



Original Article

Wrist fracture detection using self-supervised learning methodology

Sachin Ramdas Thorat, MIM.¹, Davendranath G. Jha, PhD.¹, Ashish K. Sharma, PhD.², Dhanraj V. Katkar, BE.³

¹Department of Data Science and Technology, K.J. Somaiya Institute of Management, Somaiya Vidhyavihar University, Mumbai, Maharashtra, ²Department of Computer Science and Engineering, IIT Banaras Hindu University, Varanasi, Uttar Pradesh, ³Department of Computer Science, Mumbai University Mumbai, Maharashtra, India.

***Corresponding author:**

Sachin Ramdas Thorat,
Research Scholar,
Department of Data Science
and Technology, K.J. Somaiya
Institute of Management,
Somaiya Vidhyavihar
University, Mumbai,
Maharashtra, India.

Sachin.thorat@Somaiya.edu

Received: 19 December 2023
Accepted: 12 February 2024
Epub Ahead of Print: 26 March 2024
Published: 30 April 2024

DOI

10.25259/JMSR_260_2023

Quick Response Code:



ABSTRACT

Objectives: This study aimed to assist radiologists in faster and more accurate diagnosis by automating bone fracture detection in pediatric trauma wrist radiographic images using self-supervised learning. This addresses data labeling challenges associated with traditional deep learning models in medical imaging.

Methods: In this study, we trained the model backbone for feature extraction. Then, we used this backbone to train a complete classification model for classifying images as fracture or non-fracture on the publicly available Kaggle and GRAZPERDWRI-DX dataset using ResNet-18 in pediatric wrist radiographs.

Results: The resulting output revealed that the model was able to detect fracture and non-fracture images with 94.10% accuracy, 93.21% specificity, and an area under the receiver operating characteristics of 94.12%.

Conclusion: This self-supervised model showed a promising approach and paved the way for efficient and accurate fracture detection, ultimately enhancing radiological diagnosis without relying on extensive labeled data.

Keywords: Bone fracture, Diagnostic accuracy, Medical imaging, Radiography images, Self-supervised machine learning

INTRODUCTION

Fractures caused by various factors, such as falls, accidents, and sports injuries, have profound implications for individuals' health and quality of life. To prevent further complications and disabilities, timely diagnosis and treatment are crucial. In addition, the rising global population challenges ensuring universal access to essential medical services. Fractures are expected to surge with aging populations disproportionately impacting low- and middle-income countries.^[1]

Conventionally, fractures are diagnosed through radiological imaging techniques such as radiographs, magnetic resonance imaging (MRI), and computed tomography (CT) scans. While radiographs are cost-effective and widely used, CT and MRI scans provide more detailed fracture visualization. However, the availability of experienced radiologists varies across health-care facilities leading to disparities in accurate diagnosis.^[2] This inconsistency in expertise could potentially jeopardize timely and effective patient care.

How to cite this article: Thorat SR, Jha DG, Sharma AK, Katkar DV. Wrist fracture detection using self-supervised learning methodology. J Musculoskelet Surg Res. 2024;8:133-41. doi: 10.25259/JMSR_260_2023

Moreover, the medical imaging landscape is evolving rapidly, marked by a projected 30% increase in medical imaging studies by 2025. Consequently, radiologists face mounting pressures, as evidenced by the need for an average radiologist to interpret an image every 3–4 s to manage clinical demands.^[3] This increased workload contributes to diagnostic delays and errors.^[4] This urgency for accurate and efficient diagnosis drives the need for artificial intelligence (AI) integration in medical imaging workflows.^[5]

Integrating AI holds great promise in addressing these challenges. The AI can augment radiologists' diagnostic capabilities by identifying anomalies and generating reports allowing radiologists to focus on intricate cases and improving overall diagnosis speed and accuracy.^[6]

However, developing AI models for medical imaging is constrained by the need for high-quality labeled data due to the labor-intensive nature of the annotation.^[7] Curating labeled datasets for every medical imaging task becomes impractical within the realm of healthcare automation possibilities.^[8] This predicament underscores the necessity of innovative strategies that leverage unlabeled data for training.

