International Journal of Academic Health and Medical Research (IJAHMR)
ISSN: 2643-9824
Vol. 6 Issue 8, August - 2022, Pages: 82-87

Bone Abnormalities Detection and Classification Using Machine Learning Techniques – Literature Review

Alaa M. A. Barhoom[1], Mohammed Rasheed J. AL-Hiealy[2], and Samy S. Abu-Naser[3]

[1,2]University Malaysia of Computer Science & Engineering (UNIMY), Cyberjaya, Malaysia.
[3]Faculty of Engineering and Information Technology, Al-Azhar University, Gaza, Palestine.
[1]Corresponding author: p05210001@student.unimy.edu.my

Abstract. *Bone problems is meant to be any of the injuries that affect human bones. Bones injuries are a major cause of abnormalities of the human skeletal system. Although physical injury, causing fracture, controls over disease, nonetheless fracture is one of a few common causes of bone problems. Bone fracture can circuitously cause death, as fractures and their associated difficulties can in some situation produce a descending spiral in health. About 20% of the hip fracture patients passed away within a year of the fracture. The goal of the study was to carry out a literature review of bone abnormalities that used machine learning as a mean of detection and classification of bone abnormalities, by reviewing technique which included study reference, approaches, dataset, programming language used, and the top accuracy achieved. 17 studies published between 2001 and 2020 were designated for analysis. Seven of which employ X-ray images, five employ CT images and five employ MRI Images. The outcomes indicated that amongst the classification methods studied, artificial neural networks, Convolutional Neural Network and Residual neural network extensively used with average accuracy between 73.40% and 94.71%. Most of the designated studies employed X-ray images. The result is that it is possible in the future to develop a deep learning model to identify bone abnormalities with more accurate outcomes, by adjusting the methods used and adding more actual processes.*

1. INTRODUCTION

Bone disease, any of the diseases or injuries that affect human bones. Diseases and injuries of bones are major causes of abnormalities of the human skeletal system. Although physical injury, causing fracture, dominates over disease, fracture is but one of several common causes of bone disease, and disease is in fact a common cause of fracture [1,2].

Just like other tissues and organs in our body, bones can be affected by medical conditions too. These include things like fractures, signs of wear and tear, inflammations, and cancer. Injuries and fractures are common in younger people. As we grow older, diseases like osteoporosis and osteoarthritis are more likely to develop. Various tests and examinations can be used to find out what is causing problems like pain or difficulties moving. If the doctor thinks you may have a particular problem, various imaging techniques and other tests might be considered [3].

Fractures are by far the biggest problem caused by bone disease. They are common and can be quite debilitating. In many cases they are the first sign of the disease in patients. An estimated 1.5 million individuals suffer a fracture caused by bone disease annually. In fact, fractures are the most common musculoskeletal condition requiring hospitalization among Medicare enrollees (who are age 65 and older). Approximately one-third (500,000) of fracture patients are hospitalized, while many more suffer fractures that do not require or result in hospitalization. The risk of fracture increases dramatically with age in both sexes, both because bones become more fragile and the risk of falling increases. Roughly one in four (24 percent) women aged 50 or older fall each year, compared to nearly half (48 percent) of women aged 85 or older; comparable figures for men are 16 percent and 35 percent. These falls can result in fractures, and when they do the fracture is almost always caused by low bone mass or osteoporosis [4-6].

Bone disease can indirectly lead to death, as fractures and their associated complications can in some cases trigger a downward spiral in health. In one population sample, the risk of mortality was four times greater among hip fracture patients during the first 3 months after the fracture than was the comparable risk among fracture-free individuals of similar gender and age. Approximately 20 percent of the hip fracture patients died within a year of the fracture [7].

In this paper, several algorithms are reviewed of deep learning techniques used in the medical field, especially the detection of diseases in the various medical images of humans. The goal is to contribute in the future to building an effective model for detecting some of human bone abnormalities based on X-rays to help the doctor in decision-making easily, accurately, and quickly.

2. REVIEW OF LITERATURE

Researchers have written numerous papers to employ artificial intelligence, expert systems, and neural networks in detecting some of human bone abnormalities based on X-rays to help the doctor in decision-making easily, accurately, and quickly.

In the normal situation, hospitals use MRI and CT scans often give better pictures of them. Still, x-rays are fast, easy to get, and cost less than other scans.

Several methods and models have been proposed that have contributed to increasing the accurately of the diagnosis.

