doi.org/10.54473/IJTRET.2024.9105
International Journal of Trendy Research in Engineering and Technology
Volume 9 Issue 1 February 2025
ISSN No. 2582-0958

AUTOMATED FEMORAL HEAD SEGMENTATION USING RCNN: A ROBUST APPROACH WITH COCO DATASET AND ADVANCED AUGMENTATION TECHNIQUES

Syed Osama Ali Shah[1*]
Alishba Eman[2]
Bahria University Karachi[1] , Bahria University Health Sciences Campus[2]
*Corresponding author – e-mail: s.osamaali72@gmail.com
Received 20 October 2024 Received in revised form 22 October 2024 Accepted 23 October 2024
Available Online 26 October 2024

ABSTRACT

Segmentation of the femoral head is crucial in orthopedic diagnostics, facilitating the identification and management of diseases such femoroacetabular impingement, avascular necrosis, and osteoarthritis. This study presents an automated method for femoral head segmentation utilizing a Region-based Convolutional Neural Network (RCNN) model, applied to a dataset of 2,035 femoral head pictures. Pre-processing techniques, such as auto-orientation, scaling to 640x640 pixels, contrast enhancement using contrast stretching, greyscale conversion, and noise reduction, were utilized to standardize the input data.

Data augmentation was executed through horizontal flipping, 90-degree rotations, and modifications to hue, saturation, brightness, blurring, and noise levels to enhance model robustness. The RCNN model underwent training for 300 epochs, attaining exceptional outcomes with a mean Average Precision (mAP) of 99.5%, precision of 99.7%, and recall of 100%, so illustrating the model's efficacy in precisely segmenting the femoral head.

Notwithstanding the favorable results, the dataset's restricted size may hinder the model's generalizability to varied real-world applications. Future endeavours will concentrate on augmenting the dataset and incorporating multimodal vision methodologies to guarantee conformity with contemporary healthcare demands and norms. These advances will improve the system's clinical utility and offer a more thorough solution for orthopedic imaging and diagnostic processes.

This research advances the field of medical picture segmentation by providing a scalable and efficient method for automated femoral head segmentation, with potential for further enhancement to satisfy real-time clinical requirements.

Keywords: Femoral head segmentation, Region-based Convolutional Neural Network (RCNN), multiscale geometric embedded CNN (MsgeCNN),Coco Dataset

I.INTRODUCTION

Segmentation of the femoral head is an essential task in orthopedic imaging, crucial for identifying diseases such as femoroacetabular impingement, avascular necrosis, and osteoarthritis. Precise segmentation of the femoral head enables accurate measurements, supports pre-surgical planning, and improves automated diagnostic systems. Nevertheless, attaining reliable and uniform outcomes in segmentation continues to pose a difficulty owing to discrepancies in imaging circumstances, anatomical structures, and noise.

Recent improvements in machine learning and computer vision have yielded increasingly sophisticated techniques to enhance segmentation jobs, with COCO (Common Objects in Context) segmentation demonstrating significant efficacy. This research concentrates on utilizing COCO segmentation methodologies for the automatic delineation of the femoral head, employing a dataset including 2,035 images. The objective is to improve the precision and dependability of femoral head segmentation by incorporating pre-processing and augmentation methods intended to address picture variability.

In the pre-processing pipeline, multiple stages are executed to standardize image input. This entails auto orienting the images, scaling them to a uniform 640x640 resolution, implementing contrast enhancement by contrast stretching, and converting them to greyscale, with an additional null filter used to diminish background noise. This preprocessing guaranteed consistency and enhanced the efficacy of the segmentation method.

Moreover, picture augmentation techniques are utilized to enhance the dataset's diversity and resilience, which is essential for training deep learning models to generalize well. Augmentations comprise horizontal flipping, 90-degree rotations (both clockwise and anticlockwise), and adjustments to hue (-15° to +15°), saturation (-25 to +25), brightness (-15 to +15), blur (from -10 to +10 with a maximum of 0.5px), and noise introduction up to 0.1% of the image pixels. These augmentations replicate real-world variability, aiding in the prevention of model overfitting and enhancing the trained model's resilience to imaging anomalies.

