Face Recognition on Low-Resolution Image using Multi Resolution Convolution Neural Network and Antialiasing Method

Mario Imandito[1], Suharjito[2]

Computer Science Department, BINUS Graduate Program-Master of Computer Science
Bina Nusantara University, Jakarta, Indonesia 11480

Abstract—Video surveillance applications usually take pictures of faces that have a low resolution (12x12) due to distance, lighting and shooting angles Most of face recognition algorithm have the poor performance accuracy and poor identify face on low resolution. Based on the problem, identifying the face of the query in low resolution, based on high resolution (64x64) proves to be a huge challenge. The aim of this research is to develop a new model for face recognition of low-resolution image in order to increase the accuracy of recognition. A Multi-Resolution Convolutional Neural Network (MRCNN) is proposed to address the problem. First, Antialiasing is used in preprocessing phase, then use MRCNN to extract the feature of the image. LWF (Labeled Face in Wild) will be used to evaluate the model. The result of this study is increasing the accuracy of face recognition on low-resolution image compared to the previous MRCNN model.

Keywords—*Face recognition; low resolution; convolutional neural network; antialiasing*

I. INTRODUCTION

Security in one of the most concern problem in almost every application. Face recognition is one of many samples of security method. The existing face recognition system such as Principal Component Analysis (PCA) [1], Linear Discriminant Analysis (LDA) [2], and the most popular Super Resolution (SR) [3] have achieved satisfactory performance, in case that the face images that collected are in high resolution and are well aligned. But in case of video surveillance system, as the face target are far away from the cameras, the captured facial images are usually in low resolution (32x32). This affects the accuracy of facial recognition system. This issue of face recognition is known as low-resolution face recognition (LRFR). Besides, video surveillance system usually used to identify someone in a secure area such as workspace, data center etc. To identify person in low resolution based on high resolution image proves to be a huge challenge.

Because of that, some research in face recognition for low resolution image has been done. In 2014 [4] proposed a fusion method where they take several video frames as an input and fuse it into one image in order to derive rich feature. At the same year, [5] proposed Multi-resolution Feature Fusion (MFF). At their research, they employ a Gabor – feature hallucination method to estimate the higher resolution Gabor features from low resolution (LR) from the LR Gabor Features followed by feature extraction. The result of their research is

outperforming the previous research on the ORL and FERET databases. Still at 2014, [6]proposed Coupled Basis & Distance (CBD) method. They are matching biometric data from disparate domains. Min-Chun Yang in 2015 [7] joint face recognition and hallucination algorithm based on sparse representation. Instead performing recognition as standard approaches, their model can learn a person-specific face hallucination with recognition guarantees. Where the result of the experiment outperformed the previous method. In 2016 [8] proposed a method using Low-Resolution Convolutional Neural Network (CNN) by proposing an appropriate network architecture. The input is from low-resolution video face recognition with a manifold-based track comparison strategy. In the next year in 2017 [9] proposed a multi-resolution convolutional neural network (MRCNN). This model is proposed in order to study the consistent feature representation from high-resolution and low-resolution face images.

The main problem of this research is a low-resolution image as the input process. The state-of-the-art method [9] works on feature extraction phase, and Bicubic as their preprocessing phase. Their experiment result showed great value in recognition rate. But as mentioned before, the state-of-the-art use Bicubic where Bicubic is an old method and the output is not good enough. There is better method in preprocessing phase. This study replaces Bicubic as their preprocessing phase with hallucination as proposed by [10] hopefully can increase the face recognition performance. This method is proved in their research have better result than Bicubic. And with the better input, the accuracy of face recognition will increase.

In this paper, MRCNN [9] method is proposed to achieve better accuracy of face recognition of low-resolution image. Anti-alias [11] is used to generate a high resolution (HR) image from low resolution (LR) probe and MRCNN is carefully designed to learn the consistent feature representations of both the gallery images and the generated ones. This paper uses Anti-alias, Anti-alias produce better image rather than Bicubic that is used in state of the art [9].

