2. Taekwondo Poomsae-3 Movement Identification by using CNN

by Novie Theresia Br. Pasaribu, Erwani Merry, Kalya Icasia, Jordan Eliezer

Taekwondo Poomsae-3 Movement Identification by using CNN

Novie Theresia Br. Pasaribu[1][a], Erwani Merry[1][b], Kalya Icasia[1][c], Jordan Eliezer[1][d],
Che-Wei Lin[2][e] and Febryan Setiawan[2][f]

[1]*Bachelor Program in Electrical Engineering, Maranatha Christian University, Jl. Surya Sumantri 65, Bandung, Indonesia*
[2]*Department of Biomedical Engineering, College of Engineering, National Cheng Kung University, Tainan 701, Taiwan*

novie.theresia@eng.maranatha.edu, erwani.ms@eng.maranatha.edu, 1722002@eng.maranatha.edu, 1722004@eng.maranatha.edu, lincw@mail.ncku.edu.tw, febryans2802.wtmh@gmail.com

Keywords: Taekwondo, Poomsae, OpenPose, Convolutional Neural Network (CNN).

Abstract: Taekwondo competition consists of three elements: Poomsae, sparring, and breaking. Poomsae is a set of movements consisting of punches and kicks and focuses on technique, breathing ability, balance, coordination and concentration. OpenPose is a library that can represent the human body, whether it's body, head, hands or feet. OpenPose is widely used in research related to the classification of human movement. This research will analyze the effect of OpenPose as an input for the identification of Taekwondo movements, especially in Poomsae-3 using a CNN. The image data of taekwondo movements were taken from video recordings. In this research, three types of classification models were designed to identify taekwondo movements, Model-1 with input in the form of taekwondo movement images, Model-2 with input in the form of taekwondo movement images and OpenPose, and Model-3 with input in the form of OpenPose. From the test results, Model-3 with input in the form of an OpenPose keypoint obtained the best accuracy results (amounting to 99.39%) compared to Model-2 (97.68%) and Model-1 (79.86%). For further research development, the identification of taekwondo movements can be done for input in the form of video, and analysis of the significant keypoints for the taekwondo movement.

1 INTRODUCTION

Taekwondo is one of the martial arts sports that was formed in 1962 (Kil, 2006). Taekwondo nourishes the mind and body by creating harmony between physical and mental exercises through the use of hands and feet. The word taekwondo derived from Tae stands for basic kick, the word kwon means "boxing that blocks and punches" (Kil, 2006). So if Taekwondo translates into philosophy, it is behavior instilled through discipline and training of body and mind (Kil, 2006). With the development of this sport, taekwondo is often competed. The formal Taekwondo competition consists of three elements: Poomsae, sparring, and breaking (Kil, 2006). Poomsae (shape) is a combination of basic movements such as blocking, punching, and kicking. Poomsae is a set of movements consisting of punches and kicks and focusing on technique, breathing ability, balance, coordination, and concentration (Lawofthefist.com, n.d.). In Taekwondo there are eight Poomsae. Each poomsae level is more complex than the previous poomsae. Therefore in doing a good poomsae techniques are required with accuracy and good coordination, because if the movement is not in accordance with the actual movement then the meaning will be different.

For the estimation of human poses in two dimensions in real time on an image using the Convolutional Neural Network (CNN) has been

---

[a] https://orcid.org/0000-0001-7774-9675
[b] https://orcid.org/0000-0003-3720-3584
[c] https://orcid.org/0000-0002-3198-2815
[d] https://orcid.org/0000-0002-3265-0476
[e] https://orcid.org/0000-0002-1894-1189
[f] https://orcid.org/0000-0002-3671-9127

proposed by Cao et al, by extracting the 2D frame from the video using the Red Green Blue (RGB) camera (Cao et al., 2017). OpenPose is an open library that can represent the human body, consisting of bodies, heads, hands, and feet with a total of 135 keypoints, and the input can be in the form of images, videos, or real-time videos(G. Hidalgo, Z. Cao, T. Simon, S.-E. Wei, H. Joo, Sheikh, n.d.).

OpenPose is widely used as a form of step extraction feature and utilized as a classification. Researchers previously used OpenPose to classify several movements, such as Korean heart pose, child pose, KungFu crane, and so on with ResNet-50, ResNeXt-101, and PNASnet-5 architectures (Lau, 2019). The classification is based on 3 different datasets, namely original image, original image with OpenPose skeleton, and skeleton OpenPose only. From testing result of PNASnet-5 architecture with original image has an accuracy of 83% (Lau, 2019). In Deep Learning, Convolutional Neural Network (CNN) is a class of neural networks that is commonly applied to analyzing visual images.

