A Deep Composite Predict Network for Eye Disease

Weiyi Kong\textsuperscript{1a}, Menglong Yang\textsuperscript{2} and Qinzheng Huang\textsuperscript{1*}

\textsuperscript{1}Electrical Information Engineering, Southwest Minzu University, Chengdu, Sichuan, 610041, China
\textsuperscript{2}Aeronautics and Astronautics, Sichuan University, Chengdu, Sichuan, 610041, China
\textsuperscript{a}21500004@swun.edu.cn  \textsuperscript{*}Weiyi Kong: vechiry@163.com

Abstract. Different areas of eyes reflect different states of internal organs of human body. That can assist to detect the disease internal organs, facilitate the prevention and cure by analyzing the image of human eyes. In this paper, we propose a new algorithm, Deep Composite Predict Network (DCP-Net), combined deep learning technique with traditional Chinese medicine theory, which can segment interesting areas of the human eyes, detect the disease eyes texture and then predict internal lesion organ. Aimed at helping doctors to diagnose quickly. The proposed algorithm consists three sub-modules, i.e., multiple segmentation sub-module, fast detection sub-module and predict eye disease module. The multiple segmentation module is solved by the improved retraining U-network, which can obtain an accurate eye to segment interesting areas. The fast detection module is based on enhanced Yolo framework. The predication disease method combined Chinese traditional theory with computer vision technology. Finally, our network can accurately detect the disease of human eyes and predict lesions of internal organs, which can help doctors quickly diagnose lesion organs by human eyes.

1. Introduction

Biometric identification is one of the most reliable identification technology recently [27, 28]. In biometric identification field, human eye research has been a hot spot because it is unique, stable and anti-counterfeiting [18]. As early as 1987, ARAN SAFIR, an ophthalmologist, proposed the concept of using iris automatic recognition. Nevertheless, the technique developed slowly until neural network appeared, many researchers combined biometric identification problem with computer vision technology that have done a lot of effective and fruitful work [22, 23]. Hence, how to combine biomedical theory with machine learning technology has become a hot topic in the fields of medicine and computer. In addition, Professor Peng JingShan discovered the correlations between difference regions in eyes and human inner organs based on the Traditional Chinese Medicine Theory(TCM) [10], so he divided the eye into twelve regions corresponding to different internal organs, which is the same as western medicine called eye diagnosis method. TCM proves that human eyes connect with cerebrospinal nerve vector, which means if human body have organ lesions that will appear in the eye-spot, vasoconstriction, color, and other changes. Previous medical technology used traditional image segmentation technology [2, 16, 17], cannot accurately segment interesting textures of eye region, neither detect eyes lesion categories or predict lesion location. To solve these problems, we innovation combine the TCM with deep learning method and propose a deep composite predict network (DCP-Net) which is a groundbreaking medical application in artificial intelligence. As far as know, we are
the first team to use deep learning algorithm in computer vision work on eye disease prediction problems.

Our contributions including: The first is that we segment interesting areas in human eyes accurately (99.97%) and obtain multiple eye textures with subpixels. The second is that we present a framework to detect diseases in the human eyes that based on a deep learning algorithm. The last is that we combine TCMT with the above algorithms to present the deep composite predict network (DCP-Net), which can predict the health of internal organs through diseased eyes. In conclusion, we construct an innovative disease prediction system by deep learning technology and biomedical theory.

2. Related work

With the emerge of convolutional neural networks (CNNs), near radiologist level performance can be achieved in automated medical image analysis tasks including cardiac segmentation and cancerous lung nodule detection. High representation power, fast inference, and filter sharing properties have made CNNs the standard for image segmentation. In CNNs, more and more models are proposed to be used to segment medical images, such as full convolutional network FCN [8]. These network models [6, 13, 19] have good expressive ability and can be used to roughly segment some large medical images, but the results are not good for some small medical images [29]. Other researchers have used neural networks for detection tasks [25, 27] such as detecting vehicles and detecting faces. With the development of modern deep Convnets, object detectors like R-CNN showed dramatic improvements in accuracy. R-CNN adopted a region proposal-based strategy in which each proposal was scale-normalized before classifying with a Convnet. Recent and more accurate detection methods like Fast R-CNN and Faster R-CNN advocate using features computed from a single scale, that offers a good trade-off between accuracy and speed. Multiscale detection, however, still performs better, especially for small objects. However, few methods are used to detect medical images because detect network work on these medicine task cannot have good results. In this paper, we propose an innovative algorithm work on eye diseases, called deep Composite predict network (DCP-Net). DCP-Net including three sub-modules. The first module segment eye diseases with three levels that can output interesting eye area. The second module detect disease boxes in real time. The thirdly module predict human inner organ disease connect to the first two sub-modules.

