2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)

Facial Image Denoising Using Convolutional Autoencoder Network

Naing Min Tun
Bauman Moscow State Technical University
Moscow, Russia
naingminhtun52@gmail.com

Alexander I. Gavrilov
Bauman Moscow State Technical University
Moscow, Russia
alexgavrilov@mail.ru

Nyan Linn Tun
Bauman Moscow State Technical University
Moscow, Russia
nyanlin54@gmail.com

***Abstract*—Noise effects can interfere the face recognition process in outdoor conditions. Therefore, image denoising topic is the classical issue in the field of image processing and computer vision subjects. In this paper, we show that the solution of denoising process using the autoencoder networks based on the ORL face database. The proposed method can support face recognition systems designed for use in an outdoor environment as the preprocessing stage and it can provide the effective results after training process.**

Keywords—Autoencoder, image denoising, face recognition, convolutional neural networks

I. Introduction

Face recognition for using a video surveillance system is the most popular topic of security systems because of its advantages such as the lack of user interaction. Interference may occur in the camera's area when face recognition applications are used in outdoor conditions. Noises can cause the unwanted impacts such as blurred objects, artifacts, unseen lines and disturbs background scenes. Even if the most effective deep learning methods can get the unsatisfied results when there is interference in the images or video frame that they want to recognize [1]. As a fact, it is necessary to reduce interference from face images if we want to obtain reliable results using modern face recognition methods.

The process of removing random noise and reserving the original image is the interesting topic in computer vision and image processing fields. Over the past decades, there are many methods designed to reduce noise in the digital image processing. In [2], the authors proposed a review of methods used in this context area. These methods can be categorized based on their natures of techniques. These methods are statistical based, median based, derivative based, fuzzy logic based and morphological based. Noisy models are also needed in digital image processing for the exploration of image noise reduction techniques. In paper [3], several noisy models are presented. These models are Gaussian Noise Models, White Noise, Brownian Noise (Fractal Noise), Impulse Valued Noise (Salt and Pepper Noise), Periodic Noise, Quantization noise, Speckle Noise, Photon Noise (Poisson Noise), Poisson-Gaussian Noise, Structured Noise, Gamma Noise and Rayleigh noise.

Recently, deep learning methods are widely used in many fields, such as robotics, self-driving cars, automatic text generation, automatic machine translation, automatically adding sounds to movies and image recognition, etc. Example of face recognition system using combination of feature extraction method and deep neural network and was proposed in paper [4]. Image denoising based on deep learning techniques have attracted much attention in recent 20 years [5]. The first deep convolutional neural network was applied for image denoising purpose in 2015 [6]. After then, deep networks were widely used in various applications such as speech denoising [7] and image restoration [8]. Multiple convolutional and deconvolutional networks were used for the purpose of suppressing the noise and recovering the high-resolution image [8]. Moreover, there are a lot of deep learning techniques for noisy image denoising which are comprehensively described in the paper [9]. Autoencoder is a type of deep learning methods and it can efficiently use for the image denoising purpose. Denoising autoencoders using convolutional layers have better performance for their ability of exploiting robust spatial correlations.

In this paper, we used convolutional autoencoder method for image denoising in the ORL database. The rest of this paper is composed as following: section II presents related work associated with the context of topic. Section III describes autoencoder method. Section IV discusses our experimental results. Conclusion is presented in section V.

II. Related Work

In the field of deep learning, image denoising is building block style of neural networks. Vincent et al [10] introduced denoising autoencoders as an extension to classic autoencoders for the purpose of extracting and reconstructing robust features of noisy input image. Its objective function was to learn representations of the input that are robust to small irrelevant changes in input.

Jain et al. [11] carried out image denoising using convolutional networks. Their experiments have shown that utilizing the convolutional networks method can provide better performance than the wavelet and Markov random field (MRF) methods. Cho [12] used a deep learning model called Boltzmann machines, can perform image denoising in the event of high noise levels better than the classical denoising autoencoders. I. Ali et al. [13] performed image denoising on RGB color image dataset by using autoencoder. They presented encoding and decoding process CIFAR-10 image dataset using 2D convolutional neural network.

L. Gondara [14] used autoencoders that use convolutional layers training on small sample data size to efficiently use for noise reduction on medical images. D. Lee et al. [15] proposed

978-1-7281-4590-7/20/$31.00 ©2020 IEEE

convolutional denoising autoencoder model using chest radiograms training data. They compared the performance of their model with conventional denoising algorithms including median filter, total variation (TV) minimization, and non-local mean (NLM) algorithms.

