creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

공학박사 학위논문

Deep Feature Fusion Network for Computer-aided Diagnosis of Glaucoma using Optical Coherence Tomography

광간섭단층영상 기반 녹내장 컴퓨터 보조진단 시스템을 위한 특징 융합 심층신경망

2017 년 8 월

서울대학교 대학원
협동과정 바이오엔지니어링 전공
유 병 욱

Ph. D. Dissertation

Deep Feature Fusion Network for Computer-aided Diagnosis of Glaucoma using Optical Coherence Tomography

BY

BYEONG WOOK YOO

AUGUST 2017

INTERDISCIPLINARY PROGRAM IN BIOENGINEERING

THE GRADUATE SCHOOL

SEOUL NATIONAL UNIVERSITY

ABSTRACT

Deep Feature Fusion Network for Computer-aided Diagnosis of Glaucoma using Optical Coherence Tomography

Byeong Wook Yoo
Interdisciplinary Program in Bioengineering
The Graduate School
Seoul National University

Glaucoma has been able to be diagnosed noninvasively by analyzing the optic disc thickness with the development of optical coherence tomography. However, it is essential to maintain proper intraocular pressure through early diagnosis of glaucoma. Therefore, it is required to develop a computer-aided diagnosis system to accurately and objectively analyze glaucoma of early stage. In this paper, we propose deep feature fusion network for realizing computer-aided system which can accurately diagnose early glaucoma and verify the clinical efficacy through performance evaluation using patient images. Deep feature fusion network is analyzed by fusing features which are extracted by

feature-based classification used in machine learning and by deep learning in deep neural network.

Deep feature fusion network is deep neural network composed of heterogeneous features extracted through image processing and deep learning. The area and depth features of optic nerve defects related to glaucoma were extracted by using traditional image processing methods and the features related to distinction between glaucoma and normal subjects were extracted from the middle layer output of the deep neural network. Deep feature fusion network was developed by fusing extracted features.

We analyzed features based on image processing using thickness map and deviation map of retinal nerve fiber layer and ganglion cell inner plexiform layer in order to extract features related to the area of the optic nerve defects. Optic nerve defects were segmented in each deviation map by three criteria and the area of the defects was calculated about 69 glaucoma patients and 79 normal subjects. The performance of the severity indices calculated by defects area was evaluated by the area under ROC curve (AUC). There were significant differences between glaucoma patients and normal subjects in all severity indices (p

< 0.0001) and correctly distinguished between glaucoma patients and normal subjects (AUC = 0.91 to 0.95). This suggests that the area features of optic nerve defects can be used as an objective indicator of glaucoma diagnosis.

We analyzed features based on another image processing using retinal nerve fiber layer thickness map and deviation map to extract the features related to the depth of the optic nerve defects. Depth related index was developed by using the ratio of the optic nerve thickness of the normal to the optic nerve thickness in the optic nerve defects analyzed by the deviation map. 108 early glaucoma patients, 96 moderate glaucoma patients, and 111 severe glaucoma patients were analyzed by using depth index and the performance was evaluated by AUC. There were significant differences between the groups in the index ($p < 0.001$) and the index discriminated between moderate glaucoma patients and severe glaucoma patients (AUC = 0.97) as well as early glaucoma patients and moderate glaucoma patients (AUC = 0.98). It was found that the depth index of the optic nerve defects were a significant feature to distinguish the degree of glaucoma.

Two methods were used to apply thickness map to deep learning. One method is deep learning using randomly distributed weights in LeNet and the other method is deep learning using weights pre-trained by other large image data in VGGNet. We analyzed two methods for 316 normal subjects, 226 glaucoma patients of early stage, and 246 glaucoma patients of moderate and severe stage and evaluated performance through AUC for each groups. Deep neural networks learned with LeNet and VGGNet distinguished normal subjects not only from glaucoma patients (AUC = 0.94, 0.94), but also from glaucoma patients of early stage (AUC = 0.88, 0.89). It was found that two deep learning methods extract the features related to glaucoma.

Finally, we developed deep feature fusion network by fusing the features extracted from image processing and the features extracted by deep learning and compared the performance with the previous studies though AUC. Deep feature fusion network fusing the features extracted in VGGNet correctly distinguished normal subjects not only from glaucoma patients (AUC = 0.96), but also from glaucoma patients of early stage (AUC = 0.92). This network is superior to the previous study (AUC = 0.91, 0.82). It showed excellent performance in

distinguishing early glaucoma patients from normal subjects particularly.

These results show that the proposed deep feature fusion network provides higher accuracy in diagnosis and early diagnosis of glaucoma than any other previous methods. It is expected that further accuracy of the features will be improved if additional features of demographic information and various glaucoma test results are added to deep feature fusion network. Deep feature fusion network proposed in this paper is expected to be applicable not only to early diagnosis of glaucoma but also to analyze progress of glaucoma.

Keywords: Deep feature fusion network, Glaucoma, Optical coherence tomography, Retinal nerve fiber layer thickness map, Deep learning

Student number: 2010-23352

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

General Introduction

1.1. GLAUCOMA

Glaucoma is an ophthalmic disease caused by the degeneration of retinal ganglion cells. Glaucoma causes visual field defect as it progresses and ends up with blindness(1). 55% of glaucoma patients lost sight in one eye and 18.1% lost sight in both eyes(2). In 2010, there will be 60.5 million people with glaucoma and 79.6 million by 2020(3). The number of glaucoma patients that know to have glaucoma is very small because the symptoms don't represent definitely until glaucoma progresses to an advanced stage. So, periodic tests for glaucoma are needed because it is difficult to treat when symptoms show up.

Glaucoma can be classified into two types: open-angle glaucoma (OAG) and angle-closure glaucoma (ACG). In patients with OAG, there is clogged with a part of the trabecular meshwork and increased resistance to aqueous outflow. In this case, the angle between iris and cornea is open. On the other hand, ACG is a disease caused by the closing of the angle between iris and cornea(1, 4). Factors of OAG include increased cup disk ratio (CDR), CDR asymmetry, disc hemorrhage and increased intraocular pressure(5). There is a high likelihood that

OAG will occur due to the family history of glaucoma, black people and age. Systemic or topical corticosteroids may cause OAG, prescribing should be careful(6). ACG occurs due to the shallower anterior chambers of eye, which is common in women, aged and Asian(4).

Glaucoma is a disease that occurs as the damage progresses slowly and continuously, so it is difficult to confirm at the initial stage. So, it is important to perform various tests for accurate diagnosis of glaucoma. The glaucoma tests are intraocular pressure test, optic nerve head test, visual field test and anterior chamber angle examination. The intraocular pressure test is measurement using an anterior tonometer and glaucoma is diagnosed when it exceeds 21mmHg(7). The optic nerve head test examines the optic disc where the optic nerves gather out of the eyeball. As the glaucoma progresses, the optic disc is pressed and the shape of optic disc changes and the depression of optic disc increases(8). The optic nerve head is measured by the ophthalmoscope. The visual field test is a test of pushing the button when the object or light is recognized. As the glaucoma progresses, the visual field to be viewed becomes smaller and the change of the visual field is examined(9). The

gonioscopy examines the anterior chamber angle at the junction of the iris and the cornea. It examines whether the angle of the anterior chamber is blocked or deteriorates as the glaucoma progresses(10).

Patients with glaucoma know that they have glaucoma as their vision field becomes narrower. But when the glaucoma can be diagnosed with a visual field test, 30 ~ 50 % of the retinal ganglion cells are damaged(11). So, it is important to diagnose early glaucoma because glaucoma is a disease which may reach to blindness. However, it is very challenging to diagnose glaucoma of early stage because various tests are needed to diagnose glaucoma. Therefore, there are many studies to develop and improve various tests for early glaucoma diagnosis (12-15).

1.2. OPTICAL COHERENCE TOMOGRAPHY

The optical coherence tomography (OCT) is a device that examines the microstructure and thickness of the retina and optic nerve and diagnoses glaucoma. OCT obtains a high-resolution tomographic image using the time difference and magnitude of light reflected from a desired tissue(16). OCT is commonly used to measure retina and anterior segment eye because the eye has a structure with an optically clear pathway(17, 18). OCT is a very effective screening device for examining glaucoma because it examines the structure and thickness of the optic nerve with high resolution and objectively analyzes the statistically abnormal portion of the average thickness of normal people(19).

OCT is classified into time domain OCT (TD-OCT) and frequency domain OCT (FD-OCT). And FD-OCT is divided into spectral domain OCT (SD-OCT) and swept source OCT (SS-OCT). TD-OCT is designed to analyze the reflected region according to time as the reference mirror moves. Currently, FD-OCT has been actively studied because of low signal sensitivity and slow scan speed limit of TD-OCT(20). FD-OCT is a device using the principle of the interferometer in the same way, but it

is designed to acquire depth information by detecting the wavelength components of the scattered signal in the tissue and performing Fourier transformation(21). SD-OCT is a device obtaining images by using a spectrometer as a photodetector and SS-OCT is a device obtaining images by using one photodetector while changing the central wavelength of the light source itself(22).

SD-OCT can be measured at 5 ~ 7 μm as a current device, so it has many advantage in glaucoma diagnosis because it can analyze abnormality and progress of retinal layer which was not seen in TD-OCT(23). SD-OCT can acquire 3D data through fast capture time and further analyze retinal nerve fiber layer (RNFL) and retinal ganglion cell layer (RGCL) defects as it is able to analyze a wide range of desired area tightly. SD-OCT has improved accuracy and reproducibility in the diagnosis of glaucoma and analysis of progress, but it requires additional processing including image processing technology(24).

1.3. OBJECTIVE THESIS

Glaucoma is a disease that damages the optic nerve and results in the loss of vision field. Glaucoma is caused by putting pressure on the optic nerve through intraocular pressure increase and blood supply blockage. Chronic glaucoma is rarely seen and difficult to treat when symptoms are shown. Therefore, it is very important to diagnose glaucoma earlier. Analyzing the optic nerve loss is needed to diagnose glaucoma accurately. So, it is possible to analyze the structure and thickness of the optic nerve precisely and find the glaucoma early by measuring the depth of the optic nerve through OCT. The development of OCT measuring technology has improved the depth and structure to measure optic nerve and the resolution of measured images, but the method of accurately analyzing the glaucoma using OCT image is still insignificant. Thus, it is important to extract meaningful features related to glaucoma using images through OCT and accurately classify glaucoma patients and normal people into the features.

Thus, the objective of this thesis was to extract useful features which are relevant to glaucoma using images acquired

with optical coherence tomography and develop the classifier that discriminates clearly between normal and glaucoma patients using these features. Chapter 2 calculated localized retinal nerve fiber layer (RNFL) defects on the ganglion cell inner plexiform layer (GCIPL) and RNFL deviation maps and extracted features as severity indices according to probability level. Chapter 3 extracted feature as RNFL defect depth percentage index on the RNFL thickness deviation map using the proportion of the RNFL defect depth. And Chapter 4 extracted features using deep learning which only allows a computer to distinguish between normal and glaucoma patients on RNFL thickness deviation maps and classified two groups using deep feature fusion network developed by the fusion of features in Chapter 2, 3 and 4.

CHAPTER 2

Feature Extraction for Glaucoma Diagnosis 1. Severity Index of Macular GCIPL and Peripapillary RNFL Deviation Maps

2.1 INTRODUCTION

Glaucoma is a progressive optic neuropathy characterized by the loss of retinal ganglion cells (RGCs) presenting as structural changes of the optic nerve head (ONH) and retinal nerve fiber layer (RNFL) with corresponding visual field(VF) defects(25). For prevention of functional vision loss and maintenance of the quality of life of glaucoma patients, it is essential to detect early change by glaucomatous RGC loss.

Previous study has reported that peripapillary RNFL (pRNFL) thickness measurement by Cirrus high-definition (HD) spectral-domain (SD) Optical Coherence Tomography (OCT) (software version 5.1.0.96; Carl Zeiss Meditec, Dublin, CA, USA) is highly effective for detection of localized RNFL defect on red-free RNFL photography(26). The pRNFL deviation map (Cirrus OCT software version 3.0; Carl Zeiss Meditec) has been shown to offer a higher diagnostic ability for detection of localized RNFL defects than other conventional pRNFL maps such as the clock-hour map or quadrant map(27-29).

Recently the ganglion cell analysis (GCA) algorithm was incorporated into Cirrus OCT with the newer software

version 6.0 to allow for successful and reproducible segmentation of inner macular layers (GCIPL: a combination of the ganglion cell layer [GCL] and the inner plexiform layer [IPL]) (30, 31); its glaucoma diagnostic ability, significantly, has been shown to be comparable to those of the pRNFL and ONH parameters (32, 33).

