

SYSTEM AND METHOD FOR CLASSIFYING EYE IMAGES

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims priority to U.S. Provisional Application No. 62/810,440 filed February 26, 2019, the contents of which are incorporated by reference.

TECHNICAL FIELD

[0002] The present disclosure relates generally to ophthalmic diagnostics, and more specifically to a system and method for classifying eye images to assist in automated retinal health screening that uses a deep learning technique.

BACKGROUND

[0003] As of 2017, an estimated 425 million people had diabetes worldwide, with type 2 diabetes making up about 90% of the cases. This represents 8.8% of the adult population, with equal rates in both women and men. Trends suggest that rates will continue to rise along with related ailments. Diabetic retinopathy, also known as diabetic eye disease, is a medical condition in which damage occurs to the retina due to diabetes. It is a leading cause of blindness in developed countries.

[0004] Diabetic retinopathy is the result of damage to the small blood vessels and neurons of the retina. Diabetic retinopathy affects up to 80 percent of those who have had diabetes for twenty years or more. At least 90% of new cases could be reduced with proper treatment and monitoring of the eyes. The longer a person has diabetes, the higher his or her chances of developing diabetic retinopathy. Each year in the United States, diabetic retinopathy accounts for 12% of all new cases of blindness. It is also the leading cause of blindness in people aged 20 to 64. Early diagnosis of diabetic retinopathy is critical as it generally asymptomatic in early stages.

[0005] Diabetes also increases the risk of having glaucoma. The most common type, open-angle glaucoma, is believed to result from the slow exit of aqueous humor through the trabecular meshwork. In closed-angle glaucoma the iris blocks the trabecular meshwork.

[0006] Another common disease of the eye is age-related macular degeneration (AMD). It can lead to blurred or no vision in the center of the visual field. Macular degeneration typically occurs in older people. Genetic factors and smoking also play a role. It is due to damage to the macula of the retina.

[0007] In ophthalmology, these are the three main eye diseases that ultimately lead to blindness: Age-Related Macular Degeneration (AMD), Diabetic Retinopathy (DR) and glaucoma. With limited options for treatment, early diagnosis is critical to prevent or reduce progression of the diseases.

[0008] Machine learning and deep learning are the two emerging techniques that are used for screening eye images to assist in a diagnosis of retinal diseases. In conventional and current computer methods using machine learning algorithms, the user must define features of the images for the method to identify for automatic diagnosis of the diseases. Convolutional neural networks (CNN) are feed forward neural networks, inspired by biological visual systems. When applied to image recognition, these networks allow for a learning of the intermediate representations of objects in the images which are smaller and which can be generalized for similar objects. A CNN can be made up of several convolution layers, including pooling layers and which are generally followed by a multilayer perceptron-type classifier, the output of one layer being connected to the input of the next. Deep learning CNN models have been developed for their potential to screen, classify and diagnose retinal diseases.

[0009] For example, International PCT Publication WO2018045363 A1 relates to a screening method for automated detection of vision-degenerative diseases from colour fundus images. This method comprises obtaining a dataset of fundus images, using a custom deep learning network to process the dataset of fundus images, and providing, as an output, a function for use in diagnosing a vision-degenerative disease. A computing device disclosed in this PCT application comprises memory storing a representation of the function produced by the first method, and one or more processors configured to use the function to assist in the diagnosis of the vision-degenerative disease. A second method is for determining a likelihood that a patient's eye has a vision-degenerative disease, the method comprising obtaining a fundus

image of the patient's eye, processing the fundus image using the function obtained from the first method, and, based on the processing of the fundus image, providing an indication of the likelihood that the patient's eye has the vision-degenerative disease. However, this disclosure utilizes machine learning techniques as opposed to deep learning methods and fails to refer to an abnormal classification of retinal defects.

[0010] US patent No. US 9008391 B1 relates to systems and methods for processing retinal images for screening of diseases or abnormalities. The systems and methods aid in screening, diagnosis and/or monitoring of medical conditions. The design includes a system and method for accessing retinal images related to a patient, each retinal image including a plurality of pixels. For each of the retinal images, the system designates a first set of the plurality of pixels as active pixels including interesting retinal image regions and computes a first vector of numbers comprising pixel-level descriptors for each of the active pixels. The system computes a second vector of numbers for each of the retinal images and provides a second classification using supervised learning. However, this disclosure requires the extraction of various computed descriptors and features with multiple pre-processing techniques applied on the input image such as median filtered image, normalization, thresholding and masking.

[0011] US patent application US 20150110348 A1 relates to systems and methods for automated detection of regions of interest in retinal images. Systems and methods aid in screening, diagnosis and/or monitoring of medical conditions. The systems and methods may allow, for example, for automated identification and localization of lesions and other anatomical structures from medical data obtained from medical imaging devices, computation of image-based biomarkers including quantification of dynamics of lesions, and/or integration with telemedicine services, programs, or software. However, this disclosure requires the extraction of various computed descriptors and features with multiple pre-processing techniques applied on the input image such as median filtered image, normalization, thresholding and masking.

[0012] CN patent application CN 108095683 A relates to a method and device for processing eye fundus images based on deep learning. The method for processing eye fundus images based on deep learning comprises the following steps: obtaining

multiple original sample eye fundus images and corresponding lesion grade parameters; performing image pre-processing operation and image data enhancing operation on the multiple original sample eye fundus images in order to obtain multiple training sample eye fundus images; performing deep-network training on the multiple original sample eye fundus images and the corresponding lesion grade parameters in order to obtain an eye fundus lesion grade recognition model; and inputting test eye fundus images into the eye fundus lesion grade recognition model in order to obtain lesion grade recognition results. However, this disclosure only identifies diabetic retinopathy (DR) class as abnormal and has multiple pre-processing techniques applied on the input image such as histogram equalization, colour and brightness normalization, contrast and uniformity of exposure.

[0013] The above conventional models and methodologies are used for classifying the images but do not provide CNN models with higher accuracy and sensitivity. Further, these conventional methodologies use many parameters for analysis and are trained using smaller number of image sources with a limited variation in image size, image resolution and image quality.

[0014] A need, therefore, exists for an improved robust CNN model, system and methodology that provides higher accuracy, sensitivity and specificity using deep learning technique to automatically classify eye images as normal or abnormal that may be one of Age-Related Macular Degeneration (AMD), Diabetic Retinopathy (DR) and glaucoma.

SUMMARY

[0015] The following summary is provided to facilitate an understanding of some of the innovative features unique to the disclosed embodiment and is not intended to be a full description. A full appreciation of the various aspects of the embodiments disclosed herein can be gained by taking into consideration the entire specification, claims, drawings, and abstract as a whole.

[0016] Embodiments include a system and method that automatically classifies eye images of Age-Related Macular Degeneration (AMD), Diabetic Retinopathy (DR) and glaucoma subjects as abnormal class and eye images of healthy subjects as normal

class using a trained deep learning Convolutional Neural Network (CNN).

