PLOS ONE

Prediction of the ectasia screening index from raw Casia2 volume data for keratoconus identification by using convolutional neural networks

Maziar Mirsalehi[ID][1*], **Benjamin Fassbind**[1], **Andreas Streich**[2], **Achim Langenbucher**[1]

1 Department of Experimental Ophthalmology, Saarland University, Homburg, Germany, **2** Department of Computer Science, Eidgenössische Technische Hochschule, Zürich, Switzerland

* mami00016@stud.uni-saarland.de(MM)

Abstract

Purpose Prediction of the ectasia screening index, an estimator provided by the Casia2 instrument for identifying keratoconus, from raw optical coherence tomography data using convolutional neural networks.

Methods Three convolutional neural networks models (ResNet18, DenseNet121 and EfficientNetB0) were employed to predict the ectasia screening index. Mean absolute error was used as the performance metric for predicting the ectasia screening index by the adapted convolutional neural network models on the test set. Scans with an ectasia screening index value higher than a certain threshold were classified as *Keratoconus*, while the remaining scans were classified as *Not Keratoconus*. The architectures' performance was evaluated using metrics such as accuracy, sensitivity, specificity, positive predictive value and F1 score on data collected from patients examined at the eye clinic of the Homburg University Hospital. The raw data from the Casia2 instrument, in 3dv format, was converted into 16 images per examination of one eye. For the training, validation and testing phases, 3689, 1050 and 1078 scans (3dv files) were selected, respectively.

Results In the prediction of the ectasia screening index, the mean absolute error values for the adapted ResNet18, the adapted DenseNet121 and the adapted EfficientNetB0, rounded to two decimal places, were 7.15, 6.64 and 5.86, respectively. In the classification task, the three networks yielded an accuracy of 94.80%, 95.27% and 95.83%, respectively; a sensitivity of 92.07%, 94.64% and 94.17%, respectively; a specificity of 96.61%, 95.69% and 96.92%, respectively; a positive predictive value of 94.72%, 93.55% and 95.28%, respectively; and a F1 score of 93.38%, 94.09% and 94.72%, respectively.

Conclusions Our results show that the prediction of keratoconus based on the ectasia screening index values estimated from raw data outperforms previous approaches using processed data. adapted EfficientNetB0 outperformed both the other adapted models and those in state-of-the-art studies, with the highest accuracy and F1 score.

Introduction

Keratoconus describes a disorder of the eye characterised by a cone-shaped cornea with thinning and steepening, which typically affects both eyes of a patient with varying degrees of severity and occurs in both males and females [1]. Keratoconus affects about 1 in every 2000 individuals in the general population [2].

There are two main types of corneal imaging: corneal topography and corneal tomography. In corneal topography, the shape of the anterior part of the cornea is shown but in corneal tomography a three-dimensional image of the whole cornea is shown. Optical Coherence Tomography (OCT) is a corneal tomography technique that assesses the delay of reflected infrared light from the anterior segment by comparing it to a reference reflection. This tomography technique is classified into two types: Fourier domain, which uses a stationary mirror and time domain, which adjusts the position of a reference mirror. Another corneal tomography technique is Scheimpflug imaging where a rotating camera is used to produce cross-sectional images [3].

Artificial Intelligence (AI) enables machines to perform tasks associated with human cognition like writing, speaking and seeing. AI can be used in medical specialties dealing with image analysis like ophthalmology. Machine learning is a subset of AI that enables the machine to learn in order to develop its performance. Deep learning, a specialised branch of machine learning, improves the effectiveness of motion recognition, image and speech [4].

In this study, the neural networks were used to predict the Ectasia Screening Index (ESI) of a given scan automatically. This approach is a regression task since the output of the networks is a numerical value. Also, the scans were classified into two classes, *Keratoconus* and *Not Keratoconus*. The *Keratoconus* class represents ectasia and the *Not Keratoconus* class indicates suspicion of ectasia or no ectasia pattern. This approach has an advantage over other approaches where the output is discrete and belongs to a class. With this approach, if two scans are in the *Keratoconus* class, the severity of ectasia can be compared between them by comparing the predicted ESI provided by the architecture.

The ESI values, which are computed by the instrument's software, are used as labels for training the Convolutional Neural Networks (CNNs). The objective of this study is to estimate the ESI values directly from the raw data produced by the Casia2 instrument. By using raw data, we ensure that the underlying physical information remains consistent, even if the software version changes in the future. In general, data can be utilised as preprocessed data or as raw data. Preprocessed data are altered by software and the details of these modifications may not always be transparent. Moreover, changes in software versions can lead to variations in how data are preprocessed and affect the consistency of results. In contrast, raw data remain unaltered by external software. Therefore, raw data retain their original form across different software versions. This stability in raw data can offer a more consistent and reliable foundation for analysis and model training. This approach differs from training CNNs on the OCT images produced by the software, as those images have already undergone post-processing steps such as noise reduction and filtering. Therefore, training on the raw data is not a redundant task; it allows us to develop a model that learns directly from unaltered input, making the estimation process more robust and independent of software-specific image modifications.

The ESI values provide an indication of whether the eye is clinically healthy, affected by keratoconus, or shows signs suggestive of keratoconus. They assist clinicians in diagnosing the severity of corneal ectasia and in determining the appropriate timing for intervention. Accurate estimation of the ESI is therefore clinically significant, as it supports early detection and management of keratoconus, potentially preventing disease progression and preserving visual function.

