UNDERGRADUATE RESEARCH IN NATURAL AND CLINICAL SCIENCE AND TECHNOLOGY (URNCST) JOURNAL
Read more URNCST Journal articles and submit your own today at: https://www.urncst.com

REVIEW OPEN ACCESS

Low-Cost Fine-Tuning of Data-Efficient Image Transformers on Knee X-Ray Imaging for Osteoarthritis Detection

Shamil Canbolat, BEng Student [1]*, Kevin Sogoli, BSc Student [2]

[1] Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada L8S 4L7
[2] School of Interdisciplinary Science, McMaster University, Hamilton, Ontario, Canada L8S 4L7

*Corresponding Author: canbols@mcmaster.ca

Abstract
Introduction: Technological advancements in artificial intelligence within the field of medicine, specifically skeletal pathology, have witnessed exponential growth in recent years. Researchers have trained deep learning models on radiographs to improve the detection of diseases. There is precedence for low data training on musculoskeletal imaging tasks, and "specialized" medical imaging-related tasks in general have exhibited high performance on low amounts of data. Thus, the Data-efficient Image Transformer (DeiT) model has potential to surpass conventional convolutional models in detecting musculoskeletal diseases due to its ability to extract relevant features with scarce data.
Methods: This study utilizes a DeiT model pre-trained on ImageNet 2012. The model was fine-tuned on labelled knee X-rays of patients with and without osteoarthritis using the PyTorch library. Fine-tuning and testing were done on a Google Colab notebook using a T4 GPU. A hyperparameter sweep cycling through different dropout values, optimizers, and image input sizes was tested. Results were recorded using multiple accuracy metrics.
Results: The DeiT-B 384 model with unspecified dropout had the highest scores out of all the model variations. The DeiT's composite performance in diagnosing knee osteoarthritis exceeded that of convolutional models. Fine-tuning resulted in accuracies within the standard established by current literature for the 384 model, but not for the 224 model. This suggests that larger input models trained on lower quality images performed better than smaller input models trained on higher quality images.
Implications: Fine-tuning can be an alternative to training from scratch for medical imaging. Future studies should expand their scope by taking additional patient details into account to increase diagnostic accuracy and provide local and individualized patient care. The absence of these variables in our model's training potentially limited its accuracy. Current real-time diagnostic errors are less than even the best performing computer vision models, so accuracy must significantly improve before there is incentive to adopt these methods.

Keywords: DeiT; data-efficient; transformer; ViT: fine-tuning; osteoarthritis; knee; x-ray

Introduction

Technological advancements in artificial intelligence (AI) within the field of medicine have witnessed exponential growth in recent years. Specifically in skeletal pathology, researchers have employed deep learning (DL) models to improve the detection of bone and joint diseases using imaging data [1, 2]. This new approach is especially useful when using imaging to detect musculoskeletal diseases. In osteoporosis diagnosis, gathering imaging data like radiographs may be less costly and more accessible than the principal technique of dual-energy X-ray absorptiometry [3]. Likewise, for osteoarthritis, X-ray imaging is the primary diagnostic technique. This method is prone to error as it relies on human interpretation of subtle and complex image features [4]. Osteoarthritis is a joint disease caused by the degeneration of cartilage, leading to pain and reduced mobility [5]. Given that osteoarthritis and many other musculoskeletal diseases can be remedied through early detection [6, 7], increasing checkup ease and availability through AI automation is an important area of research.

An early example of DL that healthcare workers have successfully employed for diagnosis is Convolutional Neural Networks (CNNs). CNNs were pivotal in computer vision for being able to learn patterns in images autonomously instead of relying on predefined features. Their deep feed-forward architecture can learn intricate and abstract details, making them well-suited for interpreting medical imaging for disease detection [8]. Multiple studies using CNNs for osteoporosis diagnosis recorded high sensitivity and specificity [9]. The vast majority of studies on the use of AI in knee osteoarthritis classification use a CNN coupled with X-ray imaging data [6].

While the transformer model [10] brought forth significant advancements in machine learning, it had few applications in computer vision due to the limitations of quadratic attention. The subsequent introduction of the Vision Transformer (ViT) revolutionized image processing, circumventing quadratic attention by splitting the image into patches. Unlike CNNs, ViTs have minimal inductive biases and their class tokens serve as superior global feature extractors compared to global average pooling. ViTs beat previous models on multiple benchmarks with fewer computational resources [11]. When evaluated on diagnosing osteoporosis in knee X-rays, ViTs outperformed CNNs on multiple metrics including accuracy and F1-Score [12].

