IOP Conference Series: Earth and Environmental Science

PAPER • OPEN ACCESS

Rapid classification based on image using residual neural networks

IOP Conf. Series: Earth and Environmental Science **644** (2021) 012026 doi:10.1088/1755-1315/644/1/012026

Rapid classification based on image using residual neural networks

H Setiawan[1]*, R Millinath[2]

[1]Politeknik Aceh Selatan, Department of Computer Engineering, Tapaktuan, Indonesia
[2]Hochschule Mittweida University, Mittweida, German

*Email: herysetiawan88@poltas.ac.id

Abstract. Every human being is created with a certain gender, female or male. Gender differences provide opportunities for mating and producing offspring for regeneration. With the development of the world of information technology, the recognition of gender based on digital photos can be done for the needs of biodata, permission to access public toilets and others. The process of sex classification based on photographs with the use of residual neural networks has been carried out in this study. The process consists of several stages, namely the learning process of the features of each image that has been classified in a male or female class. The next process was carried out by resnet classification of 3,354 pictures (jpg) of men (1414 files) and women (1940 files). The data divided into 2 parts, 80% for training, 20% for testing data. The results of total images of 588 from total available data obtained an accuracy rate of 89.49%.

1 Introduction

The human brain has the ability to recognize and distinguish faces from one person to another relatively quickly and easily. Human face recognition is one of the fields that is quite developed today. The application of facial recognition can be applied in the field of security (security system) such as access to enter the room[2]. One part of face recognition that has been developed at this time is gender recognition. The similarity between gender recognition and face recognition lies in the process of feature extraction[3]. However, it is a little different in the classification process. During this time to count the number of people who are male or female who come to a store or public institution is still manually, so it takes longer. To facilitate what advertisements are displayed on electronic billboards in public places or alongside a road can be adjusted according to the sex of the person who passes the ad. So as to simplify and speed up processing time software for sex recognition is based on facial images.

Difficulty in the process of gender recognition is mainly due to the complexity of the face conditions, such as image position, lighting and expression of different images that have high dimensions so that it must go through the process of compression / extraction before processing the data with the classification method[4]. One can improve the classification performance of CNNs by enriching the diversity and specificity of theseconvolutional filters through deepening the network [5]. effective way to solve these problems suggested in is Residual Networks (ResNets).

The main ResNet paper [6] authors have suggested different configurations of ResNets with 18, 34, 50, 101, and 152 layers. One could describe ResNets as multiple basic blocks that are serially connected to each other and there are also shortcut connections parallel to each basic block and it gets added to its output. This layer has no "Batch Normal" and "ReLU," in other words, the information produced by this layer is the original information after filtering the input image, then it is used to estimate the residual information by feeding back to the middle and the last layer[7].

2 Materials and Methods

2.1 Data sheet

There are only 2673 images consisting of 1418 male and 1912 female files. This is a manually collected and cleaned dataset. From the total data, we divide into 2 parts consisting of 80% for training, 20% for testing. So, instead of doing all the data augmentations offline, in our implementations we used an online version of it. Whenever a new batch arrives, the researcher pass all images in that batch from a random transformation unit. This unit first flips the image horizontally with probability 0:5, and then with some probability p, it randomly crops the image to a 224x224 pixels. Figure 3 shows how this unit works on a sample image as shown in Figure 1.

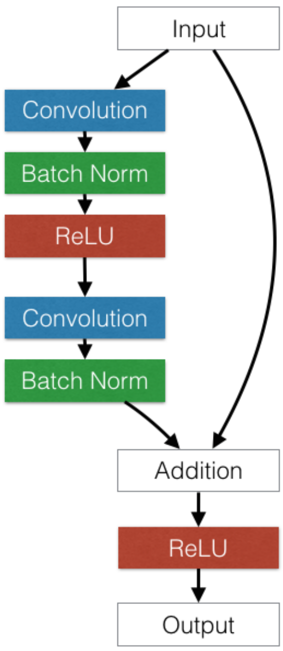

Figure 1. RestNet basic block

2.2 Network Design

The ResNet model introduced in [4] is our starting point for the network design. This model is specifically designed for images in ImageNet and accepts images with size 224ˇ 224 and classifies them in 2 categories, male or female class and then give it to the training model, or just skipping the first layer and insert the original image as the input of the second convolutional layer, and then fine tuning a few of the last layers to get higher accuracy. However, since in this research were interested in comparing ResNet34 models with their equivalent ConvNets, we had to design and train our models from scratch (although, we might get worse accuracies because of lack of computational resources). In this section we are using Residual Network with 34 hidden units, all we have to do is to add a layer at the end on the residual network to transform the dimension of the residual network to the required output. In our case, it is to the 2 possible outputs. Training using 4 epochs (4 cycles through all our data) and use the lr_find method to find the optimum learning rate.

3 Result and Discussion
From the experimental results, a typical graph is chosen a higher learning rate for which the loss is minimal. The higher learning rate makes sure that the machine ends up learning faster. The result shown in figure 2.