Self-supervised learning emerges as a promising paradigm to address this data scarcity issue. By enabling models to learn representations from unlabeled data, self-supervised learning sidesteps the need for extensive labeled datasets. This approach is particularly appealing for tasks such as fracture detection where labeled data is limited. The research presented in this paper explores the application of self-supervised learning techniques to fracture detection within the medical domain. The objective is to develop an automated system capable of aiding radiologists in accurately classifying wrist radiographic images as fracture or non-fracture employing a self-supervised learning model ensemble. Through this study, we aimed to contribute to the advancement of fracture detection by harnessing the potential of self-supervised learning ultimately enhancing the efficiency and accuracy of radiological diagnosis without being contingent on abundant labeled data.

The subsequent sections of the paper detail the methodology, experimental results, discussion, and conclusion elaborating on the proposed approach and its implications for medical imaging.

The results of this research contribute to advancing fracture detection methodologies offering potential applications in clinical practice and patient care.

Fracture detection approaches

Deep learning has gained significant attention in medical imaging due to its ability to extract complex image features automatically. Deep learning algorithms, particularly convolutional neural networks (CNNs), have revolutionized

medical image analysis. These algorithms can automatically extract complex features from medical images making them suitable for fracture detection tasks.^[9]

Several approaches have been proposed for fracture detection using conventional methods. A computer-aided diagnosis system for detecting wrist fractures from radiographic images was developed achieving an accuracy of 86.4% using traditional machine learning techniques including support vector machines (SVMs) and decision trees.^[10]

A fractal-based method for rib fracture detection in CT scans was introduced by Yao *et al.* with the calculation of the fractal dimension of the fractured ribs resulting in a sensitivity of 93%.^[11] In 2016, a method based on a CNN for detecting fractures in wrist radiographs was proposed in a study achieving an accuracy of 92.5%.^[12] In 2020, a CNN was utilized for identifying hip fractures in radiographic images leading to an accuracy of 94.2% and a precision of 96.3%.^[13]

These approaches demonstrate the potential of supervised learning in fracture detection but highlight the reliance on labeled data, which may limit scalability and generalization. Furthermore, while effective, these methods often require handcrafted features and extensive manual labeling making them less scalable and potentially less accurate in complex cases.

These deep-learning approaches typically require large amounts of labeled data for training, which can be a limiting factor in the medical domain due to the need for expert annotations. Training supervised learning models requires expensive data labeling, and even then, they can struggle with challenges such as poor generalization, false relationships, and vulnerabilities to manipulation.^[14]

Self-supervised learning

Self-supervised learning has gained traction as an alternative approach to leverage the abundance of unlabeled data for training deep models. Self-supervised learning offers a solution to the labeled data scarcity problem. By leveraging the inherent structure or information within the unlabeled data, self-supervised learning algorithms can train deep learning models in a semi-supervised or unsupervised manner.^[15]

One popular self-supervised learning technique is contrastive learning, which learns representations by maximizing agreement between augmented views of the same image and minimizing agreement between views from different images.^[15] It aims to embed augmented versions of the same sample close to each other while pushing away embeddings from different samples. By formulating pretext tasks that generate supervisory signals from the data itself, self-supervised learning enables the models to learn meaningful representations without manual annotations.^[14]

This approach could address the scarcity and cost of labeled fracture datasets.

Furthermore, integrating self-supervised learning with transfer learning techniques has demonstrated even more significant performance gains. Leveraging pre-trained models on large-scale datasets, the learned representations can transfer the learned representations to the fracture detection task with limited labeled data.^[16] This transfer-learning approach has shown improved fracture detection accuracy and efficiency, particularly in scenarios with limited annotated fracture images.

Recent studies have explored applying self-supervised learning techniques for fracture detection in medical imaging. Cho *et al.* proposed a self-supervised framework based on contrastive learning to detect fractures in radiographic images.^[17] Their results demonstrated improved fracture detection performance compared to traditional methods. Similarly, another study utilized a self-supervised approach based on rotation prediction to detect fractures in CT scans achieving competitive results.^[18] Doersch *et al.* introduced the concept of pretext tasks, such as predicting image rotations, to learn representations from natural images.^[19] Inspired by this, recent studies have applied self-supervised learning to medical images for various tasks.

A self-supervised approach for skin lesion classification using dermoscopic images was introduced achieving competitive results compared to supervised methods.^[20] Furthermore, self-supervised learning study for lung nodule detection in CT scans leads to improved performance over traditional supervised models.^[21] These studies highlight the potential of self-supervised learning to capture clinically relevant features without extensive manual annotation.