The superiority of the transformer model for this task is due to the attention mechanism emphasizing image regions that are relevant for classification. This is seen with the heat maps generated by ViTs, which were more accurate in identifying areas associated with osteoporosis compared to CNNs [12]. Despite this, the characteristics that make ViTs excel also make them extremely data demanding. ViTs were only able to outperform leading CNNs at around 100 million training images. By using knowledge distillation and learning to match the output of a "teacher" CNN model, the Data-efficient Image Transformer (DeiT) achieved comparable results to the ViT while only training on ImageNet-1k (around 1.2 million images) [13].

Despite this innovation, medical imaging related tasks still face low data availability due to the cost of scans or the niche nature of specific medical fields. However, this issue can be remedied by replacing the classification head of a DL model that has already been trained on a large general dataset in order to optimize it towards a more specific medical task. This process, known as fine-tuning, significantly reduces the required training data and time.

The applications of fine-tuning in creating a unified standard for visual representation accuracy are explored by Zhai et al. in their paper on the Visual Task Adaptation Benchmark (VTAB) [14]. The prevailing ResNet50-v2 model, pre-trained on ImageNet and fine-tuned for each of the two included medical imaging-related (i.e. "specialized") tasks, achieved Top-1 accuracies of 78.0 and 87.3. Specialized tasks have the smallest disparity in performance when trained on all samples versus just 1000 samples [14]. Furthermore, there is precedence for training on a thousand samples or less; one literature review found that studies using AI for diagnosing osteoporosis utilized an average of 726 images [9].

Considering that this study works with data scarcity and that the minimal inductive bias of image transformers facilitates transfer learning, DeiTs may exhibit the best general classification ability so far. We hypothesize that the DeiT's composite performance in diagnosing knee osteoarthritis, measured in accuracy, will exceed that of convolutional models tested on the same dataset. We further hypothesize that its performance will be comparable to the accuracy of other DeiT model variants on ImageNet.

Methods

We utilized the Knee Osteoarthritis (KOA) Severity Grading Dataset [15] for this study. All images underwent a histogram equalization process to increase contrast. The formula for this equalization [16] was written as a Python script. In addition, randomly color-inverted images were reverted back using Paint.NET. The Random and Shutil Python libraries handled data splitting, while the OpenCV and NumPy libraries handled graphical preprocessing like equalization and resizing (see Figure 1). The scripts for these operations are available on GitHub [17].

The original dataset was split into five folders from "0" to "4" in order of increasing severity, where the first two groups are non-osteoarthritic and the next three groups are osteoarthritic. This was made into a binary split by randomly selecting 500 images for the non-osteoarthritic folder (250 from original "0" and "1" folders each) and 498 images for the osteoarthritic folder (166 from each of the remaining folders) in the training set. In the same fashion, 100 images were selected for the non-osteoarthritic folder and 99 images for the osteoarthritic folder in the test set. We named the non-osteoarthritic folders "0" and the osteoarthritic folders "1". Due to the abundance of training data, there were no concerns regarding the use of a conventional training and test split.

Among the original DeiT variants, DeiT-B (base) distilled models were selected due to the size of the base model and the distillation feature that seem to significantly and independently improve performance [13]. While the larger, 384x384 input image size models also had better performance, our dataset only came in either 224x224 or 299x299 pixel size, and so the images would require resizing, thus decreasing their quality and potentially leading to worse performance. We decided to use both input image sizes with their respective models to test out which of these factors would influence performance the most.

