International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 10 | October - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Canine Vertebral Column Segmentation with UNet Models: A Positive Approach to Radiographic Image Analysis

Rajneesh Kumar[1], S. D. Samantaray[1], Kavita[2], Arup Kumar Das[2]

[1]Department of Computer Engineering, College of Technology, GBPUA&T, Pantnagar, 263145, India
[2]Department of Surgery and Radiology, College of Veterinary Sciences, GBPUA&T, Pantnagar, 263145, India

---***---

Abstract – Degenerative diseases of the vertebral column, such as spondylosis, are not uncommon in dogs. Detecting these diseases can be challenging and is often overlooked due to lack of effective treatments. The intricacy of medical images poses difficulties in disease identification. Our approach focuses on segmenting the vertebral column from X-ray images of dogs, facilitating targeted disease detection and improving accuracy. Conventional segmentation methods are ineffective for medical images due to presence of occlusions, and pixel classification methods require the laborious creation of masks. By training our UNet models, we successfully achieved precise segmentation of the vertebral column from radiographic images, significantly reducing the time required for generating ground truth masks. Our UNet models effectively enhanced the isolation of the vertebral column from X-ray images, thereby optimizing the accuracy of degenerative disease detection methods for the vertebral column.

Key Words: segmentation, medical imaging, UNet, radiographs, computer vision, canine

1.INTRODUCTION

Medical imaging plays a crucial role in the field of medicine by generating a significant amount of data. With the advent of computer technology, the ability to load onto computers has led to the emergence of automatic analysis techniques. Deep Learning in the field of medical imaging can have long-term significant effects on individuals [1]. Few areas such as radiology, pathology, ophthalmology, and dermatology have received more attention in the deep learning field due to the pattern recognition nature of the diagnostic tasks in the fields [2]. Proper analysis of these medical images can lead to better analysis and understanding, helping in improved diagnostic capabilities and prognostic evaluation.

Tasks related to medical imaging include registration, classification, localization, detection, and segmentation.

Segmentation is considered one of the most challenging tasks in medical imaging. Segmentation tasks can be divided into three main categories: region-based, boundary-following, and pixel classification methods based on the primary mode of operation described in Table 1 [3], [4], [5].

Table -1: Segmentation Methods Based on the primary mode of Operations

Method	Criteria	Examples
Region-Based	Intensity, texture, Statistical properties	Thresholding, Clustering
Boundary- Following	Boundaries and Contours	Canny edge detection, snakes, level-sets
Pixel Classification	Pixel level association	Semantic and instance segmentation

Inhomogeneity, low contrast, imaging noise, and occlusions in medical images complicate the application of region-based or boundary-following methods on our task [6]. The pixel classification method best suits the underlying task.

Diseases of the vertebral column such as Intervertebral Osteochondrosis, spondylosis deformans, uncovertebral arthrosis, osteoarthritis, and Diffuse Idiopathic Skeletal Hyperostosis (DISH), can be identified and classified using radiographic images [7], [8].

The goal of segmentation in medical images is to extract anatomical parts or organs from an image [9]. Segmentation of the vertebral column on radiographic images is crucial for classifying and detecting diseases of the vertebral column.

By segmentation of vertebral column, researchers can focus on the region of interest, allowing for more accurate measurements and assessments. This will improve diagnostic ability and treatment planning.

Neural networks, particularly Convolutional Neural Networks (CNN), have been widely used for tasks related to imaging. The complexity of medical imaging makes it difficult to manually extract features from the images. Convolutions are capable of automatic feature extraction, which makes them suitable for medical imaging tasks including segmentation [1].

UNet is one of the best algorithms for medical image segmentation. UNet uses convolution for feature extraction in local and global contexts. The scarcity of annotated medical images makes UNet a very good model for segmentation because it can perform well on very little data [10]. UNet++ (Nested UNet) [11] and Unet3+ [12] are a few modifications made to the existing UNet model, giving state-of-the-art results. Transformers have been widely used in Natural Language Processing (NLP) tasks. UNet Transformer (UNETR) combines UNet with transformers for segmentation purposes [13].

The segmentation of the vertebral column is a crucial step in the detection and classification of these diseases. Therefore, accurate segmentation is necessary for further analysis. Once accurate segmentation is performed, classification algorithms can be applied to study and differential diagnosis.

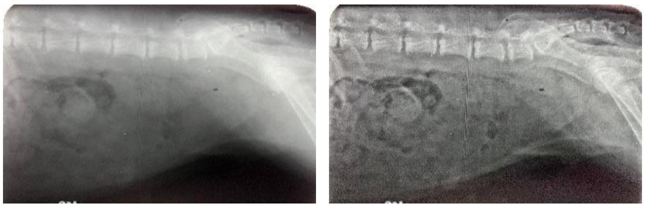
2. Methodology

This study is based on pixel classification of radiographic images which is a supervised learning problem, where pixels related to the vertebral column and the background are classified. Training a model for pixel classification requires a dataset comprising input radiographic images along with their corresponding masks. The input images and their corresponding masks are then fed as input to UNet models for training which are then evaluated on various metrics.

