Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IMPACT FACTOR: 7.056

IJCSMC, Vol. 9, Issue. 9, September 2020, pg.22 – 26

Age Unconstrained Photo ID Validation Using Contextual Generative Adversarial Networks

Nikhil Singh Rathaur[1]; Shubham[2]

[1]Computer Science & Engineering - National Institute of Engineering Mysore, India

[2]Computer Science & Engineering - National Institute of Engineering Mysore, India

[1] nikhilrathaur1998@yahoo.com; [2] shubham7070078010@gmail.com

DOI: 10.47760/IJCSMC.2020.v09i09.002

Abstract— We propose a solution to handle this problem. We are extending upon the paper where Contextual Generative Adversarial Networks are used to generate a band of aged images from one single image. The images of different age groups will be generated from one single picture but with the same context. We are using the resulting images for face validation as a photo verification system. Our algorithm will use the band of generated images of different age groups and will match it with the current face of the person. If the match score goes beyond a limit, the photo identity will be validated, or else a warning will be triggered. This system will play an important role in places where there's a need for higher security and better ways of validation.

Keywords— Generative Adversarial Networks, Face Verification, Neural Networks, Face Aging, Biometric

I. Introduction

Generating some kind of data of the same probability distribution as that of input data is what stands for Generative Adversarial Networks[1] in precise terms. We have used a GAN variant called Contextual Generative Adversarial Network[2] for face aging prediction. We take an input image and generate images of 7 different age groups from that particular image. We then try to find the best match between the input image and the 7 generated images. If this match score passes a certain minimum score, the ID proof holding the input image is validated, else it isn't. We have generated match scores in two levels — pixel level and feature level. For pixel level matching, we have used the concepts of Mean Square Error (MSE) and Structural Similarity Index (SSIM). At feature level, we have used the concept of Scale Invariant Feature Transformation (SIFT)[3] . The detailed pipeline of this process is given in Fig.1. We shall now explain each component individually.

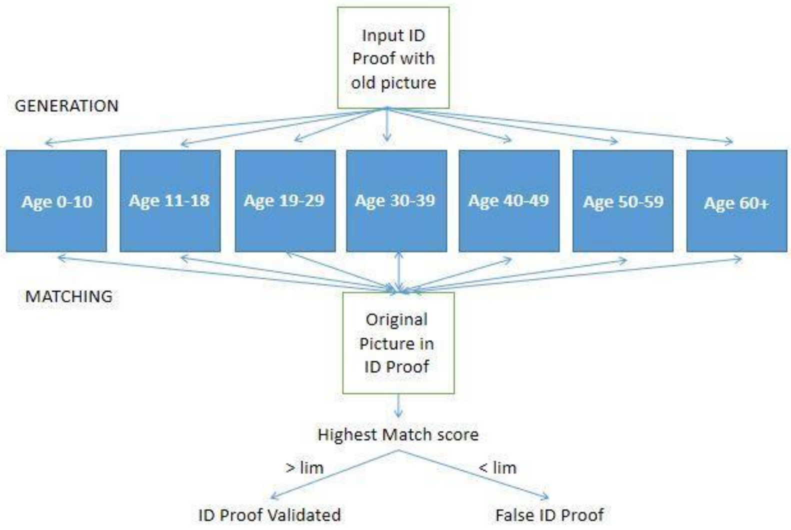

Fig. 1 Complete pipeline of system flow. Here 'lim' is the min score above which image is validated

II. CONTEXTUAL GENERATIVE ADVERSARIAL NETWORK

The Contextual GAN was an improvement upon the old Conditional GAN[4] . According to Conditional GAN, aging over a face picture is performed keeping 2 constraints. First is identity consistency and second is the appearance changes. The constraint of identity preservation is important keeping in mind the cross-age verification which is similar to the process that we are going to perform. Other is appearance changes which includes skin texture changes and shape alterations. Here output is given with the same context as the original picture. Contextual GAN improved over this by taking the context from your current picture and feature processing from the older picture. It means that after the aging is performed over the face, the picture context (including the background, face orientation and emotion) is taken from your current live picture. Contextual GAN is specifically incredibly apt in our case because while performing the pixel level subtractions and matching using MSE and SSIM, the context pixels will have no effect or disturbance.

