2018 IEEE Recent Advances in Intelligent Computational Systems (RAICS) | December 06 - 08, 2018 | Trivandrum

Missing Child Identification System using Deep Learning and Multiclass SVM

Pournami S. Chandran[1], Byju N B[2], Deepak R U[3], Nishakumari K N[4], Devanand P[5], Sasi P M[6]

Centre for Development of Advanced Computing (C-DAC)
Thiruvananthapuram, India
pournami@cdac.in[1], byjunb@cdac.in[2], deepakru@cdac.in[3], nishakn@cdac.in[4], devanand@cdac.in[5], sasipm@cdac.in[6]

***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 multitude of children available, with the help of face recognition. The public can upload photographs of 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 photo with best match will be selected from the database of missing children. For this, a deep 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. The Convolutional Neural Network (CNN), a highly effective deep learning technique for image based applications is adopted here for face recognition. Face descriptors are extracted from the images using a pre-trained CNN model VGG-Face deep architecture. Compared with normal deep learning applications, our algorithm uses convolution network only as a high level feature extractor and the child recognition is done by the trained SVM classifier. Choosing the best performing CNN model for face recognition, VGG-Face and proper training of it results in a deep learning model invariant to noise, illumination, contrast, occlusion, image pose and age of the child and it outperforms earlier methods in face recognition based missing child identification. The classification performance achieved for child identification system is 99.41%. It was evaluated on 43 Child cases.**

Keywords— Missing child identification, face recognition, deep learning, CNN, VGG-Face, Multi class SVM.

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 populous country in the world and children represent a significant percentage of 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 an 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 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 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. RELATED WORKS

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 gives better performance in face recognition than handcrafted features.

In [4], missing child identification is proposed which employees principal component analysis using Eigen vectors is used for face recognition system.

FindFace is a website that lets users search for members of the social network VK by uploading a photograph [5]. FindFace 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 percent.

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 public.

III. WORK FLOW OF FACE RECOGNITION

Here we propose a methodology for missing child identification which combines facial feature extraction based on deep learning and matching based on support vector machine. The proposed system utilizes face recognition for missing child identification. This is to help authorities and parents in missing child investigation. The architecture of the proposed frame work is given below,

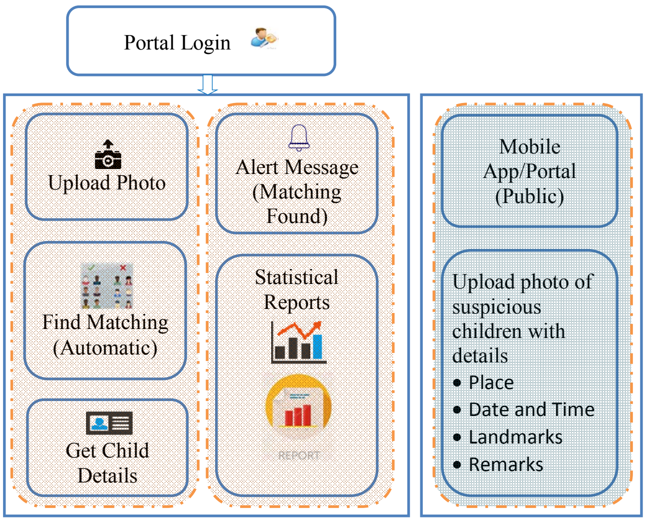

Fig. 1. Architecture of proposed child identification system

It consists of a national portal for storing details of 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. 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. The system also generates various statistical reports.

The public can upload photo of any suspicious child at any time into the portal with details like place, time, landmarks and remarks. The photo uploaded by the public will be automatically compared with photos of the registered missing children and if a matching photo with sufficient score is found, then an alert message will be sent to the concerned officer. The message will also be visible in the message box of the concerned officer login screen. The portal for the public can also be maintained as a mobile app, where he or she can upload photo of suspicious children with details. In the mobile app, location of the person updating the photo will also be automatically recorded.

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

In the following sections the paper details the work flow for child matching methodology. The flow chart of the automatic child face identification methodology is as shown in Fig 2.

