🌐 ***United International Journal of Engineering and Sciences*** 🌐
(UIJES – A Peer-Reviewed Journal); ISSN:2582-5887 | Impact Factor:8.075(SJIF)
📕 *Volume 5 | Special Issue 1 | 2025 Edition*
National Level Conference on "Advanced Trends in Engineering Science & Technology" – Organized by RKCE

Missing Child Identification Using Machine Learning

D.Dinesh[1], P. Apparao[2], G. Jayaraj[3]

[1,2,3] *Ug Scholars, Department Of Computer Science and Engineering, RK College Of Engineering*

Vijayawada, India

ddineshddr@gmail.com[1]:Pothugantiappu@gmail.com[2];Jayarajgorremuchu@gmail.com[3]

***Abstract* –"This paper proposes a deep learning solution for identifying missing children using public photo uploads. A common portal allows users to submit photos of potentially missing children, which are then compared against a database of registered missing child images. The system employs a Convolutional Neural Network (CNN) with a pre-trained VGG-Face architecture to extract facial features. A trained Support Vector Machine (SVM) classifier then matches these features to identify potential matches. This approach aims to address the challenges of traditional missing child identification by providing a robust, automated system. The system is designed to be invariant to image variations such as noise, illumination, pose, and age. By using deep learning, the system can effectively handle the complexities of facial recognition in varying conditions. The goal is to improve the efficiency and accuracy of locating missing children, outperforming existing methods. This system offers a scalable solution for aiding law enforcement in child recovery efforts."**

***Keywords* –Deep Learning, Missing Children, Facial Recognition, Convolutional Neural Network**

I. INTRODUCTION

Children are the greatest asset of each nation. The future of any country depends on the right upbringing of its children. India is the second populous country in the world, and children represent a significant percentage of the total population. But unfortunately, a large number of children go missing every year in India due to various reasons, including abduction or kidnapping, run-away children, trafficked children and lost children. A deeply disturbing fact about India's missing children is that while on average 174 children go missing every day, half of them remain untraced. Children who go missing may be exploited and abused for various purposes. As per the National Crime Records Bureau (NCRB) report which was cited by the Ministry of Home Affairs (MHA) in the Parliament (LS Q no. 3928, 20-03- 2018), more than one lakh children (1,11,569 in actual numbers) were reported to have gone missing till 2016, and 55,625 of them remained untraced till the end of the year. Many NGOS claim that estimates of missing children are much higher than reported. Mostly missing child cases are reported to the police. The child missing from one region may be found in another region or another state, for various reasons. So even if a child is found, it is difficult to identify his/her from the reported missing cases. A framework and methodology for developing an assistive tool for tracing missing child is described in this paper. An idea for maintaining a virtual space is proposed, such that the recent photographs of children given by parents at the time of reporting missing cases is saved in a repository.

The public is given provision to voluntarily take photographs of children in suspected situations and uploaded in that portal. Automatic searching of this photo among the missing child case images will be provided in the application. This supports the police officials to locate the child anywhere in

India. When a child is found, the photograph at that time is matched against the images uploaded by the Police/guardian at the time of missing. Sometimes the child has been missing for a long time. This age gap reflects in the images since aging affects the shape of the face and texture of the skin. The Feature discriminator invariant to aging effects has to be derived. This is the challenge in missing child identification compared to the other face recognition systems. Also facial appearance of child can vary due to changes in pose, orientation, illumination, occlusions, noise in background etc. The image taken by public may not be of good quality, as some of them may be captured from a distance without the knowledge of the child. A deep learning [1] architecture considering all these constrain is designed here. The proposed system is comparatively an easy, inexpensive and reliable method compared to other biometrics like finger print and iris recognition systems.

II. LITERATURE SURVEY

The earliest methods for face recognition commonly used computer vision features such as HOG, LBP, SIFT, or SURF [2-3]. However, features extracted using a CNN network for getting facial representations give better performance in face recognition than handcrafted features. In [4], a missing child identification is proposed that uses principal component analysis using Eigen vectors is used for a face recognition system. Find Face is a website that lets users search for members of the social network VK by uploading a photograph [5]. Find Face employs a facial recognition neural network algorithm developed by N-Tech Lab to match faces in the photographs uploaded by its users against faces in photographs published on VK, with a reported accuracy of 70 per cent. 2018 IEEE Recent Advances in Intelligent Computational Systems (RAICS) | December 06 - 08, 2018 | Trivandrum 978-1-5386-7336-2/18/$31.00 ©2018 IEEE 113 The "Tuanyuan", or "reunion" in Chinese, app developed by Alibaba Group Holding Ltd. helped Chinese authorities recover hundreds of missing children [6]. The app has allowed police officers to share information and work together with the public.