Relevant literature was searched that used machine learning models, especially deep learning models, in detecting human diseases within different medical images of patients, to answer the questions: What are the best machine learning models proposed by researchers for detecting human diseases in medical images (such as X-ray images) in terms of the accuracy of the results? Why did the accuracy of the results differ in the models proposed by the researchers? What type and volume of data are used in the proposed deep learning models, and does it relate to the accuracy of the results? using some keywords like deep Learning, machine learning, disease detection, diagnosis, artificial neural networks, convolutional neural networks, best accuracy, techniques, trained models, pre-trained models, method, x-ray, dataset, classification, and others through useful databases to search for journals and articles such as: Google Scholar, ResearchGate, and others, published between 2002 and 2020, and to narrow the search are Boolean operators used, then the list of references for each research was checked to find other sources that used deep learning models to detect diseases in radiographs of human patients, important sources are called after an initial evaluation is done through reading to ensure that the research answers one of the previous questions and its credibility.

The review is organized according to the Machine learning (ML) algorithms used for the diagnosis, which include Convolutional Neural Network (CNN) search and Residual neural network (RNN).

3. CONVOLUTIONAL NEURAL NETWORK (CNN)

Esteva et al. [8] works in Object classification, localization, and detection, refer to identifying the type of an object in an image, the location of objects present, and both type and location simultaneously, The first working core with GPU powered deep learning approach, in 2012, by ImageNet Large-Scale Visual Recognition Challenge, the challenge, which included many researchers aspiring to improve and develop computer vision, this resulted to implementing deep learning models, all of which were successful and accurate results, especially in the medical field and disease detection, the accuracy of classification and detection of these models was similar to the results of the diagnosis of the ordinary doctor, and sometimes exceeded their levels, the greatest credit is due to the Convolutional Neural Network (CNN), which has the ability to extract differences and analyse data from a huge set of data that divided tasks and was able to collect similar images and other tasks, (CNN) type of deep learning algorithm which hardcodes translational invariance, a key feature of image data. But the challenges of the medical field need more work with and development of different models of deep learning.

Cernazanu et al. [9] were able to train a convolutional neural network on a dataset consisting of medical images such as X-ray images and CT images containing specific diseases, using some operations such as convolutional, fully connected layers, and pooling, the convolutional neural network receives the images as input, then it will convert images to flattened vectors in the end the softmax layer represent the elements of the output vector, which actually represents the probability of detecting the disease in the images, during the training process, the internal parameters of the network layers are iteratively adjusted to improve accuracy. Typically, lower layers learn simple image features edges and basic shapes which influence the high-level representations, so that the training outputs are the answer to the question: is there a disease or not? For example, normal or abnormal.

Sitaula et al. [10] in light of the spreading epidemic of Corona virus disease, a group of researchers retrained a deep learning model based on a convolutional neural network called VGG-16 by using the attention module, that can capture the spatial relationship between the ROIs in CXR images, which could identify the likely regions of COVID-19's effect in the human lungs, with appropriate convolution layer 4th pooling layer, they designed a novel deep learning model to perform fine-tuning in the classification process, three sets of images were used to make the evaluation process, and the results were satisfactory, and we can rely on deep learning models to detect infection with the Corona virus within lung images.The researchers relied on VGG16 model for two important reasons: firstly, it extracts the features at low-level by using its smaller kernel size, secondly, it has a better feature extraction ability for the classification of COVID-19 CXR images, used the pre-trained weight of ImageNet. It helps to overcome the over-fitting problem as they had limited amount of COVID-19 CXR images for training purpose, used the four main building blocks in model: Attention module, Convolution module, FC-layers, and Softmax classifier, after training the model on the dataset with a total of 445 images divided into three training and testing categories, the accuracy of the model work was satisfactory by 79.58%. This indicates success VGG-16 Model Learning and training on a small and accurate set of data.

Esteva et al. [11], have examine the strength of deep learning approaches for pathology detection in chest radiographs to explore the ability of CNN learned from a non-medical dataset to identify different types of pathologies in chest x-rays, after tested algorithm on a 433-image dataset, the best performance was achieved using CNN and GIST features, they got an accuracy rate 0.87-

0.94 for the different pathologies. The results demonstrate the feasibility of detecting pathology in chest x-rays using deep learning approaches based on non-medical learning.

