International Journal of Application on Sciences, Technology and Engineering
(IJASTE)
Volume 1, Issue 4, 2023. ISSN:2987-2499

RECOGNITION OF WORKOUT EXERCISE BASED ON IMAGE PROCESSING USING CNN MOBILENETV2 AND EFFICIENTNETB3

Andrew Hendisutio[1], Meirista Wulandari, Wahidin Wahab

[1] Faculty of Engineering, Electrical and Electronics Engineering, Tarumanagara University
Email : andrew.525200011@stu.untar.ac.id
[2] Faculty of Engineering, Electrical and Electronics Engineering, Tarumanagara University
Email : [2] meiristaw@ft.untar.ac.id
[3] Faculty of Engineering, Electrical and Electronics Engineering, Tarumanagara University
Email : [3] wahidinwahab@ft.untar.ac.id

Submitted: 26-09-2023, Revised: 27-10-2023, Accepted: 08-12-2023

ABSTRACT

Exercise is crucial for maintaining a healthy and fit body, yet many individuals struggle to track their workout progress effectively. This paper explores the application of deep learning techniques, specifically Convolutional Neural Networks (CNNs), to recognize and classify common exercise moves such as push-ups, pull-ups, sit-ups, leg raises, and squats. The goal is to provide individuals with a tool to assist in counting repetitions and sets, thereby enhancing their exercise experience. The study employs two well-known CNN models, MobileNetV2 and EfficientNetB3, trained on a custom dataset consisting of 2,500 exercise move images. The dataset includes various training-to-testing data ratios, ranging from 90:10 to 50:50. The models are evaluated based on their accuracy in classifying exercise moves, and confusion matrices are generated to analyze their performance further.

1. Introduction

Exercise is an important habit for having a good, healthy and fit body. Exercise is part of lifestyle habits that could change unhealthy lifestyle into healthy lifestyle. According a research analysis done by UNICEF in 2018, 20% school-aged children, 14.8% adolescents, and 35.5% adults in Indonesia were living with overweight lifestyle [1]. To reduce the percentage of overweight person, people participate in physical activities, such as normal workout, going to the gym, playing sports, calisthenics, etc. Exercise workout such as push-up, pull-up, sit-up, and squat have been part of many fitness programs to achieve the said body goals.

To achieve a beneficial workout, each person requires to have consistency in mind for each exercise workout. The benefit can be felt with *hypertrophy* in the trained muscle. Training program need to be in 3 – 4 sets from the range of 6 – 15 repetitions of exercise move (can be push-up, sit-up, pull-up, squats, etc.)[2]. After exercising for a while, keeping up for counting repetitions can be hard, so using another method to keep count for each repetition and set could make each person have a better exercise overall performance.

Keeping count for exercise is uncomfortable, therefore each person doing the exercise need help or assistance to have a comfortable workout. There is already a tool such as smart watch to see how far a person has run in a period of time and measuring how fast the heart beats after exercising. There hasn’t been enough assistance in helping for individual exercises moves, so using technology could create smart tool for assisting individuals who have hard time to motivate themselves after a while of exercising.

Modern technologies could play an important role in providing an alternative method to keep count for exercising. For example, a research for making an accelerometer sensor to recognize and count repetition for workout type using CNNs [3], another using computer vision to count repetition of push-up [4]. Using computer vision is popular in recognizing action for sports and the data is from realistic sports videos [5][6]. The data obtained would help individuals who are training or exercising to achieve better comfortability in each exercise.

After reviewing some research works, computer vision could be used for AI's deep learning. Trained CNN Model could recognize exercise moves (specially push-up, sit-up, pull-up, leg raises, and squats) after deep learning from custom dataset that have been prepared. The model used to recognize exercises moves are MobileNetV2 and EfficientNetB3 [7][8]. The two model are specialized to perform in image classification and could be utilized for classifying exercise moves. In the following chapter, the two models are going to be compared for optimal output of classifying exercises.

2. Theory

2.1. Exercise and Achieving Hypertrophy

Exercise is a physical activity performed by individuals to improve or maintain a good and healthy body. Hypertrophy is the definition of the increasing fiber or whole muscle, and is due to increase the size of pre-existing muscle fibers [9]. Achieving hypertrophy, in the context of muscle training and resistance training, involves in increasing the size and strength of muscle fibers [2]. To achieve hypertrophy, individuals are required to do a consistent and structured exercise, including a specific number of sets (groups of repetitions) and repetitions (a number of times an exercise is performed within each set) [2, p. 2]. Exercise moves included in this research are one of the most basics moves that can be implemented in many exercise program, such as push-up, pull-up, leg raises, russian twist, and squats.