DOI Link: https://doi.org/10.26685/urncst.789

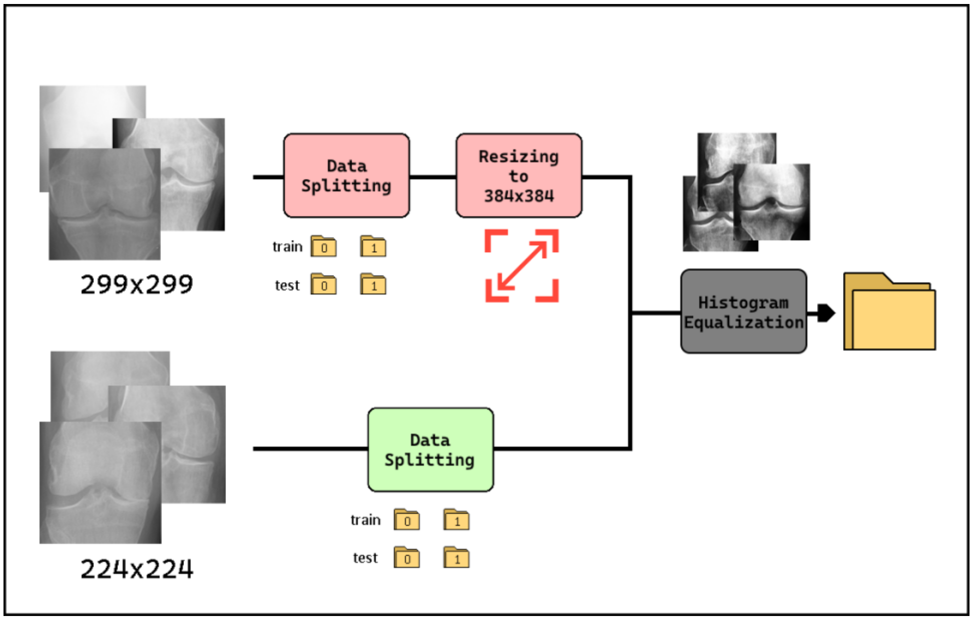

Figure 1. Data Splitting and Preprocessing Methods. Both sized sets equalized and split into "train" and "test" folders with "0" and "1" subfolders. 229x229 resized to match model input size at 384x384. Figure produced using Paint.NET.

Fine-tuning and testing were done on a Google Colab notebook [18] with a T4 GPU using the PyTorch Python library. To avoid overfitting on training data for a dataset of this size, we fine-tuned the model with 32 training steps, a batch size of 32, and a learning rate of 0.0001. More training steps would not have improved accuracy as training loss plateaued with these parameters (see Figure 2). At first, we faced the issue of class overfitting as the model only predicted one category or the other. We fixed this by setting label smoothing to 0.1.

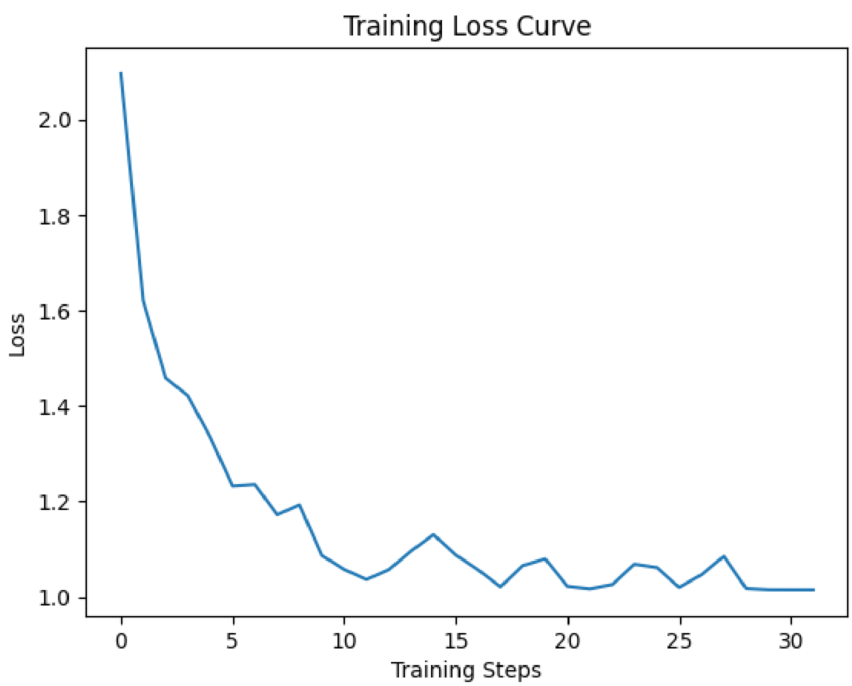

Figure 2. Training Loss Curve for a Training Run on the Best Performing Parameter Suite. Graph displays a training run for the DeiT-B distilled 384 model with Adam optimizer and no dropout. Figure produced using PyTorch in Google Colab.

Our hyperparameter sweep cycled through 0.2, 0.5, 0.8, and default hidden dropout and attention probability dropout (for default, we did not specify any value). We also cycled between the Stochastic Gradient Descent (SGD) optimizer with 0.9 momentum and the Adam optimizer. Finally, we tried each of these parameter combinations out on both the DeiT-B distilled 384 [19] and DeiT-B distilled 224 [20] variants and compared their performance.

For each hyperparameter combination, we fine-tuned the model three times and tested each iteration, measuring accuracy and precision, recall, and F1 scores. Afterwards, we took the average of each combination's three results for each metric and selected the combination with the highest accuracy as the best for each of the two DeiT variants. The Top-1 score is compared to accuracy, as they are equivalent in binary classification tasks. Recall scores were later omitted as they were found to consistently match the accuracy. These scores were compared to CNN models trained on the same KOA dataset derived from the paper by Mohammad et al. [16].