The suggested methodology seeks to tackle the issues of femoral head segmentation, offering a more efficient and precise alternative. This study enhances existing research by using a comprehensive augmentation and pre-processing method in the segmentation task, facilitating the development of superior diagnostic tools in orthopedics.

II.LITERATURE REVIEW

The automated segmentation of the femoral head is essential in medical imaging, especially for the diagnosis and treatment of disorders like osteonecrosis and hip osteoarthritis. Segmentation of anatomical structures, such as the femoral head, enhances the precision of clinical decision-making, facilitating more accurate measurements and improved pre-surgical planning. In recent years, deep learning, especially convolutional neural networks (CNNs) and U-Net architectures, has transformed medical picture segmentation, delivering superior performance across several segmentation tasks.

Convolutional neural networks (CNNs) have garnered significant interest for their capacity to represent intricate anatomical structures in medical imaging. Convolutional Neural Networks, in conjunction with region-based segmentation methods, have effectively addressed the issues associated with medical imaging, including inconsistent image quality, noise, and intricate anatomical structures. Li et al. [1] introduced a multiscale geometric embedded CNN (MsgeCNN) that has exhibited potential efficacy in segmenting femoral heads, illustrating the model's capacity to handle complex medical images with considerable precision. The multiscale approach of the model facilitates the acquisition of several levels of feature representation, allowing for the learning of both intricate and broad anatomical information. This is especially beneficial for segmentation jobs with small and unevenly shaped anatomical structures, such as the femoral head.

In a separate study, Ebsim et al. [2] created an automated approach employing U-Nets to segment hip osteophytes, commonly observed in osteoarthritis cases impacting the femoral head region. The U-Net architecture, recognized for its capacity to capture contextual and locational information via its contracting and expansive pathways, demonstrated superior efficacy in osteophyte detection. This study emphasizes the adaptability of U-Nets in medical segmentation tasks and their significance in orthopedic imaging, where precise segmentation of anatomical anomalies is essential for diagnosis and treatment.

Besides CNNs and U-Nets, various other deep learning models have demonstrated significant potential. Yun et al. [3] introduced a fully automated technique for segmenting the femur and femoral neck with CT images. Their model integrated sophisticated preprocessing and augmentation methods, including rotation, scaling, and elastic deformations, to enhance model resilience. Augmentation approaches, especially in medical imaging where data is sometimes scarce, are essential for enhancing model generalization. This method underscores the need of data diversity and augmentation in attaining precise segmentation outcomes, particularly when training datasets are limited in size.

Employing advanced augmentation techniques is a crucial strategy for enhancing the performance of segmentation models. Transformations like as rotation, scaling, and elastic deformations can substantially enhance dataset variability, hence improving the model's generalization to unfamiliar images. Yun et al. [3] illustrated the efficacy of data augmentation in their approach to femoral and femoral neck segmentation, wherein preprocessing procedures improved the model's overall accuracy. Pemmaraju et al. [4] emphasized the significance of synthetic data production to enhance the training process. Their research utilized cascaded neural networks for multi-organ segmentation, demonstrating that the generation of synthetic data for uncommon anatomical structures, such as the femoral head, can significantly enhance segmentation accuracy. The findings indicate that enhancing limited and diverse datasets with synthetic images significantly improves model robustness, resulting in superior performance in real-world clinical applications.

The COCO (Common Objects in Context) dataset is extensively utilized in computer vision applications for object detection and segmentation. Although COCO is not explicitly designed for medical imaging, its comprehensive annotations and object segmentation features provide it a beneficial resource when modified for medical applications. The COCO dataset has been employed to train algorithms for object segmentation, and its use in medical imaging—specifically for femoral head segmentation—can be advantageous. The adaptation process necessitates meticulous calibration due to the anatomical intricacies of medical images, which differ from the object recognition issues for which COCO was initially developed.

