Computers in Biology and Medicine 90 (2017) 23–32

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

# Automated arteriole and venule classification using deep learning for retinal images from the UK Biobank cohort

R.A. Welikala [a,*], P.J. Foster [b,c], P.H. Whincup [d], A.R. Rudnicka [d], C.G. Owen [d], D.P. Strachan [d], S.A. Barman [a], on behalf of the UK Biobank Eye and Vision Consortium [1]

[a] School of Computer Science and Mathematics, Kingston University, Surrey, KT1 2EE, United Kingdom
[b] NIHR Biomedical Research Centre, Moorfields Eye Hospital, London, EC1V 2PD, United Kingdom
[c] UCL Institute of Ophthalmology, London, EC1V 9EL, United Kingdom
[d] Population Health Research Institute, St. George's, University of London, London, SW17 0RE, United Kingdom

ARTICLE INFO

*Keywords:*
Retinal images
Arteriole/venule classification
Deep learning
Convolutional neural networks
UK Biobank
Epidemiological studies

ABSTRACT

The morphometric characteristics of the retinal vasculature are associated with future risk of many systemic and vascular diseases. However, analysis of data from large population based studies is needed to help resolve uncertainties in some of these associations. This requires automated systems that extract quantitative measures of vessel morphology from large numbers of retinal images. Associations between retinal vessel morphology and disease precursors/outcomes may be similar or opposing for arterioles and venules. Therefore, the accurate detection of the vessel type is an important element in such automated systems. This paper presents a deep learning approach for the automatic classification of arterioles and venules across the entire retinal image, including vessels located at the optic disc. This comprises of a convolutional neural network whose architecture contains six learned layers: three convolutional and three fully-connected. Complex patterns are automatically learnt from the data, which avoids the use of hand crafted features. The method is developed and evaluated using 835,914 centreline pixels derived from 100 retinal images selected from the 135,867 retinal images obtained at the UK Biobank (large population-based cohort study of middle aged and older adults) baseline examination. This is a challenging dataset in respect to image quality and hence arteriole/venule classification is required to be highly robust. The method achieves a significant increase in accuracy of 8.1% when compared to the baseline method, resulting in an arteriole/venule classification accuracy of 86.97% (per pixel basis) over the entire retinal image.

## 1. Introduction

Inspection of the retinal blood vessels enables a direct and non-invasive view of the blood circulatory system, with images being easily captured using fundus photography. There has been considerable interest in using retinal vessel size and shape as a marker of vascular health status. The morphology of retinal vessels has been prospectively associated with cardiovascular and systemic diseases [1–4]. Associations between retinal vessel morphology and disease precursors/outcomes may be similar or opposing for arterioles and venules. For example, hypertension and atherosclerosis may have different effects in retinal arterioles and venules resulting in a decreased arteriole to venule width (AVR) [5]. Automated systems (e.g. QUARTZ [6,7] and VAMPIRE [8–10]) that extract quantitative measures of vessel morphology from large numbers of retinal images are needed to power these biomarker discovery studies, in which the automated classification of arterioles and venules is an essential element.

The appearance of arterioles and venules in retinal images are very similar. However, they can be differentiated in general using the following features as documented by Kondermann [11]:

- Arterioles are brighter than venules.
- Arterioles are thinner then neighbouring venules.
- The central reflex (the light reflex of the inner parts of the vessels) is wider in arterioles and smaller in venules.

* Corresponding author.
*E-mail address:* R.Welikala@kingston.ac.uk (R.A. Welikala).
[1] Members of the UK Biobank Eye and Vision Consortium are listed before References.

http://dx.doi.org/10.1016/j.compbiomed.2017.09.005
Received 1 June 2017; Received in revised form 5 September 2017; Accepted 5 September 2017

- Arterioles and venules usually alternate near the optic disc before branching out; that means near the optic disc one arteriole is usually next to two venules and the other way round.

There are challenges in building a robust vessel classification system. There is intra-image and inter-image variance in respect to colour, contrast and luminosity. The size and colour of vessels changes as they move away from the optic disc due to changes in levels of oxygenation. In the periphery (far away from the optic disc) vessels become so thin that are almost indistinguishable.

A number of automated methods have been reported in retinal arteriole/venule (a/v) classification. These can be divided into two broad categories; feature based and graph based methods. The majority of methods start with the segmentation of the vasculature followed by the creation of the vessel centrelines and then the removal of bifurcation and crossover points to create vessel segments. The centreline pixels are classified as arteriole or venule, and this information is then used to award the whole vessel segment as either arteriole or venule. Alternatively, some methods avoid pixel classification and make a direct decision on the a/v status of a vessel segment.

The most popular approaches are feature based methods. Kondermann [11] addressed intra-image variability by approximating the background and removing this from the image. Classification of the centreline pixels was then achieved using colour features from a square region centred on the target pixel and a neural network with one hidden layer of 40 neurons. Grisan [12] divided the image into four quadrants with the assumption that each quadrant had at least one arteriole and one venule. Within a quadrant, centreline pixels were classified using fuzzy clustering based on RGB (red, green, blue) and HSL (hue, saturation, lightness) features measured from a local circular region centred on the pixel. The use of quadrants exploited the local nature of a/v classification and thereby addressed the issue of intra and inter-image variability. Saez [13] and Vazquez [14] rotated the four quadrants in steps of 20° with the aim of fulfilling the assumption that each quadrant should contain at least one arteriole and one venule. Both performed K-means clustering based on vessel profiles and their RGB and HSL colour information. Vazquez [14] improved performance with the use of a tracking procedure to connect segments that belong to the same vessel based on finding the path of minimal cost between two points using image information.

Other feature based methods include the following. Fraz [15] avoided the computational expense of the quadrant based approach and addressed intra and inter-image variability using shade correction and image normalization. Classification of the centreline pixels was performed using the ensemble classifier of bagged decision trees and pixel based, profile based and segment based features from both RGB and HSI (hue, saturation, intensity) colour spaces. Relan [16] developed a method with a feature vector based on colour information calculated from a circular neighbour around the target pixel. Classification was performed using a Least Square Support Vector Machine, and showed that the method performed well using a small training dataset. This is the latest method to be incorporated into the VAMPIRE software [8–10]. Niemeijer [17] proposed methodology that labelled centreline pixels using a 27-D feature vector and the linear discriminant classifier. Pixel and vessel profile based features were used based on RGB and HSI colour information. Performance was improved using the prior knowledge that arterioles and venules usually come in pairs, thus an iterative approach was used to match arteriole venule pairs. Xu [18] proposed novel features for a/v classification which included first and second order texture features measured from the vessel profile and an image patch around the target pixel. A k-nearest neighbour was used for classification.

Graph based approaches make up a smaller section of the literature. Rothaus [19] and Dashtbozorg [20] used the centreline of the vasculature to produce a planar graph, in which each link corresponded to a vessel segment and the nodes represented the branches or crossings of the vessel segments. Using information on the node (e.g. number of links connected to the node, orientation of each link etc.), the node type was determined. Once all nodes on the graph had been identified then all links that belong to a particular vessel could be identified. Rothaus [19] required a few manually labelled vessel segments and then this rule based method was used to propagate vessel labels through the vascular graph. Dashtbozorg [20] separately performed a/v classification for every centreline pixel using a 30-D feature vector based on colour information and a linear discriminant analysis classifier; this was combined with the graph based labelling to achieve the final classification. Estrada [21] did not use a rule based method to determine the type of nodes of the planar graph. Instead the method took an optimization based approach where it determined the most likely set of a/v labels by efficiently searching through the space of possible a/v labelled trees. A global likelihood model that used the features of local growth, overlap and colour was used to determine the quality of each directed tree (a/v labelled tree).

There are many other notable a/v classification methods [22–25]. Existing methods can heavily rely on hand crafted features and thereby can be limited to what humans perceive as interpretable differences between arterioles and venules. Graph based approaches can struggle when sections of the vasculature cannot be segmented, resulting in unreliability when linking vessel segments. Therefore, there is scope for improvement in a/v classification.

