Edinburgh Research Explorer

Deep Transfer Learning for Improved Detection of Keratoconus using Corneal Topographic Maps

Deep Transfer Learning for Improved Detection of Keratoconus using Corneal Topographic Maps

Ali H. Al-Timemy[1] · Nebras H. Ghaeb[1] · Zahraa M. Mosa[2] · Javier Escudero[3]

Abstract

Clinical keratoconus (KCN) detection is a challenging and time-consuming task. In the diagnosis process, ophthalmologists must revise demographic and clinical ophthalmic examinations. The latter include slit-lamb, corneal topographic maps, and Pentacam indices (PI). We propose an Ensemble of Deep Transfer Learning (EDTL) based on corneal topographic maps. We consider four pretrained networks, SqueezeNet (SqN), AlexNet (AN), ShuffleNet (SfN), and MobileNet-v2 (MN), and fine-tune them on a dataset of KCN and normal cases, each including four topographic maps. We also consider a PI classifier. Then, our EDTL method combines the output probabilities of each of the five classifiers to obtain a decision based on the fusion of probabilities. Individually, the classifier based on PI achieved 93.1% accuracy, whereas the deep classifiers reached classification accuracies over 90% only in isolated cases. Overall, the average accuracy of the deep networks over the four corneal maps ranged from 86% (SfN) to 89.9% (AN). The classifier ensemble increased the accuracy of the deep classifiers based on corneal maps to values ranging (92.2% to 93.1%) for SqN and (93.1% to 94.8%) for AN. Including in the ensemble-specific combinations of corneal maps' classifiers and PI increased the accuracy to 98.3%. Moreover, visualization of first learner filters in the networks and Grad-CAMs confirmed that the networks had learned relevant clinical features. This study shows the potential of creating ensembles of deep classifiers fine-tuned with a transfer learning strategy as it resulted in an improved accuracy while showing learnable filters and Grad-CAMs that agree with clinical knowledge. This is a step further towards the potential clinical deployment of an improved computer-assisted diagnosis system for KCN detection to help ophthalmologists to confirm the clinical decision and to perform fast and accurate KCN treatment.

Keywords AlexNet · Computer-aided diagnosis (CAD) · Corneal topographic maps · Deep learning · Keratoconus · Probability fusion · Transfer learning

Introduction

Keratoconus (KCN) is a non-inflammatory disease that can cause protrusion and thinning at the thinnest location of the cornea [1]. This may initiate blurred vision and high degree of astigmatism, potentially leading to vision loss if it is not detected and treated at an early stage. The loss of thickness of the cornea is a result of lack of some structural components, such as collagen fibrils [2]. The main signs of keratoconic cornea are having high keratoconic indices, myopia, and irregular astigmatism [3], caused by the changes in the geometry of the cornea [4]. The fundamental reasons of having KCN are unknown; however, ophthalmologists associate it with eye rubbing, systematic disease, and genetic inheritance. KCN progression can be fast or slow, and it may stop at certain stage [5].

Early detection and intervention of KCN will prevent patients from requiring complicated interventions, such as penetrating keratoplasty or corneal graft, which may lead to complications. Misdiagnosis or late detection of KCN may, in some extreme cases, cause vision loss [6]. Suspected eye with KCN is difficult to detect; therefore, a thorough

✉ Ali H. Al-Timemy
  ali.altimemy@kecbu.uobaghdad.edu.iq

✉ Javier Escudero
  javier.escudero@ed.ac.uk

1 Biomedical Engineering Department, Al-Khwarizmi College of Engineering, University of Baghdad, Jaderiyah, Baghdad, Iraq

2 College of Pharmacy, Uruk University, Baghdad, Iraq

3 School of Engineering, Institute for Digital Communications, The University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, UK

investigation is needed, before refractive surgery [7]. Hence, there is a need to diagnose a suspected KCN accurately to promote good case management and a primary treatment by surgery, if needed. This, in turn, reduces vision deterioration and keeps the refractive error within an acceptable range so that it can be corrected, after surgery, with simple methods such as vision lenses.

In order to achieve an accurate diagnosis, it is important to examine the anterior surface as well as the posterior surface of the cornea when an ophthalmologist needs to diagnose KCN [8]. The most commonly used devices for measuring the parameters of anterior corneal surface are topographical imagining (Scheimpflug or Placido rings) and optical coherence tomography (OCT), providing topographic maps of the corneal surface and corneal thickness [6, 8, 9]. The generated corneal topographic maps, commonly known as four refractive maps [10], are utilized for KCN detection and include sagittal (SAG), corneal thickness (CT), elevation front (EF), and elevation back (EB) maps.

The process of early detection and diagnosis of KCN is very challenging task since it requires great experience and knowledge and it also needs significant period of training. The KCN detection procedure is also time-consuming and complicated for junior ophthalmologists, and even for expert ophthalmologists. This is because it involves assessing the patient eye history, examining the topographic maps alongside derived indices, and, in some cases, utilizing the results of examinations from other instruments such as OCT to reach the final decision. The objective of this laborious and complex process is to provide an accurate decision towards the right treatment [11]. In particular, the ophthalmologist starts the process of KCN detection by subjectively examining the patient for eye redness, blurred vision, and having an itchy eye. Then, if needed, the ophthalmologist goes on to examine the eye with a Pentacam device (Scheimpflug imagining) resulting in four corneal topographic maps and derived Pentacam Indices (PI). A set of clinical signs and features has to be investigated in each of the four maps (SAG, CT, EF, and EB) of anterior corneal surface. Figure 1 shows examples of the 4 corneal topographic maps, for normal and KCN cases, with some of the main clinical investigations performed by a trained ophthalmologist to investigate KCN. Examples of the clinical features include (1) the irregularity of the bowtie shape and the bowtie's axis angle in the sagittal map, where bowtie is used to assess the astigmatism type, as a sign of KCN; (2) the thinnest location point at 6 mm of the corneal diameter and point of the centre of the CT map; and (3) checking if the bowtie shape is open or close in the EF and EB maps. In addition, the ophthalmologist examines other measurements obtained by the Pentacam device, i.e. PI of the anterior surface of the cornea, Keratometry readings, astigmatism degree, thinnest location at 6 mm, pachymetry apex, anterior corneal depth (ACD), and eccentricity. Other details are also examined such as colour distribution of each

Fig. 1 Example of the 4 corneal topographic maps, for normal (left) and KCN (right) cases, with the main clinical investigations performed by the ophthalmologist to check features and signs in the corneal topographic maps. SAG sagittal, CT corneal thickness, EF elevation front, EB elevation back, NOR normal, KCN keratoconus
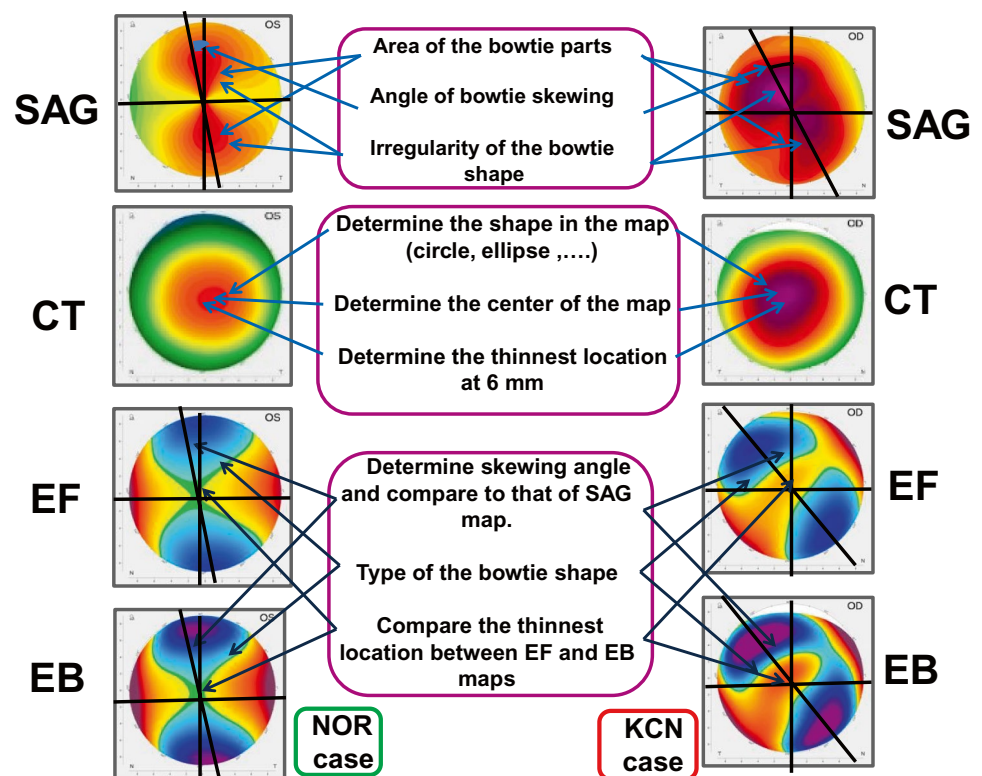
map. Then, it is decided if the patient has a KCN or not. In some cases, when the examination of the suspected case is not conclusive, other modalities are needed (e.g., OCT and specular microscopy) and consultation with additional ophthalmologists may also be required. This process requires considerable amount of training and great experience, and it is doctor-dependent and subjective [1]. The full details of the complicated clinical KCN detection procedure are explained in Sinjab [12].