[0017] In one embodiment, there is provided a system for classifying eye images, that can comprise: an input for receiving at least one eye image; a storage medium configured to store instructions defining a deep learning convolutional neural network (CNN) trained on a dataset of a plurality of eye images, wherein the CNN is trained using an ADAM optimization algorithm and blindfold cross-validation; a processor for execution of the deep learning convolutional neural network to automatically classify the at least one eye image as either normal class or abnormal class, wherein the normal class indicates the absence of an eye disease or condition and abnormal class indicates the presence of an eye disease or condition; and an output to communicate the class of the at least one image to a user.

[0018] In one embodiment, the plurality of images is comprised of at least 2,500 eye images.

[0019] In one embodiment, the eye disease or condition can be age-related macular degeneration (AMD), diabetic retinopathy (DR) and/or glaucoma.

[0020] In one embodiment, the CNN can comprise a convolution layer, a pooling layer and a fully connected layer.

[0021] In one embodiment, the CNN can comprise a series of convolution and pooling layers.

[0022] In one embodiment, the convolution layer can comprise a first, second, third and fourth convolution layer.

[0023] In one embodiment, the first convolution layer can comprise eight feature maps (3x3) ReLu, eight feature maps (5x5) ReLu, and eight feature maps (7x7) ReLu concatenated together.

[0024] In one embodiment, the second convolution layer can comprise sixteen

feature maps (3x3) ReLu.

[0025] In one embodiment, the third convolution layer can comprise sixteen feature maps (3x3) ReLu.

[0026] In one embodiment, the fourth convolution layer can comprise thirty-two feature maps (3x3) ReLu.

[0027] In one embodiment, the pooling layer can comprise a first, second and third pooling layer.

[0028] In one embodiment, the first, second and third pooling layers can reduce the spatial dimension of the input images to half the input size.

[0029] In one embodiment, the CNN can comprise a flattening step to convert the three-dimensional images to two-dimensional images.

[0030] In one embodiment, the fully connected layer can comprise a first, second and third fully connected layer.

[0031] In another embodiment, there is provided a computer-implemented method for classifying eye images that can comprise: inputting at least one eye image into the system disclosed herein for processing; and obtaining an image classification result of either normal class or abnormal class based on the processing of the at least one eye image.

[0032] In another embodiment, there is provided a computer readable storage medium, storing non-transitory instructions for controlling an automated processor to perform the computer-implemented method disclosed herein.

[0033] In another embodiment, there is provided a method of training a convolutional neural network to classify a fundus image as one of normal and abnormal, the method comprising steps of: obtaining a first plurality of fundus images with features of an eye condition or disease; obtaining a second plurality of fundus images with normal

features; processing the first plurality of fundus and the second plurality of fundus through the convolutional neural network to identify features in the first plurality of fundus images as characteristic of an ailment; wherein the convolutional neural network is comprised of a convolution layer, a pooling layer and a fully connected layer, wherein the CNN is trained using an ADAM optimization algorithm and validated by blindfold cross-validation.

[0034] In one embodiment, the eye condition or disease is at least one of age-related macular degeneration (AMD), diabetic retinopathy (DR) and/or glaucoma.

[0035] In one embodiment, the convolution layer comprises a first, second, third and fourth convolution layer.

[0036] In one embodiment, the first convolution layer comprises eight feature maps (3x3) ReLu, eight feature maps (5x5) ReLu, and eight feature maps (7x7) ReLu concatenated together.

[0037] In one embodiment, the second convolution layer comprises sixteen feature maps (3x3) ReLu.

[0038] In one embodiment, the third convolution layer comprises sixteen feature maps (3x3) ReLu.

[0039] In one embodiment, the fourth convolution layer comprises thirty-two feature maps (3x3) ReLu.

[0040] In one embodiment, the pooling layer comprises a first, second and third pooling layer.

[0041] In one embodiment, the first, second and third pooling layers reduce the spatial dimension of the input images to half the input size.

[0042] In one embodiment, the convolutional neural network comprises a flattening step to convert the three-dimensional output to two-dimensional output.

[0043] In one embodiment, the fully connected layer comprises a first, second and third fully connected layer.

BRIEF DESCRIPTION OF THE DRAWINGS

[0044] The summary above, as well as the following detailed description of illustrative embodiments, is better understood when read in conjunction with the appended drawings. For the purpose of illustrating the present disclosure, exemplary constructions of the disclosure are shown in the drawings. However, the disclosure is not limited to specific systems, methods and instrumentalities disclosed herein.

[0045] **FIG. 1** is an illustration of a block diagram of a system for automated classification of eye images using deep learning CNN model, in accordance with the disclosed embodiments.

[0046] **FIG. 2** is an illustration of a flow chart pertaining to the architecture of the deep learning CNN model and process for eye image classification, in accordance with the disclosed embodiments.

[0047] **FIG. 3** is an illustration of a graph showing accuracies versus epochs for training and validation of CNN model in accordance with the disclosed embodiments.

DEFINITIONS

[0048] Reference in this specification to "one embodiment/aspect" or "an embodiment/aspect" means that a particular feature, structure, or characteristic described in connection with the embodiment/aspect is included in at least one embodiment/aspect of the disclosure. The use of the phrase "in one embodiment/aspect" or "in another embodiment/aspect" in various places in the specification are not necessarily all referring to the same embodiment/aspect, nor are separate or alternative embodiments/aspects mutually exclusive of other embodiments/aspects. Moreover, various features are described which may be

exhibited by some embodiments/aspects and not by others. Similarly, various requirements are described which may be requirements for some embodiments/aspects but not other embodiments/aspects. Embodiment and aspect can in certain instances be used interchangeably.

[0049] The terms used in this specification generally have their ordinary meanings in the art, within the context of the disclosure, and in the specific context where each term is used. Certain terms that are used to describe the disclosure are discussed below, or elsewhere in the specification, to provide additional guidance to the practitioner regarding the description of the disclosure. It will be appreciated that the same thing can be said in more than one way.

[0050] Consequently, alternative language and synonyms may be used for any one or more of the terms discussed herein. Nor is any special significance to be placed upon whether or not a term is elaborated or discussed herein. Synonyms for certain terms are provided. A recital of one or more synonyms does not exclude the use of other synonyms. The use of examples anywhere in this specification including examples of any terms discussed herein is illustrative only, and is not intended to further limit the scope and meaning of the disclosure or of any exemplified term. Likewise, the disclosure is not limited to various embodiments given in this specification.

[0051] The term “ADAM optimization algorithm” or “Adam” refers to a replacement optimization algorithm for stochastic gradient descent for training deep learning models. Adam combines the best properties of the AdaGrad and RMSProp algorithms to provide an optimization algorithm that can handle sparse gradients on noisy problems

[0052] The term “ailment” refers to a disease, illness or medical condition. An ailment can be, for example, an ophthalmic ailment such as age-Related Macular degeneration (AMD), diabetic retinopathy (DR) and glaucoma.

[0053] The term “deep learning” refers to a refinement of artificial neural network (“ANN”), consisting of more than one hidden layer that permits higher levels of

abstraction and improved predictions from data. A “deep learning model” refers to classification models that can include deep learning neural network models.

