Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 21 No. 8 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v21i08.55419

PAPER

Efficient Deep Learning for Radiographic Body Part Classification

Hanan Sabbar[1](✉), Hassan Silkan[1], Khalid Abbad[2], El Mehdi Bellfkih[3], Imrane Chems Eddine Idrissi[4]

[1]LAROSERI, Computer Science Department, Chouaib Doukkali University, El Jadida, Morocco

[2]LSIA, Mathematics and Computer Science Department, Sidi Mohamed Ben Abdellah University, Fès, Morocco

[3]LAMS, Mathematics and Computer Science Department, Hassan II University, Casablanca, Morocco

[4]LTIM, Mathematics and Computer Science Department, Hassan II University, Casablanca, Morocco

sabbar.h@ucd.ac.ma

ABSTRACT

The growing demand for automated classification of medical X-ray images has driven the development of efficient deep learning models. This study introduces YOLOv8n-cls, a lightweight and accurate solution for body part classification in radiographic images, achieving a top-1 accuracy of 99.4%, outperforming several state-of-the-art models. A dataset of 22,500 images spanning 11 anatomical categories was constructed from public (MURA) and private DICOM sources, including previously underrepresented regions to improve generalizability. A comprehensive pre-processing pipeline comprising image conversion, windowing, and text removal via Keras OCR was employed, alongside various data augmentation techniques to enhance model robustness. Comparative evaluation with other deep learning models (EfficientNet B4, B2, B6, B0, DaViT Base, EfficientNetV2 XL) confirmed the superiority of YOLOv8n-cls in terms of both accuracy and computational efficiency. Model performance was assessed using top-1 accuracy, precision, recall, F1 score, loss curves, and confusion matrices. This work contributes a scalable, high-performing approach tailored for clinical deployment and highlights the importance of incorporating diverse anatomical coverage and pre-processing strategies for reliable X-ray classification.

KEYWORDS

medical imaging, X-ray body part classification, deep learning, medical image pre-processing, YOLOv8n-cls

1 INTRODUCTION

Recent advancements in deep learning have revolutionised medical imaging by enabling highly accurate automation of diagnostic and classification processes. Among the various diagnostic tools used in clinical practice, radiographic imaging is particularly well positioned to benefit from these innovations. One fundamental application in this domain is the classification of anatomical regions within

Sabbar, H., Silkan, H., Abbad, K., Bellfkih, E.M., Idrissi, I.C.E. (2025). Efficient Deep Learning for Radiographic Body Part Classification. *International Journal of Online and Biomedical Engineering (iJOE)*, 21(8), pp. 120–137. https://doi.org/10.3991/ijoe.v21i08.55419

Article submitted 2025-03-09. Revision uploaded 2025-04-23. Final acceptance 2025-05-03.

radiographic images. This task is crucial for supporting disease diagnosis, surgical planning, and medical education. Reliable classification not only facilitates the proper routing of medical images to the relevant clinical specialists but also improves the integration of imaging data into electronic health records (EHRs), thereby optimising the efficiency of clinical workflows [1].

Beyond medical imaging, deep learning has also demonstrated significant efficacy across other domains. For example, it has been successfully applied in evaluating students' cognitive abilities in online education environments [2] and in classifying user feedback in video games using advanced feature extraction techniques [3]. These applications highlight the adaptability and performance of deep learning models, reinforcing their relevance for complex tasks such as anatomical region classification in radiographic images.

Traditional classification methods have relied on manual annotation or rule-based algorithms, which are time-consuming, prone to human error, and lack scalability [4]. The emergence of convolutional neural networks (CNNs) has brought powerful tools for automating these processes. Architectures such as EfficientNet [5] and Vision Transformers [6] have shown remarkable performance in image classification tasks. However, their high computational cost makes them less suitable for resource-constrained clinical environments, highlighting the need for lightweight yet accurate deep learning architectures.

To address these limitations, the you only look once (YOLO) family of models has been widely adopted in computer vision tasks due to its real-time performance and compact architecture [7]. In particular, the YOLOv8n-cls model demonstrates a favourable balance between accuracy and computational efficiency. While YOLO has been successfully applied in natural image tasks, its use in medical imaging, specifically for body part classification in radiographs, remains underexplored.

In this study, we evaluate the performance of YOLOv8n-cls on a comprehensive dataset composed of 22,500 radiographic images covering 11 body part categories. This dataset combines public data from MURA [8] with private DICOM-formatted images, including several under-represented anatomical regions, thereby improving the model's generalisation. A rigorous pre-processing pipeline was applied, consisting of image format conversion, windowing, and text removal using Keras OCR. Additionally, data augmentation techniques such as flipping, translation, rotation, and brightness adjustments were used to enhance robustness.

The performance of YOLOv8n-cls is benchmarked against state-of-the-art models such as EfficientNet B0, B2, B4, B6, EfficientNetV2 XL, and DaViT Base. Evaluation metrics include top-1 accuracy, training and validation losses, and confusion matrices, which together provide comprehensive insights into the model's capabilities and limitations.

Contributions of this study are summarised as follows:

- A new, diverse dataset of radiographic images was constructed, incorporating under-represented body part categories to improve model generalisation.
- A detailed pre-processing strategy was implemented, combining windowing and automatic text removal.
- YOLOv8n-cls is systematically compared to leading deep learning models, demonstrating superior accuracy (Top-1 Accuracy of 99.4%) and computational efficiency.

- The model is proposed as a scalable and reliable solution for automated body part classification in clinical imaging workflows.

2 RELATED WORK

The classification of radiographic body parts has evolved significantly with the integration of deep learning, yet persistent gaps in anatomical coverage and computational efficiency remain critical challenges. Early approaches relied on manual annotation or rule-based systems [1], but these methods struggled with scalability and inter-observer variability, particularly for rare or overlapping anatomical regions [9]. The advent of CNNs marked a paradigm shift, with architectures such as EfficientNet [5] achieving state-of-the-art performance on standardised datasets such as MURA, which focuses predominantly on upper extremities [10]. However, the narrow anatomical scope of these datasets limited their clinical utility, highlighting the need for models that generalise across a broader range of anatomical structures.

