

Detection of Keratoconus through YOLOv8 Region of Interest Preprocessing and Pre-trained Convolutional Neural Networks Using 2D Images

© Muhammed Sideeq Anwar^{1,2}, © Emre Özbilge¹

¹Department of Computer Engineering, Cyprus International University Faculty of Engineering, Nicosia, North Cyprus

²Department of Computer Education, Tishk International University Faculty of Education, Erbil, Iraq

Abstract

BACKGROUND/AIMS: This study presents a methodology for detecting keratoconus using pre-trained convolutional neural network (CNN) models. Five models were evaluated, namely, Xception, InceptionV3, ResNet152, InceptionResNetV2, and EfficientNetV2S.

MATERIALS AND METHODS: Model performance was assessed in two stages: in the initial stage, raw image data were used with a YOLOv8 object detector to extract the region of interest (ROI) around the eyes, and the subsequent stage involved training the pre-trained CNN model through transfer learning to classify the extracted eye region image as normal, mild, or advanced keratoconus.

RESULTS: The results showed that the Xception and InceptionResNetV2 models achieved accuracies of 93.80% and 94.23%, respectively, with ROI cropping, outperforming the other CNN architectures. Without ROI preprocessing, their accuracy decreased to 91.43% and 91.45%, respectively, highlighting the importance of targeted image cropping.

CONCLUSION: Additional metrics corroborated these findings and demonstrated improved diagnostic capabilities when trained using extracted ROI data. This methodology demonstrates the potential of targeted image preprocessing and transfer learning to enhance early detection and management of keratoconus.

Keywords: Keratoconus, CNN, transfer learning, YOLOv8

INTRODUCTION

Keratoconus is a visual impairment characterised by the thinning and protrusion of the cornea into a conical shape. It typically affects both eyes, with one eye often being impacted more severely. The condition usually manifests between late adolescence and early adulthood, progressing over ten years or more. Initially, the visual impairments were corrected using spectacles or soft contact lenses. As the disease progresses, rigid gas-permeable or scleral lenses may become necessary. Severe cases may require corneal transplantations. Corneal collagen

cross-linking aims to halt or decelerate keratoconus progression, potentially averting transplant surgery.¹ Symptoms include blurred or distorted vision, sensitivity to bright light and glare, frequent changes in eyeglass prescriptions, and sudden deterioration of vision or cloudiness. Rapid vision deterioration, particularly irregular astigmatism, warrants prompt ophthalmological evaluation. The precise aetiology remains unknown; however, it is hypothesised to involve genetic and environmental factors. Risk factors include family history, vigorous eye rubbing, and conditions such as retinitis pigmentosa, Down syndrome, and connective tissue disorders.^{1,2}

To cite this article: Anwar MS, Özbilge E. Detection of keratoconus through YOLOv8 ROI preprocessing and pre-trained convolutional neural networks using 2d images. Cyprus J Med Sci. 2025;10(Suppl 1):50-54

ORCID IDs of the authors: M.S.A. 0009-0008-9444-6404; E.Ö. 0000-0002-2295-752X.



Corresponding author: Emre Özbilge
E-mail: eozbilge@ciu.edu.tr
ORCID ID: orcid.org/0000-0002-2295-752X

Received: 03.12.2024
Accepted: 27.01.2025
Publication Date: 04.06.2025



Copyright © 2025 The Author. Published by Galenos Publishing House on behalf of Cyprus Turkish Medical Association.
This is an open access article under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

Complications of keratoconus include rapid corneal swelling (hydrops) leading to sudden vision loss and scarring, as well as progressive corneal scarring requiring transplant surgery, particularly in advanced cases. However, convolutional neural network (CNN) algorithms typically analyse entire ocular images, including nonessential structures, resulting in redundant data processing that may diminish their ability to accurately detect keratoconus-related abnormalities. By isolating the region of interest (ROI), our objective was to facilitate the processing for subsequent algorithms, enhancing their capacity to identify relevant features associated with keratoconus pathology. Traditional methods, such as support vector machines (SVM) and random forests (RF), rely on manually engineered features, which may fail to capture the nuanced patterns of keratoconus progression. To define the ROI, we employed YOLOv8, which is recognised for its high object detection accuracy. This ensures that classification algorithms focus on diagnostic challenges without extraneous elements.^{3,4} This preprocessing step enables algorithms to discern differences in crucial features, potentially achieving a higher accuracy in keratoconus detection. By eliminating redundant information, we reduce the computational burden, resulting in a shorter execution time. We aimed to improve both the accuracy and efficiency of the classification algorithms, benefiting ophthalmology diagnostic practices and patient outcomes.

