Received: 24 September 2024 | Revised: 20 November 2024 | Accepted: 11 December 2024 | Published online: 30 January 2025

Artificial Intelligence and Applications
2025, Vol. 00(00) 1–10
DOI: 10.47852/bonviewAIA52024425

BON VIEW PUBLISHING

REVIEW

Deep Learning Techniques in DICOM Files Classification: A Systematic Review

Vicent Mabirizi[1,*], Simon Kawuma[2], Deborah Natumanya[3] and William Wasswa[4]

[1]*Department of Information Technology, Kabale University, Uganda*

[2]*Department of Software Engineering, Mbarara University of Science and Technology, Uganda*

[3]*Department of Computer Science, Mbarara University of Science and Technology, Uganda*

[4]*Department of Biomedical Science and Engineering, Mbarara University of Science and Technology, Uganda*

Abstract: The digital imaging and communications in medicine (DICOM) format is a widely adopted standard for storing medical imaging data, integrating both image and metadata critical for clinical diagnostics. However, its complexity poses challenges for deep learning applications, particularly in extracting and processing this dual-layered data. This review analyzes 23 peer-reviewed studies published between 2014 and 2024, sourced from PubMed, Google Scholar, PLOS, Science Direct, and IEEE databases. Guided by Arksey and O'Malley's scoping methodology, the review reveals that existing deep learning techniques typically rely on converting DICOM images into simpler formats like JPEG, TIF, or PNG for classification, a process that often results in metadata loss and reduced classification accuracy. Frameworks such as MONAI, NVIDIA Clare, SimpleITK, and OpenCV facilitate direct DICOM processing but face limitations, including overfitting, challenges with data heterogeneity, and inefficiencies in handling large datasets. This review emphasizes the urgent need for developing a robust convolutional neural network architecture capable of directly processing DICOM data to preserve metadata integrity and enhance predictive performance, paving way for more reliable and scalable medical imaging solutions.

Keywords: DICOM image processing, deep learning in radiology, convolutional neural network, medical imaging frameworks, medical metadata preservation, scalable image analysis models

1. Introduction

Deep learning (DL) has revolutionized the field of artificial intelligence (AI), offering significant advancements in tasks such as image classification, object detection, and natural language processing [1]. In the domain of medical imaging, DL techniques have emerged as a transformative tool, aiding in the diagnostic and analysis of anatomical structure to support clinical decision-making [2–5]. These methods have shown remarkable promises in handling large volumes of medical data with high accuracy and speed, making them particularly valuable in resource-constrained settings where traditional diagnostic tools and skilled professionals are limited [6].

A critical component of medical imaging is the digital imaging and communication in medicine (DICOM) format, the industry standard for storing and transmitting medical images [7]. DICOM files integrate both image data and metadata, providing essential information about patient demographics, imaging settings, and diagnostic context. However, unlike standard image formats such as JPEG, TIF, and PNG, the complexity of DICOM files presents unique challenges for DL applications, particularly in metadata extraction and image classification [8–10].

*__Corresponding author:__ Vicent Mabirizi, Department of Information Technology, Kabale University, Uganda. Email: vmabirizi@kab.ac.ug

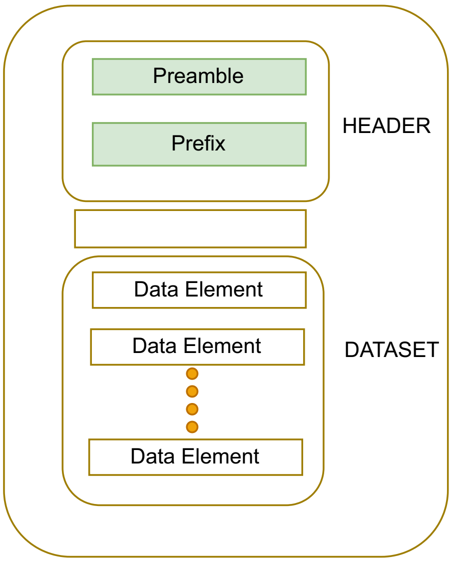
Several foundation DL techniques and frameworks, including CNN, have been adopted to process medical images. CNNs excel in feature extraction and classification but face limitations when applied directly to DICOM files. This is largely due to pre-processing pipelines that convert DICOM files into simpler formats ([11–13]), resulting in metadata loss and reduced model accuracy [14]. Existing frameworks such as MONAI [15], NVIDIA Clare [16], SimpleITK [17], and OpenCV [18] have attempted to address this by enabling direct DICOM processing. However, these tools are often constrained by issues like overfitting to specific datasets [19], inability to handle DICOM format variability, and inefficiencies in processing large datasets, thereby limiting their scalability and reliability in clinical applications [20].

This review examines the current state of DL applications in DICOM file processing, highlighting the challenges, gaps, and limitations of existing techniques. Furthermore, it emphasized the need for a DL model specifically designed to handle DICOM files without requiring format conversion. Such a model would preserve metadata integrity, enhance classification accuracy, and overcome the constraints of existing methodologies. By synthesizing findings from 23 peer-reviewed studies, this review aims to provide a comprehensive understanding of the field and propose directions for future research in medical imaging.

2. The DICOM File Structure

The DICOM format was designed by the Electronic Manufacture Association [21] and the American College of Radiology in the PlayStation3.10 (PS3.10) specification for media storage and file format for media interchange [22]. It consists of a header and the image data sets wrapped in a single file, as shown in Figure 1.

Figure 1
DICOM file structure



The header stores the patient's demographic information, image dimensions, color space, matrix size, acquisition parameters for the imaging study [21, 23], and other additional non-intensity information to support image display on the computer. Below the header, there is an attribute (7FE0) that contains the image pixel intensity. Both the header and the 7FEO are stored in a series of 1s and 0s [7]. The header information is encoded together with the actual image to enable the computer to recognize the imaging study and support image display, ensuring proper documentation for medicolegal purposes.

