Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2025. Vol. 5. No. 1; e-ISSN: 2808-4519

Detection of the Precision of the Basic Movement of the Line using the Convolutional Neural Network (CNN)

Abraham Makaborang [1]*, Pingky Alfa Ray Leo Lede[2], Eben Panja [3]

[1,2,3] *Informatics Engineering Study Program, Wira Wacana Christian University Sumba*
abrahammakaborang669@gmail.com[1*], pingky.leo.lede@unkriswina.ac.id[2], ebenpanja@unkriswina.ac.id [3]

Abstract

Payeti Christian Junior High School in Kambera District, East Sumba Regency, NTT, still makes marching exercises (PBB) an important part of extracurricular activities and character development of students. However, the evaluation of the basic movements of the United Nations is still carried out manually by observing the movements of the students one by one. This method is inefficient and tends to be subjective. Common mistakes found include not standing upright, head down, misaligned legs, and hands that are not straight when in perfect posture. This study aims to develop an automatic detection system using Convolutional Neural Network (CNN) to evaluate the accuracy of UN basic movements through static imagery. This system is expected to provide a faster, objective, and documented evaluation. The implementation results show a very high level of accuracy, with model confidence above 92% to 99.99%. Out of 200 test data, only one was misclassified. The accuracy and loss graphs show the stability of the model without overfitting, with a final validation accuracy of 98.44%.

Keywords: *Line-Up Rules (UN), Convolutional Neural Network (CNN), Body Pose*

1. Introduction

Technology is rapidly evolving and has become an integral part of various aspects of human life, including in the world of education, training, and skills evaluation. Advances in the field of artificial intelligence (AI), especially in image processing and pattern recognition, have allowed the creation of systems that are able to mimic human ability to recognize, classify, and assess visual objects. One of the methods that is widely used in digital image processing is the Convolutional Neural Network (CNN). CNN is an artificial neural network architecture specifically designed to recognize visual patterns in image data. This technology has been widely applied in various fields such as facial recognition, object detection, motion analysis, and security systems. Through convolution and pooling layers, CNN is able to identify important features of the image, which are then further processed to produce an output in the form of predictions or classifications.

The Row-Marching Regulation (PBB) is a form of training that hones physical and mental aspects with the main goal of forming discipline, cohesiveness, and the ability to follow instructions simultaneously and appropriately. This exercise is commonly applied in institutions such as the military, police, schools, and youth organizations as part of character building programs and discipline training. Broadly speaking, the United Nations is classified into three types, namely the unarmed movement (abadan), the armed movement, and the group formation movement. Examples of the UN's basic movements include perfect posture, facing right or left, turning right, walking in place, and firm steps. Each of these movements is carried out based on commands that must be followed uniformly by all participants. The Row-Marching Regulation (PBB) is a form of training that hones physical and mental aspects with the main goal of forming discipline, cohesiveness, and the ability to follow instructions simultaneously and appropriately. This exercise is commonly applied in institutions such as the military, police, schools, and youth organizations as part of character building programs and discipline training. Broadly speaking, the United Nations is classified into three types, namely the unarmed movement (abadan), the armed movement, and the group formation movement. Examples of the UN's basic movements include perfect posture, facing right or left, turning right, walking in place, and firm steps. Each of these movements is carried out based on commands that must be followed uniformly by all participants.

Based on the results of interviews conducted at one of the junior high schools, namely Payeti Christian Junior High School located in Kambera District, East Sumba Regency, East Nusa Tenggara, together with sports teachers, information was obtained that marching exercises are still an important part of extracurricular activities and character development of students. However, the coaching teacher said that the evaluation process of the basic movements of the United Nations is still carried out manually, namely by directly observing the movements of students one by one during exercises and when conducting practical exams. This method is considered less effective, especially when the number of participants is large enough that it is difficult for the instructor to provide corrections individually. In addition, the evaluations carried out tend to be subjective, depending on the perception and experience of each coach, so the assessment standards can vary. In the interview, it was also explained that many students still make common mistakes in UN practice, both when standing and moving. Some examples of mistakes that are often found when conducting student practice exams in the basic UN movements

