Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Haonan Feng Date

Retinal Image Classification for Diagnosis of Vogt-Koyanagi-Harada Disease

By

Haonan Feng

Master of Science in Public Health

Biostatistics and Bioinformatics

_____

Hao Wu, Ph.D.
(Thesis Advisor)

_____

Tianwei Yu, Ph.D.
(Reader)

Retinal Image Classification for Diagnosis of Vogt-Koyanagi-Harada Disease

By

Haonan Feng

B.S.E

Sun Yat-Sen University

2015

Thesis Committee Chair: Hao Wu, Ph.D.

Reader: Tianwei Yu, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Biostatistics

2017

Abstract

Retinal Image Classification for Diagnosis of

Vogt-Koyanagi-Harada Disease

By Haonan Feng

Vogt-Koyanagi-Harada (VKH) disease is a disease with four stages and each stage has different symptoms. The accurate identification of different stage is of increasing importance in diagnosis and treatment of VKH disease. In this work, Sunset-Glow Fundus (SGF) appearance and Serous Retinal Detachments (SRD), two of the most significant symptoms shown in the retinal images when VKH disease develops, are chosen as the feature to identify the stage of VKH disease. To extract the information of these features precisely, two image preprocessing methods are applied at the beginning. In the feature extraction section, the distribution of color intensity, Fast Fourier Transformation, and the Grey-Level Co-occurrence method will be introduced to construct the feature matrix for model training. A k-nearest neighbor algorithm and an optimized support vector machine are employed as the classifiers. A sample dataset consisting of 144 retinal images is used for identifying four different stage of VKH disease. With optimized model combined with two SVM classifiers, the classification accuracy of 86.8% will be obtained.

Retinal Image Classification for Diagnosis of Vogt-Koyanagi-Harada Disease

By

Haonan Feng

B.S.E

Sun Yat-Sen University

2015

Thesis Committee Chair: Hao Wu, Ph.D.

Reader: Tianwei Yu, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Biostatistics

2017

Acknowledgements

I would like to thank the department of Biostatistics and Bioinformatics at Emory University for their support and guidance over the past two years. I would especially like to thank Dr. Hao Wu, my thesis advisor, for his help in leading me into this research area and providing opportunities to let me work on such interesting and attractive projects. Also, I would like to thank him for all the support and feedback that he gave me on this thesis.

I would like to thank Dr. Tianwei Yu for taking time to read my thesis and giving me comments through the process.

Lastly, I am especially grateful for the support and encouragement from my family and friends throughout my study at Emory.

Introduction

Vogt-Koyanagi-Harada (VKH) disease is a disease with bilateral diffuse granulomatous uveitis, which relates to vitiligo, alopecia, and central nervous system [1]. Originally, VKH disease can be identified as two separate illnesses: Vogt-Koyanagi syndrome [2], which featured by its chronic severe anterior uveitis, alopecia, poliosis, and Sugiura's sign; and Harada's disease [3], which featured by the bilateral exudative uveitis and pleocytosis of cerebrospinal fluid. Due to the overlapping symptoms between these two illnesses, they were combined in 1932. VKH disease will affect many body systems like eyes, ears, skin, the covering of the brain and meninges [4]. The manifestations of this illness are bilateral, diffuse uveitis and a rapid loss of vision. The incidence of VKH disease depends on the geographic location and the ethnicity encountered. In general, VKH tends to affect races with dark pigments, such as Asian. The incidence accounts for 6.8% to 9.2% of uveitis cases in Japan, while only 1% to 4% in the United States [5]. The majority cases are found during the second and fifth decades of life. Moreover, women have been reported as being more affected than men, but the influence is varied and dependent on the population. Unfortunately, the pathogenesis of VKH disease is still unknown. The possibility deduced by theories and observation is that an autoimmune reaction developed by T-cell for the antigens related to melanin. Besides, the retinal pigment epithelium is also suspected as an important reason for the development of disease [6]. To better understand the disease, four stages have been defined by different clinical features, including prodromal stage, uveitis stage, chronic stage and recurrent stage.

During these four stages, the most significant feature shown in retinal image are the Sunset-Glow Fundus (SGF) appearance. Although not all the cases of VKH disease will develop the SGF, this

symptom had already been proved to have a high correlation with the existence of VKH disease. Therefore, it has been regarded as one of the essential ocular findings for diagnosis and study of VKH disease [7]. The SGF, which is led by the occurrence of choroidal depigmentation, will appear in the chronic stage several months after onset, and continue until the end of the recurrent stage. It is typified by a pale disc with a bright red choroid, and the appearance will last several months [8]. When patients are in the recurrent stage, the color of fundus will become darker than in the chronic stage, which can be observed by radiologist easily. In addition, the other significant feature shown in the retinal image is the serous retinal detachments appear in the uveitis stage. In this stage, patients will have blurred visual acuity in both eyes and following two features will be shown in the retinal image. The first feature is the thickening of the posterior choroid, hyperemia, and edema of the optic disk. The second one is the striped wrinkles shown at the central of the choroid, which called serous retinal detachment (SRD) [9, 10].

Based on different features of the retinal image shown in different stages, it is possible to construct Computer-Aided Diagnosis system (CADx) based on the retinal image to help classify different stages by using image recognition algorithm. CADx is the system that facilitates doctors and radiologists for interpretation of medical images and makes precise diagnoses. It is a concept established by combining the diagnoses from both radiologists and computer algorithms. Based on this principle, the performance of computer algorithms does not need to be better than or even comparable to the performance of clinician. It serves more like a complementation and reference for radiologists and physicians. Nowadays, various of CADx have been employed in hospitals and clinics for assisting physicians in the detection of many diseases, especially cancers and other chest diseases [11]. As an interdisciplinary technology, CADx combines elements of

artificial intelligence, machine learning, and computer vision with radiological image processing. The originate study of CADx can be traced back to 1960s, but the rapid development was trigger until 1980s at the Kurt Rossmann Laboratories for Radiologic Image Research in the Department of Radiology at the University of Chicago [11]. The basic research began from the environment of picture archiving and communication systems (PACS), which is a convenient medical image technology for storage, transition and access economically, especially from different sources [12]. Images and reports provided by hospitals and clinics around the world can be transmitted digitally through the PACS. PACS is primarily designed as a system for the management of radiologist images for hospitals, however, it is difficult to extend the system to clinical diagnosis at that time.

The concept of automated diagnosis or automated computer diagnosis, which was created in the 1960s, was then be reintroduced and promoted [13, 14] and the concept of CADx was provided. Unlike CADx, the automated computer diagnosis system committed to developing an automated diagnosis system that can fully replace the diagnosis of radiologists, which, in other words, means the requirement of accuracy and performance to be very high. This concept is much more popular than CADx in the beginning. Therefore, a prototype of automated diagnosis system for the identification of lung nodules based on chest images was developed in the mid-1980s by the Rossmann Laboratories. However, the researcher found that even though the accuracies of the prediction results are reasonably high (in some cases, the accuracy can reach 85%-90% [15]), there is usually high false positive rate. Although the higher the performance of the computer, the better the overall performance of the algorithm, researchers in general agree that the

performance of automated diagnosis system will not be higher than that of clinicians. Thus, CADx was emphasized and developed rapidly.