2.1. Reading DICOM file information

Header information: The header information is constant, standardized, and presented using a sequence of tags organized in collections of data elements such as zeros and ones (e.g., 0110). Extracting header information helps to uncover relevant insights about the image and requires third-party software. Due to an increased utilization of DICOM files in medical imaging, several free software packages for extracting header information have been developed, for example DicomWorks as used by Yaneva-Sirakova [24], XnView [25], and ImageJ [26]. With the help of the header information, dataset can be extracted.

Dataset: To extract patient data from a DICOM file, DICOM Unique Identifiers (UIDs) and header information are required. The patient's biodata and imaging study are encoded within the image header. Therefore, this information can be extracted by anyone with access to UIDs. However, this poses a risk to patient privacy especially when the file is shared over the internet.

3. Research Methods

This section outlines the strategies and methods used to identify suitable studies for this review. The selected articles were carefully chosen and analyzed to find the recent advancements in the application of DL in DICOM processing.

3.1. Scoping review methodology

In this review, we adopted the scoping methodology proposed by Arksey and O'Malley [27] which consists of six steps, namely (1) formulation of the research question, (2) identifying relevant studies, (3) selecting studies, (4) charting data, (5) collecting, summarizing, and reporting findings, and (6) consulting experts [28]. This methodology was chosen because it provides researchers with a quick overview of the existing literature on the topic [29] and accommodates the inclusion of various study types thereby reducing bias in the study conclusions [30].

3.2. Identification of studies

To obtain studies for inclusion in the review, an automatic search was conducted to retrieve relevant studies published between 2014 and 2024. The search was limited to articles published in the past 10 years to ensure the inclusion of studies that reflect current knowledge, latest advancements, methodologies, and practices in the field of DICOM processing. The primary search returned a total of 287 papers (excluding duplicates); however, only 23 met the inclusion criteria.

The review aimed to explore the current state of DL in DICOM processing and classification, identify gaps within existing DL-based models for DICOM processing, and propose a solution that addresses these limitations.

3.3. Search strategy

The search was conducted in August 2024 on five selected publically available databases: PubMed, Google Scholar, PLOS, Science Direct, and IEEE. These were considered because of their extensive integration and indexing capabilities. To find articles relevant to the study, a total of three search terms were used, namely "Deep Learning AND DICOM", "DICOM Processing", "DICOM Classification OR DICOM processing". Each search term was entered in PubMed, Google Scholar, PLOS, Science Direct, and IEEE databases to find articles suitable for our study. A Boolean search method was employed, using "OR" to include articles related to each concept individually and intersection "AND" to combine key concepts thereby focusing on the primary objective of this review. Additionally, the reference lists of selected studies were reviewed to identify additional studies to include in the review. The retrieved articles from all five databases were imported into the Mendeley referencing management tool (Mendeley Desktop v1.19.8) and were independently assessed for inclusion based on their titles, abstracts, and full texts. Any disagreements regarding the inclusion of studies were resolved through discussion among authors until a consensus was reached.

3.4. Inclusion criteria

Only studies that met the following criteria were included.

1) Explicitly described the ML/DL techniques used to process DICOM files

2) Published between 2014 and 2024
3) Paper available in English language
4) Explicitly utilized dataset in DICOM files
5) Clearly described the outcome of DICOM file processing.

Criteria (1) ensure that only studies applying machine learning and DL in their methods were included. Criteria (2) and (3) were established to ensure the inclusion of only recent studies published in English within the last ten years, excluding any papers published prior to 2013. This approach was adopted to maintain the relevance and currency of the review. Criteria (4) and (5) were considered to ensure that only studies that utilized DICOM files in model training, testing, and validation with clear performance results are considered.

3.5. Exclusion criteria

Any studies that did not meet the inclusion criteria listed in Section 3.4 were excluded from the study. The selected studies were assessed to confirm that they reported primary outcomes. Thus, review papers, position papers, protocols, and formative studies were excluded. Geographical location and study design were never considered as exclusion criteria.

3.6. Data extraction and analysis

The extracted data from the included studies were as follows: the techniques applied, the major objectives, the proposed solutions for processing DICOM files, and the performance outcomes.

3.7. Quality appraisal

To assess the quality and relevance of each selected study, a set of criteria were used, guided by PRISMA [31]. To measure the extent to which a study is appropriate and capable of yielding results suitable for the scope of inquiry, the following questions were used;

1) **QA1:** Y (Yes), the dataset used in the study contained DICOM files, P (Partially) –the dataset used in the study contained DICOM files but format conversion was done, N (No) – the dataset used in the study contained images of other formats.
2) **QA2:** Y (Yes) – the study fully applied DL techniques for image pre-processing and classification, N (No) – the study did not apply DL techniques for image pre-processing and classification.

The grading techniques were established as follows: Yes = 1, Partial = 0.5, and No = 0 or unknown. A study that met all the criteria was given a rating of 1. If a study met some of the criteria, it was given a rating of 0.5, while a study that did not meet any of the criteria was given a rating of 0. All studies with scores between 2 and 4 points, based on the criteria, were considered.

