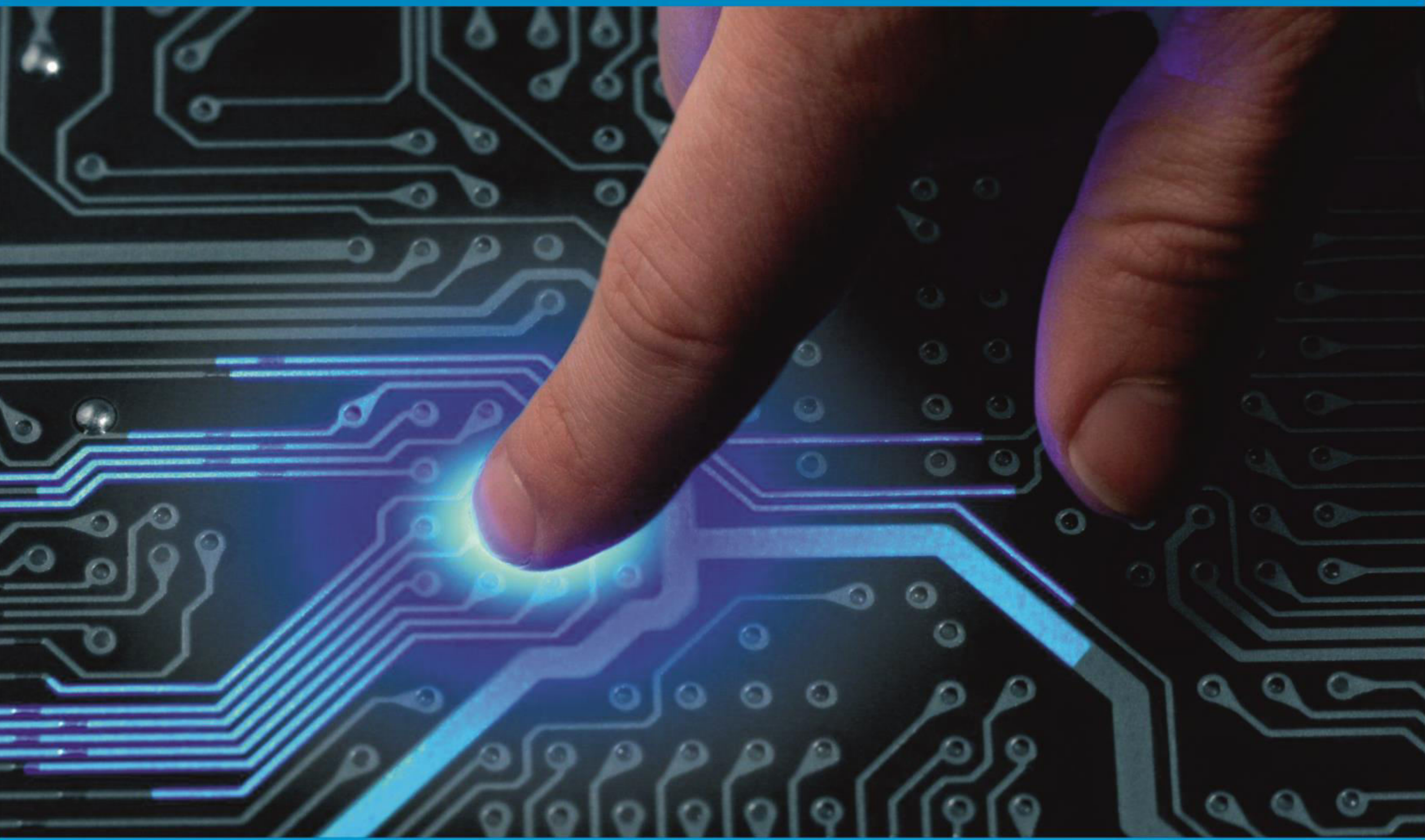
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 3, March 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903076 |

Musculoskeletal Abnormality Detector Using VGG16 Model

Mohamed Arif V A, Nithesh V, Prakash R, Manish M, Ms. Keerthika J

UG Student, Dept. of CSE., Sri Eshwar College of Engineering, Coimbatore, India

UG Student, Dept. of CSE., Sri Eshwar College of Engineering, Coimbatore, India

UG Student, Dept. of CSE., Sri Eshwar College of Engineering, Coimbatore, India

UG Student, Dept. of CSE., Sri Eshwar College of Engineering, Coimbatore, India

Assistant Professor, Dept. of CSE., Sri Eshwar College of Engineering, Coimbatore, India

