Missing child identification system using deep learning and multiclass SVM

Mrs. A. Yamuna[1], K. Devi[2]

[1]Assistant Professor, Dept of MCA, Audisankara College of Engineering and Technology (AUTONOMOUS), Gudur, AP, India.

[2]PG Scholar, Dept of MCA, Audisankara College of Engineering and Technology (AUTONOMOUS), Gudur, AP, India.

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.

1. 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 Journal of Information and Computational
Science 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.

2. LITERATURE SURVEY

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

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

[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. Still-to-video face recognition (FR) plays an important role in video surveillance, allowing to recognize individuals of interest over a network of video cameras. Watchlist screening is a challenging video surveillance application, because faces captured during enrollment (with still camera) may differ significantly from those captured during operations (with surveillance cameras) under uncontrolled capture conditions (with variations in, e.g., pose, scale, illumination, occlusion, and blur). Moreover, the facial models used for matching are typically designed a priori with a limited number of reference stills. In this paper, a multi-classifier system is proposed that exploits domain adaptation and multiple representations of face captures. An individual-specific ensemble of exemplar-SVM (e-SVM) classifiers is designed to model the single reference still of each target individual, where different random subspaces, patches, and face descriptors are employed to generate a diverse pool of classifiers. To improve robustness of face models, e-SVMs are trained using the limited number of labeled faces in reference stills from

the enrollment domain, and an abundance of unlabeled faces in calibration videos from the operational domain. Given the availability of a single reference target still, a specialized distance-based criteria is proposed based on properties of e-SVMs for dynamic selection of the most competent classifiers per probe face. The proposed approach has been compared to reference systems for stillto- video FR on videos from the COX-S2V dataset. Results indicate that ensemble of e- SVMs designed using calibration videos for domain adaptation and dynamic ensemble selection yields a high level of FR accuracy and computational efficiency.

[3] C. Geng and X. Jiang, "Face recognition using sift features", IEEE International Conference on Image Processing (ICIP), 2009.

Scale Invariant Feature Transform (SIFT) has shown to be a powerful technique for general object recognition/detection. In this paper, we propose two new approaches: Volume-SIFT (VSIFT) and Partial-Descriptor-SIFT (PDSIFT) for face recognition based on the original SIFT algorithm. We compare holistic approaches: Fisher face (FLDA), the null space approach (NLDA) and Eigenfeature Regularization and Extraction (ERE) with feature based approaches: SIFT and PDSIFT. Experiments on the ORL and AR databases show that the performance of PDSIFT is significantly better than the original SIFT approach. Moreover, PDSIFT can achieve comparable performance as the most successful holistic approach ERE and significantly outperforms FLDA and NLDA.

3. PROPOSED SYSTEM

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.

The proposed system is comparatively an easy, inexpensive and reliable method compared to other biometrics like finger print and iris recognition systems. features extracted using a CNN network for getting facial representations gives better performance in face recognition than handcrafted features This is to help authorities and parents in missing child investigation

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.

Convolutional Neural Network (ConvNet/CNN):

A Convolutional Neural Network

(ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The preprocessing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are handengineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

Image Dimensions = 5 (Height) x 5 (Breadth) x 1

(Number of channels, eg. RGB)

In the above demonstration, the green section resembles our 5x5x1 input image, I. The element involved in carrying out the convolution operation in the first part of a Convolutional Layer is called the Kernel/Filter, K, represented in the color yellow. We have selected K as a 3x3x1 matrix.

VGG Neural Networks:

While previous derivatives of AlexNet focused on smaller window sizes and strides in the first convolutional layer, VGG addresses another very important aspect of CNNs: depth. Let's go over the architecture of VGG:

- Input. VGG takes in a 224x224 pixel RGB image. For the ImageNet competition, the authors cropped out the center 224x224 patch in each image to keep the input image size consistent.

- Convolutional Layers. The convolutional layers in VGG use a very small receptive field (3x3, the smallest possible size that still captures left/right and up/down). There are also 1x1 convolution filters which act as a linear transformation of the input, which is followed by a ReLU unit. The convolution stride is fixed to 1 pixel so that the spatial resolution is preserved after convolution.