MATERIALS AND METHODS

Dataset

In this study, we used the pediatric trauma wrist radiograph dataset from Kaggle and GRAZPERDWRI-DX dataset to

train a machine-learning model to detect wrist fractures. The Kaggle dataset contains 20,327 annotated images, which were divided into a training set of 16,327 images and a testing set of 4000 images.

The GRAZPEDWRI-DX dataset is a large and diverse collection of pediatric wrist radiographic images that are publicly available for free download. The dataset consists of 20,327 images from 6091 patients who were treated at the University Hospital Graz between 2008 and 2018. The images were annotated by pediatric radiologists, and the dataset is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. GRAZPEDWRI-DX is used for backbone training, and the Kaggle dataset is used for model training.

Our method harnesses the power of self-supervised learning algorithms, precisely the contrastive learning approach. Simplified contrastive learning (SimCLR) is a way to teach a computer to understand images without showing it any labeled examples. It does this by training the computer to find image similarities and differences. The learned representations can then be used for other tasks, such as image classification, without having to train the computer from scratch.

SimCLR consists of two main components:

1. Base encoder network: This network extracts features from the input images.
2. Contrastive loss function: This function trains the model by maximizing the similarity between the features of augmented versions of the same image and minimizing the similarity between the features of different images.

Base encoder networks such as ResNet-18 are typically pre-trained CNNs [Figures 1 and 2].

Data pre-processing

As a part of the pre-processing phase, uniformity is ensured by resizing all images to dimensions of 512×512 . Normalization is employed to standardize pixel values across the dataset, which aids in enhancing model performance during training and evaluation.

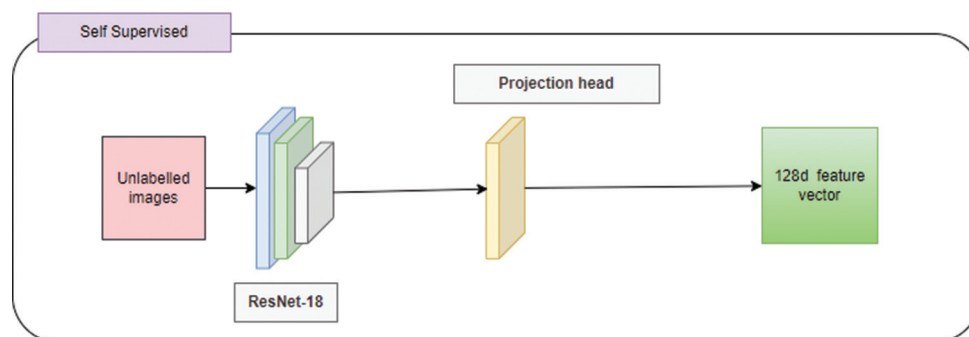


Figure 1: Self-supervised methodology – The unlabeled images are trained using ResNet-18 using a projection head to develop a 128d feature vector.

Data augmentation and pair generation

To enrich the diversity of our dataset, we employ data augmentation techniques, which encompass a range of random transformations such as rotations, translations, flips, and brightness adjustments. These augmented versions of the images serve as inputs to the subsequent stages of our approach.

For training, we generate positive and negative image pairs. Creating this paired dataset constitutes a pivotal step in our self-supervised learning framework. Positive pairs consist of augmented images originating from the same underlying image, capturing its inherent features under diverse transformations. Negative pairs are formed by combining augmented images from different original images. This dataset of image pairs is subsequently utilized to train our fracture detection model.

Architecture

We employ a ResNet-18 deep CNN as the backbone for feature extraction [Figure 3]. The SimCLR model consists of a ResNet-18 feature backbone (with the classification head removed) and a linear projection head. After training the contrastive model, we would discard the projection head and add an appropriate classification head for the downstream task. The training procedure relies heavily on producing augmented

pairs for each input data point. Those pairs, called positive samples, are the basis for the contrastive learning objective, where we optimize our model to produce representations that achieve high similarity between the positive pairs and low similarity with all the other augmented data points (negatives).

SimCLR projection head

We incorporate the SimCLR algorithm’s projection head mechanism to enhance feature extraction [Figure 4]. The projection head is integrated into the backbone neural network, ResNet-18 architecture. This augmented backbone learns to extract pertinent features from radiographic images, effectively representing both fractured and non-fractured images. The Cosine similarity function is calculated as

$$\begin{aligned} \text{sim}(u,v) &= \frac{U^T V}{\|U\| \|V\|} \\ &= \hat{U}^T \hat{V} \end{aligned}$$

sim: Cosine similarity, u and v: Vectors whose similarity to be found, U and V: Matrices that hold the vectors u and v as rows, T: Transpose operation.

