Journal of Theoretical and Applied Information Technology
30th June 2023. Vol.101. No 12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

LOW RESOLUTION FACE RECOGNITION USING COMBINATION OF GPEN SUPER RESOLUTION AND FACENET

KHAIRUL SHALEH ZEBUA[1], IMAN HERWIDIANA KARTOWISASTRO[2,3], GEDE PUTRA KUSUMA[4]

[1,4]Computer Science Department, BINUS Graduate Program – Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia, 11480

[2]Computer Science Department, BINUS Graduate Program - Doctor of Computer Science, Bina Nusantara University, Jakarta, Indonesia, 11480

[3]Computer Engineering Department, Faculty of Engineering, Bina Nusantara University, Jakarta, Indonesia, 11480

E-mail: [1]khairul.zebua@binus.ac.id, [2,3]imanhk@binus.edu, [4]inegara@binus.edu

ABSTRACT

Face Recognition has been one of the most active research areas in computer vision nowadays. It has widely applied in various applications in real human activity and has reached super good performance in term of accuracy. Deep learning approach significantly improves the accuracy. But it's mostly evaluated in High-Resolution (HR) images. Nevertheless, the Low-Resolution Face Recognition (LRFR) remains a challenging part. Recognizing face of images in low-resolution (LR) scenarios could possibly decrease the accuracy of predictions due to that LR images usually lack discriminative details. One of the solutions commonly used to improve accuracy by applying concept of Super Resolution (SR) to produce the better quality of image from LR to HR. This research aimed to solve the issue in LRFR problem to get the better accuracy by using combination of SR method with GPEN and FR with FaceNet. The model is evaluated on labelled faces in the wild (LFW) dataset. LFW dataset is filtered to person that has at least five images of faces per label so that found 423 classes. The filtered LFW dataset with 423 classes then augmented with horizontal flip to generate more dataset as input to FaceNet model. Accuracy of training, testing and validation are captured and presented at this report. The results are compared with bicubic interpolation method with data augmentation and without data augmentation. Based on that, with data augmentation the accuracy is getting better. In result, this combination method can produce training accuracy score of 82.8%, validation accuracy score of 66,6% and testing accuracy score of 69% with data augmentation.

Keywords: *Super Resolution; Face Recognition; Generative Adversarial Network; GPEN Model; FaceNet Model*

1. INTRODUCTION

A lot of research has been done over the past few years to improve accuracy of Face Recognition (FR). FR algorithms have reached near perfect with mostly over 99% accuracy achieved so far [1] evaluated in LFW dataset [2]. Deep convolutional neural network model contributes to the super good performance face recognition recently [3] [4] [5]. The face recognition problem such as handling variations in pose, age, illumination, expression and heterogenous face matching are mostly solved with new approach models [6]. Nevertheless, the good performance accuracy face recognitions are mostly evaluated in frontal face and with High Resolution (HR) images. The fact, HR images can be easily identified but it would be different for Low Resolution (LR) image, the accuracy in Low Resolution Face Recognition (LRFR) would significantly drop [7]. LR mode in FR task remains a challenging part to be solved [8] [9] [10]. There are always possibility images to produce low-resolution. Lack of angle, distances of object so that produce low-quality image [11].

The simple basic idea how to resolve the LRFR problem is by doing image super-resolution (SR) to generate better quality image [12] from the low-quality image. Several previous works methods have been developed to address this idea like Bicubic Interpolation [19], Convolutional Neural Network (CNN) [13], Generative Adversarial Network (GAN) [14]. But GAN has been the popular one in bringing the concept of image SR method [15] compared with other previous method [16]. This work is based on prior work publish in [17] for FR task by combining with SR method published in [18] and compares both previous SR using the bicubic interpolation with GAN Prior Embedded Network (GPEN), which is evaluated on LFW dataset.

The dataset is divided into three categories of data: training, validation, and testing. By filtering LFW dataset with person has at least five images of face, so there are 423 classes in classification. More data to be fed should increase the accuracy of prediction. The data is augmented with flip horizonal method to generate more training data. All Dataset is detected and cropped by using MTCNN method to eliminate other attributes at images out of face. The output of this cropped images is presented with 64x64 pixel as High-Resolution (HR) Image. Then, to generate the Low-Resolution, the HR images are down scaled to 25% of previous sizing to be 16x16 pixel as known Low-Resolution (LR) Image.

