ChildGAN: Face Aging and Rejuvenation to Find Missing Children

Rohit Das
Computer Science and Information Engineering
National Taiwan Normal University
Taipei City, Taiwan
61047086s@ntnu.edu.tw

黃意婷
Computer Science and Information Engineering
National Taiwan Normal University
Taipei City, Taiwan
61047024S@ntnu.edu.tw

Judith Lin 林于立
Computer Science and Information Engineering
National Taiwan Normal University
Taipei City, Taiwan
judithlin@csie.ntnu.edu.tw

***Abstract*—Child-face aging and rejuvenation have amassed considerable active research interest.All current face aging approaches based on generative adversarial networks (GANs) focus on adult images or long-term aging. We take this opportunity to introduce ChildGAN(1) that generates visually realistic images for enhanced face-identification accuracy while preserving the identity. This is the final project for Artificial Neural Network course.**

***Index Terms*—GAN, ChildGAN, Face Recognition, Age Estimation, Gender Preservation**

I. Introduction

According to a global report on human trafficking, most trafficking victims are women and children.The news said that approximately 6000 children disappear in the Taiwan every year. A large number of child trafficking victims are very young.As time passes, their faces may appear to be quite different from the photos provided by their parents or relatives, and many times the face images are obstructed or blurry.The face transfiguration makes it difficult to find the victims. This work can help law-enforcement agencies to trace missing children.In this paper, the proposed model is based on a variational auto-encoder (VAE) with a generative adversarial network(GAN), which can simultaneously age and rejuvenate the face in the image space.

II. Architecture

Like any GAN model, we have 3 parts. Encoder, Generator and Discriminator. Fig. 1 shows the architecture of the model.

Encoder: This will downsample the image to 32,768 feature vectors. At the end of the node, we have Self Attention Block which basically helps in maintaining and keeping the utmost details like ear,nose,mouth which might change when age progresses.

Generator: Upsamples the feature vector and synthesizes new face features. Discriminator: Tries to recognize the image generated and compare with the input image. The main goal is face recognition and age progression. Overall the model is learning 3 aspects.

1) Synthesize face (Encoder/Decoder)
2) Face Features Recognition (Discriminator)
3) Age Based Feature Generation + Recognition (Encoder/Decoder/Discriminator)

III. Loss Function

Like every machine learning model, we have loss function. The loss functions mentioned here are quite simple and straightforward.

Adversarial Loss: Double Bar E is the Euclidean space, x is input data, c is concatenated conditional vector of age(L) and gender(G) Lambda gp is the penalty co-efficient to penalize the gradients phi.The generator G and discriminator AD img stimulates generated results to be indistinguishable from real images.

$$\begin{aligned} L_{AD_{\text{img}}} =& -\mathbb{E}_{x,c\sim P_{\text{data}}(x,c)}[AD_{\text{img}}(x,c)] \\ &+ \mathbb{E}_{x,c\sim P_{\text{data}}(x,c)}[(AD_{\text{img}}(G(E(x),c)))] \\ &+ \lambda_{gp}\mathbb{E}_{\hat{x}\sim P_{\hat{x}}}[(|\left|\bigtriangledown_{\hat{x}AD_{\text{img}}}(\hat{x})\right||_2-1)^2] \end{aligned} \tag{1}$$

Perceptual Loss: The deep-feature consistency principle is applied to ensure that VAE, with adversarial learning (output) and its corresponding input images, has a similar deep feature with a much clearer and more natural nose, eyes, teeth, and hair texture, with much fewer artifacts, even on obstructed and low-resolution faces. Channel(C),Width(W),Height(H) dimension of input image,phi is the feature map.

$$L_\Phi = \frac{1}{C_\gamma W\gamma H\gamma}\|\Phi_\gamma(x)-\Phi_\gamma(\hat{x})\|_2^2 \tag{2}$$

Kullbck-Leibler Divergence Loss: One of the prime loss function in VAE(Variational Auto Encoder). The function