In summary, this study proposes a two-stage keratoconus detection approach using YOLOv8 for ROI extraction and CNN-based classification to improve the diagnostic performance. This study aimed to develop an advanced methodology for detecting keratoconus, a progressive corneal disorder, using pre-trained CNNs and targeted image preprocessing techniques. By employing a two-stage process, using YOLOv8 for eye region extraction and leveraging transfer learning on extracted regions, this approach highlights the role of ROI cropping in improving diagnostic accuracy. This study demonstrates how integrating ROI-based preprocessing with robust CNN models, such as Xception and InceptionResNetV2, can significantly enhance the detection accuracy and other diagnostic metrics. This innovation has the potential to facilitate earlier detection and better management of keratoconus, improving patient outcomes. The main contributions of this study are as follows:

Significant accuracy improvement through region of interest extraction: Utilizing YOLOv8 for precise eye-region extraction enhances the CNN classifier accuracy.

Enhanced diagnostic metrics: Integrating ROI preprocessing with transfer learning improved recall, precision, and F1-score, emphasising the benefits of focusing on diagnostically relevant regions.

Efficient two-stage detection pipeline: The proposed methodology reduces redundant data processing and computational load while maintaining high performance, enabling more efficient early detection of keratoconus in clinical settings.

The remainder of this paper is organised as follows. We begin with an introduction (section 1) that establishes the context and purpose of the investigation. The literature review (section 2) examines current research in the field, identifying gaps and significant findings that informed our work. The methodology (section 3) explains the study design, data collection, and analytical methods employed. We then present the experimental results (section 4), discuss our findings, and elucidate

their significance. Finally, we reflect on the challenges encountered and propose areas for future research (section 5).

Literature Review

Keratoconus is an eye condition that involves thinning and conical bulging of the cornea, creating challenges for analysis and treatment. With technological advances, machine learning (ML) and CNN have become valuable in clinical photograph evaluation, offering more accurate detection of keratoconus. In this literature review, we studied ML and CNN strategies to identify keratoconus and discussed their benefits, drawbacks, and functionality in scientific settings.

Traditional Machine Learning-Based Approaches

In early keratoconus detection, researchers used conventional devices to develop algorithms, such as SVM, multilayer perceptrons, and RF, focusing on features from corneal topography or tomography images.⁵ Applied an SVM to corneal topography data to differentiate normal eyes from keratoconus-affected eyes, achieving promising accuracy. However, traditional methods such as SVM and RF rely on manually engineered features, which may fail to capture nuanced, non-linear patterns of keratoconus progression.⁶

Deep Learning-Based Approaches

CNN models have revolutionised medical image analysis by automatically extracting features from raw images, thereby eliminating the need for manual feature extraction. They have shown excellent performance in detecting keratoconus, particularly in anterior segment optical coherence tomography (AS-OCT) and corneal tomography images.⁷ Developed a CNN architecture for screening keratoconus using AS-OCT images, achieving high sensitivity and specificity.⁸

Transfer learning has become prevalent in keratoconus detection. This technique involves fine-tuning pre-trained CNN models for specific tasks and leveraging knowledge from large datasets. It enhances model performance and reduces the need for extensive labeled data. Researchers are exploring innovative deep learning architectures for keratoconus detection, such as attention mechanisms and generative adversarial networks (GANs), to improve diagnostic accuracy.^{8,9} Recent advances have demonstrated the effectiveness of combining attention mechanisms with neural networks, including GANs, to enhance corneal image analysis. In¹⁰, a U-Net model with edge and spatial attention mechanisms improved OCT image segmentation accuracy, achieving a Dice score of 94.99%.¹¹ used attention-enhanced CNNs to focus on retinal fundus images for glaucoma classification, highlighting important regions, such as the optic disc.¹² proposed a constrained GAN for medical image enhancement, incorporating structural and illumination constraints with attention mechanisms to enhance corneal images, outperforming traditional methods in NIQE and PIQE metrics. A semi-supervised multi-scale self-transformer-GAN was proposed in¹³, to segment corneal ulcers from slit-lamp images, capture long-range dependencies, and enhance performance in labeled and unlabelled datasets.¹⁴ Combined local extrema information and quantised Haralick features from fundus images, with a long short-term memory network to classify diabetic retinopathy symptoms and capture texture and multiresolution details, while analysing retinal vasculature and hard-exudate patterns with high precision.