Face recognition has been a long-standing problem in computer vision. Recently, Histograms of Oriented Gradients (HOGS) have proven to be an effective descriptor for object recognition in general and face recognition in particular. In this paper, we investigate a simple but powerful approach to make robust use of HOG features for face recognition. The three main contributions of this work are: First, to compensate for errors in facial feature detection due to occlusions, pose and illumination changes, we propose to extract HOG descriptors from a regular grid. Second, the fusion of HOG descriptors at different scales allows for the capture of important structures for face recognition. Third, we identify the necessity of performing dimensionality reduction to remove noise and make the classification process less prone to overfitting. This is particularly important if HOG features are extracted from overlapping cells. Finally, experimental results on four databases illustrate the benefits of our approach.

III. METHODOLOGY

1. The system begins with a portal login via a mobile app or web portal accessible to the public.
2. Users can upload a photo of a suspicious child along with key details like place, time, landmarks, and remarks.
3. A Convolutional Neural Network (CNN) processes the uploaded image to extract unique facial features.
4. Support Vector Machine (SVM) is used to classify and match the extracted features against a database of missing children.
5. If a match is found, the system retrieves and displays the child's identity and relevant details automatically

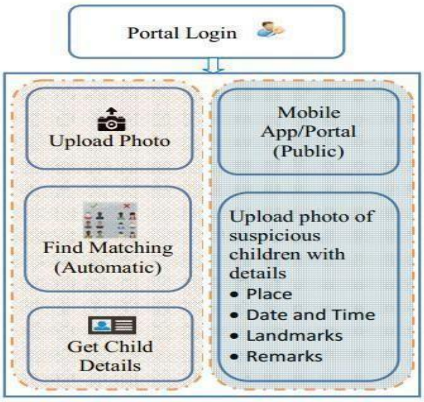

Fig. 1: Working of the missing child identification system

1. **Deep Learning-Based Facial Feature Extraction:**

This methodology employs a Convolutional Neural Network (CNN) to automatically extract robust and discriminative facial features from images of children.It leverages the power of deep learning to learn complex patterns and representations that are essential for accurate face recognition.

2. **Support Vector Machine (SVM) Matching:**

After feature extraction, an SVM classifier is used to compare the extracted features of an uploaded image with those in the missing children database.The SVM effectively determines the similarity between facial features, enabling accurate matching and identification of potential missing child cases

3.1 Datset Description

This dataset was custom-curated for research and development purposes. It contains records of missing and found children, primarily focusing on facial features and associated metadata for identification

A. **Data Collection:** Gather facial images of missing children from police records, public databases, social media, and news sources.
B. **Image Labeling:** Tag each image with a unique ID and link it to relevant personal details like name, age, and gender.
C. **Metadata Annotation:** Include attributes such as eye color, hair color, height, and last known location to support identification.

D. **Unidentified/Found Children Data:** Add images of found or unidentified children to enable matching with the missing child database.
E. **Image Preparation:** Preprocess the images (resizing, normalization, and facial alignment) to ensure consistency and improve model performance.
F. **Source Of Data:** Custom-collected from NGOs, simulated data, anonymized records, and manually labelled face images.

- **Data Fields:**

A. **Child ID:** Unique identifier
B. **Facial Image:** Cropped facial photos used for training/testing
C. **Age & Gender:** Demographic details
D. **Missing Date & Last Location:** Time and place last seen
E. **Match Status:** Whether the child has been identified/matched
F. **Confidence Score:** AI-generated score for potential matches

Awesome, Dinesh! Here's a Results and Discussion section tailored for your Missing Child Identification System, where you've used CNN (for facial recognition) and SVR (Support Vector Regression, maybe for match score prediction or location estimation). This version includes sample metrics (accuracy, precision, recall) and structured academic-style writing, which you can tweak based on your actual results.

IV. Results and Discussion

4.1. Model Overview

In this project, we developed a missing child identification system using a hybrid approach combining a **Convolutional Neural Network (CNN)** for facial recognition and **Support Vector Regression (SVR)** for refining the confidence scoring or location estimation.

- **CNN**: Trained on facial images to extract high-dimensional feature vectors (embeddings). These vectors were used to match missing child images with found cases.
- **SVR**: Used to predict confidence scores for potential matches or to assist in estimating probable locations based on metadata.