III. Research Methodology

A. *Noise Models*

When the face recognition system is used in real time at the outdoor conditions, the noise effects can impact the systems. Noise produces undesirable outcomes such as artifacts, unrealistic edges, blurred things and disturbs background scenes. The more the noise effects are eliminated from the image, the higher accuracy rate we can get. For this reason, we used three types of noise models for experiments. It is essential for further processing.

1) *Gaussian noise*

Gaussian noise is statistical noise having a probability density function (PDF) equal to that of the normal distribution. This noise is presented as the following equation:

$$PG(z)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

Where z is the grey level, μ denotes the mean of normal distribution, σ presents the standard deviation.

2) *Salt & Pepper noise*

Salt and Pepper noise is one of the most popular noises and a sparsely happening white and black pixels on images. Image pixel values are replaced by corrupted pixel values either 255 'or' 0. The probability density function (PDF) of this noise is shown in Fig. 1.

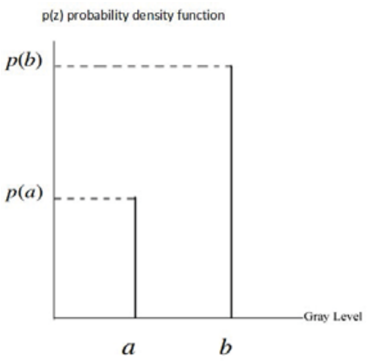

Fig. 1. The PDF of Salt & Pepper noise

$$P(z)=\begin{cases} Pa\ for\ \ z=a \\ Pb\ for\ z=b \\ 0\ otherwise \end{cases}$$

Fig. 1 presents the PDF of Salt & Pepper noise, if μ is zero and σ^2 is 0.05.

3) *Poisson noise*

Poisson noise is a type of noise which is also known as shot noise or photon noise. This noise can be modeled by a Poisson process. It is related to the measurement of light and independence of photon detections [16]. The process of photon counting is presented by the discrete probability distribution over a time interval t as follow:

$$\Pr(N=k)=\frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

Where λ is the expected number of photons per unit time or mean number of occurrences, N is the number of photons measured by a sensor element.

B. *Convolutional autoencoder*

An autoencoder is a type of unsupervised machine learning algorithms that tries to encode the input images and decode it back using some number of bits from the latent space representation. It is used for the purpose of data compression or denoising in most applications. In this paper, we used autoencoder for image noise reduction purpose. In general, it is composed of two main parts: an encoder and a decoder. The encoder maps the input images into some kind of hidden representation and the decoder reconstructs the original input images from the codes generated by the encoder. The typical workflow of convolutional autoencoder network is shown in Fig. 2.

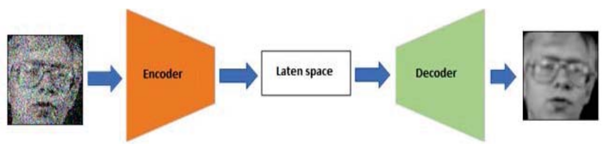

Fig. 2. Example of denoising Autoencoder network.

As the mathematical expression, an autoencoder takes the inputs $x_1, x_2, x_3, ..., x_n$ from unlabeled training input dataset (domain X) and produces the denoising images $x_1', x_2', x_3', ..., x_m'$ as the output (domain F). A general description of the autoencoder can be seen as follows:

$$\phi : X \to F$$
$$\psi : F \to X$$
$$\phi, \psi = \underset{\phi,\psi}{\arg\min} \left\| X - (\psi \circ \phi) X \right\|^2$$

Where, ϕ denotes the encoder function that maps the input image data X to a latent space (hidden representations) F. The latent space is the smallest part in the autoencoder network, and this is called a bottleneck. ψ denotes the decoder function that maps the latent space F to the output.

The encoder network of an autoencoder can be represented by the standard neural network function through an activation function as follow:

$$z = \sigma(Wx + b)$$

where, z is the element of latent space. Similarly, the decoder network can be represented in the same form, but with different weight, bias, and activation function being used.

$$x' = \sigma'(W'z + b')$$

Where, σ', W', b' of the decoder network may be unrelated σ, W, b of the encoder network.

Autoencoders are trained to minimize the loss function, which can be written in term of neural network function using the standard back propagation method. This can be written as follows:

$$L(x, x') = \|x - x'\|^2 = \left\|x - \sigma'\left(W'\left(\sigma(Wx + b)\right) + b'\right)\right\|^2$$

Where, x is usually averaged over some input training set.