This research will analyze the effect of OpenPose as an input for the identification of Taekwondo movements, especially in Poomsae-3 using a CNN.

2 METHODS

The following will be discussed about experiment scenario and design system of taekwondo Poomsae-3 movement identification using CNN.

2.1 Experiment Scenario

The experiment scenario was conducted by the Taekwondo Activity Unit, which was conducted by six respondents. Each respondent performs the same 4-set of poomsae-3 movements, which are recorded in the form of motion videos. The resulting video is 640 x 480 pixels in MP4 format and the video speed is 30 fps. The video was taken using cameras (there are two cameras) placed in the front position (Camera-1) and the rear position (Camera-2) of the respondent. Here is a layout of the camera placement and respondent (Figure 1).
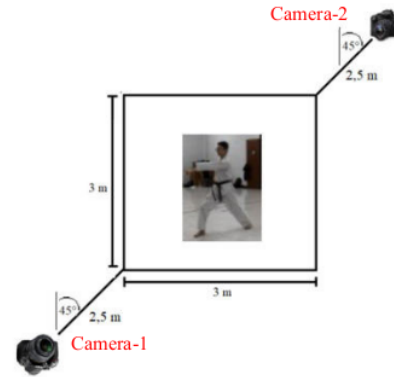
Figure 1: Layout experiment scenario.

2.1.1 Poomsae-3 Movement

In this research, the poomsae to be classified was Poomsae-3, formally known as Taegeuk Sam Jang (Kukkiwon, 2012). Generally, poomsae is trained by students in a group of 6 to rise to level 5. Poomsae is a combination of attack and defense movement and movement must be done vigorously, such as fire (Kukkiwon, 2012). The order of movement Poomsae-3 can be seen in Figure 2.

Figure 2: Poomsae-3 movement sequence.

From the sequence of movements, the Poomsae-3 movements are grouped into eight groups of movements, there are Chumbi, Are Maki, Ap Chagui, Montong Dubong Jirugui, Sonnal An Chigui, Jan Sonnal Maki, Montong Baro Jirugui, Montong An Maki, can be seen in Figure 3.

Figure 3: Poomsae-3 movement (Gimnasio Lee, 2014).

2.1.2 Poomsae-3 Dataset

Of the eight groups of Poomsae-3 movements, the research was separated between the left and right movements, so there are 15 classes of movement to be identified (see Table 1). The position and direction of movement obtained from Camera-1 and Camera-2 result in different appearances. In order for each movement obtained from the camera to show a detailed and clear posture of movement, the dataset used in this research is separated between the movement group and the imagery of which camera is used, see Table 1.

Table 1: Poomsae-3 movement class.

No	Class	Cam-1	Cam-2
1	Chumbi	v	
2	Are Maki (Left)		v
3	Are Maki (Right)	v	
4	Ap Chagui (Right)		v
5	Ap Chagui (Left)	v	
6	Montong D. Jirugui (Left)	v	
7	Montong D. Jirugui (Right)	v	
8	Sonnal An Chigui (Right)		v
9	Sonnal An Chigui (Left)	v	
10	Jan Sonnal Maki (Right)	v	
11	Jan Sonnal Maki (Left)	v	
12	Montong An Maki (Right)		v
13	Montong An Maki (Left)	v	
14	Montong. B Jirugui (Left)		v
15	Montong B. Jirugui (Right)		v

2.2 Design System Identification

In this research, three types of classification models were designed to identify taekwondo movements Poomsae-3. Broadly speaking the design of the Poomsae-3 taekwondo movement identification system using CNN, AlexNet architecture (see Figure 4). Details of each process will be explained in a later explanation.
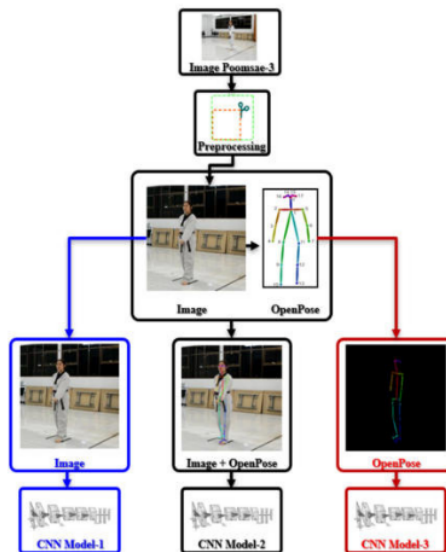

Figure 4: Design system identification Poomsae-3 Movement.

