Pakistan Veterinary Journal

ISSN: 0253-8318 (PRINT), 2074-7764 (ONLINE)
DOI: 10.29261/pakvetj/2025.138

RESEARCH ARTICLE

Deep Learning Model for CT-based Adrenal Gland Volume Determination and Normal Reference Definition in Dogs

So-Hyeon Park[1†], Hyunwoo Cho[2†], Kichang Lee[1] and Hakyoung Yoon[1*]

[1]Department of Veterinary Medical Imaging, College of Veterinary Medicine, Jeonbuk National University, Iksan-si, Republic of Korea; [2]Department of Electronic Engineering, Sogang University, Seoul, Republic of Korea
*Corresponding author: knighttt7240@gmail.com
†These authors contributed equally to this work and share first authorship

ARTICLE HISTORY (24-865)

Received: December 31, 2024
Revised: March 12, 2025
Accepted: March 14, 2025
Published online: March 28, 2025

Key words:
Artificial intelligence Model
Adrenal volume
Adrenal segmentation
Canine
Computed tomography

A B S T R A C T

Adrenal gland size is linked to its function, disease status, and tumor malignancy, if any, making accurate measurement of its size essential. However, measuring adrenal gland length is prone to errors, and volume is a reliable indicator of its size. Manual volume measurement is time-consuming and is usually inaccurate. Therefore, this study aimed to develop an artificial intelligence (AI) model for direct adrenal gland volume measurement in computed tomography (CT) images. Post-contrast CT images of 250 dogs were segmented. Of these, 200 scans were randomly selected for training and 50 for validation. A deep learning model, based on Swin-Transformers and several processing techniques, was developed. Computed tomography images of 239 dogs were used for normal reference definition, with adrenal gland volume was determined on the basis of the absence of adrenal gland lesions supported by clinical and laboratory data. The mean (±SD) Dice Similarity Coefficient (DSC) of adrenal gland segmentation was 0.885±0.075, which is slightly lower than other abdominal organs of dogs, most probably due to the small size, varied shapes, and overlapping with surrounding tissue. Agreement analysis between manual voxel counts and the AI model showed an interclass correlation coefficient of 0.957 ($P<0.001$). Adrenal gland volume correlated positively with body weight (BW; $r=0.821$, $P<0.001$) and age ($r=0.147$, $P<0.05$), and negatively with body condition score (BCS; $r=-0.233$, $P<0.001$). The relationship was represented by the regression equation: adrenal volume=-0.51xBCS+0.033×BW+0.015×age+0.373 ($\mathrm{adj}R^2=0.72$, $P<0.001$). No correlation was found between adrenal gland volume and sex of dogs. In conclusion, an AI model was developed to directly measure adrenal gland volume from CT images of dogs, which would potentially aid in adrenal disease screening.

To Cite This Article: Park SH, Cho H, Lee K and Yoon H, 2025. Deep Learning Model for CT-based Adrenal Gland Volume Determination and Normal Reference Definition in Dogs. Pak Vet J, 45(1): 320-327. http://dx.doi.org/10.29261/pakvetj/2025.138

INTRODUCTION

The size of the adrenal gland, together with hormone tests, is an important indicator of adrenal gland diseases and is associated with the functionality and malignancy of adrenal tumors (Besso *et al.*, 1997). A small adrenal gland may suggest Addison's disease, while an enlarged adrenal gland may raise the suspicion of pituitary-dependent hypercortisolism or presence of tumors. Adrenal gland diseases are relatively common in dogs, with most adrenal tumors being malignant (de Bruin *et al.*, 2009).

The adrenal glands exhibit diverse shapes, typically peanut-shaped on the left side and V-shaped on the right. Measuring only the thickness or length of the gland may lead to inaccuracies, making volumetric measurements necessary for precise assessment of its enlargement.

Imaging examination is crucial for the early diagnosis and assessment of the status of abnormalities of the gland. Ultrasonography, a non-invasive technique, allows for the evaluation of adrenal glands without anesthesia. However, it has limitations in visualizing the entire adrenal glands due to surrounding tissues, and obtaining a midsagittal view may result in overestimating adrenal gland size (Swepson *et al.*, 2022). Moreover, ultrasound measurements can vary depending on operator's skill and quality of the image.

Computed tomography (CT) provides simultaneous evaluation of both adrenal glands, including attenuation values that indicate tissue density and contrast enhancement. It offers the advantage of repeatable adrenal gland measurements and volume calculation. Additionally, previous research has demonstrated a strong correlation between CT findings and histopathology of the adrenal gland in dogs (Gregori *et al.*, 2015). However, direct volume measurement can be time-consuming, and there is low intra- and inter-observer variability in measuring the length and diameter of the caudal pole of the adrenal gland (Perfetti *et al.*, 2021). Incidentally, enlarged adrenal glands may go unnoticed by radiologists on CT images due to their small volume (Kim *et al.*, 2023). Investigations on deep learning models seem necessary to improve clinical decision-making by allowing early detection of adrenal diseases through fast and accurate measurements, ensuring consistent monitoring with reliable and reproducible results, aiding in tracking disease progression and treatment effectiveness, improving diagnostic accuracy by minimizing human error, and supporting personalized treatment to enable tailored care for affected animals. A deep learning model is a complex computational model that processes large amounts of information. It can recognize complex pictures, text, sounds, and other data patterns to produce accurate insights and predictions.

Recent studies in humans have utilized deep learning model to recognize adrenal glands in CT images and directly calculate their volume (Kim *et al.*, 2023; Robinson-Weiss *et al.*, 2023; Li *et al.*, 2024). Compared to humans, animals have diverse species with wide organ size variation, even within the same species, which makes it challenging to train veterinarians in deep learning models. Therefore, research on deep learning models related to adrenal gland volume estimation in veterinary medicine is limited. While earlier studies suggested no differences in adrenal volume among various body weight classes (Bertolini *et al.*, 2006), recent studies have shown a correlation between body weight and adrenal volume in dogs (Bertolini *et al.*, 2008; Swepson *et al.*, 2022). Despite this, there is a lack of deep learning research in veterinary medicine for direct measurement of adrenal gland volume in CT images, as existing studies rely on manually delineating regions of interest on all sections of CT scans (Bertolini *et al.*, 2006; Swepson *et al.*, 2022).

