

Article

A Deep Feature Fusion of Improved Suspected Keratoconus Detection with Deep Learning

Ali H. Al-Timemy^{1,*}, Laith Alzubaidi^{2,3}, Zahraa M. Mosa⁴, Hazem Abdelmotaal⁵, Nebras H. Ghaeb¹, Alexandru Lavric⁶, Rossen M. Hazarbassanov^{7,8}, Hidenori Takahashi⁹, Yuantong Gu^{2,3} and Siamak Yousefi^{10,11,*}

- ¹ Biomedical Engineering Department, Al-Khwarizmi College of Engineering, University of Baghdad, Baghdad 10011, Iraq
 - ² School of Mechanical, Medical, and Process Engineering, Queensland University of Technology, Brisbane, QLD 4000, Australia
 - ³ ARC Industrial Transformation Training Centre—Joint Biomechanics, Queensland University of Technology, Brisbane, QLD 4000, Australia
 - ⁴ Department of Physics, College of Science, Al-Nahrain University, Baghdad 64021, Iraq
 - ⁵ Department of Ophthalmology, Assiut University, Assiut 71526, Egypt
 - ⁶ Computers, Electronics and Automation Department, Stefan cel Mare University of Suceava, 720229 Suceava, Romania
 - ⁷ Medical School, Universidade Anhembi Morumbi, São Paulo 03101-001, Brazil
 - ⁸ Department of Ophthalmology and Visual Sciences, Paulista Medical School, Federal University of São Paulo, São Paulo 04021-001, Brazil
 - ⁹ Department of Ophthalmology, Jichi Medical University, Tochigi 329-0431, Japan
 - ¹⁰ Department of Ophthalmology, University of Tennessee Health Science Center, Memphis, TN 38163, USA
 - ¹¹ Department of Genetics, Genomics, and Informatics, University of Tennessee Health Science Center, Memphis, TN 38163, USA
- * Correspondence: ali.althimemy@kecbu.uobaghdad.edu.iq (A.H.A.-T.); siamak.yousefi@uthsc.edu (S.Y.)



Citation: Al-Timemy, A.H.; Alzubaidi, L.; Mosa, Z.M.; Abdelmotaal, H.; Ghaeb, N.H.; Lavric, A.; Hazarbassanov, R.M.; Takahashi, H.; Gu, Y.; Yousefi, S. A Deep Feature Fusion of Improved Suspected Keratoconus Detection with Deep Learning. *Diagnostics* **2023**, *13*, 1689. <https://doi.org/10.3390/diagnostics13101689>

Academic Editor: Jae-Ho Han

Received: 1 March 2023

Revised: 6 May 2023

Accepted: 8 May 2023

Published: 10 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Detection of early clinical keratoconus (KCN) is a challenging task, even for expert clinicians. In this study, we propose a deep learning (DL) model to address this challenge. We first used Xception and InceptionResNetV2 DL architectures to extract features from three different corneal maps collected from 1371 eyes examined in an eye clinic in Egypt. We then fused features using Xception and InceptionResNetV2 to detect subclinical forms of KCN more accurately and robustly. We obtained an area under the receiver operating characteristic curves (AUC) of 0.99 and an accuracy range of 97–100% to distinguish normal eyes from eyes with subclinical and established KCN. We further validated the model based on an independent dataset with 213 eyes examined in Iraq and obtained AUCs of 0.91–0.92 and an accuracy range of 88–92%. The proposed model is a step toward improving the detection of clinical and subclinical forms of KCN.

Keywords: convolutional neural networks; keratoconus; feature fusion; transfer learning; deep learning; machine learning

1. Introduction

Keratoconus (KCN) is a non-inflammatory disease that causes thinning and eventual bulging of the cornea affecting one or both eyes [1], and can lead to blindness if not detected and treated at an early stage. KCN detection at an early stage may improve treatment outcomes and prevent development of more advanced stages. Clinical diagnosis of KCN involves the subjective evaluation of the patient's history, demographic characteristics, and corneal topographic maps [2,3]. While diagnosing established KCN cases can be more straightforward, detecting subclinical forms of KCN remains a challenge.

Machine learning (ML) algorithms [4] have been used as a basis for creating several different automated models, including support vector machines (SVM) [5], decision trees [6],

and neural networks [7], which hold promise for the early detection and identification of subclinical forms of KCN, as well as established KCN. These methods are typically combined with hand crafted features [8] or machine generated features from corneal topography. In contrast, deep learning (DL) methods provide a way of both performing feature extraction and classification of images; thus, reducing the need for hand crafted features. DL methods [9–12] have previously been applied to corneal topography [9,13–17] and optical coherence tomography (OCT) imaging [18] for KCN detection.

One such DL model [15] was developed for KCN detection in patients with a history of refractive surgery from Orbscan corneal maps. A high accuracy of the model was obtained for distinguishing normal from established KCN; however, their model did not include the analysis of suspected KCN. Using a different approach, Al-Timemy et al. [2] developed an ensemble of DL of four corneal maps and Pentacam indices (PI) classifiers for established KCN detection versus normal cases using Pentacam corneal topographic maps. The developed ensemble achieved an accuracy of 95% based on AlexNet. The accuracy was improved to 98.3%, with ensemble-specific combinations of corneal maps' classifiers and PI. In a third approach, Chen et al. [19] utilized a VGG16 convolutional neural network (CNN) model to detect established KCN versus healthy eyes. Their method was applied to the four topographic maps and achieved an accuracy of 97.85% on the testing set when concatenating all maps. However, no suspect KCN group was included. Abdelmotaal et al. [14] also developed a CNN architecture to discriminate suspect, established KCN, and normal eyes. They obtained an accuracy of 98%, for discriminating normal, suspect KCN, and KCN, based on analyzing four corneal maps.

