ISSN 2706-655X Vol.5, Issue 5: 35-41, DOI: 10.25236/IJFET.2023.050506

Shooting Posture Correction System Based on Deep Learning

Yuyang Wu

The United World College of the Atlantic, Wales,UK

Abstract: *In view of the lack of effective guidance when people practice shooting posture at home, I designed this model. By importing the video captured by the camera into the model, the model will track the athletes' shoulder, elbow, wrist and other key points in the video, and then calculate the bending Angle of the elbow, and finally judge whether the athletes' shooting posture is standard. The experiment proves that this model has reached the expected goal.*

Keywords: *Deep learning, Mediapipe*

1. Introduction

With the development of the times, the pace of social life is getting faster and faster. The way of people's life is getting less and less healthy. Sometimes, people don't even have any time to eat meals and sleep, not to mention that they don't have any time to do exercise.

Because of COVID-19 outbreak in the large scale of China, lockdown policy has been established in many areas and cities. During this time period, people occasionally have this good chance to have do physical exercise at home, since they cannot go to the gym or basketball court. In order to make sure that people can have a high quality of exercising at home, this application has therefore been created.

This aim of producing this product is to satisfy young people's needs. Young people don't have enough time to do exercise, therefore, we should make sure every one exercises effectively. This calculator can analysis the shooting video and it can automatically. This helps the player correct his posture effectively, and make the exercise itself meaningful. Comparing with asking a professional basketball player, this one saves the resources of humans.

2. Related Work

In order to solve the posture problems we faced while practicing shooting basketball, we create a model that can assess and correct shooting posture. This model is mainly consisted of deep learning framework TensorFlow and human posture recognition framework Mediapipe, and it can finally classify and correct shooting posture.

To begin with, we use TensorFlow[1] to build a deep neural network model which is centered on convolutional network. Since our input data is ordered video stream, importing LSTM (Long short-term memory) network into this model and finishing our sorting mission by using sigmoid function. While training this model, using Xavier initialization of to initialize the weights in the model. Then, two typical basketball athletes' shooting videos were used as our database. In the end, the binary-cross-entropy function is used as the loss function, the Adam gradient descent method is used to realize the updating of parameters in the model, and the accuracy of model classification is used as the standard to evaluate the effectiveness of the model, so as to realize the purpose of identifying and classifying the typical shooting posture of athlete.

In addition, by watching and analyzing shooting videos from professional basketball players, we figured out that whether the shooting posture is standard depending on the angle between upper and lower arms is between 50 degrees and 85 degrees or not. Therefore, we added mediapipe framework to build human posture recognition model, then track and observe the player's angle between the upper and lower arms in the video and judge whether the posture is standard.

ISSN 2706-655X Vol.5, Issue 5: 35-41, DOI: 10.25236/IJFET.2023.050506

2.1. Convolutional Layer

Convolutional layer is one of the most important part in the convolutional neural network, and the characteristic of convolutional layer is to extract main elements from the input pictures.

$$S(i,j) = (I * K)(i,j) = \sum\sum I(i,j)K(m,n) + b \quad (1)$$

In the formula, m,n represent magnitude of the convolutional kernel, and b is the bias.

$$a = f\big(s(i,j)\big) \quad (2)$$

f is an activation function.

2.2. Maxpooling Layer

Pooling layers will run a subsampling processing to the pictures which are learned by convolutional layers. The maxpooling layer will output the maximum value in the certain range.

$$p = \max(a) = \max(f(s(i,j)) \quad (3)$$

2.3. LSTM (Long Short-Term Memory)

LSTM network is shown below. LSTM[2] is a recurrent neural network which is used to deal with sequential data , as shown in Figure 1.

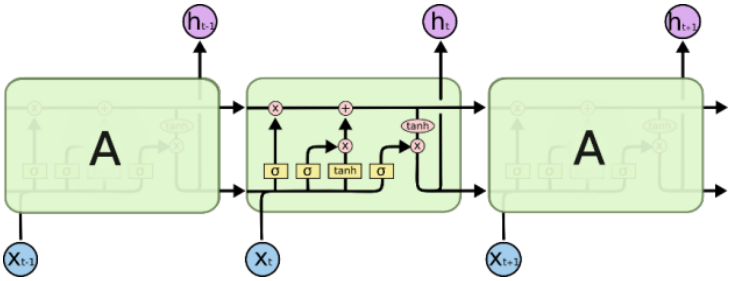

Figure 1: LSTM network structure.

