



# A deep learning framework for histopathological classification of canine testicular tumours

Lorenzo Riccio<sup>a,b,\*</sup>, Luigi Rosati<sup>b,c,\*\*</sup>, Maria De Falco<sup>b,d</sup>, Evaristo Di Napoli<sup>a</sup>,  
Pietro Formisano<sup>e</sup>, Elia Formisano<sup>f,g,1</sup>, Orlando Paciello<sup>a,h,1</sup>

<sup>a</sup> Department of Veterinary Medicine and Animal Productions, University of Naples Federico II, Via Federico Delpino, 1, Naples, NA 80137, Italy

<sup>b</sup> Department of Biology, University of Naples Federico II, Via Vicinale Cupa Cintia, 21, Naples, NA 80126, Italy

<sup>c</sup> Centro Interdipartimentale di Ricerca "Ambiente" (CIRAM), University of Naples Federico II, Via Tarsia, 31, Naples, NA 80134, Italy

<sup>d</sup> Biostructures and Biosystems National Institute, Via dei Carpegna, 19, Roma, RM 00165, Italy

<sup>e</sup> Department of Translational Medicine, Federico II University of Naples, Via S. Pansini 5, Naples 80131, Italy

<sup>f</sup> Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht 6200 MD, the Netherlands

<sup>g</sup> Maastricht Centre for Systems Biology (MaCSBio), Faculty of Science and Engineering, Maastricht University, Maastricht 6200 MD, the Netherlands

<sup>h</sup> Istituto Zooprofilattico Sperimentale del Mezzogiorno, Portici, NA 80055, Italy

## ARTICLE INFO

### Keywords:

Digital pathology  
Whole-slide imaging  
Deep learning  
Convolutional neural networks  
EfficientNet  
Grad-CAM  
Canine testicular tumours

## ABSTRACT

Testicular neoplasms are common in dogs, and their histopathological classification may be challenging in selected cases, particularly when morphological patterns overlap, with potential implications for diagnostic consistency and case management. In this study, we propose an artificial intelligence (AI)-based computational pathology approach for the automatic classification of the three main canine testicular neoplasms (Seminoma, Leydig cell tumour, and Sertoli cell tumour) from digitised haematoxylin and eosin-stained Whole-Slide Images (WSIs). To this end, 200 histological WSIs from 200 male dogs with histologically diagnosed testicular tumours were collected from the archives of the DIPSA laboratory, University of Naples Federico II. We fine-tuned the EfficientNet-B4 convolutional neural network (CNN) on tiles from WSIs and evaluated performance using 5-fold cross-validation with WSI-level partitioning. In addition, we implemented an attention-based multiple instance learning (ABMIL) framework as a complementary weakly supervised approach for WSI-level classification. The fine-tuned EfficientNet-B4 model provided the strongest tile-level discrimination, while Logistic Regression with mean probability aggregation achieved the highest WSI-level performance among the evaluated WSI-level approaches (accuracy  $0.84 \pm 0.07$ ). Furthermore, the ABMIL framework showed competitive WSI-level performance, reaching an accuracy of  $0.80 \pm 0.04$ . Finally, Grad-CAM heatmaps revealed that the fine-tuned CNN focused on histologically relevant tumour regions. Overall, our findings support the feasibility of AI-assisted WSI-level histopathological classification of canine testicular tumours and provide a proof-of-concept for the potential development of decision-support tools for veterinary pathologists.

## 1. Introduction

Among canine reproductive neoplasms, testicular tumours represent a frequent diagnostic finding, particularly in older dogs (Grieco et al., 2008; Mosier, 1989; Riccio et al., 2026a, 2026b). The most common histological testicular tumour types include Seminomas, Leydig cell tumours, and Sertoli cell tumours (Orlandi et al., 2022). The transition

from conventional microscopy to digitised WSIs has accelerated the development of Digital Pathology, creating new opportunities for computational image analysis, algorithmic decision-support, and data sharing (Aeffner et al., 2019; Ancheta et al., 2024; Prasad et al., 2025). In this context, AI systems represent a potential innovation for veterinary pathology, offering tools for the reproducible analysis of specific morphological patterns (Fotouh, 2025; Rogers et al., 2024; Xiao et al.,

\* Corresponding author at: Department of Veterinary Medicine and Animal Productions, University of Naples Federico II, Via Federico Delpino, 1, Naples, NA 80137, Italy.

\*\* Corresponding author at: Department of Biology, University of Naples Federico II, Via Vicinale Cupa Cintia, 21, Naples, NA 80126, Italy

E-mail addresses: [lorenzo.riccio@unina.it](mailto:lorenzo.riccio@unina.it) (L. Riccio), [luigi.rosati@unina.it](mailto:luigi.rosati@unina.it) (L. Rosati).

<sup>1</sup> Co-last authors

2025). These systems can assist pathologists and researchers by identifying image regions relevant to classification and supporting the interpretation of digitised histological slides (Ahmed et al., 2022; Campanella et al., 2019). Among feature-learning methods, deep learning approaches, particularly convolutional neural networks (CNNs), have been widely applied to digital slides because they can learn hierarchical representations and detect morphological variations that may elude manual assessment (Amaral et al., 2024; Gertych et al., 2019; Ponzio et al., 2019; Sarvamangala and Kulkarni, 2022). Across species and tissue contexts, AI-based computational pathology has been applied to tasks ranging from the classification of human breast and lung tumours to the localisation of diagnostically informative regions in canine cutaneous mast cell tumours (Aubreville et al., 2020; Coudray et al., 2018; Sajiv and Ramkumar, 2023a, 2023b, 2023c). Beyond histopathological classification, AI-based approaches have also been applied across different areas of oncology, including image enhancement, tumour detection and early diagnosis, as well as personalised treatment planning (Balakrishna et al., 2025; Kumar et al., 2026; Sajiv and Meenakshisundaram, 2025; Sivakumar et al., 2023). Despite these advances, the implementation of AI-based tools in veterinary pathology remains limited, mainly due to the lack of annotated image datasets and standardised validation protocols (Abels et al., 2019; Battazza et al., 2024; Prasad et al., 2025; Zuraw and Aeffner, 2022). In light of these considerations, in our study we applied these approaches specifically to canine testicular pathology by integrating complementary tile-level and WSI-level classification strategies within the same validation framework. This design allowed local morphological discrimination to be evaluated alongside the diagnostically more relevant slide-level classification, while Grad-CAM analysis provided an assessment of the histological regions contributing to CNN predictions. Therefore, the aims of this study were to: (1) develop and evaluate a reproducible artificial intelligence framework for the automated histopathological classification of the three main canine testicular tumours from whole-slide images; (2) compare complementary tile-level and slide-level modelling strategies; and (3) assess CNN interpretability through Grad-CAM.

2. Materials and methods

2.1. Ethics approval and sample collection

All procedures were conducted in accordance with European Directive 2010/63/EU and were approved by the Ethical Animal Care and Use Committee of the University of Naples Federico II (Protocol No. PG/2023/0029409; approval date: 10 March 2023). Samples were collected exclusively during routine castration procedures; no additional procedures were performed for research purposes, and no animals were euthanised for the purposes of this study.

2.2. Histopathological processing and WSI acquisition

Haematoxylin and eosin (H&E) stained slides from cases previously diagnosed by three veterinary pathologists as Seminoma, Leydig cell tumour, and Sertoli cell tumour were retrieved from the archives of the DIPSA Laboratory, Department of Veterinary Medicine and Animal Reproduction, University of Naples Federico II. For all cases, testes were sectioned longitudinally and fixed in 10% neutral-buffered formalin (05–01005Q, BioOptica, Milan, Italy) for 72 h. Following fixation, representative tissue samples were trimmed, placed in histological cassettes, routinely processed, dehydrated, and embedded in paraffin wax (Power et al., 2026). Sections of 3–4 µm thickness were subsequently obtained and stained with H&E for histopathological evaluation (Piegari et al., 2024; Riccio et al., 2025). All slides were stained using a Leica Autostainer XL to ensure consistent staining quality (d’Aquino et al., 2023; Zappulli et al., 2026) and subsequently digitised at high resolution using an EpreDia Panoramic Scan scanner at 20 × magnification (d’Aquino et al., 2026). The digitised slides were organised into

category-specific directories as follows: Seminoma (label = 0), Leydig cell tumour (label = 1), and Sertoli cell tumour (label = 2). The dataset consisted of 200 whole-slide images (WSIs), one per animal (Seminoma n = 75, Leydig cell tumour n = 75, Sertoli cell tumour n = 50).

