IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11,Iss 12, Dec 2022

Missing Child Identification System using HOG and KNN Machine Learning Classifier

Guide: Dr. Prabagar S
Computer Science and Engineering
Presidency University
Bangalore, India
parbagar.s@presidencyuniversity.in

M Shahid Afrid
Computer Engineering
Presidency University
Bangalore, India
201910101510@presidencyuniversity.in

P Taraka Sai Tarun
Computer Engineering
Presidency University
Bangalore, India
201910101541@presidencyuniversity.in

B Praneeth Kumar Reddy
Computer Engineering
Presidency University
Bangalore, India
201910101513@presidencyuniversity.in

M Govardhan Reddy
Computer Engineering
Presidency University
Bangalore, India
201910102066@presidencyuniversity.in

Abstract— **In India, a countless number of children are reported missing every year. Among the missing child cases, a large percentage of children remain untraced. This paper presents a novel use of deep learning methodology for identifying the reported missing child from the photos of a multitude of children available, with the help of face recognition. The public can upload photographs of a suspicious child into a common portal with landmarks and remarks. The photo will be automatically compared with the registered photos of the missing child from the repository. Classification of the input child image is performed and the photo with the best match will be selected from the database of missing children. For this, the machine learning model is trained to correctly identify the missing child from the missing child image database provided, using the facial image uploaded by the public. In this system, for face detection, we are using a HOG (Histogram of oriented Gradient) based face detector which gives more accurate results rather than other methodologies like Haar Cascade and CNN.**

Keywords—Children, missing, photographs, Deep Learning, Face Recognition, identify, Machine Learning.

I. INTRODUCTION

Children are the greatest asset of each nation. The future of any country depends upon the right upbringing of its children. India is the second-most 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, 20-032018), 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. The 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 him/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 are saved in a repository.

II. Proposed System

Face recognition has been a rapidly growing and intriguing region progressively applications. A huge number of face recognition calculations have been produced a long time ago. Here we propose a methodology for missing child identification that combines facial feature extraction based on the HOG algorithm and matching based on KNN. The proposed system utilizes face recognition for missing child identification. This is to help authorities and parents in the missing child investigation.

III. METHODS USED

A. *Data Collection*

The first step in developing an automatic missing child identification system is to collect data. In this case, the data will consist of images of missing children along with their personal details such as name, age, gender, etc.

B. *Pre-processing*

Once the data is collected, it needs to be pre-processed. This includes cleaning the data, removing noise, and standardizing the format of the data.

C. *Feature Extraction*

The next step is to extract the features from the images. The proposed system uses the HOG (Histogram of Oriented Gradients) algorithm for missing child face detection. This algorithm detects facial features such as the shape of the eyes, nose, and mouth, and extracts them as features.

D. *Face Recognition*

After extracting the features, the ML algorithm KNN classifier is used for face recognition. The KNN classifier compares the extracted facial features of the missing child with the features of the children in the database. If a match is found, the system provides the missing child information, and an alert message is sent to the concerned authorities.

E. *Storage*

If the missing child's facial features do not match with the features in the database, then the system stores the child's information in the storage server as a new entry.

F. *Web Application*

Finally, the proposed system is designed as a web application so that the public or the authorities can search for missing children by submitting their profile picture.