To the best of our knowledge, this is the first time that raw OCT data have been used for a regression task to predict the ESI for the purpose of keratoconus diagnosis. Below we briefly review the current neural network-based approaches to automatically identify keratoconus.

State of the art

Zhang et al. [5] explored keratoconus diagnosis by employing the CorNet model. The model was trained and evaluated with a dataset of 1786 raw data from the Corvis ST (Oculus, Wetzlar, Germany). Corvis ST is a non-contact device that measures corneal biomechanics by recording dynamic deformation following a rapid air-puff excitation. Keratoconus was diagnosed by using clinical signs such as stromal thinning, Fleischer's ring and a central K-value greater than 47 dioptres, in addition to other indicators. The CorNet model achieved an accuracy of 92.13%, sensitivity of 92.49%, specificity of 91.54%, Positive Predictive Value (PPV) of 94.77% and an F1 score of 93.62% on the validation set.

Ruiwei Feng et al. [6] introduced a deep learning method named KerNet for identifying keratoconus and sub-clinical keratoconus using raw data from the Pentacam HR system (Oculus, GmbH, Wetzlar, Germany). This system includes a rotating Scheimpflug camera, which gathers three-dimensional data of the cornea and a software which is designed to analyse and display the data. The corneal data, exported from the Pentacam HR system, comprised five numerical matrices for each sample. These matrices were considered as five two-dimensional image slices, representing the front and back surface curvatures, the front and back surface elevations and the pachymetry of the eye. 854 samples were used as dataset. KerNet employed a specialised architecture with five branches to handle the matrices individually as input to identify features, which are subsequently combined for prediction. The model achieved an accuracy of 94.74%, with a sensitivity of 93.71%, PPV of 94.10% and an F1 score of 93.89%.

Schatteburg et al. [7] introduced a plan for using CNNs for keratoconus diagnosis based on the ESI from data of the SS-1000 Casia OCT Imaging System. The dataset sourced from over 1900 patients and included three-dimensional OCT images of both the anterior and posterior cornea, together with parameters calculated by the Casia software. However, the study did not include evaluation metrics.

Fassbind et al. [8] focused on identifying abnormalities such as keratoconus by employing CorNeXt as a CNN model. In this study, cornea topography maps from the Casia2 anterior OCT instrument were used. The used CorNeXt model is based on the ConvNeXt [9] CNN architecture. To employ ConvNeXt for corneal disease classification, modifications to the architecture were implemented. Measurements of axial refractive power, as well as the elevation of the cornea's front and back surfaces and its thickness were taken from the scan for every individual cornea and five related maps were created and displayed as grayscale images. ConvNeXt was adapted to include all cornea data by stacking these maps into a five-channel pseudo-image. The dataset included a total of 2182 scans (1552 scans for training, 388 scans for validation and 242 scans for test). The model achieved a sensitivity of 98.46% and a specificity of 91.96% in distinguishing healthy from pathological corneas. For the labeled class of keratoconus, it reached 92.56% accuracy, 84.07% sensitivity, 100% specificity and a 91.34% F1 score.

Materials and methods

Convolutional neural network

Artificial Neural Networks (ANNs) mimic the brain's processing through nodes and weighted connections, learning via adjustable weights during training [10]. CNNs, a specialised form

of ANN, are designed for image data, using convolutional layers with convolution kernels, referred to as filters, to detect features and generate feature maps [11,12]. For this study, three CNN models were selected.

Quality criteria

In this study, Mean Squared Error (MSE) is used as a loss function for the regression task. MSE is a derivable criterion and having a derivable criterion is essential for gradient descent algorithms, which are used universally to adjust weights in neural networks during training. MSE is defined as Eq (1), where N signifies the number of actual values, which is equivalent to the number of predicted values; y_i represents the actual value at position i and $\hat{y}_i$ represents the predicted value at the same position [13].

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i|^2 \tag{1}$$

To compare the performance of different prediction models, Mean Absolute Error (MAE) is used, as this measures the average absolute difference between the actual values and the predicted values by the model [11]. Eq (2) illustrates the MAE computation, where N, y_i and $\hat{y}_i$ retain the same meanings as in Eq (1) [13].

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{2}$$

Key metrics for evaluating a binary classifier are derived from the four entries in the confusion matrix. They are crucial for assessing the classifier's performance. True Positive (TP) signifies the count of correctly classified positive samples, such as images with keratoconus correctly identified as having keratoconus. True Negative (TN) represents the count of correctly classified negative samples, like images without keratoconus correctly identified as not having keratoconus. False Positive (FP) refers to the count of samples that have been incorrectly classified as positive; that is, in our case, images without keratoconus mistakenly identified as having keratoconus. False Negative (FN) indicates the count of samples that have been incorrectly classified as negative, such as images with keratoconus incorrectly identified as not having keratoconus. Fig 1 shows the confusion matrix.

In this study, the metrics below are used to assess how effectively the architectures classify the data into two different categories [14].

Accuracy measures the proportion of correctly classified samples out of the total number of samples in the test dataset. Accuracy is calculated as [14]

$$\textbf{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{3}$$

Sensitivity is the proportion of correctly identified positive samples out of all actual positive samples, calculated as [14]:

$$\textbf{Sensitivity} = \frac{TP}{TP + FN} \tag{4}$$

Fig 1. Confusion matrix. *Abbreviations:* TN = True Negative, FP = False Positive, FN = False Negative, TP = True Positive.

