Pre-trained Convolutional Neural Networks in the Assessment of Bone Scans for Metastasis

Vincent Peter C. Magboo, MD, MS[1, 2], Ma. Sheila A. Magboo, MS[1]

[1] *Department of Physical Sciences and Mathematics, University of the Philippines Manila*
[2] *Section of Nuclear Medicine, University of Perpetual Help Medical Center*
E-mail address: vcmagboo@up.edu.ph, mamagboo@up.edu.ph

ABSTRACT

Background:

Numerous applications of artificial intelligence have been applied in radiological imaging ranging from computer-aided diagnosis based on machine learning to deep learning using convolutional neural networks. One of the nuclear medicine imaging tests being commonly performed today is bone scan. The use of deep learning methods through convolutional neural networks in bone scintigrams has not been fully explored. Very few studies have been published on its diagnostic capability of convolutional neural networks in assessing osseous metastasis.

Objective:

The aim of our study is to assess the classification performance of the pre-trained convolutional neural networks in the diagnosis of bone metastasis from whole body bone scintigrams of a local institutional dataset.

Methods:

Bone scintigrams from all types of cancer were retrospectively reviewed during the period 2019-2020 at the University of Perpetual Help Medical Center in Las Pinas City, Metro Manila. The study was approved by the Institutional Ethical Review Board and Technical Review Board of the medical center. Bone scan studies should be mainly for metastasis screening. The pre-processing techniques consisting of image normalization, image augmentation, data shuffling, and train-test split (testing at 30% and the rest (70%) was split 85% for training and 15% for validation) were applied to image dataset. Three pre-trained architectures (ResNet50, VGG19, DenseNet121) were applied to the processed dataset. Performance metrics such as accuracy, recall (sensitivity), precision (positive predictive value), and F1-scores were obtained.

Results:

A total of 570 bone scan images with dimension 220 x 646 pixel sizes in .tif file format were included in this study with 40% classified with bone metastasis while 60% were classified as without bone metastasis. DenseNet121 yielded the highest performance metrics with an accuracy rate of 83%, 76% recall, 86% precision, and 81% F1-score. ResNet50 and VGG19 had similar performance with each other across all metrics but generally lower predictive capability as compared to DenseNet121.

Conclusion:

A bone metastasis machine learning classification study using three pre-trained convolutional neural networks was performed on a local medical center bone scan dataset via transfer learning. DenseNet121 generated the highest performance metrics with 83% accuracy, 76% recall, 86% precision and 81% F1-score. Our simulation experiments generated promising outcomes and potentially could lead to its deployment in the clinical practice of nuclear medicine physicians. The use of deep learning techniques through convolutional neural networks has the potential to improve diagnostic capability of nuclear medicine physicians using bone scans for the assessment of metastasis.

Keywords: *Deep Learning, Convolutional Neural Networks, Transfer Learning, Bone Metastasis*

INTRODUCTION

In the recent years, numerous applications of artificial intelligence (AI) have been applied in radiological imaging ranging from computer-aided diagnosis based on machine learning to deep learning using convolutional neural networks (CNN) [1]. Deep learning techniques have been studied for various potential applications such as data acquisition, image reconstruction and image registration, image segmentation, image classification and lesion segmentation [2]. The seeming endless applications in radiological imaging being considered a data-rich medical specialty have been made possible due to the advances and widespread availability not only in hardware but in software as well [3]. However, for many radiologists and nuclear medicine physicians, the term AI appears to be a blackbox with doubts on its interpretability and perceived as a threat to their clinical practice [4].

One of the nuclear medicine imaging tests commonly performed today is bone scan. Its primary indication is to detect the presence of osseous metastasis which would then suggest that the cancer has reached its advanced stage with a median survival of a few months and having limited appropriate therapies [5, 6]. Many cancers, like breast, prostate, and lung malignancies are known to spread to the bones. In 25-40% of advanced breast cancer patients, bones are usually the first site of distant metastasis [7]. In prostate cancer, the metastatic deposits in the axial skeleton can cause pain, debility and/or functional impairment impacting the quality of life of the patients [8].