The research and development of CADx are based on three foundations from the beginning to the present time. First, the foundations of the basic strategies, including techniques and algorithms used for the development of detection and classification in medical images, are based on the understanding of medical image reading by radiologists [16]. The second one is the measurement of the success of the CADx. One of the best evaluations of success will be the daily use of commercialized CADx in clinical diagnosis at different hospitals, institutions and even industries around the world. The third is to enhance the acceptance of CADx for diagnosis and stimulate the distribution of CADx research at different academic and industrial institutions. After 30 years' development, nowadays a large number of commercial CADx systems for detection had been widely used for assisting clinical diagnosis, especially in the detection of cancers. Also, the accuracy of the CADx had been greatly improved on the detection of various cancers in breast [17], lung [18], and colon [19], according to the latest commercial CADx claimed. However, the CADx based on retinal image for eye illnesses detection still has some technical difficulties and not be widely use, the development of CADx for retinal image is necessary.

Image recognition algorithm is the core of the CAD system. Usually, the image recognition algorithm in CAD systems for different medical images will have huge difference due to the specific features and characteristic of the images. Even for retinal image recognition, the

detection algorithm will vary since features will be shown in different structures like blood vessel, optic disc, macula, etc. Even though, the common procedures to construct a detection model are similar, which includes image preprocessing, feature extraction, modeling training, and evaluation. Due to the difference of light exposure and contrast resulted from cameras and lens, the basic image preprocessing to normalize the picture, and then crop them to the same size for convenience in the next step is necessary. Previous methodological development in this field includes color normalization and contrast enhancement for the non-uniform illumination [20], and extraction of blood vessel for further detection by using either matched filter method [21, 22] or optimized Gabor filter [23]. Vessel extraction based on Green channels have been proved to have the best result because of its highest contrast. After preprocessing, the image will be used to extract features and then be transformed into an efficacious data structure for model training. Histogram-based approach is one of the easiest for feature extraction. In [24], Zheng et al. used the histogram and a Dynamic Time Warping comparison technique to classify 144 retinal images with Age-related Macular Degeneration (AMD) or Normal tags after extracting blood vessel. The green channel and the saturation channel in HSI transformation were used, and the algorithm obtained 82%-86% accuracy from the ten-fold cross validation. Beside this, 2D-Gabor transformation or discrete wavelet transformation (DWT) are also good methods of feature extraction, especially for texture detection. According to the work of Khademi et al. [25], the average accuracy of DWT model can go over to approximately 85%.

In this work, 144 retinal images will be used to construct a prediction model in order to diagnosis the Sunset-Glow Fundus (SGF) Appearance. All the 144 retinal images had been tagged as 4 types, which are normal, early, middle and late stages, corresponding to the prodromal stage,

uveitis stage, chronic stage and recurrent stage, respectively. The method section will introduce further backgrounds and detailed methods of image preprocessing, feature extraction and model training. Result section presents cross-validation based prediction results with respect to all the models. In the conclusion section, we will summarize the model developed in this work, and discuss the future potentials.

Method

In the light of symptoms mentioned in papers above [6-10], two features will be utilized to construct classifiers. The first one is the change of fundus colors. Usually, the color of the normal fundus is yellow and it will maintain until SGF appears. As time goes, the fundus will turn to dark red in the recurrent stage. Also, the retinal images in Uveitis stage are brighter than the prodromal or normal stage, shown Figure 1. Another symptom is that the serous retinal detachment presents only in uveitis stage, but then vanishes rapidly in the chronic stage. The feature of this symptom shown in the retinal image is obvious. For the images of other three stages, the fundus is smoother. However, in the images of uveitis stage, the serous retinal detachment will be presented in the fundus like wrinkles.

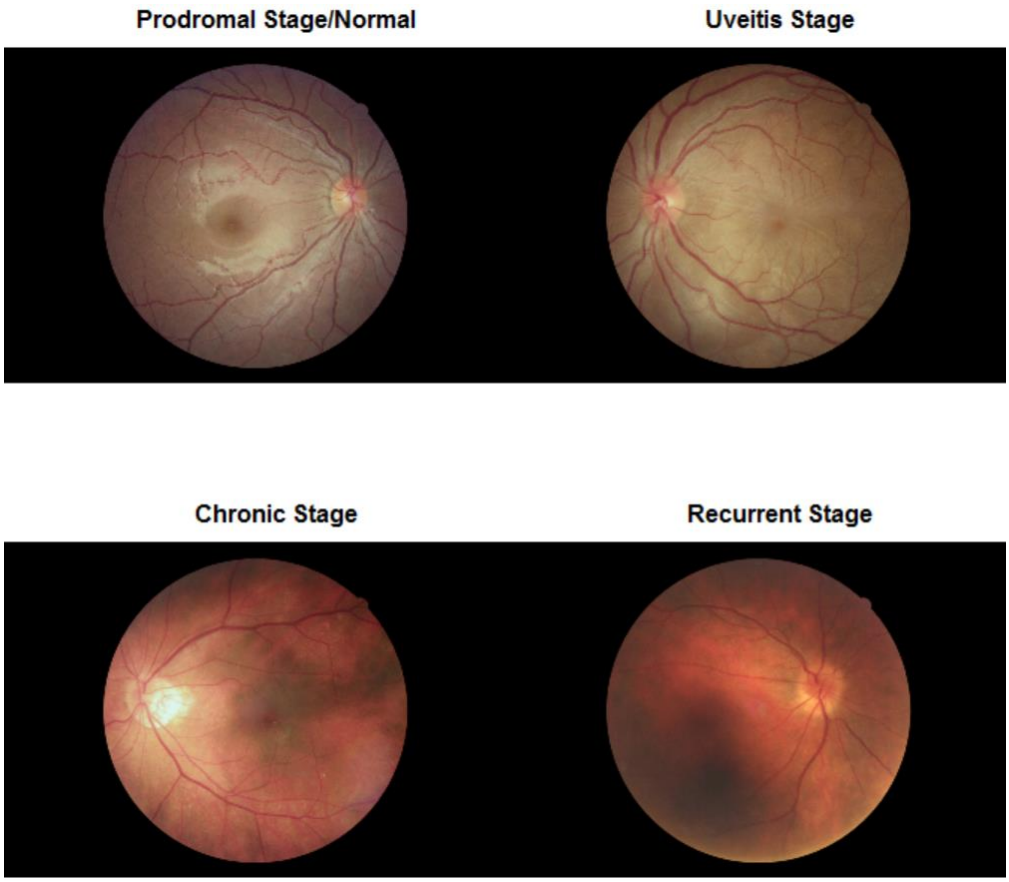

Figure 1. Overview of Retinal Images in Four Stages of VKH disease

Based on the two features shown above, the classifier can be constructed by either using statistics created by the distributions of color intensities or using texture analysis to detect wrinkles. Consider that wrinkles will only appear in uveitis stage, the classifier made from texture analysis alone may not have good predictive performance in middle and late stages. Here we will also consider combining these two classifiers into one model by using a hierarchical test procedure.