ABSTRACT: The monotonous routine of medical image analysis has always led to work debility for many medical practitioners under tight time constraints. Medical image analysis is susceptible to errors, which increases the chances of a wrong treatment being recommended. While the improvisation of complex deep learning models has reached performance beyond human capability in some computer vision works, widespread adoption in the bio-medical field has been held back, among other factors, by poor model accountability and a lack of high-quality labelled data. This paper provides a model analysis and visualisation system for analysing a deep convolutional neural network's feature extraction mechanism and applies it to abnormality detection using the musculoskeletal radiograph dataset (MURA, Stanford). The proposed framework provides a mechanism for interpreting VGG16 deep learning architectures. It aims to provide a deeper insight about the paths of feature generation and reasoning within a VGG16 model architecture. When evaluated on MURA at abnormality detection, the model interpretation framework has been shown that it is able to identifying limitations in the reasoning of an ImageNet architecture applied to radiography, which can in turn be improved through model interpretation and visualisation.

KEYWORDS: Bone, Radiographs, Muscularity, VGG16, Neural Network Algorithm

I. INTRODUCTION

Radiography may be a medical imaging technique that uses X-rays to get images of human anatomy. X-rays have now been used for quite 100 years for various applications like detection of fractures or detection of obstruction of the duodenum. The rays penetrate the thing of interest, say a wrist or an elbow, and are attenuated consistent with the density of the fabric they travel through. Thus, denser materials, like bones, present a better attenuation than softer materials, like muscles or organs. The density is usually indicated in images as various levels of brightness (dense) or darkness (soft), although in some cases the dimensions are reversed. It is important to note that X-ray images are two-dimensional projections of three-dimensional bodies, and thus in some cases, two or more images are acquired with different angles of view like posterior-anterior (e.g. Figure 1(a)) or lateral views (e.g. Figure 1(b)).Metallic objects appear much brighter and denser than bones. Radiographic images are generally analyzed by certified radiologists or specialists reporting radiologists. On a typical day, one analysis is performed under a decent time constraint alongside other heavy workloads. It has been reported that employment fatigue can cause errors in interpretation. Therefore, it's attractive to explore computational approaches which will support radiologists within the analysis of X-ray images. Recent advances within the areas of machine learning, computer vision, and particularly deep learning model development have reported results where the computational models have surpassed human performance on many predictive tasks. In 2012, the Alexnet won the ImageNet Large Scale Visual Recognition Competition (ILSVRC) by a large margin against the runner-up. More experienced approaches have improved the efficiency obtained with Alexnet. VGG16 developed by Karen Simonyan, has outperformed human prediction which brought remarkable results for the ImageNet Challenge. Despite their superior performance, widespread adoption of those complex deep learning models has been held back primarily thanks to poor model interpretability, the need for a really sizable amount of annotated cases to coach the models, and of sophisticated computational resources. The lack of model transparency of its predictions, especially, has led to challenges in obtaining regulatory approval to deploy in life-critical applications for healthcare. The contribution of this paper is two-fold: (1) The proposed framework allows a systematic approach for

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903076 |

analyzing the fracture in the X-Rays using the VGG16 model; (2) The effectiveness of the approach is evaluated using MURA, showing that the framework is capable of identifying failures.

II. RELATED WORK

A model for classifying and localizing masses on mammogram images using a convolutional neural network. Class activation maps were calculated using a VGG16 CNN architecture. The model utilized transfer learning and fine-tuned on pre-trained CNNs of image patches that are cropped. The CNN is fed a complete mammography image as input, with patch images and class activation maps computed for localizing abnormalities. Different accessible convolutional neural network architectures were explored and tested in the study based on document strengths such as top-1 accuracy, top-5 accuracy, model size, number of parameters, and network depth, and this was documented. The single models and ensembles were built using the identified CNN architectures. The MURA dataset was preprocessed to create a structure and format that could be used for training and validation of various CNN models. In the MURA dataset, the empirical evaluation results of the various models were reported and compared to the radiologist's efficiency. We did more data-augmentation at the training stage because using only horizontal flip and rotation will cause overfitting. So, we further do random-sized crop in the pre-processing stage. In other words, each input image is resized to 256 × 256 and then random-cropped to 224 × 224 before feeding into the network. At test time, we do center crop instead. We use the weights pre-trained on Imagenet and tune the hyper-parameters. After trying different batch sizes 8, 16, 32, 64, 128 and the corresponding learning rate, we chose to use batch size 16 and an initial learning rate of 0.0001 that is decayed by a factor of 10 each time the validation loss plateaus after an epoch as the hyperparameters, which perform the highest accuracy and AUC of ROC on validation.