This study aimed to develop an AI model that can measure adrenal gland volume quickly and accurately in order to facilitate the early diagnosis of adrenal disorders, which are relatively common and often discovered incidentally in dogs. Attempts were also made to determine reference ranges for adrenal gland volume in clinically healthy dogs based on body weight. Additionally, correlations of body weight (BW), body condition score (BCS), sex, and age of dogs with adrenal gland volume were also explored. It was hypothesized that a deep learning model can accurately measure adrenal gland volume from CT images and that adrenal gland volume would show reasonable correlations with body weight, body condition score (BCS), sex, and age in dogs.

MATERIALS AND METHODS

Deep learning model development

CT dataset, manual segmentation and pre-processing: The CT scans from 250 dogs (during the period from 2020 to 2023) across multiple centers were used with the dataset split 75:25, resulting in 200 dogs for training and the remaining 50 for validation. In total, 363,058 slices were analyzed. Dogs with history of increased appetite, higher water consumption, increased urination, or a pot-bellied appearance were excluded. Similarly, animals with an isolated increase in ALP among liver enzyme activities or a Na:K ratio lower than 27:1 were also excluded. Only cases with no evidence of adrenal gland enlargement, as confirmed by three veterinary imaging residents, were included.

On each slice of post-contrast axial CT images (Fig. 1A), adrenal segmentation was performed by drawing the region of interest along the boundaries of both adrenal glands (Fig. 1B), using the MediLabel software, excluding the phrenico-abdominal vein and caudal vena cava adjacent to the adrenal glands. The accuracy of the region of interest (ROI) delineation was verified using the reconstructed dorsal view (dorsal views before and after segmentation are shown in Fig. 1D and Fig. 1E, respectively) and the 3D-rendered view for the left (Fig. 1C) and the right adrenal gland (Fig. 1F). For deep learning model training, scans were pre-processed with voxel interpolation and intensity normalization. Unit voxel size was chosen as $x = 0.75\text{mm},\ y = 0.75\text{mm},\ z = 0.75\text{mm}$. Hounsfield unit (HU) values were clipped to a range of -50 to 500 and normalized to 0–1 to enhance contrast.

Model architecture and implementation: A Swin-transformer-based model (SwinUNETR-V2) was used for adrenal gland segmentation, which is a state-of-the-art model for CT segmentation (He *et al.*, 2023). The input volume patch size was $x = 128\ \text{voxels},\ y = 128\ \text{voxels},\ z = 128$ voxels with one input and two output channel dimensions (background and adrenal glands). Intermediate feature dimension was set to 96, and block depths included four stages with one Swin-transformer layer per stage. Transformer heads were set to 3, 6, 12, and 24 for each stage, respectively. Instance normalization was applied to each layer, as described earlier (Huang and Belongie, 2017). The residual convolutional block from SwinUNETR-V2 was incorporated, as described previously (He *et al.*, 2023). The overall architecture of the model is shown in Fig. 2.

Deep learning model was implemented using the PyTorch and MONAI frameworks, as described earlier (Cardoso *et al.*, 2022). The loss function combined dice loss and focal loss to address class imbalance caused by the small size of the adrenal gland (Ma *et al.*, 2021). Basically, the dice loss function is widely adopted in medical image segmentation tasks. However, because the adrenal gland occupies a relatively small portion of the entire volume, the network cannot sufficiently focus on this small region using dice loss alone. Therefore, the focal loss function, a weighted version of cross-entropy, was incorporated to encourage the model to better focus on the small foreground object. The combined loss function was determined by using the following equations:

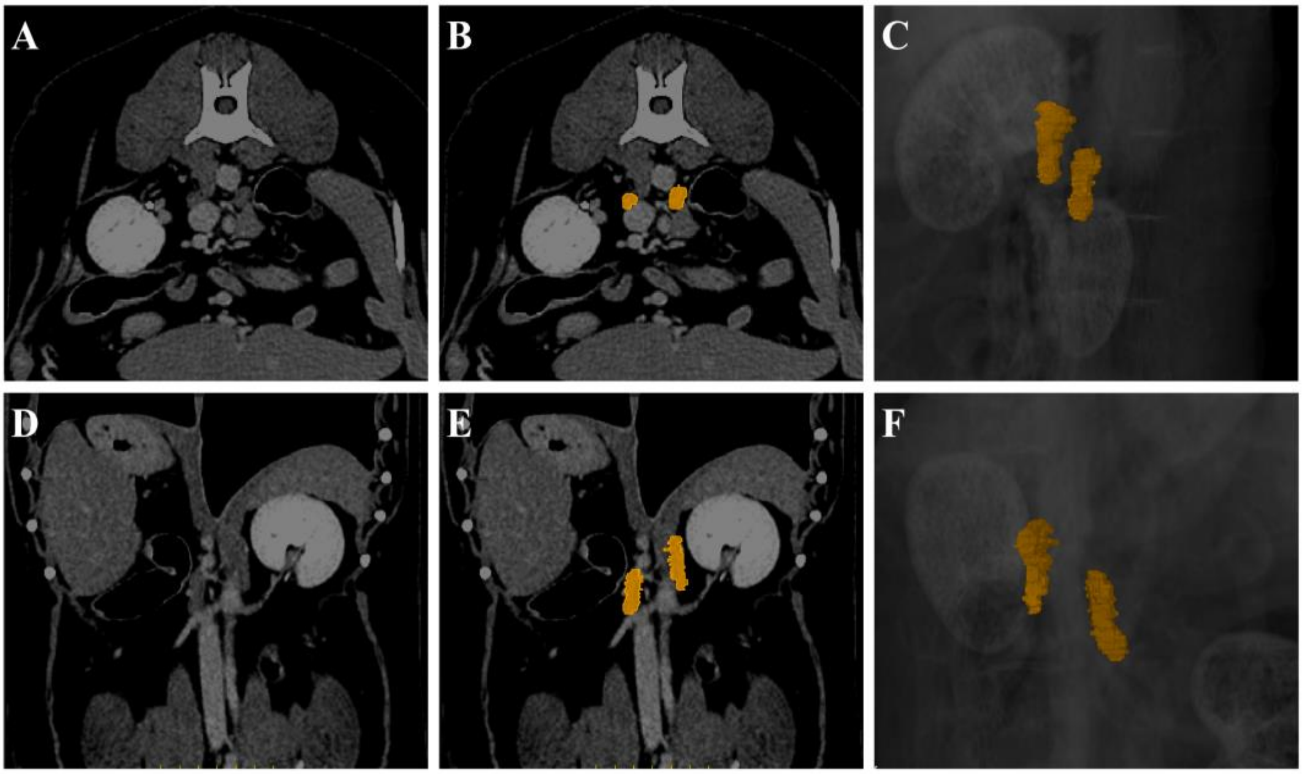

Fig. 1: Examples of the manually segmented adrenal glands. A and B are axial views, with A before segmentation and B after segmentation of adrenal glands shown in yellow color. D and E are dorsal views, with D before segmentation and E after segmentation of adrenal glands shown in yellow color. C and F are 3D rendered images, with C focusing on the left adrenal gland and F focusing on the right adrenal gland.

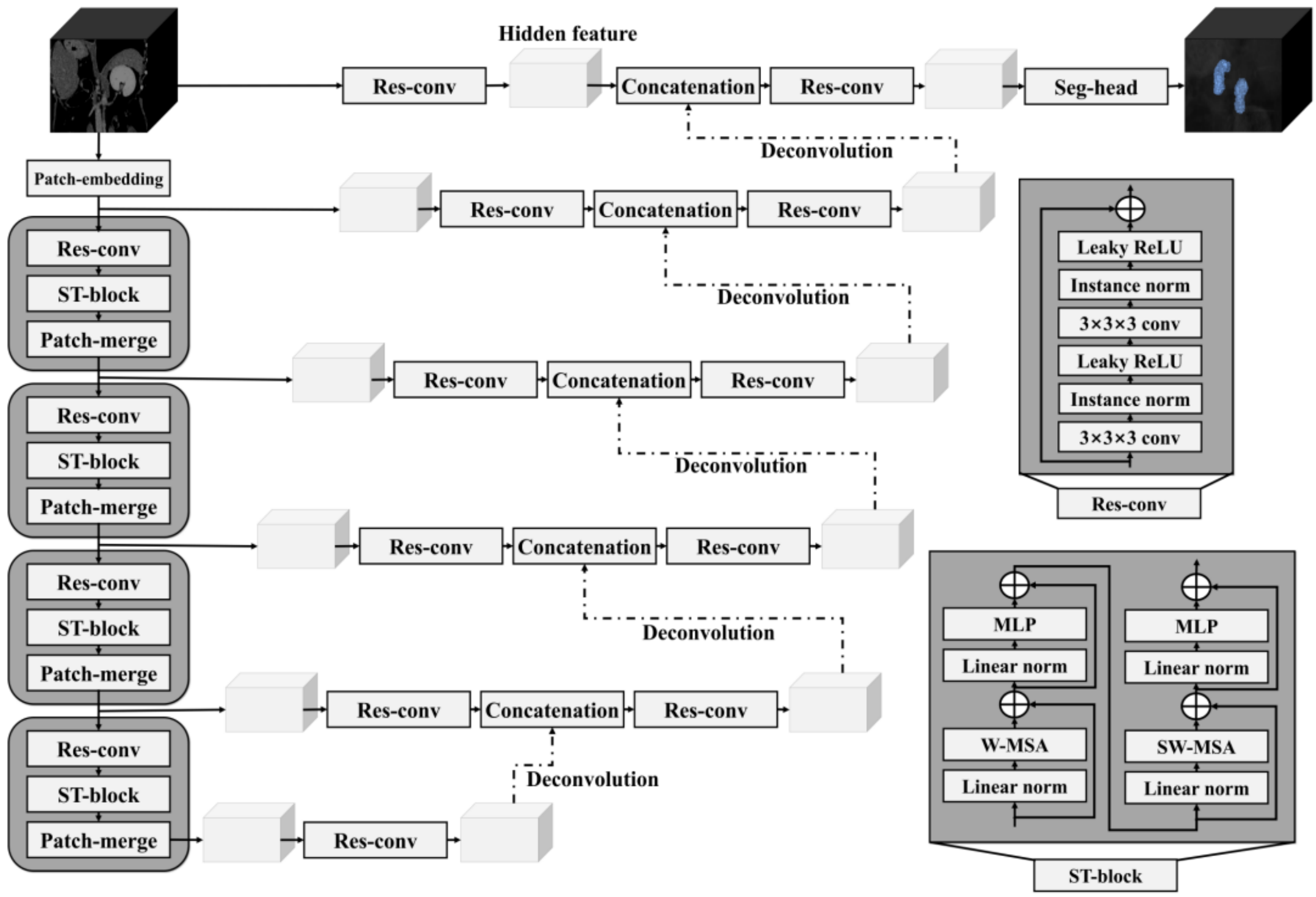

Fig. 2: Graphical representation of the deep learning model. "Res-conv" denotes residual convolutional block, and "ST-block" represents the Swin-Transformer block. The input volume was processed through encoder and decoder structures, and the processed features were finalized into segmentation result on the segmentation head (i.e., Seg-head).

(1) $L_{Focal} = -\sum_{c=1}^{C}\sum_{i=1}^{N} t_{i,c}\left(1 - f(s)_{i,c}\right)^{\gamma}\log\left(f(s)_{i,c}\right)$

(2) $L_{Dice} = 1 - \frac{2\sum_{c=1}^{C}\sum_{i=1}^{N} t_{i,c} f(s)_{i,c}}{\sum_{c=1}^{C}\sum_{i=1}^{N} t_{i,c}^2 + \sum_{c=1}^{C}\sum_{i=1}^{N} f(s)_{i,c}^2 + \epsilon}$

(3) $L = L_{Focal} + L_{Dice}$

where $t_{i,c}$ is ground truth, $f(s)_{i,c}$ is predicted probability of the deep learning model, and γ is focusing parameter to address class imbalance. In this study, γ was emprically set to 2 by experiment. The model was optimized using the AdamW optimizer with a learning rate of $1e^{-4}$ and weight decay of $1e^{-5}$ for 100,000 steps (Loshchilov and Hutter, 2019).

