


**ORIGINAL ARTICLE**

# Artificial intelligence as a tool to aid in the differentiation of equine ophthalmic diseases with an emphasis on equine uveitis

Anna May<sup>1</sup>  | Stefan Gesell-May<sup>2</sup> | Tobias Müller<sup>3</sup> | Wolfgang Ertel<sup>3</sup>

<sup>1</sup>Equine Hospital, Ludwig Maximilians University Munich, Munich, Germany

<sup>2</sup>Centre for Equine Ophthalmology, Equine Hospital in Parsdorf, Vaterstetten, Germany

<sup>3</sup>Institute for Artificial Intelligence, Ravensburg-Weingarten University, Doggenriedstrasse, Germany

**Correspondence**

Anna May, Equine Hospital, Ludwig Maximilians University Munich, Veterinärstrasse 13, 80539 Munich, Germany.

Email: [Anna.May@pferd.vetmed.uni-muenchen.de](mailto:Anna.May@pferd.vetmed.uni-muenchen.de)

**Summary**

**Background:** Due to recent developments in artificial intelligence, deep learning, and smart-device-technology, diagnostic software may be developed which can be executed offline as an app on smartphones using their high-resolution cameras and increasing processing power to directly analyse photos taken on the device.

**Objectives:** A software tool was developed to aid in the diagnosis of equine ophthalmic diseases, especially uveitis.

**Study design:** Prospective comparison of software and clinical diagnoses.

**Methods:** A deep learning approach for image classification was used to train software by analysing photographs of equine eyes to make a statement on whether the horse was displaying signs of uveitis or other ophthalmic diseases. Four basis networks of different sizes (MobileNetV2, InceptionV3, VGG16, VGG19) with modified top-layers were evaluated. Convolutional Neural Networks (CNN) were trained on 2346 pictures of equine eyes, which were augmented to 9384 images. 261 separate unmodified images were used to evaluate the performance of the trained network.

**Results:** Cross validation showed accuracy of 99.82% on training data and 96.66% on validation data when distinguishing between three categories (uveitis, other ophthalmic diseases, healthy).

**Main limitations:** One source of selection bias for the artificial intelligence presumably was the increased pupil size, which was mainly present in horses with ophthalmic diseases due to the use of mydriatics, and was not homogeneously dispersed in all categories of the dataset.

**Conclusions:** Our system for detection of equine uveitis is unique and novel and can differentiate between uveitis and other equine ophthalmic diseases. Its development also serves as a proof-of-concept for image-based detection of ophthalmic diseases in general and as a basis for its further use and expansion.

**KEYWORDS**

artificial intelligence, blindness, equine uveitis, horse, machine/deep learning, ophthalmology

The abstract is available in German in the Supporting Information section of the online version of this article.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Equine Veterinary Journal* published by John Wiley & Sons Ltd on behalf of EVJ Ltd.

## 1 | INTRODUCTION

Equine ophthalmology is a highly specialised field in veterinary medicine, and detection as well as differentiation of equine ophthalmic diseases can be challenging for veterinarians.

Equine recurrent uveitis (ERU) can be a devastating ophthalmic disease leading to chronic or recurring inflammatory bouts and blindness, which has an enormous emotional and financial impact on the horse industry.<sup>1,2</sup> There are different types of uveitis. Anterior uveitis is often associated with more severe clinical signs and pain, whereas posterior uveitis may go unnoticed until the eye is severely damaged or destroyed.<sup>3</sup>

Because early detection and treatment of uveitis is critical to preserve the affected eyes, image analysis using deep learning (artificial intelligence) has potential as an additional tool for horse owners and veterinarians to help detect uveitic conditions in horses.

In human medicine, artificial intelligence and machine learning, especially deep learning algorithms have made vast advances in automatically diagnosing diseases.<sup>4-7</sup> Machine learning algorithms are trained to recognise patterns similar to the way veterinarians or human doctors look at them. The main difference is that algorithms need to see many concrete examples of the conditions to distinguish different categories and generalise the patterns on new examples which makes it especially useful in routine diagnostics with repetitive actions. As the information needs to be digitised for the computer to learn, machine learning can be extremely helpful in areas with sufficient amounts of digitised information available, such as detecting diseases based on CT scans or MRI images.