Computer-based methods are useful to aid in the diagnosis of KCN. Hybrid computer-aided diagnosis (CAD) for KCN detection was proposed in Issarti et al. [10], based on a neural network, mathematical model, and a Grossberg-Runge Kutta architecture to detect clinical and suspect KCN. The accuracy of the hybrid CAD system was 96.56% compared to 79.00% for the topographical KCN classification. CT, EF, and EB maps were utilized without using the SAG map. In addition, the model was developed for the right eye only and it is not known if the results can be generalized to detect KCN in the other eye. Machine learning (ML) may become a crucial tool to aid the ophthalmologist for a better KCN detection, with corneal topographic maps [11, 13]. It is worth mentioning that ML has utilized features and indices, extracted semi-manually with image processing from the topographic maps, to classify KCN with ML classifiers [11]. The features and indices are variable for each map and for each eye disease. Furthermore, extracting features from each individual map may not be an easy task since the number of maps is more than four for some systems.

In recent years, deep learning (DL) methods are becoming rapidly adopted in healthcare sectors to play an important role as an assistive diagnostic method for their reliable accuracy and precision [14]. Convolutional neural network (CNN) is a DL architecture inspired from the model of visual cortex [14] and is applied to healthcare sector [15–17]. DL methods can help ophthalmologists for improving the quality of patients' care [18, 19], because of its ability to extract features which may not be detectable by human expert as well as no subjectivity in its decision making.

Deep learning can aid ophthalmologist by improving diagnosis and suggesting personalized treatments [20] in addition to save the ophthalmologist's time in examining maps and images of the diseased eyes. Despite the fact that the general clinical acceptance of the 'black-box' model is still an existing challenge [19], FDA approval has recently been given to an ophthalmic device,[1] the IDx-DR [21], that integrates a DL model for detection of diabetic retinopathy. This shows the potential of using DL in clinical

practice when embedded into ophthalmic devices. Continuing development and validation is being conducted aiming at advancing the clinical care of patients with other ophthalmic diseases with DL [22]. Therefore, DL has the potential to improve clinical decision support and modernize clinical ophthalmology practice in the upcoming years [19].

KCN detection with CNN has previously been investigated [23]; for instance, a CNN was used in Lavric and Valentin [24] to detect KCN using synthetic corneal topographic maps generated by SyntEyes KCN model. The obtained accuracy was equal to 99%. However, the performance was evaluated entirely with synthetic maps. The main challenges with training CNN include requiring large number of images to train the network, a common challenge in medical community. Moreover, the process of data collection, training, and tuning the network parameters is time-consuming. To tackle these challenges, pretrained CNNs on the ImageNet dataset of million images to classify large number of classes, such as AlexNet (AN) [25] and SqueezeNet (SqN) [26], ShuffleNet (SfN) [27], and MobileNet-v2 (MN) [28], can be utilized to learn a new task by fine tuning of the last fully connected layers with the process of transfer learning. Transfer learning is faster than designing a CNN and training it from scratch [29], where the new network can be used to classify small dataset of images without overfitting the model.

Few studies utilized transfer learning for KCN detection with corneal topographic maps. In Abdülhüssein et al. [30], VGG-16, a pretrained CNN, was utilized to detect individual topographic maps. The results of the classification accuracy were 88.8%, 98.9%, 94.8%, and 94.5% for SAG, EF, EB, and CT maps, respectively. It should be noted that the performance was evaluated on training and testing sets without a validation set. Furthermore, all maps were not used together to evaluate the performance. Kamiya et al. [31] proposed a system for KCN grade detection based on the ResNet-18 network. They separately trained six neural networks to classify each of the six maps; then, the average of KCN grades of 1–4 for the six networks was taken to classify four KCN grades on a dataset of 543 eyes. The grade accuracy for four KCN grades was equal to 87.4% using six color-coded maps and 99% accuracy for KCN versus normal. However, SAG map was not used in their work. In addition, fivefold cross validation was used to evaluate the system performance which may indicate that only training and testing sets were utilized to train and test the networks without the use of a validation set, needed to optimize the network. From previous literature, either only 2-way data split was used, or no decision fusion was proposed to have the output decision from all four corneal topographic maps and the PI. Furthermore, topographic maps from a single eye side, either left or right, were utilized; having a system that can detect KCN in both sides would lead to a more difficult

---

task, since the features extracted by CNN may be different for each eye side. However, this could be desirable in order to have a system that could work with images from either the left or right-side eye.

In this paper, we propose an ensemble of deep transfer learning (EDTL) with four probability fusion methods, including majority voting, averaging, product, and median, to obtain a decision from each topographic map and PI, then combine the probabilities to decide what the output of the suspected case is. Our work is motivated by the clinical need to help in the detection of KCN. This study contributes a classification pipelines that considers images from right and left eyes in contrast with previous studies that analyzed only one side. Furthermore, we also seek to validate the clinical relevance of the features extracted by the DL classifiers by visualizing learnable filters for KCN and control cases and a Grad-CAM of finally chosen classifier. Methodologically speaking, we propose a transfer learning strategy which is adopted to tackle the challenge of overfitting for the

small privately collected dataset of KCN and normal cases. To the best of our knowledge, this is the first work to adopt multi information fusion to combine the decisions from the corneal topographic maps and the PI, as a CAD tool to assist the ophthalmologist for KCN detection, saving the time and efforts, compared to clinical KCN detection steps, described in Sinjab [12]. The proposed system is validated with four types of pretrained transfer learning networks, SqN, AN, SfN, and MN, and we employ a 3-way data split on 2136 corneal topographic maps.

Methodology

Proposed Ensemble of Deep Transfer Learning

Figure 2 presents the block diagram of our proposed EDTL method with corneal topographic maps and PI classifiers. It illustrates the main phases of the system, training,

Fig. 2 Block diagram of the proposed ensemble of deep transfer learning (EDTL) for KCN detection. **A** Training phase, where the training set will be used to train the 5 networks/classifier and the validation set will be used to decide when to stop training; Input size is $227 \times 227 \times 3$ for SqN and AN, $224 \times 224 \times 3$ for MN and SfN. **B** Testing phase where the unseen testing set will be used for performance evaluation. Orange dashed rectangle indicates that classifiers are trained and tested with either with SqN or AN or MN or SfN, one at a time. Red rectangles around the maps illustrate samples of KCN maps, while green rectangles represent samples of normal maps. DL deep learning; PI Pentacam indices; NOR normal; KCN keratoconus; LRSGD-PI Logistic Regression with Stochastic Gradient Descent classifier for PI; K_1, K_2, K_{max}, and K_{mean} keratometry readings; Pack pachymetry apex; Astig astigmatism degree; ACD anterior corneal depth; P_{NOR} and P_{KCN} classifier output probability for NOR and KCN classes, respectively
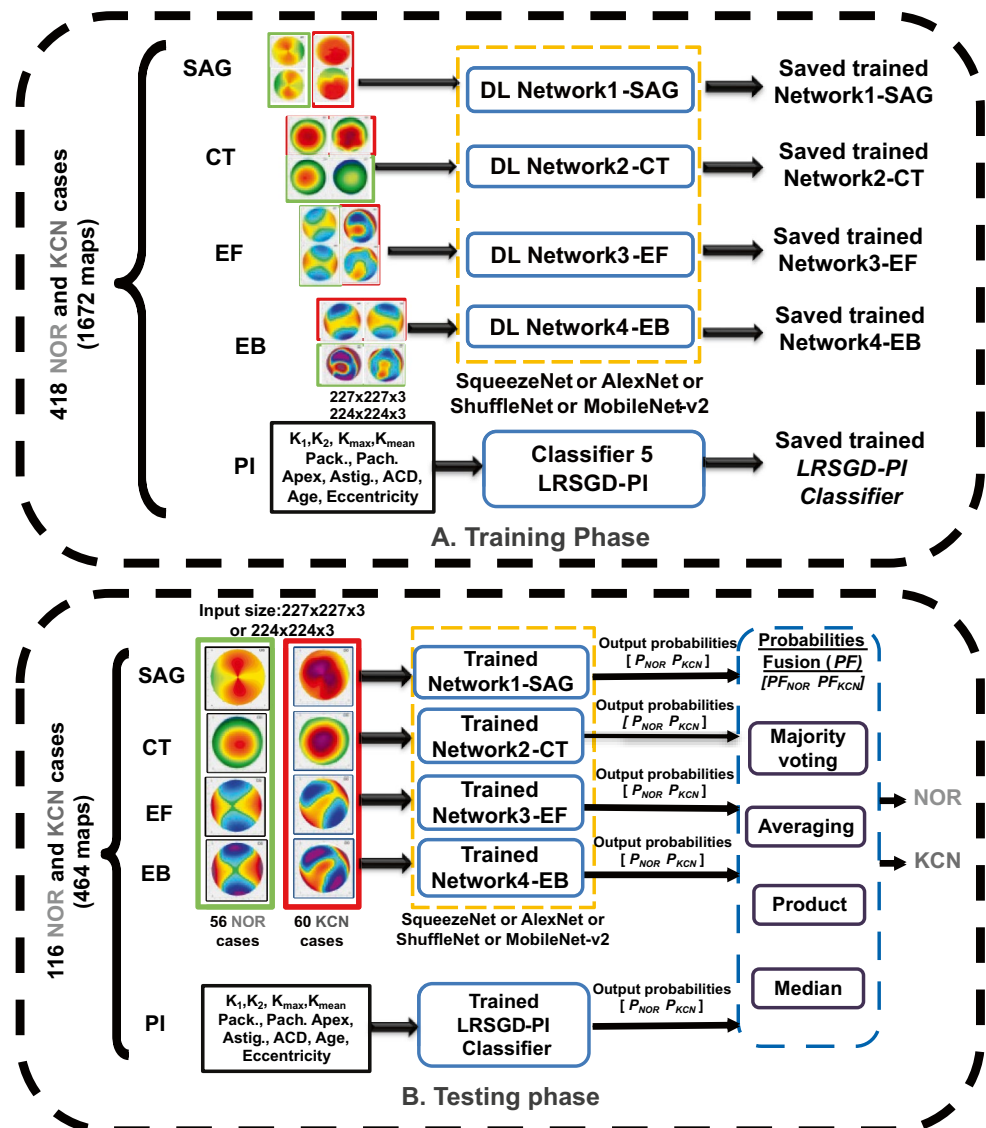
and testing phases. Four types of pretrained deep learning networks were used in this study including SqueezeNet (SqN) [26], AlexNet (AN) [25], ShuffleNet (SfN) [27], and MobileNet-v2 (MN) [28]. We modified the last fully connected classification layers to detect normal and KCN. These networks were used to classify the corneal topographical maps (i.e., SAG, CT, EF, and EB). A logistic regression classifier with Stochastic Gradient Descent optimizer (LRSGD-PI) is used to classify the Pentacam indices (PI). We will test all pretrained DL networks with the proposed EDTL method, one at a time, and evaluate their performance, as shown in Fig. 2.