To overcome these limitations, researchers have explored hybrid datasets and multimodal fusion techniques. For instance, Pelka et al. [11] proposed a method to combine automatically generated image keywords with radiographs to enable enriched multi-modal image representation for improved body part classification, while Pham et al. [12] developed the DICOM Imaging Router, a deep learning system for automatic classification of five anatomical regions in DICOM X-ray scans. Their MobileNet-V1 model achieved 98.5% precision and 98.1% F1-score, demonstrating the value of dataset diversification. However, these models remain limited both in the number of anatomical regions covered and in their applicability to real-time clinical use.

Class imbalance and limited annotations remain major challenges, particularly when integrating rare anatomical regions. Self-supervised learning (SSL) frameworks, such as MUSCLE [13], have addressed these issues by utilising contrastive learning on unlabelled X-rays, reducing annotation requirements by 70% while enhancing cross-domain generalisation. Similarly, a one-class classification approach using the One-Class Support Vector Machine (OCSVM) was proposed to tackle data imbalance issues in medical imaging [14]. This method achieved 92% accuracy in classifying highly imbalanced datasets, including space-occupying kidney lesions and breast cancer metastasis prediction, demonstrating its effectiveness compared to traditional binary classification models. These approaches align with broader trends in federated learning [15], which enable collaborative model training across institutions without data sharing, offering a critical advantage for expanding anatomical coverage while preserving patient privacy. However, these techniques are rarely applied to body part classification in radiographic images.

The integration of 3D imaging data has further complicated body part classification, as traditional 2D CNNs struggle with volumetric context. A hybrid CNN-LSTM framework was developed to partition whole-body CT scans into five consecutive anatomical regions, achieving accuracies of 97.3% and 98.2% for CT and MRI scans [16]. Similarly, a two-step framework for body part detection using EfficientNet and HRNet for pose estimation was proposed, achieving a Probability of Correct Keypoint (PCK) of 90.13% and a mean average precision (mAP) of 89.80%

for body part detection in non-medical images [17]. These methods demonstrate the potential of combining spatial and contextual features, but their complexity and reliance on multi-stage architectures limit their scalability for real-world clinical applications.

Data pre-processing and augmentation remain pivotal for robust classification. Chlap et al. [18] conducted a comprehensive review of 65 distinct data augmentation techniques used in medical imaging, categorising them into 11 different purposes. These techniques include spatial transformations, contrast adjustments, noise addition, deformations, data mixing, filtering, multi-scale approaches, and meta-learning, among others, demonstrating their role in improving model generalisability. Concurrently, adversarial training techniques such as WGAN-GP [19] have been used to address data imbalance in medical imaging by generating synthetic images, improving classification accuracy. In histopathological breast cancer data, WGAN-GP boosted accuracy from 84.25% to 95.40%, mitigating overfitting in small datasets and rare anatomical classes. Despite their utility, these methods are underutilised in body part classification tasks and often lack integration with lightweight, real-time models.

A CNN-based X-ray body part classification model was proposed [20], achieving 97.38% accuracy using FastAI and TensorFlow. The approach involved pre-processing steps such as normalisation, windowing, text removal, and data augmentation. The study demonstrated the effectiveness of CNNs in automating X-ray analysis and emphasised the importance of expanding anatomical coverage, aligning with the objectives of our work. However, the study remained limited to a smaller set of anatomical regions and did not include comparisons with more efficient models such as you only look once.

Recent advancements in lightweight architectures have prioritised computational efficiency without compromising accuracy. YOLOv8 [21], a model from the YOLO family, has demonstrated real-time inference capabilities while maintaining precision in tasks such as lesion detection [22] and multi-organ localisation [23]. However, its application to body part classification remains largely underexplored, particularly for novel regions such as mammography views and cervical spine projections. The present study addresses this gap by evaluating the performance of YOLOv8n-cls for comprehensive body part classification in radiographic images.

Despite significant progress in the field, current studies continue to exhibit limitations such as inadequate anatomical coverage, restricted support for real-time applications, and high computational requirements. To address these challenges, our study introduces a novel framework based on the YOLOv8n-cls architecture for body part classification in radiographic images. The proposed approach offers three main contributions: (1) an expansion to 11 anatomical categories, including six novel regions not present in the widely used MURA dataset; (2) the implementation of dynamic data augmentation techniques to improve generalisation under varied imaging conditions; and (3) the integration of DICOM metadata to reduce ambiguity in images with overlapping anatomical regions. These innovations collectively aim to enhance both the classification performance and the computational efficiency of the model, facilitating its adoption in real-world clinical environments.

3 MATERIALS AND METHODS

This section presents the methodology used for automated classification of anatomical regions in radiographic images with the YOLOv8n-cls model. The process comprises four core stages: dataset construction, image pre-processing, model training, and performance evaluation. We compiled a dataset of 22,500 radiographic images from both public repositories, including the MURA dataset, and private DICOM-formatted clinical sources to ensure diversity and representativeness.

The pre-processing stage involved converting image formats, applying windowing, and removing embedded text using Keras OCR. To enhance model generalisation, we applied data augmentation techniques such as flipping, rotation, translation, and brightness/contrast adjustments. We trained the model using optimised hyperparameters and applied stratified sampling to split the dataset into training (70%), validation (15%), and test (15%) sets, ensuring balanced class distribution. Figure 1 illustrates the complete workflow. The following subsections detail each stage to provide a structured and transparent overview of the approach.