III. PROPOSED ALGORITHM

A. *Source of Data:*

The musculoskeletal radiographic dataset for this research is made up of 14,863 upper extremity studies from 12,173 patients, with a total of 40,561 multi-view radiographic images from Stanford Hospital. The dataset consists of study types of the finger, elbow, humerus, forearm. shoulder, wrist and hand. Here, we have examined specifically on elbow, finger, humerus and forearm. The studies are labelled abnormal or normal by radiologists manually, and split into training and validation sets for evaluation.

B. *Existing System:*

The typical way of analysing an X-ray of a human is manually done by the Radiologist in-order to detect the abnormality present in it. Usually the Radiologist spot the abnormality using the bright light. The radiologists achieved their highest performance on either wrist studies or humerus studies, and their lowest performance on finger studies. The VGG16 also achieved its highest performance on elbow studies and its lowest performance on finger studies. Detecting abnormality in these musculoskeletal radiographs is a very tedious task. If this process in automatic detection of abnormality can be introduced, it would be really helpful for further treatment and diagnosis. In this regard, MURA dataset was published with 40,561 images from 14,863 studies. In this paper, a neural network is designed to classify normal and abnormal condition and compared the result with VGG16 architecture.

C. *Proposed System:*

Using machine learning algorithms for medical imaging has been a go-to practice in past two years, Since the only issue we are facing now to this modern approach is lack of datasets for making machine's understand complexities in images. For the issue of detecting fractures in a bone Stanford open-sources a dataset named MURA which has been a large dataset of X-ray images of upper body extremity. We have tried to use that dataset and train a machine learning algorithm for detecting fractures in images. With this approach with a decent training time for our convolutional neural network this approach for a medical image seems promising.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903076 |

IV. SIMULATION RESULTS

In this study, training and test data use features created by board-certified radiologists from the Stanford Hospital at the time of clinical radiographic interpretation between 2001 and 2012. Model performance was evaluated on elbow data after training on 5 epochs.

Evaluating the model performance of elbow test set, it shows accuracy greater the 75% in detecting abnormalities. The amount of loss it has produced is significantly reduced as it nears around 50%. This loss can be reduced further when trained more. The following graph shows the accuracy and loss of the model performance on elbow data.

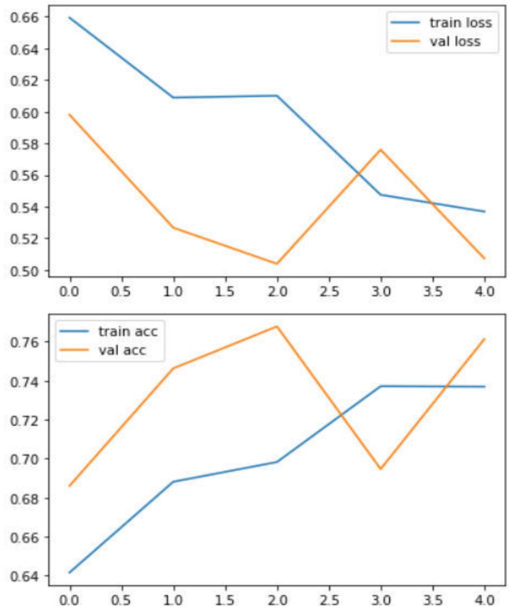


Figure 2: Graph on accuracy and loss on elbow data.

Model performance was evaluated on fingers test data set after training on 7 epochs. On evaluating these data, fingers model performance was not that promising, with VGG16 model showing less than 75% of accuracy and few losses as well as compared with elbow data set. The following graph represents the loss and accuracy of the finger data set on evaluation.

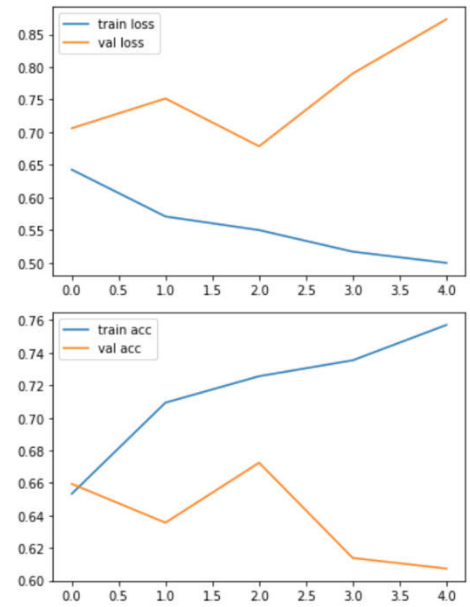


Figure 3: Graph on accuracy and loss on finger data.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903076 |

After training on 5 epochs on humerus test data the model performance was evaluated. The result shows that this performance was quite fair as accuracy hits 56% and loss has 80%. The VGG16 model can be further trained to increase their model accuracy. The graph tells us the variation it occurred during each epoch of the model.