Results

We compiled the results of our sweeps in Tables 1 and 2. DeiT models that trained with the Adam optimizer outperformed DeiT models that trained on the SGD optimizer, regardless of their input image size (224 or 384 pixels). We set different dropout parameters to better explore differences among the DeiT models (Tables 1 and 2). From these, we observed that the DeiT-B distilled 224 model with 0.8 dropout had the highest scores from among the same model with different dropout values (see Table 2). However, the DeiT-B distilled 384 model with unspecified dropout had the highest scores out of all the model variations. We then plotted our top-performing 224 and 384 models against other DeiT models to examine whether our fine-tuning stayed on par with current models (see Figure 3). To observe how this model compared to other DL models fine-tuned on the same data, we plotted a bar graph that compared the values of the top performers from the DeiT-B distilled 224 and 384 models to CNN models trained on the same dataset (see Figure 4).

Table 1. Comparison Between DeiT-B Distilled 384 Model Trained on Adam Vs SGD Optimizer With Variable Dropouts

Optimizer	**Dropout**	**Testing Accuracy**	**Precision**	**F1-Score**
*Adam	unspecified	0.8425	0.8459	0.8421
Adam	0.2	0.8141	0.8162	0.8137
Adam	0.5	0.8191	0.8276	0.8178
Adam	0.8	0.8258	0.8283	0.8254
SGD	unspecified	0.8308	0.8314	0.8307
SGD	0.2	0.8174	0.8192	0.8171
SGD	0.5	0.8174	0.8177	0.8174
SGD	0.8	0.8225	0.8227	0.8224

All values are averaged from three sweeps. All iterations were trained on equal parameters except for indicated changes in dropout. SGD momentum was 0.9. All values are in percentages. * = top performer in all values.

Table 2. Comparison Between DeiT-B Distilled 224 Model Trained on Adam Vs SGD Optimizer With Variable Dropouts

Optimizer	**Dropout**	**Testing Accuracy**	**Precision**	**F1-Score**
Adam	unspecified	0.7772	0.7786	0.7769
Adam	0.2	0.7739	0.7748	0.7737
Adam	0.5	0.7789	0.7803	0.7786
*Adam	0.8	0.7806	0.7816	0.7803
SGD	unspecified	0.7169	0.7213	0.7155
SGD	0.2	0.7471	0.7495	0.7463
SGD	0.5	0.7353	0.7385	0.7344
SGD	0.8	0.7102	0.7103	0.7102

All values are averaged from three sweeps. All iterations were trained on equal parameters except for indicated changes in dropout. SGD momentum was 0.9. All values are in percentages. * = top performer in all values.