2.1. Segment Neural Network

Convolutional neural networks (CNNs) have achieved great success on many object recognition problems in computer vision community [9, 15, 24, 29]. Many researchers followed this trend and proposed to utilize various CNNs for learning feature representations in the application of liver and lesion segmentation. For example, that can be used in CT areas about medical and can also be used in image processing segmentation [4, 25]. During the 2017 ISBI Li TS challenge, Han, proposed a 2.5D with 24-layers FCN model to segment liver tumors, where the residual block is employed as the repetitive building blocks and the U-Net [28] connection was designed across the encoding part and decoding part. 2.5D refers to using 2D convolutional neural network with the input of adjacent slices from the volumetric images. A large number of previous studies have shown that this convolution-based neural network is suitable for the field of medical images. However, this method also has disadvantages, such as the difficulty in image processing with tiny texture details and the imprecision of some medical image segmentation results. In this paper, we use improved deep convolutional u-network to train the system to obtain segmentation interesting eye area (which is to separate the inner texture region of the human eye image from the non-texture region of the eyelid).
Fig. 1: This diagram shows the model that divides the region of interest of the human eye. First, the one-dimensional image of human eyes is mapped to the three-dimensional space dimension, and then features of different levels are extracted through deconvolution at different expansion rates. Then, deconvolution is carried out at a symmetric rate, and the up-sampling operation is carried out. Finally, the predicted image of human eyes background (the second layer of gray image) and the final segmentation image (the top-level segmentation image) are output.

2.2. Detect Framework

Image features are important for detection [27]. For the detection of diseased human eyes, the darknet used in the series of yolo algorithms having been proposed to solve the detection problem, which has clear structure and good real-time performance [5, 14]. The first detection generation network mainly extracts image features through the convolution layer and uses the full connection layer to predict the class and coordinates of the b-box. Yolo network was inspired by development of GoogleNet work, however, different from inception modules of GoogleNet, 1 1 dimension reduction layer and 3×3 convolution layers are used in the YoloV1, and it has 24 convolution layers and 2 full connection layers. Fast Yolo has 9 convolution layers, finally outputs 3 tensors with 7 times. Yolo network has fewer of these filters, so they [3, 11, 12] have fast detection capability. However, there are some disadvantages about it, i.e. the object detection accuracy of Yolo is lower than that of other state-of-the-art object detection systems, which produces a positioning error and cannot detects small objects. To solve the above problems, the researchers improved the second-generation network. The second detection network-YoloV2, eight tricks method was proposed to in terms of precision was improved, mainly in terms of speed of the network architecture of proposed a more concise Darknet-19 to reduce the network parameters, speed, and the trained classifier of YoloV1 conducted to improve the process of feature extraction, which improve the precision of the network, and then in Darknet-19 testing network which constructed on the basis of the second generation network learning Faster-R-NN prior box of RPN network strategy. Based on the Yolo algorithm development, we use a new detection framework which better than YoloV1 and YoloV2, propose three scales to obtain detection details in eye textures.

3. Deep Composite Predict Network(DCP-Net)

Our DCP-Net including three sub-modules, i.e., multiple segmentation sub-module, fast detection sub-module and predict disease module. Finally, the DCP-Net can detect disease eyes accurately and predict inner organs in a real time. Even more noteworthy is our DCP-Net can use to assist the doctor to judge the disease quickly. The process we shown in Fig.2.
3.1. Multiple Segmentation sub-module

In our work, the first step is to set model of data background segmentation and normalization processing, input images for the model, the role is to make the image prediction interest area and background-independent distribution, the establishment of a basic prediction of human eye interesting area model. The specific implementation is as follows, human eye region is transformed from the rectangular coordinate system to the polar coordinate system model.

\[ I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \]  

And,

\[
\begin{align*}
  x(r, \theta) &= (1 - r)x_p(\theta) + rx_f(\theta) \\
  y(r, \theta) &= (1 - r)y_p(\theta) + ry_f(\theta),
\end{align*}
\]

where \( I(x, y) \) represents the rectangular coordinate system of the initial human eye region, \( I(r, \theta) \) represents the corresponding normalized polar coordinate system. \((x_p, y_p)\) represents the pupil boundary coordinate vector of human eye, \((x_f, y_f)\) is the coordinate vector representing the boundary of the region of interest. This step is prepared for the part of the partition localization of the experiment.