include body positions that are not perpendicular, heads that are lowered, legs that are too wide or not aligned, and hands that are not straight or less firm when doing perfect postures.
To answer these challenges, an automatic evaluation system based on artificial intelligence technology is needed that is able to analyze the movements of participants visually and objectively. One promising approach is the use of *Convolutional Neural Network* (CNN), a *Deep Learning* method that has proven effective in recognizing visual patterns in images. CNN works by extracting important features from images, then classifying or evaluating the data based on previously learned patterns. In addition, a number of previous studies have also developed CNN models for the recognition of human movements in various fields. For example, pose *estimation* technology such as that used in the analysis of pencak silat movements has been used to analyze posture in the field of sports. On the other hand, studies that use CNN in the analysis of athletic movements have shown its effectiveness in assessing sports techniques such as running and long jumps. However, these studies have not specifically targeted motion detection within the United Nations which has unique patterns and strict standards. Therefore, this study aims to fill the gap by developing a CNN model specifically designed to assess the accuracy of UN movements.
This research aims to develop a CNN-based system that is able to automatically detect the accuracy of the basic movements of the United Nations through imagery. The system will be trained using *a standardized UN movement imagery dataset, then classified whether x to support the teacher or instructor assessment process, especially in the context of a line-up practical exam.*

2. Page layout

2.1. Line-Marching Regulation (UN)

The Line-Marching Regulation (PBB) is a regulation of marching procedures that are manifested in the form of physical training necessary to instill habits and corporal spirit in military life which is directed to the formation of a soldier attitude with character and body that is firm, agile, fosters discipline, high loyalty, togetherness and a sense of responsibility so that always prioritizes the interests of duty over individual interests [1].

2.2. Digital Image

Digital imagery is a two-dimensional visual representation of the three-dimensional real physical forms of an object that has gone through a processing process in a computer device, where the image is obtained from various types of digital devices such as cameras, scanners, or other sensors, thus producing visual data that can be analyzed, modified, and utilized in various information and communication technology applications [2]

2.3. Digital Image Processing

Digital image processing is a science that studies matters related to improving the quality of an image (increasing contrast, color change, image restoration), image transformation (translation, transformation rotation, scale, geometric), selecting optimal feature images for analysis purposes, storing data that was previously reduced and compression, data transmission, and data processing time [3]

2.4. Artificial Intelligence (AI)

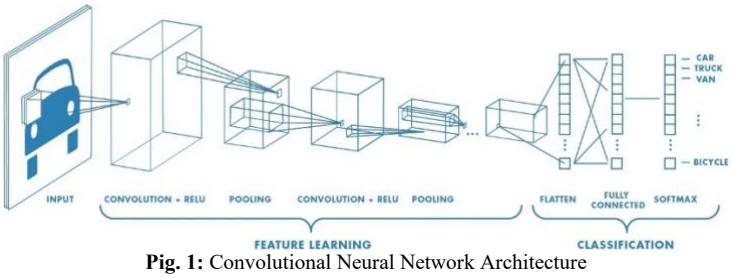
Artificial Intelligence (AI) is a technology in machines that is designed and developed with artificial intelligence capabilities, so that it allows the machine to carry out various tasks that usually require human intelligence, such as recognizing patterns, making decisions, solving problems, and interacting with the environment, without requiring manual assistance or direct intervention from humans, with the aim of improving efficiency, accuracy, and speed in completing various jobs in various fields such as industry, education, health, and information technology [4]

2.5. Deep Learning

Deep Learning (DL) is a technique in Neural Network (NN) that utilizes special methods such as the Restricted Boltzmann Machine (RBM) to speed up the learning process in a network that has many layers, usually more than seven, so that it can reduce the time needed for the training process because it solves the problem of gradient loss in the reverse propagation process more effectively[4]

2.6. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is one of the algorithms *Deep Learning* which is widely used to solve image classification problems because of its ability to achieve a high level of accuracy even though it only uses *Preprocessing* or minimal segmentation [5].
According to Wahyuni & Sulaeman (2022), CNN is a trainable architecture and consists of several stages. The inputs and outputs of each stage consist of several arrays commonly called feature maps. Each stage consists of three layers, namely convolution, layer activation function and Pooling Layer. Figure 2.1 is the Convolutional Neural Network architecture network.