Zhang et al. [5] investigated the capability of deep learning models to enhance inter-reader reliability in femoral head volume segmentation. Their findings indicate that models trained on varied datasets such as COCO can markedly surpass conventional approaches, especially in diminishing variability across radiologists in manual segmentations. This corresponds with the objectives of employing COCO-inspired methodologies in this research, aimed at leveraging the advantages of COCO segmentation to enhance femoral head segmentation.

The future of segmentation models depends on their capacity to manage escalating complexity, with breakthroughs in deep learning, particularly generative adversarial networks (GANs), set to expand these limits. Generative Adversarial Networks (GANs), when integrated with conventional designs such as U-Net, have exhibited enhanced precision in segmentation tasks. Dong et al. [6] integrated U-Net with GANs to improve the automatic segmentation of medical images, especially for intricate anatomical features. Their findings indicated that this hybrid method was superior at distinguishing complex features, such as the curved structures of the femoral head, which are sometimes difficult to record accurately using standard techniques alone.

Furthermore, Ditmer [7] created a fully automated method for the detection and segmentation of the proximal femur in radiography images. Their research illustrated the feasibility of deep learning models in practical orthopedic applications, highlighting the significance of automation in improving clinical operations. The use of such methodology to femoral head segmentation, in conjunction with previously stated techniques, may yield substantial enhancements in automated diagnostic instruments for orthopedic care.

III.METHODOLOGY

This research technique employs a Region-based Convolutional Neural Network (RCNN) model on a collection of femoral head pictures to create an accurate and efficient automatic segmentation system. The procedure consists of multiple phases: data preparation, pre-processing, augmentation, model training utilizing the RCNN model, and performance assessment. The subsequent stages are delineated below:

1. Dataset Overview

This study utilized a total of 2,035 pictures of femoral heads. The photos were obtained from various orthopedic imaging datasets, guaranteeing extensive representation for quality, illumination, orientation, and noise levels. The dataset comprised images of both healthy and diseased femoral heads, ensuring a balanced collection for model training and assessment.

2. Preprocessing

The photos underwent pre-processing to achieve homogeneity and improve the model's performance through input standardization. The subsequent pre-processing procedures were executed:

- Auto-orientation: All photos were automatically adjusted to rectify misalignments and maintain uniform orientation.
- Resizing: Images were adjusted to a consistent dimension of 640x640 pixels, preserving aspect ratios and minimizing processing expenses.
- Contrast Adjustment: Contrast stretching was utilized to automatically enhance the contrast in the pictures, hence increasing the visibility of the femoral head in low-contrast areas.
- Greyscale Conversion: All photos were transformed to greyscale, streamlining the input data while preserving essential structural information required for precise segmentation.
- Filter Null: A filter null approach was employed to eliminate background noise and extraneous information, hence boosting attention on the femoral head region.

3. Data Augmentation

Extensive data augmentation approaches were employed to enhance the model's variability and robustness, hence preventing overfitting and improving generalization. The subsequent augmentations employed are :

- Horizontal Flip: Randomly flipping photos horizontally to emulate various orientations.
- Rotation: 90-degree rotations (both clockwise and anticlockwise) to add variety in image orientation.
- Hue Adjustments: Modifications in hue ranging from -15° to +15° to replicate various lighting conditions.
- Saturation Modifications: Modifications in saturation ranging from -25 to +25 to compensate for color discrepancies in the photos.
- Brightness Modifications: Brightness alterations ranging from -15 to +15 to replicate exposure discrepancies.
- Blurring: A range of -10 to +10, with a maximum of 0.5 pixels, to replicate noise and indistinct images.
- Noise Addition: A maximum of 0.1% of pixels in each image were modified by introducing random noise to replicate low-quality imaging settings.

