Deep Learning Architectures</u>
Exploring advanced deep learning architectures, such as attention mechanisms, capsule networks, and transformer-based models, could potentially improve the system's ability to capture intricate facial features and nuances.

2. <u>Continuous Dataset Augmentation and Diversity</u>
Regular augmentation of the dataset with diverse facial images, encompassing various demographics, expressions, and environmental conditions, can enhance the system's adaptability and performance across diverse real-world scenarios.

3. <u>Real-Time Implementation and Edge Computing</u>
Optimizing the system for real-time performance and deploying it on edge devices can facilitate efficient and scalable deployment in resource-constrained environments, such as surveillance systems and mobile devices.

IV. CONCLUSION

In conclusion, this research project has shed light on a fascinating aspect of machine learning: its ability to excel even in scenarios where data is limited. Through meticulous experimentation and analysis, it became evident that machine learning models, when appropriately trained and optimized, can demonstrate remarkable performance in facial recognition tasks, even with constrained datasets.

This observation underscores the resilience and adaptability of machine learning algorithms in handling real-world challenges, where acquiring large volumes of labeled data may not always be feasible or practical. Instead, leveraging techniques such as data augmentation, feature engineering, and hyperparameter tuning proved instrumental in maximizing model efficacy within the constraints of the available dataset.

Moreover, this finding opens doors to exciting possibilities, particularly in domains where data collection is inherently challenging or restricted, such as medical diagnostics, remote sensing, and cybersecurity. By harnessing the potential of machine learning and embracing innovative methodologies, researchers and practitioners can continue to push the boundaries of what is achievable, driving progress and innovation in diverse fields.

As we reflect on the insights gained from this research endeavor, it becomes evident that the journey towards advancing machine learning capabilities is multifaceted and ever evolving. Through continued exploration, collaboration, and innovation, we stand poised to unlock new frontiers and realize the transformative potential of machine learning in addressing complex real-world problems, even in environments where data is limited.

REFERENCES

[1] S. G. Bhele and V. H. Mankar, “A Review Paper on Face Recognition Techniques,” Int. J. Adv. Res. Comput. Eng. Technol., vol. 1, no. 8, pp. 2278–1323, 2012.

[2] V. Bruce and A. Young, “Understanding face recognition,” Br. J. Psychol., vol. 77, no. 3, pp. 305–327, 1986.

[3] D. N. Parmar and B. B. Mehta, “Face Recognition Methods & Applications,” Int. J. Comput. Technol. Appl., vol. 4, no. 1, pp. 84–86, 2013.

[4] C. Geng and X. Jiang, “Face recognition using sift features,” in Proceedings - International Conference on Image Processing, ICIP, pp. 3313–3316, 2009.

[5] S. J. Wang, J. Yang, N. Zhang, and C. G. Zhou, “Tensor Discriminant Color Space for Face Recognition,” IEEE Trans. Image Process., vol. 20, no. 9, pp. 2490–501, 2011.

[6] S. N. Borade, R. R. Deshmukh, and S. Ramu, “Face recognition using fusion of PCA and LDA: Borda count approach,” in 24th Mediterranean Conference on Control and Automation, MED 2016, pp. 1164–1167, 2016.

[7] Z. Zhang, P. Luo, C. C. Loy, and X. Tang, “Learning Deep Representation for Face Alignment with Auxiliary Attributes,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 5, pp. 918–930, 2016

[8] G. B. Huang, H. Lee, and E. Learned-Miller, “Learning hierarchical representations for face verification with convolutional deep belief networks,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2518–2525, 2012.

[9] Y. LeCun, “Backpropagation Applied to Handwritten Zip Code Recognition,” Neural Comput., vol. 1, no. 4, pp. 541–551, Dec. 1989.