Convergence
A Journal for Young Researchers

ORIGINAL REVIEW

Received: 2-16-26; Accepted: 7-19-26;
Managing Editor(s): Jonas Katona;
Action Editor(s): Amy Nguyen

Enhancing the Clinical Reliability of AI Models in Medical Imaging through Domain-Specific Pre-Training

Indraneel Mandal

La Martinière for Boys, Kolkata, West Bengal, India

Abstract

Background: Artificial Intelligence (AI) has been integrated into many sectors, leading to improvements in accuracy and efficiency. AI has demonstrated improvements in image classification, object detection, and image segmentation tasks in which they can outperform conventional approaches under controlled conditions. Healthcare is one of the most critical domains because even minor diagnostic errors can significantly affect patient safety and treatment planning. Therefore, reliability is one of the most important factors for any AI system operating in the healthcare sector. Despite this, many AI models, which are primarily pre-trained on natural images, are used in various healthcare applications, which may limit their clinical reliability and generalizability.

Methods: This manuscript presents a qualitative narrative review of existing literature, comparing general-purpose pre-training strategies and domain-specific pre-training strategies across four major imaging modalities: chest X-rays, magnetic resonance imaging, computed tomography, and ultrasound. Relevant studies were identified through searches across major databases, including PubMed, IEEE Xplore and arXiv.

Results: The reviewed studies generally suggest that AI models pre-trained on domain-specific data demonstrate an improved performance, robustness, and clinical applicability across multiple medical imaging modalities when compared with AI models that are pre-trained on natural image datasets.

Conclusion: The reviewed evidence suggests that domain-specific pre-training represents a promising strategy for improving the clinical reliability and generalizability of AI systems used in medical imaging. Future research should focus on multimodal data integration, cross-institutional validation, real-world clinical testing, and systematic comparisons with alternative learning paradigms such as transfer learning and meta-learning.

Keywords: domain-specific pre-training, AI in medical imaging, transfer learning, deep learning in radiology, modality-aware learning, chest X-ray (CXR), CT imaging (computed tomography), magnetic resonance imaging (MRI), ultrasound

1. Introduction

Medical imaging is vital for modern clinical diagnosis, allowing for non-disruptive imaging of anatomical structures, cells and tissues, required for treatment, prognosis, and disease mapping. Imaging modalities such as chest x-ray (CXR), computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound are widely used in the healthcare sector and are a critical component of the diagnostic workflow.

In recent years, AI, particularly deep learning (DL), has been a significant tool for medical image analysis, allowing noticeable improvements in diagnostic accuracy, efficiency, and clinical progress (Litjens et al., 2017). The success of DL in general computer vision, which includes tasks such as image classification, object detection, and object recognition using natural-image datasets, has accelerated the adoption of general-purpose foundational AI models in medical imaging. These AI models are mostly pre-trained on large-scale natural image datasets. Large-scale natural-image datasets include ImageNet, which is the most widely used dataset for pre-training DL models in computer vision. These AI models in medical imaging have shown promising accomplishment across a wide range of general visual tasks. As medical images are modality-specific, they differ greatly from natural images and thus require trained and expert clinical data interpretation. Consequently, AI models trained on non-medical data may not be generalized optimally for medical imaging tasks, potentially limiting their clinical reliability and robustness (Kelly et al., 2019; Roberts et al., 2021). Previous studies suggest that models pretrained on natural-image datasets may not adequately capture clinically relevant anatomical and pathological features, which can reduce performance on certain medical imaging tasks (Tajbakhsh et al., 2016; Litjens et al., 2017).

All imaging modalities face challenges and difficulties and thus demands domain-aware learning. MRI offers superior soft-tissue contrast through multiple imaging sequences, thereby enabling comprehensive assessment of neurological and musculoskeletal abnormalities, but requires an understanding of complex tissue-specific signal intensity variations. CT provides high-resolution volumetric imaging for applications such as oncology and trauma assessment. Meanwhile, ultrasound presents challenges such as speckle noise, operator dependency, and variable image quality. However, general foundational AI models are still used for many of these imaging tasks despite these modality-specific challenges.

The process of training an AI model on a large dataset to learn general feature representations before fine-tuning it on a specific task is referred to as pre-training. In medical imaging, general-purpose pre-training using natural-image datasets such as ImageNet often provides limited benefits due to the fundamental differences between natural and medical images (Raghu et al., 2019). The limited transferability of natural-image pre-training has been associated with differences in feature distributions and optimization behavior between medical imaging datasets and non-medical datasets. In natural-image datasets such as ImageNet, convolutional neural networks (CNNs) primarily learn edges, textures, colors, and object-level semantic features. However, medical imaging tasks oftentimes depend upon subtle grayscale intensity variations, tissue morphology and pathology-specific structures that differ greatly from natural images (Raghu et al., 2019; Tajbakhsh et al., 2016). As a result, pretrained representations derived from natural image datasets may converge towards clinically irrelevant features, thereby resulting in weaker generalization under cross-institutional or modality-shifted conditions. Domain-specific pre-training addresses these limitations by exposing AI models to medically relevant anatomical and pathological patterns during representation learning, thereby improving robustness to modality-specific variability. This review examines how different pre-training strategies influence the performance, robustness, and clinical reliability of AI models across various medical imaging modalities.

This paper aims to evaluate how domain-specific pre-training influences the clinical reliability, robustness and generalizability of AI models across four major medical imaging modalities: CXR, CT, MRI, and ultrasound. Furthermore, it critically compares domain-specific and general-purpose pre-training strategies, reviews the mechanisms underlying their performance differences and hence evaluates their implications for the development of clinically reliable AI systems in medical imaging.

2. Model architectures and the pre-training mechanisms in medical imaging

2.1. Comparative Framework and Feasibility Index

In the context of this narrative review, mathematical or architectural derivation in detail is beyond the scope. The key concept is to properly understand how different model architectures interact with pre-training strategies and how this interaction influences the performance across medical imaging tasks. As far as capturing spatial hierarchies, contextual information, and modality-specific patterns, these architectures differ.

Figure 1 illustrates a clear conceptual distinction between general-purpose and domain-specific pre-training strategies. While the general-purpose models transfer representations learned from the natural image datasets, domain-specific approaches learn modality relevant features (anatomical and pathological) directly from the medical imaging data, which improves both robustness and clinical reliability to a great extent.