Post-processing and volume measuring: Segmented adrenal volumes were calculated by converting voxel counts into actual volumes using voxel spacing data. The conversion was conducted using the following formula:

$$Volume = N_{AG} \times S_X \times S_Y \times S_Z \quad (4)$$

where N_{AG} is the number of voxels classified into adrenal glands and S_X, S_Y, S_Z are the unit voxel sizes of the preprocessed data (0.75mm). A smoothing process was applied using 3D slicer software to represent realistic gland surfaces.

Model accuracy: The Dice similarity coefficient (DSC) was employed to assess the consistency between manual and automated segmentation. The DSC was calculated by using the following formula:

$$\text{DSC} = \frac{2 * \textit{Area of overlap}}{\textit{Sum of region segmentations}}$$

Moreover, sensitivity, specificity, and mean absolute error (MAE) were also measured and have been described in the "results" section.

Establishment of a normal reference range for adrenal gland volume:

Computed Tomography dataset: Of the original 250 dogs, 239 were included, excluding those with ambiguous boundaries between the liver and the adrenal glands. Dogs were classified into five groups according to their body weight (Table 1) and normal reference range for adrenal gland volume was determined for each body weight group.

Statistical analysis: Data are shown as mean±SD. Kolmogorov-Smirnov test and Shapiro-Wilk tests were applied to assess normal distribution. ANOVA was applied to compare adrenal gland volume according to body weight, while the Kruskal-Wallis test was used to compare adrenal volume according to neutering status of dogs ((intact female, neutered female, intact male, neutered male). T-test was applied for comparison of adrenal volumes between two sexes (male vs female) and two neutering groups (intact vs neutered), while partial correlation and regression analyses were used to determine relationships between BCS, BW, age and adrenal volume. Pearson's correlation coefficients were computed between left and right adrenal volumes. Multiple regression analysis was conducted using BW and age as independent variables, and adrenal gland volume as the dependent variable. To validate the regression model, the residuals were examined for normality and homoscedasticity. Adjusted R^2 was used to evaluate model fit, and variance inflation factors (VIFs) were checked to ensure no multicollinearity among the predictors. The Durbin-Watson statistic was calculated to evaluate the presence of autocorrelation in the residuals. A Durbin-Watson statistic value close to 2 indicated that there were no significant autocorrelations in the residuals, confirming the independence of errors in the regression model. Confidence intervals for the regression coefficients were calculated at a 95% confidence level to assess the precision of the estimates.

Table 1: Adrenal gland volumes (Means ± SD) for class of body weight groups and side effects

Parameter	Body weight level (kg)	Volume (cm^3)	95% reference intervals (cm^3)	Scheffe
BW group** (n=239)	A; ≤ 2.5 (n=32)	0.22±0.10	0.19-0.26	A, B vs C vs D vs E
	B; >2.5-5 (n=80)	0.33±0.11	0.31-0.36	
	C; >5-10 (n=62)	0.48±0.16	0.44-0.53	
	D;>10-20 (n=38)	0.67±0.31	0.57-0.77	
	E; >20-40 (n=27)	1.11±0.36	0.96-1.25	
Side[NS] (n=478)	Left (n=239)	0.52±0.32		
	Right (n=239)	0.49±0.35		

** Significant intergroup differences on ANOVA (P<0.001). [NS]Non-significant intergroup differences on Pearson's correlation coefficients.

RESULTS

Total adrenal gland volume agreement analysis: The statistical comparison of total adrenal volume between the deep learning model and the manual voxel counting method showed high similarity based on the Pearson's correlation coefficients (Fig. 3). The intraclass correlation coefficient (ICC) for perfect agreement between the two methods was 0.957 (95% CI: 0.924-0.975, P<0.001).

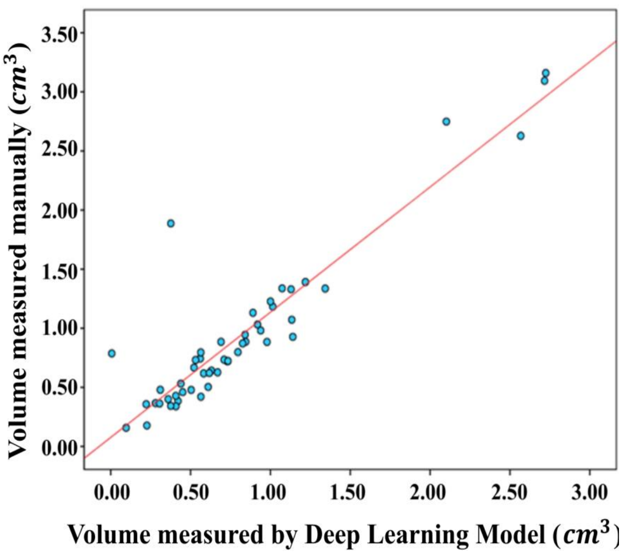

Fig. 3: Pearson correlation analysis between volume measurements from manual method and deep-learning model. The manually estimated volume and the directly measured volume using the deep learning model exhibited a strong Pearson's correlation coefficient (r=0.922, adjR^2=0.851, P<0.001).