In the past few years there has been a shift towards methods employing deep learning approaches across a variety of fields. Deep learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks. Deep referring to the number of layers in the network. The driving force behind deep learning is that we now have fast enough computers and large enough datasets to actually train large neural networks. Neural network results get better with more data and larger models, but in turn require more computation to train. In addition to scalability, another benefit of deep learning models is their ability to perform automatic feature extraction from raw data, also called feature learning. Automatically learning features at multiple levels of abstraction allow a system to learn complex functions mapping the input to the output directly from data, without depending completely on human-crafted features. A convolutional neural network (CNN) is a powerful deep learning technique whose architectures make the explicit assumption that the inputs are images. Since winning the ImageNet [26] competition in 2012 with AlexNet [27], CNNs have gained wide popularity in computer vision. The highest performing algorithms in the recent Kaggle competition [28], which completed in July 2015, all used CNNs to identify signs of Diabetic retinopathy in retinal images. The application of CNNs on retinal images continues to grow with multiple recent studies [29–36]. To the best of our knowledge deep learning has yet to be applied to tackle a/v classification.

In this paper, we present a convolutional neural network architecture designed for the automated classification of arterioles and venules in retinal images. Implementing deep learning enabled complex patterns to automatically be learnt from the data; thus, avoiding the use of hand crafted features. The classification was performed across the entire vasculature in the retinal images and not simply concentric areas centred on the optic disc or a limited number of pre-specified retinal vessel locations. The method has been developed and evaluated using retinal images from UK Biobank which is a large population-based cohort study. The proposed method will replace the current a/v classification method [15] in QUARTZ (QUantitative Analysis of Retinal vessel Topology and siZe) [6,7]. QUARTZ is a retinal image analysis system developed by our research group, capable of processing large numbers of retinal images and obtains quantitative measures of vessel morphology to be used in epidemiological studies [37].

## 2. Materials

UK Biobank [38] contains a very large retinal image repository (135,867 images from 68,549 participants at baseline examination) in a

middle-aged population-based cohort study. The sizable amount of health data recorded make it a powerful research resource for epidemiological studies. The entire UK Biobank retinal image dataset has now been processed by QUARTZ, which incorporated our proposed a/v classification algorithm. The retinal images are macular centred and colour; captured with a non-mydriatic fundus camera (Topcon 3DOCT-1000 Mk 2) with a 45° field-of-view and saved in PNG format with a resolution of 2048 × 1536 pixels. The UK Biobank study was approved by the Northwest Region NHS research ethics committee.

For training and evaluating a/v classification, 100 retinal images were used. As will be explained in section 3.4, this equated to 835,914 centreline pixels from 5881 vessel segments. These 100 images were randomly selected from 96,518 out of 135,867 images in the UK Biobank dataset that had been automatically labelled by QUARTZ as being of adequate image quality for use in epidemiological studies [39]. From our experience, most large retinal datasets used in epidemiological studies (e.g. UK Biobank and EPIC-Norfolk [40]) contain large amounts of poor quality images caused by a variety of issues. In an attempt to reduce wastage, QUARTZ was designed to be capable of extracting useful information from some poor quality images [39]. This meant images deemed as adequate by QUARTZ still possessed considerable variability in image quality (see Fig. 1). This made this a challenging dataset and hence a/v classification was required to be highly robust.

## 3. Method

### 3.1. Overview

The vascular network was segmented from the retinal image and the vascular skeleton consisting of centreline pixels was constructed. Vessel segments were generated by the search and removal of bifurcations and crossover points. For each centreline pixel in a vessel segment, a small square colour image patch centred on the target pixel was provided as the input for the convolutional neural network (CNN) in order to assign an arteriole or venule label to the target pixel. The labels for the centreline pixels were then used to award the whole vessel segment as arteriole or venule using a voting strategy.

The proposed methodology has the same initial framework as work previously published by our research group [15], thus section 3.2 and 3.3 shall be kept brief. The new addition shall be detailed in sections 3.4–3.7 which comes in the form of a convolutional neural network to replace an ensemble classifier of bagged decision trees which used hand crafted features from the RGB and HSI colour spaces.

### 3.2. Vessel segmentation and segment extraction

An unsupervised vessel segmentation approach based on a multi-scale line detector [41] was implemented on the inverted green channel of the coloured retinal images. This was based on the approximation of retinal vessel segments being piecewise linear. The average grey-level of the pixels along a line passing through the target pixel was calculated at 12 different orientations spaced by 15° each. For the target pixel, the line-strength was obtained from the orientation with the largest value subtracted by the average grey-level of a square sub-window centred at the target pixel. As this was a multiscale approach, lines of different lengths (15,17,19,21 pixels) were applied and a linear combination of responses resulted in a final line-strength image. This was followed with the application of a hysteresis thresholding based morphological reconstruction to produce the segmented vasculature.

A morphological thinning operation was applied to the segmented vasculature to create one pixel wide vessel centrelines. Vessel segments were created from these centrelines by first removing spurs (≤10 pixels), followed by the removal of bifurcation and crossover points, and ending with the removal of short segments (≤15 pixels). Fig. 2 illustrates the outputs of this procedure.

### 3.3. Pre-processing

The absolute colour of the blood in the vessels varies within the images (intra-image) and across the subjects (inter-image). To reduce intra-image variation shade correction was performed, which was accomplished by estimating the background image and subtracting that from the original image. The background image was obtained by mean filtering the original image with a 151 × 151 pixel kernel. Image normalization was then used to reduce inter-image variation. In which the histogram was calculated and the median value was found (belonged to the background of the retinal image). Then a simple linear intensity transformation set the median value to unified centre (0.5 in this study) and standardized the intensity around this value. These pre-processing steps were performed for each of the RGB colour channels individually. The results of pre-processing are shown in Fig. 3.

### 3.4. Data preparation and augmentation

This was a two-class problem and the methodology was designed so that the classifier would label each vessel centreline pixel as belonging to an arteriole or a venule. CNNs require a large amount of labelled data for training. As explained in Section 2, 100 retinal images were used in this study. The extracted vessel segments from these images were manually labelled as arteriole or venule by two human observers (clinicians). The labelled data by the first observer was used as the reference standard. The second observer labelled a random subset of five images; and a high agreement of 98.84% was achieved between the two observers. All centreline pixels in a segment took up the manual label awarded to the segment. This equated to 835,914 centreline pixels (5881 vessel segments in total) possessing a manual label. This labelled data was then randomly divided into 50 training images, 15 validation images and 35 testing images corresponding to 424,122, 115,262, 314,409 centreline

