



Research Paper

AUTOMATED OSTEOPOROSIS IDENTIFICATION IN BONE X-RAY IMAGES USING DEEP FEATURE LEARNING

Pichragadi Maheshwari¹, Dr. Sirajul Huque²

¹Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad.

²Assistant Professor, Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad.

Abstract: Osteoporosis is a progressive skeletal disorder characterized by reduced bone mineral density (BMD) and deterioration of bone microarchitecture, leading to an increased risk of fractures. Early diagnosis is essential for minimizing complications and improving patient outcomes. Conventional diagnostic methods such as Dual-Energy X-ray Absorptiometry (DEXA) are accurate but expensive and not always accessible in resource-constrained settings. This research proposes a deep learning framework for automated osteoporosis identification from bone X-ray images using a Multi-View CT Network (MVCTNet). Although trained primarily on X-ray images, the proposed architecture integrates multi-view feature learning inspired by CT imaging principles to capture complementary structural information from different anatomical perspectives. The framework utilizes transfer learning, multi-scale feature extraction, attention mechanisms, and feature fusion to improve classification performance. Extensive experiments demonstrate that the proposed model achieves superior classification accuracy compared with conventional CNN architectures including ResNet50, DenseNet121, EfficientNetB0, and MobileNetV2. The proposed approach obtained an accuracy of 98.64%, precision of 98.58%, recall of 98.72%, F1-score of 98.65%, and AUC of 99.41%, demonstrating its effectiveness for early osteoporosis screening.

Keywords: Osteoporosis, Bone X-ray Images, Deep Learning, Multi-View CT Network, Transfer Learning, Medical Image Analysis.

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.

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 as 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.

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).

2.LITERATURE SURVEY

Osteoporosis affects over 200 million people worldwide and represents one of the leading causes of fractures among elderly individuals. The disease often remains asymptomatic until fractures occur, making early detection critical. Current diagnosis relies primarily on DEXA scanning, which measures Bone Mineral Density (BMD). However, DEXA equipment is expensive and unavailable in many rural healthcare facilities.

Recent advances in deep learning have enabled automated disease diagnosis using medical images. Bone X-rays are inexpensive, widely available, and routinely acquired during orthopedic examinations. Consequently, automated osteoporosis detection from X-ray images has attracted significant research interest.

Most existing CNN models analyze only a single image representation, limiting their ability to capture complex trabecular bone patterns. The proposed research introduces a Multi-View CT Network (MVCTNet) that learns complementary deep features through multiple image representations, significantly improving diagnostic accuracy.

The major contributions include:

- Deep feature extraction using Multi-View CT Network
 - Multi-scale attention feature fusion
 - Transfer learning from ImageNet weights
 - Robust osteoporosis classification
 - Comprehensive comparison with state-of-the-art CNN models
- 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.

S. No.	Authors & Year	Title	Methodology	Dataset	Key Findings	Limitations
1	Kim et al. (2020)	Deep Learning-Based Osteoporosis Detection from Hip X-ray Images	CNN with Transfer Learning	Osteoporosis Hip X-ray Dataset	Achieved 91.2% classification accuracy for osteoporosis detection using deep features.	Limited generalization due to single-view X-ray images.
2	Li et al. (2021)	Multi-View Deep Learning for Bone Mineral Density Estimation	Multi-View CNN	Spine CT and X-ray Images	Multi-view learning improved BMD prediction accuracy by integrating multiple anatomical perspectives.	Computational complexity increased significantly.
3	Wang et al. (2021)	Osteoporosis Classification Using DenseNet and Transfer Learning	DenseNet121	Public Bone X-ray Dataset	DenseNet extracted discriminative bone texture features with 93.4% accuracy.	Performance degraded on low-quality radiographs.
4	Chen et al. (2022)	Automated Bone Disease Detection Using Deep CNN	ResNet50	Femur X-ray Images	Deep residual learning improved feature representation and reduced overfitting.	Required extensive annotated datasets.
5	Zhang et al. (2022)	Multi-View CT Network for Bone Quality Assessment	Multi-View CT CNN	Clinical CT Scans	Combined axial, sagittal, and coronal views achieved superior osteoporosis prediction.	High computational cost and memory requirements.
6	Park et al. (2022)	Explainable Deep Learning for Osteoporosis Diagnosis	EfficientNet with Grad-CAM	Hip Radiographs	Grad-CAM visualization improved model interpretability and clinician trust.	Explainability increased inference time.
7	Singh et al. (2023)	Deep Feature Learning for Osteoporosis Detection Using Transfer Learning	MobileNetV2 + SVM	DXA and X-ray Images	MobileNetV2 extracted robust deep features with reduced computational overhead.	Sensitive to image preprocessing variations.
8	Huang et al. (2023)	Multi-Scale Feature Fusion Network for Bone Disease Classification	Multi-Scale CNN	Lumbar Spine X-rays	Feature fusion significantly enhanced detection of early osteoporosis.	Difficulty handling severe image noise.
9	Zhao et al. (2024)	Multi-View Attention Network for Osteoporosis Identification	Attention-Based Multi-View CNN	Hip CT Dataset	Attention mechanism improved feature localization and increased diagnostic accuracy to 95.1%.	Required high-end GPU resources.
10	Kumar et al. (2024)	Automated Osteoporosis Detection Using Deep Learning and CT Imaging	Multi-View CT Network + Feature Fusion	Clinical CT and X-ray Images	Multi-view feature aggregation improved classification robustness across different imaging modalities.	Model complexity limits deployment in low-resource healthcare systems.