The use of CNN in bone scintigrams has not been fully explored. Very few studies have been published on the diagnostic capability of CNN in assessing osseous metastasis. Papandrianos et al developed a robust CNN architecture for bone metastasis diagnosis using whole - body scan images with an impressive classification accuracy of 92.50% besting other popular and well-known CNN architectures for medical imaging like ResNet50, VGG16, MobileNet, and DenseNet [6]. Using a meticulous exploration of CNN hyperparameter selection and fine-tuning, Papandrianos et al, applied a CNN model for the classification of bone scans for metastasis among prostate cancer patients. The model yielded classification testing accuracy of 97.38% outperforming VGG16, ResNet50, GoogleNet, and MobileNet [9]. In another study involving 14,972 bone lesions from whole - body bone scans the authors compared a 2D CNN based on the RestNet50 architecture with InceptionV3, VGG16, and DenseNet169. Results showed their CNN model bested other pre-trained architectures with an average sensitivity, specificity, accuracy, positive predictive value, and negative predictive value for all visible bone lesions at 81.30%, 81.14%, 81.23%, 81.89%, and 80.61%, respectively [10]. In a masteral thesis by Dang, the author designed a CNN to classify hotspots in bone scintigram for bone metastasis with a testing accuracy rate of 89% [11].

OBJECTIVE

The aim of our study is to assess the classification performance of the pre-trained convolutional neural networks in the diagnosis of bone metastasis from whole body bone scintigrams of a local institutional dataset. The performance metrics include accuracy, precision, recall and F1-score.

MATERIALS AND METHODS

Bone scintigrams during the period 2019-2020 at the University of Perpetual Help Medical Center in Las Pinas City, Metro Manila were retrospectively reviewed. The study was approved by the Institutional Ethical Review and Technical Review Boards of the medical center and was conducted in accordance with the Declaration of Helsinki for the ethical conduct of research involving human participants. The machine learning pipeline for this study is shown in Figure 1.

Characteristics of Dataset

Bone scintigrams from all types of cancers during the period 2019-2020 of the medical center were included in the study. The bone scan studies should be mainly for metastasis screening. Other indications of bone scans such as assessment of metabolic bone disease, osteomyelitis versus cellulitis, loosening of implants/ prosthesis, identification of primary bone tumors, etc. were excluded in the study.

Bone scan images consisted of whole body anterior and posterior views with 1024 x 256 pixel resolution. All bone scan procedures were performed with a Siemens

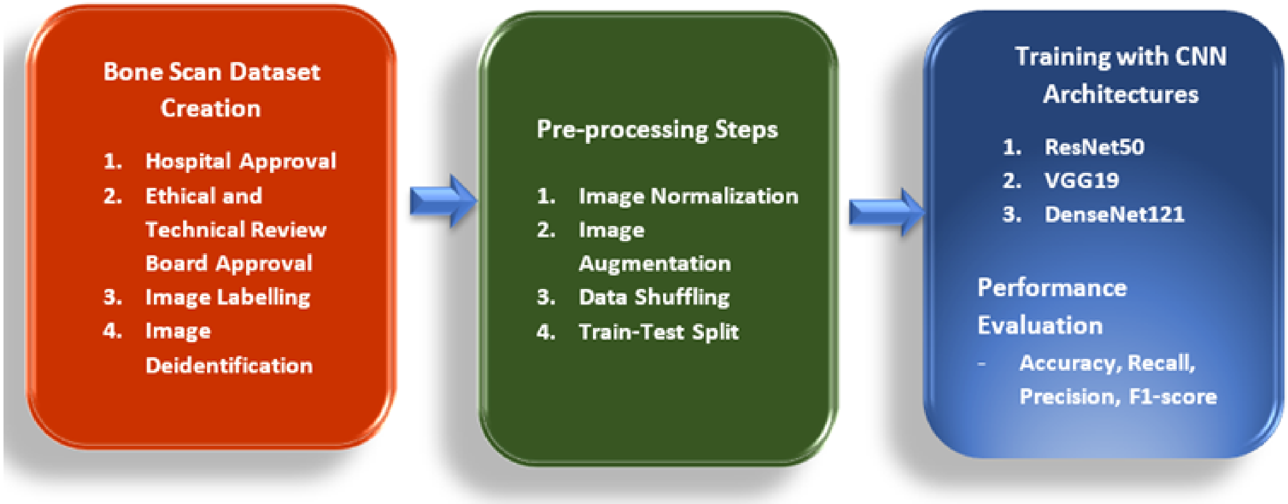

FIGURE 1. Machine Learning Pipeline for Bone Metastasis Study

gamma camera Symbia S series with two heads, with low energy high resolution collimators (LEHR), and with SyngoVE32B software. Bone scans were obtained 3 hours after the intravenous administration of 25 - 30 mCi of technetium-99m methylene diphosphonate (Tc-99m MDP) using a low-energy high-resolution collimator, matrix size of 1024 x 256, an acquisition time of 15 – 20 cm/min and photon energy centered on the 140-keV photo-peak with a symmetrical 20% energy window [12, 13]. The scanning procedure was in accordance with the guidelines set by the European Association of Nuclear Medicine and Society of Nuclear Medicine.