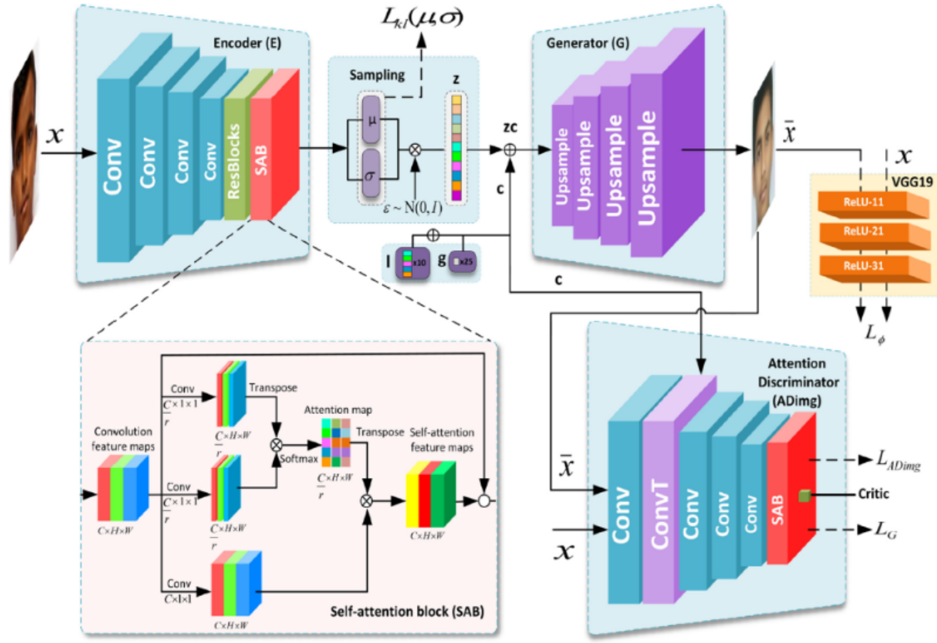

Fig. 1. ChildGAN(1) Model Architecture

regularizes Z vector(32,786 features vector) using reparameterization trick after normalization. For input image x the E network output the mean and covariance.

$$L_{kl}(\mu, \sigma) = -\frac{1}{2}\sum_k (1 + \log(\sigma_k) - \mu_k^2 - \sigma_k) \tag{3}$$

Total Variational Loss: This function ensures measurable continuity and smoothness in the generated image to avoid noisy and over pixelated results.This is basically Sum of absolute difference for adjacent pixel values in the generated image.

$$L_{tv} = \sum_{c=1}^{C}\sum_{w=1}^{W}\sum_{h=1}^{H} |(\bar{x})_{w+1,h} - (\bar{x})_{w,h}|^2 + |(\bar{x})_{w,h+1} - (\bar{x})_{w,h})|^2 \tag{4}$$

Full Objective Function: Matching up all the equations we get the Full Objective Function. The goal of ChildGAN is to minimize the objective function for the discriminator and Encoder as well as Generator. The L represents the set of 1-Lipschitz constraint.

$$\min_{\|AD_{\text{img}}\|_L \leq 1} \mathcal{L}_{AD_{\text{img}}} = \lambda_{adv1}\mathcal{L}_{AD_{\text{img}}} \tag{5}$$

$$\min_{E,G} \mathcal{L}_{EG} = \lambda_p \mathcal{L}_\Phi + \lambda_{kl}\mathcal{L}_{kl}(\mu, \sigma) + \lambda_{adv2}\mathcal{L}_G + \lambda_{tv}\mathcal{L}_{tv} \tag{6}$$

IV. Dataset

Since the aim is to generate a series of photos to simulate how a child looks like when they grow up, it is crucial that the dataset includes photos of people of different races and in different ages. While most child abductions happen pre-pubescent, the dataset should also have sufficient photos of children. Therefore, the author of this paper created a dateset called Multi-Racial Child Dataset (MRCD).

MRCD dataset contains 64,965 face images of four races (East-Asians, Africans, Caucasians, and Indians) and five age groups ([0-3], [4-5], [9-12], [13-16], and [17-20]). The majority of the images are selected from publicly available datasets. In order to keep the image samples evenly distributed among the four races and the five age groups, images of under represented races and age groups are selected into MRCD through web-crawling. Table I shows statistics of image samples in MRCD dataset and Fig. 2 is some example of face images in MRCD dataset.