2.3. Tile extraction and preprocessing

From each WSI, 2500 non-overlapping tiles (380 × 380 px) were randomly sampled from tumour-containing regions of interest (ROIs) manually delineated by a trained researcher under veterinary pathologist supervision, resulting in a total of 5.0 × 10<sup>5</sup> tiles before quality control (maximum yield: 200 × 2500). After multi-step quality control, 91.6% of the tiles were retained, yielding 4.58 × 10<sup>5</sup> tiles for model development and validation. For readability, aggregate tile counts in Table 1 are reported as rounded values. Table 1 summarises the dataset composition and preprocessing statistics. Quality control (QC) and preprocessing were performed using OpenSlide and OpenCV libraries. Otsu’s thresholding (Otsu, 1979) was applied for tissue detection, followed by exclusion of tiles dominated (>85%) by near-white (intensity >240) or near-black (intensity <15) pixels. Tiles were retained only if they exhibited sufficient contrast (grayscale SD ≥15), appropriate mean intensity (20–235), inter-channel colour variation (SD ≥10), and detectable structural content (≥1% edge pixels via the Canny detector) (Canny, 1986). Tiles passing QC underwent Reinhard colour normalisation using a representative reference tile in LAB colour space to minimise staining variability across slides (Reinhard et al., 2001).

2.4. Feature extraction and classifier evaluation

Tile embeddings were extracted from normalised tiles using an ImageNet-pretrained EfficientNet-B4 backbone (Tan and Le, 2019) as a frozen feature extractor. After removal of the final classification layer, 1, 792-dimensional tile embeddings were obtained from the penultimate convolutional block. All classification analyses were performed using a stratified 5-fold cross-validation scheme at the WSI level. In each iteration, one fold was held out for testing, while the remaining four folds were used for model development. Within the training partition, 10% of the WSIs were randomly selected in a stratified manner as an internal validation subset. All tiles originating from the same WSI were retained within the same partition throughout the cross-validation procedure, thereby preventing tile-level data leakage between training, validation, and test sets. The internal validation subset was used to monitor MLP training, whereas final performance was computed exclusively on the held-out test fold. Tile embeddings were standardised using z-score parameters estimated on the training set and applied to the corresponding validation and test sets. For qualitative visualisation of class separability, LDA (Fisher, 1936; Rao, 1948) was fitted on the training-set embeddings and applied to the corresponding validation and test sets.

The extracted tile embeddings were used to train three supervised

**Table 1**  
Dataset composition and preprocessing summary. Tiles were retained after QC if Canny edge pixels ≥ 1%, mean intensity ∈ [20, 235], grayscale s.d. ≥ 15, inter-channel s.d. ≥ 10, and near-white/near-black pixel fraction < 85%. Aggregate tile counts are rounded for readability.

| Diagnostic category | No. of WSIs | No. of animals | Tiles (pre-QC) (max 2500 per WSI) | Tiles (post-QC) (total) | Retention (%) | Tiles per WSI (post-QC mean) |
|---------------------|-------------|----------------|-----------------------------------|-------------------------|---------------|------------------------------|
| Seminoma            | 75          | 75             | 187,500                           | 174,000                 | 92.8          | 2320                         |
| Leydig cell tumour  | 75          | 75             | 187,500                           | 168,000                 | 89.6          | 2240                         |
| Sertoli cell tumour | 50          | 50             | 125,000                           | 116,000                 | 92.8          | 2320                         |
| Total               | 200         | 200            | 500,000                           | 458,000                 | 91.6          | 2290                         |

classifiers: 1) Logistic Regression (LR) with class balancing (Hosmer et al., 2013); 2) a Multilayer Perceptron (MLP) with a 256-neuron hidden layer, ReLU, dropout, and a three-neuron output layer (Rumelhart et al., 1986); and 3) a Random Forest (RF) with 200 trees and balanced class weights (Breiman, 2001). As a complementary WSI-level baseline, slide-level predictions were obtained by aggregating tile-level class probabilities from the LR classifier. For each held-out WSI, tile probabilities were computed with the fitted LR model. WSI-level probabilities were derived using either mean pooling across all tiles or top-k pooling ( $k = 50$ ), in which the 50 tiles with the highest classification confidence were selected and their class-probability vectors were averaged to obtain the WSI-level prediction. The value of  $k$  was selected a priori and was not tuned on the validation set.

## 2.5. Fine-tuning of EfficientNet-B4

EfficientNet-B4 was also fine-tuned end-to-end on the tile-level dataset. The ImageNet-pretrained backbone (Deng et al., 2009) was adapted by replacing the output layer with a three-neuron classifier corresponding to the diagnostic classes and adding a dropout layer ( $p = 0.4$ ). Data augmentation included random horizontal and vertical flips, 45° rotations, and mild colour jitter affecting contrast, saturation, brightness, and hue. Tiles were normalised using ImageNet mean and standard deviation statistics. Under the same cross-validation design, class imbalance was addressed through inverse-frequency weighting in the cross-entropy loss. Optimisation was performed using Adam (Kingma and Ba, 2014) with a learning rate of  $1 \times 10^{-4}$ , batch size of 4, and up to 20 epochs. Early stopping (patience = 3) was based on the accuracy measured on the internal validation subset.

## 2.6. Attention-based multiple instance learning

An attention-based multiple instance learning (ABMIL) framework was implemented to perform weakly supervised WSI-level classification from fixed EfficientNet-B4 tile embeddings. Each WSI was treated as a collection of image tiles and assigned a single slide-level diagnosis, without requiring individual tile annotations. Tile representations were obtained using the frozen ImageNet-pretrained EfficientNet-B4 feature extractor. Tile embeddings were standardised using z-score parameters estimated from the training WSIs within each fold and applied to the corresponding validation and test WSIs. Under the same cross-validation design, the ABMIL architecture comprised a learnable projection block, followed by an attention branch producing instance-level attention weights. These weights quantified the relative contribution of individual tiles to the final WSI-level prediction. Attention-weighted aggregation yielded a slide-level representation, which was passed to a linear classifier for three-class prediction. A maximum of 1200 tiles per WSI were randomly sampled at each training epoch. For validation and testing, up to 2000 QC-passed tiles per WSI were selected once at random within each fold using the fold-specific random seed. Class imbalance was handled using inverse-frequency weighting in the cross-entropy loss. The model was optimised with AdamW for up to 30 epochs, with early stopping (patience = 6) based on averaged F1-score measured on the internal validation subset.

## 2.7. Model explainability

Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2017) was applied to the fine-tuned EfficientNet-B4 model to visualise the image regions contributing most to classification. For quantitative evaluation, a stratified random subset of 150 correctly classified test-set tiles (50 per class) was annotated with ROIs by a trained researcher under veterinary pathologist supervision, blinded to model predictions and Grad-CAM overlays. The analysis was restricted to correctly classified tiles in order to assess whether successful model predictions were supported by attention to histologically relevant

regions. Because this was a single-annotator analysis, no inter-observer agreement was assessed; therefore, the quantitative Grad-CAM analysis should be considered exploratory. ROI annotations were used to compute Intersection-over-Union (IoU) between the expert ROI and the Grad-CAM map thresholded at the 70th percentile, Pointing Game Accuracy, and Top 20% Energy inside the ROI. As a sanity check, the Grad-CAM pipeline was repeated after model weight randomisation, and the same IoU metric was recomputed.

## 2.8. Computational framework and reproducibility

All procedures were implemented in Python 3.10 using OpenSlide, NumPy, OpenCV, PyTorch, and LMDB. Model training and inference were performed on a workstation equipped with an NVIDIA GeForce RTX 3050 GPU (6 GB GDDR6). Random seeds were fixed across tile extraction, fold partitioning, model training, and evaluation to improve reproducibility. Custom scripts for tile extraction, preprocessing, model training, and evaluation are available upon reasonable request.

## 2.9. Statistical analysis

The individual animal, represented by a single WSI, was considered the independent sampling unit; tile-level predictions from the same WSI were therefore treated as correlated observations. Model performance on the held-out test folds was evaluated using accuracy, macro-averaged F1-score, macro-averaged area under the receiver operating characteristic curve (ROC-AUC), and macro-averaged average precision (AP-macro). Fold-wise performance estimates were summarised as mean  $\pm$  standard deviation (SD). For WSI-level analyses, out-of-fold predictions were pooled, and 95% confidence intervals (95% CIs) were estimated using 10,000 stratified bootstrap resamples at the WSI level.

