Identifying lost children using Deep Learning

Amisha Pupala   
Department of Computer Engineering  
Ramrao Adik Institute of TechnologyNavi Mumbai, India  
amishapupala20@gmail.com

Mrs. Smita Bharne  
Department of Computer Engineering  
Ramrao Adik Institute of TechnologyNavi Mumbai, India  
smita.bharne@rait.ac.in

Samruddhi Mokal  
Department of Computer Engineering  
Ramrao Adik Institute of TechnologyNavi Mumbai, India  
samruddhimokal10@gmail.com

Neha Pandit  
Department of Computer Engineering  
Ramrao Adik Institute of TechnologyNavi Mumbai, India  
panditneha124@gmail.com

*Abstract*— A countless number of children go missing in India every year. And a large number of children remain untraced due to the problem in the identification of children using the photograph. As the state-of-the-art missing children identification systems using face recognition fall short in identifying children at later stages because of aging. This paper presents a novel use of face recognition with face aging to overcome the limitation of existing systems. We have made a portal where the public can upload images of missing children and can also search for a lost child. We have used the Age C-GAN algorithm for face aging and the FaceNet algorithm for face feature extraction and face recognition.

Keywords—component, formatting, style, styling, insert (key words)

# Introduction

Children represent a significant percentage of the total population of India. Unfortunately, many children go missing every year due to various reasons like kidnapping, trafficked children and lost children, run-away children, and abduction. According To the National Crime Record Bureau 2019 report, the number of missing children has increased by 8.9% since 2019 giving a total of 73,137 children reported missing last year. Many NGOs claim that the actual number of missing children cases is much more than reported. A child lost in one location can be found in some other location. Even if the child is found it gets difficult to recognize him/her. The face of Face recognition is thus the most promising biometric technology for recognizing missing children. But if the child is found after many years from the date when he/she is lost the state-of-art face recognition systems fall short to identify children from the photograph because a child's face undergoes many temporal variations due to aging like a change in skin texture, change in the shape of the face, color, weight, facial hair growth, etc. Hence as the time difference between a probe image and a true mate image gets larger, the performance of the face recognition algorithm decrease, and thus the search gets harder. Thus the primary goal of the proposed system is to generate an age-progressed image of the child rather than enhancing the face recognition performance. Thus this system enhances the ability of face matches to identify and locate children who are lost at a young age by aging face features to reunite them back with their families. Our system will be helpful to the police and higher authorities for tracking down missing children quickly.

# Literature Survey

In paper [1] the author has used CNN and VGG face technologies for the identification of missing children. The faces of children are detected and features are extracted by using CNN. CNN is an important tool that works appropriately with image data. Blocks of convolutional, ReLU pooling layers, and fully connected layers are composed in CNNs or ConvNets. For classification, they used SVM Classifier. But these technologies slower the process of classification and extraction of faces from images.

In paper [2] a Face-to-face matching method is used which is based on IPE-based blind computing which is used to restrict access to photos. Using Multitask FaceNet Model, the system can safely run and provides privacy for both the unsure and missing children parts. To provide a more effective face recognition module, the output of images with encryption schemes needs to be improved.

In paper [3] Three Convolutional neural network algorithms were using which are: VGG16, ResNet50, and Mobile-FaceNet. From which Mobile-FaceNet has the smallest model size and provides higher accuracy than VGG16 and ResNet50 and VGG16 has the largest model in size.

In paper [4] the author has used DeepId and modified Inception structures which contain different types such as DeepId2, DeepId3 modules which give the best performance with higher accuracy when there is very little difference between a wide module and deep module. however, the mutual information goes down with additional branches such as DeepId3 in a wide module.

In paper [5] the author had used Age-cGAN to generate artificial images within different age categories. The face recognition method used by this system is the Identity preserving latent vector optimization method used to conserve the original photo's identity while face aging which can be improved by combining Pixelwize and Identity-Preserving methods into one singular method.

