Deep learning-based system for automatic identification of benign and malignant eyelid tumours

Fan, W.L.; Jager, M.J.; Dai, W.W.; Heindl, L.M.

Clinical science

Deep learning-based system for automatic identification of benign and malignant eyelid tumours

Wanlin Fan ,[1] Martine Johanna Jager ,[2] Weiwei Dai,[3] Ludwig M Heindl [1,4]

[1]Department of Ophthalmology, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany
[2]Department of Ophthalmology, Leiden University Medical Center, Leiden, The Netherlands
[3]Changsha Aier Eye Hospital, Hunan, China
[4]Center for Integrated Oncology (CIO), Aachen-Bonn-Cologne-Duesseldorf, Cologne, Germany

Correspondence to
Prof.Dr.Dr. Ludwig M Heindl; ludwig.heindl@uk-koeln.de and Dr. Weiwei Dai; daiweiwei@aierchina.com

WD and LMH contributed equally.

WD and LMH are joint senior authors.

Received 6 January 2025
Accepted 10 April 2025
Published Online First 10 May 2025

ABSTRACT

Aims Our aim is to develop a deep learning-based system for automatically identifying and classifying benign and malignant tumours of the eyelid to improve diagnostic accuracy and efficiency.

Methods The dataset includes photographs of normal eyelids, benign and malignant eyelid tumours and was randomly divided into a training and validation dataset in a ratio of 8:2. We used the training dataset to train eight convolutional neural network models to classify normal eyelids, benign and malignant eyelid tumours. These models included VGG16, ResNet50, Inception-v4, EfficientNet-V2-M and their variants. The validation dataset was used to evaluate and compare the performance of the different deep learning models.

Results All eight models achieved an average accuracy greater than 0.746 for identifying normal eyelids, benign and malignant eyelid tumours, with an average sensitivity and specificity exceeding 0.790 and 0.866, respectively. The mean area under the receiver operating characteristic curve (AUC) for the eight models was more than 0.904 in correctly identifying normal eyelids, benign and malignant eyelid tumours. The dual-path Inception-v4 network demonstrated the highest performance, with an AUC of 0.930 (95% CI 0.900 to 0.954) and an F1-score of 0.838 (95% CI 0.787 to 0.882).

Conclusion The deep learning-based system shows significant potential in improving the diagnosis of eyelid tumours, providing a reliable and efficient tool for clinical practice. Future work will validate the model with more extensive and diverse datasets and integrate it into clinical workflows for real-time diagnostic support.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Eyelid tumours, which are common in ocular oncology, can be benign or malignant. Malignant types, such as basal cell carcinoma, require early detection for a better prognosis. Deep learning has shown promise in medical imaging, but previous studies of eyelid tumours have mainly focused on Asian populations with limited sample diversity.

WHAT THIS STUDY ADDS

⇒ This study introduces a novel deep-learning system tailored to Caucasian populations for distinguishing normal, benign and malignant eyelid tumours from clinical images. The dual-path Inception-v4 model demonstrated high diagnostic accuracy, achieving an AUC of 0.930 and F1-score of 0.838, significantly enhancing the performance of traditional models.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study develops a valuable diagnostic tool, especially in regions with limited access to experienced ophthalmologists. Its application in smartphones could revolutionise early screening and detection practices, ultimately leading to early intervention and improved patient outcomes.

INTRODUCTION

Eyelid tumours are the most common ocular tumours,[1] and account for 5%–10% of skin tumours.[1 2] Eyelid tumours can be classified as benign and malignant, of which 80%–90% are benign.[3 4] Malignant tumours include basal cell carcinoma (BCC), squamous cell carcinoma (SCC), sebaceous cell carcinoma and melanoma. Of these, BCC is the most common eyelid malignancy, accounting for 80%–90%. SCC accounts for approximately 5%–10% of periocular malignancies.[5] Sebaceous gland carcinomas (SGC) account for 1.0%–5.5% of all eyelid malignancies and were previously thought to be more common in Asians, but recent studies have shown a more equal incidence in all races.[6] Melanoma accounts for less than 1% of all eyelid malignancies.[7] Cook *et al* reported that if malignant eyelid tumours are detected at their earliest stages (with a depth of skin invasion ≤0.76 mm), the estimated 5-year survival rate can exceed 99%.[2] Therefore, early recognition and treatment are crucial for malignant eyelid tumours.