TABLE I
Statistics of in-the-wild multi-racial child faces in MRCD dataset

Dataset	Race				Total
	Asian	African	Caucasian	Indian	Images
CACD	-	-	6,137	-	6,137
MorphII	-	3,330	-	-	3,330
FFHQ	-	3,996	1,322	7,934	13,252
UTKFace	1,197	410	2,251	716	4,574
IMDB	-	-	1,482	-	1,482
MegaAsian	6,641	-	-	-	6,641
ICD	-	-	-	13,561	13,561
Web crawling	5,557	8,292	1,493	826	16,168
Total	17,211	13,354	19,297	15,103	64,965

As shown in Table II, training and testing datasets are then selected from MRCD dataset, making sure that all the training images and testing images are evenly distributed in

(a) Asian (b) African (c) Caucasian (d) Indian

Fig. 2. Exemplar blurred face images of MRCD dataset.

different races. There are sufficient training result for each race; therefore, we can ensure the model would not be bias against any of the races.

TABLE II
TRAINING SET AND TESTING SET STATISTICS

Dataset	Asian	African	Caucasian	Indian	Total
Training set	16,350	12,686	18,332	14,348	61,716
Testing set	861	668	965	755	3,249
Total	17,211	13,354	19,297	15,103	64,965

V. EVALUATION

Because the main goal of ChildGAN is to help find missing children, the evaluation of the model should take into account whatever may prevent parents or law-enforcement from finding the missing children. One issue of finding missing children is that some missing children's family have difficulty providing high resolution photos; another issue is that some children might wear glasses which obstruct their faces. Therefore, while evaluating the model, we focus on the following two aspects:

1) The performance of the model on low resolution face images.
2) The performance of the model on face images with obstructions.

We compare this model with other face aging models (AIM, CAPAVAE, CAAE, AcGAN, IPCGAN, and Ablation Study) under the above mentioned aspects. Other than that, we also put this model into use and generate a few simulations of face images from missing children in Taiwan.

Fig. 3 shows the comparison of performance across different models. The leftmost column is the photo taken before the child is missing; the rightmost column is the photo taken after the same child is reunited with the family. The columns in between shows the age-progressed and age-regressed images. From observation, we can see that the age change is not visible in AIM and AcGAN models. Also, some of the model does not preserved the skin-tone, making the simulated photos unrealistic.

Fig. 4 shows the comparison of performance accross different models. The leftmost column is the photo taken before the child is missing; the rightmost column is the photo taken after the same child is reunited with the family. The columns in between shows the age-progressed and age-regressed images. From observation, we can see that AIM removes the sunglasses. CPAVAE, CAAE, Ablation Study and ChildGAN remove the tinted glasses but keep the glasses characteristics while AcGAN and IPCGAN keep the sunglasses characteristics.

It is worth noticing that because the photo taken before the child was missing does not show any glasses, none of the models generate simulated age-progressed face images with glasses. However, some children's eyesight deteriorates as they get older. The absent of the glasses might cause issues in recognizing the missing children.

In order to test how well this model works on Taiwanese missing children, we first randomly selected eight children from the Missing Children Data Resource Center website [2] and downloaded their photos. The photos are cropped into square images with faces in the center and fed into the model. The generalization result is shown in Fig. 5.

After observing the result, we found the though generalized images are recognizably face images, each of them shows glitches. Some of the images are even blurry. Nevertheless, we can see visible differences between different age groups, which means age-progressing and age-regressing processes have taken effect.

VI. CONCLUSION

This project demonstrate the ability of ChildGAN model (along with MRCD dataset) to adapt face images of a wide range of races and preserve the distinctive characteristics of each race. Though some generalizing results still show glitches, the result is quite promising. ChildGAN model and benchmark datasets should drive face-aging and cross-age face recognition research toward finding missing children, as well as numerous real-world applications, e.g., border control, passport verification, access control, forensic science, and underage driver's licenses.

During the training process, we also got some failed experiment result. As shown in Fig. 6, the face images generated are severely distorted, and especially so in age groups above 9 years old.