2.4. Dense Layer

The dense layer is shown below. There are a number of neurons in the dense layers, and each neuron will connect with every its preceding neurons , as shown in Figure 2.

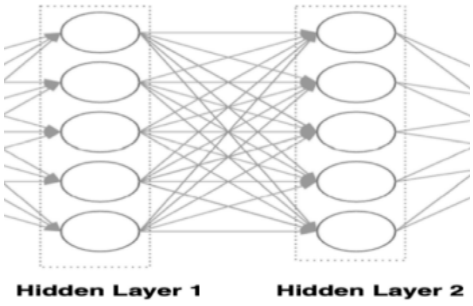

Figure 2: Dense layer.

2.5. MediaPipe

MediaPipe[3] is a human posture recognition framework. MediaPipe is a framework for building machine learning pipelines for processing video, audio, and other time series data. We use MediaPipe to recognize the angle between the tester's upper arm and lower arm, and analyzing whether the posture is correct.

ISSN 2706-655X Vol.5, Issue 5: 35-41, DOI: 10.25236/IJFET.2023.050506

3. Experiment Design

3.1. Basketball Shooting Posture Classification Model

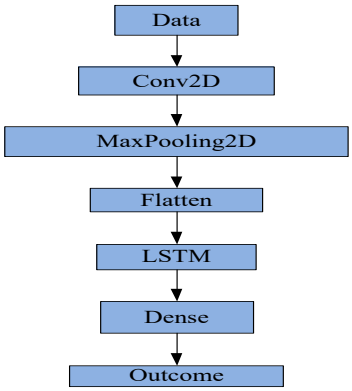

Figure 3: Classification model structure.

Classification model structure is shown above. When we import a data, it goes through the convolutional layer first, and the data was extracted through the convolutional layer. There are two convolution kernel, and the size of convolution kernel is 2×2. To solve the problems of gradient explosion and gradient disappearance, we use ReLU (Linear rectification function) to activate the function. Then in the maxpooling layer, a deeper analysis was set up. In this layer, the size of pooling layer is 2×2. Next, multiple dimension data experiences a process of flatten. Through this process, the data becomes to one dimension database. Since the data we imported was video stream, we then use LSTM to analyze our data. Finally, in the dense layer, each neuron is full connected with every neuron in the previous layer. The dimension of the outcome data is one- dimension. It is a two-classify model, so we should use sigmoid function as the activation function in this layer, and finally output the classification result , as shown in Figure 3.

3.2. Experiment Process.

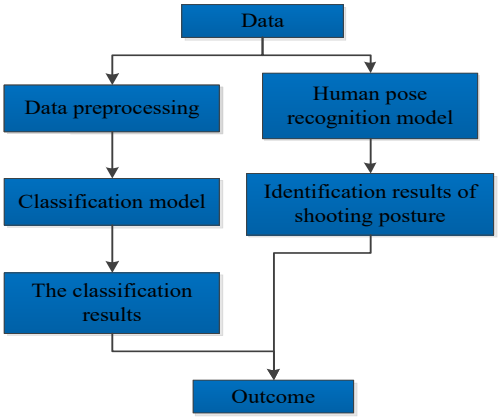

Figure 4: Experiment processing.

Experiment processing is shown above. Through inputting out video data, this procedure will first start to preprocess the data. By capturing the characteristics of the shooter in the video, we are able to compare it with the two professional basketball players. With the help of the classification model, we can classify the results, and decide the shooter's posture is similar with which player. As that same time, this procedure can also recognize human posture while shooting the ball. Through this model, we have identification results of shooting posture, which later helps a lot with the correction process , as shown in Figure 4.

ISSN 2706-655X Vol.5, Issue 5: 35-41, DOI: 10.25236/IJFET.2023.050506

3.3. Experiment Steps

3.3.1. Data Preprocessing

When a video was input, the procedure will start preprocessing. First, identify the numbers of the frames in the video, and then artificially decide the step length, which means after one step length, we extract one frame, eight frames in total in one video, and analyze them.

3.3.2. Classification Model

The structure of classification model is shown in figure 3. While compiling model, we use binary-crossentropy[4] as the loss function in this model. When training the model, the Xavier initialization is used to initialize the weights in the model. Then, we use the shooting videos of two typical basketball players as our database. Finally, the Adam gradient descent method is used to update the parameters of the model, and the classification accuracy of the model is taken as the criterion to evaluate the validity of the model. The mini batch size of input data is 4, and the epoch is 4 in this training network, then save the training model as CNN[5]-LSTM.h5 file. At last, input the test video so that we can test this model.