# 3. Results

## 3.1. Histopathological characterisation

Histopathological examination revealed distinct morphological features among the three tumour types. Seminomas consisted of relatively uniform germ cells arranged in lobules separated by fibrous septa (black arrow). Leydig cell tumours were composed of polygonal interstitial cells with eosinophilic cytoplasm (asterisk). Sertoli cell tumours were characterised by elongated columnar cells arranged in tubular structures (black arrowhead). Representative H&E-stained regions for each tumour type are shown in Fig. 1A–C.

## 3.2. Feature-based and tile-level classification performance

The evaluated models showed discrimination among tumour types. For qualitative assessment of class separability, LDA projections of the EfficientNet-B4 tile embeddings for the training, validation, and test sets are shown in Fig. 2A–C, indicating broadly preserved separation among the three tumour classes across data partitions. Mean row-normalised confusion matrices for MLP, Logistic Regression, and Random Forest across the five held-out WSI test folds are shown in Fig. 2D–F. Class-specific recall ranged from 0.54 to 0.71 for MLP, 0.66–0.78 for Logistic Regression, and 0.72–0.81 for Random Forest. Among the feature-based classifiers, Random Forest achieved the highest mean cross-validated accuracy (0.76). A summary of tile-level model performance is reported in Table 2.

When feature-based classifiers trained on deep embeddings were replaced by the fine-tuned EfficientNet-B4 end-to-end model, classification performance improved markedly across all tile-level metrics (Fig. 3). The network achieved a mean cross-validated tile-level accuracy of  $0.82 \pm 0.03$ , with F1-macro = 0.81, AUC-macro = 0.95, and AP-macro = 0.91. Class-specific recall was 0.82 for Seminoma, 0.78 for Leydig cell tumour, and 0.86 for Sertoli cell tumour, confirming

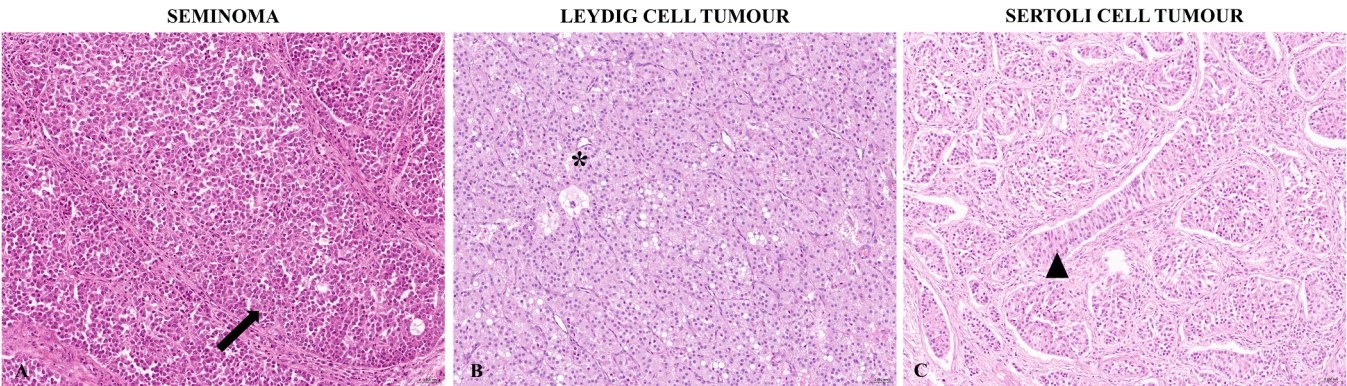


**Fig. 1.** Representative histopathological features of the three main canine testicular tumour types on H&E-stained sections. (A) Seminoma, showing lobular architecture with relatively uniform germ cells separated by fibrous septa (black arrow). (B) Leydig cell tumour, composed of polygonal interstitial cells with eosinophilic cytoplasm (asterisk). (C) Sertoli cell tumour, characterised by tubular arrangements of elongated columnar neoplastic cells (black arrowhead). Original magnification, 20 ×; scale bar = 50 μm. H&E: haematoxylin and eosin.

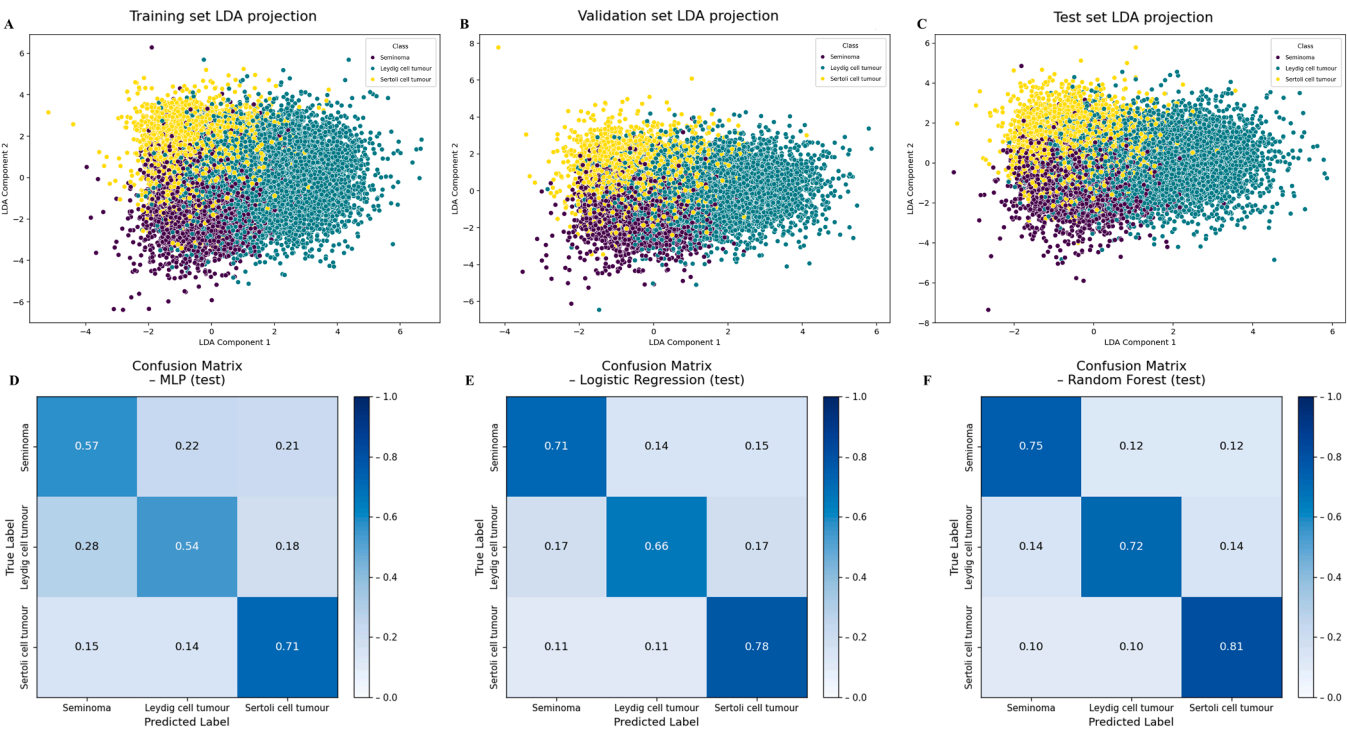


**Fig. 2.** (A–C) LDA projections of EfficientNet-B4 tile embeddings from a representative cross-validation iteration, shown for the training, validation, and held-out WSI test sets, respectively. The LDA transformation was fitted on the training-fold embeddings and applied to the corresponding validation and test sets. (D–F) Mean row-normalised confusion matrices across the five held-out WSI test folds for MLP (D), Logistic Regression (E), and Random Forest (F), in the class order Seminoma, Leydig cell tumour, and Sertoli cell tumour. LDA: linear discriminant analysis; MLP: multilayer perceptron; WSI: whole-slide image.

