Classification of Human Bones Using Deep Convolutional Neural Network

Nitesh Pradhan¹*, Vijaypal Singh Dhaka² and Himanshu Chaudhary³
¹Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur-303007
²Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur-303007
³Department of Electronics and Communication Engineering, Manipal University Jaipur, Jaipur-303007
Email: *nitesh.pradhan@jaipur.manipal.edu

Abstract. In human body, there are total 206 types of different bones. Each bone has its own importance. It is very important to correctly identify human bone and then suggest treatment. To classify the human bones, we will use Musculoskeletal Radiographs (MURA) dataset. MURA dataset is one of the largest public radiographic image datasets. MURA dataset contains total 40,005 x-ray images of 14,052 patients, in which 36,808 images use as a training set and rest 3197 images use the testing set. These all images belong to seven different categories of bones such as finger, elbow, hand, forearm, humerus, wrist and shoulder. This paper aims to present a novel classification method using a deep convolutional neural network (DCNN). Dataset is freely available at https://stanfordmlgroup.github.io/competitions/mura.

1. Introduction
A machine can be learned in two different ways either as supervised learning or unsupervised learning. In supervised learning, already we have some labeled data through which we train our machine and perform classification/prediction based on training [7, 8]. In unsupervised learning, we have unlabeled data. In our case, we used supervised learning to train our machine.

Classification is a concept of supervised machine learning in which the computer learns from the labeled data [1] we input and then uses this learning to classify new data given to it. Classification is used to predict the category of data we provide [3]. This classification can be done on binary input meaning 2 categories as well as multiple categories. An example of this would be a classic Titanic question. In which we the computer learns from the given data and then tries to predict whether a person would have died or survived the Titanic crash based on the independent variables including age, gender, class etc. The classification has many applications ranging from medicine to marketing [4].

The convolutional neural network plays an important role in case of 2D/3D object prediction/classification [11]. The deep convolutional neural network also used to classify the biological images [1]. In the medical field, some task such as brain tumor segmentation, skin lesion segmentation etc can be easily performed by deep convolutional neural network [4, 8].

2. Dataset
For the classification of human bones, we require some x-ray images of the different body part. For that region, we used Musculoskeletal Radiographs (MURA) dataset. This MURA dataset contains seven different categories of human bones belonging to Elbow, Hand, Wrist, Shoulder, Finger, Humorous and Forearm. Each category of bone has some x-ray images with respect to the number of patients shown in below table.
Table 1: Classification of x-ray images in different categories

| Study            | Total No. of Patients | Total x-ray Images |
|------------------|-----------------------|--------------------|
| ELBOW Bone       | 1863                  | 5689               |
| HAND Bone        | 2104                  | 6312               |
| WRIST Bone       | 3474                  | 10502              |
| SHOULDER Bone    | 2867                  | 8571               |
| FINGER Bone      | 2031                  | 6293               |
| HUMERUS Bone     | 719                   | 1202               |
| FOREARM Bone     | 994                   | 1436               |
| **Total**        | **14052**             | **40005**          |

2.1. Training and Testing Set
To classify/prediction the category of bone, we require training and testing dataset [5]. So MURA set also contain some training and testing set in the ratio of 90:10. It means if we have total x-ray images 40,005 then 36,808 images treated as training set and remaining 3,197 x-ray images treated as test set images. Same as x-ray images, patients are also divided into training set and test set shown in table 2.

Table 2: Division of dataset based on training and testing set

| Study            | Training Set | Test Set |
|------------------|--------------|----------|
|                  | No. of Patients | No. of x-ray images | No. of Patients | No. of x-ray images |
| ELBOW Bone       | 1711         | 5233     | 152       | 456      |
| HAND Bone        | 1945         | 5835     | 159       | 477      |
| WRIST Bone       | 3267         | 9881     | 207       | 621      |
| SHOULDER Bone    | 2694         | 8052     | 173       | 519      |
| FINGER Bone      | 1865         | 5795     | 166       | 498      |
| HUMERUS Bone     | 587          | 887      | 132       | 315      |
| FOREARM Bone     | 865          | 1125     | 129       | 311      |
| **Total**        | **12934**    | **36808**| **1118**  | **3197** |

3. Model Architecture and Training
When we talk about classification, then classification is dividing into two different categories named as the binary classifier and categorical classifier. Binary classifier, classify given data in two class/category only such as yes and no or like and dislike etc. Categorical classifier, classify the given data into more than two classes/categories. Deep convolutional neural network (DCNN) model plays an important role in the case of classification/prediction [10,11].