First, in the training phase shown in Fig. 2A, four networks and a PI classifier (LRSGD-PI) are trained individually, utilizing training set while the validation set is used to optimize the network. It is important to mention here that 3-way data split is performed in this study into training, validation, and testing sets. The trained classifiers (deep networks with transfer learning for each type of map, plus the LRSGD-PI classifier) are saved for later use to evaluate their performance on the unseen testing maps in the testing phase.

During the testing phase illustrated in Fig. 2B, the normal and KCN images are introduced to the saved networks in the form of sets of four maps. We obtain the output probability of each trained classifier, in sets of vectors $P_{net} = [P_{NOR}, P_{KCN}]$, where P_{NOR} is classifier output probability for normal and P_{KCN} is classifier output probability for KCN. The PI related to each subject will be introduced to the LRSGD-PI. A key element of the proposed EDTL method is the introduction of probability fusion step where all output probabilities from all four networks and PI classifier are concatenated $P = [P_{net-SAG}, P_{net-CT}, P_{net-EF}, P_{net-EB}, P_{LRSGD-PI}]$, then fused as will be explained next.

In order to fuse the continuous output of the 5 networks/classifier illustrated in Fig. 2B, we utilized four probability fusion methods, described in Polikar [32]: majority voting (MV), averaging, product, and median. An illustrative example of the probability fusion methods with true numbers is displayed in Fig. 3. We apply the predefined fixed rules, without any additional modification or fine-tuning. The result of the fusion method (PF) will be calculated to obtain $PF = [P_{NOR}, P_{KCN}]$ where PF_{NOR} is the result of the probabilities fusion for the normal class and PF_{KCN} is the result of the probabilities fusion for the KCN class. The output probabilities of all classifiers are considered with equal weights. For each case introduced to the five trained classifiers, the values of (PF_{NOR}) is compared against (PF_{KCN}). If $PF_{NOR} > PF_{KCN}$, then the output of the case will be normal class. Instead, if $PF_{NOR} < PF_{KCN}$, then the output will be KCN. The evaluation of the proposed EDTL
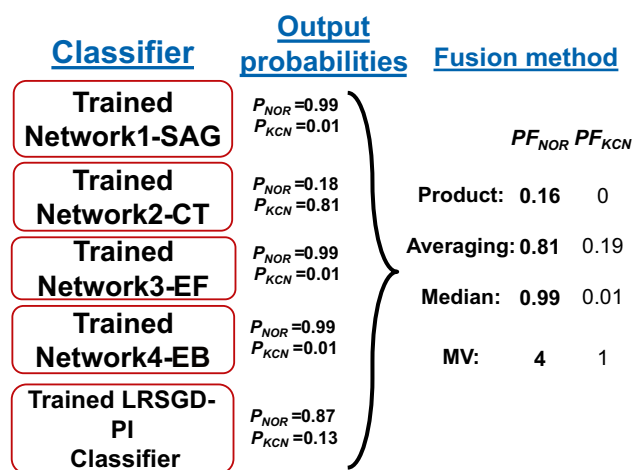

Fig. 3 Illustrative example of the process of probability fusion. The output class is shown with bold. PF_{NOR} and PF_{KCN} results of the probabilities' fusion for NOR and KCN classes, respectively. MV majority voting

method will be presented in 'Evaluation of the Proposed EDTL Method'.

Deep Transfer Learning Networks via SqN, AN, SfN, and MN

Transfer learning of pretrained CNN can help to address the challenge of having a reduced dataset to train a CNN and can help to avoid overfitting [16, 17]. Moreover, it reduces the time required for fine tuning the network [15]. In transfer learning, a pretrained CNN is used where early network layers learn low-level features such as edges and colours, while the last layers learn class-specific features [29]. Then, the early layers are then transferred and the last layers are replaced to match the size of the new dataset [14], two classes in our work, normal and KCN. Figure 4 shows the process of transfer learning applied to the challenging task of KCN detection.

In the current study, four pretrained networks were utilized with the proposed EDTL method, SqN [26], AN [25], SfN [27], and MN [28]. They have been applied in ophthalmic field; for instance, AN was used to improve diabetic retinopathy detection [33] and SqN was used for classification of cataracts [34]. The two deep learning networks are pretrained on over 1 million images (ImageNet dataset) and can classify them into 1000 classes. These networks will take an input image of size $227 \times 227 \times 3$ for AN and SqN, while the rest has an input of $224 \times 224 \times 3$ and produce a class label as an output with probabilities output for each class. The last fully connected layers of SqN and AN have been replaced to match the number of classes in this study, i.e. two classes (normal and KCN), see Table 1.

Fig. 4 The process of transfer learning applied to the problem of KCN detection, modified from [15]
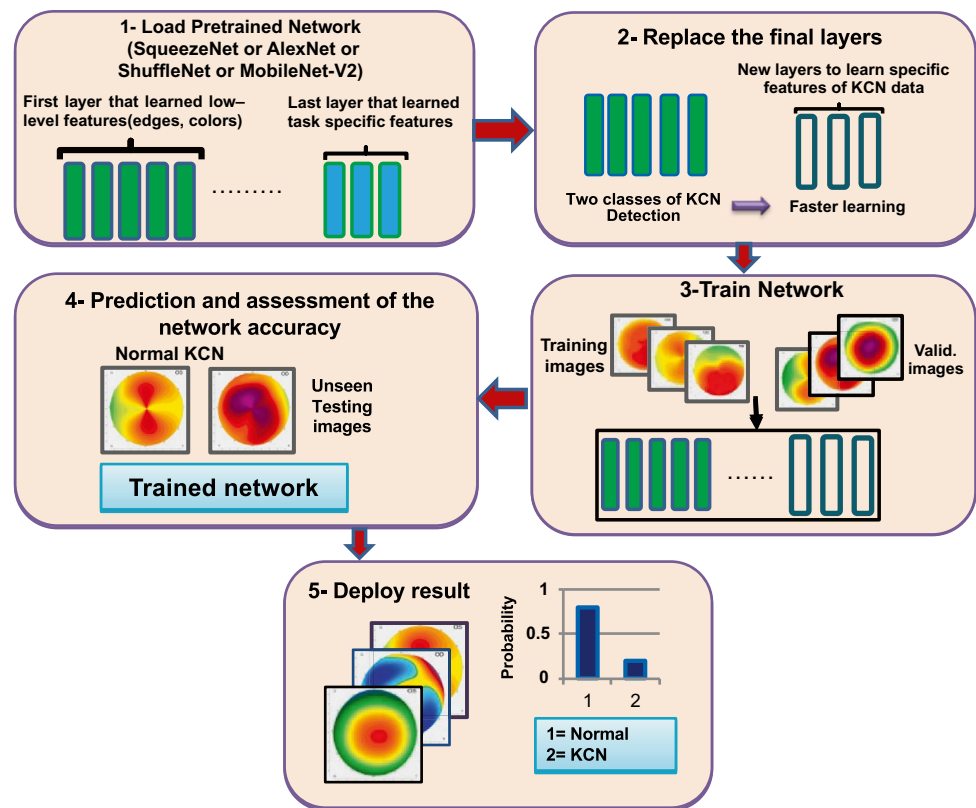
Data Collection

The dataset used in this study is based on pre-assessed KCN patients and normal participants referred to Al-Amal ophthalmic centre in Baghdad, Iraq. Each participant had undergone sets of measurements. One of these measurements is the topographical images using the Pentacam Scheimpflug measurements (Oculus GmbH, Wetzlar, Germany). The number of collected cases is 444 (226 right eye cases, 218 left eye cases). The collected maps from the Pentacam are

Table 1 The default and the new sizes of the last output layers of the networks used in the study. A. The five modified layers of SqN. B. The three modified layers of AN, SfN, and MN

No	Layer name	Default output size	New output size
A			
1	Convolution	$14 \times 14 \times 1000$	$14 \times 14 \times 2$
2	ReLU	$14 \times 14 \times 1000$	$14 \times 14 \times 2$
3	Average pooling	$1 \times 1 \times 1000$	$1 \times 1 \times 2$
4	Softmax	$1 \times 1 \times 1000$	$1 \times 1 \times 2$
5	Classification output	1000 classes	Normal and KCN
B			
1	Fully connected	$1 \times 1 \times 1000$	$1 \times 1 \times 2$
2	Softmax	$1 \times 1 \times 1000$	$1 \times 1 \times 2$
3	Classification output	1000 classes	Normal and KCN

the four standard topographical refractive maps which are SAG, CT, EF, and EB maps for a diameter of 8 mm. PI of anterior surface of the cornea were also included for all cases, which are keratometry readings (mean, maximum, K_1, and K_2), astigmatism degree, thinnest location at 6 mm, pachymetry apex, anterior corneal depth (ACD), and eccentricity; These indices have been utilized clinically in KCN detection alongside the corneal maps [9, 11]. For more information about the four refractive maps, the reader is referred to Sinjab [12]. The experiment was done according to declaration of Helsinki [35], and its later amendments and the privacy of all subjects were preserved.