Since benign eyelid tumours and malignant eyelid tumours that appear on the eyelid are sometimes similar, it is often difficult for the examining physician to distinguish between them.[8] Unlike many skin lesions in other parts of the body, the anatomy of the eyelid is complex; therefore, diagnosing eyelid tumours requires an experienced ophthalmologist. Although over 200 000 ophthalmologists are currently employed globally, there is a shortage of ophthalmologists in both developing and developed countries, especially experienced ophthalmologists, in particular, may hamper the early detection of eyelid malignancies.[9] Therefore, there is an urgent need for a simple diagnostic tool for early screening of eyelid tumours.

Deep learning (DL) has emerged as a powerful tool in medical image analysis, offering the potential

To cite: Fan W, Jager MJ, Dai W, *et al*. *Br J Ophthalmol* 2025;**109**:1050–1055.

for high accuracy in diagnosing various conditions. Recently, it has been reported that artificial intelligence has achieved a high level of accuracy in the automated detection of a wide range of diseases from clinical images.[10–13] In the field of ophthalmology, many studies have developed DL-based systems that can accurately detect ophthalmic diseases such as cataracts,[14] diabetic retinopathy,[15] age-related macular degeneration[16 17] and glaucoma.[17–19] A few studies have already been reported on DL-based systems that detect eyelid tumours accurately.[20 21] However, these studies have mainly focused on Asians. No similar studies have been found for Caucasians, the group of people with the highest incidence of eyelid tumours.

In this study, we propose a DL-based system explicitly designed to identify eyelid tumours based on ocular image data of Caucasian. This system aims to support ophthalmologists by providing a second opinion and reducing the subjective variability in tumour classification.

METHODS AND MATERIALS

Image dataset

This study was a single-centre retrospective and prospective study for developing diagnostic tools. The study retrospectively collected patients who underwent an ophthalmologic examination from 1 January 2017 to 30 December 2023, at the Eye Center of the University Hospital Cologne (Cologne, Germany). We collected 1251 unilateral ocular photographic images (1122 participants), including 846 photos from 831 patients with eyelid tumours and 405 photos from 291 healthy individuals. Images were taken at different locations, such as outpatient clinics, inpatient wards and operating rooms, so the lighting and background of the images are inconsistent, reflecting our dataset's richness and diversity. We included only cases with single lesions of eyelid tumours visible on eyelid skin. Tumours primarily involving the conjunctiva or originating conjunctival with secondary eyelid skin involvement were excluded, as were cases with multiple eyelid lesions. Two experienced ophthalmologists independently reviewed all cases according to defined inclusion and exclusion criteria to ensure accuracy and minimise selection bias. The dataset included photographs of normal eyelids, benign and malignant eyelid tumours and was randomly divided into a training and validation dataset in a ratio of 8:2. No overlap was allowed between the training and validation datasets, and the three types of photos, normal eyelids and benign and malignant eyelid tumours were included in both datasets at the same ratio.

Development of an algorithm for eyelid tumour detection

This study investigates the application of multiple convolutional neural networks (CNNs) for automated detection of benign and malignant eyelid tumours. Histopathological diagnosis serves as the ground truth. Our findings indicate that anatomical structures and colour intensity are crucial for distinguishing tumours, and neural networks should extract and learn these features to perform better. To effectively guide the network in learning, we preprocess the ocular surface images to highlight anatomical structures and design a dual-path network to extract structural and intensity features separately.

Feature enhancement

Many preprocessing methods for retinal images are adapted to enhance the structural features of eyelid tumours.[22] Initially, a Gaussian filter estimates the background intensity of the colour image. The original colour image is subtracted from the estimated background and undergoes brightness normalisation and contrast enhancement. The factors are scaled and shifted to limit the total intensity. The whole process is as follows :

$$x_n^{eh} = \alpha\left(x_n - G_\sigma * x_n\right) + \gamma$$

The original image x_n is processed using a Gaussian kernel, G_σ , with a variance of σ, through a convolution operation $(*)$. This step smoothens the image, reducing noise and highlighting important features. Through empirical analysis, optimal values for these factors are determined: α=4, γ=0.5. In previous retinal image studies[2], it is expected to set σ=r/30, where r represents the radius of the Field of View of the retinal image. However, through experimental analysis in this study, we have determined that setting σ=512/30 results in a more pronounced enhancement of structural features.