However, to date, little of the ability of the macular GCIPL deviation map in Cirrus OCT for detection of localized RNFL defects has been demonstrably determined. Therefore, this study was undertaken to compare the ability of the macular GCIPL deviation map with that of the pRNFL deviation map for detection of localized RNFL defects shown on red-free RNFL photography.

2.2. METHODS

2.2.1. Study subjects

Glaucomatous eyes with localized RNFL defects and normal control eyes satisfying the eligibility criteria were enrolled from the Glaucoma Clinic of Seoul National University Hospital during the period from December 2012 to May 2013. In cases where both eyes were eligible, one eye was randomly selected.

The subjects were all aged 18 years or older. All had a best-corrected visual acuity (BCVA) of 20/40 or better, a spherical-equivalent refractive error within ± 5.00 diopters (D), astigmatism within ± 3.00 D, an open anterior chamber angle and high-quality red-free RNFL photography. Eyes with a history of ocular or systemic diseases possibly affecting the peripapillary area (e.g. large peripapillary atrophy, chorioretinal coloboma, peripapillary staphyloma) or macula area (e.g. epiretinal membrane, age-related macular degeneration, macular oedema), amblyopia, uveitis, intra-ocular surgery (excepting uncomplicated cataract surgery), diabetes or any other ocular or systemic diseases affecting RNFL thickness or VF (e.g. retinal

vein occlusion, ischemic optic neuropathy) were excluded. OCT images with a signal strength ≤ 6 or with visible eye motion or blinking artifacts were excluded as well.

The glaucomatous eyes had localized RNFL defects on red-free RNFL photography with asymptomatic-to-moderate glaucomatous VF loss (mean deviation (MD) > -12 dB). Glaucomatous VF loss was defined as a pattern standard deviation (PSD) outside the 95% normal limits, glaucoma hemifield test results outside the normal limits and/or a cluster of at least three points with a p-value <0.05 on the pattern deviation plot, 1 of each with p < 0.01 affecting the same hemifield; also, the cluster could not be contiguous with the blind spot and could not cross the horizontal midline, on two consecutive VF tests.

Preperimetric glaucomatous eyes were defined as those having a localized wedge-shaped RNFL defect clearly visible on red-free RNFL photography with normal standard automated perimetry (SAP) results in at least two tests. Normal VF was defined as MD and PSD within 95% confidence limits and a glaucoma hemifield test result within the normal limits.

Normal control eyes were defined as those having an intra-ocular pressure (IOP) ≤ 21 mmHg with no history of increased IOP, showing an absence of glaucomatous disc appearance, no visible RNFL defect on red-free RNFL photography and a normal SAP result.

The study adhered to the tenets of the Declaration of Helsinki and was approved by the Institutional Review Board of Seoul National University Hospital. Informed consent was obtained from all of the subjects.

2.2.2. Red-free RNFL photography

The subjects underwent a comprehensive ophthalmic examination including a medical history review; measurement of BCVA; slit-lamp biomicroscopy; Goldmann applanation tonometry (GAT); gonioscopy; dilated fundoscopic examination with a 90 (D) lens; and stereoscopic disc photography (SDP), red-free RNFL photography and SAP (Humphrey Field Analyzer II; Carl Zeiss Meditec). Red-free RNFL photography was obtained after dilation of the pupil using a digital fundus camera system (CF-60UVi/D60; Canon, Inc, Tokyo, Japan) with a green filter inserted to enhance the RNFL(34). Images were saved in

a 1600 x 1216-pixel digital imaging and communications in medicine format and were stored in the picture archiving communication system (PACS) of Seoul National University Hospital.

Localized RNFL defects on red-free RNFL photography were defined as having a width at a 1-disc-diameter distance from the edge of the disc larger than that of a major retinal vessel, diverging in an arcuate or wedge shape and reaching the edge of the disc(35). Two masked glaucoma specialists independently evaluated the red-free RNFL photography without knowledge of clinical information such as OCT or VF test results; discrepancies were resolved by adjudication of a third glaucoma specialist.

2.2.3. Cirrus OCT imaging

Using Cirrus OCT (software version 6.0), OCT images were acquired by macular scan (macular cube 200 x 200 protocol) and pRNFL scan (optic disc cube 200 x 200 protocol) subsequent to pupil dilation. The macular GCIPL thickness within a 6 x 6 x 2 mm (14.13 mm^2) elliptical annulus around the fovea was measured and computed by GCA algorithm embedded in Cirrus OCT software version 6.0. The annulus cube was of 1

mm inner vertical diameter, 4 mm outer vertical diameter, 1.2 mm inner horizontal diameter and 4.8 mm outer horizontal diameter, excluding the central portions of the fovea where the layers are thin and difficult to defect(30). GCIPL thickness was then analyzed according to eight parameters: average, minimum and in six sectors (superonasal, superior, superotemporal, inferotemporal, inferior, and inferonasal). This computation method has been described in detail in previous reports(30, 31). The pRNFL thickness within a 3.46-mm diameter circle (256 A-scan) automatically positioned around the optic disc was measured and analyzed in 17 parameters: average, four quadrants (superior, inferior, temporal, nasal) and 12 clock-hour sectors(36). The RNFL thickness in a 6 x 6 mm^2 area around the optic disc cube was measured by 200 x 200 axial scans (pixels) for generation of the deviation map

On the basis of a comparison with the built-in internal normative database, the GCIPL and pRNFL thickness values were analyzed and then represented on color-coded deviation maps composed of 50 x 50 superpixels (200 x 200 pixels). The uncolored (grey color) superpixels indicated the normal range, whereas yellow- or red- colored superpixels indicated

abnormality at the 5% or 1% level, respectively. The GCIPL deviation map represented the OCT enface image of the annular cube (between the inner and outer rings) of the macula excluding the central fovea. The pRNFL deviation map represented the OCT enface image of the optic disc that showed the boundaries of the cup, disc and 3.46-mm diameter circle.

2.2.4. Deviation map analysis protocol

In this study, the GCIPL and pRNFL measurement data were exported as image files (file format: JPEG) using the built-in export function of Cirrus OCT. The deviation map for each file was retrieved using a customized image processing program written in Matlab R2012a (The Mathworks, Inc., Natick, MA, USA). Then, the number of color-coded (yellow or red) abnormal superpixels on the GCIPL and pRNFL deviation map corresponding to the location of localized RNFL defects visible on the red-free RNFL photographs was calculated using the same customized Matlab program. The three criteria for significantly aberrant superpixels on the GCIPL and pRNFL deviation map corresponding to RNFL defect were determined arbitrarily as follows:

Critetion 1 : Cluster of contiguous yellow superpixels

≥ 3 including red superpixel ≥ 1

Critetion 2 : Cluster of contiguous yellow superpixels

≥ 5 including red superpixel ≥ 3

Critetion 3 : Cluster of contiguous yellow superpixels

≥ 10 including red superpixel ≥ 5

In the case of the macular GCIPL deviation map, we excluded from the analysis 1 superpixel around the inner circle of the GCIPL scan area considering the possibility of artifact. With respect to the pRNFL deviation map, we excluded the nasal quadrant (90°) centered on the pRNFL calculation circle (3.46-mm diameter) as well as three superpixels around the disc margin, in the light of the frequency of glaucomatous damage and the possibility of artifact (Figure 2.1).

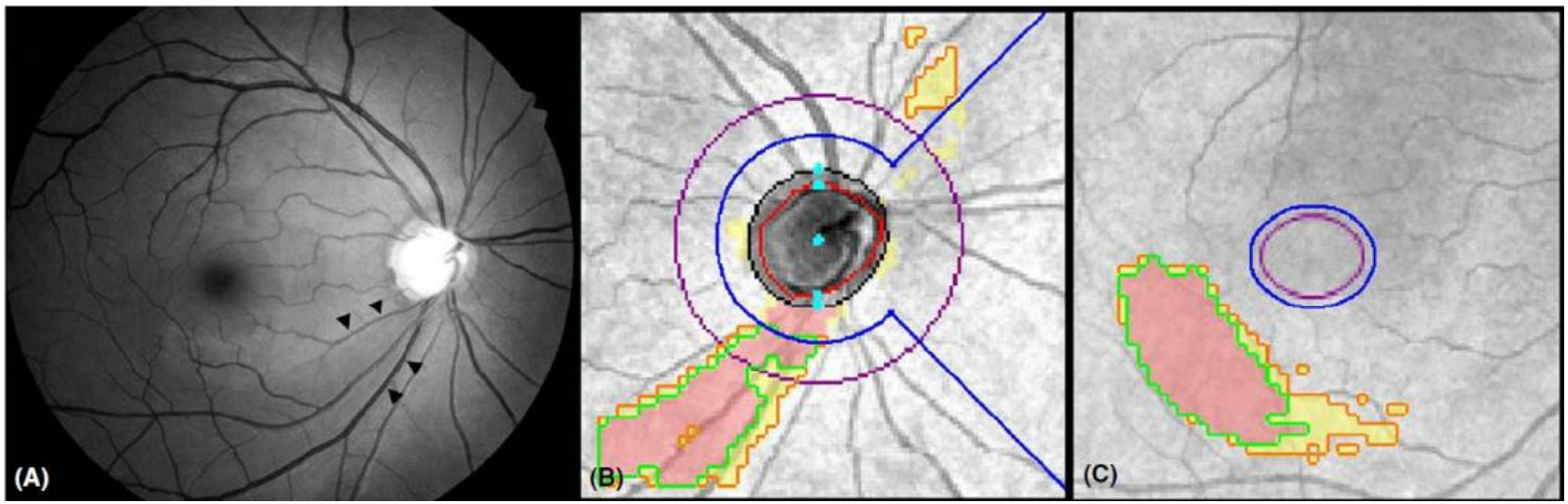

Figure 2.1. Calculation of aberrant deviation-map superpixels corresponding to localized retinal nerve fiber layer (RNFL) defect using customized Matlab program. (A) Red-free RNFL photography showing inferotemporal localized RNF L defect (arrowheads). (B) Aberrant pRNFL deviation-map superpixels corresponding to photographic RNFL defect: nasal 90° centered on the pRNFL calculation circle and three superpixels near the disc

margin were excluded from the analysis (blue line). The superpixels coded in red and yellow were recognized and calculated by the customized image processing program (red superpixels for the area bounded by the green line, yellow superpixels for the area bounded by the orange line). (C) Aberrant GCIPL deviation-map superpixels corresponding to photographic RNFL defect: 1 superpixel near the inner ring (border) of the GCIPL scan area (purple line) was excluded from the analysis (blue line). The superpixels coded in red and yellow were recognized and calculated by the customized image processing program (red superpixels for the area bounded by the green line, yellow superpixels for the area bounded by the orange line).

The number of aberrant superpixels was calculated according to the severity index (S1, S2, or S3), which was determined in relation to the probability level. S1 represented the number of yellow and red superpixels for the $p < 0.05$ probability level; S2 indicated the number of only red superpixels for the $p < 0.01$ probability level. As the red superpixels represented a more significant abnormality than the yellow superpixels, we double-weighted the red superpixels. Therefore, S3 represented the number of red superpixels multiplied by 2 plus the number of yellow superpixels.

2.2.5. Statistical analysis

Statistical analysis were performed using SPSS version 19.0 (SPSS Inc., Chicago, IL, USA) and MedCalc 12.3.0 (MedCalc Software, Mariakerke, Belgium). The Student t-test was used to compare the continuous variables between the normal control group and the glaucoma group. Additionally, the Pearson chi-square test was used to compare the two groups' categorical variables. Pearson correlation analyses were performed to calculate the correlations of the deviation map results (the number of aberrant superpixels represented as a severity index for each probability level: S1, S2, S3) between

the GCIPL and pRNFL. The diagnostic abilities of the deviation map algorithms in discriminating glaucomatous eyes with localized RNFL defects from normal control eyes were evaluated by computing the areas under the receiver operating characteristic curves (AUROCs) and comparing the results(37). A $p-value < 0.05$ was considered statistically significant.

2.3. RESULTS

The final study sample included 148 eyes of 148 subjects (69 eyes of 69 subjects with localized RNFL defects and 79 eyes of 79 normal control subjects). Among those 69 eyes with localized RNFL defects, 51 showed perimetric glaucoma (40 eyes: early stage (MD > −6 dB) of VF loss; 11 eyes: moderate stage (−12 dB < MD ≤ −6 dB) of VF loss) and 18 showed preperimetric glaucoma. The ocular and demographic characteristics of the subjects are presented in Table 2.1. Between the two groups, there were no significant differences in mean age, BCVA or spherical equivalent of refraction (all $p > 0.05$, Student t−test). However, the sex distribution showed a significant discrepancy, in that more females were included in the glaucoma group (64%) than in the normal control group (47%) ($p = 0.039$, chi− square test). Moreover, the VF indices (MD, PSD) did significantly differ ($p < 0.0001$, Student t−test).