Gulshan et al. [12] they developed the EyePACS-1 data set consisted of 9963 images from 4997 patients (mean age, 54.4 years; 62.2% women; prevalence of RDR, 683/8878 fully gradable images [7.8%]); the Messidor-2 data set had 1748 images from 874 patients (mean age, 57.6 years; 42.6% women; prevalence of RDR, 254/1745 fully gradable images [14.6%]). For detecting RDR, the algorithm had an area under the receiver operating curve of 0.991 (95% CI, 0.988-0.993) for EyePACS-1 and 0.990 (95% CI, 0.986-0.995) for Messidor-2. Using the first operating cut point with high specificity, for EyePACS-1, the sensitivity was 90.3% (95% CI, 87.5%-92.7%) and the specificity was 98.1% (95% CI, 97.8%-98.5%). For Messidor-2, the sensitivity was 87.0% (95% CI, 81.1%-91.0%) and the specificity was 98.5% (95% CI, 97.7%-99.1%). Using a second operating point with high sensitivity in the development set, for EyePACS-1 the sensitivity was 97.5% and specificity was 93.4% and for Messidor-2 the sensitivity was 96.1% and specificity was 93.9%.

Moreno et al. [13] presented first findings towards assessing how computer vision, natural language processing and other systems could be correctly embedded in the clinicians' pathway to better aid on the fracture detection task. We present some initial experimental results using publicly available fracture datasets along with a handful of data provided by the National Healthcare System from the United Kingdom in a research initiative call. Results show that there is a high likelihood of applying trans-fer learning from different existing and pre-trained models (VGG16, Resnet50, InceptionV3) to the new records provided in the challenge, and that there are various ways in which these techniques can be embedded along the clinicians' pathway, The accuracy rate By VGG16 model was 92.7 % in one of the stages.

Moran et al. [14] introduced model to classify regions in periapical examinations according to the presence of periodontal bone destruction. This study considered 1079 interproximal regions extracted from 467 periapical radiographs. This data was annotated by experts and used to train a ResNet and an Inception model, which were after evaluated with a test set. Inception presented the best results and an impressive rate of correctness even on the small and unbalanced dataset. The final accuracy, precision, recall, specificity, and negative predictive values are 0.817, 0.762, 0.923, 0.711, and 0.902, respectively. The ROC and PR curves also demonstrate the good performance of both models. These results suggest that the evaluated CNN model can be used as a clinical decision support tool to diagnose periodontal bone destruction in periapical exams.

El-Saadawy et al. [15] presents a method for detecting the fractures in the seven extremity upper bones (shoulder, humerus, forearm, elbow, wrist, hand, and finger) using X-ray images. A two-stage classification method based on MobileNet network is proposed. Enhanced X-ray image is fed into the first stage to detect bone type. Thereafter, the bone image is directed according to the result of the first stage to one of seven classifiers (one for each bone type) to detect the abnormality in the bone. MURA dataset is utilized as a performance dataset and average accuracy 73.42% has been achieved after merging the two classification stages.

Ananthu et al. [16] employed in many medical imaging applications for the diagnosis of diseases. However, one of the key issues is the limited availability of microscopic images for training the models. To overcome this difficulty, transfer learning techniques are put forward, researchers present a comparative analysis of different transfer learning models like MobileNet to detect acute lymphocytic leukemia (ALL) from blood smear cells. All models were trained on ALL-IDB2 dataset and achieved an accuracy of 97.88%, from MobileNet model.

Wang et al. [17] tried in this search to improve and speed up bone age assessments by using different object detection methods to detect and segment bones anatomically important for the assessment and using these segmented bones to train deep learning models to predict bone age. A dataset consisting of 12811 X-ray hand images of persons ranging from infant age to 19 years of age was used. In the first research question, compared the performance of three state-of-the-art object detection models: Mask R-CNN, Yolo, and RetinaNet, was selected the best performing model, Yolo, to segment all the growth plates in the phalanges of the dataset, proceeded to train four different pre-trained models: Xception, InceptionV3, VGG19, and ResNet152, using both the segmented and unsegmented dataset and compared the performance, achieved good results using both the unsegmented and segmented dataset, although the performance was slightly better using the unsegmented dataset. The analysis suggests that we might be able to achieve a higher accuracy using the segmented dataset by adding the detection of growth plates from the carpal bones, epiphysis, and the diaphysis. The best performing model was Xception, which achieved a mean average error of 1.007 years using the unsegmented dataset and 1.193 years using the segmented dataset.