# Proposed Work

It is known that the performance of the face recognition algorithm is affected due to aging. It is still a problem to consider the aging parameter under consideration for face recognition as most of the algorithm fails to detect the face of the child after a few years. A child’s face undergoes various changes including facial hair, weight, the texture of the skin, shape, and size of the face, etc. Several studies were performed to analyze the extent to which facial aging affects the performance of face recognition algorithms. Two conclusions can be drawn from these studies:

1. The performance of traditional face recognition algorithms decreases with an increase in time-lapse.
2. Performance of face recognition and detection algorithms decreases more rapidly in the case of younger individuals as compared to older individuals.

Hence it is important to consider the age progression of the face for enhancing the performance of the face recognition algorithm, especially when enrolled at a young age.

In particular, our contributions are as follows:

* A portal for finding missing children that compare the image of the missing child with the images of the already registered missing children.
* To take into consideration many kinds of variations in face images which are taken under uncontrolled conditions such as change of expression, illumination, pose, change in lighting condition, presence of noise, and blurred image.
* Using Age Conditional Generative Adversarial Network to reconstruct high-quality synthetic images of the required age group preserving the original person’s identity. Using methodology like Facenet and C-Gan the process of recognition gets easier which will help the authorities to find the missing children quickly and solve many cases which have been untraceable because of the above issues.

Thus the proposed system presents a novel deep learning methodology to find the missing children using the photo of the children currently available using face recognition. Whenever anyone finds any suspicious child on the road he/she can click the photograph of the child and upload it into the common portal with landmarks and remarks. The public can upload the photograph of the missing child into the portal. Then the photo will automatically be compared with the already registered images in the repository. If no matching record is found then the image will be age progressed and again the age-progressed image will be compared with the registered photos. For this, a deep learning model Aging C-GAN is used. And the face recognition is done using FaceNet. Thus the proposed system outperforms the earlier methods in face recognition based on missing child identification.

As mentioned earlier face aging model is based on an age-conditional generative adversarial network, a model for synthesis and generation of human faces of the required age category.

Hence the task of our cGAN network is to generate the image of a child at different ages based on the input image and target age. The structure of the model consists of three parts: encoder, age condition, and decoder. The encoder is used for feature extraction of the input face image. We are using three groups of convolution layers and nine residual blocks. Each convolution layer consists of a ReLU non-linearity layer, spatial batch normalization layer stride, and a convolution layer. After this, we get a 56 feature map of the image. Since there is a huge range of images of a large number of age groups in our dataset we are dividing the image into y number of groups which is represented as a one-hot age vector. Each block represents a specific age group. The task of the decoder is to convert the face to generate an image of the required age group based on the features.

# Proposed Methodology

## Age C-GAN Algorithm

**Conditional Generative Adversarial Network is conditioned on extra information and is based on the idea of Vanilla GAN which allows us to control the output of the generator network. In this, we have to give extra information y as an additional layer to the generator. As compared to vanilla GAN, In vanilla GANs we cannot control the category of the retrieved images as it can learn only one category, however, in CGAN we can generate images of the specific category using condition y. y can we any data depending on the application for which we are building the model for like integer data or class label. Thus CGAN can be used to generate models with different categories and conditions.**

##### 

**The Face Aging C-GAN has four networks:**

**Encoder: Encoder is a deep convolution neural network used to generate the latent vector of images.**It is used to learn the inverse mapping of input image and age condition with the latent vector Z.

* It generates a latent vector of 100 dimensions of the input face images of dimension(64,64,3). There are 2 dense layers and 4 convolution blocks.
* All the convolution blocks except the first layer have a convolution layer which is followed by a batch normalization layer and then an activation function.

**Generator Network:** It is a deep convolutional network that takes a face image and condition vector and tries to generate realistic images of the face.

* It is a CNN having upsampling, convolutional, and dense layers which take a condition vector which is additional information added to the network which is age, and a latent vector to generate a realistic image of a dimension of (64, 64, 3). The condition vector is the additional information that is provided to the network. For the Age-cGAN, this will be the age.

**Discriminator Network:**It is CNN that is used to discriminate between the real image and the fake image. It contains several convolutional blocks which contain a batch normalization layer, convolutional layer, and an activation function.

