DEEP LEARNING-BASED AUTOMATED OSTEOPOROSIS DETECTION FROM BONE X-RAY IMAGES

S. Neha Jabeen[1], Dr. Afshan Fatima[2]
[1]PG Scholar, Department of CS, Shadan Women's College of Engineering and Technology, Hyderabad, njabeens2112@gmail.com
[2]Assoc. Professor, Department of CSE, Shadan Women's College of Engineering and Technology
afshafatima0889@gmail.com

ABSTRACT

Osteoporosis is a skeletal condition that is hard to diagnose before symptoms appear. Due to cost and safety concerns, current skeletal disease screening techniques, such dual-energy X-ray absorptiometry, are only employed in certain situations after symptoms appear. In terms of early treatment and cost, early identification of osteopenia and osteoporosis employing alternative methods for comparatively frequent tests is beneficial. Deep learning-based osteoporosis detection techniques for a variety of modalities have been proposed in numerous recent papers, with excellent results. However, because these investigations include laborious procedures like physically cropping a region of interest or diagnosing osteoporosis rather than osteopenia, their clinical applicability is limited. In this work, we describe a classification task that uses computed tomography (CT) to diagnose osteopenia and osteoporosis. We also suggest a multi-view CT network (MVCTNet) that uses two images from the original CT scan to automatically classify osteopenia and osteoporosis. The MVCTNet extracts different features from the pictures produced by our multi-view settings, in contrast to earlier approaches that use a single CT image as input. The MVCTNet consists of three task layers and two feature extractors. Using the photos as distinct inputs, two feature extractors utilize dissimilarity loss to discover distinct features. The two feature extractors' features are used by the target layers to learn the target task, which they then aggregate. We employ a dataset with 2,883 patients' CT images classified as normal, osteopenia, and osteoporosis for the tests. Based on the quantitative and qualitative assessments, we also note that the suggested approach enhances the performance of every experiment.

1. INTRODUCTION

Generally speaking, a sender and receiver must first Osteoporosis is a skeletal condition that weakens bones and reduces calcium mass, which raises the risk of fractures. One of the main causes of fractures is osteoporosis, which is hard to identify before symptoms appear. Because of these factors, more than half of fracture patients have never been promptly diagnosed with osteoporosis. Many experts predict that as the population ages, there will be more osteoporotic individuals, which will result in higher medical expenses, injuries, and mortality rates. Prevention by early diagnosis and treatment of osteoporosis is crucial, and early identification and therapy are sufficient to prepare for these risks. DXA, or dual-energy X-ray absorptiometry, is frequently used to assess the risk of osteoporosis. Nevertheless, DXA can result in needless exposure to the other organs surrounding the test site and is rarely carried out until symptoms (such as a fracture) manifest. For these reasons, other common methods must be used for early diagnosis. For instance, computed tomography (CT) is a common procedure with comparatively little risk. Experts can also utilize CT to evaluate the risk of osteoporosis. Patients with normal, osteopenia, and osteoporosis CT scans are displayed in Fig. 1. The differences between the three situations are depicted in the photographs in the areas indicated by red rectangles. Osteopenia and osteoporosis exhibit texture variations with a decrease in the Hounsfield unit (HU), while the typical case has a homogeneous texture. Osteoporosis and osteopenia can be analyzed due to these CT characteristics. This can lessen the social and financial impact of osteoporotic fractures and offer a chance for early treatment. Consequently, we plan to diagnose osteoporosis using CT scans. We perform an osteoporotic classification task on a reasonably simple-to-obtain sagittal CT picture. We collected our dataset for the job from patients who had DXA and contrast-enhanced abdominal CT images. In Section III-E, we also give specifics about the dataset. Recent advancements in deep learning have made it possible for some studies to detect osteoporosis with significant success. However, the lack of data and manual procedures, which require laborious effort from radiology specialists, continue to pose serious obstacles. Yasaka demanded that the vertebrae be clipped out of the CT scans. The hip joint regions on the X-ray pictures were manually cropped. Additionally, they only used CT images that were cropped inside a specific HU range. Medical professionals employ a certain HU range to diagnose osteoporosis because HU has a wider pixel range than a picture. However, due to incomplete or superfluous information, deep neural networks may deteriorate osteoporosis diagnosis performance.