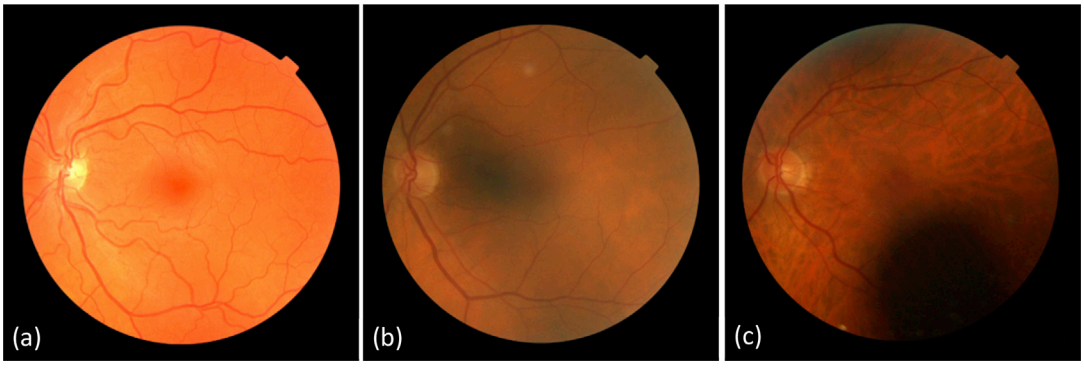


**Fig. 1.** Retinal images of differing quality deemed as adequate by QUARTZ. © UK Biobank.

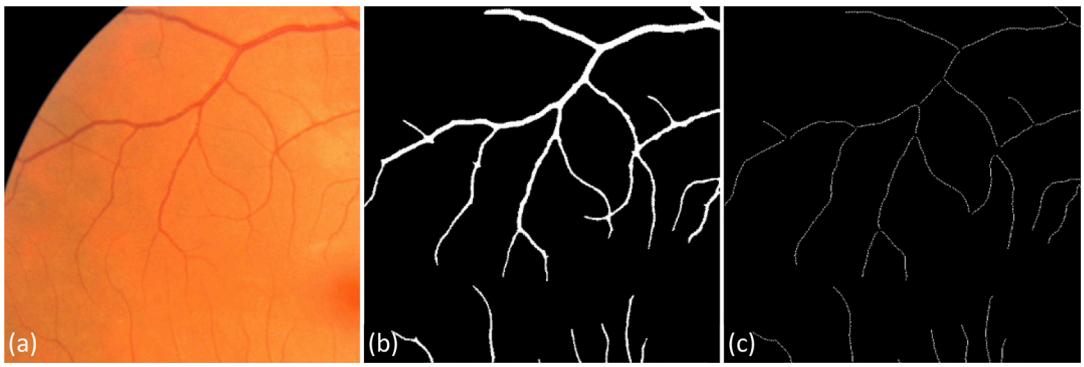


**Fig. 2.** Zoom-in regions: (a) colour retinal image, (b) segmented vasculature, (c) extracted vessel segments. © UK Biobank. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

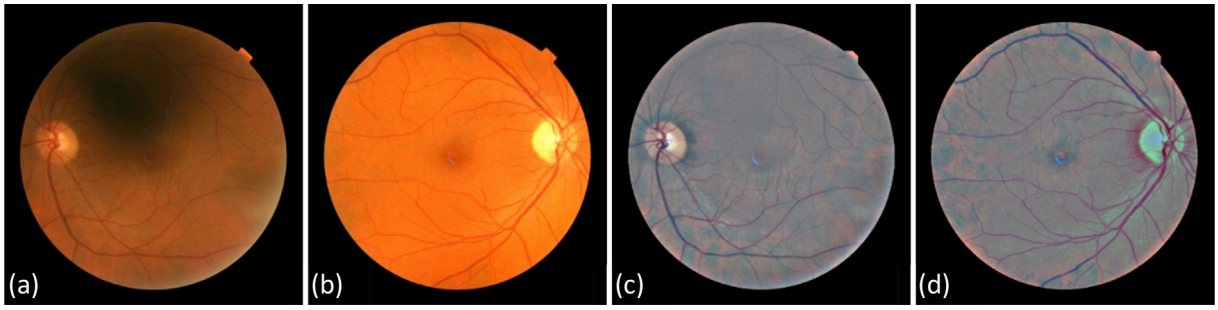


**Fig. 3.** (a)–(b) Original retinal images, (c)–(d) pre-processed images. © UK Biobank.

pixels respectively. The training data was balanced (validation data was also balanced) to ensure the CNN has no bias to either class (arteriole or venule). The test data remained unchanged to represent the real-world scenario.

For each centreline pixel in a vessel segment, a 31 × 31 pixel colour (pre-processed) image patch centred on the target pixel was provided and the intention was to use convolutional neural networks (CNN) in order to assign an arteriole or venule label to the target pixel (examples shown in Fig. 4). With deep learning, the size of the network means that overfitting can become a significant problem, even with 400,000 + training examples. Artificially enlarging the training dataset, known as data augmentation, was applied to reduce overfitting. This came in the form of image translations and horizontal reflections, achieved by extracting random 25 × 25 pixel colour patches (with random horizontal reflections) from the 31 × 31 patches and training our network (as explained in section 3.6) on these extracted patches. This increased the size of the training set by a factor of 72, although this included some repetition due to adjacent centreline pixels' local regions having large overlaps.

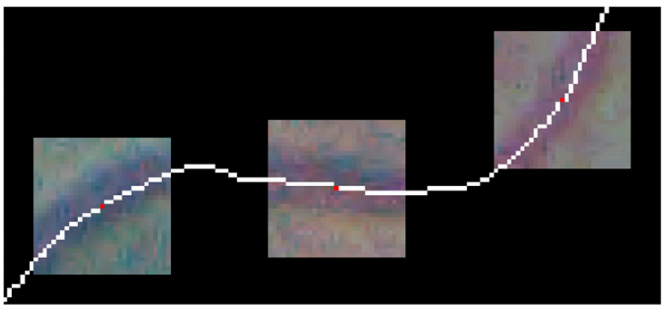

**Fig. 4.** Centreline of a vessel and examples of 31 × 31 pixel colour (pre-processed) image patches created for different centreline pixels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 3.5. Network architecture

Convolutional neural networks are very similar to regular neural networks: they are made up of neurons that have learnable weights and biases (parameters), with the neurons arranged in layers. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity (activation function: tanh, sigmoid, ReLU). However, CNN architectures make the explicit assumption that the inputs are images, to exploit the strong spatially local correlation present in natural images. Instead of a neuron being connected to every neuron in the previous layer, each neuron connects to only a local region (receptive field of the neuron) from the previous layer. Thus, each neuron only learns weights for inputs from its receptive field, vastly reducing the amount of parameters in the network. A layer of a CNN has neurons arranged in 3 dimensions: width, height and depth. Note that the receptive field of a neuron is small spatially (along width and height), but extends through the full depth of the previous layer. All neurons at a single depth (referred to as depth slice) use the same weights (parameter sharing). Thus, the output of a layer can in each depth slice be computed as a convolution of the neuron's weights with the input volume. This is why it is common to refer to the sets of weights as a filter, that is convolved with the input. The depth corresponds to the number of filters in the layer. Stanford University [42] provide a thorough introduction to CNNs.

The paragraph above describes convolutional layers; however, there are other layers used to build CNN architecture. Typical layers are: convolutional, ReLU, cross channel normalization, pooling, fully-connected (exactly as in regular neural networks), dropout and softmax. The ReLU (rectified linear units) activation function introduces nonlinearity to the system and is preferred over the tanh and sigmoid functions due to the computational efficiency it brings to training. Pooling, cross channel normalization and dropout all help against overfitting. In addition, pooling is used to reduce the number of parameters in the network. The softmax layer follows the final fully-connected layer

and produces a probability distribution over the output class labels.

We propose a CNN designed to label the input ($25 \times 25 \times 3$ extracted image patch) with the class label of arteriole or venule. The proposed architecture contained 6 layers with weights, the first three were convolutional and the remaining three were fully-connected. The three convolutional layers had 24, 48 and 48 filters of size $3 \times 3 \times 3$, $3 \times 3 \times 24$ and $3 \times 3 \times 48$ respectively. The three fully-connected layers contained 85, 85 and 2 neurons respectively, the final layer contained 2 neurons as there were two classes. The ReLU layer followed every convolutional layer and the first two fully-connected layers. Cross channel normalization (window channel size = 5) followed all the convolutional layers. Overlapping max-pooling followed the first and third convolutional layers. Dropout (probability = 0.5) followed the first two fully-connected layers. The final fully-connected layer was followed by a softmax layer which outputs a probability score for the arteriole class and the venule class. Fig. 5 illustrates the order of these layers in our network. Table 1 provides additional details (e.g. stride, zero-padding etc.) on the convolutional, max-pooling and fully connected layers of our CNN and Fig. 6 provides a schematic overview. Overall our CNN architecture amounted to 141,439 parameters.