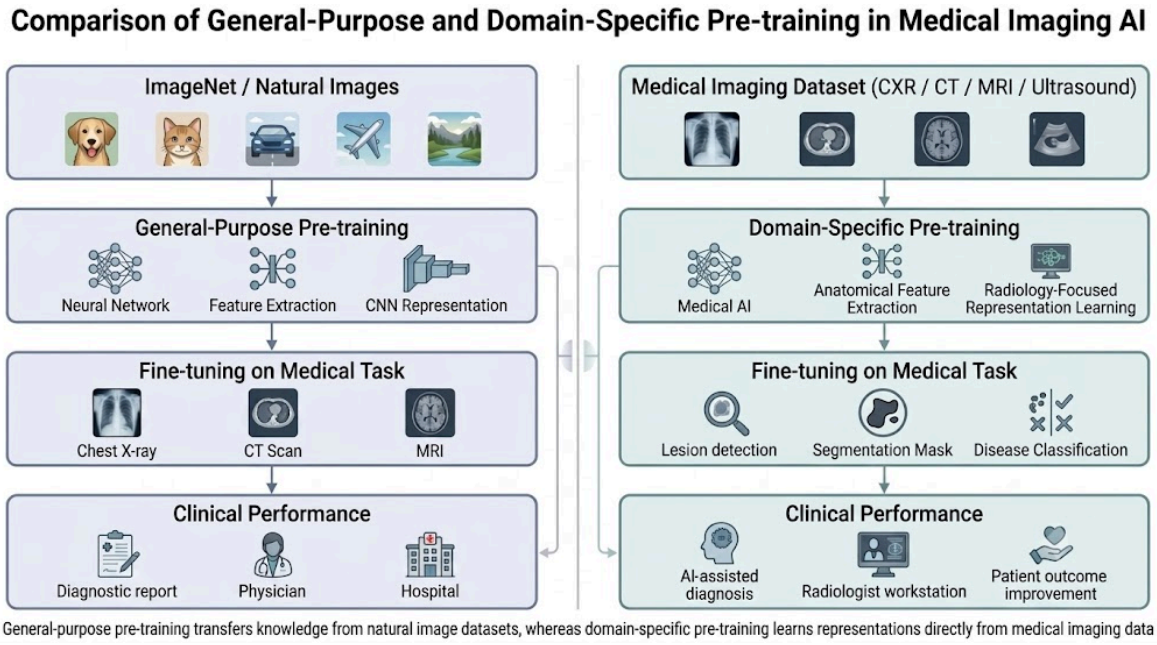

Figure 1: Conceptual comparison of general-purpose and domain-specific pre-training pipelines in medical imaging.

General-purpose pre-training transfers representations which are learned from the natural image datasets whereas the domain-specific pre-training learns representations directly from the medical imaging data before task-specific fine-tuning.

Encoder-decoder architectures such as U-Net have become very important for medical image segmentation, which involves identifying and precisely delineating anatomical structures or pathological regions within medical images. This is because they combine contextual feature extraction with precise spatial localization. As shown in Figure 2, skip connections preserve fine-grained anatomical information while integrating higher-level semantic features, thereby making U-Net highly effective for biomedical segmentation tasks (Ronneberger et al., 2015).

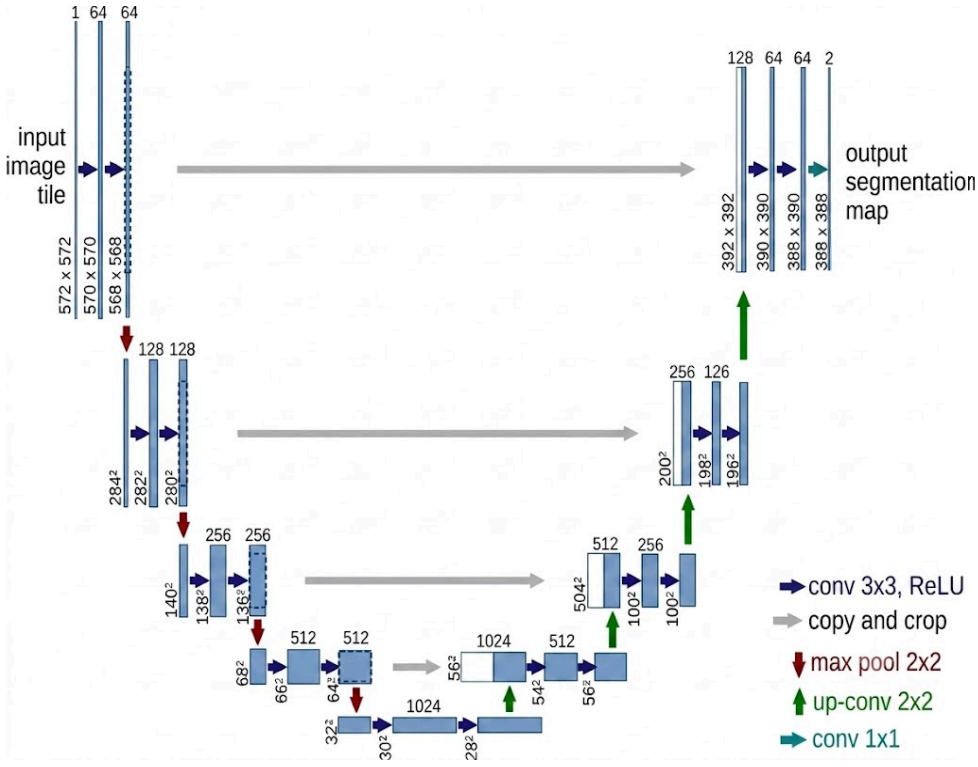

Figure 2: The U-Net architecture, an encoder–decoder network designed for biomedical image segmentation, in which the contracting path captures contextual information and the expanding path enables precise localization through skip connections (Ronneberger et al., 2015).

During the process of pre-training, model optimization typically reduces a classification loss function such as cross-entropy loss through gradient descent-based learning (Goodfellow et al., 2016; LeCun et al., 2015). The learned weights steadily encode hierarchical feature representations, beginning with low-level spatial filters in earlier convolutional layers and slowly moving towards higher level semantic abstractions in the deeper layers (Zeiler & Fergus, 2014; LeCun et al., 2015). For classification tasks, optimization is commonly performed using the categorical cross entropy

loss functions, which measures the differences between the predicted probabilities and the ground truth levels.

In medical imaging, the transferability of these learned representations depends strongly upon the resemblance between the source and target domains (Tajbakhsh et al., 2016; Raghu et al., 2019). When pre-trained on natural image datasets, optimization trajectories may become biased toward non-clinical visual features, limiting the ability of the model to identify pathology-specific structures (Raghu et al., 2019; Tajbakhsh et al., 2016). This representation mismatch becomes especially important in volumetric imaging modalities such as CT and MRI, where the contextual continuity across slices and subtle tissue-intensity relationships are clinically meaningful. Domain-specific pre-training somewhat addresses this issue by disclosing the optimization process to the medically relevant spatial and anatomical patterns during representation learning (Azizi et al., 2021; Raghu et al., 2019).

Consequently, the latent feature space learned by the model becomes more parallel with the clinically meaningful variations instead of generic object-recognition features. On the other hand, the domain-specific pre-training enables models to learn the feature representations which are directly aligned with medical imaging features. For example, 3D CNNs which are trained on volumetric CT data can properly capture inter-slice spatial continuity that is clinically very important for anatomical interpretation (Çiçek et al., 2016). Similarly, MRI-focused models can learn relationships across multiple imaging sequences, thereby improving tissue characterization (Isensee et al., 2021). Ultrasound-specific models have also been shown to improve robustness to speckle noises and operator-dependent variability when trained on domain relevant datasets (Liu et al., 2019).