4. RESIDUAL NEURAL NETWORK (RNN)

Wang et al. [17] proposed a new residual neural network to identify the pathological type of lung cancer via CT images. Due to the low amount of CT images in practice explored a medical-to-medical transfer learning strategy by pre-trained on public medical images dataset luna16, and then fine-tuned on our intellectual property lung cancer dataset collected in Shandong Provincial

Hospital. Data experiments show that method achieves 85.71% accuracy in identifying pathological types of lung cancer from CT images and outperforming other models trained with 2054 labels that method performs better than AlexNet, VGG16 and DenseNet, which provides an efficient, non-invasive detection tool for pathological diagnosis.

5. SUMMARY OF THE PREVIOUS STUDIES

Table 1 presents a summary of the most important previous studies that were discussed, where the reference, machine learning methodology, detailed information about the dataset used, the programming language used, if any, and the best accuracy obtained.

Table 1: Summarizes of the most important discussed previous studies

Study Reference	Methods	Dataset			Programming Language	Best Accuracy
		Provider	Size	Attributes		
(Russakovsky, O. et al., 2015)	CNN	Flickr, other search engines.	200,000	MRI images	-	94.7
(Cernazanu-Glavan, C., & Holban, S., 2013)	CNN	DICOM	16,384	X-ray images	Python	Good
(Esteva, A. et al., 2017)	CNN	ISIC	129,450	X-ray images	-	72.1
(Sitaula, C., Hossain, M.B., 2021)	CXR, CNN	GitHub, RSNA	3,901	X-ray images	Python	79.58
(Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen & H. Greenspan., 2015)	CNN, GIST	Sheba Medical Center.	433	CT images	-	94.0
(Dimitris, K., Ergina, K., 2017)	RNN	PubMed Central.	164,614	MRI images	Caffe framework	85.83
(Korolev, A. Safiullin, M. Belyaev & Y. Dodonova., 2017)	VoxCNN, RNN	ADNI	231	MRI images	-	88.0
(N. Dhungel, G. Carneiro and A. P. Bradley., 2017)	mResNet, CNN	INbreast	410	CT images	-	79.0
(Wang, Shudong, Dong, Liyuan, Wang, Xun and Wang, Xingguang., 2020)	RNN	LIDC/IDRI	2,054	CT images	-	85.71
(Gulshan V, Peng L, Coram M, et al., 2016)	CNN	EyePACS	9,963	X-ray images	Python	92.5
(Moreno-García, Carlos, et al., 2020)	CNN	NHS, MURA	256,000	X-ray images	-	92.7
(MORAN, Maira Beatriz Hernandez, et al., 2002)	CNN	Local	467	MRI images	-	92.3
(Moran, Maira, et al., 2021)	CNN	Sirona Heliodent Plus	112	CT images	Python	73.3
(El-Saadawy, Hadeer, et al., 2020)	CNN	MURA	40,561	X-ray images	-	73.42
(Ananthu, K. S., et al., 2021)	CNN	ALL-IDB2	260	CT images	-	92.88
(Westerberg, Erik., 2020)	CNN	KAGGLE,	12,811	X-ray images	Python	93.4
(Ding, Yiming, et al., 2019)	CNN	ADNI	2,149	MRI images	Python	92.0

6. RESULTS AND DISCUSSION

After studying and observing many recent studies that preceded us in this field, it turned out that deep learning technologies can learning from complex data that huge and huge volumes of data, unlike any other technology for image recognition and categorization, ensures high processing speed and accurate results.

The most famous areas of applications of deep learning are medical diagnosis that directly serve humanity, through use it in health informatics, biomedicine, and magnetic resonance image MRI analysis, also diagnosis, classification, prediction, and detection of bone diseases in x-ray images.

And by accurately diagnosing the x-ray image of the bones well, quickly, and effectively, this may spare humans many risks and burdens, and may avoid humans the risk of death.

So, we can say that deep learning has surpassed the usual and old diagnostic method through automatic learning and training from data set and update data by himself.

7. CONCLUSION

It is possible to build a model for the automatic detection of bone malformations using deep learning technology and connecting it to a large dataset of images then train the models on them, then do a test evaluation on the models and gives the results with high accuracy, reduces the cost, and saves time, where the Human bone x-ray images are applied in the models and completing the training and evaluation stage, the model can give the results as classification of human bone types, so that the model is trained on 14 types of upper bones of humans (for example), and two other groups to classify the result to normal or abnormal in each type of bone.. This leads to the fact that deep learning technology can solve the most complex medical problems, which is the early detection of diseases, to avoid humans from the scourge of misdiagnosis and high cost. In addition to proving that we can develop and provide better deep learning models compared to the previously established models.