Image Preprocessing

The original retinal images from different cameras usually have different sizes, color intensity distributions, and contrasts. We first remove the black backgrounds from the image, and then resize the image to be a square of 500 by 500 pixels.

The recognition of fundamental anatomies such as blood vessel, optic disc and macular are often necessary before identifying pathological structures [24]. For some diseases, the reference points for diagnostic model training can be provided by the knowledge of shape and location of these structures. In our work, segmentation of these structures is important to eliminate noises, since structures like blood vessel could be recognized as features in the fundus. In addition, these structures could affect the color distribution of fundus and lead to imprecise result. It has been established that a better classifier will be constructed by focusing only on the color distribution or feature of the pure fundus background [24].

To extract the blood vessel and other structures, a 2D-Median Filtering algorithm as proposed by J. Turkey [26] and improved by Simon Perreault [27] were utilized as normalization and noise reduction before the detection. This algorithm has been widely used in image preprocessing. Assume that an image has m × n pixels and the windows *w* for the median filter is a square with a width equal to $2r + 1$. As shown in Figure 2, for the pixel $X_{i,j}$, a squared window (shown as yellow) will be constructed. The median of pixel intensities in the whole square will be defined as the smoothed value of $Y_{i,j}$, the pixel of the output image at the same location. Since the contrast of the images are affected by the source of light when taking pictures, such normalization procedure will smooth out the different level of contrast in an image [28]. We

further normalize the color intensity values by shifting the median to 0.5, and rescale to make the IQR in all images to be 0.1. This step is necessary to make the contrasts and exposure levels in all images homogeneous. These two steps (within image and cross-image normalization) will improve the results in downstream analyses.

$$Y_{ij} = median\{X[p,q], (p,q) \in w, 0 < p,q < 2r+1\} \quad (\mathbf{1})$$

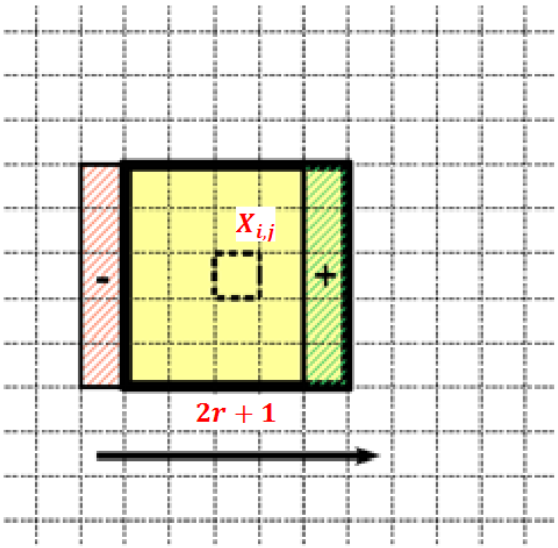

Figure 2. Illustration of the Median Filter Algorithm [27]

Next, some retinal structures such as blood vessel, macular, and optic disc should be removed from consideration[29]. This step is necessary because these structures will bring noises to the later feature extraction step. Due to the highest contrast, green channel will be used for structure detection. Based on observation, it is visually obvious that the blood vessels appear darker than the fundus, while the macular and optic disc are brighter. This means that the color intensity values are lower for blood vessel, and higher for macular and optic disc. Thus, an easy approach to detect these structures is to apply thresholds on the color intensity values, and call regions with very low values blood vessels, and regions with high values macular and optic disc. We will

take quantiles from color distribution as thresholds (details are provided in the results section). After removing the structures, the retinal images will be used as the training dataset for the color distribution classifier. Although other researches [20-23] have developed more advanced method to more precisely detect the retinal structures, we believe they will have little impact on downstream classification. Therefore, we continue to use this simple method for its computational efficiency. For serous retinal detachment, we found that the fundus' color has significant impact on the texture analysis for detecting the "wrinkles". Thus, further preprocessing is applied to remove the fundus before texture analysis. Here we again apply a threshold (based on quantiles) on colors to remove the fundus. After that, only serous retinal detachments part of the image is retained for texture analysis. Examples of images after preprocessing are provided in the Result section.

Therefore, we will have two kinds of retinal images after the image preprocessing. The first one are the retinal images without blood vessels, macular, and optic disc. These images will be constructed as the dataset for color information extraction. The second one are the retinal image without any structures but just left the serous retinal detachments. After merging the prodromal and uveitis stage and named it as PU stage, the classifier for three stages (PU stage, chronic stage, and recurrent stage) will be provided since the most significant feature for these three stages are the difference of color distributions.

Feature Extraction

There are two symptoms shown in the retinal image when patients have VKH disease, the first one is the SGF appearance in the chronic stage, which makes the color of fundus become brightly red. This symptom will be maintained in the recurrent stage with a darker red fundus. In

addition, obvious wrinkles (serous retinal detachment) will be shown in fundus when patients are in uveitis stage, but then the wrinkles will be faded at the chronic and recurrent stage. Based on these two symptoms, two methods will be conducted to extract the features for classification. The first method is to extract the information from color distribution, while the second one is to detect the serous retinal detachment by using texture analysis skills, including Fast Fourier Transformation (FFT), and Grey-Level Co-occurrence Matrix (GLCM.)

- *Color Distribution*

For the SGF symptom, extract features from the color distribution will be a reasonable method and it can be proved by comparing the color distributions of different stages in RGB and HSV channel. If the difference can be observed, then the training dataset can be constructed by using the information of color distribution. The requirement of choosing statistics is to represent the whole color distribution significantly. The first choice is the combination of mean and standard deviation of each channel in both RGB and HSV, which is the summary of the whole distribution and efficiently decrease the size of the dataset. The other choice is the combination of quantile with interval 0.1.

- *Fast Fourier Transformation (FFT)*

Fourier Transformation is one of the most important mathematical tools and had been widely applied to Electronic signal detection, image procession and texture analysis. The main idea of Fourier Transformation is to decompose a function of time into the frequencies. Therefore, it is also called the representation of frequency domain for the original function. In the image

processing, Fast Fourier Transformation [30] is an essential algorithm to extract features, especially border and texture, from the original image. The main idea is developed from the discrete Fourier transformation. In this algorithm, the original image, which is a two-dimensional matrix, will be regarded as a discrete function $f(j,k)$, and then be decomposed into a series of periodic functions such as sine and cosine functions. Based on the Euler's formula, the trigonometric function can be transformed into an exponential term with plural, so the equation of discrete Fourier transformation is,

$$F(m,n) = \sum_{j}^{N_1} \sum_{k}^{N_2} f(j,k) e^{-i2\pi(\frac{mj}{N_1}+\frac{lk}{N_2})} \qquad (\mathbf{2})$$

Where $f(j,k)$ is the intensity of the pixel (j,k) in the original image (spatial domain), the exponential term is the basic function corresponding to each pixel (m,n) in the frequency domain (Fourier domain) [31]. Since the procession of the original formula in the computer is a N^2 complexity question, Fast Fourier Algorithm was then produced to improve the efficiency of the discrete Fourier Transformation. The outcome of the FFT is the frequency domain from the original image. Each pixel in the frequency domain is in the interval between zero and one, which can be shown as the brightness in the image. Usually, the brighter the pixels shown in the frequency domain means the higher the contrast of the original image. In addition, the amplitude and the wavelength can be decomposed from the function of original image $f(j,k)$ and then be displayed on the frequency domain. After setting specific threshold, the frequency and orientation of the signal in the original image can be shown by the distribution of bright pixel in the frequency domain. In fact, if the intensities in 2-dimensional image are regarded as the third dimension, then the Fourier transformation is a mathematical tool to extract information of gradient in the third dimension and present it in the frequency domain [32]. It means that, if the

original image is smooth enough, the frequency domain will be totally black, while if a signal or wave shown in the original image, its gradient will be detected as an evidence of wave and presented as bright pixels in the frequency domain.