2.1. Dataset

The dataset was collected from the radiological image (analog) repository of the Department of Surgery and Radiology, College of Veterinary Sciences, GBPUA&T, Pantnagar. Over 2000 X-ray images of dogs between the years 2002-2023 were examined. The main objective behind the examination of the X-rays was to find the presence of spondylosis in the vertebral column of dogs. Based on the objective 410 images were carefully chosen consisting of X-rays with the presence of spondylosis in them and some normal X-rays for comparison.

The analog X-rays were converted to digital images frame by frame with the help of an iPhone13 camera by placing the camera parallel to the illuminator and the film. A good amount of contrast in the image was lost in the process of digitalization as shown in Fig -1: (a). The contrast was corrected by using a Contrast Limited Adaptive Histogram Equalization (CLAHE) method as shown in Fig -1: (b), which can be accessed using a tool called ImageJ, created by the National Institute of Health (NIH) for processing medical images.



(a) (b)

Fig -1: (a) Image taken using the iPhone13 camera, (b) Image after applying CLAHE to enhance image contrast.

The ground truth masks were prepared by a radiologist in the department. LabKit extension of the ImageJ tool was used for creating ground truth masks for the X-ray images. As shown in Fig -2 the area belonging to the vertebral column is marked in LabKit and the remaining area is considered the background. The annotation is then saved in .jpeg format as the ground truth mask

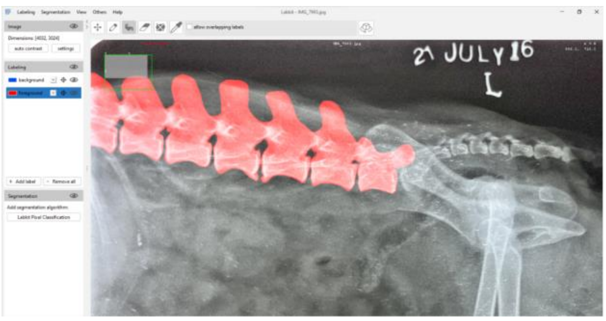


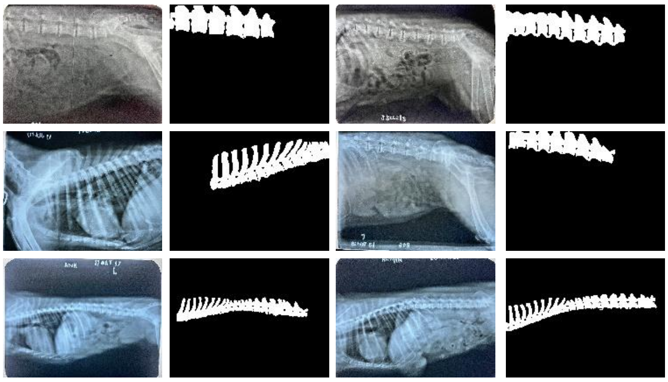
Fig -2: An annotated X-ray image opened on Labkit.

A sample of data along with the ground truth masks is given in Figure 3.

2.2. Segmentation Models

1. UNet: UNet is a fully convolutional model developed for biomedical image segmentation. It consists of a contracting path called the encoder, and an expansive path called the decoder. Because the UNet architecture requires

Fig -3: Sample X-ray images and their corresponding masks.



less sample data than any other architecture, it is suitable for medical image segmentation [10].

2. UNet ++: Also known as the nested UNet model, improves upon UNet by incorporating nested skip pathways. The dense skip connections in UNet++ preserve spatial localization and provide precise segmentation boundaries [11].

3. UNet3+: It is an extension of the original UNet architecture with the introduction of the concept of multi-skip pathways in the contracting path. This allows the model to use-reuse features from different levels through multi-skip connections [12].

4. UNet Transformer (UNETR): UNETR is an approach that combines the strength of both the UNet and Transformers. The self-attention mechanism of transformers allows it to capture the global context whereas UNet captures local information. Combined, they capture both local and global features to improve segmentation [13].

2.3. Evaluation Metrics

1. Loss: We use Total Loss which is a combination of Dice Loss and Categorical Focal Loss.

$$Loss = Dice\ Loss + (1 \times Categorical\ Focal\ Loss) \quad (1)$$

2. Jaccard Index: Jaccard measures how well the overlap between the predicted and ground truth masks.

$$Jaccard = \frac{TP}{(TP+FN+FP)} \quad (2)$$

3. Dice Score: Dice Score considers the ratio of the intersection area to the total area covered by both masks.

$$Dice = \frac{2\times TP}{(2\times TP+FN+FP)} \quad (3)$$

2.4. Training Process

The dataset was divided into train, test and validation set, where 70 percent of the data was used for training and 15 percent for validation and another 15 percent of the data for testing.