4. RCNN Segmentation Model

The primary segmentation task was performed via an RCNN (Region-based Convolutional Neural Network) model. RCNNs have significant efficacy in object detection and segmentation tasks; in this study, the model was modified to accurately recognize and segment the femoral head. The principal elements of the RCNN model pertinent to this investigation encompass:

- The Region Proposal Network (RPN) inside the RCNN architecture detects regional proposals in pictures, concentrating on areas that are probable to encompass the femoral head.
- Feature Extraction: Convolutional layers extract advanced features from images, which are subsequently utilized for segmentation.
- The RCNN employs bounding box regression to enhance the localization of the femoral head and utilizes mask prediction to produce precise segmentation of the femoral head region.
- The model was refined for this medical application by transfer learning from a pre-trained RCNN model, allowing the network to acquire both general characteristics from extensive datasets and femoral head-specific features from the orthopedic imaging dataset.

5. Instruction

The model underwent training with the curated dataset and augmentation methods for a total of 300 epochs. A batch size of 32 was utilized, and the Adam optimizer was implemented to enhance convergence speed. The learning rate was constantly modified through a learning rate scheduler to enhance performance and guarantee consistent advancements in segmentation accuracy.

6. Assessment Metrics

The performance of the RCNN model was assessed using the following segmentation metrics:

- Intersection over Union (IoU): This metric quantifies the overlap between the predicted segmentation mask and the ground truth mask, offering a definitive measure of segmentation accuracy.
- Dice Coefficient: The Dice coefficient served as a similarity parameter to assess the alignment of the anticipated segmentation mask with the real boundaries of the femoral head.
- Precision and Recall: These measures were employed to evaluate the model's capability to accurately identify the femoral head (precision) and reduce false negatives (recall).

The validation dataset was utilized to calculate these measures, guaranteeing a thorough assessment of the model's correctness and resilience.

Post-processing

Post-processing techniques were utilized to enhance the segmentation results. Morphological techniques, including erosion and dilation, were utilized to

International Journal of Trendy Research in Engineering and Technology
Volume 9 Issue 1 February 2025
ISSN No. 2582-0958

eliminate noise and improve the definition of segmentation borders. This post-processing guaranteed that the resultant segmentation masks were appropriate for clinical applications and further analysis.

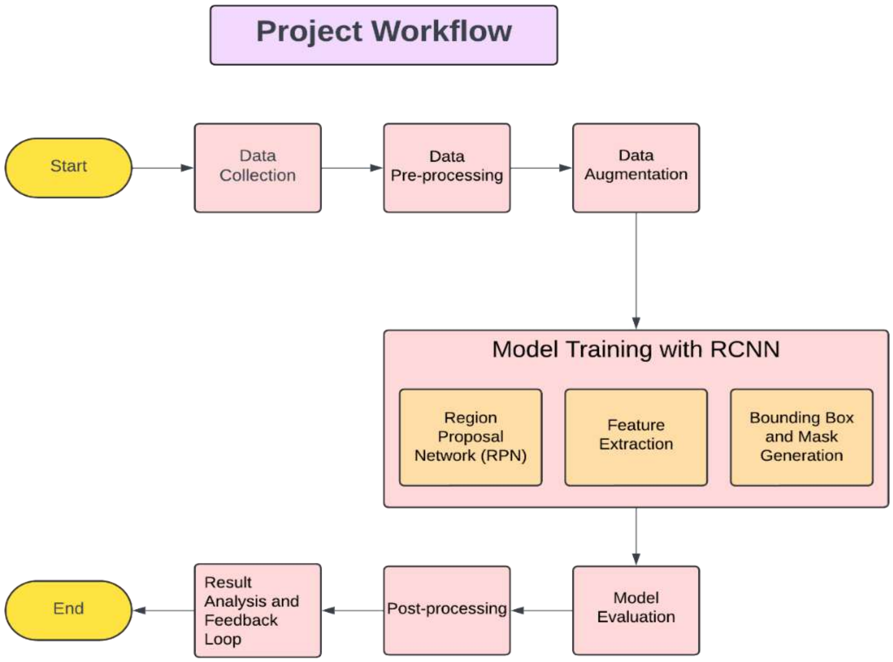

Fig 1.0: Model Workflow

This figure explains the working of this model and step by step operations of it.

