

Diagnosis of Polypoidal Choroidal Vasculopathy From Fluorescein Angiography Using Deep Learning

Yu-Yeh Tsai¹, Wei-Yang Lin¹, Shih-Jen Chen^{2,3}, Paisan Ruamviboonsuk⁴, Cheng-Ho King¹, and Chia-Ling Tsai⁵

¹ Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi Taiwan

² Department of Ophthalmology, Taipei Veterans General Hospital, Taipei, Taiwan

³ School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan

⁴ Department of Ophthalmology, Rajavithi Hospital, Bangkok, Thailand

⁵ Computer Science Department, Queens College, CUNY, Queens, New York, USA

Correspondence: Chia-Ling Tsai, Computer Science Department, Queens College, CUNY, Queens, New York, USA.
e-mail: ctsai@qc.cuny.edu

Received: October 4, 2021

Accepted: January 7, 2022

Published: February 3, 2022

Keywords: polypoidal choroidal vasculopathy; fluorescein angiography; deep learning; screening; segmentation

Citation: Tsai YY, Lin WY, Chen SJ, Ruamviboonsuk P, King CH, Tsai CL. Diagnosis of polypoidal choroidal vasculopathy from fluorescein angiography using deep learning. *Transl Vis Sci Technol.* 2022;11(2):6. <https://doi.org/10.1167/tvst.11.2.6>

Purpose: To differentiate polypoidal choroidal vasculopathy (PCV) from choroidal neovascularization (CNV) and to determine the extent of PCV from fluorescein angiography (FA) using attention-based deep learning networks.

Methods: We build two deep learning networks for diagnosis of PCV using FA, one for detection and one for segmentation. Attention-gated convolutional neural network (AG-CNN) differentiates PCV from other types of wet age-related macular degeneration. Gradient-weighted class activation map (Grad-CAM) is generated to highlight important regions in the image for making the prediction, which offers explainability of the network. Attention-gated recurrent neural network (AG-PCVNet) for spatiotemporal prediction is applied for segmentation of PCV.

Results: AG-CNN is validated with a dataset containing 167 FA sequences of PCV and 70 FA sequences of CNV. AG-CNN achieves a classification accuracy of 82.80% at image-level, and 86.21% at patient-level for PCV. Grad-CAM shows that regions contributing to decision-making have on average 21.91% agreement with pathological regions identified by experts. AG-PCVNet is validated with 56 PCV sequences from the EVEREST-I study and achieves a balanced accuracy of 81.132% and dice score of 0.54.

Conclusions: The developed software provides a means of performing detection and segmentation of PCV on FA images for the first time. This study is a promising step in changing the diagnostic procedure of PCV and therefore improving the detection rate of PCV using FA alone.

Translational Relevance: The developed deep learning system enables early diagnosis of PCV using FA to assist the physician in choosing the best treatment for optimal visual prognosis.

Introduction

Age-related macular degeneration (AMD) is a type of degeneration that gradually leads to significant loss of central vision. The wet form of AMD includes typical AMD with choroidal neovascularization (CNV) and polypoidal choroidal vasculopathy (PCV), which is also a type of neovascular condition associated with abnormal branching vascular network and aneurysmal dilations, referred to as *polyps*. In

the Asian population, 50% of exudative maculopathy involves PCV. Although treatment with antivascular endothelial growth factors by intraocular injection can maintain or improve visual function for patients with either typical AMD or PCV, early diagnosis to differentiate between both diseases may help decide if combination with photodynamic therapy should be applied to those with PCV.¹

Fluorescein angiography (FA) is the major imaging modality for diagnosis and treatment of retinal disorders. Rapid-sequence photographs of the retina are

captured after intravenous injection of the sodium fluorescein dye to reveal fine details of the retina, such as the retinal vessels. When a wet AMD is imaged with FA, some features like nodular hyperfluorescence or massive block fluorescence caused by blood might suggest PCV, but these leaking patterns have low specificity because of similar presentation by occult CNV.² Optical coherence tomography (OCT) and OCT angiography have been studied recently for diagnosing PCV, but the sensitivity varies from 30% to 83%^{3–6} because it lacks the capability of displaying temporal leaking information, which is an important diagnostic feature for differentiating subtypes of neovascular AMD. For this reason, OCT is mainly used for assessment of PCV disease activities before the treatment and during the follow-up period. Indocyanine green angiography (ICGA) remains the gold standard for diagnosing PCV, despite its invasive nature, because of the visualization of choroidal vessels by the near infrared light penetration through pigment epithelium layer and the high protein bound with low dye diffusion. However, FA is widely used as a routine image study for exudative AMD, and not every medical center has the ICGA equipment to further confirm the condition of exudative AMD. If using only FA, clinicians can misdiagnose PCV as occult CNV for 90% of the cases.² For this reason, it is important that a clinician can confidently diagnose the condition of exudative AMD as CNV or PCV using FA alone to avoid incorrect treatment leading to poor visual prognosis.¹

Research on deep learning (DL) for diagnosis of PCV is yet very limited in the literature. Xu et al.⁷ applied a convolutional neural network (CNN) to color fundus photographs (CFP) and OCT separately to learn features and applied fully connected network on combined features for classification of CNV and PCV. Chou et al.⁸ performed a similar study but included manually determined OCT biomarkers instead of the whole OCT images. Yang et al.⁹ and Kim et al.¹⁰ both applied a public-available DL platform (AutoML; Google Inc., Mountain View, CA, USA) to screen PCV using ICGA images. Ma et al.¹¹ and Hwang et al.¹² applied well-known CNN models to differentiate PCV from other forms of AMD using OCT. None of the work involves using DL to make diagnosis of PCV using FA.