Processing and analysis of images via computer to understand their contents is called 'computer vision.' It comprises many different techniques such as image segmentation to divide a picture into related areas, the recognition and tracking of objects in pictures, and the classification of whole pictures. A further development of this technology is the recognition of anomalies with detection of patterns that do not fit into a given figure, eg detection of distinctive features in mammography.<sup>4</sup> Deep learning tools analyse radiographs and highlight potentially relevant regions so the radiologist can concentrate on pre-selected pictures.<sup>4</sup> Convolutional Neural Networks (CNN) represent a deep learning concept which is able to determine and recognise patterns by assigning importance (learning weights) to various aspects in images. Thereby it is able to differentiate them from one another.<sup>5</sup>

CNN are widely used in image classification due to their reduced complexity, less training parameters and adaptability compared to other neural networks.<sup>8</sup> In contrast to the other mentioned techniques, direct classification of pictures with CNNs has a low initial effort to obtain objective and provable data for this test (establishing a ground truth). Moreover, CNNs are able to distinguish between more than two categories, which makes this method the preferred one.

Other examples for deep learning in human medicine comprise similar programmes to the one described in this study. In human

ophthalmology some research groups concentrate on the examination of retinal pathologies, such as diabetic retinopathy,<sup>6-8</sup> age-related macular degeneration or glaucomatous optic neuropathy.<sup>9,10</sup> In classification of skin cancer, the most common human malignancy, which is primarily diagnosed visually, deep neural networks (neural networks with many layers) achieved performance comparable to board-certified dermatologists.<sup>11</sup> In veterinary medicine most artificial intelligence programmes are related to academic projects or are used in histopathology. One group is working on artificial intelligence recognising mitotic figures in mast cell tumours as an indication to evaluate tumour stages in dogs.<sup>12,13</sup>

For horses, there is an artificial intelligence system to predict the need for surgery in colic horses.<sup>14</sup> A research group from Sweden is developing a machine learning based pain expression recognition programme.<sup>15</sup>

In this study, different Convolutional Neural Network (CNN) models were evaluated in order to determine the CNN that detects uveitis in equine eye pictures with the highest possible accuracy.

The aim of this study was to develop an accessible smartphone tool for horse owners, which is able to distinguish between uveitis, other ophthalmic diseases and healthy eyes and achieves an accuracy of at least 95% on validation data. Our goal was to create a programme that can differentiate between normal and abnormal conditions and may help in decision-making of whether to treat the horse as an emergency.

## 2 | MATERIALS AND METHODS

### 2.1 | Technology

After assessing given computer vision tools in the context of equine ophthalmology, the following ideas were evaluated for this particular application:

#### 2.1.1 | Segmentation of the picture and classification through a separate neural network

Segmentation of each individual image is very time-consuming in the preparation step and as this method only considers geometric shape, other features of the disease (eg colour changes) would be missed. Therefore, this technology was dismissed.

#### 2.1.2 | Procedure to detect anomalies that differ from healthy eyes

The system learns how healthy eyes look and detects eyes that do not look normal. As this tool can only differentiate between healthy and not healthy and further classification is not possible, this technology was also dismissed.

### 2.1.3 | Direct classification of the overall picture with convolutional neural networks (CNN)

These tools have already shown their potential in similar studies as they are flexible enough to be used in various applications. For this approach, data has to be pre-classified in order to define a 'ground truth' for training and evaluation, but it requires significantly less preparation than the segmentation method. As CNN met the requirements for evaluating equine eye images, they were used in this study.

## 2.2 | Data

Photographs from various angles were taken when the horses arrived at the equine ophthalmology referral centre, Equine Hospital in Parsdorf to imitate the situation owners or general practitioners would find before examination. Some horses had been treated by referring veterinary surgeons prior to admission and this potentially included administration of mydriatics. After photography, horses were examined to verify findings in the pictures. We then dilated the pupil with mydriatics (Tropicamide) and horses were assessed via routine direct (WelchAllyn® direct ophthalmoscope), indirect ophthalmoscopy (HEINE Omega 500 LED indirect binocular ophthalmoscope and HEINE indirect ophthalmoscopy lens 20D), slit lamp biomicroscopy (Keeler PSL Classic LED) and tonometry (Icare® Tonovet). Sedation or regional anaesthesia were used when necessary. Ophthalmologic findings and diagnoses were obtained by a board-certified internal medicine specialist and a specialist in equine ophthalmology. Only horses for which the significant ophthalmologic findings were also visible in the photographs were included and photographs were classified based on the findings. This procedure of classification according to ophthalmologic findings was used to avoid differences in the categorisation of the horses.

Photographs of eyes classified as 'healthy' were taken from 86 horses. These eyes showed no pathological ophthalmic findings. Of these horses, 118 eyes were used and 668 pictures were taken into account for this study.