Specificity measures the proportion of correctly classified negative samples out of all samples classified as negative [14]:

$$\textbf{Specificity} = \frac{TN}{TN + FP} \tag{5}$$

PPV is defined as the proportion of correctly classified samples relative to all samples predicted to belong to the positive class [14]:

$$\textbf{PPV} = \frac{TP}{TP + FP} \tag{6}$$

As we are using a threshold on the estimated ESI, a high sensitivity or specificity can be trivially achieved at the cost of a useless low value of the respective other metric. The F1-score finds a balance between these two metrics. The F1 score is defined as [14]:

$$\textbf{F1} = 2 \times \frac{PPV \times Sensitivity}{PPV + Sensitivity} \tag{7}$$

Furthermore, the F1 score has an advantage when dealing with imbalanced datasets, where one class significantly outnumbers the other. In such cases, metrics like accuracy, sensitivity and specificity may not effectively measure how well the model distinguishes between classes. Therefore, the F1 score can be used because it provides a more balanced evaluation of the model's performance.

The Receiver Operating Characteristic (ROC) curve was analysed to find the best trade-off between sensitivity and specificity for predictions by identifying the optimal threshold, which is the point that maximises the difference between the true positive rate (sensitivity) and the false positive rate (1-specificity). Following this, the predicted ESI values were classified into positive and negative classes based on the optimal threshold to compute the confusion
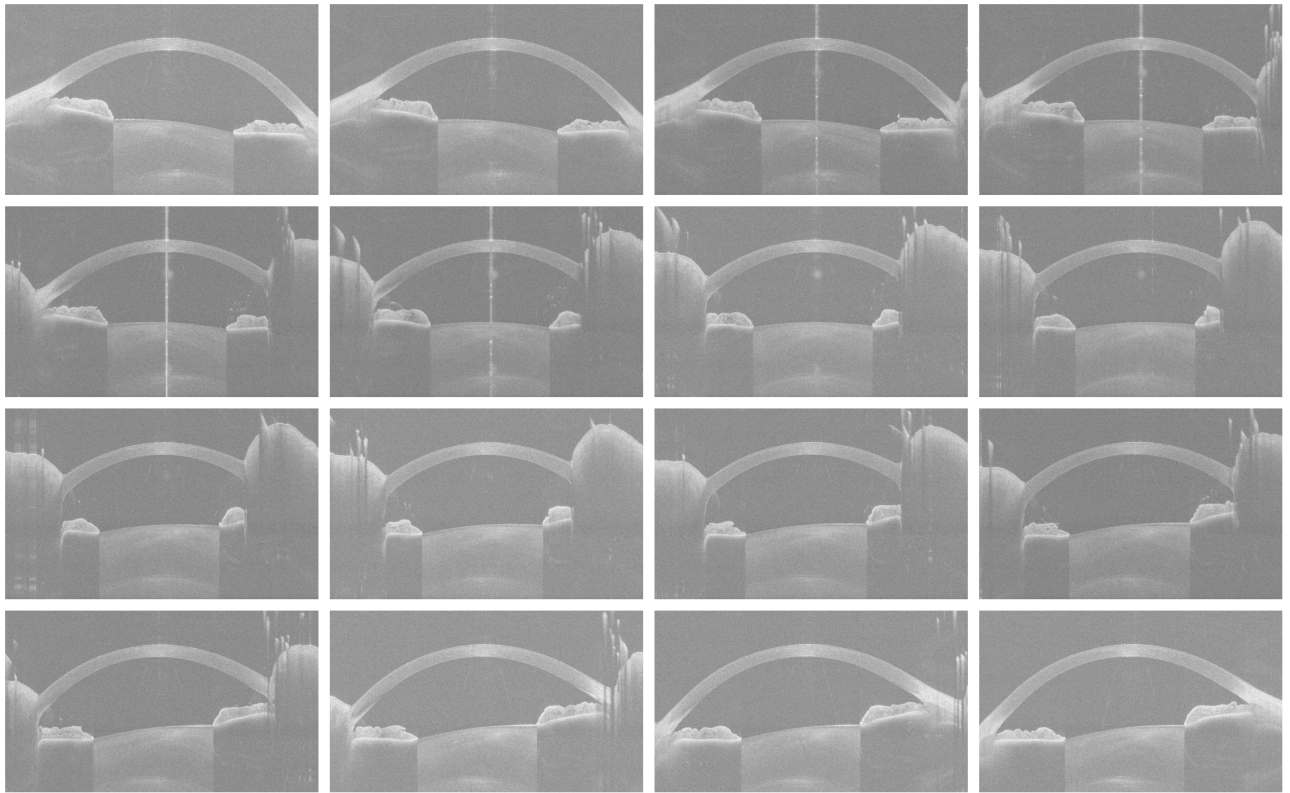
matrix values. The predicted ESI values that are equal to or exceed the threshold are considered as *Keratoconus* which indicate the presence of ectasia and those below the threshold are categorised as *Not Keratoconus* which indicate suspicion of ectasia or no ectasia pattern.