4. Results

4.1. Search results

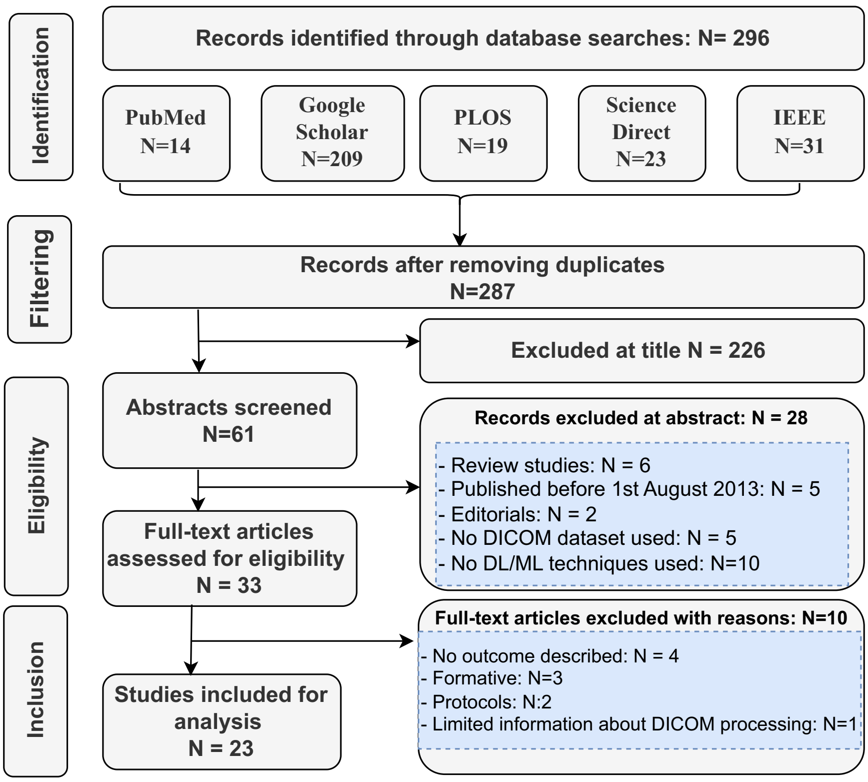
Figure 2 presents the search results, screening, exclusion, and final inclusion in the study.

4.2. Survey of existing literature

Fajar et al. [11] applied histogram equalization and a trainer to reconstruct 3D images from DICOM files of MRI brain images. A visual quality score (mean opinion score (MOS)) between 3 and 5 was obtained for different images.

H. H. Pham et al. [32] presented a method for classifying body parts from DICOM files using CNNs, which achieved 95% precision

Figure 2
Search results



on X-ray images. In their experiment, DICOM files were first converted to PNG format before being subjected to the classification algorithm.

Mamdouh et al. [13] proposed a method for converting 2D DICOM images to 3D images using Seg3D2 and ImageVis3D techniques. Although no quantifiable results are presented in their findings, the authors claim a successful conversion.

Vallez et al. [33] used transfer learning techniques such as AlexNet, VGG, Inception V3, ResNet, and GoogleNet to develop a web viewer system for interpreting WSI DICOM images. The experimental results produced 96% accuracy in interpreting DICOM files.

Kathiravelu et al. [8] proposed a processing pipeline using machine learning techniques to convert DICOM files into standard formats like JPEG and PNG. In their experiment, X-ray images were used, and conversion accuracy of 96% was achieved.

Kawuma et al. [34] proposed a method for diagnosing and classifying TB chest X-ray images for children under 15 years. In their method, transfer learning models – VGG16, VGG19, ResNet50, and Inception V3 – were trained and fine-tuned to form a scratch layer tailored to the local dataset collected from Mbarara Regional Referral Hospital. The model yielded an overall accuracy of 88.23% on the local dataset and 50% accuracy on a multicenter dataset.

Alshmrani et al. [35] applied VGG19 to develop a method for classifying multi-class lung diseases using X-ray images. The experimental results showed successful classification, with 96.48% accuracy, 93.75% recall, and 97.59% precision.

Basaia et al. [36] conducted an automated classification of Alzheimer's disease and cognitive impairment using CNN. The classification was done using MRI images. The method was evaluated, and 75% classification accuracy was achieved using a testing dataset.

Kuraparthi et al. [37] proposed a method for brain tumor classification in MRI images using a deep convolutional neural network. In their experiment, transfer learning models, including AlexNet, VGG16, and ResNet50, were applied. The proposed approach achieved an overall classification accuracy of 97.89%.

Jia and Chen [38] proposed a method for brain tumor identification and classification using a dataset of MRI images. Their approach utilized a support vector machine for classification, focusing on distinguishing tumor types with high precision. The performance of their method was evaluated, achieving an impressive classification accuracy of 98.51%.

Zeimarani et al. [39] developed a method for breast lesion classification using a dataset of ultrasound images. They implemented a custom-built CNN specifically tailored for this task. Their approach was evaluated for its classification performance and achieved an accuracy of 85.98%, demonstrating the potential of CNN models in enhancing the accuracy of breast lesion diagnosis.

R. Kumar et al. [40] proposed a method for classifying carotid artery media thickness using ultrasound images. In their method, a custom-built CNN was developed and tested on ultrasound images. The model demonstrated strong performance, achieving an accuracy of 89.1%, with a specificity of 88% and a sensitivity of 89%.

Liebenlito et al. [41] developed a method to classify tuberculosis (TB) and pneumonia in human lungs using chest X-ray images. They employed a CNN with optimized hyperparameters to enhance classification accuracy. The model's performance was evaluated using area under the curve (AUC), achieving 86% for TB and 96% for pneumonia classification.

Masood et al. [42] proposed a method for recognizing and classifying brain tumors using MRI image datasets. They employed a custom mask region-based CNN with a DenseNet40 backbone architecture. The model achieved a segmentation accuracy of 96.3% and classification accuracy of 98.34%, demonstrating its effectiveness in both segmenting brain tumor regions and classifying tumor types.

Yimer et al. [43] developed a method for classifying multiple lung diseases, including lung cancer, pneumonia, TB, pneumothorax, and chronic obstructive pulmonary disease, using chest X-ray images. In their setup, Xception model, built on a CNN architecture was used. The model attained an overall accuracy of 97.3%, with a sensitivity of 97.2% and a specificity of 99.4%.

Ibrahim et al. [44] proposed a method for pneumonia classification using chest X-ray images. The classification was performed using the AlexNet CNN architecture. The model achieved an accuracy of 93.4%, with a sensitivity of 98.18% and a specificity of 98.18%.

Cao et al. [45] conducted a systematic evaluation of the performance of several state-of-the-art object detection and classification methods for computer-aided diagnosis of breast lesions using ultrasound images. The study compared multiple CNN architectures, including ZFNet, VGG16, AlexNet, GoogleNet, ResNet, and DenseNet. The best-performing models achieved an average precision rate of 96.89%, an average recall rate of 67.23%, and an $F1$ score of 79.38%.

