Facial Recognition for Humanitarian Efforts: A Deep Learning based Solution for Missing Person Identification

Dhruv Limbani
Student
Department of Data Science and Business Systems, School of Computing,
SRM Institute of Science and Technology, Kattankulathur 603203, Chennai, India

Abhishek Barhate
Student
Department of Data Science and Business Systems, School of Computing,
SRM Institute of Science and Technology, Kattankulathur 603203, Chennai, India

Hemavathi D
Associate Professor
Department of Data Science and Business Systems, School of Computing,
SRM Institute of Science and Technology, Kattankulathur 603203, Chennai, India

Abstract—Our work offers a novel method for identifying missing persons that makes use of recent developments in face recognition technology. Conventional techniques depend on figuring out how far apart picture embeddings are from one another by calculating distances between the embeddings, which can be computationally demanding, particularly when dealing with big datasets. Our process involves first extracting faces from images using OpenCV's face extractor module, and then using VGGFace to turn those faces into embeddings. However, we employ Principal Component Analysis (PCA) to further reduce their dimensionality to 128-dimensional vectors rather than directly comparing embeddings. On the basis of these reduced embeddings, a Dense Neural Network is then trained. With this method, the traditional O(n) time for classification is greatly reduced to O(1). Our tests show that our methodology outperforms numerous other methods and reaches an excellent accuracy of 96%. This simplified method gives encouraging outcomes in situations involving missing persons identification in addition to increasing efficiency.

Index Terms—Face Extraction, Face Recognition, CNN, Deep Learning, PCA, Missing person Identification

I. INTRODUCTION

Missing persons represent a complex social issue with far-reaching consequences for communities, families, and law enforcement. This phenomenon includes circumstances in which people become unreachable for a variety of reasons, leaving behind uncertainty and sorrow. The human cost of missing person identification and the urgent need for prompt action highlight how urgent it is to address this issue. For a number of reasons, it is imperative that missing people be found quickly and accurately. Reuniting families and communities with their loved ones is the main way it gives them comfort. Timely identification also contributes to public safety by helping law enforcement agencies resolve cases effectively.

Governments tackle the issue of missing person identification in a variety of ways. Conventional approaches entail legal processes including reporting missing persons, sending out notices, and working with law enforcement. These methods depend on labor-intensive processes, community involvement, and legal protocols. Even if they provide a feeling of community and public participation, they could not have the quickness and accuracy required in urgent situations.

Technological advancements, especially in AI and computer vision, have brought about a new era for missing person identification. AI-based methods greatly improve the speed and accuracy of identification operations. They include facial recognition, picture analysis, and data matching algorithms [1,2,7,8,11,12]. The traditional methodologies are being revolutionized by these computational technologies, which create new avenues for dependable and effective solutions.

Conventional AI techniques for identifying missing persons encounter significant challenges, primarily stemming from computational inefficiencies and restrictive timing requirements. Systems that depend on vector distance comparisons between face embeddings face computational complexity because the comparison procedure grows naturally with dataset size, leading to an O(n) time complexity [1,3,7]. In real-time applications, this computing overhead causes delays and inefficiencies in the identification process. Furthermore, some techniques need that the subject's face remain in the frame for a minimum amount of time, usually minutes or several seconds, in order for recognition to take place. Such strict time constraints place real-world constraints on the prompt identification of missing people in dynamic or transient contexts. These underlying constraints highlight the urgent need for more effective and flexible methods of identifying missing persons, underscoring the importance of developments in deep learning and image analysis technologies.

Our suggested methodology, which seamlessly integrates cutting-edge technology to achieve greater versatility and efficiency, transforms the identification process. Our solution achieves better performance and functionality by using a Dense Neural Network for classification, Principal Component Analysis (PCA) for dimensionality reduction, VGGFace for

embedding construction, and OpenCV's face extractor module for extraction of multiple faces from the given image. The advantages of our suggested system are as follows:

- It eliminates the need to store raw photos after embedding production, lowering storage overhead and addressing privacy problems.
- It runs in constant time complexity ($O(1)$), avoiding computational constraints associated with prior techniques.
- Ultimately, our system performs exceptionally well in facial recognition tasks, recognizing several people from the same image at different viewing angles and without time limits, which makes it more applicable in dynamic contexts like video surveillance.