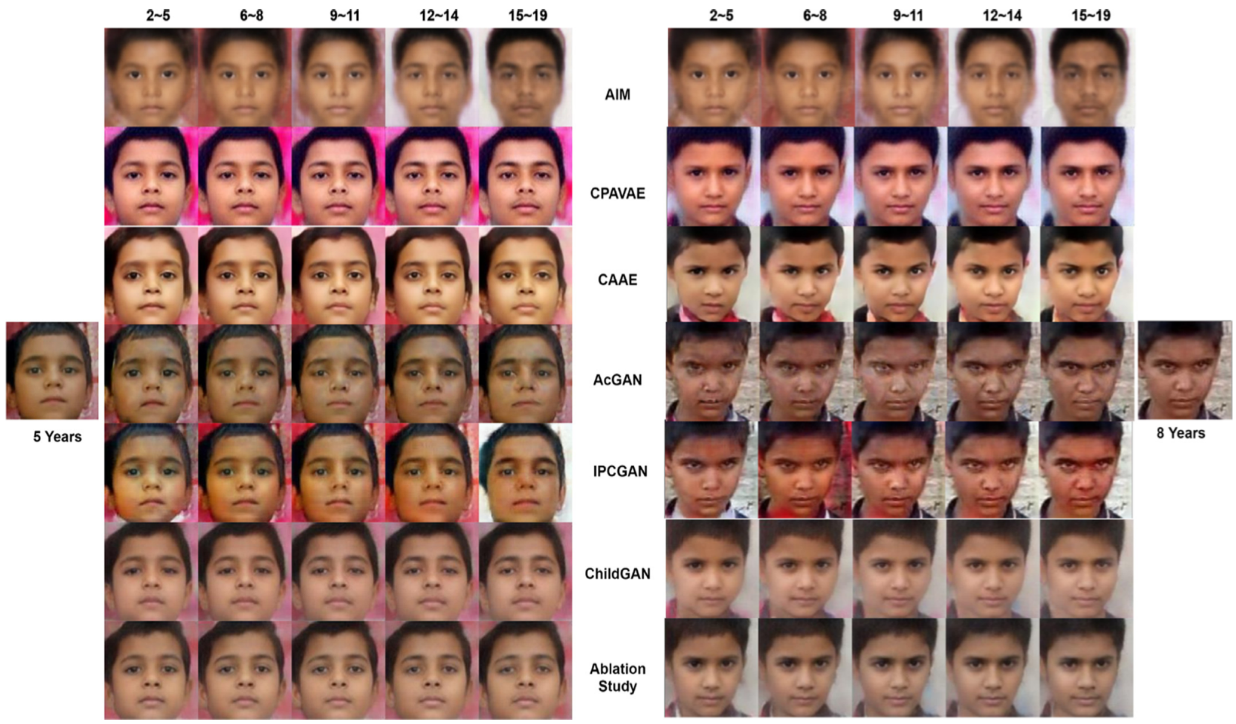

Fig. 3. Comparison of performance across different models, of simulating face images from low resolution images.