### 3.6. Details of learning

The parameters (weights and biases) in the convolutional and fully-connected layers were trained by backpropagation and gradient descent (with the cross-entropy cost function) so that the class scores that the CNN computed were consistent with the manual a/v labels in the training set ($25 \times 25 \times 3$ extracted image patches). Stochastic gradient descent was used with a mini-batch size of 128 samples, momentum of 0.9 and $L_2$regularization (weight decay) of 0.0005. The weights in each layer were initialised from a Gaussian distribution with a mean of 0 and a standard deviation of 0.01. The initial biases were set to 0. The learn rate was initialised at 0.01 and was divided by 10 every 10 epochs. Training was run for a maximum of 30 epochs. The training set was shuffled prior to being used. Training took 32 h on a NVIDIA Quadro K1100 M 2GB GPU.

To further combat overfitting on the training set, the performance of the model on the validation set was recorded at every epoch. The accuracy for the validation set peaked at epoch 14 with 83.46%. Therefore, the parameters of the CNN at epoch 14 was chosen as the final model. Fig. 7 plots the model accuracy for the training set and the validation set; it is evident that after epoch 14 slight overfitting began to occur (the accuracy for the validation set stops increasing whilst the accuracy for the training set continues to increase).

The trained CNN had learnt filters which can be considered as automatically learnt features. Fig. 8 displays the automatically learnt features from the first convolutional layer. Notice how the first layer of the network had learnt some filters for capturing edge features. These primitive features were then processed by deeper network layers, which combine the early features to form higher-level image features. These higher-level features are better suited for recognition tasks because they combine all the primitive features into a richer image representation.

**Table 1**
Details of the convolutional, max-pooling and fully-connected layers of proposed CNN architecture.

| Layer | Size | Filter size | Stride | Zero-Padding |
|---|---|---|---|---|
| Input | $25 \times 25 \times 3$ | – | – | – |
| Convolutional | $23 \times 23 \times 24$ | $3 \times 3 \times 3$ | 1 | 0 |
| Max-Pooling | $11 \times 11 \times 24$ | $3 \times 3$ | 2 | 0 |
| Convolutional | $11 \times 11 \times 48$ | $3 \times 3 \times 24$ | 1 | 1 |
| Convolutional | $11 \times 11 \times 48$ | $3 \times 3 \times 48$ | 1 | 1 |
| Max-Pooling | $5 \times 5 \times 48$ | $3 \times 3$ | 2 | 0 |
| Fully-Connected | 85 | – | – | – |
| Fully-Connected | 85 | – | – | – |
| Fully-Connected | 2 | – | – | – |

### 3.7. Vessel labels

The final label of the vessel centreline pixel was obtained by the trained CNN's prediction of ten different $25 \times 25$ colour image patches extracted from the $31 \times 31$ colour image patch. Four corner patches and the centre patch as well as their horizontal reflections (hence ten patches overall); averaging the predictions made by the CNN's softmax layer on the ten patches. This procedure was followed for both the validation and the testing phase.

The soft labels assigned to the centreline pixels were regarded as a vote for the label of the complete vessel segment, and the mean of these votes was assigned as the label for the entire vessel segment. This procedure was performed across the entire vasculature in the retinal images.

## 4. Results

To evaluate the proposed system, the true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) were calculated for each vessel type (see Table 2) and the performance metrics have been calculated separately for arterioles and venules (see Table 3). Evaluation of the system was performed on both a centreline pixel basis and a segment basis.

In the first instance, the methodology was evaluated (using the test data) in terms of a centreline pixel basis prior to the implementation of the voting strategy for the segment label. This gives a direct sight into the classifier's performance. These results are presented in Table 4 along with the performance of the baseline method in which we implemented the approach by Fraz [15] on the exact same UK Biobank data. The proposed CNN achieves an accuracy of 82.26% (pixel basis).

Following the segment voting strategy the accuracy on a pixel basis increases to 86.97% (consider all centreline pixels within the segment take on the label of the segment) and the accuracy on a segment basis is 85.24%. These results are listed in Table 5. For comparison, the performance was also assessed over a specific region-of-interest (ROI), 0.5–1.0 optic disc diameters from the optic disc boundary. These results are listed in Table 6. The classification of vessel segments into arterioles and venules is shown in Fig. 9, where the red colour segments represent arterioles and blue colour segments represent venules.

The reported results for other a/v classification methods are presented in Table 7. To enable a better comparison to these methods, the


**Fig. 5.** Layers of proposed CNN architecture.

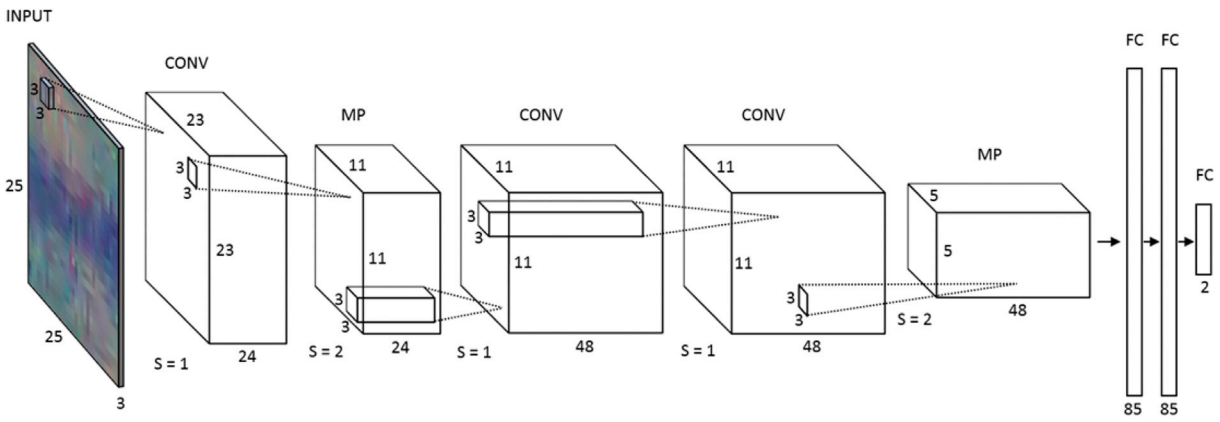


**Fig. 6.** Schematic overview of proposed CNN architecture with respect to the convolutional (CONV), max-pooling (MP) and fully-connected (FC) layers. S = stride.

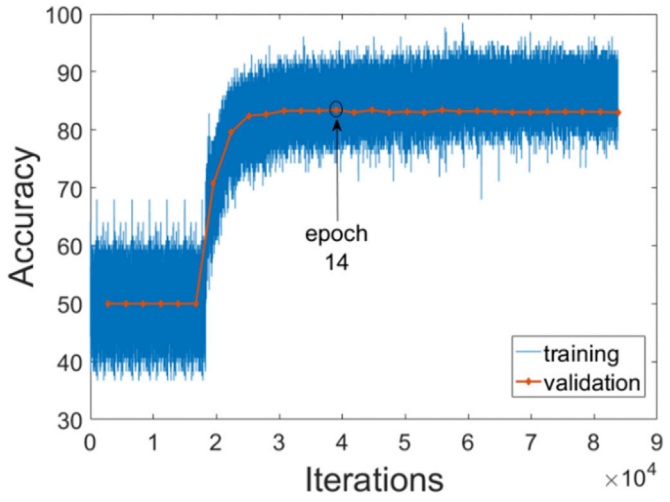


**Fig. 7.** Plot of model accuracy on the training set (recorded at every iteration) and validation set (recorded at every epoch). Each evaluation of the gradient using the mini-batch is an iteration. The full pass of the training over the entire training set using mini-batches is an epoch.