Data

In this study, the data were obtained from patients examined at the eye clinic of the Homburg University Hospital, between February 01, 2021 and September 01, 2023. The data were anonymised at the source and were transferred to us for further processing on October 02, 2023. We were freed from the requirement for ethics approval for the data by the ethics committee of the Saarland medical council (registration number 157/21). Age and sex were not considered important. The instrument, Cornea/Anterior Segment OCT Casia2 from Tomey Corporation, made in Japan, was used for data acquisition from patients. This instrument uses optical coherence tomography with a 1310 nm wavelength laser to measure different parameters, such as corneal thickness, the depth from the anterior surface of the cornea to the anterior surface of the crystalline lens and the depth from the posterior surface of the cornea to the anterior surface of the crystalline lens. The scan range is 13 mm in depth and 16 mm in diameter. The Casia2 instrument has two modes available: 'Anterior Segment mode' and 'Lens mode'. In Anterior Segment mode, high-sensitivity measurements of the cornea, angle and intraocular lens can be performed, but it does not allow visualisation of the posterior lens. Lens mode provides a simultaneous view of the entire area from the cornea to the posterior lens. Since in this study the visualisation of the posterior lens is not important for the detection of keratoconus, the Anterior Segment mode was selected. The Casia2 instrument produces raw data after measurement, which is in the format of 3dv. Each 3dv file related to the corneal map is 36.6 MB in size. For each 3dv file there is an xpf file that contains metadata about the measurement, including the examined eye (left or right), date and time of the examination and the exam protocol name. For each measurement, the ESI is stored in a csv file, which can be exported from the Casia2 instrument's software. Ectasia screening identifies keratoconus by independently analysing the shapes of the anterior and posterior cornea. The final diagnosis is based on the results from both assessments. For the anterior cornea, the evaluation focuses on spherical, asymmetry and regular astigmatism components of Fourier analysis. For the posterior cornea, the evaluation focuses on the steepest point of instantaneous power, as well as the asymmetry, regular and higher-order irregular astigmatism components of Fourier analysis. If the analysis area is insufficient for either cornea, the result for that cornea will be marked as 'N/A'. The final diagnosis is determined by the higher score from either assessment; if both are 'N/A', the final result will also be 'N/A'. If the ESI result ranges from 0 to 4, no ectasia pattern is detected. If the the ESI result is between 5 and 29 suggests a suspicion of ectasia and a result between 30 and 95 indicates clinical ectasia.

We used a Python [15] script to extract 16 images from raw data (3dv file) which originally were stored in a 16-bit unsigned integer format. Each image, with a resolution of 800 pixels in width and 1464 pixels in height, was then saved as a grayscale PNG file. Fig 2 shows a series of 16 resized images of a left eye with an ESI of 0, where the height has been reduced to one-third of the original dimension by using a Python script to better represent the realistic shape of the eye. The image preprocessing involved cropping 25% from the left side of the images and 25% from the right side of the images to exclude unnecessary eyelid areas and 60% from the bottom of the images to remove regions that did not cover the cornea. After that, the images were resized to a dimension of 224×224 pixels.

Fig 2. Resized images of a left eye with an ESI of 0. *Abbreviation:* ESI = Ectasia Screening Index.

Experimental design and implementation

Since CNNs are suited for detecting objects within images [12], three models (ResNet18, DenseNet121 and EfficientNetB0) were selected based on their performance in the field. ResNet was examined on ImageNet and CIFAR-10 [16], DenseNet was tested on CIFAR-10, CIFAR-100, SVHN and ImageNet [17] and EfficientNet was evaluated on ImageNet and transfer learning datasets, including CIFAR-10, CIFAR-100, Birdsnap, Stanford Cars, Flowers, FGVC Aircraft, Oxford-IIIT Pets and Food-101 [18].

ResNet18 is a variant of the residual network architecture. In residual networks, short-cut connections are used to bypass one or more layers and implement identity mapping which allow their outputs to be summed with the outputs of the intermediate layers [16]. DenseNet121 belongs to the dense convolutional network series. In this type of neural networks, all layers are connected directly with each other which allow them to receive additional inputs from preceding layers and propagate their feature maps to subsequent layers. Unlike residual networks, features are concatenated rather than summed before being forwarded to the subsequent layer [17]. EfficientNetB0 is part of the EfficientNet series. In EfficientNet, the depth, width and resolution of the network are uniformly scaled by a specific set of scaling coefficients [18].

All CNN models were trained from scratch using Python and the PyTorch library [19] on a system equipped with an 11th Gen Intel(R) Core(TM) i7-11700@2.5 GHz processor, 32 GB of RAM and a 64-bit operating system with an x64-based processor. The input images were 16-channel, whereas standard pretrained models are designed for 3-channel

Red-Green-Blue input. Although it is technically possible to adapt pretrained networks to accept multi-channel inputs, for example by averaging pretrained weights across channels, such modifications introduce complexity and may reduce the benefit of transfer learning, especially when the additional channels contain modality-specific information not represented in natural images. Therefore, all architectures were trained from scratch. The training proceeded for 100 epochs, during which the validation MSE became stable. The data were divided into disjoint training, validation and test datasets to ensure that the architectures were trained on one subset, evaluated on another to detect overfitting (where the architecture fails to apply its learned patterns from training data to unseen data [20]) and finally tested on a separate unseen subset to assess their ability to perform on new data. The batch sizes for the training, validation and test sets were set to 64. From a total of 15457 3dv files, 5817 were selected for training, validation and testing. The files not chosen were excluded due to defects on the cornea, such as keratoplasty. During the training phase, 3689 scans (stored as 3dv file) were used. This represents approximately 63.42% of the total dataset. Similarly, the validation phase involved 1050 scans (accounting for around 18.05% of the total) and the testing phase consisted of 1078 scans (accounting for around 18.53% of the total).