MATERIALS AND METHODS

Dataset

Sample input images for keratoconus in the normal, mild, and advanced cases are shown in Figure 1. A total of 1,000 images were obtained and separated into severe, mild, and normal eye condition categories. These images were collected, under a non-disclosure agreement, at the North Eye Center in Iraq with a non-disclosure agreement. Mild and advanced cases accounted for 70% of the total samples, with approximately 35% for each of these classes, while 30% were normal cases. Half of the images were used to train the object detection model in stage 1, and the other half to train the classification model in stage 2. The datasets for each stage were split into training, validation, and test sets at ratios of 70%, 20%, and 10%, respectively. The performance of these stages was evaluated using a test set to evaluate the validity of both models.

To improve the performance of both models, data augmentation, including contrast adjustment, rotation, linear translation, zooming, and flipping, was applied to enhance the model generalization and mitigate overfitting. This augmentation process was applied only to the training datasets before training, and each method was randomly applied to the original images with a certain probability. This process was repeated six times to generate different variations in the original image. The original images were retained without modifying the training dataset.

Keratoconus Diagnosis System

The proposed keratoconus detection method was implemented in two stages (Figure 2). The raw facial image was presented to the trained

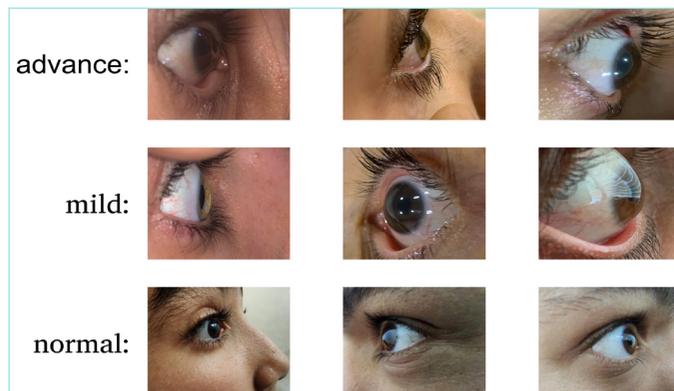


Figure 1. Sample keratoconus image for advance, mild and normal eye conditions.

YOLOv8 model, and the extracted eye region images were fed to the pre-trained ImageNet model for keratoconus stage classification (normal, mild, or advanced). Both models were trained using separate datasets. However, during inferencing, they were combined sequentially, with the YOLOv8 object detector's output becoming the classifier model's input, constituting the complete pipeline for keratoconus identification.

YOLOv8 Object Detection Model

YOLOv8, introduced by Ultralytics,¹⁵ significantly enhances end-to-end object detection. It employs an efficient backbone architecture for feature extraction and integrates a Feature Pyramid Network to improve detection across multiple scales. YOLOv8 adopts an anchor-free approach that directly predicts the center, width, and height of bounding boxes, simplifying detection by eliminating the predefined anchor boxes. This method enhances speed and accuracy. The algorithm processes the entire image in a single pass, consistent with its "you only look once" principle, and applies non-maximum suppression to remove overlapping boxes, retaining only the most confident detections. The YOLOv8 model was trained to detect eye regions in facial images, and the extracted regions were then presented to the CNN classifier for keratoconus stage identification.

Transfer Learning for Classification

Transfer learning has been widely applied in deep learning. The weights of a pre-trained CNN model's weights, learned on the ImageNet challenge dataset, were used for the custom task by replacing the classifier head, which allowed the transferred model to learn a new task. This method is more robust and accurate than a deep learning model trained from scratch. Five pre-trained CNN models (Xception, InceptionV3, ResNet152, InceptionResNetV2, and EfficientNetV2S) were evaluated. Classifier heads were replaced to adapt the models to keratoconus diagnostic tasks. A global average pooling layer flattened the extracted features, followed by two fully connected dense layers (512 and 256 nodes) and one output dense layer with three a softmax activation to learn the keratoconus stages. The models were trained using the Adam optimiser, with a 0.01 learning rate up to 100 epochs.