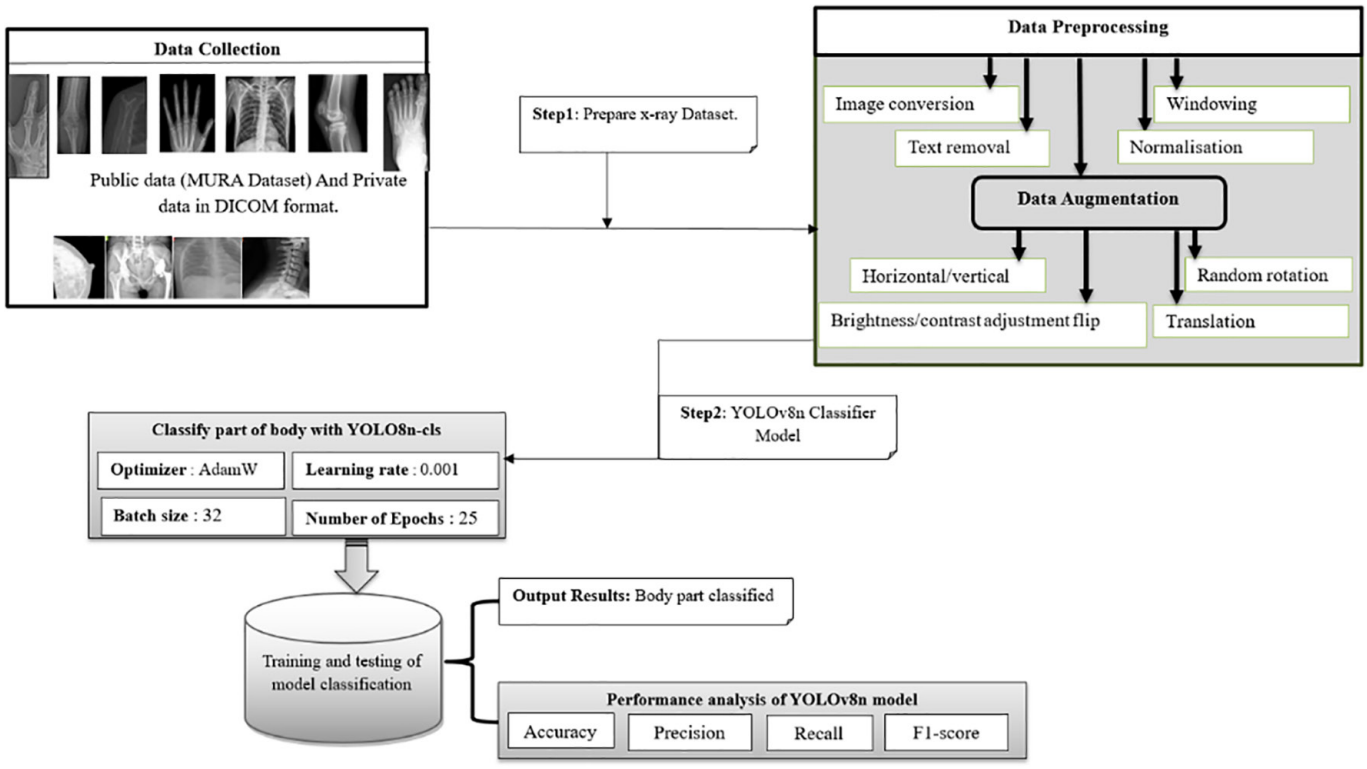


Fig. 1. Flow chart of proposed method

3.1 Data collection

To build a diverse and representative dataset for radiographic classification, we adopted a multifaceted approach. The dataset combines two primary sources: the publicly available MURA dataset from Kaggle (http://stanfordmlgroup.github.io/

competitions/mura), which includes radiographs of various body parts, and private DICOM-format data, representing real-world clinical cases. In total, our dataset consists of 22,500 images across 11 anatomical categories.

These combined sources provided a comprehensive dataset for model training and validation. Extensive manual classification was required since the images were not pre-labelled. Some images contained multiple anatomical features simultaneously, complicating classification. To ensure better consistency, we merged certain categories, such as "FOREARM" and "HAND" and "THORAX (LATERAL)" and "LUMBAR", due to their anatomical overlap in radiographs. This approach reduced ambiguity and enhanced model robustness by minimising uncertain classifications.

Table 1 provides a clear and updated overview of the dataset size and its structured distribution across training, validation, and test sets. Using stratified sampling, the dataset maintains balanced class proportions, enabling consistent preprocessing, effective management, fair evaluation, and improved model generalisation across all anatomical categories.

Table 1. Distribution of images per body part across dataset splits

Body Part	Training Set	Validation Set	Test Set	Total
Abdominal	1400	300	300	2000
Elbow	1400	300	300	2000
Feet	1400	300	300	2000
Forearm and Hand	2100	450	450	3000
Humerus	1400	300	300	2000
Shoulder	1190	255	255	1700
Mammography	1400	300	300	2000
Thorax (Lateral and Lumbar)	1400	300	300	2000
Thorax (Frontal)	1400	300	300	2000
Hips	1400	300	300	2000
Cervical (Head or neck)	1260	270	270	1800
Total	**17,150**	**3,375**	**3,375**	**22,500**

3.2 Pre-processing data

The pre-processing step includes image conversion, text removal, and data augmentation to optimise input quality and enhance model training. These processes ensure that the dataset is standardised, free of unnecessary textual elements, and sufficiently diverse to improve the generalisation capabilities of the deep learning model.

Image format conversion and windowing. The dataset initially consisted of images in the DICOM format, which is widely used in medical imaging due to its rich metadata. One of the primary challenges in processing DICOM images is their large size; some radiographs can reach 3000×4000 pixels, requiring significant memory capacity for storage and processing. Directly feeding these high-resolution images

into deep learning models can lead to memory constraints, slowing down training and increasing computational costs.

To ensure compatibility with deep learning algorithms, these images were converted to PNG format using the Pydicom library for extraction and PIL (Python Imaging Library) for conversion. This process also included windowing, which adjusts greyscale levels based on tissue densities of interest.

We utilised Pydicom to extract images from DICOM files and PIL to convert pixel data into PNG format. Additionally, windowing was applied, a crucial step in optimising greyscale levels to enhance anatomical or pathological structures. Windowing adjusts brightness and contrast, ensuring that the final PNG images retain key diagnostic features, ultimately improving the quality of deep learning model inputs.

