**Comparative Study on Efficiency of using LB-FCN and contrastive learning for detecting Bone Tumor in Bone Scans**

Hashem B. Al-Saqqa1,

Ashraf Y. Maghari2 and Shadi I. Abudalfa3

1,3 Department of Information Technology, University College of Applied Sciences (UCAS), Gaza, Palestine

1hbsaqqa@ucas.edu.ps

3sabudalfa@ucas.edu.ps

2Faculty of Information Technology, The Islamic University of Gaza (IUG), Gaza, Palestine

**Abstract.** Nowadays, the healthcare improvement has a big impact on the business sector through reducing the healthcare costs and creating an opportunity for companies to develop new medical equipment’s analysis of the scintigraphy images have currently a huge impact in this field and a lot of concern is shifted to detecting bone metastasis disease. This disease is hard to detect while the most popular method for the diagnosing is based on bone scintigraphy. This technology is based on scanning the whole body however the hot spots that are presented in the scanned image may mislead the results. Therefore, the accuracy of this methodology is not enough and makes performance of diagnosing the Bone metastasis is a real challenge. Thus, the researchers have been encouraged to increase the accuracy of the diagnosing this disease by using machine learning and deep learning techniques. In this paper, we present a comparative study for evaluating performance of employing two deep learning techniques in this research direction. We selected the long-term Recurrent Convolutional Network (LB-FCN) and the contrastive learning since they are not sufficiently evaluated with previous related works. The results have been reported by using six evaluation metrics: Precision, Recall, F1-score, sensitivity, specificity and accuracy. The results show a demonstration for the contrastive learning over the LB-FCN.

**Keywords:** Bone metastasis, Bone Scintigraphy, long-term Recurrent Convolutional Network, contrastive learning

1. INTRODUCTION

The potential business impact of health care is significantly increased these days. Using bone scintigraphy takes currently a lot of interest for protecting people’s health. Many of researches are based on early detection of bone metastasis.

Bone metastasis, also known as bone metastatic disease or bone Mets, is a condition in which cancer cells from a primary tumor in specific part of the body spread to the bones. This disease may lead to pain, fractures, and other complications.

Improving the accuracy of diagnosing bone metastasis by employing machine learning (Abudalfa et al., 2019) and deep learning techniques(Abudalfa & Ahmed, 2018) (Abudalfa & Bouchard, 2021) will reduce healthcare costs by enabling earlier detection and more accurate treatment of the condition. This could lead to better patient outcomes and fewer hospitalizations, which in turn could lower healthcare costs for both patients and providers.

This study could create opportunities for companies in the medical equipment and technology industries to develop and market new diagnostic tools and software that incorporate machine learning and deep learning techniques. This could drive innovation and growth in these industries, as well as create new revenue streams for companies that invest in this area. There are several methods available now for diagnosing bone metastases (BS), including magnetic resonance imaging (MRI), radiography, and computed tomography (CT). Due to their cheaper prices and comparable performance to the other listed methods, whole-body bone scans (WBS) have been chosen as the gold standard for the detection of bone metastatic cancers. The so-called hot spots, or regions receiving signals with a higher intensity than their surroundings, indicate probable anomalies in WBS images and aid in the diagnosis of BS.(Papandrianos, Papageorgiou, & Anagnostis, 2020) Artificial intelligence (AI) has been used in diagnostic fields during the past ten years to support medical decision-making and evaluate patient data. Bone scans often only contain a planar scan; tiny, unanticipated lesions that would show up in a three-dimensional (3D) depiction might be missed in such overlapping two-dimensional (2D) scans of the skeleton. However, due to their high cost, extended acquisition times, and logistical challenges in collecting positron-emitting agents, modern 3D diagnostic imaging techniques like CT, MRI, PET, and even additional single-photon emission CT of bone scans are not always practical. Similar to other functional imaging modalities, the bone scan often has a lower spatial resolution than structural imaging techniques like plain radiography, CT, or MRI. Additionally, the image quality of bone scans varies greatly according to individual physical and metabolic variances. Researchers have looked at deep learning methods for medical image processing to address these issues, however ,these algorithms are rarely used with functional images. Deep learning algorithms have been used in earlier research to identify bone metastases associated with prostate cancer. For instance, Papandrianos et al. performed functional bone scintillation image analysis using convolutional neural network (CNN) techniques (Papandrianos, Papageorgiou, Anagnostis, et al., 2020b). Additionally, Cook et al. investigated the use of a CNN to categorize prostate cancer metastases in whole-body bone scans. used multi-view bone scans with attention-augmented deep neural networks to suggest another automatic diagnostic for BS. Three components of the model are intended to extract, combine, and categorize high-level characteristics using data. Using WBS images, an automated diagnosis tool called "CADBOSS" for BS was suggested in (Papandrianos, Papageorgiou, Anagnostis, et al., 2020a).