All models were trained using the Adam optimizer with a batch size of eight and a learning rate determined by the optimizer itself. The training process involved iterating the dataset for 100 epochs.

A brief methodology for the training process is given in Fig -4. The images along with the corresponding ground truth masks were split into train, test and validation sets. UNet segmentation models are applied to the dataset and their performance is measured using the metrics given.

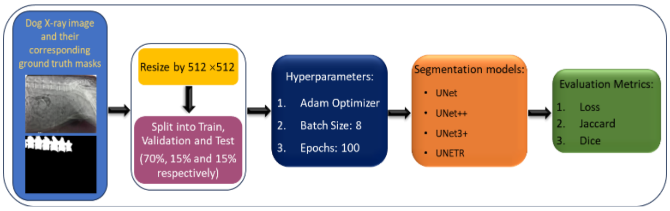


Fig -4: Training Process for UNet models

2.5. Results and Discussions

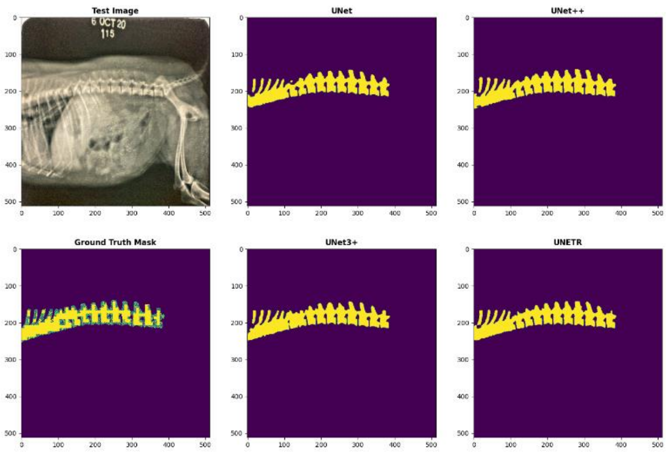
The four UNet models were evaluated on three key performance metrics namely Loss, Jaccard Index, and Dice Score. These models are essential for assessing the effectiveness of models in the context of segmentation task.

Table -2 displays the performance metrics for the four UNet models, on train, valid and test datasets. The UNETR model shows best result on the test dataset.

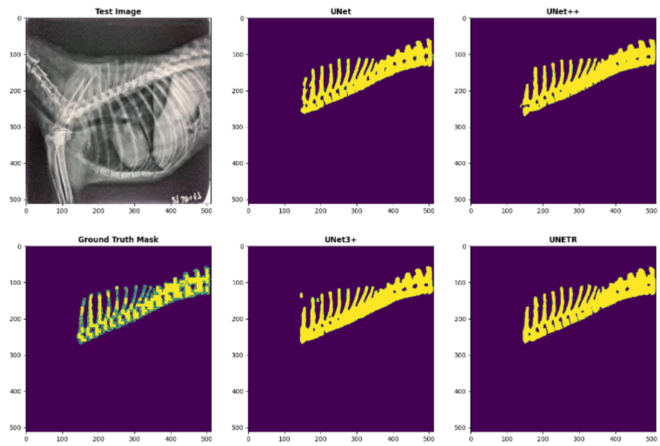
Table -2: Performance of UNet models on various evaluation metrics

Metrics	Dataset	UNet	UNet++	UNet3+	UNETR
Loss	Train	0.5367	0.5320	0.5324	0.5332
	Valid	0.5649	0.5628	0.5625	0.5578
	Test	0.5582	0.5600	0.5594	0.5526
Jaccard	Train	0.9592	0.9632	0.9630	0.9623
	Valid	0.9420	0.9426	0.9435	0.9452
	Test	0.9462	0.9448	0.9459	0.9497
Dice	Train	0.9792	0.9813	0.9812	0.9808
	Valid	0.9701	0.9704	0.9709	0.9718
	Test	0.9722	0.9714	0.9721	0.9741

Visual inspection reveals that the UNETR model exhibits a slightly higher level of accuracy compared to other UNet models as shown in Fig 5- (a) and (b).



(a)



(b)

Fig -5: Segmentation result on test images of various UNet models

The plotted confusion matrix (Fig - 6) demonstrates the classification performance of the UNet models in distinguishing between pixels representing the vertebral column and those constituting the background in the image. Notably, the UNet++ model achieved the highest classification accuracy, correctly identifying 91.25 percent of the pixels associated with the vertebral column. Following closely, the UNETR model demonstrated a classification accuracy of 90.13 percent.