2. EXISTING SYSTEM OF PROJECT

- Numerous studies in photo-domain computer vision have suggested effective techniques, such as self-supervised and semi-supervised learning, that make use of multiple perspectives of the same instance in distinct domains. Inspired by these, we suggest a straightforward yet effective technique known as a multi-view CT network that makes advantage of two views produced by the CT domain knowledge (different HU clip settings in the same CT picture).

- Three target layers that learn the classification tasks and two feature extractors that extract various characteristics from two perspectives make up the DL.
- Additionally, features from the feature extractors are aggregated by the main task layer, one of the three task layers. We develop a task loss utilizing multi-class and ordinal classification to categorize osteoporosis risk. We also employ a dissimilarity loss that makes use of the outputs of the feature extractors.

Existing System Disadvantages:

- Limited in learning complex, high-dimensional data representations.
- Dependency on pre-trained models and potential domain shift issues.
- Sample inefficiency, requiring a large number of samples for effective learning.
- Computationally expensive, especially for longer sequences.

3. RELATED WORK

According to a review of the literature, attempts to classify bone radiographs can be broadly classified into three categories: convolutional neural networks, hand-crafted features, and hybrid approaches (combining both). To improve the orientation selectivity of the current dual-tree wavelet transform for OS identification, Oulhaj et al. [16] suggest using an anisotropic discrete dual-tree wavelet transform. This approach improves the classification of trabecular bone by enabling more accurate and effective feature extraction from X-ray images. This method works well since it allows for a multiresolution and multi-orientation analysis of the images due to the transforms' orientation and scale sensitivity. Su and others In order to take advantage of the complementing qualities of Gabor filters and Local Binary Patterns (LBP), [17] present a unique handcrafted feature set. Making use of the scale and direction selectivity of Gabor filters provide a reliable feature set for efficiently quantifying X-ray images when used with the localization feature of LBP features. Multifractals can be used to describe the trabecular bone texture in bone radiographs, according to Palanivel et al. [19]. In order to quantify global regularity, they first compute exponents and then Hausdorff dimensions. Lacunarity is then used as a quantitative metric to detect diseases in the bone structure, with encouraging outcomes. To identify OS in patients, Wang et al. [25] have used a thorough method that makes use of both local and global data taken from chest X-ray pictures. After using a landmark detection approach, the authors extract pertinent areas of interest (ROI) to produce a set of potential sub-images for pathology identification. Convolutional neural networks (CNNs) and a self-attention mechanism are then used in a hybrid design to identify pathology in the photos. The authors claim that, in contrast to many vision tasks, OS detection presents substantial problems, with traditional CNNs having limited usefulness. Although a more complete examination of bone mineral density (BMD) is thought to be required for a successful diagnosis, the incorporation of the attention mechanism produces encouraging quantitative results. Similar approaches to segmenting airways have recently been investigated, utilizing various data augmentation techniques to extract sub-pictures from a series of photos [28]. Gabor filters were used by Mebarkia et al. [26] to extract a feature set that included components from Local Phase Quantization (LPQ) and the Histogram of Oriented Gradients (HOOG). They performed feature fusion at both the score and decision levels, producing intriguing results, and used a bat-based optimization technique to fine-tune the parameters of the Gabor filters. For OS detection using X-rays, Wani et al. [24] provide a novel method based on convolutional neural networks (CNNs). The authors address the detection of OS as a three-class problem: normal, osteopenia (which indicates a potential risk for developing OS), and OS. They do this by using transfer learning with different CNN architectures, such as AlexNet [29], VGG [30], and ResNet [31]. The effectiveness of using deep learning techniques for this assignment is covered in great detail by the authors. Furthermore, according to the authors, the main goal of the research is to present a novel dataset intended to verify the efficacy of the suggested techniques.Nemati et al. [23] showed that the visual resemblance between normal and osteoporotic images makes them unsuitable for efficient processing using convolutional neural networks (CNNs), making the direct use of raw radiography images for OS diagnosis problematic. CT images were used by Yijie et al. [22] to identify OS. UNet was used to segment the pictures, while Dense-121 was used to calculate BMD. The authors showed a strong connection between the computed and ground truth values, suggesting a high degree of confidence in the use of CT images for OS identification. For opportunistic OS screening, Pan et al. [27] attempted to build a deep learning model using chest CT scans. The fusion feature images of the lumbar 1 and lumbar 2 vertebral bodies were examined, and the results were favourable for identifying OS in the patients. To identify OS in the photos, Jae et al. [18] used the so-called single column and multiple column approaches.