IV.RESULTS

The suggested femoral head segmentation model, employing an RCNN architecture, exhibited outstanding performance across all evaluation metrics. The outcomes derived from the model assessment on the validation set are as follows:

- The model attained a Mean Average Precision (mAP) of 99.5%. This signifies that the anticipated segmentation masks closely aligned with the ground truth masks, exhibiting minimal inconsistencies in the identification of the femoral head.
- The model achieved an accuracy of 99.7%. The elevated precision score indicates that the model produced little false positive predictions, successfully identifying femoral head areas in nearly all instances.
- The model achieved an impeccable recall score of 100%. This indicates that the model successfully identified all instances of the femoral head within the sample, with no omissions in segmentation.

Fig 2.0: Results

This figure displays the results of our model and it’s performance

The results confirm the efficacy of the pre-processing and data augmentation strategies, as well as the RCNN model's capacity to generalize effectively across diverse visual situations. The elevated precision and recall metrics demonstrate that the model effectively identifies all pertinent femoral head structures while maintaining a low incidence of false positives, rendering it very dependable for clinical applications in orthopedic imaging.

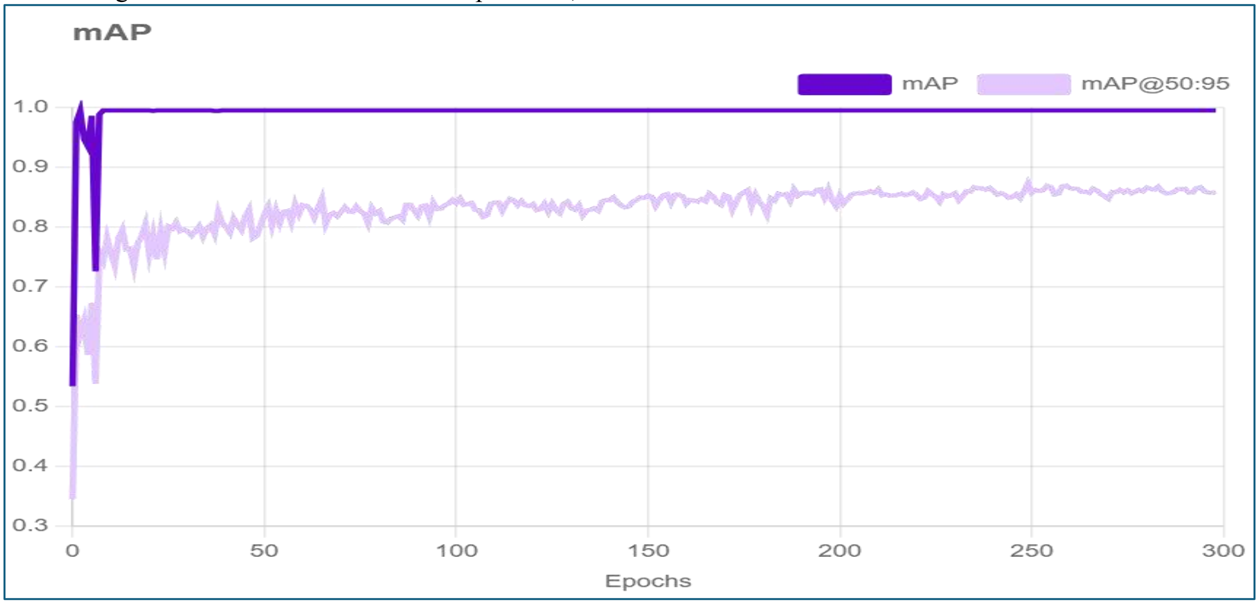

Fig 3.0: Epochs

This Figure provides the details about the model's performance across epochs

The outstanding mAP score underscores the model's efficacy in accurately delineating the femoral head's limits, rendering this methodology an invaluable asset for automated diagnostic systems and pre-surgical planning in orthopedics.

Fig 4.0: Object Loss

These figures provide us detail about the loses that occur during the training

V.CONSTRAINTS

Although the study's results are encouraging, specific limitations must be addressed to improve the model's resilience and usefulness in real-world clinical contexts. A primary restriction is the dataset's size and diversity. Despite the utilization of 2,035 images for training and evaluation, this dataset may not comprehensively represent the extensive diversity inherent in real-world orthopedic imaging. Disparities in imaging modalities, patient demographics, disease states, and equipment quality may lead to inconsistencies when the model is utilized across various clinical settings.

