IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 2, July 2023

Performance Evaluation of MobileNetV2 CNN Architecture in Localized Datasets

Shem L. Gonzales[1] and Rosemarie C. Arcaya[2]

Surigao del Norte State University, Surigao City, Philippines

slgonzales@ssct.edu.ph[1] and rarcaya@ssct.edu.ph[2]

Abstract: *Facial features of humans are part of physiological characteristics that are the basis for identity verification. Face Recognition technology has expanded into a more thorough process to increase the accuracy rate in recognizing individuals. This paper assessed the performance of face recognition in Deep Convolutional Neural Networks with the application of localized datasets. Five (5) people were used as a class, with thirty (30) images per person. A total of 150 images were gathered from photo albums and collections. The images were preprocessed with some of the basic image processing techniques, including histogram equalization cropping, and resizing, before training using MobileNetv2 Pre-trained CNN architecture. Earlier layers of this architecture were used as feature extractors. The final 2-3 layers were fine-tuned following the number of classes. During the training, images were divided into 80% for training datasets, and 20% for testing and validation datasets. The graphical illustration showed an accuracy rate of 90% as well as the generation of a confusion matrix. The results indicate that MobileNetV2 is a promising CNN architecture that can be used in Face Recognition Technology with localized datasets*

Keywords: Face Recognition, Deep Learning, Image Processing, MobileNetV2

I. INTRODUCTION

Image Processing is the new method of extracting important information from the picture which can be useful for recognition and verification [1][2]. Through state-of-the-art of Deep Learning (DL), this technology brought substantial evidence of an exponential improvement for convolutional neural network models [3] in classification problems. When an image of a person is taken, the face is the most exposed part of the human body. This gives an advantage for biometric technology to use face images since it carries a huge amount of information for individual identification [4]. In face recognition, it has been applied to thoroughly examine, process, and abstract physiological values which are factors for human identity recognition. These values have been fed to a machine learning or deep learning process for the development of a face recognition software framework. [5].

Different techniques were used for face recognition development. [6] designed a real-time face recognition system using CNN. In the study [7], they mentioned Principal Component Analysis, Artificial Neural Networks, graph matching, and Hidden Markov Models as means for face recognition systems [8][9][10]. Some researchers also applied the concepts of transfer learning in pre-trained models for easier application and training reduction [11][12] [13]. All of these methods showed unquestionable performance as to the values in terms of accuracy, specificity, and efficiency.

All of the face recognition processes require datasets that support the training. These datasets are divided into training, testing, and validation. Most of the studies used images downloaded from online database providers. In this study, fewer images with five (5) local personalities as classes were used from their photo album collection and on Facebook. The images of each class were composed of 30 images per class with a total number of 150 images. These images are processed using transfer learning and fine-tuning method.

MobileNetV2 was selected due to the fewer parameter and smaller computational complexity advantage [14][15][16][17]. The architecture was fine-tuned to its best accomplishment. The performance of the trained model was observed from the learning and loss curves and the generation of the confusion matrix after the training process. The model was tested with the function created to evaluate the recognition percentage accuracy.

The goal of this study is to determine the performance of MobileNetV2 architecture using photo collections from local personalities. Specifically, it aimed to:

- Generate the learning and loss curves from the training process;
- Plot the confusion matrix; and
- Determine the accuracy of the network.

Figure 1 illustrates the stages of the development of face recognition using MobileNetv2 CNN architecture. The first stage is the collection of images from selected local personalities. The images were preprocessed following the specific dimension needed by the network. In the first stage, the images were split into training and testing. 80% of the images are used for the training dataset, and the rest are for the testing dataset. Feature extraction was then performed by architecture with its built-in layers. The fully connected layer was replaced by a new fully connected layer with 5 classes. The last layer was also replaced by a new output layer. After the network performed the training process, it generated a training model. The last stage is the performance evaluation. This stage initiated the learning curves as well as the plotting of the confusion matrix