Of the 444 cases, 266 cases bilateral with normal topography were included (subjects aged 31.4 ± 9.2 years, mean $\pm$ Standard Deviation, SD), with no other ocular, subjectively and slit lamp abnormality symptoms. The normal cases included 219 normal and 47 forme fruste cases where the readings were similar to normal and the disease is stopped and did not progress. Forme fruste cases can be regarded as a cornea with no abnormal findings which is confirmed by both corneal topography and slit-lamp examinations [36].

Other cases included 178 keratoconic eyes (subjects aged 32.95 ± 10.86 years, mean $\pm$ SD). The keratoconic eyes included 152 KCN and 26 ectasia cases. The decision of the KCN was supported also by subjective and slit lamp examinations. The classification of the two groups was screened

by ophthalmologist and ophthalmology supervisor to reach an optimum decision. Any patient with other ocular diseases was excluded from the study. Details of the two groups with their Pentacam indices are given in Table 2. For some subjects, there were more than one set of maps, but taken from different angles. Each additional set was diagnosed, and it was included in the study.

Since KCN is a bilateral disease and also because of the limited sample size, both right and left side cases were included which is a more difficult classification problem than that of a single sided eye. It is expected that the system to be able to detect KCN in both eyes, but probably with less performance than that of a single eye KCN detection system.

Evaluation of the Proposed EDTL Method

In the subsequent subsections, the evaluation of the proposed EDTL method will be presented. We include information about data collection, data augmentation, and preprocessing. The two key parts, training and testing phases, will also be presented. The proposed method will be tested with all four types of pretrained networks SqN, AN, SfN, and MN, one at a time. Matlab 2019b software (MathWorks, USA) was used to perform the analyses with Deep Learning Toolbox [29]. The analysis was performed on a single core, 2.6 GHz Core *i5* computer with 16 GB RAM.

Data Augmentation and Preprocessing

Since there are 266 normal and 178 KCN cases, there is a case of class imbalance where the samples of one class in a dataset outnumber the samples of the other class. The class imbalance is a common issue with DL problems [37], and it can influence the results [10]. To deal with this issue, we

Table 2 The Pentacam indices of the cases accompanying the topographic maps for normal and KCN groups (mean ± SD) used for the LRSGD-PI classifier

Group	Normal (266)	KCN (178)
K_{mean} (D)	43.47 ± 1.48	45.98 ± 3.74
K_{max} (D)	44.87 ± 1.6	51.73 ± 5.2
Pachymetry$_{min}$ (µm)	536.3 ± 34.0	471.54 ± 49.8
Astig. (D)	− 1.55 ± 1.45	− 1.73 ± 3.4
K_1 (D)	42.6 ± 1.58	44.58 ± 3.62
K_2 (D)	44.4 ± 1.6	47.57 ± 4.3
Eccentricity	0.43 ± 0.14	0.34 ± 0.63
Pachymetry apex	540.4 ± 34.9	484.2 ± 48.9
Anterior corneal chamber depth (ACD)	3.14 ± 0.4	3.19 ± 0.6

KCN keratoconus, K_{mean} mean keratometry, K_{max} maximum keratometry, *Astig* anterior corneal astigmatism, *D* dioptre

exploited data augmentation, a common procedure to deal with class imbalance for DL [37]. Augmentation can be done where the minority class is augmented through operations such as rotation, scaling, translation, and flipping of images [37].

In this study, scaling and translation of the topographic maps can change the diagnosis of a specific KCN or normal case as discussed with the ophthalmology supervisor. In addition, rotation of the maps by a certain angle in either clockwise or anticlockwise directions can also affect the diagnosis of the case since the skew angle of certain maps (SAG, EF, and EB) is one of essential points in KCN detection. Therefore, augmentation via rotation was excluded. However, flipping can be done around the *y*-axis of each map. Vertical flipping around the *y*-axis is used in this study to perform augmentation of the KCN class to balance the data to that of normal class, as agreed with the ophthalmology supervisor. Augmentation was performed on randomly selected 90 cases from the 178 KCN cases by vertical flipping. The KCN group size after augmentation became 268 KCN cases (178 + 90 augmented cases). Thus, 268 KCN cases and 266 normal cases will be used for performance evaluation, which makes the total number to be 534 cases; each has four topographic maps (2136 maps).

To tackle the issue of class imbalance for the Pentacam indices, we utilized Synthetic Minority Over-Sampling Technique (SMOTE) [38] to equalize the number of classes, by adding extra random 90 cases, similar to that of the augmentation of the four corneal maps. In SMOTE, the size of the majority class is kept the same, whereas *k*-nearest neighbour (kNN) is utilized in to create synthetic instances to oversample the minority classes [38]. In this study, Weka filter 'SMOTE' [39] was used where the number of *k*-nearest neighbour was equal to 5.

To preprocess the topographic maps, first we extract each topographic map without the map measurements and colour code. Then, all the 2136 maps were resized with the default Matlab Bicubic interpolation (weighted average of pixels in the nearest 4-by-4 neighbourhood), with antialiasing filter, to $227 \times 227 \times 3$ or $224 \times 224 \times 3$, to match the input size of the pretrained networks.

Training Phase

In order to train our DL networks (SqN, AN, SfN, and MN) and the PI with the proposed EDTL method, the dataset of 534 cases were randomly divided into 3-way data split, training set (66%, 355 cases), validation set (12%, 63 cases), and testing set (22%, 116 cases). It should be noted that the augmented images and PI with SMOTE were only used during the training phase and the unseen test set consisted only of non-augmented cases. The training set is used to train the

classifiers, and validation set is used to decide when to stop the training based on the validation accuracy, the loss value, and testing accuracy.

The pretrained networks (SqN, AN, SfN, and MN) were imported with the add-on Matlab explorer, then the last layers were replaced with new layers to match our KCN detection problem of two classes, KCN and normal as explained in 'Deep Transfer Learning Networks via SqN, AN, SfN, and MN'. All topographic maps were pre-processed to match the input size as explained in 'Data Augmentation and Preprocessing'.

Four SqN networks were trained on each of the four maps (SAG, CT, EF, and EB), using the training set. In addition, LRSGD-PI classifier was implemented in Weka [39], and it was utilized as the PI classifier, see Fig. 2, with utilizing training and validation sets. After training was finished, the four trained SqN classifiers were saved to be used later in the evaluation phase. The same was repeated for AN, SfN, and MN, and the five networks/classifier were saved to be used in the evaluation phase including the LRSGD-PI classifier. The network optimizer Stochastic Gradient Descent with Momentum (SGDM) was utilized for optimizing the loss function for the pretrained networks. Various iterations and epoch numbers have been investigated and performed in order to train the networks and to find the optimal epoch number. For the best trained networks, the training time for each network of SqN was between 6 and 12 min, while it was shorter for AN of about 4–8 min and longer for SfN, and MN because of the big size of these networks. The number of epochs for SqN was equal to 3, 11, 6, and 5 for SAG, CT, EF, and EB maps, respectively while it was equal to 4, 3, 4, and 5 for AN.

Testing Phase

To evaluate the performance of the proposed EDTL method, the unseen test set (56 normal and 60 KCN cases, each has 4 topographic maps and the PI) is used. First, the four trained SqN are tested by introducing each case, then, the output probabilities for each map, and also, the PI classifiers are calculated. Afterwards, we utilized the probability fusion step to combine the five output probabilities vectors. The product of the probabilities is then obtained, $PF = [PF_{NOR}, PF_{KCN}]$, and the case output is determined either normal or KCN based on the comparison of PF_{NOR} and PF_{KCN} as explained in 'Proposed Ensemble of Deep Transfer Learning'. The same steps above were repeated for AN, SfN, and MN. To examine the robustness of our proposed method with other classifiers' combinations, different fusion of three classifiers were investigated with the probabilities fusion to find the best three classifiers' fusion. The results will be presented and discussed in the next sections.

The performance of the proposed EDTL was evaluated with the following classification performance measures which include accuracy, sensitivity, specificity, and precision, given by

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

$$Sensitivity = \frac{TP}{TP + FN} \qquad (2)$$

$$Specificity = \frac{TN}{TN + FP} \qquad (3)$$

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

where TP (True Positive) is the number of KCN cases identified correctly, TN (True Negative) is the number of normal cases identified correctly, FN (False Negative) is the number of KCN cases classified as normal, and FP (False Positive) is the number of normal cases classified as KCN. We also calculated the confusion matrix for the two classes (normal and KCN). In a confusion matrix, the diagonal line represents the correctly classified cases, while the off diagonal represents the wrong classifications.