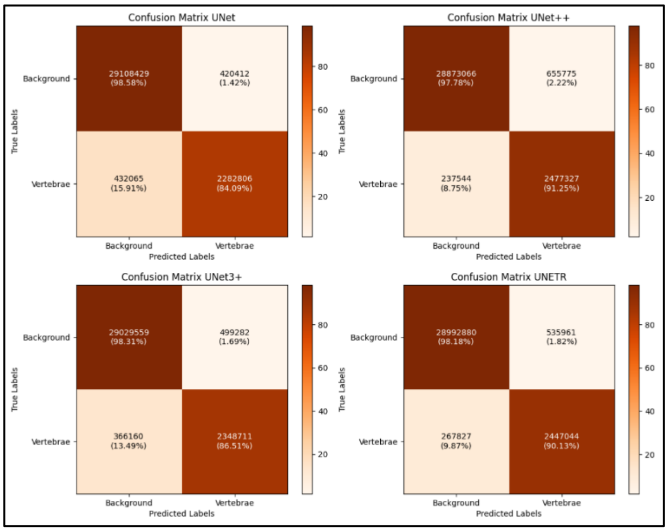


Fig -6: Confusion matrix showing the pixels in the image correctly classified into background and the vertebral column.

3. CONCLUSIONS

The segmentation of the region of interest has promising potential in enhancing disease detection from medical images. Through our training of UNet models, we achieved highly accurate segmentation of the vertebral column from X-ray images. This will significantly contribute to the development of robust disease detection models specifically tailored for the vertebral column of dogs.

REFERENCES

1. Moorthy, J., and Gandhi, U.D.: A Survey on Medical Image Segmentation Based on Deep Learning Techniques. Big Data Cogn. Comput. (2022), 6(4), 117. https://doi.org/10.3390/bdcc6040117
2. Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., Liu, Y., Topol, E., Dean, J., Socher, R.: Deep learning-enabled medical computer vision. Digit. Med. (2021), 4(1), 5. https://doi.org/10.1038/s41746-020-00376-2
3. Hesamian, M.H., Jia, W., He, X., Kennedy, P.: Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. J. Digit. Imaging. (2019), 32(4), 582–596. https://doi.org/10.1007/s10278-019-00227-x
4. Suganyadevi, S., Seethalakshmi, V., Balasamy, K.: A review on deep learning in medical image analysis. Int. J. Multimed. Info. Retr. (2022), 11(1), 19–38. https://doi.org/10.1007/s13735-021-00218-1
5. Withey, D.J., Koles,Z.J.: A Review of Medical Image Segmentation: Methods and Available Software, Int. J. Bioelectromagn. (2008), pp. 125-148.
6. Elnakib, A., Gimel'farb, G., Suri, J.S., El-Baz, A.: Medical Image Segmentation: A Brief Survey, Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies: Volume two (2011), pp.1–39. https://doi.org/10.1007/978-1-4419- 8204-9_1
7. Kranenburg, H.J.C., Voorhout, G., Grinwis, G.C., Hazewinkel, H.A., Meij, B.P.: Diffuse idiopathic skeletal hyperostosis (DISH) and spondylosis deformans in pure bred dogs: a retrospective radiographic study. Vet. J. (2011), 190(2), e84–e90. https://doi.org/10.1016/j.tvjl.2011.04.008
8. Resnick, D.: Degenerative diseases of the vertebral column. Rad. (1985), 156(1), 3-14.
9. Elyan, E., Vuttipittayamongkol, P., Johnston, P., Martin, K., McPherson, K., Jayne, C., Sarker, M.K.: Computer vision and machine learning for medical image analysis: recent advances, challenges, and way forward. Artif. Intell. Med. (2022), 2(1). https://doi.org/10.20517/ais.2021.15
10. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In Med. Image Comput. Comput. Assist. Interv. (MICCAI) (2015), pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
11. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In Deep Learn. Med. Image. Anal. Multimodal Learn. Clin. Decis. Support (2018), pp. 3-11. https://doi.org/10.1007/978-3-030-00889-5_1
12. Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.W., Wu, J.: UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation. In ICASSP 2020 - 2020 IEEE Int. Conf. on Acoust., Speech and Signal Process (ICASSP) (2020), pp. 1055–1059. https://doi.org/10.1109/ICASSP40776.2020.9053405
13. Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H.R., Xu, D.: UNETR: Transformers for 3D Medical Image Segmentation. In 2022 IEEE/CVF Winter Conf. App. of Comput. Vis. (WACV) (2022), Waikoloa, HI, USA, pp. 1748–1758. https://doi.org/10.1109/WACV51458.2022.00181