3.3.3. Human Pose Recognition Model

In this model, the procedure is able to recognize human pose in the video. It can analyze the angle of upper arm and the lower arm, and show the specific data at a particular time of shooting. MediaPipe helps create this method of analysis, and this framework realizes three function:

1) We can recognize players' posture.

2) Analyzing the joints position of the players' arms.

3) Find the angle we need.

4. Experiment and Analysis

4.1. Classification Model Performance

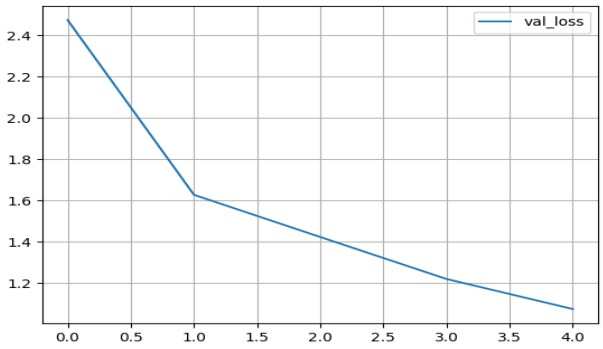

Figure 5: Graph for validation loss along training.

When the learning rate is 0.01, the training process of the classification model is shown below, and the x-axis represents epoch number, and the y-axis represent loss or accuracy , as shown in Figure 5.

While number of epoch increasing, the model became more and more accurate. As figure 6 is shown below, with the gradient descent, the validation loss keeps decreasing from 0 epoch to 4 epoch. Comparatively, the accuracy of this model is increasing.

ISSN 2706-655X Vol.5, Issue 5: 35-41, DOI: 10.25236/IJFET.2023.050506

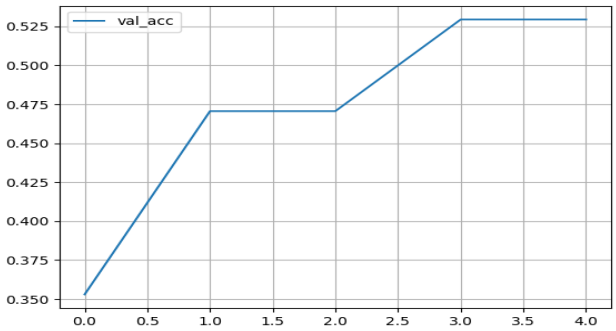

Figure 6: Graph for validation accuracy along training.

4.2. Human Posture Recognition Model

We use MediaPipe framework to create an atmosphere to analyze player's posture.

The first test is shown below. When we trained this model, we separate the shooting posture into two categories: "Curry" and "Klay Thompson", since they are the top shooters in the world. Then, while testing this model, we input a video of Curry's shooting shots. This trained model then gives a feedback that this video is belongs to the "Curry", and figures out that his shooting posture is standard.

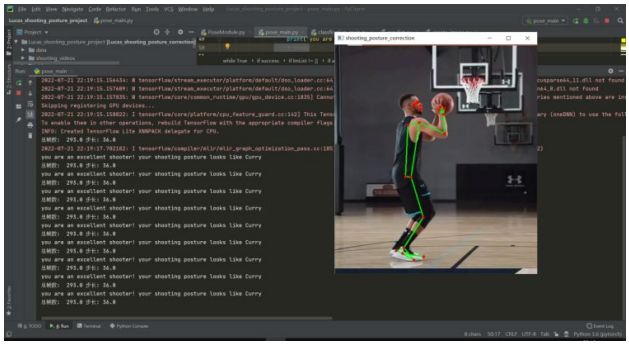

Figure 7: Testing I, screenshot.

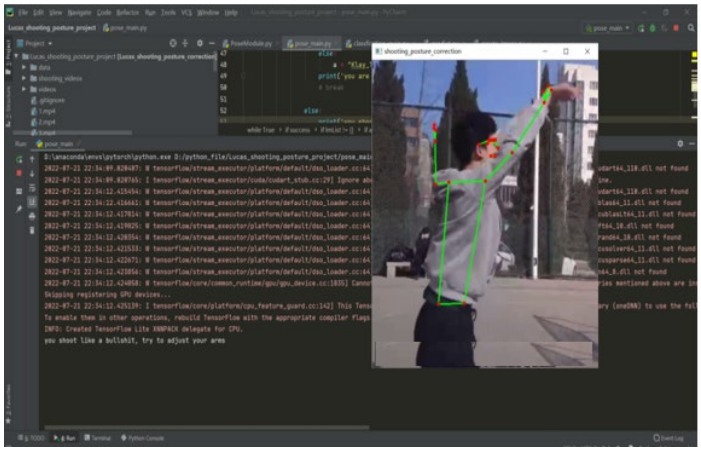

Figure 8: Testing II, screenshot.