Visualization

To better understand the networks capability in capturing clinically relevant features from the topographic maps, (1) we plotted the learnable filters for AN and SqN and Grad-CAM for SqN. The learnable filters will be plotted for 'pool1' layer in SqN and AN for the best single map performed, for normal and KCN cases, as examples. Furthermore, we also plotted Gradient-weighted Class Activation Map (Grad-CAM) [40], which has been used in the medical field [41, 42] to better understand the capability of the SqN in capturing the relevant features from corneal maps. Expert ophthalmology supervisor will examine the learnable filters of SqN and AN and also the Grad-CAM for normal and KCN cases, to check if there is any relevant clinical features in the visualizations of the networks.

Results

Testing of Individual Maps with the Unseen Testing Set

Table 3 displays the results for each individual input to the network (SAG, CT, EF, and EB maps) when evaluated on the unseen test set, for all deep transfer learning networks

Table 3 The classification accuracy rates in percentage estimated on the testing set (116 cases) of the four individual trained networks and LRSGD-PI classifier, as detailed in Fig. 2, for SqN, AN, SfN, and MN. The best map is shown with bold

No	Network name	Classification accuracy (%) for input map				Average of 4 maps' accuracy (%)	LRSGD-PI classifier accuracy (%)
		SAG	CT	EF	EB		
1	SqN	90.5	84.5	**93.1**	88.8	89.2	93.1
2	AN	**95.7**	83.6	87.0	93.1	89.9	
3	SfN	85.34	85.3	86.2	**87**	86	
4	MN	87	80.2	**91.4**	87	86.4	

SAG sagittal, *CT* corneal thickness, *EF* elevation front, *EB* elevation back

investigated in this study: SqN, AN, SfN, and MN. EF map was the best performer map for SqN (93.1%) and MN (91.4%), while SAG map was the best map for AN (95.7%). CT map was the topographic map with the lowest classification accuracy. As for the 5th classifier—LRSGD-PI, displayed in Fig. 2—the classification accuracy was equal to 93.1%. The average of the classification accuracy for all four maps' classifiers, illustrated in Table 3, was equal to 89.9% for AN compared to 89.2% to SqN and 86.4% for MN.

Testing of the Proposed EDTL Method with Multiple Probability Fusion Methods

In Table 4, the classification accuracy is illustrated for the proposed EDTL for all networks when using the unseen testing set with all types of fusion methods. It is noteworthy to mention that fusion techniques do not require fine-tuning and, therefore, they were applied directly to the test set. MV and median fusion methods outperformed other fusion methods for AN, SfN, and MN, while for SqN, averaging and product were the best performers. The highest best fusion accuracies were 93.1% for SqN, 94.8% for AN, 87.9% for SfN, and 90.5% for MN.

An example of the proposed EDTL's output can be seen in Fig. 5 for probabilities fusion methods for a normal and KCN cases. For the normal case, the probability fusion of CT map (low performance map in Table 3) indicted the case is KCN. However, the EDTL predicted the case as normal with all fusion rules. For the KCN case, all fusion methods detected the case as KCN.

The next part of the analysis was to investigate different classifiers' combinations with the proposed EDTL to find the results for best fusion of three network/classifier among the five trained classifiers, i.e. SAG, CT, EF, and EB, and LRSGD-PI classifier. We evaluated different combinations of three networks (odd number) with different probability fusion methods, for SqN, AN, SfN, and MN, since median and MV methods require odd number of probabilities to get a decision. The LRSGD-PI classifier was included in all combination as it is an essential part of the process of

clinical KCN detection [9]. The accuracy results are illustrated in Table 5.

In Table 5, classifiers tested with SqN for EF, EB maps, and the LRSGD-PI classifier achieved an accuracy of 93.1% better than other combinations, for all fusion methods. The number of misclassifications was equal to eight KCN cases out of 60 KCN cases being misclassified as normal, for probabilities' fusion of the best three maps. All 56 normal cases were classified correctly. As for AN result (Table 5), the fusion of SAG and EB maps and also LRSGD-PI classifier outperformed other combinations with an accuracy of 98.3% with product fusion. It is worth to mention that the best three combinations achieved an accuracy of 98.3% better than using all classifiers with AN (94.8%; Table 4), and also better than the average of 5 classifiers (89.9%; Table 3), in addition to an improvement of around 8% (Table 3) for MN and SfN (SAG, EF, and LRSGD-PI), which shows the improvements with the proposed EDTL method. From Table 5, the EB and the LRSGD-PI classifier are shared among the best classifiers for AN and SqN and SAG and EF for SfN and MN.

We further investigate the performance of the 3 classifier combinations, when using the unseen testing set, with SqN and AN (best performers in Table 4), with other parameters precision, sensitivity, specificity, and accuracy where the results are illustrated in Table 6, respectively. In Table 6, the results were similar for all fusion methods where all normal cases were classified correctly and 8 KCN cases

Table 4 The classification accuracy rates in percentage of the EDTL method for each type of the SqN, AN, SfN, and MN transfer learning networks (Fig. 2b) for fusion of all outputs, for the testing set with four probability fusion methods

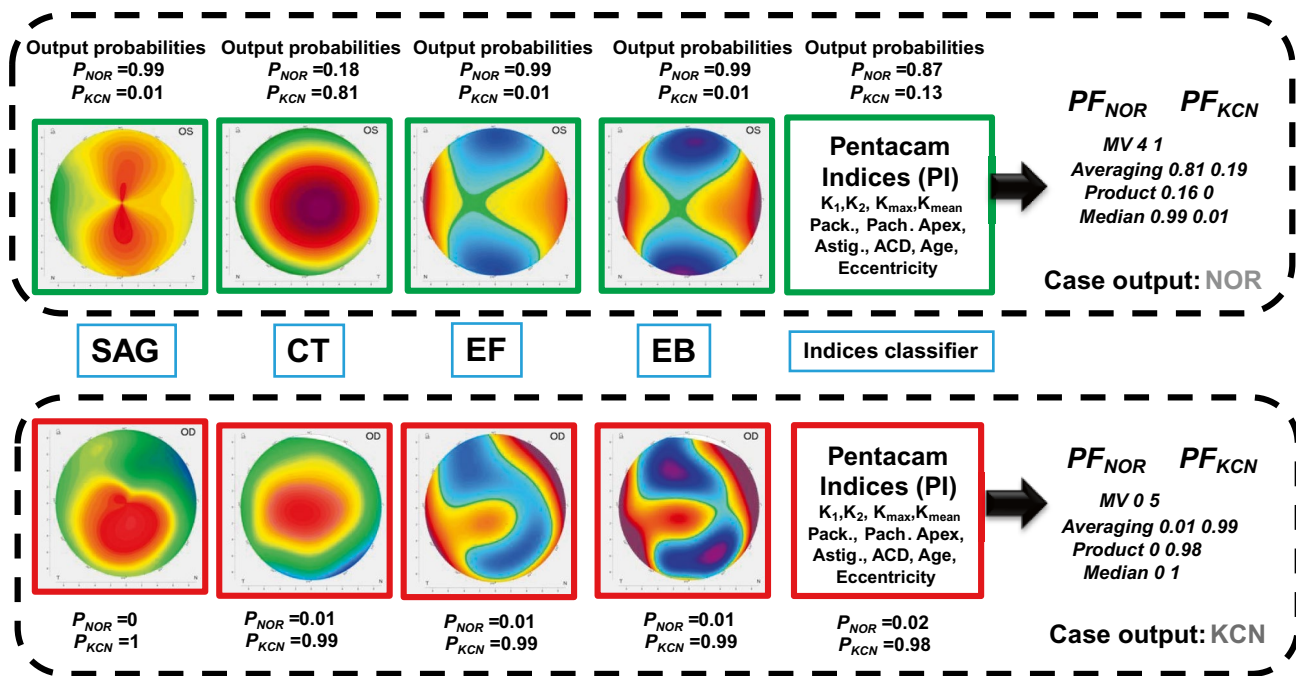| Transfer learning network	Classification accuracy (%) for probability fusion method			
	MV	Averaging	Product	Median
SqN	92.2	93.1	93.1	92.2
AN	94.8	94.0	93.1	94.8
SfN	87.9	87.1	87.9	87.9
MN	90.5	88.8	88.8	90.5

Fig. 5 An example of the EDTL output with different fusion methods for normal and KCN cases

were misclassified as normal, see corresponding confusion matrix in Table 7 for SqN. As for AN, the product fusion method outperformed other fusion methods, for the case of 3 classifier' combinations. The confusion matrix in Table 7 shows that when using AN, only 2 KCN cases were misclassified as normal and all normal cases were classified correctly with an accuracy of 98.3%, which shows the importance of our proposed EDTL method.

Learnable Filters and Grad-CAM Visualizations

Figure 6 shows examples of learnable filters for 'pool1' layer in SqN for EF map (best single map performed for SqN; Table 3) for a normal and KCN cases, while in Fig. 7, examples of learnable filters for 'pool1' layer in AN for SAG map for a normal and KCN cases are shown.

As for the Grad-CAM visualization, Fig. 8 shows the visualization of Grad-CAM for a normal case for the 4 refractive maps, SAG, CT, EF, and EB maps, respectively. In Fig. 9, the Grad-CAM plots are shown for the 4 refractive maps of a KCN case.