4. PROPOSED SYSTEM

- Professional clinicians in the medical world use CT to diagnose illnesses by keeping an eye on a particular HU range. This enables people to concentrate more on regions they wish to view, such organs or bones. Nevertheless, this might give neural networks inadequate data, which could impair their functionality.
- We use two distinct HU ranges to give the neural network additional information in order to solve this problem. We clamp the CT scan HU inside particular limits based on the various settings given the labeled CT scan {ct, y}.
- As a result, each CT configuration can yield two CT images with distinct texture information. The photos, which may be viewed as the several views for the same instance, are then used as input for the MVCTNet.

Proposed System Advantages:

- Understanding complex patterns at different levels
- Pre-trained CNNs adapt well to new tasks with limited data.
- Learned features can often be repurposed for improved performance in related tasks.

5. MODULES

1. Data Augmentation

By making altered versions of already-existing data samples, data augmentation is a technique used to artificially increase the size and diversity of the training dataset. It lessens overfitting and enhances model generalization, particularly when the dataset is small. Transformations like rotation, flipping, scaling, cropping, brightness correction, contrast improvement, and noise addition can all be considered forms of augmentation in image-based applications. By simulating various real-world scenarios and viewpoints, these changes make the model more resilient and able to adjust to new data. All things considered, data augmentation is essential for enhancing model stability and accuracy.

2. Data Preprocessing

Before raw data is fed into the model, it must be cleaned and prepared, a process known as data preprocessing. It guarantees that the dataset is accurate, reliable, and appropriate for training. Resizing photos to a consistent size, standardizing pixel values, eliminating noise, and, if required, changing image formats are typical preprocessing methods. Label encoding and one-hot encoding are sometimes used with categorical data. Faster convergence and better model performance result from this stage, which guarantees that the input data is in a format compatible with the deep learning or machine learning framework. The accuracy and dependability of the model's predictions are directly impacted by appropriate preprocessing.

3. Data Split

Splitting the dataset into distinct subsets for testing, validation, and training is known as data splitting. Depending on the quantity of the data and the needs of the model, the dataset is usually divided in a ratio like 70:20:10 or 80:10:10. The model is trained using the training set; hyperparameters are adjusted and overfitting is avoided using the validation set; and the final performance of the trained model is assessed using the test set. By ensuring that the model is trained on a subset of the data and tested on untested samples, this procedure guarantees an objective evaluation of the model's capacity for generalization.

4. Model Training

The process of putting preprocessed and supplemented data into a selected neural network or machine learning algorithm in order to extract patterns and relationships from the data is known as model training. In order to minimize the loss function during training, the model modifies its internal parameters (weights and biases) using optimization methods such stochastic gradient descent or the Adam optimizer. To make sure the model learns efficiently and avoids overfitting, its performance is tracked on the validation set. Until the model attains acceptable accuracy or the loss stabilizes, the training process is repeated. To guarantee accurate and dependable forecasts in practical situations, proper model training is crucial.

5. Prediction

Using the trained model to draw conclusions from fresh, unobserved data is the prediction step. Depending on the application, the model uses the correlations and patterns it discovered during training to forecast results like segmentations, classifications, or numerical values. Accuracy, precision, recall, and other performance indicators are then measured by comparing the predicted outcomes with the ground truth, if it is known. The prediction module can be included into a real-time system or application for practical deployments, enabling effective and automated decision-making based on input data. This stage illustrates the efficacy and practicality of the model.

6. TECHNIQUES USED IN PROJECT

6.1. EXISTING TECHNIQUE: -

The broad topic of deep learning (DL) includes a variety of techniques and structures designed to extract representations from data. Convolutional Neural Networks (CNNs) for image recognition, Recurrent Neural Networks (RNNs) for sequential data tasks such as natural language processing, and variants such as Long Short-Term Memory networks (LSTMs) and Gated Recurrent Units (GRUs) for long-range dependencies are some of the popular techniques. Furthermore, Autoencoders develop effective representations by compressing and rebuilding data, whereas Generative

Adversarial Networks (GANs) create new instances of data. By utilizing pre-trained models, transfer learning maximizes resources, and attention mechanisms improve performance by concentrating on particular input components.

