ORIGINAL ARTICLE

A Deep Learning-based Approach in Automated Detection of Genu Valgum

[1]Satyake Bakshi, [2]Sathya A

ABSTRACT

This paper investigates an attempt to detect genu valgums in patients. A typical diagnosis of genu valgum is based on the observation of how the alignment of the knees are during standing or if legs have different lengths. This system is an attempt to computationally draw a line between what kind of knee is a "genu valgum" and a "normal" knee. This system would also focus on an approach to increase the accuracy of the model to correctly classify this shape "disorder". Authors would rely on the creation of our dataset from Physiobank (modified) test it on a live subject to establish the standard of performance.

Keywords: Artificial neural network, Convolutional neural networks, Genu valgum, Knock knees.

INTRODUCTION

Bowlegs (genu varum) and knock knees (genu valgum) are usual in growing children that affect the lower portions.[1] Most of these problems represent normal physiological development and usually self-correct with changing time.[1] These cases typically do not require treatment. However, it is important for the physician to correctly differentiate between physiological deformities and those which would be pathological.[1] Without proper differentiation between the two, a treatment cannot be planned. Most inexperienced physicians would have trouble identifying the two if they both look identical.

In sports, knock knees is a cause for concern for most athletes. As the athletes grow physically, they begin to exert more force on the knees.[2] But the force developed can be more in one compartment than the other. In the case of valgus knees, this force occurs at the "lateral compartment" which can cause stretching of the medial ligaments.[2] For females usually, they have wider hips which causes a larger increase of the angle at which the femur directs towards the centreline of the body.[2] Athletes have learned to manage the behavior of such problems by using ankle bands/cuffs just above the knee during a squat session.[2]

Often it becomes quite unnecessary to undergo surgery for knock knees as it usually subsides over time or it might not be a cause of concern in milder cases. This designed system can help a physician draw a correct boundary between what type of knees should require surgery and which type of knee should not require surgery but should still be classified as knock knees for medical assessment purposes.

This system can also be used as a screening tool by sports instructors to identify the severity of knock knees in them which could impair running and cause knee pain if left untreated for long.

A search of the literature yielded a lot of previous studies which had been done earlier by making use of neural networks. Neural networks had been used for predicting the movement of stock markets.[3] Neural networks have also been and are currently used in many domains of diagnostic medicine including cancer detection.[4] Pedestrian detection systems have also been tested and tried by making use of convolutional neural networks (CNN).[5] Skin cancer based detection systems have also been developed by making use of CNN to perform dermatologist level decision making.[6] A study which was done to evaluate the best approach of classification in the case of medical images showed that the use of multilayer perceptron typically yielded higher accuracy than naïve Bayes classifiers.[7]

The focus here is in investigating which images that are fed into the system are correctly classified under the label of knock knees to evaluate the performance of our network. As stated in[7] that medical legal partnership (MLP) yielded higher accuracy for categorization of medical images, we use the same algorithm but in a different classification task. The dataset is a curated list of images of patients suffering from genu valgum which have been generated from various sources of which few

[1,2]Student

[1]Department of Systems and Computer Engineering, Carleton University, Ottawa-Carleton Institute for Biomedical Engineering, Ottawa, Canada

[2]Department of Sensor and Biomedical Technology, Vellore Institution of Technology, Vellore, Tamil Nadu, India

Corresponding Author: Satyake Bakshi, Student, Department of Systems and Computer Engineering, Carleton University, Ottawa-Carleton Institute for Biomedical Engineering, Ottawa, Canada, e-mail: satyakeb@gmail.com

of them have been obtained from Physiobank which a very good repository for valid medical data. However, the major goal is the correct decision-making process on the unknown image which would be feed from a live volunteered subject.

Our classification problem would be binary for the scope of this paper, as in two cases "normal" knee and "knock knees" also called genu valgum. This second class "normal" could also be a classification between bow knees or any other types of physiological shape disorder being considered in which case it would yield a multiclass problem.

MATERIALS AND METHODS

A tuned CNN has been used. The algorithm here is the "Adam" algorithm which is a technique of performing stochastic gradient descent on the CNN network. The process is explained in the next section.

Before the CNN can be initialized there are a few considerations mainly relating to the number of images in the training and test set, what would be the learning parameter, what should be the batch size and what would be the ideal epoch for this case.