To ensure the proposed method is better than previous work, the LR image is upscaled with bicubic interpolation as well for comparison. Bicubic interpolation as simple way to upscale the image with lower computation but still left blurry. Is the proposed method better and worth to be implemented? Is the augmentation data can improve the accuracy? To answer that the upscaling outputs of SR are compared in accuracy in FaceNet. The proposed method by combining SR method using GPEN and FaceNet is expected to produce better in accuracy.

2. RELATED WORKS

Super Resolution. The previous SR task methods using bicubic interpolation [35] can generate the higher resolution image in minimum computation. But blurring is still a serious problem in quality of output images [19]. This challenge boosts enthusiasm of researchers to propose new way how to solve the issue. Singe Image Super-Resolution using a Generative Adversarial Network (SRGAN) method was proposed in 2017 by employing a deep residual network (ResNet) with skip connection and diverge from MSE [39]. SRGAN improves the quality of images. But in next year, the enhancement for SRGAN was proposed [37]. Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) improve the quality of images through three key components. First, in network architecture ESRGAN introduced the Residual-in-Residual Dense Block (RRDB) without batch normalization. Second, by idea from relativistic GAN, discriminator predicts relative realness rather than the absolute value. And third, improving the perceptual loss with utilizing features before activation, so that produce stronger provision for brightness consistency and texture recovery. GAN can produce better results in super-resolution concept, but not often generate over-smoothed output. GAN Prior Embedded Network (GPEN) proposed a method with first learning a GAN for high-quality image to be embedded into a U-shaped DNN and do fine-tuning the GAN prior embedded DNN. GPEN can improve the quality of output images also can minimize the over-restoration process.

Face Recognition. As growth of CNN methods, it performs well in FR task in HR images [20] [21] [22], since then, the CNN methods are developed widely to solve the FR problems. Recent FR methods [17] [30] [33] [34] has produced accuracy over 99% evaluated on LFW dataset.

Low Resolution Face Recognition. In 2018 LRFR task using CNN [8] was introduced to solve the face recognition in low-resolution by using Deep Coupled ResNet (DCR) that can produce 96.6% in accuracy at 16 x 16 pixel. Another research in 2019, proposed architecture of two deep convolutional neural networks with combination of Super-Resolution CNN and Face Recognition [36]. The summary evaluation results of related works are shown in Table 1.

Table 1: Summary of Related Works

Publication	Method	Problem	Evaluation Result
[39]	Super-Resolution GAN	Super Resolution	Set5 (PSNR:29.40, SSIM: 0.8472, MOS 3.58) Set 14 (PSNR:26.02, SSIM: 0.7397, MOS: 3.72)
[37]	Enhanced Super-Resolution GAN	Super Resolution	Set14 (PSNR: 20.35, Percpetual Index: 1.98) Set14 (PSNR: 30.50, Percpetual Index: 3.64)
[38]	Attribute Embedding Super Resolution	Super Resolution	PSNR: 21.82, SSIM: 0.62
[18]	GAN Prior Embedded Networks	Super Resolution	PSNR: 20.80, Fid: 31.72, LPIPS: 0.346
[17]	FaceNet	Face Recognition	LFW: 99.63% YTF: 95.12%
[30]	ArcFace	Face Recognition	LFW: 99.83% YTF: 98.02%
[33]	SFace	Face Recognition	LFW: 99.60%
[34]	AdaCos	Face Recognition	LFW: 99.73%
[8]	CNN with Deep Couple ResNet	Low Resolution Face Recognition	LFW 8x8: 93.6%, LFW 12x12: 95.3%, LFW 16x16: 96.6%, LFW 20x20: 97.3% LFW 112x96:98.7% SCface (d1: 73.7%, d2: 93.5%, d3: 98.0%)
[36]	Two Branch CNN	Low Resolution Face Recognition	Feret 6x6: 81.4%, Feret 12x12: 92.1%, Feret 24x24: 96.7% LFW 8x8: 76.3% MBGC 12x12: 68.64%

3. THEORY AND METHODS

3.1. Low Resolution

As low-resolution is still one of the main issues in face recognition problems, a lot of researches trying to propose new model to solve it. LR images has lack discriminative details that not easy to be recognized even in human eyes. The image lower than 32 x 32 pixel is classified as low-resolution image [23]. This research uses 16 x 16 pixel image as LR input for SR task with 4x upscaling size to 64 x 64 pixel output as HR image to be fed to FR system using FaceNet.