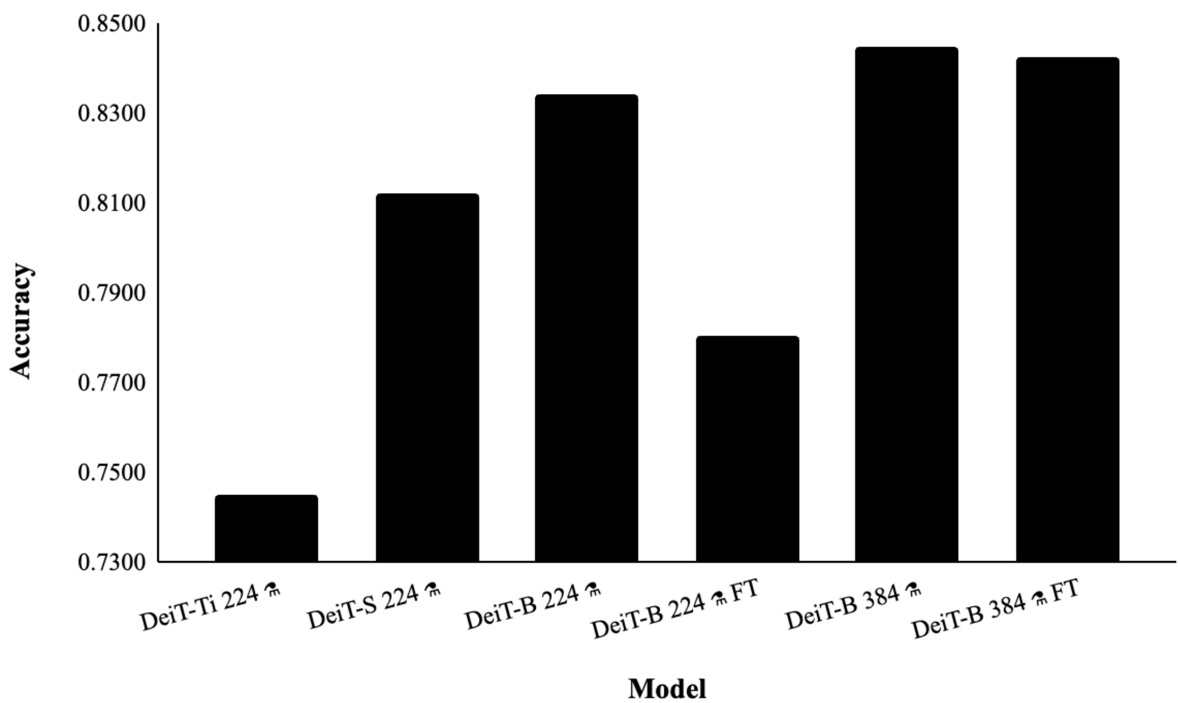

Figure 3. Comparison Between DeiT Models. The symbol ⚗ indicates that the model is distilled. FT means that the model was fine-tuned on the KOA dataset [16]. The data for DeiT-Ti 224 ⚗, DeiT-S 224 ⚗, DeiT-B 224 ⚗, and DeiT-B 384 ⚗ were taken from a study done by Touvron et al. [13]. The DeiT-B 224 ⚗ FT model is the model trained on the Adam optimizer with 0.8 dropout. The DeiT-B 384 ⚗ FT model is the model trained on the Adam optimizer with unspecified dropout. The Y-axis was zoomed in from 0.73 to 0.85 to better display discrete differences between the models. Results for the fine-tuned models are averaged from 3 sweeps. Figure produced using Microsoft Excel.

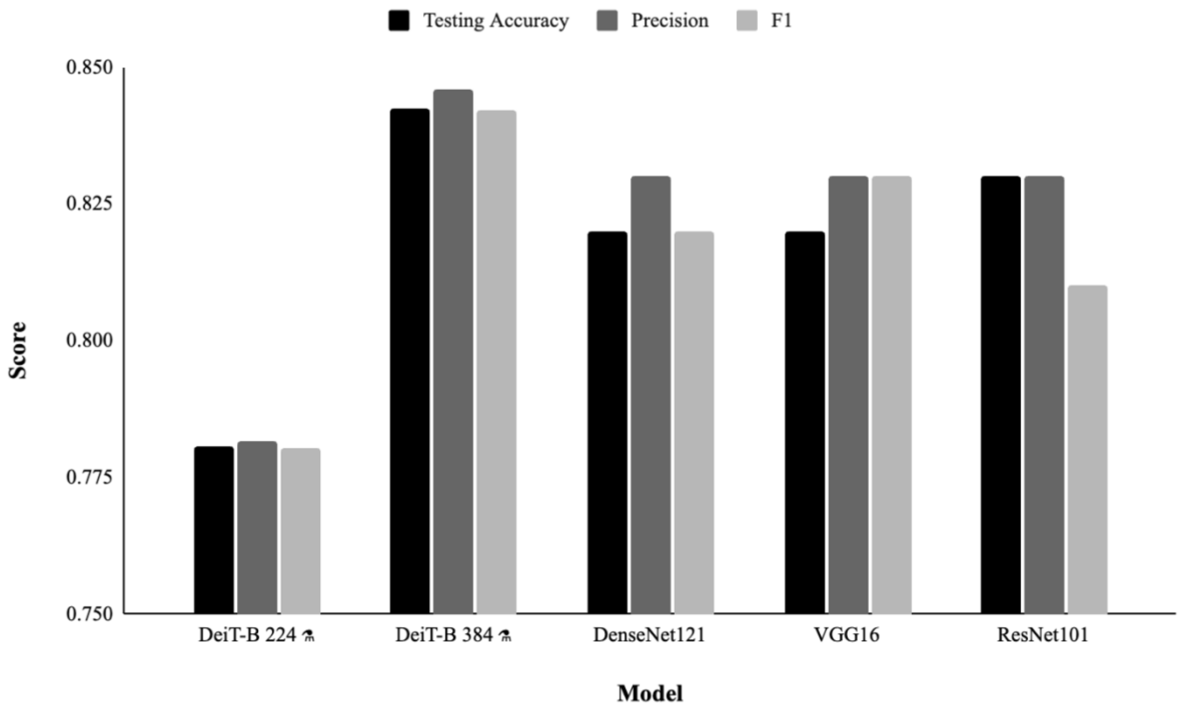

Figure 4. Comparison Between DeiT Models and CNN Models Trained on the Same Dataset. DenseNet121, VGG16, and ResNet101 results were taken from a study done by Mohammad et al. [16]. The symbol ⚗ indicates that the model is distilled. The DeiT-B 224 ⚗ model was trained with the Adam optimizer with 0.8 dropout. The DeiT-B 384 ⚗ model was trained with the Adam optimizer with unspecified dropout. The Y-axis was zoomed in to 0.75 to 0.85 to better display discrete differences between the models. Results for the DeiT models are averaged from 3 sweeps. Since recall values were omitted for the DeiT models, they were also omitted for the CNN models. Figure produced using Microsoft Excel.