Now the modification which has been done to the CNN is that instead of supplying user-defined parameters to the CNN, A choice of parameters have been given to the CNN, and the CNN is supposed to make the "judgment" call on what parameters to pick and what would give the best accuracy and performance. The GridSearch technique has been used from Keras library in this case for the selection of the best hyperparameters in this model.

There is still a problem with the existing dataset firstly the images which have been chosen in the dataset are not of the same size and also the size of the database is very limited (20–40 images) in either of the training and test sets. There would be a problem of overfitting if the dataset is too small meaning the neural network can "remember" the training set very well but yields poor performance and accuracy on the test set.[8] To deal with this, a method called image augmentation has been performed which basically applies functions like "translation", "rotation", "shearing", "zooming in," "zooming out", etc. to create a greater number of sample images. This is very beneficial since the number of images can be augmented from the parent dataset by applying these transformations. This has been performed in the image preprocessing step.

One important point here is that the image has initially been manually separated into two folders characterizing the "knock" and "normal" knees.

Figure 1 gives the entire process of the modified network.

More literature on the working of CNN networks can be referred form this article here (Shin et al.). The steps involved in the network have been described in the representation above.

RESULTS AND EVALUATION

Figure 2 shows sample test set which has been used. The parameters are chosen by "Research" which yielded the best accuracy of 79% for this dataset which can be improved by supplying a greater number of epochs or changing the batch size. However, this is completely at the discretion of the physician or the ML expert considering the application to focus in.
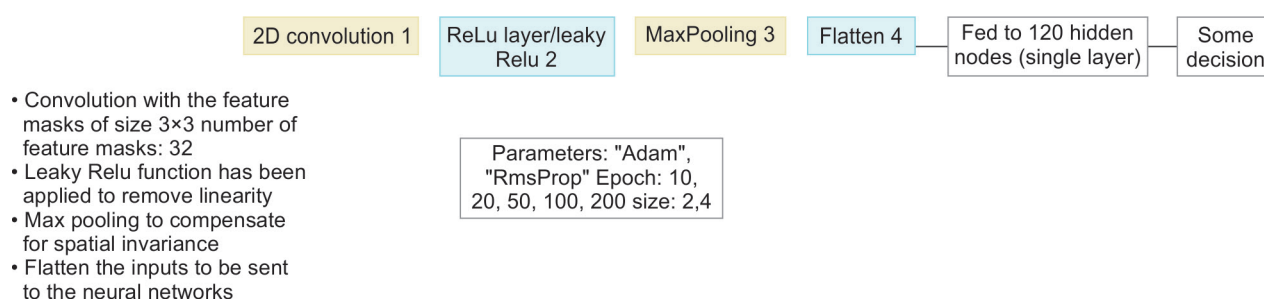
| 2D convolution 1 | ReLu layer/leaky Relu 2 | MaxPooling 3 | Flatten 4 | Fed to 120 hidden nodes (single layer) | Some decision |

- Convolution with the feature masks of size 3×3 number of feature masks: 32
- Leaky Relu function has been applied to remove linearity
- Max pooling to compensate for spatial invariance
- Flatten the inputs to be sent to the neural networks

Parameters: "Adam", "RmsProp" Epoch: 10, 20, 50, 100, 200 size: 2,4

Fig. 1: Modified CNN

Genu knee 26   Genu knee 27   Genu knee 28   Genu knee 29   Genu knee 30

Fig. 2: Sample test set (All copyrights of individual images acknowledged)