The remainder of this paper is organized as follows. In Section II, related work in face recognition on low-resolution image. Section III, Research method will be explained that was used in this study. Section IV will explain the result of the study. Section V is Conclusion and future works.

II. RELATED WORKS

As we know that security is so much important in application development. Because if there is no security, the data can be harmed. Because this motivation, may research try to build a strong security method. And one of security system is face recognition. In last decade of research, face recognition has been solved. But as far as technology goes, there are comes a new challenge in face recognition. One of them is face recognition on low resolution image. This kind of problem make researcher challenged to solve the problem.

In 2014 [4] make a research in surveillance system. They see that the captured faces are often very small resolution. In their research for pre-processing phase, they use Histogram Equalization for reducing illumination variation. After that, the images will be fused using curvelet feature. In order to enhance face feature, they proposed a super-resolution based on face recognition algorithm. They use 2 methods here. First, they use of sparsity signal representation to train low-resolution image. The second method is Eigen-subspace feature of human face. Both of high-resolution face image are then combined into one image with pixel by pixel decision making. After combining all blocks together, the final enhanced face image will be used for recognition.

In the same year [5] the do the research on face recognition on low resolution as well. As same as the previous research above, they focus on feature fusion method. They proposed Multi-resolution Feature Fusion (MFF). The research starts with Gabor Wavelet for extracting local features following by Canonical Correlation Analysis (CCA) for measuring the linear relationship between two multidimensional variables, Generalized Canonical Analysis (GCCA) and Generalized Canonical Projective Vector (GCPV). After that MFF method for face recognition presented. Different with above [6] they focus on the problem of comparing a low-resolution image with the high-resolution one. The previous coupled mapping methods do not fully exploit the high-resolution information, or they do not simultaneously use samples from both domains during training. Because of that, they proposed Coupled Basis & Distance (CBD) that learns coupled distance metrics. The method learns coupled distance metrics in two steps. In addition, they propose to jointly learn wo semi-coupled bases that yield optimal representation. In particular, the high-resolution images are used to learn a basis and distance that result in increased class-separation. The low-resolution images are used to learn a basis and distance metric that map to low-resolution data to their class-discriminated high-resolution pairs. And at the end, the two-distance metrics are refined to simultaneously enhance the class separation of both high-resolution class-discriminant and low-resolution projected images.

In 2016 [8] as same as [4], they do their research on surveillance system. They proposed Convolutional Neural Network (CNN) for the low-resolution video face recognition. They transfer the success of CNN on high-resolution image into low-resolution image by proposing an appropriate network architecture. Their focus is on efficiency of track matching strategy. Because related literature employs an effective but inefficient many-to-many comparison. Instead, they reduce the necessary comparisons by defining a fixed number of local patches in the face descriptor set and show that low numbers of patches are sufficient for superior comparison result.

Next research in 2017 [12] proposed a cluster-based regularized simultaneous discriminant analysis (C-RSDA) based on SDA. Next year in 2018 [13] proposed a method called low-rank representation and locality-constrained regression (LLRLCR) to learn occlusion-robust representation features. Last research found in 2019 proposed by [3] they called SSR2 (Sparse Signal Recovery) for single-image super-resolution on faces with extreme low resolution.

Based on the literature review result on Table I, the goal of the research is to increase the recognition rate. Our research focus is on CNN method and the model in [9] is the state-of-the-art of our research. We found that the state-of-the-art use Bicubic for the preprocessing phase. Where there is better method that has better result than Bicubic. And from literature review, we will use Eigen transformation Hallucination that proposed by [10] to replace Bicubic. Another research from [7] that uses Hallucination also shows better result than Bicubic.