The second step is to train improved a deep encode and decode network, which shape like letter big U. And interesting area frequency weight and segment background boundary area with interesting area center are defined as follows,

\[ E = \sum_{x \in \Omega} w(x) \log(p_i(x)), \]

And,

\[ w(x) = w_0(x) + \exp \left( - \frac{d_1(x) + d_2(x)}{2\sigma^2} \right) \cdot w_0, \]

where \( w_0 \) represents frequency weight, \( d_1 \) represents the nearest distance from the pixel in the background with no interest to the inner interest area, \( d_2 \) represents the second near distance to the boundary. Respectively, in our interest area segmentation experiment, \( w_0 \) sets to 10 and the standard deviation sets to 5. The preliminary background prediction segmentation experiment is shown in Fig.3. In order to expand dataset and enhance the robustness about the final model, we disturbed input training set by image enhancement method, we build model on images width elastic respectively, then uses the random rotation on them, including the level of translation, vertical translation, image scaling, image of the color of the overall transformation, flip horizontal operation, upside down specific operation parameters for the rotation range sets to 0.02, width disturbance factor sets to 0.5, the zoom disturbance coefficient sets to 0.5. Specifically, the high dimensional images obtained in the previous step are sampled at multiple levels to obtain more internal features of the images, which are mainly
divided into three layers, all of which are through multiple 3×3 cascade convolution. Then, the inner
texture features of the image are output through the sampling operation on the deconvolution of
symmetric expansion frequency, in which the middle layer will output the background prediction
image of the image, and the final depth segmentation model will output the segmentation image of the
region of interest of the human eyes, which is used in the detection sub-module below. In our work,
we trained thirty thousand times eye disease dataset and test on twenty percent dataset, accuracy of
the segment interest region of disease eyes dataset is 99.97%, which also provides a feasible model for
the next sub-module for detection. The segment frame work we shown in Fig.1.

Fig. 3: The results about interesting area of human eyes

3.2. Fast Detection sub-module
In this section, we use the output model from segmentation and for detect eye disease in real time, we
adopt anchor-box method to label six types eye diseases. The k-means-based clustering method is used
to predict the location of human eye diseases. The specific operation is shown in Fig.4.

Fig. 4: Fast Detection sub-module framework.

3.2.1. Clustering of Eye Disease Labels
In this part, we explained how to handle the eye disease label. First to tag is used to detect eye disease,
we adopt the method of K-means clustering, then label the application of clustering, we set up six
types of common eye disease and represents them use number 0 until 5, i.e., 0 means color spots(CS),
1 means color in plaque(PQ), 2 means black eye spots(BS), 3 means black eye patches(BP), 4 means
blood-shot eyes(BS), 5 means abnormal texture lines(AL). Finally, we shown the results of detection
of eye diseases. In part 4, we analyze the performance of detection network. The next step is to test the
diseased eyes with good labels, we adopt the RPN method is a kind of sliding window object detector,
it has nothing to do with the class in the operation of RPN, use a small child to evaluate network on
the sliding window with a dense, in single dimension convolution feature mapping operations, using
binary classification and boundary box of regression, we use a 3×3 convolution layer and two same
sizes 1×1 convolution, reference for classification and regression of anchor box respectively determine
the object standard and boundary box of return target (the eye lesions) in order to cover different types
of pathological changes, the anchor has multiple forward dimension and aspect ratio Among them, we
use the k-means function as the loss function of the clustering center. K-means clustering is an iterative technique to segment an image into K-pieces. First, randomly select k objects from n data objects as clustering center. Second, for other objects, according to their similarity (distance) to the cluster center, they are assigned to the cluster that is most similar to them. Third, calculate the center of each obtained new cluster (mean of all objects). And then, repeat steps of second and third until convergence (clustering no longer changes). Distance calculation adopt absolute deviation or deviation square between pixel and cluster center algorithm. (Deviation method use pixel color, brightness, texture, position, or a weighted combination which K is random). If represents a sample set, and each sample contains D-dimensional vectors, the objective function of the algorithm is

$$J = \sum_{i=1}^{K} \sum_{\mathbf{x}_i \in C_i} \|\mathbf{x}_i - C_i\|^2,$$

(5)