[0054] The term “convolutional neural network” (“CNN”) is as conventionally used in the technical field and generally refers to powerful tools for computer vision tasks, whereby deep learning CNNs can be formulated to automatically learn mid-level and high-level abstractions obtained from raw data such as images. A convolution layer can contain one or more convolution kernels, which each have an input matrix, which can be the same, but which have different coefficients corresponding to different filters. Each convolution kernel in a layer produces a different output map such that the output neurons are different for each kernel. The convolutional networks can also include local or global “pooling” layers which combine the neuron group outputs of one or more output maps. The combination of the outputs can consist, for example, in taking the maximum or average value of the outputs of the group of neurons, for the corresponding output, on the output map of the “pooling” layer. The “pooling” layers make it possible to reduce the size of the output maps from one layer to the other in the network, while improving the performance levels thereof by making it more tolerant to small deformations or translations in the input data.

[0055] The term “big data” refers to large data sets that can be analyzed computationally to reveal patterns, trends, and associations, especially relating to human physiological responses, human behavior and interactions. This can include, for example, a databases of electronic health records from a hospital or clinic.

[0056] The term “computer learning” or “machine learning” refers to an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

[0057] The term “fundus” or “fundus of the eye” refers to the interior surface of the eye opposite the lens and includes the retina, optic disc, macula, fovea, and posterior pole. The fundus can be examined by ophthalmoscopy and/or fundus photography.

[0058] The term “module” refers to a self-contained unit, such as an assembly of electronic components and associated wiring or a segment of computer software, which itself performs a defined task and can be linked with other such units to form a larger system.

[0059] The term “processor” includes any suitable hardware and / or software system, mechanism or component that processes data, signals, or other information. The processor may include a general-purpose central processing unit, a multi-processing unit, a dedicated circuit that implements a specific function, or other system. The process need not be limited to geographic locations or have time limits. For example, the processor can perform functions in “real time”, “offline”, “batch mode”, and the like. Some of the processing can be performed at different times and places by another (or the same) processing system. Examples of processing systems can include servers, clients, end-user devices, routers, switches, network storage, and the like. The computer can be any processor that communicates with memory. The memory is any suitable processor readable storage medium, such as random access memory (RAM), read only memory (ROM), magnetic or optical disk, or other tangible medium, suitable for storing instructions to be executed by the processor.

[0060] The term “features” refers to the hidden signatures present in images.

[0061] The term “ReLU” refers to Rectified Linear Unit that is an activation function used avoid gradient exploding during training, whereby Gradient exploding refers to a model that is not trained or converging.

[0062] The patients, individuals or subjects of the invention system and method are, in addition to humans, veterinary subjects in the context of this disclosure herein. Such subjects include livestock and pets as well as sports animals such as horses, greyhounds, and the like.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0063] The particular configurations discussed in the following description are non-

limiting examples that can be varied and are cited merely to illustrate at least one embodiment and are not intended to limit the scope thereof.

[0064] While the invention is primarily described for monitoring and predicting an eye ailment such as age-related macular degeneration (AMD), diabetic retinopathy (DR) and glaucoma, it is understood that the invention is not so limited and can be used to assist with other endeavors that require analyzing fundus images, monitoring a person's ophthalmic health and/or diagnosing a person for vision disorders.

[0065] Disclosed herein are a system and method that provides a user, such as a clinician or ophthalmologist, with assistance in determining the likelihood of an individual having normal eye condition or disease. The system and method can use a trained deep learning convolutional neural network (CNN) to automatically classify and distinguish healthy (normal class) and pathological (abnormal class) eye images from one another. The system and method can also distinguish between the classified pathological (abnormal class) eye images as one of Age-Related Macular Degeneration (AMD), Diabetic Retinopathy (DR) or glaucoma.

[0066] The disclosure herein provides a highly discriminative system and method for distinguishing healthy eye images from pathological eye images using a trained deep learning convolutional neural network (CNN). The system and method can accurately and sensitively discriminate minute features of eye conditions or diseases. In particular, the system and method can detect symptomatic pathologies from an eye image and can be implemented on any type of device with computational power, such as a laptop or smartphone.

[0067] Consequently, the disclosure provides an all-purpose solution for eye condition or disease detection, and the classification results attained indicate high accuracy, sensitivity and specificity of the disclosed system and method in accurately and efficiently determining the likelihood of an individual having an eye condition or disease.

[0068] Accordingly, in one embodiment the system and method can automatically

classify an eye image as either normal class or abnormal class. In particular, the system and method disclosed herein can automatically classify an abnormal class eye image as indicating an individual has or is at risk of having age-related macular degeneration (AMD), diabetic retinopathy (DR) and/or glaucoma.

[0069] In one embodiment, the system and method disclosed herein can automatically classify age-related macular degeneration (AMD), diabetic retinopathy (DR) and/or glaucoma in all pathological/developmental stages of the indicated eye condition or disease as abnormal class. Thus, the system and method disclosed herein is capable of identifying and indicating if an individual has (or is at risk of having) age-related macular degeneration (AMD), diabetic retinopathy (DR) and/or glaucoma in all pathological stages. In particular, age-related macular degeneration (AMD) can be discriminated into early AMD, intermediate AMD, late AMD and Dry AMD; Diabetic retinopathy can be discriminated into mild non-proliferative diabetic retinopathy, moderate non-proliferative diabetic retinopathy, severe non-proliferative diabetic retinopathy, and proliferative diabetic retinopathy; and Glaucoma can be discriminated into primary open-angle glaucoma, primary angle-closure glaucoma, secondary glaucoma and developmental glaucoma.

[0070] In some embodiments, the system and method disclosed herein reaches at least 96% accuracy, at least 97% sensitivity and at least 92% specificity, thus demonstrating high clinical applicability for automated classification of eye images and detection of eye conditions or disease. In one embodiment, the system and method can attain an accuracy of 96.31%, a sensitivity of 97.96% and a specificity of 92.67%. In particular, the accuracy (ACC), sensitivity (SEN) and specificity (SPE) measurements can be determined in relation to the classification results of the input images as follows:

$$SEN = \frac{T_{+ve}}{T_{+ve} + F_{-ve}} \quad (1)$$

$$SPE = \frac{T_{-ve}}{T_{-ve} + F_{+ve}} \quad (2)$$

$$ACC = \frac{T_{+ve} + T_{-ve}}{T_{+ve} + T_{-ve} + F_{+ve} + F_{-ve}} \quad (3)$$

wherein, T_{+ve} = true positive, F_{+ve} = false positive, F_{-ve} = false negative and T_{-ve} = true negative.

[0071] As will be appreciated, the system and method described herein can be implemented on a programmable computer using a combination of both hardware and software. Various aspects can be implemented on programmable computers, each computer including a one or more input unit, a data storage medium, a hardware processor and an output unit or communication interface. It should be appreciated that the use of terms such as servers, services, units, modules, interfaces, portals, platforms, or other systems formed from computing devices is deemed to represent one or more computing devices having at least one processor configured to execute software instructions stored on a computer readable storage medium. For example, a server can include one or more computers operating as a web server, database server, or other type of computer server in a manner to fulfil described roles, responsibilities, or functions.