**Table 2**  
Performance of the evaluated models under 5-fold cross-validation. Feature-based models are evaluated at tile-level; EfficientNet-B4 results are reported at tile-level.

| Model                        | Approach      | Accuracy (mean ± SD) | F1-macro    | AUC-macro   | AP-macro    | Class-wise recall (Seminoma; Leydig cell tumour; Sertoli cell tumour) |
|------------------------------|---------------|----------------------|-------------|-------------|-------------|-----------------------------------------------------------------------|
| MLP                          | Feature-based | 0.64 ± 0.06          | 0.63 ± 0.07 | 0.84 ± 0.05 | 0.82 ± 0.06 | 0.57 ± 0.09 / 0.54 ± 0.10 / 0.71 ± 0.07                               |
| Logistic Regression          | Feature-based | 0.71 ± 0.05          | 0.70 ± 0.05 | 0.89 ± 0.03 | 0.87 ± 0.04 | 0.71 ± 0.07 / 0.66 ± 0.08 / 0.78 ± 0.06                               |
| Random Forest                | Feature-based | 0.76 ± 0.05          | 0.74 ± 0.04 | 0.91 ± 0.03 | 0.90 ± 0.03 | 0.75 ± 0.06 / 0.72 ± 0.07 / 0.81 ± 0.05                               |
| EfficientNet-B4 (fine-tuned) | End-to-end DL | 0.82 ± 0.03          | 0.81 ± 0.06 | 0.95 ± 0.02 | 0.91 ± 0.03 | 0.82 ± 0.04 / 0.78 ± 0.05 / 0.86 ± 0.04                               |

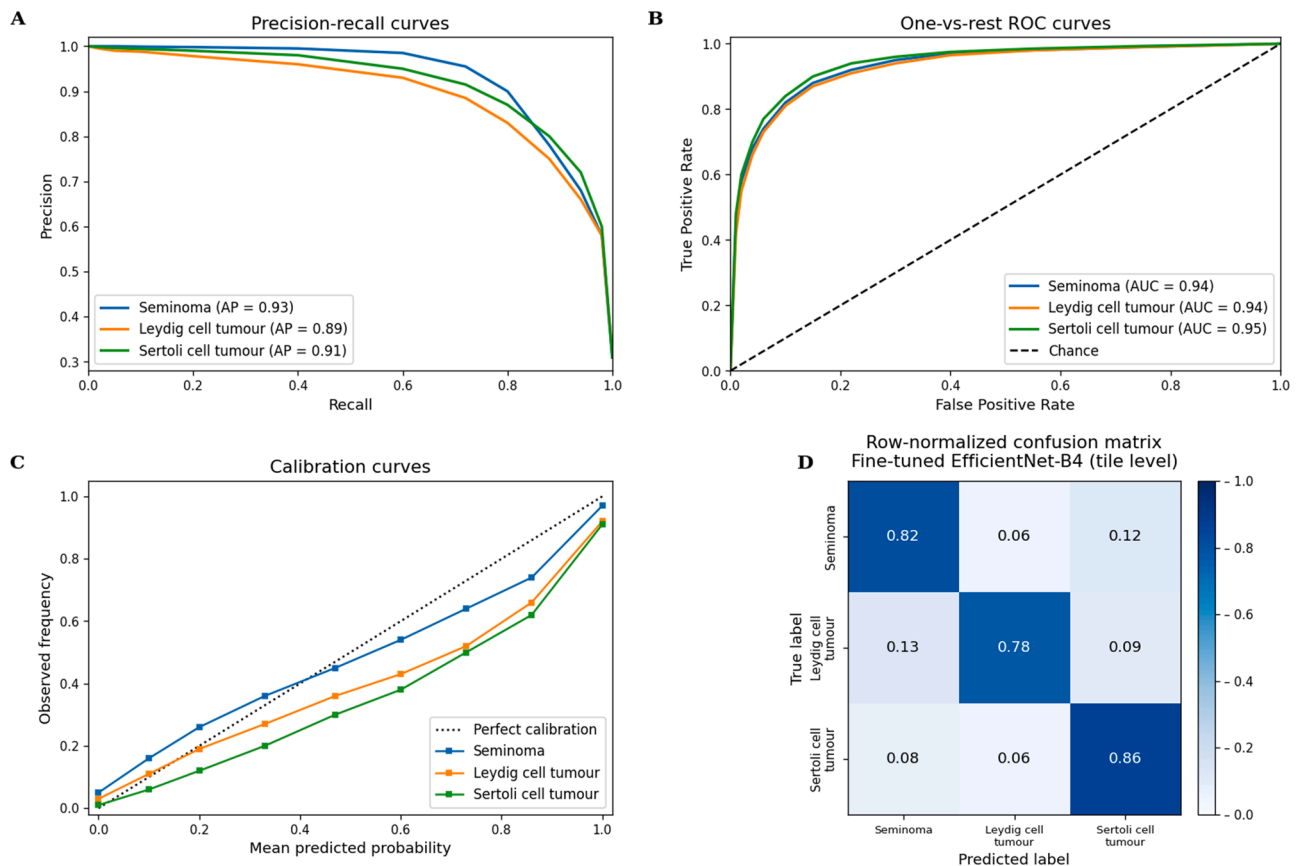


**Fig. 3.** Tile-level performance of the fine-tuned EfficientNet-B4 model. (A) Precision-recall curves for the three tumour classes. (B) One-vs-rest ROC curves for the three tumour classes. (C) Class-wise calibration curves showing the agreement between predicted probabilities and observed outcomes. (D) Row-normalised confusion matrix for the fine-tuned EfficientNet-B4 model at tile level, aggregated across the held-out test folds. ROC: receiver operating characteristic.

consistent inter-class discrimination. Precision–recall and ROC analyses (Fig. 3A, B) demonstrated stable performance across the three diagnostic categories, with AP values ranging from 0.89 to 0.93 and AUC values from 0.94 to 0.95. The calibration curves (Fig. 3C) suggested broadly consistent agreement between predicted probabilities and observed outcomes across the three tumour classes. The normalised confusion matrix is shown in Fig. 3D. Compared with the best feature-based approach (Random Forest), the end-to-end model achieved higher overall accuracy (0.82 vs 0.76), better inter-class balance, and lower variance across folds (SD = 0.03 vs 0.05). Overall, the fine-tuned EfficientNet-B4 provided the strongest tile-level performance.

### 3.3. WSI-level classification performance

WSI-level evaluation was performed separately by aggregating tile probabilities at the slide level. Using mean pooling, Logistic Regression achieved a WSI-level accuracy of  $0.84 \pm 0.07$ , an F1-macro of  $0.80 \pm 0.10$ , an AUC-macro of  $0.90 \pm 0.06$ , and an AP-macro of  $0.84 \pm 0.08$ . Using top-k pooling (k = 50), Logistic Regression achieved an accuracy of  $0.79 \pm 0.11$ , an F1-macro of  $0.75 \pm 0.10$ , an AUC-macro of  $0.90 \pm 0.06$ , and an AP-macro of  $0.85 \pm 0.08$ . The ABMIL model achieved a mean WSI-level accuracy of  $0.80 \pm 0.04$  and a F1-macro of  $0.76 \pm 0.06$ . In the pooled ABMIL analysis, AP values were 0.81 for Seminoma, 0.92 for Leydig cell tumour, and 0.63 for Sertoli cell tumour (Fig. 4A), while the corresponding one-vs-rest AUCs were 0.92, 0.92, and 0.84, respectively (Fig. 4B). Calibration curves indicated moderate agreement between predicted probabilities and observed frequencies overall, with less stable calibration for Sertoli cell tumour (Fig. 4C). The row-normalised confusion matrix showed class-specific recall of 0.82 for Seminoma, 0.83 for Leydig cell tumour, and 0.70 for Sertoli cell tumour,

indicating comparatively lower sensitivity for Sertoli cell tumour (Fig. 4D). WSI-level probability-aggregation results, together with ABMIL performance, are summarised in Table 3. Analysis of pooled out-of-fold predictions provided complementary estimates of case-level performance with 95% CIs. Logistic Regression with mean pooling achieved an accuracy of 0.85 (95% CI: 0.80–0.89), F1-macro of 0.80 (95% CI: 0.73–0.86), AUC-macro of 0.90 (95% CI: 0.86–0.94), and AP-macro of 0.85 (95% CI: 0.79–0.90). Logistic Regression with top-k pooling achieved an accuracy of 0.79 (95% CI: 0.74–0.85), F1-macro of 0.76 (95% CI: 0.69–0.82), AUC-macro of 0.90 (95% CI: 0.86–0.94), and AP-macro of 0.84 (95% CI: 0.78–0.90). For ABMIL, pooled out-of-fold performance was 0.80 for accuracy (95% CI: 0.75–0.86), 0.77 for F1-macro (95% CI: 0.71–0.83), 0.89 for AUC-macro (95% CI: 0.85–0.93), and 0.79 for AP-macro (95% CI: 0.71–0.86). Among the evaluated WSI-level approaches, Logistic Regression with mean pooling showed the highest numerical accuracy and F1-macro.