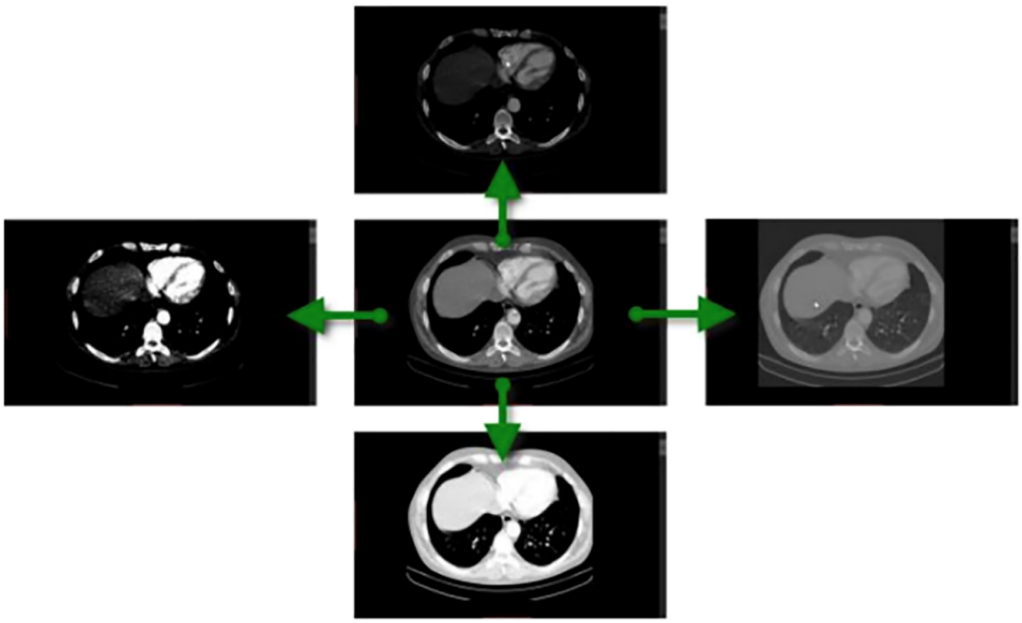


Fig. 2. Windowing effect

The Figure 2 demonstrates the impact of windowing on radiological visualisation, showing how adjusting window centre and width enhances specific structures, maintaining essential diagnostic details. This conversion and windowing process ensures that deep learning models receive optimised images for accurate analysis. By applying windowing, the desired brightness and contrast from the X-ray machine are preserved critical for reliable medical diagnostics.

Text removal process. Radiographic images often contain overlaid text, such as hospital identifiers, patient information, and acquisition parameters, which, if retained, could introduce bias and lead to overfitting in deep learning models. To ensure the model focuses solely on anatomical structures, we implemented an automated text removal process using Keras OCR, a deep learning-based tool capable of detecting and eliminating textual elements from images. The process involved three key steps:

1. **Text detection:** Keras OCR scanned each radiograph and identified textual elements. The model generated bounding boxes around the detected text, providing precise localisation.

2. **Text removal:** A masking technique was applied to erase the text while preserving the underlying anatomical structures. In cases where text overlapped with important regions, interpolation techniques were used to reconstruct the original pixel values, ensuring minimal visual distortion.
3. **Manual validation:** Each cleaned image was manually reviewed to verify that all text was completely removed while maintaining diagnostic integrity. This ensured that no anatomical features were altered during the process.

a) Before b) After

Fig. 3. Text removal for radiographic images

Figure 3 illustrates the impact of text removal, comparing an original X-ray with overlaid text to the cleaned version. By pre-processing images in this manner, we enhanced the robustness, fairness, and reliability of the classification model, ensuring that predictions were based on actual radiographic features rather than unintended correlations.

By integrating these pre-processing steps, we ensured that the datasets used for training our models were of the highest possible quality, thereby maximising the accuracy and reliability of AI-driven diagnostics.

3.3 Image augmentation

To enhance model generalisation and improve classification accuracy, several data augmentation techniques were applied:

- **Fliplr (Flip Left-Right)):** Images were flipped horizontally, increasing variability and enabling the model to recognise symmetrical structures.
- **Flipud (Flip Up-Down):** Images were flipped vertically to diversify data and help the model learn features from various orientations.
- **Mosaic:** Four training images were combined into one, simulating complex anatomical interactions for better scene understanding.
- **Hsv_h (Brightness and Contrast Adjustment):** Adjustments improved the model's robustness against variations in lighting and contrast.
- **Translate:** Images were shifted along the x and y axes, adding spatial variability and improving the model's ability to handle positional shifts.
- **Rotate:** Random rotations within a specified range helped the model identify body parts regardless of the angle of capture.

These techniques were applied with specific parameters as shown in Table 2 to enrich the dataset and enhance model performance.

Table 2. Data augmentation techniques and their applied parameters

Technique	Value
Fliplr	0.3
Flipud	0.3
Mosaic	0.05
Hsv_h	0.05
Translate	0.2
Rotate	0.3

Figure 4 shows an example of augmentation techniques applied to radiographic images, showcasing used methods to enhance dataset diversity.

Fig. 4. The effect of augmentation for body part classification

3.4 Computer system configuration

In our study of body part classification models using YOLOv8n-cls, we trained the model using the Ultralytics, Pydicom PIL and Keras OCR libraries in Python. The training process was accelerated with a high-performance computing setup, featuring an NVIDIA GeForce RTX 4070 Ti GPU with 37,690 MB of memory and 7,680 CUDA cores, supported by an Intel Core i7-10700 CPU, 48 GB of RAM and CUDA-optimised GPU integration. This setup significantly improved the efficiency and speed of handling computationally intensive tasks associated with the large data set and numerous network parameters.

3.5 Model optimisation

To achieve optimal performance in the body parts classification task, the YOLOv8n-cls model was trained using carefully tuned hyperparameters. These parameters were selected to balance computational efficiency and model accuracy, ensuring robust learning from the dataset. The key hyperparameters used in the training process are outlined in Table 3.

Table 3. Optimization of hyperparameters used in the training of the YOLOv8n-cls network

Parameter	Value
Epochs	25
Batch size	32
Image size	640×640
Optimizer	AdamW
Learning Rate	1.0×10^{-3}
Momentum	0.9
Weight decay	0.0005

The use of the AdamW optimiser, a learning rate of 0.001, and a momentum of 0.9 provided a balance between convergence speed and generalisation. Additionally, the image size of 640×640 ensured sufficient detail in the input data, while a batch size of 32 and a weight decay of 0.0005 helped mitigate overfitting. These choices facilitated effective training of the YOLOv8n-cls model, resulting in accurate and reliable classification outcomes.

3.6 YOLOv8n-cls based model

For the classification of body parts in radiographic images, the YOLOv8n-cls model was chosen due to its speed and efficiency. This lightweight version of YOLOv8 is optimised for classification tasks, offering high accuracy while maintaining a compact architecture. Experimental comparisons between YOLOv8n (nano) and YOLOv8x (extra-large) revealed similar precision, leading to the selection of the nano version for its superior speed and lower computational requirements. This decision is based on the assumption that the observed limitations in accuracy stem from the dataset's size and diversity rather than the model's capacity.