TABLE. I. RELATED WORK COMPARISON

Author	Method	Dataset	Result
[10]	Eigen transformation based on Hallucination method.	XM2VTS	Increasing recognition rate
[4]	Fusion method from video surveillance system.	FRGC, AR, ScFace and Curtin Face.	Increasing recognition rate
[5]	Multi-resolution Feature Fusion (MFF)	ORL and FERET	Increasing recognition rate
[6]	Coupled Basis & Distance (CBD)	ScFace and Multi-PIE (Pose, Illumination, and Expression)	Increasing recognition rate
[7]	Joint method Hallucination and Face recognition	CMU MULTI-PIE and FRGC	Increasing recognition rate
[8]	Low-Resolution Convolutional Neural Network (CNN)	YTF	Increasing recognition accuracy
[9]	multi-resolution convolutional neural network (MRCNN)	CMU-PIE and Extended Yale B	Increasing recognition rate
[12]	Cluster-based regularized simultaneous discriminant analysis (C-RSDA)	Feret, LFW, SCFace	Increasing recognition rate
[13]	Low-rank representation and locality-constrained regression (LLRLCR)	AR face and Extended Yale B	Increasing recognition rate
[3]	SSR2 (Sparse Signal Recovery)	MPIE	Increasing recognition rate

Multi-resolution Convolutional Neural Network (MRCNN) was proposed by [9] to study the consistent feature representation from high resolution and low-resolution face images. First, the corresponding labelled multi-resolution face images are utilized to train the MRCNN model. After that process, the trained model is used as the feature extractor in order to obtain features for the targets in the gallery and query images respectively. Finally, the nearest neighbor is applied as the classifier for the purpose of final identification. The preprocessing phase of this method using Bicubic interpolation to ensure that the network inputs are of similar size. For the feature extraction, they use CNN. The input or MRCNN architecture is the mixed gray facial images from the gallery images and the generated gallery images. After the features are obtained, the cosine distance is used to measure the similarity between the probe feature and the gallery feature.
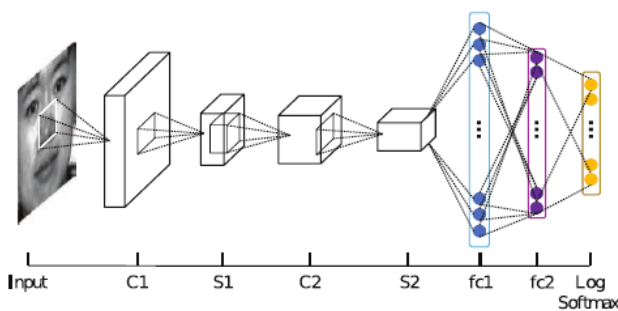

Fig. 1. The Architecture of MRCNN [9].

Fig. 1 shows the architecture of MRCNN. The mixed gray facial image from gallery image and the generated high-resolution image are the input for the network of MRCNN above. For extract local information they use two convolutional layers following by max-pooling layers. For the activation function, they use ReLu function. And after the convolutional operation, they obtained 64 vectorized feature maps. A fc1 or Fully Connected layer is used to fuse the global information and the output of it are used to the features. Finally, from the fig x, the combination between fc2 and the Log SoftMax layer is used as classifier in order to compute the probability that the input belongs to a certain class.

III. RESEARCH METHODS

A. Research Steps

The research steps to build a new model of face recognition method by Multi Resolution CNN and Anti-alias are: model development, model implementation such as training and testing model, and analysis the result. The flow this study is shown in Fig. 2. The research begins with define research background and the scope of the result. Literature study is to know more about face recognition on low resolution image and the state of the art for this research. Besides that, literature review can be used for guiding of the methods to be used.

The second step is data collection. In this study the data of face was collected from internet. LWF (Labeled Face in Wild) was used to develop the model of low-resolution face recognition. Based on literature review, the research can define what model will be made. Third step is model implementation including training and testing model. The fourth step is analysis the result, what the benefits and limitation compared with the previous research. And the final step is conclusion.