The application of SimCLR involves augmenting the dataset of fractured reports with specific transformations. Pairs of augmented data points are then created, and the SimCLR

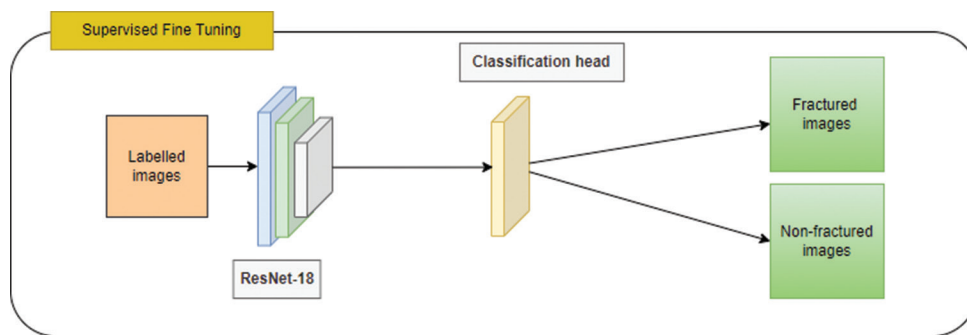


Figure 2: Supervised fine tuning – The labeled images use the learned representations to detect fracture and non-fracture images.

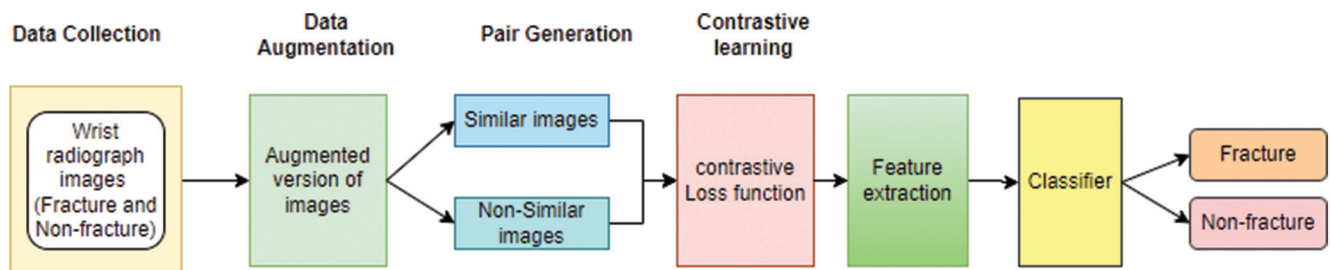


Figure 3: Architecture of self-supervised learning – Data collection, data augmentation, and pair generation followed by contrastive learning for feature extraction to classify fracture and not-fracture images.