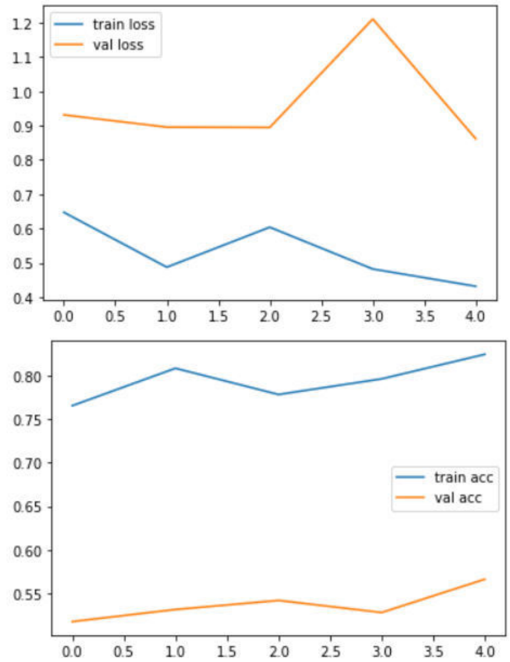


Figure 4: Graph on accuracy and loss on humerus data.

The forearm test data set's model performance was evaluated. The evaluation shows that the dataset is trained with 75% accuracy. The graph is represented below

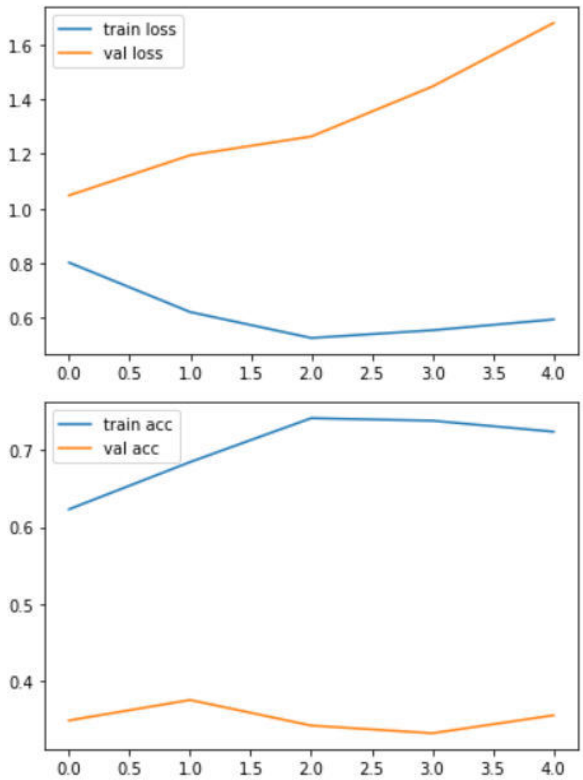


Figure 5: Graph on accuracy and loss on humerus data.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903076 |

V. CONCLUSION AND FUTURE WORK

The VGG16 model, employing deep transfer learning to optimize training on limited data, detected abnormalities in humerus radiographs with 75% accuracies, allowing for these models to serve as useful initial screening tools to prioritize studies for expedited review. The deployed models can be retrained on new data on a regular basis to boost accuracies on complex pathologies, normal variants, and site-specific variations. Finger radiograph performance was less promising, likely due to the limitations of significant inter-radiologist variance. The causes of this variance should be investigated further using machine learning methods, which may lead to effective remediation.

REFERENCES

1. Rajpurkar, P; Irvin, J; Ball, R L; Zhu, K; Yang, B; Mehta, H; Duan, T; Ding, D; Bagul, A; Langlotz, C.P; et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNet algorithm to practicing radiologists.
2. Gulshan, V; Peng, L; Coram, M; Stumpe, M.C; Wu, D; Narayanaswamy, A; Venugopalan, S; Widner, K; Madams, T; Cuadros, J; Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus PhotographsAccuracy of a Deep Learning Algorithm for Detection of Diabetic RetinopathyAccuracy of a Deep Learning Algorithm for Detection of Diabetic Retinopathy.
3. Thomas Davenport and Ravi Kalakota. The potential for artificial intelligence in healthcare. Future Healthcare Journal, 6, 2019.
4. United States Bone and Joint Initiative: The Burden of Musculoskeletal Diseases in the United States (BMUS). Available online: http://www.boneandjointburden.org (accessed on 14 June 2019).
5. Mormont, R; Geurts, P; Marée, R. Comparison of deep transfer learning strategies for digital pathology.
6. David Walsh, "Recent progress in radiography," The Hospital, vol. 27, no. 680, pp. 57, Oct 1899.