2.2.1 Preprocessing

From the video dataset obtained, a frame will be taken depicting each movement. Then the imagery in the left and right crop so that from the size of 640x480 pixels, to the size of 480x480 pixels. After that it will be resized to a size of 227x227 pixels. Specifically for Model-2 and Model-3, a dataset rendered with OpenPose is used. For Model-2, a combined dataset of the original image and the OpenPose image is used. For Model-3, the keypoints image dataset from OpenPose rendering is used (Figure 5).
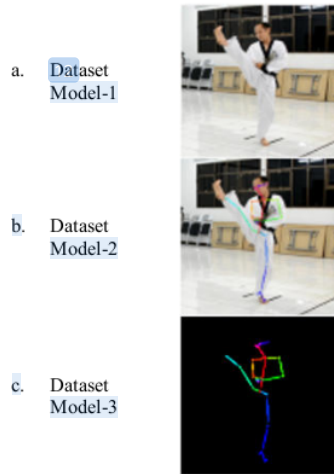

a. Dataset Model-1

b. Dataset Model-2

c. Dataset Model-3

Figure 5: Example of dataset model for Ap Chagui Class (a. Model-1, b. Model-2, c. Model-3).

2.2.2 OpenPose

In this research, three types of data sets will be used in the form of images (Dataset Model-1), images and skeleton OpenPose (Dataset Model-2) and Skeleton OpenPose without background (Dataset Model-3). After the preprocessing process, the next is to perform the OpenPose process, to obtain the Model-2 Dataset and the Model-3 Dataset. Then the dataset is separated between the training and testing datasets. In this research, used 6-fold Cross Validation. For each model, there are 360 training data sets and 45 test data sets.

2.2.3 CNN Model

The CNN model that will be used in this research is AlexNet. AlexNet is a Neural Network created by Alex Kriszhevsky and his colleagues in 2012 (Krizhevsky et al., 2017), and published with Illya

Sutskever and Geoffrey Hinton. This network has eight layers, seen in Figure 6. The first five layers are convolutional neural layers that act as filters for capturing features on the input (Khan et al., 2018), and followed by the max-pooling layer, and the last three layers are fully connected layers that multiply inputs by weight plus bias vector. In addition, the output layer used has a non-saturation ReLU activation function (Krizhevsky et al., 2017) and architecture as follows (Figure 6):
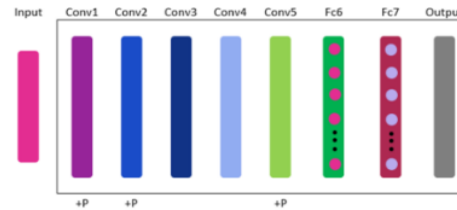

Figure 6: AlexNet architecture.

The first step will be the training process and validation of CNN architecture with different datasets. After the training model is obtained then testing is conducted. The last step done is confusion matrix calculation.

In the confusion matrix, there are six terms used to represent the results of classification, namely True Positive (TP) with positive observation value and positive prediction value, True Negative (TN) with negative observation value and negative prediction value, False Positive (FP) with negative observation value but positive prediction value, and False Negative (FN) with positive observation value but negative prediction value. Values TP, TN. FP, and FN obtained can be used to calculate accuracy values, sensitivity values (recalls), and precision values that can be calculated with the following formulas (Manliguez, 2016).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \; x \; 100\% \quad (1)$$

$$Sensitivity = \frac{TP \; total}{TP \; total \; + \; FN} \; x \; 100\% \quad (2)$$

$$Precision = \frac{TP \; total}{TP \; total \; + \; FP} \; x \; 100\% \quad (3)$$

3 RESULTS AND DISCUSSION

The following are the training and testing results of the identification of the Taekwondo Poomsae-3 movement using CNN.

3.1 CNN Training & Validation Result

The following training results & validation from each model get the following results (Table 2).

Table 2: Training & Validation Accuracy.

Model	Accuracy each Fold (%)						AVG (%)
	1	2	3	4	5	6	
1	80.00	78.30	70.00	80.00	77.70	80.00	77.69
2	97.70	97.30	100.00	96.60	100.00	96.60	98.07
3	98.30	100.00	100.00	95.00	100.00	94.60	98.00

In Table 2, the accuracy of the Model-2 is 98.07%, then the Model-3 is 98.00% and the last is the Model-1 by 77.69%.

3.2. CNN Testing Result

After the training process & validation continued the testing process. From the test results obtained accuracy results Model-1 (Figure 7), Model-2 (Figure 8), and Model-3 (Figure 9). From the test results of Model-1, the lowest accuracy is in Montong An Maki class (Left) which is 62.5%, and the highest accuracy is 93.75% in Ap Chagui (Right) and Sonnal An Chigui (Right) classes.