3.2. MTCNN

Mutli-Task Cascaded Convolutional Neural Networks (MTCNN) is a method used to detect faces and facial landmarks on images [24]. MTCNN method produces the precision detection face in unconstrained environment [25]. MTCNN task purpose to ensure the background image and other attributes exists at image would not affect the face while being recognized.

The architecture of MTCNN pipeline has three stages of CNN. First stage is called Proposal Network (P-Net) to get the candidate windows in a box vector. Then estimated bounding box vector used to calibrate the candidates. Next the non-maximum suppression (NMS) employed to merge highly overlapped candidates. The second stage called Refine Network (R-Net) which all candidates are fed to another CNN and rejects many false candidates. Calibrate with bounding box regression and merge the NMS candidate. And the third stage called O-Net to describe face in more detail that output five facial landmark's position. This research utilizes MTCNN to detect face area and crop it.

3.3. GAN Prior Embedded Network Super Resolution

Super resolution (SR) is used to increase the quality image from LR image to HR image. The concept of SR method is to generate the new size of image without losing details pixel in image. SR task is not just to scaleup the size of image. Moreover, to find the missing pixel value so that improve the quality.

GAN contributes to improve quality of images in SR concept. GAN initially introduced in the publication on 2014 [26]. The basic concept of GAN consists of 2 neural networks: Generator (G) and Discriminator (D). G network generates fake image while D network for learning purpose for image and compete each other. D keeps learning on deciding is the input image real or fake.

GPEN Super Resolution (GPEN SR) with high-quality face image learning and embed a GAN prior network into a U-shaped DNN with set of low-quality face images. The design of GAN blocks is easily to be used for fine-tuning further [18]. GPEN SR is inspired from StyleGAN [31] [32], uses mapping network to project latent code z into a less entangled space $w \in W$. The intermediate code w then broadcasted to each GAN block. Total numbers of GAN blocks at GPEN is equal to the number of skipped feature maps extracted in the U-shaped DNN and this is related to the resolution of input face image. GPEN SR model is not fully convolutional, the low-quality image is recommended to be resized first to desired resolution using simple bilinear interpolar before inputted to GPEN. But at this experiment, the LR images are directly as input for GPEN without resizing to desired resolution HR image.

3.4. FaceNet

FaceNet is a method of deep neural network to extract features from the image of face. FaceNet was published in 2015 by researcher at Google [17] and at that time it has reached a new record accuracy with 99.63% evaluated on LFW dataset [30] and 95,12% accuracy on YouTube Faces DB (YTF) dataset. FaceNet takes an image of the person face as input and embed all the important information into the 128-element of vector. The vectors extracted from same person's images would be very close to each other, while distance between vectors extracted from the different person's images would be much more. FaceNet trains triple images named with Anchor, Positive and Negative. The Anchor and Positive images are expected to the same person while the Negative from different person. This well-trained CNN produces unified embedding for each given face and this way to ensure that not any two different faces have embedding closer than alpha. Furthermore, Inference task is performed to get the best match to the given face. Triplet loss concept learning is illustrated at Figure 1.

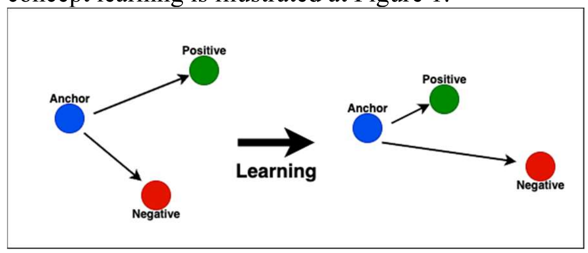

Figure 1: Triplet loss reduce the distance between anchor and positive and increase the distance between anchor and negative [17]

4. METHODS

4.1. Dataset

This research is evaluated on LFW dataset (*Labelled Faces in the Wild*) [30]. The LFW dataset is commonly used to evaluate the face recognition tasks. It's a database of face photographs collected from public web, contains 13,233 in total of images from 5,749 person. There are 1680 of the people pictured have two or more distinct photos in the dataset. There are now four different sets of LFW images including the original and three different types of "aligned" images. This research takes the original as dataset.