Labelling of Images

The interpretation of bone scan images was performed by a board-certified nuclear medicine physician with almost 25-year clinical experience in bone scan interpretation. Quality assurance of all the images before its inclusion in the machine learning pipeline was made. A pre-processing approach was also done to remove artifacts (non-osseous uptake such as urine contamination, site of tracer injection etc) in the original images. Images with medical devices such as implant, catheters etc were also excluded in the study to avoid interference with image interpretation. Additionally, all included bone scan whole body images underwent deidentification procedure resulting to cropped images for inclusion in the study. Images were then classified into two classes: (1) with scintigraphic evidence of bone metastasis and (2) without scintigraphic evidence of bone metastasis. The following were used as the criteria for the scintigraphic evidence of bone metastasis: (a) based on the typical patterns of tracer uptake seen in metastasis, (b) interval appearance of new bone lesions that cannot be ruled out as malignant in follow-up scans, (c) presence of flare phenomenon on scans, (d) when the accompanying medical records and radiological reports (CT scan, radiographs, MRI, PET/CT, bone alkaline phosphatase elevation) indicate bone destruction, and (e) when bone lesions appeared enlarged after at least 3 months follow-up. On the other hand, following criteria were used to classify bone scan as without scintigraphic evidence of bone metastasis: (a) bone lesions confirmed to be traumatic in origin, (b) lesions whose tracer uptake appeared around the bone joint, (c) lesions which the accompanying radiological studies indicate non-osseous metastasis and (d) equivocal lesions which lack definitive evidence of metastasis. Figure 2 shows a sample image indicating with and without bone metastasis

Pre-Processing Techniques

Numerous pre-processing techniques were applied to image dataset. These consisted of (a) image normalization using Min-Max normalization, (b) image augmentation, c) data shuffling for random order of the datset, and (d) data train-test split. The dataset was split in three parts: testing (30%) and the rest (70%) is split 85% for training and 15% for validation. The following geometric augmentations were applied to all images: (1) zoom range, (2) horizontal flipping, (3) rotation range, (4) translation, and (5) shear range.