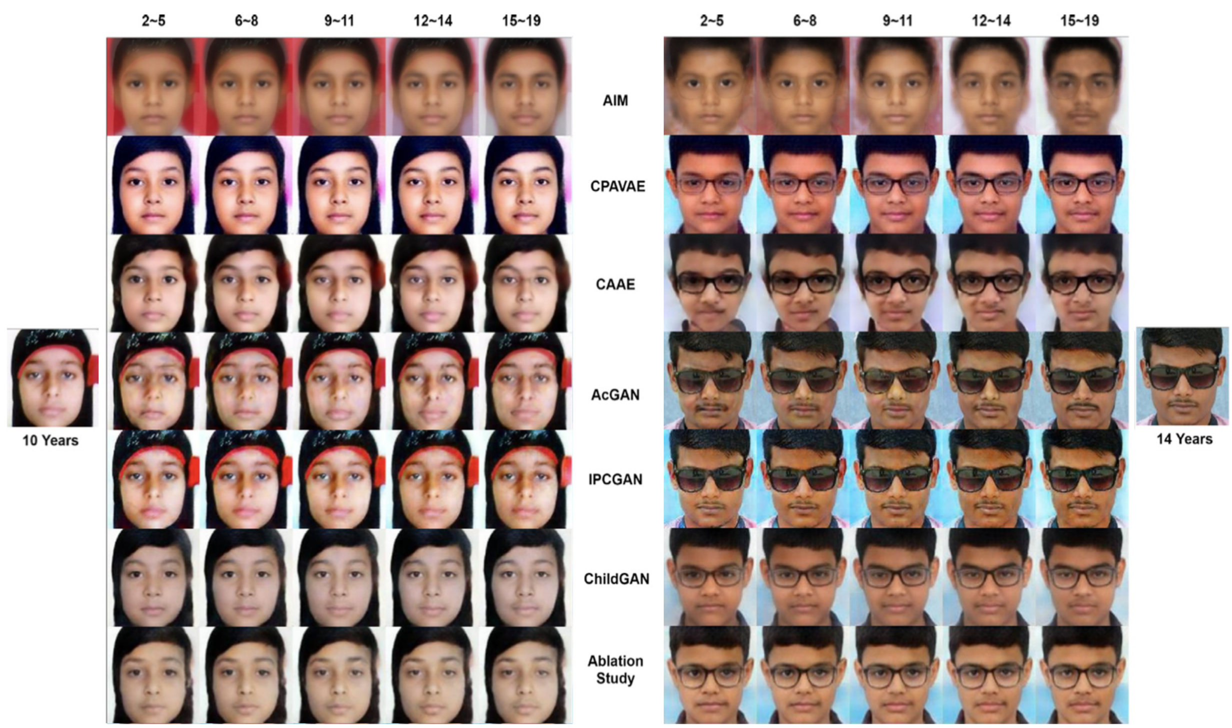

Fig. 4. Comparison of performance across different models, of simulating face images from obstructed face images.

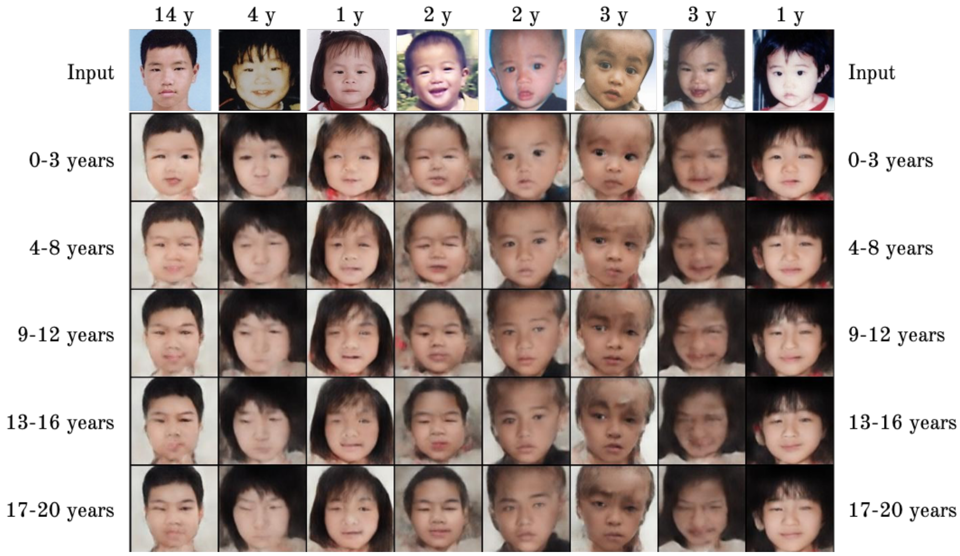

Fig. 5. Applying ChildGAN on Taiwanese missing children

Fig. 6. Example of failed experiments.

ACKNOWLEDGMENT

This experiment and its all relevant data and information is taken and motivated from the paper(1) ”Praveen Kumar Chandaliya, Neeta Nain, ChildGAN: Face aging and rejuvenation to find missing children, Pattern Recognition, Volume 129, 2022, 108761, ISSN 0031-3203”

REFERENCES

[1] Praveen Kumar Chandaliya, Neeta Nain, ChildGAN: Face aging and rejuvenation to find missing children, Pattern Recognition, Volume 129, 2022, 108761, ISSN 0031-3203

[2] Missing Children Data Resource Center: https://www.missingkids.org.tw