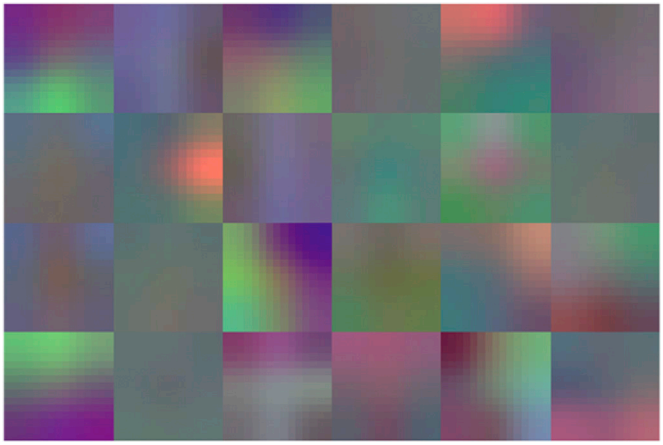

**Fig. 8.** 24 filters of size 3 × 3 × 3 (resized here for visualisation) learnt by the first convolutional layer on the 25 × 25 × 3 input image patches.

proposed method was also evaluated on the publicly available DRIVE [43] dataset (results included in Table 7). The methodology remained unchanged; hence, the same CNN architecture was used. This was enabled by resizing the DRIVE images from 565 × 584 pixels to 1464 × 1513 pixels to match the size of the circular field-of-view of the UK Biobank images. To maximise the performance the CNN was retrained on the DRIVE dataset (following manual labelling and data augmentation), with the 40 DRIVE images being split into: 25 training images, 5 validation images and 10 testing images. The need for validation data and the large training data requirements of deep learning meant that it was not practical to use the original split of the DRIVE dataset (20 training and 20 testing images). The parameters of the CNN at epoch 9 was chosen as the final model. The classification of vessel segments into arterioles and venules is shown in Fig. 10.

**Table 2**
Vessel classification measures, calculated separately on a centreline pixel basis and on a segment basis.

| Measure | Description of identification |
|---|---|
| $TP_{Venule}$ | System = venule, observer = venule |
| $FP_{Venule}$ | System = venule, observer = artery |
| $TN_{Venule}$ | System = artery, observer = artery |
| $FN_{Venule}$ | System = artery, observer = venule |
| $TP_{Arteriole}$ | System = artery, observer = artery |
| $FP_{Arteriole}$ | System = artery, observer = venule |
| $TN_{Arteriole}$ | System = venule, observer = venule |
| $FN_{Arteriole}$ | System = venule, observer = artery |

**Table 3**
Performance metrics for vessel classification, calculated separately on a centreline pixel basis and on a segment basis. Also, calculated separately for venules using $TP_{Venule}$, $FP_{Venule}$, $TN_{Venule}$, $FN_{Venule}$ and arterioles using $TP_{Arteriole}$, $FP_{Arteriole}$, $TN_{Arteriole}$, $FN_{Arteriole}$.

| Measure | Description |
|---|---|
| Sensitivity | TP/(TP + FN) |
| Specificity | TN/(TN + FP) |
| Positive Predictive Value | TP/(TP + FP) |
| Negative predictive Value | TN/(TN + FN) |
| Positive Likelihood Ratio | Sensitivity/(1-Specificity) |
| Negative Likelihood Ratio | (1-Senitivity)/Specificity |
| Accuracy | (TP + TN)/(TP + FP + TN + FN) |

**Table 4**
Performance of the proposed a/v classification method and the baseline method on UK Biobank data, over entire images on a per pixel (centreline) basis prior to the segment voting strategy.

| Method | Measure | Arteriole | Venule |
|---|---|---|---|
| Fraz [15] | Sensitivity | 72.98% | 75.08% |
| | Specificity | 75.08% | 72.98% |
| | Accuracy | **74.16%** | **74.16%** |
| Proposed | Sensitivity | 82.71% | 81.91% |
| | Specificity | 81.91% | 82.71% |
| | Accuracy | **82.26%** | **82.26%** |

**Table 5**
Performance of the proposed a/v classification method on UK Biobank data after the segment voting strategy, over the entire image. Fraz [15] achieved an accuracy of 74.91% (pixel basis) and an accuracy of 71.37% (segment basis) over the entire image.

| Measure | Pixel | | Segment | |
|---|---|---|---|---|
| | Arteriole | Venule | Arteriole | Venule |
| Sensitivity | 86.07% | 87.67% | 85.14% | 85.32% |
| Specificity | 87.67% | 86.07% | 85.32% | 85.14% |
| Positive Predictive Value | 84.56% | 88.92% | 81.23% | 88.49% |
| Negative Predictive Value | 88.92% | 84.56% | 88.49% | 81.23% |
| Positive Likelihood Ratio | 6.9815 | 6.2938 | 5.7987 | 5.7402 |
| Negative Likelihood Ratio | 0.1589 | 0.1432 | 0.1742 | 0.1725 |
| Accuracy | **86.97%** | **86.97%** | **85.24%** | **85.24%** |

**Table 6**
Performance of the proposed a/v classification method on UK Biobank data after the segment voting strategy, over the ROI (0.5–1.0 optic disc diameters from the optic disc boundary).

| Measure | Pixel | | Segment | |
|---|---|---|---|---|
| | Arteriole | Venule | Arteriole | Venule |
| Sensitivity | 86.64% | 87.66% | 86.90% | 90.79% |
| Specificity | 87.66% | 86.64% | 90.79% | 86.90% |
| Positive Predictive Value | 86.65% | 87.65% | 90.00% | 87.90% |
| Negative Predictive Value | 87.65% | 86.65% | 87.90% | 90.00% |
| Positive Likelihood Ratio | 7.0214 | 6.5590 | 9.4345 | 6.9287 |
| Negative Likelihood Ratio | 0.1525 | 0.1424 | 0.1443 | 0.1060 |
| Accuracy | **87.17%** | **87.17%** | **88.89%** | **88.89%** |

## 5. Discussion and conclusion

In this paper, we have presented a highly robust a/v classification method based on the use of deep learning. This comprises of a convolutional neural network whose architecture contains six learned layers: three convolutional and three fully-connected. Complex patterns were automatically learnt from the data, avoiding the use of hand crafted features. The method was developed for use on the UK Biobank retinal image dataset and produces an accuracy of 86.97% (per pixel basis) for a/v classification over the entire image. The performance increases to 87.17% (per pixel basis) when evaluated over a ROI, the increase is only slight which indicates the performance is stable across the entire retina (normally a significant increase is expected for the ROI). These figures represent a high level of performance when considering the challenging nature of the UK Biobank dataset in terms of the considerable variability in image quality. Likelihood ratios (listed in Tables 3 and 4) were also high and confirmed the high reliability of our proposed method.

The main contribution of this paper is the novel application of CNNs to tackle a/v classification in retinal images. This sees the proposed CNN labelling small image patches containing sections of individual vessels. Human observers use the aspects of size, brightness and colour, relative to neighbouring vessels and also track the routes of vessels to determine the a/v status. Therefore, human observers would struggle to determine the a/v status of a random selection of these image patches. Conventionally CNN models are designed to classify images in which different classes display apparent visual differences (e.g. diabetic retinopathy grades, breed of dog etc.). Our method further demonstrates the capabilities of CNNs, effectively classifying image patches when the differences are not so apparent. Thus, avoiding the need to identify neighbouring vessels or correctly linking vessel segments which can be