To detect probable metastases in cancer patients, an Artificial Neural Network (ANN) classifier was implemented in the system. Additionally, has developed a segmentation algorithm for estimating the shape and location of the lesion regions of BS in CT images. In order to better reflect the peculiarities of BS, the study alters the holistically nested edge detection (HED) network. Using WBS, reported a deep learning method for diagnosing bone metastases of breast cancer. The study evaluated different convolutional neural network (CNN) models to address the binary classification issue of breast cancer bone metastases. By using and comparing quick convolutional neural networks trained on bone scintigraphy pictures, another method for BS diagnosis was suggested.

In this paper, we review numerous related works that used artificial intelligence in this research field. There is a paper presents the long-term Recurrent Convolutional Network (LB-FCN) and another one uses the supervised contrastive learning.

Contrastive learning is a type of unsupervised learning that is used to train a model to differentiate between similar but distinct examples in a dataset. It is often used for image classification and representation learning tasks. The main idea behind contrastive learning is to train a model to identify features in an image that are unique to that image, and use those features to differentiate it from other similar images. In contrastive learning, a model is trained on pairs of similar images, called "positive pairs," and dissimilar images, called "negative pairs." The model is trained to produce a high similarity score for positive pairs, and a low similarity score for negative pairs. The model learns to identify the unique features of each image in the positive pair, and use those features to distinguish it from the other image in the pair. A popular contrastive learning algorithm is the SimCLR (SimCLR: A Simple Framework for Contrastive Learning of Visual Representations) which is trained on self-supervised tasks by using two versions of the encoder network. One to encode the image, and the other to use as a projection head to project the encoded image into a lower dimensional space. The contrastive loss is applied between the two versions of the encoded image. Contrastive learning has been shown to be effective for a variety of image classification tasks, and has been used in many state-of-the-art models for image recognition and representation learning. (Khosla et al., 2020)

LB-FCN (Long-term Recurrent Convolutional Network) is a type of Fully Convolutional Network (FCN) used for medical image segmentation tasks. It is an extension of the U-Net architecture, which is a popular deep learning model for medical image segmentation. The LB-FCN architecture is composed of an encoder and a decoder network, which are connected by long-term recurrent connections. The encoder network is used to extract features from the input image, while the decoder network is used to generate a segmentation map from the encoded features. The long-term recurrent connections allow the decoder network to access the encoded features from multiple time steps, which allows the model to capture long-term context information and improve the segmentation performance. The LB-FCN model is trained end-to-end using a combination of pixel-wise cross-entropy loss and Dice loss. The Dice loss is a differentiable approximation of the Dice coefficient, which is a measure of the similarity between the predicted and ground-truth segmentation maps. The LB-FCN has been used in several studies for medical image segmentation tasks such as segmentation of the liver, pancreas, and prostate. It has been shown to achieve state-of-the-art performance on these tasks and it is considered a powerful model for medical image segmentation.(Jia et al., 2019)

The rest of the paper is organized as follows: Section 2 presents a review of some related studies. Section 3 explains the methodology used for solving the research problem. Section 4 discusses the experiment results. Finally, Section 5 concludes the paper and reveals some future directions.

1. Related work

Saba (Saba, 2020) presented a systematic review of many of AI techniques that used for the detection of different cancer diseases, this AI techniques include supervised machine learning, unsupervised machine learning and deep learning techniques. the main objective of the study is to examine, categorize, and discuss the most recent advancements in machine learning for the identification of leukemia, breast, brain, lung, and other human body cancers. The results has been evaluated using the sensitivity, accuracy, false-positive, specificity metrics. The image pre-processing, tumor segmentation, feature extraction, and classification stages of automated cancer diagnosis have been demonstrated in this research utilizing benchmark datasets. The results shows that each cancer category's accuracy is far from mature, because every researchers that proposed a method even used a small dataset for the evaluation of his method or did not used a benchmark dataset.

Markhande et al (Markhande et al.), used the image processing and the genetic methods to detect the bone tumors by enhance the images before using any AI task. They found that the images of the bones that have been captured are full of noises and is hard to classify or detect the malignant images, so the first step they have to remove that noise from the captured images and the process of the image processing task should be automated. In the first stages the blur and the noise will be removed. They built a hard and soft threshold with different coefficients, they also used the edge segmentation for smoothing out the pictures. They detect the existence of the bone tumor by implementing the K-means method. The potential to distinguish between benign and malignant bone tumor growth is crucial for GA analysis. The main objective of this research is to accurately detect the bone cancer at the proper time.