These peculiarities can be utilized to identify wrinkles. If the retinal image, or says, the fundus is smooth and has no wrinkles, the frequency domain will have little bright pixels and be depicted as nearly black. However, the retinal image with wrinkles will be transformed to a frequency domain with bright pixels. The more wrinkles showed in the fundus, the more pixels will be presented in the frequency domain. Moreover, the brightness of each pixel in the frequency domain can be regarded as the significance of the wrinkles shown in the fundus. Since the size of the original retinal image is too large to identify wrinkles, the preprocessed retinal image will be cropped into small pieces before doing Fast Fourier transformation. The threshold will then be defined to filter those extremely high or low frequencies for noise reduction. The number of pixels in the frequency domain which fulfill the threshold will then be utilized as the statistics for model training.

- *Grey-Level Co-occurrence Matrix (GLCM)*

Gray-Level Co-Occurrence Matrix, another common algorithm for texture detection, will also be used in this work for comparison. This method, which also called Grey-Tone Spatial-Dependence Matrix, was produced by Haralick R.M in the 1970s [33] and had been widely used in texture detection. Based on a series of second order texture calculations, this theory considers the relationship between groups of two pixels (usually neighboring) in the original image, as the Figure 3(a) below shown; while the first order measurements are the statistics calculated from

the original image directly, such as variance and mean. The method of third order measurement was provided by Akino et.al in 2003, but its utility still needs to be proofed [34].

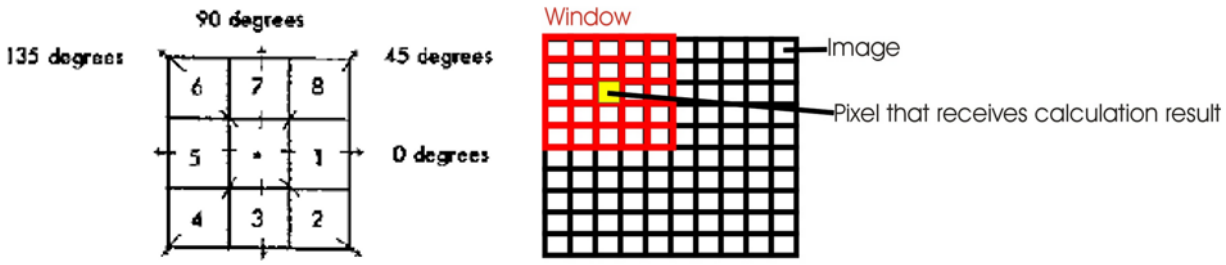

Figure 3. Illustration of GLCM matrix construction and feature extraction[33]

The GLCM is a tabulation of how often a different combination of pixel brightness values, or says gray levels, occur in an image [35]. Suppose the gray levels of each pixel in original images are known and quantized. Let $L_x = \{1,2,\dots,N_x\}$ be the notation of horizontal domain, $L_y = \{1,2,\dots,N_y\}$ be the notation of vertical domain. Assume that the texture context information can be adequately specified by the matrixes of relative frequency P_{ij} with which two neighboring pixels separated by distance *d*, one pixels with grey levels *i* and the other one with grey level *j*. After knowing the location of these two pixels as *(k,l)* and *(m,n)* respectively, the calculation of GLCM can be defined by formulas [Equation 3] in four different angular. Since the formula of GLCM contain the distance and angular between two pixels, the information of spatial relationship for each pair of pixels will be measured.

$$P(i,j,d,0°) = \#\{((k,l),(m,n)) \in (L_y * L_x) * (L_y * L_x) | k-m=0, |l-n| = d\}$$
$$P(i,j,d,45°) = \#\{((k,l),(m,n)) \in (L_y * L_x) * (L_y * L_x) | |k-m| = d, |l-n| = d\}$$
$$P(i,j,d,90°) = \#\{((k,l),(m,n)) \in (L_y * L_x) * (L_y * L_x) | |k-m| = d, l-n = 0\}$$
$$P(i,j,d,135°) = \#\{((k,l),(m,n)) \in (L_y * L_x) * (L_y * L_x) | |k-m| = d, |l-n| = d\}$$

Where # denotes the number of elements in the set. **(3)**

Once the GLCM had been constructed, all the texture features can be extracted from these matrixes. In Haralick's paper [33], there are totally 14 measurements of the textural feature being defined. The most common five statistics include angular second moment, contrast, inverse difference moment, entropy and correlation, which are given below [Equation 4]. To calculate these features, the window sizes should be first defined. As the Figure 5(b) shown, these statistics will be calculated under the range of window and the result will be received by the central pixel of the window. The other pixels which are also in the window will be called as the reference pixels. The angular second moment measures the "energy" store in the image, which means the window is ordered when high values of ASM occur. However, the entropy, which is the opposite of energy, measures the inhomogeneity of the window. The Contrast uses the difference of gray level between two pixels as the weight and thus can show the contrast of the window. The higher value the contrast has, the clears the window will be. The inverse different moment can be a measurement of similarity among pixels in the window. As for the correlation, it usually represents the consistency of texture. When the value of correlation calculated from one angular is much higher than the value from matrixes of other angular, then the direction of texture in this window will be close to this angular [36].

Equation 4. GLCM Statistics

Angular Second Moment	$f_1 = \sum_i \sum_j \{P(i,j)\}^2$
Contrast	$f_2 = \sum_i n^2 \sum_{\lvert i-j \rvert = n} P(i,j)$
Inverse Difference Moment	$f_3 = \sum_i \sum_j \frac{P(i,j)}{1+(i-j)^2}$
Entropy	$f_4 = \sum_i \sum_j P(i,j) log P(i,j)$

Correlation	$f_5 = \sum_i \sum_j \frac{ijP(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$

Where $\mu_x, \mu_y, \sigma_x, \sigma_y$ are the means and standard deviation of p_x, p_y

Referring to the equation 3, for a chosen distance *d* we have 4 angular and can construct 4 grey-level co-occurrence matrixes. Therefore, a set of 56 statistics can be constructed (14 statistics in each angular matrix). By calculating the mean and range of these 14 statistics, a vector with 28 statistics will then be provided as the features of a retinal image. Since some statistics are highly correlated among each other, a feature-selection procedure will be needed to increase accuracy [15]. In addition, the size of the window and the number of grey-level will strongly affect the quality of these statistics. A smaller size of the window will have a more precise extraction with less noise, but the definition of "small" is a subjective judgment, it depends on the size of texture shown in the window, which result in difficulty in choosing a suitable size of the window. For the number of grey-level, the common original image in the real world is 256. Although more levels usually mean more accurate texture information, the increased computational cost is incredible. Therefore, a reduction of the grey level will usually be processed before doing GLCM calculation. But this also results in the loss of information [16].