Table 1 presents the distribution of 3dv files which were used for training, validation and testing. The dataset is categorised based on the ESI, with a threshold of 30, as determined by the Casia2 instrument. An ESI of 30 or greater indicates the *Keratoconus* class, which signifies clinical ectasia.

An ESI below 30 classifies the files as *Not Keratoconus* class, indicating either a suspicion of ectasia or no ectasia pattern detected.

Every set of 16 images from a single 3dv file was stacked together. These stacked images were fed into the architectures, with the first convolutional layer modified to accept a 16-channel input. The fully connected layer for the output was also modified to produce a single output. Additionally, an extra fully connected layer was included to process the combined features which integrates one feature from the architecture and two features representing the eye parameters (right eye and left eye). This formed a combined feature vector of dimension three, which was passed through a final linear layer to yield the predicted ESI. Each ESI value was used as the label for a set of 16 stacked images in the adapted CNN models. For the training process, MSE was used as the loss function to minimise prediction errors. Adam is a favoured optimiser for training deep neural networks due to its quicker convergence compared to stochastic gradient descent [21]. Based on [21], AdamW converges faster and generalises better than Adam. In the experiments, the model parameters were optimised using the AdamW optimiser with a learning rate of 0.01 and a weight decay of 0.05. Moreover, a scheduler was implemented to adjust the learning rate on a plateau, with a reduction factor of 0.1 and a patience of 10 epochs.
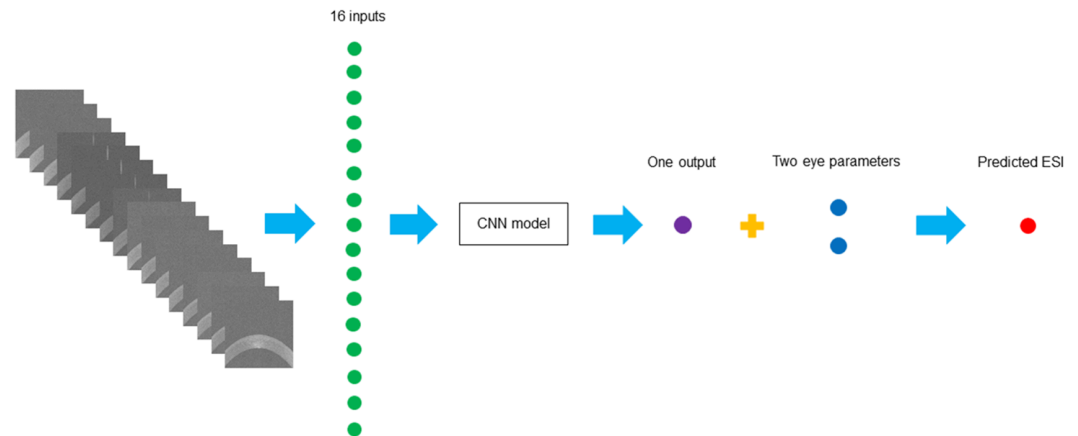
Fig 3 illustrates the workflow for predicting the ESI by using the adapted CNN models.

Table 1. Dataset distribution of 3dv files and classes.

Dataset	Total	*Keratoconus* class	*Not Keratoconus* class
Train	3689	1486 (40%)	2203 (60%)
Validation	1050	405 (39%)	645 (61%)
Test	1078	429 (40%)	649 (60%)

Fig 3. Workflow diagram for predicting the ESI. *Abbreviations:* ESI = Ectasia Screening Index, CNN = Convolutional Neural Network.

https://doi.org/10.1371/journal.pone.0311036.g003

Results and discussion

Table 2 presents the MAE and MSE values, rounded to two decimal places, derived from the evaluation of the adapted ResNet18, the adapted DenseNet121 and the adapted EfficientNetB0 on the test dataset.

Fig 4 shows Kernel Density Estimates (KDEs) of errors between the predicted ESIs and the actual ESIs for the adapted ResNet18, the adapted DenseNet121 and the adapted Efficient-NetB0. These KDE plots represent the distribution of errors, where the error is determined by subtracting the actual ESI from the predicted ESI.

Table 3 provides a summary of the frequency of errors within specified error ranges for the adapted CNN models.

Fig 5 illustrates the correlation between the actual ESIs and the architecture predictions for the adapted ResNet18, the adapted DenseNet121 and the adapted EfficientNetB0, respectively.