There were 221 horses in the 'uveitis' group with 244 eyes taken into account and 720 pictures of these eyes were used. Clinical signs on these pictures led to the suspected diagnosis of classic, insidious or posterior form of ERU. As the diagnosis of ERU is also based on epidemiological data, a clear diagnosis was not possible just by looking at two-dimensional images. Therefore, inclusion criteria

were typical findings of inner eye (anterior chamber, iris, lens, vitreous body, fundus) involvement such as fibrin or flare in the anterior chamber, miosis, inflammatory deposits on the anterior or posterior lens capsule or in the vitreous body, cataract, retinal ablation (seen as irregularities in the pupil), as well as vitreal cellular infiltrate.

The 'other diseases' group comprised different ophthalmic conditions such as glaucoma, hyphaema (not in terms of serohaemorrhagic effusion, but massive bleeding resulting from putative trauma), different types of keratitis, ulcers, endothelitis, iris coloboma, and neoplasia. This group consisted of 346 horses of which 350 eyes and 1219 pictures were taken into account.

During the development of the neural network and its implementation in a smartphone app the dataset was expanded continuously (Table 1). In total, 2346 training images (90% of the dataset) were used. The data were expanded to 9384 images using augmentation, which is the process of increasing the amount of data by adding slightly modified copies of already existing data. Augmentation was only applied to the training set and did not affect the validation set.

At first, differentiation between uveitis and healthy eyes was prioritised. With the collection of more images, a third category named 'other diseases' was taken into account.

### 2.3 | Training of the artificial intelligence tool using machine learning

In this study, different Convolutional Neural Network (CNN) models were tested and compared to each other to find the CNN that detects uveitis with the highest possible accuracy. The choice of a suitable CNN architecture was based on two criteria:

- the accuracy on ImageNet data. ImageNet is a large visual database designed for use in visual object recognition software research. It is often used as a common benchmark for comparison.<sup>16</sup>
- the size of the network with regard to mobile use.

To find the most suitable CNN for this purpose four base neural networks of different sizes (listed from small to large) with modified top-layers were evaluated<sup>17</sup>:

- MobileNetV2
- InceptionV3
- VGG16
- VGG19

TABLE 1 Number of pictures used in the different categories

|                       | Date            | ERU/uveitis | Healthy | Other diseases | Total |
|-----------------------|-----------------|-------------|---------|----------------|-------|
| Before start of study | 5 February 2020 | 115         | 18      | 262            | 395   |
| First training        | 3 March 2020    | 172         | 214     | 272            | 658   |
| Second training       | 29 April 2020   | 720         | 668     | 1219           | 2607  |

Abbreviation: ERU, equine recurrent uveitis.

To analyse the sections of the image the artificial intelligence programme puts emphasis on, the open-source library tf-explain (<https://github.com/sicara/tf-explain> with the implementation of Grad CAM: <https://arxiv.org/abs/1610.02391>) was used, which visualises the neural activations within the network.

## 2.4 | Cross validation to evaluate machine learning models

Cross validation, in the context of neural network training, is a set of methods to validate the predictive performance of a model by splitting the dataset into training data (90% of the data) which the model has seen, and unknown validation data (10% of the data) which the model first sees during prediction in the validation stage. In the used k-fold variant, this is repeated multiple (k) times, where the validation subset differs on each repetition and the model is trained on the remaining subset.<sup>8</sup> The goal is to get a reliable measure of the predictive performance on many unseen samples without losing too much data for training, which is especially useful on smaller datasets.<sup>18</sup>

In this study, the trained model was also evaluated with a limited data sample. Of the total 2607 images, 261 random images were excluded for validation. Each neural network validated using this method was trained 10 times with the random subset for validation being different on each run, which consisted of 100 epochs. An epoch refers to one cycle through the whole training dataset.

For the first training with 658 total images, the dataset was reduced to two classes (uveitis and healthy). The second training with the final dataset (2607 images) was optimised and all three classes (uveitis, other diseases and healthy) were included.

## 2.5 | Loss and accuracy

A loss function describes the classification error of the model on the training and validation samples, respectively. In this study, categorical crossentropy was used to assess the models classification performance. Categorical crossentropy is a loss function that is used in multi-class classification tasks. These are tasks where an example can only belong to one of the possible categories, and the model must decide which one. The loss value was minimised during training, hence the lower this value got, the better the prediction of the model fitted the ground truth, which was the expected result based on the diagnosis of the ophthalmologists.

Accuracy is a metric that was measured after a training pass to evaluate the model on a coarser level. This measurement yielded how many samples were correctly classified and not how much the erroneously classified samples differed from the ground truth (Table 2).

## 2.6 | Data acquisition for the smartphone app

The horses were taken inside for ophthalmic examination and acquisition of photographs in order not to have any sun reflection on the cornea. Images were taken in different environments and with various light sources, as well as in complete darkness. Pictures in this study were meant to resemble normal imperfect photographs taken by owners or veterinarians in suboptimal settings in order to test the apps capability to detect ophthalmic findings under real life conditions.