This research was based on [9] where we replace Bicubic as their preprocessing phase with Eigen transformation-based hallucination. In Fig. 3 shows the proposed model that was used in this study. After preprocessing phase using the mentioned method above, the feature extraction phase using Multi Resolution Convolutional Neural Network and the output will be used as classifier to compute probability that the input belongs to a certain class.

B. Proposed Model

The First step of this model is data preparation, HR images with 250x250 are resize into 32x32 and down sampled into LR images 12x12. After the LR images are created, then upscaled using Anti-alias and called generated HR or HR'. The LR image and HR' images furthermore will be blended into one image. And the blended images will be used as train and validate the CNN model.

In this model, the method of feature extraction used CNN. The inputs are mixed gray scaled face image between HR' image and HR image. The input size is 1x32x32. Two convolutional layers, followed by Max-pooling layers, are used to extract local information. The kernel size is 5×5. ReLU function is used as the activation function. After the convolutional operation, 64 feature maps of size 5×5 are obtained. Those maps are vectorized and a fully connected layer (fc1) is used to fuse the global information to form a global feature and the outputs of the fc1 layer are taken as the features. Finally, the combination of fc2 layer and the Log SoftMax layer is used as the classifier to compute the probability that the input belongs to a certain class.

C. Evaluation Methods

LWF (Labeled Face in Wild) [14] contains 5.749 persons and 13.223 images. Each person has variety number of images. This study will run 2 training model. The first one is training model for classification into 2 classes (top 2 person with the most images which are George W Bush with 530 images and Collin Powel with 236 images) and the second one is training model for classification into 143 classes (person with images more than 10 images). This study will use 70%, 20%, and 10% images for training data, validation data, and testing data respectively for each person. For evaluating the proposed model, this study used accuracy and execution time. The result of evaluation will be compared with the state of the art of this model [9].
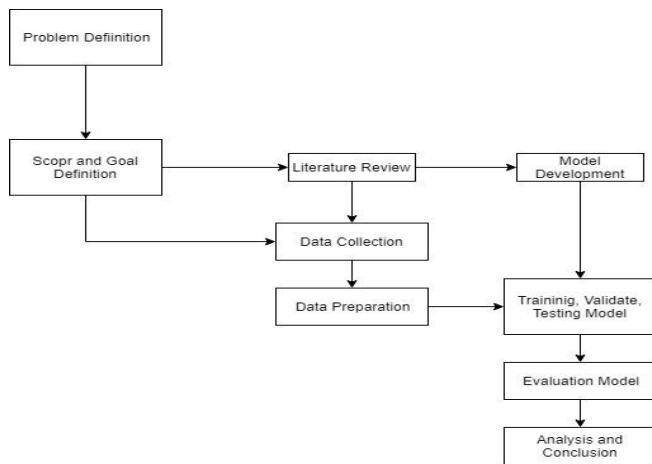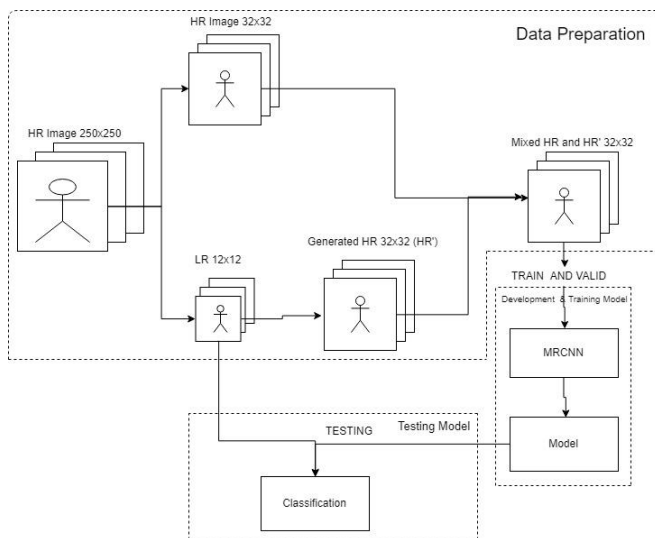
Fig. 2. Research Steps.