The architecture of YOLOv8 (see Figure 5) includes key components such as the backbone, neck, and head, tailored for classification tasks. The backbone extracts essential features from input images using convolutional layers, while the neck enhances feature representation through feature pyramids. Finally, the head performs classification by mapping features to output labels. This streamlined design ensures efficiency and robustness for real-time applications.

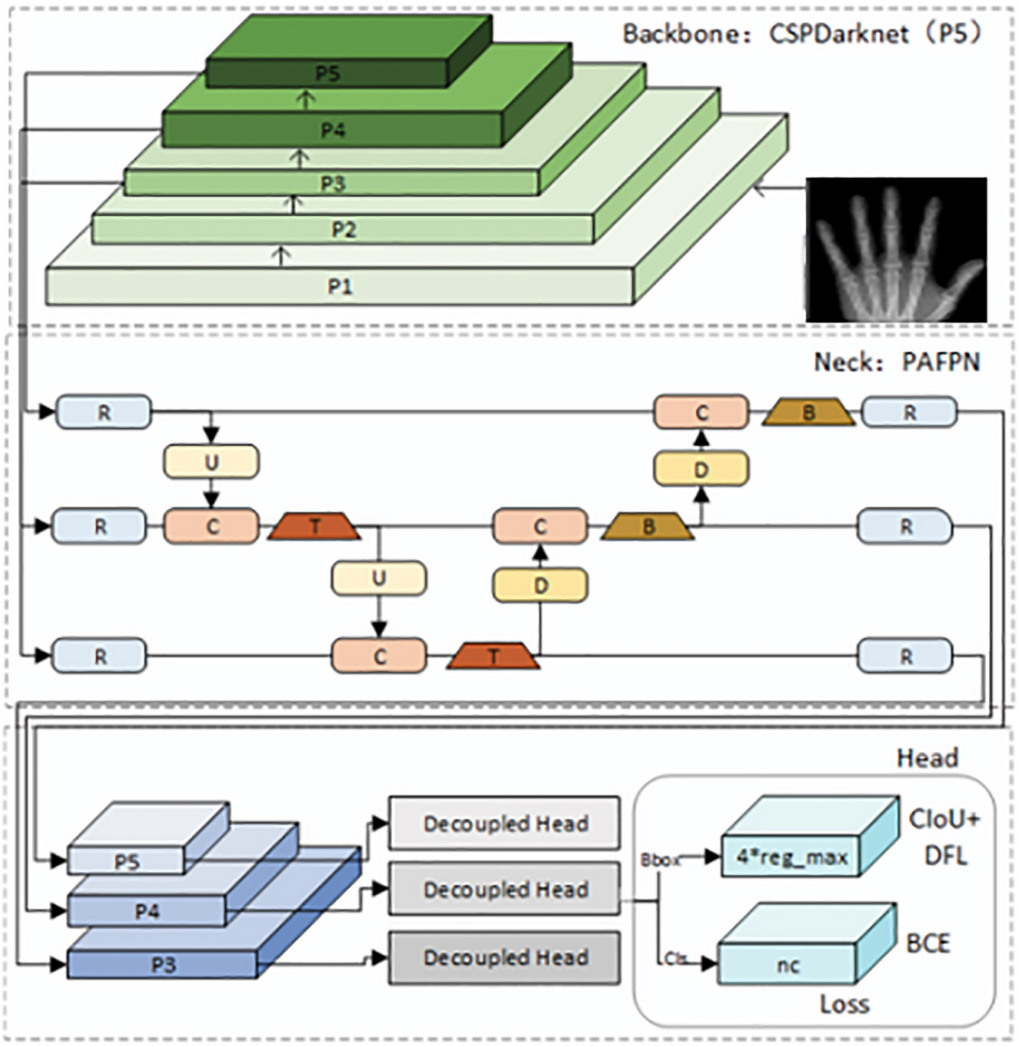


Fig. 5. The structure of YOLOv8n network

Note: The backbone, neck, and head networks are from top bottom.

3.7 Performance evaluation metrics

To evaluate the YOLOv8n-cls model's performance in body part classification, we employed standard evaluation metrics, including top-1 accuracy, precision, recall, and F1 score. These metrics quantify how accurately the model predicts and offer valuable insights into its reliability across various anatomical categories. We also used a confusion matrix to analyse prediction outcomes in detail, highlighting both correct classifications and misclassifications. Additionally, we tracked training and validation losses throughout the learning process to evaluate the model's ability to generalise to unseen data. Together, these evaluation tools provide a comprehensive

assessment of the model's accuracy and robustness, supporting its potential deployment in clinical environments. The definitions of the key metrics used appear below.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \times 100 \quad (3)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

4 RESULTS AND DISCUSSION

This section presents a comprehensive evaluation of the YOLOv8n-cls model for body part classification, highlighting both its performance metrics and the underlying causes of classification errors. The evaluation leverages a range of indicators, including Top-1 accuracy, training and validation losses, confusion matrices, and class-wise precision, recall, and F1 scores. These metrics enable a multi-dimensional analysis of the model's effectiveness, robustness, and potential suitability for clinical deployment.

As illustrated in Figure 6, the YOLOv8n-cls based model achieved a Top-1 accuracy of 99.4% at the 19th epoch, demonstrating high precision and stability in identifying anatomical regions. This performance ceiling is likely attributable to the limited dataset diversity, indicating that expanding the dataset to include a larger dataset, varied imaging conditions, and rare anatomical cases could improve generalisation.