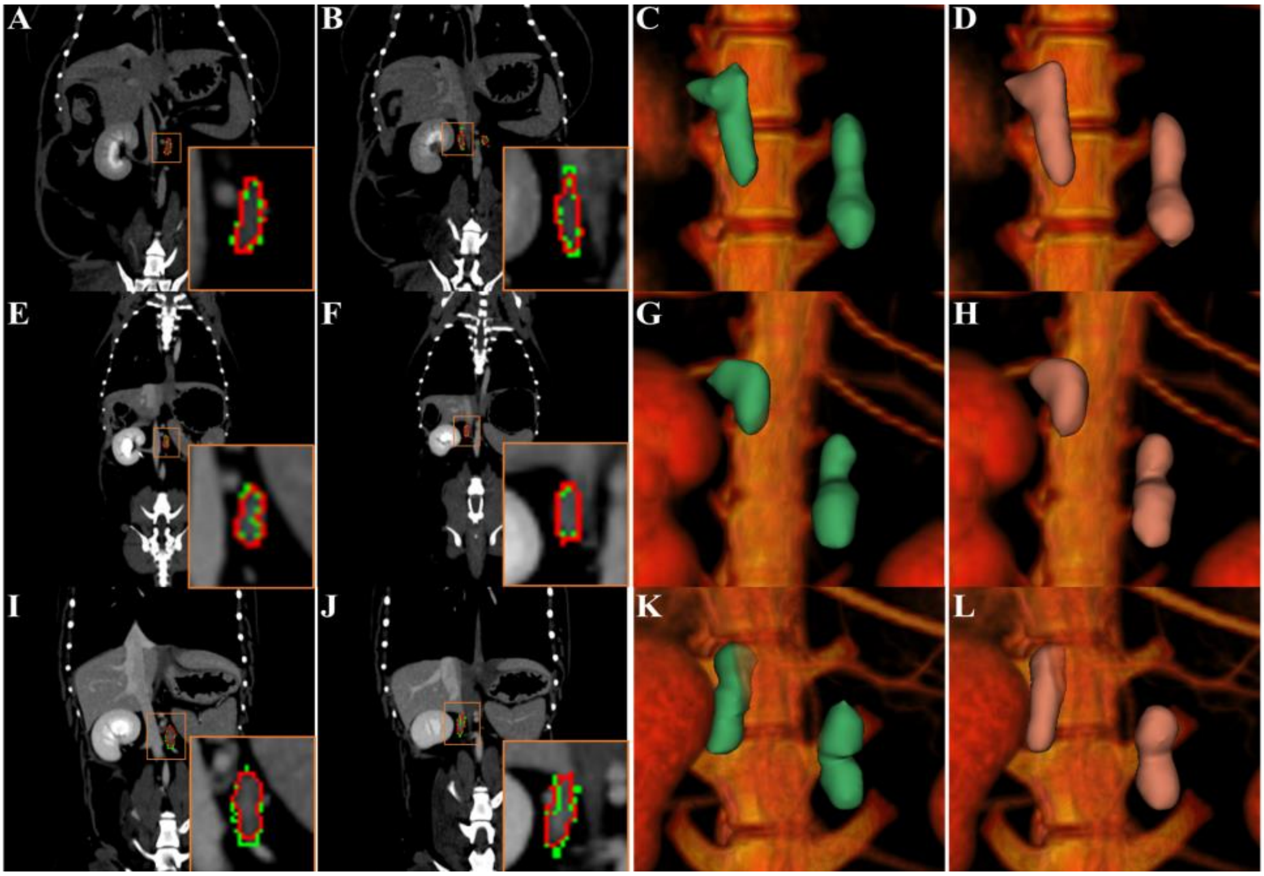

Fig. 4: Samples of the automatically segmented results produced by the deep learning model. Green and red contours represent ground-truth and predicted adrenal glands, respectively (Left gland: A, E, I; Right gland: B, F, J). 3D renderings of the ground-truth (C, G, K) and prediction (D, H, L) are also shown. Orange rectangles in A, B, E, F, I, J highlight the magnified areas.

Accuracy and the time consumption for direct estimation of adrenal gland volume: In order to assess the model's performance, the DSC was calculated for 50 test sets comparing manual and automated segmentations. The DSC was 0.885±0.075. Fig. 4 illustrates the adrenal glands identified by the manual method and the automated method. Figs. 4A, 4E, and 4I represent the left adrenal gland, while Figs. 4B, 4F, and 4J show the right adrenal gland. Figs. 4C, 4G, and 4K depict the manual method, and Figs. 4D, 4H, and 4L represent the automated method in 3D rendering. The sensitivity and specificity of the estimations of adrenal gland volume by automated method were 0.759±0.152 and 0.999±0.001, respectively. The measured MAE was 0.004±0.002. These quantitative assessments consistently indicate the high accuracy of the trained model.

The average time required for direct estimation of adrenal gland volume in the validation set was 2.53 seconds per dog, whereas the manual segmentation approach required approximately 25 minutes per dog. Thus, the time required for the estimation of adrenal gland volume by manual method was about 592 times higher than that required by direct method.

Adrenal gland volume in normal dogs: Both the left and right adrenal glands were visible in all 239 dogs, yielding a total of 478 glands. The differences in the volumes of the left and right adrenal glands were statistically non-significant. The mean adrenal gland volumes for each BW group are summarized in Table 1. The adrenal gland volume increased with BW, and ANOVA revealed a significant difference (P<0.001) between the BW groups, except between groups A (BW≤2.5) and B (BW>2.5-5.0), as shown in Table 1. The Pearson correlation coefficient (r) between adrenal gland volume and BW was 0.821 (P<0.001).

The relationship of BW and age with adrenal volume was expressed by the following regression equation: adrenal volume=0.033×BW+0.014×age+0.092(adjR^2=0.70, R^2=0.70; P<0.001) (Fig. 5A). The correlation between adrenal gland volume and BW was represented by the following formula: adrenal gland volume=0.032×BW+0.213 (adjR^2=0.67, R^2=0.67, P<0.001).

The data on mean age, BW, and adrenal volume of neutered females, intact females, neutered males, and intact males are presented in Table 2. The Kruskal-Wallis test indicated statistically non-significant difference in mean adrenal gland volume among the dogs of four groups. The t-test comparing adrenal volumes between neutered and intact dogs, as well as between female and male dogs, showed non-significant differences.

The correlation between age of the dog and adrenal volume was evaluated using Pearson's correlation analysis, revealing a weak positive correlation (r=0.147, P<0.05). The correlation between BCS and volume, while controlling for the influences of BW and age, was assessed using partial correlation analysis, which showed a weak negative correlation (r=-0.233, P<0.001). This relationship was expressed in the following regression equation: Adrenal volume=-0.51xBCS+0.033×BW+ 0.015×age+0.373 ($adjR^2$^=0.72, P<0.001), as shown in Fig. 5B.