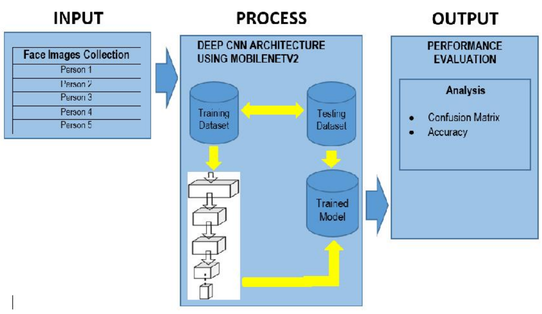

Figure 1. Conceptual Framework

II. METHODOLOGY

2.1 General Approach

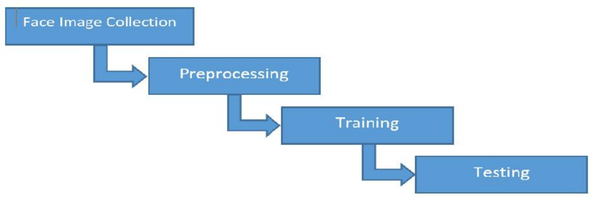

Figure 2. Implementation Flow

2.2 Face Image Collection

Figure 3. Samples of acquire face images

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, July 2023

Sample images will be collected from five (5) different local sources. These selected/identified local personalities served as classes with thirty (30) face images per class for a total of 150 images. The collected images were used as the dataset needed for the training. The images were taken from a photo album collection and the web (Facebook) in JPEG image file format. The datasets were split into 2. 80% of the images were used for training, and 20% were used for testing. Training datasets were fed to MobileNetv2 CNN architecture to obtain the model.

2.3 Image Preprocessing

The collected images came in different dimensions. Initially, these images were cropped to isolate and focus only on the face section of the image using a photo editor. The cropped face images were resized to 224x224 square pixels needed for training in the CNN architecture. The images were applied by an image augmenter to help improve the performance of the network. [19].

2.4 MobileNetV2 Pre-trained CNN Architecture

With less operations and memory use while maintaining or improving accuracy performance compared to existing Deep CNN architectures, MobileNetv2 was created to support the development of mobile vision applications. [16][17]. Despite using deep separable convolution in a similar manner to MobileNetv1, this new architecture added two new features: linear bottlenecks between layers, and shortcut connections between bottlenecks. [18] as shown in figure 4.

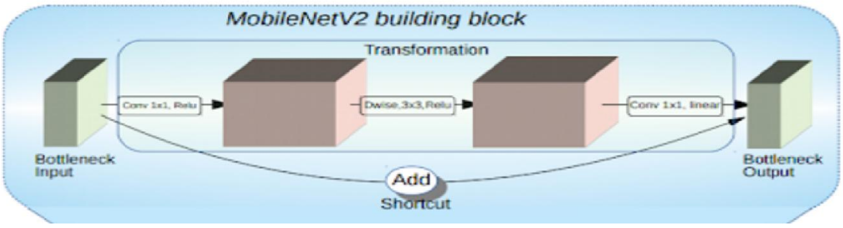

Figure 4. MobileNetV2 block illustration
(Sandler et al., 2018)

2.5 Fine Tuning

Table 1. Parameter configuration in Matlab.

Parameters	Configuration Settings
Minimum batch size	5
Maximum epoch	80
Initial learn rate	.00001
Image shuffling	every-epoch
Learn rate schedule	Piecewise
Verbose	False
Training function	Adam
Weight learn rate factor	20
Bias learn rate factor	20

Table 1. Parameter configuration in Matlab.

III. RESULTS AND DISCUSSIONS

3.1 Training and Loss Curves

Figure 5 presented the monitoring of the training process. This development demonstrated the network's improvement at each iteration. Following the configuration values corresponding to the parameter name in table 1, the figure indicated the network's improvement as it iterates to every epoch [20]. Figure 6 shows the generated confusion matrix indicating an accuracy of 90%.