Neural network design

The study compares four state-of-the-art CNN architectures: VGG16, ResNet50, Inception-v4 and EfficientNet-V2-M. These neural networks are tested on the original ocular surface. A dual-path network with a CNN architecture is designed to extract different features better. Two identical CNN networks are built as feature extractors, sharing parameters. The original red-green-blue (RGB) image is input into the RGB path and the preprocessed image into the enhanced image path. The embeddings of the two paths are fused and passed through two fully connected layers to obtain a three-dimensional vector representing the tumour discrimination score.

Model training

Fourfold cross-validation is used to develop the model and identify the optimal model. All networks are initialised with pretrained parameters from ImageNet, inspired by transfer learning. All images are scaled to 512×512 dimensions for input into the network. Data augmentation methods such as horizontal flipping and random rotation (−10° to 10°) are applied during training to improve the model's robustness and mitigate overfitting risks. The models are developed on a server with 8 NVIDIA A800 GPUs, with a mini-batch size of 32 per GPU. The ADAM optimizer, an initial learning rate of 0.0016, β1 of 0.9, β2 of 0.999 and a weight decay of 0.05 are used. Each network is trained for 50 epochs, with the F1 value of the validation set serving as the reference for model saving.

Heatmap generation

This study employed Gradient Weighted Class Activation Maps (Grad-CAM) to visualise the CNN model's decision-making process.[23] Grad-CAM generates heat maps highlighting the image regions with the most significant impact on classification, which are associated with distinguishing benign and malignant eyelid tumours.

Statistical analysis

Python V.3.7.3 (Wilmington, Delaware) and MATLAB R2016a (https://www.mathworks.com/) were used for all statistical analyses. The performance of the DL model was evaluated using accuracy, sensitivity, specificity and receiver operating characteristic (ROC) curves. We calculated the area under the curve (AUC) and 95% CIs. The distributions of benign and malignant eyelid tumours were calculated, and the pie plot was drawn using ChiPlot (https://www.chiplot.online/).

RESULTS

Demographics of participants

This study used 449 photographs of 434 benign cases, 397 photographs of 397 malignant lesions and 405 photographs

Table 1 Characteristics of the datasets

	Case set		Control set
	Malignant tumour	Benign tumour	Tumour-free
Photographs	397	449	405
Participants	397	434	291
Age (SD)	71.4 (13.4)	55.1 (20.1)	50.6 (21.1)
Gender			
Female (%)	220 (55)	236 (54)	128 (44)
Male (%)	177 (45)	198 (46)	163 (56)
Side			
Right eye	191	200	94
Left eye	206	219	83
Both	0	15	114

of 291 healthy cases. All 1251 photographs were used to train and validate the DL system. The mean age (SD) of the patients with malignant eyelid tumours was 71.4±13.4 years, including 177 (45%) males and 220 (55%) females. The mean age (SD) of patients with benign eyelid tumours was 55.1±20.1 years, including 198 (46%) males and 236 (54%) females. In contrast, the mean age of healthy individuals was (SD) 50.6±21.1 years, with 163 (56%) males and 128 (44%) females (table 1). Benign tumours were diagnosed histopathologically in 434 subjects and malignant tumours in 397 patients. The most common benign eyelid tumour was a nevus (95/449, 21%), followed by chalazion (89/449, 20%), cyst (70/449, 16%) and papilloma (52/449, 12%) (figure 1A). The most common malignant tumour of the eyelid in the dataset was BCC (319/397, 80%), followed by SCC (62/397, 16%), melanoma of the Eyelid (8/397, 2.0%) and Merkel cell carcinoma (3/397, 0.8%) (figure 1B). Two hundred benign tumours were located in the right eyelid, 219 in the left eyelid and 15 were located bilaterally. For malignant tumours, 191 cases were located in the right eyelid and 206 in the left eyelid.

Performance of multiple CNNs

In terms of performance in identifying a normal eyelid, benign and malignant eyelid tumours, the eight models (four deep neural networks —VGG16, ResNet50, Inception-v4 and EfficientNet-V2-M—along with their corresponding dual-path networks) achieved an average accuracy greater than 0.746, with average sensitivity and specificity exceeding 0.790 and 0.866, respectively. The average AUC for the eight models was greater than 0.904 in correctly identifying normal eyelids, benign and malignant eyelid tumours. The dual-path Inception-v4 network demonstrated the highest performance, with an AUC of 0.930 (95% CI 0.900 to 0.954) and an F1-score of 0.838 (95% CI 0.787 to 0.882). These results highlight that the enhancement of structural features in the photographs through advanced data processing methods played a critical role in improving the network's performance. This enhancement, when combined with the dual-path network, significantly boosted the benchmark network's diagnostic accuracy, particularly for the VGG16 network, where the accuracy improved notably from 0.746 (95% CI 0.680 to 0.789) to 0.790 (95% CI 0.737 to 0.837) (table 2). Figure 2 illustrates the ROC and precision-recall curves for the eight models across three categories: normal, benign and malignant. In identifying malignant eyelid tumours, the optimised dual-path Inception-v4 network achieved an AUC of 0.930 (95% CI 0.865 to 0.948), a sensitivity of 0.923 (95% CI 0.859 to 0.976) and a specificity of 0.963 (95% CI 0.916 to 1.000).