Jabeen et al. [46] developed a method for breast cancer classification using ultrasound images. In their experiment, the DarkNet-53 CNN architecture was used to perform classification task. The model demonstrated outstanding performance, achieving an accuracy of 99.1%.

Yimer et al. [43] developed a DL model named AM_DenseNet for the classification of chest X-ray images, employing the DenseNet121 architecture. The model's performance was evaluated using the AUC, achieving a value of 85.37%.

Mamdouh et al. [47] developed a 3D model for converting DICOM files using the Visualization Toolkit (VTK) library. The visual quality of the resulting 3D model was assessed using the MOS, ranging from 1 to 5. Similarly, Nguyen et al. [48] reconstructed 3D objects from a 2D DICOM dataset using a similar tool, also evaluating the visual quality with the MOS scale. Another study by Van Sinh et al. [49] applied the VTK library to construct models from X-ray CT and MRI images, with the visual quality rated using the same MOS scale. These studies emphasize the utility of the VTK library in generating high-quality 3D visualizations from medical imaging data.

4.3. Summary of reviewed methods

All studies (presented in Table 1) included in this review were either moderate or of high performance with a score ranging between 50% (2) and 95% (3.8).

4.4. Inferences and analysis of reviewed methods

The reviewed methods reveal a wide range of techniques and models applied to various datasets, specifically for disease classification tasks. The methods employed different DL architecture, including VGG variants, AlexNet, ResNet, Inception V3, and other third-party libraries for DICOM file conversion, each selected based on the classification needs and imaging modalities such as MRI, X-ray, and ultrasound. Some of the key inferences drawn from the review include:

1) **Compression and conversion impact on detail preservation:** Converting DICOM image formats, like JPEG and PNG for model compatibility led to a significant metadata loss, impacting the classification performance. Models for tasks like lung and brain tumor classification experienced up to 5%

Table 1
Deep learning methods for classifying DICOM files

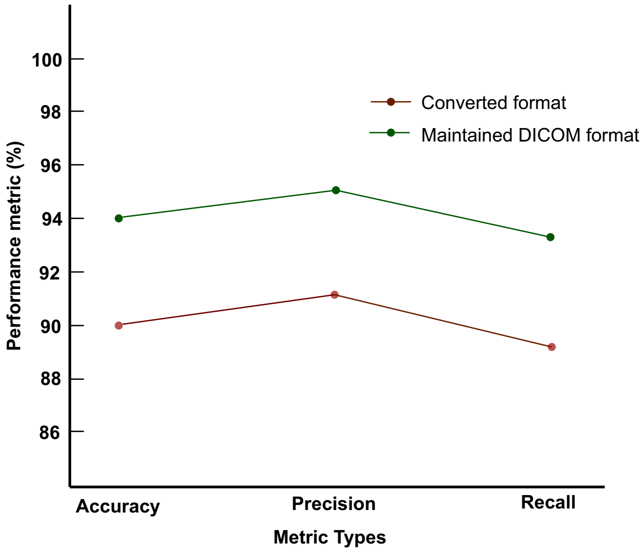
Author	Goal of study	Dataset	Techniques applied	Applied on DICOM or converted image	Performance
Fajar et al. [11]	Reconstructing a 3D image from DICOM file	MRI brain images	Histogram equalization and trainer	DICOM	Visual quality (Mena opinion score: 1 to 5)
H. H. Pham et al. [32]	Classifying body parts from DICOM X-ray scans	16,093 X-ray images	CNNs	Converted to PNG	Precision (95%)
Mamdouh et al. [13]	Converting 2D DICOM images to 3D	412 liver and kidney ultrasound images	Seg3D2 and ImageVis3D	Converted	N/A
Vallez et al. [33]	Web viewer system for interpreting WSI DICOM images	51,000 WSI images	AlexNet, VGG, Inception V3, ResNet, GoogleNet	DICOM	Accuracy (91%)
Kathiravelu et al. [8]	Building a DICOM of enabling ML and processing pipeline in research clusters	989 X-ray images	Machine learning	DICOM	Accuracy (96%)
Hnatkova et al. [12]	Converting DICOM files to 3D image.	CT and MRI images	Mimics software	Converted	Smooth surface accuracy (0.05–0.1 mm)
Kawuma et al. [34]	Diagnosis and classification of TB Chest X-ray images of children less than 15 years.	474 X-ray images	VGG16, VGG19, ResNet50, InceptionV3 and Scratch model	Converted	Precision (50%) Accuracy (50%) Recall (50%)
Alshmrani et al. [35]	Multi-class lung disease classification using CXR images	47089 X-ray images	VGG19	Converted	Accuracy (96.48%) Recall (93.75%) Precision (97.56%)
Basaia et al. [36]	Automated classification of Alzheimer's disease and mild cognitive impairment	1409 MRI images	CNN	Converted	Accuracy (75%)
Kuraparthi et al. [37]	Brain tumor classification of MRI images using deep CNN.	253 MRI images	AlexNet, VGG16, and ResNet50	DICOM	Accuracy (97.87%)
Jia and Chen [38]	Brain tumor identification and classification.	1500 MRI images	Support Vector Machine	DICOM	Accuracy (98.51%)
Zeimarani et al. [39]	Breast lesion classification	641 Ultrasound images	Custom-built CNN	Converted	Accuracy (85.98%)
R. Kumar et al. [40]	Classifying carotid artery media thickness of ultrasound images	501 Ultrasound images	Custom-built CNN	Converted	Accuracy (89.1%) Specificity (88%) Sensitivity (89%)
Liebenlito et al. [41]	Classify tuberculosis and pneumonia in human lungs based on Chest X-ray images	4273 Chest X-ray images	CNN hyperparameter	DICOM	Area under the curve value (86% for TB and 96% for pneumonia)
Masood et al. [42]	Recognizing and classifying brain tumor from MRI images	3317 MRI images	Custom mask region-based CNN with densenet40 backbone architecture	Converted	Accuracy (96.3% for segmentation & 98.34% for classification)

(Continued)

Table 1
(*Continued*)