The paper is organized as follows: An overview of the literature on ongoing related research efforts can be found in section II. Section III gives a summary of the dataset. The Face extraction and embedding generation process, along with the associated modules involved, are explained in sections IV and V. In Section VI, a DNN architecture for classification is presented. Section VII provides an explanation of the dataset analysis and training-testing set development processes. The preprocessing and dimensionality reduction procedures are emphasized in section VIII. Section IX discusses the analysis and findings of the experiment. The suggested model is contrasted with models from previous works in section X. Section XI provides a summary and conclusion of the study, while section XII describes the work's potential next steps.

II. RELATED WORKS

Deep learning techniques for image recognition have grown significantly and quickly in recent years. There are many different techniques and approaches used in the field of facial recognition and identification of missing persons. The following works demonstrate the different approaches employed in the current field:

In order to combine two face recognition techniques, Neel Ramakant Borkar et al. [1] proposed a hybrid face recognition algorithm that integrates (PCA) principal component analysis and (LDA) linear discriminant analysis. The algorithm demonstrated accuracy ranging from 91% to 96%. Additionally, they suggested computing the eigenvectors required for the PCA and LDA methods using the Jacobi approach.

Yue Luo, et al. [2] employed an improvised CNN model with the main process of face image preprocessing, face image feature extraction, acquisition of test sample features, face image feature classification, and recognition results. The number of convolution operations, the number of pooling operations, and the three aspects of the face expression picture pixels were all subjected to the cumulative sum operation in the convolution activation function. The operation was then combined with the pooling operation. The weight and bias updates in the convolution, pooling, and classification loss functions were incorporated, and two processes of convolution and pooling operations were taken into consideration in the classification loss function.

Rama Devi P, et al. [3] investigates the application of Siamese neural networks for facial picture similarity, using their own dataset for Indian Faces. The study quantifies how different two photos are from one another. They created a new dataset to test for facial detection using a neural network with a simple architecture capable of one-shot learning.

A PCA (Principle Component Analysis) based facial recognition system was suggested by Maliha Khan et al. [4]. Reducing the substantial quantity of data storage needed for a feature space and making effective use of the available space is their primary goal when employing PCA. They suggested creating broad 1-D pixel vectors from 2-D images that provide the compact primary components of the space function, which is subsequently filtered using PCA.

Vandana S. Bhat, et al. [5] employed a detection system which combined neural networks with Gabor filters. It processes both face and non-facial templates and creates a feature set using a Gabor filter. They employ inverse rapid fourier transforms to translate the image into the frequency domain. These domains are fed into a neural network with scaled conjugate training along with a feature set.

Shivam Singh et al [6] proposed an automated facial recognition system. This system is based on face detection, feature extraction and recognition algorithms. They have used KLT algorithm, Viola-Jones algorithm for facial detection and Haar Cascading Classifier and PCA for feature selection and recognition.

S. Ayyappan et al. [7] have proposed a system which uses CNN algorithm for feature extraction and Haar-Cascading Classifier along with Adaboost for facial recognition. They have also integrated web-scraping into the system for faster matches and increased efficiency of finding correct matches.

KH Teoh et al. [8] discusses the use of deep learning techniques in facial detection over the Haar Cascading Classifiers that are used traditionally. They have implemented a CNN classifier to their self made dataset to develop a real time image detection and saw an increase in accuracy over the haar cascading classifier results.

Klemen Grm et al. [9] introduced a novel deep pair-wise similarity learning (DPSL) strategy tailored for overcoming data scarcity issues in deep hierarchical models, particularly in the context of face recognition. Unlike traditional approaches, DPSL operates directly on image pairs to learn pair-wise similarities, thereby bypassing the need for intermediate feature representations. By leveraging image pairs as input, DPSL effectively increased the amount of training data available for model training, which is crucial for models with large parameter sizes. The efficacy of the proposed strategy was demonstrated through the development of a pose-invariant face recognition model named Pose-Invariant Similarity Index (PISI), with comparative experimental results presented on the FERET and IJB-A datasets.

M. Geetha et al. [10] in their work focused on enhancing the accuracy of existing face recognition systems by leveraging Support Vector Machine (SVM) and Eigenface algorithms. An approach similar to Eigenface was employed for facial feature

extraction via facial vectors, and datasets were trained using the SVM algorithm for face classification and detection. Their methodology aimed to expedite face recognition processes, particularly in applications such as online exam monitoring.