Where $C_i$ is the clustering center of class samples, c represents the class i sample set, and K represents the number of sample categories. The process of K-means algorithm used in our work, first, Set the number of categories k of samples, initialize the clustering center, and then set the number of iterations $l = 0$. second, for each sample in the sample set, it is classified into the category to which the nearest cluster center belongs.

When the situation is as follows:

$$\|\mathbf{x}_p - C_i^{(l)}\| < \|\mathbf{x}_p - C_j^{(l)}\|, i = 1, 2, \ldots, K,$$

(6)

in which $i \neq j, \mathbf{x}_p \in C_j^{(l)}$, we update the clustering process as:

$$C_i^{(l+1)} = \frac{1}{N_i} \sum_{\mathbf{x} \in C_i^{(l)}} \mathbf{x}, \quad i = 1, 2, \ldots, K,$$

(7)

where $N_i$ represents the number of class i samples.

If the above conditions are not met,

$$C_i^{(l+1)} \neq C_i^{(l)} (i = 1, 2, \ldots, K),$$

(8)

until $l = l + 1$ while back to the forward step. Otherwise, stop the iteration and end the algorithm.

3.3. Predict Disease Module

In the module of eye disease prediction, our algorithm is based on the segmentation results above, and the detection network cascade. In the eye disease theory of traditional Chinese medicine [27], the human eye is divided into 12 fan-shaped regions, and the lesions in each region correspond to the pathological changes of human internal organs [29]. Generally speaking, if lesions occur in the corresponding regions of the human eye, the corresponding internal organs may have lesions. In the previous submodule, we use the data set of the established tag to get the region of interest for detection through deep segmentation sub-network. Next, we used the third generation of yolo series detection framework to detect the texture of human eye lesions, the improved yolo algorithms which based on darknet framework. We use yolov3 framework because of its clear structure and good real-time performance in this part of the detection and prediction cascade module. The loss function consists of four parts: coordinate prediction, box confidence prediction with target, box confidence prediction without target and category prediction. which $\lambda_1$ represents the coordinate error, $\lambda_2$ represents the predict error with object boxes $\lambda_3$ represents the predict error within object boxes and $\lambda_4$ represents the classification error. Here, we break it down in detail,

$$\lambda_1 = \lambda_{box} \sum_{i=0}^{B} \sum_{j=0}^{P} \left[ (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] + \lambda_{reg} \sum_{i=0}^{B} \sum_{j=0}^{P} \left[ (\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \right],$$

(9)

in which, the above equation is the coordinate prediction error, including the loss of the predicted central coordinate and the loss of the predicted boundary box height, $\lambda$ used to determine which kind of lesion the j category prior box corresponds to $o$ represents the target of lesions.
The above equation represents prediction of the confidence of the detection box containing the target.

$$\lambda_2 = \sum_{i=0}^{s^2} \sum_{j=0}^{B} \sum_{l_j} [(e_i - \hat{e}_i)^2],$$  \hspace{1cm} (10)$$

The above equation represents prediction of confidence of detection frame without lesion targets.

$$\lambda_3 = \lambda_{nou} \sum_{i=0}^{s^2} \sum_{j=0}^{B} (p_i(e) - \hat{p}_i(e))^2,$$  \hspace{1cm} (11)$$

The above equation represents the category of lesions used for prediction and lo is used to determine whether there is a lesion within the prior box.

$$\lambda = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4,$$  \hspace{1cm} (13)$$

Finally, we add all sub-losses together to detect eye diseases. After accurate segmentation and detection of human eye disease, we will describe how to predict disease in the human internal organs. Since the research focus of this paper is to use the deep learning technology of image processing to process human eye images, we do not elaborate on the medical theory, interested scholars can read by themselves. In traditional Chinese medicine [7, 21], each region corresponds to different internal organs of the human body. If we detect any region with abnormal texture changes, it means that the internal organs of the corresponding region have lesions. Area 1 represents the heart, area 2 represents the face, area 3 represents the throat, area 4 represents the upper back, area 5 represents the lower back, area 6 represents the pelvis, area 7 represents the lower abdomen, area 8 represents the upper abdomen, area 9 represents the chest, area 10 represents the neck, and area 11 represents the physiological brain. The regional division is shown in the Fig.5.