Model Training

To solve this pattern recognition problem and construct the model, machine learning algorithms were utilized in the model training. In this study, Support Vector Machine (SVM) will be the primary algorithm for model training and prediction. In addition, K-nearest neighbor algorithm will also be tried and used as a comparison toward SVM. Both these two algorithms are the core

algorithm of the supervised learning model. Given a set of training data, usually a set of vectors contain features or statistics as well as the known outcome, the algorithm will construct a model. After training, the model can be used for prediction. Accuracy of the prediction model, which can be assessed based on training data, will be the primary measurement for the model performance.

K-nearest neighbor algorithm (KNN) is one of the earliest algorithms for pattern recognition, classification, and regression [37]. Given a universe *U* of objects, a subset *A* is commonly defined by specifying the objects of the universe that belong to group *A*. The definition for the $x \in U$ is,

$$U_A(x) = \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases} \qquad (\mathbf{5})$$

Now assume that a new object is added into the universe *U*, which is notated as $\hat{x}$. In order to predict whether this new object belongs to the subset *A* or not, the definition of the distance between x_i and x_j is necessary. The Euclidean distance is the most common distance used in KNN,

$$distance(\boldsymbol{x_i}, \boldsymbol{x_j}) = \sqrt{(x_{i1} - x_{j1})^2 + \cdots + (x_{in} - x_{jn})^2} \qquad (\mathbf{6})$$

In which, $\boldsymbol{x_i} = (x_{i1}, \ldots x_{in})$ is a vector with n dimensions that belong to the universe *U*. By using this formula, the distance among the new object $\hat{x}$ and each x_i can be calculated. If we set value *k* as the number of objects that near the new object $\hat{x}$, and create a new subset of all this *k*-nearest neighbor as K, then by counting the number of objects in this subset which also belong to subset *A*, we can finally predict that the new object $\hat{x}$ belongs to subset *A* if the number of those

neighbor objects which belong to subset *A* is larger than that not belong to subset *A*. Figure 6 is the illustration of this algorithm.

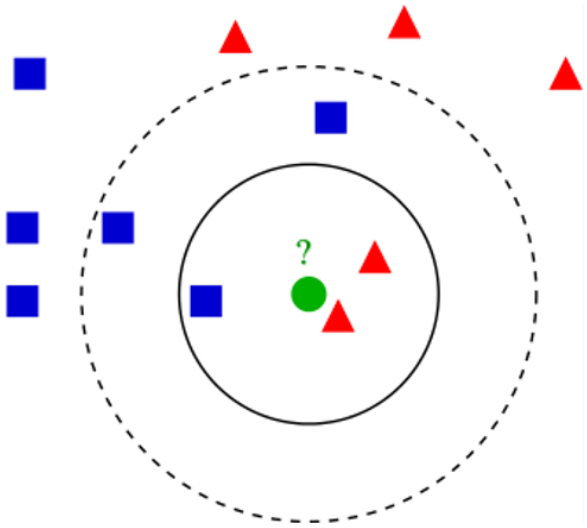

Figure 6. Illustration of K-NN Algorithm [38]

KNN algorithm is developed from the fundamental thought of classification and thus is easy to be understood. Due to its simplicity and high efficiency but still have good performance in classification, KNN is one of the most popular algorithms in machine learning [38]. Though, since it is a parametric algorithm, the difficulty of this algorithm is the role to define the parameter *k*, that is, the number of neighbor objects which will be used to decide the category of the new object. From the figure 6, it is obvious that different k value will result in different results of the new object. Generally, the larger the k value we use, the more information we will have in constructing the model and thus result in a more accurate result. But it should be balanced with the efficiency. What's more, KNN may not be sensitive enough when identifying the object with high dimensions.

Support Vector Machine, on the other hand, is famous for its rigorous mathematical logic and deduction. The binary SVM was first suggested by Vapnik and widely used in classification [39].

Although the Euclidean distance is also applied in SVM, unlike KNN, the distance will be used to construct a boundary that best separate objects in different categories, as figure 7 shown. Actually, the best boundary here is the vertical line of the longest distance between the center of these two groups.

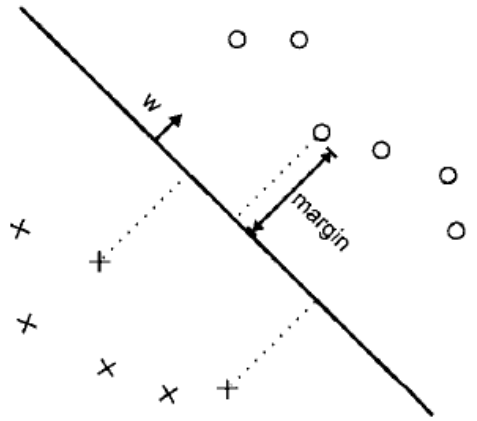

Figure 7. SVM Algorithm in Constructing Linear Boundary[40]

Given a training set of N objects $\{x_k, y_k\}_{k\in N}^{\mathrm{N}}$, in which the x_k means the vector of features or statistics of object *k* and the y_k means the category, or says pattern, of the object *k*, as we have defined in the KNN section. The support vector machine method is to construct a linear classifier with the form as,

$$y(x) = sign\left[\sum_{k=1}^{N} \alpha_k y_k \Psi(x, y_k) + b\right] \qquad (7)$$

In which, α_k means the positive real constants while *b* is the real constant. For $\Psi(x, y_k)$, this is a kernel function depend on the structure of dataset. Usually, the linear kernel function will be $\Psi(x, y_k) = x_k^T x_k$ [41]. For the training set which cannot be separated by linear boundary, SVM provides a reasonable strategy that can completely solve this problem theoretically. By mapping

the whole training set to a higher dimension space, a hyperplane will then be constructed as the classifier boundary.

Theoretically, SVM is a powerful, efficient, and robust algorithm. The procedure of projecting the dataset to high dimension space extend the applicable scope of this algorithm, and provide a higher sensitivity toward different structures of the dataset, compared with the original linear format. However, the disadvantages of SVM also result from the projection. For the original training dataset with high dimensions, the number of the dimension in projection space will be dramatically large, which may lead to low efficiency in model training. Also, defining a desired number of dimension in projection space is a difficult issue.

After model construction, a reasonable measurement for efficacy comparison among each classifier is necessary. In this paper, accuracy will be the only measurement. Assume the classifier is designed for identifying *k* categories. The number of objects in each category will be notated as n_i while the successfully identified objects in each category will be notated as s_i. Thus, the accuracy formula is straightforward for single layer classifier, that is,

$$accuracy = \frac{s_1 + s_2 + \cdots + s_k}{n_1 + n_2 + \cdots + n_k} \qquad (\mathbf{8})$$

As for the multiple layer classifiers, the specific formula developed from the basic accuracy calculation will be shown later.