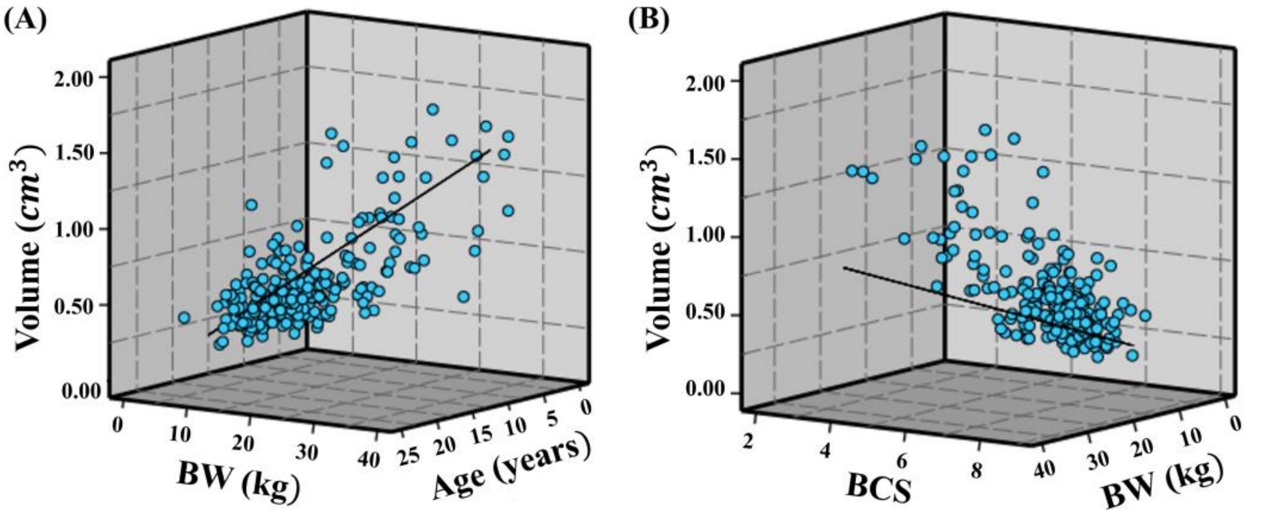

Fig. 5: Scatter plots from the multiple regression analysis are presented. A: Multiple regression analysis of the relationship between BW, Age, and adrenal gland volume, expressed by the equation: adrenal volume=0.033xBW+0.014xAge+0.092 (adjR^2=0.70, P<0.001). B: Multiple regression analysis of the relationship between BW, BCS, and adrenal gland volume, expressed by the equation: Adrenal volume=-0.51xBCS+0.033xBW+0.015xAge+0.373 (adjR^2=0.72, P<0.001). BW, body weight; BCS, body condition score.

Table 2: Data for each group, such as the number of dogs, mean age, mean BW, and mean adrenal volume

Median (IQR)	Neutered female	Intact female	Neutered male	Intact male
Number of dogs (n)	86	24	113	16
Age (years)	9.75(7.0-12.0)	8.04(3.21-11.50)	9.00 (5.33-11.08)	7.50 (5.29-9.09)
BW (kg)	4.55(2.85-8.80)	7.25(2.85-16.50)	5.35 (3.80-9.20)	17.78 (5.30-21.45)
Volume (cm^3) *	0.38(0.29-0.51)	0.44 (0.27-0.74)	*0.40 (0.30-0.57)*	*0.73 (0.35-0.90)*

* Non-significant intergroup differences on Kruskal-Wallis test.

Residual analysis confirmed that the regression model met assumptions of linearity, normality, and homoscedasticity. Additionally, VIFs for all predictors were 1.00, suggesting no concerns with multicollinearity. The confidence intervals for the regression coefficients were narrow and did not include zero, indicating precise and statistically significant estimates for all predictors. There was no difference in adrenal gland volume before and after the smoothing process. Additionally, statistically non-significant difference was observed when the spacing was fixed at 0.75 during post-processing compared to situation when it was not fixed.

DISCUSSION

In this study, a Swin-Transformer-based segmentation model was employed for accurate segmentation of the adrenal glands. However, convolutional neural networks (CNNs), such as nnU-Net (Isensee *et al.*, 2021), have been widely used. Recent studies have demonstrated that Transformer-based segmentation models can be highly effective for various medical images segmentation tasks, particularly when large-scale datasets are available (Hatamizadeh *et al.*, 2021). Given that we collected a relatively large dataset, this study adopted a Transformer-based approach for adrenal gland segmentation, achieving high accuracy in volume determination. Nevertheless, as indicated in a recent comparative study (Isensee *et al.*, 2024), the performance superiority between CNN-based and Transformer-based models depends on specific tasks and dataset size. Thus, a comprehensive comparison involving different models and varying dataset sizes would be valuable to explore in future research.

In the present study, the deep learning model showed high similarity with the manual method in segmentation and volume measurement of the adrenal gland (r=0.922, P<0.001). Only two data points showed low correlation: where most of the right adrenal gland was overlapped by the liver. In previous studies, the Dice Similarity Coefficient (DSC) was 0.915 in a deep learning model for kidney recognition in dog CT images (Ji *et al.*, 2022), and 0.93 in a model for eye recognition (Park *et al.*, 2021). In our study, DSC was slightly lower (0.885±0.075) than other organs, most probably because the adrenal glands are smaller, vary in shape, and are often overlapped by the surrounding tissues.

In order to address the factors contributing to the relatively lower DSC observed in this study, two potential strategies can be proposed. First, the implementation of a multi-scale approach could help overcome the challenge posed by the small size of the adrenal glands. By utilizing multiple scales, the model could more effectively capture both fine details and larger anatomical structures, which is particularly important when working with small and variable-sized organs such as the adrenal glands. Second, expanding the dataset to include a broader and more diverse population of dogs, encompassing a wider range of adrenal gland sizes and shapes, would enhance ability of the model to generalize across individual variations. This would likely improve robustness and overall performance of the model. We believe that the adoption of these strategies would address some of the limitations encountered in this study and could lead to substantial improvements in the DSC in future iterations of the model. In a recent study of deep learning model for adrenal gland detection in human medicine, the DSC was 0.7009, and the ICC between the manual and automatic methods was 0.91 (Kim *et al.*, 2023). Compared to clinical medical imaging results, our model achieved a higher DSC.