Discussion

A new method for KCN detection with EDTL method was proposed and validated in this study based on corneal topographic maps dataset and derived PI. From the results

illustrated in Table 3 for testing each individual map, with the trained SqN and AN, SAG map was the best performer for AN and the second best performer for SqN and MN. We utilized 3-way data split into training, validation, and testing sets in our study. The training set was used to train and optimize the networks, where the validation set is used to decide when to stop training, while the unseen test set was used for performance evaluation of our proposed EDTL. In Lavric and Valentin [24], 3-way data split was used and achieved an accuracy of 99.3%. However, their work was based on synthetic maps generated by a SyntEyes model [43].

Learnable filters showed that discriminative features can be observed, similar to the features of clinical importance illustrated in Fig. 1, such as regularity of bowtie shape (Fig. 6A and B). The ophthalmology supervisor confirmed clear area of the superior and inferior bowtie (Fig. 6A), while a tongue-like bowtie shape is starting to be isolated (Fig. 6B) for KCN case. Very clear bowtie shape, that can be used to estimate the weight of astigmatism, can be observed for the normal case (Fig. 7A) and irregular and total inferior bowtie shape can be seen (Fig. 7B), as confirmed by ophthalmology supervisor, which reflects an advanced KCN case. Overall discussions with the ophthalmology supervisor on Figs. 6 and 7 showed that there are many important clinical features, obtained with SqN and AN, that are not seen in the original image such as irregularity of the bowtie shape. AN and SqN are simple networks and require low computational power—even allowing running on a CPU machine—which

Table 5 The results of proposed EDTL for SqN, AN, SfN, and MN, in terms of accuracy (%), for different probability fusion methods of the three classifiers. The results for the best classifiers' combination are shown in bold font

Net	Classifiers' combinations	Accuracy (%) for different probability fusion methods			
		MV	Averaging	Product	Median
SqN	Net.1-SAG, Net.2-CT and LRSGD-PI class	92.2	91.4	89.7	92.2
	Net.1-SAG, Net.3-EF and LRSGD-PI class	93.1	92.2	92.2	93.1
	Net.1-SAG, Net.4-EB and LRSGD-PI class	91.4	91.4	92.2	91.4
	Net.2-CT, Net.3-EF and LRSGD-PI class	92.2	92.2	91.4	92.2
	Net.2-CT, Net.4-EB and LRSGD-PI class	92.2	91.4	90.5	92.2
	Net.3-EF, Net.4-EB and LRSGD-PI class	**93.1**	**93.1**	**93.1**	**93.1**
AN	Net.1-SAG, Net.2-CT and LRSGD-PI class	96.6	96.6	95.7	96.6
	Net.1-SAG, Net.3-EF and LRSGD-PI class	94.8	94.8	93.1	94.8
	Net.1-SAG, Net.4-EB and LRSGD-PI class	**96.6**	**97.4**	**98.3**	**96.6**
	Net.2-CT, Net.3-EF and LRSGD-PI class	89.7	91.4	91.4	89.7
	Net.2-CT, Net.4-EB and LRSGD-PI class	94.0	92.2	92.2	94.0
	Net.3-EF, Net.4-EB and LRSGD-PI class	94.0	92.2	92.2	94.0
SfN	Net.1-SAG, Net.2-CT and LRSGD-PI class	93.1	90.5	89.7	93.1
	Net.1-SAG, Net.3-EF and LRSGD-PI class	**94.0**	**94.0**	**93.1**	**94.0**
	Net.1-SAG, Net.4-EB and LRSGD-PI class	92.2	89.7	89.7	92.2
	Net.2-CT, Net.3-EF and LRSGD-PI class	87.9	87.1	86.2	87.9
	Net.2-CT, Net.4-EB and LRSGD-PI class	87.9	87.1	87.1	87.9
	Net.3-EF, Net.4-EB and LRSGD-PI class	88.8	90.5	90.5	88.8
MN	Net.1-SAG, Net.2-CT and LRSGD-PI class	90.5	89.7	88.8	90.5
	Net.1-SAG, Net.3-EF and LRSGD-PI class	**94.0**	**93.1**	**94.0**	**94.0**
	Net.1-SAG, Net.4-EB and LRSGD-PI class	91.4	90.5	90.5	91.4
	Net.2-CT, Net.3-EF and LRSGD-PI class	89.7	88.8	88.8	89.7
	Net.2-CT, Net.4-EB and LRSGD-PI class	87.1	87.1	87.9	87.1
	Net.3-EF, Net.4-EB and LRSGD-PI class	90.5	89.7	89.7	90.5

Net network, *ind.* indices, *class.* classification, *MV* majority voting

may be relevant for resource-limited settings as well as performing well on the problem of KCN detection.

Grad-CAM visualizations illustrated that the activation maps point towards areas of clinical interest in the KCN case depicted in Fig. 9, including the localized composition of conus (cornea) in Fig. 9A; the thinning area that starts to clear in Fig. 9B; and the localized tongue position in Fig. 9C. An ophthalmologist supervisor confirmed the clinical relevance of these features. As for Fig. 8 for the normal case, the activations were away from the 6-mm central clinical circle, which is the main area of clinical interest. Therefore, the locations are not pathologically driven in the case of Fig. 8.

It is worth to mention here that the average of classification accuracies for using each individual map illustrated in Table 3 was equal to 89.9% for AN compared to 89.2% for SqN and 86% for SfN and MN. When applying our proposed EDTL for each case of the unseen testing set, the

Table 6 The results of proposed EDTL for SqN and AN, in terms of precision, sensitivity, specificity, and accuracy for the best 3 combination (A. SqN: Net.3-EF, Net.4-EB, and LRSGD-PI classifier, B. AN: of Net.1-SAG, Net.4-EB, and LRSGD-PI classifier) with different fusion methods

	Performance measure	MV	Averaging	Product	Median
A. SqN	Precision (%)	100	100	100	100
	Sensitivity (%)	86.6	86.6	86.6	86.6
	Specificity (%)	100	100	100	100
	Accuracy (%)	93.1	93.1	93.1	93.1
B. AN	Precision (%)	98.3	100	100	98.3
	Sensitivity (%)	95	95	96.6	95
	Specificity (%)	98.2	100	100	98.2
	Accuracy (%)	96.6	97.4	98.3	96.6

Table 7 Confusion matrix for the best 3 probabilities' fusion combinations with product fusion for A. SqN, EF, and EB maps and also the LRSGD-PI classifier. B. AN, SAG, and EB maps and the LRSGD-PI classifier

A				B			
Actual	KCN	NOR		Actual	KCN	NOR	
Predicted				Predicted			
KCN	52	0		KCN	58	0	
NOR	8	56		NOR	2	56	

Fig. 6 Learnable filters from the first 'pool1' layer of SqN for Elevation Front (EF) map for **A** normal case and **B** KCN case. Discriminative images with clear clinical features are enclosed with yellow squares
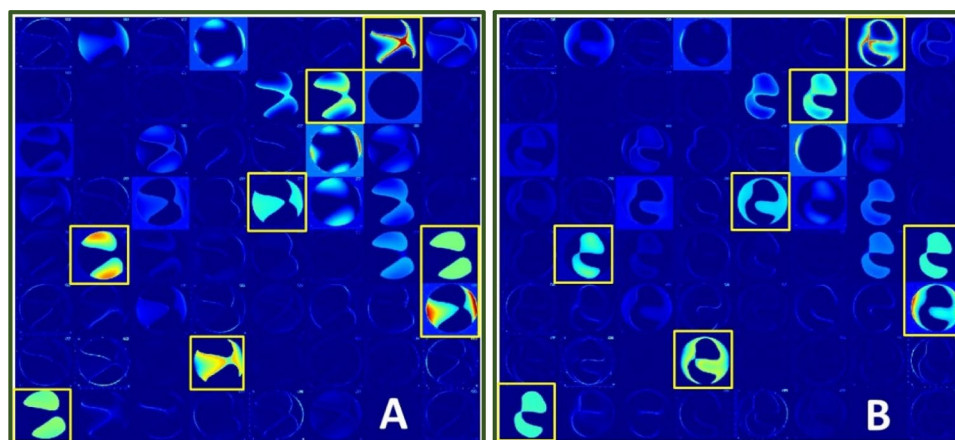
classification accuracy increased by approximately 5% for AN and 4% for SqN and MN, as illustrated in Table 4. Since four fusion methods were included, the median and MV fusion methods outperformed other fusion methods for AN.

The second part of the analysis was to investigate different input maps' combinations with the proposed EDTL to find the results for best fusion of three classifiers among the trained classifiers explained in 'Proposed Ensemble of Deep Transfer Learning'. In Table 5, classification accuracy of 94% was obtained with the proposed EDTL method using SfN and MN with SAG and EF maps and LRSGD-PI compared to 86% (Table 3) when using the proposed method on all four input maps and LRSGD-PI classifier. As for the AN result displayed in Table 5, SAG, EB maps, and PI achieved an accuracy of 98.3% with the proposed EDTL method against 94.8% when using all classifiers. It should be noted that the average of 4 individual maps was equal to 89.9% (Table 3), which shows the importance of the proposed EDTL method. This is potentially due to the fact that CT map has small values of output probabilities for KCN class compared to that of the other maps, as the performance of CT map was low (Table 3). Adding the probabilities of the CT map to the other probabilities in the fusion step reduced the fused probability value (PF_{KCN}), thus reducing classification accuracy. This may explain why including CT map in the full set reduced the classification accuracy compared to the best three classifiers.

It is noteworthy to mention that 100% specificity rate was obtained in this study, for both AN and SqN (Table 6), where all normal cases were detected correctly. This is potentially useful to the ophthalmologists since it may allow them to label the KCN cases, with our proposed EDTL method trained on an external dataset, where they could be very precise that a case is KCN when the proposed EDTL tells them so.