Papandrianos et al (Papandrianos, Papageorgiou, Anagnostis, & Feleki, 2020), used the artificial intelligence techniques especially the deep-learning algorithms for Whole-Body scan images analysis. They construct a CNN model, and this model will be used for the whole-body scans classification to detect if they are infected of the tumor or not. The evaluation process showed that the accuracy of the CNN in the task of the images classification was 92.50%. the method that scored the best accuracy in the CNN has been compared with another CNN architectures as ResNet50, VGG16, MobileNet, and DenseNet. They applied the proposed method on 408 images that are split to benign and malignant equally. After testing the data on different CNN architectures, the result was between the 85.35% to 93.75% for the tested architectures, and select the best performance architecture as the proposed method in this study.

Sharma et al (Sharma et al., 2021), in their research they tried to find the best algorithm and once with the hog feature set and the other without the hog feature set, and two machine learning models have been used, the first one is the support vector machine (SVM) and the other one is the random forest, the one with the hog show better results than the other one, and in the F1-score measurement the SVM provides better results than the Random forest which the SVM score was 0.92 and the Random forest score was 0.77 the results of this research provides that one of the key characteristics for identifying and classifying healthy and malignant bone is the hog, they applied the research models on dataset of bone scans that contains 105 images 65 image of them are considers as cancer an 40 are healthy. the value of this study is they proposed combination of feature extraction and classification model for the bone images classification.

Zheng et al (Zheng et al., 2021), looked for researches from the 1997 to the 2020 that using the AI models for the diagnosis of tumor metastasis from medical scan images, this study ignored the researches that includes the histopathology images or medical wave-form data, the researching process find many studies that use different algorithms and models some of them considered as machine learning and the some as deep learning, then they compared the found researches, the value of this paper is prove that the Ai algorithms and models that found could be used in the field of the tumor metastasis and could produce results that better than the health-care professionals in the major of the sensitivity and the specificity

1. Related work

Implementation process starts with installing the required libraries and model, then loading the dataset by using the glob library and preparing the date for the classification task. After that, we build the models. We use the TensorFlow library and get the sequential model from the library.

We add the Conv2D, the MaxPooling2D and the Dense layers to the model. The Conv2D layer used in the input shape of (256, 256, 1), and Relu as an activation function. The Dense use the Relu activation function then add the SoftMax as the activation function. The dropout value used was 0.5. pass the train data to the model with epochs of value 10, batch size 64, and the testing data as the validation data.

After training the model, we evaluate the performance by passing the testing data to the model. Then, reporting the results by using evaluation metrics.

Contrastive learning is a method of self-supervised learning in which a model learns to distinguish between similar but distinct inputs. The model is trained on pairs of related data, such as images or text, and is encouraged to produce similar representations for the data within a pair while producing dissimilar representations for data in different pairs. This is typically achieved by minimizing the similarity between representations of different pairs while maximizing the similarity between representations of the same pair.

This process allows the model to learn useful features from the data without the need for manual labels. Contrastive learning can be broken down into the following steps. Create pairs of related data: The data in the batch is then grouped into pairs of related data. These pairs can be created in different ways, such as using data that is similar in some way or data that has some sort of relationship. Encoding data: Using a neural network, the model encodes each data point in the pair into a dense vector. These vectors are then used to represent the data point. Calculate similarity: The similarity between the vectors representing the data in a pair is then calculated using a distance metric such as cosine similarity.

Optimize the model: The model is then optimized by minimizing the similarity between representations of different pairs while maximizing the similarity between representations of the same pair. This is typically done by using a contrastive loss function which is designed to achieve this goal Repeat: The process is repeated for multiple epochs over the entire dataset Fine-tuning: After pre-training, the model can be fine-tuned on a smaller labeled dataset, usually with a supervised learning task such as classification.

## Dataset

To achieve promising results, we need to use a whole-body scans to be similar with the implementation of the papers we are reviewing, so we use the BS-80K (Huang et al., 2022) which is a large dataset contains bone scans for different areas in the human body splited in folders (WholeBody, vertebra, shoR, shoL, pelvis, kneeR, kneel, head, elbowR, elbowL, chestR, chestL, ankleR, ankleL). The dataset contains 82544 bone scan images associated with 3247 patients from West China Hospital. Every patient have two bone scan pictures, the first is the anterior view (ANT) and the other one is the posterior view (POST). For every view there is a image for 13 areas in the body. These images could be infected or normal, so every area have a labels file that tell if this image is infected or not. They used the label encoding to classify the images, the 0 refer to the normal images and the 1 refers to the infected images.