REFERENCES

1. TOC. (2018). 10 DIFFERENT TYPES OF BONE DISEASES TO WATCH OUT FOR in Bone Care. Total Orthopaedic Care. Retrieved October 24, 2021, from https://www.toc.md/2018/08/03/10-different-types-of-bone-diseases-to-watch-out-for/bone.
2. Bauer, G. C.H. (2019). bone disease. Encyclopedia Britannica. Retrieved October 24, 2021, from https://www.britannica.com/science/bone-disease.
3. InformedHealth. (2017). Understanding tests used to detect bone problems. Institute for Quality and Efficiency in Health Care. Retrieved October 24, 2021, from https://www.ncbi.nlm.nih.gov/books/NBK279413/.
4. Riggs, B., & Melton, L. (1995). The worldwide problem of osteoporosis: insights afforded by epidemiology. Bone, 17(5), S505–S511.
5. Chrischilles, E. A. (1991). A Model of Lifetime Osteoporosis Impact. Archives of Internal Medicine, 151(10), 2026.
6. Winner, S. J., Morgan, C. A., & Evans, J. G. (1989). Perimenopausal risk of falling and incidence of distal forearm fracture. BMJ, 298(6686), 1486–1488.
7. Leibson, C. L., Tosteson, A. N. A. et al. (2002). Mortality, Disability, and Nursing Home Use for Persons with and without Hip Fracture: A Population-Based Study. Journal of the American Geriatrics Society, 50(10), 1644–1650.
8. Esteva, A., Chou, K., Yeung, S. et al. (2021). Deep learning-enabled medical computer vision. Npj Digital Medicine, 4(1).
9. CERNAZANU-GLAVAN, C., & HOLBAN, S. (2013). Segmentation of Bone Structure in X-ray Images using Convolutional Neural Network. Advances in Electrical and Computer Engineering, 13(1), 87–94.
10. Sitaula, C., & Hossain, M. B. (2021). Attention-based VGG-16 model for COVID-19 chest X-ray image classification. Applied Intelligence, 51(5), 2850-2863.
11. Esteva, A., Kuprel, B. et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118.
12. Gulshan, V., & Peng, L. et al. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA, 316(22), 2402.
13. Moreno-García, Carlos, et al. (2020). Assessing the clinicians pathway to embed artificial intelligence for assisted diagnostics of fracture detection. CEUR Workshop Proceedings.
14. Moran, M. B. H., & Faria, M. et al. (2020). On using convolutional neural networks to classify periodontal bone destruction in periapical radiographs. 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2036–2039.
15. El-Saadawy, H., & Tantawi, M. et al. (2020). A Two-Stage Method for Bone X-Rays Abnormality Detection Using MobileNet Network. Advances in Intelligent Systems and Computing, 372–380.
16. Ananthu, K. S., & Krishna Prasad, P. et al. (2021). Acute Lymphoblastic Leukemia Detection Using Transfer Learning Techniques. Intelligent Sustainable Systems, 679–692.
17. Wang, S., & Dong, L. (2020). Classification of pathological types of lung cancer from CT images by deep residual neural networks with transfer learning strategy. Open Medicine, 15(1), 190–197.
18. Russakovsky, O., & Deng, J. et al. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3), 211–252.

19. Bar, Y., & Diamant, I. et al. (2015). Chest pathology detection using deep learning with non-medical training. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), 294–297.
20. Dimitris, K., Ergina, K.(2017). Concept detection on medical images using Deep Residual Learning Network. In Working notes of conference and labs of the evaluation forum. Springer.
21. Dhungel, N., Carneiro, G., & Bradley, A. P. (2017). Fully automated classification of mammograms using deep residual neural networks. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), 310–314.
22. Moran, M., & Faria, M. et al. (2021). Classification of Approximal Caries in Bitewing Radiographs Using Convolutional Neural Networks. Sensors, 21(15), 5192.
23. Westerberg, E. (2020). AI-based Age Estimation using X-ray Hand Images . A comparison of Object Detection and Deep Learning models (Dissertation). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:bth-19598.
24. Ding, Y., & Sohn, J. H. et al. (2019). A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18F-FDG PET of the Brain. Radiology, 290(2), 456–464.