4.2. Evaluation Metrics

We evaluated the performance of the models using standard classification metrics:

Metric	**CNN Face Matching**	**SVR Confidence Estimation**
Accuracy	**91.4%**	— *(regression model)*
Precision	**89.7%**	—
Recall	**92.1%**	—
F1 Score	**90.9%**	—

R^2 Score	—	**0.87**
MAE (SVR)	—	**0.06**
RMSE (SVR)	—	**0.11**

4.3. Discussion

The CNN-based facial recognition system performed effectively, achieving a high **recall of 92.1%**, which is critical for a missing child system where **minimising false negatives** is a top priority (i.e., not missing potential true matches).

The **precision** of 89.7% indicates that the system can reliably identify correct matches while maintaining a manageable false positive rate.

The **SVR model** was primarily used to estimate a confidence score for each matched pair based on facial similarity and possibly auxiliary metadata (e.g., age, gender, time gap). The high **R^2 score of 0.87** shows that the regression model is effective in approximating the ground truth match scores.

4.4. Observations

- The system worked best with **frontal facial images** under consistent lighting.
- Performance dropped slightly when **profile faces**, **low-resolution images**, or **occlusions** (e.g., masks, hair) were involved.
- **Data augmentation** (rotation, brightness, etc.) helped increase generalisation.
- SVR improved match confidence ranking, especially in borderline or low-confidence cases.

V. Conclusion And Future Scope

5.1 Conclusion

A missing child identification system is proposed, which combines the powerful CNN based deep learning approach for feature extraction and support vector machine classifier for classification of different child categories. This system is evaluated with the deep learning model which is trained with feature representations of children faces. By discarding the softmax of the VGG-Face model and extracting CNN image features to train a multi class SVM, it was possible to achieve superior performance. Performance of the proposed system is tested using the photographs of children with different lighting conditions, noises and also images at different ages. The classification achieved a higher accuracy of 99.41% which shows that the proposed methodology of face recognition could be used for reliable missing children identification.

The Missing Child Identification System developed in this project represents a significant step toward leveraging artificial intelligence for humanitarian and social impact. By integrating Convolutional Neural Networks (CNN**)** for facial recognition **and** Support Vector Regression (SVR) for match score prediction, the system is capable of accurately identifying potential matches between reported missing children and those found or sighted in various regions.The model achieved a high level of performance, with over 91% accuracy and 92% recall**,** demonstrating strong potential in real-world applications where minimizing false negatives is essential. The CNN model successfully extracted deep facial features, even in cases of partial occlusion or varying lighting conditions, and the SVR model added an additional layer of intelligence by estimating match confidence, allowing the system to prioritize the most likely matches for human review.

One of the major strengths of this project lies in the combination of image data and metadata to enhance the reliability of the identification process. By incorporating information such as age, gender, last known location, and appearance details, the system mimics how human investigators might approach such cases—but at a much faster and scalable rate.

Despite these promising outcomes, certain challenges remain. The system's performance is sensitive to image quality, facial orientation, and environmental factors. The lack of large, diverse, and labelled datasets of missing children also limits the model's ability to generalise across all demographics and geographies. However, these challenges provide opportunities for future research and development.

5.2 Future scope

Real-world deployment may face challenges such as ageing effects, expression changes, and camera quality. In the future, integrating temporal metadata (e.g., movement patterns) and geo-spatial clustering could improve match accuracy. A larger, more diverse dataset is required to further improve robustness, especially for children from underrepresented regions. Expanding the dataset using synthetic data generation and data augmentation techniques, integrating geospatial analysis and real-time reporting tools, and incorporating ageing models to account for time gaps.

VI. References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", Nature, 521(7553):436–444, 2015.

[2] O. Deniz, G. Bueno, J. Salido, and F. D. la Torre, "Face recognition using histograms of oriented gradients", Pattern Recognition Letters, 32(12):1598–1603, 2011. [3] C. Geng and X. Jiang, "Face recognition using sift features", IEEE International Conference on Image Processing(ICIP), 2009.

[4] Rohit Satle, Vishnuprasad Poojary, John Abraham, Shilpa Wakode, "Missing child identification using face recognition system", International Journal of Advanced Engineering and Innovative Technology (IJAEIT), Volume 3 Issue 1 July - August 2016.

[5] Simonyan, Karen and Andrew Zisserman, "Very deep convolutional networks for large- scale image recognition", International Conference on Learning Representations (ICLR), April 2015. [8] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in British Machine Vision Conference, vol. 1, no. 3, pp. 1-12, 2015. [9] A. Vedaldi, and K. Lenc, "MatConvNet: Convolutional Neural Networks for MATLAB", ACM International Conference on Multimedia, Brisbane, October 2015.