(a)

(b)

(c)

(d)

(e)

Figure 1 Images from Kaggle Dataset includes (a)Push-up, (b)Pull-up, (c)Squats, (d)Leg Raises, and (e)Russian Twist [10]

2.2. Image Processing
Image Processing is a field that mixes computer science and engineering, manipulating digital image and extracting information from those images, or enhancing the quality of visual they give. Techniques used in the research such as enhancement, segmentation, and recognition. The purpose of using image processing is to improve data quality given to model, giving higher detail and so making higher precision in the image recognition overall [11].

2.3. MobileNetV2
MobileNetV2 is a CNN (Convolutional Neural Network) architecture designed for efficient and lightweight deep learning applications [12], especially on mobile and embedded devices. MobileNetV2's architecture includes depth-wise separable convolutions, which reduce the number of parameters and computations, making it suitable for real-time image analysis and classification.

2.4. EfficientNetB3
EfficientNetB3 is part of EfficientNet family of CNN Models known for their remarkable performance in image classification tasks. EfficientNetB3 is known to achieved much better accuracy and efficiency that other models from Convolution Neural Network [8]. It represents scale of model complexity, with larger versions such as B3 having more parameters and higher accuracy.

2.5. Deep Learning
Deep learning is a subset of machine learning that focuses on training artificial neural networks with multiple layers (deep neural networks) to perform tasks such as image and speech recognition, natural language processing, and more. Deep learning models, like CNNs, uses custom datasets that the user inputs to it, automatically learning features from the custom dataset.

3. Method
The block diagram of recognition and classification for exercise moves is displayed in **Figure 2**. The block includes input, the processing, and output. The input is the exercise image of moves. The process is classifying the moves from each other. It used two models from CNNs family to be compared: MobileNetV2 and EfficientNetB3. The program produced the accuracy value for each moves as the output. The accuracy value from the two methods was compared to each other. The process was simulated by using Pycharm in Python language and Intel Processor Core i3.

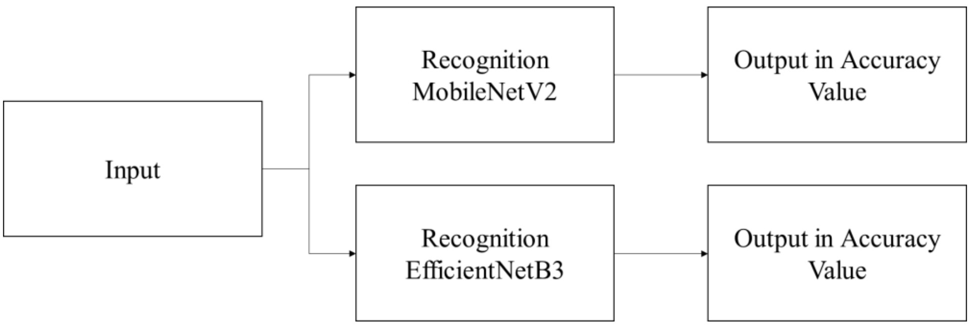


Figure 2 block diagram of recognition for workout exercise

4. Results

In this research, the data obtained to simulate the model classifying exercise moves is from Kaggle, a website and community dedicated to deep learning and giving datasets. The dataset is split into 5 moves; push-up, pull-up, russian twist, leg raises, and squats. Each classes contains 500 images, totalling about 2500 images used for training, validating, and testing the model [10]. The ratio for the research is configured from 90:10, 80:20, 70:30, 60:40, to 50:50.