Pig. 1: Convolutional Neural Network Architecture

2.6.1 Convolution Layer

Convolution Layer is the initial stage in the feature learning process at CNN. In this layer, a convolution operation is performed between the input data and the kernel derived from the filter. The convolution operation itself is the process of multiplication between two matrices, and then the results are summed to produce a feature map that is a representation of the features of the input.

2.6.2 Pooling Layer

According to[4], CNN has two main approaches to shrink the dimension of input volume, namely by *Convolution Layer* use *Stride* >1 and *Pooling Layer*. The Pooling Layer is placed after the Convolution Layer and functions to reduce the dimension of the feature map using a certain size filter with a certain stride. Two commonly used methods in pooling are *Average pooling functions to take the average value of the filter area, while max pooling chooses the highest value in that area*. The use of pooling gradually helps to reduce the output size, number of parameters, and calculations on the network, so that it can control overfitting. For example, in the 2x2 max pooling technique with a stride of 2, the system will select the highest value of each block of 2x2 pixels to proceed to the next process:

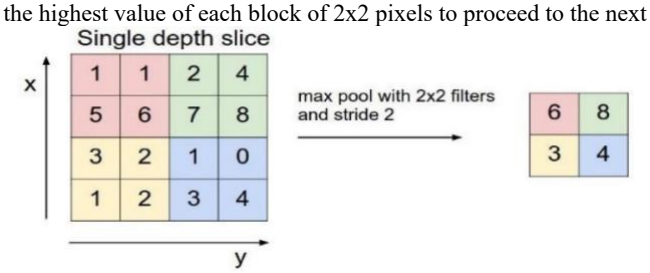


Fig. 2: *Max Pooling*

2.6.3 Fully Connected Layer

According to [4], *The Fully Connected layer is a part of the neural network where all the neurons of the previous layer are fully connected and used in the final classification process. This layer is similar to conventional neural networks, both in the form of single-layer networks and multilayer perceptron (MLP). Before entering the classification stage, the feature map resulting from the feature learning process is still in the form of a multi-dimensional array, so it needs to be converted first into a one-dimensional vector through a process known as flattening.*

2.6.4 Confusion Matrix

The Confusion Matrix can be used to analyze the potential of the classifier perfectly. All diagonal elements show correctly classified results. The incorrectly classified result is represented on the diagonal of the *Confusion Matrix*. Therefore, the best classifier will have a confusion matrix with only the diagonal element and the other elements set to zero. *The Confusion Matrix* generates the actual value and the prediction value after the classification process.

2.6.5 TensorFlow

TensorFlow is a software library developed by the Google team to support various activities in the field of machine learning and artificial neural networks. This platform combines the concept of computational algebra with optimization techniques in the compilation process, so that it is able to present the calculation of mathematical expressions in an efficient, systematic, and structured manner[4]

2.6.6 Python

Python is a versatile programming language that is interpretive and designed with the principle of high code readability. The language is known for having a simple yet robust syntax, and offers extensive functionality through a large and comprehensive standard library, supporting efficient and flexible program development[6].

3. Research Methods

3.1. Research Flow

This study uses the Waterfall method approach, which is a sequential and systematic system development method from the initial stage to the end.

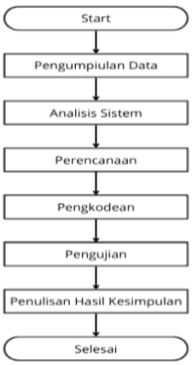


Fig. 3: Research flow

1. System Analysis

After the data is collected, a system analysis is carried out to identify the needs of the users. This analysis aims to determine what features and functions must be possessed by the system in order to be able to answer the problems found.

2. Planning

At this stage, the results of the analysis are used to design the system as a whole. Planning includes designing system workflows and creating CNN models. The goal is for the development of the system to run in a directional manner and in accordance with the needs that have been analyzed.

3. Coding

This stage is the implementation of the planning results, where the system design is developed using the python programming language. Each designed feature begins to be realized in the form of program code. This stage requires precision so that the system can run properly and efficiently.

4. Testing

The developed system then goes through the testing stage to ensure that each feature works according to design. This process also aims to identify possible errors (bugs) and evaluate the level of accuracy of the system. If obstacles or inconsistencies are found, it is necessary to make repairs before the system is implemented thoroughly.

