iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: https://doi.org/10.22214/ijraset.2023.57436

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

Face Recognition using SVM with the Help of Kernel Characteristics Function

A. Sri Varshini[1], Y. Srija[2], Y. Srikanth[3], H. Srikanth[4], D. Srikar[5], P. Srikar[6], Ragipati Karthik[7]
[1, 2, 3, 4, 5, 6]*B.Tech, (AIML), School of Engineering MallaReddy University Hyderabad, India*
[7]*Professor, School of Engineering, AIML, Department MallaReddy UniversityHyderabad, India*

Abstract: ***This project focuses on building a face recognition system using the Labeled Faces in the Wild (LFW) dataset. The dataset includes diverse facial images of individuals across various age groups, such as kids, women, men, and the elderly. The model employs a Support Vector Machine (SVM) with a radial basis function (RBF) kernel for effective classification. Preprocessing techniques like normalization, grayscale conversion, and face alignment enhance the dataset's quality. The system's performance is assessed using accuracy, precision, recall metrics, and a Receiver Operating Characteristic (ROC) curve. The project aims to create a reliable and versatile face recognition solution applicable in security, authentication, and surveillance scenarios.***

I. INTRODUCTION

Face recognition technology has developed as a critical tool with applications ranging from security to user authentication. Using the well-known Labelled Faces in the Wild (LFW) dataset, this study investigates the beneficial combination of Support Vector Machines (SVM) and dimensionality reduction approaches for facial identification. The dataset's numerous facial photos make it an excellent standard for assessing the robustness of facial recognition algorithms.

Our method requires building a full pipeline that incorporates Principal Component Analysis (PCA) for dimensionality reduction and SVM for classification. Through thorough preparation of the LFW dataset and its partition into different training and testing sets, the objective is not just to achieve high accuracy but also to solve issues caused by the curse of dimensionality.

We use common measures such as accuracy, classification reports, and confusion matrices to evaluate our model. Furthermore, color-coded visualisations examine situations when the model's predictions differ from actual results, providing useful information for prospective changes. This rigorous review sets the ground for a detailed examination of the model's performance in facial recognition tasks.

In the final section of our research, we widen the scope to include Receiver Operating Characteristic (ROC) curve analysis. We create ROC curves and compute areas under the curve for each class using a One-Vs-Rest classification approach using SVM decision function scores. This method gives a more nuanced insight of the model's discriminatory power across different face identities, expanding our understanding of its performancelandscape.

II. LITERATURE REVIEW

Face recognition has made substantial advances in algorithmic sophistication and dataset variety. Because of its capacity to handle high-dimensional data and display strong performance in classification settings, Support Vector Machines (SVMs) have frequently been used in previous research for facial recognition tasks. SVMs have proved to be helpful in correcting issues caused by changes in lighting, face emotions, and positions. However, when presented with datasets containing a significant number of facial traits, these studies frequently suffer from the curse of dimensionality.

Dimensionality reduction approaches, such as Principal Component Analysis (PCA), have proved crucial in addressing the issues associated with high-dimensional face data. PCA assists in collecting the most significant characteristics while rejecting those that are less useful, boosting computing efficiency and model interpretability. Prior research has shown that PCA can improve the performance of facial recognition systems, particularly in circumstances with limited computational resources.

The Labelled Faces in the Wild (LFW) dataset has been used to assess the robustness of facial recognition models. While prior research employed the LFW dataset well, they frequently focused on particular factors such as algorithmic performance or dataset properties. Few have thoroughly merged SVMs with PCA, addressing the subtle balance between dimensionality reduction and classification accuracy on this difficult dataset.

©IJRASET: All Rights are Reserved

Our suggested technique, which uses the LFW dataset, combines SVMs and PCA in a pipeline for facial recognition tasks, in contrast to the literature that already exists. Our strategy aims to use the advantages of both approaches by merging them, improving the model's capacity to manage intricate face changes. The comprehensive assessment measures, including as confusion matrices, accuracy, and classification reports, offer a nuanced picture of the model's performance and enable a straight comparison with previous research.