The objective of this study is to investigate the efficacy of diagnosing PCV using FA alone by leveraging the technology of DL. Challenges of our work come from the high degree of variation in lesion appearances caused by various speed of circulation of the dye, nonstandard protocol in sequence image acquisition, coexistence of other medical conditions, and low availability of clinical images. To facilitate this study, we developed two attention-gated DL networks

to exploit subtle visual cues of choroidal circulation on FA images for diagnosis of PCV. The first network differentiates PCV from other types of wet AMD, mainly choroidal neovascularization (CNV), and the second network determines the extent of PCV, including both polyps and the branching vascular network, for cases determined as PCV. Both models were assessed at both image-level and patient-level.

Methods

Data Description

Our FA datasets are collected from multiple sources with different imaging protocols: 56 PCV sequences from EVEREST-I study,¹³ 45 PCV and 70 CNV sequences from Taipei Veterans General Hospital (TVGH, Taipei, Taiwan), and 63 PCV sequences from Rajavithi Hospital (RH, Bangkok, Thailand). All sequences are treatment-naïve (no prior treatment), with the disease condition confirmed by standard ICGA imaging.

For all 56 PCV sequences from EVEREST-I, the ICGA and FA images were captured individually or by simultaneous acquisition mode of scanning laser ophthalmoscope (Heidelberg Engineering Inc., Heidelberg, Germany), with fluorescein dye and indocyanine green dye injected separately or together at the same time. For images taken in the simultaneous mode with FA and ICGA shown side by side, the FA and ICGA components are separated as two images for future processing. Each patient has FA images taken at five different time points (i.e., 1.5, 3, 5, 10, and 20 minutes). The set of 115 FA sequences from the Department of Ophthalmology, TVGH, is composed of 31 classic CNV, 39 occult CNV, and 45 PCV. Each sequence has five or more images, taken within 20 minutes of injection. All 63 sequences from the Department of Ophthalmology, RH, are PCV, with each having three images taken mainly in the first five minutes. In total, there are 164 PCV cases with 847 images and 70 CNV cases with 898 images. No image contains traceable patient information or a hospital-specific code.

Attention-Gated Convolutional Neural Network for Screening

Classic machine learning approach requires manual determination of biomarkers or features to be computed, which is termed *feature engineering*. A classifier is built to learn from a training set containing features with known class labels. The recognition power of a classifier mostly relies on the discriminative power of the chosen features. In the era of DL, feature

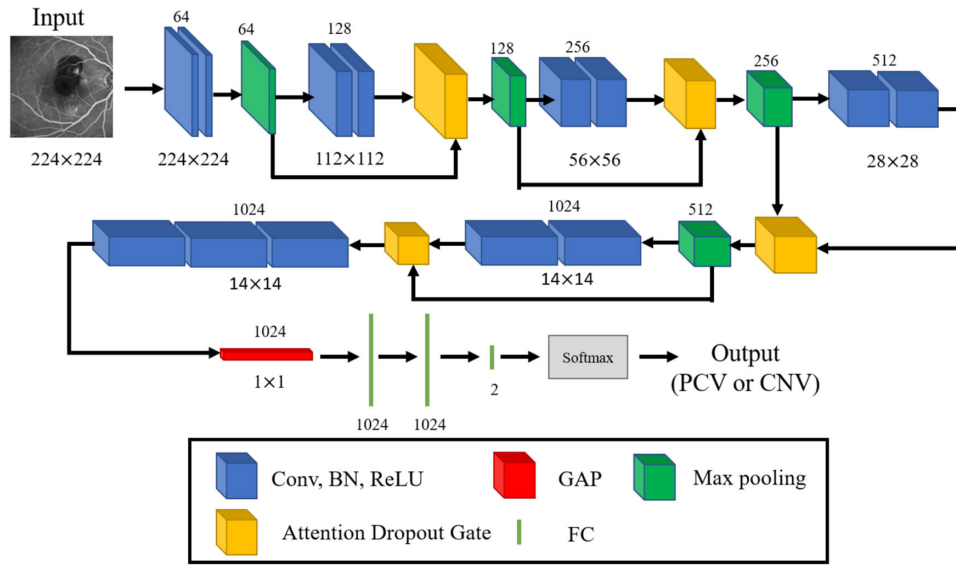


Figure 1. Architecture of AG-CNN.

engineering is fully automated as part of the machine learning process. CNN is a DL model that imitates the central nervous system. It has multiple intermediate layers positioned between the input (the image) and output layers, allowing each level to learn to extract features of increasing abstraction from its input signal during training.

Different from most of the works on PCV in the literature, we design our own CNN with the attention mechanism, named attention-gated CNN (AG-CNN). The architecture is shown in Figure 1. AG-CNN takes one single FA image (of any time point) downsampled to the size of 224×224 and produces the probability of the image being PCV. This is image-level classification. The probabilities of images from the same FA sequence are averaged to produce the patient-level result. If the value is greater than 50%, the patient is classified as PCV. Because AG-CNN works on individual images, not a complete sequence, there is no assumption on the number of images per sequence to accommodate various imaging protocols.