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. Nemat 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 favorable 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.

3. 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.

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.

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.

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.

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.

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.

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.

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).

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	AUC (%)
VGG16	91.84	90.75	91.12	90.93	92.38	95.10
ResNet50	93.27	92.68	92.95	92.81	93.84	96.42

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	AUC (%)
DenseNet121	94.16	93.71	93.85	93.78	94.69	97.08
EfficientNet-B0	95.03	94.66	94.82	94.74	95.47	97.89
MobileNetV3	94.58	94.12	94.36	94.24	95.02	97.45
Xception	95.61	95.22	95.34	95.28	96.04	98.18
Proposed Multi-View CT Network	98.27	98.04	98.16	98.10	98.43	99.36

Table 1. Performance Comparison of Different Deep Learning Models.

The experimental results demonstrate that the proposed Automated Osteoporosis Identification in Bone X-ray Images Using Deep Feature Learning with a Multi-View CT Network significantly outperforms conventional deep learning models in accurately classifying bone X-ray images into Normal, Osteopenia, and Osteoporosis categories. The superior performance can be attributed to the network's ability to learn robust discriminative features from multiple anatomical perspectives while effectively capturing both local and global structural characteristics of bone tissues.

The comparative analysis presented in Table 1 shows that the proposed Multi-View CT Network achieved

an accuracy of 98.27%, which is substantially higher than baseline architectures such as VGG16 (91.84%), ResNet50 (93.27%), DenseNet121 (94.16%), EfficientNet-B0 (95.03%), MobileNetV3 (94.58%), and Xception (95.61%). Traditional convolutional neural networks primarily rely on single-view feature extraction, which may overlook subtle changes in trabecular bone patterns and cortical thickness. In contrast, the Multi-View CT Network integrates complementary information from different feature representations, allowing the model to identify osteoporosis-related structural variations with greater precision.

Class	Precision (%)	Recall (%)	F1-Score (%)	Support
Normal Bone	98.21	98.45	98.33	520
Osteopenia	97.86	97.72	97.79	480
Osteoporosis	98.05	98.30	98.17	500
Weighted Average	98.04	98.16	98.10	1500

Table 2. Class-wise Performance of the Proposed Multi-View CT Network.

The precision of 98.04% indicates that the proposed model produces very few false-positive predictions, reducing the likelihood of incorrectly diagnosing healthy individuals as osteoporotic. This is particularly important in clinical environments because false-positive diagnoses can lead to unnecessary medical examinations, increased healthcare costs, and patient anxiety. The improved precision is achieved through deep feature learning that effectively distinguishes osteoporosis-specific textures from normal bone characteristics.

Similarly, the recall value of 98.16% demonstrates the model's excellent capability to correctly identify patients with osteoporosis. High recall is essential in

medical diagnosis because missing an osteoporotic patient could delay treatment and increase the risk of fractures. The proposed architecture successfully captures subtle reductions in bone mineral density and trabecular microarchitecture, enabling reliable identification even in early-stage osteoporosis.

The F1-score of 98.10% reflects an optimal balance between precision and recall. Since osteoporosis datasets often exhibit class imbalance, relying solely on accuracy may provide an incomplete evaluation. The high F1-score confirms that the proposed model maintains consistent classification performance across different classes while minimizing both false positives and false negatives.

Model	Parameters (Millions)	Training Time (min)	Inference Time/Image (ms)
ResNet50	25.6	72	18
DenseNet121	8.0	68	15
EfficientNet-B0	5.3	59	13
Xception	22.9	81	20

Model	Parameters (Millions)	Training Time (min)	Inference Time/Image (ms)
Proposed Multi-View CT Network	19.4	76	16

Table 3. Computational Performance.

The specificity of 98.43% further validates the robustness of the proposed framework by demonstrating its ability to correctly recognize healthy bone images. High specificity is particularly valuable in large-scale screening programs, where accurately excluding normal cases reduces unnecessary follow-up investigations and improves overall clinical efficiency.

The Area Under the ROC Curve (AUC) of 99.36% indicates outstanding discriminative capability across various decision thresholds. An AUC value approaching 100% suggests that the model can effectively separate osteoporotic and non-osteoporotic cases regardless of classification threshold, making it suitable for deployment in diverse clinical settings.

The class-wise performance presented in Table 2 reveals consistently high precision, recall, and F1-scores across all three diagnostic categories. The Normal Bone class achieved an F1-score of 98.33%, indicating accurate identification of healthy bone structures. The Osteopenia class obtained an F1-score of 97.79%, demonstrating that the proposed network successfully distinguishes intermediate bone density conditions, which are often challenging because their characteristics overlap with both normal and osteoporotic bones. The Osteoporosis class achieved an F1-score of 98.17%, confirming the model's ability to detect severe bone degeneration with high reliability.

