Lightweight Classification of Canine Eye Diseases

Isselmou Abdarahmane[1] **Peter M. Roth** [2,3]
isselmou.abdarahmane@fhwn.ac.at peter.roth@webster.ac.at
[1]Medial University of Vienna
[2]University of Veterinary Medicine, Vienna
[3]Webster Vienna Private University

Abstract

Eye diseases in dogs are visually similar and difficult to distinguish without professional examination. Furthermore, assessing the severity of such conditions – and in particular determining whether immediate veterinary attention is required – poses a significant challenge for pet owners. To address this problem, we aim to assist pet owners in performing an initial triage of canine eye conditions with minimal technical expertise required. Given a photograph of the affected eye taken with a smartphone, we provide an automated preliminary assessment indicating whether a veterinary visit is advisable. To this end, we employ a computationally efficient convolutional neural network (CNN) to classify the images, identifying potential conditions and reporting the result to the user.

1 Introduction and Problem Statement

Veterinary ophthalmology is a well-studied field, e.g., [4, 5, 9, 13], However, as eye diseases in dogs are visually similar, they are difficult to distinguish laypersons. Furthermore, assessing the severity of such conditions – and in particular determining whether immediate veterinary attention is required – poses a significant challenge for pet owners. To address this problem, we aim to assist pet owners in performing an initial triage of canine eye conditions with minimal technical expertise required. Given a photograph of the affected eye taken with a smartphone, we provide an automated preliminary assessment indicating whether a veterinary visit is advisable. To this end, we employ a computationally efficient convolutional neural network (CNN) to classify the images, identifying potential conditions and reporting the result to the user. Specifically, we target four common canine ocular conditions: corneal oedema, episcleral hyperaemia, epiphora, and cherry eye.

The automated analysis of medical images using deep learning has seen considerable progress in recent years, with convolutional neural networks (CNNs) achieving strong performance across a range of diagnostic tasks, including the classification of diabetic retinopathy, skin lesions, and pathological findings in radiology [6, 10]. More recently, these methods have begun to be applied in veterinary medicine, where they offer the potential to assist practitioners and pet owners alike [1].

In the context of ocular disease, deep learning-based approaches have been explored primarily for human ophthalmology, with well-established benchmarks for conditions such as glaucoma and age-related macular degeneration [10]. Comparable work in veterinary ophthalmology remains sparse. A closely related study [7] developed CNN-based models to classify the severity of corneal ulcers in dogs from photographic images, using GoogLeNet, ResNet, and VGGNet architectures fine-tuned on ImageNet, achieving classification accuracies above 90 %. More recently, a U-Net-based approach was proposed for the segmentation and diagnosis of multiple canine ocular conditions from smartphone and camera images [2], demonstrating the feasibility of image-based triage tools for pet owners.

The Third Austrian Symposium on AI and Vision (AIROV26).

Lightweight CNN architectures such as MobileNetV2 [12] have been specifically designed for deployment on resource-constrained devices, making them well suited for mobile applications. Transfer learning from large-scale datasets such as ImageNet has been shown to yield strong results even when task-specific training data is limited [11], which is particularly relevant in veterinary contexts where large labeled datasets are rarely available.

In this work, we adopt MobileNetV2 as a backbone and fine-tune it on the *DogEyeSeg4* dataset for the classification of four canine ocular conditions. To the best of our knowledge, this represents one of the first attempts to apply mobile-efficient deep learning specifically to the triage of canine eye diseases from smartphone images. The rest of the paper is organized as follows: Sec. 2 describes the relevant technical implementation. Sec. 3 presents the experimental setup and results, and Sec. 4 summaries the findings and outlines directions for future work.

2 Implementation Details

The envisioned application requires a solution that is accurate enough to provide a meaningful initial assessment, yet lightweight enough to run on a standard smartphone without specialized hardware. To meet these constraints, the classification system is implemented in Python and built upon a MobileNetV2 backbone [12] pre-trained on ImageNet [3]. The backbone weights are kept frozen during training, and only the custom classification head is trained, which reduces the number of trainable parameters and mitigates overfitting on the small available dataset. The design prioritizes computational efficiency, requiring no specialized hardware for either training or inference, and is therefore well suited for eventual deployment as a mobile application.