Each blue block in Figure 1 is a convolution layer that applies a various number of mathematical filtering operations (known as *convolution*) to detect discriminative features. Batch normalization (BN) and rectified linear activation function are performed; the former improves the stability for faster convergence, and the latter introduces nonlinearity in the model for learning of more complex functions. The loss function is cross entropy, the optimizer is Adam, the learning rate is 0.0001, the number of epochs is 300, and the batch size is 32. Feature maps of a convolution layer are downsampled using max pooling so features of higher-level representation can be detected in the successive layer. The attention layer is an adaptive mask learned during

the training process to regulate the attention of the network. We implement Attention Dropout Gate¹⁴ so the network can also pay attention to areas being dropped with certain probability to avoid overfitting. The drop rate is 0.25 and the drop threshold is 0.8. Global average pooling (GPA) is applied to the last convolution layer; feature maps are each reduced into a single number through averaging, resulting in 1024 features for each input image. Fully-connected layer is performed to do classification, and the outcome of SoftMax are the disease probabilities, adding up to 100%.

To provide the visual explanation of the network model, we apply the concept of gradient-weighted class activation mapping (Grad-CAM)¹⁵ on AG-CNN to highlight the regions in the image that drive the prediction of a given class of object, so different maps are produced for different classes. In the map,

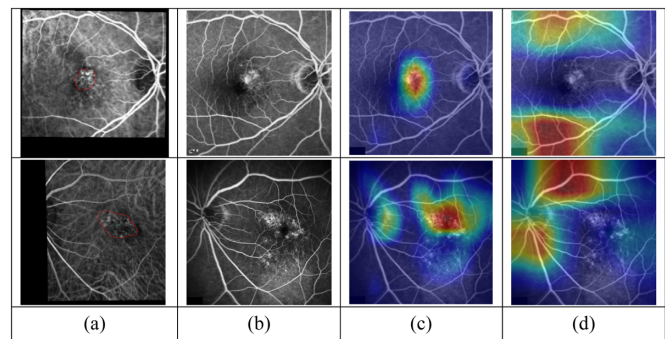


Figure 2. Grad-CAM for explainability of AG-CNN in two cases with PCV. The first row is case 1 and the second row is case 2. (A) ICGA images show delineation of PCV. (B) FA images. (C) Grad-CAM produced from FA by AG-CNN, our proposed DL model. (D) Grad-CAM produced from FA by ResNet-50.

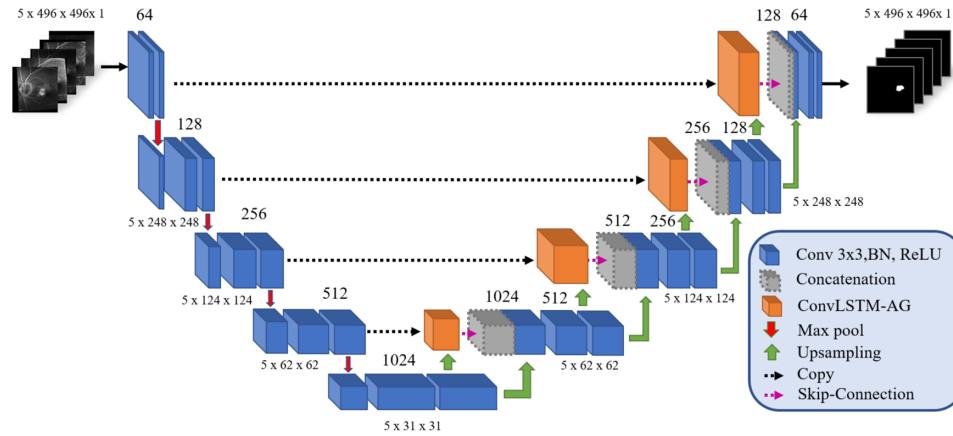


Figure 3. Architecture of AG-PCVNet.

each pixel is assigned a value between [0,1], indicating its contribution. Grad-CAM can facilitate clinical translation if the learning process is pathology driven, not imaging device driven. As shown in Figure 2(c) is the Grad-CAM generated by AG-CNN, and (d) is generated by ResNet-50,¹⁶ which is one of the publicly available networks used in other studies.^{7,11,12} AG-CNN focuses on the region that has a greater overlap with biomarkers of PCV identified by the specialists, whereas ResNet-50 provides much lower explainability for its decision.

Attention-Gated PCVNet for Segmentation

Our proposed method for PCV segmentation is based on the U-Net architecture,¹⁷ which has proven to be effective for semantic segmentation. The name of U-Net comes from its U-shape; the left arm is the encoder for extracting features and the symmetric right arm is the decoder for precise localization. Figure 3 shows our modified architecture, named Attention-Gated PCVNet, which is attention U-Net with ConvLSTM¹⁸ in the decoding path.

Attention-gated PCVNet (AG-PCVNet) takes a sequence of aligned images and outputs a sequence of segmented images. The segmentation is the delineation of the PCV lesion, including branching vascular network and polyps. This is image-level segmentation. The novelty of AG-PCVNet is its ability to work on a sequence end-to-end so the temporal context can also be considered, because ConvLSTM is a type of recurrent neural network for spatiotemporal prediction. We apply ConvLSTM to the attention maps, so the attention map can be determined by the input feature maps and the attention map of the previous frame. Same as for AG-CNN, at each level, the attention map is applied to the input feature maps in the decoding path. The loss

function is a combination of binary cross entropy and dice loss. The latter is found to be effective for training with small datasets.¹⁹ The optimizer is Adam, the learning rate is 0.0005, the number of epochs is 1600, and the batch size is five sequences. For the patient-level segmentation, we take the segmentation result of the last image frame for the given sequence because the network retains the spatiotemporal information from the complete sequence when segmenting the last image frame.