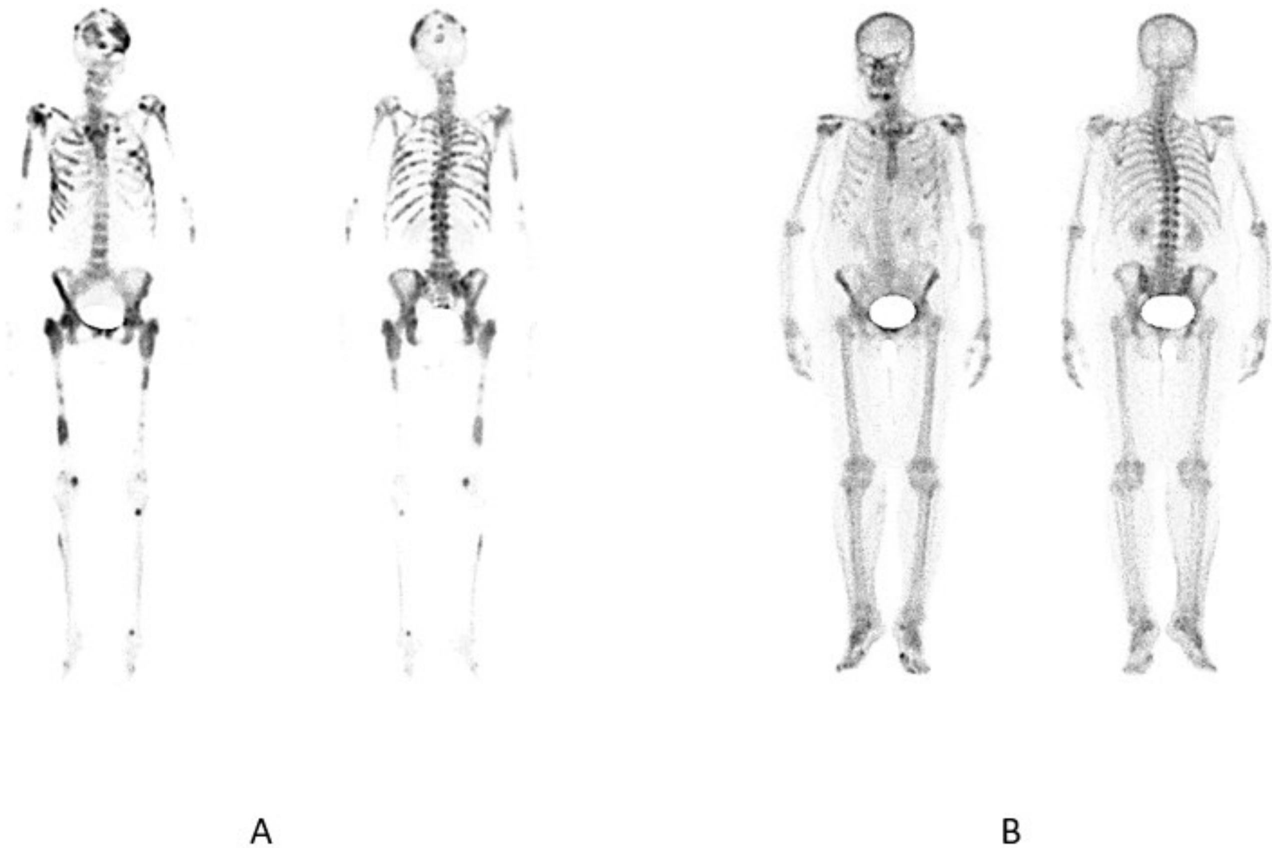

FIGURE 2. Sample bone scan image (A) with bone metastasis and (B) without bone metastasis

Architecture of the Convolutional Neural Networks

Popular pre-trained CNNs typically used in medical imaging were applied to the processed dataset. These include ResNet50, VGG19, and DenseNet121. ResNet50 is a CNN model consisting of 50 layers version of ResNet (Residual Neural Network) trained on ImageNet database. Its architecture consists of sequences of convolutional blocks with average pooling and uses softmax at the last layer for classification [14, 15, 16, 17]. VGG-19 is one of the VGG (Visual Geometry Group) based architectures with 19 connection layers, including 16 convolution layers and 3 fully connected layers. The convolution layers extract features of the input images, the fully connected layers with softmax make the final classification and uses Maxpooling instead of average pooling for downsampling to reduce volume size prior to classification [15, 17, 18, 19]. DenseNet (Dense Convolutional Neural Network) is another type of CNN architecture commonly employed for visual object recognition. DenseNet121 consists of 121 layers with parameters of more than 8 million, divided into DenseBlocks. The layers between the blocks are called transition layers and uses a batch normalization for down-sampling and employ softmax activation function in the last fully connected layer for the classification [16, 20, 21]. Huan et al have reported the advantages of DenseNet121 as follows: alleviation of the vanishing-gradient problem, strengthening of the feature propagation, encourage feature reuse, and substantial reduction of the number of parameters [20]. The architecture of these pre-trained CNN's are shown in Figures 3 - 6.

All simulation experiments were performed in Kaggle as it supports free use of NVIDIA TESLA P100 GPUs. Keras 2.6.0, TensorFlow 2.6.0., and python language 3.7.10 were utilized in all simulations.

Performance Metrics

Performance of the pre-trained architectures in the classification of bone metastasis, accuracy, recall (sensitivity), precision (positive predictive value), and F1 scores were computed.