The input images are rescaled to a uniform size of 224×224 pixels with three color channels (RGB), matching the input format expected by MobileNetV2. The frozen backbone extracts a feature vector of dimensionality 1280 via Global Average Pooling, which collapses the final $7 \times 7 \times 1280$ feature map into a compact representation. This feature vector is then passed through a custom classification head consisting of two fully connected layers with ReLU activations, batch normalization, and Dropout regularization, before a final Softmax layer produces a probability distribution over the four target classes: corneal oedema, episcleral congestion, epiphora, and cherry eye.

The network is trained using the Adam optimizer [8] with a learning rate of 0.0001 and a batch size of 12. To prevent overfitting, early stopping is applied, terminating training once no further improvement on the validation set is observed; the maximum number of epochs is set to 250. The available data is partitioned into $70\ \%$ for training, $10\ \%$ for validation, and $20\ \%$ for testing.

3 Experimental Results

To demonstrate the functionality of the proposed approach, we run it on a publicly available dataset, namely *DogEyeSeg4_dataset1*[1]. This dataset comprises 145 images at a resolution of 320×320 pixels, depicting the following conditions:

- Corneal oedema
- Episcleral hyperaemia
- Epiphora (excessive tearing)
- Cherry eye (nictitating gland prolapse)

The results achieved over the course of training are shown in Fig. 1. The left panel displays the trajectory of the loss function – the optimization criterion being minimized – while the right panel shows the development of classification accuracy. In both cases, it is evident that performance on the independent validation data, which was not accessible to the model during training, improves with increasing epochs before eventually plateauing. By employing early stopping, training is automatically halted at this point, as no further improvement on the validation data is expected.

[1] https://pmc.ncbi.nlm.nih.gov/articles/PMC11467576/

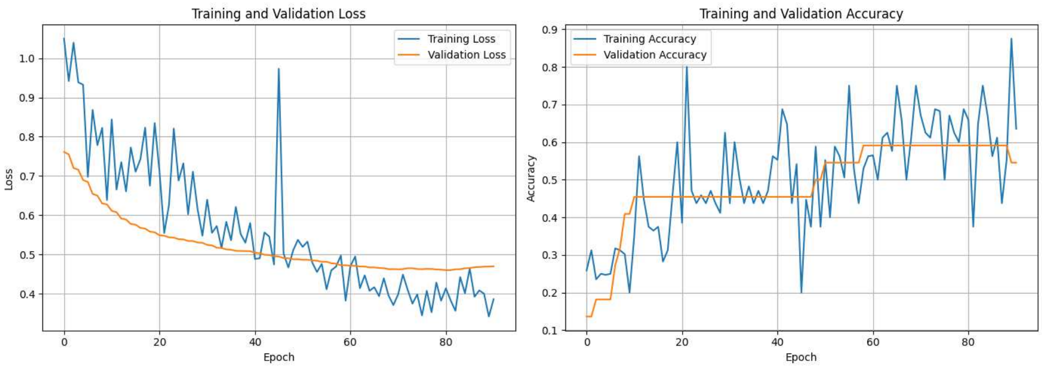


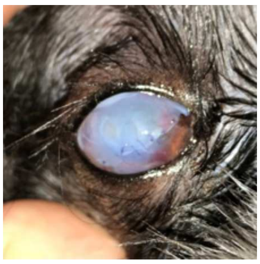
Figure 1: Loss curves and classification accuracy over training epochs.

The class-specific results are presented in Table 1. Three metrics are reported: the *Precision* describes the proportion of correctly classified positive images among all positively predicted images, indicating how reliable a positive prediction is. The *Recall* gives the proportion of correctly identified positive images among all truly positive images, measuring how completely the model detects each class. The *F1-score* is the harmonic mean of precision and recall, combining both metrics into a single value that accounts for both false positives and false negatives. Additionally, Fig. 2 shows illustrative examples of correctly and incorrectly classified images. From the results in Table 1 and the training curves in Fig. 1, it can be seen that the detection accuracy needs to be increased and the training behavior is still somewhat noisy. Both aspects can be addressed by extending the training set.

Table 1: Quantitative results per class.