Result

The dataset used in this paper consists 144 retinal images from patients with VKH disease and the size of each image is 1600*1066. In these 144 retinal images, the four stages of VKH disease

had already been separated and tagged, of which 62 are in the prodromal stage, 29 from uveitis stage, 19 from the chronic stage and 34 from the recurrent stage.

Image preprocessing are applied before feature extraction. As discussed, since two symptoms will be used in this paper, there are two methods of image preprocessing. One is for the color distribution and the goal is to remove all structures, while the other is for wrinkle detection and thus would like to use processed image without any structures and fundus but wrinkles. From the Figure 4(a) shown below, the color of the blood vessel is darker than other structures and fundus, which means the color intensity of vessel tend to be low and thus can be detected in the low quantile of color distribution. For optic disc extraction, the same method will be applied on the value channel in HSV image, since the optic disc in the value channel tends to brighter than other areas, the high quantile rather than the low one will be used to detect it. In addition, a transformation from RGB image to HSV image is needed before the detection of the optic disc because the original image only has the RGB channel. In order not to remove too much information from the original image, quantiles of the intensity values will be chosen as thresholds. In this work, the quantile for detecting the vessel were set as lower than 10%, while the optic disc will be 95% and above. Figure.4(d) shows the result of structure detection and extraction. The Figure.4(c) will be the image processed by 4(d). Generally, the performance of image processing proves the efficacy of median filter algorithm and thus this algorithm will be used in the paper for image preprocessing.

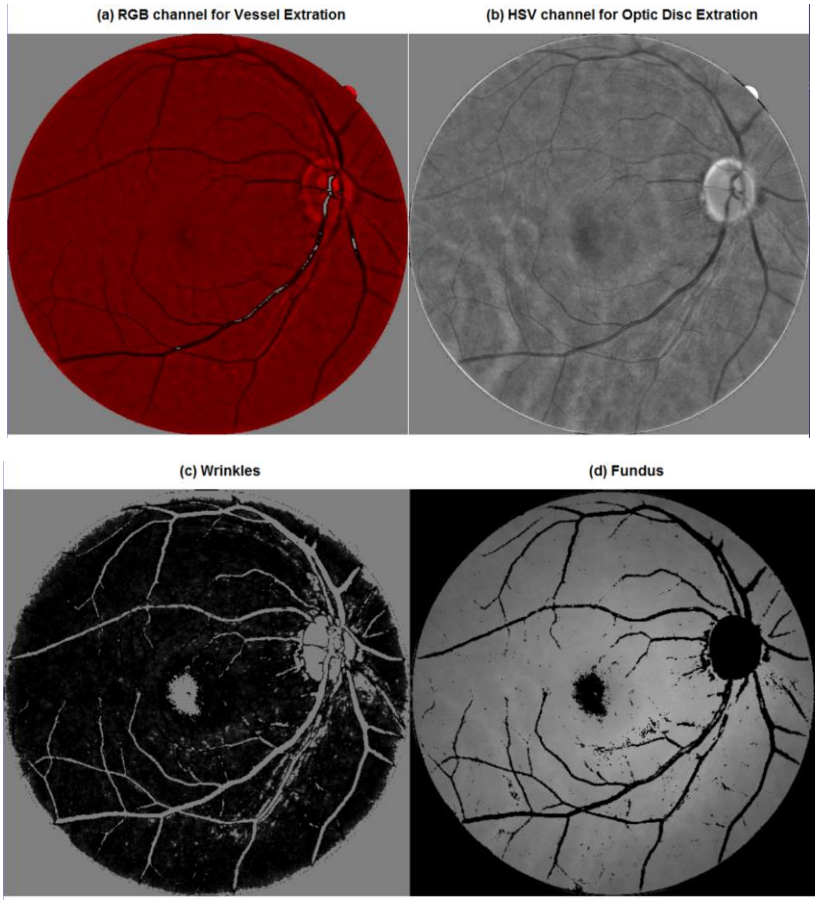

Figure 4. Retinal Image: (a) Green Channel; (b) Value Channel; (c) Wrinkles (d) Structure-removed Fundus

For feature extraction, the method developed from the color distribution will be assessed at first. Figure 5 below shows the color distribution from some examples in different stages. Here the overall distribution of all RGB intensities are shown. The differences in color distributions are visually obvious, especially between the chronic/recurrent stage and other stages.