Interpreting CNN models

We generated heatmaps to show how CNN models make classification decisions. The heatmaps showed areas important for classifying tumours and matched the actual tumour locations closely, regardless of tumour size, location or shape. Because the highlighted regions matched the real tumour sites closely, CNN models appear capable of accurately identifying areas important for tumour diagnosis (figure 3).

DISCUSSION

In this study, we evaluated the performance of a DL system in identifying tumour-free eyelids, benign and malignant eyelid tumours from clinical images captured by digital cameras. All algorithms showed an average accuracy above 0.798, an average sensitivity and specificity above 0.898 and 0.942, respectively, and an average AUC greater than 0.921. The best two-path Inception-v4 network achieved an AUC of 0.908 (95% CI 0.865 to 0.948), a sensitivity of 0.923 (95% CI 0.859 to 0.976) and a specificity of 0.963 (95% CI 0.916 to 1.000) for identifying eyelid malignant tumours. The system's ability to differentiate between normal eyelids, benign and malignant eyelid tumours

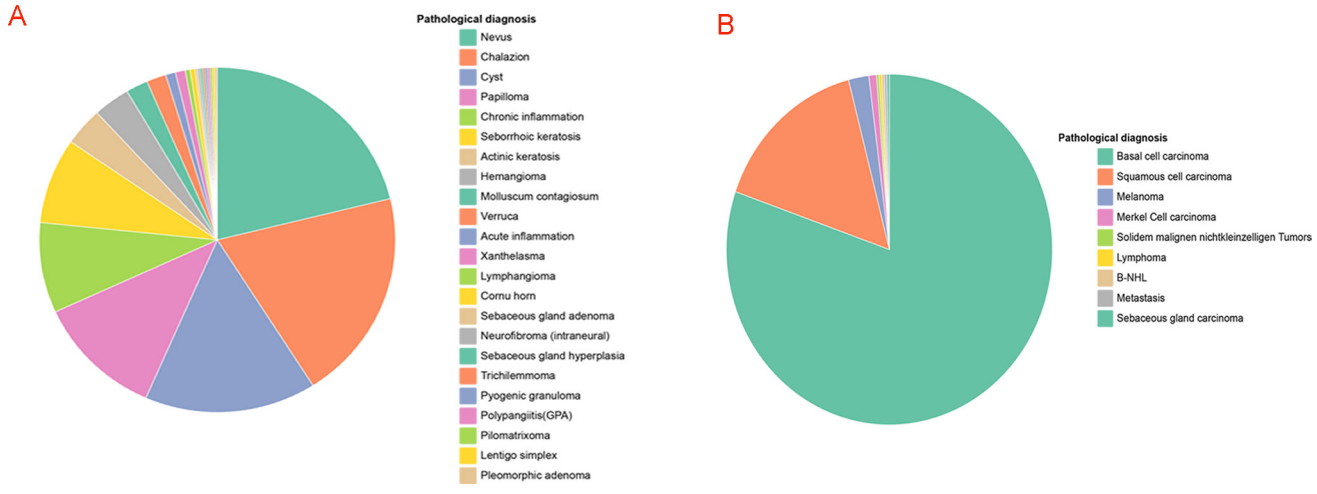

Figure 1 Distribution of benign and malignant eyelid tumours. (A) Distribution of benign eyelid tumours; (B) distribution of malignant eyelid tumours. B-NHL, B-cell Non-Hodgkin Lymphoma; GPA, granulomatosis with polyangiitis.