The four refractive maps (SAG, CT, EF, and EB) are clinically the first line of decision that should be reviewed by the ophthalmologist to recognize the patient state: normal or KCN. Each one of the four maps has an important role that may be changed according to the patient complicated state. The SAG map appears abnormal when values of keratometry exceed 46 D. The CT is considered as abnormal when the thickness is less than 450 μm; the elevation maps (EF and EB) are considered as abnormal when the thickness exceeds 15 μm. These values are clinical standards and are interchanged; it is not easy to be defined or appears in one case study. This may be the reason behind the variable accuracies

Fig. 7 Learnable filters from the first 'pool1' layer of AN for Sagittal (SAG) map for **A** normal case and **B** KCN case. Discriminative images are enclosed with yellow squares
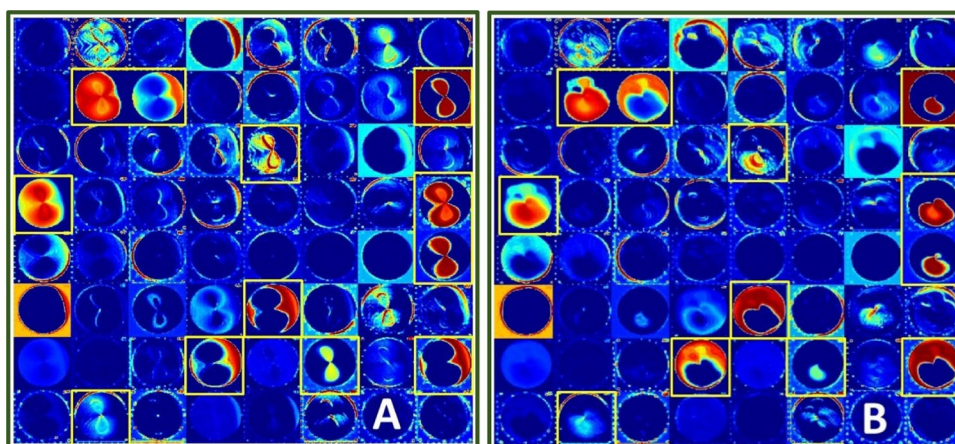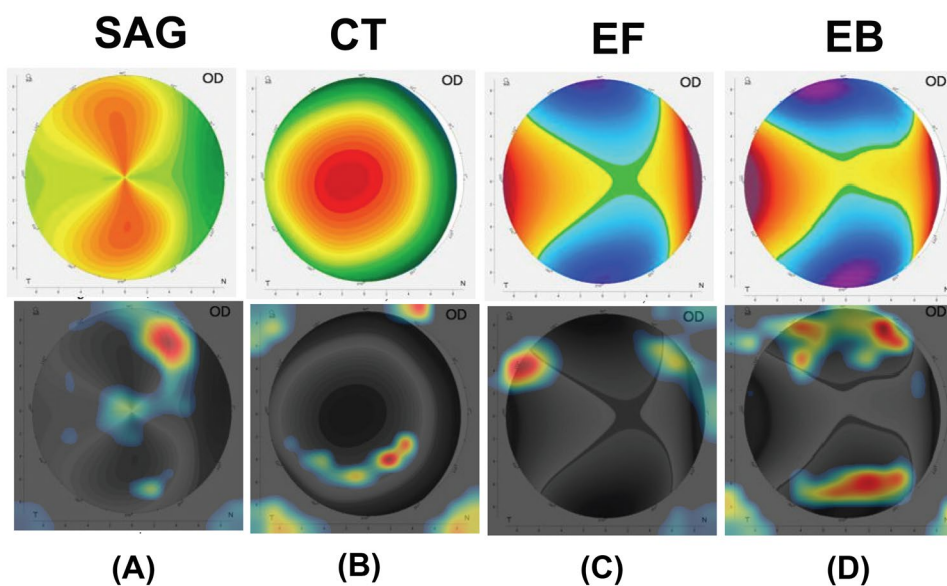
Fig. 8 Grad-CAM plot of *NOR* case for the 4 refractive maps: (**A**) SAG, (**B**) CT, (**C**) EF, and (**D**) EB

(A) (B) (C) (D)

when utilizing different maps, but still the accuracy is above 94.8% when fusing the outputs.

SqN and AN are trained on wide range of many general natural images, including animals and everyday objects from the ImageNet dataset [25]. The networks learned rich feature representations from these images. The early network layers learned low-level features, such as edges, borders, curves, and colours, from the general images [29]. In contrast, the last layers learned class-specific features of the corneal topographic maps for KCN detection. In our case, this corresponds to features related to the corneal topographic maps for KCN detection. Thus, the networks tested in our study can detect general shapes and patterns in the images of the corneal topographic maps such as the bowtie shape in the SAG map, circles and curves in the CT map, and curves and islands shapes in EF and EB maps (Fig. 1). This is potentially the reason of the good performance of the networks tested in this study on the corneal topographic maps even though the original pre-trained networks were not trained on them.

Fig. 9 Grad-CAM plot of *KCN* case for the 4 refractive maps: (**A**) SAG, (**B**) CT, (**C**) EF, and (**D**) EB
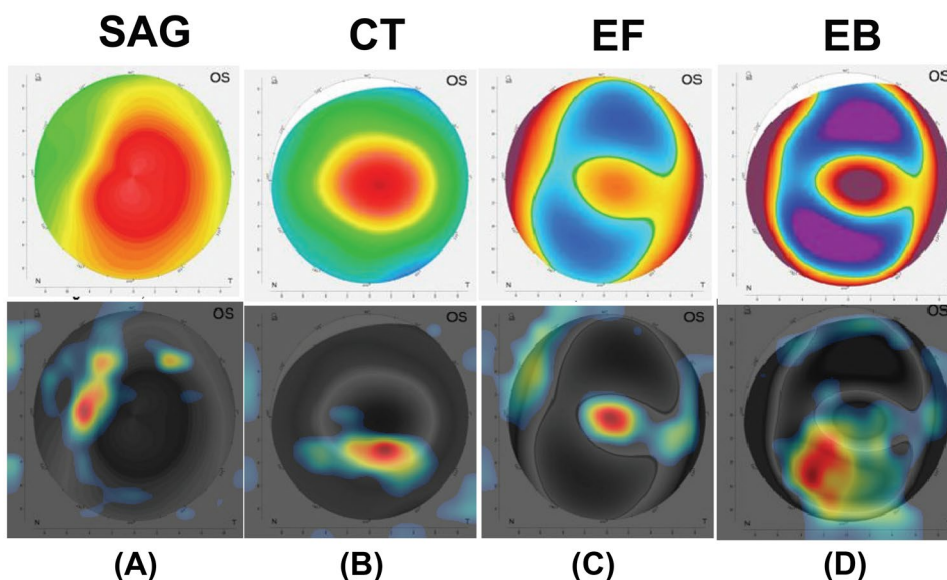

(A) (B) (C) (D)

Table 8 Comparison to the previous literature on KCN detection from corneal topographic images

Study	Device name	Dataset/number of maps	Evaluation method	Network used	Accuracy
Lavric and Valentin [24]	Synthetic maps	SyntEyes and SyntEyes KTC models/1 map	Training/validation/testing	KeratoDetect	99.3%
Zéboulon et al.[42]	Orbscan	3000 cases/4 maps	Ten-fold CV	CNN	99.3%
Kamiya et al. [31]	CASIA	543 cases/6 maps	Five-fold CV	ResNet-18	99%
Kuo et al. [44]	Tomy TMS-4 Topographer	354 cases/1 map	Training, test, and subclinical test	VGG16 InceptionV3 ResNet152	931% 93.1% 95.8%
This study	Pentacam	534 cases/4 maps	Training/validation/testing	EDTL with AlexNet and product fusion	98.3%

CV cross-validation

*SyntEyes and SyntEyes KTC [43] models

Unlike our work where SAG map was one of the best performer maps, it achieved the lowest performance in Abdülhüssein [30], compared to the other three maps, i.e. CT, EF, and EB. Moreover, SAG map was one of the best 3 classifiers when tested with our proposed method using AN (98.3% accuracy for SAG, EB maps, and PI), and also SqN, SfN, and MN which show the importance of this map for KCN detection. This map was not utilized in previous studies [10, 31].

In this study, data was included from both groups, normal and KCN from both right and left eyes, and our trained networks achieved an accuracy of 98.3% for the best three maps tested with the proposed EDTL method. In Issarti et al. [10], 96.56% accuracy was obtained with their proposed CAD system trained with topographic maps from only right eye. Having a trained network to detect KCN versus normal from topographic maps in both eye sides is a more complicated and challenging classification task than having maps that belong to one eye side, as KCN characteristics and features in the four topographic maps for one side are different from that of the other side, such as skewing angle of the bowtie shape.

Table 8 displays a comparison with the previous literature for KCN detection with deep learning applied to corneal topographic maps. It is noteworthy to mention that the comparison is made for illustrative purposes, since each study utilized different image database and different type of device to obtain the corneal maps.

The sample size in the current study was limited to 534 cases (with augmentation) of both left and right eyes. This number is still small despite being larger than previous studies for KCN detection with machine learning (40 right eye cases only) [11]. The proposed EDTL was evaluated with SqN, AN, SfN, and MN. Other DL network designs such as Capsule network proposed in Sabour et al. [45] and VGG [44] can also be explored in a future study.