**FaceNet:**It is a face recognition model that learns the difference between the original image x(input image) and the generated image x’. it is used to recognize the person's identity in the image. We are using a ResNet2

For FaceNet Inception ResNet, ResNet-50, or pre-trained Inception can be used. We are using the ResNet2 model without fully connected layers. FaceNet is used to identify the person’s identity in the given input image. The extracted embeddings for the original input image and the reconstructed image can be found by calculated Euclidean distance between the embeddings.

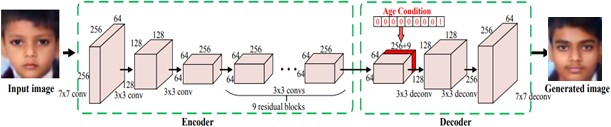


Fig. 1. Layers used in Aging cGAN

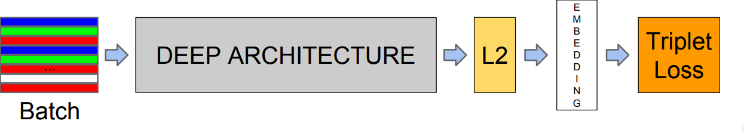
Stages of Age-cGAN:

Age-cGAN has four networks and it is trained in three stages:

1. **Conditional GAN training:** In this stage, we have trained the generator network which generates the blurred images of the face after training and the discriminator network.
2. **Initial latent vector approximation:** In this stage, we have trained the encoder network to approximate the latent vector using both the real images and the generated images which generates the latent vector from learned distribution.
3. **Latent vector optimization:** In this stage, we have optimized both the encoder and the generator network.

## FaceNet

We are using FaceNet in our work for feature extraction and face recognition. We have first passed all the images through Multi-Task Cascaded Convolutional Neural Network (MTCNN) for finding and extracting faces from the images. Then we have used FaceNet for extracting high-quality features from children's faces which we have used to train a face identification system. It uses Convolution layers to learn embeddings from the face directly which are used for face verification and face recognition. These embeddings are of 128 dimensions and insert them into feature space such that the squared distance between faces of the same person is less whereas the squared distance between images to two different people is large.

Fig. 2. Model architecture of FaceNet.

# Dataset

## Dataset for Face Recognition

We have used two separate datasets for the training of face recognition and the Age Progression model. For face recognition, we have created our dataset by collecting images of children from various sources. We have created a directory structure where each child to be recognized is having a dedicated directory with their images saved in it. The photos in the directory provide a range of orientation, sizes, lighting conditions. If an image contains multiple faces then the face having the highest probability will be considered for further steps. We are using this dataset as the basis for our classifier which is trained on a training dataset only.

## IMDB Wiki-Crop Dataset

For training the Age progression model we have used the IMDB WIKI-crop dataset which is the largest publically available dataset of human faces. It contains about 400 thousand images where each image is labeled in the format gender\_age\_name with all the metadata information like dob, the year when the photo was taken, gender, face location, face, and secondary face score, etc in .mat file. The dataset contains images of all the age categories from 0 to 100 years of age which is not required for the model. So we have first filtered the dataset and considered the images of children in the age group from 0-15 years of age and have deleted the rest of the images. The dataset contains images with huge variations in facial expression, illumination, resolution, occlusion, pose, etc. The dataset also provides the corresponding landmarks. This dataset can be used for a variety of tasks including age estimation, age progression, landmark localization, progression/ regression, face detection, etc.



Fig. 3. Sample images of different individuals from IMDB Wiki-Crop dataset face dataset.

# System Design

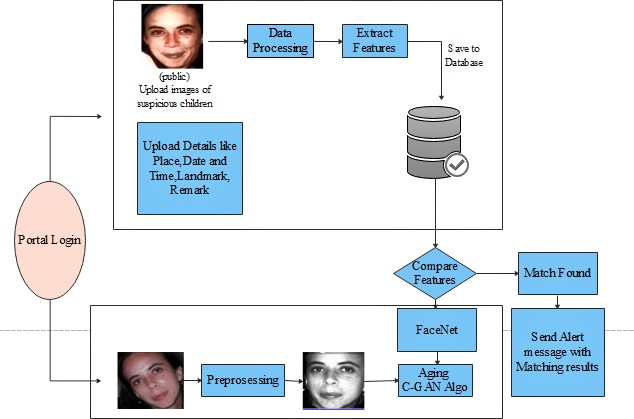


Fig. 4. System design for lost child idetification using FaceNet and C-GAN

The proposed system consists of a portal where the public can store the details with a photograph of the missing child. The public here refers to the parent of the missing child or the police officer. By using the photograph the public can find the matching children from the already stored data. The system will prompt the most matching cases after applying a face recognition algorithm to the uploaded image. Once the matching is done then the concerned person will get the details of the child.