Furthermore, we investigate Receiver Operating Characteristic (ROC) curve analysis, an aspect that has been neglected in previous studies. This research offers a more thorough evaluation across various face identities and sheds light on our model's discriminative skills.

In conclusion, while previous research has established the framework for face recognition technology, our method provides a comprehensive integration of SVMs, PCA, and in- depth performance assessments on the difficult LFW dataset. By drawing comparisons, we hope to highlight the progress made by our suggested technique and add to the changing field of facial recognition research.

III. PROBLEM STATEMENT

Applications ranging from user identification to security systems now require facial recognition technologies. Even with significant advancements, there are still difficulties in creating reliable and effective face recognition models. When addressing high-dimensional facial data, existing methods frequently struggle with the curse of dimensionality, which can result in increased computing complexity and possible performance reduction. Moreover, although Support Vector Machines (SVMs) have demonstrated efficacy in classification assignments, there is still room for improvement in their amalgamation with dimensionality reduction methodologies, most notably Principal Component Analysis (PCA), when applied to intricate datasets such as Labelled Faces in the Wild (LFW).

The LFW dataset presents special difficulties for face recognition models because of its diversity and complexity in the actual world. The dataset's variations in lighting, positions, and facial expressions call for a complex method to produce accurate and dependable identification. Recent studies frequently lacks a thorough integration of PCA and SVMs designed to handle the complexities of the LFW dataset.

Given these difficulties, this study addresses two issues: first, the need for an efficient combination of SVMs and PCA to handle high-dimensional facial data; and second, the requirement for a comprehensive assessment of the model's performance on the difficult LFW dataset. By putting forth a thorough pipeline that makes use of SVMs and PCA, our research seeks to close these gaps and offer a sophisticated response to the difficulties associated with facial recognition on intricate datasets. This work aims to provide insights that improve our understanding of facial recognition systems and open the door to improved performance in practical applications through a thorough investigation.

IV. METHODOLOGY

This study's technique is specifically designed to improve facial recognition performance by combining Principal Component Analysis (PCA) and Support Vector Machines (SVMs) in an efficient manner. We start by carefully preparing the Labelled Faces in the Wild (LFW) dataset to make sure that a variety of face identities are represented in the training and testing subsets. The informative material that is maintained is enhanced by selectively applying Principal Component Analysis to minimize the dimensionality of face characteristics. Then, an SVM classifier with a Radial Basis Function (RBF) kernel is introduced, which makes it easier to learn the discriminative patterns needed for facial recognition.

A. Architecture

The suggested face recognition system's architecture is intended to provide a systematic and efficient framework for incorporating Support Vector Machines (SVMs) and Principal Component Analysis (PCA). The system is divided into phases, each of which contributes to the ultimate objective of developing a strong and accurate facial recognition model.