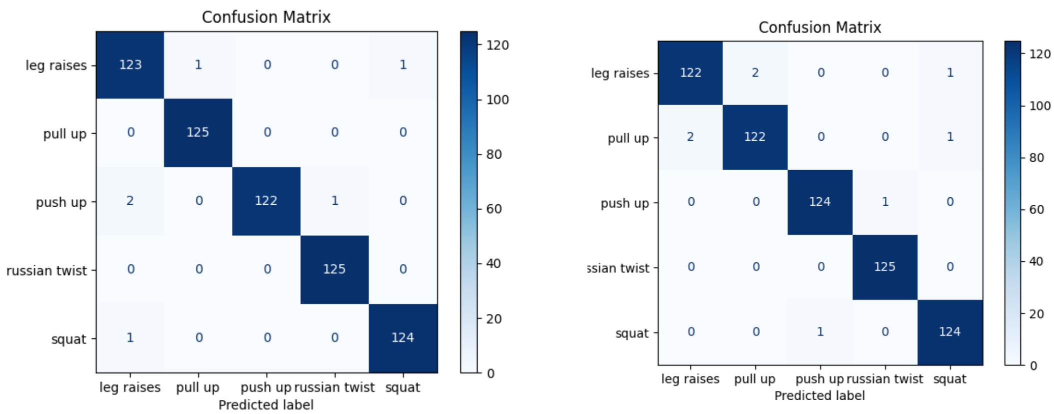
Table 1 shows the accuracy value from MobileNetV2 and EfficientNetB3.

Training Data : Test Data	Accuracy Value for MobileNetV2	Accuracy Value for EfficientNetB3
50 : 50	99.04%	98.72%
60 : 40	98.8%	98.8%
70 : 30	98.4%	98.67%
80 : 20	98.8%	97.6%
90 : 10	99.2%	97.6%

Table 1 Result of training and testing from both models

Based on

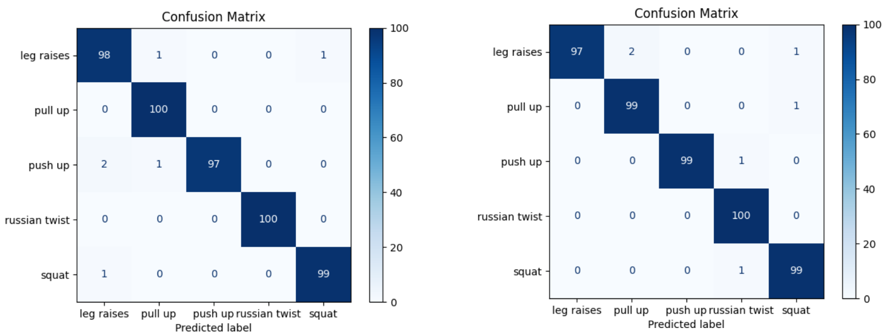
Table 1 the accuracy value output from MobileNetV2 is a bit higher compared to EfficientNetB3, with the highest accuracy from MobileNetV2 in 99.2% with 90 : 10 ratio of training data to test data, and highest accuracy from EfficientNetB3 in 98.8% with 60 : 40 ratio of training data and test data. Furthermore, there will be confusion matrices to see each result for classifying exercise workout based on data comparison of training data and testing data.



(a) (b)

Figure 3(a) confusion matrix of 50 : 50 for MobileNetV2 and (b)confusion matrix for EfficientNetB3

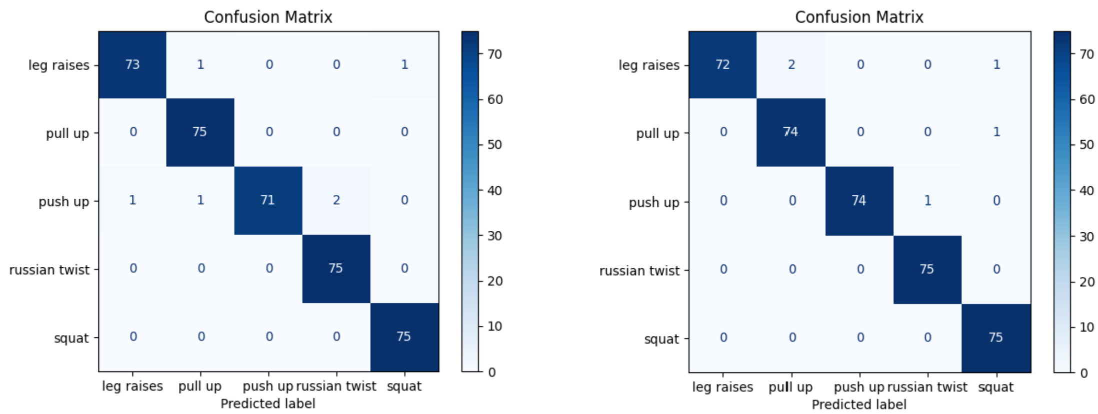
For 50:50 ratio of training data and testing data, the classification for the five exercise moves is especially good for the move russian twist, as both model is capable to recognize and classify the move for 125 images. The recognition and classifying for leg raises are the hardest for both model with 123 images being detected for MobileNetV2 and 122 images for EfficientNetB3.