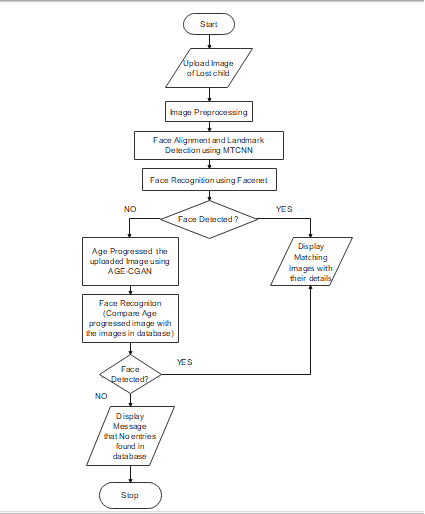


Fig. 5. Proposed Flow of the System

The proposed model is divided into various phases. First, whenever the public finds any missing or suspicious child they can upload the photograph of the child with the details like name of the child, age, place where the child was found, date and time when the child was found and remarks like what clothes the child was wearing when he was spotted or any information the child might have told to the person which can be helpful to find the child.

This uploaded image goes through an image pre-processing phase. The input to this is the raw image that is uploaded by the public and the output of this phase is an aligned frontal face image. The main purpose of this step is to align all the face images based on eye coordinates such that all the images are aligned with the same standard size.

The next stage is to extract the required face features. In this stage, we have used a Multi-task Cascaded Convolutional Networks (MTCNN) model which is a framework used for face alignment and face detection which consists of three stages of convolutional networks that can detect landmark location such as eyes, nose, and mouth and can recognize face very efficiently. The preprocessing step includes removing noise from the image, straighten the image, and detecting and cropping the face from the whole image. The extracted face features will be stored in the database for further processing.

To search for a missing child parent of the child need to upload a photograph of the child then this image will go through the face recognition algorithm after the image pre-processing and feature extraction step. If any matching photo is found in the database then the image of the child will be displayed on the screen. If no matching image is found then the uploaded image goes through the Age CGAN algorithm. The GAN algorithm is used to generate a new age-progressed image with the same statistics as the training images. As the child may be spotted after many years, the face of the child changes from the last known photograph taken before the child had gone missing.

So this model will be used to synthesis the child's face to produce an age-progressed facial image of the child. So basically this model will be used to synthesis child face to the required age category.

This face synthesized and age-progressed image goes through the FaceNet algorithm for feature extraction and face recognition. The main distinction between FaceNet and other face recognition algorithms is that FaceNet does not use any bottleneck layer for verification and recognition tasks but rather learns the mapping directly from images and creates embedding. The FaceNet is used to automatically compare the age-progressed photo of the missing child with the already present images in the database and if the photo matches with any of the photos present in the database then it will be displayed on the screen along with details like name, location and time where the child was found.

# Result and analysis

**Performance of face recognition algorithm was evaluated on child dataset before and after age progressing the child images. The Age cGAN model was trained on the IMBD Wiki-Crop dataset. The IMDB Wiki-Crop Dataset contains more than 400k images with different age groups ranging from 0-100 years. We have first filtered that dataset and have just used images from age group 0-15 years to train the model.**



Fig. 7. Sample images of different individuals from IMDB Wiki-Crop dataset face dataset.

Based on their experiments, the following conclusions are

derived:

* The accuracy of the face recognition algorithm with the images of the same age category is: train accuracy:87.38% and test accuracy is:81.07%.
* **The**Accuracy of face recognition algorithm with images if different age category: 44.76%.
* **The**Accuracy of FaceNet algorithm with images of different age category after age progressing images using Age-cGAN algorithm: 47.32%.