Table 2.2 shows the averages of the severity indices for the GCIPL and pRNFL deviation maps as determined by the number of aberrant superpixels at each probability level. Comparing the glaucoma groups and the normal control group,

Table 2.1. Demographics and ocular characteristics of normal controls and subjects with glaucoma and localized retinal nerve fiber layer (RNFL) defects.

	Normal control (n = 79)	Glaucoma (n = 69)	p-value
Age (y) (mean ± SD)	56.14 ± 12.27	58.19 ± 9.91	0.270*
Sex (male/female)	42/37	25/44	0.039†
BCVA (log MAR)	0.04 ± 0.65	0.04 ± 0.74	0.802*
SE (diopters)	-0.73 ± 1.51	-0.27 ± 1.62	0.143*
VF Index			
MD (dB)	-0.44 ± 1.90	-3.02 ± 3.21	<0.0001*
PSD (dB)	2.02 ± 1.08	5.35 ± 3.88	<0.0001*
Stage of disease			
Preperimetric		18 (26%)	
Early (MD >-6 dB)		40 (58%)	
Moderate (-12 dB < MD ≤ -6 dB)		11 (16%)	

SD = standard deviation, BCVA = best-corrected visual acuity, SE = spherical equivalent,
VF = visual field, MD = mean deviation, PSD = pattern standard deviation.
* The comparison was performed using the Student t-test.
† The comparison was performed using the chi-square test.

Table 2.2. Severity index of deviation map of macular ganglion cell−inner plexiform layer (GCIPL) and peripapillary retinal nerve fiber layer (pRNFL) computed using a customized Matlab program.

	S1 (red and yellow superpixels)			S2 (red superpixels only)			S3 (red superpixels x 2 + yellow superpixels)		
Criteria	Normal control (n = 79)	Glaucoma (n = 69)	p-value	Normal control (n = 79)	Glaucoma (n = 69)	p-value	Normal control (n = 79)	Glaucoma (n = 69)	p-value
Criterion 1									
pRNFL	16.01 ± 37.77	196.93 ± 111.12	<0.0001	4.39 ± 13.11	111.30 ± 71.57	<0.0001	20.41 ± 47.99	308.23 ± 179.35	<0.0001
GCIPL	34.66 ± 69.50	275.39 ± 163.52	<0.0001	7.76 ± 17.61	204.01 ± 151.27	<0.0001	42.42 ± 85.14	479.41 ± 311.80	<0.0001
Criterion 2									
pRNFL	10.29 ± 29.91	192.96 ± 111.13	<0.0001	4.04 ± 13.10	111.70 ± 72.38	<0.0001	14.33 ± 41.87	304.65 ± 180.17	<0.0001
GCIPL	30.22 ± 66.12	271.61 ± 166.48	<0.0001	7.13 ± 16.88	203.28 ± 151.28	<0.0001	37.34 ± 81.03	474.88 ± 314.79	<0.0001
Criterion 3									
pRNFL	8.10 ± 27.35	191.57 ± 110.79	<0.0001	3.57 ± 12.85	110.03 ± 71.30	<0.0001	11.67 ± 39.30	301.59 ± 179.20	<0.0001

GCIPL	26.14 ± 61.91	270.28 ± 167.39	<0.0001	6.24 ± 16.32	202.68 ± 151.56	<0.0001	32.38 ± 76.32	472.96 ± 315.91	<0.0001

S1 = number of yellow and red superpixels with probability level of $p < 0.05$, S2 = number of only red superpixels with probability level of $p < 0.01$, S3 = number of red super pixels x 2 + number of yellow superpixels, Criterion 1 = Cluster of ≥3 contiguous yellow superixels including ≥1 red superpixel; Criterion 2 = Cluster of ≥5 contiguous yellow superixels including ≥3 red superpixel; Criterion 1 = Cluster of ≥10 contiguous yellow superixels including ≥5 red superpixel.

* Comparison was performed using Student t-test

there were statistically significant differences in all three severity indices (S1, S2, S3) for each criterion (all $p < 0.0001$, Student t-test). The severity indices of both deviation maps generally decreased according to the change of criterion from 1 to 3. However, the S2 (number of red superpixels) of both maps showed relatively similar values regardless of the criterion.

With the GCIPL deviation map algorithm, the AUROCs in discriminating the glaucomatous eyes with localized RNFL defects from the normal control eyes ranged from 0.910 to 0.931 and for the pRNFL deviation map algorithm, from 0.934 to 0.950, according to the criteria and probability levels. In a comparison of all of the corresponding severity indices, those for the pRNFL deviation map showed larger AUROCs than those for the GCIPL deviation map. However, the differences were not statistically significant (all $p > 0.05$; Figure 2.2.). According to the different criteria, the AUROCs for S2 (0.931 for Criterion 1, 0.924 for Criterion 2 and 0.916 for Criterion 3) and S3 (0.922 for Criterion 1, 0.917 for Criterion 2 and 0.910 for Criterion 3) of the GCIPL deviation map were significantly larger than those for S1 (all $p < 0.05$); there was no significant difference between S2 and S3 (all $p > 0.05$). By contrast, there were no significant differences

found among the three severity indices of the pRNFL deviation map for any of the criteria (all $p > 0.05$). When matched for fixed specificity, the sensitivities of the severity indices of the pRNFL deviation map were generally higher than those of the GCIPL deviation map, excepting S2 for criteria 2 and 3 (specificity $\geq$ 95%; Table 2.3).

To investigate the correlation of the severity indices in the deviation map algorithm between the GCIPL and pRNFL, Pearson correlation analysis was performed. Strongly significant correlations were observed in the three indices, regardless of the criterion (all $p < 0.0001$). S3 showed the strongest correlations between the GCIPL deviation map and the pRNFL deviation map (coefficient of determination (R2) = 0.642 for Criterion 1, R2 = 0.653 for Criterion 2 and R2 = 0.654 for Criterion 3; all $p < 0.0001$) compared with S2 (R2 = 0.627 for Criterion 1, R2 = 0.629 for Criterion 2 and R2 = 0.626 for Criterion 3; all $p < 0.0001$) and S1 (R2 = 0.615 for Criterion 1, R2 = 0.632 for Criterion 2 and R2 = 0.638 for Criterion 3; all $p < 0.0001$; Table 2.4).

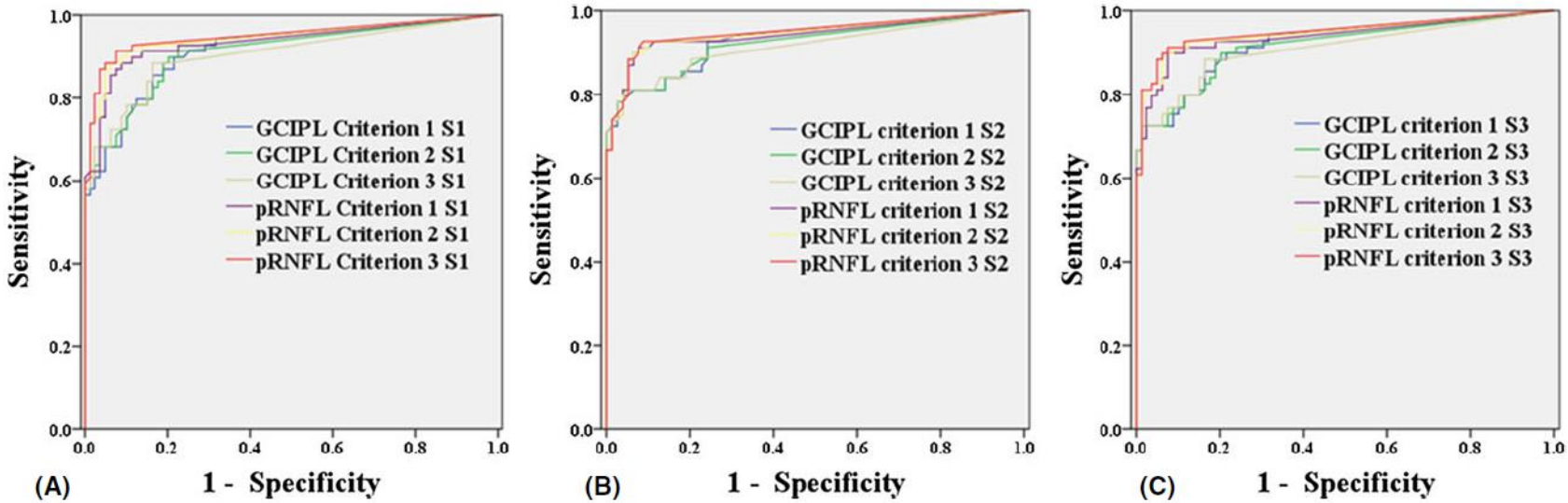

Figure 2.2. AUROC of severity indices on deviation map for macular ganglion cell–inner plexiform layer (GCIPL) and peripapillary retinal nerve fiber layer (pRNFL) by defined criteria. (A) AUROC of S1 on both GCIPL and pRNFL deviation map. (B) AUROC of S2 on both deviation maps. (C) AUROC of S3 on both deviation maps.

Table 2.3. Areas under receiver operating characteristic curves (AUROCs) and sensitivities at fixed specificities for severity indices on deviation map of macular ganglion cell–inner plexiform layer (GCIPL) and peripapillary retinal nerve fiber layer (pRNFL).

	AUROC ± SE (95% CI)			Sensitivity			
				Specificity ≥ 80%		Specificity ≥ 95%	
	GCIPL	pRNFL	p-value*	GCIPL	pRNFL	GCIPL	pRNFL
Criterion 1							
S1	0.915 ± 0.024 (0.868-0.961)	0.934 ± 0.023 (0.889-0.961)	0.468	87.0	91.3	62.3	75.4
S2	0.931 ± 0.022 (0.888-0.974)	0.943 ± 0.022 (0.900-0.986)	0.633	85.5	92.8	79.7	81.2
S3	0.922 ± 0.023 (0.878-0.967)	0.938 ± 0.022 (0.895-0.982)	0.526	87.0	92.8	72.5	79.7
Criterion 2							
S1	0.910 ± 0.025 (0.860-0.959)	0.944 ± 0.021 (0.903-0.985)	0.133	88.4	92.8	68.1	79.7
S2	0.924 ± 0.024 (0.877-0.971)	0.946 ± 0.021 (0.906-0.987)	0.310	85.5	92.8	81.2	79.7
S3	0.917 ± 0.025 (0.868-0.965)	0.946 ± 0.021 (0.906-0.987)	0.184	88.4	92.8	72.5	82.6
Criterion 3							

S1	0.904 ± 0.027 (0.851-0.957)	0.949 ± 0.020 (0.910-0.989)	0.063	88.4	92.8	68.1	87.0
S2	0.916 ± 0.026 (0.865-0.967)	0.949 ± 0.020 (0.891-0.989)	0.159	85.5	92.8	81.2	78.3
S3	0.910 ± 0.026 (0.858-0.962)	0.950 ± 0.020 (0.891-0.989)	0.093	88.4	92.8	72.5	82.6

AUROC = area under receiver operating characteristic curve, SE = spherical equivalent, S1 = number of yellow and red superpixels with probability level of $p < 0.05$, S2 = number of only red superpixels with probability level of $p < 0.01$, S3 = number of red super pixels × 2 + number of yellow superpixels, , Criterion 1 = Cluster of ≥3 contiguous yellow superixels including ≥1 red superpixel; Criterion 2 = Cluster of ≥5 contiguous yellow superixels including ≥3 red superpixel; Criterion 1 = Cluster of ≥10 contiguous yellow superixels including ≥5 red superpixel.

* Comparison was performed by the method of DeLong et al. (1988).

Table 2.4. Correlation of severity index of deviation map algorithm between macular ganglion cell−inner plexiform layer (GCIPL) and peripapillary retinal nerve fiber layer (pRNFL) by Pearson correlation analysis.