Vijay Deep et al. [11] explored a multimodal approach for missing people identification. Their work employed filters like name, age group and location in addition to a person's image to aid the identification process. The work compared Haarcascade, LDA(Linear Discriminant Analysis) and SVM with their proposed model which achieved 96% accuracy.

Chahrazed Rouabhia et al. [12] developed a new similarity measure called "weighted matrix distance" for human face recognition. Instead of using only vectors, their newly developed measure used matrices and weights in addition to vectors by employing 2DPCA (two-dimensional principal component analysis). The proposed measure was tested on ORL, Yale, PF01 and a subset of FERET image databases.

III. Dataset Description

The Facial Recognition Dataset utilized in this study was Pins Face Recognition dataset which was collected from Pinterest, a popular image-sharing platform, and curated specifically for celebrity identification purposes. Although the main goal is to develop a face recognition system for missing person identification, this dataset was selected for its richness in class diversity and variability in facial angles, which are crucial factors in robust model training and evaluation. A total of 105 different celebrities are included in the collection; each is represented by a different number of photos that highlight their facial traits. 17,534 face photos altogether from a variety of angles, resolutions, and facial emotions are included in the dataset. With at least 80 facial photos linked to each celebrity in the dataset, there is more than enough data for thorough model training and assessment.

IV. Faces Extraction

To effectively detect faces in images, OpenCV's "face_cascade.detectMultiScale" method combines cascade classifiers, sliding window techniques, and Haar-like features [13]. When it comes to identifying patterns like mouths, noses, and eyes, Haar-like features serve as the foundation, and cascade classifiers use a hierarchical framework to gradually improve the detection process while drastically lowering computational load. Integral images speed up feature computation and help assess the rectangle sum regions that are important for face feature detection. The system thoroughly looks for possible face candidates by using a sliding window approach at various image scales, providing robustness against scale fluctuations. The function returns the bounding box coordinates enclosing the faces it has successfully spotted.

V. Pre-Trained VGGFace Descriptor as an Embedder

The model architecture shown in Fig. 1 is a deep convolutional neural network (CNN) architecture; more precisely,

it is a facial recognition-specific VGGNet variation known as VGGFace [14]. Following are the components and their roles in the network:

1) Input Layer: This layer represents the input to the network, which consists of images with dimensions 224x224 pixels and three color channels (RGB).

2) Convolutional Layers: These layers use learnable filters to extract features from the input image by performing convolutions on it. The filters look for different aspects and patterns in the input image, like edges, textures, and forms. A rectified linear unit (ReLU) activation function, which adds non-linearity to the network, comes after each convolutional layer.

3) Zero Padding Layers: Additional border pixels with a value of zero are added around the input image using zero padding layers. During convolution procedures, padding aids in the preservation of spatial dimensions and information at the image's borders.

4) Max Pooling Layers: By choosing the maximum value inside each pooling region, max pooling layers downsample the feature maps that are produced from convolutional layers. This procedure helps with translation invariance and lowers computing cost by shrinking the spatial dimensions of the feature maps while maintaining significant features.

5) Dropout Layers: Dropout layers are a type of regularization that randomly removes a portion of the neurons during training in order to minimize overfitting. This lessens the network's dependency on particular neurons and forces it to learn more reliable and generalizable properties.

6) Flatten Layer: The multi-dimensional feature maps are reshaped into a one-dimensional vector by the flatten layer, which gets the data ready for input into fully connected layers. In our work this output is further preprocessed and used as an input for fully connected DNN.

This VGG Face Embedder converts (224, 224) sized images into an embedding vector of length 2622.

VI. DNN for Classification

The proposed DNN architecture as shown in Fig. 2 follows a sequential model design, where layers are stacked sequentially on top of each other. The 128-dimensional vectors produced by PCA dimensionality reduction are sent to the input layer. Higher-level features are gradually extracted from the input data by the succeeding dense layers. By regularizing the learning process, dropout layers are strewn across the network to prevent overfitting. The model is able to predict the class labels of the input data because the last dense layer generates the classification probabilities for each of the 105 classes.