1. Model accuracy

|  |  |  |
| --- | --- | --- |
| **Model** | **Age Category** | **Accuracy** |
| Face Recognition | Same | 81.07% |
| Face Recognition | Different | 44.76% |
| Face Recognition on Age Progressed images | Different | 47.32% |

Based on the experiments, the following conclusions are derived:

* The model is correctly recognizing the misaligned faces and images clicked in an uncontrolled environment like the change of expression, illumination, pose, change in lighting condition and presence of noise.
* The average decrease in the performance of face recognition is more than 40% due to age variation in the faces of the children.
* Intra-person variations like expression and pose are degrading the performance of the system is similar to face aging. Also, large self-occlusion, blur, and misaligned faces are slightly degrading the performance of the model.

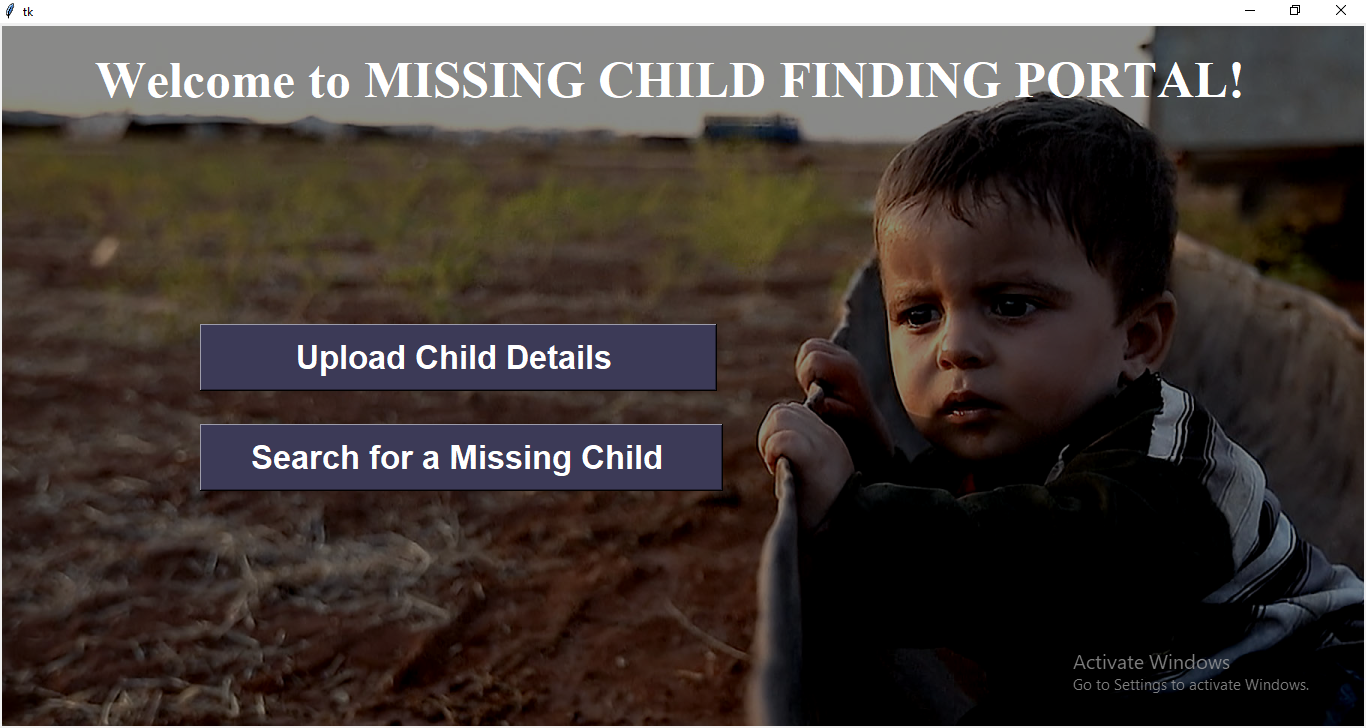


Fig. 6.a. Welcome page of a missing child finding portal

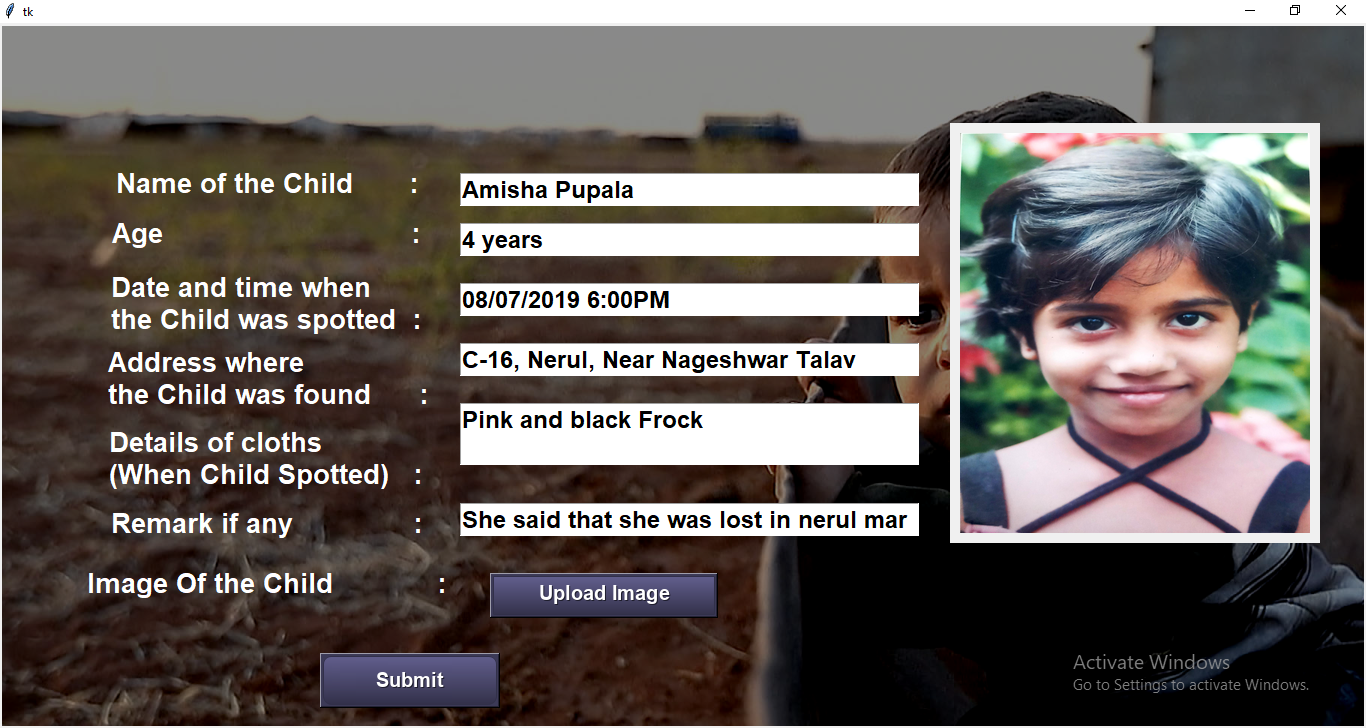