Horses were restrained with a halter and the camera positioned so the eye was on full screen. Most horses tolerated the procedure very well. Pictures were taken with the app Camera+2 (©LateNightSoft 2018-2020) so the LED light of the device was permanently activated in order not to scare the horses with a sudden flash of light. In horses that did not tolerate the camera next to their eye very well and moved a lot, freeze frame pictures were obtained

| Subset for validation     | Training |          | Validation |          |
|---------------------------|----------|----------|------------|----------|
|                           | Loss     | Accuracy | Loss       | Accuracy |
| 0                         | 0.002766 | 99.88%   | 0.004837   | 96.17%   |
| 1                         | 0.000144 | 99.90%   | 0.000253   | 96.17%   |
| 2                         | 0.000121 | 99.91%   | 0.067921   | 96.17%   |
| 3                         | 0.019623 | 99.90%   | 0.000000   | 97.70%   |
| 4                         | 0.002281 | 99.88%   | 0.000002   | 97.70%   |
| 5                         | 0.025071 | 99.49%   | 0.000033   | 95.02%   |
| 6                         | 0.002434 | 99.91%   | 0.002526   | 98.46%   |
| 7                         | 0.004406 | 99.72%   | 0.000050   | 96.53%   |
| 8                         | 0.000007 | 99.93%   | 0.000000   | 96.55%   |
| 9                         | 0.058943 | 99.65%   | 0.000071   | 96.17%   |
| <b>Average</b>            | 0.011580 | 99.82%   | 0.007569   | 96.66%   |
| <b>Standard deviation</b> | 0.017847 | 0.14     | 0.020175   | 0.95     |

**TABLE 2** Cross validation results: 2nd training, 3 labels, 'VGG19Large', after Convolutional Neural Network (CNN) fine tuning; ERU, Equine Recurrent Uveitis

from videos taken with an iPhone 7 plus (Apple). With a constant light source, details were more visible and the horse accepted the owner taking the picture much better because there was no flash.

### 3 | RESULTS

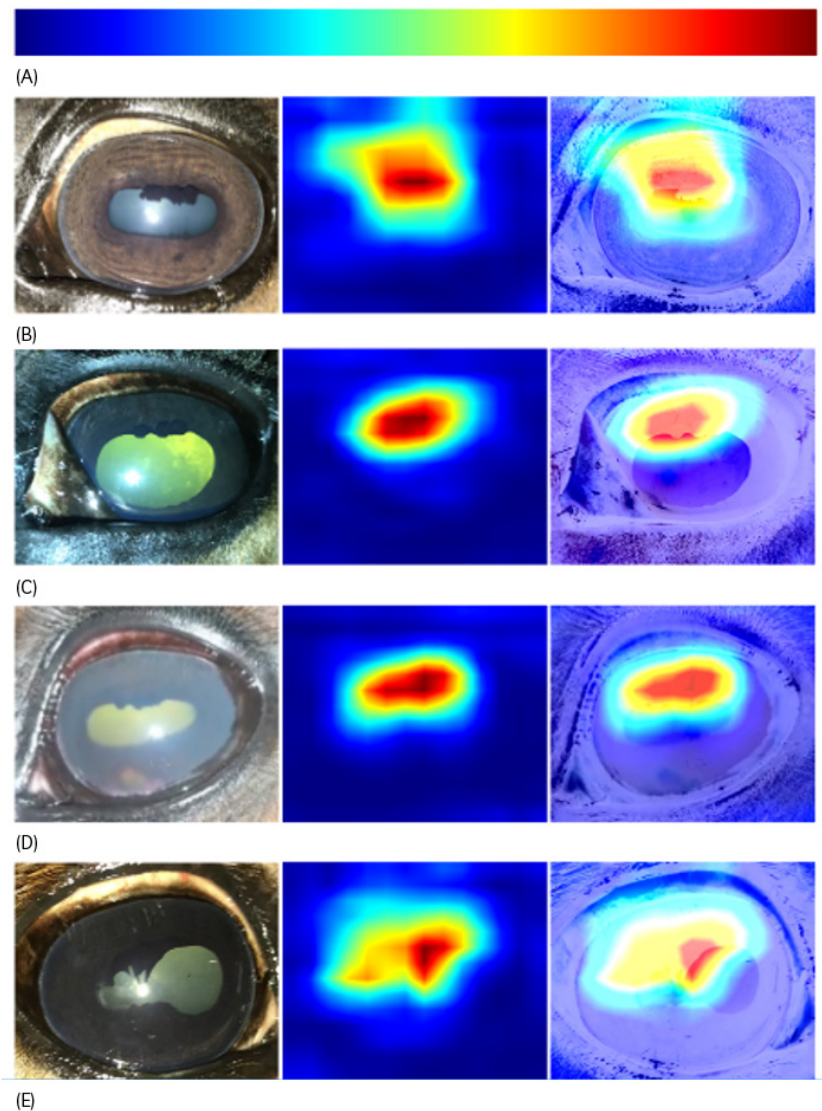
#### 3.1 | Training of the artificial intelligence tool using machine learning