Author	Goal of study	Dataset	Techniques applied	Applied on DICOM or converted image	Performance
Yimer et al. [43]	Multiple lung diseases classification from Chest X-ray images	12149 Chest X-ray	Xception model of CNN	DIOCM	Accuracy (97.3%) Sensitivity (97.2%) Specificity (99.4%)
Ibrahim et al. [44]	Pneumonia classification from Chest X-ray images	5856 Chest X-ray images	AlexNet	Converted	Accuracy (93.4%) Sensitivity (89.18%) Specificity (98.92%)
Cao et al. [45]	Systematic evaluation of the performance of several existing state-of-the-art object detection and classification methods for breast lesion CAD	579 Ultrasound images	AlexNet, ZFNet, VGG16, GoogleNet, ResNet, DenseNet	Converted	Average precision rate (96.89%) Average recall rate (67.23%) *F*1score (79.38%)
Jabeen et al. [46]	Breast cancer classification from ultrasound images	780 Ultrasound images	DarkNet-53	Converted	Accuracy (99.1%)
Yimer et al. [43]	Developing a deep learning model (AM_DenseNet) for classifying Chest X-ray images	112120 Chest X-ray images	DenseNet121	Converted	Area under the curve (85.37%)
Mamdouh et al. [47]	Building a 3D model for converting DICOM file.	473 DICOM images	Visualization Toolkit(VTK) library	DICOM	Visual quality (Mena opinion score: 1 to 5)
Nguyen et al. [48]	Reconstructing the shape of 3D objects from 2D DICOM datasets	489 DICOM images	Visualization Toolkit(VTK) library	DICOM	Visual quality (Mena opinion score: 1 to 5)
Van Sinh et al. [49]	To construct 3D models from digital and cross-sectional imaging data	X-ray CT and MRI images	Visualization Toolkit(VTK) library	DICOM	Visual quality (Mena opinion score: 1 to 5)

Figure 3
Impact of image format on classification performance metrics



decrease in accuracy due to compression. To establish this variation, we plotted the loss across methods utilizing converted datasets with an average drop of around 4% as illustrated in Figure 3. This highlights the importance of maintaining DICOM format integrity for critical diagnostic tasks.

2) **Variability in model choice based on dataset size:** Complex architectures like Inception V3 and DenseNet perform better with larger datasets, especially for multi-class disease classification. However, custom-built CNNs often suffice for smaller datasets, as shown by models handling fewer than 500 images. From Figure 4, the statistical regression on dataset size versus accuracy indicates a positive correlation, with large datasets (over 10,000 images) yielding up to a 5% improvement in model accuracy when complex architectures are applied.

3) **Model performance across modalities:** The performance of models varies significantly depending on the image modality

Figure 4
Variability in model choice based on dataset size

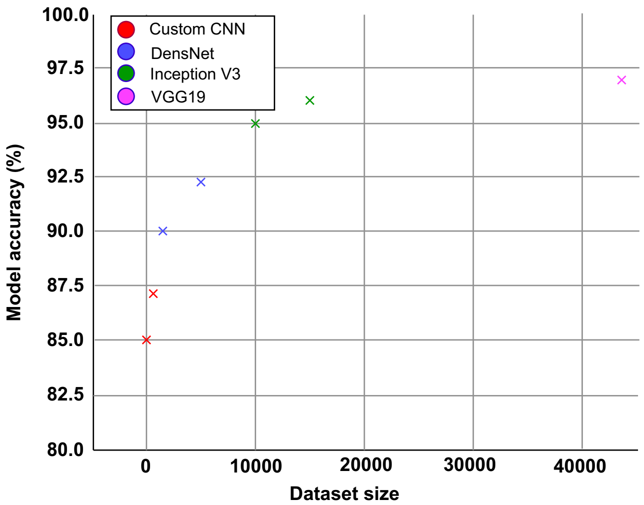
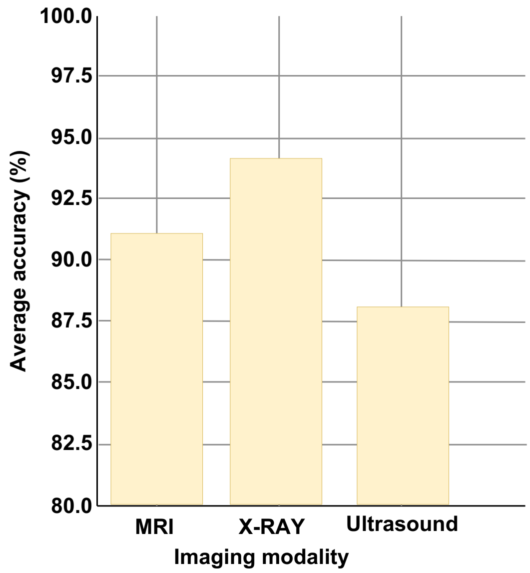


Figure 5
Model performance across modalities



and dataset size. For example, models for MRI images, such as ResNet50, and SVM, achieved high accuracy (up to 98.5%), indicating MRI's higher feature consistency for tasks like brain tumor classification. On the other hand, models for chest X-rays, like VGG19 and DenseNet121, are preferred for lung disease classification due to their robust performance in handling large datasets. To quantitatively establish this comparison, average accuracy and recall across different modalities were computed. For instance, from Figure 5, MRI-based models had an average accuracy of approximately 91% while for X-ray images, it reached 94%, demonstrating MRI's reliability for specific tasks but highlighting X-ray's broader applicability.

5. Discussion

The analysis of existing studies reveals that many approaches rely on a processing pipeline that converts DICOM images into standard formats such as PNG, TIF, and JPEG before subjecting them to classification or segmentation networks. While this conversion facilitates the use CNNs for image classification, it introduces critical challenges. Specifically, this process often results in the loss of essential DICOM metadata [50], including clinical parameters like slice thickness, image resolution, and spacing between slices [14]. The absence of this metadata compromises the accuracy and reliability of diagnostic predictions, as it removes crucial contextual information required for informed clinical decisions. Furthermore, the resizing and compression associated with these standard formats can lead to a loss of image details, negatively impacting the precision of disease classification.