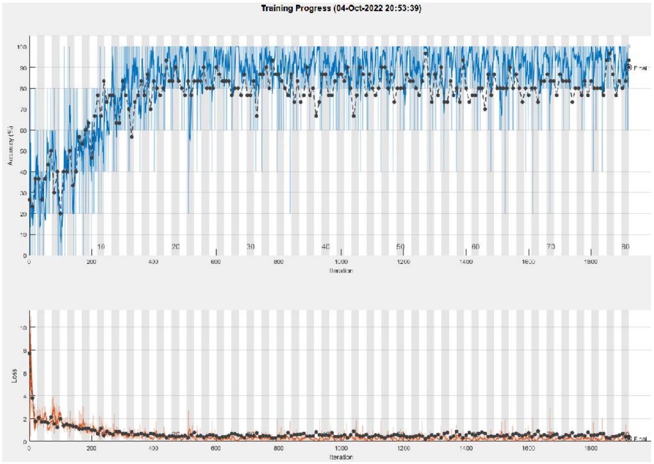

Figure 5. Learning and loss curves during the training

3.2 Confusion Matrix

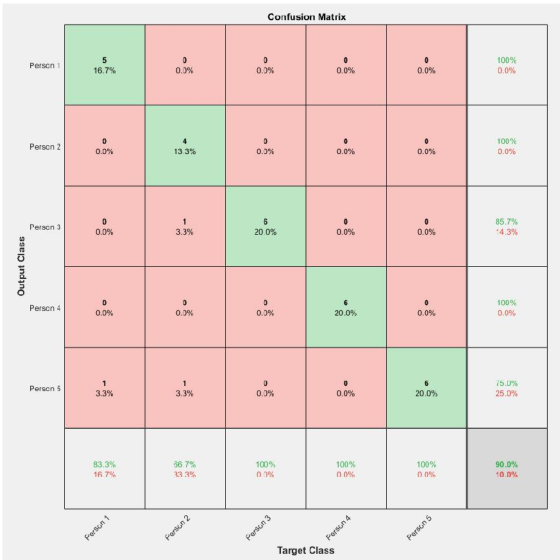

Figure 6. Generated confusion matrix after the training.

3.3 Evaluation

Figure 7. Recognition output of MobileNetV2.

Figure 7 demonstrated how successfully MobileNetV2 was able to identify photos from the test datasets. Even though several photographs had low accuracy predictions, the precise person was nonetheless recognized.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, July 2023

IV. CONCLUSION

This paper examined the performance of MobileNetv2 Convolutional Neural Network architecture with images from local personalities using Matlab. The evaluation showed that the model resulted in a 90% accuracy coming from generated training and confusion matrix. Although some of the images in the datasets have low resolution, images from test datasets were still recognized and showed the best percentage accuracy.

REFERENCES

[1]. Silva, E., & Mendonça, G. (2005). Digital Image Processing. The Electrical Engineering Handbook. Pages 891-910. https://doi.org/10.1016/B978-012170960-0/50064-5.

[2]. Jin, K.H., McCann, M.T., Froustey, E. & Unser, M. (2017) Deep Convolutional Neural Network for Inverse Problems in Imaging. IEEE Transactions on Image Processing, vol. 26, no. 9, pp. 4509-4522. https://doi.org/10.1109/TIP.2017.2713099.

[3]. Choudhary, K., DeCost, B., Chen, C. et al. (2022). Recent advances and applications of deep learning methods in materials science. npj Comput Mater 8, 59. https://doi.org/10.1038/s41524-022-00734-6

[4]. Boussaad, L., & Boucetta, A. (2022). An effective component-based age-invariant face recognition using Discriminant Correlation Analysis. J. King Saud Univ. Comput. Inf. Sci., 34, 1739-1747.J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, Jan. 1959.