The effectiveness of architectures depend strongly upon the alignment between the training data and the target clinical task. In the case where the pre-training data closely matches the modality and task distribution, the models tend to perform well. This allows them to learn meaningful representations. Conversely, when there is a mismatch in the domain, the performance degrades which leads to poor generalization, increased false positives and unreliable clinical predictions.

Therefore, it is the combination of architecture design and domain-specific pre-training that largely determines the success of AI systems in the field of medical imaging rather than the architecture alone (Raghu et al., 2019; Tajbakhsh et al., 2016; Azizi et al., 2021).

The advantages of domain-specific pre-training are also influenced by the dataset size and task complexity. Raghu et al. (2019) showed that ImageNet pre-training may provide modest improvements when medical datasets are small, but the advantage decreases gradually as task-specific medical datasets become larger. This indicates that the transfer learning effectiveness not only depends on architecture selection, but also on the dataset scale, modality complexity, and feature similarity between source and target domains.

3. Methodology

Data sources. This study is a narrative review that synthesizes existing literature on the role of domain-specific pre-training of AI models in medical imaging. Relevant studies were identified through searches conducted on databases including PubMed, IEEE Xplore, and arXiv.

Search strategy. Keywords were used including, “medical imaging”, “transfer learning”, “domain-specific pre-training”, “MRI deep learning”, “CT segmentation”, “ultrasound AI”, and “chest x-ray deep learning”. Widely cited papers were identified

through database citation metrics and the examination of reference lists from highly relevant review articles and foundational studies in medical imaging AI. Full-text versions of potentially relevant studies were manually reviewed to assess the methodological quality, empirical validation, and relevance to the objectives of this review.

Inclusion and exclusion criteria. Peer-reviewed studies and widely cited research papers that provided empirical evaluation of AI models in medical imaging were prioritized. Included studies directly compared domain-specific pre-training with general-purpose pre-training under comparable experimental settings, including image classification, segmentation, and detection tasks. Widely cited papers were selected when they represented foundational contributions to the field or were frequently referenced in subsequent medical imaging literature. Studies that mainly discussed theoretical concepts without empirical validation or did not clearly describe the datasets, evaluation metrics, or model training procedures were excluded from detailed analysis. Studies which compared domain-specific pre-training with general-purpose pre-training were included and discussed on the basis of modality-specific challenges throughout this paper. Methodological clarity was assessed based on the availability of information regarding datasets, model architectures, evaluation protocols, and the reported performance metrics. Studies published between 2015 and 2023 were selected to reflect recent developments in DL.

Study scope and analysis. This review is qualitative in nature and does not perform a formal meta-analysis. Instead, performance metrics reported in the original studies were interpreted within the context of their datasets and experimental conditions. Area Under the Receiver Operating Characteristic Curve (AUROC) was used to assess the discriminative performance in classification tasks because it measures an AI model's ability to distinguish between positive and negative cases across different decision thresholds. However, AUROC alone may not be sufficient for direct cross-study comparisons because datasets, disease prevalence, class distributions, and evaluation protocols can differ substantially between studies. The Dice Similarity Coefficient (DSC) was used to evaluate overlap accuracy in segmentation tasks. Additional metrics such as sensitivity and specificity were considered where reported. Performance comparisons across studies were interpreted cautiously because differences in datasets, class distributions, disease prevalence, evaluation protocols, and experimental settings can substantially influence reported performance metrics, including AUROC and DSC. MRI sequence terminology is also discussed throughout this review, where T1-weighted images primarily provide anatomical detail, T2-weighted images highlight fluid-containing structures and pathology and the contrast-enhanced T1-weighted (T1ce) images improve visualization of lesions and tumor boundaries (Bitar et al., 2006).

4. Results

The findings across the reviewed studies generally indicated that domain-specific pre-training was associated with improved performance, robustness, and clinical applicability of the AI models across different imaging modalities. Table 1 summarizes the key imaging modalities, associated AI tasks, and their corresponding challenges. Tables 2–5 further compare representative studies across CXR, CT, MRI, and ultrasound imaging modalities, highlighting the differences in robustness, sensitivity, and generalizability between domain-specific and general-purpose pre-trained AI models.

Table 1: Overview of various imaging modalities

Modality	Key AI tasks	Example domain-specific AI models	Challenges addressed
CXR	Pneumonia detection, multi-disease classification	CheXNet	Overlapping anatomical structures, subtle radiographic findings
CT	Lesion detection, volumetric segmentation	3D CNN, 3D U-Net	Volumetric context, scanner variability
MRI	Tumor segmentation, tumor classification	nnU-Net (self-configuring U-Net framework)	Multi-sequence data, tissue-contrast variability
Ultrasound	Organ segmentation, echocardiography analysis	Ultrasound-specific CNNs	Speckle noise, operator dependency

Abbreviations: AI, artificial intelligence; CXR, chest x-ray; CT, computed tomography; 3D, 3-dimensional; CNN, convolutional neural network; MRI, magnetic resonance imaging; nnU-Net, no-new-Net.

5. Hardware workings of the various medical imaging modalities and their limitations

The CXR imaging is primarily based on the photons passing through the tissues of various densities present (Bushberg et al., 2012; Herring, 2020). While this method is fast and cost-effective, there are some limitations, which include the X-rays having many 2D projections with multiple overlapping anatomical structures, which limit the depth perception and also the sensitivity factors to subtle pathologies such as early COVID-19 infiltrates and other kinds of severe lung diseases (Jacobi et al., 2020; Herring, 2020).

CT is widely used currently in the healthcare industry. The CT scans reconstruct multiple X-ray projections into volumetric 3-dimensional (3D) representations using various advanced tomographic reconstruction algorithms and therefore it offers superior resolution and also shows a proper and a better density variation (Bushberg et al., 2012; Kalender, 2011), but it also

has a major limitation that is it exposes the patient to a higher amount of radiation doses which is extremely harmful (Brenner & Hall, 2007).

MRI relies completely on the nuclear magnetic resonance principles to capture the tissue-specific contrast through variations in the relaxation timings (T1, T2) and the pulse sequences (Brown & Semelka, 2010; Westbrook et al., 2018). MRI also provides an exceptionally soft tissue contrast but on the other hand it suffers from extremely high cost, long scan times and also scanner-dependent variability (Brown & Semelka, 2010; Westbrook et al., 2018).

Ultrasound imaging, uses high frequency sound waves to generate real-time images (Hoskins et al., 2010). The advantage of this is that it is safe, portable and cheaper compared to the other imaging modalities but it is highly sensitive to speckle noises and operator dependency (Hoskins et al., 2010; Noble & Boukerroui, 2006).