While the previous studies utilized Pentacam maps, other attempts have been developed on synthetic data. In one such study [20], the authors developed a deep learning model based on CNN synthesized corneal maps and achieved a high accuracy, but the images were synthetic and no suspected KCN was included in their dataset.

Detecting subclinical KCN is challenging for expert clinicians, as there may be no clear manifestation of signs. This problem is even more challenging for less experienced corneal experts. Hence, DL models may play an important role in detecting suspect (subclinical) KCN [21].

In this study, our aim was to improve both the accuracy and robustness of detecting subclinical forms of KCN, based on an ensemble of DL features, which is an emerging approach in which the features are extracted from a pretrained DL network and fed into a machine learning classifier, such as SVM, to make the final classification. This approach has previously been applied to detect two corneal classes (normal and KCN) with four corneal topographic maps and Alexnet with SVM [22], and has also been examined in three-class detection (normal, KCN, and suspect KCN) with seven corneal topographic maps and EfficientNet-b0 architecture and SVM [23].

In order to harvest the power of both DL features and ML classifier, we propose a novel fusion DL method based on Xception [24] and InceptionResNetV2 [25] to achieve a three-class clinical suspect KCN detection. Two datasets of 4752 images were used in this study. A dataset of 4113 images, collected in Egypt, were utilized to develop the fusion-based model, and validated with an independent test set of 639 images acquired in Iraq.

2. Materials and Methods

2.1. Datasets

Two datasets were used in this study: the first dataset was acquired with the Pentacam instrument in Egypt and the second dataset was collected with the Pentacam instrument in Iraq. Both studies were conducted according to the ethical codes of the declaration of Helsinki.

The first dataset included images from normal, suspect, and established keratoconus, which were also utilized in Abdelmotaal et al. [14]. It included refractive corneal maps, as well as sagittal (Sag), corneal pachymetry (CorPac) and elevation front (ElvF) maps,

acquired from the Pentacam instruments (Oculus GmbH, Wetzlar, Germany). This original study was a retrospective review of the Pentacam four-map selectable display images of non-consecutive refractive surgery candidates, patients with unilateral or bilateral keratoconus, and patients with subclinical keratoconus. All study participants provided written consent. Two experienced corneal specialists independently classified the anonymized images as keratoconus, subclinical keratoconus, or normal. The keratoconus class (KCN) included those with a clinical diagnosis of keratoconus (e.g., the presence of a central protrusion of the cornea with Fleischer ring, Vogt striae, or both, as determined by slit-lamp examination) or an irregular cornea (as determined by distorted keratometry mires or distortion of retinoscopic red reflex, or both). The keratoconus class also included the subsequent topographic findings, as summarized by Piñero and colleagues: a focal steepening located in a zone of protrusion, surrounded by concentrically decreasing power zones; focal areas with dioptric (D) values > 47.0 D; inferior–superior (I–S) asymmetry measured as >1.4 D; or angling of the hemimeridians in an asymmetric or broken bowtie pattern, with skewing of the steepest radial axis. The suspect keratoconus class (suspect) included subtle corneal topographic changes within the aforementioned keratoconus abnormalities, with the absence of slit-lamp or visual changes typical of keratoconus (former fruste, or asymptomatic, keratoconus). The traditional class (normal) included refractive surgery candidates and subjects applying for contact lens fitting with a refractive error of 8.0 D sphere, with 3.0 D of astigmatism, and without clinical, topographic, or tomographic signs of keratoconus or subclinical keratoconus. After classification, the labeled images were then reviewed by a 3rd party corneal specialist who identified images with conflicting labels and adjudicated their classes by consensus. The two raters' first-round group labels were withheld during adjudication. The dataset is comprised of 4113 images from 1371 eyes of the: normal = 500, suspect KCN = 500, and KCN = 371, and each has 3 corneal maps.

The second dataset included corneal maps from normal and KCN subjects. It was originally collected at Al-Amal eye Clinical in Baghdad Iraq, and was used in previous study as well [2]. The diagnosis of the eyes was performed by 2 experts, and patients with other ocular diseases were excluded from the study. The dataset was composed of the Pentacam Scheimpflug measurements, acquired from 213 eyes (114 normal and 99 KCN). Each eye had the 3 refractive corneal maps described in [26], which made the total images 639 images (213×3 maps).

2.2. Image Dataset Preprocessing Pipeline

For the 3-class dataset, all maps have black numbers and annotations generated by the measurement equipment. Figure 1 shows examples of the 3 corneal maps for the 3 classes investigated in this study, i.e., normal, suspect, and KCN. All 4113 images were denoised to remove the annotations. For the independent test set, the images were clean. All images were subsequently cropped from the standard refractive maps image acquired from the Pentacam. All images from the 3-class and independent 2-class datasets ($4113 + 639 = 4752$) were resized to match the input of the 2 DL networks used in this study (Xception and InceptionResNetV2), i.e., $299 \times 299 \times 3$.

2.3. Deep Transfer Learning

Transfer learning is a technique to learn from a large dataset and then transfer the knowledge to a small dataset [27]. One of the most popular sources of transfer learning is the ImageNet [28] dataset. This dataset has 1000 classes of natural images, including various animals, flowers, objects, etc., with millions of images. Several successful applications used the ImageNet dataset as a source of transfer learning to address the issue of small datasets, including medical applications. ImageNet dataset can be useless if it does not have features relevant to the target dataset, in our case, KCN detection with corneal topography. For instance, the ImageNet dataset is a color dataset that cannot be helpful for grayscale medical images such as X-Ray, CT, MRI, etc. There is a clear mismatch between the learning features. However, our task has features relevant to the ImageNet dataset. Therefore, in

this paper, two pre-trained models (Xception [24] and InceptionResNetV2 [25]) were used for this task. These two models were chosen based on their high performance with the ImageNet dataset. Both of them have an input size of $299 \times 299 \times 3$. The workflow is described in Figure 2. To begin, the pre-trained model was loaded. Then, the last layers were replaced by new layers to learn the new features of the target dataset. The first layers learn the low-level features, such as color and edges, while the last layers typically learn the task-specific features. Therefore, the last layers are typically replaced by appropriate domain-related layers to learn the related task, which were normal, suspect, and KCN for this study, rather than 1000 classes of the initial ImageNet dataset. Next, the fine-tuned model was trained on a small dataset and tested.