Studies indicate that methodologies that preserve DICOM metadata alongside image data tend to achieve superior precision in diagnostic tasks compared to those that rely on converted image formats like PNG, JPEG, and TIF. For instance, Stiefel et al. [51] demonstrated the significant advantages of maintaining metadata integrity in improving diagnostic accuracy.

In addition, several methods have adopted third-party libraries, such as VTK, and foundation models, such as NVIDIA Clare, to process DICOM images directly, eliminating the need for format conversion. While these libraries are designed for DICOM image handling, they face limitations in specific scenarios. A key issue is their tendency to overfit to specific datasets, where models perform well on training data but fail to generalize effectively to diverse datasets [19]. For example, Kawuma et al. [34] reported that a model achieving 88.23% accuracy on local dataset saw its accuracy drop to 50% when tested on multicenter datasets. This lack of generalization restricts their applicability in real-world clinical settings, where variability in DICOM image composition arises from differences in scanning protocols, imaging conditions, and equipment used [20].

Moreover, foundation models often encounter difficulties in handling large datasets, a critical requirement for machine learning applications. The size and complexity of these datasets frequently hinder effective processing and classification, impacting the scalability of these models in medical imaging contexts.

These observations underscore the need for developing a DL architecture that process DICOM images natively while preserving essential metadata. Such an architecture would ensure the retention of critical clinical information and enhance diagnostic reliability. Furthermore, improving the generalization of these models to accommodate the variability of DICOM datasets is imperative. Addressing overfitting and incorporating robust data handling technique as parallel data processing and advanced data augmentation strategies could improve the challenges associated with large datasets. These innovations would contribute to higher accuracy and scalability, enabling the deployment of DL models in diverse clinical environment.

6. Conclusion

This review has presented a summary of ten years of published studies in application of DL techniques in DICOM file classification. The finding should assist researchers in understandings the benefits and limitations of existing techniques to guide the development of better solutions aimed at improving medical imaging results, specifically those that apply to DICOM images.

Although the current advancements in DL provide valuable tool for medical image analysis, there is significant room for improvement in handling the unique challenges posed by DICOM files. Developing a more robust, sophisticated, and flexible model with the abilities to directly process DICOM images without compromising the image composition will be a crucial step towards advancing the application of AI in medical diagnosis.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors. This review is part of a PhD research project approved by the Mbarara University of Science and Technology Research Ethics Committee (Protocol number: MUST-2024-1656), supervised by Dr. Eng. William Wasswa and Dr. Simon Kawuma.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Vicent Mabirizi: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Simon Kawuma:** Conceptualization, Validation, Writing – review & editing, Supervision. **Deborah Natumanya:** Project administration. **William Wasswa:** Validation, Resources, Writing – review & editing, Supervision.

References

[1] Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, *10*(2), 1–37. https://doi.org/10.2478/jagi-2019-0002

[2] Bhatt, C., Kumar, I., Vijayakumar, V., Singh, K. U., & Kumar, A. (2021). The state of the art of deep learning models in medical science and their challenges. *Multimedia Systems*, *27*(4), 599–613. https://doi.org/10.1007/s00530-020-00694-1

[3] Castiglioni, I., Rundo, L., Codari, M., Di Leo, G., Salvatore, C., Interlenghi, M., . . . , & Sardanelli, F. (2021). AI applications to medical images: From machine learning to deep learning. *Physica Medica*, *83*, 9–24. https://doi.org/10.1016/j.ejmp.2021.02.006

[4] Shamshad, F., Khan, S., Zamir, S. W., Khan, M. H., Hayat, M., Khan, F. S., & Fu, H. (2023). Transformers in medical imaging: A survey. *Medical Image Analysis*, *88*, 102802. https://doi.org/10.1016/j.media.2023.102802

[5] Simon, K., Vicent, M., Addah, K., Bamutura, D., Atwiine, B., Nanjebe, D., & Mukama, A. O. (2023). Comparison of deep learning techniques in detection of sickle cell disease. *Artificial Intelligence and Applications*, *1*(4), 228–235. https://doi.org/10.47852/bonviewAIA3202853

[6] Vicent, M., Simon, K., & Yonasi, S. (2022). An algorithm to detect overlapping red blood cells for sickle cell disease diagnosis. *IET Image Processing*, *16*(6), 1669–1677. https://doi.org/10.1049/ipr2.12439

[7] Varma, D. R. (2012). Managing DICOM images: Tips and tricks for the radiologist. *Indian Journal of Radiology and Imaging*, *22*(01), 4–13. https://doi.org/10.4103/0971-3026.95396

[8] Kathiravelu, P., Sharma, P., Sharma, A., Banerjee, I., Trivedi, H., Purkayastha, S., . . . , & Gichoya, J. W. (2021). A DICOM framework for machine learning and processing pipelines against real-time radiology images. *Journal of Digital Imaging*, *34*(4), 1005–1013. https://doi.org/10.1007/s10278-021-00491-w

[9] Prevedello, L. M., Halabi, S. S., Shih, G., Wu, C. C., Kohli, M. D., Chokshi, F. H., . . . , & Flanders, A. E. (2019). Challenges related to artificial intelligence research in medical imaging and the importance of image analysis competitions. *Radiology: Artificial Intelligence*, *1*(1), e180031. https://doi.org/10.1148/ryai.2019180031

[10] Willemink, M. J., Koszek, W. A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., . . . , & Lungren, M. P. (2020). Preparing medical imaging data for machine learning. *Radiology*, *295*(1), 4–15. https://doi.org/10.1148/radiol.2020192224

[11] Fajar, A., Sarno, R., Fatichah, C., & Fahmi, A. (2022). Reconstructing and resizing 3D images from DICOM files. *Journal of King Saud University-Computer and Information Sciences*, *34*(6), 3517–3526. https://doi.org/10.1016/j.jksuci.2020.12.004

[12] Hnatkova, E., Kratky, P., & Dvorak, Z. (2014). Conversion of 2D medical scan data into 3D printed models. *Advances in Environmental Sciences, Development and Chemistry*, *2*, 315–318.