6.2. PROPOSED TECHNIQUE USED OR ALGORITHM USED:

"MVCTNet" may refer to a specialized architecture created for particular computer vision tasks or applications if it is in fact a CNN model. Because CNNs can automatically build hierarchical representations from image data, they are well-known for their efficacy in tasks like segmentation, object identification, and picture classification. We train MVCTNet's feature extractors to conduct the osteoporosis categorization task and capture the features of various viewpoints. In order to do this, we extract representations, which are the inputs of the task layers and our dissimilarity loss and the outputs of the global average pooling (GAP) layer of the feature's extractors. In our tests, we use ResNet-18 [21] and EfficientNet-b0 [22], which have been pre-trained on the ImageNet dataset, as the two task layers and feature extractors (Ga, Gb).

7. RESULTS

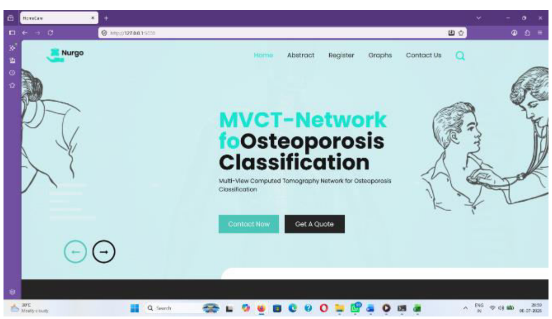

Fig: 7.1 Home page

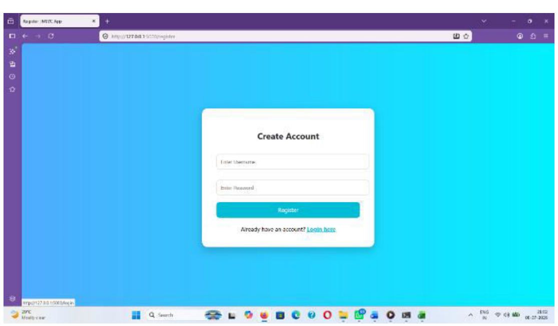

Fig:7.2 Registration Page

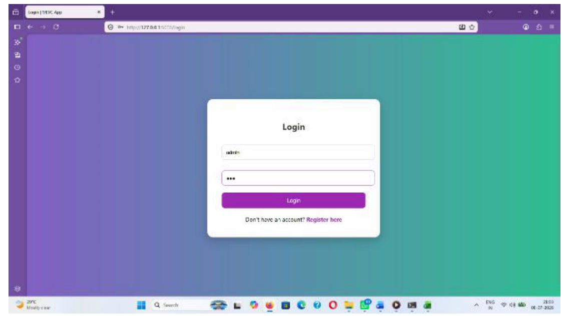

Fig: 7.3 Login Page

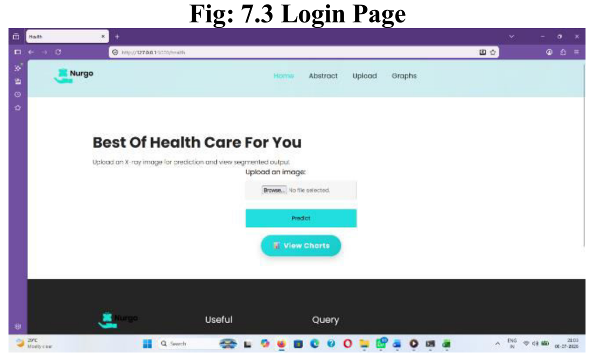

Fig 7.4 X-Ray Image Upload and Osteoporosis prediction

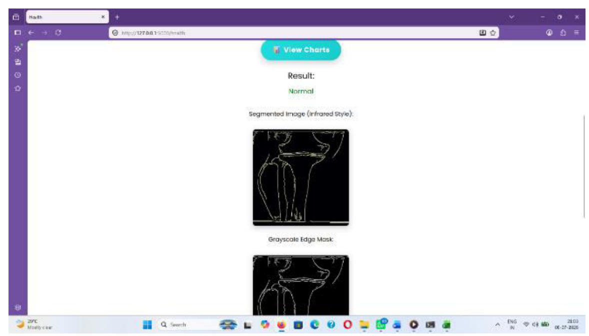

Fig 7.5 Normal Prediction Result Page

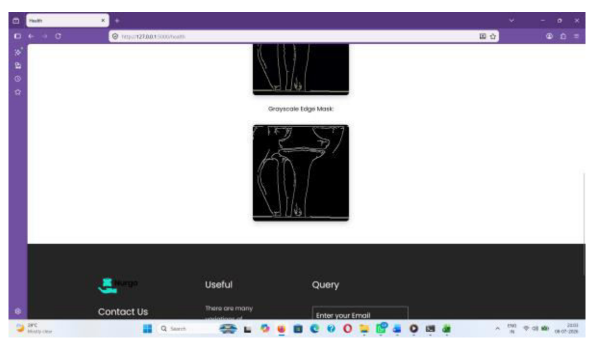

Fig 7.6 Result Page

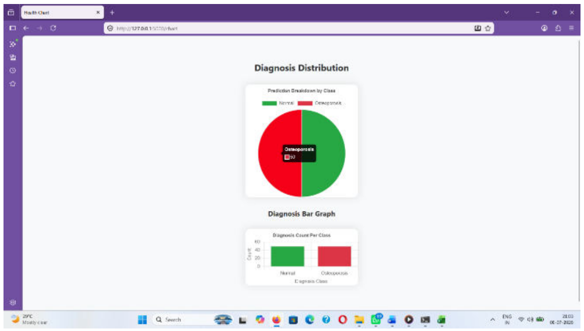

Fig 7.7 Diagnosis Distribution for Normal

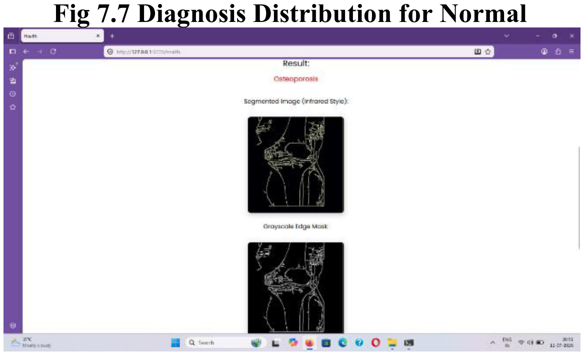

Fig 7.8 Osteoporosis Result Page

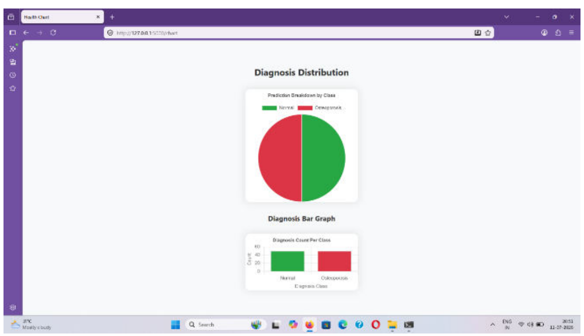

Fig 7.9 Diagnosis Distribution for Osteoporosis

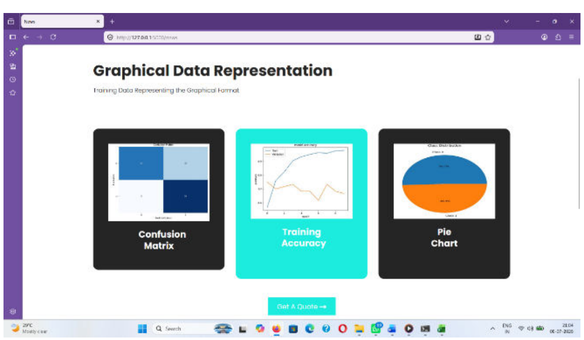

Fig 7.10 Graphical Data Representation

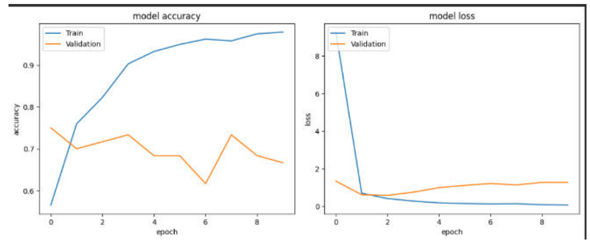

Fig 7.11 Model Accuracy

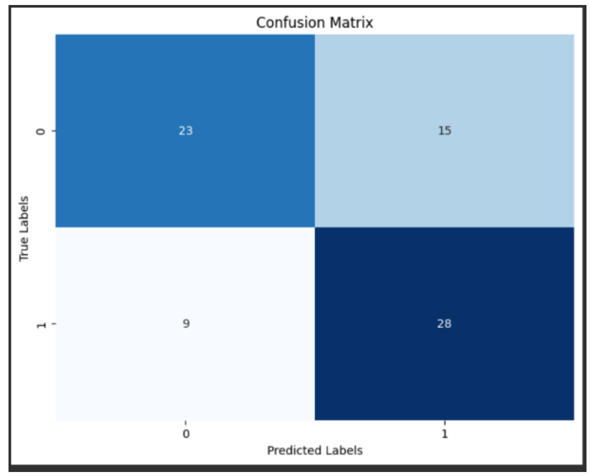

Fig 7.12 Confusion Matrix

8. CONCLUSION