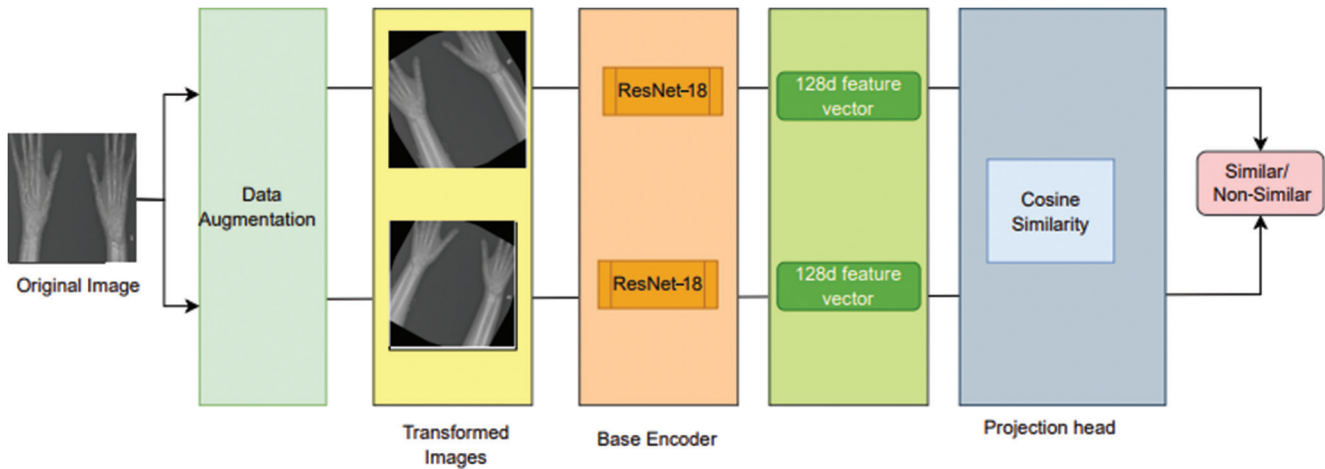


Figure 4: Simplified contrastive learning – Project head is used to augment images. The transformed images using base encoder ResNet-18 to develop 128d feature vectors.

algorithm is employed to learn and maximize the similarity of pairs from the same class while minimizing the similarity of pairs from different classes. This learning process ensures that representations are robust and discriminative, aligning well with fracture detection requirements.

The learned representations from the SimCLR algorithm serve as the basis for training a fracture detection classifier. The classifier’s training is executed on a separate held-out test set containing fractured images.

Loss function

SimCLR uses a contrastive loss function to measure the similarity between two views of an image. The goal is to maximize the similarity between positive pairs (two views of the same image) and minimize the similarity between negative pairs (two views of different images). The model’s objective is achieved through a contrastive loss function – NT-Xent Loss, which penalizes errors in predicting similarities. The NT-Xent loss function for a positive pair is calculated as

$$l_{i,j} = -\log \frac{\exp\left(\frac{\text{sim}(z_i, z_j)}{\tau}\right)}{\sum_{k=1}^{2N} [k \neq i] \exp\left(\frac{\text{sim}(z_i, z_k)}{\tau}\right)}$$

l: Loss value, *i* and *j*: Indexes referring to a specific positive pair of examples within a batch, *sim*: Cosine similarity function, *exp*: Exponential function, *z*: Vectors produced by the neural network for each image in the batch, *N*: Total number of images in a batch, *T*: Temperature hyperparameter, *k*: Index that iterates over all the examples in a batch (of size 2*N*) excluding the current positive example.

Model training

The model is trained using a stochastic gradient descent (SGD) optimizer, which is an iterative algorithm that updates the model parameters by moving in the direction of the negative gradient of the loss function. During training, the weights of the network are adjusted based on the contrastive loss. This loss function measures both the similarity between positive pairs (representing related data points) and the dissimilarity between negative pairs (representing unrelated data points). By minimizing the contrastive loss, the network learns to encode similar data points close together in the feature space, while pushing dissimilar data points apart. This process ultimately helps the network learn meaningful representations of the data.

Accuracy charts

Accuracy charts are used to visualize the performance of a model on a given task. The charts are explained below [Figures 5 and 6].

Area under the receiver operating characteristics (AUROC)

The classification head was trained for 40 epochs. The entire model was then trained from epoch 40 to 50 results.

Fine tuning

After training a neural network on a supervised task, the base network and projection head is fine-tuned using labeled data for a specific task. This helps to adapt the learned representations to the new task. The network is fine-tuned on labeled data to learn the specific features of fractures that are important for detection.

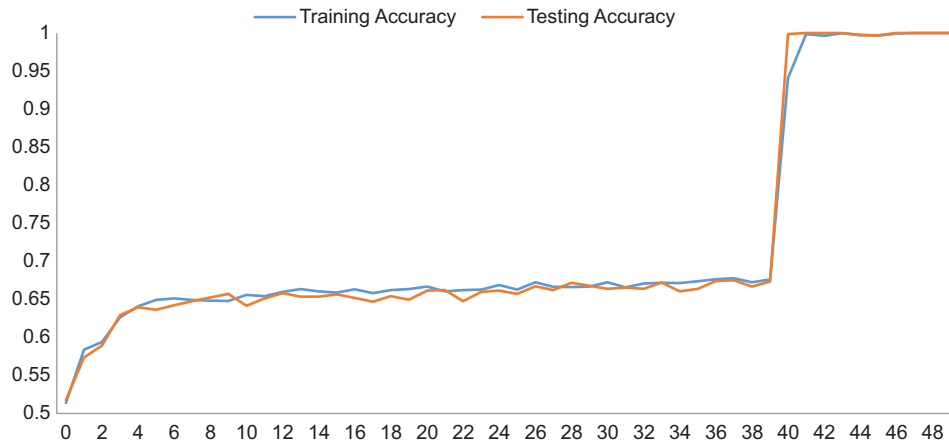


Figure 5: Plots of accuracy versus the number of epochs obtained while training a linear classifier on labeled training data.