Discussion

We confirmed our hypothesis that the DeiT's composite performance in diagnosing knee osteoarthritis exceeded that of convolutional models on the same data. We also found that our fine-tuning resulted in accuracies within the standard established by current literature for the 384 model, while the 224 performed below the current standard (see Figure 1). While exploring this phenomenon, we discovered that among the variable DeiT models, the 384 models did better in general compared to the 224 models, showing that the bigger DeiT with lower-quality images was capable of performing better than the smaller DeiT with higher-quality images. This could indicate that the general structure of the knee joint is more indicative of osteoarthritis than intricate details like osteophytes or subchondral cysts [21], which rely on image quality for visibility.

Our results exceed what previous literature has achieved because although other DeiT models have gotten Top-1 accuracies in the high 80s [13], their results are not applicable to skeletal disease diagnoses as they were trained to classify general datasets like ImageNet. Other DL models, like ViTs and CNNs, scored Top-1 accuracies ranging merely within the low 60s when trained to diagnose osteoporosis [12]. In contrast, our model achieved accuracies within the 80s showing DeiT's potential to overcome traditional models within the field of medical image diagnosis.

As an aside, we were intrigued by the optimizers' performances. While the SGD optimizer often generalizes better, there are some reasons the Adam optimizer may have performed better. A smaller dataset may have more variable gradient updates. The Adam optimizer's gradient smoothing mechanisms, in addition to its insensitivity to less than ideal parameter choices, of which we may have selected some, may have allowed it to better address these situations. We hope future research can further explore this.

Our study explores the gap in literature for vision transformers in medicine. Advancements in this subfield can help healthcare providers avoid missing diagnostically relevant points of interest on X-rays during human validation [4, 21], potentially saving patients from serious injury that may require high-cost treatments later on. Although our study showed promising results, there are a few limitations. It is known that information beyond physical bone appearance is crucial in diagnosing musculoskeletal diseases [22]. The KOA dataset takes important risk factors into account by outlining general inclusion criteria that accept all ethnic minorities (with a focus on African Americans) and men and women between the ages of 45 and 79 [15]. Future studies could aim to build upon these criteria and consider other factors that are just as important in osteoarthritis diagnosis, such as weight [23], diet [24], smoking [25], alcohol consumption [26], diabetes [27], hypothyroidism [28], occupation [30], and family history of bone diseases [31]. Our model did not explicitly take these factors into account.

Conclusions

Our findings suggest that fine-tuning can be an alternative to training from scratch for tasks where the latter is not viable, such as in medical imaging. The number of training images we used is reasonable for local practitioners to acquire either independently or through sharing with other hospitals. Future studies should expand their scope beyond simply analyzing images by taking other variables into account to increase diagnostic accuracy and early detection. The absence of additional variables in our model's training potentially limited its accuracy. Auxiliary patient data could be used in addition to custom training data in order to provide local and individualized patient care, addressing socioeconomic and regional gaps in healthcare. However, until accuracies reach higher levels, healthcare providers have little incentive to implement such methods. Current real-time diagnostic errors from imaging range from between 3-5% [32], while even the best performing computer vision models have error rates in the double digits. Should the model reach higher accuracy, it could provide probability scores and visual heatmaps highlighting areas of concern that radiologists may miss due to human error, thus assisting them in diagnosis. Beyond diagnosis, the model could help track the effectiveness of interventions like physical therapy, medications, or non-invasive treatments by assessing changes in knee X-rays over time.

For future research to be effective in improving diagnosis rates, we suggest that more high-quality, uniform, accessible medical imaging datasets be produced. Common issues with datasets include inconsistency, class imbalance, low resolution or repeating images, bad folder organization, random color inversion, artifacts, long and difficult download processes, and lack of response from corresponding authors. Moreover, current literature does not target challenges specific to medical imaging, such as handling imbalanced datasets or capturing subtle and localized features (e.g., microfractures, bone density variations), all of which significantly affect the model's performance in this field [13].

Generally, future studies on the usage of vision transformers in medical imaging should use heat maps to identify relevant parts of the image that the attention mechanism focuses on. This could later be used along with image annotation tools like Labelme to guide the model and improve accuracy. More specifically, for data-efficient models, we suggest that a study be conducted similar to VTAB wherein these models are evaluated on a multitude of solely medical image-related tasks. This study could rank the performance of newer DeiT editions that have since been released such as DeiT III and CaiT.

List of Abbreviations

AI: artificial intelligence
CNN: convolutional neural network
DeiT: data-efficient image transformer
DeiT-B: DeiT base
DL: deep learning
KOA: knee osteoarthritis
SGD: stochastic gradient descent
ViT: vision transformer
VTAB: visual task adaptation benchmark

Conflicts of Interest

The authors declare that they have no conflict of interests.