These modality-specific hardware characteristics directly influence the feature distributions available to the AI models, thereby affecting the effectiveness of different pre-training strategies and contributing to variations in model generalizability across the imaging modalities.

6. Chest X-Ray Imaging

One of the most widely used diagnostic modalities commonly used in clinical practices is the CXR. This is generally used for the detection and also monitoring of pulmonary and cardiovascular diseases. Since the cost is low, the outcome is prompt, and it is also broadly available, chest radiology is considered to be the primary and the first imaging tool for detecting conditions such as pneumonia, tuberculosis, lung cancer, COVID-19, pleural effusion, and heart failure.

CXRs present multiple analytical challenges for both automated systems and AI models because the images are two-dimensional (2D), with overlapping anatomical structures and subtle intensity variations that often require expert interpretation. Consequently, CXR analysis remains a complex task, and several studies suggest that domain-specific pre-training enables AI models to learn radiographic features that are more relevant for clinical interpretation than representations learned from natural-image datasets (Raghu et al., 2019; Litjens et al., 2017).

One of the most prominent examples of such training in the case of CXR imaging is CheXNet, proposed by Rajpurkar et. al. (2017). Beyond pneumonia, various recent studies and literature have demonstrated the effectiveness of the usage of domain-specific AI models in detecting COVID-19 related things, tuberculosis cavitation and lung nodules as well, which are often subtle and easily missed by the general-purpose AI models. Since it was trained on large-scale expertly labelled X-ray datasets, CheXNet achieved state-of-the-art performance on multiple thoracic diseases, including pneumonia, with AUROC scores typically in the low-to-mid 0.80s on the ChestX-ray14 test set, comparable to radiologist-level performance (Rajpurkar et al., 2017).

It was also found that CheXNet illustrated improved diagnostic performances in several reviewed parameters as compared to the general CNNs. On the other hand, the performances of general-purpose CNNs pre-trained on natural images were stated to show limited generalizability in certain clinical settings (Zech et al., 2018). Zech et al. (2018) also showed that the AI models that were trained on CXRs from a single institution were unable to generalize to external datasets. Especially when pretrained on non-medical images.

These findings suggest that the performance gaps between domain-specific and general-purpose pre-training are strongly associated with feature transferability and the dataset biases. These differences are likely related to the limited transferability of features learned from natural-image datasets to chest radiography, particularly when disease manifestations are subtle (Zech et al., 2018; Roberts et al., 2021). In chest radiography, models sometimes rely on false correlations such as hospital specific markers, scanner artefacts, or even acquisition patterns instead of the pathology-specific anatomical features (Zech et al., 2018; Roberts et al., 2021). This issue becomes very important especially during external validation, where the models trained under narrow institutional distributions may experience substantial drops in performances.

Table 2: Chest X-Ray AI Model Comparison

Model	Pre-training	Task	Performance	Notes
CheXNet	CXR dataset	Pneumonia classification	AUROC ≈ 0.82–0.85	Evaluated on ChestX-ray14 dataset (Rajpurkar et al., 2017)
CNN	ImageNet	Pneumonia classification	Performance varies across datasets	Reduced external generalizability reported (Zech et al., 2018)
CheXNet	CXR dataset	Multi-disease detection	Strong performance across multiple thoracic diseases	Supported by multiple studies

Note: Domain-specific models such as CheXNet demonstrate improved robustness and generalizability compared to general-purpose CNNs, particularly under cross-institutional evaluation settings.

Abbreviations: CXR, chest x-ray; AUROC, Area Under the Receiver Operating Characteristic Curve; CNN, convolutional neural network

Collectively, these findings suggest that domain-specific pre-training may contribute to an improved diagnostic performance, robustness, and generalizability in CXR analysis when compared to general-purpose pre-training methods.

However, the magnitude of these benefits may vary greatly depending on the dataset characteristics, model architecture, and evaluation settings. There are more recent studies that comparatively evaluated and found that the models pretrained on large-scale CXR datasets most of the time performed much better when compared to the general-purpose vision models of comparable sizes in both the classification and segmentation tasks (Irvin et al., 2019).

Domain-specific pre-training may also improve model interpretability by encouraging AI systems to focus on clinically relevant anatomical regions rather than non-diagnostic image artefacts. Several explainability studies using gradient-based localization techniques have shown that domain-trained AI models more consistently highlight pulmonary opacities, nodules, and other diagnostically relevant thoracic regions (Rajpurkar et al., 2017; Roberts et al., 2021).

These findings indicate that both the foundational and the more recent studies consistently report potential benefits of using domain-specific pre-training for CXR analysis. Taken together, the evidence which is available suggests that the modality aware pre-training may play a crucial role in improving the clinical reliability of AI systems used in chest radiography. The importance of domain-specific pre-training is further supported by the availability of large-scale medical imaging datasets. For example, the CheXpert dataset contains 224,316 chest radiographs from 65,240 patients, providing substantial modality-specific data for representation learning and model development (Irvin et al., 2019). The DL models which were trained on CXR images specifically achieved a higher multi-pathology classification performance when compared to the general foundation AI models.

Table 2 summarizes representative comparisons between CXR specific models and ImageNet-pretrained CNNs across pneumonia classification and multi-disease detection tasks.

7. Computed Tomography Imaging

CT imaging plays a vital role in day to day clinical workflows, especially in oncology and trauma assessment tests (Kalender, 2011; Bushberg et al., 2012). CT scan imaging is also widely used in the detection of pulmonary embolism, strokes, intracranial hemorrhages, and COVID-19 related abnormalities (Rubin et al., 2015; Kwee & Kwee, 2020). CT scans provide a high-resolution 3D volumetric data which properly capture finely grained anatomical structures and details across a wide range of slices (Kalender, 2011). These characteristics demand AI models that are capable of understanding complex 3D spatial relationships, tissue-density variations, and subtle pathological patterns (Çiçek et al., 2016; Litjens et al., 2017).

CT imaging represents one of the most complex modalities for reliable automated medical analysis. General purpose vision models trained on two dimensional (2D) natural images often struggle to interpret volumetric CT data effectively because they are designed primarily for planar image representations rather than volumetric feature learning (Raghu et al., 2019; Tajbakhsh et al., 2016). These models therefore lack intrinsic mechanisms for learning spatial continuity across adjacent slices and may thereby fail to properly capture clinically relevant 3D contextual information (Çiçek et al., 2016; Isensee et al., 2021).

In contrast, the AI models trained on volumetric CT data demonstrate highly improved sensitivity and robustness in the clinical detection and segmentation tasks (Çiçek et al., 2016; Setio et al., 2016). Previous studies demonstrate that CT specific models which are trained on volumetric data improve cancer detection sensitivity and lesion localization to a great extent when compared to 2D vision models pretrained on natural images (Çiçek et al., 2016; Setio et al., 2016). Several studies show high improvements in sensitivity for lesion detection when CT-specific 3D architectures are used in

comparison to 2D vision models (Setio et al., 2016).