The smallest possible network MobileNetV2 proved to be insufficient for this purpose as validation accuracy only reached 39%. Inception V3 was a larger network that experienced overfitting, a phenomenon where a model learns the given training dataset almost perfectly but misclassifies unseen or new pictures and is therefore less able to generalise to new data and has a poor validation score. As a result, two much larger networks with a different base architecture, VGG16 and VGG19,<sup>16</sup> were evaluated. These networks proved to be more suitable to analyse the data, with a varying accuracy between 93% and 96.8%.

With better generalisation as a result of augmenting the training data, the validation accuracy improved, which was the desired effect. Analysis of the important sections for the CNN revealed that the artificial intelligence used different parameters than the human eye. The section of activation was shown as a heat map (Figure 1A-E) and the colour pattern showed the activation (low = dark blue to high = dark red). The visualisation displayed that most activation took place in the dorsal aspect of the eye, as particularly the yellow and red colours were displayed there. The analysed section covered the upper part of the iris and its margins (Figure 1E). Therefore, the CNN seemed to be primarily analysing the transition zones between pupil and iris and dorsal part of the cornea and sclera, rather than the details in the inner eye.

#### 3.2 | Cross validation to evaluate machine learning models

The algorithm was validated using the validation data, meaning a subset of pictures of the whole dataset. For the final evaluation



**FIGURE 1** Examples of original images, 'heat map' and superimposed view: A, Used colour scale 'jet'—left: low activation to right: high activation. B, Healthy eye. C, Eye with uveitis: mild corneal opacity, greenish appearance of the inner eye. D, Eye with uveitis: moderate to severe corneal opacity, miosis. E, Eye with uveitis: posterior synechia



custom variants of the VGG19 base architecture were chosen.<sup>17</sup> In the first run with two classes (uveitis, healthy) on the 'VGG19Small'-network accuracy was 99.46% ( $\pm$ SD 0.17) for the training (347 images) and 97.15% ( $\pm$ SD 1.79) for the validation data (39 images).

In the second training with three classes (uveitis, other diseases and healthy) the network we called 'VGG19Large' showed an accuracy of 97.86% ( $\pm$ SD 0.25) on the training (2346 images) and 92.29% ( $\pm$ SD 1.32) on the validation data (261 images) after 100 epochs. After fine tuning this network for another 50 epochs the accuracy surpassed the desired 95% (99.82% ( $\pm$ SD 0.14) for training data and 96.66% ( $\pm$ SD 0.95) for validation data, respectively; Table 2).

The difference between 'VGG19Small' and 'VGG19Large' is construction of the top layer which was modified and enlarged to improve learning capacity.

### 3.3 | The smartphone app

To demonstrate proof-of-concept of this artificial intelligence, a web application (app) was developed which offers a quick analysis of equine eyes on different mobile devices, such as Apple iPhone and Samsung mobile phones. The 'Equine A-Eye' comprises different functions, which can be used by the horse owner or veterinarian. At the time of completion of the study the web app 'Equine A-Eye' can be accessed at the URL <http://equine-a-eye.anirec.de/>. The main function of the web application is to take a photograph with the

smartphone camera or choose a photo from the storage of the device, to have the image analysed by the neural network and instantly get the result of this analysis on the device. Figure 2 shows the recording/selection of the image. The camera preview is shown in the centre of the display and the record button is situated underneath (Figure 2A). After a successful picture has been taken, it can be selected (Figure 2B). The image is then classified and the probability for each label can be displayed (Figure 3). Additional information can also be presented (heat map, superimposed view) to make the decision of the CNN more transparent for the user and to give the probability for the given suspected diagnosis (Figure 3B).

In this study the programme was able to differentiate between 'uveitis', 'healthy' and 'other diseases' based on the training with pictures of equine ophthalmic diseases. The quality of a photograph taken by a smartphone was sufficient for the programme to distinguish between these three conditions. As the artificial intelligence programme has been trained with specific image sections (close-up view), it is important to develop a guide for the owner on how to obtain adequate images in order to get as many details as possible in the picture in full resolution.