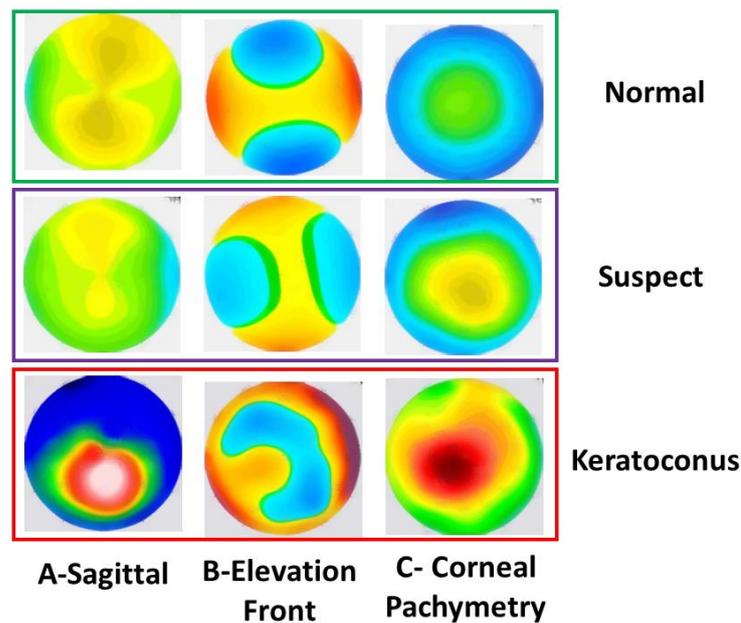


Figure 1. Samples of the input maps for 3 classes (normal, suspect, and keratoconus). (A) sagittal map; (B) elevation front map; and (C) corneal pachymetry map.

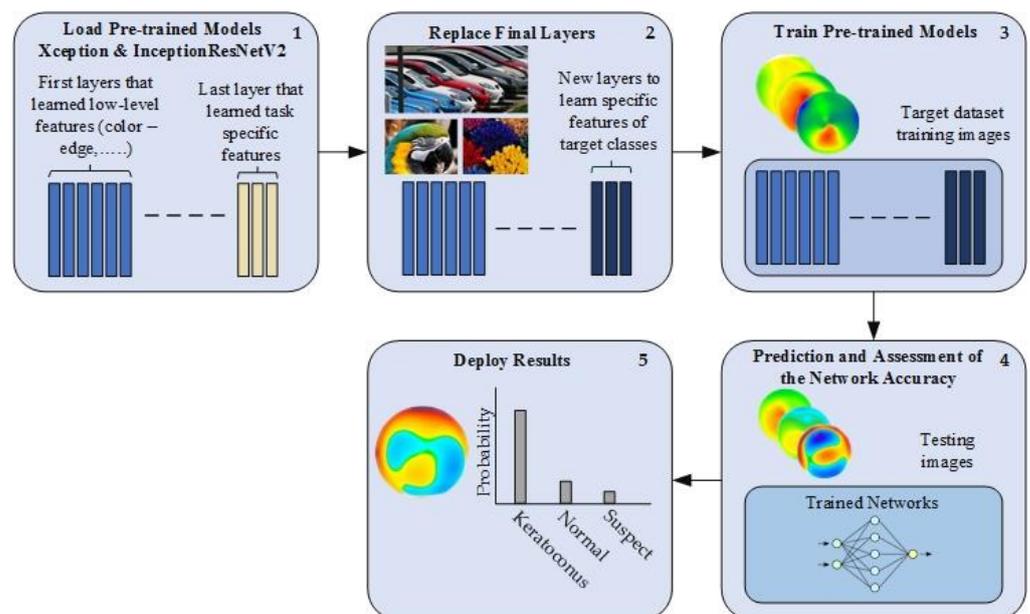


Figure 2. Transfer learning workflow.

2.4. Feature Fusion and ML Classifiers

Machine learning (ML) classifiers' performance may be improved using descriptive features. Therefore, in this paper, the feature fusion technique [29] has been employed to improve the results of the individual models (Figure 3). This technique provides a compact representation (fusing) of multiple features extracted from different sources, and thus may improve the performance.

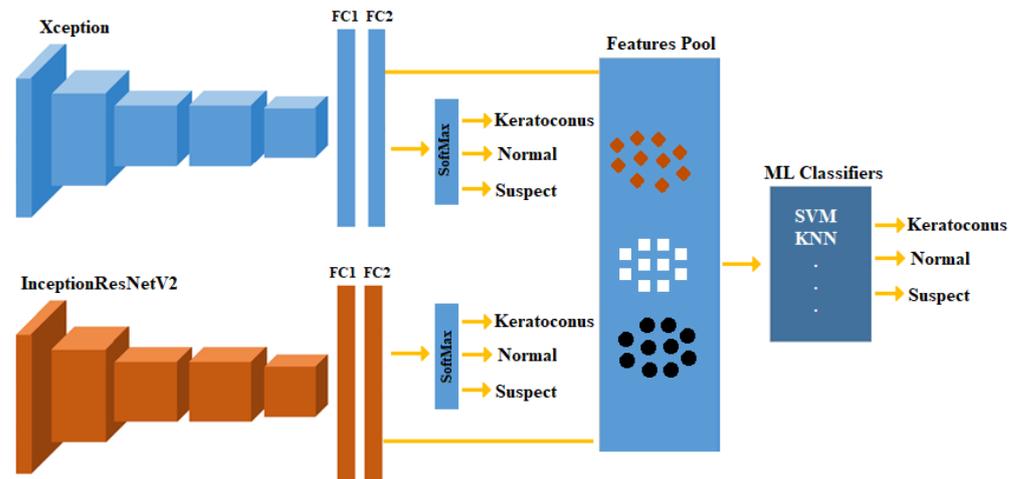


Figure 3. Block diagram of the proposed fusion model workflow.