- Fully-Connected Layers. VGG has three fully-connected layers: the first two have 4096 channels each and the third has 1000 channels, 1 for each class.

- Hidden Layers. All of VGG's hidden layers use ReLU (a huge innovation from AlexNet that cut training time). VGG does not generally use Local Response Normalization (LRN), as LRN increases memory consumption and training time with no particular increase in accuracy.

- The Difference. VGG, while based off of AlexNet, has several differences that separates it from other competing models:

- Instead of using large receptive fields like AlexNet (11x11 with a stride of 4), VGG uses very small receptive fields (3x3 with a stride of 1). Because there are now three ReLU units instead of just one, the decision function is more discriminative. There are also fewer parameters (27 times the number of channels instead of AlexNet's 49 times the number of channels).

- VGG incorporates 1x1 convolutional layers to make the decision function more nonlinear without changing the receptive fields.

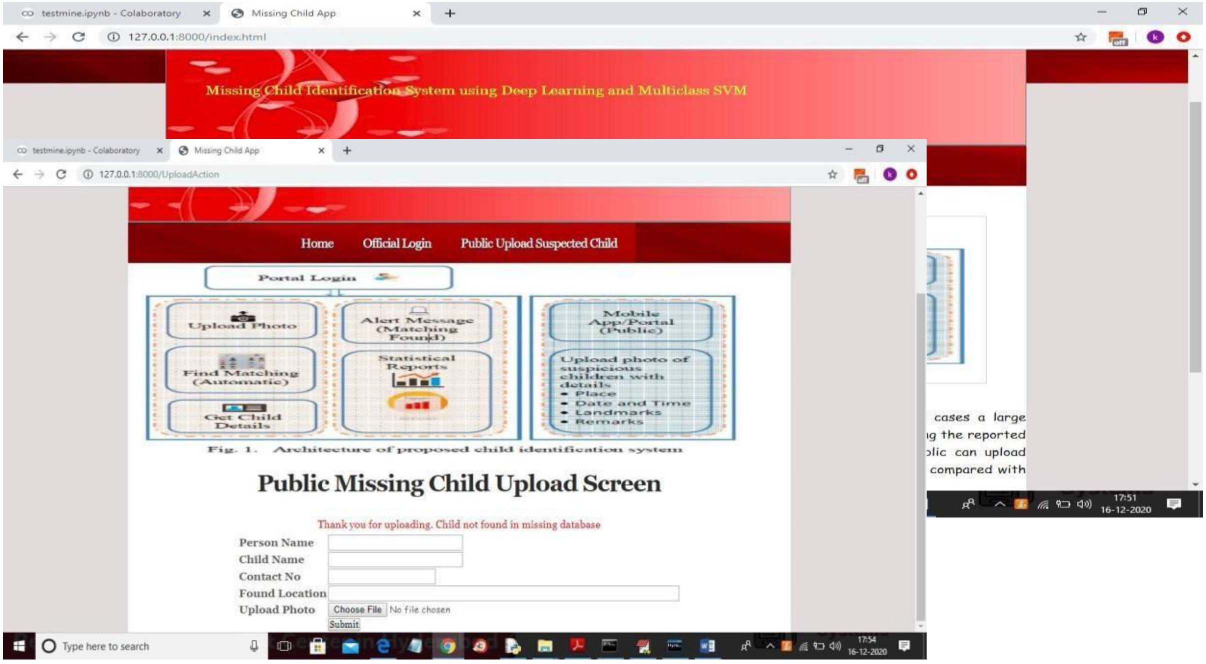
- The small-size convolution filters allows VGG to have a large number of weight layers; of course, more layers leads to improved performance. This isn't an uncommon feature, though. GoogLeNet, another model that uses deep CNNs and small convolution filters, was also showed up in the 2014 ImageNet competition.