Results

All sequences are involved in the development of the network for screening, which is AG-CNN. We partition the collection to three sets: training, validation and test, each containing images from all three centers. The model is trained with the training set, fine-tuned with the validation set for selection of hyper-parameters, and tested with the test set. Table 1 shows the distributions of images in the three sets. Data augmentation is performed on the minority class of the training set to balance the two classes. All images of a sequence go to the same set to make sure fair testing. Please note that having sequences with highly similar image frames does not give our DL model an extra advantage, because

Table 1. Data Distribution for Training, Validation, and Testing of AG-CNN for Classification

| Disease | Training Set | Validation Set | Test Set |
|---------|--------------|----------------|----------|
| PCV | 514 (102) | 176 (33) | 157 (29) |
| CNV | 559 (42) | 176 (14) | 163 (14) |

Each set is recorded as the number of images with the number of patients in the parentheses.

Table 2. Accuracies of Different Models for Classification of PCV in Both Image-Level and Patient-Level

| | Validation | | | Test | | |
|-----------------------|------------|--------|---------|--------|--------|---------|
| | PCV | CNV | Average | PCV | CNV | Average |
| Model (Image-level) | | | | | | |
| ResNet-50 | 0.7671 | 0.6648 | 0.7159 | 0.7261 | 0.5644 | 0.6438 |
| AG-CNN | 0.8125 | 0.7614 | 0.7869 | 0.8280 | 0.6749 | 0.75 |
| Model (Patient-level) | | | | | | |
| ResNet-50 | 0.9394 | 0.6154 | 0.8478 | 0.8276 | 0.7857 | 0.8140 |
| AG-CNN | 0.8485 | 0.8462 | 0.8478 | 0.8621 | 0.7857 | 0.8372 |

AG-CNN is our proposed model. The PCV detection accuracy shows the sensitivity, and the CNV detection accuracy shows the specificity of a model.

such images provide no additional information in the learning process, but they serve the purpose of data augmentation for imbalanced data.

We compare AG-CNN with the popular network model ResNet-50 and human expert performance documented in the literature. Table 2 shows the classification accuracies of different models in both image-level and patient-level. The PCV detection accuracy is computed as the ratio of the PCV detected to the total PCV images/sequences. This is equivalent to the sensitivity measure, because PCV is considered a positive case. Similarly, the CNV detection accuracy is the ratio of the CNV detected to the total CNV images/sequences. This is equivalent to the specificity measure. Our proposed model, AG-CNN, outperforms ResNet-50 with an average test accuracy of 75.0% at the image-level and 83.72% at the patient-level for both conditions. If considering only PCV, AG-CNN achieves test accuracy of 86.21% at patient-level, which far exceeds human expert performance, because clinicians can misdiagnose PCV as occult CNV for 90% of the cases,² if only using FA. If considering only CNV, classic and occult have the test accuracy of 65.2% and 50%, respectively, at the image-level.

We measure the explainability of the model by computing the dice similarity score (DSC) between the ground truth and the part of the Grad-CAM with values greater than 0.5 using the EVEREST-I dataset. Our dice scores are 0.1966 and 0.2191 for validation and test, respectively, whereas ResNet-50 has 0.1371

and 0.0799. Please note that DSC is only good for making comparison between models but is not a fair measure for explainability, because the model does not learn from the ground truth, and it might only need part of the lesion with strong features for making the decision. It is also possible that the model realizes potential biomarkers not yet investigated.

We evaluate the performance of the network for segmentation, AG-PCVNet, using 56 FA sequences from the EVEREST-I dataset only, because sequences from TVGH and RH are lacking pixel-level ground truth needed for segmentation. The ground truth of the EVEREST-I dataset was provided by the reading center and had been reported for segmentation of polyps in ICGA.²⁰ The ground truth annotation is transferred from the ICGA sequence to the corresponding FA sequence using Edge-driven DBICP²¹ to align images. We perform 5-fold validation for segmentation on the 56 PCV sequences because of low number of sequences involved. The dataset is divided into five almost equal portions; each portion is used in turn for validation whereas the other four portions are used for training. The final reported error is the average of the five validation errors.

We compare the performance of AG-PCVNet with the standard U-Net, attention-gated U-Net (AG-U-Net), which is AG-PCVNet without ConvLSTM, and our network pretrained with ICGA sequences, with the measures of sensitivity, specificity, balanced accuracy (average of sensitivity and

Table 3. Image-Level Segmentation Performance of Different Models

| Model (Image-Level) | Sensitivity | Specificity | Bal. Acc. | DSC |
|----------------------|--------------|--------------|--------------|--------------|
| U-Net | 0.3632/0.127 | 0.9872/0.009 | 0.6752/0.060 | 0.3566/0.064 |
| AG-U-Net | 0.3193/0.063 | 0.9914/0.002 | 0.6553/0.032 | 0.3690/0.070 |
| AG-PCVNet | 0.6054/0.136 | 0.9469/0.027 | 0.7760/0.579 | 0.4009/0.089 |
| AG-PCVNet pretrained | 0.5438/0.127 | 0.9694/0.009 | 0.7566/0.060 | 0.4325/0.119 |

AC-PCVNet is our proposed model. The numbers are recorded as mean/SD of 5-fold cross validation.

Bal. Acc., Balanced Accuracy.