Table 2 Performance of the different models in the validation dataset

Deep neural network	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Precision (95% CI)	Recall (95% CI)	F1 (95% CI)
VGG16	0.904 (0.876 to 0.928)	0.746 (0.694 to 0.798)	0.790 (0.691 to 0.889)	0.976 (0.939 to 1.000)	0.747 (0.691 to 0.799)	0.750 (0.699 to 0.797)	0.737 (0.680 to 0.789)
ResNet50	0.922 (0.894 to 0.947)	0.794 (0.746 to 0.843)	0.932 (0.867 to 0.985)	0.962 (0.916 to 1.000)	0.792 (0.741 to 0.841)	0.792 (0.743 to 0.841)	0.792 (0.741 to 0.839)
Efficient-Net	0.916 (0.883 to 0.943)	0.806 (0.754 to 0.851)	0.877 (0.795 to 0.944)	0.939 (0.884 to 0.986)	0.804 (0.751 to 0.851)	0.808 (0.757 to 0.852)	0.805 (0.752 to 0.850)
Inception-V4	0.922 (0.893 to 0.947)	0.790 (0.734 to 0.839)	0.918 (0.850 to 0.974)	0.937 (0.877 to 0.978)	0.792 (0.738 to 0.838)	0.789 (0.737 to 0.836)	0.790 (0.737 to 0.836)
Dual VGG16	0.931 (0.907 to 0.953)	0.790 (0.738 to 0.839)	0.939 (0.881 to 0.987)	0.866 (0.789 to 0.929)	0.811 (0.762 to 0.856)	0.784 (0.734 to 0.832)	0.788 (0.737 to 0.837)
Dual ResNet50	0.923 (0.896 to 0.946)	0.806 (0.753 to 0.854)	0.844 (0.758 to 0.918)	0.976 (0.935 to 1.000)	0.806 (0.756 to 0.854)	0.808 (0.757 to 0.853)	0.801 (0.746 to 0.847)
Dual Efficient-Net	0.916 (0.883 to 0.945)	0.814 (0.761 to 0.858)	0.961 (0.912 to 1.000)	0.914 (0.854 to 0.971)	0.811 (0.761 to 0.858)	0.812 (0.762 to 0.858)	0.812 (0.759 to 0.856)
Dual Inception-V4	0.930 (0.900 to 0.954)	0.839 (0.790 to 0.883)	0.923 (0.861 to 0.975)	0.963 (0.913 to 1.000)	0.838 (0.788 to 0.884)	0.838 (0.788 to 0.882)	0.838 (0.787 to 0.882)

AUC, area under the receiver operating characteristic curve.

with high accuracy could significantly aid in early diagnosis and treatment planning.

An epidemiological survey by Levinkron *et al*[24] showed that out of 1423 tissue samples of eyelid tumours from Israel, 1210 (85%) were benign, and 213 (15%) were malignant/premalignant lesions. In the study by Sendul *et al*[25] revealed that among 428 eyelid lesions diagnosed by histopathology in Turkey, 373 (87%) were benign and 55 (13%) were malignant. The top three malignant eyelid tumours reported in the literature were BCC, SCC and SGC, and the top three benign eyelid tumours were SCP, seborrheic keratosis and nevus.[1 26–29] This is in general agreement with the statistical findings of our dataset, and some of these differences, especially in benign tumours, can be related to different regions and ethnicities. Similar to previous studies, we also observed an older age of onset of eyelid malignant tumours than benign tumours in our data.[1 25 30]

Compared with histopathology, the gold standard, oculoplastic specialists, general ophthalmologists and ophthalmology residents diagnose eyelid tumours with accuracies of 87.9%, 90.0% and 85.8%, respectively.[31] In contrast, DL models have demonstrated good performance, with reported accuracies ranging from 81.8% to 95.8% in various studies.[20 21 31] Adamopoulos *et al*[32] used 143 photographic images from the Eye Clinic of the University Hospital of Heraklion, Crete, Greece, to differentiate between BCCs. However, the sample size in that study was too small, and there was only one category of eyelid malignancies, BCC, so the model may be unable to identify other eyelid malignancies. In addition, the model's sensitivity, specificity and AUC were not analysed in this study. The model developed by Li[20] *et al* identified 1417 eyelid tumours in China as benign or malignant. However, they trained the model to localise the tumours before performing benign-malignant identification of the tumours, with an average accuracy of 76%, resulting in 24% of the tumours not being correctly identified as benign or malignant due to mislocalisation. Hui *et al*[21] developed and validated the model using 339 precisely located photographs of Chinese individuals with benign and malignant eyelid tumours. However, the study's sample size was small, and as in the above study, no healthy individuals were included as control photos. Therefore, our study included a larger number of benign and malignant tumour samples and a comparable number of healthy control group samples. Second, in addition to pinpointing the tumours before developing and validating the model, each of our photographs included only the unilateral periocular region, which allowed for better detection of small and difficult to detect tumours in the periocular region. Caucasians have a higher incidence of skin cancer, especially melanoma, largely due to their lower melanin content, which increases sensitivity to ultraviolet light.[33] This study is the first to develop an eyelid tumour detection model for a large sample of Caucasians, and the sample not only contained a variety of eyelid benign and malignant tumours but also included control samples from healthy individuals.