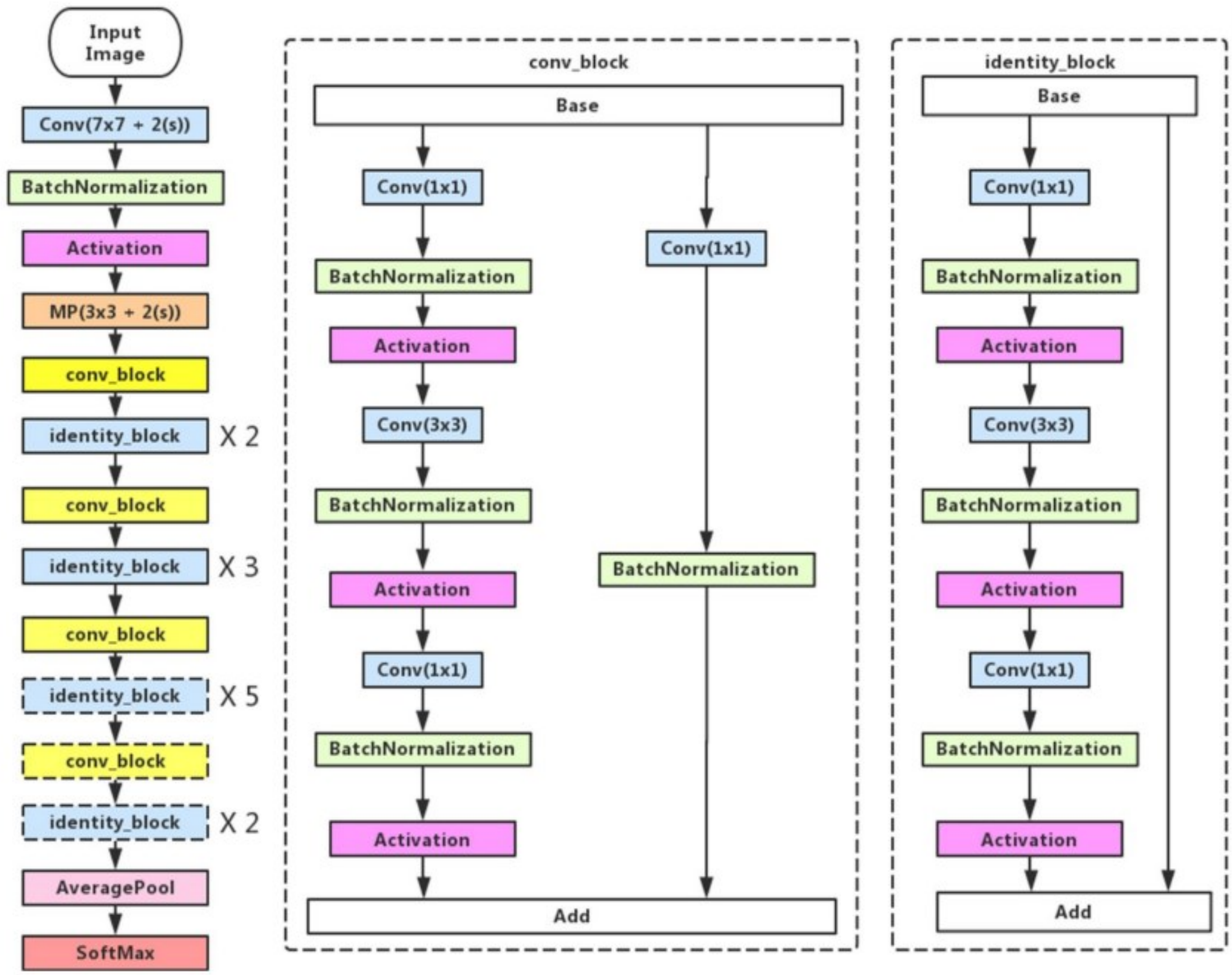

FIGURE 3. ResNet50 Architecture (Source:[14])

FIGURE 4. VGG19 Architecture (Source:[19])

RESULTS

Included in this study were 570 bone scan images with dimension 220 x 646 pixel sizes in .tif file format of which 228 (40%) images were classified with bone metastasis while 342 (60%) images were classified as without bone metastasis. Majority of our cases were females at 68% with breast cancer as the most common type of malignancy. The clinical characteristics of bone scan patients is seen in Table 1.

DenseNet121 yielded the highest performance metrics with an accuracy rate of 83%, 76% recall (sensitivity), 86% precision (positive predictive value), and 81% F1 - score. ResNet50 and VGG19 had similar performance with each other across all metrics but generally lower predictive capability as compared to DenseNet121. The performance metrics of each architecture are shown in Figure 6.