Table 4. Segmentation Performance Measured in DSC of Individual Image Frames

| Model | 1.5 min. | 3 min. | 5 min. | 10 min. | 20 min. |
|----------------------|----------|--------|--------|---------|---------|
| U-Net | 0.3479 | 0.3568 | 0.3791 | 0.3591 | 0.340 |
| AG-PCVNet pretrained | 0.4329 | 0.4330 | 0.4320 | 0.4325 | 0.4325 |

AG-PCVNet is our proposed model.

specificity), and DSC. For the pretrained model, training using ICGA images takes 1600 epochs, and tuning using FA images takes 300 epochs. Table 3 shows the results. Our network model pretrained with ICGA sequences gives the best average performance of 0.4325 for DSC. For cases with larger ground truth areas, the DSC tends to be higher, with the maximum reaching 0.88. Because there are many more cases with small ground truth areas, the average is brought down to 0.4325.

We also compute the DSC of individual frames, shown in Table 4, to study the effect of information passing among frames for an improved performance. Compared to U-Net, our network model achieves a more consistent image-level segmentation with a difference of 0.001 in DSC value between the best and the worst time frames.

Discussion

Our study demonstrated the efficacy of using deep learning model for diagnosis of PCV from FA alone for patients with exudative AMD. Although recent deep learning models trained from fundus color photographs and OCT can also offer a prediction of PCV, diagnosis of PCV from FA not only allows additional identification check point but also provides segmentation of the total lesions.

We compared different algorithms of deep learning for screening. AG-CNN substantially outperforms the generic ResNet-50 in both accuracy and explainability. AG-CNN performs slightly better in patient-level than in image-level for PCV but substantially better in patient-level for CNV with an improvement of about 11%. Although image-level success is important for the development of a DL model, this experiment stresses the importance of patient-level analysis in a clinical setting because not one single image in a temporal sequence, like FA, contains all information needed for the most accurate diagnosis. When results from all images are integrated in any fashion, which can be as simple as averaging, the accuracy improves, compared to the image-level.

Grad-CAM allows visualization of where AG-CNN is looking to make sure the decision is medically sound. Examples are shown in Figure 4 to Figure 6. Figure 4 shows an example of a PCV case being successfully classified in both image-level and patient-level. The polyp, shown in ICGA, is at the center under the massive submacular hemorrhage with mild leakage at late phase FA. Grad-CAM of AG-CNN highlights the area of interest with leakage at different times for the correct diagnosis of PCV. Figure 5 shows a case with polyps underneath the pigment epithelial detachment (PED) in ICGA whereas FA shows occult

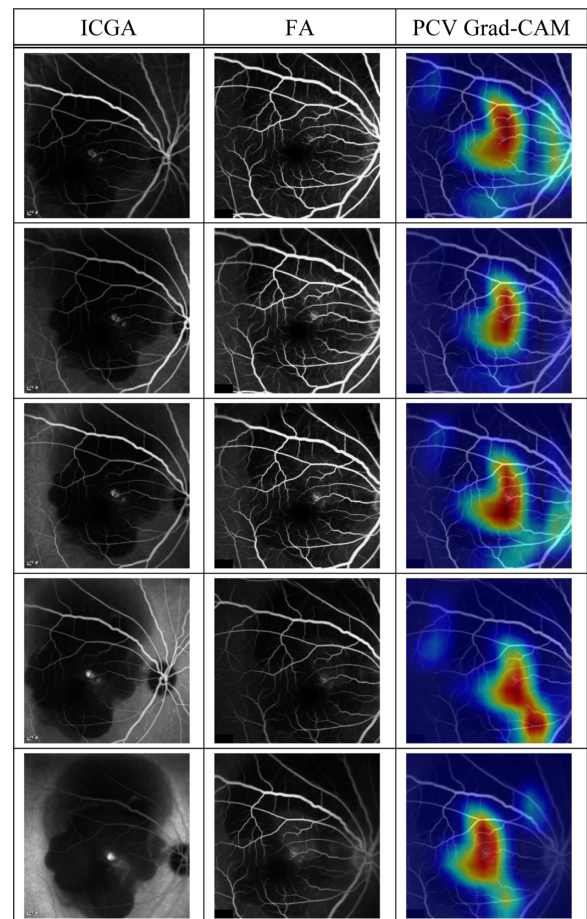


Figure 4. An FA sequence of PCV correctly classified at both image-level for all images and patient-level for the sequence in an eye with PCV with hemorrhagic PED and subretinal hemorrhage. Rows are images taken at 1.5, 3, 5, 10, and 20 minutes, respectively.