Fig. 6.b. Page where the public can upload details of a suspicious or lost child

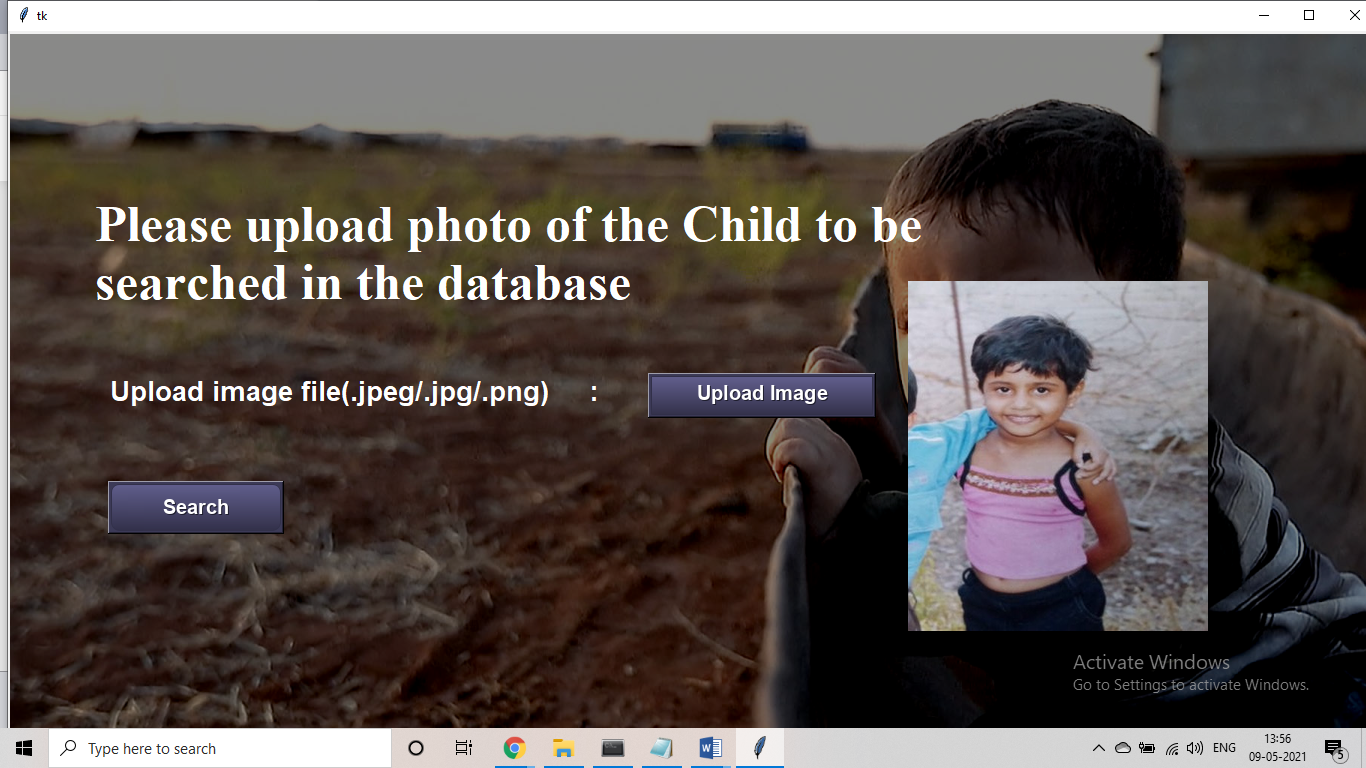


Fig. 6.c. Page to search for a lost child using his/her photograph

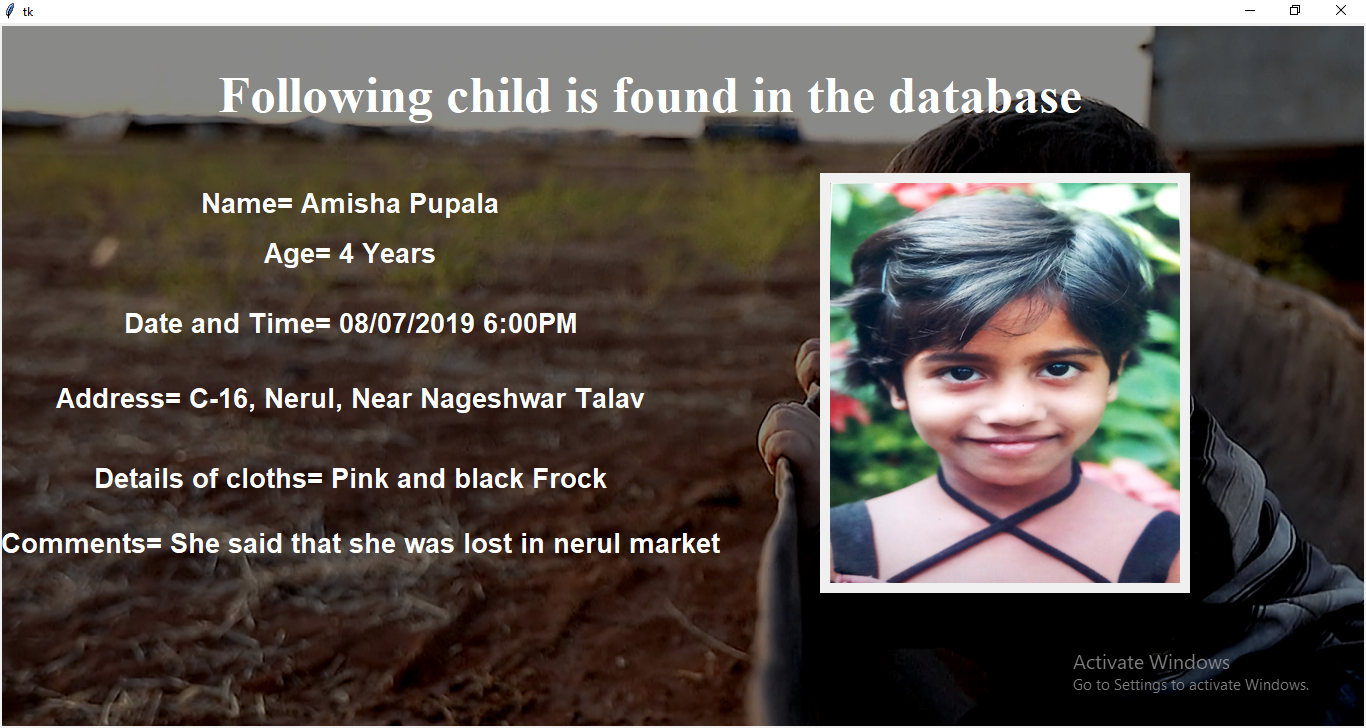


Fig. 6.d. Details of the child found in the database

## Limitations

* Training of the deep learning model which requires a large dataset remains a challenge when training with a small dataset hence accuracy is compromised.
* The age-progressed images get blurred which eventually affects the performance of the model.
* The performance of the face recognition model heavily depends on the accuracy of age-progressed images which are obtained.
* Due to the long duration of code execution and heaviness of the model we have trained the model for 300 epochs due to which the performance of the model is compromised

## Suggested Solutions

* Training of the deep learning model which requires a large dataset remains a challenge when training with a small dataset hence accuracy is compromised.