3.1 Deep Convolutional Neural Network
The concept behind the deep convolutional neural network is to make the computer learn things like our brain does. It too has a neural network with neurons exactly like our brain. These neurons have different weights depending upon that neurons significance in classifying an object [6]. These weights get updated as our computer learns from the given data. These weights are then passed through an activation function and the output is given. An example of this would be a dog or cat classification problem. In which the machine will take thousands of images of dogs and cats and learn from them what features both have that differentiates them from one another and then predict whether the given photo is of a cat or a dog. DCNN is widely used on almost every image related problem because of its amazing accuracy that blows its competition out of water [9].
In the deep convolutional neural network, mainly three layers are used namely convolutional layer, pooling layer and flatten layer [2]. Basically, all these three layers are used to extract useful information from an image. The convolutional layer is treated as a filter that is applied to the image and get the feature map. This feature map is the reduce size (in term of dimension) from the original image. Now on this feature map, we can apply a different type of pooling such as max pooling, min pooling and avg pooling etc based on your application. So by applying to pool on feature map we get pooled feature map (i.e. more reduce size in term of dimension). After applying the pooling layer, perform flattering on the pooled feature map. This flattening is the important feature of an image and treated as input layer neuron of an artificial neural network.

As mentioned above, MURA dataset contains total 40,005 x-ray images of 14,052 patients, in which 36,808 images use as a training set and rest 3197 images use as a testing set. This MURA dataset contains seven different category of human bones belong to Elbow, Hand, Wrist, Shoulder, Finger, Humorous and Forearm. Our task is to identify the class/category of a particular image. To classification, the model architecture uses four convolutional layers, four pooling layers, one flattened layer, and two hidden layers with dropout rate 0.2. The complete architecture of DCNN is described in table 3.

| Layer (type)            | Output Shape            | Param # |
|-------------------------|-------------------------|---------|
| conv2d_4 (Conv2D)       | (None, 62, 62, 32)      | 896     |
| max_pooling2d_4 (MaxPooling2) | (None, 31, 31, 32)      | 0       |
4. Result
When we train our network by 36808 x-ray images with different epochs then we got the different accuracy of the training set. After training, we test our trained network by 3197 x-ray images and found good accuracy. The accuracy of training and testing set with respect to multiple epochs are mentioned in table 4.

| No. of epoch | Training Accuracy (in %) | Testing Accuracy (in %) |
|--------------|--------------------------|-------------------------|
| 3            | 86.23                    | 85.11                   |
| 7            | 91.62                    | 90.40                   |
| 12           | 93.37                    | 91.24                   |
| 18           | 94.23                    | 91.37                   |

As we can see in table 4, in 18 epochs we got 94.17% accuracy it means if we give an image of bone to the machine, 94.17% chance that the machine will predict the correct class/category of bone.
5. Conclusion
We successfully applied deep convolutional neural network for classification of human bone with four convolutional layers, four pooling layers, one flattened layer, and two hidden layers with dropout rate 0.2. We run this model architecture for multiple epochs and got the best accuracy 94.23% for the training set and 91.37% for the test set.

In this paper, we used only seven different classes/categories of bones such as Elbow, Hand, Wrist, Shoulder, Finger, Humorous and Forearm. In the further study, the number of classes/categories can be increased to predict/classify the different types of human bones.

6. Acknowledgements
We would like to acknowledge the Stanford Program for Artificial Intelligence in Medicine and Imaging for clinical dataset infrastructure support.

References
[1] Carlos, A., AndréLuis, D., Fábio, H., André, C., and Leon “Deep learning for biological image classification” Expert Systems with Applications 2017, pp. 114–122
[2] Grewal, Monika, Srivastava, Muktabh Mayank, Kumar, Pulkit, and Varadarajan, Srikrishna “Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in ct scans” 2017, arXiv preprint arXiv:1710.04934.
[3] A. Rajkomar, S. Lingam, A. G. Taylor, M. Blun, and J. Mongan, “Highthroughput classification of radiographs using deep convolutional neural networks” J. Digit. Image 2017 vol. 30, no. 1, pp. 95–101.
[4] Sergio, P., Adriano, P., and Carlos, A “Brain Tumor Segmentation using Convolutional Neural Networks in MRI Images” IEEE Transactions on Medical Imaging 2016, pp. 1240-1251.
[5] H.-C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures dataset characteristics and transfer learning" IEEE Trans. Med. Image May 2016, vol. 35, no. 5, pp. 1285-1298.
[6] Nima, T., Jae, Y., Suryakanth, R., Todd, H., and Michael, B “Convolutional Neural Networks for Medical Image Analysis: Fine Tuning or Full Training” IEEE Transactions on Medical Imaging 2016, pp. 1299-1312.
[7] Criminisi, A “Machine learning for medical images analysis” Medical Image Analysis 2016, pp. 91–93.
[8] Jafari, M.H., Karimi, N., Nasr-Esfahani, E., Samavil, S., Soroushmehr, S.M.R., Ward, K. and Najarian, K “Skin Lesion Segmentation in Clinical Images Using Deep Learning” 23rd International Conference on Pattern Recognition (ICPR) 2016, pp. 337-342.
[9] W. Shen, M. Zhou, F. Yang, C. Yang, J. Tian, "Multi-scale convolutional neural networks for lung nodule classification", in Information Processing in Medical Imaging June 2015 vol. 24, pp. 588-599.
[10] A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet classification with deep convolutional neural networks", Proc. Adv. Neural Inf. Process. Syst 2012, pp. 1097-1105.

[11] R. Socher, B. Huval, B. Bath, C. D. Manning, and A. Y. Ng, "Convolutional-recursive deep learning for 3D object classification" in Proc. Adv. Neural Inf. Process. Syst 2012, pp. 656–664.