### 3.4. Grad-CAM visualisation

Grad-CAM heatmaps were computed for correctly classified tiles in the test set to visualise the regions that most influenced the model's predictions. Representative examples are shown in Fig. 5A–F, including the original H&E tiles for Seminoma, Leydig cell tumour, and Sertoli cell tumour (Fig. 5A–C) and the corresponding Grad-CAM overlays (Fig. 5D–F). Heatmap colour gradients represented the intensity of model attention, ranging from very high (red) to very low (blue). The highlighted regions corresponded to morphological structures relevant to tumour classification. In Seminoma (Fig. 5D), activations were mainly localised in dense clusters of germ cells and areas of nuclear crowding. In Leydig cell tumour (Fig. 5E), the network focused on large polygonal

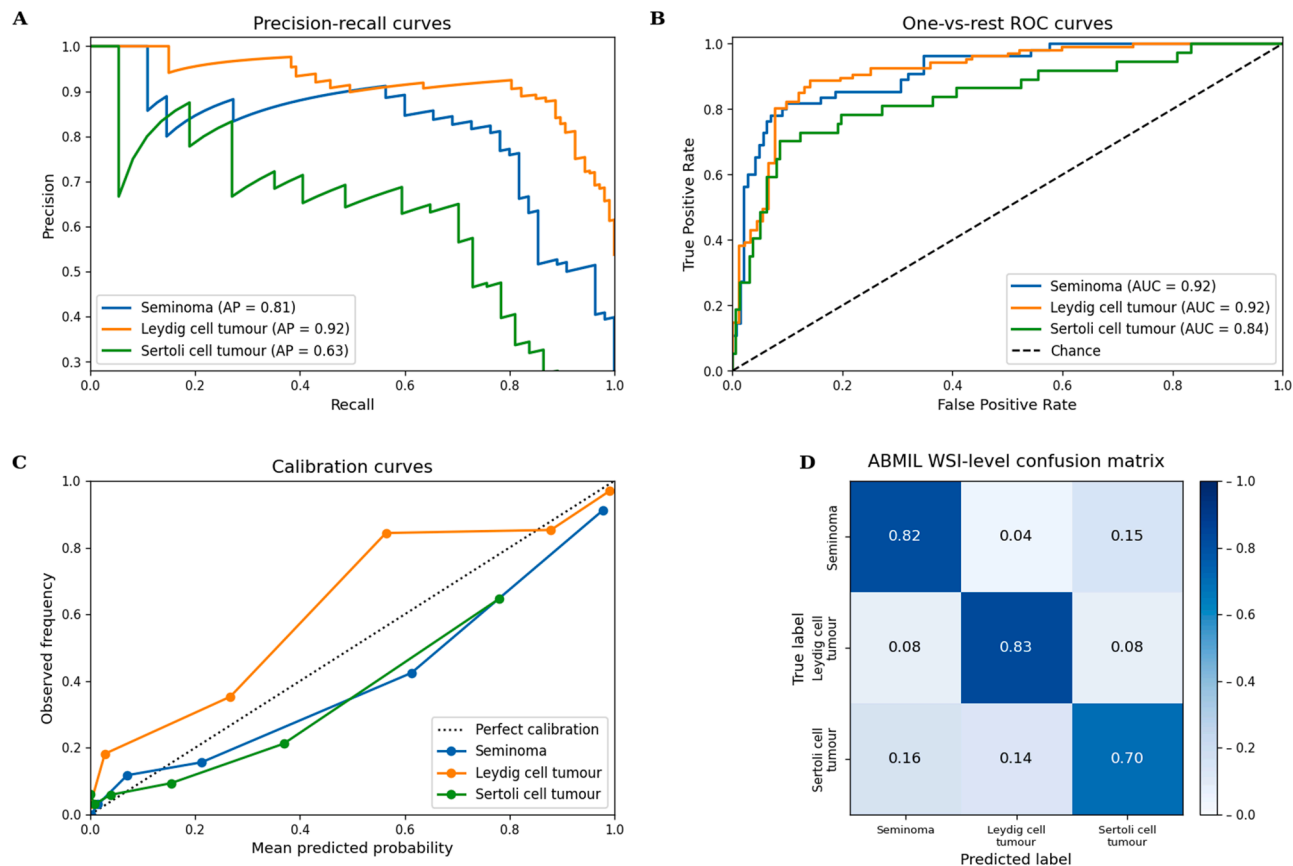


**Fig. 4.** WSI-level performance of the ABMIL model based on pooled out-of-fold predictions from the five held-out test folds. (A) Class-wise precision–recall curves for the three tumour classes. (B) Class-wise one-vs-rest ROC curves. (C) Class-wise calibration curves showing the agreement between predicted probabilities and observed outcomes. (D) Row-normalised confusion matrix derived from pooled out-of-fold WSI-level predictions. ABMIL: attention-based multiple instance learning; ROC: receiver operating characteristic; WSI: whole-slide image.

**Table 3**  
WSI-level performance of probability-aggregation and ABMIL approaches. Values are mean  $\pm$  SD across five test folds, followed by pooled out-of-fold estimates with 95% bootstrap confidence intervals.

| Model / aggregation                         | Input                      | Accuracy (mean $\pm$ SD)            | F1-macro (mean $\pm$ SD)            | AUC-macro (mean $\pm$ SD)           | AP-macro (mean $\pm$ SD)            |
|---------------------------------------------|----------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Logistic Regression (mean pooling)          | Tile probabilities         | 0.84 $\pm$ 0.07<br>0.85 (0.80–0.89) | 0.80 $\pm$ 0.10<br>0.80 (0.73–0.86) | 0.90 $\pm$ 0.06<br>0.90 (0.86–0.94) | 0.84 $\pm$ 0.08<br>0.85 (0.79–0.90) |
| Logistic Regression (top-k pooling, k = 50) | Tile probabilities         | 0.79 $\pm$ 0.11<br>0.79 (0.74–0.85) | 0.75 $\pm$ 0.10<br>0.76 (0.69–0.82) | 0.90 $\pm$ 0.06<br>0.90 (0.86–0.94) | 0.85 $\pm$ 0.08<br>0.84 (0.78–0.90) |
| ABMIL                                       | EfficientNet-B4 embeddings | 0.80 $\pm$ 0.04<br>0.80 (0.75–0.86) | 0.76 $\pm$ 0.06<br>0.77 (0.71–0.83) | 0.91 $\pm$ 0.03<br>0.89 (0.85–0.93) | 0.84 $\pm$ 0.05<br>0.79 (0.71–0.86) |

tumour cells with abundant eosinophilic cytoplasm and well-defined cell borders. In Sertoli cell tumour (Fig. 5F), activations were predominantly distributed along tubular and trabecular structures, with elongated, parallel-oriented nuclei located near luminal surfaces, corresponding to the typical organisation of Sertoli cells. An exploratory quantitative analysis performed on an annotated subset ( $n = 150$  tiles) was consistent with these observations, showing an average IoU of  $0.47 \pm 0.08$ , Pointing Game Accuracy =  $0.86 \pm 0.05$ , and Top 20% Energy inside ROI =  $0.73 \pm 0.07$  (Table 4). The randomisation sanity checks yielded IoU of  $0.09 \pm 0.04$ , supporting that the Grad-CAM activation patterns were related to learned morphological representations rather than random activations.