## 4 | DISCUSSION

The developed deep learning software is a simple tool, which is able to detect changes due to uveitis and other ophthalmic diseases in the equine eye. It can achieve high classification accuracy due to

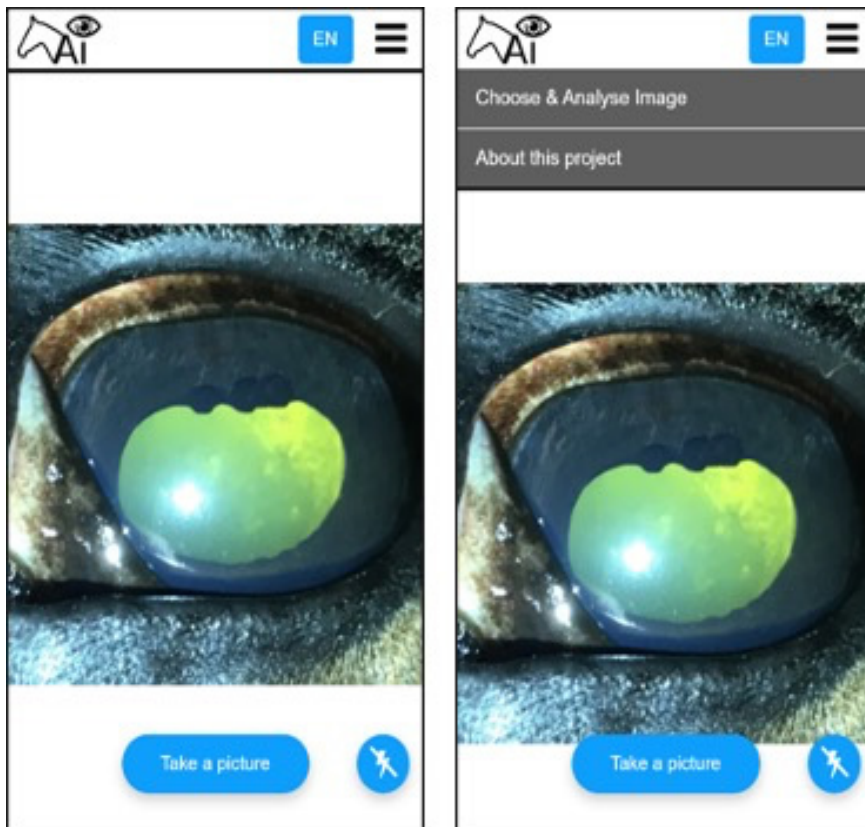
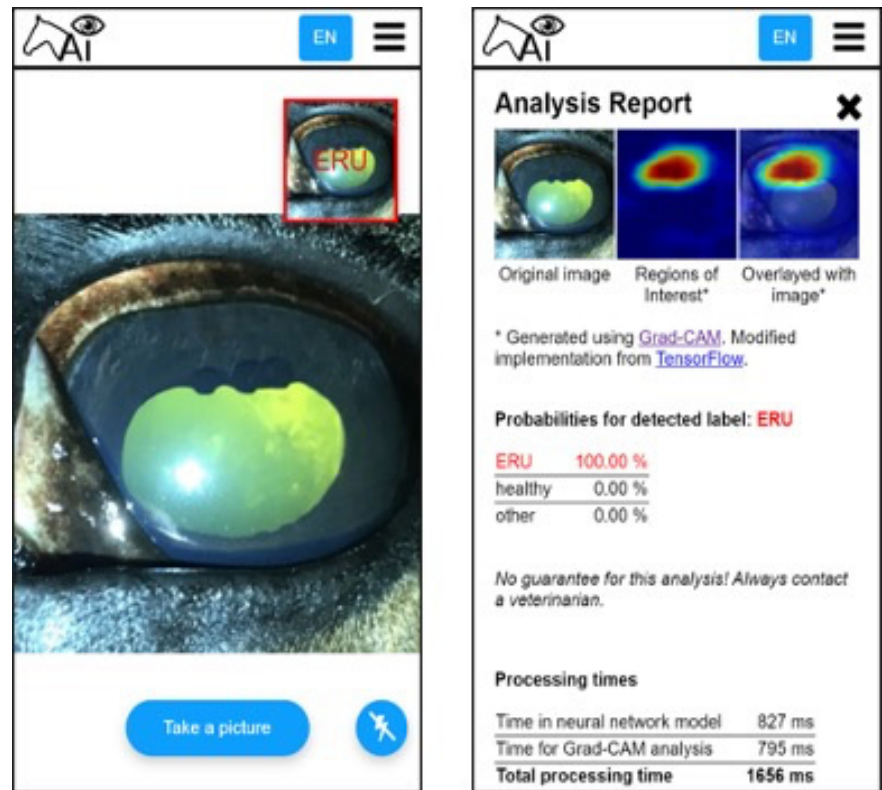


FIGURE 2 Smartphone app: recording of the image and selection in the menu

**FIGURE 3** Smartphone app: display of the classification results and detailed report (right box) showing original picture on the left as well as the relevant areas in the middle and these areas superimposed with the original picture on the right. The analysis report also gives probabilities for the given diagnosis ('uveitis' = Equine Recurrent Uveitis in the example)



many of the pathological changes being visible, even in photographs. The high accuracy in the cross validation process is surprising given the fact that the programme concentrates on areas of the equine eye that are less important to the human examiner.