IV. IMPLEMENTATION

A. System Architecture

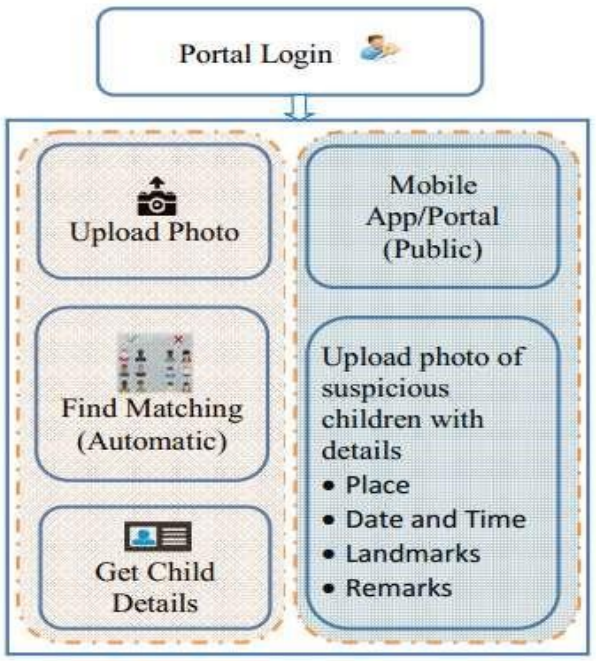

Figure 1.1 System Architecture

The above figure 1.1 shows us the systems architecture which consists of a national portal for storing details of the missing child along with the photo. Whenever a child missing is reported, along with the FIR, the concerned officer uploads the photo of the missing child into the portal. The public can search for any matching child in the database for the images with them. The system will prompt the most matching cases. Once the matching is found, the officer can get the details of the child.

B. Pre-processing

Pre-processing input raw images in the context of face recognition involves acquiring the face region and standardizing images in a format compatible with the HOG methodology employed. The photographs of missing children acquired by a digital camera or mobile phone are taken and categorized into separate cases for creating the database of the face recognition system. The face region in each image is identified and cropped for getting the input face images.

C. Upload Photo

It consists of a national portal for storing details of the missing child along with the photo. Whenever a child missing is reported, along with the FIR, the concerned officer uploads the photo of the missing child into the portal. The public can upload photos of any suspicious child at any time into the portal with details like place, time, landmarks, and remarks. The photo uploaded by the

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 12, Dec 2022

users will be automatically compared with photos of the registered missing children and if a matching photo with a sufficient score is found, then an alert email will be sent to the concerned officer. The message will also be visible in the message box of the concerned officer's login screen.

D. Search

Whenever users upload a photo of a suspected child, the system generates a template vector of the facial features from the uploaded photo. If matching is found in the repository, the system displays the most matched photo and pushes a message to the concerned Officer portal or email the alert message of the matching child. Similarly, the Officer can check for any matching with the database at any time using the proposed system.

V. ALGORITHM

A. Histogram Oriented Gradients

HOG (Histogram Oriented Gradients): The most popular way for face and object detection, in general, is using HOG classifiers. HOG stands for Histogram of Oriented Gradients. The crux of the matter is in finding appropriate feature descriptors for an image, be it faces or other objects. The Histogram of Oriented Gradients (HOG) is a function descriptor used primarily for object recognition in image processing. A function descriptor is a representation of an image or an image patch that by extracting valuable information from it, simplifies the image.

The theory behind the descriptor histogram of directed gradients is that the distribution of intensity gradients or edge directions will define the appearance and shape of local objects within an image. The x and y derivatives of an image (Gradients) are helpful because due to sudden change in amplitude, the magnitude of gradients is high around edges and corners and we know that edges and corners pack a lot more object shape details than flat regions. Therefore, the gradient path histograms are used as properties of this descriptor.

Steps to calculate HOG descriptors for face detection:

1. The picture is divided into 8 to 8 cell blocks, and for each 8 to 8 cell block, a histogram of gradients is measured.
2. A vector of 9 buckets (numbers) corresponding to angles from 0 to 180 degrees is basically a histogram (20-degree increments).
3. The values of these 64 cells (8X8) are binned into these 9 buckets and cumulatively inserted.
4. This restricts 64 values to 9 values, in principle.

B. K-Nearest Neighbour (KNN)

K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on the Supervised Learning technique. K-NN algorithm assumes the similarity between the new case/data and available cases and puts the new case into the category that is most similar to the available categories. K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm. So, the K-NN algorithm can be used for the Classification problems. Here KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

The K-NN working can be explained on the basis of the below algorithm:

1. For implementing any algorithm, we need dataset. So, during the first step of KNN, we must load the training as well as test data.
2. Next, we need to choose the value of K i.e., the nearest data points. K can be any integer.
3. For each point in the test data do the following – • 3.1 − Calculate the distance between test data and each row of training data with the help of any of the methods namely: Euclidean, Manhattan, or Hamming distance. The most commonly used method to calculate distance is Euclidean. • 3.2 − Now, based on the distance value, sort them in ascending order.
   • 3.3 − Next, it will choose the top K rows from the sorted array.
   • 3.4 − Now, it will assign a class to the test point based on the most frequent class of these rows.
4. End.