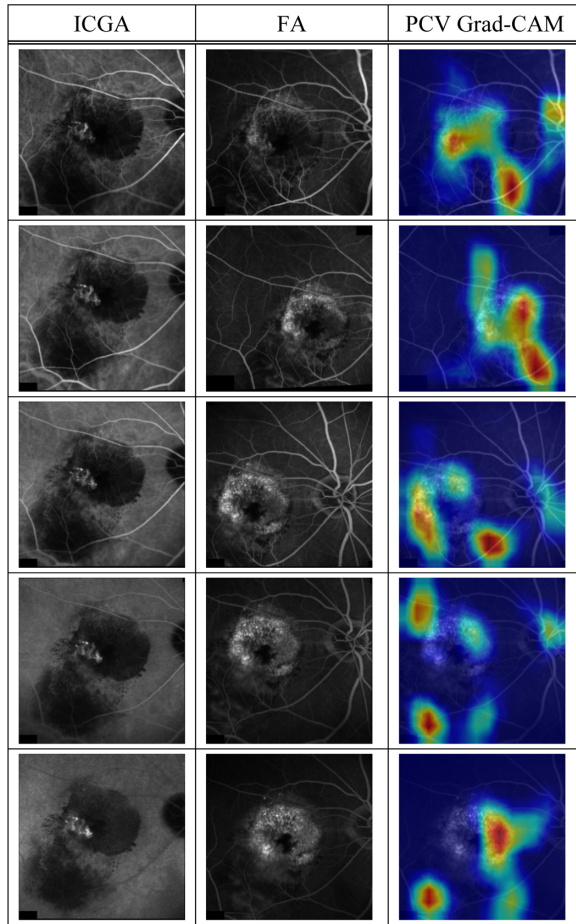


Figure 5. An FA sequence of PCV correctly classified at both image-level for all images and patient-level for the sequence in a patient with leaking fibrovascular PED. Rows are images taken at 1.5, 3, 5, 10, and 20 minutes, respectively.

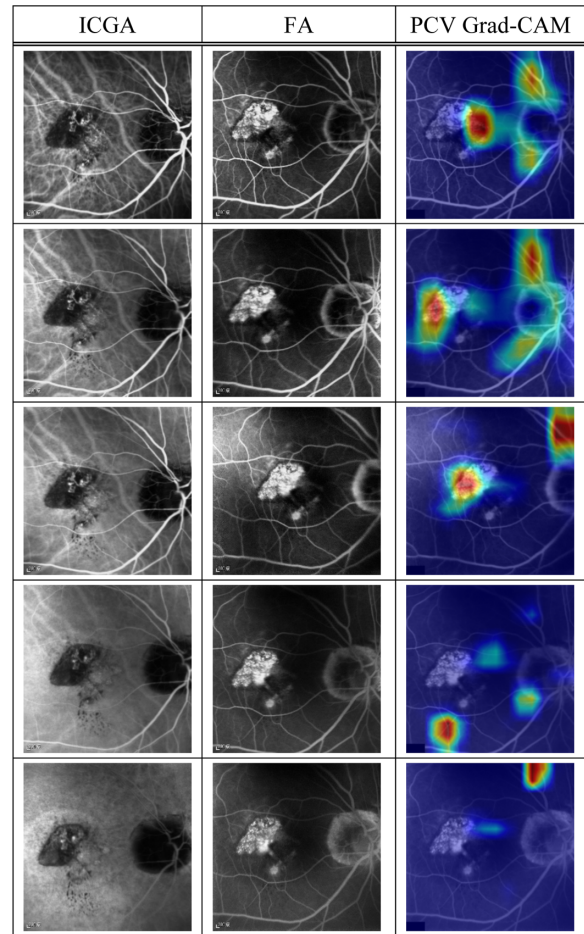


Figure 6. An FA sequence of PCV having the appearance of classic CNV. Rows are images taken at 1.5, 3, 5, 10, and 20 minutes, respectively. Images taken at 1.5 and 3 minutes are correctly classified as PCV but were classified as CNV at 5, 10, and 20 minutes. The sequence is incorrectly classified at patient-level because the average probability is below 50% for PCV.

CNV with fibrovascular PED. Although AG-CNN correctly classifies this case as PCV at the image and patient-levels, the Grad-CAM does not focus as well on the lesion for all images of the sequence. Grad-CAM may also facilitate error analysis. Figure 6 shows another case of PCV underneath a shallow trapezoid PED whereas FA shows early hyperfluorescence with late leakage confined to the area of PED. AG-CNN correctly classifies this case as PCV in the first two images but classifies this case as CNV at the patient-level because the last three images are misclassified, which brings the average probability for PCV to below 50%.

Unfortunately, Grad-CAMs for CNV are not as informative, because they are more like Grad-CAMs generated by ResNet-50. This might explain the lower accuracy for CNV (see Table 2) and can be attributed to insufficient number of cases for multiple types of CNV with substantial appearance variation. However, this study should not be confused with our earlier work

on segmentation of classic CNV,²² which achieved an average segmentation accuracy of 83.26% when given only cases of classic CNV. The current study does not include the same range of disease conditions.

PCV is known to be easily misdiagnosed as occult CNV in FA.² However, to the best of our knowledge, no studies in the literature of ML for PCV diagnosis report the composition of subtypes (classic vs. occult) of the CNV dataset for classification of PCV, even for CFP and OCT modalities. In the early phase of the current study, we tested AG-CNN with only 25 CNV sequences—19 classic and six occult (with a ratio of close to 3:1). We achieved average accuracies of 88.59% and 89.69% for validation and test, respectively. After expanding the dataset to include 31 classic and 39 occult CNV (with a ratio close to 3:4), the accuracies are reduced to 78.69% and 75.0% respectively, as reported in Table 2. The reason for the reduced

performance is that occult CNV has an appearance closer to PCV than to classic CNV. This demonstrates the importance of data composition for development of an DL algorithm for a clinical application, because such development is data-driven; if the algorithm was developed with mostly classic CNV cases, the performance can degrade substantially for a mix of CNV subtypes in a typical clinical setting.

Recent studies on diagnosing PCV using OCT, combined with fundus photograph or optical coherence tomography angiography (OCTA), have reported expert performance of between 82% and 83%⁴⁻⁶ sensitivity based on features of sharply peaked PED, hyper-reflective ring, and complex retinal pigment epithelium (RPE) elevation. Without the diagnostic features predetermined, AG-CNN can boost expert performance from 10%² to 86.21% at the patient-level using FA images alone.