5. Writing Conclusions

After the system is tested and declared to be running well, a final report is prepared containing the entire series of system development activities, from data collection to testing. At this stage, conclusions are also drawn about the results of system development and suggestions for further development in the future.

3.2. Data Collection Methods

The data collection method was used to obtain the information needed in designing the Precision Detection system for Basic Line-Marching Movement Poses Using Convolutional Neural Network (CNN) at Payeti Christian Junior High School. Data collection is carried out through several techniques, including:

1. Observation

The observation technique is carried out by directly observing conditions in the field to identify problems that exist at Payeti Christian Junior High School. This process also includes visual documentation in the form of photographs as a support to obtain a clearer picture of the situation and activities taking place at the research site.

2. Interview

The interview was conducted to dig up information in depth from relevant sources, especially the Row-Marching Rules trainer. This method is carried out through a direct question and answer session to obtain the data needed to formulate problems and achieve research objectives comprehensively.

3. Documentation

The dataset in this study was collected using the documentation method, which is the process of taking data using image or video media at the research location. This method is used to record various UN movements directly from the object that performs the movement. The data collected was in the form of images or videos of UN movements such as perfect attitude, respect, rest in place, forward foot, left and right side, and left and right side-footed. Each image or frame of the video is then analyzed using PoseNet to extract the key points of the human body. These key points are then used as inputs for the training of the Convolutional Neural Network (CNN) model so that the system can recognize and evaluate the appropriateness of the pose of each UN movement. Data collection is carried out in an environment that supports sufficient lighting and a simple background, to maximize the accuracy of pose detection. All documentation data is stored in a standard format and grouped by movement type to simplify the process of labeling and model training.

3.3. System Workflow

Fig. 4: System Workflow

Figure 2 is the workflow of the basic line-marching movement posture detection system starting with the user opening the application and selecting the type of movement they want to test, such as perfect posture, respect, proper rest, left and right sassy. After selecting a gesture, the system will activate the camera to detect the user's pose. The detection process is carried out using *the PoseNet* algorithm which extracts body keypoints such as the head, shoulders, hands, and feet. The *keypoint* data is then sent to a *Convolutional Neural Network* (CNN)-based classification model for analysis. The CNN model will compare the user's pose with the correct reference pose data. If the pose is appropriate, the system provides feedback that the movement is correct. If it doesn't match, the system will give a sign that the movement is still wrong. This process ends after the detection results are displayed to the user, and the user can choose to repeat the process or try other moves.

3.4. Implementation

Fig. 5: Implementation

Figure 3 shows the results of the implementation of the line-line basic motion positivity detection system using the Convolutional Neural Network (CNN) method. The system successfully identifies different types of UN movements with a high degree of accuracy, as indicated by the confidence value of each prediction. Some of the movements that were successfully detected included Lancang Kanan (91.63%), Rest on the Spot (70.91%), Respect (63.58%), Perfect Attitude (99.99%), and Lancang Kiri (93.97%). Each prediction is displayed directly on the input image with information on the movement name and the model's confidence level. These results indicate that the CNN model used is able to distinguish body poses fairly accurately despite variations in individual backgrounds, lighting, and postures.

3.5 Loss Function and Accuracy Function Graph

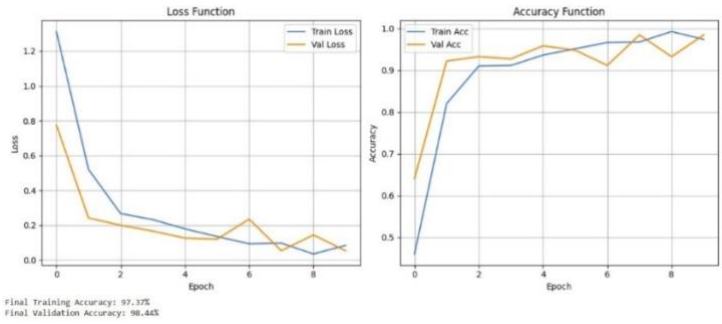


Fig. 6: Loss Function *and* Accuracy Function *Graph*

Figure 6 shows a graph of loss and accuracy during the model training process in 10 epochs. The graph on the left depicts a decrease in the value of the loss in the training and validation data. A fairly sharp decline occurred until the 4th epoch, then the value was stable and remained low, which indicates that the model has learned well without overfitting. Meanwhile, the graph on the right shows a significant increase in accuracy from epoch 0 to epoch 3, then continues to increase gradually closer to a value of 1.0 or 100%. At the end of the training, the model obtained an accuracy of 97.37% for the training data and 98.44% for the validation data, demonstrating the model's ability to generalize very well to data that had never been used before.