In this scenario, we trained each model (Xception and InceptionResNetV2 here) separately, then used those models to extract features from the new training and testing subsets. The extracted features from models are fused in one pool to train conventional ML classifiers. Several ML classifiers were trained and tested, including decision trees (DT), quadratic discriminant analysis, SVMs, and K-nearest neighbor (KNN). Using the classification learner application on Matlab, the best performing ML model was selected and the results were reported.

2.5. Training the DL Models

All used datasets were divided into three sets: training, validation, and testing. The percentage of split was equal to 80:10:10 (training, validation, and testing). The used datasets were trained in two different scenarios as follows:

- A. Scenario 1 (S1—2-class): training the models with two classes of normal and KCN.
- B. Scenario 2 (S2—3-class): training the models with three classes of normal, suspect, and KCN.

To train the networks in this study, the following settings were used. The training options included an Adam optimizer, a mini-batch size of 15, max epochs set at 100, shuffled for every epoch, and the initial learning rate was 0.001. The processor properties used in this experiment were Intel (R) Core i7/32 GB/1 TB/Nvidia RTX A3000 12 GB. Matlab 2022a was used to develop the experiments.

2.6. Performance Evaluation

Both the pre-trained models and the ML classifiers were evaluated based on accuracy, specificity, recall, precision, and F1-score metrics. These evaluation metrics were computed based on the TN, TP, FN, and FP values. The TN and TP implied precisely categorized negative and positive instances, whereas FN and FP denoted misclassified positive and negative cases, respectively. Every evaluation metric equation is presented as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (1)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

$$\text{F1}_{\text{score}} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

In order to examine the network activations, Grad-Cams [30] were also plotted to see if the network was looking at areas of clinical importance, which is usually examined by the ophthalmologists. In addition, the learnable filters at different layers of the developed network were examined. The results were organized based on the three maps: (1) Sag, (2) ElvF, and (3) CorPac, with a test set and both scenarios, S1 (2 class, normal vs. KCN) and S2 (3 class, normal vs. suspect vs. KCN).

3. Results

The first dataset included 4113 corneal maps from 1371 eyes. A total of 500 eyes were normal, 500 suspect, and 371 KCN. The mean age of the participants was (normal: 36.50 ± 9.50 years, suspect: 31.80 ± 8.30 years and KCN = 31.5 ± 8.20 years, mean \pm Standard Deviation, SD).

The second dataset included 639 corneal maps from 213 eyes. A total of 114 eyes were normal, and 99 KCN. The mean age of all participants was 31.4 ± 9.2 years for the normal eyes and 32.95 ± 10.86 years for KCN, mean \pm SD).

3.1. Scenario 1: 2-Class Problem (Normal and KCN)

Once we trained, tuned, and tested the model on 80/10/10 (training, validation, and testing) split of the dataset from Egypt, the Xception model generated an accuracy of 99.0% (Table 1). Only two eyes with KCN were misclassified as normal based on the Xception model (Figure 4). Once we extracted the feature using the Xception model and trained an SVM classifier, we obtained a perfect classification with 100% accuracy. As training and testing datasets were selected from the same dataset, we expected the accuracy to be high. As for the ElvF and CorPac maps, CorPac achieved the best performance with 100% accuracy with SVM.

Table 1. Performance metrics of the Xception model applied on Sag, ElvF, and CorPac maps based on scenario S1 (2 class), without and with SVM.

Corneal Map	DL Model	Accuracy	Specificity	Sensitivity	Precision	F1-Score
Sag	Xception	99.0	98.0	100	98.2	99.12
	Xception + SVM	100	100	100	100	100
ElvF	Xception	95.9	92.0	99.1	93.7	96.3
	Xception + SVM	95.9	92.1	99.1	93.7	96.4
CorPac	Xception	100	100	100	100	100
	Xception + SVM	100	100	100	100	100

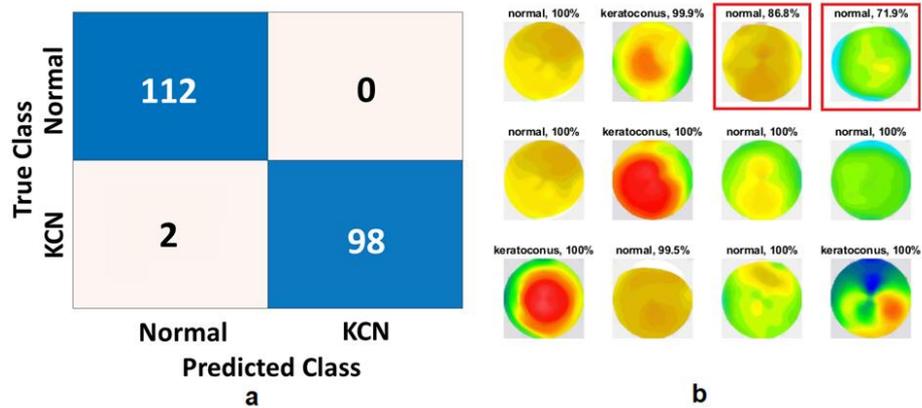


Figure 4. Xception model results for the Sag map with S1 (2 class). (a) Confusion matrix, (b) prediction of test set, where the red boxes are misclassified by the model.