DL models are accurate for medical diagnosis, but they often lack clear explanations, which limits their use in clinics.[34 35] To solve this problem, we used Grad-CAM to show important areas the model uses to make decisions. The heatmaps clearly showed the eyelid tumour areas and matched the real tumour locations well for both benign and malignant tumours (figure 3). These images make the model easier to understand and help doctors trust DL systems for diagnosing eyelid tumours. Although this study demonstrates the great potential of DL systems in distinguishing between malignant and benign eyelid tumours, our study has some limitations. First, our system was developed based on Caucasians in the German region. Thus, more

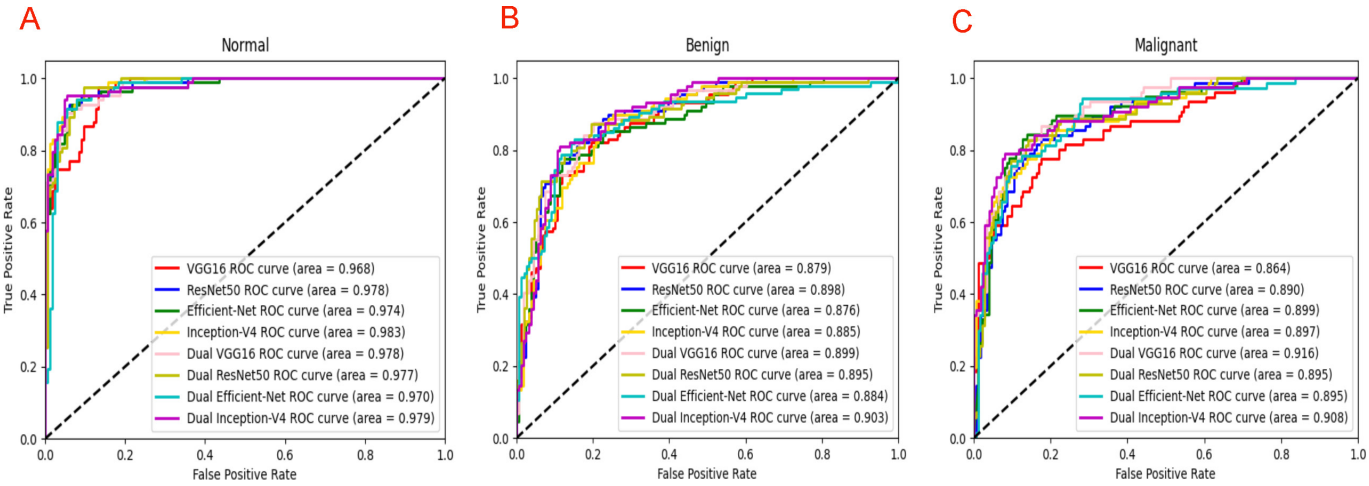

Figure 2 Performance of eight deep learning models across three categories of data. (A) The receiver operating characteristic (ROC) curves of the deep learning models in the normal test set. (B) The ROC curves of the deep learning models in benign tumour test set. (C) The ROC curves of the deep learning models in malignant tumour test set. AUC, area under the ROC curve.

samples from different regions and ethnicities must be included to enrich the diversity of the samples. Further validation of the effectiveness in other ethnic populations is needed. Second, the model in this study can identify benign and malignant tumours but cannot diagnose specific types of tumours. Therefore, we need to collect more images of different kinds of eyelid tumours to develop a model that can precisely diagnose tumour types. Our dataset focuses on skin-surface eyelid tumours, excluding cases primarily involving the conjunctiva, which may limit the model's applicability. Additionally, the system assigns a single label per image, making it unable to differentiate multiple coexisting lesions. Future work should explore multi-modal DL to integrate conjunctival lesion analysis and adopt object detection or segmentation techniques to identify multiple tumours within a single image. Third, the tumours in this study have been accurately localised, and the reliability of the automatic localisation still needs to be verified. Therefore, we need to include more photos of accurately localised tumours in the future to achieve more accurate automatic localisation. Finally, the visual transformer (ViT) model was not evaluated due to computational power limitations; only the CNN model was evaluated. Future studies should consider ViT and other advanced architectures to determine their effectiveness in classifying eyelid tumours.