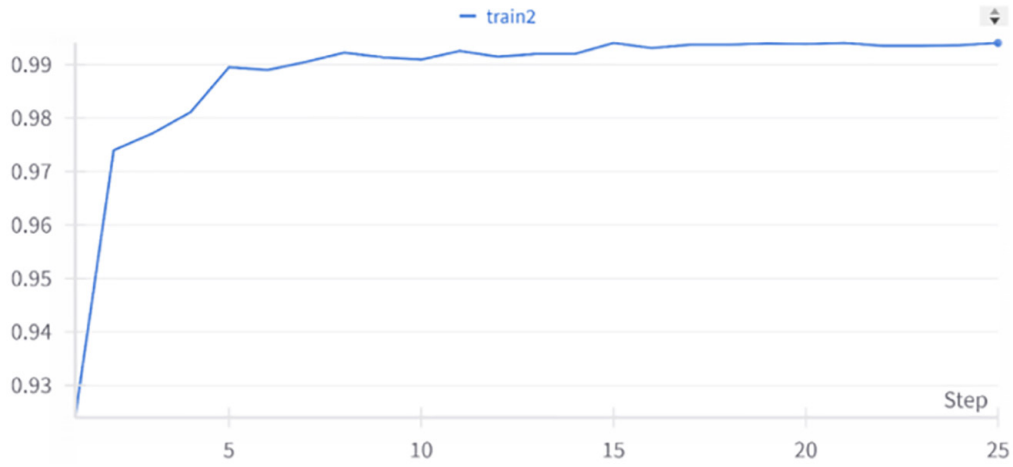
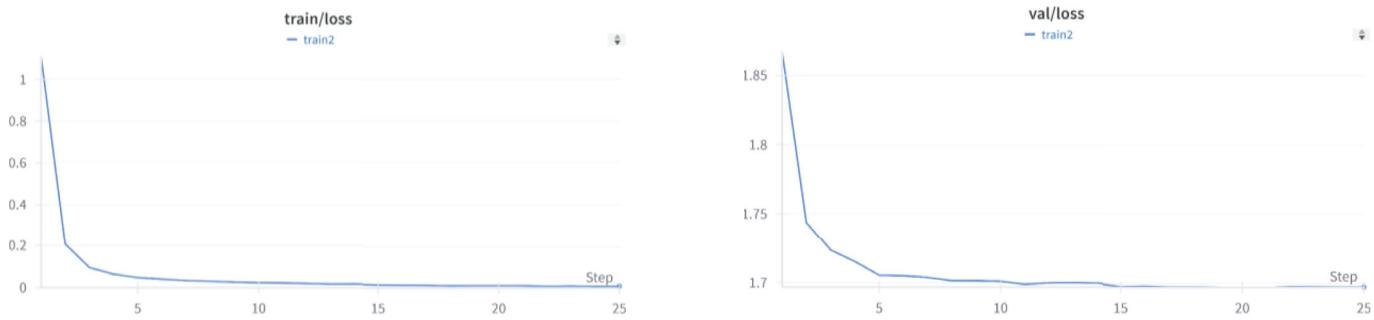


Fig. 6. Top-1 accuracy achieved by YOLOv8n-cls in body part classification

Figure 7 presents the training and validation loss curves, which serve as key indicators of the learning process. The training loss (see Figure 7a) decreased steadily, reflecting the model's ability to fit the data effectively. Meanwhile, the validation loss (see Figure 7b) remained consistently low, confirming the model's generalisation capability on unseen samples. This consistent convergence trend, accompanied

by minimal overfitting, validates the robustness of the training strategy and the appropriateness of the selected hyperparameters.



a) Training Loss **b) Validation Loss**

Fig. 7. Training and validation loss during model training

The confusion matrix in Figure 8 offers a detailed view of the model's classification outcomes across 11 body part categories. Diagonal dominance confirms accurate predictions, while off-diagonal entries reveal misclassifications. Analysis of the confusion matrix reveals that some misclassifications occurred due to overlapping body parts in a single image. For example, the elbow was occasionally misclassified as the humerus, and the hand was confused with the foot, suggesting that overlapping radiographic features and similar orientations posed classification challenges. These results emphasise the need to explore alternative strategies such as multi-label classification or incorporating spatial context.

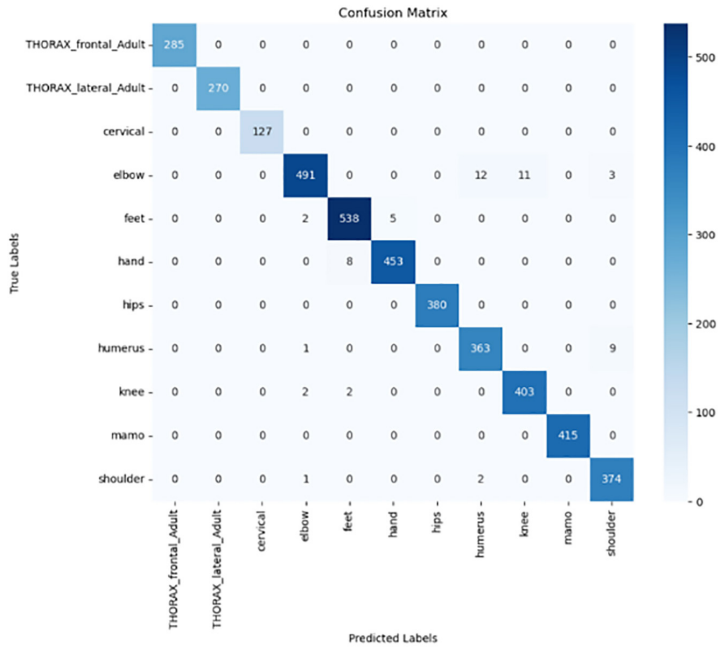


Fig. 8. Confusion matrix for YOLOv8n-cls body part classification

In addition to Top-1 accuracy, we reported precision, recall, and F1-score for each anatomical category to ensure greater statistical rigor. These metrics reflect the model's ability to minimise both false positives and false negatives, offering a more comprehensive evaluation of classification performance. For instance, the thorax

frontal and thorax lateral categories achieved 100% precision and 100% recall, demonstrating perfect identification. In contrast, the elbow category exhibited slightly lower recall, likely due to anatomical overlap with neighbouring regions such as the humerus and shoulder. A summary of measurements by class is shown in Table 4 below:

Table 4. Class-wise performance metrics of the YOLOv8n-cls model on the body part classification task

Class	Precision (%)	Recall (%)	F1-Score (%)
Thorax_frontal	100	100	100
Thorax_lateral	100	100	100
Cervical	100	99.22	99.61
Elbow	98.79	94.97	96.84
Feet	98.18	98.72	98.44
Hand	98.91	98.26	98.59
Hips	100	100	100
Humerus	96.29	97.32	96.8
Knee	97.34	99.02	98.17
Mamography	100	100	100
Shoulder	96.64	99.2	97.91

5 COMPARATIVE ANALYSIS OF YOLOV8N-CLS AGAINST ALTERNATIVE MODELS

Several other models were tested to assess their performance in radiographic image classification. As outlined in Table 5, models such as EfficientNet B4, EfficientNet B2, EfficientNet B6, EfficientNet B0, DaViT Base, EfficientNetV2 XL, and EfficientNetV2 L were evaluated. While these models achieved strong top 1 accuracy, with EfficientNet B4 reaching 96.7%, none surpassed YOLOv8n-cls, which achieved a remarkableTop-1 accuracy of 99.4%.