In this study, we introduced a straightforward yet effective architecture, MVCTNet, for the osteoporosis detection job in sagittal CT scans. For the job, we gathered a sizable dataset of 2,883 individuals who had contrast-enhanced abdomen CT scans. Furthermore, as shown in Table 1, we completed the work utilizing the MVCTNet without the manual procedures that were previously employed for automatic osteoporosis diagnosis. The MVCTNet consistently performed better than alternative baseline techniques in multi-class classification and ordinal regression tasks, according to both quantitative and qualitative studies. Additionally, we verified that the MVCTNet's components are useful for diagnosing osteoporosis based on the ablation research. Although there are benefits to our MVCTNet, choosing a CT slice in a real-world medical setting can provide extra difficulties. To address these issues, we will provide a more comprehensive version of our model in further work, such as a 3D medical picture model.

9. FUTURE ENHANCEMENT

In the future, the research could look at a number of ways to increase the precision and usefulness of the "Multi-View Computed Tomography Network for Osteoporosis Classification." First, adding more sophisticated deep learning architectures, including ensemble models or attention mechanisms, may improve the model's capacity to identify intricate correlations in multi-view CT images. In order to improve the robustness and generalization of the model, the project might also benefit from a more diverse dataset that includes a wider range of patient demographics, diseases, and imaging devices. A more thorough understanding of osteoporosis may result from the integration of longitudinal data and temporal elements, which may offer insights into the course of the disease.

REFERENCES