SUPPORT VECTOR MACHINE(SVM)

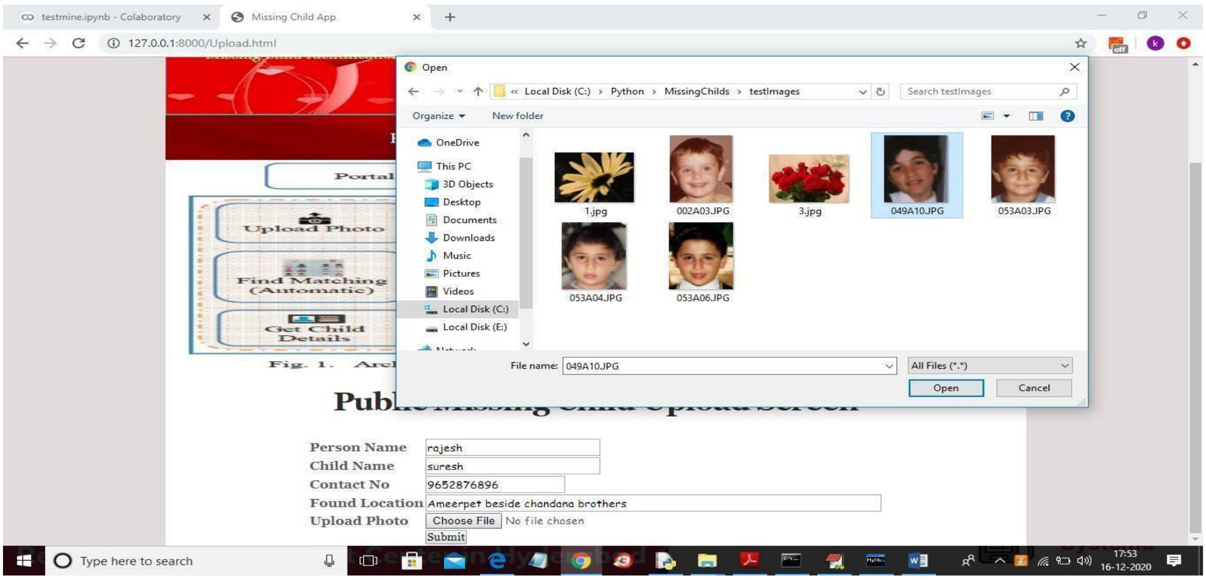
"Support Vector Machine" (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges.

However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very well (look at the below snapshot). The SVM algorithm is implemented in practice using a kernel. The learning of the hyperplane in linear SVM is done by transforming the problem using some linear algebra, which is out of the scope of this introduction to SVM. A powerful insight is that the linear SVM can be rephrased using the inner product of any two given observations, rather than the observations themselves.