Condition	Precision	Recall	F1-Score
Corneal Edema	0.71	0.68	0.69
Episcleral Congestion	0.89	0.86	0.86
Epiphora	0.59	0.59	0.59
Cherry Eye	0.74	0.82	0.78

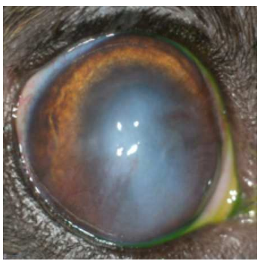
(a) Epiphora/**Epiphora**



(b) Cor. Edema/**Cor. Edema**

(c) Cor. Edema/**Epiphora**

(d) Cherry Eye/**None**



(e) Cor. Edema/**Cor. Edema**

(f) Cor. Edema/**Cor. Edema**

Figure 2: Illustrative classification results: true label / predicted label.

4 Conclusion

The goal of this study was to evaluate whether a lightweight classification system could be applied to automatically triage canine eye diseases, enabling pet owners without veterinary expertise to determine whether a visit to a veterinary practice is necessary. Our results demonstrate that the proposed approach achieves promising results across four common canine ocular conditions, confirming that the approach is generally feasible. However, as the available dataset is limited, the current performance does not yet meet the threshold required for clinical deployment. The next steps therefore include extending the dataset and applying domain-specific data augmentation to improve training robustness. In parallel, a modular client-server architecture will be developed, in which a lightweight mobile application captures an eye photograph alongside a short symptom survey, securely transmits the data to a centralized backend for joint image-survey classification, and returns an urgency-aware recommendation to the user.

Acknowledgments and Disclosure of Funding

This work was supported by the State of Lower Austria through the project *HOLSTEIN (Holistic approach to the sustainable provision of livestock health in Lower Austria).*

References

[1] J. Alves, T. Banzato, C. Bertram, S. Boroffka, B. Broux, F. Cian, A. Csomos, R. Drees, S. Goericke-Pesch, R. Gutierrez-Quintana, M. Hafner, A. Haghofer, R. Klopfleisch, J.-G. Kresken, I. Lautenschlaeger, I. Lieske, S. Meller, K. Mie, J. Nessler, C. Ober, G. Scharf, D. Scharner, M. Schmidt, A. Sewell, A. Tipold, T. Wrzesinski, and F. K. Zeugswetter. Review of applications of deep learning in veterinary diagnostics and animal health. *Frontiers in Veterinary Science*, 12:1511522, 2025.

[2] M. Buric, S. Grozdanic, and M. Ivasic-Kos. Diagnosis of ophthalmologic diseases in canines based on images using neural networks for image segmentation. *Heliyon*, 10(19):e38287, 2024.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A large-scale hierarchical image database. In *Conference on Computer Vision and Pattern Recognition*, 2009.

[4] D. W. Esson and S. Calvarese, editors. *Clinical Atlas of Canine and Feline Ophthalmic Disease*. Wiley-Blackwell, Hoboken, NJ, 2022.

[5] D. Gould and G. J. McLellan, editors. *BSAVA Manual of Canine and Feline Ophthalmology*. British Small Animal Veterinary Association, Quedgeley, Gloucester, 2014.

[6] D. S. Kermany, M. Goldbaum, W. Cai, C. C. S. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, J. Dong, M. K. Hewett, J. Dong, B. Ziyech, P. Shi, and K. Zhang. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5):1122–1131, 2018.

[7] J. Y. Kim, H. E. Lee, Y. H. Choi, S. J. Lee, and J. S. Jeon. CNN-based diagnosis models for canine ulcerative keratitis. *Scientific Reports*, 9(1):14209, 2019.

[8] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.

[9] M. Linek, editor. *Ophthalmologie auf den Punkt gebracht: Ein Leitfaden für die Kleintierpraxis*. VBS-Vet-Verl., Babenhausen, 1 edition, 2008.

[10] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017.

[11] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio. Transfusion: Understanding transfer learning for medical imaging. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

[12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen. MobileNetV2: Inverted residuals and linear bottlenecks. In *Conference on Computer Vision and Pattern Recognition*, pages 4510–4520, 2018.

[13] I. Walde, B. Nell, E. H. Schäffer, and R. Köstlin, editors. *Augenheilkunde: Lehrbuch und Atlas*. Schattauer GmbH, Stuttgart, 3 edition, 2008.