[1] K. H. Nam, I. Seo, D. H. Kim, J. I. Lee, B. K. Choi, and I. H. Han, ‘‘Machine learning model to predict osteoporotic spine with Hounsfield units on lumbar computed tomography,’’ J. Korean Neurosurgical Soc., vol. 62, no. 4, pp. 442–449, Jul. 2019.

[2] K. Yasaka, H. Akai, A. Kunimatsu, S. Kiryu, and O. Abe, ‘‘Prediction of bone mineral density from computed tomography: Application of deep learning with a convolutional neural network,’’ Eur. Radiol., vol. 30, pp. 3549–3557, Jun. 2020.

[3] S. Rastegar, M. Vaziri, Y. Qasempour, M. R. Akhash, N. Abdalvand, I. Shiri, H. Abdollahi, and H. Zaidi, ‘‘Radiomics for classification of bone mineral loss: A machine learning study,’’ Diagnostic Interventional Imag., vol. 101, no. 9, pp. 599–610, Sep. 2020.

[4] B. Zhang, K. Yu, Z. Ning, K. Wang, Y. Dong, X. Liu, S. Liu, J. Wang, C. Zhu, Q. Yu, and Y. Duan, ‘‘Deep learning of lumbar spine X-ray for osteopenia and osteoporosis screening: A multicenter retrospective cohort study,’’ Bone, vol. 140, Nov. 2020, Art. no. 115561.

[5] I. M. Wani and S. Arora, ‘‘Computer-aided diagnosis systems for osteoporosis detection: A comprehensive survey,’’ Med. Biol. Eng. Comput., vol. 58, pp. 1873–1917, Sep. 2020.

[6] A. D. Smith, ‘‘Screening of bone density at CT: An overlooked opportunity,’’ Radiology, vol. 291, no. 2, pp. 368–369, May 2019.

[7] T. Sözen, L. Özíşík, and N. C. C. Başaran, ‘‘An overview and management of osteoporosis,’’ Eur. J. Rheumatology, vol. 4, no. 1, p. 46, 2017.

[8] J. Wu, Y. Qu, K. Wang, and Y. Chen, ‘‘Healthcare resource utilization and direct medical costs for patients with osteoporotic fractures in China,’’ Value Health Regional Issues, vol. 18, pp. 106–111, May 2019.

[9] P. J. Mitchell, ‘‘Fracture liaison services: The UK experience,’’ Osteoporosis Int., vol. 22, no. 3, pp. 487–494, Aug. 2011.