4.2. Pre-processing

The LFW dataset is filtered to select person that has at least five images faces. Based on that, there are 423 people in total, so there are 423 classes at classification.

The dataset is split into three categories: 60% as training dataset, 20% as validation dataset and the rest 20% as testing dataset. As Comparison further, the train dataset is augmented using flip horizontal to generate more training dataset by twice. Output of flip horizontal function is shown at Figure 2.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

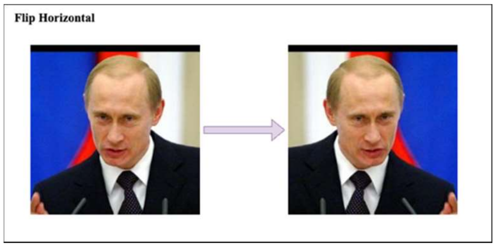

Figure 2: The horizontal flip original image

Every single image from both of sources then detected and cropped by using MTCNN method and resize the face to HR size with 64 x 64 pixel. They are used as the base image HR with size 64 x 64 pixel and as input for downscaling and upscaling to evaluate the method. The pipeline task of MTCNN detect and crop face is shown at Figure 3. Meanwhile, the example results of cropped face at images are shown at Figure 4.

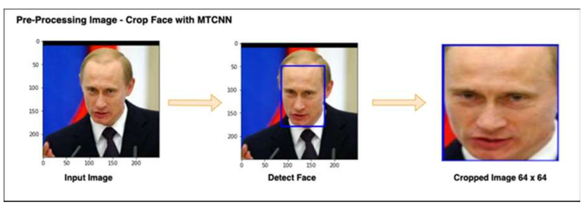

Figure 3: Pre-Processing Image – Crop the face with MTCNN

Figure 4: The Example Results Cropped Face Of Image LFW Dataset By Using MTCNN.

Next the cropped HR image size 64 x 64 pixel is downscaled to 25% by bicubic to generate low-resolution image so that produces images with size 16 x 16 pixel (LR image). The output results of LR images downscaling process are shown at Figure 5.

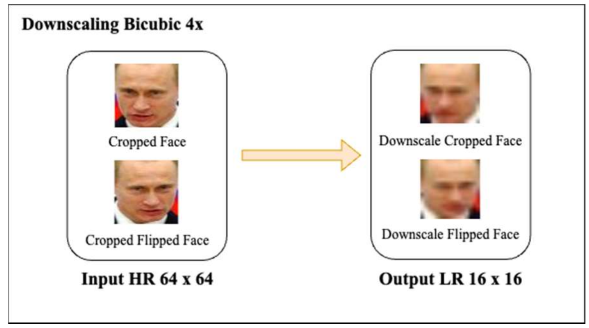

Figure 5: Downscaling Output of 25% Using Bicubic

4.3. Bicubic Interpolation

Bicubic Interpolation is the basic way to resize images. The input of this process is LR images. The LR images are upscaled so that produces HR images with 64 x 64 pixel. The output results of upscaling 4x by bicubic process are shown at Figure 6.

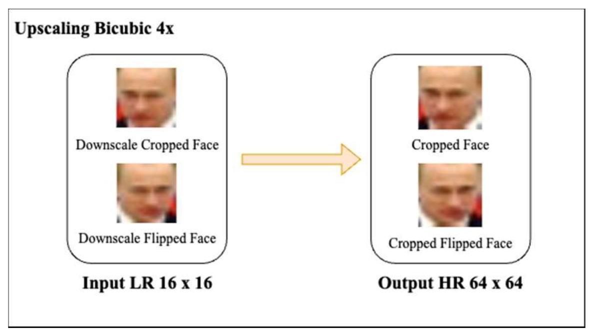

Figure 6: The Output Results Of Upscaling 4x By Bicubic Interpolation

4.4. GPEN Super Resolution

Super Resolution method that used at this research is GAN Prior Embedded Network (GPEN). The aligned cropped LR face image with 16 x 16 pixel is as input for GPEN Super Resolution to generate new image with 4x upscaling HR 64 x 64 pixel. The output results of upscaling 4x by GPEN Super Resolution is shown at Figure 7.