4. Experiment

In our work, through many experiments and improvements, we finally obtained a model that can be used to segment and detect human eyes, which can achieve the prediction of organ diseases.

4.1. Dataset details

We evaluate our method on the database of patients with eye diseases in real hospitals. Due to protecting patients' privacy, we could not specify the information on hospitals and patients. The platform for implementing the disease detection algorithm is the TensorFlow of Pycharm.

The results of the algorithm can be used to detect eye diseases and predict human lesions. We also use the segmented network on ISIB cell and Iris dataset and use the traditional method on our dataset. All the segmentation accuracy is shown in Table 1. Data Augmentation We use the method of data augmentation [25] to make up for the shortcomings of fewer data. The specific methods are data perturbation enhancement, grayscale processing, and data distortion. All the above methods can improve the robustness of the training model.

4.2. Segment Sub-module detail

Our segmentation approach is based on the well-known convolution neural network, u-network, which is widely used in medical image processing and it also have been recently employed to solve problems from both the computer vision [26]and medical image analysis fields [20]. We use improved u-network to segment the interesting area textures in human eyes. even if the dataset is a few for training. The architecture of U-network consists of two parts, included down-sample on the left and down sample on the right. The left down-sampling can improve accuracy, while the right up-sampling can extract complex features. Among, Down-sampling section uses the maximum pooling with a step size of $2 \times 2$. At each step, the feature channels are doubled, which reduces the number of feature channels by half.
At the same time, the corresponding clipped feature graph in the cascade contraction path is shown. Up-sampling part is a typical convolutional network, which made up of two important $3\times3$ convolution kernels and the activation part uses the modified linear element function leaky-relu. We used leaky-relu for the activation function selection, because it can activate efficiently and is more robust [1]. Through thirty thousand times experiments, we have obtained a good segmentation model, which can segment the interesting area of the disease human eyes. It was found that after 1000 pieces of training in each round, the experiment of dividing the interest area with the training step length of 5 reduced the loss to about 0.1280 and the segmentation accuracy to about 0.99, and the interest area of human eyes could be completely divided. The detail shown in the Table 1.

![Graph of 12 sectional areas indicating organ health](image)

**Fig. 5:** This is a graph of 12 sectional areas, where the disease box is, indicating that the internal organs are not healthy. As shown in the figure, this person has only 3, 8, 9 areas with a small area of detection box, while other areas, especially 10, 11 areas, account for a large proportion, which is probably caused by neck and head injuries.

**Table 1:** Based on the comparison precision experiments of different segmentation networks, it is obvious that the precision of our deep improved segment network is higher than other advanced medical segment framework.

| Method       | Data       | Segment accuracy |
|--------------|------------|------------------|
| Unet         | ISBI cell  | 0.9203           |
| Unet         | Iris       | 0.9723           |
| Unet         | Eye interest area | 0.7863       |
| Gabor and SVM| Eye interest area | 0.7400       |
| Raw img and SVM | Eye interest area | 0.7600       |
| **Ours**     | **Eye interest area** | **0.9997** |