[13] Mamdouh, R., El-Bakry, H. M., Riad, A. L. A. A., & El-Khamisy, N. (2020). Converting 2D-medical image files “DICOM” into 3D-models, based on image processing, and analysing their results with python programming. *WSEAS Transactions on Computers*, *19*(2), 10–20. https://doi.org/10.37394/23205.2020.19.2

[14] Chen, Y., Liu, Z., Xu, H., Darrell, T., & Wang, X. (2021). Meta-baseline: Exploring simple meta-learning for few-shot learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 9062–9071.

[15] Cardoso, M. J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., . . . , & Feng, A. (2022). Monai: An open-source framework for deep learning in healthcare. *arXiv Preprint:2211.02701*. https://doi.org/10.48550/arXiv.2211.02701

[16] Kumar, A., & Panicker, M. R. (2024). AI enabled high frame rate portable ultrasound imaging pipeline: Prototype implementation with GPU acceleration. In *2024 IEEE Ultrasonics, Ferroelectrics, and Frequency Control Joint Symposium*, 1–5. https://doi.org/10.1109/UFFC-JS60046.2024.10794017

[17] Mitura, J., Chrapko, B. E., & Bachanek-Mitura, O. (2024). MedVoxelHD: Improved CUDA-accelerated morphological Hausdorff distances in medical image analysis. *SoftwareX*, *26*, 101744. https://doi.org/10.1016/j.softx.2024.101744

[18] Adusumalli, H., Kalyani, D., Sri, R. K., Pratapteja, M., & Rao, P. P. (2021). Face mask detection using OpenCV. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks*, 1304–1309. https://doi.org/10.1109/ICICV50876.2021.9388375

[19] Roelofs, R., Shankar, V., Recht, B., Fridovich-Keil, S., Hardt, M., Miller, J., & Schmidt, L. (2019). A meta-analysis of overfitting in machine learning. *Advances in Neural Information Processing Systems*, *32*, 1–11.

[20] Clunie, D. A. (2021). DICOM format and protocol standardization—A core requirement for digital pathology success. *Toxicologic Pathology*, *49*(4), 738–749. https://doi.org/10.1177/0192623320965893

[21] Farahat, Z., Hasni, M., Megdiche, K., Souissi, N., & Ngote, N. (2020). Comparative study of DICOM files handling software’s: Study based on the anatomage table. In *Innovation in Information Systems and Technologies to Support Learning Research: Proceedings of EMENA-ISTL 2019, 7*, 390–399. https://doi.org/10.1007/978-3-030-36778-7_43

[22] Kalen, J. D., Clunie, D. A., Liu, Y., Tatum, J. L., Jacobs, P. M., Kirby, J., . . . , & Doroshow, J. H. (2021). Design and implementation of the pre-clinical DICOM standard in multi-cohort murine studies. *Tomography*, *7*(1), 1–9. https://doi.org/10.3390/tomography7010001

[23] Nord, M., Webster, R. W., Paton, K. A., McVitie, S., McGrouther, D., MacLaren, I., & Paterson, G. W. (2020). Fast pixelated detectors in scanning transmission electron microscopy. Part I: Data acquisition, live processing, and storage. *Microscopy and Microanalysis*, *26*(4), 653–666. https://doi.org/10.1017/S1431927620001713

[24] Yaneva-Sirakova, T., Zlatancheva, G., Karamfiloff, K., Traykov, L., Petrov, I., & Vassilev, D. (2023). The role of periprocedural hemodynamic variables during carotid stenting for the mid-term general mortality in advanced age patients. *Folia Medica*, *65*(6), 902–908. https://doi.org/10.3897/folmed.65.e100100

[25] Sun, X., Li, L., Bo, L., Wu, X., Wei, Y., & Li, B. (2024). Automatic software vulnerability classification by extracting vulnerability triggers. *Journal of Software: Evolution and Process*, *36*(2), e2508. https://doi.org/10.1002/smr.2508

[26] Schroeder, A. B., Dobson, E. T., Rueden, C. T., Tomancak, P., Jug, F., & Eliceiri, K. W. (2021). The ImageJ ecosystem: Open-source software for image visualization, processing, and analysis. *Protein Science*, *30*(1), 234–249. https://doi.org/10.1002/pro.3993

[27] Maggio, L. A., Larsen, K., Thomas, A., Costello, J. A., & Artino Jr, A. R. (2021). Scoping reviews in medical education: A scoping review. *Medical Education*, *55*(6), 689–700. https://doi.org/10.1111/medu.14431

[28] Peters, M. D., Marnie, C., Colquhoun, H., Garritty, C. M., Hempel, S., Horsley, T., . . . , & Tricco, A. C. (2021). Scoping reviews: Reinforcing and advancing the methodology and application. *Systematic Reviews*, *10*, 263. https://doi.org/10.1186/s13643-021-01821-3

[29] Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods*, *5*(4), 371–385. https://doi.org/10.1002/jrsm.1123

[30] Westphaln, K. K., Regoeczi, W., Masotya, M., Vazquez-Westphaln, B., Lounsbury, K., McDavid, L., . . . , & Ronis, S. D. (2021). From Arksey and O’Malley and Beyond: Customizations to enhance a team-based, mixed approach to scoping review methodology. *MethodsX*, *8*, 101375. https://doi.org/10.1016/j.mex.2021.101375

[31] Mustapha, I., Ali, M., Khan, N., & Sikandar, H. (2023). The impact of industry 4.0 on innovative organisations, a thematic review using the PRISMA statement 2020. *International Journal of Interactive Mobile Technologies*, *17*(9), 88–105. https://doi.org/10.3991/ijim.v17i09.39465

[32] Pham, H. H., Do, D. V., & Nguyen, H. Q. (2021). DICOM imaging router: An open deep learning framework for classification of body parts from DICOM X-ray scans. *arXiv Preprint:2108.06490*. https://doi.org/10.48550/arXiv.2108.06490