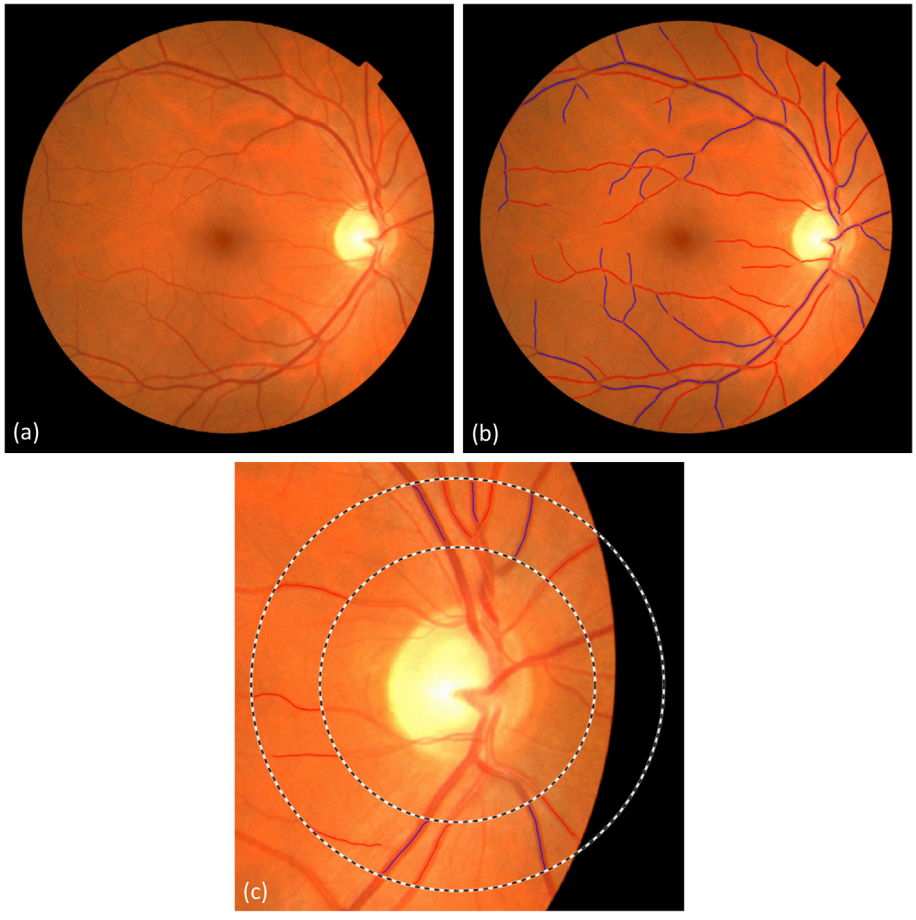


**Fig. 9.** Classification of arterioles and venules on UK Biobank data. (a) Original retinal image. Classification results for the: (b) entire image and (c) ROI. (c) © UK Biobank.

**Table 7**
Reported results for other automated a/v classification methods and the results of the proposed method on the publicly available DRIVE [43] dataset. OD = optic disc, ODD = optic disc diameter, ODR = optic disc radius. Saez [13] reports an arteriole and venule sensitivity of 78.19% and 87.90% respectively. Niemeijer [17] reports an area under the ROC curve of 0.84.

| Method | Dataset | Year | Accuracy | Level | Location |
|---|---|---|---|---|---|
| Kondermann [11] | Private | 2007 | 95.32% | Pixel | Within 3.0 ODD from OD centre |
| Grisan [12] | Private | 2003 | 87.58% | Segment | 4 quadrants, 5 largest vessels in each |
| Saez [13] | Private | 2012 | – | Segment | 1.5–2.5 ODR from OD centre |
| Vazquez [14] | VICAVR | 2013 | 87.68% | Segment | 1.5–3.0 ODR from OD centre |
| Fraz [15] | Private | 2014 | 83% | Pixel | Entire image |
| Relan [16] | DRIVE | 2014 | 89.4% | Segment | 0.5–1.0 ODD from OD boundary |
| Niemeijer [17] | Private | 2011 | – | Pixel | 1–1.5 ODD from OD centre |
| Xu [18] | DRIVE | 2017 | 92.3% | Pixel | Entire image |
| Dashtbozorg [20] | DRIVE | 2014 | 87.4% | Pixel | Entire image |
| Estrada [21] | DRIVE | 2015 | 91.7% | Pixel | Entire image |
| | | | 93.5% | Segment | Entire image |
| Proposed method | DRIVE | – | 91.97% | Pixel | Entire image |
| | | | 91.27% | Segment | Entire image |
| | | | 91.99% | Pixel | 0.5–1.0 ODD from OD boundary |
| | | | 90.77% | Segment | 0.5–1.0 ODD from OD boundary |

unreliable when sections of the vasculature cannot be segmented due to variability in the image quality of this dataset.

Significant improvements were demonstrated when comparing our method to the baseline performance, increasing classification accuracy from 74.16% to 82.26% (per pixel basis, entire image, prior to voting strategy), an increase of 8.1%. The baseline performance implemented the method developed by Fraz [15] on the exact same retinal images from UK Biobank data used in this paper. Thus, we can state the use of convolutional neural networks is more effective than the ensemble classifier of bagged decision (used with colour features). Whilst Fraz [15] does not present the highest accuracy in Table 7, it was the only other method designed for use with images of variability quality from a large cohort study (EPIC-Norfolk [40]) and therefore was a suitable choice as the baseline method. Fraz [15] reports a classification accuracy of 83% for images from the EPIC-Norfolk [40] dataset which contains retinal images similar to UK Biobank dataset in respect to image quality. However, their evaluation process requires adjustment to be suitable for comparison purposes because their method uses the right eyes of the participants as training data and the left eyes of those same participants are used as testing data.

Table 7 presents the reported results of other a/v classification methods. Impressive results by Estrada [21] and Xu [18] were reported, both achieved classification accuracies of +90%. The results reported by Kondermann [11] have been considerably boosted using manual vessel segmentation, use of automatic segmentation sees the results deteriorate by 10%. Amongst the datasets used by these methods [18,21] was the publicly available DRIVE dataset [43], in which images are of good quality with respect to not suffering from poor illumination, camera artefacts, obstruction from eyes lashes, etc. To enable a direct comparison to these methods, our proposed methodology was also evaluated on the DRIVE dataset, achieving a similar classification rate (+90%) to the state-of-the-art methods. This is impressive considering the training of the CNN would have been limited by the small size of the DRIVE dataset (despite the use of data augmentation). Also, whilst the DRIVE dataset images are of good quality with respect to illumination etc., their low image resolution of 565 × 584 pixels is another limiting factor.

Numerous CNN architectures in respect to the number of layers, the type of layers, the number of neurons, the number of filters, etc. were explored with their performances being compared on the validation set. It was important to develop a model complex enough to tackle the problem, without being overly complex to avoid overfitting; this was ascertained by comparing the performance on the training set to that on the validation set. Rotation and scaling were explored for data augmentation to further enlarge the training set, this helped to combat overfitting and allowed more complex models to be investigated. Despite this, no improvements could be achieved, likely due to interpolation overly altering pixel intensities which is problematic considering this CNN was designed to be very reactive to small pixel distribution patterns. Other exploration included trying alternatives to the pre-processed image patch inputs. This included original image patches, rotating the patch according the vessel orientation and varying the size of the patch according to the vessel width (with the patch being resized afterwards).

The future plan would be to further take advantage of the scalability of deep learning. There is the potential to improve results by using more data, multiple GPUs and more complex (deeper) models. We have data available in terms of UK Biobank retinal images; our limitation is that hand labelling those images is time-consuming. A possibility would be crowdsourcing with the prerequisite of the individuals having an ophthalmic background to ensure data of high quality. Producing a single trained CNN that performs well on all retinal image datasets is another aspect that warrants investigation. The CNN trained on UK Biobank data

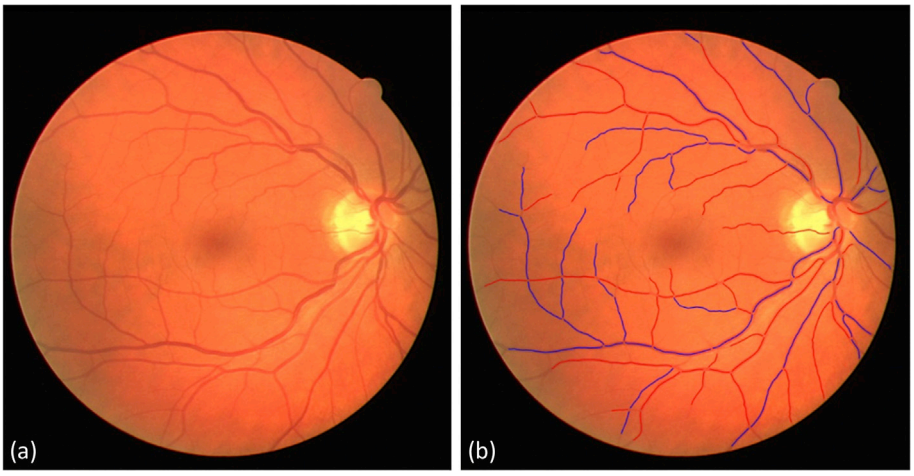


**Fig. 10.** Classification of arterioles and venules on DRIVE data. (a) Original retinal image. (b) Classification results.

was not effective on the DRIVE dataset (despite resizing) and required retraining on DRIVE data, although preliminary investigations show it to be very effective on the similar EPIC-Norfolk [40] dataset. Finally, the possibility of deep learning completely taking over the whole procedure is also viable in the future, directly performing segmentation of the background, arterioles, venules and the optic disc.