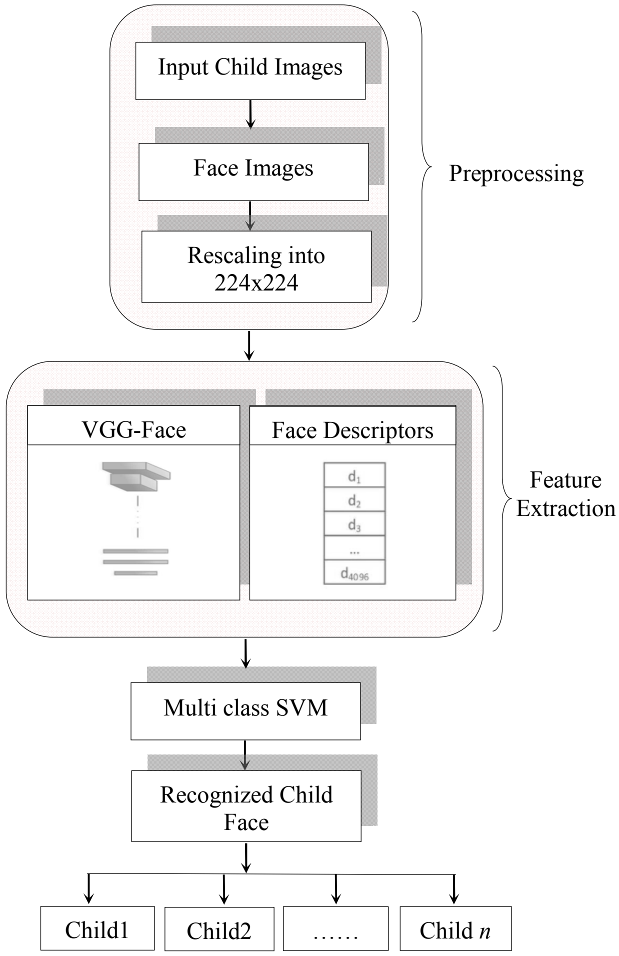

Fig. 2. Software Flow of face recognition system

Images of reported missing children are saved in a repository and the face area is selected for cropping to obtain input face images. Learned features from a Convolutional Neural Network (CNN), a specific type of deep learning algorithm, are used for training a multi class SVM classifier. This machine learning approach is used to correctly label the child using the name indicated in the database provided by the concerned authority.

IV. CONVOLUTIONAL NEURAL NETWORKS(CNN)

Convolutional Neural Networks (CNNs) are essential tools for deep learning methods and are more appropriate for working with image data [7]. CNNs or ConvNets are composed of series of interconnected layers and these layers consist of repeated blocks of convolutional, ReLU (rectified linear units), pooling layers and fully connected layers.

Convolutional layer convolves the input face image data with different kernels to produce activation maps or feature maps representing low level features like edges or curves. This feature map is given to next convolutional layer producing activations which represent high level features indicating landmarks in face. The Convolutional layer basically defines a set of filter weights which are updated during network training.

ReLU followed by each convolutional layer introduces nonlinearity in the system. This layer applies the function f(x) = max(0,x) to the input data of the layer.

The pooling layers merge similar features into one by down sampling with suitable size. The basic idea behind pooling layer is that the relative position to other feature is more important than the exact location of a specific feature. It reduces the dimensions of feature maps and network parameters.

The final layer called fully connected layer outputs the number of classes. There are several fully-connected layers converting the 2D feature maps into a 1D feature vector, for further feature representation.

A. VGG-Face CNN descriptor

A very deep CNN called VGG-Face network [8] is used for face recognition and its architecture is given in full detail in Fig 3. The CNN architecture comprises 11 blocks, each containing a linear operator followed by one or more non-linearities such as ReLU and max pooling. The first eight such blocks are said to be convolutional as the linear operator is a bank of linear filters (linear convolution). It uses filters of size 3x3 with stride and pad of 1, throughout the network. All the convolution layers are followed by a rectification layer (ReLU). Max pooling layers used only 2x2 size with stride 2. The last three blocks are fully connected layers, they are the same as a convolutional layer, but the size of the filters matches the size of input data, such that each filter provides representative data from the entire image. Output of the first two FC layers are 4096 dimensional and the last FC layer has 2622 dimensions followed by L-dimensional metric embedding. Optimization is done by stochastic gradient descent using mini-batches of 64 samples and momentum coefficient of 0.9