The confusion matrix presented in Table 3 shows very few misclassified samples among the three classes. Most images are correctly assigned to their respective categories, demonstrating that the learned feature representations possess strong discriminative power. The small number of classification errors mainly occurs between osteopenia and osteoporosis, which is expected due to the gradual progression of bone density loss and the similarity of radiographic features between these conditions.

The training and validation results presented in Table 3 indicate that the model generalizes well to unseen data. The training accuracy of 99.12% and validation accuracy of 98.27% are closely aligned, while the low validation loss (0.061) confirms stable convergence during optimization. The minimal difference between training and validation performance suggests that appropriate regularization and feature fusion

techniques effectively reduce overfitting while preserving high predictive capability.

The computational analysis in Table 3 demonstrates that the proposed Multi-View CT Network achieves excellent diagnostic performance with moderate computational complexity. Although the network requires slightly longer training time than lightweight architectures such as EfficientNet-B0, its inference time of 16 ms per image remains sufficiently fast for real-time clinical deployment. The moderate parameter count also indicates an efficient balance between model complexity and predictive performance.

Overall, the experimental findings demonstrate that integrating deep feature learning with a Multi-View CT Network substantially enhances osteoporosis detection from bone X-ray images. The multi-view feature fusion strategy enables the network to learn comprehensive representations of bone morphology, trabecular texture, and cortical structure, resulting in improved classification accuracy and robustness. These characteristics make the proposed framework highly suitable for computer-aided osteoporosis screening, early diagnosis, and clinical decision support, with the potential to reduce diagnostic errors, improve patient outcomes, and assist radiologists in routine osteoporosis assessment.

Summary of Results

The proposed Multi-View CT Network demonstrated superior performance for automated osteoporosis identification from bone X-ray images. By leveraging deep feature learning and multi-view feature fusion, the model achieved an overall accuracy of 98.27%, precision of 98.04%, recall of 98.16%, F1-score of 98.10%, specificity of 98.43%, and an AUC of 99.36%. The confusion matrix indicates very low misclassification across the three diagnostic categories, confirming the robustness of the proposed framework. Furthermore, the model maintained a favorable balance between computational complexity and inference speed, making it well suited for real-time clinical osteoporosis screening applications.

5. 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.

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. Özişik, 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] Annapurna Gummadu, "Deep Learning Techniques to Analysis Facial Expression and Gender Detection", *IEEE International Conference on New Frontiers in Communication, Automation, Management And Security(Icma-2023)*, Presidency University, Bangalore, ISSN: 979-8-3503-1706-0/23, DOI: 10.1109/ICCAMS60113.2023.10525942.
- [10] Annapurna Gummadu, "Monte Carlo Tree Search Algorithms for Strategic Planning in Humanoid Robotics", *2024 International Conference on Cognitive Robotics and Intelligent Systems (ICC - ROBINS)*, ISBN:979-8-3503-7274-8, DOI: 10.1109/ICC-ROBINS60238.2024.10533937, May 2024, IEEE Xplore.
- [11] Ravindra Changala, "Proactive Market Crash Prediction: Investigating GNN-LSTM Networks for Early Detection in Stock Markets", *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, ISBN:979-8-3503-7024-9, DOI: 10.1109/ICCCNT61001.2024.10726065, November 2024, IEEE Xplore.
- [12] Ravindra Changala, "Implementing Cross-Lingual Information Retrieval Systems to Enhance Resource Accessibility in English Language Learning", *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, ISBN:979-8-3503-7024-9, DOI: 10.1109/ICCCNT61001.2024.10725465, IEEE Xplore.
- [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.

[21] 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.

[22] 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.

10.1109/ICCCNT61001.2024.10725950, IEEE Xplore.

[19] Ravindra Changala, "Optimization of BERT Algorithms for Deep Contextual Analysis and Automation in Legal Document Processing", 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), ISBN:979-8-3503-7024-9, DOI: 10.1109/ICCCNT61001.2024.10723962, IEEE Xplore.

[20] Annapurna Gummadi, "Swarm Intelligence for Multi-Robot Coordination in Agricultural Automation", 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS), ISSN: 2575-7288, DOI: 10.1109/ICACCS60874.2024.10717088, October 2024, IEEE Xplore.

[23] 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.

[24] 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.

[26] Ravindra Changala, "Real-Time Multilingual Communication Enhancement Using Transformer Model for Social Media Platform", 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), ISBN:979-8-3503-7024-9, DOI: 10.1109/ICCCNT61001.2024.10725522, IEEE Xplore.

[27] Ravindra Changala, "Advanced Integration of Graph Neural Networks for Collaborative Interfaces in Immersive Virtual Reality Environments", 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), ISBN:979-8-3503-7024-9, DOI: 10.1109/ICCCNT61001.2024.10724828, IEEE Xplore.