To distinguish between healthy and uveitic eyes, veterinarians will consider abnormalities within the inner eye including effusions in the anterior chamber and in the vitreous body, irregularities in the pupil and a turbid greenish appearance of the inner eye. While in eyes categorised as 'uveitis' or 'other diseases' the artificial intelligence tool focuses on the outer structures of the eye, there is significant activation in the inner eye in 'healthy' eyes. This suggests that the clear homogenic appearance of the inner eye is a major feature for the artificial intelligence tool to categorise an image as 'healthy'. In most of the used images there is significant variation in the dorsal part of the picture (corneal opacities of varying degrees) so the programmes focus on that section seems plausible. Another explanation for the focus on the dorsal aspect of the eye would be that the artificial intelligence focuses on the transition from iris to inner eye and the difference in colour on that level. The important role of the iris margins became especially clear in the last picture as the programme focused on its irregular shape (Figure 1B-E).

One main limitation is the fact that the programme is not able to look at the posterior segment of the eye. Only if irregularities in the pupil are visible in the picture, can it be concluded that there is posterior ophthalmic pathology. In cases of emergency (eg corneal ulcer, acute uveitis) the tool is still useful, as it is capable to detect most ophthalmic features associated with these diseases. Emergencies of the posterior segment (eg retinal detachment, inflammation of the optic nerve) will be missed by the tool in most cases.

If owners use the application in the field there might be a possible responsibility concern if eye conditions are misdiagnosed as the app misinterprets findings resulting in treatment being initiated too late. The tool can therefore only be used as an addition to the examination of a veterinarian. Owners have to be made aware that in case of doubt a veterinarian should be called to thoroughly assess the eye. Possible scope of usage would be the evaluation of emergency situations in regions with no direct access to veterinary care and in after hour situations where veterinarians are not readily available. Inexperienced veterinarians may use the app as a further diagnostic tool to reach a tentative diagnosis and as an aid in differentiating between various ophthalmic conditions.

To evaluate the performance of the tool with pictures that were not taken under perfect conditions, different light settings were taken into account. The horses were taken inside a building but there was light exposure of different kinds. The main limitation for the quality of the pictures was direct sunlight when structures of the eye were obscured and there was a light reflection on the cornea.

As the LED light reflection is visible on all images to a certain degree, it probably does not cause any selection bias and does not cause the programme to focus on it. Eye lashes and other foreign particles in the image may be reasons for wrong interpretation of individual pictures, but do not cause any bias due to their rareness in the dataset. Iris colour, which is recognisable in the highly activated part, does also not produce bias as there is variation within all categories ranging from dark to light brown, and a large colour spectrum is covered in all classes.

The size of the pupil is one aspect that may serve as a selection bias as this feature is not homogeneously dispersed in all data sets

of the different categories. Dilation of the pupil using mydriatics is often used in equine ophthalmologic examination. Therefore categories 'uveitis' and 'other diseases' include more pictures with dilated pupils, whereas the 'healthy' category usually has normal pupils because they did not receive medication. Thus a wide pupil may seem to signal an indication for a diseased eye, while it is only a side effect of the ophthalmologic examination. In this study, we took pictures of eyes with pathological findings before using mydriatics ourselves and included eyes that were medicated prior to arrival at the hospital. This fact is especially important as the tool aims to analyse pictures taken by owners and veterinarians in the field possibly before use of mydriatics. One potential interpretation of the activations in [Figure 1](#) could be the CNN focusing on the distance between eyelid and margin of the iris to distinguish between a normal or widened pupil, which would support the theory of the bias.

In this case, however, the CNN would most likely interpret the shape of the pupil rather than the margin of the iris. Furthermore, the activations in [Figure 1](#) are in the same area regardless of the pupil size and shape and the images were classified correctly. The image in [Figure 1D](#) would have been categorised as 'healthy' if pupil size was the determining factor which also applies to many images in the 'uveitis' group, where pupils are not dilated or the horse is displaying miosis. Another advantage of our study design is the inclusion of the 'other diseases' category which also contains predominantly eyes with dilated pupils so that the programme does not tend to consider dilated pupils unique to 'uveitis'.