layer	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
type	input	conv	relu	conv	relu	mpool	conv	relu	conv	relu	mpool	conv	relu	conv	relu	conv	relu	mpool	conv
name	–	conv1_1	relu1_1	conv1_2	relu1_2	pool1	conv2_1	relu2_1	conv2_2	relu2_2	pool2	conv3_1	relu3_1	conv3_2	relu3_2	conv3_3	relu3_3	pool3	conv4_1
support	–	3	1	3	1	2	3	1	3	1	2	3	1	3	1	3	1	2	3
filt dim	–	3	–	64	–	–	64	–	128	–	–	128	–	256	–	256	–	–	256
num filts	–	64	–	64	–	–	128	–	128	–	–	256	–	256	–	256	–	–	512
stride	–	1	1	1	1	2	1	1	1	1	2	1	1	1	1	1	1	2	1
pad	–	1	0	1	0	0	1	0	1	0	0	1	0	1	0	1	0	0	1

layer	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
type	relu	conv	relu	conv	relu	mpool	conv	relu	conv	relu	conv	relu	mpool	conv	relu	conv	relu	conv	softmx
name	relu4_1	conv4_2	relu4_2	conv4_3	relu4_3	pool4	conv5_1	relu5_1	conv5_2	relu5_2	conv5_3	relu5_3	pool5	fc6	relu6	fc7	relu7	fc8	prob
support	1	3	1	3	1	2	3	1	3	1	3	1	2	7	1	1	1	1	1
filt dim	–	512	–	512	–	–	512	–	512	–	512	–	–	512	–	4096	–	4096	–
num filts	–	512	–	512	–	–	512	–	512	–	512	–	–	4096	–	4096	–	2622	–
stride	1	1	1	1	1	2	1	1	1	1	1	1	2	1	1	1	1	1	1
pad	0	1	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0

Fig. 3. VGG-Face network architecture

V. PREPROCESSING

Preprocessing input raw image in the context of face recognition involves acquiring the face region and standardizing images in a format compatible with the CNN architecture employed. Each CNN has a different input size requirement. The photographs of missing child acquired by a digital camera or mobile phone are taken and categorized into separate cases for creating the database of face recognition system. The face region in each image is identified and cropped for getting the input face images. The cropped face images are resized to 224x224 because VGG-face network can process only RGB images in this particular size. The input to the deep network is fixed sized image with mean face image, computed from all the training set images, subtracted.

VI. EXTRACTION OF FACIAL FEATURES

VGG-Face is trained to recognize the 2622 identities and other classes can't be identified using this. But the activation vectors extracted from VGG-Face architecture can be used as the feature representations to classify each child category. The last classification layer is removed and extracts the 4K dimensional features from the first fully connected layers. The resulting feature vector is normalized by dividing each component by the L2 norm of this 4096 dimensional vector. Thus the pre-trained CNN VGG-Face is made to perform as an automatic facial feature extractor for training the classifier.

VII. MULTI CLASS SVM CLASSIFIER

Each face image corresponds to a child and child face recognition is considered as an image category classification problem. The task is to classify input image uploaded by the public into one of the given category based on the image representation. Basically CNN architecture consists of computational layers for feature extraction and a classifier layer at the final stage. The VGG-face CNN model employs the softmax activation function for labeled class prediction, suggesting the class each image belongs to. The softmax in the CNN layers is replaced with a multi class SVM trained with feature vector array from each image. One-versus-rest linear SVM classifier is used and is trained on the dataset. Extracted feature vector array is used to train this classifier.

VIII. RESULTS AND DISCUSSIONS

The face identification algorithm is implemented using MATLAB 2018a platform. The experiments are carried on Microsoft Windows 7, 64 bit Operating System with Intel core i7, 3.60GHz processors having 32GB RAM. For dealing with CNN architectures additional processing capability is needed. Use of GPU is recommended for training the models and Nvidia GeForce TitanX 12GB graphics card is used.