From the test results of Model-2, the lowest accuracy is in Montong An Maki (Right) class which is 90.48%, and the highest accuracy is 100% in Chumbi, Are Maki (Left), Are Maki (Right), Ap Chagui (Left), Ap Chagui (Right), Jan Sonnal Maki (Left), and Jan Sonnal Maki (Right) classes.

From the test results of Model-3, the lowest accuracy is 97.37% in the Ap Chagui (Right), Ap Chagui (Left) and Montong An Maki (Right) classes and the highest accuracy is 100% in other classes.

In Table 3, it can be seen that from all tests, Model-3 showed the highest accuracy, sensitivity, and precision results compared to Model-1 and Model-2, which is 99.39% (Accuracy), 97.78% (Sensitivity) and 97.78% (Precision).

Table 3: Testing Accuracy, Sensitivity & Precision.

Model	Accuracy (%)	Sensitivity (%)	Precision (%)
1	79.86	41.67	61.51
2	97.68	86.67	87.78
3	99.39	97.78	97.78

In Model-3, which is an OpenPose image without the original image, it has the highest accuracy, because the features used in this model focuses on OpenPose without considering the background, body shape and clothing color of the taekwondo player, etc.
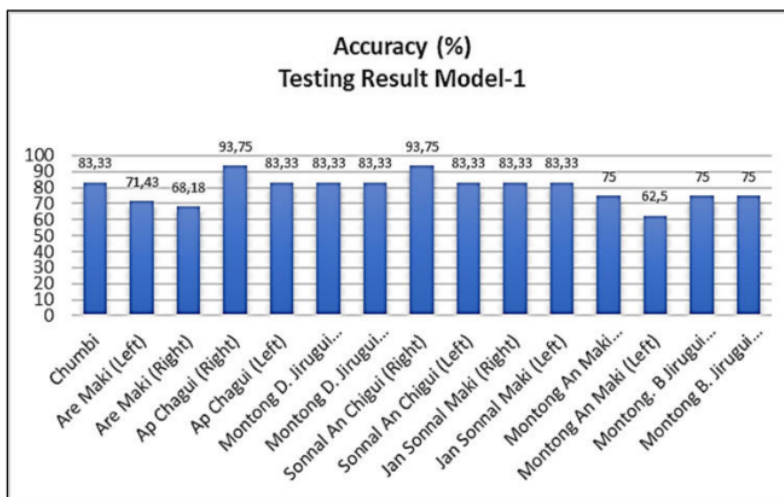

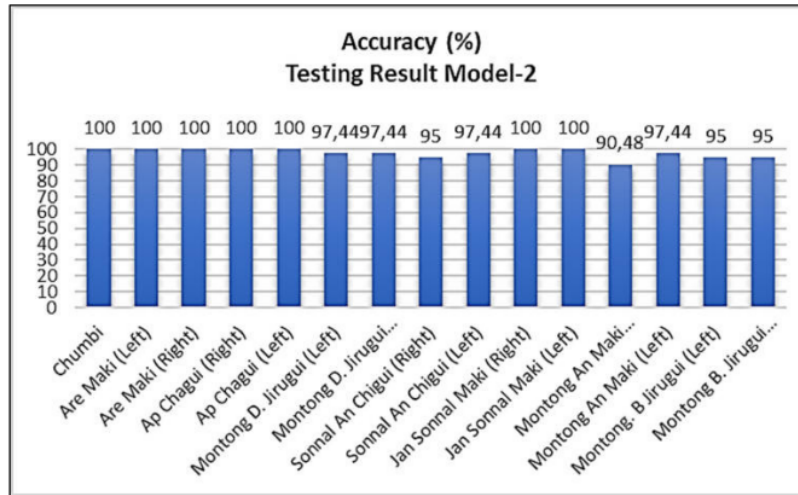Figure 7: Accuracy Testing Model-1.
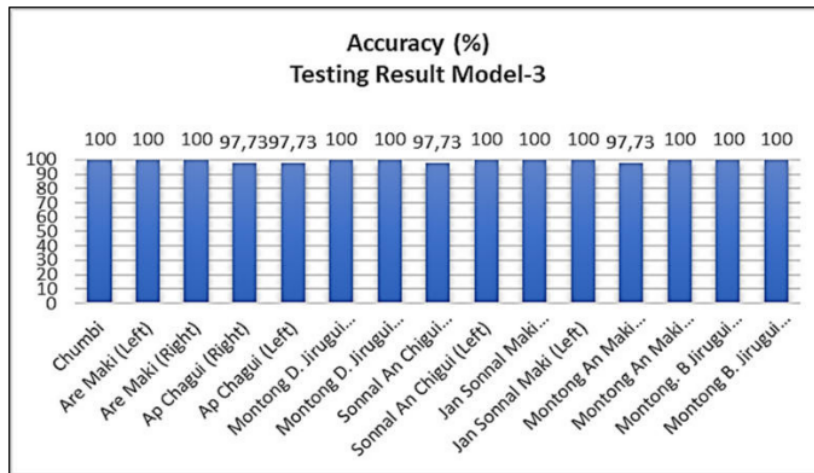
Figure 8: Accuracy Testing Model-2.