This substantial performance gap highlights the superiority of YOLOv8n-cls for clinical applications requiring high precision. Unlike EfficientNet and DaViT-based models, which rely on larger and more complex architectures, YOLOv8n-cls achieves superior results with a lightweight yet highly efficient design, making it well-suited for real-time medical imaging tasks.

Table 5. Other experiments on body part classification

Model	Top-1 Accuracy (%)
YOLOv8n-cls based model	99.4
EfficientNet B4	96.7
EfficientNetV2 XL	96.4
EfficientNet B6	95.7
EfficientNetV2 L	96.2
EfficientNet B2	95.8
DaViT Base	92.3
EfficientNet B0	77.3

The Top-1 accuracy curves in Figure 9 further illustrate this advantage. YOLOv8n-cls demonstrates rapid convergence, high stability, and superior final accuracy, outperforming other architectures, particularly in early training phases. The EfficientNet models, while achieving high accuracy, required more training epochs to stabilize, and some models, such as EfficientNet B0 and DaViT Base, struggled to reach the same level of precision.

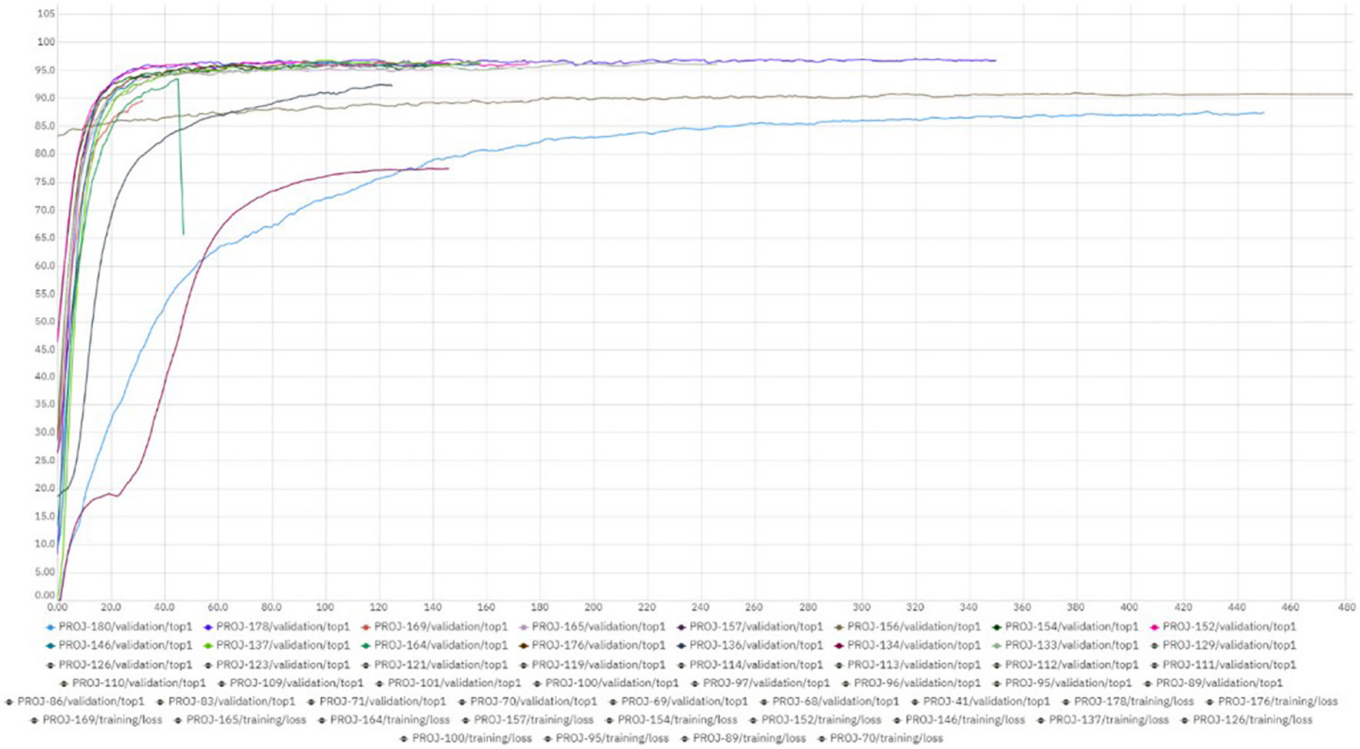


Fig. 9. The top-1 accuracy curves for multiple trained models

These results confirm that YOLOv8n-cls not only surpasses existing state-of-the-art models in accuracy but also offers computational efficiency, making it a robust solution for automated classification in radiographic imaging.

6 CONCLUSION

This study introduced an efficient deep learning framework based on the YOLOv8n-cls model for body part classification in radiographic images. By leveraging a diverse dataset that combines public and private sources, along with a robust preprocessing pipeline including image windowing, Keras OCR-based text removal, and data augmentation techniques, the model demonstrated high performance and strong generalisation ability.

The YOLOv8n-cls model achieved a Top-1 accuracy of 99.4 percent, outperforming several advanced architectures such as EfficientNet and DaViT. Evaluation metrics, including accuracy, precision, recall, F1 score, and confusion matrices, confirmed the model's effectiveness in reliably classifying anatomical regions while maintaining computational efficiency suitable for real-time clinical integration. Minor misclassifications were primarily due to overlapping anatomical features, suggesting opportunities for further improvement in handling complex image patterns.

In future work, the framework can be extended by incorporating multi-label classification strategies, enriching the dataset with additional rare and diverse anatomical cases, and including advanced statistical evaluations such as confidence intervals and standard deviation. These improvements are expected to further enhance the model's diagnostic performance and support broader clinical adoption.

The YOLOv8n-cls model presents a reliable, scalable, and high-performing solution for automated body part classification and offers promising potential for improving clinical workflows and decision-making in radiology.