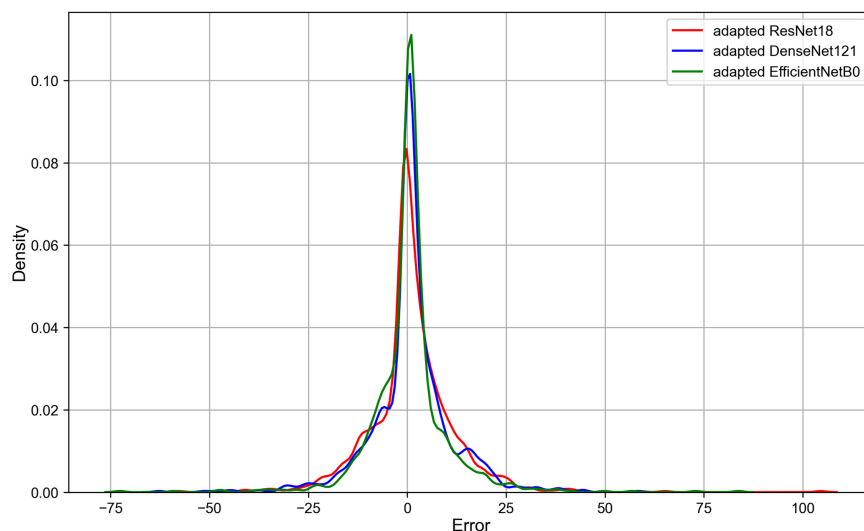Fig 6 shows the confusion matrices for each of the CNN architectures tested.

Fig 7 illustrates the ROC curves for the three adapted CNN models. The ROC curve analysis is based on the predicted ESI values. The optimal classification thresholds, determined using Youden's Index and rounded to two decimal places, were 33.23 for the adapted ResNet18, 30.61 for the adapted DenseNet121 and 32.12 for the adapted EfficientNetB0. These values correspond to the points on each curve that maximise the trade-off between sensitivity and specificity.

Table 4 presents a comparison of classification performance metrics for the adapted ResNet18, the adapted DenseNet121 and the adapted EfficientNetB0 (rounded to four decimal places) with three models of CorNet [5], KerNet [6] and CorNeXt [8] on the test set.

Table 2. Test set MAE and MSE performance of the adapted CNN models. *Abbreviations:* MAE = Mean Absolute Error, MSE = Mean Squared Error, CNN = Convolutional Neural Network.

CNN architecture	MAE	MSE
adapted ResNet18	7.15	122.04
adapted DenseNet121	6.64	110.33
adapted EfficientNetB0	**5.86**	**101.05**

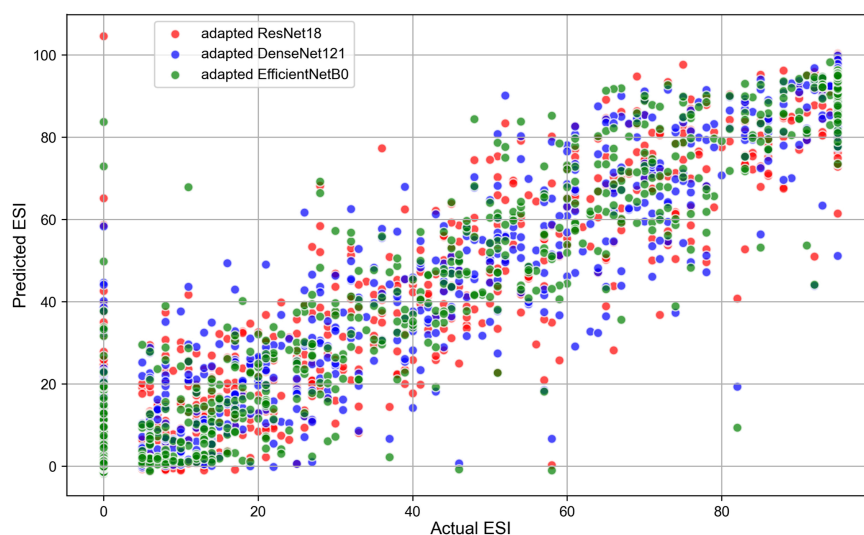https://doi.org/10.1371/journal.pone.0311036.t002

Fig 4. KDEs of errors between the predicted ESIs and the actual ESIs for the different CNN architectures.
Abbreviations: KDE = Kernel Density Estimate, ESI = Ectasia Screening Index, CNN = Convolutional Neural Network.