Experimental Results and Discussions

The performance of the pre-trained models was evaluated with and without ROI extraction using the YOLOv8 object detector. Table 1 presents the results of the pre-trained CNN models when raw facial images were directly inputted into them. Table 2 illustrates the results when the eye region was initially extracted using YOLOv8 and subsequently presented to the classification models. This is evidenced by the consistent improvement in the model performance when

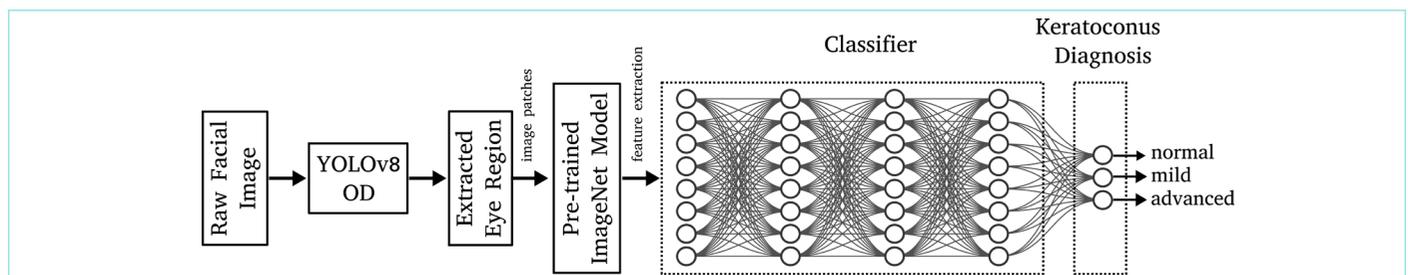


Figure 2. Proposed keratoconus diagnostic system through detecting image patches from the facial image and then classifying input image patches (eye regions) with the stage of keratoconus disease.

Table 1. Comparison results of CNN models without performing YOLOv8 detection.

Model	Accuracy	Recall	Precision	F1-score
Xception	91.43%	89.02%	89.05%	89.04%
InceptionV3	87.98%	80.13%	80.11%	80.11%
Resnet152	85.18%	79.43%	76.22%	77.79%
InceptionResNetV2	91.45%	88.23%	88.08%	88.16%
EfficientNetV2S	90.55%	86.78%	83.79%	86.42%

CNN: Convolutional neural network.

Table 2. Comparison results of CNN models after performing YOLOv8 detection.

Model	Accuracy	Recall	Precision	F1-score
Xception	93.80%	90.98%	89.05%	90.00%
InceptionV3	90.15%	82.99%	81.61%	82.29%
Resnet152	87.51%	81.46%	77.98%	79.68%
InceptionResNetV2	94.23%	89.82%	89.64%	89.73%
EfficientNetV2S	93.11%	87.93%	84.97%	86.42%

CNN: Convolutional neural network.

employing YOLOv8 object detection. With ROI extraction, the Xception model demonstrated a 2.37% improvement in accuracy, rising from 91.43% to 93.80%, whereas InceptionResNetV2 showed the highest accuracy at 94.23%. The recall metric of Xception improved from 89.02% to 90.98%. InceptionV3, which initially demonstrated an accuracy of 87.98% without an object detector, improved to 90.15% after eye-region extraction. These improvements suggest that ROI extraction facilitates focusing on sample characteristics, thereby enhancing its ability to generalize and identify key data patterns.

However, not all the models exhibited the same trend. InceptionResNetV2 demonstrated a modest increase in accuracy from 91.45% to 94.23% and a slight improvement in recall and the F1 scores. This indicates that although YOLOv8 object detection, enhances the model performance, the extent of its effect may depend on the model's inherent architecture. ResNet152 exhibited one of the largest improvements in accuracy, increasing from 85.18% to 87.51%. Overall, extracting the eye region appears to be valuable for enhancing model performance, particularly for increasing recall and precision. The results in Table 2 highlight the efficacy of YOLOv8 preprocessing in improving classification performance. Among the evaluated models, InceptionResNetV2 demonstrated the highest accuracy, whereas ResNet152 demonstrated modest gains.

CONCLUSION

Despite significant advancements, challenges persist in developing and applying ML and CNN for keratoconus detection. Limited annotated datasets, variability in imaging protocols, and generalisability to diverse populations pose constraints. The capacity to train deep learning models and their integration into clinical workflows warrants further investigation. ML and CNN approaches demonstrate considerable potential for enhancing the early diagnosis and management of keratoconus. Continued research on data standardisation, model interpretation, and clinical validation is essential to translate this technology into routine clinical practice, benefiting keratoconus patients globally. In refining keratoconus diagnostic techniques, we encountered the following obstacle: the incorporation of extraneous data into ocular images. Traditional CNN algorithms exhibit limitations

owing to redundant information from nontrivial sources, which affects accuracy and efficiency. To address this, an object detection method that emphasises a ROI extraction from ocular images is proposed. Future integration of explainable AI methods could improve clinical adoption by enabling transparent decision making for keratoconus diagnosis.