In conclusion, this paper has demonstrated a deep learning approach for the automatic classification of arterioles and venules across the entire retinal image for images of variable quality. The method was incorporated into QUARTZ and was used to process the entire UK Biobank retinal image dataset with the output currently being used in epidemiological studies.

## Members of the UK Biobank Eye and Vision Consortium

Prof Tariq Aslam, Prof Sarah Barman, Prof Paul Bishop, Mr Peter Blows, Dr Catey Bunce, Dr Roxana Carare, Prof Usha Chakravarthy, Miss Michelle Chan, Mrs Antonietta Chianca, Dr Valentina Cipriani, Prof David Crabb, Mrs Philippa Cumberland, Dr Alexander Day, Miss Parul Desai, Prof Bal Dhillon, Prof Andrew Dick, Prof Paul Foster, Dr John Gallacher, Prof David (Ted) Garway-Heath, Mr Srini Goverdhan, Prof Jeremy Guggenheim, Mrs Priyal Gupta, Prof Chris Hammond, Dr Ruth Hogg, Prof Anne Hughes, Mr Pearse Keane, Prof Sir Peng Tee Khaw, Mr Anthony Khawaja, Mr Gerassimos Lascaratos, Prof Andrew Lotery, Prof Phil Luthert, Dr Tom Mac-Gillivray, Dr Sarah Mackie, Prof Keith Martin, Ms Michelle McGaughey, Dr Bernadette McGuinness, Dr Gareth McKay, Mr Martin McKibbin, Dr Danny Mitry, Prof Tony Moore, Prof James Morgan, Ms Zaynah Muthy, Mr Eoin O'Sullivan, Prof Chris Owen, Mr Praveen Patel, Dr Tunde Peto, Prof Jugnoo Rahi, Dr Alicja Rudnicka, Miss Carlota Grossi Sampedro, Mr David Steel, Mrs Irene Stratton, Mr Nicholas Strouthidis, Prof Cathie Sudlow, Dr Caroline Thaung, Miss Dhanes Thomas, Prof Emanuele Trucco, Mr Adnan Tufail, Prof Stephen Vernon, Mr Ananth Viswanathan, Miss Cathy Williams, Dr Katie Williams, Prof John Yates, Dr Max Yates, Dr Jennifer Yip, Dr Haogang Zhu.

## Conflict of interest

None declared.

## Acknowledgements

This research has been conducted using the UK Biobank resource under application number 522. The UK Biobank Eye and Vision Consortium is supported by funding from The Special Trustees of Moorfields Eye Hospital NHS Foundation Trust, and at the NIHR Biomedical Research Centre at Moorfields Eye Hospital and UCL Institute of Ophthalmology. The image analysis work was supported by the Medical Research Council Population and Systems Medicine Board (MR/L02005X/1), the British Heart Foundation (PG/15/101/31889), and Fight for Sight (1477/8).

## References

[1] N. Cheung, S.M. Saw, F.M. Islam, S.L. Rogers, A. Shankar, K. Haseth, et al., BMI and retinal vascular caliber in children, Obesity 15 (2007) 209–215.
[2] C.G. Owen, A.R. Rudnicka, C.M. Nightingale, R. Mullen, S.A. Barman, N. Sattar, et al., Retinal arteriolar tortuosity and cardiovascular risk factors in a multi-ethnic population study of 10-year-old children; the Child Heart and Health Study in England (CHASE), Arterioscler. Thromb. Vasc. Biol. 31 (2011) 1933–1938.
[3] T.Y. Wong, F.M. Islam, R. Klein, B.E. Klein, M.F. Cotch, C. Castro, et al., Retinal vascular caliber, cardiovascular risk factors, and inflammation: the multi-ethnic study of atherosclerosis (MESA), Invest. Ophthalmol. Vis. Sci. 47 (2006) 2341–2350.
[4] T.Y. Wong, R. Klein, A.R. Sharrett, B.B. Duncan, D.J. Couper, J.M. Tielsch, et al., Retinal arteriolar narrowing and risk of coronary heart disease in men and women: the Atherosclerosis Risk in Communities Study, JAMA J. Am. Med. Assoc. 287 (2002) 1153–1159.
[5] J.J. Kanski, B. Bowling, Clinical Ophthalmology: a Systematic Approach, Elsevier Health Sciences, 2011.
[6] M.M. Fraz, R.A. Welikala, A.R. Rudnicka, C.G. Owen, D.P. Strachan, S.A. Barman, QUARTZ: quantitative Analysis of Retinal Vessel Topology and size - an automated system for quantification of retinal vessels morphology, Expert Syst. Appl. 42 (2015) 7221–7234.
[7] R.A. Welikala, M.M. Fraz, S. Hayat, A.R. Rudnicka, P.J. Foster, P.H. Whincup, et al., Automated retinal vessel recognition and measurements on large datasets, in: Engineering in Medicine and Biology Society, EMBC, Annual International Conference of the IEEE, 2015, pp. 5239–5242.
[8] T.J. MacGillivray, J.R. Cameron, Q. Zhang, A. El-Medany, C. Mulholland, Z. Sheng, et al., Suitability of UK Biobank retinal images for automatic analysis of morphometric properties of the vasculature, PLoS One 10 (2015), e0127914.
[9] A. Perez-Rovira, T. MacGillivray, E. Trucco, K.S. Chin, K. Zutis, C. Lupascu, et al., VAMPIRE: vessel assessment and measurement platform for images of the REtina, in: Engineering in Medicine and Biology Society, EMBC, Annual International Conference of the IEEE, 2011, pp. 3391–3394.
[10] E. Trucco, L. Ballerini, D. Relan, A. Giachetti, T. MacGillivray, K. Zutis, et al., Novel VAMPIRE algorithms for quantitative analysis of the retinal vasculature, in: Biosignals and Biorobotics Conference, 2013, pp. 1–4.
[11] C. Kondermann, D. Kondermann, M. Yan, Blood vessel classification into arteries and veins in retinal images, Medical Imaging, Int. Soc. Opt. Photonics (2007), 651247–651247.
[12] E. Grisan, A. Ruggeri, A divide et impera strategy for automatic classification of retinal vessels into arteries and veins, in: Engineering in Medicine and Biology Society, Annual International Conference of the IEEE, 2003, pp. 890–893.
[13] M. Saez, S. González-Vázquez, M. González-Penedo, M.A. Barceló, M. Pena-Seijo, G. Coll de Tuero, et al., Development of an automated system to classify retinal vessels into arteries and veins, Comput. Methods Programs Biomed. 108 (2012) 367–376.
[14] S.G. Vázquez, B. Cancela, N. Barreira, M.G. Penedo, M. Rodríguez-Blanco, M. Pena Seijo, et al., Improving retinal artery and vein classification by means of a minimal path approach, Mach. Vis. Appl. 24 (2013) 919–930.
[15] M.M. Fraz, A.R. Rudnicka, C.G. Owen, D.P. Strachan, S.A. Barman, Automated arteriole and venule recognition in retinal images using ensemble classification, in: 9th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP), 2014, pp. 194–202.
[16] D. Relan, T. MacGillivray, L. Ballerini, E. Trucco, Automatic retinal vessel classification using a least square-support vector machine in VAMPIRE, in: Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE, 2014, pp. 142–145.
[17] M. Niemeijer, X. Xu, A.V. Dumitrescu, P. Gupta, B. van Ginneken, J.C. Folk, et al., Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs, IEEE Trans. Med. Imaging 30 (2011) 1941–1950.
[18] X. Xu, W. Ding, M.D. Abràmoff, R. Cao, An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image, Comput. Methods Programs Biomed. 141 (2017) 3–9.
[19] K. Rothaus, X. Jiang, P. Rhiem, Separation of the retinal vascular graph in arteries and veins based upon structural knowledge, Image Vis. Comput. 27 (2009) 864–875.
[20] B. Dashtbozorg, A.M. Mendonça, A. Campilho, An automatic graph-based approach for artery/vein classification in retinal images, IEEE Trans. Image Process. 23 (2014) 1073–1083.
[21] R. Estrada, M.J. Allingham, P.S. Mettu, S.W. Cousins, C. Tomasi, S. Farsiu, Retinal artery-vein classification via topology estimation, IEEE Trans. Med. Imaging 34 (2015) 2518–2534.
[22] H. Yu, S. Barriga, C. Agurto, S. Nemeth, W. Bauman, P. Soliz, Automated retinal vessel type classification in color fundus images, Proc. SPIE 8670 (2013), 86700P-86700P.
[23] Q. Mirsharif, F. Tajeripour, H. Pourreza, Automated characterization of blood vessels as arteries and veins in retinal images, Comput. Med. Imaging Graph. 37 (2013) 607–617.
[24] Q.P. Lau, M.L. Lee, W. Hsu, T.Y. Wong, Simultaneously identifying all true vessels from segmented retinal images, IEEE Trans. Biomed. Eng. 60 (2013) 1851–1858.
[25] Q. Hu, M.D. Abràmoff, M.K. Garvin, Automated construction of arterial and venous trees in retinal images, J. Med. Imaging 2 (2015), 044001–044001.
[26] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, L. Fei-Fei, Imagenet: a Large-scale Hierarchical Image Database, CVPR, 2009, pp. 248–255.
[27] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. (2012) 1097–1105.
[28] Kaggle, Inc. Diabetic Retinopathy Detection 2015, February 21 2017.
[29] M.D. Abràmoff, Y. Lou, A. Erginay, W. Clarida, R. Amelon, J.C. Folk, et al., Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning, Invest. Ophthalmol. Vis. Sci. 57 (2016) 5200–5206.
[30] V. Gulshan, L. Peng, M. Coram, M.C. Stumpe, D. Wu, A. Narayanaswamy, et al., Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, JAMA 316 (2016) 2402–2410.
[31] P. Burlina, K.D. Pacheco, N. Joshi, D.E. Freund, N.M. Bressler, Comparing humans and deep learning performance for grading AMD: a study in using universal deep features and transfer learning for automated AMD analysis, Comput. Biol. Med. 82 (2017) 80–86.
[32] R. Tennakoon, D. Mahapatra, P. Roy, S. Sedai, R. Garnavi, Image quality classification for DR screening using convolutional neural networks, in: Proceedings of the Ophthalmic Medical Image Analysis International Workshop, 2016, pp. 113–120.