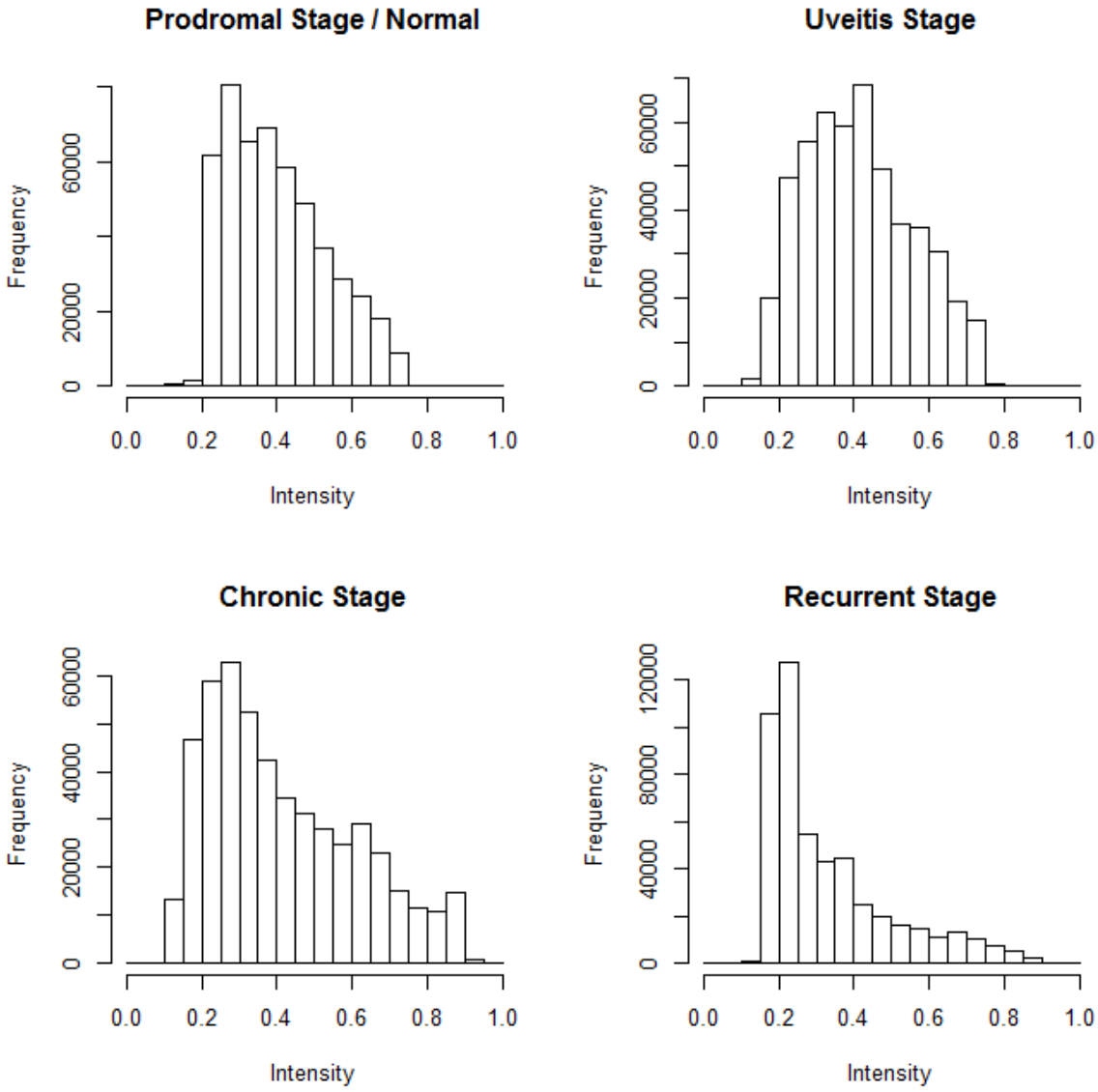

Figure 5. Overall Distribution of all RGB intensities in Each Stage

Further discoveries show that the differences of color distribution are even clearer on each separate channel. The chronic and recurrent stage tend to have distributions with right skewness, and the tail of chronic stage is thicker than the recurrent stage. While the distributions of prodromal and uveitis stage are more uniform and difficult to be distinguished. To predict based

on the distributional properties of the intensities, we obtain a vector of statistics including mean and standard deviation in each channel and use these as features for prediction.

For texture analysis to detect serous retinal detachment (wrinkles), six statistics will be chosen for GLCM results, including homogeneity, contrast, dissimilarity, entropy, second moment and correlation. The mean and standard deviation of each statistic will be calculated and used as the features. For FFT, after setting window size as 50, the whole image will be separated into 100 small pieces, and the FFT algorithm will be processed in each piece to get the frequency domain. In order to reduce the noise from the fundus, a threshold to filter the high and low frequencies is needed. Based on the observation, the value of frequencies which lower than 2 and higher than 20 will be blocked, and the summation of frequencies which left in the frequency domain will be utilized to represent the features.

The features gathered from color distribution, GLCM and FFT methods will be used to train the SVM and KNN classifier, and the classifier accuracy shown below will be used as a measurement for comparison. KNN classifiers were firstly be dropped due to its bad performance. The maximum accuracies of all KNN classifiers are lower than 60%. It is not surprising since KNN tend to have bad performance in the complex model, especially when the number of variables is very large. Then, SVM will be used to test in either color distribution, GLCM or FFT method. From the table 3 shown below, the color distribution method had the highest accuracy, even in the cross-validation test. From the all 144 images, 13 out of 23 mismatch images are from the identification of uveitis stage and chronic stage. This result is predictable since the color distribution method does not consider the appearance of serous retinal detachment (wrinkles) in uveitis stage, and this is also the reason why texture analysis method will be tested in this paper. For GLCM and FFT method, although they have similar results in

prediction accuracy, the performance of GLCM in cross-validation is even worse than KNN method. Thus, the good performance of GLCM without cross-validation could be from overfitting. Based on these results, the FFT method is then deemed to be the most accurate classifier in texture analysis.

Table 1. The Prediction result of Two Classifiers

C / Real	PU	Chronic	Recurrent
PU	87	0	4
Chronic	2	12	5
Recurrent	3	0	31

F / Real	Prodromal	Uveitis
Prodromal	59	3
Uveitis	7	22

Table 2. The Prediction result of Two Classifiers by Cross-Validation

C / Real	PU	Chronic	Recurrent
PU	87	0	4
Chronic	2	11	6
Recurrent	7	1	26

F / Real	Prodromal	Uveitis
Prodromal	58	4
Uveitis	10	19

To improve the accuracy of identification, a double layer hierarchical classifier combining with color distribution and FFT method will then be constructed. Since the good performance of color distribution in classifying the sunset glow fundus and that of FFT method in detecting the serous retinal segmentation, the design of double layer classifier is to combine the prodromal and uveitis stage and denoted it as PU stage, then the color distribution method will be used as the first classifier (C-classifier) to identify 3 stages: PU, chronic and recurrent stage. Objects in the PU stage will be classified whether in the prodromal or uveitis stage by using the FFT method as the second classifier (F-classifier). The result of this new classifier is more sensitive than before. Overall within the whole dataset, the accuracy of C-classifier and F-classifier are 90.3% and

89.1% respectively. This good performance can even be observed in the cross-validation test, with 86.1% accuracy in C-classifier and 84.6% accuracy in F-classifier, which are shown above (Table 1 & 2). The accuracy after combination will be the same formula shown in the method section, but the number of successful identification should be calculated by using a successful number of PU stage times the accuracy of F-classifiers, so the accuracy will be 83.7% without CV, and 76.8% with CV. In fact, this accuracy calculation is not precise enough because of the unbalanced data structure. The ensemble study model can potentially be applied for improvement in the future work.

Table 3. Classification Result for Different Combination of Features

Feature	Classification Accuracy	Classification Accuracy by Cross-Validation
GLCM / FFT / Color + KNN	< 60%	< 46%
GLCM + SVM	67.4%	36.1%
FFT + SVM	67.3%	50% - 54%
Color + SVM	80.6%	76.4%
Color + FFT + SVM	86.8%	82.1%

Conclusion and Discussion

In this work, we present an effective classifier construction method in retinal image diagnosis to identify different stages of VKH disease. The classifier is constructed by Support Vector Machine (SVM) and contains two layers. The first layer is an SVM classifier made from features

of color distribution to identify the PU stage, chronic stage, and the recurrent stage; while the second layer is also an SVM classifier but just for classification of prodromal stage and uveitis stage by using fast Fourier transformation (FFT) to detect serous retinal segmentation and extract information. The performance of this double-layer classifier is satisfactory in the original 144 retinal images database and the highest accuracy is around 90%. The method that used FFT for detection of serous retinal segmentation and the utilized of color distribution could be applied to other retinal image classification projects for accuracy improvement, especially in wrinkle recognition area.

Although there are some advantages in this work, we have to admit that there are also some limitations. Due to the small sample size with only 144 retinal images, the performance of machine learning model still not reach our expectation and thus further work on image collection is necessary. Moreover, the algorithm of feature extraction is still needed to improve, especially for FFT method. Although the FFT method can detect the wrinkles by analysis the frequencies of the original image, the observation of frequency domain shows that many bright pixels are irregular distributed, even when the wrinkles in the original image are regular. That may because the fundus we use, which has already been preprocessed, are still not even and smooth, thus lead to high-frequency noise. Therefore, further research for the image preprocessing and FFT is needed. Finally, if the size of the training data is large enough, convolution neural network method in deep learning can potentially be used to eliminate the manual feature extraction step. Deep learning is a popular and active research field nowadays, and many results have shown its strong performances. However, to assess whether it can perform well in small training set will be our future work.