**Table 2:** The standard precision comparison method to evaluate the detection accuracy on dataset

| Method       | Backbone  | Size  | FPS  | AP   | AP-50 | AP-75 | AP-s | AP-m | AP-l |
|--------------|-----------|-------|------|------|-------|-------|------|------|------|
| YoloV3       | Darknet 53| 320   | 40M  | 27%  | 51%   | 29%   | 11.5% | 30.5% | 43%  |
| YoloV3       | Darknet 53| 608   | 20M  | 33%  | 57.9% | 34.4% | 18%  | 35.3% | 41%  |
| SSD          | VGG-16    | 300   | 43M  | 25%  | 43.1% | 25.6% | 6.6% | 25.5% | 41.3%|
| SSD          | VGG-16    | 512   | 22M  | 29%  | 48.5% | 30%   | 10.9%| 31.5% | 43.5%|
| RefineDet    | VGG-16    | 320   | 39M  | 29.2%| 49.5% | 31.2% | 10.1%| 31.5% | 44.5%|
| RefineDet    | VGG-16    | 512   | 23M  | 31%  | 54.5% | 35.2% | 16.1%| 36.5% | 44.3%|
| LPF          | VGG-16    | 300   | 77M  | 32%  | 51.5% | 33.2% | 12.5%| 34.5% | 46.3%|
| **Ours**     | **Darknet 53** | **608** | **20M** | **36%** | **60%** | **38%** | **20.6%** | **37%** | **46%** |
4.3. Deep Composite Predict Performance

In this section we evaluate object detection by the Average Precision (AP) and PASCAL-style AP (at a single IoU threshold of 0.5). We also report AP on objects of small, medium, and large sizes, which named AP-s, AP-m, and AP-l. This is the standard method to evaluate the detection accuracy. The detail we shown in Table 2. In our work, we design three scales detect sub-module, we input images size 416×416×32 and then use conv 32×3×3 add conv 64×3×3 with stride of 2. The scale one level including residual block 1×64 that batchsize 208, 208, 64, and then connect conv 128×3×3×3 that batch size 104, 104, 128, the second level in scale one including residual block 2×128 and connect conv 256×3×3. The first scale is aim at detect small objects, second scales aim at detect medium objects and thirdly scale level is aim at detect big objects. The first two scales have residual blocks and conv blocks, and thirdly scale has convolution 255×1×1 with batch size of 13, 13, 255. Finally, we detect results of eye diseases boxes. The final images of segmentation and detection of eye diseases. Detection model can detect the pathological changes of people eyes about six types, we through to the confidence level and the adjustment of the threshold got more groups of test results, the training of new lesions detection model we can efficiently detect lesions in place of the human eye and eye partition table by mathematical modeling and detection lesion location model integration mode, can predict diseased organs, triage again to a specific doctor’s clinic effective to improve the speed of triage, effective auxiliary medical care. As shown in the Table 3, we report our detect running time comparison to other methods, as you can see, our detection speed of eye disease is faster, which is also due to the high precision segmentation of eye interest area.

| Method      | MAP-50 | Run-Time(ms) |
|-------------|--------|--------------|
| SSD 321     | 45.4   | 61           |
| DSSD 321    | 46.1   | 85           |
| R-FCN       | 51.9   | 85           |
| SSD5 13     | 50.4   | 125          |
| DSSD 513    | 53.3   | 156          |
| **Our Detect Sub-net** | **60** | **51**       |

5. Conclusion

This paper proposes a comprehensive algorithm that can predict the health of human internal organs by segmentation and detection of human eye diseases. i.e., Deep Composite Predict Network (DCP-Net). For such small data sets of images as human eyes, the multi-dimensional depth extraction image feature method and the rapid disease detection model are adopted to obtain the result of 99.97% segmentation accuracy in the human eye disease database, and the human eye disease can be detect in real time. We are the first team to apply the depth segmentation and detection technology on the human eye database. Finally, based on the theory that TCM can predict visceral organ lesions through eye diseases, we established a dense and deep composite network of 12 fan-shaped sections and cascades, and finally output the image of detecting human eye diseases. In terms of eye disease detection, we are also the first team to combine the theories of traditional Chinese and western medicine with deep learning.

**References**

[1] Abokyi, S., Owusumensah, J., Osei, K.A.: Caffeine intake is associated with pupil dilation and enhanced accommodation. Eye 31(4), 615–619 (2017)
[2] Babenko, B., Yang, M., Belongie, S.: Visual tracking with online multiple instance learning pp. 983–990 (2009)
[3] Babenko, B., Yang, M., Belongie, S.: Robust object tracking with online multiple instance learning. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(8), 1619–1632 (2011)
[4] Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39(12), 2481–2495 (2017)

[5] Clark, R., Blundell, J., Dunn, M., Erichsen, J.T., Giardini, M., Got- tlob, I., Harris, C., Lee, H., Mcilreavy, L., Olson, A., et al.: The potential and value of objective eye tracking in the ophthalmology clinic. Eye 33(8), 1200–1202 (2019)

