**Final Project Report**

**Musculoskeletal Radiographs Abnormality Detection**

**Introduction:**

HealthCare, it is very demanding field and requires upgraded and latest technologies which is beneficial for all the stakeholders. There are many fundamentals included in the term related to “HealthCare”. Musculoskeletal (bone) X-ray is an essential tool in diagnosing the abnormalities. In recent years, deep learning algorithms have increasingly been applied in musculoskeletal radiology and have produced remarkable results. Here in this study, we used **MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs,** containing 40,561 images from 14,863 studies, where each study is manually labeled by radiologists as either normal or abnormal. To evaluate models robustly and to get an estimate of radiologist performance.

The aim of this study is to investigate new model architectures and deep transfer learning to improve the performance in detecting abnormalities of upper extremities in limited data environments and allow them to serve as initial screening tools to prioritize studies for expedited review. Focus is placed on the MURA Elbow dataset, where model had the lowest performance. Improvements using this approach should translate into performance improvements when deployed to other sites.

A picture containing text, different, colorful, several

Description automatically generated

Increasing radiologist workloads and primary care radiology services make it relevant to explore the use of artificial intelligence (AI) and particularly the deep learning to provide diagnostic assistance to radiologists with a better and improved quality for primary care.For abnormality detection, image classification is required, and it is possible to create with the help of Neural Network models. One of the most popular deep neural networks is the Convolutional Neural Network (CNN). We used CNN network here as it gets the highest accurate results while deploying the models to the real world. We also show that our model produces high quality image samples. DenseNet-169: Deep Learning with Depth wise Separable Convolutions models were implemented as a baseline model here on the “Elbow” radiographs.

**Related Work:**

We worked on the dataset of MURA (Musculoskeletal Radiographs Abnormality Detection) is a large dataset of bone X-rays. MURA is an Algorithms are tasked with determining whether an X-ray study is normal or abnormal. MURA is one of the largest public radiographic image datasets.

We started with the research on some study paper for the abnormality detection of bones in the upper body parts. Then we decided to work on the “Elbow”, but the difficult task was we need to load the images of “Elbow” only. We worked on the images to identify the elbow images. We implemented our baseline model “DenseNet-169”, we were looking forward to getting the better prediction accuracy and so that we research the possible ways for model implementation and evaluation to get better accuracy. By using the deep learning, we decided to implement on simple CNN. But the result was not satisfactory. We tried Auto Encoder just as a practice and we implemented it and we got an average accuracy, but the thing is in the actual scenario the Auto Encoder is not very suitable for the image processing. We were looking for a model which gives a better accuracy, we did a lot of research along with the model trials. At the end, got an idea from one of the study papers and we implemented “Multi-Scale Convolution Neural Network (MSCNN) with fully connected network”. We made it as our final model of this study.

**Methods:**

**Dataset Access:**

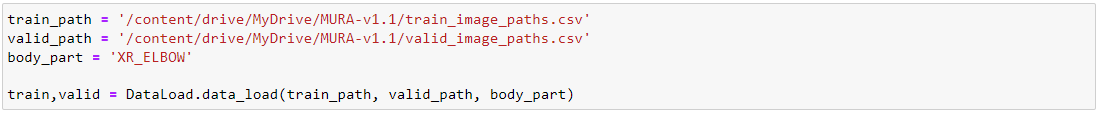
In the very first place we didn’t have a dataset access, so we requested for the access and got approval for the same. After that we downloaded the MURA dataset.

**Load data:**

We used “Google Collaboratory Pro” to run the python notebook and for that we loaded our data into the Google Drive. It was a large dataset, so it took more than 5 hours to load into google drive.

**Data Preprocessing:**

MURA dataset contains 7 body parts to detect abnormalities. We worked on “Elbow” and to do so, we fetch the images of elbow. We passed a string which contains “Elbow” from the image path of test label and validation label, which will fetch the elbow images.



Then we split the data into two parts; one is “training set” and other is “validation set”. Where training set contains 70% and validation set contains 30% from the whole data set.



We normalize all the images by giving a fixed size for the images. We gave pixel size of 177 x 224 for all Elbow images.

**Exploratory Data Analysis:**

Eventually we didn’t have any specific Exploratory Data Analysis, but we load train and validation data by using the flag 0 or 1 for the Elbow. Here, 0 contains no abnormality detection and 1 contains abnormality detection.

**Create Baseline Model:**

We created DenseNet-169 baseline model for training. We wanted to get a better accuracy than the baseline model, we research for the models to implement and evaluate. We interested in a model that give high accuracy compared to the baseline model.

**Model used for Evaluation on Training Data:**

By having a research, code implementation, and model execution we finalize our main model for this study was Multi-Scale Convolution Neural Network (MSCNN) with fully connected network. We got this idea from one of the study papers where they mentioned that MSCNN is the model which gives better, and highest accuracy compared to other deep learning models.

**Results:**

For our final model, we have got the average model prediction accuracy 60% while implementing the MSCNN model with 15 Epochs. All the implemented and evaluated models are listed below with the accuracy we got.

|  |  |
| --- | --- |
| **Model** | **Accuracy** |
| DenseNet-169 | 50.53% |
| Simple Convolutional Neural Network (CNN) | 67% |
| Auto Encoder | 75% |
| Multi-Scale Convolution Neural Network (MSCNN) with fully connected network | 60% (15 Epochs) |

**Discussion:**

Initially when we started to work on this study, we needed to explore more about model and their prediction accuracy. At very first we tried Dense-Net 169 model, but it comes with only 50% accuracy, so we decided to make it as baseline. But for the final model, we tried and implemented various models like Xception, Simple CNN, Auto Encoder but we did not get any satisfactory results from that. It was very challenging for us to decide and evaluate our final model. By taking a reference of one study, we implemented MSCNN with fully connected network and we ran only 15 Epochs because it is very time-consuming process.

**Conclusion:**

We got an average accuracy and predictive model from this dataset for elbow by studying some research papers and implementing our ideas for the model selection. We selected MSCNN fully connected network model and where we tried different features, different ideas so that we can get better result. At the end we got an improved result, but it does not meet our actual goal. We will try and work on our final model by training it in a different way and by using different features to get higher accuracy.

**Contribution:**

To accomplish the project, we worked as a team. We all contributed equally in terms of research, browsing the ideas, implement and evaluation of the models, exploring features and models, all the study related documents.

**References:**

:

* [Data source](https://stanfordmlgroup.github.io/competitions/mura/)
* [Research paper](https://www.mdpi.com/1424-8220/20/11/3153)
* [Research Paper2](https://arxiv.org/abs/1712.06957)
* [AI for Medical Diagnosis Course](https://www.coursera.org/learn/ai-for-medical-diagnosis/home/welcome)

**Appendices:**

Simple CNN:

Graphical user interface, application, table

Description automatically generated

- Loss and Accuracy graph

Chart, line chart, histogram

Description automatically generated

Auto Encoder:

A picture containing table

Description automatically generated

MSCNN:

Table

Description automatically generated