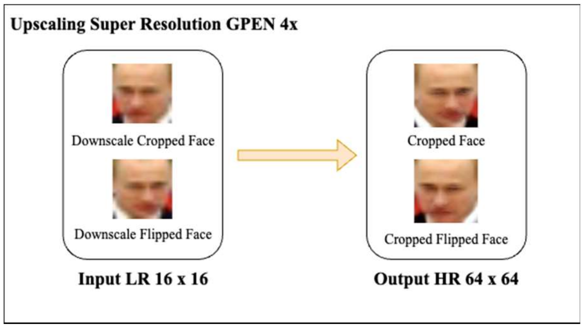

Figure 7: The Output Results Of Upscaling 4x By GPEN Super Resolution

4.5. FaceNet

FaceNet is facial recognition system that was proposed by Google Researchers in 2015. FaceNet approach to generate a high-quality face mapping from the images using deep learning architectures such as ZF-net and Inception Network. It used a method called triplet loss as a loss function to train this architecture. FaceNet employs end-to-end learning in its architecture.

All sources of dataset are fed as input to Face Recognition system by FaceNet. Support Vector Machine (SVM) is used to build the model and measure the accuracy.

4.6. Evaluation Process

Support Vector Machine (SVM) as one of the most popular Supervised Learning algorithms which commonly used for classification as well as regression problems to find a hyperplane in an N-dimensional space that distinctly classifies the data points [40]. N represents the number of features. This research used SVM for classification to predict the accuracy score of the model. Output of the accuracy training and validation are captured to have compared each other to determine the best method. Based on that a Classification report is generated to measure the quality of predictions. It displays the precision, recall, F1-Score and support score for the model. The precision can show accuracy of positive predictions. Ability of a classifier not to label an instance positive that actually negative measured by precision score. Ability of classifier to find all positive instances is measured by recall score. The recall score shows the percent of the positive cases that caught. The F1 score is used to measure of positive predictions were correct. It's a weighted harmonic mean of precision and recall; the best score is 1.0 while the worst is 0.0. The support shows the number of actual occurrences of the class in the dataset.

5. RESULT AND DISCUSSION

In this research, the LFW dataset with 423 classes is augmented by horizontal flip to compare the results with the non-augmented data. The HR images with 64 x 64 pixel are cropped using MTCNN as input for HR image. Then the HR image is scaled down with 25% by using bicubic to produce the LR images with 16 x 16 pixel. LR images are scaled up with ratio 4x back by using bicubic interpolation and compared the results with ratio 4x by using GPEN SR. There are eight images as input for FR task in Facenet in total.

The dataset with additional augmentation process produces the better performance in higher accuracy compared with without data augmentation process. As deep learning requires more data in process of learning. The more data fed, the better accuracy resulted. This research proof that in augmentation dataset, at four inputs to FaceNet (HR images, LR images, upscaled 4x bicubic interpolation, upscaled GPEN SR) can get higher accuracy compared to less data (without augmentation).

HR images are easily recognized by human eyes. In FaceNet it reaches over 98% in both cases without data augmentation and with data augmentations. Without data augmentation the accuracy is 98.1% but increased with data augmentation. With data augmentation accuracy in HR images reaches 98.4%.