IV. Results and Discussion

A. Data

We used ORL face database for our experiments. This database is composed of 400 images with the size of 112x92 and 256 grey levels per pixel. There are 40 persons in this database and each person belongs to 10 different facial images which were taken at various times, altering lighting, facial expressions and wearing glasses or not. All the images were taken contrary a dark uniform background with the subjects in a frontal and upright position. Examples of images of the ORL face database are shown in Fig. 3.

Fig. 3. Example of ORL face database after face detection.

B. Experimental setup

First of all, all images data were preprocessed before autoencoder modelling by resizing to 128 x 128 for reducing computational resource. Then we divided all these images into training and testing sets by 75% and 25%, respectively. And then noise was added with noise factors value 0.2 to these training and test sets using NumPy's random normalization function. All images in both training and testing datasets have been clipped between 0 and 1. The results of before adding noise and after adding noise results can be seen in Fig. 4.

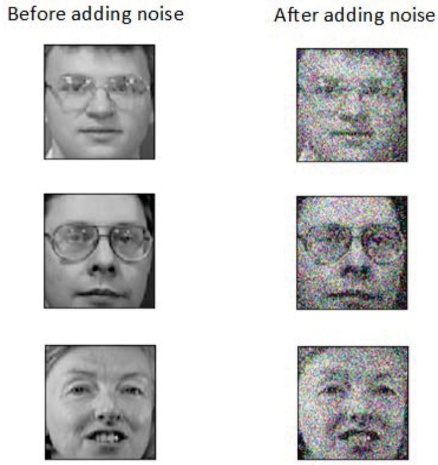

Fig. 4. Before and after conditions of noise adding process

After adding noise process, the autoencoder model was trained on the training and testing datasets. The architecture of autoencoder was composed of 2 2D convolutional layers for encoding part and 2 2D deconvolutional layers for decoding part. The architecture of autoencoder model can be seen in the following figure:

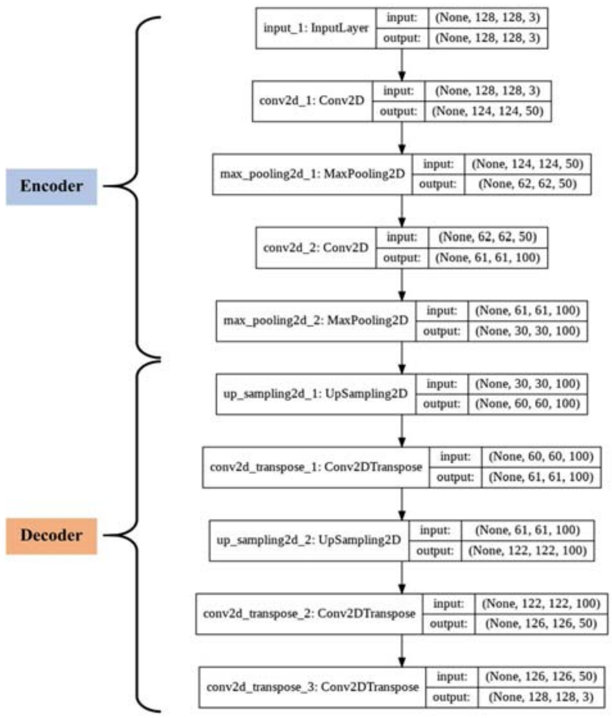

Fig. 5. Architecture of 2D convolutional denoising autoencoder.

Implementation of this autoencoder model was performed using Keras with TensorFlow backend on Google Colab GPU runtime. Losses per epoch in learning process were described in Fig. 6.

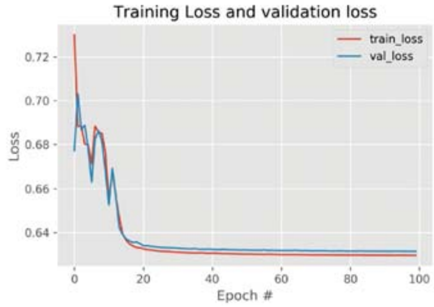

Fig. 6. Training and validation losses per epoch.

Experiments were conducted with three types of noise models (Gaussian, Salt and pepper, Poisson) in this article for the testing of proposed autoencoder model. Different parameters of these noise models are shown in table 1.