7 REFERENCES

[1] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017. https://doi.org/10.1016/j.media.2017.07.005

[2] L. Zhao, "Use of a deep learning approach for the evaluation of students' online learning cognitive ability," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 18, no. 12, pp. 58–74, 2023. https://doi.org/10.3991/ijet.v18i12.41093

[3] A. Al Tawil, H. Fathi, S. Alzoubi, A. Shaban, and L. H. Almazaydeh, "Advanced feature extraction and machine learning techniques for classifying steam game feedback," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 1, pp. 107–125, 2025. https://doi.org/10.3991/ijim.v19i01.51237

[4] Goli Arji *et al.*, "A systematic literature review and classification of knowledge discovery in traditional medicine," *Computer Methods and Programs in Biomedicine*, vol. 168, pp. 39–57, 2019. https://doi.org/10.1016/j.cmpb.2018.10.017

[5] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the International Conference on Machine Learning (ICML)*, 2019, pp. 6105–6114.

[6] A. Dosovitskiy *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2021. https://doi.org/10.48550/arXiv.2010.11929

[7] J. Farooq *et al.*, "An improved YOLOv8 for foreign object debris detection with optimized architecture for small objects," *Multimedia Tools and Applications*, vol. 83, pp. 60921–60947, 2024. https://doi.org/10.1007/s11042-023-17838-w

[8] P. Rajpurkar *et al.*, "Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists," *PLoS Medicine*, vol. 15, no. 11, p. e1002686, 2018. https://doi.org/10.1371/journal.pmed.1002686

[9] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2097–2106. https://doi.org/10.1109/CVPR.2017.369

[10] P. Rajpurkar *et al.*, "MURA: Large dataset for abnormality detection in musculoskeletal radiographs," *arXiv preprint arXiv:1712.06957*, 2018. https://doi.org/doi:10.48550/arXiv.1712.06957

[11] O. Pelka, F. Nensa, and C. M. Friedrich, "Variations on branding with text occurrence for optimized body parts classification," in *Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 890–894. https://doi.org/10.1109/EMBC.2019.8857478

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

[13] Weibin Liao *et al.*, "Muscle: Multi-task self-supervised continual learning to pre-train deep models for x-ray images of multiple body parts," in *International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2022*, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, Eds., vol. 13438, Springer, Charm, 2022, pp. 151–161. https://doi.org/10.1007/978-3-031-16452-1_15

[14] L. Gao, L. Yang, D. Arefan, and S. Wu, "One-class classification for highly imbalanced medical image data," in *Proc. SPIE 11318, Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications,* 2020. https://doi.org/10.1117/12.2551389

[15] H. R. Roth, N. Rieke, S. Albarqouni, and Q. Li, "Guest editorial special issue on federated learning for medical imaging: Enabling collaborative development of robust AI models," *IEEE Transactions on Medical Imaging*, vol. 42, no. 7, pp. 1914–1919, 2023. https://doi.org/10.1109/TMI.2023.3278528

[16] Z. Ouyang, P. Zhang, W. Pan, and Q. Li, "Deep learning-based body part recognition algorithm for three-dimensional medical images," *Medical Physics*, vol. 49, no. 5, pp. 3067–3079, 2022. https://doi.org/10.1002/mp.15536

[17] M. B. Gamra, M. A. Akhloufi, C. Wang, and S. Liu, "Deep learning for body parts detection using HRNet and EfficientNet," in *17th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2021, pp. 1–8. https://doi.org/10.1109/AVSS52988.2021.9663785

[18] M. Cossio, "Augmenting medical imaging: A comprehensive catalogue of 65 techniques for enhanced data analysis," *arXiv preprint arXiv:2303.01178,* 2023. https://doi.org/10.48550/arXiv.2303.01178

[19] M. Ryspayeva, "Generative adversarial network as data balance and augmentation tool in histopathology of breast cancer," in *Proc. 2023 IEEE International Conference on Smart Information Systems and Technologies (SIST)*, 2023, pp. 99–104. https://doi.org/10.1109/SIST58284.2023.10223577

[20] S. R. Reeja *et al.*, "X-ray body part classification using custom CNN," *EAI Endorsed Transactions on Pervasive Health and Technology,* vol. 10, 2024. https://doi.org/10.4108/eetpht.10.5577

[21] M. Sohan, T. Sai Ram, and C. V. Rami Reddy, "A review on yolov8 and its advancements," in *International Conference on Data Intelligence and Cognitive Informatics*, Springer, 2024, pp. 529–545. https://doi.org/10.1007/978-981-99-7962-2_39

[22] M. Rašić, M. Tropčić, P. Karlović, D. Gabrić, M. Subašić, and P. Knežević, "Detection and segmentation of radiolucent lesions in the lower jaw on panoramic radiographs using deep neural networks," *Medicina*, vol. 59, no. 12, p. 2138, 2023. https://doi.org/10.3390/medicina59122138

[23] L. Du *et al.*, "Automatic body region localization in 3D-CT images based on the improved YOLO model," in *MATEC Web of Conferences*, 2022, vol. 355, p. 03022. https://doi.org/10.1051/matecconf/202235503022

8 AUTHORS

Hanan Sabbar is currently a student researcher at the LAROSERI Laboratory, specializing in medical imaging at the Faculty of Science, Chouaib Doukkali University, El Jadida, Morocco. Her study interests include medical image analysis, deep learning, machine learning, and bioinformatics (E-mail: sabbar.h@ucd.ac.ma).

Hassan Silkan is a research professor at the LAROSERI Laboratory, Chouaib Doukkali University, El Jadida, Morocco. His study focuses on artificial intelligence, advanced signal processing, and image analysis, notably applying AI techniques for

the detection, segmentation, and classification of medical images (E-mail: silkan.h@ucd.ac.ma).

Khalid Abbad is a research professor at the Intelligent Systems and Applications Laboratory, Sidi Mohamed Ben Abdellah University, Fès, Morocco. His work focuses on artificial intelligence, advanced signal processing, and medical image analysis, with an emphasis on machine learning applications (E-mail: khalid.abbad@usmba.ac.ma).

El Mehdi Bellfkih is a PhD holder and researcher at Hassan II University, specializing in coding theory and artificial intelligence. His research focuses on developing innovative solutions in coding and AI fields (E-mail: elmehdi.bellfkih@gmail.com).

Imrane Chems Eddine Idrissi is a PhD holder and researcher at Hassan II University, Morocco. His research focuses on machine learning and error-correcting codes, with applications to real-world problems (E-mail: imran.chems@gmail.com).