VII. Training, Testing and Validation Sets and Analysis

A train-test split ratio of 80:20 is used in the training and testing phase, with an extra 20% of the training data set aside for validation. This distribution guarantees that the model receives enough training data, and it also sets aside a
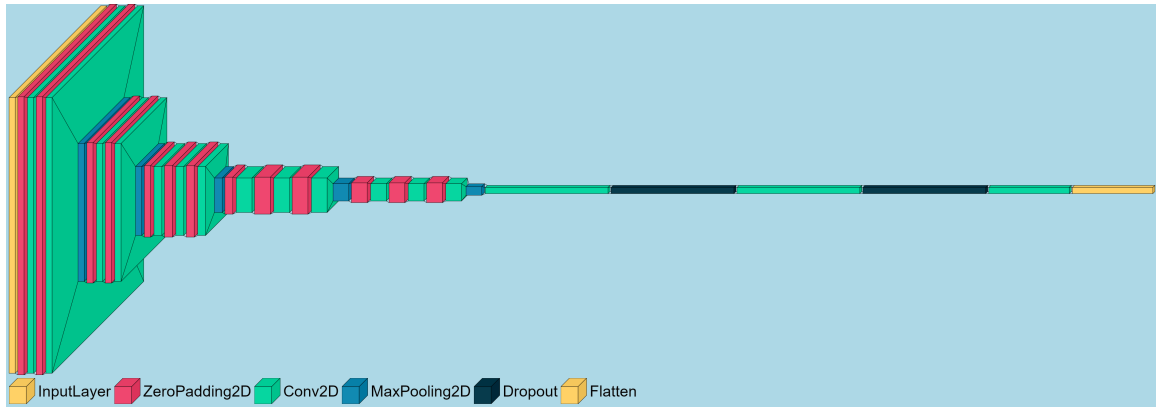
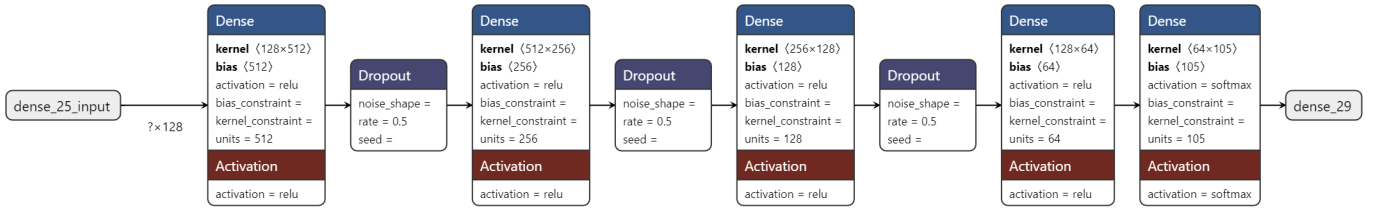Fig. 1. VGGFace Embedder Architecture

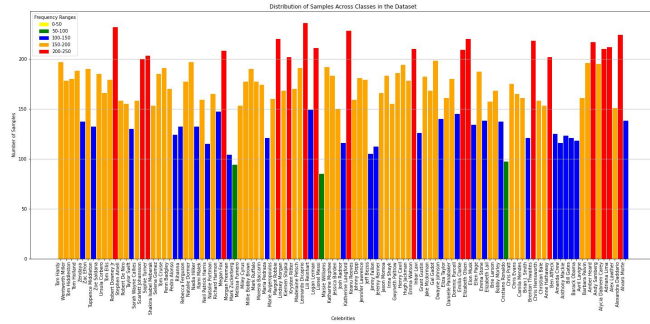Fig. 2. DNN Architecture for Probabilistic Classification
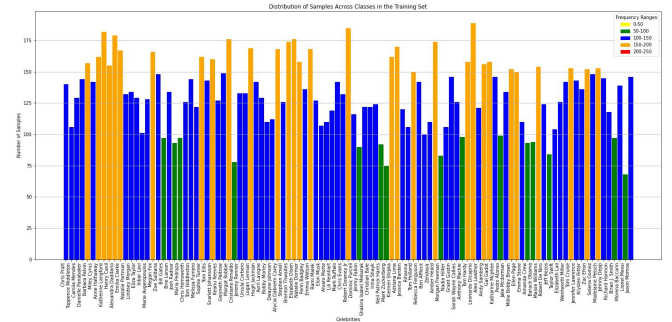

Fig. 3. Class Distribution for entire Dataset

Fig. 4. Class Distribution for Train Dataset

specific amount for validation in order to track results and avoid overfitting.