[10] P. A. Bromiley, E. P. Kariki, J. E. Adams, and T. F. Cootes, ‘‘Classification of osteoporotic vertebral fractures using shape and appearance modelling,’’ in Proc. Int. Workshop Comput. Methods Clin. Appl. Musculoskeletal Imag. Cham, Switzerland: Springer, 2017, pp. 133–147.

[11] N. Yamamoto, S. Sukegawa, A. Kitamura, R. Goto, T. Noda, K. Nakano, K. Takabatake, H. Kawai, H. Nagatsuka, K. Kawasaki, Y. Furuki, and T. Ozaki, ‘‘Deep learning for osteoporosis classification using hip radiographs and patient clinical covariates,’’ Biomolecules, vol. 10, no. 11, p. 1534, Nov. 2020.

[12] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, ‘‘Momentum contrast for unsupervised visual representation learning,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 9726–9735.

[13] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, ‘‘A simple framework for contrastive learning of visual representations,’’ in Proc. 37th Int. Conf. Mach. Learn., vol. 119, Jul. 2020, pp. 1597–1607.

[14] J.-B. Grill, F. Strub, F. Altche, C. Tallec, P. Richemond, E. Buchatskaya, C. Doersch, B. A. Pires, Z. Guo, M. G. Azar, and B. Piot, ‘‘Bootstrap your own latent—A new approach to self-supervised learning,’’ in Proc. Adv. Neural Inf. Process. Syst., vol. 33, 2020, pp. 21271–21284.

[15] K. Sohn, D. Berthelot, N. Carlini, Z. Zhang, H. Zhang, C. A. Raffel, E. D. Cubuk, A. Kurakin, and C.-L. Li, ‘‘FixMatch: Simplifying semisupervised learning with consistency and confidence,’’ in Proc. Adv. Neural Inf. Process. Syst., vol. 33, 2020, pp. 596–608.

[16] S.-H. Choi, D.-H. Hwang, D.-H. Kim, S.-H. Bak, and Y. Kim, ‘‘Efficient osteoporosis prediction using a pair of ensemble models,’’ J. Korea Soc. Comput. Inf., vol. 26, no. 12, pp. 45–52, 2021.

[17] K. Chaitanya, E. Erdil, N. Karani, and E. Konukoglu, ‘‘Contrastive learning of global and local features for medical image segmentation with limited annotations,’’ in Proc. Adv. Neural Inf. Process. Syst., vol. 33, 2020, pp. 12546–12558.

[18] D. Zeng, Y. Wu, X. Hu, X. Xu, H. Yuan, M. Huang, J. Zhuang, J. Hu, and Y. Shi, ‘‘Positional contrastive learning for volumetric medical image segmentation,’’ in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2021, pp. 221–230.

[19] Z. Li, Z. Cui, S. Wang, Y. Qi, X. Ouyang, Q. Chen, Y. Yang, Z. Xue, D. Shen, and J.-Z. Cheng, ‘‘Domain generalization for mammography detection via multi-style and multi-view contrastive learning,’’ in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2021, pp. 98–108.

[20] P. Yang, Z. Hong, X. Yin, C. Zhu, and R. Jiang, ‘‘Self-supervised visual representation learning for histopathological images,’’ in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2021, pp. 47–57.

[21] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Deep residual learning for image recognition,’’ in Proc. IEEE Conf.

Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.

[22] M. Tan and Q. Le, ‘‘EfficientNet: Rethinking model scaling for convolutional neural networks,’’ in Proc. Int. Conf. Mach. Learn., 2019, pp. 6105–6114.

[23] Z. Niu, M. Zhou, L. Wang, X. Gao, and G. Hua, ‘‘Ordinal regression with multiple output CNN for age estimation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 4920–4928.

[24] W. Cao, V. Mirjalili, and S. Raschka, ‘‘Rank consistent ordinal regression for neural networks with application to age estimation,’’ Pattern Recognit. Lett., vol. 140, pp. 325–331, Dec. 2020.

[25] H. Zhu, H. Shan, Y. Zhang, L. Che, X. Xu, J. Zhang, J. Shi, and F.-Y. Wang, ‘‘Convolutional ordinal regression forest for image ordinal estimation,’’ IEEE Trans. Neural Netw. Learn. Syst., vol. 33, no. 8, pp. 4084–4095, Aug. 2022.

[26] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, ‘‘Grad-CAM: Visual explanations from deep networks via gradient-based localization,’’ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 618–626.