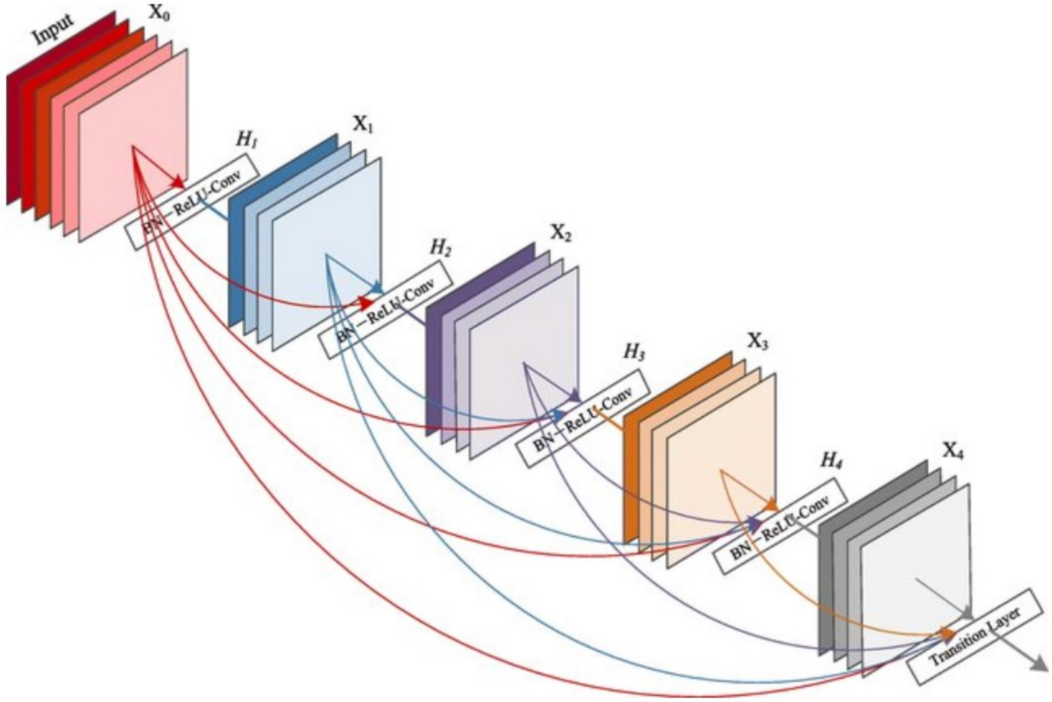

FIGURE 5. DenseNet121 Architecture (Source:[20])

DISCUSSION

We have conducted simulation experiments on a bone scan dataset of a local medical center. All bone scan images underwent data cleaning for quality assurance. This was followed by application of pre-processing techniques in an attempt to increase the predictive capability of CNN models as well as avoidance of potential overfitting. The use of geometric augmentation is a standard practice in machine learning as it improves the performance of CNN models in image classification. In our study, DenseNet121 showed the highest predictive performance as compared to the VGG19 and ResNet50. As compared to other studies reported in the literature showing good performance of various pre-trained architectures in the classification of osseous metastasis from bone scans, our results are generally lower [6,9,10,19]. This could be brought about by the relatively small number of images included in our dataset. Nonetheless, our results are similar to that of Papandrianos et al [6], which showed DenseNet121 with the highest performance.

Pre-trained CNN models are trained on numerous datasets with various categories of images. However, these are not trained on radiological images. However, transfer learning techniques can still employ these pre - trained models on a variety of computer vision problems. This is more particularly prominent in areas with limited resources (dataset and computing resources). The availability of labelled radiological images for deep learning studies are very few. In the Philippines, we believe this study is the first its kind in applying machine learning techniques in nuclear medicine images.

While the pre-trained CNN models have fairly - satisfactory performance metrics rates suitable for use in clinical practice, DenseNet121 obtained the highest predictive capability for classifying osseous metastasis. Hence, DenseNet121 can be tapped by nuclear medicine physicians as a decision support tool in the interpretation of bone scintigrams. Complementing bone scan (with well-known excellent sensitivity) with the use of these CNN models due to its positive predictive value highlights the potential utility of CNN models in the clinical practice of nuclear medicine physicians.

CONCLUSIONS

A bone metastasis machine learning classification study was performed on a local medical center bone scan dataset via transfer learning. Three pre-trained convolutional neural networks were assessed for its capability to detect bone metastasis from bone scans. DenseNet121 generated the highest performance metrics with 83% accuracy, 76% recall, 86% precision

TABLE 1. Clinical Characteristics of Bone Scan Patients

Clinical Features	Percentage (%)
Male	**32%**
Female	**68%**
Age (years)	**46.55 ± 8.12**
Type of Malignancy	
Breast Cancer	**70 %**
Prostate Cancer	**15 %**
Lung Cancer	**2 %**
GI Malignancies (Colon,	**2 %**
Rectal, Anal, Appendiceal)	**1 %**
Other Malignancies	
Labelled Images (570 images)	
With bone metastasis	**228 (40%)**
Without bone metastasis	**342 (60%)**