3.6 Model Evaluation Results

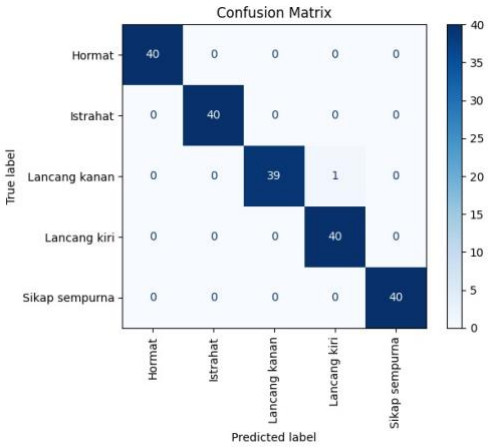


Fig.7: Model Evaluation Results

Figure 7 shows the results of the evaluation of the classification model in the form of a confusion matrix for five types of movements, namely Respect, Rest, Right-Leaning, Left-leaning, and Perfect Attitude. Based on the confusion matrix, it can be seen that the model has a very high accuracy in classifying each movement. Out of a total of 200 test data, there was only 1 prediction error, namely 1 Lancang Kanan movement data which was classified as Lancang Left. The remaining 199 data were correctly classified according to their original labels. In general, these results show that the model performs very well in recognizing and classifying each type of movement. This is also reflected in the diagonal value of the confusion matrix which is all high (as many as 40 for each class, except for Lancang right which is 39), and the value outside the diagonal which is close to zero.

4. Conclusions

A detection system has been successfully designed and developed in this study of CNN-based basic UN motion pose accuracy that is able to classify static images with high accuracy. The system shows optimal performance with 97.37% training accuracy and 98.44% validation, as well as a high confidence value in each movement. Evaluation through the confusion matrix showed excellent classification results, with minimal errors. This system has proven to be effective in automatically detecting participants' poses and can support the UN training process objectively and documented. For further development, it is recommended that the system include more variations of UN movements and use video data to understand the sequence of movements dynamically. Field trials with various conditions and the development of more interactive interfaces, such as auto-grading features and corrective suggestions, are also needed to improve the reliability and widespread utilization of the system.

.

References

[1] P. S. Santosa *et al.*, "Inculcating Discipline Values Through Marching Rules," *TRIMAS J. Inov. and Service. To Masy.*, vol. 3, no. 1, pp. 52–56, 2023, doi: 10.58707/trimas.v3i1.370.

[2] K. Khairunnisak, H. Ashari, and A. P. Kuncoro, "Forensic Analysis to Detect the Authenticity of Digital Images Using the NIST Method," *J. Resist. (Sist. Computer)*, vol. 3, no. 2, pp. 72–81, 2020, doi: 10.31598/jurnalresistor.v3i2.634.

[3] N. Z. Munantri, H. Sofyan, and M. Y. Florestiyanto, "Digital Image Processing Application for Tree Age Identification," *Ict*, vol. 16, no. 2, p. 97, 2020, doi: 10.31315/telematics.v16i2.3183.

[4] Mr. Reza *et al.*, "Artificial Intelligence: Image Processing & Application with Python," *Sem. Nas. Service. Mass. LPPM UMJ*Pp. 1–8, 2022, [Online]. Available: http://jurnal.umj.ac.id/index.php/semnaskat

[5] A. Jinan, B. H. Hayadi, and U. P. Utama, "Classification of Rice Plant Diseases Using Convolutional Neural Network Method through Leaf Imagery (Multilayer Perceptron)," *J. Comput. Eng. Sci.*, vol. 1, no. 2, pp. 37–44, 2022.

[6] A. N. Syahrudin and T. Kurniawan, "Inputs and Outputs in the Python Programming Language," *J. STMIK Python Programming Basics*, no. June 2018, pp. 1–7, 2018, [Online]. Available: https://www.researchgate.net/publication/338385483