https://doi.org/10.1371/journal.pone.0311036.g004

Table 3. **Frequency of errors for CNN architectures within specified ranges.** *Abbreviation:* CNN = Convolutional Neural Network.
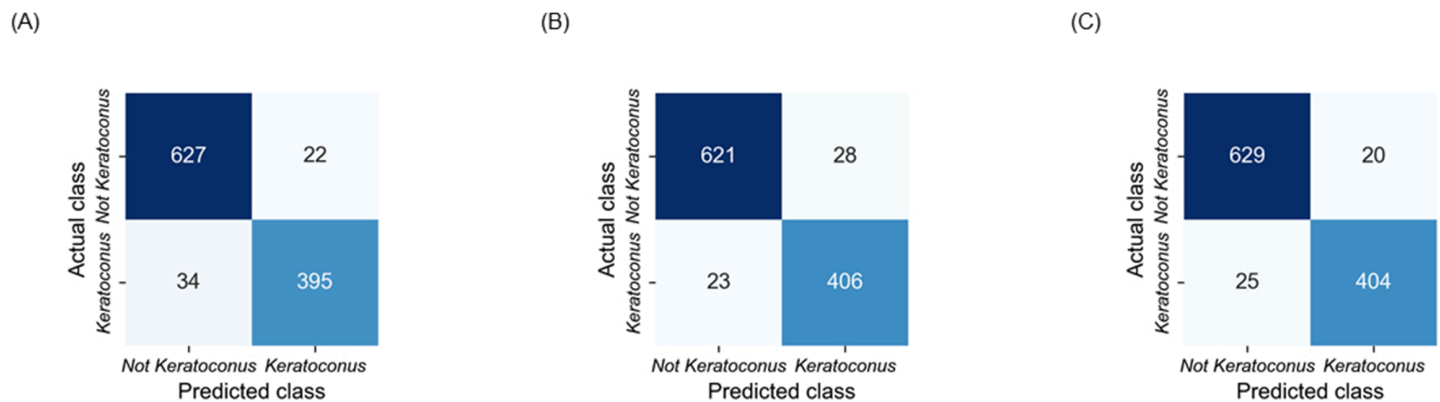
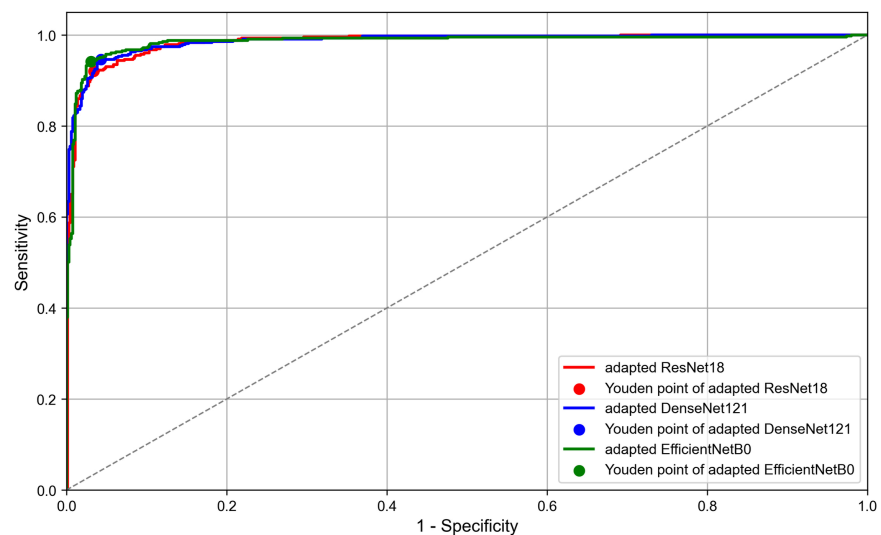| CNN architecture	Error range							
	below -10	-10 to -5	-5 to -2	-2 to 0	0 to 2	2 to 5	5 to 10	above 10
adapted ResNet18	124	87	74	232	133	134	134	160
adapted DenseNet121	107	100	62	142	267	138	121	141
adapted EfficientNetB0	86	113	89	147	298	149	86	110

https://doi.org/10.1371/journal.pone.0311036.t003

Fig 5. Correlation between the actual ESIs and the architecture predictions for the different CNN architectures.
Abbreviations: ESI = Ectasia Screening Index, CNN = Convolutional Neural Network.

https://doi.org/10.1371/journal.pone.0311036.g005

(A)

(B)

(C)

Fig 6. Confusion matrices of CNN architectures. (A) adapted ResNet18. (B) adapted DenseNet121. (C) adapted EfficientNetB0. Abbreviation: CNN = Convolutional Neural Network.

Fig 7. ROC curves for the three CNN architectures. *Abbreviations:* ROC = Receiver Operating Characteristic, CNN = Convolutional Neural Network.

Table 4. Evaluation metrics for CNN architectures. *Abbreviations:* CNN = Convolutional Neural Network, PPV = Positive Predictive Value.

CNN architecture	Metrics				
	Accuracy	Sensitivity	Specificity	PPV	F1 Score
adapted ResNet18	0.9480	0.9207	0.9661	0.9472	0.9338
adapted DenseNet121	0.9527	0.9464	0.9569	0.9355	0.9409
adapted EfficientNetB0	**0.9583**	**0.9417**	**0.9692**	**0.9528**	**0.9472**
CorNet [5]	0.9213	0.9249	0.9154	0.9477	0.9362
KerNet [6]	0.9474	0.9371	None	0.9410	0.9389
CorNeXt [8]	0.9256	0.8407	1	None	0.9134

This study explored the use of three CNN architectures (adapted ResNet18, adapted DenseNet121 and adapted EfficientNetB0) for predicting the ESI by using raw data from the Casia2 instrument.

Based on the performance metrics presented in the Table 2, the adapted EfficientNetB0 showed the best performance in predicting the ESIs on the test dataset. According to Fig 4, the peak around 0 indicates that most predictions from all three architectures (the adapted ResNet18, the adapted DenseNet121, and the adapted EfficientNetB0) are very close to the actual ESI values. Also, the plots are centered around zero, which indicates that the errors are symmetrically distributed on either side of the zero error line. Moreover, the adapted EfficientNetB0 model has the highest peak, which indicates that it has the highest proportion of predictions with smaller errors compared to the other two architectures. Additionally, all architectures show very low densities of extreme errors (far from zero) which is consistent with Fig 5. According to Table 4, the adapted EfficientNetB0 achieved higher accuracy and F1 score in distinguishing between *Keratoconus* and *Not Keratoconus* classes compared to the two other adapted CNN models and the CorNet, KerNet and CorNeXt models. The higher accuracy and F1 score rates observed for the adapted EfficientNetB0 emphasises the potential of this CNN architecture in distinguishing between *Keratoconus* and *Not Keratoconus* classes based on the raw data from the Casia2 instrument.