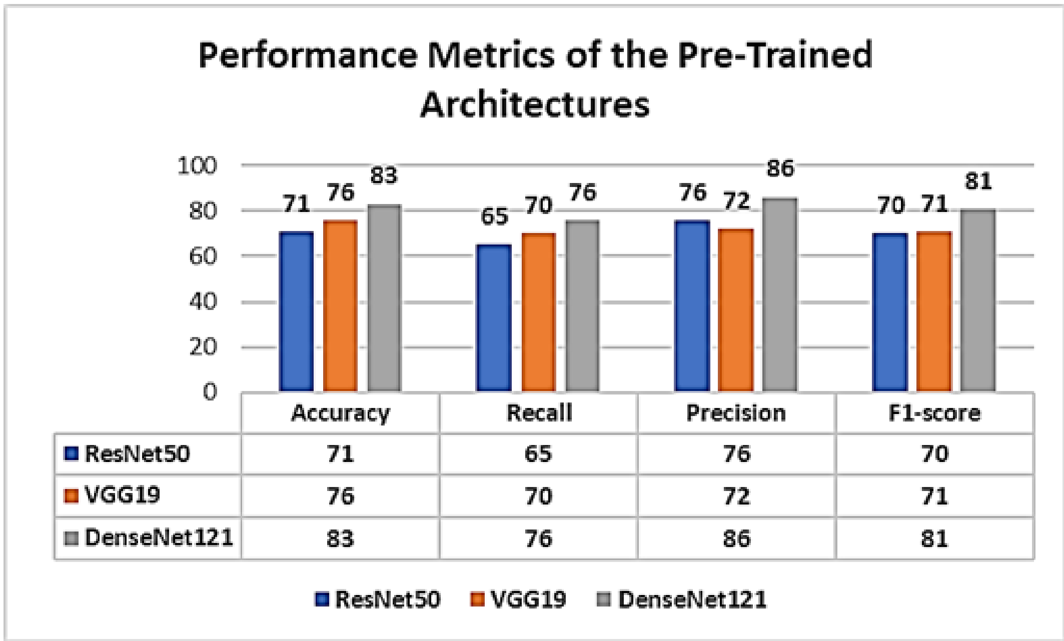

FIGURE 6. Forest Performance Metrics of the Pre-Trained Architectures

and 81% F1-score. Our simulation experiments generated promising outcomes and potentially could lead to its deployment in the clinical practice of nuclear medicine physicians. The use of deep learning techniques through convolutional neural networks has the potential to improve diagnostic capability of nuclear medicine physicians using bone scans for the assessment of metastasis.

RECOMMENDATIONS

Addition of more bone scan images is highly recommended to improve the performance of the neural network models as deep learning techniques generally require huge amount of images. Additionally, it is highly recommended to do more simulations with other pre-trained architectures, using different learning rates and different types of optimizers.

REFERENCES

1. Mun SK, Wong KH , Lo SCB, Li Y, Bayarsaikhan S. "Artificial Intelligence for the Future Radiology Diagnostic Service". Frontiers in Molecular Biosciences, Vol 7, 2021. https://www.frontiersin.org/articles/10.3389/fmolb.2020.614258.
2. Wang S, Cao G, Wang Y, Liao S, Wang Q, Shi J, Li C, Shen D. "Review and Prospect: Artificial Intelligence in Advanced Medical Imaging". Frontiers in Radiology, Vol 1, 2021. https://www.frontiersin.org/articles/10.3389/fradi.2021.781868