4 RESULTS

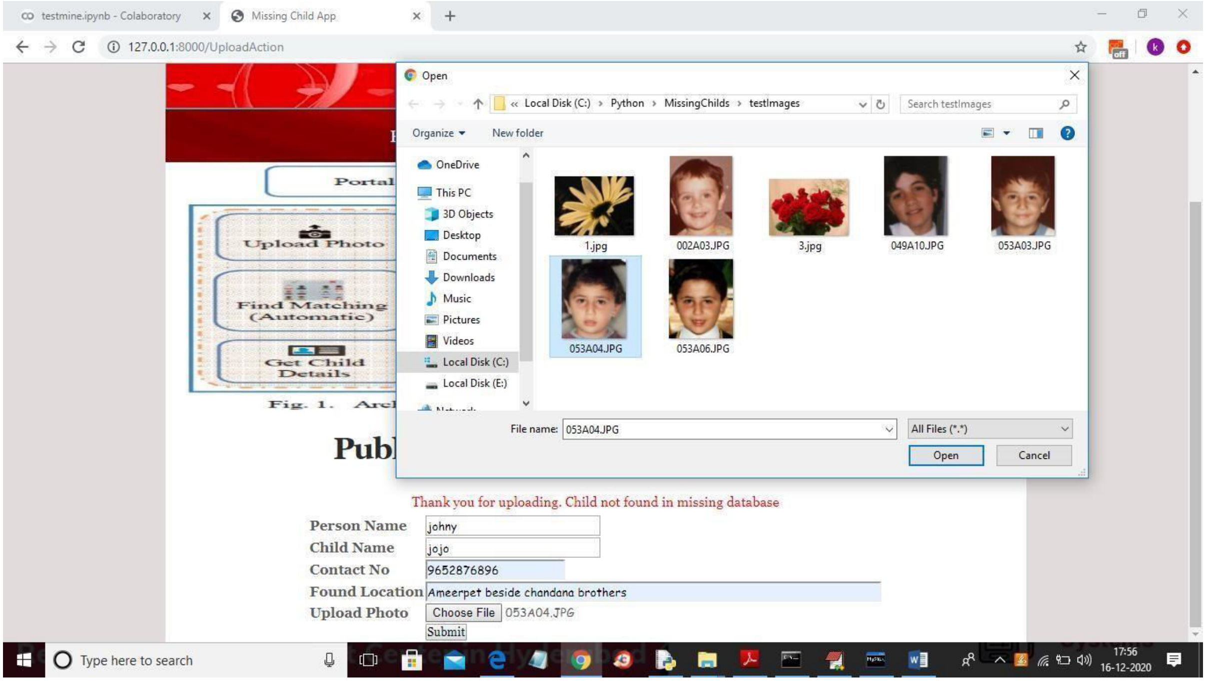


In above screen public can click on 'Public Upload Suspected Child' link to get below page and to add missing child details

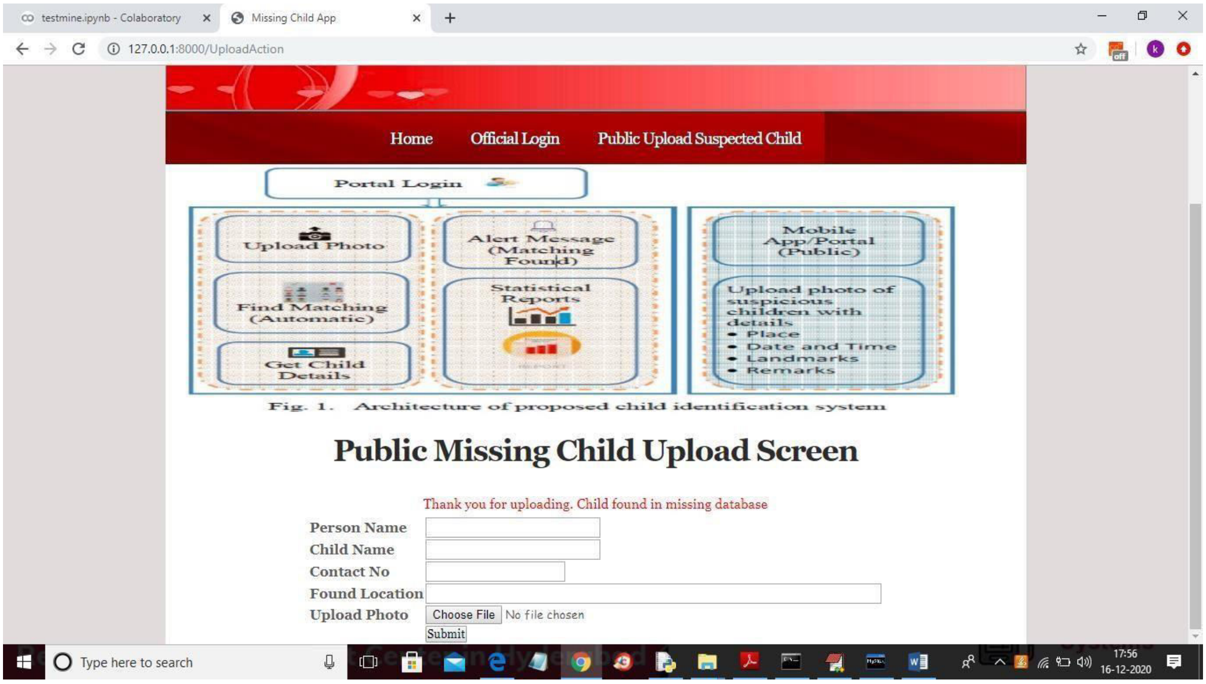


In above screen public will enter suspected child details and then upload photo and then click on 'Submit' button and to get below result