3.2. Scenario 2: 3-Class Problem (Normal, Suspect, and KCN)

The DL model was trained, tuned, and tested based on the 80/10/10 split of the Egypt dataset, and the Xception model generated an accuracy of 99.7% (Table 2). The InceptionResNetV2 achieved a lower accuracy of 97.0%. Like the 2-class problem, the model with feature extraction and fusing using a decision tree obtained a perfect classification accuracy of 100%. The confusion matrix for feature fusion based on the 3-class problem (S2) is shown in Figure 5

Table 2. Results of the Sag map on the S2 (3 class) scenario, with test set with a decision tree classifier.

DL Model	Accuracy	Specificity	Recall	Precision	F1-Score
Xception	99.7	99.5	100	99.1	99.6
InceptionResNetV2	97.0	95.3	100	92.4	96.0
Feature Fusion + Decision Tree	100	100	100	100	100

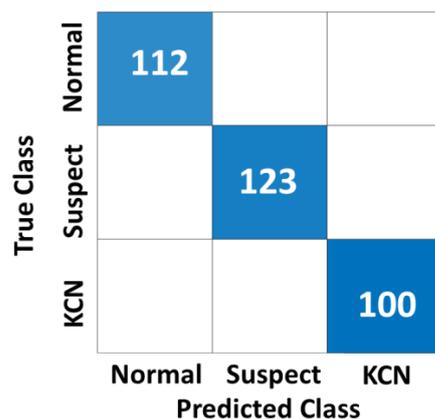


Figure 5. Confusion matrix for feature fusion with a decision tree for the Sag map with S2 (3 class), with an AUC of 0.99.

The ElvF map will be considered in this section. Scenario 2 was first evaluated by training testing both models (Xception and InceptionResNetV2) as listed in Table 3. The Xception model achieved higher results than the InceptionResNetV2 by obtaining an accuracy of 98.5%, specificity of 98.6%, recall of 98.3%, precision of 97.5%, and F1 score of 97.9%. There are eight samples that were misclassified by the Xception model. The InceptionResNetV2 achieved an accuracy of 97.0%, specificity of 95.9%, recall of 99.1%, precision of 92.9%, and F1-score of 95.9%. There were 12 samples misclassified by the InceptionResNetV2 model.

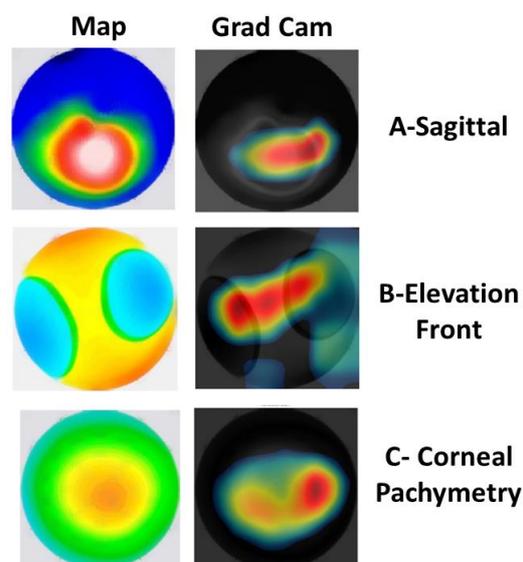
Table 3. Results of the ElvF map on S2 (3 class) scenario with the test set.

DL Model	Accuracy	Specificity	Recall	Precision	F1-Score
Xception	98.5	98.6	98.3	97.5	97.9
InceptionResNetV2	97.0	95.9	99.1	92.9	95.9
Feature Fusion+ SVM	99.1	99.0	99.1	98.3	98.7

In order to overcome the misclassified samples, features that were extracted by both models have been combined to train several ML classifiers. The highest results were achieved by two classifiers: SVM and K-nearest neighbor. SVM was chosen to be listed as shown in Table 3. SVM achieved accuracy of 99.1%, specificity of 99.0%, recall of 99.1%, precision of 98.3%, and F1 score of 98.7%.

Although the feature fusion with SVM improved the results compared to both models, four samples were misclassified, as shown in Figure 6. Lastly, the visualization of the Xception model's focus to make decisions is shown in Figure 7B. From the obtained results, we can conclude that the model identified the appropriate place on the corneal map on which to base the decision, which validates the high achieved results.

True Class	Normal	120	1	
	Suspect	1	119	
	KCN		2	97
		Normal	Suspect	KCN
		Predicted Class		

Figure 6. Confusion matrix of SVM for the ElvF map with the proposed DL Fusion model.**Figure 7.** Original and class activation maps (CAMs), corresponding to different corneal maps with the Xception model for (A) the Sag map with KCN, (B) the ElvF map with a normal eye and (C) the CorPac map with KCN. The heatmaps show where the model focuses to make the decision.

3.3. Visualization of the CAMs

In order to validate the clinical relevance of the models, class activation maps (CAMs) were generated to show the regions of the corneal map that were more important for the model to make decision. Figure 7 shows the CAMs generated based on three different corneal maps including Sag, ElvF, and CorPac maps. Different CAMs suggested that the model considers clinically relevant regions of the cornea for making KCN decisions. This suggests that the model identifies the relevant spot upon which to make a decision, which, in turn, results in a correct prediction.

3.4. Features Extracted by the Xception Model

To further understand how feature extraction, based on the Xception model, impacts the final KCN decision, and to assess the clinical relevance of the features, we visualized the filters that the Xception model learned from input corneal maps. Figure 8 shows the sets of filters that were learned from ElvF map. These features were essentially learned from the first convolutional layers. The shape of filters highlights the clinical relevance of the features.