	GCIPL S1 versus pRNFL S1		GCIPL S2 versus pRNFL S2		GCIPL S3 versus pRNFL S3	
Criteria	R^2	p-value	R^2	p-value	R^2	p-value
Criterion 1	0.615	<0.0001	0.627	<0.0001	0.642	<0.0001
Criterion 2	0.632	<0.0001	0.629	<0.0001	0.653	<0.0001
Criterion 3	0.638	<0.0001	0.626	<0.0001	0.654	<0.0001

R^2: Coefficient of determination by Pearson correlation analysis

2.4. DISCUSSION

The present study was designed with the main objective of evaluating the diagnostic ability of the GCIPL deviation map in Cirrus OCT for detection of localized RNFL defects. The results showed that the GCIPL deviation map offers favorable, comparable-to-pRNFL diagnostic performance in the detection of such defects on red-free RNFL photography.

In a large proportion of glaucoma patients, RNFL damage is known to precede noticeable change of the ONH and VF defect(38, 39). Therefore, RNFL assessment is quite helpful for early diagnosis of glaucoma. There are various techniques available for detecting RNFL defects; they include fundus examination, red-free RNFL photography, Heidelberg Retina Tomography (HRT) and OCT. Given the recent advances in OCT technology, SD-OCT offers objective, reproducible real-time measurement of RNFL thickness with a fast scan speed(23, 40, 41). Notably, Cirrus OCT, a commercial SD-OCT platform, generates a pRNFL deviation map that outperforms the conventional clock-hour map or quadrant map in RNFL defect detection(27-29). For example, Hwang et al. (2013) reported

that for 295 eyes with early-stage glaucoma (MD > -6.0 dB), the pRNFL deviation map by Cirrus OCT showed a lower frequency of misidentification of photographic RNFL defects than the clock-hour map(26).

Meanwhile, the GCA algorithm, recently incorporated into Cirrus OCT, provides the macular GCIPL deviation map by analyzing the GCIPL thickness for each superpixel in an area around the fovea(30, 31). Although there have been a few studies on the diagnostic ability of the GCIPL deviation map for discriminating glaucomatous eyes from normal control eyes, they have shown only whether the GCIPL deviation map detects glaucomatous damage or not(32, 33, 42). For example, Sung et al. (2013) have reported that the GCIPL deviation map showed similar discrimination ability between normal controls and patients with preperimetric or early glaucoma and compared with pRNFL deviation map by quantifying abnormal superpixels(42). However, to our knowledge, there has been no study evaluating the ability of GCIPL deviation map to identify localized RNFL defect and comparing with that of pRNFL deviation map.

In the present study, for the purposes of a quantitative analysis on the ability to detect localized RNFL defect, we

digitized the deviation map results so that the number of abnormal superpixels could be counted by a customized image processing program. Moreover, we used completely automated image processing program for counting the number of abnormal superpixels, which entailed the advantage of eliminating the counting errors and interobserver variability that would be incurred when counting manually or using software requiring manual determination of the boundaries of deviated superpixels.

In our results, all of the severity indices (S1, S2, and S3), which represented the number of abnormal superpixels counted by the image processing program, were larger in glaucomatous eyes with RNFL defects than in normal control eyes (Table 2.1). This means that the results of the deviation map corresponded well to the RNFL defects visible on red-free RNFL photography.

As the scan area of the GCIPL deviation map was limited to the parafoveal region and included only about 50% of RGC population(43), it might not detect RNFL defects located far from the fovea. Therefore, we assumed that the ability of the GCIPL deviation map for detection of localized RNFL defects might prove inferior to that of the pRNFL deviation map that computes data on the full 360-degree peripapillary region, including a total

sampling of the RGC axons. We found that the AUROCs for detection of localized RNFL defects were larger on the pRNFL deviation map than on the GCIPL deviation map. However, the differences were not statistically significant. This was contrary to our earlier expectation (Table 2.3). Determining the exact mechanism involved will require further investigation stratified by the peripapillary widths and locations of RNFL defects.

We found that S2, which indicated the number of red superpixels (abnormal at the probability level < 0.01) and S3, which weighted the number of red superpixels, showed significantly larger AUROCs than S1. Therefore, we suggest observing red superpixels more closely than yellow ones on the deviation map would be better for GCIPL deviation map evaluation.

In addition, the values of sensitivities at fixed specificities (specificity > 95%) in this study generally were not high, which was ranged from 62.3% to 81.2% on the GCIPL deviation map and from 75.4% to 87.0% on the pRNFL deviation map. We speculated that this was due to the fact that we included patients with mostly early stage of glaucoma (40 eyes; 58%), even with preperimetric stage of glaucoma (18 eyes; 26%). Because the

diagnostic performance of imaging device is highly affected by the severity of disease(44), the relatively low diagnostic performance may be related to the early stage of glaucomatous damage in study subjects.

All of the three severity indices were found to be significantly and closely correlated between the two deviation maps (GCIPL and pRNFL) (Table 2.4). That is, the more abnormal superpixels on the GCIPL deviation map corresponded to RNFL defects, the more abnormal superpixels on the pRNFL deviation map also corresponded to RNFL defects. Considering the pathophysiology of the glaucomatous optic neuropathy that primarily affected the RGC and their axons (RNFL), we cautiously speculated that this result was due to the fact that the (glaucomatous) RNFL defect, presenting as abnormal superpixels on the pRNFL deviation map, was mostly accompanied by the RGC loss presenting as abnormal superpixels on the GCIPL deviation map. In this context, we supposed that abnormal lesions corresponding to RNFL defect on the two deviation maps would tend to coincide. In other words, glaucomatous structural damage in these two anatomical lesions

would be detected by both GCIPL and pRNFL deviation maps, respectively.

In this study, the assessment of diagnostic ability was performed under a case-control design, which included eyes with a photographically identifiable (noticeable) RNFL defect (the cases) and normal eyes with a perfectly normal RNFL (the controls). Because the RNFL defects met certain levels of depth and width to qualify as defects on red-free RNFL photography, distinctly glaucomatous eyes with prominent structural changes were enrolled in the present study. Thus, in this study, the diagnostic performance of the deviation map for both scans might have been overestimated relative to real clinical practice, in which very-early-stage glaucomatous eyes with ambiguous RNFL defects are encountered. However, in present study, among 69 eyes with localized RNFL defects, 18 (26.1%) showed asymptomatic VF loss (preperimetric glaucoma) and 40 (58.0%) early-stage VF loss (MD > -6 dB). Considering that 58 of the 69 eyes (84.1%) showed early-stage VF loss, we supposed that the influence of subject selection on the study results was not significant.

The present study has several limitations. First, as aforementioned, we integrated all of the subjects within a preperimetric-to-moderate range of glaucoma. Because the diagnostic ability was influenced by the disease severity, for further, high-precision investigation of the diagnostic accuracies of those deviation maps, subgroup analysis according to disease severity, with a large number of cases, will be mandatory. Second, we excluded diffuse RNFL atrophies and ambiguous RNFL defects from consideration, enrolling only patients with localized RNFL defects and clear margins. Considering, however, that glaucoma patients presenting with solely localized RNFL defects represent only a subset of the total glaucoma patient population, for more comprehensive evaluation of the utility of the GCIPL deviation map for detection of RNFL defects in the clinical field, further study involving defects without clear margins is required. Third, we excluded eyes with a history of ocular or systemic disease possibly affecting the peripapillary area or macula area, and all of our OCT images were obtained after dilation of the pupil. However, in the real-world clinical setting, OCT is frequently performed on eyes with several minor lesions in the macula or peripapillary area without pupil dilation. Therefore, the findings of this study, in the context of real clinical practice, might be

limited. Fourth, the criteria for significantly aberrant superpixels corresponding to RNFL defects used in this study were arbitrary, lacking any anatomical explanation for the cut-off number of superpixels. To compensate for this weakness, we established not a single criterion but rather three criteria for the diversity and the possibility. Consequently, all the analysis showed consistent results, regardless of the criterion.

These limitations notwithstanding, the present study remains significant for its first quantitative comparison of the diagnostic abilities of the GCIPL and pRNFL deviation maps in detecting localized RNFL defects. A quantitative analysis of the respective deviation map algorithms using a customized image processing program showed comparable diagnostic performances for early-to-moderate-stage glaucomatous eyes, though the AU- ROC of the pRNFL deviation map showed numerically better values than that of the GCIPL.

CHAPTER 3

Feature Extraction for Glaucoma Diagnosis 2. RNFL Defect Depth Index of Thickness Deviation Map

3.1. INTRODUCTION

Glaucomatous optic neuropathy is characterized by progressive injury of the optic nerve and retinal nerve fiber layer (RNFL)(38), the RNFL being one of the most important structural parameters for determinations of the presence and progression of glaucoma(45, 46). Because the RNFL is a 3-dimensional (3-D) structure, features and progression patterns of RNFL defects are characterized by depth (the vertical element on the z-axis(45, 47-50)) and area or angular width (the horizontal element on the x-y axis(47, 50-53)). Therefore, evaluation of the 2 parameters has been the focus of numerous studies(45, 47-53). However, depth is particularly difficult to measure accurately with the currently established methods: RNFL photography is limited by subjectivity and its qualitative nature, and optical coherence tomography (OCT)-derived circumpapillary RNFL (cpRNFL) thickness is limited by its questionable ability for differentiating the severity of localized RNFL defects(45-49).

Recently, Cirrus high-definition (HD) OCT (Carl Zeiss Meditec, Inc., Dublin, CA), one of the commercially available

fourth-generation spectral-domain OCT devices, was found to enable volumetric analysis of the RNFL by using RNFL thicknesses at each superpixel of a color-coded 3-D thickness deviation map(51). Whereas the angular width and area of defects recently have been evaluated for this method(51-53), little is known about the depth.

The main purpose of this study was to quantify the depth of localized and diffuse RNFL defects on a Cirrus HD OCT deviation map according to their severity on red-free fundus photography. Additionally, its discriminating ability was compared with that of cpRNFL thickness.

3.2. METHODS

This investigation is based on the RNFL Defect Evaluation Study, an ongoing study of glaucoma and healthy individuals at Seoul National University Hospital (SNUH). This study was conducted in accordance with the ethical standards of the Declaration of Helsinki (1964) and was approved by the institutional review board of SNUH. Informed consent was obtained from all of the subjects.

3.2.1. Subjects

All of the subjects, both glaucomatous and healthy, were enrolled consecutively between July 2009 and December 2012 at SNUH. The glaucoma subjects visited the Glaucoma Clinic of SNUH, and the healthy subjects visited the SNUH Outpatient Department for regular health check-ups, spectacle prescriptions, and minor ocular- surface diseases (e.g., dry eye syndrome) to provide data on RNFL thicknesses representing the upper 95th percentile range of age-matched healthy subjects.

All of the subjects underwent a full ophthalmic examination, including measurement of visual acuity and

refraction, intraocular pressure (IOP) assessment, gonioscopy, fundus examination, standard automated perimetry (Humphrey C 30-2 SITA-Standard visual field [VF]; Carl Zeiss Meditec, Inc. Dublin, CA), disc stereophotography and red-free fundus photography using a digital fundus camera (VX-10, Tokyo, Japan), and Cirrus HD OCT. The OCT and photography were conducted within a 2-month period.

The glaucomatous eyes included in the study had 1 or more of the forms of glaucoma that cause optic disc changes with the VF defect typical of the disease. Glaucomatous optic disc change was defined as either a cup-to-disc asymmetry between fellow eyes of more than 0.3, rim narrowing, notching, excavation, or RNFL defect. Glaucomatous VF defect was defined as manifesting a pattern standard deviation outside the 95% normal confidence limits, glaucoma hemifield test results outside the normal limits, a cluster of 3 or more nonedged points with a P value of less than 0.05 on a pattern deviation plot on 2 consecutive standard automated perimetry assessments (Humphrey C 30-2 SITA standard VF; Carl Zeiss Meditec, Inc.), or a combination thereof(54). Only reliable perimetric results with less than 20% fixation loss and less than 33% false-positive

or false-negative results were included in the analysis. The healthy eyes had a normal optic disc appearance and VF on standard automated perimetry without glaucomatous optic disc or VF change as described above, no history of IOP more than 21 mmHg, and open angles by gonioscopy.

All of the subjects were 30 years of age or older and had a visual acuity of 20/40 or better and a spherical refractive error within the +3.0- to -6.0-diopter range. The following exclusion criteria were applied: a history of ocular intervention other than cataract extraction before or at the time of the imaging tests; any ocular or neurologic disease, including diabetes mellitus, which could have caused VF defect, disc anomaly, or increased IOP; and poor photographic quality or unacceptable quality of OCT image, as described below. If both eyes fulfilled the inclusion criteria, 1 eye was selected randomly.