CT-specific domain-trained models have been demonstrated to reduce false positives in lesion detection tasks, which is an extremely important and crucial factor in minimizing unnecessary follow-up procedures (Setio et al., 2016). These findings show the disadvantages of the general-purpose foundational AI models in handling volumetric medical data and also highlight the necessity of modality-aware pre-training.

A major difference between CT-specific architectures and conventional 2D CNN transfer learning approaches lies in their representations of spatial context. Standard ImageNet-pretrained CNNs process slices independently and therefore may lose inter-slice anatomical continuity, on the other hand, 3D CNNs and volumetric U-Net variants learn spatial dependencies directly across adjacent slices. These architectural differences become particularly important for detecting diffuse lesions, irregular tumor boundaries and low-contrast abnormalities that require volumetric contextual interpretation (Çiçek et al., 2016).

Table 3: Computed Tomography AI Performance

Model	Pre-training Type	Task	Performance	Notes
3D CNN	CT-specific dataset	Lesion detection	Higher sensitivity reported in controlled studies	Captures volumetric context (Setio et al. (2017))
2D CNN	ImageNet	Lesion detection	Lower sensitivity in volumetric tasks	Limited 3D understanding (Çiçek et al. (2016); Setio et al. (2017))

Note: CT specific AI models such as CNNs achieve a higher performance in volumetric lesion detection tasks because they preserve volumetric special relationships.

Abbreviations: 3D, 3-dimensional; CNN, convolutional neural network; CT, computed tomography; 2D, 2-dimensional.

Another major advantage of CT-specific domain pre-training is in its ability to account for scanner-related variability and reconstruction artefacts. The CT images vary a lot, depending a lot on the scanner manufacturers, the radiation dose protocols and the reconstruction algorithms. The pre-training of AI models with specific domain data enables the AI models to learn invariant features properly and in depth so that they can generalize across such variations, whereas general-purpose vision models often wrongly interpret reconstruction artefacts as pathological features and hence limit

their clinical reliability (Litjens et al., 2017; Roberts et al., 2021).

Domain-specific CT models are not completely without limitations. Although volumetric pre-training improves contextual understanding to a great extent, these models are computationally expensive and require mostly larger annotated datasets compared to the conventional 2D transfer learning approaches. Moreover, cross-institutional variability in CT acquisition protocols may still reduce the external generalizability despite domain-specific training. This suggests that architecture selection itself is not sufficient and must be combined with dataset diversity and thorough external validation.

Overall, the reviewed evidence highlights the limitations of general-purpose vision AI models for CT imaging and supports the use of domain-specific pre-training to improve performance, robustness, and clinical applicability. For complex applications such as cancer detection and trauma diagnosis, domain-specific CT AI models are more appropriate for achieving reliable clinical performance.

Table 3 summarizes representative comparisons between CT-specific 3D architectures and conventional ImageNet-pretrained 2D CNN approaches in volumetric lesion detection tasks.

8. Magnetic Resonance Imaging

MRI has emerged as one of the most widely used imaging tools in modern day medical diagnosis, especially in the fields of neurological imaging, oncology, and musculoskeletal imaging. Unlike other medical imaging tools such as X-ray technology-based medical imaging modalities, MRI image has very high-resolution volumetric data with very subtle contrast in soft tissues that are demonstrated through appropriate image acquisition protocols such as the T1-weighted image, the T2-weighted image, and the contrast-acquired image (T1ce image). This very multi-sequence and high-dimensional nature of the MRI data makes it in particular less acceptable for direct transfer learning from the natural image datasets. The analysis of an MRI image becomes more complex for a general AI system; this is because there is a large variation in the intensities specified to MRI image capturing tools which is again combined with a large amount of anatomical details which were absent in natural images. Due to this, MRI serves as a critical platform for evaluation of the effectiveness of domain-specific pre-training in medical imaging models. The transfer learning from the natural image datasets most of the times provide a slight benefit in capturing the modality-specific intensity variations which are inherent to MRI data (Raghu et al., 2019).

A growing body of the existing literature continuously shows that the pre-trained AI models, those that are pre-trained on specific MRI datasets are much better than the general-purpose foundational AI models of same sizes across a wide range of tasks including tumor detection, classification and segmentation. Raghu et al. (2019) investigated the effectiveness of features learnt from the datasets which contained natural images in medical imaging tasks and demonstrated that the ImageNet pre-training provided very limited benefits for applications which were based on MRI. Their findings showed that when CNN architectures of similar sizes and capacities were compared, AI models pre-trained directly on MRI datasets consistently outperformed ImageNet-pretrained counterparts across medical imaging tasks such as brain tumor detection and tumor segmentation (Raghu et al., 2019). This limitation arises because MRI data contains modality-specific features, which includes tissue contrast mechanisms, multi-sequence information, and complex anatomical structures which are not present in the natural-image datasets. Consequently, representations learned from natural images may not adequately capture clinically relevant MRI characteristics, thereby limiting the effectiveness of direct transfer learning without domain-specific exposure (Raghu et al., 2019; Tajbakhsh et al., 2016; Litjens et al., 2017). These findings regarding the limited transferability of the

natural-image pre-training are further supported by studies which demonstrate that the AI models which are trained directly on MRI data achieve superior tumor segmentation performances across multiple MRI sequences, including T1-weighted, T2-weighted, and contrast-enhanced T1-weighted (T1ce) images, when compared to the AI models initialized using natural-image datasets (Bakas et al., 2018; Isensee et al., 2021).

This advantage in the performance has been continuously observed on public brain tumor segmentation benchmarks such as the Brain Tumor Segmentation (BraTS) Challenge (Bakas et al., 2018). Across BraTS benchmark evaluations, MRI-specific training strategies such as nnU-Net have consistently reported Dice similarity coefficients exceeding 0.80 for several tumor sub-regions, depending on the segmentation task and validation protocol (Isensee et al., 2021). On the contrary, evidence from transfer-learning studies suggests that ImageNet pre-training provides only a limited number of benefits for MRI-based segmentation tasks because the learned representations do not properly capture the MRI-specific tissue characteristics and anatomical structures (Raghu et al., 2019). These findings therefore highlight the limitations of relying solely on natural-image pre-training for MRI segmentation applications.

Cross-study comparisons show that the benefits of MRI-specific pre-training can become more pronounced as task complexity increases. While ImageNet transfer learning may provide moderate initialization advantages for simpler classification tasks with limited datasets, segmentation tasks which involve heterogeneous tumor boundaries and multi-sequence fusion will generally require modality-specific representational learning. This suggests that the effectiveness of transfer learning depends greatly upon the interaction between the task complexity, dataset scale, and modality-specific structural variation (Raghu et al., 2019; Isensee et al., 2021). These differences in performance are particularly significant in clinical matters, where proper and precise tumor boundary demarcation directly influences the treatment planning and the diagnostic outcomes. In addition to these segmentation tasks, the MRI-specific foundational AI models have also demonstrated a much better performance in classification tasks like tumor grading and the detection of diseases. These tasks include stroke outcome prediction as well.