The user defined database includes 846 child face images with 43 unique children cases. Training and test set is prepared by splitting the database images. 80% of images from each child category are selected for training and 20% for testing, resulting in 677 training set images and 169 test set images. The training set and validation set consists of images of each child in the earlier days and testing is done with images of children after an age gap to evaluate the system in all conditions.

CNN implementation is based on MatConvNet package [9] with deep integration of CNN building blocks in MATLAB environment. Pre-trained VGG-Face CNN is also provided by MatConvNet. For the experiments here MatConvNet 1.0-beta25 version is downloaded and used.

The training set images are preprocessed to the size specified by the CNN architecture before passing to the CNN model. The face region is cropped within a rectangular region from every image of the acquired input database. The images fed to VGG-Face are of fixed size by rescaling to 224x224. The activations to the input image produced by the first fully connected layer of the VGG-Face network architecture is taken as the CNN Feature descriptor. The normalized feature vector, each having a length of 4096, is used for training the SVM classifier for classifying the image of face and recognizes the child.

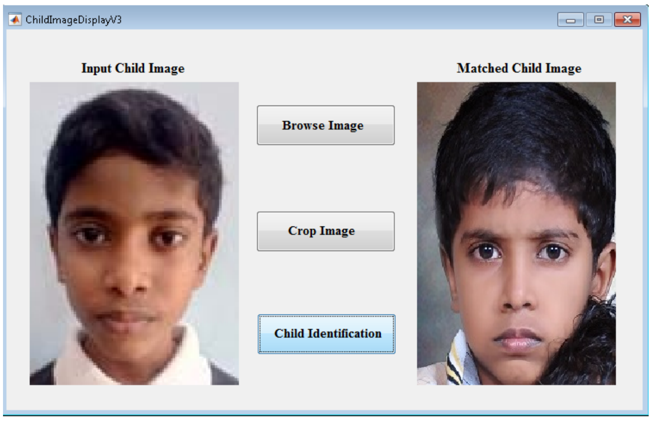

Fig. 4. GUI for child identification showing an input image and matched output image in the database

To assess the flexibility of face recognition deep architecture against variations in image quality, artificially degraded images are created. Images obtained by changing noise level, brightness, contrast, lighting conditions, obstructions, blur, aspect ratio and face positions are used for testing the child identification system.

Face identification accuracy is computed as the ratio of correctly identified face images to the total number of child face images in the test set.

$$Accuracy = \frac{Correctly\ recognized\ face\ images}{Total\ number\ of\ child\ face\ images} \quad (1)$$

The computed recognition accuracy of the multi class SVM using learned features from CNN is 99.41%.

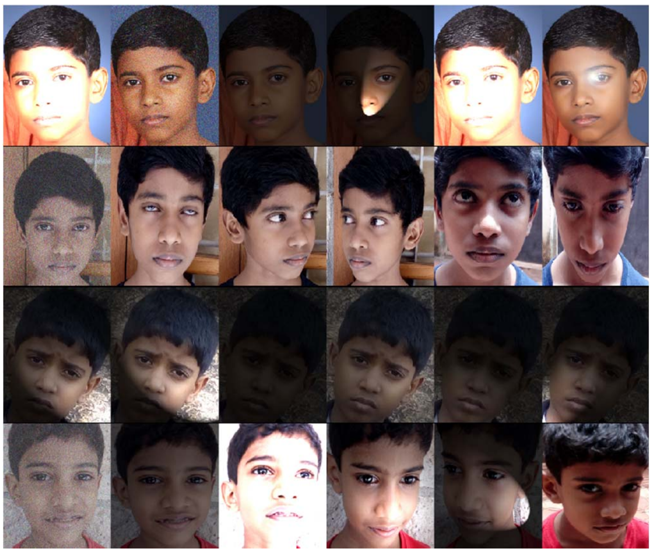

Fig. 5. Images with variations correctly classified by the system

IX. 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 of children. 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.

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] https://en.wikipedia.org/wiki/FindFace

[6] https://www.reuters.com/article/us-china-trafficking-apps/mobile-app-helps-china-recover-hundreds-of-missing-children-idUSKBN15J0GU

[7] 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.