[5]. Khan, M., Chakraborty, Astya S., & Khepra R. (2019). Face Detection and Recognition Using OpenCV,"2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS). pp. 116-119,

[6]. PranavK, B., & Manikandan, J. (2020). Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks. Procedia Computer Science, 171, 1651-1659. https://doi.org/ 10.1109/ICCCIS48478.2019.8974493

[7]. Dhawle T., Ukey, U., & Choudante, R. (2020). Face Detection and Recognition using OpenCV and Python. International Research Journal of Engineering and Technology

[8]. Yan, K., Huang, S., Song, Y. Lium, W., & Fan, N. (2017) Face recognition based on convolution neural network. 36th Chinese Control Conference (CCC), pp. 4077-4081, https://doi.org/10.23919/ChiCC.2017.8027997.

[9]. Mahalingam, G. & Kambhamettu, C. (2011). Video Based Face Recognition Using Graph Matching. In: Kimmel, R., Klette, R., Sugimoto, A. (eds) Computer Vision – ACCV 2010. ACCV 2010. Lecture Notes in Computer Science, vol. 6494. https://doi.org/10.1007/978-3-642-19318-7_7

[10]. Biswas, D., Jayan, S., Nadipalli, S. S. B., & S. R. (2022) Attendance Tracking with Face Recognition Through Hidden Markov Models. International Conference on Electronics and Renewable Systems (ICEARS), pp. 1640-1645. https://doi.org/10.1109/ICEARS53579.2022.9751888.

[11]. Wang, Q., & Guo, G. (2019). Benchmarking Deep Learning Techniques for Face Recognition. Journal of Visual Communication and Image Representation. Volume 65. https://doi.org/10.1016/j.jvcir.2019.102663

[12]. Khan, S., Ahmed, E., Javed, M.H., & Shah, Syed, A.S. (2019). Transfer Learning of a Neural Network Using Deep Learning to Perform Face Recognition. 1-5. http://doi.org/10.1109/ICECCE47252.2019.8940754

[13]. Luttrell, J., Zhou, Z., Zhang, C., Gong, P., & Zhang, Y. (2017) Facial Recognition via Transfer Learning: Fine-Tuning Keras_vggface. International Conference on Computational Science and Computational Intelligence (CSCI), pp. 576-579, https://doi.org/10.1109/CSCI.2017.98.

[14]. Huang, M.L., Liao, Y.C., (2022) A Lightweight CNN-Based Network on COVID-19 Detection using X-ray and CT Images. Computers in Biology and Medicine, Vol. 146, https://doi.org/10.1016/j.compbiomed.2022.105604.

[15]. Kamarudin, M.H. & Ismail Z.H. (2022). Lightweight Deep CNN Models for Identifying Drought Stressed Plant. The 9th AUN/SEED-Net Regional Conference on Natural Disaster. doi:10.1088/1755-1315/1091/1/012043

[16]. Sandler, M., et al., (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. Computer Vision and Pattern Recognition. https://doi.org/10.48550/arXiv.1801.04381

[17]. Dong, K., et al., (2020).MobileNEtV2 Model for Image Classiifcation. International Conference on Information Technology and Computer Application. https:/doi.org/10.1109/ITCA52113.2020.00106

[18]. Prasetyo, E., Purbaningtyas, E., Adityo, R.D., Suciati, N., & Fatichah, C. (2022). Combining MobileNetV1 and Depthwise Separable Convolution Bottleneck with Expansion for Classifying the Freshness of Fish Eyes. Information Processing in Agriculture. Vol. 9, pp 485-496.https://doi.org/10.1016/j.inpa.2022.01.002.

[19]. Shorten, C., Khoshgoftaar, T.M., (2019) A survey on Image Data Augmentation for Deep Learning. J Big Data. https://doi.org/10.1186/s40537-019-0197-0

[20]. Monitor Custom Training Loop Progress. Retrieved from https://www.Mathworks.com/Help/Deeplearning/ug/Monitor-Custom-Training-Loop-Progress.Html.

[21]. C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, “Rotation, scale, and translation resilient public watermarking for images,” IEEE Trans. Image Process., vol. 10, no. 5, pp. 767-782, May 2001.