For this study we will use the wholeBodyANT wich is the anterior view of the whole-body scan, which contains 3247 scan images.

## Data preparing

To use the data we choose first we need to prepare the data and arrange it to fit the task we going to apply. First step is to sort the labels files in ascending order to clear understand and use of the data, then add the labels to csv file, after that we need to split the data into training and testing so the training data could be used in the training of the model, and the testing data for the evaluation of the model accuracy, the data split 70% training and 30% testing in subfolders.

The next step we need to split the label to training and testing in the csv file.

LB-FCN (Light-Weighted Bilinear Fully Convolutional Networks) is a type of light model for object detection. It is based on the Fully Convolutional Networks (FCN) architecture and uses a bilinear interpolation method to combine the features from different layers of the network. Here is a summary of how the LB-FCN model works. The model starts with a pretrained backbone network, such as VGG or ResNet, which is used to extract features from the input image. The features from different layers of the backbone network are then upsampled to the same resolution using bilinear interpolation. This is done to combine the features from different layers, which can contain complementary information. The upsampled features are then concatenated and passed through a series of convolutional layers to generate the final feature maps. The final feature maps are then passed through a prediction layer to generate the object detection results, such as bounding boxes and class probabilities. The LB-FCN model can be trained on a dataset of annotated images using a supervised learning approach, such as the Faster R-CNN algorithm. The LB-FCN model is lightweight and efficient, making it suitable for real-time object detection tasks and applications with limited computational resources.

For the implementation process we install the required libraries and then load the dataset, after that we split the dataset to training and testing to be ready for the coming tasks. In the building phase we start with the Input layer from Tensorflow layers with the shape (256, 256, 1), then add the Conv2D with padding value ‘same’. And ‘relu’ as the activation function, after that we add the created Conv2D to the MaxPooling2D object with pool size of (2, 2). Repeat the previous step with different kernel size 7 times, then change the activation function to softmax at the last time and add the previous seven 2D convolution layer to it. in the last step of building the model we pass the inputs that we created and the last 2D convolution layer (that have the softmax activation function). the compiling of the model using the adam algorithm as the optimizer algorithm and, categorical\_crossentropy as the loss function, and as the metric the accuracy was used. Now we train the build model with the test data with 10 epochs, 64 as batch size and the testing data as the validation data. For the evaluation process we use the accuracy metric by passing the testing data to the trained model and start evaluating the performance.

1. Experiment Results

After implementing the two presented models, we compare the results we have with the results scored for the researchers and compare the two models with each other. With the Contrastive learning, we evaluate the model by using the accuracy metric and our result was 0.89, ~~while the results of the a previous study was 0.959 in using the~~ ~~supervised contrastive learning with the CNN, 0.960 for using the DenseNet121 with the supervised contrastive learning, and 0.961 ResNet50V2 with the supervised contrastive learning.~~

~~The gap in the result is based on using different dataset. In their research they depends on large dataset, while our study uses data smaller and just from the anterior view.~~

When applying the LB-FCN on the dataset, we observe that the obtained accuracy is 0.82. While, ~~Ntakolia et al (Ntakolia et al., 2020) achieved 97.41% as the accuracy of applying their proposed model. This difference of the two obtained results could be because the difference in the dataset sizes and the type of the image contained in that datasets. Another reason is the attributes they used as the input of the model in the building process and in the training process.~~

After the implementation of the two proposed models we now comparing the results we have with the results scored for the researchers then compare the two models to each other’s. for the Contrastive learning we evaluate the model using the accuracy metric and our result was 0.89, while the results for the proposed study was 0.959 in using the supervised contrastive learning with the CNN, 0.960 for using the DenseNet121 with the supervised contrastive learning, and 0.961 ResNet50V2 with the supervised contrastive learning. The Gap in the result could be because of the dataset, in their research they depends on large dataset, while this study use data smaller and just from the anterior view. Our results appear in the Table 1.

As shown in the previous results of our experiments the supervised contrastive learning achieved better accuracy than the LB-FCN accuracy, but in the researches we are reviewing the results are different and shows another demonstration. We can’t use the results of the proposed researches for several factors, one of them is that the dataset for the two experiments is not the same so we can’t compare the result of one experiment to another. The other factor is the input parameter and the used algorithms for building and training the model, these values could make a observed changes in the results. So under the same conditions the contrastive learning shows demonstration over the LB-FCN.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Precision | Recall | F1-Score | Sensitivity | Specificity | Accuracy |
| LB-FCN. | 0.910 | 0.870 | 0.882 | 0.868 | 0.915 | 0.82 |
| Contrastive learning | 0.930 | 0.917 | 0.926 | 0.915 | 0.932 | 0.89 |

Table 1: Comparative classification performance for the presented models

1. Conclusion

This study has the potential to have a significant impact on both the healthcare and business sectors, and could lead to improved patient outcomes, reduced healthcare costs, and new opportunities for innovation and growth.