4. Discussion

Artificial Intelligence systems are increasingly being integrated into computational pathology workflows, with growing evidence of their

value for histopathological image analysis and classification (Campanella et al., 2019; Sarvamangala and Kulkarni, 2022; Wulcan et al., 2025). Recent studies have further extended AI-based image analysis to disease-specific applications in veterinary histopathology, highlighting its emerging role in diagnostic support (Zuraw and Aeffner, 2022). In this context, the present study demonstrated the feasibility of an AI-assisted computational pathology framework for the classification of Seminoma, Leydig cell tumour, and Sertoli cell tumour from digitised WSIs. By integrating complementary tile-level and WSI-level classification strategies within the same validation framework, the study enabled the assessment of both local morphological discrimination and whole-slide diagnostic classification. Overall, our results indicated that model performance depended on the level at which classification was assessed. End-to-end fine-tuning of EfficientNet-B4 improved tile-level discrimination compared with the feature-based approaches, whereas at the WSI level, aggregation of Logistic Regression probabilities provided the highest numerical case-level performance among the

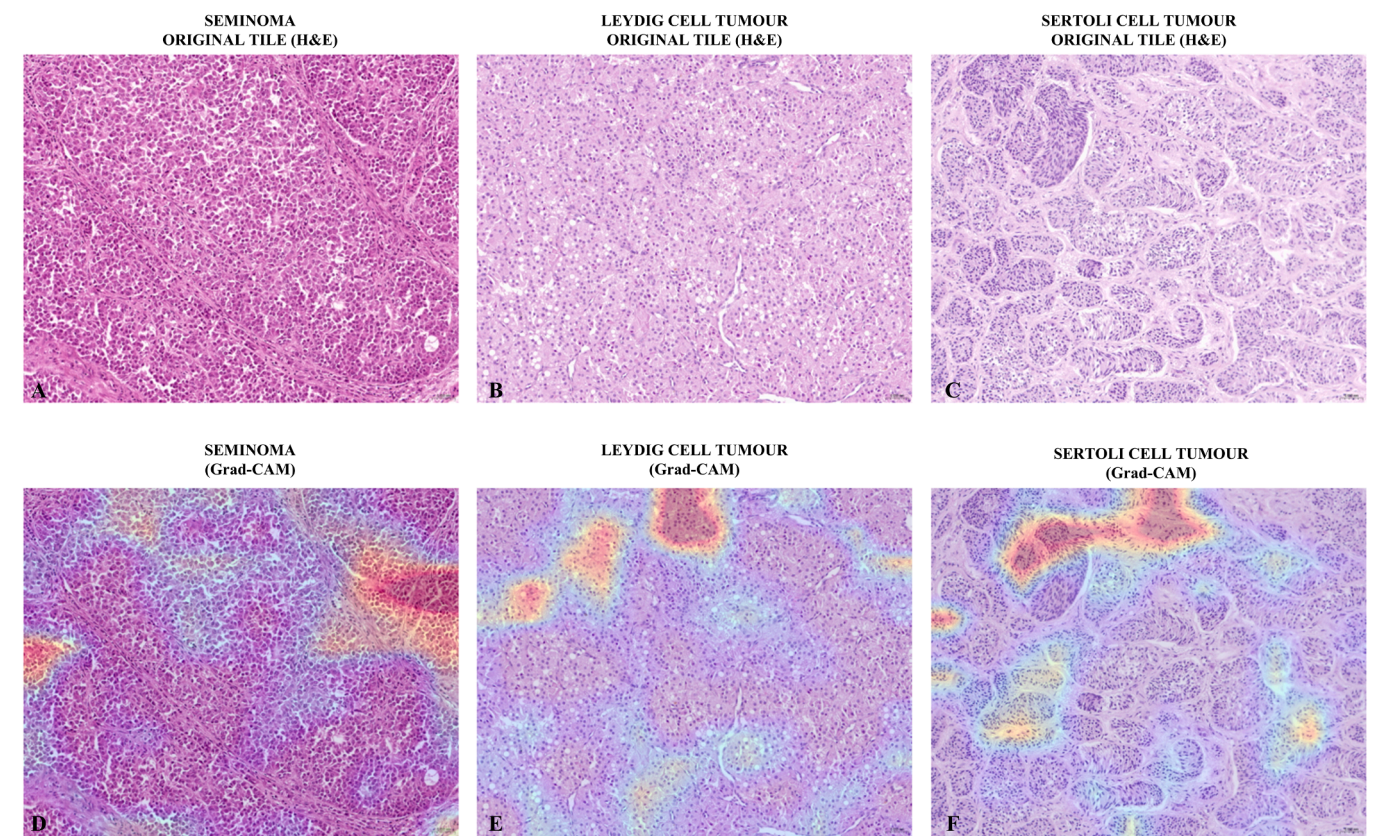


**Fig. 5.** Representative Grad-CAM visualisations for correctly classified tiles from the fine-tuned EfficientNet-B4 model. (A–C) Original H&E tiles for Seminoma, Leydig cell tumour, and Sertoli cell tumour. (D–F) Corresponding Grad-CAM overlays highlighting the regions that contributed most to the model predictions. Heatmap colours range from high activation (red/yellow) to low activation (green/blue). Grad-CAM: gradient-weighted class activation mapping; H&E: haematoxylin and eosin.

**Table 4**  
Quantitative Grad-CAM assessment on annotated tiles. Values are mean  $\pm$  SD across tumour classes (50 tiles per class). ROI, region of interest.

| Metric                                     | Definition                                                                                              | Mean $\pm$ SD   |
|--------------------------------------------|---------------------------------------------------------------------------------------------------------|-----------------|
| IoU (70th-percentile threshold)            | Intersection-over-Union between the Grad-CAM mask thresholded at the 70th percentile and the expert ROI | 0.47 $\pm$ 0.08 |
| Pointing Game Accuracy                     | Proportion of tiles in which the maximum activation lies within the ROI                                 | 0.86 $\pm$ 0.05 |
| Top 20% Energy inside ROI                  | Fraction of Grad-CAM energy within ROI                                                                  | 0.73 $\pm$ 0.07 |
| Randomised IoU (70th-percentile threshold) | IoU after model weight randomisation using the same 70th-percentile threshold                           | 0.09 $\pm$ 0.04 |

evaluated strategies. These findings supported the potential of AI-based methods as decision-support tools in veterinary histopathology, while emphasising the importance of evaluating performance at the WSI level.

The difference between tile-level and WSI-level performance also highlighted the importance of integrating information across multiple tissue regions when approaching histopathological classification at the case level. Rather than relying on isolated image patches, WSI-level aggregation incorporated predictions derived from a broader representation of tumour morphology. Similar case-level aggregation strategies have been reported in veterinary histopathology. [Fragoso-Garcia et al. \(2023\)](#) combined patch-wise predictions for slide-level classification of canine skin tumours, whereas [Defourny et al. \(2026\)](#) aggregated tile-level CNN predictions to obtain biopsy-level diagnoses in feline intestinal disease.

To further investigate which tissue regions contributed to the CNN predictions, Grad-CAM heatmaps were evaluated. Class-specific activation patterns spatially coincided with characteristic histomorphological features of the three tumour types, including germ-cell-rich regions in Seminoma, polygonal eosinophilic neoplastic cells in Leydig cell tumours, and tubular or trabecular architecture in Sertoli cell tumours ([Samek et al., 2021](#); [Selvaraju et al., 2017](#)). This spatial concordance, further supported by the exploratory quantitative analysis, suggested that model inference was informed, at least in part, by diagnostically relevant architectural and cytomorphological features.

In addition to the tile-level CNN framework, we explored a complementary weakly supervised ABMIL strategy operating directly at the WSI level ([Ilse et al., 2018](#); [Lu et al., 2021](#)). Under the same 5-fold WSI-level cross-validation design, ABMIL showed competitive WSI-level performance, supporting the feasibility of slide-level classification using attention-based aggregation of deep morphological embeddings without tile-level supervision ([Ilse et al., 2018](#); [Lu et al., 2021](#)). Notably, within the present single-centre dataset, simple WSI-level aggregation of tile probabilities from Logistic Regression achieved higher numerical case-level accuracy than ABMIL. This suggested that, under relatively homogeneous acquisition conditions, averaging tile-level predictions may provide an effective slide-level decision rule in this setting.

Taken together, these findings emphasised that tile-level and WSI-level analyses provide complementary information, with the latter being more directly aligned with the diagnostic assessment performed in veterinary pathology. In this setting, ABMIL provided a complementary weakly supervised WSI-level classification strategy without requiring individual tile annotations ([Ilse et al., 2018](#); [Lu et al., 2021](#)). However, several limitations should be acknowledged, including the single-centre

design, the homogeneous staining and scanning conditions, and the absence of an independent external validation cohort. Despite these limitations, to the best of our knowledge, this study represents the first application of an AI-based computational pathology framework to the histopathological classification of canine testicular tumours.

From a veterinary pathology perspective, the WSI-level results more closely reflect the diagnostic unit routinely evaluated by pathologists than predictions generated from individual image tiles. The ability of the proposed framework to distinguish among the three major canine testicular tumour types at the slide level therefore provides initial evidence that computational approaches may support histopathological assessment at the case level. In this context, AI should be considered a potential complementary decision-support tool rather than a replacement for pathological evaluation. Its future integration into digital pathology workflows could provide additional support in cases showing overlapping morphological features, while the spatial information provided by approaches such as Grad-CAM may help relate model predictions to histologically recognisable tumour patterns. Therefore, given the single-centre design of this dataset and the absence of external validation, this AI-based framework should currently be regarded strictly as an experimental proof-of-concept for research use, rather than a clinical decision-support tool or automatic diagnostic system. Future studies should focus on expanding multi-institutional datasets to reduce bias and improve generalisation, and on including a blinded ring test involving veterinary pathologists with varying levels of experience.