Multiple studies and literature have reported that the MRI-specific representational learning improves the disease classification and also the tumor grading performances when compared to the general-purpose foundational vision AI models, especially when the model sizes are the same (Raghu et al., 2019; Litjens et al., 2017). These improvements show the importance of AI models learning MRI-specific anatomical priors, which the general vision AI models often fail to capture effectively and properly. Other important and crucial factors influencing the MRI performances are scanner variability and acquisition heterogeneity. MRI images vary greatly across the scanner manufacturers, field strengths and patient demographics. Domain-specific pre-training with proper and wide MRI datasets exposes the AI models to the scanner and acquisition variability during learning, which leads to improved robustness and cross-institutional generalization, in contrast, the general-purpose models often fail to adapt reliably and effectively to the unseen MRI distributions (Raghu et al., 2019; Roberts et al., 2021). Such robustness is extremely crucial for the safety and the reliability factor of clinical deployment of the MRI based AI models.

Some studies also indicate that domain-specific pre-training itself does not completely eliminate generalization failures. The differences in MRI acquisition protocols, magnetic field strengths, and institutional imaging practices may still produce significant performance variability during external validation. Therefore, although MRI-specific pre-training improves representation quality, robust clinical deployment additionally requires heterogeneous datasets, standardized evaluation protocols and rigorous external benchmarking as well (Roberts et al., 2021).

Overall, the extensive empirical evidence confirms that the MRI-based tasks substantially benefit from domain-specific pre-training with MRI-specific data when the AI models of the similar sizes are compared. These findings properly highlight that the MRI's modality-specific complexity requires proper and appropriate learning strategies and that the reliable clinical deployment of AI in MRI cannot rely solely on the general-purpose vision pre-training.

Table 4 summarizes representative comparisons between MRI-specific architectures and ImageNet-pretrained CNN approaches across segmentation tasks on BRATS and related benchmarks.

Table 4: Magnetic Resonance Imaging AI Performance

Model	Pre-training Type	Task	DSC / AUROC	Notes
nnU-Net	MRI dataset	Brain tumor segmentation	DSC ≈ 0.80+ to 0.90+	BRaTS benchmark (Isensee et al., 2021)
CNN	ImageNet	Brain tumor segmentation	Typically, lower than MRI specific models	Limited transferability (Raghu et al. (2019))

MRI specific AI models have frequently been reported to achieve a higher performance than ImageNet pre-trained AI models because they learn modality specific tissue contrast and multi-sequence dependencies. Abbreviations: DSC, Dice Similarity Coefficient; AUROC, Area Under the Receiver Operating Characteristic Curve; nnU-Net, no-new-Net; MRI, Magnetic Resonance Imaging; BRaTS, Brain Tumor Segmentation Challenge; CNN, convolutional neural network.

9. Ultrasound Imaging

Ultrasound imaging, due to its non-invasiveness, real-time acquisition and low cost, is being widely adopted across a wide range of clinical domains (e.g., abdominal imaging, obstetrics and cardiology). Other important applications include vascular assessment and echocardiography assessments as well. However, ultrasound imaging carries a number of complex and unique challenges, e.g., low signal-to-noise ratio, speckle noises, operator dependency, and considerable variability in image qualities.

The general foundational vision AI models which are trained on natural image datasets may misinterpret speckle noises and acquisition artefacts as meaningful features, thereby leading to unreliable predictions (Xie et al., 2018; Chen et al., 2019). On the other hand, domain-specific pre-training on large-scale ultrasound data enables AI models to learn robust clinically relevant patterns despite the substantial noise and operator variability (Xie et al., 2018; Chen et al., 2019).

Multiple surveys in the literature have shown that DL models pre-trained on domain-specific ultrasound data consistently yield strong performance compared to general foundational AI models with comparable parameter sizes (Xie et al., 2018;

Chen et al., 2019).

Performance metrics which include accuracy (the proportion of correctly classified cases), F1 score (the harmonic mean of precision and recall) and robustness under noisy imaging conditions generally favor the domain-trained ultrasound AI models. To be specific, under noisy and operator-dependent imaging conditions, AI models pre-trained on domain-specific ultrasound data most of the time outperform general-purpose foundational AI models of comparable size (Xie et al., 2018; Chen et al., 2019).

Table 5: Ultrasound AI Performance

Model	Pre-training Type	Task	Performance	Notes
Ultrasound-specific CNN	Ultrasound dataset	Clinical tasks	Improved robustness under noise conditions	Domain adaptation beneficial (Xie et al. (2018); Chen et al. (2019))
CNN	ImageNet	Clinical tasks	Performance sensitive to noise and variability	Less robust (Xie et al. (2018); Chen et al. (2019))

Note: Studies reviewed in this paper suggest that domain-specific ultrasound AI models provide a greater robustness to noise and operator variability than general-purpose AI models.

Abbreviations: CNN, convolutional neural network.

Unlike CT and MRI, ultrasound imaging introduces a significant operator-dependent variability during image acquisition. Consequently, performance differences between domain-specific pre-training and general-purpose pre-training are not only influenced by representation learning, but also by acquisition inconsistency. Domain-specific ultrasound models are generally considered better at learning invariant features under noisy imaging conditions, while natural-image pretrained models may incorrectly amplify speckle artefacts or acquisition distortions during feature extraction. These differences are very critical when AI is taken and applied under real-world clinical settings, especially when ultrasound images vary significantly across devices and patient populations. This significant performance gap between general-purpose foundational vision AI models and ultrasound-specific AI models accentuates the need for modality-aware training in AI models in the healthcare field for trustworthy, accurate, and reliable results.

The literature surrounding ultrasound AI remains comparatively less standardized than MRI and CT research. Dataset sizes are often smaller, annotation quality varies substantially, and benchmarking protocols are less uniform across the studies. As a result, although current evidence strongly supports the advantages of ultrasound-specific pre-training, the direct comparison across studies remains difficult because of the differences in acquisition settings, devices, and operator expertise.

Table 5 summarizes representative comparisons between ultrasound-specific CNNs and ImageNet-pretrained CNN models across noisy and operator-variable clinical imaging conditions.

10. Discussion

Across the various imaging modalities discussed in this review, the literature generally suggests that the AI models which are pre-trained on domain-specific medical datasets are associated with improved performance, robustness, and clinical applicability when compared to general-purpose foundational AI models (Tajbakhsh et al., 2016; Raghu et al., 2019; Zech et al., 2018; Roberts et al., 2021). Various findings from medical imaging studies suggest that modality specific details and pathologies are often absent in the natural image datasets, which may contribute to reduced performance in certain clinical tasks. Collectively, the reviewed studies indicate that the alignment between the pre-training dataset and the target imaging modality plays a vital role in determining feature transferability and downstream clinical performance (Tajbakhsh et al., 2016; Raghu et al., 2019; Zech et al., 2018; Oakden-Rayner, 2020; Roberts et al., 2021). Sometimes, general vision AI models may also have difficulty adapting to variations in image acquisition conditions, image reconstruction algorithms, and patient demographics as well. These variances are more effectively accommodated by domain-specific pre-training, which may contribute to more reliable clinical predictions.