The dataset included in this study predominantly comprises photos from regulated sources and may not entirely reflect the intricacies present in standard clinical practice. Factors include disparate resolutions, picture noise, and divergent imaging techniques among hospitals and imaging centers can affect the model's effectiveness. The constrained size and scope of this dataset may result in overfitting or diminished generalization in practical applications, particularly with rare illnesses, anatomical anomalies, or subpar imaging data.

To mitigate these problems and ensure the model conforms to contemporary clinical standards, future efforts must prioritize the augmentation of the dataset. Integrating a broader and more varied collection of images—encompassing numerous imaging modalities, patient states, and clinical settings—would improve the model's capacity for effective generalization. Furthermore, incorporating additional pathological instances, photos with diverse noise levels, and images captured from various perspectives and orientations will replicate the complexity encountered in real-world applications. This would enhance the model's resilience and reliability across varied environments.

Moreover, although the existing RCNN model exhibits exceptional performance, subsequent versions could enhance their efficacy by incorporating more sophisticated features, such as multi-scale feature extraction or attention methods. These enhancements may enhance the model's concentrate on pivotal areas of interest and augment its capacity to manage complex scenarios. Integrating contemporary necessities such as real-time processing capabilities, interoperability with diverse medical imaging systems, and adaptability to distinct anatomical locations would be crucial for rendering the model a viable instrument in clinical applications.

In conclusion, although the present findings are promising, expanding the dataset and enhancing the model's capabilities will be essential for its compatibility with the varied and intricate requirements of real-world orthopedic imaging.

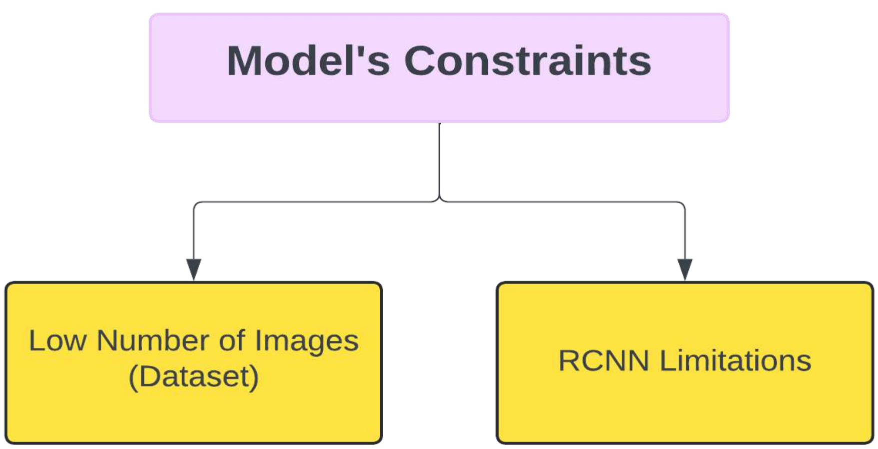

Fig 5.0: Constraints

This figure explains the constraints of our project

VI.FUTURE SCOPE

Future prospects exist for improving the existing femoral head segmentation model. A key avenue for future research is the incorporation of multimodal vision techniques that amalgamate data from many imaging modalities or disparate databases. Utilizing a multi-model framework, the model could integrate not only the existing single-source dataset but also supplementary medical imaging modalities, including MRI, CT scans, and X-rays. This would enhance the model's ability to generalize and yield more thorough segmentation results, which is essential in clinical decision-making.

Furthermore, augmenting the quantity of training images is crucial for developing a more resilient and dependable system. An expanded dataset would more adequately encompass the variety of real-world medical images, incorporating varied patient demographics, imaging modalities, and clinical states. This would mitigate the risk of overfitting and enhance the model's capacity to generalize across various clinical environments. Expanding the dataset and integrating multi-modal inputs could enhance segmentation performance, according to contemporary healthcare norms and ensuring applicability in real-time clinical settings.