MAIN POINTS

- Novel dataset was collected in this study.
- Two stages deep learning based approach was developed to detect the keratoconus eye disease.
- Comparing several deep learning architectures to evaluate the performance of deep learning models.
- Compare the performance of the deep learning models with and without using object detection algorithm.

ETHICS

Ethics Committee Approval: Not available.

Informed Consent: Prior to participation, all individuals were fully informed about the objectives and procedures of the study and voluntarily provided verbal consent to the researchers. Furthermore, ethical approval and authorization for data collection were granted by North Eye Center.

Footnotes

Authorship Contributions

Concept: M.S.A., E.Ö., Design: M.S.A., E.Ö., Data Collection and/or Processing: M.S.A., Analysis and/or Interpretation: M.S.A., Literature Search: M.S.A., E.Ö., Writing: M.S.A., E.Ö.

DISCLOSURES

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study had received no financial support.

REFERENCES

1. Santodomingo-Rubido J, Carracedo G, Suzaki A, Villa-Collar C, Vincent SJ, Wolffsohn JS. Keratoconus: an updated review. *Contact Lens and Anterior Eye*. 2022; 45(3): 101559.
2. Zhang X, Munir SZ, Sami Karim SA, Munir WM. A review of imaging modalities for detecting early keratoconus. *Eye*. 2021; 35(1): 173-87.
3. Reis D, Kupec J, Hong J, Daoudi A. Real-time flying object detection with YOLOv8. *arXiv preprint arXiv:2305.09972*. 2023 May 17.
4. Terven J, Córdova-Esparza DM, Romero-González JA. A comprehensive review of yolo architectures in computer vision: from yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction*. 2023; 5(4): 1680-716.
5. Arbelaez MC, Versaci F, Vestri G, Barboni P, Savini G. Use of a support vector machine for keratoconus and subclinical keratoconus detection by topographic and tomographic data. *Ophthalmology*. 2012; 119(11): 2231-8.
6. Lavric A, Popa V, Takahashi H, Yousefi S. Detecting keratoconus from corneal imaging data using machine learning. *IEEE Access*. 2020;8:149113-21.
7. Michl M, Fabianska M, Seeböck P, Sadeghipour A, Najeeb BH, Bogunovic H, et al. Automated quantification of macular fluid in retinal diseases and their response to anti-VEGF therapy. *British Journal of Ophthalmology*. 2022; 106(1): 113-20.
8. Lavric A, Valentin P. KeratoDetect: keratoconus detection algorithm using convolutional neural networks. *Comput Intell Neurosci*. 2019; 2019(1): 8162567.
9. Al-Timemy AH, Ghaeb NH, Mosa ZM, Escudero J. Deep transfer learning for improved detection of keratoconus using corneal topographic maps. *Cognitive Computation*. 2022; 14(5): 1627-42.
10. Jocher G, Chaurasia A, Qiu J. Ultralytics YOLO (Version 8.0. 0)[Computer software]. Available from: <https://github.com/ultralytics/ultralytics>. 2023.
11. Karn PK, Abdulla WH. Advancing ocular imaging: a hybrid attention mechanism-based u-net model for precise segmentation of sub-retinal layers in OCT images. *Bioengineering*. 2024; 11(3): 240.
12. Cho YS, Song HJ, Han JH, Kim YS. Attention mechanism-based glaucoma classification model using retinal fundus images. *Sensors*. 2024; 24(14): 4684.
13. Ma Y, Liu J, Liu Y, Fu H, Hu Yan, Cheng J. Structure and illumination constrained GAN for medical image enhancement. *Transactions on Medical Imaging*. 2021; 40(12): 3955-67.
14. Wang T, Wang M, Zhu W, Wang L, Chen Z, Peng Y, et al. Semi-MsST-GAN: a semi-supervised segmentation method for corneal ulcer segmentation in slit-lamp images. *Front Neurosci*. 2022; 15: 793377.
15. Ashir AM, Ibrahim S, Abdulghani M, Ibrahim AA, Anwar MS. Diabetic retinopathy detection using local extrema quantized haralick features with long short-term memory network. *Int J Biomed Imaging*. 2021; 2021: 6618666.