The reviewed literature also indicates that the magnitude of improvement associated with domain-specific pre-training is not uniform across all imaging modalities. Simpler classification tasks with limited datasets might sometimes still benefit moderately from ImageNet initialization, especially when computational resources or annotated medical data are limited (Tajbakhsh et al., 2016). However, as task complexity increases, specifically in volumetric segmentation, multi-sequence analysis, or noisy imaging environments, the limitations of natural-image transfer learning becomes more apparent. Across the studies included in this review, CT and MRI applications generally reported larger improvements than CXR classification tasks because volumetric imaging and multi-sequence analysis require richer modality-specific feature representations. Similarly, ultrasound imaging benefited from domain-specific pre-training through improved robustness to speckle noise, operator dependency and acquisition variability. These findings suggest that the advantages of domain-specific pre-training become more pronounced as imaging complexity and contextual requirements increase.

The domain-specific image datasets have also shown considerable scope for further improvement. There are widely used image datasets that are not fully expert-verified by radiologists, which may result in residual annotation variability (Rajpurkar et al., 2017). In addition, some image datasets also lack adequate demographic diversity, which may raise concerns regarding fairness and model generalizability (Oakden-Rayner, 2020; Roberts et al., 2021). To overcome these constraints, continued collaboration between clinicians, radiologists, and AI researchers is essential. Domain-specific AI models are not intended to replace doctors or radiologists. Instead, they serve as decision-support tools that can enhance human-AI collaboration. When these AI models are developed using appropriately curated datasets and rigorous validation, they may function as a reliable second reader, helping to reduce radiologists' workload and support clinical

decision-making. Nevertheless, the reviewed literature suggests that robust external validation and diverse, expert-annotated datasets remain essential for achieving clinically reliable AI systems. Though there are substantial advantages of domain-specific pre-training, there are certain limitations as well.

One major challenge is the limited availability of large-scale, high-quality annotated medical datasets, because expert annotation is both time-consuming and also very expensive. Moreover, many existing datasets lack sufficient demographic diversity, which may lead to bias and limit the generalizability of AI models across different populations (Roberts et al., 2021; Oakden-Rayner, 2020). Furthermore, variability in imaging protocols, scanner types, and acquisition settings across institutions can affect model robustness despite domain-specific pre-training. The computational costs related to training large-scale domain-specific models also remain high, thereby posing practical constraints for widespread adoption.

When compared with alternative approaches such as transfer learning and emerging techniques such as meta-learning, domain-specific pre-training often provides better alignment between learned representations and modality-specific clinical features. Transfer learning enables adaptation of previously learned representations to new tasks using limited labelled data, whereas meta-learning aims to improve rapid generalization across multiple tasks and domains (Pan & Yang, 2010; Hospedales et al., 2021). However, several studies suggest that modality-specific pre-training may capture clinically better relevant imaging characteristics when sufficient domain data are available (Raghu et al., 2019; Tajbakhsh et al., 2016). Future research should therefore explore hybrid strategies that combine domain-specific representation learning with adaptable learning frameworks to improve both scalability and generalization (Hospedales et al., 2021).

Recent literature suggests that self-supervised learning may become an important complementary strategy in medical imaging. Unlike conventional supervised pre-training, self-supervised approaches learn representations directly from large volumes of unlabelled medical data through pretext or proxy tasks, thereby reducing dependence on expensive expert annotations (Azizi et al., 2021; Taleb et al., 2020). Several studies have reported that self-supervised domain-specific pre-training can improve robustness, feature transferability, and downstream task performance, particularly in data-limited clinical environments (Azizi et al., 2021; Taleb et al., 2020). However, external validation procedures and benchmarking frameworks remain insufficiently standardized across institutions, and further comparative research is required before these approaches can be integrated reliably into routine clinical workflows (Roberts et al., 2021).

Overall, the findings synthesized in this review suggests that the effectiveness of pre-training strategies depends on the interaction between imaging modality, dataset characteristics, task complexity and evaluation methodology. Rather than relying on a single universal pre-training strategy, future medical imaging systems are likely to benefit from modality-aware learning frameworks supported by diverse datasets, rigorous external validation, and clinically meaningful evaluation protocols

11. Future Directions

Future research in medical imaging should focus not only on improving predictive performance but also on enhancing clinical reliability, external validation, and explainability across multiple healthcare settings. One important direction involves the development of large-scale multi-institutional datasets that encompass demographic diversity, heterogeneous scanner protocols, and geographically distributed patient populations. Such datasets would help reduce

dataset-specific bias while improving model robustness during external validation (Roberts et al., 2021).

Another important research direction involves the development of multimodal learning frameworks that combine information from multiple imaging modalities, electronic health records, pathology reports and other relevant clinical metadata as well. Existing literature suggests that multimodal integration may improve contextual reasoning and reduce reliance on isolated image features, thereby potentially enhancing diagnostic consistency in complex clinical settings (Litjens et al., 2017).

Future work should also prioritize explainability and uncertainty estimation in clinical AI systems. Although high AUROC or DSC values may indicate strong benchmark performance, these metrics alone do not guarantee clinically reliable decision-making. Explainable AI techniques, saliency mapping methods, and uncertainty aware prediction frameworks may assist clinicians in interpreting model outputs more effectively while identifying predictions that require additional clinical review.

Greater systematic comparisons are also needed between domain-specific pre-training, transfer learning, meta-learning, and self-supervised learning under comparable evaluation conditions. Current literature often evaluates these approaches on different datasets and experimental settings, making direct comparison more difficult. Benchmarking frameworks that incorporate common datasets, external validation protocols, and clinically meaningful evaluation metrics would thereby improve reproducibility and facilitate more reliable assessment of these learning strategies.

Finally, future clinical deployment research should prioritize prospective validation in real-world clinical environments rather than relying primarily on retrospective benchmark datasets. Several studies have shown that AI systems performing strongly in controlled experimental settings may sometimes fail under real clinical variability because of distribution shifts, demographic imbalance, or institutional differences (Roberts et al., 2021). Consequently, clinically deployable AI systems are likely to require continuous monitoring, recalibration, and collaborative oversight involving clinicians, radiologists and AI researchers to ensure sustained safety, reliability, and clinical effectiveness.

12. Conclusion

Domain-specific pre-training plays a vital role in improving the performance, robustness, and clinical applicability of AI models across various medical imaging modalities. The effectiveness of pre-training strategies depends on factors including modality characteristics, dataset scale, task complexity and external validation conditions as well. When compared to general-purpose vision AI models pre-trained on natural-image datasets, the domain-specific pre-training often enables the learning of representations that are more closely aligned with clinically relevant anatomical and pathological features.