Future research could explore the applicability of other CNN models beyond the ones evaluated in this study to further enhance performance metrics.

Conclusions

To the best of our knowledge, this study is the first to use raw OCT data from the Casia2 instrument to predict the ESI. In conclusion, the adapted EfficientNetB0 outperformed the adapted ResNet18, the adapted DenseNet121 and the models in state-of-the-art studies in distinguishing between *Keratoconus* and *Not Keratoconus* classes. This highlights the effectiveness of this CNN architecture in improving diagnostic accuracy and F1 score based on raw data from the Casia2 instrument and suggests its significant potential for enhancing ophthalmological evaluations.

Author contributions

Conceptualization: Maziar Mirsalehi.

Data curation: Maziar Mirsalehi, Benjamin Fassbind.

Formal analysis: Maziar Mirsalehi.

Funding acquisition: Achim Langenbucher.

Investigation: Maziar Mirsalehi.

Methodology: Maziar Mirsalehi.

Project administration: Achim Langenbucher.

Software: Maziar Mirsalehi, Benjamin Fassbind.

Supervision: Andreas Streich, Achim Langenbucher.

Validation: Maziar Mirsalehi.

Visualization: Maziar Mirsalehi.

Writing – original draft: Maziar Mirsalehi.

Writing – review & editing: Benjamin Fassbind, Andreas Streich, Achim Langenbucher.

References

1. Santodomingo-Rubido J, Carracedo G, Suzaki A, Villa-Collar C, Vincent SJ, Wolffsohn JS. Keratoconus: An updated review. Cont Lens Anterior Eye. 2022;45(3):101559. https://doi.org/10.1016/j.clae.2021.101559 PMID: 34991971

2. Rabinowitz YS. Keratoconus. Surv Ophthalmol. 1998;42(4):297–319. https://doi.org/10.1016/s0039-6257(97)00119-7 PMID: 9493273

3. Fan R, Chan TC, Prakash G, Jhanji V. Applications of corneal topography and tomography: A review. Clin Exp Ophthalmol. 2018;46(2):133–46. https://doi.org/10.1111/ceo.13136 PMID: 29266624

4. Almodin E, Nassaralla BA, Sandes J. Keratoconus: A comprehensive guide to diagnosis and treatment. Springer Nature; 2022.

5. Zhang P, Yang L, Mao Y, Zhang X, Cheng J, Miao Y, et al. CorNet: Autonomous feature learning in raw Corvis ST data for keratoconus diagnosis via residual CNN approach. Comput Biol Med. 2024;172:108286. https://doi.org/10.1016/j.compbiomed.2024.108286 PMID: 38493602

6. Feng R, Xu Z, Zheng X, Hu H, Jin X, Chen DZ, et al. KerNet: A novel deep learning approach for keratoconus and sub-clinical keratoconus detection based on raw data of the pentacam HR system. IEEE J Biomed Health Inform. 2021;25(10):3898–910. https://doi.org/10.1109/JBHI.2021.3079430 PMID: 33979295

7. Schatteburg J, Langenbucher A. Protocol for the diagnosis of keratoconus using convolutional neural networks. PLoS One. 2022;17(2):e0264219. https://doi.org/10.1371/journal.pone.0264219 PMID: 35180279

8. Fassbind B, Langenbucher A, Streich A. Automated cornea diagnosis using deep convolutional neural networks based on cornea topography maps. Sci Rep. 2023;13(1):6566. https://doi.org/10.1038/s41598-023-33793-w PMID: 37085580

9. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2022. p. 11976–86.

10. Grossi E, Buscema M. Introduction to artificial neural networks. Eur J Gastroenterol Hepatol. 2007;19(12):1046–54. https://doi.org/10.1097/MEG.0b013e3282f198a0 PMID: 17998827

11. Li Z, Liu F, Yang W, Peng S, Zhou J. A survey of convolutional neural networks: Analysis, applications, and prospects. IEEE Trans Neural Netw Learning Syst. 2022;33(12):6999–7019. https://doi.org/10.1109/tnnls.2021.3084827

12. Taye MM. Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. Computation. 2023;11(3):52. https://doi.org/10.3390/computation11030052

13. Qi J, Du J, Siniscalchi SM, Ma X, Lee C-H. On mean absolute error for deep neural network based vector-to-vector regression. IEEE Signal Process Lett. 2020;27:1485–9. https://doi.org/10.1109/lsp.2020.3016837

14. Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P, et al. On evaluation metrics for medical applications of artificial intelligence. Sci Rep. 2022;12(1):5979. https://doi.org/10.1038/s41598-022-09954-8 PMID: 35395867

15. Python W. Python releases for windows; 2021. p. 24.

16. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR); 2016. p. 770–8. https://doi.org/10.1109/cvpr.2016.90

17. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4700–8.

18. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning; 2019. p. 6105–14.

19. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A. Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems; 2019. p. 32.

20. Ying X. An overview of overfitting and its solutions. J Phys: Conf Ser. 2019;1168:022022. https://doi.org/10.1088/1742-6596/1168/2/022022

21. Zhou P, Xie X, Lin Z, Yan S. Towards understanding convergence and generalization of AdamW. IEEE Trans Pattern Anal Mach Intell. 2024;46(9):6486–93. https://doi.org/10.1109/TPAMI.2024.3382294 PMID: 38536692