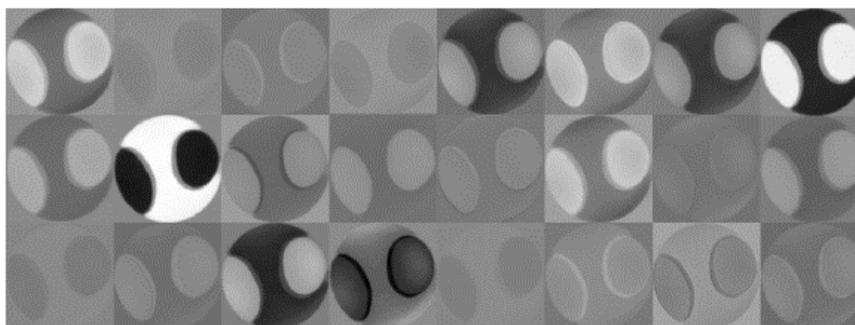


Figure 8. Learned filters by Xception model for the ElvF map.

3.5. Validating the Models Based on an Independent Dataset from Iraq

To assess how the model is generalizable to unseen corneal maps from different settings and ethnicities, we re-evaluated the performance of the model based on an independent dataset. This dataset was collected from an eye clinic in Baghdad, Iraq, which represents corneal data from different settings with likely different settings. Table 4 shows the accuracy metrics of the 2-class problem based on three different corneal maps.

Table 4. The results of all maps for the independent test set from Iraq, with a feature fusion of Xception and InceptionResNetV2.

Corneal Map	Accuracy	Specificity	Sensitivity	Precision	F1-Score	AUC
Sagittal	89.2	90.3	87.8	88.7	88.3	0.91
Elevation Front	91.7	90.6	92.9	89.3	91.0	0.96
Pachymetry	87.7	84.3	92.9	79.8	85.8	0.92

Figures 9–11 show the confusion matrices of the fused CNN model to discriminate normal from KCN based on three different maps: Sag, ElvF, and CorPac, respectively. The model based on corneal ElvF map generated the highest AUC of 0.96, with an accuracy of 91.7%.

True Class	Normal	103	12
	KCN	11	87
		Normal	KCN
		Predicted Class	

Figure 9. Confusion matrix of the independent test set for the *Sag* map, where the AUC is equal to 0.91.

True Class	Normal	107	7
	KCN	11	92
		Normal	KCN
		Predicted Class	

Figure 10. Confusion matrix of the independent test set for the *ElvF* map, where the AUC is equal to 0.96.

True Class	Normal	108	6
	KCN	20	79
		Normal	KCN
		Predicted Class	

Figure 11. Confusion matrix of the independent test set for the *CorPac* map, where the AUC is equal to 0.92.

4. Discussion

We developed several conventional CNN models, based on state-of-the-art architectures, including Xception and InceptionResNetV2, which were able to detect corneas with suspect or established KCN. We observed that the accuracy of these CNN architectures to identify normal eyes from eyes with established KCN based on different corneal maps, reached around 99%. A similar efficacy was observed in the same CNN models for distinguishing normal eyes, eyes with suspect KCN, and eyes with established KCN. While these results were highly promising, the testing subset was selected from the same dataset as the model training.

To assess the generalizability, we evaluated the accuracy of the CNN models based on another independent dataset collected from a Pentacam instrument in another country, with likely different device settings. For the S1 (two-class) problem, we achieved an AUC and accuracy of 0.99 and 100, respectively, with the CorPac map. For the S2 (three-class) problem, we obtained an accuracy of 100 with the Sag map. Our findings suggest that our CNN models are generalizable to unseen corneal maps from other instruments from other countries with likely different settings, and corneal maps from different populations.

With our DL fusion models (Xception and InceptionResNetV2), different classifiers were utilized including DT, quadratic discriminant analysis, KNN, and SVM. SVM and DT

were the best performing classifiers compared to the other ML models, with an accuracy of 99.1 and 100, respectively.

Examining Table 1, it can be noted that the CorPac map generated the highest performance. For testing with the independent test dataset, the ElvF map was the best map, with an AUC of 0.96 and accuracy of 91.7% (Table 4).

Based on the results in Tables 3 and 4, the InceptionResNetV2 model has lower performance than that of Xception model. In the case of InceptionResNetV2, the low number of parameters could be due to a few reasons. One possibility is that the model was fine-tuned or pruned to reduce its size and complexity. Another possibility is that the dataset used to train the model was relatively small or simple, and therefore, a smaller model was sufficient to achieve good performance as in the case of Xception model.

A comparison of our model with other models previously published is shown in Table 5. It should be noted that suspect KCN was not included in many previous studies [2,13,15,19,22]. To tackle that, we developed and validated our proposed method on the three-class problem, including suspect KCN. For example, the model previously introduced by Al-Timemy et al. [23] includes an ensemble of deep transfer learning to combine the four topographic maps with Pentacam indices, whereas our proposed fusion model dealt with each topographic map individually, and achieved good performance for each map.

Table 5. The previous literature on KCN detection with deep learning.

Study	Imaging Device	Number of Classes	Number of Maps	Network Used	Comments	Accuracy
Zéboulon et al. [15]	Orbscan	2	3000	CNN	No suspect	98.3%
Al-Timemy et al. [2]	Pentacam	2	534	Ensemble of Alexnet	No suspect	95–98.3%
Chen et al. [19]	Pentacam	2	1926	VGG16	No suspect	97.85%
Lavric and Valentin [20]	SyntEyes and SyntEyes KTC models/1 map		3000	KeratoDetect	No suspect synthetic maps	99.3%
				VGG16		93.1%
Kuo et al. [13]	Tomy TMS-4 Topographer	2	354	InceptionV3	No suspect	93.1%
				ResNet152		95.8%
Firat et al. [22]	Pentacam	2	628	ALexNet and SVM	No suspect	98.53%
This study	Pentacam	3	4113	Fusion of Xception and InceptionResNetV2	-	97–100%

As shown in the learnable filters displayed in Figure 8, the (Xception) model learned very strong features from the ElvF map. Additionally, the results from the GradCam, shown in Figure 7B, illustrate that the model is focusing on an area at the center, which is the clinically relevant region. The accuracy obtained with the Xception model was equal to 98.5%. This may explain the good performance achieved with the Xception model.