In conclusion, the study shows that the DL system we developed can distinguish between malignant and benign aspects of eyelid tumours. This system is expected to help doctors and suspected patients proactively screen for eyelid tumours and detect malignant tumours early. We plan to apply the system to smartphones in the future, which will help more patients with early screening and detection of eyelid tumours as well as help clinicians with diagnosis and treatment.

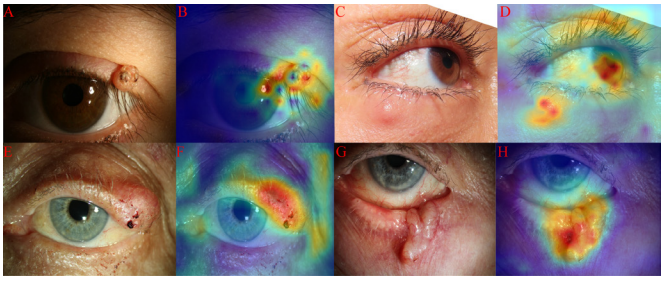

Figure 3 Photographic images and corresponding heatmaps of benign and malignant eyelid tumours. (A–D) Benign eyelid tumour: (A, B) Nevus. (C, D) Chalazion. (E–H) Malignant eyelid tumour: (E, F) Merkel cell carcinoma. (G, H) Basal cell carcinoma.

Contributors WF, WD and LMH designed the study; LMH provided data. WF and WD analysed and interpreted the data; WF drafted the manuscript; MJJ, WD and LMH revised the manuscript. WD and LMH are responsible for the overall content as the guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval The study followed the principles of the Declaration of Helsinki and received ethical approval by the Ethics Committee of the University of Cologne (23-1486-retro). The study was a retrospective data analysis, and the Ethics Committee did not consider that informed patient consent was required.

Provenance and peer review Not commissioned; externally peer-reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Wanlin Fan http://orcid.org/0000-0001-7143-6707
Martine Johanna Jager http://orcid.org/0000-0003-2261-3820
Ludwig M Heindl http://orcid.org/0000-0002-4413-6132

REFERENCES

1 Deprez M, Uffer S. Clinicopathological features of eyelid skin tumors. A retrospective study of 5504 cases and review of literature. *Am J Dermatopathol* 2009;31:256–62.
2 Cook BE, Bartley GB. Treatment options and future prospects for the management of eyelid malignancies: an evidence-based update. *Ophthalmology* 2001;108:2088–98.
3 Tesluk GC. Eyelid lesions: incidence and comparison of benign and malignant lesions. *Ann Ophthalmol* 1985;17:704–7.
4 Wang L, Shan Y, Dai X, *et al*. Clinicopathological analysis of 5146 eyelid tumours and tumour-like lesions in an eye centre in South China, 2000–2018: a retrospective cohort study. *BMJ Open* 2021;11:e041854.
5 Cook BE, Bartley GB. Epidemiologic characteristics and clinical course of patients with malignant eyelid tumors in an incidence cohort in Olmsted County, Minnesota. *Ophthalmology* 1999;106:746–50.
6 Wu A, Rajak SN, Huilgol SC, *et al*. Re-evaluating the epidemiology of cutaneous sebaceous carcinoma. *Australas J Dermatol* 2020;61:e135–6.
7 Boulos PR, Rubin PAD. Cutaneous melanomas of the eyelid. *Semin Ophthalmol* 2006;21:195–206.
8 Huang Y-Y, Liang W-Y, Tsai C-C, *et al*. Comparison of the Clinical Characteristics and Outcome of Benign and Malignant Eyelid Tumors: An Analysis of 4521 Eyelid Tumors in a Tertiary Medical Center. *Biomed Res Int* 2015;2015:453091.