III. ARCHITECTURAL BRIEF

The primary aging network architecture of Contextual GAN is the same as the Conditional GAN. There are three primary networks in this architecture — Conditional Transformation Network, Age Discriminative Network and Transition Pattern Discriminative Network. Let's call them CTN, ADN and TPDN respectively. Here CTN is the generator, and ADN & TPDN are two discriminators. An image with an age label is given as input. The probability distribution of this data is recorded and the CTN generates image-age pairs. The ADN tries to tell if the generated image is real or fake and keeps doing this until it gets fooled by the image produced by CTN. The TPDN investigates and tries to spot if the age label generated goes along with the image generated. This loop of distinction continues until TPDN also gets fooled by the CTN image-age pair. Once the pair passes through all these networks cleanly, we get an output image-age pair that is as good as the original image-age pair. The complete architectural diagram is given in Fig.2. The mathematics behind this architecture is stated in the original paper of Contextual GANs.

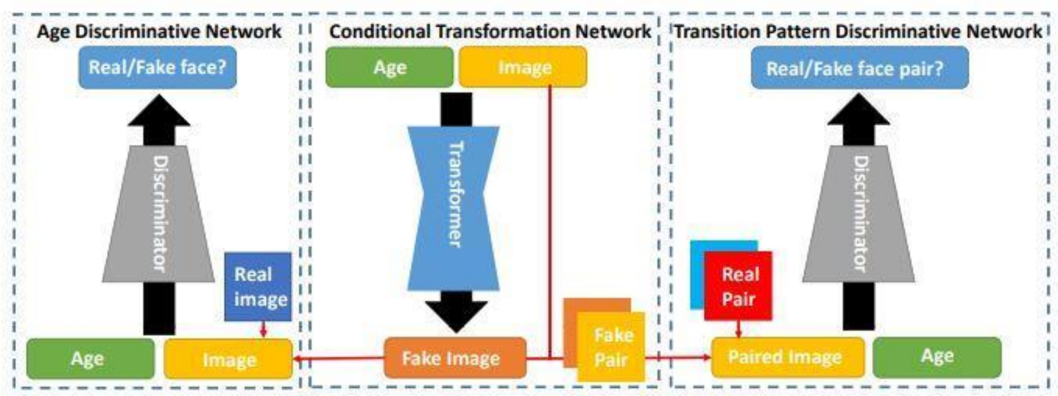

Fig.2 Complete architecture of C-GAN Network

IV. EXPERIMENTAL IMPLEMENTATION

The implementation of this C-GAN is done over the training data taken from multiple datasets including CACD, LFW and Morph. These datasets contain sequential data which gave us a total of 4047 images with equal biological distribution which includes the gender and the age. For equal age distribution and due to lack of senior citizen image dataset, we used non sequential data from wikihow dataset too which had a total of 15030 images with equal biological distribution. We now divided all these images into 7 different age categories. These categories are : 0-10, 11-18, 19-29, 29-39, 39-49, 49-59 and 60+ years.

After the data preparation, the training of this data is done. For training purposes, we found the most efficient learning rate to be 0.00015. The code used for this training is based on publicly available torch based code associated with the project on Deep Convolutional Generative Adversarial Networks[5] . The code parameters include $\beta 1$ as 0.55 and batch size as 24. The optimization method chosen is Adam method with 120 epochs. The training is done over a high performance NVIDIA-DGX GPU system.

V. RESULTS

After training the network, we obtained a band of images of 7 different age groups as depicted in the two sample outputs of the Contextual GAN model in Fig.3. One constraint that should be kept while filtering the training data is to keep an orientation constraint. The faces should be front facing and not side facing for better results.

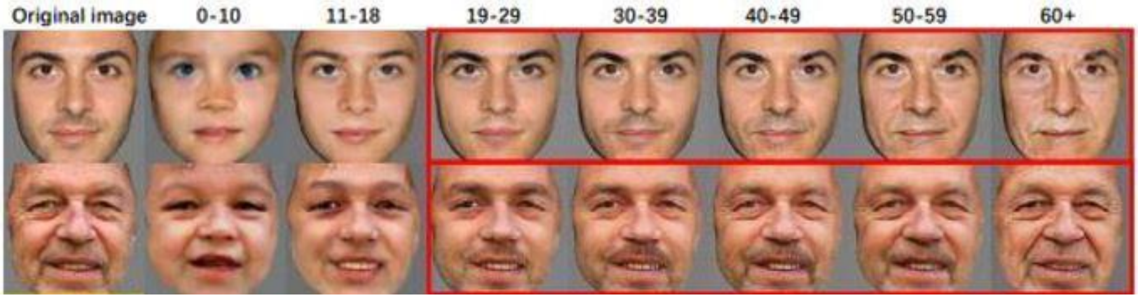

Fig.3 Output age bands from Contextual GAN

VI. AGE UNCONSTRAINED FACE VALIDATION

A. Pixel Level Validation

Now that we have age band images, we can begin the validation. First we will do the pixel level validation through Mean Square Error (MSE) scoring and Structural Similarity Index (SSIM)[6] scoring. Mean square error works with simple subtraction of pixel values (0-255). We have converted the images into a threshold of green blue for better scaling by keeping the pixel values in a short range to avoid high errors at minimum variation. A sample match comparison is done by taking the original image being matched against the generated image of 2 different age groups. More the MSE, higher the irrelevance. Similarly, we generated an SSIM score based on structural similarity i.e. location of pixels in the 2 dimensional space using the same threshold reduced images. The Highest possible SSIM value is 1. These outputs results are shown in Fig.4. We can observe clearly that images that are more similar to each other have a higher SSIM and lower MSE. For validation, a limit can be put accordingly on how strict relevance is being required to clear the validation test.