## 5. Conclusions

The present study supports the feasibility of AI-based histopathological classification of canine testicular tumours from WSIs. The fine-tuned EfficientNet-B4 achieved the strongest tile-level performance, whereas Logistic Regression with slide-level probability aggregation yielded the highest WSI-level accuracy among the evaluated WSI-level approaches. While this study demonstrates the technical feasibility of WSI-level classification, the reliance on a single-centre dataset and the lack of external validation preclude any immediate clinical applicability. Consequently, this work must be regarded as a foundational proof-of-concept rather than a diagnostic system. Future studies must prioritise validation across diverse, multi-institutional datasets with heterogeneous scanning protocols, alongside blinded ring tests involving veterinary pathologists with varying expertise. From a One Health perspective, integrating molecular and clinical data with morphological features may further advance precision pathology and translational insight into shared oncogenic mechanisms in human and veterinary medicine.

## CRedit authorship contribution statement

**Orlando Paciello:** Writing – review & editing, Supervision, Resources, Conceptualization. **Evaristo Di Napoli:** Resources. **Maria De Falco:** Writing – review & editing. **Elia Formisano:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Conceptualization. **Pietro Formisano:** Writing – review & editing. **Luigi Rosati:** Writing – review & editing, Supervision. **Lorenzo Riccio:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

## Ethics declaration

This study was conducted in accordance with the following guidelines for animal welfare and/or reporting: European Directive 2010/63/EU. This study was approved by the Ethical Animal Care and Use Committee of the University of Naples Federico II. (Approval No. PG/2023/0029409)

## Ethics statement

Ethical approval was granted by the Ethical Animal Care and Use Committee of the University of Naples Federico II (Protocol No. PG/2023/0029409; approval date: 10 March 2023). All samples were obtained during routine castration procedures, with no additional procedures or euthanasia performed for the purposes of this study.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors used generative AI tools for language refinement and clarity improvement. The authors reviewed and edited the content as necessary.

## Funding

The authors received no specific funding for this work.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors thank Aleksey Costagliola for critical technical support in slide scanning and image digitisation. The authors also gratefully acknowledge the Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, and its staff for the welcoming and stimulating research environment provided during the research stay. Lorenzo Riccio particularly acknowledges Ana-Maria Goré, Jelle Bonthuis, and Marcelo Malisano Barreto Filho for their friendship and valuable personal support during his time in Maastricht.

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## References

- Abels, E., Pantanowitz, L., Aeffner, F., Zarella, M.D., van der Laak, J., Bui, M.M., Vemuri, V.N.P., Parwani, A.V., Gibbs, J., Agosto-Arroyo, E., Beck, A.H., Kozlowski, C., 2019. Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. *J Pathol* 249, 286–294. <https://doi.org/10.1002/path.5331>.
- Aeffner, F., Zarella, M.D., Buchbinder, N., Bui, M.M., Goodman, M.R., Hartman, D.J., Lujan, G.M., Molani, M.A., Parwani, A.V., Pantanowitz, L., 2019. Introduction to digital image analysis in whole-slide imaging: a white paper from the Digital Pathology Association. *J Pathol Inform* 10, 9. <https://doi.org/10.4103/jpi.jpi.82.18>.
- Ahmed, A.A., Abouzid, M., Kaczmarek, E., 2022. Deep learning approaches in histopathology. *Cancers* 14, 5264. <https://doi.org/10.3390/cancers14215264>.
- Amaral, C.L., Langohr, I.M., Giaretta, P.R., Ecco, R., 2024. Digital pathology and artificial intelligence in veterinary medicine. *Braz J Vet Pathol* 17, 147–151. <https://doi.org/10.24070/bjvp.1983-0246.v17i3p147-151>.
- Ancheta, K., Le Calvez, S., Williams, J., 2024. The digital revolution in veterinary pathology. *J Comp Pathol* 214, 19–31. <https://doi.org/10.1016/j.jcpa.2024.08.001>.
- Aubreville, M., Bertram, C.A., Marzahl, C., Gurtner, C., Dettwiler, M., Schmidt, A., Bartenschlager, F., Merz, S., Fragos, M., Kershaw, O., Klopffleisch, R., Maier, A., 2020. Deep learning algorithms out-perform veterinary pathologists in detecting the mitotically most active tumor region. *Sci Rep* 10, 16447. <https://doi.org/10.1038/s41598-020-73246-2>.
- Balakrishna, R., Patra, S., Kaur, C., Aljawarneh, N.M., G., S., R., D., 2025. Optimizing personalized cancer treatment plans using deep reinforcement learning in IoT-enabled healthcare systems. In: 2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE), pp. 1–8. IEEE. <https://doi.org/10.1109/ICICKE65317.2025.11136234>.
- Battazza, A., Brasileiro, F.C.S., Tasaka, A.C., Bulla, C., Ximenes, P.P., Hosomi, J.E., Rocha, N.S., 2024. Integrating telepathology and digital pathology with artificial intelligence: an inevitable future. *Vet World* 17, 1667–1671. <https://doi.org/10.14202/vetworld.2024.1667-1671>.