Fig. 3 shows the distribution of samples across all the classes in the entire dataset. It is observed that the distribution is not uniform across the classes and ranges from 50 to 250 for all the classes. Hence, we have employed stratification which reduces the possibility of introducing bias and guarantees that all classes are sufficiently represented in both the training and validation sets by maintaining the distribution of classes within each subset proportionate to the original dataset. When class imbalances occur, stratification plays a critical role in maintaining the distributional integrity of the dataset and producing more consistent results for model training and assessment. Fig. 4 and Fig. 5 shows the distribution of samples across all the classes in training and testing set respectively after applying the stratified sampling.
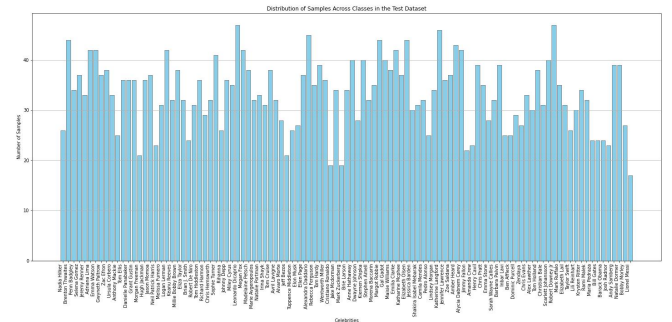

Fig. 5. Class Distribution for Test Dataset

As it is observed from the charts in Fig. 4 and Fig. 5, the number of samples in each class in training dataset ranges roughly from 65 to 190. While for the testing dataset, there are around 15 to 50 samples in each class since the train-test ratio was set to 80:20 as previously mentioned.

VIII. PREPROCESSING AND DIMENSIONALITY REDUCTION

A. Preprocessing with Standard Scaler

In the preprocessing stage, the dataset i.e. the embedding vectors of all the images undergoes standardization using the Standard Scaler technique. To ensure that every feature is on the same scale, this technique entails converting the features to have a mean of zero and a standard deviation of one. For machine learning algorithms, especially those that are sensitive to feature scaling, standardization is crucial because it keeps features with bigger scales from taking over the learning process. The convergence of optimization algorithms can be accelerated and the performance and stability of the next modeling phases can be improved by standardizing the characteristics. This preprocessing step reduces the impact of feature scale discrepancies, which improves the interpretability and efficacy of the model.

B. Dimensionality Reduction with Principal Component Analysis (PCA)

After normalization, Principal Component Analysis (PCA) is used to reduce the length of embedding vectors without losing useful information [15]. PCA is an effective method that maintains the greatest variance in the original dataset while converting high-dimensional data into a lower-dimensional representation. Particularly in situations when the original feature space is big or noisy, PCA enables more effective model training and inference by identifying the most important patterns and minimizing redundancy in the data. In addition to increasing computational efficiency, lower dimensionality helps lessen the negative effects of dimensionality, which improves the resilience and generalizability of the model. Leveraging PCA as a preprocessing step allows us to extract essential features and streamline the modeling process, ultimately contributing to the overall efficacy and performance of the face recognition system for missing person identification.

IX. EXPERIMENTAL RESULTS AND ANALYSIS

Fig. 6 and Fig. 7 depict the accuracy and loss curves for training and validation sets respectively. It is clear from the figures that the model exhibits superior performance, including lower loss and higher accuracy, on the validation set compared to the training set. This event indicates strong learning capacity and successful adaptation to new cases, suggesting that the model generalizes effectively to unseen data.

The model yielded an overall accuracy of 96.62% on the test set, with corresponding macro average precision, recall, and F1-score values of 96.72%, 96.60%, and 96.59%.