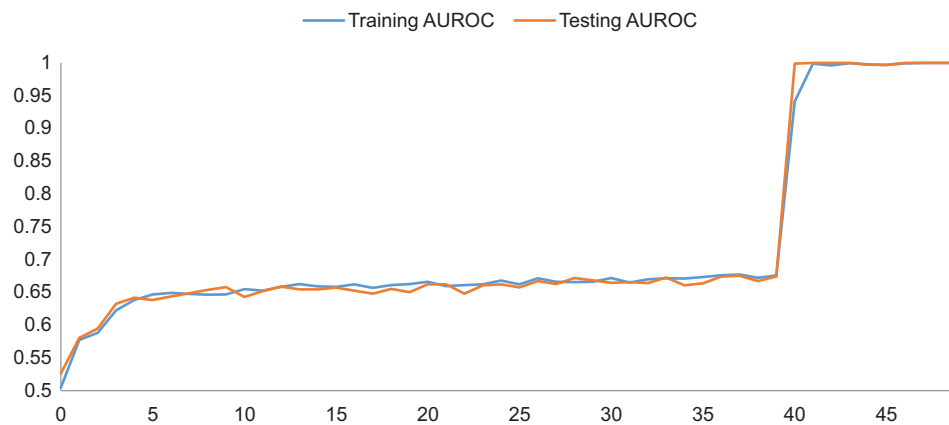


Figure 6: Plots of area under the receiver operating characteristics (AUROC) versus the number of epochs obtained while training a linear classifier on labeled training data.

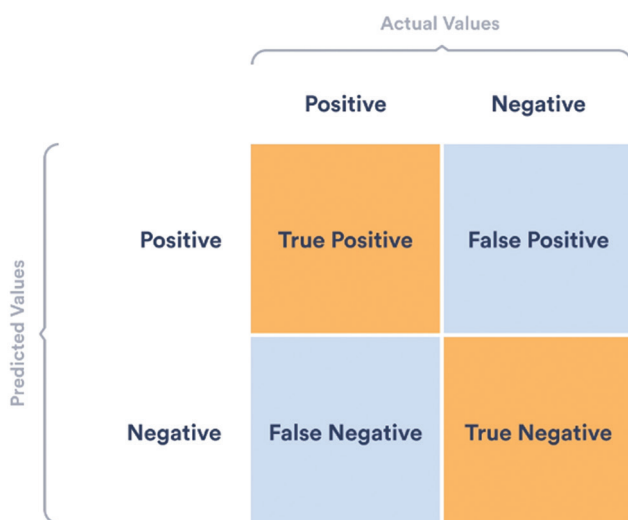


Figure 7: Confusion matrix for bone fracture detection.

Evaluation metrics

The evaluation measures of the models are calculated using accuracy, precision, recall, F1-score, AUC, and ROC [Figure 7]. These metrics are evaluated based on the confusion matrix of each model.

These metrics comprehensively assess the model’s performance across key fracture detection accuracy and robustness aspects. The confusion matrix is a tool that provides a detailed overview of the performance of a classifier model. Accuracy can be misleading, especially for imbalanced datasets. The figure illustrates the different types of predictions (true positive, true negative, false positive, and false negative) in the context of bone fracture detection.

Accuracy is the percentage of correct predictions made by a model, calculated by dividing the number of correct predictions by the total number of predictions.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

TP: True positive, TN: True negative, FP: False positive, FN: False negative.

Recall, also known as sensitivity, is the percentage of positive cases correctly identified by a model, calculated by dividing the number of true positives by the total number of positive cases.

$$Recall = \frac{TP}{(TP + FN)}$$

The F1 score is a metric that combines precision and recall into a single measure of model performance. It is calculated by taking the harmonic mean of precision and recall, which gives equal weight to both metrics.

$$F1 - Score = \frac{(2 \times (Precision \times Recall))}{(Precision + Recall)}$$

F1 - Score: Metric used to evaluate performance of a model.