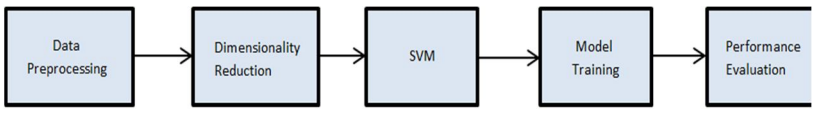


Fig 1.Architecture

1) *Data Preprocessing:* The data preparation stage ensures that the dataset is standardised and representative for future model training. Normalisation and grayscale conversion are performed on raw face photos from the Labelled Faces in the Wild (LFW) collection. This procedure improves the model's generalisation skills across a wide range of face traits, resulting in consistent input for later phases.
2) *Dimensionality Reduction:* Principal Component Analysis (PCA) is used to minimize the dimensionality of facecharacteristics while keeping crucial information. PCA extracts the most significant information by translating high- dimensional face traits into a lower-dimensional space. The selection of 150 primary components finds a compromise between computational efficiency and information retention.
3) *Support Vector Machine:* The SVM stage's goal is to categorise face characteristics based on previously learnt discriminative patterns. To achieve successful recognition, an SVM classifier with a Radial Basis Function (RBF) kernel is used. To address possible discrepancies in the distribution of face identities, class weights are adjusted to ensure equitable representation across different classes.
4) *Model Training:* The goal of model training is to teach the SVM to recognise and discriminate facial identities based on previously learnt discriminative patterns. The SVM is trained using PCA-transformed features and labels from the LFW dataset's selected training subset. This procedure is critical for optimizing the performance of the SVM in subsequent recognition challenges.
5) *Performance Evaluation:* In the performance assessment step, the trained model's effectiveness and accuracy are evaluated. On the testing subset of the LFW dataset, the SVM model is thoroughly assessed, and metrics like accuracy, precision, recall, and F1-score are calculated. This thorough assessment sheds light on the model's performance and capabilities for a range of face identities.

B. Design

The data flow diagram (DFD) that is being shown here provides a visual depiction of the complex information flow that takes place within a facial recognition system. From the first collection of unprocessed face photos to the last assessment of the performance of the trained model, this figure offers an organized perspective of the data flow across several phases.

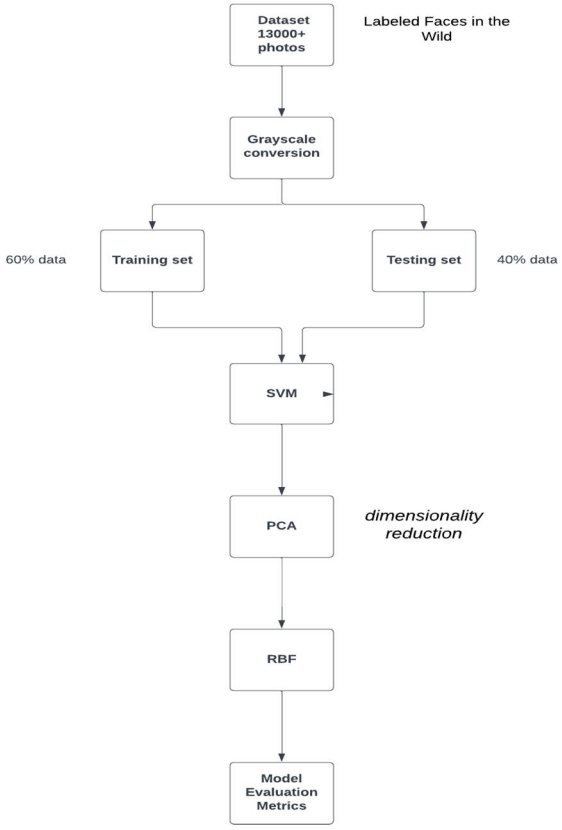


Fig 2 Design

V. EXPERIMENTAL RESULTS

Using Support Vector Machines (SVM) with a Radial Basis Function (RBF) kernel and Principal Component Analysis (PCA) for dimensionality reduction, we implemented a facial recognition pipeline in this project. With a minimum of 60 facesper individual, the Labelled Faces in the Wild (LFW) dataset is the one used.

A 90-10 split ratio was used to divide the dataset into training and testing sets. The chosen dataset was used to train the SVM model, which performed well in identifying faces. 150 main components were extracted using PCA during the training procedure, producing a well-generalized model.