(a) (b)

Figure 4(a) confusion matrix of 60 : 40 for MobileNetV2 and (b)confusion matrix for EfficientNetB3

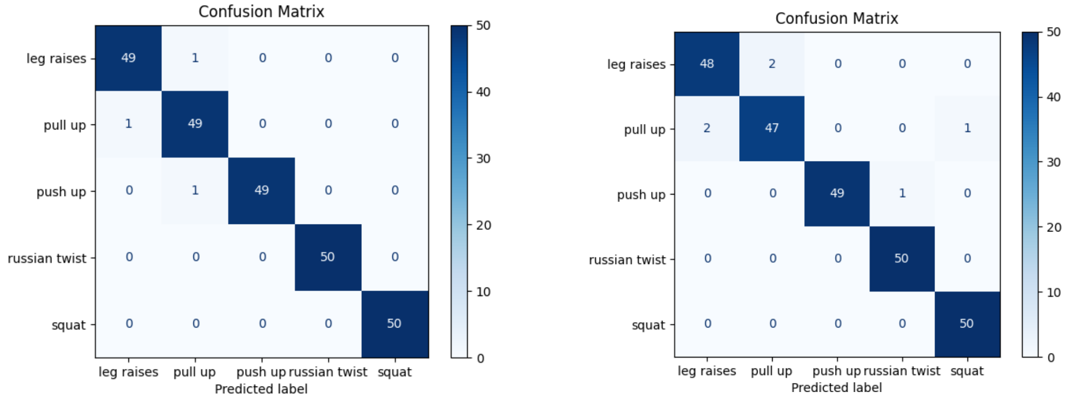
For 60:40 ratio of training data and testing data, the classification for the five exercise moves is especially good for the move russian twist, as both model is capable to recognize and classify the move for 100 images. The recognition and classifying for leg raises are the hardest for both model with 98 images being detected for MobileNetV2 and 97 images for EfficientNetB3.



(a) (b)

Figure 5 (a) confusion matrix of 70 : 30 for MobileNetV2 and (b)confusion matrix for EfficientNetB3

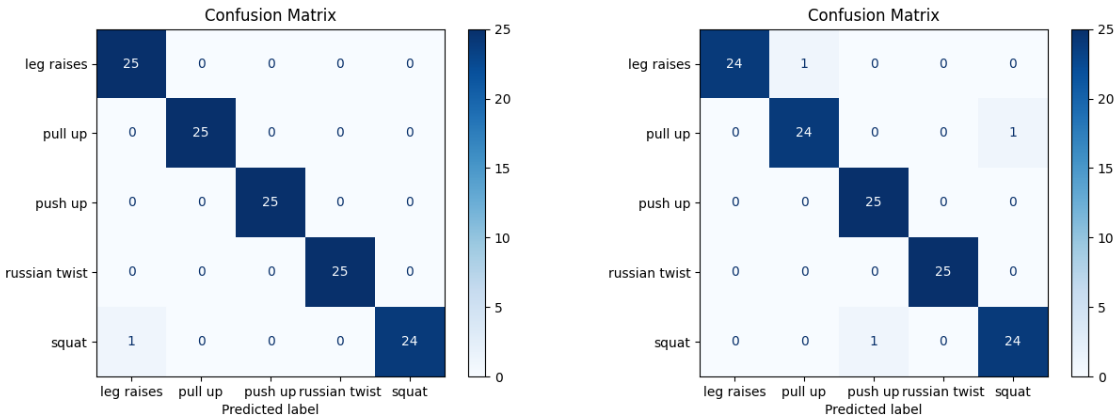
For 70:30 ratio of training data and testing data, the classification for the five exercise moves is especially good for the move russian twist and squat, as both model is capable to recognize and classify the move for 75 images. The recognition and classifying for leg raises and push-up are the hardest for both model with 73 images for leg raises; 71 images for push-up being detected for MobileNetV2 and 72 images for leg raises; 74 images for push-up for EfficientNetB3.