9 Resnikoff S, Felch W, Gauthier T-M, *et al*. The number of ophthalmologists in practice and training worldwide: a growing gap despite more than 200,000 practitioners. *Br J Ophthalmol* 2012;96:783–7.
10 Zhang K, Liu X, Shen J, *et al*. Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. *Cell* 2020;182:1360.
11 Shi Z, Miao C, Schoepf UJ, *et al*. A clinically applicable deep-learning model for detecting intracranial aneurysm in computed tomography angiography images. *Nat Commun* 2020;11:6090.
12 Li Z, Guo C, Nie D, *et al*. Deep learning from 'passive feeding' to 'selective eating' of real-world data. *NPJ Digit Med* 2020;3:143.
13 Li Z, Guo C, Nie D, *et al*. Automated detection of retinal exudates and drusen in ultra-widefield fundus images based on deep learning. *Eye (Lond)* 2022;36:1681–6.
14 Keenan TDL, Chen Q, Agrón E, *et al*. DeepLensNet: Deep Learning Automated Diagnosis and Quantitative Classification of Cataract Type and Severity. *Ophthalmology* 2022;129:571–84.
15 Takahashi H, Tampo H, Arai Y, *et al*. Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy. *PLoS One* 2017;12:e0179790.
16 Peng Y, Dharssi S, Chen Q, *et al*. DeepSeeNet: A Deep Learning Model for Automated Classification of Patient-based Age-related Macular Degeneration Severity from Color Fundus Photographs. *Ophthalmology* 2019;126:565–75.
17 Ting DSW, Cheung CY-L, Lim G, *et al*. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA* 2017;318:2211–23.
18 Rasheed HA, Davis T, Morales E, *et al*. DDLSNet: A Novel Deep Learning-Based System for Grading Funduscopic Images for Glaucomatous Damage. *Ophthalmol Sci* 2023;3:100255.
19 Muhammad H, Fuchs TJ, De Cuir N, *et al*. Hybrid Deep Learning on Single Wide-field Optical Coherence tomography Scans Accurately Classifies Glaucoma Suspects. *J Glaucoma* 2017;26:1086–94.
20 Li Z, Qiang W, Chen H, *et al*. Artificial intelligence to detect malignant eyelid tumors from photographic images. *NPJ Digit Med* 2022;5:23.
21 Hui S, Dong L, Zhang K, *et al*. Noninvasive identification of Benign and malignant eyelid tumors using clinical images via deep learning system. *J Big Data* 2022;9:84.
22 Graham B. 22. Kaggle diabetic retinopathy detection competition report. University of Warwick; 2015.
23 Selvaraju RR, Cogswell M, Das A, *et al*. Grad-cam: visual explanations from deep networks via gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV); 618–26. Venice.
24 Levinkron O, Schwalb L, Shoufani A, *et al*. Comparison of the clinical characteristics of benign and malignant eyelid lesions: an analysis of 1423 eyelid lesions, compared between ophthalmology department and plastics department. *Graefes Arch Clin Exp Ophthalmol* 2024;262:615–21.
25 Sendul SY, Akpolat C, Yilmaz Z, *et al*. Clinical and pathological diagnosis and comparison of benign and malignant eyelid tumors. *J Fr Ophtalmol* 2021;44:537–43.
26 Yu S-S, Zhao Y, Zhao H, *et al*. A retrospective study of 2228 cases with eyelid tumors. *Int J Ophthalmol* 2018;11:1835–41.
27 Shields JA, Shields CL. Sebaceous adenocarcinoma of the eyelid. *Int Ophthalmol Clin* 2009;49:45–61.
28 Silverman N, Shinder R. What's New in Eyelid Tumors. *Asia Pac J Ophthalmol (Phila)* 2017;6:143–52.
29 Yin VT, Merritt HA, Sniegowski M, *et al*. Eyelid and ocular surface carcinoma: diagnosis and management. *Clin Dermatol* 2015;33:159–69.
30 Xu XL, Li B, Sun XL, *et al*. Eyelid neoplasms in the Beijing Tongren Eye Centre between 1997 and 2006. *Ophthalmic Surg Lasers Imaging* 2008;39:367–72.
31 Lee MJ, Yang MK, Khwarg SI, *et al*. Differentiating malignant and benign eyelid lesions using deep learning. *Sci Rep* 2023;13:4103.
32 Adamopoulos A, Chatzopoulos EG, Anastassopoulos G, *et al*. Eyelid basal cell carcinoma classification using shallow and deep learning artificial neural networks. *Evolving Systems* 2021;12:583–90.
33 You W, Henneberg R, Coventry BJ, *et al*. Cutaneous malignant melanoma incidence is strongly associated with European depigmented skin type regardless of ambient ultraviolet radiation levels: evidence from Worldwide population-based data. *AIMS Public Health* 2022;9:378–402.
34 Thomas SM, Lefevre JG, Baxter G, *et al*. Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer. *Med Image Anal* 2021;68:101915.
35 Rudin C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nat Mach Intell* 2019;1:206–15.