The findings presented for CXR, CT, MRI and ultrasound applications highlight the potential advantages of modality-aware learning strategies for supporting proper and reliable medical image analysis. However, clinically reliable deployment requires more than just improved model performance alone. Diverse and representative datasets, rigorous external validation, standardized benchmarking frameworks, explainable AI methods and prospective real-world evaluation remain essential for improving safety and generalizability.

Future research should focus on multimodal data integration, cross-institutional validation, improved dataset diversity,

systematic comparisons between domain-specific pre-training and alternative learning paradigms, and the development of clinically deployable AI systems that support effective human-AI collaboration as well. Taken together, these efforts may contribute to the development of more reliable, transparent and clinically useful AI systems for medical imaging.

13. References

1. Litjens, G., Kooi, T., Ehteshami Bejnordi, B., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

2. Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems, 32*. https://proceedings.neurips.cc/paper/2019/hash/eb1e78328c46506b46a4ac4a1e378b91-Abstract.html

3. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C. P., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv*. https://doi.org/10.48550/arXiv.1711.05225

4. Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., & Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLOS Medicine, 15*(11), e1002683. https://doi.org/10.1371/journal.pmed.1002683

5. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R. L., Shpanskaya, K., Seekins, J., Mong, D. A., Halabi, S. S., Sandberg, J. K., Jones, R., Larson, D. B., Langlotz, C. P., Patel, B. N., Lungren, M. P., & Ng, A. Y. (2019). CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of the AAAI Conference on Artificial Intelligence, 33*(1), 590–597. https://doi.org/10.1609/aaai.v33i01.3301590

6. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods, 18*(2), 203–211. https://doi.org/10.1038/s41592-020-01008-z

7. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R. T., Berger, C., Ha, S. M., Rozycki, M., Prastawa, M., Alberts, E., Lipková, J., Freymann, J. B., Kirby, J. S., Wiest, R., & Jambawalikar, S. R. (2018). Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. *arXiv*. https://doi.org/10.48550/arXiv.1811.02629

8. Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging, 35*(5), 1299–1312. https://doi.org/10.1109/TMI.2016.2535302

9. Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P. K., Jacobs, C., van Riel, S. J., Wille, M. M. W., Naqibullah, M., Sánchez, C. I., & van Ginneken, B. (2016). Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks. *IEEE Transactions on Medical Imaging, 35*(5), 1160–1169. https://doi.org/10.1109/TMI.2016.2536809

10. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, & W. Wells (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* (Lecture Notes in Computer Science, Vol. 9901, pp. 424–432). Springer. https://doi.org/10.1007/978-3-319-46723-8_49

11. Liu, S., Wang, Y., Yang, X., Lei, B., Liu, L., Li, S. X., Ni, D., & Wang, T. (2019). Deep learning in medical ultrasound analysis: A review. *Engineering, 5*(2), 261–275. https://doi.org/10.1016/j.eng.2018.11.020

12. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine, 17*(1), Article 195. https://doi.org/10.1186/s12916-019-1426-2

13. Oakden-Rayner, L. (2020). Exploring large-scale public medical image datasets. *Academic Radiology, 27*(1), 106–112. https://doi.org/10.1016/j.acra.2019.10.006

14. Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A. I., Etmann, C., McCague, C., Beer, L., Weir-McCall, J. R., Teng, Z., Gkrania-Klotsas, E., Rudd, J. H. F., Sala, E., & Schönlieb, C.-B. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence, 3*(3), 199–217. https://doi.org/10.1038/s42256-021-00307-0

15. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells III, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Lecture Notes in Computer Science, Vol. 9351, pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28

16. Azizi, S., Mustafa, B., Ryan, F., Beaver, Z., Freyberg, J., Deaton, J., Loh, A., Karthikesalingam, A., Kornblith, S., Chen, T., Natarajan, V., & Norouzi, M. (2021). Big self-supervised models advance medical image classification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 3458–3468). IEEE. https://doi.org/10.1109/ICCV48922.2021.00346

17. Hospedales, T. M., Antoniou, A., Micaelli, P., & Storkey, A. J. (2022). Meta-learning in neural networks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44*(9), 5149–5169. https://doi.org/10.1109/TPAMI.2021.3079209

Acknowledgements

The author of this paper acknowledges the contributions of all the researchers whose work formed the foundation of this review paper and also extends his sincere gratitude to his mentors and peers for their valuable guidance and feedback all throughout this research process.

Author Biography

Indraneel Mandal is a 17-year-old student from India with a very strong interest in artificial intelligence, machine learning, deep learning, cybersecurity, robotics and software development. He is a Certified Full-Stack Developer and has completed an Ethical Hacking and Cybersecurity certification program under the Ministry of Micro, Small and Medium Enterprises (M.S.M.E.), Government of India. He has also completed a work experience program at Amazon Web Services (AWS), where

he served as the Team Lead, strengthening his technical, collaborative, and leadership skills. Indraneel has also actively participated in various other science and technology competitions, thereby earning multiple gold medals and other accolades. He is a regional gold medalist at the World Robot Olympiad (WRO) and received the Best Start-up Idea Award at the WRO Nationals for excellent innovation, marketing, and business presentation skills. He was also a finalist in the International Space Settlement Design Competition (ISSDC) which was held at NASA's Kennedy Space Center, where he collaborated with multiple international teams to develop large-scale engineering solutions for future space settlements which further strengthened his engineering, research, leadership, and thinking skills. Driven by a passion for innovation, Indraneel enjoys conducting research and developing technology-driven solutions to real-world problems. He aspires to work on artificial intelligence and cybersecurity in the future, while also contributing to the development of responsible and impactful technologies across diverse domains.

Mentor Contribution Statement

Indraneel Mandal, author of "Enhancing the Clinical Reliability of AI Models in Medical Imaging through the Use of Domain-Specific Pre-Training across Modalities," independently developed and finalized his paper under the mentorship of **Dr. Siddharth Krishnan**, Assistant Professor in the Department of Computer Science at the University of North Carolina at Charlotte. Throughout the research process, Dr. Krishnan's role involved broadly guiding Indraneel through the principles of scientific research and academic writing while providing mentorship on the organization and development of the manuscript. His guidance included refining the research topic and scope, identifying relevant studies, strengthening critical analysis, and enhancing the clarity and scientific rigor of the writing. The mentoring sessions helped Indraneel independently apply these research and writing principles throughout the development of his paper. In addition, Indraneel received constructive feedback from Dr. Krishnan on the manuscript during the mentoring sessions. Through this mentorship, Indraneel independently conducted the review, synthesized and analyzed the scientific evidence, prepared the manuscript and tables, incorporated reviewers' feedback, and completed the final revisions for publication.