Fig. 3. Proposed Model

IV. RESULTS AND DISCUSSION

Fig. 4 shows data pre-processing using Antialiasing method.

Based on Fig. 4, it can be seen that the input model are the HR image using 250x250 size. These images were resize into 32x32 as HR Image. The input images were down sampling into LR image with size 12x12. After that the LR image will be built using the antialiasing method to be generated HR with a size of 32x32. Then the resized HR image from input and generated HR is combined as an MRCNN modeling input in the training and testing process.

Fig. 4 shows example of data pre-processing using Bicubic that has been proposed in [9]. This study used 648 face images for 2 classes. The evaluation result for 2 classes faces recognition were shown in Table II.

Based on Table II, the accuracy of MRCNN model using Antialiasing method for pre-processing was greater than MRCNN model that using Bicubic interpolation method. Bicubic interpolation method acquired 61% in accuracy on low resolution face recognition. While Antialiasing acquired

70% in accuracy of on low resolution face recognition. The differences of this accuracy result were 15%, it indicates that the performance of the proposed method is outperform than the state-of-the-art model [9].

Table III shows the execution time comparison between the proposed model and the state of the art. The execution time was count in generating image from LR into generated HR.

As shown in Table III, Bicubic took less time to generated HR image from LR image. Bicubic interpolation method needed 2,199 second in generating HR image from LR. Antialiasing method used 3,434 for generating HR image from LR image. It means that Bicubic interpolation method is faster than Antialiasing method that used an Eigen transformation based on hallucination method.
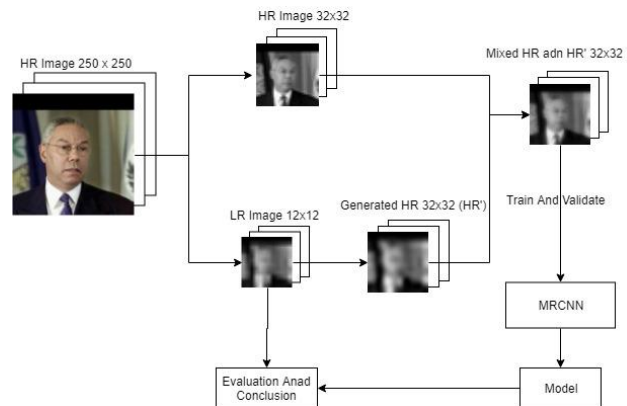

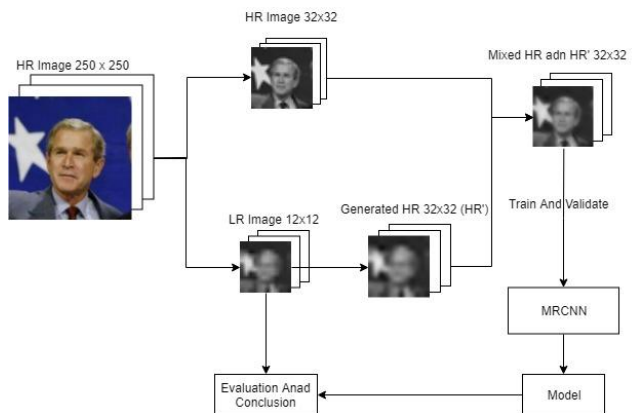Fig. 4. Modelling step using Anti-alias for 2 Classes image