- Breiman, L., 2001. Random forests. *Mach Learn* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Campanella, G., Hanna, M.G., Geneslaw, L., Mirafior, A., Werneck Krauss Silva, V., Busam, K.J., Brogi, E., Reuter, V.E., Klimstra, D.S., Fuchs, T.J., 2019. Clinical-grade computational pathology using weakly supervised deep learning on whole-slide images. *Nat Med* 25, 1301–1309. <https://doi.org/10.1038/s41591-019-0508-1>.
- Canny, J., 1986. A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell* 8, 679–698. <https://doi.org/10.1109/TPAMI.1986.4767851>.
- Coudray, N., Ocampo, P.S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyo, D., Moreira, A.L., Razavian, N., Tsirigos, A., 2018. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 24, 1559–1567. <https://doi.org/10.1038/s41591-018-0177-5>.
- d'Aquino, I., Riccio, L., Piegari, G., Ambrosio, N., Longobardi, C., Ciarcia, R., Cortese, L., Di Napoli, E., Paciello, O., Russo, V., 2026. Association between urinary cotinine and whole-slide digital cytomorphometric alterations in the oral mucosa of tobacco smoke-exposed cats. *Vet Sci* 13, 354. <https://doi.org/10.3390/vetsci13040354>.
- d'Aquino, I., Piegari, G., Miletto, G., Sannino, E., Costanza, D., Meomartino, L., Paciello, O., 2023. Morphometrical and immunohistochemical evaluation of kidney as an indirect parameter to estimate age in puppies in veterinary forensic pathology. *Animals* 13, 2665. <https://doi.org/10.3390/ani13162665>.
- Defourny, S.V.P., Sabattini, S., Wright, M., Faroni, E., Crosby-Durrani, H., Ricci, E., Ressel, L., 2026. Application of convolutional neural networks for histopathological diagnosis of feline low-grade T-cell lymphoma and lymphoplasmacytic enteritis in intestinal biopsies. *Front Vet Sci* 13, 1851689. <https://doi.org/10.3389/fvets.2026.1851689>.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE. <https://doi.org/10.1109/CVPR.2009.5206848>.
- Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems. *Ann Eugen* 7, 179–188. <https://doi.org/10.1111/j.1469-1809.1936.tb02137.x>.
- Fotouh, A., 2025. Digitalization of veterinary pathology in the era of artificial intelligence: a comprehensive review. *Benha Vet Med J* 49, 100–110. <https://doi.org/10.21608/bvmj.2025.416499.1980>.
- Fragoso-Garcia, M., Wilm, F., Bertram, C.A., Merz, S., Schmidt, A., Donovan, T., Klopfeisch, R., 2023. Automated diagnosis of 7 canine skin tumors using machine learning on H&E-stained whole slide images. *Vet Pathol* 60, 865–875. <https://doi.org/10.1177/03009858231189205>.
- Gertych, A., Swiderska-Chadaj, Z., Ma, Z., Ing, N., Markiewicz, T., Cierniak, S., Knudsen, B.S., 2019. Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides. *Sci Rep* 9, 1483. <https://doi.org/10.1038/s41598-018-37638-9>.
- Grieco, V., Riccardi, E., Greppi, G.F., Teruzzi, F., Iermano, V., Finazzi, M., 2008. Canine testicular tumours: a study on 232 dogs. *J Comp Pathol* 138, 86–89. <https://doi.org/10.1016/j.jcpa.2007.11.002>.
- Hosmer Jr., D.W., Lemeshow, S., Sturdivant, R.X., 2013. third ed. *Applied Logistic Regression*. John Wiley & Sons.
- Ilse, M., Tomczak, J., Welling, M., 2018. Attention-based deep multiple instance learning. *International Conference on Machine Learning*. PMLR, pp. 2127–2136.
- Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kumar, U.P., Lakshmi, E.V., Kumar, G.S., Bhattacharjee, P., Gokulakrishnan, S., Sajiv, G., 2026. Contrast enhancement in mammogram images using CLAHE and CNN-based approaches. 2026 6th International Conference on Expert Clouds and Applications (ICOECA). IEEE, pp. 756–761. <https://doi.org/10.1109/ICOECA68095.2026.11485574>.
- Lu, M.Y., Williamson, D.F., Chen, T.Y., Chen, R.J., Barbieri, M., Mahmood, F., 2021. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng* 5, 555–570. <https://doi.org/10.1038/s41551-020-00682-w>.
- Mosier, J.E., 1989. Effect of aging on body systems of the dog. *Vet Clin North Am Small Anim Pract* 19, 1–12. [https://doi.org/10.1016/S0195-5616\(89\)50001-9](https://doi.org/10.1016/S0195-5616(89)50001-9).
- Orlandi, R., Vallesi, E., Boiti, C., Polisca, A., Bargellini, P., Troisi, A., 2022. Characterization of testicular tumor lesions in dogs by different ultrasound techniques. *Animals* 12, 210. <https://doi.org/10.3390/ani12020210>.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 9, 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>.
- Piegari, G., d'Aquino, I., Salanti, G.V., Romano, V., Miletto, G., Sannino, E., Paciello, O., 2024. Pathological changes and sodium rhodizonate test as tools for investigating gunshot wounds in veterinary forensic pathology. *Animals* 14, 2913. <https://doi.org/10.3390/ani14192913>.
- Ponzio, F., Urgese, G., Ficarra, E., Di Cataldo, S., 2019. Dealing with lack of training data for convolutional neural networks: the case of digital pathology. *Electronics* 8, 256. <https://doi.org/10.3390/electronics8030256>.
- Power, K., Riccio, L., Sciliano, A., Santorufu, L., Rosati, L., Landi, S., Maisto, G., 2026. Impact of conventional vs. biodegradable and compostable microplastics on *Eisenia fetida* S.: an ecopathological approach. *Environ Toxicol*. <https://doi.org/10.1002/tox.70030>.
- Prasad, D., Khandeshi, A., Sartin, S., Jain, R., Dahdaleh, N., Lesniak, M., Ahuja, C., 2025. Will AI become our Co-PI? *npj Digit Med* 8, 440. <https://doi.org/10.1038/s41746-025-01859-w>.
- Rao, C.R., 1948. The utilization of multiple measurements in problems of biological classification. *J R Stat Soc Ser B* 10, 159–193. <https://doi.org/10.1111/j.2517-6161.1948.tb00008.x>.
- Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P., 2001. Color transfer between images. *IEEE Comput Graph Appl* 21, 34–41. <https://doi.org/10.1109/38.946629>.
- Riccio, L., Chianese, T., Mileo, A., Balsamo, S., Sciarillo, R., Gatta, R., Capaldo, A., 2025. Changes of CB1 receptor expression in tissues of cocaine-exposed eels. *Animals* 15, 1734. <https://doi.org/10.3390/ani15121734>.
- Riccio, L., De Falco, M., Piscopo, M., Lettieri, G., Musella, F., Trinchese, G., Ambrosio, N., Spada, A., Di Napoli, E., Paciello, O., Rosati, L., 2026a. Decoding canine male reproductive dysfunction: a multi-level assessment of sperm quality in dogs from the 'Land of Fires' (Campania, Italy). *Anim. Reprod. Sci.* 294, 108335. <https://doi.org/10.1016/j.anireprosci.2026.108335>.
- Riccio, L., De Falco, M., Trifuoggi, M., Cuculo, C., Conte, F., Chianese, T., Rosati, L., 2026b. Testicular concentrations of metals and metalloids in dogs from the 'Land of Fires' (Campania Region, Italy): a preliminary morphological, chemical and molecular study. *Environ Toxicol Pharmacol*, 105099. <https://doi.org/10.1016/j.etap.2026.105099>.
- Rogers, L., Galezowski, A., Ganshorn, H., Goldsmith, D., Legge, C., Waine, K., Davies, J. L., 2024. The use of telepathology in veterinary medicine: a scoping review. *J Vet Diagn Invest* 36, 490–497. <https://doi.org/10.1177/10406387241241270>.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature* 323, 533–536. <https://doi.org/10.1038/323533a0>.
- Sajiv, G., Meenakshisundaram, N., 2025. An improved diagnostic process to identify breast cancer in earlier stages using deep learning mechanism. In: 2025 8th International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1637–1642. IEEE. <https://doi.org/10.1109/ICOEI65986.2025.11013687>.
- Sajiv, G., Ramkumar, G., 2023a. Deep learning based breast cancer classification using artificial neural network on histopathological images. In: 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), pp. 1–7. IEEE. <https://doi.org/10.1109/ICONSTEM56934.2023.10142911>.
- Sajiv, G., Ramkumar, G., 2023b. An efficient machine learning model for breast cancer categorization using logistic regression on histopathological images. In: 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), pp. 1–7. IEEE. <https://doi.org/10.1109/ICONSTEM56934.2023.10142781>.
- Sajiv, G., Ramkumar, G., 2023c. Analysis of breast cancer recognition in histopathological images using convolutional neural network. In: 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), pp. 1–8. IEEE. <https://doi.org/10.1109/ICECONF57129.2023.10084065>.
- Samek, W., Montavon, G., Lapuschkin, S., Anders, C.J., Müller, K.-R., 2021. Explaining deep neural networks and beyond: a review of methods and applications. *Proc IEEE* 109, 247–278. <https://doi.org/10.1109/JPROC.2021.3060483>.
- Sarvamangala, D.R., Kulkarni, R.V., 2022. Convolutional neural networks in medical image understanding: a survey. *Evol Intell* 15, 1–22. <https://doi.org/10.1007/s12065-020-00540-3>.
- Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* 618–626. <https://doi.org/10.1109/ICCV.2017.74>.
- Sivakumar, A., S. L.S., Suganthi, M., Gunasundari, B., Sivamurugan, S., G. S., 2023. A robust and novel hybrid deep learning based lung nodule identification on CT scan images. In: 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), pp. 1–8. IEEE. <https://doi.org/10.1109/ICECONF57129.2023.10083700>.
- Tan, M., Le, Q.V., 2019. EfficientNet: rethinking model scaling for convolutional neural networks. In: *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pp. 6105–6114. PMLR.
- Wulcan, J.M., Giarretta, P.R., Fingerhoo, S., de Brot, S., Crouch, E.E.V., Wolf, T., Keller, S.M., 2025. Artificial intelligence-based quantification of lymphocytes in feline small intestinal biopsies. *Vet Pathol* 62, 139–151. <https://doi.org/10.1177/03009858241286828>.
- Xiao, S., Dhand, N.K., Wang, Z., Hu, K., Thomson, P.C., House, J.K., Khatkar, M.S., 2025. Review of applications of deep learning in veterinary diagnostics and animal health. *Front Vet Sci* 12, 1511522. <https://doi.org/10.3389/fvets.2025.1511522>.
- Zappulli, V., Moccia, V., Torrigiani, F., Molinari, A., Detillo, P., Gola, C., Bussolati, B., 2026. Non-toxic acid-free glyoxal fixative for veterinary gross specimen preservation, histopathology, immunohistochemistry, and molecular analysis. *Vet Pathol* 63, 282–295. <https://doi.org/10.1177/03009858251372572>.
- Zuraw, A., Aeffner, F., 2022. Whole-slide imaging, tissue image analysis, and artificial intelligence in veterinary pathology: an updated introduction and review. *Vet Pathol* 59, 6–25. <https://doi.org/10.1177/03009858211040484>.