3.2.2. Red-free fundus photography imaging

Red-free fundus photographs were obtained after maximum pupil dilation. Sixty-degree wide-angle views of the optic disc, carefully focused on the retina using the built-in split-line focusing device and centered between the fovea and

the optic disc, were obtained and reviewed on a liquid crystal display monitor. Two independent observers evaluated the photographs in a random order and masked fashion. They classified the red-free RNFL photographs of the glaucomatous eyes as showing one of the following: no visible RNFL defect, localized defect, diffuse defect, or ambiguous information. This method has been described in detail elsewhere(47, 48). After excluding subjects with ambiguous information or no visible defect, those classified by both observers as having either diffuse or localized RNFL defects were included in this study. Subjects about whom the 2 observers disagreed were classified as ambiguous and were excluded from further analysis. In the case of glaucoma subjects having RNFL defects satisfying the above-noted criteria in both the superior and inferior hemifields, 1 hemifield was selected randomly to avoid bias in the depth determination(52).

The RNFL defect severity was classified into 3 grades based on an evaluation of the brightness and texture of the RNFL and the degree of blood vessel obscuration (Figure 3.1): grade 1 (G1), mild defect; grade 2 (G2), moderate defect; and grade 3 (G3), severe defect. Grade 1 indicates that fine RNFL striations

are visible and only large vessels are clearly visible, G2 indicates that RNFL striations are barely detectable and even small vessels are clearly visible, and G3 indicates that RNFL texture is not at all visible. This method has been described in detail elsewhere(45, 47, 48). Interobserver discrepancies were resolved by adjudication of a third experienced observer. The measurements showed good interobserver and intraobserver agreement (intraobserver intraclass correlation coefficient values 0.88, 0.87; interobserver intraclass correlation coefficient, 0.84).

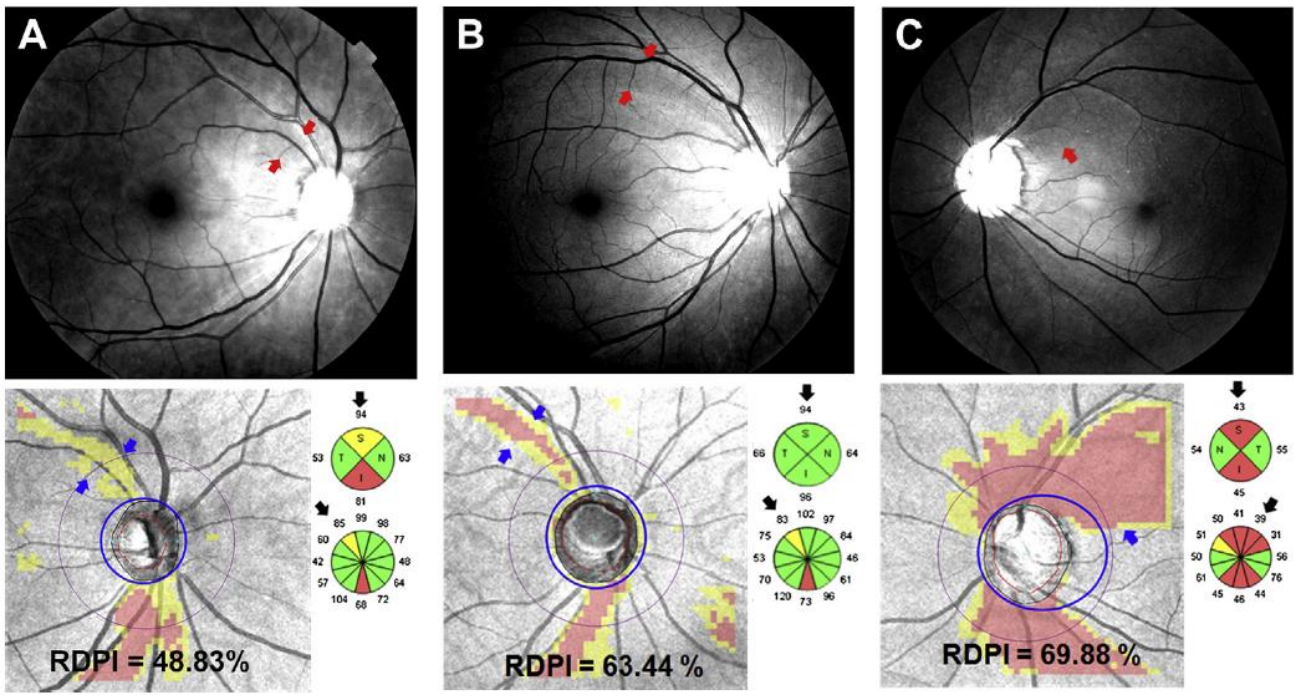

Figure 3.1. Retinal nerve fiber layer (RNFL) defects shown on the superior hemifield on red-free fundus photographs (red arrows), in Cirrus high-density optical coherence tomography-derived circular diagrams of circumpapillary

RNFL (cpRNFL) thicknesses (black arrows), and on RNFL thickness deviation maps (blue arrows). Cases of (A) grade 1 (G1), (B) grade 2 (G2), and(C) grade 3 RNFL defects showing mild, moderate, and severe RNFL loss, respectively. The RNFL defect depth percentage index (RDPI) values using the RNFL thickness deviation map showed notable differences among the 3 groups, whereas the cpRNFL thicknesses of the G1 and G2 groups did not show remarkable differences. I = inferior; N =− nasal; S = superior; T = temporal

3.2.3. Optical coherence tomography retinal nerve fiber layer imaging

The principles of spectral-domain OCT have been discussed in detail elsewhere(55, 56). After pupillary dilation to a minimum diameter of 5 mm, an optic disc cube scan in a 6×6-mm^2 para- papillary area (200×200 pixels) was obtained using Cirrus HD OCT (version 6.0). Each pixel was coded either yellow or red if the RNFL measurement was less than the lower 5% or 1% percentile range, respectively, on the RNFL thickness deviation map.

The system also calculates the RNFL thickness at each point on a set-diameter (3.46-mm) circle consisting of 256 A-scans that were positioned automatically around the optic disc, thus generating a circular diagram of average RNFL thickness of four 90° RNFL quadrants and twelve 30° RNFL clock-hour sectors. Only scans of good image quality, as defined by the following criteria, were used: signal strength of 6 or more (maximum, 10), no algorithm segmentation failure in the peripapillary RNFL extracted B-scan, and no saccadic or motion

artifacts as detected by line- scanning ophthalmoscopy overlaid with OCT enface images.

3.2.4. Measuring depth of retinal nerve fiber layer defects on cirrus high-definition optical coherence tomography derived deviation map

For both the glaucoma and healthy subjects, RNFL thicknesses of 200×200 pixels on the RNFL thickness map were obtained as an Excel file (Microsoft Corp. Redmond, WA) using the Advanced Extraction analysis ability of Cirrus HD OCT. Then, thicknesses of 50×50 superpixels were obtained by averaging those of 4×4 pixels and were superimposed on those of the RNFL thickness deviation map by 1 examiner masked to the subjects' identities and test results using Matlab R2012a (The MathWorks, Inc. Natick, MA; Figure 3.2). We found that 2 superpixels of the 4 lateral (right, left, upper, and lower) areas of the deviation map were not assessable because they were not coded in red or yellow on areas corresponding to the RNFL defects (Figure 3.2A). Therefore, 46×46 superpixels of RNFL thicknesses and deviation map images were obtained after removing the 2 superpixels of the 4 lateral areas, and those

located within the blue empty circle, representing the margin of the optic disc or b zone parapapillary atrophy, if present, were removed (Figure 3.2B). An observer masked to the subjects' identities and test results outlined the margin with reference to the optic disc photographs. The left-eye data were converted to the right-eye format.

The inclusion criteria for RNFL defects on the deviation map were as follows(51, 52): 10 or more superpixels colored red or yellow extending 3 or more superpixels away from the margin of the optic disc or b zone parapapillary atrophy and 1 or more superpixels located within the region between the margin of the optic disc or b zone parapapillary atrophy and 1.73 mm from the disc center. An RNFL defect depth percentage index (RDPI) was derived for measurement of the proportion of the RNFL defect depth (as a percentage) according to the following equation: 100×(1-[summation of thicknesses of RNFL defects {red or yellow superpixels}/summation of RNFL thicknesses of upper 95th percentile range of age-matched healthy subjects in areas corresponding to RNFL defects]) (Figure 3.2B). If there were 2 or more isolated defects meeting the above-noted

inclusion criteria, only the defect having the largest number of superpixels was determined automatically by the Matlab program.

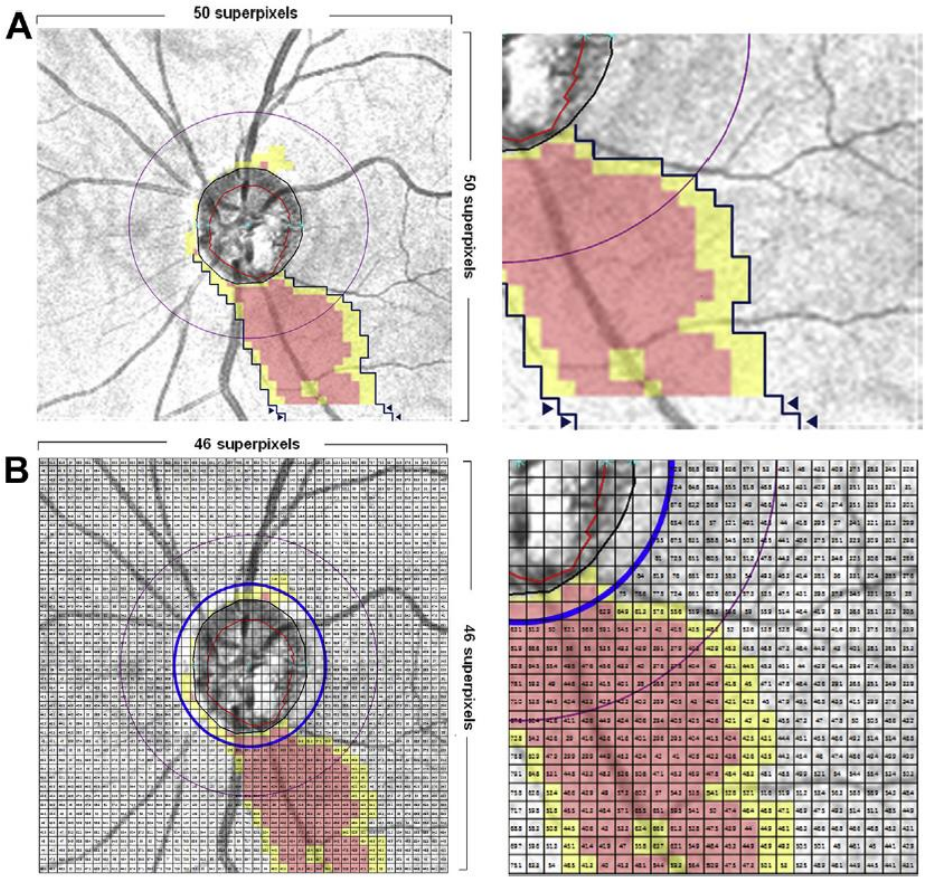

Figure 3.2. Measurement of the depth of retinal nerve fiber layer (RNFL) defect on a spectral-domain optical coherence tomography (OCT)-derived RNFL thickness deviation map. A magnified view of the inferotemporal quadrant of the image is shown on the right. A, Two superpixels of the 4 lateral (right, left, upper, and lower) areas of the deviation map composed of 50 x 50 superpixels were not assessable. Note that 2 superpixels corresponding to the RNFL defect were not coded in red or

yellow (blue arrowheads). B, After removing the 2 superpixels of the 4 lateral areas, 46 x 46 superpixels of RNFL thicknesses superimposed on the deviation map were obtained, and those corresponding to optic disc and βzone parapapillary atrophy (blue empty circle) on the deviation map were removed. The RNF L defect depth percentage index values were derived by the equation 100 x (1 − [summation of RNFL thicknesses of RNFL defects {red or yellow superpixels}/summation of thicknesses of upper 95^{th} percentile range age−matched healthy subjects in areas corresponding to RNFL defects]).

3.2.5. Data analysis

The regression model analyses were used to estimate the upper 95th percentile RNFL thickness values of the age-matched healthy subjects using the same format as that used for the normative Cirrus HD OCT reference values(57). The 1-way analysis of variance (ANOVA) was used to compare the OCT-derived parameters according to the severity of the RNFL defect on red-free fundus photography. Results with statistical significance also were provided after Bonferroni correction based on the number of comparisons. Receiver operating characteristic curves were calculated to compare the usefulness of RDPI and cpRNFL thickness in discriminating RNFL defect severity on red-free fundus photography. An area under the receiver operating characteristic curve (AUROC) was compared using the method described by DeLong et al(37). All of the analyses were performed using Medcalc software version 10.0 (Medcalc Software, Mariakerke, Belgium) and SPSS for Windows version 20.0 (SPSS, Inc., Chicago, IL). A value of $P < 0.05$ was accepted as statistically significant.