The LR images significantly drop in accuracy. LR images are not easy to be recognized even with human eyes. In LR images 16 x 16 pixel without data augmentation the accuracy is 66.2% while with data augmentation can get better accuracy in 76.4%. Upscaling LR images with 4x by using bicubic interpolation can increase the accuracy but the results output images are still blurry, although the accuracy is better compared with LR images with lower resolutions. The experiment in upscaling 4x LR images by using GPEN Super Resolution produces better performance compared with the bicubic interpolation. Without data augmentation by upscaling 4x LR images with GPEN SR can produce 74.2% in accuracy. But in augmentation data the result gets better. Images with augmentation data that upscaled 4x LR images with GPEN Super-Resolution method in FaceNet can produce 82.8% in accuracy. The combination GPEN SR and FaceNet produces the best

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

performance in accuracy based on experiment at this research. Classifications model by using SVM to measure the accuracy training and accuracy validation metrics score are summarized in Table 2.

Table 2: Summary Of Evaluation Results at FaceNet

Images sources as Input		Accuracy	
		Training	**Validation**
Without Data Augmentation	HR Image 64x64	98.1%	92.8%
	LR Image 16x16	66.2%	49.9%
	Upscaling 4x Bicubic Interpolation	69.2%	51.9%
	Upscaling 4x GPEN Super Resolution	74.2%	53.8%
With Data Augmentation	HR Image 64x64	98,4%	97.7%
	LR Image 16x16	76.4%	59.6%
	Upscaling 4x Bicubic Interpolation	80.1%	64.8%
	Upscaling 4x GPEN Super Resolution	**82.8%**	**66.6%**

To measure the quality of prediction, classification report testing is generated for the result of upscaling 4x LR images using GPEN SR and FaceNet classification.The best method by using combination of upscaling 4x LR images by using GPEN SR and FaceNet experiment evaluated on LFW dataset with 423 classes at this research can produce Accuracy Testing score 69%. Precision, Recall, F1-Score and Support of the method are shown at Table 3.

Table 3: Classification Report Test For Combination Upscaling 4x GPEN Super Resolution and FaceNet

	Precision	**Recall**	**F1-Score**	**Support**
0	0.58	0.88	0.70	8
1	0.57	0.80	0.67	5
2	0.00	0.00	0.00	2
3	1.00	0.50	0.67	2
4	0.67	0.67	0.67	3
5	1.00	0.67	0.80	3
6	0.00	0.00	0.00	2
7	0.00	0.00	0.00	2
8	1.00	1.00	1.00	2
9	0.65	0.94	0.77	16
10	0.00	0.00	0.00	3
…	…	…	…	…
…	…	…	…	…
413	0.00	0.00	0.00	2
414	0.00	0.00	0.00	2
415	1.00	1.00	1.00	2
416	0.75	1.00	0.86	3
417	0.75	1.00	0.86	3
418	0.75	1.00	0.86	3
419	1.00	1.00	1.00	2
420	1.00	0.83	0.91	6
421	1.00	1.00	1.00	4
422	1.00	1.00	1.00	2
Accuracy			0.69	2394
Macro avg	0.56	0.49	0.50	2394
Weighted avg	0.65	0.69	0.64	2394

In summary, by generating more dataset with data augmentation, accuracy is getting better. The combination method upscaling 4x GPEN Super Resolution produce better accuracy compared with the method using upscaling 4x Bicubic Interpolation although it's still not close to HR image accuracy score. Upscaling 4x Bicubic Interpolation with data augmentation produces testing accuracy score of 66%. The summary testing accuracy score are shown at Table 4.

Table 4: Summary Of Testing Accuracy Score

Images Sources as input		Testing Accuracy Score
Without Data Augmentation	HR Image 64x64	96%
	LR Image 16x16	52%
	Upscaling 4x Bicubic Interpolation	55%
	Upscaling 4x GPEN Super Resolution	57%
With Data Augmentation	HR Image 64x64	97%
	LR Image 16x16	61%
	Upscaling 4x Bicubic Interpolation	66%
	Upscaling 4x GPEN Super Resolution	**69%**

6. CONCLUSION AND FUTURE WORK

This research performs combination method GPEN Super Resolution and FR method with FaceNet to solve the LRFR problem as best method compared with traditional way using bicubic interpolation. The FaceNet evaluation results are presented in Table 2. As more data fed, the more accuracy resulted. By reproducing the training dataset through augmentation flip horizontal can increase the training accuracy to 82.8% and produce better testing accuracy to 69% as well.