Because the EVEREST-I dataset is complete with both FA and ICGA images for all sequences, we are able to study the effect of transfer learning from ICGA to FA for segmentation. For AG-PCVNet, the performance is boosted slightly in the DSC if the network is pretrained with ICGA sequences for a higher specificity but dropped slightly in sensitivity, as shown in Table 3. The improvement of 3.16% in DSC with pretraining is statistically significant ($P < 0.05$). In other words, ICGA and FA do share some common features such that features learned in ICGA can guide the network when being fine-tuned with FA sequences for better performance.

We also examined the agreement between the size of the detected region and the size of the ground truth PCV with a Bland-Altman plot, which shows the difference between two areas as a function of the mean of the two areas. The difference is computed by subtracting the ground truth area from the detected area. As shown in Figure 7, AG-PCVNet tends to produce an area larger than the ground truth, since there are 16 more images (of 280) with the area difference above 0.

When a retina specialist examines a condition of exudative AMD, characteristics of the fluorescein leakage pattern of a complete FA sequence should be considered, because not one single image captures all the visual clues for making a proper diagnosis. As shown in Table 4, U-Net performs better in five-minute on average, which might not be true for all sequences. Figure 8 shows three image frames from the same sequence, but U-Net has the highest DSC score for the image at 1.5 minutes. The choice of the best image frame can be patient-dependent for U-Net. On the contrary, AG-PCVNet retains information from earlier frames in the same sequence to achieve a more consistent image-level prediction—all

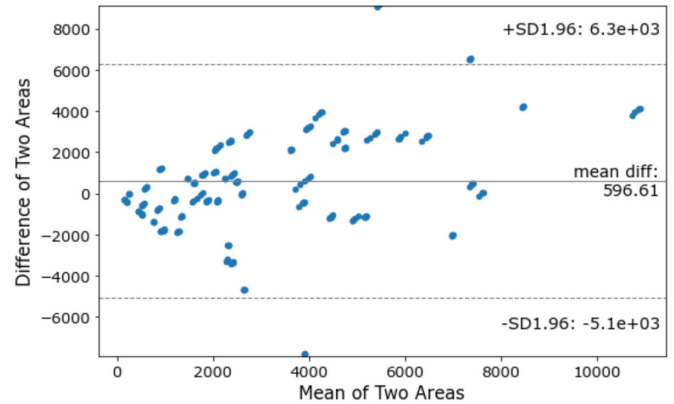


Figure 7. Bland-Altman plot. The plot shows the agreement between the size of the ground truth PCV and the size of the segmented area by AG-PCVNet. The areas are expressed in pixel.

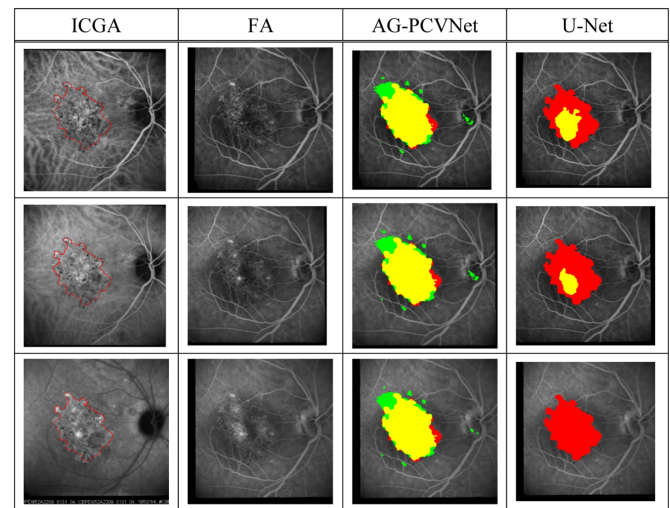


Figure 8. Segmentation result. The third column contains outcomes of AG-PCVNet, and the last column contains outcomes of U-Net. The first row is for 1.5 minutes, and second row for 3 minutes and the third row is for 5 minutes. Yellow patches are correct segmentation, red patches are undersegmentation, and green patches are oversegmentation. AG-PCVNet achieves an average DSC of 0.88, whereas U-Net achieves only 0.16 for this case.

phases have a DSC score very close to 0.4325, as shown in Table 4. Figure 8 also shows very consistent segmentation outcomes produced by AG-PCVNet in various phases.

For the future work, screening and segmentation should be combined so one single end-to-end network can take an aligned sequence and explore spatiotemporal information for both screening and segmentation to produce a segmented region if the case is confirmed as PCV. The data collection should be expanded to contain more cases of CNV of various types to improve the detection rate of CNV and more annotated PCV cases to train the component for segmentation to bring

the model one step closer to becoming a useful diagnostic tool in clinical practice.

Conclusions

The developed software provides a means of performing screening and segmentation of PCV on FA images for the first time. Screening of FA exemplified in our study far exceeds expert performance and achieves an accuracy of 86.21% at the patient-level by deep learning. This study is a promising step in supplementing the diagnostic procedure of PCV and therefore improving the detection rate of PCV using FA. In addition, explainability of the deep network computation offered by Grad-CAM can potentially shed light on novel biomarkers associated with pathophysiology of PCV.

Acknowledgments

The authors thank the investigators in the EVEREST trial for sharing the image data: Adrian Koh, Won Ki Lee, Lee-Jen Chen, Hakyoung Kim, Timothy Lai and Tock-Han Lim.