In this study, the deep learning model obtained adrenal volumes from CT images 592 times faster than the manual method (2.53 sec V 25.0 minutes), significantly reducing clinicians' time. Adrenal gland volume is a reliable indicator for the early diagnosis of adrenal diseases, offering an advantage over length measurements. However, this study showed that the manual calculation of the volume of a single adrenal gland required approximately 25 minutes, making it impractical for clinical use. In contrast, deep learning can determine the adrenal gland volume in just 2.53 seconds, enabling measurements in a larger number of dogs. This enhances the efficiency of clinical practice, allows for earlier detection of diseases, facilitates individualized treatment plans and supports consistent monitoring, thereby improving its clinical value. Further studies are suggested to reduce computational complexity and improve execution time by employing lightweight deep neural networks and model compression techniques (Cho *et al.*, 2023, 2024). These techniques could accelerate the deep learning models, enabling quasi-realtime analysis of CT images.

In this study, adrenal gland volume showed a significant positive correlation with BW ($r=0.821$; $P<0.05$), consistent with previous studies (Swepson *et al.*, 2022; Büttelmann *et al.*, 2023). Except for groups A (BW ≤2.5kg) and B (BW >2.5-5.0kg), adrenal gland volume differed significantly among all BW groups, possibly because most dogs (19 of 32 dogs) in Group A weighed more than 2kg, while 12 of 80 dogs in Group B fell within the 2.5kg to 3.0kg range. Given that the adrenal gland is a very small organ, this may have played a significant role. However, considering that the average adrenal volume in Group A is $0.22\,cm^3$ and in Group B it is $0.33\,cm^3$, if smaller dogs had been included in group A, there would have been statistically significant difference between dogs of groups A and B.

To the best of our knowledge, there have been no studies examining the relationship between BCS and adrenal gland volume in dogs. In this study, a weak negative correlation was found between BCS and adrenal gland volume. This suggests that for two dogs of the same body weight, a small-breed dog with a high BCS and a large-breed dog with a low BCS, the adrenal volume would be higher in the large-breed dog with a lower BCS. Therefore, adrenal gland volume is not influenced solely by body weight, and BCS should also be considered.

A previous study on 48 dogs (Bertolini *et al.*, 2006) has shown that the left adrenal gland has a higher volume than the right. However, in the present study, non-significant difference was observed in the volume of the left and right adrenal glands. This discrepancy could be attributed to differences in study design, sample size, or the specific population of dogs studied. Additionally, factors such as breed, age, or health conditions of dogs could also influence adrenal gland size, which may vary across different studies. Further research with a larger sample size and a more diverse population is needed to confirm these findings.

The relationship between age and adrenal gland size remains controversial in veterinary ultrasound research. Previous studies involving 49 (Bertolini *et al.*, 2008) and 62 dogs (Barthez *et al.*, 1995) did not find any correlation between age and adrenal gland size on ultrasound. However, other studies have identified a weak positive correlation between age and adrenal gland thickness (De Chalus *et al.*, 2013; Bento *et al.*, 2016). In our study, adrenal gland volume increased with age, similar to the findings in human studies, and it is thought that adrenal cortex atrophy occurs with age, while compensatory hyperplasia results in increased adrenal gland volume (Hornsby, 2002; Meier *et al.*, 2007). The weak correlation observed in this study suggests that BCS and age alone may not be a reliable predictor of adrenal gland size in dogs. However, when considered together with BW, BCS and age could contribute to a more accurate estimation of adrenal volume, highlighting the complexity of the relationship between these factors.

In this study, we present a regression model using BW, age, and BCS, which demonstrated high predictive power and statistical reliability. By comparing the measured adrenal gland volume with the expected volume for individual subjects based on the regression model, this model could enhance confidence in the process of diagnosing adrenal diseases based on CT findings and clinical information.

In veterinary medicine, the relationship between sex and adrenal gland size is also debated. Some studies report that adrenal gland is thicker in males than in females (Mogicato *et al.*, 2011), while others claim the opposite (De Chalus *et al.*, 2013). Some research suggests that sex is not related to adrenal size (Bento *et al.*, 2016), while in our study, no correlation was observed between sex or neutering status of dogs and adrenal volume. This discrepancy may be explored further in future research.

This study had certain limitations. Due to the retrospective nature of this study, definitive diagnostic methods for adrenal gland disease, such as Low Dose Dexamethasone Suppression Test (LDDST) or Urinary Cortisol Creatinine Ratio (UCCR), were not available for all patients. This means that some dogs with subclinical adrenal gland disease may have been included, potentially affecting the results. Similarly, the dataset is limited to dogs without evidence of adrenal gland enlargement, which may not fully represent the diversity of adrenal gland conditions observed in clinical practice. While the model performs well on the dataset used, its generalizability to other breeds, sizes, or conditions of dogs is not fully explored. Further validation on a more diverse dataset would strengthen conclusions of the study. Furthermore, the deep learning model, while effective, is computationally complex. Thus, future work is also suggested to reduce computational complexity and improve execution time, which would be beneficial for real-time clinical applications.

Conclusions: In conclusion, the deep learning model established in our study could aid veterinarians in efficiently estimating adrenal gland volume from canine CT views. Additionally, this study also offers a reference range for adrenal gland volume in healthy dogs, accounting for BW, age, and BCS, which can be utilized in the clinical assessment of adrenal glands.

Conflict of interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethics statement: The animal study protocol was reviewed and approved by the Institutional Animal Care and Use Committee of Jeonbuk National University, Iksan, Republic of Korea (approval No. NONJBNU 2021-0156).

Acknowledgments: The authors would like to thank clinicians of Veterinary Medical Imaging Department of the Teaching Hospital of Jeonbuk National University, Iksan, Republic of Korea for their assistance with the manual segmentation of CT images used in this study.

Authors contribution: SP and HY conceived the idea, designed the study and drafted the manuscript. SH and HC acquired data. SH, HC, KL, and HY analyzed and interpreted the data. All the authors contributed to the preparation of the manuscript and approved the final version.

REFERENCES

Barthez PY, Nyland TG and Feldman EC, 1995. Ultrasonographic evaluation of the adrenal glands in dogs. J Am Vet Med Assoc 207:1180–83.