Fig. 5. Modelling step using Bicubic for 2 Classes image

TABLE. II.   COMPARATION OF ACCURACY RATE

Method	Accuracy Rate (%)
Bicubic	61
Antialiasing	**70**

TABLE. III.   COMPARATION OF EXECUTION TIME

Method	Time (sec)
Bicubic	2.199
Antialiasing	3.434

V. Conclusion and Future Works

This paper proposes a new Multi Resolution Convolutional Neural Network model with Antialiasing techniques as a pre-processing method. The results of the comparison of this pre-processing method with the Bicubic method shows that the antialiasing pre-processing method produces better accuracy than the Bicubic pre-processing method. The proposed model can increase the accuracy by 15% in face recognition using the multi-resolution convolutional Neural Network method. However, the antialiasing pre-processing method uses a longer processing time compared to the Bicubic pre-processing method. Therefore, for further research will be directed to obtain a hybrid pre-processing method that can adapt to facial image characteristics that will be recognized based on differences in resolution and lighting.

Acknowledgment

References

[1] S. K. Dandpat and S. Meher, "Performance improvement for face recognition using PCA and two-dimensional PCA," in International Conference on Computer Communication and Informatics, 2013.

[2] M. Kan, S. Shan, H. Zhang, S. Lao and X. Chen, "Multi-view discriminant analysis," IEEE transactions on pattern analysis and machine intelligence, vol. 38, pp. 188--194, 2015.

[3] R. Abiantun, F. Juefei-Xu, U. Prabhu and M. Savvides, "SSR2: Sparse Signal Recovery for Single-Image Super-Resolution on Faces with Extreme Low Resolutions.," Pattern Recognition., vol. 90, pp. 308-324., 2019.

[4] X. Xu, W. Q. Liu and L. Li, "Low resolution face recognition in surveillance systems.," Journal of Computer and Communications, vol. 2, pp. 70-77, 2014.

[5] H. K. Pong and K. M. Lam, "Multi-resolution feature fusion for face recognition.," Pattern Recognition, vol. 47(2), pp. 556-567, 2014.

[6] P. Moutafis and I. A. Kakadiaris, "Semi-coupled basis and distance metric learning for cross-domain matching: Application to low-resolution face recognition.," In Biometrics (IJCB), 2014 IEEE International Joint Conference on, pp. (pp. 1-8), 2014.

[7] M. C. Yang, C. P. Wei, Y. R. Yeh and Y. C. F. Wang, "Recognition at a long distance: Very low resolution face recognition and hallucination.," In Biometrics (ICB), 2015 International Conference on, pp. (pp. 237-242), 2015, May.

[8] C. Herrmann, D. Willersinn and J. Beyerer, "Low-resolution convolutional neural networks for video face recognition..," In Advanced Video and Signal Based Surveillance (AVSS), 2016 13th IEEE International Conference on, pp. (pp. 221-227), 2016.

[9] C. Ding, T. Bao, S. Karmoshi and M. Zhu, "Low-resolution face recognition via convolutional neural network.," In Communication Software and Networks (ICCSN), 2017 IEEE 9th International Conference on, pp. (pp. 1157-1161), 2017, May.

[10] X. Wang and X. Tang, "Face hallucination and recognition.," In International Conference on Audio-and Video-Based Biometric Person Authentication, pp. pp. 486-494, 2003, June.

[11] B. Dai, R. Zhuo, S. Yin, M. Lv, R. Hong, Q. Wang, and X. Wang, "Ultrafast imaging with anti-aliasing based on optical time-division multiplexing," Optics letters, vol. 41, pp. 882--885, 2016.

[12] Y. Chu, T. Ahmad, G. Bebis and L. Zhao, "Low-resolution face recognition with single sample per person.," Signal Processing, vol. 141, pp. 144-157, 2017.

[13] G. Gao, Z. Hu, P. Huang, M. Yang, Q. Zhou, S. Wu and D. Yue, "Robust low-resolution face recognition via low-rank representation and locality-constrained regression.," Computers & Electrical Engineering, vol. 70, pp. 968-977., 2018.

[14] G. B. Huang, V. Jain and E. Learned-Miller, "Unsupervised Joint Alignment of Complex Images," in In 2007 IEEE 11th International Conference on Computer Vision , 2007.