Reference

1. Keino, H., H. Goto, and M. Usui, Sunset glow fundus in Vogt-Koyanagi-Harada disease with or without chronic ocular inflammation. Graefes Arch Clin Exp Ophthalmol, 2002. 240(10): p. 878-82.
2. Harada, E., Acute diffuse choroiditis. Acta Soc Ophthalmol Jpn, 1926. 30: p. 356-378.
3. Hayasaka, S., H. Okabe, and J. Takahashi, Systemic Corticosteroid Treatment in Vogt-Koyanagi-Harada Disease. Graefes Archive for Clinical and Experimental Ophthalmology, 1982. 218(1): p. 9-13.
4. Bordaberry, M.F., Vogt-Koyanagi-Harada disease: diagnosis and treatments update. Current Opinion in Ophthalmology, 2010. 21(6): p. 430-435.
5. Liu, X.Y., et al., Features of Optical Coherence Tomography for the Diagnosis of Vogt-Koyanagi-Harada Disease. Retina-the Journal of Retinal and Vitreous Diseases, 2016. 36(11): p. 2116-2123.
6. Sehu, K.W. and W.R. Lee, Ophthalmic pathology: an illustrated guide for clinicians. 2012: John Wiley & Sons.
7. Read, R.W., et al., Revised diagnostic criteria for Vogt-Koyanagi-Harada disease: report of an international committee on nomenclature. Am J Ophthalmol, 2001. 131(5): p. 647-52.
8. Albert, D.M. and F.A. Jakobiec, Principles and practice of ophthalmology. Vol. 1. 2000: WB Saunders Company.
9. Moorthy, R.S., H. Inomata, and N.A. Rao, Vogt-Koyanagi-Harada syndrome. Survey of Ophthalmology, 1995. 39(4): p. 265-292.

10. Rao, N.A., Pathology of Vogt–Koyanagi–Harada disease. International ophthalmology, 2007. 27(2-3): p. 81-85.

11. Doi, K., Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. Computerized Medical Imaging and Graphics, 2007. 31(4-5): p. 198-211.

12. Choplin, R.H., J.M. Boehme, and C.D. Maynard, Picture Archiving and Communication-Systems - an Overview. Radiographics, 1992. 12(1): p. 127-129.

13. Lodwick, G.S., et al., Computer Diagnosis of Primary Bone Tumors. Radiology, 1963. 80(2): p. 273-275.

14. Kruger, R.P., et al., Automated Radiographic Diagnosis Via Feature Extraction and Classification of Cardiac Size and Shape Descriptors. Ieee Transactions on Biomedical Engineering, 1972. Bm19(3): p. 174-&.

15. Chan, H.P., et al., Image feature analysis and computer-aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography. Med Phys, 1987. 14(4): p. 538-48.

16. Doi, K., Computer-Aided Diagnosis in Medical Imaging: Achievements and Challenges. World Congress on Medical Physics and Biomedical Engineering, Vol 25, Pt 5, 2009. 25: p. 96-96.

17. Dean, J.C. and C.C. Ilvento, Improved cancer detection using computer-aided detection with diagnostic and screening mammography: Prospective study of 104 cancers. American Journal of Roentgenology, 2006. 187(1): p. 20-28.

18. Armato, S.G., et al., Lung cancer: Performance of automated lung nodule detection applied to cancers missed in a CT screening program. Radiology, 2002. 225(3): p. 685-692.

19. Yoshida, H., et al., Computer-aided diagnosis scheme for detection of polyps at CT colonography. Radiographics, 2002. 22(4): p. 963-79.

20. Foracchia, M., E. Grisan, and A. Ruggeri, Luminosity and contrast normalization in retinal images. Medical Image Analysis, 2005. 9(3): p. 179-190.

21. Zhang, B., et al., Retinal vessel extraction by matched filter with first-order derivative of Gaussian. Computers in Biology and Medicine, 2010. 40(4): p. 438-445.

22. Chaudhuri, S., et al., Detection of Blood-Vessels in Retinal Images Using Two-Dimensional Matched-Filters. Ieee Transactions on Medical Imaging, 1989. 8(3): p. 263-269.

23. Daugman, J.G., Complete Discrete 2-D Gabor Transforms by Neural Networks for Image-Analysis and Compression. Ieee Transactions on Acoustics Speech and Signal Processing, 1988. 36(7): p. 1169-1179.

24. Hijazi, M.H.A., F. Coenen, and Y.L. Zheng, Retinal Image Classification using a Histogram Based Approach. 2010 International Joint Conference on Neural Networks Ijcnn 2010, 2010.

25. Khademi, A. and S. Krishnan, Shift-invariant discrete wavelet transform analysis for retinal image classification. Med Biol Eng Comput, 2007. 45(12): p. 1211-22.

26. W., T.J., Exploratory Data-Analysis. Biometrics, 1977. 33(4): p. 768-768.

27. Perreault, S. and P. Hebert, Median filtering in constant time. Ieee Transactions on Image Processing, 2007. 16(9): p. 2389-2394.

28. Kuri, S.K. and J.V. Kulkarni, Automated Segmentation of Retinal Blood Vessels using Optimized gabor filter with local entropy thresholding. International Journal of Computer Applications, 2015. 114(11).

29. Pajak, R., Use of two-dimensional matched filters for estimating a length of blood vessels newly created in angiogenesis process. Opto-Electronics Review, 2003. 11(3): p. 237-241.

30. Loan, C.V., Back Matter, in Computational Frameworks for the Fast Fourier Transform. 1992, Society for Industrial and Applied Mathematics. p. 259-273.

31. Cuaron, A., et al., Fundamentals of Digital Processing of Clinical Images. Revista Mexicana De Radiologia, 1986. 40(4): p. 139-145.

32. Smith, S.W., The scientist and engineer's guide to digital signal processing. 1997.

33. Haralick, R.M., K. Shanmugam, and I. Dinstein, Textural Features for Image Classification. Ieee Transactions on Systems Man and Cybernetics, 1973. Smc3(6): p. 610-621.

34. Akono, A., et al., Nouvelle m@ thodologie d'évaluation des paramètres de texture d'ordre trois. International Journal of Remote Sensing, 2003. 24(9): p. 1957-1967.

35. Hall-Beyer, M., GLCM texture: a tutorial. National Council on Geographic Information and Analysis Remote Sensing Core Curriculum, 2000.

36. Selvathi, D., N. Prakash, and N. Balagopal, Automated detection of diabetic retinopathy for early diagnosis using feature extraction and support vector machine. 2012.

37. Altman, N.S., An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. American Statistician, 1992. 46(3): p. 175-185.

38. Keller, J.M., M.R. Gray, and J.A. Givens, A Fuzzy K-Nearest Neighbor Algorithm. Ieee Transactions on Systems Man and Cybernetics, 1985. 15(4): p. 580-585.

39. Lili, X. and L. Shuqian. Support vector machine based method for identifying hard exudates in retinal images. in 2009 IEEE Youth Conference on Information, Computing and Telecommunication. 2009.

40. Tong, S. and E. Chang. Support vector machine active learning for image retrieval. in Proceedings of the ninth ACM international conference on Multimedia. 2001. ACM.

41. Suykens, J.A.K. and J. Vandewalle, Least squares support vector machine classifiers. Neural Processing Letters, 1999. 9(3): p. 293-300.