Hyper parameters

- Learning rate: 0.001–0.01
- Optimizer: Stochastic gradient descent
- Loss function: NT-Xent loss
- Similarity: Cosine similarity function
- Number of epochs: 50
- Batch size: 32
- Model architecture: ResNet-18

In our experiments, the proposed algorithm consistently achieves noteworthy results: an accuracy of 94.10%, a specificity of 93.21%, and a recall of 95.03%. This surpasses the performance benchmarks of conventional fracture detection methods.

RESULTS

The model was able to detect fracture and non-fracture images with 94.10% accuracy, 93.21% specificity, and an AUROC of 94.12%.

DISCUSSION

The model achieved promising results, demonstrating its potential as a decision support tool for radiologists in detecting wrist fractures.

The self-supervised approach offers a valuable alternative to traditional supervised learning, especially when labeled data is limited.

Observations

- Self-supervised learning is a promising approach for fracture detection. Self-supervised learning methods have shown promise in detecting fractures in various medical images, including radiographs.
- Self-supervised learning methods are more scalable and cost-effective than traditional methods. Self-supervised learning methods do not require labeled data, which can be expensive and time-consuming to collect.
- Self-supervised learning methods can be used to automate fracture detection. This could free up radiologists to focus on more complex cases and help to reduce the time it takes to diagnose fractures.
- Self-supervised learning methods can be used to improve the accuracy of fracture detection.
- Self-supervised learning methods can learn representations of medical images that are more robust to variations in pose and contrast than traditional methods, leading to earlier diagnosis and treatment of fractures and improving patient outcomes.

However, challenges such as dataset variability, generalizability to different imaging modalities, and interpretability of learned features still need to be addressed. These studies suggest that self-supervised learning is a promising approach for fracture detection. However, more research is needed to improve the accuracy and robustness of self-supervised learning methods for fracture detection.

Limitations

- Limited data understanding: Self-supervised models learn from unlabeled data, which may lack the specific context or medical expertise needed to identify and classify fractures accurately. This can lead to misinterpretations and missed diagnoses.
- False positive/negative rates: Without the guidance of labeled data, self-supervised models can struggle to differentiate between fractures, leading to both false positives (misdiagnosing fractures) and false negatives (missing actual fractures).
- Generalization issues: Models trained on unlabeled data may not generalize well to unseen scenarios or variations in fracture presentation.

The research provides recommendations for future improvements and potential research directions.

CONCLUSION

Applying self-supervised learning techniques in fracture detection within the medical domain has emerged as a

promising approach. The utilization of contrastive learning, predictive coding, and transfer learning strategies has shown significant improvements in accuracy, reducing the reliance on large labeled datasets. By leveraging unlabeled data, self-supervised approaches have the potential to overcome the limitations of traditional supervised methods, especially in cases where labeled data is scarce or expensive to obtain.

The findings of this research pave the way for the broader adoption of self-supervised learning in the field of fracture detection, enhancing accuracy with minimal reliance on labeled data.

Future direction

Further research and exploration of self-supervised learning methods hold great potential for advancing fracture detection systems, enhancing diagnostic accuracy, and improving patient outcomes. Future research should focus on investigating different pretext tasks and architectural designs that best suit the task of fracture detection. In addition, transfer learning techniques that leverage self-supervised features for downstream tasks could further enhance the clinical applicability of self-supervised learning in fracture detection.

AUTHORS' CONTRIBUTIONS

SRT conceptualized the idea, studied the literature review, proposed methodology, trained the model, and drafted the first manuscript. DGJ provided critical supervision, helped prepare the manuscript, and reviewed the final draft. AKS did data analysis and performed a quality assessment. DVK edited the manuscript and critical review during the writing of the manuscript. All authors participated in the writing of and made essential contributions to the content and similarity index.

ETHICAL APPROVAL

The authors confirm that the ethical policies of the journal, as noted on the journal's author guidelines page, have been adhered to. No ethical approval was required as data was used from public sources. This article contains no studies with human participants performed by any authors.

DECLARATION OF PATIENT CONSENT

Patient's consent is not required, as there are no patients in this study.

USE OF ARTIFICIAL INTELLIGENCE (AI)-ASSISTED TECHNOLOGY FOR MANUSCRIPT PREPARATION

The author(s) confirms that there was no use of Artificial Intelligence (AI)-Assisted Technology for assisting in the

writing or editing of the manuscript and no images were manipulated using the AI.

CONFLICTS OF INTEREST

There are no conflicting relationships or activities.