While we used relatively large datasets of corneal maps from two datasets collected from two different countries, additional datasets with larger number of corneal maps from other ethnicities/races are desirable to further validate the proposed CNN models. Additionally, our CNN models are limited to corneal data from Pentacam. Therefore, the use of innovative device-agnostic CNN models is warranted to exploit data from other corneal topography instruments, such as CASIA, Orbscan, and Galilei.

5. Conclusions

DL fusion-based methods were proposed and developed for KCN detection. While single CNN architecture achieves good performance, the fusion of two DL models (Xception and InceptionResNetV2) with features from these models and an SVM classifier is particularly suited to meet the challenge of identification of earlier stage, clinical suspect KCN, distinct from established KCN. We validated our proposed method on a three-class dataset and achieved an accuracy of 97–100%. Furthermore, independent evaluation of a two-class dataset from a different country showed an accuracy of 88–92% and an AUC 0.91–0.96 for the three corneal maps.

Author Contributions: Conceptualization, A.H.A.-T., L.A. and S.Y.; methodology, A.H.A.-T. and L.A.; software, A.H.A.-T. and L.A.; validation A.H.A.-T. and L.A.; formal analysis, A.H.A.-T., Z.M.M. and L.A.; investigation, A.H.A.-T. and L.A.; resources, A.H.A.-T., L.A., N.H.G., H.A. and S.Y.; data curation, A.H.A.-T., L.A., N.H.G. and H.A.; writing—original draft preparation, A.H.A.-T., L.A., Z.M.M., N.H.G., A.L., R.M.H., H.T. and S.Y.; writing—review and editing, A.H.A.-T., L.A., Z.M.M., N.H.G., A.L., R.M.H., H.T. and S.Y.; visualization, A.H.A.-T., L.A. and Z.M.M.; supervision, S.Y.; project administration, A.H.A.-T. and S.Y.; funding acquisition, S.Y. and Y.G. All authors have read and agreed to the published version of the manuscript.

Funding: Ali H. Al-Timemy (Mentee) and Siamak Yousefi (Mentor) received the Roche Collaborative Fellowship Program Award from Association for Research in Vision and Ophthalmology (ARVO) Foundation in 2022.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and its later amendments.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The image dataset and the trained models and Matlab codes are made available from the link: https://drive.google.com/drive/folders/1RiiAsDTmpNTqcCktuZPRUmh2iEIB_ACs?usp=sharing (accessed on 28 February 2023).

Acknowledgments: Ali H. Al-Timemy acknowledges discretionary membership at Plymouth University. Yuantong Gu and Laith Alzubaidi acknowledge the support received through the following funding schemes of Australian Government: ARC Industrial Transformation Training Centre (ITTC) for Joint Biomechanics under grant IC190100020.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yousefi, S.; Yousefi, E.; Takahashi, H.; Hayashi, T.; Tampo, H.; Inoda, S.; Arai, Y.; Asbell, P. Keratoconus severity identification using unsupervised machine learning. *PLoS ONE* **2018**, *13*, e0205998. [[CrossRef](#)] [[PubMed](#)]
2. Al-Timemy, A.H.; Ghaeb, N.H.; Mosa, Z.M.; Escudero, J. Deep Transfer Learning for Improved Detection of Keratoconus using Corneal Topographic Maps. *Cognit. Comput.* **2021**, *14*, 1627–1642. [[CrossRef](#)]
3. Al-Timemy, A.; Al-Zubaidi, L.; Ghaeb, N.; Takahashi, H.; Lavric, A.; Mosa, Z.; Hazarbassanov, R.M.; Alyasseri, Z.A.A.; Yousefi, S. A device-agnostic deep learning model for detecting keratoconus based on anterior elevation corneal maps. *Investig. Ophthalmol. Vis. Sci.* **2022**, *63*, 2101-F0090.
4. Alyasseri, Z.A.A.; Al-Timemy, A.H.; Abasi, A.K.; Lavric, A.; Mohammed, H.J.; Takahashi, H.; Milhomens Filho, J.A.; Campos, M.; Hazarbassanov, R.M.; Yousefi, S. A Hybrid Artificial Intelligence Model for Detecting Keratoconus. *Appl. Sci.* **2022**, *12*, 12979. [[CrossRef](#)]
5. Lavric, A.; Popa, V.; Takahashi, H.; Yousefi, S. Detecting Keratoconus From Corneal Imaging Data Using Machine Learning. *IEEE Access* **2020**, *8*, 149113–149121. [[CrossRef](#)]
6. Twa, M.D.; Parthasarathy, S.; Roberts, C.; Mahmoud, A.M.; Raasch, T.W.; Bullimore, M.A. Automated decision tree classification of corneal shape. *Optom. Vis. Sci. Off. Publ. Am. Acad. Optom.* **2005**, *82*, 1038. [[CrossRef](#)] [[PubMed](#)]
7. Maeda, N.; Klyce, S.D.; Smolek, M.K. Neural Network Classification of Corneal Topography Preliminary Demonstration. *Investig. Ophthalmol. Vis. Sci.* **1995**, *36*, 1327–1335.
8. Mosa, Z.M.; Ghaeb, N.H.; Ali, A.H. Detecting Keratoconus by Using SVM and Decision Tree Classifiers with the Aid of Image Processing. *Baghdad Sci. J.* **2019**, *16*, 4.
9. Kamiya, K.; Ayatsuka, Y.; Kato, Y.; Shoji, N.; Mori, Y.; Miyata, K. Diagnosability of keratoconus using deep learning with Placido disk-based corneal topography. *Front. Med.* **2021**, *8*, 724902. [[CrossRef](#)]
10. Reddy, J.C.; Bhamidipati, P.; Dwivedi, S.; Dhara, K.K.; Joshi, V.; Hasnat Ali, M.; Vaddavalli, P.K. KEDOP: Keratoconus early detection of progression using tomography images. *Eur. J. Ophthalmol.* **2022**, *32*, 2554–2564. [[CrossRef](#)]
11. Alzubaidi, L.; Bai, J.; Al-Sabaawi, A.; Santamaría, J.; Albahri, A.S.; Al-dabbagh, B.S.N.; Fadhel, M.A.; Manoufali, M.; Zhang, J.; Al-Timemy, A.H. A survey on deep learning tools dealing with data scarcity: Definitions, challenges, solutions, tips, and applications. *J. Big Data* **2023**, *10*, 1–82.
12. Albahri, A.S.; Duhaim, A.M.; Fadhel, M.A.; Alnoor, A.; Baqer, N.S.; Alzubaidi, L.; Albahri, O.S.; Alamoodi, A.H.; Bai, J.; Salhi, A. A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Inf. Fusion* **2023**, *96*, 156–191. [[CrossRef](#)]
13. Kuo, B.-I.; Chang, W.-Y.; Liao, T.-S.; Liu, F.-Y.; Liu, H.-Y.; Chu, H.-S.; Chen, W.-L.; Hu, F.-R.; Yen, J.-Y.; Wang, I.-J. Keratoconus screening based on deep learning approach of corneal topography. *Transl. Vis. Sci. Technol.* **2020**, *9*, 53. [[CrossRef](#)] [[PubMed](#)]
14. Abdelmotaal, H.; Mostafa, M.M.; Mostafa, A.N.R.; Mohamed, A.A.; Abdelazeem, K. Classification of color-coded Scheimpflug camera corneal tomography images using deep learning. *Transl. Vis. Sci. Technol.* **2020**, *9*, 30. [[CrossRef](#)] [[PubMed](#)]