* The age-progressed images get blurred which eventually affects the performance of the model.
* The performance of the face recognition model heavily depends on the accuracy of age-progressed images which are obtained.
* Due to the long duration of code execution and heaviness of the model we have trained the model for 300 epochs due to which the performance of the model is compromised

# Conclusion

In this work, a system is proposed for an age-invariant face recognition system for finding missing children. This system reduces the manual work of scanning through all the photographs of the lost children to find a match. We have tried to solve the problem of age variation by age progressing the images of the child using the Aging cGAN deep learning algorithm. The model improves the performance of the FaceNet model which we have used for face recognition from 44.76% to 47.32%. These results suggest that age progressing the faces of the children before face recognition enhances the ability of the system to identify the lost children who are possible victims of child abduction and trafficking. This will be beneficial for solving cases of untraced missing children and thus will help to reduce the increasing number of missing child cases.

##### References

1. S. Chandran, Pournami Balakrishnan, Byju Rajasekharan, Deepak N Nishakumari, K Devanand, P M Sasi, P. (2018). “Missing Child Identification System Using Deep Learning and Multiclass SVM”.
2. Xin Jin, Shiming Ge, Chenggen Song, Xiaodong Li, Jicheng Lei, Chuanqiang Wu, and Haoyang Yu “Double-Blinded Finder: A Two-SidePrivacy-Preserving Approach for Finding Missing Children”(2020).
3. Shun Lei Myat Oo, Aung Nway Oo University of Information Technology, Yangon, Myanmar “ Child Face Recognition with Deep Learning”(2019).
4. Chang Shu School of Communication and Information Engineering, University of Electronic Science and Technology of China Chengdu, China “Optimizing deep neural network structure for face recognition” (2017).
5. Grigory Antipov, Moez Baccouche, Jean-Luc Dugelay ”Face Aging with Conditional Generative Adversarial Networks” (2017) arXiv:1702.01983v2