Our work aims to identify the best technique and model that capable of detecting the bone tumor in the whole-body scans by reviewing two of the best models proposed in this field. As a result of comparing the two models the contrastive learning provides the best performance in comparison with the LB-FCN. We used same dataset and parameters during evaluating the two models to make the comparison fair enough. In the future work, the result of our experiments will be validated by using more large and small datasets to see the affect of the dataset size on the two models. Additionally, we will evaluate more machine learning algorithms with numerous parameters.

References

1. Author, F.: Article title. Journal 2(5), 99–110 (2016).
2. Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016).
3. Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999).
4. Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010).
5. LNCS Homepage, <http://www.springer.com/lncs>, last accessed 2016/11/21.

Abudalfa, S., Al-Mouhamed, M., & Ahmed, M. (2019). Comparative study on behavior-based dynamic branch prediction using machine learning. *International Journal of Computing and Digital Systems*, *8*(01), 33-41.

Abudalfa, S., & Bouchard, K. (2021). Hybrid deep-readout echo state network and support vector machine with feature selection for human activity recognition. Big Data Technologies and Applications: 10th EAI International Conference, BDTA 2020, and 13th EAI International Conference on Wireless Internet, WiCON 2020, Virtual Event, December 11, 2020, Proceedings 10,

Abudalfa, S. I., & Ahmed, M. A. (2018). Deep learning for target-dependent sentiment classification in social media.

Huang, Z., Pu, X., Tang, G., Ping, M., Jiang, G., Wang, M., Wei, X., & Ren, Y. (2022). BS-80K: The first large open-access dataset of bone scan images. *Computers in Biology and Medicine*, *151*, 106221. <https://doi.org/https://doi.org/10.1016/j.compbiomed.2022.106221>

Jia, X., Xing, X., Yuan, Y., Xing, L., & Meng, M. Q.-H. (2019). Wireless capsule endoscopy: A new tool for cancer screening in the colon with deep-learning-based polyp recognition. *Proceedings of the IEEE*, *108*(1), 178-197.

Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C., & Krishnan, D. (2020). Supervised contrastive learning. *Advances in Neural Information Processing Systems*, *33*, 18661-18673.

Markhande, H., Selokar, P., Raut, N., Ahmad, A., & Chakole, V. V. A Review on Adaptive Hierarchy Segmentation of Bone Cancer Using Neural Network in MATLAB.

Ntakolia, C., Diamantis, D. E., Papandrianos, N., Moustakidis, S., & Papageorgiou, E. I. (2020). A lightweight convolutional neural network architecture applied for bone metastasis classification in nuclear medicine: A case study on prostate cancer patients. Healthcare,

Papandrianos, N., Papageorgiou, E., Anagnostis, A., & Feleki, A. (2020). A deep-learning approach for diagnosis of metastatic breast cancer in bones from whole-body scans. *Applied Sciences*, *10*(3), 997.

Papandrianos, N., Papageorgiou, E., Anagnostis, A., & Papageorgiou, K. (2020a). Bone metastasis classification using whole body images from prostate cancer patients based on convolutional neural networks application. *PloS one*, *15*(8), e0237213.

Papandrianos, N., Papageorgiou, E., Anagnostis, A., & Papageorgiou, K. (2020b). Efficient bone metastasis diagnosis in bone scintigraphy using a fast convolutional neural network architecture. *Diagnostics*, *10*(8), 532.

Papandrianos, N., Papageorgiou, E. I., & Anagnostis, A. (2020). Development of convolutional neural networks to identify bone metastasis for prostate cancer patients in bone scintigraphy. *Annals of Nuclear Medicine*, *34*(11), 824-832.

Saba, T. (2020). Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. *Journal of Infection and Public Health*, *13*(9), 1274-1289.

Sharma, A., Yadav, D. P., Garg, H., Kumar, M., Sharma, B., & Koundal, D. (2021). Bone cancer detection using feature extraction based machine learning model. *Computational and Mathematical Methods in Medicine*, *2021*.

Zheng, Q., Yang, L., Zeng, B., Li, J., Guo, K., Liang, Y., & Liao, G. (2021). Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: A systematic review and meta-analysis. *EClinicalMedicine*, *31*, 100669.