15. Zéboulon, P.; Debellemanière, G.; Bouvet, M.; Gatinel, D. Corneal topography raw data classification using a convolutional neural network. *Am. J. Ophthalmol.* **2020**, *219*, 33–39. [[CrossRef](#)] [[PubMed](#)]
16. Feng, R.; Xu, Z.; Zheng, X.; Hu, H.; Jin, X.; Chen, D.Z.; Yao, K.; Wu, J. 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. Heal. Informatics* **2021**, *25*, 3898–3910. [[CrossRef](#)]
17. Fassbind, B.; Langenbucher, A.; Streich, A. Automated cornea diagnosis using deep convolutional neural networks based on cornea topography maps. *Sci. Rep.* **2023**, *13*, 6566. [[CrossRef](#)]
18. Kamiya, K.; Ayatsuka, Y.; Kato, Y.; Shoji, N.; Miyai, T.; Ishii, H.; Mori, Y.; Miyata, K. Prediction of keratoconus progression using deep learning of anterior segment optical coherence tomography maps. *Ann. Transl. Med.* **2021**, *9*, 1287. [[CrossRef](#)]
19. Chen, X.; Zhao, J.; Iselin, K.C.; Borroni, D.; Romano, D.; Gokul, A.; McGhee, C.N.J.; Zhao, Y.; Sedaghat, M.-R.; Momeni-Moghaddam, H. Keratoconus detection of changes using deep learning of colour-coded maps. *BMJ Open Ophthalmol.* **2021**, *6*, e000824. [[CrossRef](#)]
20. Lavric, A.; Valentin, P. KeratoDetect: Keratoconus Detection Algorithm Using Convolutional Neural Networks. *Comput. Intell. Neurosci.* **2019**, *2019*, 8162567. [[CrossRef](#)]
21. Maile, H.; Li, J.-P.O.; Gore, D.; Leucci, M.; Mulholland, P.; Hau, S.; Szabo, A.; Moghul, I.; Balaskas, K.; Fujinami, K. Machine Learning Algorithms to Detect Subclinical Keratoconus: Systematic Review. *JMIR Med. Inform.* **2021**, *9*, e27363. [[CrossRef](#)] [[PubMed](#)]
22. Firat, M.; Çankaya, C.; Çınar, A.; Tuncer, T. Automatic detection of keratoconus on Pentacam images using feature selection based on deep learning. *Int. J. Imaging Syst. Technol.* **2022**, *32*, 1548–1560. [[CrossRef](#)]
23. Al-Timemy, A.H.; Mosa, Z.M.; Alyasseri, Z.; Lavric, A.; Lui, M.M.; Hazarbassanov, R.M.; Yousefi, S. A Hybrid Deep Learning Construct for Detecting Keratoconus From Corneal Maps. *Transl. Vis. Sci. Technol.* **2021**, *10*, 16. [[CrossRef](#)] [[PubMed](#)]
24. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
25. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
26. Sinjab, M.M. *Step by Step®: Reading Pentacam Topography: Basics and Case Study Series*; Jaypee Brothers Medical Publishers: New Delhi, India, 2015; ISBN 9351523977.
27. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaria, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 1–74. [[CrossRef](#)]
28. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]
29. Mungoli, N. Adaptive Ensemble Learning: Boosting Model Performance through Intelligent Feature Fusion in Deep Neural Networks. *arXiv* **2023**, arXiv:2304.02653.
30. Dhinakaran, P.; Shivanandappa, M.G.; Devnath, S. Keratoviz-A multistage keratoconus severity analysis and visualization using deep learning and class activated maps. *Int. J. Electr. Comput. Eng.* **2023**, *13*, 920–935. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.