Ethics Approval and/or Participant Consent

Consent from study participants was gained by those who made the original datasets.

Authors' Contributions

SC: Drafted the manuscript, contributed to much of introduction, methods and conclusion, wrote code for preprocessing and fine-tuning, managed repository, created figure 1, and gave final approval of the version to be published.
KS: Drafted the manuscript, contributed to much of results and discussion, collected the bulk of the data, produced all other figures and tables, and gave final approval of the version to be published.

Acknowledgements

This journal article was written under the supervision of Amel Sassi of the University of Toronto department of Biomedical Engineering. We thank her for her general guidance, writing assistance, and technical support. We also thank an anonymous Google employee and an anonymous forum member for helping us troubleshoot our code.

Funding

This study was not funded.

References

[1] Mishra S, Dash A, Jena L. Use of deep learning for disease detection and diagnosis. Bio-inspired Neurocomputing. 2020 Jul 22;181–201. https://doi.org/10.1007/978-981-15-5495-7_10

[2] Khalifa M, Albadawy M. Artificial Intelligence for clinical prediction: Exploring key domains and essential functions. Computer Methods and Programs in Biomedicine Update. 2024;5:100148. https://doi.org/10.1016/j.cmpbup.2024.100148
[3] Jang R, Choi JH, Kim N, Chang JS, Yoon PW, Kim C-H. Prediction of osteoporosis from simple hip radiography using deep learning algorithm. Scientific Reports. 2021 Oct 7;11(1). https://doi.org/10.1038/s41598-021-99549-6
[4] Zhang L, Wen X, Li J, Xu J, Yang X, Li M. Diagnostic error and bias in the department of radiology: A pictorial essay. Insights into Imaging 2023;14. https://doi.org/10.1186/s13244-023-01521-7
[5] Hunter DJ, Felson DT. Osteoarthritis. BMJ. 2006 Mar 16;332(7542):639–42. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1403209/
[6] Ahmed HA, Mohammed EA. Using Artificial Intelligence to classify osteoarthritis in the knee joint: Review. NTU Journal of Engineering and Technology. 2022 Sep 9;1(3). https://doi.org/10.56286/ntujet.v1i3.155
[7] Albuquerque GA, Carvalho DDA, Cruz AS, Santos JPQ, Machado GM, Gendriz IS, et al. Osteoporosis screening using machine learning and electromagnetic waves. Scientific Reports. 2023 Aug 8;13(1):12865. https://doi.org/10.1038/s41598-023-40104-w
[8] Indolia S, Goswami AK, Mishra SP, Asopa P. Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. Procedia Computer Science. 2018;132:679–88. https://doi.org/10.1016/j.procs.2018.05.069
[9] Gao L, Jiao T, Feng Q, Wang W. Application of artificial intelligence in diagnosis of osteoporosis using medical images: A systematic review and meta-analysis. Osteoporosis International. 2021 Feb 27; https://doi.org/10.1007/s00198-021-05887-6
[10] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention Is All You Need [Internet]. arXiv.org. 2017. [cited 2024 Dec 16]. Available from: https://arxiv.org/abs/1706.03762
[11] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale [Internet]. arXiv.org. [cited 2024 Dec 16]. Available from: https://www.arxiv.org/abs/2010.11929v2
[12] Sarmadi A, Razavi ZS, van Wijnen AJ, Soltani M. Comparative analysis of vision transformers and convolutional neural networks in osteoporosis detection from X-ray images. Scientific Reports. 2024 Aug 3;14(1). https://doi.org/10.1038/s41598-024-69119-7
[13] Touvron H, Cord M, Douze M, Massa F, Sablayrolles A, Jégou H. Training data-efficient image transformers & distillation through attention [Internet]. arXiv.org. [cited 2024 Dec 16]. Available from: https://arxiv.org/abs/2012.12877v2
[14] Zhai X, Puigcerver J, Kolesnikov A, Ruyssen P, Riquelme C, Lucic M, et al. The Visual Task Adaptation Benchmark [Internet]. arXiv. [cited 2024 Dec 16]. Available from: https://arxiv.org/abs/1910.04867
[15] Chen P. Knee Osteoarthritis Severity Grading Dataset [Internet]. Mendeley Data. [cited 2024 Dec 16]. Available from: https://data.mendeley.com/datasets/56rmx5bjcr/1
[16] Mohammed AS, Hasanaath AA, Latif G, Bashar A. Knee Osteoarthritis Detection and Severity Classification Using Residual Neural Networks on Preprocessed X-ray Images. Diagnostics (Basel). 2023 Apr 10;13(8):1380. https://doi.org/10.3390/diagnostics13081380
[17] Canbolat S, Sogoli K. DeiT-Osteoarthritis Detection [Internet]. GitHub. [cited 2024 Dec 16]. Available from: https://github.com/shamilemir/DeiT-osteoarthritis-detection
[18] Canbolat S, Sogoli K. Google Colab [Internet]. Google. [cited 2024 Dec 26]. Available from: https://colab.research.google.com/drive/1-a5yPyNA2cOykydEG49GiJ2Wr4C56Zvw?usp=sharing
[19] Hugging Face - DeiT-B distilled 384 [Internet]. Facebook. [cited 2024 Dec 26]. Available from: https://huggingface.co/facebook/deit-base-distilled-patch16-384/tree/main
[20] Hugging Face - DeiT-B distilled 224 [Internet]. Facebook. [cited 2024 Dec 26]. Available from: https://huggingface.co/facebook/deit-base-distilled-patch16-224
[21] Lloyd-Jones G. Imaging of Musculoskeletal Disorders. Radiology Masterclass 2020. https://www.radiologymasterclass.co.uk/tutorials/musculoskeletal/imaging-joints-bones/osteoarthritis#top_6th_img
[22] Donald JJ, Barnard SA. Common patterns in 558 diagnostic radiology errors. Journal of Medical Imaging and Radiation Oncology 2012;56:173–8. https://doi.org/10.1111/j.1754-9485.2012.02348.x
[23] Allen KD, Thoma LM, Golightly YM. Epidemiology of osteoarthritis. Osteoarthritis and Cartilage 2021;30:184–95. https://doi.org/10.1016/j.joca.2021.04.020
[24] Bliddal H, Leeds AR, Christensen R. Osteoarthritis, obesity and weight loss: evidence, hypotheses and horizons - a scoping review. Obesity Reviews. 2014;15:578–86. https://doi.org/10.1111/obr.12173