Bento PL, Center SA, Randolph JF, *et al.*, 2016. Associations between sex, body weight, age, and ultrasonographically determined adrenal gland thickness in dogs with non-adrenal gland illness. J Am Vet Med Assoc 248:652–60.

Besso JG, Penninck DG and Gliatto JM, 1997. Retrospective ultrasonographic evaluation of adrenal lesions in 26 dogs. Vet Radiol Ultrasound 38:448–55.

Bertolini G, Furlanello T, De Lorenzi D, *et al.*, 2006. Computed tomographic quantification of canine adrenal gland volume and attenuation. Vet Radiol Ultrasound 47:444–48.

Bertolini G, Furlanello T, Drigo M, *et al.*, 2008. Computed tomographic adrenal gland quantification in canine adrenocorticotroph hormone-dependent hyperadrenocorticism. Vet Radiol Ultrasound 49:449–53.

Büttelmann G, Harder LK, Nolte I, *et al.*, 2023. Impact of body weight and sex in selected dog breeds on the canine adrenal gland dimensions measured by computed tomographic imaging. BMC Vet Res 19:99.

Cardoso MJ, Li W, Brown R, *et al.*, 2022. MONAI: An open-source framework for deep learning in healthcare. arXiv:2211.02701. https://doi.org/10.48550/arXiv.2211.02701.

Cho H, Kim D, Chang S, *et al.*, 2024. A system-on-chip solution for deep learning-based automatic fetal biometric measurement. Expert Syst Appl 237:121482.

Cho H, Song I, Jang J, *et al.*, 2023. A lightweight deep learning network on a system-on-chip for wearable ultrasound bladder volume measurement systems: Preliminary Study. Bioengineering 10:525.

de Bruin C, Meij BP, Kooistra HS, *et al.*, 2009. Cushing's disease in dogs and humans. Horm Res 71(suppl 1):140–43.

De Chalus T, Combes A, Bedu AS, *et al.*, 2013. Ultrasonographic adrenal gland measurements in healthy Yorkshire Terriers and Labrador Retrievers. Anat Histol Embryol 42:57–64.

Gregori T, Mantis P, Benigni L, *et al.*, 2015. Comparison of computed tomographic and pathologic findings in 17 dogs with primary adrenal neoplasia. Vet Radiol Ultrasound 56:153–59.

Hatamizadeh A, Nath V, Tang Y, *et al.*, 2021. Swin UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images. In: Crimi A, Bakas S (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2021. Lecture Notes in Computer Science, vol 12962. Springer, Cham. https://doi.org/10.1007/978-3-031-08999-2_22

He Y, Nath V, Yang D, *et al.*, 2023. SwinUNETR-V2: Stronger Swin Transformers with Stagewise Convolutions for 3D Medical Image Segmentation. In: de Bruijne M, Cattin PC, Cotin S, Padoy N, Speidel S, Zheng Y, Essert C (eds). Medical Image Computing and Computer-Assisted Intervention – MICCAI 2023. Lecture Notes in Computer Science, vol 14223. Springer, Cham 416–26. https://doi.org/10.1007/978-3-031-43901-8_40

Hornsby PJ, 2002. Aging of the human adrenal cortex. Ageing Res Rev 1:229–42.

Huang X and Belongie S, 2017. Arbitrary style transfer in real-time with adaptive instance normalization. arXiv:1703.06868. https://doi.org/10.48550/arXiv.1703.06868

Isensee F, Jaeger PF, Kohl SA, *et al.*, 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods 18(2): 203-211.

Isensee F, Wald T, Ulrich C, *et al.*, 2024. Nnu-net revisited: A call for rigorous validation in 3d medical image segmentation. In: Linguraru MG, *et al.* (eds), Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. MICCAI 2024. Lecture Notes in Computer Science, vol 15009. Springer, Cham. https://doi.org/10.1007/978-3-031-72114-4_47

Ji Y, Cho H, Seon S, *et al.*, 2022. A deep learning model for CT-based kidney volume determination in dogs and normal reference definition. Front Vet Sci 9:1011804.

Kim TM, Choi SJ, Ko JY, *et al.*, 2023. Fully automatic volume measurement of the adrenal gland on CT using deep learning to classify adrenal hyperplasia. Eur Radiol 33:4292–302.

Li Y, Zhao Y, Yang P, *et al.*, 2024. Adrenal volume quantitative visualization tool by multiple parameters and an nnU-Net Deep Learning Automatic Segmentation Model. J Digit Imaging Inform Med 38: 47-49. https://doi.org/10.1007/s10278-024-01158-y

Loshchilov I and Hutter F, 2019. Decoupled weight decay regularization. arXiv:1711.05101. https://doi.org/10.48550/arXiv.1711.05101.

Ma J, Chen J, Ng M, *et al.*, 2021. Loss odyssey in medical image segmentation. Med Image Anal 71:102035.

Meier JM, Alavi A, Iruvuri S, *et al.*, 2007. Assessment of age-related changes in abdominal organ structure and function with computed tomography and positron emission tomography. Semin Nucl Med. Elsevier 37:154–72.

Mogicato G, Layssol-Lamour C, Conchou F, *et al.*, 2011. Ultrasonographic evaluation of the adrenal glands in healthy dogs: Repeatability, reproducibility, observer-dependent variability, and the effect of bodyweight, age and sex. Vet Rec 168(5):130.

Park J, Choi B, Ko J, *et al.*, 2021. Deep-learning-based automatic segmentation of head and neck organs for radiation therapy in dogs. Front Vet Sci 8:721612.

Perfetti S, Diana A, Baron Toaldo M, *et al.*, 2021. CT measures of adrenal gland length and caudal pole diameter are reproducible in large breed dogs: A pilot study. Vet Radiol Ultrasound 62:402–12.

Robinson-Weiss C, Patel J, Bizzo BC, *et al.*, 2023. Machine learning for adrenal gland segmentation and classification of normal and adrenal masses at CT. Radiology 306:e220101.

Swepson R, Hosgood G and Stander N, 2022. Algorithmic estimation of adrenal volume based on body weight and CT landmarks in dogs. Vet Radiol Ultrasound 63:264–71.