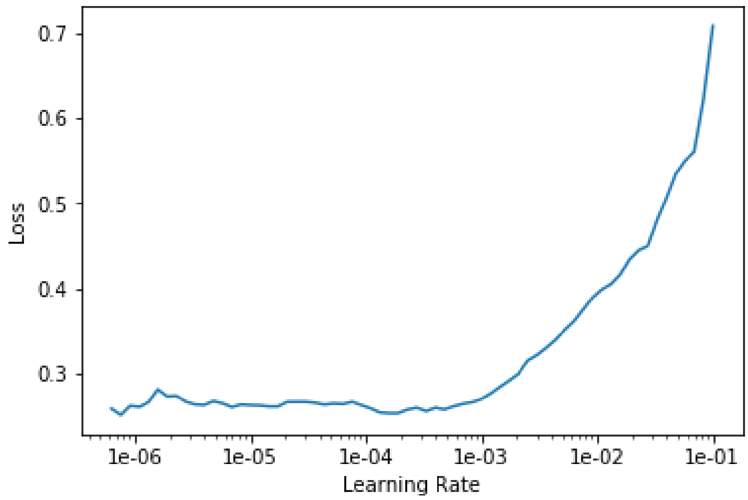

Figure 2. Optimum learning rate.

From around slice 1e-5, 1e-4 mark, we have an optimum learning rate. Now that the optimum learning rate. We will first see which were the categories that the model most confused with one another. We will try to see if what the model predicted was reasonable or not. In this case the errors look reasonable (none of the errors seems obviously naive). This is an indicator that our classifier is working correctly. From the prediction results using resnet34, there is the greatest loss value of 11.26. as shown in Figure 3.

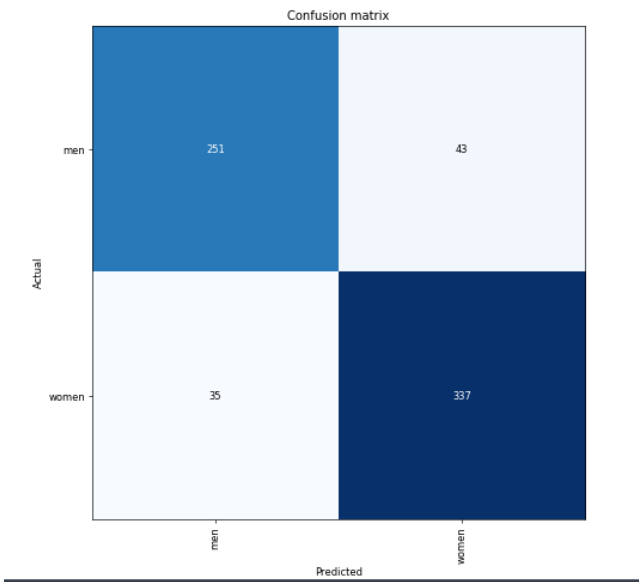

Figure 3. Confusion Matrix.

Furthermore, when we plot the confusion matrix, we can see that the distribution is heavily skewed: the model makes the same mistakes over and over again but it rarely confuses other categories. This suggests that it just finds it difficult to distinguish some specific categories between each other; this is normal behavior. From 627 samples of testing data we have tested. The results obtained are shown in Figure 6. As many as 285 male images and 372 female images, the classification achieved an accuracy of 89.50%.

4. Conclusion

In this paper, we have designed and trained a deep residual network 34 layers for rapid classification based on images. From the results of experiments that have been carried out, the accuracy reaches 89.50%.

References

[1] R. Hamzah, M. Sciences, and U. S. Alam, "INVESTIGATION OF SPEECH DISFLUENCIES CLASSIFICATION ON DIFFERENT THRESHOLD SELECTION TECHNIQUES USING ENERGY FEATURE," *MALAYSIAN J. Comput.*, vol. 4, no. 1, pp. 178–192, 2019.
[2] F. Damayanti, "Pengenalan Jenis Kelamin Berdasarkan Citra Wajah Menggunakan Metode Two-Dimensional Linear Discriminant Analysis," *Konf. Nas. Sist. Inform.*, pp. 1–6, 2015.
[3] N. Muhammad *et al.*, "Conceptual Framework for Lightweight Ciphertext Policy-Attribute Based Encryption," *MALAYSIAN J. Comput.*, vol. 4, no. 1, pp. 237–245, 2019.
[4] H. Setiawan and E. M. Yuniarno, "Biometric Recognition Based on Palm Vein Image Using Learning Vector Quantization," *2017 5th Int. Conf. Instrumentation, Commun. Inf. Technol. Biomed. Eng.*, pp. 95–99, 2017, doi: 10.1109/ICICI-BME.2017.8537770.
[5] M. S. Ebrahimi and H. K. Abadi, "Study of Residual Networks for Image Recognition," pp. 1–6, 2018, [Online]. Available: http://arxiv.org/abs/1805.00325.
[6] G. A. Kaiming,Hembury, V. V. Borovkov, J. M. Lintuluoto, and Y. Inoue, "Deep Residual Learning for Image Recognition," *Chem. Lett.*, vol. 32, no. 5, pp. 428–429, 2003, doi: 10.1246/cl.2003.428.
[7] Z. Guo, Y. Sun, M. Jian, and X. Zhang, "Deep residual network with sparse feedback for image restoration," *Appl. Sci.*, vol. 8, no. 12, 2018, doi: 10.3390/app8122417.
[8] "Men/Women Classification | Kaggle." https://www.kaggle.com/playlist/men-women-classification (accessed Sep. 02, 2019).