Face Recognition with KNN

In this proposed system, we are using the KNN algorithm for face recognition. However, the HOG feature will be detected faces from training images and crop that face detected part and convert into pixel format and store with face Lebel (like the recognized name). So, the training dataset will be trained with KNN, and when we pass the testing image as input image then KNN will compare with all datapoints like pixels then which data points are nearly matched then that the system will return that face Label as output so that we can achieve the face recognition approach.

EXPERIMENTAL RESULTS

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 12, Dec 2022

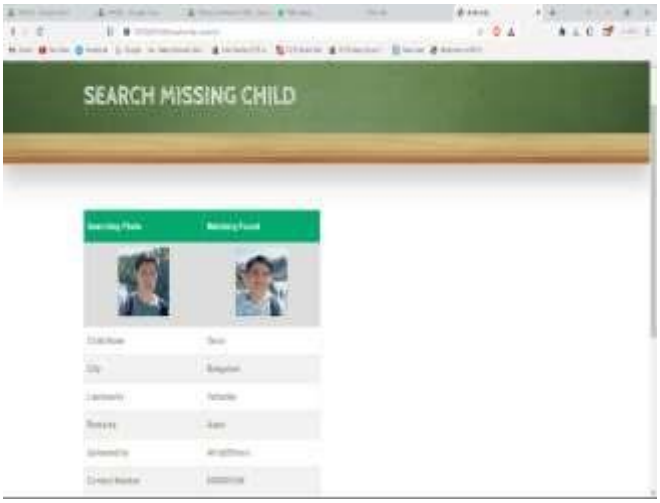

Figure 2.1 Sample Result(1)

Figure 2.2 Sample Result(2)

Both the above figures{2.1, 2.2} displays the sample outputs delivered when we uploaded different images of the same child.

VI. CONCLUSION

In conclusion, the proposed system of automatic missing child identification using ML approaches is a novel solution to the problem of missing children. The traditional process of searching for missing children is a time-consuming and effort-intensive process that often yields little success. The proposed system, on the other hand, is designed to use machine learning algorithms to automatically detect missing children from their profile pictures and match them with the database of missing children. This approach not only saves time and effort but also increases the chances of finding missing children. The web application makes it accessible to the public and the authorities, making it a useful tool for everyone. Therefore, the proposed system has the potential to be an effective solution for the issue of missing children, and it could be implemented in various countries for the benefit of society.

REFRENCES

[1]. Mr. A. David Rajkumar, Mr. R. Karthick Raja, Mr. S. Sankar Ganesh, Dr. V. R. S. Mani, "IDENTIFICATION OF MISSING PERSON IN THE CROWD USING PRETRAINED NEURAL NETWORK", IRJET, Volume: 07, Issue: 02, Feb 2020..

[2]. Ahtasham Ansari, Aditya Singh, Abhishek Sagar, Komal," Android-based Application – Missing Person", IRJET, Volume: 07, Issue: 03, Mar 2020.

[3]. Peace Muyambo, An Investigation on the Use of LBPH Algorithm for Face Recognition to Find Missing People in Zimbabwe, "INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY" (IJERT) Volume 07, Issue 07 (July 2019).

[4]. S. AYYAPPAN, Dr.S.MATILDA, "Criminals And Missing Children Identification Using Face Recognition And Web Scrapping", International Conference on System, Computation, Automation and Networking (ICSCAN),2020.

[5]. S. Chandran, Pournami & Balakrishnan, Byju & Rajasekharan, Deepak & N Nishakumari, K & Devanand, P & M Sasi, P. "Missing Child Identification System Using Deep Learning and Multiclass SVM", (2018).

[6]. Bharath Darshan Balar, D S Kavya, Chandana M, Anush E, Vishwanath R Hulipalled, "Efficient Face Recognition System for Identifying Lost People", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue5S, May 2019.