[6] Danelljan, M., Robinson, A., Khan, F.S., Felsberg, M.: Beyond correlation filters: Learning continuous convolution operators for visual tracking 9909, 472–488 (2016)

[7] Gao, X., Dai, L., Han, S., Chihlin, I., Heath, R.W.: Energy-efficient hybrid analog and digital precoding for mmwave mimo systems with large antenna arrays. IEEE Journal on Selected Areas in Communications 34(4), 998–1009 (2016)

[8] He, K., Gkioxari, G., Dollar, P., Girshick, R.: Mask r-cnn pp. 2980–2988 (2017)

[9] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arXiv: Computer Vision and Pattern Recognition (2015)

[10] Just, M.A., Carpenter, P.A.: Eye fixations and cognitive processes. Cognitive Psychology 8(4), 441–480 (1976)

[11] Li, D., Wong, K.D., Hu, Y.H., Sayeed, A.M.: Detection, classification, and tracking of targets. IEEE Signal Processing Magazine 19(2), 17–29 (2002)

[12] Li, M., Cheng, H., Yuan, Y., Wang, J., Chen, Q., Me, R., Ke, B.: Change in choroidal thickness and the relationship with accommodation following myopic excimer laser surgery. Eye 30(7), 972–978 (2016)

[13] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation pp. 3431–3440 (2015)

[14] Ma, C., Huang, J., Yang, X., Yang, M.: Hierarchical convolutional features for visual tracking pp. 3074–3082 (2015)

[15] Ma, J., Li, Y.: Soft combination and detection for cooperative spectrum sensing in cognitive radio networks pp. 3139–3143 (2007)

[16] Nusslein-Volhard, C., Wieschaus, E.F.: Mutations affecting segment number and polarity in drosophila. Nature 287(5785), 795–801 (1980)

[17] Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychological Bulletin 124(3), 372–422 (1998)

[18] Rayner, K.: Eye movements and attention in reading, scene perception, and visual search. Quarterly Journal of Experimental Psychology 62(8), 1457–1506 (2009)

[19] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation pp. 234–241 (2015)

[20] Rusu, R.B., Cousins, S.: 3d is here: Point cloud library (pcl) pp. 1–4 (2011)

[21] Salvucci, D.D., Goldberg, J.H.: Identifying fixations and saccades in eye-tracking protocols pp. 71–78 (2000)

[22] Satgunam, P., Datta, S., Sumalini, R.: Near vision in individuals with down syndrome: a vision screening study. Eye 33(8), 1254–1260 (2019)

[23] Tang, J., Liu, B.T., Ma, K.: Traditional chinese medicine. The Lancet 372(9654), 1938–1940 (2008)

[24] Tatler, B.W., Baddeley, R.J., Gilchrist, I.D.: Visual correlates of fixation selection: effects of scale and time. Vision Research 45(5), 643–659 (2005)

[25] Wang, T., Liu, M., Zhu, J., Tao, A., Kautz, J., Catanzaro, B.: High-resolution image synthesis and semantic manipulation with conditional gans pp. 8798–8807 (2018)

[26] Wang, X., Morrisnatschke, S.L., Lee, K.H.: New developments in the chemistry and biology of the bioactive constituents of tanshen. Medicinal Research Reviews 27(1), 133–148 (2007)

[27] Xiang, Y., Chen, L., Zhao, Y., Chen, W., Chen, Z., Liu, S., Jing, S., Manyande, A., Wang, P., Zhang, H., et al.: Measuring changes in schlemm’s canal and trabecular meshwork in
different accommodation states in myopia children: an observational study. Eye 34(2), 1–9 (2020)

[28] Xiong, X., Duan, L., Liu, L., Tu, H., Yang, P., Wu, D., Chen, G., Xiong, L., Yang, W., Liu, Q.: Panicle-seg: a robust image segmentation method for rice panicles in the field based on deep learning and superpixel optimization. Plant Methods 13(1), 1–15 (2017)

[29] Zhou, X., Wang, L., Zhou, X., Yu, Z.: Wavefront aberration changes caused by a gradient of increasing accommodation stimuli. Eye 29(1), 115–121 (2015)