[25] Chopp-Hurley JN, Wiebenga EG, Keller HH, Maly MR. Diet and Nutrition Risk Affect Mobility and General Health in Osteoarthritis: Data from the Canadian Longitudinal Study on Aging. The Journals of Gerontology: Series A 2019;75. https://doi.org/10.1093/gerona/glz277
[26] Felson DT, Zhang Y. Smoking and osteoarthritis: a review of the evidence and its implications. Osteoarthritis and Cartilage 2015;23:331–3. https://doi.org/10.1016/j.joca.2014.11.022
[27] To K, Mak C, Zhang C, Zhou Y, Filbay S, Khan W. The association between alcohol consumption and osteoarthritis: a meta-analysis and meta-regression of observational studies. Rheumatology International 2021;41:1577–91. https://doi.org/10.1007/s00296-021-04844-0
[28] Piva SR, Susko AM, Khoja SS, Josbeno DA, Fitzgerald GK, Toledo FGS. Links Between Osteoarthritis and Diabetes. Clinics in Geriatric Medicine 2015;31:67–87. https://doi.org/10.1016/j.cger.2014.08.019
[29] Chen S, Sun X, Zhou G, Jin J, Li Z. Association between sensitivity to thyroid hormone indices and the risk of osteoarthritis: an NHANES study. European Journal of Medical Research 2022;27. https://doi.org/10.1186/s40001-022-00749-1
[30] Wang X, Perry TA, Arden N, Chen L, Parsons CM, Cooper C, et al. Occupational Risk in Knee Osteoarthritis: A Systematic Review and Meta-Analysis of Observational Studies. Arthritis Care & Research 2020;72:1213–23. https://doi.org/10.1002/acr.24333
[31] Zhai G, Huang J. Genetics of osteoarthritis. Best Practice & Research Clinical Rheumatology 2024;38:101972–2. https://doi.org/10.1016/j.berh.2024.101972
[32] Itri JN, Tappouni RR, McEachern RO, Pesch AJ, Patel SH. Fundamentals of Diagnostic Error in Imaging. RadioGraphics 2018;38:1845–65. https://doi.org/10.1148/rg.2018180021

Article Information
Managing Editor: Jeremy Y. Ng
Peer Reviewers: Amel Sassi, Randa Mudathir
Article Dates: Received Dec 19 24; Accepted Mar 12 25; Published Apr 14 25

Citation
Please cite this article as follows:
Canbolat S, Sogoli K. Low-cost fine-tuning of data-efficient image transformers on knee x-ray imaging for osteoarthritis detection. URNCST Journal. 2025 Apr 14: 9(4). https://urncst.com/index.php/urncst/article/view/789
DOI Link: https://doi.org/10.26685/urncst.789