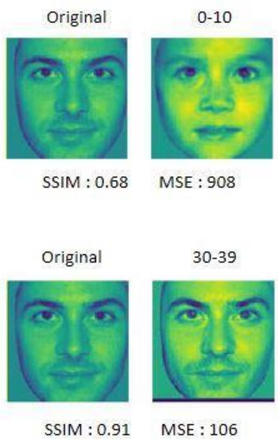

Fig.4 MSE and SSIM match score between original and generated images

B. Feature Level Validation

Another way for photo validation is feature matching technique called Scale Invariant Feature Transformation (SIFT). This method is the state-of-the-art image matching technique and is considered to be more reliable and crisp. N number of key features are identified in the input images and these features are first counted and then common features are matched and scored. A demonstration for this match over the same images is shown in Fig.5. First the key points in both the images are counted, then the features that match are counted and finally the match score of these matching features is generated. It is easily visible that 2 images that are more similar than the other pair of images have more number of matching features and hence a higher match score. Here also a minimum match limit can be set according to the strictness of the validation process.

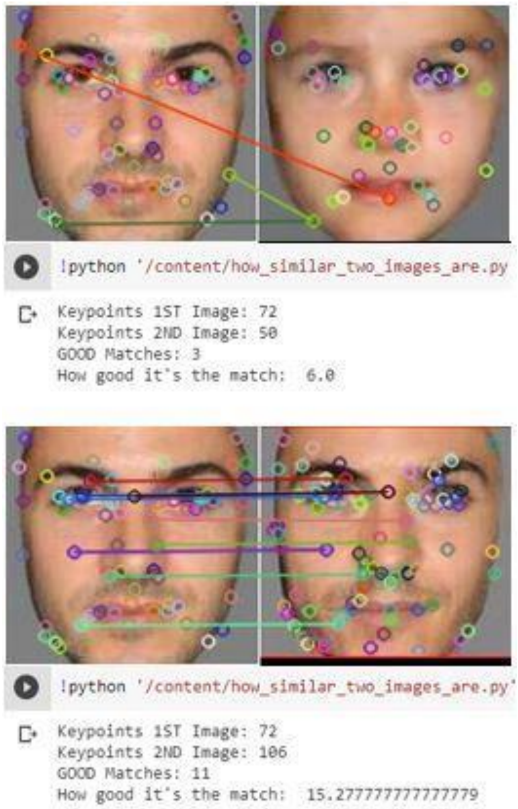

Fig.5 Identification and key point match score generation between original and generated images

VI. CONCLUSIONS

This system is a blend of several state-of-the-art techniques including Contextual GAN and match techniques like SIFT. This system solves several issues that usually occur during face verification and validation. Future improvements upon these results are possible with implementation of unconstrained orientation of faces in this system. This system is able to administer the issues in the current face recognition systems and can be used as an extension upon them for better administration.

ACKNOWLEDGEMENT

This research would not have been possible by the immense support of IIT Jodhpur as they provided us with all the resources including high performing GPUs. Big thanks to NIE Mysore Computer Science department as well for providing us the right directions during the research.

REFERENCES

[1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde Farley, Sherjil Ozair, Aaron Courville, & Yoshua Bengio. (2014). Generative Adversarial Networks.

[2] Si Liu, Yao Sun, Defa Zhu, Renda Bao, Wei Wang, Xiangbo Shu, & Shuicheng Yan. (2018). Face Aging with Contextual Generative Adversarial Nets.

[3] Li, X., Luo, J., Duan, C., Yin, P., & Zhi, Y. (2019). Portrait matching based on sift features *IOP Conference Series: Materials Science and Engineering*, *612*, 032162.

[4] Grigory Antipov, Moez Baccouche, & Jean-Luc Dugelay. (2017). Face Aging With Conditional Generative Adversarial Networks.
[5] Horé, Alain & Ziou, Djemel. (2010). Image quality metrics: PSNR vs. SSIM. 2366-2369.10.1109/ICPR.2010.579.
[6] Alec Radford, Luke Metz, and Soumith Chintala. 2015. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *Computer Science* (2015).