TABLE I. DIFFERENT PARAMETERS OF NOISE

Noise models	parameters
Gaussian	$\mu=0,\ \sigma=0.2$
Gaussian	$\mu=0,\ \sigma=0.3$
Gaussian	$\mu=0,\ \sigma=0.4$
Salt & pepper	$\mu=0,\ \sigma^2=0.1$
Salt & pepper	$\mu=0,\ \sigma^2=0.2$
Salt & pepper	$\mu=0,\ \sigma^2=0.3$
Poisson	$\lambda=1.5$
Poisson	$\lambda=1.3$
Poisson	$\lambda=1.1$

μ and σ are the mean of normal distribution and standard deviation of gaussian noise, σ^2 is the variance of Salt & pepper noise, λ is the mean of Poisson noise distribution. The implications of various noise models were presented below:

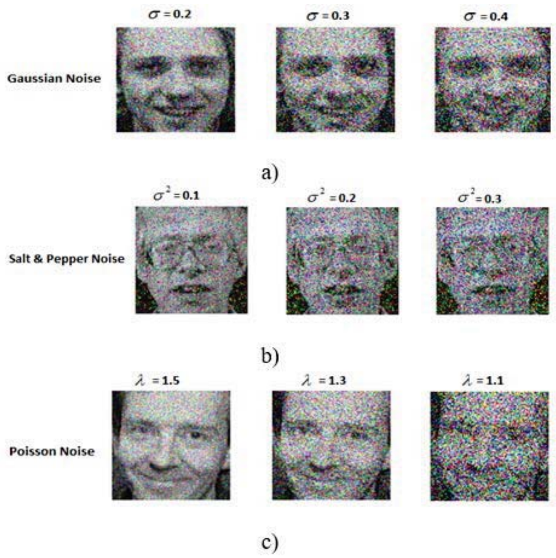

Fig. 7. a) Different values of deviation of gaussian noise, b) different values of variance of salt and pepper noise c) different values of mean of Poisson noise

The performances of convolutional autoencoder network on the different noise parameter values of Gaussian noise, Salt and Pepper noise and Poisson noise can be seen by comparing original, different noisy and denoised images in Fig. 8, Fig. 9 and Fig. 10 respectively.

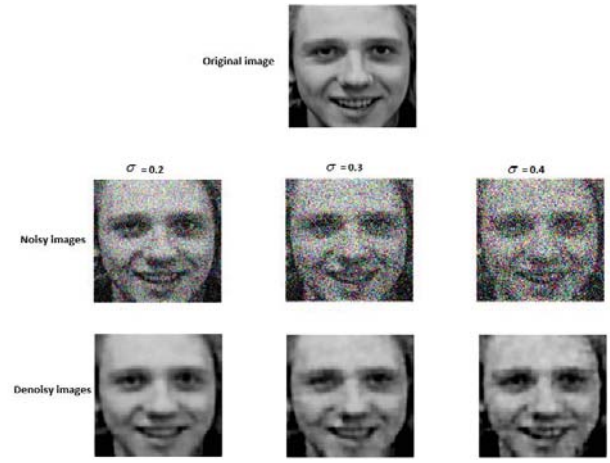

Fig. 8. Performance of convolutional autoencoder's network model on images with Gaussian noisy

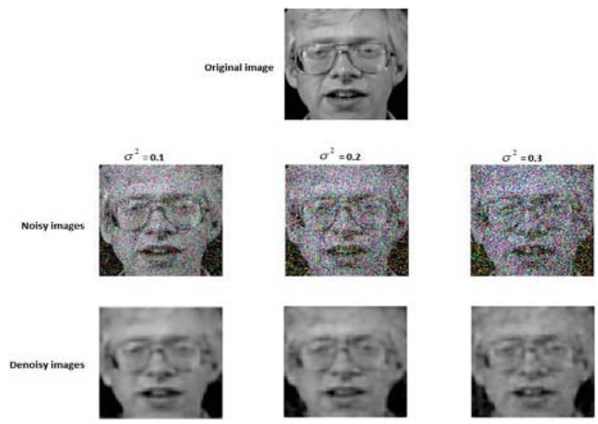

Fig. 9. Performance of convolutional autoencoder's network model on images with Salt and Pepper noisy