[0072] FIG. 1 is a representative flowchart of the system and operational relationship between both computational hardware and software components. The system **200** can comprise an input unit **204** configured to receive at least one fundus/eye image **202**, a storage medium **206** configured to store information defining the deep learning convolutional neural network (CNN) disclosed herein as software instructions for execution by a hardware computer processor **208**, and an output unit **216** configured to communicate the classification result obtained from the processor **208** to one or more users.

[0073] Further, the computer system disclosed herein can include, but is not limited to additional components. For example, the system can include one or more communication channels or interconnection mechanism such as a bus, controller, or network, that interconnects the components of the system. In various embodiments of the operating system software provides an operating environment for various software's executing in the computer system, and manages different functionalities of the components of the computer system. The communication channel(s) allow communication over a communication medium to various other computing entities. The communication medium provides information such as program instructions, or

other data in a communication media. The communication media can include wired or wireless methodologies implemented with an electrical, optical, RF, infrared, acoustic, microwave, bluetooth or other transmission media.

[0074] Accordingly, in one embodiment there is provided a system for classifying eye images that can comprise an input unit for receiving at least one eye image, a storage medium configured to store instructions defining a deep learning convolutional neural network (CNN), a processor for execution of the deep learning convolutional neural network to automatically classify the at least one eye image, and an output unit configured to communicate the result to a user.

[0075] The system described herein can implement a method for classifying eye images. Accordingly, in another embodiment there is provided a computer-implemented method for classifying eye images that can include the steps of inputting at least one eye image into the system for processing; and subsequently obtaining an image classification result of either "normal class" or "abnormal class" based on the processing of the image.

[0076] In one embodiment, the input unit **204** of the system can be any computational device or image input module capable of receiving eye images from one or more sources, such as an image capturing device or database. The image capturing device can be capable of scanning or imaging the central and peripheral retina, optic disc and macula of an individual. The eye image can be received, for example, by attaching the image capturing device to a computer or mobile device that comprises the input unit **204**. As another example, the eye image can be received from a database.

[0077] In one embodiment, the at least one eye image **202** can be a fundus image. In an embodiment, the system is capable of processing multiple fundus images at a time. In one embodiment, the fundus image is at least of a size of 180×270 pixels, preferably the fundus image is 180×270 pixels. In one embodiment, the at least one eye image **202** is not subjected to any pre-processing including but not limited to median filtering of image, histogram equalization, normalization, thresholding, masking, contrasting and uniformity of exposure.

[0078] The storage medium **206** of the system can store instructions for execution by the processor **208** to receive and process the at least one eye image using a deep learning convolutional neural network (CNN) **210** and subsequently produces a classification result of the at least one eye image as an output.

[0079] In this regard, the CNN and instructions for execution of the CNN can be in the form of a software product. The storage medium and software product stored thereon can include a number of instructions that enable the processor **208** to execute the instructions defining the CNN.

[0080] In one embodiment, the storage medium **206** can be a non-transitory computer-readable medium having computer-readable program code stored thereon, the computer-readable program code can comprise instructions that when executed by the processor **208**, cause the processor to receive and classify the at least one eye image. The processor **208** extracts one or more features from the at least one eye image using the deep learning CNN to classify the at least one eye image. The processor **208** can be, for example, any type of general-purpose microprocessor or microcontroller, a digital signal processing (DSP) processor, an integrated circuit, a field programmable gate array (FPGA), a reconfigurable processor, or any combination thereof. In one embodiment, the processor **208** can be a graphics processing unit (GPU).

[0081] Accordingly, in another embodiment there is provided a computer readable storage medium, storing non-transitory instructions for controlling an automated processor to execute the method and CNN model that can be implemented either on the system disclosed herein or another system configured to execute the instructions defining the CNN model disclosed herein on said storage medium.

[0082] The convolutional neural network (CNN) disclosed herein has been previously trained using a large dataset of eye images representing a variety of eye conditions, such as healthy or pathological, and is processed using deep learning techniques to obtain a classification result as an output. The classification result can be

communicated to a user to assist one or more users to determine the likelihood or risk of an individual having an eye condition or disease.

[0083] The CNN utilises automated feature learning to classify each input eye image. By training a CNN through deep learning techniques, the CNN can be applied to the system and method disclosed herein for accurately and sensitively discriminating symptomatic features of eye conditions or diseases for classification. Accordingly, in one embodiment the CNN disclosed herein is a deep learning CNN.

[0084] Applying a deep learning technique in training the CNN disclosed herein results in self-feature extraction that facilitates in capturing image features automatically rather than such features being pre-selected and pre-determined for a CNN to extract for classification. Specifically, the deep learning CNN disclosed herein automatically learns feature abstraction from the input images. This automatic deep learning CNN model is desirable as it reduces the necessity of the time-consuming hand-crafting of features that would otherwise be required to pre-process the images with application-specific filters or by calculating computable features. As will be appreciated, the training of a CNN cannot be updated, and any change in the algorithm or training data requires re-optimization of the entire network.

[0085] In one embodiment, the deep learning CNN has been previously trained and developed with a dataset of at least 2,500 eye/fundus images. In another embodiment, the deep learning CNN has been previously trained and developed with a dataset of at least 3,000, 3,500, 4,000, 5,000, 7,500 or 10,000 eye/fundus images. These images include a heterogeneous cohort of patients with a multitude of ophthalmic conditions or diseases such as diabetic retinopathy (DR), glaucoma, and age-related macular degeneration (AMD) at different stages and severity of disease development. Each of the images in the dataset is pre-associated with a diagnostic label of "normal" ("healthy") or "abnormal" ("pathological") by qualified clinicians and medical professionals. The dataset of images include many images varying in size, resolution and quality. For example, the dataset can contain a comprehensive set of fundus images from a variety of individuals under a variety of lighting conditions taken with varying camera models. The size of the dataset and wide variety of image size,

resolution and quality can result in a more robust deep learning CNN model.

[0086] Further, the deep learning CNN is preferably trained with the dataset of images using an adaptive learning rate optimization algorithm that has been designed specifically for training deep neural networks (e.g. adaptive moment Estimation, 'ADAM'). The 'ADAM' algorithm is for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments. The method is straightforward to implement, is computationally efficient, has little memory requirements, is invariant to diagonal rescaling of the gradients, and is well suited for problems that are large in terms of data and/or parameters (See, for example, "Adam: A Method for Stochastic Optimization" by Diederik P. Kingma and Jimmy Ba, conference paper at the 3rd ICLR 2015).

[0087] Following training of the deep learning CNN, the classification results and accompanying accuracy is preferably validated with a blindfolded cross-validation technique. In particular, the deep learning CNN is processed with a validation set of images that were not used for training and are a separate distinct dataset of images to the training dataset. The performance of the deep learning CNN using the validation dataset can be compared against the training dataset to determine the accuracy of the deep learning CNN disclosed herein, as exemplified in FIG.3.