3.3. RESULTS

Initially, 613 eyes of 613 subjects (385 glaucoma patients and 228 healthy subjects) were enrolled in the study. Of the 385 glaucomatous eyes, 9 with no visible RNFL defect and 26 with ambiguous information were excluded, thus leaving 350 with localized and diffuse RNFL defects for further evaluation by Cirrus HD OCT. Of these 350 glaucomatous eyes, 22 with unacceptable Cirrus HD OCT scans and 13 that did not meet the inclusion criteria for RNFL defect on the deviation map were excluded. Among the 228 healthy eyes, 11 with unacceptable Cirrus HD OCT scans were excluded. Therefore, a total of 532 eyes of 532 subjects (315 glaucomatous eyes with localized or diffuse RNFL defects and 217 healthy eyes) finally were selected for analysis.

The mean age of the healthy subjects, 55.0±14.2 years, was not significantly different from that of the glaucoma subjects, 56.6±11.1 years (P = 0.1332, unpaired t test; Figure 3.3, available at www.aaojournal.org; Table 1). The mean values of refraction and of optic disc area were —0.57±1.4 diopters and 2.1±0.4 mm^2, respectively, and the male-to-female ratio was

99:118. All of the healthy subjects' RNFL thicknesses on the thickness map passed a normality test ($P < 0.05$, Kolmogorov–Smirnov test); the mean thickness values are plotted in Figure 3.4A (available at www.aaojournal.org). The RNFL thicknesses decreased with age in most areas of the RNFL thickness map on regression analysis ($P < 0.05$; Figure 3.4B, available at www.aaojournal.org). No other factors, including refraction and optic disc area, showed a significant correlation with thickness in most areas of the map ($P > 0.05$; data not shown).

The glaucoma subjects' baseline characteristics according to RNFL defect severity on red-free fundus photography are provided in Table 3.1. The mean age, sex, type of glaucoma, laterality, spherical error, IOP at the time of testing, optic disc area, and signal strength of Cirrus HD OCT did not significantly differ by defect severity ($P > 0.05$). The VF mean deviation and average RNFL thickness of the Cirrus HD OCT decreased, and the VF pattern standard deviation increased as the defect severity increased. The RNFL defects more commonly were located in the superior area in G1 and in the inferior area in G3. Group 3 had a larger proportion of the diffuse defects than G1 and G2 (Table 3.1).

Table 3.1. Baseline Characteristics of Glaucoma Subjects

	All (n = 315)	Grade 1 (n = 108)	Grade 2 (n = 96)	Grade 3 (n = 111)	P Value
Age (yrs)	56.6±11.1	55.5±10.5	57.2±11.6	57.5±11.2	0.292*
Sex ratio (male/female)	164/151	59/49	47/49	58/53	0.720†
Diagnosis (n)					0.4728†
Open-angle glaucoma with normal tension	256	94	78	84	
Open-angle glaucoma with high tension	40	10	12	18	
Chronic angle-closure glaucoma	16	4	5	7	
Pseudoexfoliation glaucoma	3	0	1	2	
Laterality (right/left)	154/161	47/61	50/46	57/54	0.385†
Spherical error (diopters)	-1.0±1.9	-1.0±1.8	-1.0±1.9	-1.0±2.0	0.995*
IOP (mmHg)	13.9±2.8	13.6±2.8	13.8±2.9	14.2±2.8	0.438*
VF MD (dB)	-6.2±5.8	-3.9±3.6	-5.1±4.8	-9.5±6.9	**<0.001***
VF PSD (dB)	6.1±4.3	4.3±3.8	5.3±3.5	8.4±4.3	**<0.001***
Location of RNFL defect (superior/inferior)	156/159	78/30	46/50	32/79	**<0.0001**†
Type of RNFL defect (localized /diffuse)	235/80	87/21	79/17	69/42	**0.0009**†
Optic disc area (mm^2)	2.1±0.5	2.2±0.5	2.1±0.5	2.1±0.5	0.241*
Signal strength of Cirrus high-density OCT	7.5±0.9	7.7±0.9	7.5±0.9	7.5±0.9	0.150*
Average RNFL thickness by Cirrus high-density OCT (μm)	73.4±11.8	78.3±10.7	75.3±10.0	67.1±11.4	**<0.001***

IOP = intraocular pressure at time of testing; MD = mean deviation; PSD = pattern standard deviation; RNFL = retinal nerve fiber layer; OCT = optical coherence tomography; VF = visual field.
Significant values are in boldface.
*One-way analysis of variance with Bonferroni correction.
†Chi-square test.

Table 3.2 and Figure 3.5 show the Cirrus OCT-derived RDPI and cpRNFL thicknesses according to the RNFL defect severity on red-free fundus photography. For the superior hemifield, the RDPIs of G1, G2, and G3 were 55.9±3.3%, 62.4±1.9%, and 67.4±1.6%, respectively. Similarly, for the inferior hemifield, the RDPIs of G1, G2, and G3 were 56.0±3.4%, 64.0±3.1%, and 68.6±2.0%, respectively. The RDPIs increased with increasing RNFL defect severity, both before and after post hoc analysis with Bonferroni correction for both the superior and inferior hemifields ($P < 0.05$, 1-way ANOVA test with Bonferroni correction; Table 3.2, Figure 3.5). With regard to the cpRNFL thicknesses, significant differences were observed among the 3 subgroups in all areas except the 2- and 4-o' clock sectors ($P < 0.05$, 1-way ANOVA; Table 3.2). Post hoc analysis revealed that G2 and G3 were significantly different in all areas except the 1-, 2-, and 4-o' clock sectors ($P < 0.05$, 1- way ANOVA test with Bonferroni correction). However, G1 and G2 did not differ in any areas except the superior and inferior hemifields and the 7-, 11-, and 12-o' clock sectors ($P > 0.05$, 1-way ANOVA with Bonferroni correction).

Table 3.2. Optical Coherence Tomography–Derived Circumpapillary Retinal Nerve Fiber Layer (RNFL) Thicknesses and RNFL Defect Depth Percentage Index According to RNFL Defect Severity.

	Grade 1	Grade 2	Grade 3	*P* Value*	Post Hoc
Superior RNFL					
No.	78	46	32		
RDPI (%)	55.9±3.3	62.4±1.9	67.4±1.6	**<0.001**	G1<G2<G3
Average RNFL (μm)	78.2±10.7	74.9±9.9	67.5±12.5	**<0.001**	G1, G2>G3
Superior RNFL (μm)	94.9±16.2	87.7±13.2	70.3±15.5	**<0.001**	G1>G2>G3
10-o'clock RNFL (μm)	64.6±11.5	66.0±14.2	52.7±16.5	**<0.001**	G1, G2>G3
11-o'clock RNFL (μm)	86.8±10.8	80.5±15.2	56.0±12.0	**<0.001**	G1>G2>G3
12-o'clock RNFL (μm)	111.1±23.9	92.0±20.1	72.6±23.2	**<0.001**	G1>G2>G3
1-o'clock RNFL (μm)	94.1±22.2	90.4±19.8	82.1±20.7	**0.020**	G1>G3
2-o'clock RNFL (μm)	73.6±15.1	70.6±13.0	67.5±11.5	0.096	
Inferior RNFL					
No.	30	50	79		
RDPI (%)	56.0±3.4	64.0±3.1	68.6±2.0	**<0.001**	G1<G2<G3
Average RNFL (μm)	78.4±10.9	75.7±10.2	66.8±1.0	**<0.001**	G1, G2>G3
Inferior RNFL (μm)	92.6±11.1	83.4±13.3	65.1±13.4	**<0.001**	G1>G2>G3
4-o'clock RNFL (μm)	59.5±8.7	60.4±10.0	57.0±8.6	0.103	
5-o'clock RNFL (μm)	81.2±13.3	81.0±16.1	70.2±17.1	**<0.001**	G1, G2>G3
6-o'clock RNFL (μm)	98.0±17.4	90.8±23.0	66.6T21.0	**<0.001**	G1, G2>G3

7-o'clock RNFL (μm)	98.5±18.2	77.5±16.4	55.9T12.0	**<0.001**	G1>G2>G3
8-o'clock RNFL (μm)	64.9±15.4	57.5±14.5	51.9T14.0	**<0.001**	G1, G2> G3

RDPI = retinal nerve fiber layer defect depth percentage index; RNFL = retinal nerve fiber layer.

Data are mean ± standard deviation unless otherwise indicated.

*One-way analysis of variance with Bonferroni correction. Significant values are in boldface.

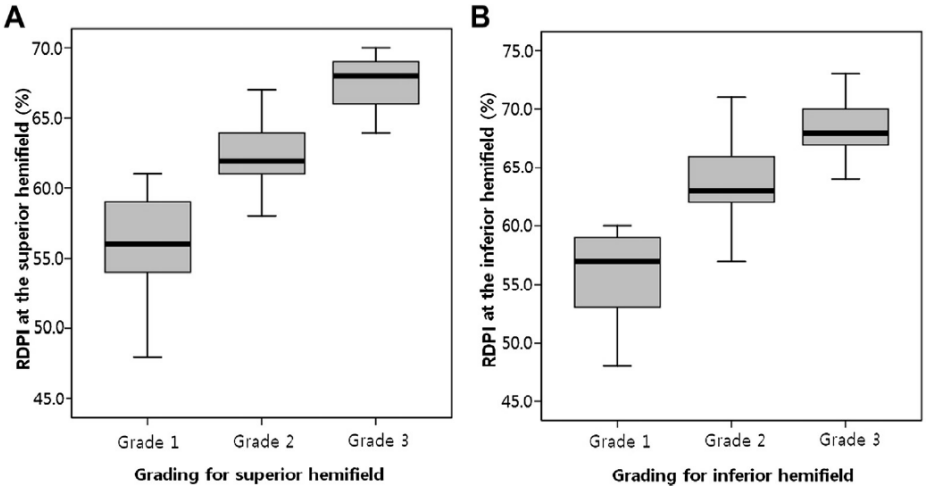

Figure 3.3. Box-and-whisker plots showing the distribution of optical coherence tomography-measured retinal nerve fiber layer (RNFL) defect depth percentage index (RDPI) values according to the severity of the RNFL defect. The RDPI value increase with increasing severity of the RNFL defect in both (A) superior and (B) inferior RNFL hemifields.

Table 3.3 provides the AUROCs of the RDPIs and cpRNFL thicknesses in the superior and inferior areas. For the superior hemifield, the cpRNFL thickness parameter with the largest AUROC for discriminating G1 from G2 was the 12-o'clock sector (0.674), and that for discriminating G2 from G3 was the 11-o'clock sector (0.904). For the inferior hemifield, the 7-o'clock sector had the largest AUROC (0.794, G1 vs. G2; 0.870, G2 vs. G3). The AUROCs of the RDPIs (0.969 and 0.975 in the superior and inferior hemifields, respectively) were larger than those of all of the cpRNFL thickness parameters for discriminating between G1 and G2 ($P < 0.05$; Table 3.3; Figure 3.1A, B). For discriminating G2 from G3, the AUROCs of the RDPIs (0.961 and 0.891 in the superior and inferior hemifields, respectively) were larger than the cpRNFL thicknesses in all areas except the inferior quadrant and the 6-, 7-, and 11-o'clock sectors ($P < 0.05$; Table 3.3). The RDPI cutoff points showing the best ability in discriminating between G1 and G2 were 60.1% and 59.8% in the superior and inferior hemifields, respectively; those for discriminating G2 from G3 were 64.5% and 66.4% in the superior and inferior hemifields, respectively.

Table 3.3. Area under the Receiver Operating Characteristic Curves of the Optical Coherence Tomography-Derived Circumpapillary Retinal Nerve Fiber Layer (RNFL) Thickness and RNFL Defect Depth Percentage Index for Discriminating of RNFL Defects According to Severity.