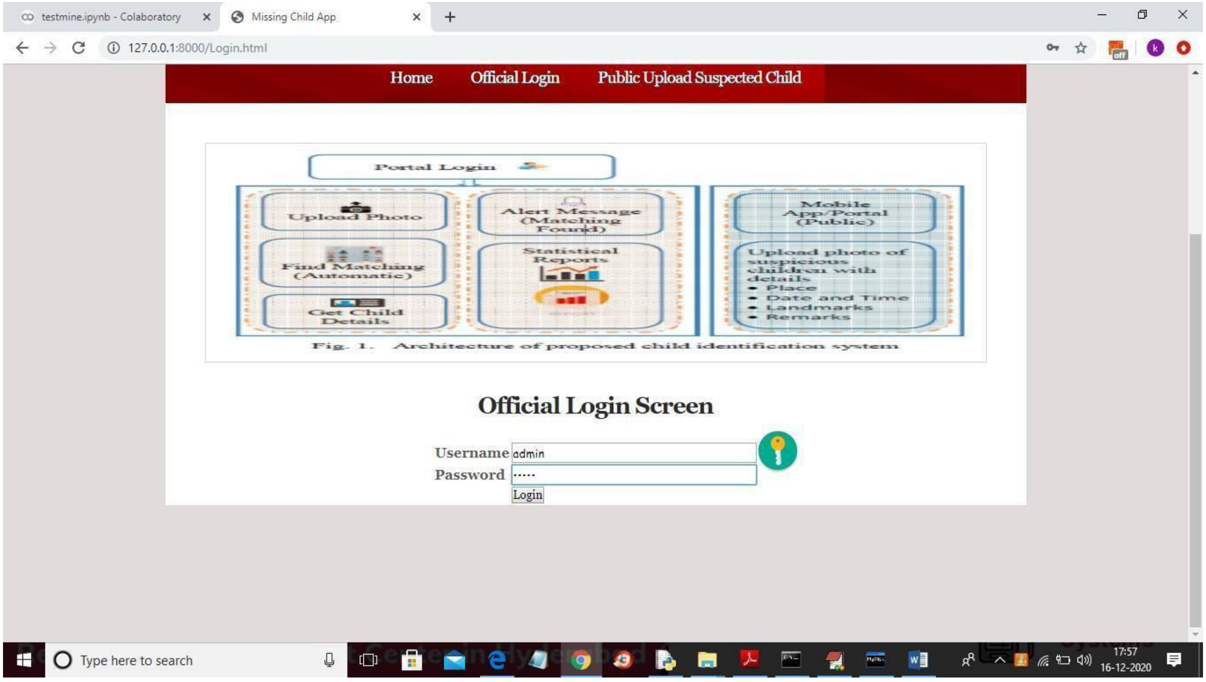
In above screen we can see child not found in missing DB and we can try with other image



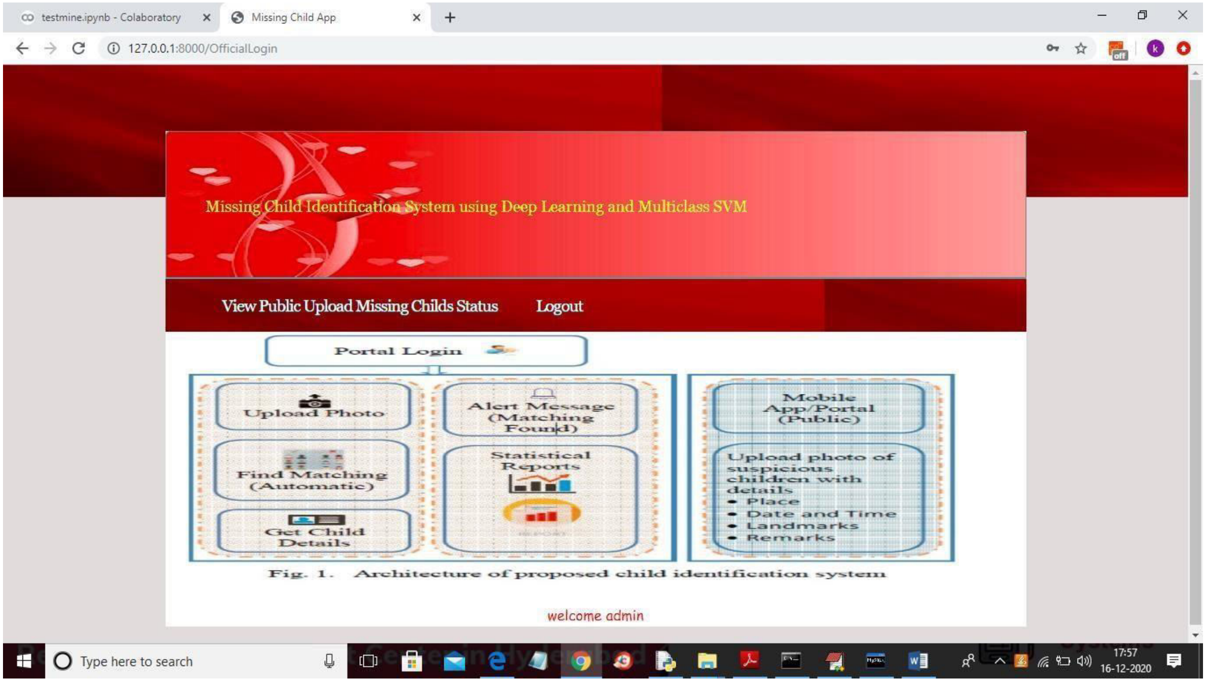
And below is the result for new above child details