The following evaluation metrics can be used to determine our face recognition model's performance:

1) *Accuracy Score:* One of the most important measures of our facial recognition model's overall performance is its accuracy. We can declare with assurance that the model successfully classified the great majority of cases properly throughout the testing phase, with an accuracy score of about 87.41%. Even though accuracy isa key indicator, for a more thorough knowledge of the model's performance, it is necessary to go further into additional assessment metrics.
2) *Classification Report:* An extensive analysis of the model's performance for every face identification class is included in the classification report. F1-score, recall, and precision are crucial measures that provide insightful information. Recallevaluates the capacity to record every positive event, precision analyses the accuracy of positive predictions, and the F1-score strikes a compromise between the two. We will be able to pinpoint individual strengths and possible areas for development across a range of face identities by analysing the categorization report.
3) *Confusion Matrix:* A thorough evaluation of the confusion matrix is required to comprehend the distribution of accurate and wrong predictions. True positives, true negatives, false positives, and false negatives are all fully defined, providing a detailed view of the model's performance. The confusion matrix is useful for detecting classes that may provide problems for the model, allowing for focused improvements.
4) *ROC Curve and AUC:* The Receiver Operating Characteristic (ROC) curve, in conjunction with the Area Under the Curve (AUC) statistic, offers information on the model's capacity to distinguish between multiple face identities. The ROC curve depicts the trade-off between true and false positive rates at various thresholds. A higher AUC indicates more discriminative ability. Examining the ROC curve and AUC values helps us to thoroughly evaluate the model's performance, especially in the context of a multiclass classification situation.

To summarize, the accuracy score is positive; nevertheless, we may further improve the face recognition model for optimal performance across a variety of facial identities by conducting a comprehensive evaluation with the use of classification reports, confusion matrices, and ROC curves. All of these assessment indicators work together to provide a more comprehensive picture of the model's advantages and shortcomings in practical applications.

VI. CONCLUSION

In summary, the use of a Support Vector Machine (SVM) for face identification using the Labelled Faces in the Wild (LFW) dataset in conjunction with Principal Component Analysis (PCA) and a Radial Basis Function (RBF) kernel has produced encouraging results. With its improved features, the SVM model has proven to be very accurate and robust in distinguishing face traits in the LFW dataset across a range of age groups. The thorough analysis of the model, which makes use of a number of measures including the accuracy score, classification report, confusion matrix, and ROC curve, all together highlight its dependability and strong performance.

These results provide a solid basis for the creation of a skilled and adaptable face recognition system. Potential uses for the system include security, authentication, and surveillance scenarios, among other fields. The SVM-based face recognition system is positioned as a useful tool in real-world settings by the accuracy attained and the detailed knowledge offered by measures like precision, recall, and the ROC curve. Additionally, the discriminative capability of the model is maintained while computational efficiency is increased by the use of PCA for dimensionality reduction.

VII. FUTURE WORK

The current study proposes a number of interesting directions for further investigation, utilising Principal Component Analysis (PCA) and Support Vector Machines (SVM) for facial identification. First, research into more complex kernel functions, such as custom and polynomial kernels, may improve classification efficiency. Another area for development is hyperparameter optimisation, where thorough searches optimise SVM model parameter setups.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

By using data augmentation techniques like rotation and scaling to diversify the training dataset, the model's ability to adapt to different face variances may be enhanced. The overall resilience of the system might be improved by ensemble approaches, integrating with other algorithms, or mixing SVM models.