[0088] As will be appreciated, the typical and conventional CNN architecture for image processing can include of a series of layers of convolution filters, interspersed with a series of data reduction or pooling layers. The convolution filters or kernels are applied to areas of the input image to detect increasingly more relevant features in the image, for example lines or circles, and then higher order features such as local and global shape and texture, both of which may represent features symptomatic of an eye disease or condition. These convolution filters are learned by the CNN from the training. The output of the CNN is typically one or more probabilities or class labels, which in the context of the present invention is normal or abnormal class.

[0089] The deep learning CNN model disclosed herein has a particular architecture that was trained and developed in order to advantageously attain the clinically relevant

accuracy, sensitivity and selectivity classification results for eye images.

[0090] In this regard, **FIG. 2** illustrates a representative internal architecture **300** of the CNN disclosed herein. Accordingly, the CNN can include at least three main layer types made up of a convolution layer, a pooling layer and a fully connected layer. The convolution and pooling layers can perform feature extraction, whereby the convolution layer detects features symptomatic of the eye condition or disease. The fully connected layers then act as a classifier on top of these features and assign a percent probability for the input image being “abnormal class” or “normal class.” Specifically, the probability can be an output with a sigmoid function and the value ranges from 0 to 1, whereby if the value is less than 0.5 then the probability is stated as “normal” and 0.5 or more then the probability is labelled as “abnormal”. The abnormal class can comprise an ailment such as diabetic retinopathy (DR), glaucoma, and age-related macular degeneration (AMD). Each ailment can be at different stages and progress and/or severity of pathological development.

[0091] In one embodiment, these three main layers can further comprise a series of convolution and pooling layers along with a concatenate step and flattening step.

[0092] In one embodiment, the at least one eye image for analysis is input **302** for processing by the CNN disclosed herein. The at least one eye image can be a fundus image and any size, but is preferably of size 180 x 270 pixels. As will be appreciated, a larger image size is preferred for input into the CNN.

[0093] In one embodiment, the CNN can include a first convolution layer **304**, **306**, **308**, a concatenate layer **310**, a first pooling layer **312**, a second convolution layer **314**, a second pooling layer **316**, a third convolution layer **318**, a third pooling layer **320**, a fourth convolution layer **322**, a flattening layer **324**, a first fully connected layer **326**, a second fully connected layer **328** and a third fully connected layer **330**. Accordingly, in one embodiment the CNN can include twelve layers.

[0094] In one embodiment, the eye image(s) is first processed by the first convolution layer **304**, **306**, **308** with different sized kernels for interpreting the input image and

can produce three differently sized groups of feature maps. The convolution layer can be performed in 3D, whereby the input image is represented as a 3D matrix with dimensions of height, width and depth, with depth corresponding to colour channels (RGB). In this regard, the feature maps or grids herein can have a specific height and width, such as 3x3, 5x5 or 7x7 and by design covers the entire depth of the input image to be 3D. In one embodiment, the first convolution layer **304**, **306**, **308** can produce eight feature maps in each size of 3x3 (**304**), a 5x5 (**306**), and a 7x7 (**308**) Rectified Linear Unit (ReLU). In applying three kernels or filters of sizes (3x3; 5x5 and 7x7), low level (gross) features of various sizes are extracted effectively from the input images.

[0095] The feature maps in the first convolution layer **304**, **306**, **308** can be concatenated together **310** for aggregation, analysis and feature extraction. The features extracted from the first convolution layer **304**, **306**, **308** can then be used for classification by subsequent layers.

[0096] After each convolution layer, a pooling layer can be performed to reduce the dimensionality of the image for classification. The pooling layer enables a reduction of the number of parameters and downsizes each feature map independently, reducing the height and width, but keeping the depth intact. The pooling layer slides a window over the input image and simply takes the max value in a window of a specific size and stride. A type of pooling that can be used is max-pooling that takes the max value in the pooling window and has no parameters.

[0097] Accordingly, in one embodiment the first, second and third pooling layers can follow each convolution layer. In one embodiment, the first, second and third pooling layers can be max-pooling layers. In one embodiment, the size of the pooling window for each pooling layer can be 2x2.

[0098] In one embodiment, the first pooling layer **312** follows the concatenate layer **310** and reduces the spatial dimension of the input images without losing any pertinent information. The number of feature maps remain the same as the concatenate step **310** and the images size is halved to have only the maximum value retained within a selected stride. For example, in an embodiment with an input image size of 180 x 270

pixels, the image will become a size of 90 x 135 pixels while retaining the same number of feature maps.

[0099] In one embodiment, the second convolution layer **314** follows the first pooling layer **312** and can produce sixteen feature maps in a size of 3x3 Rectified Linear Unit (ReLU). The output image size of the second convolution layer **314** can be a size of 88 x 133 pixels. In the second convolution layer **314**, middle level features are extracted effectively from the input images.

[00100] In one embodiment, the second pooling layer **316** follows the second convolution layer **314** and the input image size can be halved while retaining the number of feature maps. For example, in an embodiment with an input image size of 88 x 133 pixels, the image will become a size of 44 x 66 pixels while retaining sixteen feature maps. In one embodiment, the pooling window can be 2 x 2.

[00101] In one embodiment, the third convolution layer **318** follows the second pooling layer **316** and can produce sixteen feature maps in a size of 3x3 Rectified Linear Unit (ReLU). The output image size of the third convolution layer **318** can be a size of 42 x 64 pixels. In the third convolution layer **318**, middle level features are extracted effectively from the input images.

[00102] In one embodiment, the third pooling layer **320** follows the third convolution layer **318** and the input image size can be halved while retaining the number of feature maps. For example, in an embodiment with an input image size of 42 x 64 pixels, the image will become a size of 21 x 32 pixels while retaining sixteen feature maps. In one embodiment, the pooling window can be 2 x 2.

[00103] In one embodiment, the fourth convolution layer **322** follows the third pooling layer **320** and can produce thirty-two feature maps in a size of 3x3 Rectified Linear Unit (ReLU). The output image size of the fourth convolution layer **322** can be a size of 19 x 30 pixels. In the fourth convolution layer **322**, high level features are extracted effectively from the input images.

[00104] In one embodiment, the flattening layer **324** follows the fourth convolution

layer **322** and can convert the three-dimensional (3D) samples to two-dimensional (2D) samples by vectorisation. In particular, for input image into the subsequent fully connected layer the output of the final pooling layer is flattened to a vector to become the input image to the fully connected layer. Flattening is simply arranging the 3D volume of the previous convolution and pooling layers into a 2D representation. In one embodiment, the output image of the flattening step **324** comprises 18240 extracted features from the image representing hidden signatures.

[00105] Following the convolution and pooling layers, one or more fully connected layers can be added to the CNN architecture disclosed herein. In one embodiment, three fully connected layers **326, 328, 330** can be added. These fully connected layers can be used to extract higher order features present in the input image. In one embodiment, the first two fully connected layers **326, 328** can apply ReLu and the third fully connected layer **330** can apply Sigmoid activation.