At present, artificial intelligence has the ability to change and improve diagnostics in almost all areas of human and veterinary medicine.<sup>5,7,12-14</sup> With the help of artificial intelligence programmes, medicine can be made more accurate, faster and can therefore improve outcomes for human and veterinary patients. The more medical data becomes digitised and unified, the better the artificial intelligence systems can be trained to find data patterns that can be used to help analyse complex diagnostic problems.

The tool described in the current study is a knowledge- and data-intensive computer-based solution for different eye conditions of the equine patient. It has the potential to help categorise suspected diagnoses to support veterinarians who are not specialised in equine ophthalmology and may guide horse owners on when to call a veterinarian if they are confronted with equine ophthalmic problems. Further studies on the benefits and effectiveness of this tool in clinical practice and when used by owners are required.

## INFORMED CONSENT

Explicit owner informed consent for inclusion of animals in this study was not stated.

## CONFLICT OF INTERESTS

No competing interests have been declared.

## AUTHOR CONTRIBUTIONS

A. May contributed to study design, data analysis and interpretation, and preparation of the manuscript. S. Gesell-May, T. Mueller and W.

Ertel contributed to study design, study execution, and data analysis and interpretation. All authors approved the final version of the manuscript and had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

## ETHICAL ANIMAL RESEARCH

Research ethics committee oversight not currently required by this journal: procedures were non-invasive.

## PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/evj.13528>.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed in this study.

## ORCID

Anna May  <https://orcid.org/0000-0003-3038-5952>

## REFERENCES

- Gerding JC, Gilger BC. Prognosis and impact of equine recurrent uveitis. *Equine Vet J*. 2016;48(3):290–8.
- Gilger BC. Equine recurrent uveitis: the viewpoint from the USA. *Equine Vet J*. 2010;42(37):57–61.
- Gilger BC, Hollingsworth SR. Diseases of the uvea, uveitis and recurrent uveitis. In: Gilger BC, editor *Equine ophthalmology*, 3rd edition. Wiley Blackwell, 2017; p. 369–415.
- Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology*. 2019;292:60–6.
- Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *J Am Med Assoc*. 2016;316(22):2402–10.
- Quelleg G, Charriere K, Boudi Y, Cochener B, Lamard M. Deep image mining for diabetic retinopathy screening. *Med Image Anal*. 2017;39:178–93.
- Alyoubi WL, Shalash WM, Abulkhair MF. Diabetic retinopathy detection through deep learning techniques: a review. *Inform Med Unlocked*. 2020;20.
- Goodfellow I, Bengio Y, Courville A. *Deep learning*. MIT Press. 2016;329ff.
- Lee CS, Baughman DM, Lee AY. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images. *Ophthalmol Retina*. 2017;1(4):322–7.
- Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmol*. 2018;125(8):1199–206.
- Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115–8.
- Bertram CA, Aubreville M, Marzahl C, Maier A, Klopffleisch R. A large-scale dataset for mitotic figure assessment on whole slide images of canine cutaneous mast cell tumor. *Nature Sci Data*. 2019;6.
- Aubreville M, Bertram CA, Marzahl C, Gurtner C, Dettwiler M, Schmidt A, et al. Deep learning algorithms out-perform veterinary pathologists in detecting the mitotically most active tumor region. *Nature Sci Rep*. 2020;10.



14. Fraiwan MA, Abutarbush SM. Using artificial intelligence to predict survivability likelihood and need for surgery in horses presented with acute abdomen (colic). *J Equine Vet Sci.* 2020;90.
15. Andersen PH, Glerup KB, Wathan J, Coles B, Kjellström H, Broome S, et al. Can a machine learn to see horse pain? An interdisciplinary approach towards automated decoding of facial expressions of pain in the horse. In: *International Conference on Methods and Techniques in Behavioural Research, Survey and position paper 2018.*
16. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet large scale visual recognition challenge. *Int J Comput vis.* 2015;115:211–52.
17. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *ArXiv:1409.1556* 2015.
18. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the Fourteenth*

International Joint Conference on Artificial Intelligence 1995. San Mateo, CA: Morgan Kaufmann. 2(12); p.1137–43.

#### SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** May A, Gesell-May S, Müller T, Ertel W. Artificial intelligence as a tool to aid in the differentiation of equine ophthalmic diseases with an emphasis on equine uveitis. *Equine Vet J.* 2022;54:847–855. doi:[10.1111/evj.13528](https://doi.org/10.1111/evj.13528)



**imv**  
imaging

## Take back your time with Direct Radiography Perfectly designed for equine veterinary practice

- ✓ Obtain images in seconds
- ✓ Excellent image quality
- ✓ Minimal maintenance required

**ClaroX  
DR X-ray  
system**

**Cuatro  
Slate Hub  
veterinary  
DR X-ray  
system**