The bar chart of misclassifications by class is displayed in Fig. 8. It has been found that more than half of the classes
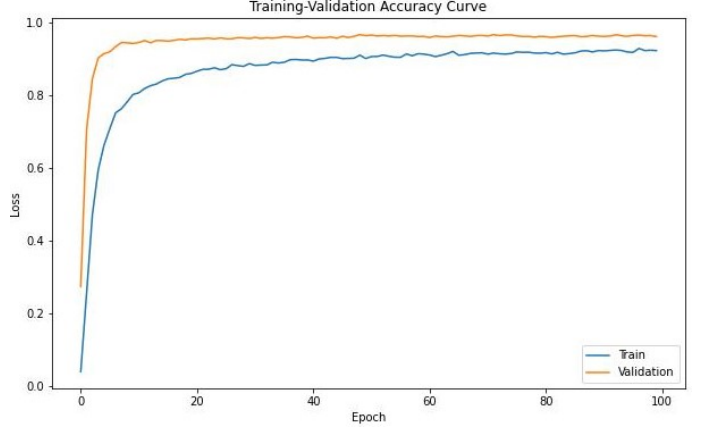

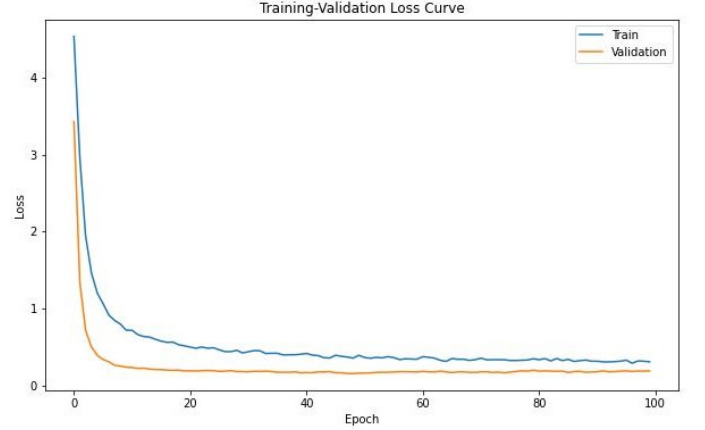Fig. 6. Training Vs Validation Accuracy
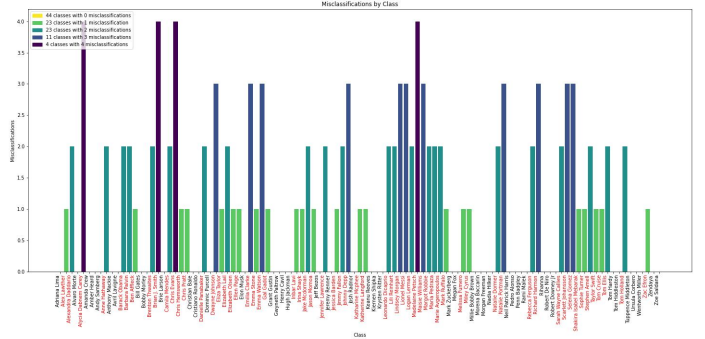

Fig. 7. Training Vs Validation Loss

Fig. 8. Number of Misclassifications for each class

have one or fewer incorrect classifications. The maximum number of misclassifications is found in four classes, each of which has four misclassifications. There are 23 classes with two incorrect classifications each, and 11 classes with three incorrect classifications each.

X. MODEL COMPARISON

The accuracy-wise performance comparison between our proposed model with other models from related works is shown in Table I.

TABLE I
PERFORMANCE COMPARISON OF PROPOSED MODEL

Sr. no	Paper/ Author	Models	Accuracy (%)
1	Neel Ramakant Borkar et al [1]	PCA	91
		LDA	94
2	Yue Luo et al.[2]	CNN	88
3	S.Ayyappan et al. [7]	Haar Cascade Classifier	90
4	KH Teoh et al.[8]	CNN	91.7
5	Vijay Deep et al. [11]	Pretrained CNN-dlib library	96
6	Chahrazed Rouabhia et al. [12]	Weighted AMD	95.5
5	**Our paper**	**VGG16 + PCA + DNN**	**96.62**

XI. CONCLUSION

In conclusion, our research endeavors culminated in the development of a sophisticated deep learning model for Missing Person Identification from images, boasting an exceptional accuracy, precision, recall, and F1 score of 96.62%. Our method outperforms some of the state-of-the-art models in related research in terms of efficiency as well as greater predictive ability [1,2,7,8]. In contrast to traditional techniques that depend on determining the distances between embeddings, our approach simplifies identification while preserving accuracy. Moreover, the fact that our model is not constrained by strict timing specifications guarantees flexibility and adaptability in real-world situations. The compact size of our deep learning model, a mere 2MB, renders it highly deployable on resource-constrained devices such as smartphones and edge computing platforms. When taken as a whole, these successes highlight how our strategy may improve search and rescue operations, support law enforcement, and eventually benefit society as a whole