ISSN 2706-655X Vol.5, Issue 5: 35-41, DOI: 10.25236/IJFET.2023.050506

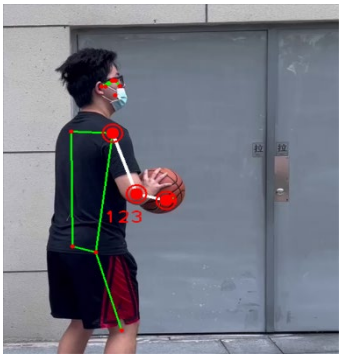

Figure 9: Testing III screenshot

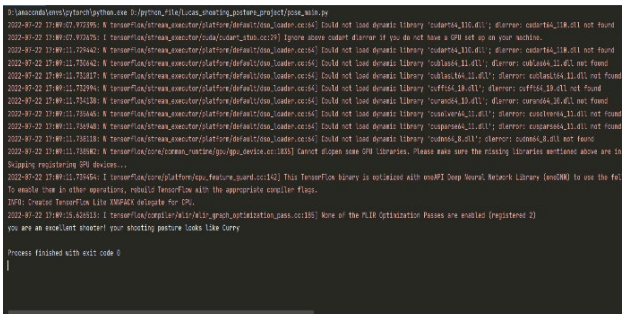

Figure 10: Testing IV screenshot.

The second test is shown below.Next time, we input another video, and this shooter is apparently not good at shooting at all. This model then successfully recognize this, and gives a suggestion to help this player to adjust his posture , as shown in Figure 7, Figure 8, Figure 9 and Figure 10.

After that, we then take videos on ourselves. Our player first tries him best to shoot as standard as possible, and the procedure successfully gives a result that our shooting posture is excellent, and divides us into “Curry”.

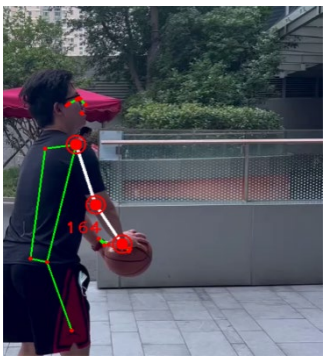

Figure 11: Testing V screenshot.

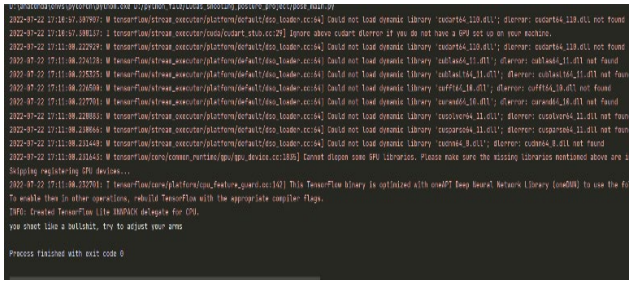

Figure 12: Testing VI screenshot

Then, our shooter stretches his arm while shooting, and the procedure also gives some suggestions , as shown in Figure 11 and Figure 12.

ISSN 2706-655X Vol.5, Issue 5: 35-41, DOI: 10.25236/IJFET.2023.050506

5. Conclusions

Shooting posture is a very important thing while practicing playing basketball, and sometimes, it is really hard for someone to pay attention to his own posture. As a result, we use MediaPipe framework to create a procedure that can help with correcting user's shooting posture, and separate his action into two types of shooting posture. However, the accuracy is not so accurate in some way, and this procedure prefer to separate use's posture into the "Curry" type, not "Thompson", though it should be half to half ideally.

This procedure can be developed in the future by adding more database, so that the accuracy can be improved.

References

[1] Abadi M., et al. "Tensor Flow: Large-Scale Machine Learning on Heterogeneous Distributed Systems." hgpu.org (2015).
[2] Hochreiter S., and J. Schmidhuber. "Long Short-Term Memory." Neural Computation 9.8(1997): 1735-1780.
[3] Lugaresi C., et al. "MediaPipe: A Framework for Building Perception Pipelines." (2019).
[4] Shannon and E. C. . "A Mathematical Theory of Communication." Bell Systems Technical Journal 27.4(1948):623-656.
[5] Fukushima K. "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position." Biological Cybernetics 36.4(1980): 193-202.