Fig. 3: Unknown images of the control sample

```
In [36]: training_set.class_indices
Out[36]: {'Genuvalgum': 0, 'Normal Knees': 1}
```
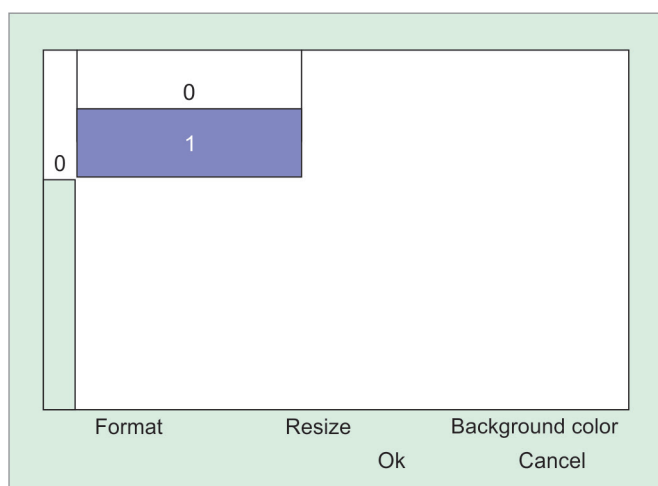
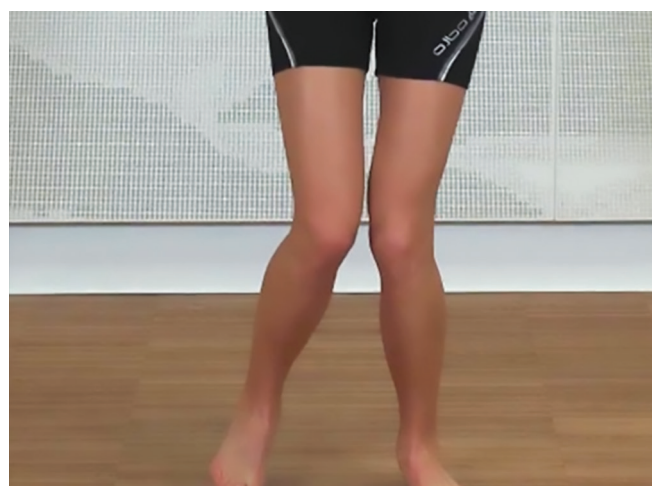

Fig. 4 Subject yielded a normal result

Fig. 5: Simulated genu valgum (from the dataset)

```
In [78]: training_set.class_indices
Out[78]: {'Genuvalgum': 0, 'Normal Knees': 1}

In [79]: if results[0][0]==1:
    ...:     print('Normal')
    ...: else:
    ...:     print('GenuValgum')
GenuValgum
```

Fig. 6: Successful outcome

Figure 3 shows unknown image which was used on the network to check the successful prediction. Subject yielded normal result (Fig. 4).

However, the subject was also made to closely simulate a genu valgum to which the resulting prediction was very successful.

Note: This study does not involve any kind of risk not greater than the daily dangers a person might face, a signed consent deemed to be not necessary by the Ethics Board.

Figure 5 shows simulated genu valgum which was done to test the validity and performance of the model. Figure 6 shows the results for this image.

CONCLUSION

Thus, the performance of the CNN has been evaluated in the automated detection of "knock knees" in patients. The results were successful for the unknown image and yielded an accuracy of 90%+. However, it did not yield good accuracy (75%) on the test set which was presented. Reasons could be due to overfitting problem which can be eliminated by choosing the parameters wisely and more analytically or by making use of grid search which has been used in this application.

This can further be modified to yield multiclass classification as decided by the physician/sport's instructor. Therefore, we conclude the successful

outcome of CNN in the classification of physiological shape disorders.

REFERENCES

1. McDade W. Bow legs and knock knees. Pediatric Clinics of North America. 1977 Nov 1;24(4):825-839.
2. 'Knock knees' a challenge for athletes, coaches. 2012 June 10. Retrieved from http://www.columbian.com/news/2012/jun/11/knock-knees-a-challenge-for-athletes-coaches/.
3. Lawrence R. Using neural networks to forecast stock market prices. University of Manitoba. 1997 Dec 12;333.
4. Egmont-Petersen M, de Ridder D, Handels H. Image processing with neural networks—a review. Pattern recognition. 2002 Oct 1;35(10):2279-2301.
5. Tomè D, Monti F, Baroffio L, Bondi L, Tagliasacchi M, Tubaro S. Deep convolutional neural networks for pedestrian detection. Signal processing: image communication. 2016 Sep 1;47:482-489.
6. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 Feb;542(7639): 115.
7. Sharma K, Kaur A, Gujral S. Brain tumor detection based on machine learning algorithms. International Journal of Computer Applications. 2014 Jan 1;103(1):7-11.
8. Tetko IV, Livingstone DJ, Luik AI. Neural network studies. 1. Comparison of overfitting and overtraining. Journal of Chemical Information and Computer sciences. 1995 Sep 1;35(5):826-833.