(a) (b)

Figure 6 (a) confusion matrix of 80 : 20 for MobileNetV2 and (b)confusion matrix for EfficientNetB3

For 80:20 ratio of training data and testing data, the classification for the 5 workout exercise is especially good once again for the move russian twist and squat moves, as both model is capable to recognize and classify the move for 50 images. The recognition and classifying for push up is now the hardest for both model with 49 images being detected for MobileNetV2 and 47 images for EfficientNetB3.



(a) (b)

Figure 7 (a) confusion matrix of 90 : 10 for MobileNetV2 and (b)confusion matrix for EfficientNetB3

Lastly, for 90:10 ratio of training data and testing data, the classification for the 5 exercise move is especially good for the move push-up and russian twist, as both model is capable to recognize and classify the move for 25 images. The recognition and classifying for squat are the hardest for both model with 24 images being detected for MobileNetV2 and 24 images for EfficientNetB3.

5. Conclusion

Based on both results, it can be seen that the model MobileNetV2 and EfficientNetB3 can recognize workout exercise with accuracy value above 90%. Also, the hardest moves that can be recognized would be leg raises and push-up, and would also be put in to consideration for the next research as the image taken is clear for input data, and background for input should be clear. Also, the accuracy value from MobileNetV2 obtained the highest accuracy in 90:10 ratio of training data to testing data with accuracy percentage of 99.2%. It can be determined that more training data, the result would be much more satisfying. In conclusion, deep learning using MobileNetV2 and EfficienNetB3 for recognizing and classifying workout exercise certainly could be implemented to have a better experience for individuals doing physical activity and better information for training later on. It could be noted that this research could be more developed for others research such as real-time tools and many more.

References

[1] UNICEF, "Landscape Analysis of Overweight and Obesity in Indonesia," 2021. Accessed: Sep. 25, 2023. [Online]. Available: https://www.unicef.org/indonesia/media/15481/file/Landscape%20analysis%20of%20overweight%20and%20obesity%20in%20Indonesia.pdf

[2] V. M. Iversen, M. Norum, B. J. Schoenfeld, and M. S. Fimland, "No Time to Lift? Designing Time-Efficient Training Programs for Strength and Hypertrophy: A Narrative Review," *Sports Medicine*, vol. 51, no. 10. Springer Science and Business Media Deutschland GmbH, pp. 2079–2095, Oct. 01, 2021. doi: 10.1007/s40279-021-01490-1.

[3] K. Skawinski, F. M. Roca, and R. Dieter, "Workout Type Recognition and Repetition Counting with CNNs from 3D Acceleration Sensed on the Chest."

[4] R. Arlin and R. Munir, "The Development of Push Up Counter Android Application with Computer Vision."

[5] K. Soomro and A. R. Zamir, "Action recognition in realistic sports videos," *Advances in Computer Vision and Pattern Recognition*, vol. 71, pp. 181–208, 2014, doi: 10.1007/978-3-319-09396-3_9.

[6] G. Li and C. Zhang, "Automatic detection technology of sports athletes based on image recognition technology," *EURASIP J Image Video Process*, vol. 2019, no. 1, Dec. 2019, doi: 10.1186/s13640-019-0415-x.

[7] Evan, M. Wulandari, and E. Syamsudin, "Recognition of pedestrian traffic light using tensorflow and SSD MobileNet V2," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Dec. 2020. doi: 10.1088/1757-899X/1007/1/012022.

[8] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," May 2019, [Online]. Available: http://arxiv.org/abs/1905.11946

[9] B. Schoenfeld *et al.*, "Resistance Training Recommendations to Maximize Muscle Hypertrophy in an Athletic Population: Position Stand of the IUSCA," *International Journal of Strength and Conditioning*, vol. 1, no. 1, Aug. 2021, doi: 10.47206/ijsc.v1i1.81.

[10] Hasyim Abdillah, "Workout/Exercise Images." Accessed: Sep. 25, 2023. [Online]. Available: https://www.kaggle.com/datasets/hasyimabdillah/workoutexercises-images

[11] V. Tyagi, *Understanding Digital Image Processing*. CRC Press, 2018. doi: 10.1201/9781315123905.

[12] R. Indraswari, R. Rokhana, and W. Herulambang, "Melanoma image classification based on MobileNetV2 network," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 198–207. doi: 10.1016/j.procs.2021.12.132.