However, this combination method is still not the highest in accuracy. FaceNet in HR images reached over 98% accuracy in this experiment. There are still ranges of gaps that can be improved in the next future work by combining the other SR method to generate the better-quality HR images from LR images. Or find another better way to improve quality of images from LR input. Compared with other SR methods, the GPEN can reduce the over-restoration process so that produce better performance in accuracy. The results show that there are some gaps in accuracy testing, validation compared with training accuracy. Since this experiment using the pre-defined model of Super Resolution GPEN without any fine-tuning done further, that's probably why the testing accuracy significantly dropped.

Future work for improvement, the SR model need more fine-tuning and training to produce the better quality of images in super-resolution process based on existing model or by combining this work with other SR model to generate the better quality of images, so the result would be better in accuracy.

Conflict of interest

The authors declare that there is no conflict of interest in this paper.

REFERENCES:

[1] Ge Wen, Huaguan Chen, Deng Cai, Xiaofei He, "Improving face recognition with domain adaptation", NeuroComputing, vol. 287, pp. 45-51, 2018.
[2] G. B. Huang, M. Mattar, T. Berg, and E. Learned-miller, "Labeled Faces in the Wild : A Database for Studying Face Recognition in Unconstrained Environments," Tech. rep., 2008.
[3] B. Amos, B. Ludwiczuk and M. Satyanarayanan, "OpenFace: A general-

purpose face recognition library with mobile applications," 2016.

[4] T. Baltrusaitis, P. Robinson and L.-P. Morency, "OpenFace: an open source facial behavior analysis toolkit," in 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), New York, 2016.

[5] Y. Wang, T. Bao, C. Ding and M. Zhu, "Face recognition in real-world surveillance videos with deep learning method," in 2017 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, China, 2017.

[6] Guodong Guo, Na Zhang, "A survey on deep learning based face recognition", Computer Vision and Image Understanding, 2019, doi: 10.1016/j.cviu.2019.102805

[7] Sisi Li, Zhonghua Liu, Di Wu, Hua Huo, Haijun Wang, Kaibing Zhang, "Low-resolution face recognition based on feature-mapping face hallucination", Computers and Electrical Engineering, vol. 101, July 2022, doi: 10.1016/j.compeleceng.2022.108136.

[8] Z. Lu, X. Jiang and A. Kot, "Deep Coupled ResNet for Low-Resolution Face Recognition," IEEE Signal Processing Letters, vol. 25, no. 4, pp. 526-530, April 2018.

[9] Z. Cheng, X. Zhu and S. Gong, "Face re-identification challenge: Are face recognition models good enough?," Pattern Recognition, 2020.

[10] Z. Cheng, X. Zhu and S. Gong, "Low-Resolution Face Recognition," in Asian Conference on Computer Vision, 2018.

[11] P. Li, L. Prieto, D. Mery and P. Flynn, "Face Recognition in Low Quality Images: A Survey," Computer Vision and Pattern Recognition, 2018.

[12] C. Herrmann, D. Willersinn and J. Beyerer, "Low-resolution Convolutional Neural Networks for video face recognition," in 2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Colorado Springs, CO, USA, 2016.

[13] T. Bao, C. Ding, S. Karmoshi, and M. Zhu, "Video-based face recognition via convolutional neural networks," Second Int. Work. Pattern Recognit., vol. 10443, p. 104430I, 2017.

[14] P. Li, L. Prieto, D. Mery and P. J. Flynn, "On Low-Resolution Face Recognition in the Wild: Comparisons and New Techniques," in IEEE Transactions on Information Forensics and Security, vol. 14, no. 8, pp. 2000-2012, Aug. 2019, doi: 10.1109/TIFS.2018.2890812.

[15] C. You, G. Li, Y. Zhang, X. Zhang and H. Shan, "CT Super-Resolution GAN Constrained by the Identical, Residual, and Cycle Learning Ensemble (GAN-CIRCLE)," vol. 39, no. 1, January 2020.

[16] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-Janua, pp. 105–114, 2017.

[17] F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 815-823, doi: 10.1109/CVPR.2015.7298682.