[00106] In one embodiment, the fully connected layers have been trained with a backpropagation algorithm, after which, half of the nodes are randomly dropped. As is readily appreciated in the art, a dropout can be applied to the CNN by a regularization technique to prevent overfitting during training, whereby at each iteration, a neuron or node is temporarily “dropped” or disabled with probability p . The hyperparameter p can be termed the dropout-rate and typically can be a number around 0.5, corresponding to 50% of the neurons or nodes being dropped out. In one embodiment, the CNN disclosed herein can comprise a dropout rate of 0.5.

[00107] The output of the final fully connected step **330** yields a binary (0, 1) output class whereby the “0” indicates a normal class and “1” indicates an abnormal class. The abnormal class can include age-related macular degeneration (AMD), diabetic retinopathy (DR) or glaucoma, as well as a value for the severity and pathological developmental stages of these eye condition or disease. In addition, the output can be accompanied by a probability indication with a sigmoid function and value ranging from 0 to 1, whereby if the value is less than 0.5 then the probability is stated as “normal” and 0.5 or more then the probability is labelled as “abnormal”.

[00108] The output result from the CNN disclosed herein can be communicated to one

or more users through the output unit **216**. Accordingly, in one embodiment the system can include an output unit **216** configured to communicate the classification result of the at least one image input into the system to the one or more users. In particular, the output unit **216** communicates or displays on a user terminal and communicative interface the CNN classification result of the at least one eye image as either normal class, abnormal class. In one embodiment, the output unit **216** can be a graphical user interface (GUI). In one embodiment, the GUI can have the facility to load and display the input image with a 'Diagnose' button/function to be pressed by the one or more user that will display the output class ('Normal' or 'Abnormal') on a 'text panel' of the GUI.

[00109] In one embodiment of the system, the one or more users of the system can comprise, one or more individuals, one or more patients, one or more ophthalmologists, one or more physicians and any other stakeholder/concerned individual. In one embodiment, the output unit **216** can be configured to communicate the classification results by the CNN in various formats. In one embodiment, the classification result can be communicated to the one or more users automatically via one or more communication channels.

[00110] The classification and output result of the system disclosed herein can be used to assist in determining the likelihood of an individual having an eye condition or disease, not necessarily for full automated diagnosis. In this regard, the system disclosed herein can also be used to indicate or determine the risk of an individual having an eye condition or disease.

[00111] There are a number of advantages that can be derived from the system and the computer implemented method disclosed herein. For example, the dependence on clinicians for diagnostics can be reduced or eliminated, whereby individuals or technicians can use the system or method disclosed herein, to attain, independent predictions on the likelihood of an individual having an eye condition or disease. Further, the system or method disclosed herein can reduce unnecessary workload on clinicians or ophthalmologists in medical settings, rather than spending time trying to diagnose potentially diseased patients out of a demographic of millions, clinicians can attend to patients already determined to be at high-risk for an eye condition or disease,

thereby focusing on providing actual treatment in a time-efficient manner.

WORKING EXAMPLES

[00112] The following non-limiting examples are provided for illustrative purposes only in order to facilitate a more complete understanding of representative embodiments now contemplated. These examples are intended to be a mere subset of all possible contexts in which the components of the formulation may be combined. Thus, these examples should not be construed to limit any of the embodiments described in the present specification, including those pertaining to the type and amounts of components of the formulation and/or methods and uses thereof

Example 1

Training and validation of the CNN model

[00113] FIG. 3 is an illustration of a graph showing accuracies versus epochs for training and validation of the CNN model and system represented by **FIG. 1** and **2**, in accordance with the disclosed embodiments.

[00114] In this representative example, the CNN model was trained, developed and tested using a minimum of 2500 images with Adam optimization algorithm to train the net. A subset of 50% of the images were used to train and develop the deep learning CNN model and the rest of the 50% of the images were used to validate and test the model for accuracy, sensitivity and selectivity.

[00115] An accuracy of 96.31%, sensitivity of 97.96% and specificity of 92.67% was obtained for the representative deep learning CNN model based upon blind fold cross validation.

[00116] The system and methods disclosed herein are advantageous to conventional models as feature extraction and selection techniques are not required through the use of use of a specifically developed twelve-layer deep CNN architecture, that has been trained on a large diverse dataset and proven to be robust by blindfold cross-

validation. Moreover, the system and methods disclosed herein do not require de-noising or other pre-processing of the input images to attain the above significant accuracy, sensitivity and selectivity % in combination.

[00117] Accordingly, the deep learning CNN applied in the system and methods disclosed herein is advantageous in providing a model that is deeper but has fewer parameters. This is indicative of a higher degree of abstraction on the images by the deep learning CNN model, which advantageously results in improved robustness against unknown or unseen images. As a consequence of the fewer parameters less storage space is required on operating computer systems, which lends itself for the deep learning CNN to be suitable for deployment in cloud or mobile electronic devices, for example mobile phone or tablet.

Example 2

Training and validation of the CNN model using Fundus Images

[00118] In this example, a database of fundus images are used to train the system to identify characteristics and differences between fundus images from normal (i.e. healthy) eyes and those with an ailment.

[00119] A total of 451 normal, 254 AMD, 409 DR and 280 glaucoma images were used for training. A total of 451 normal, 408 DR, 254 AMD and 280 glaucoma images were thereafter used for testing. Thus, the net was trained and tested using 2,788 images.

[00120] The developed convolutional neural network (CNN) was tested to automatically classify images of age-related macular degeneration (AMD), diabetic retinopathy (DR) and glaucoma as "abnormal" and images of normal subjects as "normal." The size of the input images used were 180 x 270 pixels. The CNN included three main layers: the convolution, pooling and fully connected layers with a series of convolution and max pooling steps.

[00121] In the convolution layer, the net model used different sized kernels to interpret the input. The different feature maps (3x3, 5x5, 7x7 Relu) were then concatenated

together for analysis. The extracted features were then used for classification in the subsequent layers. In the max-pooling layer, the pooling operation reduced the spatial dimension of the input sample, while retaining pertinent information. The number of feature maps remained the same. The input samples were halved by retaining only the maximum value within a selected stride. Accordingly, the image size became 90x135, while retaining 24 feature maps in this step.

[00122] In the next convolution step, the output size of convolution (88x133) was computed with the filter size of 16. In the next max-pooling step, the input samples were halved to the image size of 44x66, with 16 feature maps. In the next convolution step, the output size of convolution was 42x64, with 16 feature maps. Consequently, in the next max-pooling step, the input samples were halved to the size of 21 x32.

[00123] The next convolution step involved the computation of the output size to 19x30 with 32 feature maps. Subsequently, vectorisation was done in the flattening step, to convert 3-dimensional images to 2-dimensional images. In the fully connected layer, the CNN was trained with a back propagation algorithm, after which, half of the nodes were randomly dropped. This yielded binary (0,1) output classes whereby the images were classified accordingly.

