Bridging Species with AI: Transfer Learning for Fetlock Fracture Detection in Racehorses

Hanya Tamer Ahmed[1], Dagmar Berner[1], Kristien Verheyen[1], Francisco Llabres-Diaz[1], Vanessa G. Peter[2], Qianni Zhang[3], Yu-Mei Chang[1]

[1]Royal Veterinary College, [2] Rossdales Equine Diagnostic Centre, [3]Queen Mary University of London

Journal

Lay Summary

Equine joint (fetlock) fracture is a leading cause of injury in Thoroughbred racehorses, significantly impacting both their welfare and racing careers. Approx. 10% of racehorses sustain a fracture during training, with bone injuries occurring at a rate of 1.3 per 1000 starts in flat racing. However, despite their impact, subtle incomplete fractures can be difficult to diagnose. Assessment relies largely on X-ray images, yet variations in image quality and projection, combined with the difficulty of identifying subtle bone changes, can limit diagnostic accuracy.

Improving early and reliable fracture detection is therefore critical to support horse welfare and prevent fatal injuries. This study led by the Royal Veterinary College marks the first step in a longer term research programme focused on identifying early bone changes before they progress to serious fractures.

Methodology

Current dataset composition:

- **~4,000 human** images to enhance generalisability (public MURA & CheXpert datasets curated by Stanford University)
- **~100 equine** cases (sourced from two equine hospital archives and public databases such as PubMed).
- **~70 feline** cases (from Queen Mother Hospital for Animals) used for validation.

We developed a three stage AI pipeline that identifies the scan type, recognises the joint's position, and detects fractures. This 'hierarchical' design is what allows the model to work accurately across different species.

Module 1: Modality Identification *(Determines the imaging modality, e.g., X-rays, CT and MRI)* → **Module 2: Projection Classification** *(Classifies the projection or anatomical view of the image)* → **Module 3: Fracture Detection** *(Detects fractures based on the image input)*

Data Curation & Preprocessing:

- Preliminary dataset consisted of equine fetlock images (MRI, CT, X-rays) and a large dataset from human medical imaging (MRI, CT, X-rays). To enhance model robustness, data were collected from different manufacturers and included different projections and with varying levels of noise.
- The dataset also incorporated cases with **obvious fractures** to establish a **baseline**.
- Further validation and testing was performed using feline imaging (X-rays) datasets.

Hierarchical Classification for Equine:

- ***Modality Classification:*** ViT + ResNet model classifies the image as MRI, CT or X-ray using a specialised loss function to handle imbalanced datasets.
- ***Region Classification:*** For X-rays, a secondary classification step determines if the image is the targeted joint.
- ***View Identification:*** The model then identifies the projection dorsopalmar, lateromedial, and oblique to improve diagnostic consistency.

Fracture Detection:

- A SqueezeNet-ViT model is used for detecting fractures within the joint (fetlock).

Results

The study revealed that the AI system was able to detect and localise fractures in horses using knowledge gained from thousands of human fracture images. This approach, known as “transfer learning”, enabled the model to be trained on a large human dataset before being adapted for veterinary use.

As a result, the system achieved fracture localisation accuracy of between 71% and 84% without requiring an unrealistically large number of equine images. These findings confirm that our 'cross species' approach can provide a robust, adaptable tool for improving diagnostic speed and accuracy in veterinary clinics.

Table 1: Quantitative results of fracture detection proposed method for equine and feline datasets using IoU (Intersection over Union) and accuracy.

Equine Dataset

Dataset	IoU	Sensitivity	Specificity	Accuracy
Hospital-acquired	0.76 [0.72–0.80]	92.3% [89.5–94.1]	86.4% [82.1–88.7]	89.4% [85.8–91.4]
Literature-derived	0.71 [0.65–0.75]	84.0% [81.2–86.8]	78.9% [75.3–82.4]	81.5% [78.3–84.6]

Feline & Human Datasets

Dataset	IoU	Sensitivity	Specificity
Human	0.79 [0.75–0.82]	95.2% [93.0–97.1]	92.8% [90.4–94.9]
Feline	0.74 [0.70–0.77]	87.6% [85.0–89.8]	84.2% [81.5–86.7]

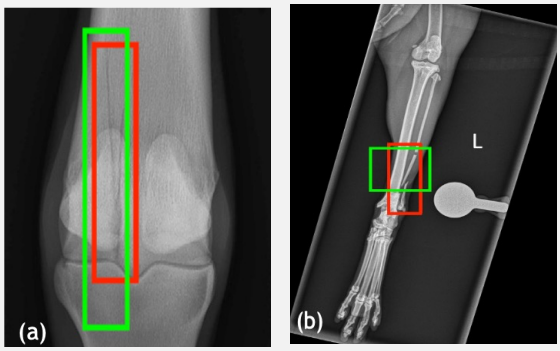

Figure 1: Qualitative results of detection proposed method (a) for equine (b) feline datasets in which green boxes are ground truth and red boxes are predicted.

Discussion and Conclusion

- The findings demonstrate the potential for AI-assisted tools to strengthen fracture diagnosis across veterinary practice.
- Faster and more reliable detection could help reduce uncertainty in clinical decision making and enable earlier treatment, with clear benefits for the welfare and recovery of racehorses and companion animals.

Future Work:

- **Focus on Complex Cases:** Emphasise detecting subtle fracture and pre-fracture abnormalities.

Acknowledgements

- This research was funded by the Horserace Betting Levy Board