Future work will focus on recruiting more sample size for different grades of KCN (mild, moderate, and severe KCN) and classify them with pretrained transfer learning networks. Graphical User Interface (GUI)-based standalone application can be developed that accepts an input of the four corneal topographic maps and PI. Then, decide if the case has a KCN or not, based on the proposed method, with a confidence level calculated from averaging the output probabilities of all maps/indices.

Conclusion

This study proposed EDTL method based on the corneal topographic maps and Pentacam index classifiers to give a decision supporting ophthalmologist's diagnosis and raising the confidence of the KCN detection. Four pretrained transfer learning networks were investigated, SqN, AN, SqN, and MN. An improved accuracy of 98.3% was obtained with AN with only two refractive maps: SAG and EB and PI classifiers. It should be noted that the accuracy of the proposed EDTL method was achieved without the aid of any other imaging modalities, slit lamp, or OCT. The outcome of this study can help in potential deployment of clinical CAD system for KCN detection that aids ophthalmologists for better KCN diagnosis and management.

Declarations

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained originally by Al-Amal ophthalmic centre at the time of data collection from all participants included in the study.

Conflict of Interest The authors declare no competing interests.

References

1. Yousefi S, Yousefi E, Takahashi H, Hayashi T, Tampo H, Inoda S, et al. Keratoconus severity identification using unsupervised machine learning. PLoS One. Public Library of Science San Francisco, CA USA; 2018;13:e0205998.

2. Accardo PA, Pensiero S. Neural network-based system for early keratoconus detection from corneal topography. J Biomed Inform Elsevier. 2002;35:151–9.

3. Belin MW, Ambrósio Jr R. Scheimpflug imaging for keratoconus and ectatic disease. Indian J Ophthalmol. Wolters Kluwer--Medknow Publications; 2013;61:401.

4. Castillo JH, Hanna R, Berkowitz E, Tiosano B. Wavefront analysis for keratoconus. Int J Keratoconus Ectatic Corneal Dis. Jaypee Brothers Medical Publishers Ltd.; 2014;3:76.

5. Karamichos D, Hjortdal J. Keratoconus: tissue engineering and biomaterials. J Funct Biomater. Multidisciplinary Digital Publishing Institute; 2014;5:111–34.

6. Gatinel D, Malet J, Hoang-Xuan T, Azar DT. Corneal elevation topography: best fit sphere, elevation distance, asphericity, toricity and clinical implications. Cornea. NIH Public Access; 2011;30:508.

7. Smolek MK, Klyce SD. Current keratoconus detection methods compared with a neural network approach. Invest Ophthalmol Vis Sci. The Association for Research in Vision and Ophthalmology; 1997;38:2290–9.

8. Aslani F, Khorrami-Nejad M, Amiri MA, Hashemian H, Askarizadeh F, Khosravi B. Characteristics of posterior corneal astigmatism in different stages of keratoconus. J Ophthalmic Vis Res. Wolters Kluwer--Medknow Publications; 2018;13:3.

9. Yoo TK, Ryu IH, Lee G, Kim Y, Kim JK, Lee IS, et al. Adopting machine learning to automatically identify candidate patients for corneal refractive surgery. NPJ Digit Med Nature Publishing Group. 2019;2:1–9.

10. Issarti I, Consejo A, Koppen C, Rozema JJ, Jiménez-garcía M, Hershko S. Computer aided diagnosis for suspect keratoconus detection. Comput Biol Med [Internet]. Elsevier Ltd; 2019;109:33–42. Available from: https://doi.org/10.1016/j.compbiomed.2019.04.024

11. Mosa ZM, Ghaeb NH, Ali AH. Detecting keratoconus by using SVM and decision tree classifiers with the aid of image processing. Baghdad Sci J. 2019;16:4.

12. Sinjab MM. Step by Step®: reading Pentacam topography: basics and case study series. Jaypee Brothers Medical Publishers; 2015.

13. Maeda N, Klyce SD, Smolek MK. Neural network classification of corneal topography preliminary demonstration. Invest Ophthalmol Vis Sci. 1995;36.

14. Rav D, Wong C, Deligianni F, Berthelot M, Andreu-perez J, Lo B. Deep learning for health informatics. IEEE J Biomed Heal Informatics. 2017;21:4–21.

15. Alzubaidi L, Fadhel MA, Oleiwi SR, Al-Shamma O, Zhang J. DFU_QUTNet: diabetic foot ulcer classification using novel deep convolutional neural network. Multimed Tools Appl. Springer; 2019;1–23.

16. Chen H, Ni D, Qin J, Li S, Yang X, Wang T, et al. Standard plane localization in fetal ultrasound via domain transferred deep neural networks. IEEE J Biomed Heal informatics IEEE. 2015;19:1627–36.

17. Norouzifard M, Nemati A, GholamHosseini H, Klette R, Nouri-Mahdavi K, Automated YS, deeptransfer learning: proposal of a system for clinical testing. GDU. Int Conf Image Vis Comput New Zeal. IEEE. 2018;2018:1–6.

18. Lopes BT, Eliasy A, Ambrosio R Jr. Artificial intelligence in corneal diagnosis: where are we? Curr Ophthalmol Rep Current Ophthalmology Reports. 2019;7:204–11.

19. Shu D, Ting W, Pasquale LR, Peng L, Campbell JP, Lee AY, et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019;103:167–75.

20. Hallak JA, Azar DT. The AI revolution and how to prepare for it. Transl Vis Sci Technol. The Association for Research in Vision and Ophthalmology; 2020;9:16.

21. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med Nature Publishing Group. 2018;1:1–8.

22. Varma R. How AI benefits patients and physicians. Ophthalmol Times [Internet]. 2018; Available from: https://www.ophthalmologytimes.com/view/how-ai-benefits-patients-and-physicians

23. Ting DSJ, Foo VHX, Yang LWY, Sia JT, Ang M, Lin H, et al. Artificial intelligence for anterior segment diseases: emerging applications in ophthalmology. Br J Ophthalmol: BMJ Publishing Group Ltd; 2020.

24. Lavric A, Valentin P. KeratoDetect: keratoconus detection algorithm using convolutional neural networks. Comput Intell Neurosci. 2019;2019:1–9.

25. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst. 2012. p. 1097–105.

26. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. arXiv Prepr arXiv160207360. 2016;

27. Zhang X, Zhou X, Lin M, Sun J. Shufflenet: an extremely efficient convolutional neural network for mobile devices. Proc IEEE Conf Comput Vis pattern Recognit. 2018. p. 6848–56.

28. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. Mobilenetv2: inverted residuals and linear bottlenecks. Proc IEEE Conf Comput Vis pattern Recognit. 2018. p. 4510–20.

29. Beale MH, Hagan MT, Demuth HB. Deep Learning Toolbox™ user's guide. MathWorks Inc. 2019;1–20.

30. Abdülhüssein NS. Building smart algorithm to extract features of topographic images of a human eye. MSc thesis, Üniversitesi Fen Bilimleri Enstitüsü, Aksaray; 2018.

31. Kamiya K, Ayatsuka Y, Kato Y, Fujimura F, Takahashi M, Shoji N, et al. Keratoconus detection using deep learning of colour-coded maps with anterior segment optical coherence tomography: a diagnostic accuracy study. BMJ Open. 2019;9:1–7.

32. Polikar R. Ensemble based systems in decision making. IEEE Circuits Syst Mag. 2006;6:21–44.

33. Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Ophthalmol Vis Sci. The Association for Research in Vision and Ophthalmology; 2016;57:5200–6.

34. Qian X, Patton EW, Swaney J, Xing Q, Zeng T. Machine learning on cataracts classification using SqueezeNet. 4th Int Conf Univers Village. IEEE; 2018. p. 1–3.

35. Association WM. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. Jama American Medical Association. 2013;310:2191–4.

36. Ueki R, Maeda N, Fuchihata M, Koh S, Kitaoka T, Nishida K. Differentiation of forme fruste keratoconus from normal cornea using parameters of corneal tomography, aberration, and biomechanics. Invest Ophthalmol Vis Sci. The Association for Research in Vision and Ophthalmology; 2014;55:3705.

37. Johnson JM, Khoshgoftaar TM. Survey on deep learning with class imbalance. J Big Data [Internet]. Springer International Publishing; 2019;6. Available from: https://doi.org/10.1186/s40537-019-0192-5

38. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321–57.

39. Witten I, Frank E. Data mining: practical machine learning tools and techniques. 2nd ed. San Francisco: Morgan Kaufmann; 2005.

40. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: visual explanations from deep networks via gradient-based localization. Proc IEEE Int Conf Comput Vis. 2017. p. 618–26.

41. Panwar H, Gupta PK, Siddiqui MK, Morales-Menendez R, Bhardwaj P, Singh V. A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. Chaos, Solitons & Fractals. Elsevier; 2020;140:110190.

42. Zéboulon P, Debellemanière G, Bouvet M, Gatinel D. Corneal topography raw data classification using a convolutional neural network. Am J Ophthalmol: Elsevier; 2020.

43. Rozema JJ, Rodriguez P, Ruiz Hidalgo I, Navarro R, Tassignon M, Koppen C. SyntEyes KTC: higher order statistical eye model for developing keratoconus. Ophthalmic Physiol Opt Wiley Online Library. 2017;37:358–65.

44. Kuo B-I, Chang W-Y, Liao T-S, Liu F-Y, Liu H-Y, Chu H-S, et al. Keratoconus screening based on deep learning approach of corneal topography. Transl Vis Sci Technol. The Association for Research in Vision and Ophthalmology; 2020;9:53.

45. Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. Adv Neural Inf Process Syst. 2017. p. 3856–66.