	Grade 1 vs. Grade 2 Area Under the Receiver Operating Characteristic Curve (Standard Error)	P Value*	Grade 2 vs. Grade 3 Area Under the Receiver Operating Characteristic Curve (Standard Error)	P Value*
Superior RNFL				
RDPI	0.969 (0.0164)	-	0.961 (0.0190)	-
Average RNFL	0.572 (0.0536)	<0.0001	0.667 (0.0670)	<0.0001
Superior RNFL	0.637 (0.0514)	<0.0001	0.819 (0.0538)	0.0102
10-o'clock RNFL	0.561 (0.0557)	<0.0001	0.758 (0.0571)	0.0006
11-o'clock RNFL	0.664 (0.0537)	<0.0001	0.904 (0.0350)	0.1342
12-o'clock RNFL	0.674 (0.0501)	<0.0001	0.723 (0.0623)	0.0001
1-o'clock RNFL	0.541 (0.0534)	<0.0001	0.624 (0.0671)	<0.0001
2-o'clock RNFL	0.543 (0.0543)	<0.0001	0.565 (0.0662)	<0.0001
Inferior RNFL				
RDPI	0.975 (0.0140)	-	0.891 (0.0318)	-
Average RNFL	0.597 (0.0681)	0.0001	0.715 (0.0458)	0.0018
Inferior RNFL	0.727 (0.0585)	0.0002	0.830 (0.0346)	0.2099

4-o'clock RNFL	0.515 (0.0676)	<0.0001	0.589 (0.0520)	<0.0001
5-o'clock RNFL	0.522 (0.0669)	<0.0001	0.682 (0.0464)	0.0005
6-o'clock RNFL	0.605 (0.0640)	<0.0001	0.793 (0.0395)	0.0506
7-o'clock RNFL	0.794 (0.0509)	0.0015	0.870 (0.0306)	0.3642
8-o'clock RNFL	0.616 (0.0659)	<0.0001	0.646 (0.0494)	<0.0001

RDPI = retinal nerve fiber layer defect depth percentage index.
*Pairwise comparison of the area under the receiver operating characteristic curve values between the RDPI and each circumpapillary RNFL thickness parameter. Significant values are in boldface.

3.4. DISCUSSION

In this study, the RDPI using the RNFL thickness deviation map showed an excellent correlation with the RNFL defect severity on red-free RNFL photography. Furthermore, this parameter demonstrated a discriminating ability superior to that of the circular diagram of the cpRNFL thicknesses, especially in discriminating between mild and moderate RNFL defects. The RDPI cutoff value for discriminating between mild and moderate RNFL defects was approximately 60%, and that for discriminating between moderate and severe defects was approximately 65%. This suggests that a novel parameter using 3-D volumetric analysis of the Cirrus HD OCT-derived RNFL thickness deviation map can be the practical standard for objective differentiation of RNFL defect depth.

For its effectiveness in quantitatively obtaining optimal RNFL thickness measurement, OCT has drawn attention as potential complement to RNFL photography, the conventional method. A recent study by Jeoung et al(48) showed that objective assessment of the severity of diffuse RNFL atrophy was possible using Stratus OCT-derived cpRNFL thickness.

However, with the third-generation time-domain Stratus OCT used in their study, 3-D volumetric assessment of the RNFL was not possible. Moreover, their analysis did not include localized RNFL defects that may be of relatively smaller area and angular width than diffuse defects(48, 49, 51). In the present study, the proportion of localized defects in the mild and moderate severity subgroups (G1 and G2) was larger than that in the severe subgroup (G3). This suggests that relatively large numbers of RNFL defects of G1 and G2 may have smaller angular widths than those of circumpapillary circular diagram, which cannot be smaller than 30° (Figure 3.1A, B). For these defects, thicknesses of healthy RNFLs can be included, thus diminishing the discriminating ability of the circular diagram. Meanwhile, a new parameter used in the present study, the RDPI, is based on each superpixel of a 6×6-mm^2 RNFL thickness deviation map with a relatively small measurement area of approximately 0.014 mm^2; thus, the possibility of including a healthy RNFL is minimized, even for small mild to moderate defects. These differences in the present study's design might have contributed to its different results compared with the study by Jeoung et al. The discriminating ability of the cpRNFL

thicknesses for mild (G1) and moderate (G2) defects in the present study (largest AUROCs, 0.674 and 0.794 for superior and inferior hemifields, respectively) were lower than those in the study by Jeoung et al (largest AUROCs, 0.820 and 0.922), whereas those for moderate (G2) and severe (G3) defects in the present study (largest AUROCs, 0.904 and 0.870) were similar to those of Jeoung et al (largest AUROCs, 0.861 and 0.899). Therefore, the advantage of the RDPI over cpRNFL thickness was notable in discriminating between mild and moderate defects, not in discriminating between moderate and severe defects. This issue is also related closely to the correlation between the area and depth of RNFL defects(51). Because the present study demonstrated that quantification of defect depth, as well as area(51), is possible using the RNFL thickness deviation map, further studies evaluating the correlation between these 2 parameters perhaps will yield clues to the sensitive parameter to detect RNFL progression according to the RNFL defect severity.

For measurement of RNFL defect depth, the RDPI was derived according to the ratio of the RNFL defect thickness to the thickness of the upper 95th percentile range of age-matched healthy subjects in areas corresponding to the RNFL defects.

This approach has advantages over the use of absolute values of RNFL thickness because age and the circumferential location of the optic nerve head, which can influence absolute thickness values as well as defect severity, can be adjusted by using RDPI.

In this study, the mean RDPIs of mild (G1) and severe (G3) RNFL defects were approximately 56% and 68%, respectively. This result is comparable with those of Quigley and Addicks(58), Hood and Kardon(59), and Hood et al(60, 61) that clinical detection of the RNFL defect was possible after loss of 50% of the neural tissue in primate eyes(58), and residual RNFL thickness after complete loss of axons is approximately 33% of the healthy thickness because of vessels and glial cells(59-61).

The ability of RDPI to identify glaucoma, as well as RNFL defect, may be of clinical importance. When compared with average cpRNFL thickness in discriminating between 217 healthy eyes and 315 glaucomatous eyes, the RDPI showed superior performance (RDPI AUROC, 0.957 vs. average cpRNFL thickness AUROC, 0.915; P = 0.0336; data not shown). This suggests that the RDPI may provide a useful alternative to determine the presence of glaucoma. Further studies using a

separate control group of healthy subjects are warranted to clarify this issue.

We used the RNFL grading system proposed by Quigley et al(45) using red-free fundus photography as the reference standard. This method has an advantage over other grading systems(62) and methods (e.g., confocal scanning laser ophthalmoscopy RNFL photography(52)) in that no reference set is required and a relatively larger scan area is obtained. Additionally, the present results showed good intraobserver and interobserver agreement in the determined defect severities, in accordance with several previous studies that have used this method(47, 48).

In conclusion, the RDPI, a novel parameter using the Cirrus HD OCT-derived RNFL thickness deviation map, can be useful in objectively quantifying RNFL defect depth. This parameter has an advantage over cpRNFL thickness in discriminating between mild and moderate RNFL defects, not in discriminating between moderate and severe defects.

CHAPTER 4

Glaucoma Classification using Deep Feature Fusion Network

4.1. INTRODUCTION

Glaucoma is a disease which is caused by abnormalities of optic nerve through elevated intraocular pressure (IOP) and impaired blood supply(38). Patients with glaucoma suffer from vision loss according to defects in the optic nerve. Many patients notice they have visual field defects after optic nerve defects have already progressed(11). Therefore, glaucoma diagnosis of early stage is important. However, it is very challenging to diagnose glaucoma of early stage because glaucoma occurs for several reasons and progresses slowly.

Spectral domain optical coherence tomography (SD-OCT) examinations is very important for glaucoma diagnosis because SD-OCT can measure the optic nerve thickness in the region of interest such as peripapillary retinal nerve fiber layer (pRNFL) and ganglion cell-inner plexiform layer (GCIPL) and can analyze whether the measured thickness is statistically abnormal compared with that of the normal person(23). However, the processing techniques such as image processing, machine learning are necessary for precise analysis of these measured results. Previous researches studied that glaucoma can be classified by machine learning using features such as thickness

average in quadrants, pRNFL symmetry, rim area, disc area, average cup-disk ratio, vertical dup-disk ration and cup volume(63-65). But, these researches used only the overall features in the thickness map and did not reflect information that affects glaucoma in the thickness map.

Therefore, deep feature fusion network (DFFN) were developed by fusing feature to analyze glaucoma accurately in this study. Deep learning was trained with the thickness map data, which allowed computer to extract meaningful features for glaucoma diagnosis directly. And DFFN and other method were evaluated by the accuracy and area under receiver operating characteristic curve (AUROC).

4.2. METHODS

4.2.1. Study subjects

All study subjects underwent a complete ophthalmic examination including visual acuity and refraction assessment, slit-lamp biomicroscopy, gonioscopy, Goldmann applanation tonometry (Haag-Streit, Koniz, Switzerland), dilated stereoscopic examination of the optic disc, digital color stereo disc photography, red-free RNFL photography, and Cirrus HD-OCT using the Optic Disc Cube 200 × 200 protocol. The patients also underwent central corneal thickness measurement (Orbscan™ 73 II, Bausch & Lomb Surgical, Rochester, NY, USA) and axial-length measurement (IOL Master™ ver. 5, Carl-Zeiss Meditec, Dublin, CA, USA) at the baseline examination.

We retrospectively recruited glaucoma patients with various severities of visual field (VF) loss representative of different degrees of glaucomatous optic neuropathy in primary open-angle glaucoma (POAG). Glaucoma was diagnosed, irrespective of untreated intraocular pressure (IOP) level, when certain characteristic changes (i.e., localized or diffuse

neuroretinal rim thinning, RNFL defect) were present in the optic disc and/or retina, as visible on stereo disc photography images or red-free RNFL photographs. The patients were divided into three groups based on their standard automated perimetry (Humphrey VF Analyzer™ 30-2 SITA-Standard strategy; Carl Zeiss Meditec, Inc., Dublin, CA, USA) results: an early glaucoma group (VF MD better than -6 dB) and a mild-or-severe glaucoma group (VF MD worse than -6 dB). A glaucomatous VF defect was diagnosed when at least two consecutive VF tests yielded a glaucoma hemifield test result outside the normal limits. The defect was diagnosed when at least two consecutive VF tests bore three or more contiguous test points of $P < 0.01$, at least one of which was of $P < 0.05$, within the same hemifield on a pattern deviation plot. These tests also required reliability indices better than 15%.

The exclusion criteria were: (1) spherical refraction less than -6 diopters (D) and greater than +3 D, (2) the existence of any type of glaucoma other than POAG, (3) diseases in the external eyes or retina, and (4) a history of intraocular surgery besides cataract surgery. We also excluded the following eyes: those with an optic disc torsion of more than 15° 16 or a tilt ratio

(minimum-to-maximum optic disc diameter) less than 0.75 17, and those with any abnormalities in the circumpapillary region, including marked peripapillary atrophy, that affected the scan ring where the OCT average RNFL thickness measurements were obtained. Whether or not to exclude the patient has been determined by an experienced examiner (YKK). In cases where both eyes were eligible, one eye was selected randomly.

In addition, the normal eyes in our study met the following criteria: (1) baseline IOP lower than 22 mmHg, and no history of IOP elevation; (2) no glaucomatous optic disc on the SDP image; (3) no RNFL defects on the red-free fundus image; 4) normal visual field on 30-2 threshold test.

4.2.2. OCT imaging

The pupil in the qualifying eye of each participant was dilated using 1% tropicamide and 2.5% phenylephrine eye drops 10–15 min prior to scanning. All scans were acquired using the Optic Disc Cube 200 × 200 protocol of Cirrus HD-OCT™; this protocol is designed to position the cube scan on the ONH, and is primarily utilized in glaucoma analysis. After the subject was seated and properly aligned, the iris was brought into view using

a mouse-driven alignment system, and the line-scanning ophthalmoscopic image was focused, adjusting for refractive error. The ONH was then centered on the live image, after which the centering (Z-offset) and enhancement (polarization) were optimized. The laser then scanned a 6 mm × 6 mm area, capturing a cube of data consisting of 200 A-scans from 200 linear B-scans (40,000 points) in about 1.5 seconds (27,000 A-scans/sec). The ONH parameters were automatically measured by a Carl Zeiss Meditec analysis algorithm developed for Cirrus HD-OCT (version 6.0). Only good-quality scans (signal strength ≥ 7, without RNFL discontinuity or misalignment, without involuntary saccade or blinking artifacts, without RNFL algorithm segmentation failure) were used in the analysis. None was removed due to ONH algorithm failure.

The pRNFL thickness values were converted to color values according to thickness on the thickness map. The pRNFL thickness values were analyzed and then represented on color-coded deviation maps composed of 50 x 50 superpixels (200 x 200 pixels) on the basis of a comparison with the built-in internal normative database. The uncolored (grey color) superpixels indicated the normal range, whereas yellow- or

red- colored superpixels indicated abnormality at the 5% or 1% level, respectively. The pRNFL deviation map represented the OCT enface image of the optic disc that showed the boundaries of the cup, disc and 3.46-mm diameter circle.