Supported by grants from the Ministry of Science and Technology, Taiwan (MOST 105-2221-E-194-049) and PSC-CUNY Research Award 64450-00-52.

Disclosure: **Y.-Y. Tsai**, None; **W.-Y. Lin**, None; **S.-J. Chen**, Bayer (C), Novartis (C), Roche (C), Allergan (C); **P. Ruamviboonsuk**, None; **C.-H. King**, None; **C.-L. Tsai**, None

References

1. Koh A, Lai TY, Takahashi K, et al. Efficacy and safety of ranibizumab with or without verteporfin photodynamic therapy for polypoidal choroidal vasculopathy: a randomized clinical trial. *JAMA Ophthalmol*. 2017;135:1206–1213.
2. Tan CS, Ngo WK, Lim LW, Tan NW, Lim TH. EVEREST study report 4: fluorescein angiography features predictive of polypoidal choroidal vasculopathy. *Clin Exp Ophthalmol*. 2019;47:614–620.
3. Carlo TE, Kokame GT, Kaneko KN, Lian R, Lai JC, Wee R. Sensitivity and specificity of detecting polypoidal choroidal vasculopathy with en face optical coherence tomography and optical coherence tomography angiography. *Retina*. 2019;39:1343–1352.
4. Cheung CMG, Lai TYY, Teo K, et al. Polypoidal choroidal vasculopathy: consensus nomenclature and non-indocyanine green angiograph diagnostic criteria from the Asia-Pacific Ocular Imaging Society PCV Workgroup. *Ophthalmology*. 2021;128:443–452.
5. Cheung CMG, Yanagi Y, Akiba M, et al. Improved detection and diagnosis of polypoidal choroidal vasculopathy using a combination of optical coherence tomography and optical coherence tomography angiography. *Retina*. 2019;39:1655–1663.
6. Chaikitmongkol V, Khunsongkiet P, Patikulsila D, et al. Color fundus photography, optical coherence tomography, and fluorescein angiography in diagnosing polypoidal choroidal vasculopathy. *Am J Ophthalmol*. 2018;192:77–83.
7. Xu Z, Wang W, Yang J, et al. Automated diagnoses of age-related macular degeneration and polypoidal choroidal vasculopathy using bi-modal deep convolutional neural networks. *Br J Ophthalmol*. 2021;105:561–566.
8. Chou Y-B, Hsu C-H, Chen W-S, et al. Deep learning and ensemble stacking technique for differentiating polypoidal choroidal vasculopathy from neovascular age-related macular degeneration. *Sci Rep*. 2021;11:7130.
9. Yang J, Zhang C, Wang E, Chen Y, Yu W. Utility of a public-available artificial intelligence in diagnosis of polypoidal choroidal vasculopathy. *Graefes Arch Clin Exp Ophthalmol*. 2020;258:17–21.
10. Kim IK, Lee K, Park JH, Baek J, Lee WK. Classification of pachychoroid disease on ultrawide-field indocyanine green angiography using auto-machine learning platform. *Br J Ophthalmol*. 2021;105:856–861.
11. Ma D, Kumar M, Khetan V, et al. Differential diagnosis between polypoidal choroidal vasculopathy (PCV) and age-related macular degeneration (AMD) using Deep Neural Network. *Invest Ophthalmol Vis Sci*. 2020;61:2024.
12. Hwang D, Choi S, Ko J, et al. Distinguishing retinal angiomatous proliferation from polypoidal choroidal vasculopathy with a deep neural network based on optical coherence tomography. *Sci Rep*. 2021;11:9275.
13. Koh A, Lee WK, Chen L-J, et al. EVEREST study: efficacy and safety of verteporfin photodynamic therapy in combination with ranibizumab or alone versus ranibizumab monotherapy in patients with symptomatic macular polypoidal choroidal vasculopathy. *Retina*. 2012;32:1453–1464.

14. Choe J, Lee S, Shim H. Attention-based dropout layer for weakly supervised object localization. *IEEE Trans Pattern Anal Mach Intell.* 2021;43:4256–4271.
15. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In *International Conference on Computer Vision*, 2017:618–626.
16. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Conference on Computer Vision and Pattern Recognition*, 2016:770–778.
17. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer Assisted Intervention*. Cham: Springer; 2015:234–241.
18. Shi X, Chen Z, Wang H, Yeung D-Y, Wong W-K, Woo W-C. Convolutional LSTM Network: a machine learning approach for precipitation now-casting. In *International Conference on Neural Information Processing Systems*. 2015:802–810.
19. Milletari F, Navab N, Ahmadi SA. Vnet: fully convolutional neural networks for volumetric medical image segmentation. In *International Conference on 3D Vision (3DV)*. 2016:565–571.
20. Lin W-Y, Yang S-C, Chen S-J, Tsai C-L, Du S-Z, Lim T-H. Automatic segmentation of polypoidal choroidal vasculopathy from indocyanine green angiography using spatial and temporal patterns. *Transl Vis Sci Technol.* 2015;4(2):7.
21. Tsai C-L, Li C-Y, Yang G, Lin K-S. The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence. *IEEE Trans Med Imaging.* 2010;29:636–649.
22. Tsai C-L, Yang Y-L, Chen S-J, Lin K-S, Chan C-H, Lin W-Y. Automatic characterization of classic choroidal neovascularization by using adaboost for supervised learning. *Invest Ophthalmol Vis Sci.* 2011;52:2767–2774.