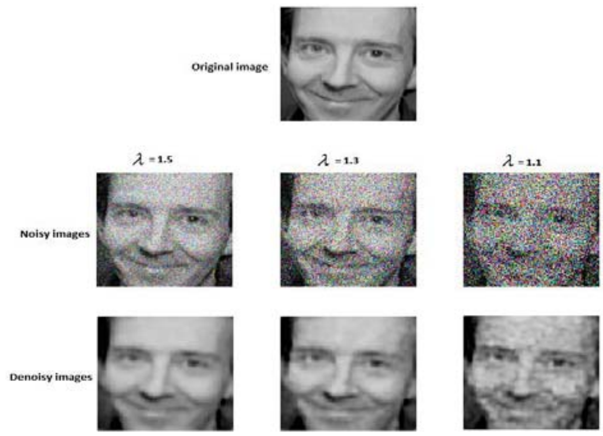

Fig. 10. Performance of convolutional autoencoder's network model on images with Poisson noise

V. Conclusion

In this paper, the experiments were conducted using the convolutional denoising autoencoder network on the ORL database. Different parameter values of Gaussian noise, Salt & Pepper noise and Poisson noise were applied for evaluating autoencoder model. Due to the experimental results, the denoising autoencoder model can be used to eliminate the average noise from facial images in face recognition systems, especially in outdoor situations. In the near future, we will apply this image denoising autoencoder model in outdoor recognition systems in real time.

Acknowledgment

We would like to thank our supervisor Alexander Gavrilov, who helps me a lot during the experiments, and friends who always advise valuable motives.

References

[1] Y. Ding, Y. Cheng, and X. Cheng, “Noise-resistant network: a deep-learning method for face recognition under noise,” Image Video Proc., 43, 2017. [Online]. Available: https://doi.org/10.1186/s13640-017-0188-z

[2] I. Irum, M. A. Shahid, M. Sharif, and M. Raza, “A Review of Image Denoising Methods,” Journal of Engineering Science and Technology Review, 8(5), pp. 41–48, 2015.

[3] Ajay Kumar Boyat and Brijendra Kumar Joshi, “A review paper: noise models in digital image processing,” Signal and Image Processing: An Int. Journal, vol. 6, no. 2, pp. 63–75, 2015.

[4] Naing Min Tun, A. I. Gavrilov, and P. Phyo Paing, “Human Face Recognition Using Combination of ZCA Feature Extraction method and Deep Neural Network,” IEEE Conf. on ICIEAM, 2019.

[5] J. Xu, L. Zhang, and D. Zhang, “External prior guided internal prior learning for real-world noisy image denoising,” IEEE Transactions on Image Processing, 27(6), pp. 2996–3010, 2018.

[6] J. Xu, L. Zhang, and D. Zhang, “Denoising convolutional neural network,” IEEE Int. Conf. on Information and Automation, pp. 1184–1187, 2015.

[7] Z. Zhang, L. Wang, A. Kai, and T. Yamada, “Deep neural network-based bottleneck feature and denoising autoencoder-based dereverberation for distant-talking speaker identification,” EURASIP Journal on Audio, Speech, and Music Processing, (1), 12, 2015.

[8] X. Mao, C. Shen, and Y.-B. Yang, “Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections,” Advances in neural information processing systems, pp. 2802–2810, 2016.

[9] Chunwei Tian, Lunke Fei, Wenxian Zheng, and Yong Xu, “Deep Learning on Image Denoising: An Overview,” Computer Vision and Pattern Recognition, 2020.

[10] P. Vincent, “Extracting and composing robust features with denoising autoencoders,” Proc. of the 25th int. conf., 2008.

[11] V. Jain and S. Seung, “Natural image denoising with convolutional networks,” Advances in Neural Information Processing Systems, pp. 769-776, 2009.

[12] Cho Kyunghyun, ”Boltzmann machines and denoising autoencoders for image denoising,” arXiv preprint arXiv:1301.3468, 2013.

[13] Irfan Ali, Haque Nawaz, S. M. Hassan, and Abdullah Maitlo, “Image Denoising with Color Scheme by Using Autoencoders,” IJCSNS Int. Journal of Computer Science and Network Security, vol. 18, no. 12, 2018.

[14] L. Gondara, “Medical image denoising using convolutional denoising autoencoders,” IEEE 16th Int. Conf. on Data Mining Workshops, 2015.

[15] Donghoon Lee, Sunghoon Choi, and Hee-Joung Kim, “Performance evaluation of image denoising developed using convolutional denoising autoencoders in chest radiography,” Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 884, pp. 97–104, 2018.

[16] Samuel W. Hasinoff, “Photon, Poisson Noise,” Computer Vision, 2014.