Furthermore, further efforts must guarantee that the system adheres to all health data standards, encompassing privacy and legal mandates, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union. Adhering to these standards is essential for the integration of the system into clinical settings and for its compliance with contemporary health demands. Facilitating interoperability with current medical infrastructure and adherence to health information exchange norms will be crucial for its extensive implementation.

In conclusion, augmenting the model via multimodal vision methodologies, enlarging the dataset, and adhering to healthcare norms will substantially improve the system's efficacy. These developments will facilitate the development of a cutting-edge, scalable solution that can address the evolving and intricate requirements of contemporary healthcare.

VI.CONCLUSION

This study introduces a highly efficient method for femoral head segmentation utilising an RCNN model, achieving outstanding results with a mean Average Precision (mAP) of 99.5%, precision of 99.7%, and recall of 100%. The implemented pre-processing and augmentation procedures, combined with the rigorous training of the RCNN model, have facilitated the system's precise delineation of the femoral head across a varied array of pictures. These findings underscore the capability of deep learning models to automate intricate medical imaging activities, thereby enhancing the efficiency and accuracy of diagnostic procedures in orthopedics.

The restricted dataset presents a difficulty for extensive generalization across various imaging settings and real-world situations. Future research should concentrate on augmenting the dataset and integrating multimodal vision methodologies to enhance model robustness and clinical relevance. Moreover, adhering to contemporary healthcare norms will be essential for implementing this technology in clinical environments.

In summary, although the existing model demonstrates exceptional accuracy and performance, enhancements like augmenting dataset diversity and including more advanced functions will facilitate the creation of a more thorough and applicable system for real-world orthopedic imaging. This study establishes a robust basis for next research in medical image segmentation, potentially improving clinical outcomes via the implementation of automated diagnostic instruments.

REFERENCES

[1]. Li et al. "MsgeCNN: Multiscale geometric embedded convolutional neural network for ONFH segmentation and grading" Medical Physics 50(2023)3788-3800 doi:10.1002/mp.16302

[2]. Ebsim, R. *et al.* (2022). Automatic Segmentation of Hip Osteophytes in DXA Scans Using U-Nets. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2022. MICCAI 2022. Lecture Notes in Computer Science, vol 13435. Springer, Cham. doi.10.1007/978-3-031-16443-9_1

[3]. Y. -J. Yun, B. -C. Ahn, M. S. Kavitha and S. -I. Chien, "An Efficient Region Precise Thresholding and Direct Hough Transform in Femur and Femoral Neck Segmentation Using Pelvis CT," in IEEE Access, 8 (2020)110048-110058, doi.10.1109/ACCESS.2020.3001578

[4]. Pemmaraju R, Kim G, Mekki L, Song DY, Lee J. Cascaded cross-attention transformers and convolutional neural networks for multi-organ segmentation in male pelvic computed tomography.Jl.Med.Imaging(Bellingham).11(2024)024009. doi: 10.1117/1.JMI.11.2.024009

[5]. Zhang L, Wells JE, Dessouky R, Gleason A, Chopra R, Chatzinoff Y, Fey NP, Xi Y, Chhabra A. 3D CT segmentation of CAM type femoroacetabular impingement-reliability and relationship of CAM lesion with anthropomorphic features. Br J Radiol. 91(2018)20180371. doi: 10.1259/bjr.20180371

[6]. Dong X, Lei Y, Wang T, Thomas M, Tang L, Curran WJ, Liu T, Yang X"Automatic multiorgan segmentation in thorax CT images using U-net-GAN",Medical-physics,46 (2019) 2157- 2168. doi.10.1002/mp.13458

[7]. Ditmer S, Dwenger N, Jensen LN, Kim H, Boel RV, Ghaffari A, Rahbek O. Fully automatic system to detect and segment the proximal femur in pelvic radiographic images for Legg-Calvé-Perthes-disease.Jl.Orthop.Res. 42(2024):1074-1085. doi:10.1002/jor.25761