Figure 9: Accuracy Testing Model-3.

4 CONCLUSIONS

Identification of Taekwondo Movement Poomsae-3 is successfully realized by using CNN, by using 3 dataset models, namely datasets in the form of image (Model-1), datasets in the form of images with OpenPose (Model-2) and dataset OpenPose (Model-3). From each CNN test, Model-3 produced the best results compared to the Model-1 and Model-2, with the accuracy of 99.39%, precision of 97.78% and sensitivity of 97.78%.

For further research development, the identification of taekwondo movements can be done for input in the form of video, and analysis of the significant keypoints of the taekwondo movement. Moreover it can be implemented to support the sport of taekwondo by helping the training process for the participants, or by assisting to calculate values in a competition.

ACKNOWLEDGEMENTS

REFERENCES

Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2D pose estimation using part affinity fields. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-January,* 1302–1310. https://doi.org/10.1109/CVPR.2017.143

G. Hidalgo, Z. Cao, T. Simon, S.-E. Wei, H. Joo, Sheikh, Y. (n.d.). *OpenPose library.* https://github.com/CMU-Perceptual-Computing-Lab/openpose

Gimnasio Lee. (2014). *TAEGUK SAM CHANG . 3º Filosofía.* https://www.gimnasiolee.com/taekwondo_html/pumse3_tkd.html

Khan, S., Rahmani, H., Shah, S. A. A., & Bennamoun, M. (2018). A Guide to Convolutional Neural Networks for Computer Vision. In *Synthesis Lectures on Computer Vision* (Vol. 8, Issue 1). https://doi.org/10.2200/s00822ed1v01y201712cov015

Kil, Y. S. (2006). *Competitive Taekwondo* (1st ed.). Human Kinetics.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90. https://doi.org/10.1145/3065386

Kukkiwon. (2012). *Kukkiwon Taekwondo Textbook* (1st ed.). Korean Book Service.

Lau, B. (2019). *OpenPose, PNASNet 5 for Pose Classification Competition (Fastai).* Towards Data Science. https://towardsdatascience.com/openpose-pnasnet-5-for-pose-classification-competition-fastai-dc35709158d0

Lawofthefist.com. (n.d.). *Your Ultimate Guide To Taekwondo Forms : Poomsae & Patterns.* https://lawofthefist.com/your-ultimate-taekwondo-poomsae-guide/

Manliguez, C. (2016). *Generalized Confusion Matrix for Multiple Classes Generalized Confusion Matrix for Multiple Classes The total numbers of false negative (TFN), false positive (TFP), and true negative (TTN) for each class i will be calculated based on the Generalized.* November, 5–7. https://doi.org/10.13140/RG.2.2.31150.51523

2. Taekwondo Poomsae-3 Movement Identification by using CNN

Publication

| 6 | philpapers.org
Internet Source | 1 % |

| 7 | Yoav Bergner, Shiri Mund, Ofer Chen, Willie Payne. "First Steps in Dance Data Science", Proceedings of the 6th International Conference on Movement and Computing - MOCO '19, 2019
Publication | 1 % |

| 8 | George Leitmann. "Adaptive control of aircraft in windshear", International Journal of Robust and Nonlinear Control, 1993
Publication | 1 % |

| 9 | I Gede Susrama Mas Diyasa, Ni Made Ika Marini Mandenni, Mohammad Idham Fachrurrozi, Sunu Ilham Pradika et al. "Twitter Sentiment Analysis as an Evaluation and Service Base On Python Textblob", IOP Conference Series: Materials Science and Engineering, 2021
Publication | 1 % |

| 10 | ijournalse.org
Internet Source | 1 % |

| 11 | www.hindawi.com
Internet Source | <1 % |

| 12 | COMPEL: The International Journal for Computation and Mathematics in Electrical | <1 % |

and Electronic Engineering, Volume 33, Issue 5 (2014-09-16)

Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

2. Taekwondo Poomsae-3 Movement Identification by using CNN

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7