[33] Vallez, N., Espinosa-Aranda, J. L., Pedraza, A., Deniz, O., & Bueno, G. (2023). Deep learning within a DICOM WSI viewer for histopathology. *Applied Sciences*, *13*(17), 9527. https://doi.org/10.3390/app13179527

[34] Kawuma, S., Kumbakumba, E., Mabirizi, V., Nanjebe, D., Mworozi, K., Oyesigye Mukama, A., & Kyasimire, L. (2024). Diagnosis and classification of tuberculosis chest X-ray images of children less than 15 years at Mbarara regional referral hospital using deep learning. *Journal of AI and Data Mining*, *12*(2), 315–324. https://doi.org/10.22044/jadm.2024.14270.2530

[35] Alshmrani, G. M. M., Ni, Q., Jiang, R., Pervaiz, H., & Elshennawy, N. M. (2023). A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. *Alexandria Engineering Journal*, *64*, 923–935. https://doi.org/10.1016/j.aej.2022.10.053

[36] Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., . . . , & Alzheimer’s Disease Neuroimaging Initiative. (2019). Automated classification of Alzheimer’s disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical*, *21*, 101645. https://doi.org/10.1016/j.nicl.2018.101645

[37] Kuraparthi, S., Reddy, M. K., Sujatha, C. N., Valiveti, H., Duggineni, C., Kollati, M., & Kora, P. (2021). Brain tumor classification of MRI images using deep convolutional neural

network. *Traitement du Signal*, *38*(4), 1171–1179. https://doi.org/10.18280/ts.380428
[38] Jia, Z., & Chen, D. (2020). Brain tumor identification and classification of MRI images using deep learning techniques. *IEEE Access*. https://doi.org/10.1109/ACCESS.2020.3016319
[39] Zeimarani, B., Costa, M. G. F., Nurani, N. Z., Bianco, S. R., Pereira, W. C. D. A., & Costa Filho, C. F. F. (2020). Breast lesion classification in ultrasound images using deep convolutional neural network. *IEEE Access*, *8*, 133349–133359. https://doi.org/10.1109/ACCESS.2020.3010863
[40] Kumar, R., Mohanty, K., Mundra, A. S., & Sai, B. G. K. (2024). Classification of carotid artery intima-media thickness ultrasound images with deep learning. In *AIP Conference Proceedings, 3075*(1), 1–12. https://doi.org/10.1063/5.0226744
[41] Liebenlito, M., Irene, Y., & Hamid, A. (2020). Classification of tuberculosis and pneumonia in human lung based on chest X-ray image using convolutional neural network. *InPrime: Indonesian Journal of Pure and Applied Mathematics*, *2*(1), 24–32. https://doi.org/10.15408/inprime.v2i1.14545
[42] Masood, M., Nazir, T., Nawaz, M., Mehmood, A., Rashid, J., Kwon, H. Y., . . . , & Hussain, A. (2021). A novel deep learning method for recognition and classification of brain tumors from MRI images. *Diagnostics*, *11*(5), 744. https://doi.org/10.3390/diagnostics11050744
[43] Yimer, F., Tessema, A. W., & Simegn, G. L. (2021). Multiple lung diseases classification from chest X-ray images using deep learning approach. *International Journal of Advanced Trends in Computer Science and Engineering*, *10*(5), 2936–2946. https://doi.org/10.30534/ijatcse/2021/021052021
[44] Ibrahim, A. U., Ozsoz, M., Serte, S., Al-Turjman, F., & Yakoi, P. S. (2024). Pneumonia classification using deep learning from chest X-ray images during COVID-19. *Cognitive Computation*, *16*, 1589–1601. https://doi.org/10.1007/s12559-020-09787-5
[45] Cao, Z., Duan, L., Yang, G., Yue, T., & Chen, Q. (2019). An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Medical Imaging*, *19*, 1–9. https://doi.org/10.1186/s12880-019-0349-x
[46] Jabeen, K., Khan, M. A., Alhaisoni, M., Tariq, U., Zhang, Y. D., Hamza, A., . . . , & Damaševičius, R. (2022). Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. *Sensors*, *22*(3), 807. https://doi.org/10.3390/s22030807
[47] Mamdouh, R., El-Bakry, H. M., Riad, A., & El-Khamisy, N. (2020). Converting 2D-medical image files “DICOM” into 3D- models, based on image processing, and analysing their results with Python programming. *WSEAS Transactions on Computers*, *19*(2), 10–20. https://doi.org/10.37394/23205.2020.19.2
[48] Tran, M. H., & Vu, H. M. Q. (2018). An improved method for building a 3D model from 2D DICOM. In *2018 International Conference on Advanced Computing and Applications*, 125–131. https://doi.org/10.1109/ACOMP.2018.00027
[49] Tran, M. H., & Vu, H. M. Q. (2016). A research on 3D model construction from 2D DICOM. In *2016 International Conference on Advanced Computing and Applications*, 158–163. https://doi.org/10.1109/ACOMP.2016.031
[50] Vahdati, S., Khosravi, B., Mahmoudi, E., Zhang, K., Rouzrokh, P., Faghani, S., . . . , & Erickson, B. J. (2024). A guideline for open-source tools to make medical imaging data ready for artificial intelligence applications: A society of imaging informatics in medicine (SIIM) survey. *Journal of Imaging Informatics in Medicine*, *37*(5), 2015–2024. https://doi.org/10.1007/s10278-024-01083-0
[51] Stiefel, M., Müller, M., Bachmann, B. I., Guitar, M. A., Nayak, U. P., & Mücklich, F. (2024). Enhancing machine learning classification of microstructures: A workflow study on joining image data and metadata in CNN. *MRS Communications*, *14*(3), 363–371. https://doi.org/10.1557/s43579-024-00549-0

How to Cite: Mabirizi, V., Kawuma, S., Natumanya, D., & Wasswa, W. (2025). Deep Learning Techniques in DICOM Files Classification: A Systematic Review. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA52024425