[33] K.K. Maninis, J. Pont-Tuset, P. Arbeláez, L. Van Gool, Deep retinal image understanding, in: International Conference on Medical Image Computing and Computer-assisted Intervention, 2016, pp. 140–148.
[34] Q. Li, L. Xie, Q. Zhang, S. Qi, P. Liang, H. Zhang, et al., A supervised method using convolutional neural networks for retinal vessel delineation, in: Image and Signal Processing, International Congress on, 2015, pp. 418–422.
[35] M.J.J.P. van Grinsven, B. van Ginneken, C.B. Hoyng, T. Theelen, C.I. Sánchez, Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images, IEEE Trans. Med. Imaging 35 (2016) 1273–1284.
[36] H.A. Leopold, J. Orchard, J. Zelek, V. Lakshminarayanan, Segmentation and feature extraction of retinal vascular morphology, SPIE Med. Imaging, Int. Soc. Opt. Photonics 101330 (2017).
[37] M. Yates, S. Hayat, R.N. Luben, R. Welikala, A. Rudnicka, C. Owen, et al., Retinal vessel morphometry associations with polymyalgia rheumatica; findings from the European Prospective Investigation of Cancer (EPIC) in Norfolk [abstract], Arthritis Rheumatol. 10 (2016).
[38] UK Biobank. http://www.ukbiobankeyeconsortium.org.uk/, 2017.
[39] R.A. Welikala, M.M. Fraz, P.J. Foster, P.H. Whincup, A.R. Rudnicka, C.G. Owen, et al., Automated retinal image quality assessment on the UK Biobank dataset for epidemiological studies, Comput. Biol. Med. 71 (2016) 67–76.
[40] EPIC-Norfolk, European Prospective Investigation of Cancer (EPIC), 2017. http://www.Srl.Cam.Ac.uk/epic/.
[41] U.T.V. Nguyen, A. Bhuiyan, L.A.F. Park, K. Ramamohanarao, An effective retinal blood vessel segmentation method using multi-scale line detection, Pattern Recognit. 46 (2013) 703–715.
[42] CS231n, Convolutional Neural Networks for Visual Recognition, 2017. http://cs231n.github.io/.
[43] J. Staal, M.D. Abràmoff, M. Niemeijer, M.A. Viergever, B. van Ginneken, Ridge-based vessel segmentation in color images of the retina, IEEE Trans. Med. Imaging 23 (2004) 501–509.

**Roshan Welikala** has received BSc, MSc, and PhD degrees all within the field of imaging science. His research interests are retinal image analysis and pattern recognition. He is currently a Research Associate and a member of the Digital Imaging Research Centre (DIRC) at Kingston University.

**Paul Foster** is Professor of Ophthalmic Epidemiology & Glaucoma Studies at the UCL Institute of Ophthalmology. His research into the epidemiology of glaucoma involves collaborative work with colleagues at Cambridge University working on the EPIC Norfolk cohort, which encompasses primary open angle glaucoma, and involves assessment of the impact of diet, lifestyle and environment on glaucoma.

**Peter Whincup** is Professor of Cardiovascular Epidemiology in the Population Health Research Institute at St George's, University of London. His research interests are in the epidemiology and prevention of coronary heart disease, type 2 diabetes and obesity and particularly in ethnic and geographic differences in these conditions and in the emergence of disease risks at different stages of the life course.

**Alicja Rudnicka** is a Reader in Medical Statistics in the Population Health Research Institute at St George's, University of London. Her recent research has focused on cardiovascular risk factors, physical activity, hierarchical modelling and complex Bayesian meta-analysis in ophthalmic epidemiology

**Christopher Owen** is Professor of Epidemiology in the Population Health Research Institute at St George's, University of London. His recent research has focused on measuring physical activity and fitness and understanding their determinants, especially in children. He also led on a retinal vessel assessment of these children, which has led to further work examining retinal vessel characteristics in other large adult population-based studies.

**David Strachan** is Professor of Epidemiology and the Director of the Population Health Research institute at St George's, University of London. His current international research collaborations are investigating combinations of genetic and environmental influences on disease, particularly respiratory and allergic conditions. Professor Strachan was a member of the UK Biobank Protocol Development Committee in 2002–2003.

**Sarah Barman** is a Professor at Kingston University. Her main area of interest in research is in the field of medical image analysis. Her work is currently focused on research into novel image analysis techniques to enable the recognition and quantification of features in ophthalmic images.