[18] Tao Yang, Peiran Ren, Xuansong Xie, Lei Zhang, "GAN Prior Embedded Network for Blind Face Restoration in the Wild," Computer Vision and Pattern Recognition, 2021, doi: 10.48550/arXiv.2105.06070.

[19] J. W. Hwang and H. S. Lee, "Adaptive Image Interpolation Based on Local Gradient Features," IEEE Signal Process. Lett., vol. 11, no. 3, pp. 359–362, 2004.

[20] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," BMVC, no. Section 3, pp. 41.1-41.12, 2015.

[21] H. M. Chang, H. Seo, and S. Bum, "A face recognition system based on convolution neural," Soft Comput., 2016.

[22] X. Tang, "DeepID3: Face Recognition with Very Deep Neural Networks," pp. 2–6, 2015.

[23] P. Li, L. Prieto, D. Mery, and P. Flynn, "On low-resolution face recognition in the wild: Comparisons and new techniques," IEEE Trans.

[24] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Yu Qiao, "Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks", Computer Vision and Pattern Recognition, 2016, doi: 10.48550/arXiv.1604.0287.

[25] K. Kranthi Kumar, Y. Kasiviswanadham, D.V.S.N.V. Indira, Pushpa Priyanka palesetti, Ch.V. Bhargavi, "Criminal face

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

identification system using deep learning algorithm multi-task cascade neural network (MTCNN)," 2021, doi: 10.1016/j.matpr.2021.06.373

[26] I. J. Goodfellow et al., "Generative Adversarial Networks," pp. 2672– 2680, 2014

[27] R. A. A. Al-falluji, A. A. Youssif, and S. K. Guirguis, "Single Image Super Resolution Algorithms : A Survey and Evaluation," Int J Adv Res Comput Eng Technol, vol. 6, no. 9, pp. 1445–1451, 2017.

[28] S. Singh and S. V. A. V. Prasad, "Techniques and challenges of face recognition: A critical review," Procedia Comput. Sci., vol. 143, pp. 536– 543, 2018.

[29] G. B. Huang, M. Mattar, T. Berg, and E. Learned-miller, "Labeled Faces in the Wild : A Database for Studying Face Recognition in Unconstrained Environments," Tech. rep., 2008.

[30] Deng, Jiankang & Guo, Jia & Zafeiriou, Stefanos. (2018). ArcFace: Additive Angular Margin Loss for Deep Face Recognition.

[31] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. ArXiv, 2018.

[32] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan.

[33] Yaoyao Zhong, Weihong Deng, Jiani Hu, Dongyue Zhao, Xian Li, Dongchao Wen, "SFace: Sigmoid-Constrained Hypersphere Loss for Robust Face Recognition," Computer Vision and Pattern Recognition, 2022, doi: 10.48550/arXiv.2205.12010.

[34] Xiao Zhang, Rui Zhao, Yu Qiao, Xiaogang Wang, Hongsheng Li, "AdaCos: Adaptively Scaling Cosine Logits for Effectively Learning Deep Face Representations," Computer Vision and Pattern Recognition, 2019, doi: 10.48550/arXiv.1905.00292

[35] M. R. Golla and P. Sharma, Performance Evaluation of Facenet on Low Resolution Face Images. Springer International Publishing, 2019.

[36] Erfan Zangeneh, Mohammad Rahmati, Yalda Mohsenzadeh, "Low resolution face recognition using a two-branch deep convolutional neural network architecture," Expert Systems with Applications, Volume 139, 2020, doi: 10.1016/j.eswa.2019.112854.

[37] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, Xiaoou Tang, "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," Computer Vision and Pattern Recognition, 2019, doi: 10.48550/arXiv.1809.00219

[38] X. Yu, B. Fernando, R. Hartley and F. Porikli, "Super-Resolving Very Low-Resolution Face Images with Supplementary Attributes," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 908-917, doi: 10.1109/CVPR.2018.00101.

[39] Ledig, C., Theis, L., Husz´ar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image superresolution using a generative adversarial network. In: CVPR. (2017).

[40] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt and B. Scholkopf, "Support vector machines," in IEEE Intelligent Systems and their Applications, vol. 13, no. 4, pp. 18-28, July-Aug. 1998, doi: 10.1109/5254.708428.