In above screen uploaded child found in database and now click on 'Official Login' link to get below login screen

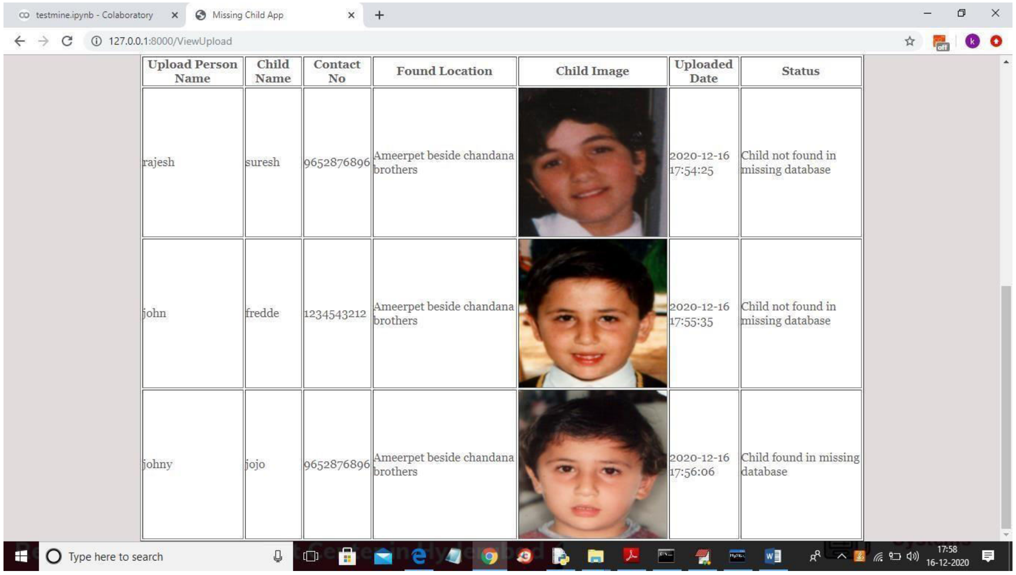


In above screen admin can login by entering username and password as 'admin' and 'admin'



and after clicking on 'Login' button will get below screen

In above screen official can click on 'View Public Upload Missing Childs Status' link to view all uploads and its result done by public.



In above screen officials can see all details and then take action to find that child

5.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.

6. 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] RohitSatle, VishnuprasadPoojary, John Abraham, ShilpaWakode, "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/uschina-trafficking-apps/mobileapp-helpschina-recover-hundreds-of-missing-childrenidUSKBN15J0GU

[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. 2018 IEEE Recent Advances in Intelligent Comp

Author's Profile:

Mrs. A. YAMUNA Currently Working as Assistant Professor in Audisankara College of Engineering and Technology AUTONOMOUS Gudur, Tirupathi (Dt), Andhra Pradesh, India.

Ms. K. DEVI is pursuing MCA from Audisankara College of Engineering and Technology, AUTONOMOUS, Gudur. Affiliated to JNTUA, Andhra Pradesh, India.