XII. FUTURE WORKS

Our research points to a number of intriguing directions for future work that should be investigated and developed. First off, adding more modalities to the images—like text or audio information about the individual in question—could enhance the model's comprehension and boost its prediction power. Second, the model's capacity to recognize complex facial traits and contextual subtleties may be further improved by utilizing cutting-edge deep learning approaches like transformer topologies and attention mechanisms. More reliable and equitable models may also be produced by looking at methods to correct for possible biases in the dataset, such as differences in image quality or imbalances in demographics.

Furthermore, expanding the model's use in dynamic contexts could involve applying it to real-time video streams or surveillance data. Finally, responsible and ethical use in practice requires working with domain experts and stakeholders to incorporate privacy protections, legal frameworks, and ethical considerations into the deployment of such models.

In future research endeavors, exploring the integration of multi-modal data, leveraging cutting-edge deep learning methodologies, addressing dataset biases, extending model applicability to real-time scenarios, and prioritizing ethical considerations might collectively advance the field of Missing Person Identification and contribute to its societal impact and ethical deployment. This comprehensive approach will not only enhance the model's performance but also ensure its responsible and ethical deployment in real-world scenarios.

REFERENCES

[1] Borkar, Neel & Kuwelkar, Sonia. (2017). Real-time implementation of face recognition system. 249-255. 10.1109/ICCMC.2017.8282685.

[2] Luo, Yue, et al. "Design of Facial Expression Recognition Algorithm Based on CNN Model." 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA). IEEE, 2023.

[3] Penikalapati, Rama & Yashashvini, R & Navyadhara, G & Ruchitha, M.. (2021). Similar Face Detection for Indian Faces using Siamese Neural Networks. 1-5. 10.1109/INCET51464.2021.9456199.

[4] Khan, Maliha, et al. "Face detection and recognition using OpenCV." 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS). IEEE, 2019.

[5] Bhat, Vandana S., and Jagadeesh D. Pujari. "An approach for face detection using artificial intelligence." 2014 International Conference on Contemporary Computing and Informatics (IC3I). IEEE, 2014.

[6] Singh, Shivam, and S. Graceline Jasmine. "Face recognition system." International Journal of Engineering Research & Technology (IJERT) 8.05 (2019): 2278-0181.

[7] Ayyappan, S., and S. Matilda. "Criminals and missing children identification using face recognition and web scrapping." 2020 International conference on system, computation, automation and networking (IC-SCAN). IEEE, 2020.

[8] Teoh, K. H., et al. "Face recognition and identification using deep learning approach." Journal of Physics: Conference Series. Vol. 1755. No. 1. IOP Publishing, 2021.

[9] Grm, Klemen, Simon Dobrisek, and Vitomir Struc. "Deep pair-wise similarity learning for face recognition." 2016 4th International Conference on Biometrics and Forensics (IWBF). IEEE, 2016.

[10] Geetha, M., et al. "Design of face detection and recognition system to monitor students during online examinations using Machine Learning algorithms." 2021 international conference on computer communication and informatics (ICCCI). IEEE, 2021.

[11] DEEP, V., KAUSHIK, G., & Dronacharya College of Engineering, Gurgaon, Haryana. (2022). Finding missing person using face recognition system - KHOJ. In International Journal of Innovative Research in Technology (Vol. 8, Issue 12, pp. 1539–1541) [Journal-article]. https://ijirt.org/master/publishedpaper/IJIRT155143_PAPER.pdf

[12] Rouabhia, Chahrazed, and Hicham Tebbikh. "Efficient face recognition based on weighted matrix distance metrics and 2DPCA algorithm." Archives of Control Sciences 21.2 (2011): 207-221.

[13] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001. Vol. 1. Ieee, 2001.

[14] Parkhi, Omkar, Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." BMVC 2015-Proceedings of the British Machine Vision Conference 2015. British Machine Vision Association, 2015.

[15] Wold, Svante, Kim Esbensen, and Paul Geladi. "Principal component analysis." Chemometrics and intelligent laboratory systems 2.1-3 (1987): 37-52.