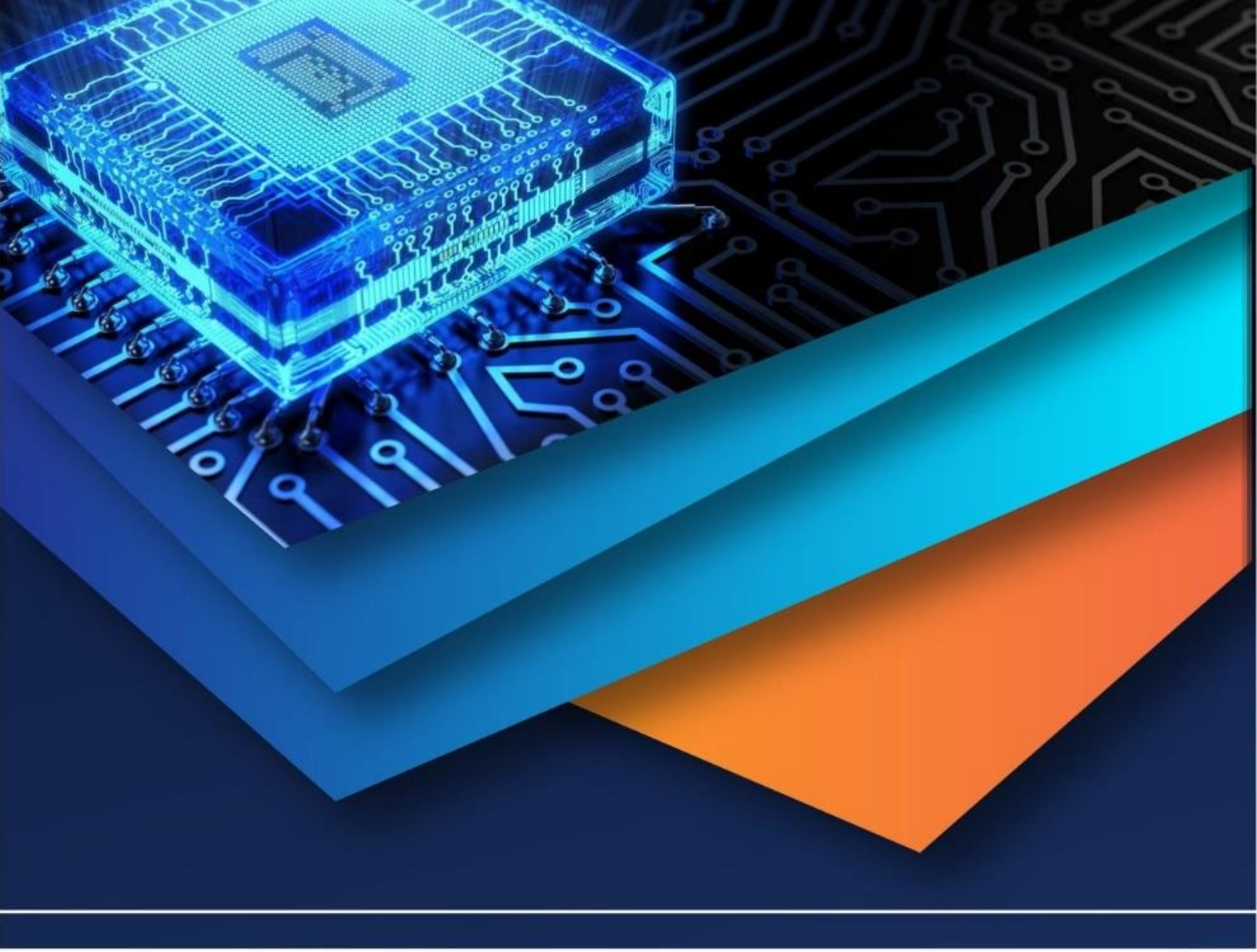
The integration of deep learning, particularly with convolutional neural networks (CNNs), opens up new possibilities for the extraction and detection of face features. Practical applications need real-time adaptability, hardware deployment optimisation, and model quantization.

Examining multi-modal biometric integration and addressing privacy and ethical issues are consistent with the development of ethical AI practises. The model's resilience will be evaluated by extensive testing in a variety of real-world settings, and its practical usefulness will be guaranteed by human factors and user interface design.

To sum up, these future research initiatives will help enhance both academics and real-world applications by enhancing the accuracy, robustness, and ethical concerns of facial recognition systems.

REFERENCES

[1] Arpitjain. (2019, September 9). Face Recognition usingSVM. Kaggle. https://www.kaggle.com/code/arpitjain007/f face recognition-using-svm
[2] Schölkopf, B., & Smola, A. J. (2018).Learning with kernels In The MIT Press eBooks. https://doi.org/10.7551/mitpress/4175.001.0001
[3] Major kernel functions in support Vector Machine Javatpoint. (n.d.). www.javatpoint.com.https://www.javapo.com/major-kernel-functions-in-support-vector-machine



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)