3. Sala E, Ursprung, S. “Artificial Intelligence in Radiology: The Computer’s Helping Hand Needs Guidance”. Radiology: Artificial Intelligence, vol 2 (6), pe200207, 2020. doi = 10.1148/ryai.2020200207.
4. Nensa F, Demircioglu A, Rischpler C. “Artificial Intelligence in Nuclear Medicine”. J Nucl Med 2019; 60:29S–37S. DOI: 10.2967/jnumed.118.220590.
5. Valenti MT, Mottes M, Carbonare LD, Feron O. “Editorial: Bone Metastases”. Frontiers in Oncology, vol 11, 3030, 2021. https://www.frontiersin.org/article/10.3389/fonc.2021.741515.
6. Papandrianos N, Papageorgiou E, Anagnostis A, Feleki A. “A Deep-Learning Approach for Diagnosis of Metastatic Breast Cancer in Bones from Whole-Body Scans”. Applied Sciences. 2020; 10(3):997. https://doi.org/10.3390/app10030997.
7. Murat K, Baha Z, Raika D, Funda T, Zehra A, Cenk S, Adam U. “Clinicopathologic features of single bone metastasis in breast cancer”. Medicine. 100, 1, e24164, 2021. DOI: https://doi.org/10.1097/MD.0000000000024164.
8. Mingyu Z, Xin L, Yuan Q, Silong H, Yingjian Z, Wentao L, et al. “Bone metastasis pattern of cancer patients with bone metastasis but no visceral metastasis”. J Bone Oncol. 2019 Apr 15. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6357895/.
9. Papandrianos N, Papageorgiou E, Anagnostis A, Papageorgiou K. “Bone metastasis classification using whole body images from prostate cancer patients based on convolutional neural networks application”. PLoS ONE 15(8): e0237213, 2020. https://doi.org/10.1371/journal.pone.0237213.
10. Liu Y, Yang P, Pi Y, Jiang L, Zhong X, Cheng J, Xiang Y, Wei J, Li L, Yi Z, Cai H, Zhao Z. “Automatic identification of suspicious bone metastatic lesions in bone scintigraphy using convolutional neural network”. BMC medical imaging, 21(1), 131, 2021. doi: https://doi.org/10.1186/s12880-021-00662-9.
11. Dang, J. Classification in Bone Scintigraphy Images Using Convolutional Neural Networks. Master’s Thesis, Lund University, Lund, Sweden, 2016.
12. den Wyngaert TV, Strobel K, Kampen WU, Kuwert T, van der Bruggen W, Mohan H, Gnanasegaran G, Delgado-Bolton R, Weber W, Beheshti M, Langsteger W, Giammarile F, Mottaghy F, Paycha F, and EANM Bone & Joint Committee and the Oncology Committee. 2016. “The EANM practice guidelines for bone scintigraphy”. European journal of nuclear medicine and molecular imaging, 43, 9, 1723–1738, 2016. doi: https://doi.org/10.1007/s00259-016-3415-4.
13. Bartel T, Kuruva M, Gnanasegaran G, Beheshti M, Cohen E, Weissman A, and Yarbrough T. “SNMMI Procedure Standard for Bone Scintigraphy 4.0”. Journal of nuclear medicine technology, 46, 4, 398–404, 2018. doi: https://tech.snmjournals.org/content/46/4/398.
14. Ji Q, Huang J, He W, Sun Y. “Optimized Deep Convolutional Neural Networks for Identification of Macular Diseases from Optical Coherence Tomography Images”. Algorithms, 12(3):51, 2019. https://doi.org/10.3390/a12030051.
15. Ikechukwu AV, Murali S, Deepu R, Shivamurthy RC. “ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images”. Global Transitions Proceedings, 2(2), Issue 2 375-381, 2021. https://doi.org/10.1016/j.gltp.2021.08.027.
16. Shazia, A., Xuan, T.Z., Chuah, J.H. et al. A comparative study of multiple neural network for detection of COVID - 19 on chest X-ray. EURASIP J. Adv. Signal Process. 2021, 50 (2021). https://doi.org/10.1186/s13634-021-00755-1.
17. Ali L, Alnajjar F, Jassmi, HA, Goocho M, Khan W, Serhani MA. “Performance Evaluation of Deep CNN - Based Crack Detection and Localization Techniques for Concrete Structures.” Sensors 21, 1688, 2021. https://doi.org/10.3390/s21051688.
18. Simonyan K, Zisserman A. “Very deep convolutional networks for large-scale image recognition”. 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc (2015), pp. 1-14.
19. Bansal, M., Kumar, M., Sachdeva, M. et al. Transfer learning for image classification using VGG19: Caltech-101 image data set. J Ambient Intell Human Comput (2021). https://doi.org/10.1007/s12652-021-03488-z.
20. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243.
21. Hasan, N., Bao, Y., Shawon, A. et al. DenseNet Convolutional Neural Networks Application for Predicting COVID-19 Using CT Image. SN COMPUT. SCI. 2, 389 (2021). https://doi.org/10.1007/s42979-021-00782-7.
22. Ntakolia C, Diamantis D, Papandrianos N, Moustakidis S, Papageorgiou E. “A Lightweight Convolutional Neural Network Architecture Applied for Bone Metastasis Classification in Nuclear Medicine: A Case Study on Prostate Cancer Patients.” Healthcare (Basel, Switzerland), 8, 4, 493, 2020. DOI: https://doi.org/10.3390/healthcare8040493.