[00124] An ADAM optimization algorithm was used and an accuracy of 96.31% was achieved, with sensitivity and specificity of 97.96% and 92.67% respectively using a blind-fold validation technique. In this example, 50% of the data was used to train and develop the deep learning model and the remaining 50% of the data was used to test the model.

Example 3

Routine Screening for Diabetic Retinopathy

[00125] In this example, a patient presents an elevated risk for developing diabetic retinopathy. Patients at risks include those with Type I diabetes and Type II diabetes. During pregnancy, diabetic retinopathy can also be a problem for women with diabetes. The longer a person has had diabetes, the higher their risk of developing some ocular problem. Screening can also be conducted on larger scales such as at

nursing homes.

[00126] A healthcare provider can use the system and method disclosed herein. This automatic deep learning CNN model reduces the necessity of the manually identifying features in affected fundus images. Further, the CNN model can be trained to further classify abnormal images according to a particular disease, such as diabetic retinopathy.

[00127] The patient can obtain a fundus image of each eye through routine procedures known and used in the field. The patient's fundus images can be uploaded into the system. The processor can classify the fundus images as either "normal" or "abnormal," wherein the normal class indicates the absence of an eye disease or condition and abnormal class indicates the presence of an eye disease or condition. If classified as abnormal, the processor can provide additional information that includes the nature of the ailment and extent of its progression. For example, the system can indicate with a high degree of probability whether the ailment should be classified as mild non-proliferative diabetic retinopathy, moderate non-proliferative diabetic retinopathy, severe non-proliferative diabetic retinopathy or proliferative diabetic retinopathy. This information can be relayed to a healthcare provider and, if necessary, the patient can be scheduled for further diagnostic testing.

OPERATING ENVIRONMENT

[00128] The system is typically comprised of a central server (residing on premise server or on cloud or both), that is connected by a data network to a user's computer. The central server can be comprised of one or more computers connected to one or more mass storage devices. The precise architecture of the central server does not limit the claimed invention. Further, the user's computer can be a laptop or desktop type of personal computer. It can also be a cell phone, smart phone or other handheld device, including a tablet. The precise form factor of the user's computer does not limit the claimed invention. Examples of well-known computing systems, environments, and/or configurations that may be suitable for use with the invention include, but are not limited to, personal computers, server computers, hand held, laptop or mobile computer or communications devices such as cell phones and PDA's,

multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like. In one embodiment, the user's computer is omitted, and instead a separate computing functionality provided that works with the central server. In this case, a user would log into the server from another computer and access the system through a user environment.

[00129] The user environment can be housed in the central server or operatively connected to it. Further, the user can receive from and transmit data to the central server by means of the Internet, whereby the user accesses an account using an Internet web-browser and browser displays an interactive web page operatively connected to the central server. The central server transmits and receives data in response to data and commands transmitted from the browser in response to the customer's actuation of the browser user interface. Some steps of the invention can be performed on the user's computer and interim results transmitted to a server. These interim results can be processed at the server and final results passed back to the user.

[00130] The invention described herein can be executed on a computer system, generally comprised of a central processing unit (CPU) that is operatively connected to a memory device, data input and output circuitry (I/O) and computer data network communication circuitry. Computer code executed by the CPU can take data received by the data communication circuitry and store it in the memory device. In addition, the CPU can take data from the I/O circuitry and store it in the memory device. Further, the CPU can take data from a memory device and output it through the I/O circuitry or the data communication circuitry. The data stored in memory can be further recalled from the memory device, further processed or modified by the CPU in the manner described herein and restored in the same memory device or a different memory device operatively connected to the CPU including by means of the data network circuitry. The memory device can be any kind of data storage circuit or magnetic storage or optical device, including a hard disk, optical disk or solid state memory. The I/O devices can include a display screen, loudspeakers, microphone and a movable

mouse that indicate to the computer the relative location of a cursor position on the display and one or more buttons that can be actuated to indicate a command.

[00131] The computer can display on the display screen operatively connected to the I/O circuitry the appearance of a user interface. Various shapes, text and other graphical forms are displayed on the screen as a result of the computer generating data that causes the pixels comprising the display screen customer's actuation of the browser user interface. Some steps of the invention can be performed on the user's computer and interim results transmitted to a server. These interim results can be processed at the server and final results passed back to the user.

[00132] The invention can also be executed on one or more servers. A server can be a computer comprised of a central processing unit with a mass storage device and a network connection. In addition a server can include multiple of such computers connected together with a data network or other data transfer connection, or, multiple computers on a network with network accessed storage, in a manner that provides such functionality as a group. Practitioners of ordinary skill will recognize that functions that are accomplished on one server can be partitioned and accomplished on multiple servers that are operatively connected by a computer network by means of appropriate inter process communication. In addition, the access of the web services can be by means of an Internet browser accessing a secure or public page or by means of a client program running on a local computer that is connected over a computer network to the server. A data message and data upload or download can be delivered over the Internet using typical protocols, including TCP/IP, HTTP, TCP, UDP, SMTP, RPC, FTP or other kinds of data communication protocols that permit processes running on two remote computers to exchange information by means of digital network communication. As a result a data message can be a data packet transmitted from or received by a computer containing a destination network address, a destination process or application identifier, and data values that can be parsed at the destination computer located at the destination network address by the destination application in order that the relevant data values are extracted and used by the destination application. The precise architecture of the central server does not limit the claimed invention. In addition, the data network can operate with several levels, such

that the user's computer is connected through a fire wall to one server, which routes communications to another server that executes the disclosed methods.

[00133] Computer program logic implementing all or part of the functionality previously described herein can be embodied in various forms, including, but in no way limited to, a source code form, a computer executable form, and various intermediate forms (e.g., forms generated by an assembler, compiler, linker, or locator.) Source code can include a series of computer program instructions implemented in any of various programming languages (e.g., an object code, an assembly language, or a high-level language such as C, C++, C#, Action Script, PHP, EcmaScript, JavaScript, JAVA, or 5 HTML) for use with various operating systems or operating environments. The source code can define and use various data structures and communication messages. The source code can be in a computer executable form (e.g., via an interpreter), or the source code can be converted (e.g., via a translator, assembler, or compiler) into a computer executable form.

[00134] The invention can be described in the general context of computer-executable instructions, such as program modules, being executed by a computer. Generally, program modules include routines, programs, objects, components, data structures, etc., that perform particular tasks or implement particular abstract data types. The computer program and data can be fixed in any form (e.g., source code form, computer executable form, or an intermediate form) either permanently or transitorily in a tangible storage medium, such as a semiconductor memory device (e.g., a RAM, ROM, PROM, EEPROM, or Flash-Programmable RAM), a magnetic memory device (e.g., a diskette or fixed hard disk), an optical memory device (e.g., a CD-ROM or DVD), a PC card (e.g., PCMCIA card), or other memory device. The computer program and data can be fixed in any form in a signal that is transmittable to a computer using any of various communication technologies, including, but in no way limited to, analog technologies, digital technologies, optical technologies, wireless technologies, networking technologies, and internetworking technologies. The computer program and data can be distributed in any form as a removable storage medium with accompanying printed or electronic documentation (e.g., shrink wrapped software or a magnetic tape), preloaded with a computer system (e.g., on system ROM

or fixed disk), or distributed from a server or electronic bulletin board over the communication system (e.g., the Internet or World Wide Web.) It is appreciated that any of the software components of the present invention can, if desired, be implemented in ROM (read-only memory) form. The software components can, generally, be implemented in hardware, if desired, using conventional techniques.