FINANCIAL SUPPORT AND SPONSORSHIP

This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Klestil T, Röder C, Stotter C. Impact of timing of surgery in elderly hip fracture patients: A systematic review and meta-analysis. *Sci Rep* 2018;18:13933.
2. Mollura D, Culp M, Pollack ER. Artificial intelligence in low- and middle-income countries: Innovating global health radiology. *Radiology* 2020;297:513-20.
3. McDonald RJ, Schwartz KM, Eckel LJ, Diehn FE, Hunt CH, Bartholmai BJ, *et al.* The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. *Acad Radiol* 2015;22:1191-8.
4. Hendriksen JM, Koster-van Ree M, Morgenstern MJ, Oudega R, Schutgens RE, Moons KG, *et al.* Clinical characteristics associated with diagnostic delay of pulmonary embolism in primary care: A retrospective observational study. *BMJ Open* 2017;7:e012789.
5. Allen B Jr., Seltzer SE, Langlotz CP, Dreyer KP, Summers RM, Petrick N, *et al.* A road map for translational research on artificial intelligence in medical imaging: From the 2018 National Institutes of Health/RSNA/ACR/the academy workshop. *J Am Coll Radiol* 2019;16:1179-89.
6. Derevianko A, Pizzoli SF, Pesapane F, Rotili A, Monzani D, Grasso R, *et al.* The use of artificial intelligence (AI) in the radiology field: What is the state of doctor-patient communication in cancer diagnosis? *Cancers (Basel)* 2023;15:470.
7. Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ, *et al.* Deep learning in medical imaging. *Neurospine* 2019;16:657-68.
8. Huang SC, Pareek A, Jensen M, Lungren MP, Yeung S, Chaudhari AS. Self-supervised learning for medical image classification: A systematic review and implementation guidelines. *NPJ Digit Med* 2023;6:74.
9. Litjens G, Kooi T, Bejnordi BE, Setio A, Ciompi F, Ghafoorian M. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60-88.
10. Hardalac F, Uysal F, Peker O, Ciceklidag M, Tolunay T, Tokgoz N, *et al.* Fracture detection in wrist X-ray images using deep learning-based object detection models. *Sensors (Basel)* 2022;22:1285.
11. Yao L, Guan X, Song X, Tan Y, Wang C, Jin C, *et al.* Rib fracture detection system based on deep learning. *Sci Rep* 2021;11:23513.
12. Zech JR, Carotenuto G, Igbinoza Z, Tran CV, Insley E, Baccarella A, *et al.* Detecting pediatric wrist fractures using deep-learning-based object detection. *Pediatr Radiol* 2023;53:1125-34.

13. Gao Yan, Soh TY, Liu N, Lim G, Ting D, Cheng TL, *et al.* Application of a deep learning algorithm in the detection of hip fractures. *iScience* 2023;29:5469-77.
14. Liu X, Zhang F, Hou Z, Zhang J. Self-supervised learning: Generative or contrastive. In: *IEEE Transactions on Knowledge and Data Engineering*. Vol. 35. 2023. p. 857-76.
15. Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. *JMLR* 2020;149:1597-607.
16. Hinterwimmer F, Consalvo S, Neumann J, Micheler C, Wilhelm N, Lang J, *et al.* From self-supervised learning to transfer learning with musculoskeletal radiographs. *Curr Direct Biomed Eng* 2022;8:9-12.
17. Cho K, Kim KD, Nam Y, Jeong J, Kim J, Choi C, *et al.* CheSS: Chest X-ray pre-trained model via self-supervised contrastive learning. *J Digit Imaging* 2023;36:902-10.
18. Wang T, Philip I. Understanding contrastive representation learning through alignment and uniformity on the hypersphere. In: *ICML'20: Proceedings of the 37th International Conference on Machine Learning*. Vol. 37. 2020. p. 9929-39.
19. Doersch C, Gupta A, Efros AA. Unsupervised visual representation learning by context prediction, 2015 IEEE International Conference on Computer Vision (ICCV). 2015. p. 1422-30.
20. Kwasigroch A, Grochowski M, Mikołajczyk A. Self-supervised learning to increase the performance of skin lesion classification. *Electronics* 2020;9:1930.
21. Liu J, Cao L, Akin O, Tian Y. Robust and accurate pulmonary nodule detection with self-supervised feature learning on domain adaptation. *Front Radiol* 2022;2:1041518.