[00135] The invention can also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules can be located in both local and remote computer storage media including memory storage devices. Practitioners of ordinary skill will recognize that the invention can be executed on one or more computer processors that are linked using a data network, including, for example, the Internet. In another embodiment, different steps of the process can be executed by one or more computers and storage devices geographically separated but connected by a data network in a manner so that they operate together to execute the process steps. In one embodiment, a user's computer can run an application that causes the user's computer to transmit a stream of one or more data packets across a data network to a second computer, referred to here as a server. The server, in turn, can be connected to one or more mass data storage devices where the database is stored. The server can execute a program that receives the transmitted packet and interpret the transmitted data packets in order to extract database query information. The server can then execute the remaining steps of the invention by means of accessing the mass storage devices to derive the desired result of the query. Alternatively, the server can transmit the query information to another computer that is connected to the mass storage devices, and that computer can execute the invention to derive the desired result. The result can then be transmitted back to the user's computer by means of another stream of one or more data packets appropriately addressed to the user's computer. In one embodiment, the relational database can be housed in one or more operatively connected servers operatively connected to computer memory, for example, disk drives. In yet another embodiment, the initialization of the relational database can be prepared on the set of servers and the interaction with the user's computer occur at a different place in the overall process.

[00136] It should be noted that the flow diagrams are used herein to demonstrate various aspects of the invention, and should not be construed to limit the invention to any particular logic flow or logic implementation. The described logic can be partitioned into different logic blocks (e.g., programs, modules, functions, or subroutines) without changing the overall results or otherwise departing from the true scope of the invention. Oftentimes, logic elements can be added, modified, omitted, performed in a different order, or implemented using different logic constructs (e.g., logic gates, looping primitives, conditional logic, and other logic constructs) without changing the overall results or otherwise departing from the true scope of the invention.

[00137] It will be appreciated that variations of the above disclosed and other features and functions, or alternatives thereof, may be desirably combined into many other different systems or applications. Also, various presently unforeseen or unanticipated alternatives, modifications, variations or improvements therein may be subsequently made by those skilled in the art which are also intended to be encompassed by the following claims.

[00138] Although embodiments of the current disclosure have been described comprehensively in considerable detail to cover the possible aspects, those skilled in the art would recognize that other versions of the disclosure are also possible.

[00139] While the present invention has been described in terms of particular embodiments and applications, in both summarized and detailed forms, it is not intended that these descriptions in any way limit its scope to any such embodiments and applications, and it will be understood that many substitutions, changes and variations in the described embodiments, applications and details of the method and system illustrated herein and of their operation can be made by those skilled in the art without departing from the spirit of this invention.

What is Claimed is:

1. A system for classifying eye images, comprising:
 - an input for receiving at least one eye image;
 - a storage medium configured to store instructions defining a deep learning convolutional neural network (CNN) trained on a dataset of a plurality of eye images, wherein the CNN has been trained using an ADAM optimization algorithm and validated by blindfold cross-validation;
 - a processor for execution of the deep learning CNN to automatically classify the at least one eye image as either normal class or abnormal class, wherein the normal class indicates the absence of an eye disease or condition and abnormal class indicates the presence of an eye disease or condition; and
 - an output to communicate the class of the at least one image to a user.
2. The system of claim 1, wherein the plurality of images is comprised of at least 2,500 eye images.
3. The system of claim 1, wherein the eye disease or condition is age-related macular degeneration (AMD), diabetic retinopathy (DR) and/or glaucoma.
4. The system of claim 1, wherein the CNN comprises a convolution layer, a pooling layer and a fully connected layer.
5. The system of claim 4, wherein the convolution layer comprises a first, second, third and fourth convolution layer.
6. The system of claim 5, wherein the first convolution layer comprises eight feature maps (3x3) ReLu, eight feature maps (5x5) ReLu, and eight feature maps (7x7) ReLu concatenated together.

7. The system of claim 5, wherein the second convolution layer comprises sixteen feature maps (3x3) ReLu.
8. The system of claim 5, wherein the third convolution layer comprises sixteen feature maps (3x3) ReLu.
9. The system of claim 5, wherein the fourth convolution layer comprises thirty-two feature maps (3x3) ReLu.
10. The system of claim 4, wherein the pooling layer comprises a first, second and third pooling layer.
11. The system of claim 10, wherein the first, second and third pooling layers reduce the spatial dimension of the input images to half the input size.
12. The system of claim 4, wherein the CNN comprises a flattening step to convert the three-dimensional output to two-dimensional output.
13. The system of claim 4, wherein the fully connected layer comprises a first, second and third fully connected layer.
14. A computer-implemented method for classifying eye images, comprising:
inputting at least one eye image into the system of claim 1 for processing; and
obtaining an image classification result of either normal class or abnormal class based on the processing of the at least one eye image.
15. A computer readable storage medium, storing non-transitory instructions for controlling an automated processor to execute the method of claim 14.

16. A method of training a convolutional neural network to classify a fundus image as one of normal and abnormal, the method comprising steps of:

- a) obtaining a first plurality of fundus images with features of an eye condition or disease;
- b) obtaining a second plurality of fundus images with normal features;
- c) processing the first plurality of fundus and the second plurality of fundus through the convolutional neural network to identify features in the first plurality of fundus images as characteristic of an eye condition or disease;

wherein the convolutional neural network is comprised of a convolution layer, a pooling layer and a fully connected layer, wherein the CNN has been trained using an ADAM optimization algorithm and validated by blindfold cross-validation.

17. The method of claim 16, wherein the eye condition or disease is at least one of age-related macular degeneration (AMD), diabetic retinopathy (DR) and/or glaucoma.

18. The method of claim 16, wherein the convolution layer comprises a first, second, third and fourth convolution layer.

19. The method of claim 18, wherein the first convolution layer comprises eight feature maps (3x3) ReLu, eight feature maps (5x5) ReLu, and eight feature maps (7x7) ReLu concatenated together.

20. The method of claim 18, wherein the second convolution layer comprises sixteen feature maps (3x3) ReLu.

21. The method of claim 18, wherein the third convolution layer comprises sixteen feature maps (3x3) ReLu.

22. The method of claim 18, wherein the fourth convolution layer comprises thirty-two feature maps (3x3) ReLu.
23. The method of claim 16, wherein the pooling layer comprises a first, second and third pooling layer.
24. The method of claim 23, wherein the first, second and third pooling layers reduce the spatial dimension of the input images to half the input size.
25. The method of claim 16, wherein the convolutional neural network comprises a flattening step to convert the three-dimensional output to two-dimensional output.
26. The method of claim 16, wherein the fully connected layer comprises a first, second and third fully connected layer.