4.2.3. Deep Feature Fusion Network

Deep Feature Fusion Network (DFFN) is machine learning algorithm that fuses features extracted by deep learning and hand crafted. DFFN was used to analyze facial recognition and answer prediction(66, 67). Glaucoma was analyzed to DFFN that fuses hand crafted features in Chapter 2, 3 and features extracted by deep learning. Deep learning algorithm was selected for convolutional neural network (CNN) which is used to analyze images and videos. In this study, glaucoma is analyzed more precisely by fusing the features selected by the computer learned to deep learning algorithm and the features selected manually through DFFN.

Hand crafted features were extracted by applying the method analyzed in Chapter 2 and Chapter 3. The features selected in Chapter 2 were analyzed in the deviation map and the thickness map. Since three severity indices extracted according

to three criteria, nine features were extracted through Chapter 2. One RNFL defect depth percentage index (RDPI) was selected because the feature extracted through Chapter 3 is the ratio of the thickness of the normal person to the thickness of the thickness map in the region detected as defects in the deviation map.

Deep learning is a machine learning algorithm that allows the computer to directly learn the images to be analyzed in the deep neural network architecture and extract the appropriate features automatically. CNN is a popular deep learning algorithm used in images and videos. It consists of convolutional layer, pooling layer and fully connected layer(68). Convolutional layer is a step of extracting the features between pixels in an image and acts as a filter for detecting features in image processing. Pooling layer selects meaningful values such as average value and maximum value among features extracted from convolutional layer. Fully connected layer combines all the features extracted from the convolutional layer and the pooling layer and analyzes the relationship between each feature. Since the deep learning algorithm is computationally expensive and the extracted features are likely to be over-fitted, CNN method is analyzed in

two ways for medical images with small data. One method is to learn from the weights set randomly with a deep learning network having small depth. Another method is to learn in a deep learning network having deeper depth where weights are calculated and analyzed for other images similar to medical images(69).

In the first method analyzing with randomly set weights, LeNet was used. LeNet networks consists of two convolutional layers, two pooling layers and three fully connected layers(70). Since the image size using this network should be 32 X 32 gray scale image, the thickness map was converted. The thickness map is a color map according to the optic nerve thickness of 0 ~ 370 μm. So, the value of each pixel of the thickness map was converted to the optic nerve thickness and this value was converted into the value of 0 ~ 255 which can represent the image. And the full size of the image was resized to 32 X 32 and made into an image that can be learned by LeNet. Since the amount of data must be large to be able to be learned better, augmentation is performed in which the amount of image data is multiplied by the image containing the Gaussian noise. And the

weights before learned was used as a random Gaussian distribution and a rectified linear unit (ReLU) was added to the convolutional layer and the fully connected layer to be learned better nonlinearly(71). The last features calculated from the fully connected layer were classified into groups using the softmax function. Learning rate and momentum parameters were selected in order to find the best learning condition. Since last fully connected layer should have three classes (normal, early glaucoma, mild or severe glaucoma), other fully connected layers had 120 and 50 features according to the changed ratio. And it was judged whether learning was successful by changing the loss value according to epoch. Learning rate parameter varied from 0.1 to 0.0001 and momentum parameter varied with 0.5, 0.7 and 0.9.

The learning method using pre-trained deep neural network is called transfer learning. This method is to train the fully connected layer or the last layer of convolutional layer in the deep neural network architecture trained for the other large amount of image data. In the transfer learning, fully connected layers that were changed in the network were trained at the learning rate which is about 1/10 of the learning rate when

training the previous data and convolutional layers were trained to about 1/100 of the learning rate. This method was analyzed to VGGNet which consists of 13 convolutional layers, 5 pooling layers and 3 fully connected layers(72). VGGNet network was pre-trained by ImageNet. ImageNet is image data most commonly used for deep learning as 100,000 image data(73, 74). Since the image size should be match the previously trained image size in order to be learned by VGGNet, the thickness map was resized into RGB image of 96 X 96 X 3 and the amount of data increased through image augmentation that adds Gaussian noise as it was trained in LeNet. Also, ReLU was used next to convolutional layer and fully connected layer to make network learning well. The final analysis was used to the softmax function and the network was trained at learning rate varied from 0.001 to 0.0001. Since last fully connected layer should have three classes (normal, early glaucoma, mild or severe glaucoma), other fully connected layers had 1024 and 120 features according to the changed ratio. And two transfer learning methods was studied and compared, one is to change fully connected layers and the other is to change convolutional layers and fully connected layers.

DFFN was analyzed by fusing features selected to deep learning and hand crafted (Figure 4.1). The features extracted by deep learning were analyzed by extracting the features of two kinds of fully connected layers in LeNet and VGGNet respectively. DFFN was analyzed by using artificial neural network (ANN) classifier with these features(67). ANN classifier is a common learning model associated with mathematical algorithm that interpret and recognize a pattern. ANN classifier has been discovered to use nonlinear data with several features and to investigate the performance of features in an automated pattern recognition system(75). ANN classifier was analyzed according to hidden layer and analyzed by 5-fold cross validation(76).

$$Softmax_n(x) = \frac{e^{x_n}}{\sum_{k=1}^{N} e^{x_k}}, n \in \{1,2,3,\dots,N\}$$

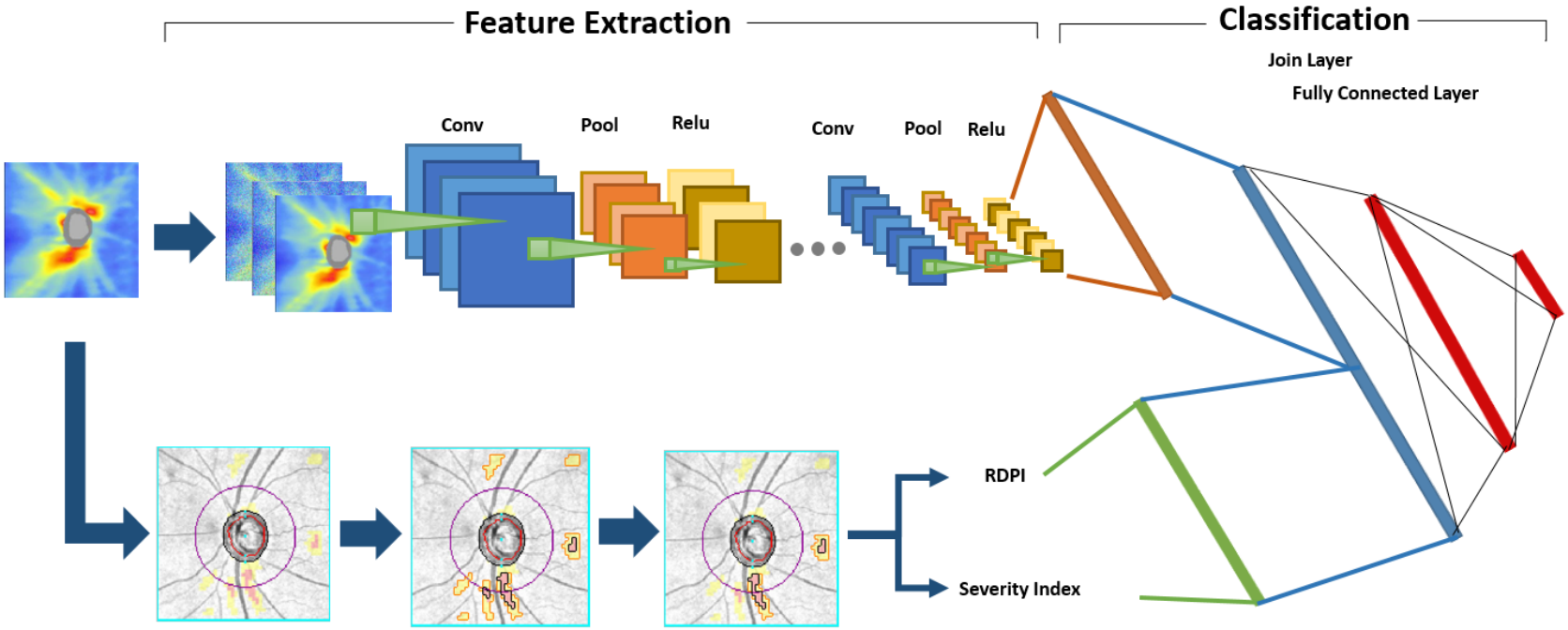

Figure 4.1. Description of Deep Feature Fusion Network

4.2.4. Statistical analysis

Statistical analysis was performed to calculate AUROC and accuracy for each method. Each analysis method was evaluated to how glaucoma group, which combines early glaucoma and mild or severe glaucoma group, and normal group are precisely distinguished as the performance of the diagnosis of glaucoma. Also, each analysis method was evaluated to how early glaucoma group and normal group are correctly classified as the performance of the early diagnosis of glaucoma. Analysis Methods included machine learning analysis with features extracted from previous studies and machine learning analysis with features extracted from Chapter 2, Chapter 3 and deep learning analysis using thickness map and DFFN fusing features extracted by deep learning analysis and from Chapter 2, Chapter 3. Deep learning analysis was evaluated about LeNet, VGGNet respectively.

Accuracy was calculated by whether the groups last selected by machine learning and deep learning were matched with ground truth groups as following equation. AUROC was calculated as the area of the receiver operating characteristic

(ROC) curve which was plotted with specificity and sensitivity calculated by following equation using the results from the analysis methods. The results analyzed by machine learning and deep learning were finally calculated by using softmax function as probability of each group and selected as the group with the highest probability. So the results analyzed to AUROC were used as the probability selecting normal group.

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ number\ of\ subject}$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive}$$

4.3. RESULTS

This study included 788 eyes of 788 subjects (472 eyes of 472 subjects with glaucoma and 316 eyes of 316 normal subjects). Among 472 eyes of glaucoma patients, 226 eyes showed glaucoma of early stage and 246 eyes showed glaucoma of mild or severe stage. The demographic and ocular characteristics of the subjects are presented in Table 4.1. The t-test was performed by combining the two groups of the three groups to see if there were significant differences between the two groups.

There were significant differences in mean age and central corneal thickness (CCT) between normal subjects and glaucoma subjects of early stage ($p = 0.007$, $p = 0.009$), but no significant differences were found between other groups. Gender and spherical equivalent did not significantly differ any groups. Intraocular pressure (IOP) and axial length showed significant differences between glaucoma subjects of early stage and mild or severe stage ($p = 0.031$, $p = 0.005$) and not between other groups. Visual field (VF) mean deviation (MD) and VF pattern

standard deviation (PSD) had significant differences among all three groups ($p < 0.001$).

When LeNet was trained from randomly distributed weights, learning rate and momentum were changed and selected parameter ranges trained well about loss and accuracy according to epoch. The loss according to epoch about each parameters was plotted in Figure 4.2. In Figure 4.1A, the learning rate is 0.1 where trained data are over-fitted because the loss converges too quickly and only local features are extracted. In Figure 4.2D, the learning rate is 0.0001 where the loss decreases slowly and was not over-fitted, but converges at a large loss, resulting in poor accuracy. In Figure 4.2B, 4.2C, learning rates were 0.01 and 0.001 where the loss decreases appropriately and converges to nearly 0. So, LeNet was trained to learning rate between 0.01 and 0.001 and was selected by learning rate with the best accuracy. For the momentum value, the degree of change of the loss value was not large according to momentum and the momentum with the best accuracy was found.

Table 4.1. Demographic Data of Study Subjects According to the Stages of Glaucoma

	Normal [A] (n = 316)	Early [B] (n = 226)	Mild-or-Severe [C] (n = 246)	*P'* [A-B]	*P''* [A-C]	*P'''* [B-C]
Age (yrs)	62.5 ± 14.9	59.1 ± 13.7	61.9 ± 15.5	0.007†	0.642†	0.039†
Gender (male / female)	158 / 158	121 / 105	130 / 116	0.416‡	0.503‡	0.880‡
Spherical equivalent (D)	-0.81 ± 2.51	-0.89 ± 2.53	-0.73 ± 2.85	0.716†	0.996†	0.993†
Intraocular pressure (mmHg)	14.6 ± 3.5	14.2 ± 3.3	14.9 ± 3.7	0.180†	0.326†	0.031†
CCT (mm)	542.5 ± 31.1	535.2 ± 32.8	537.2 ± 34.3	0.009†	0.056†	0.519†
Axial length (mm)	24.5 ± 1.71	24.3 ± 1.69	24.8 ± 2.11	0.178†	0.063†	0.005†
VF MD (decibels)	-0.2 ± 0.8	-6.2 ± 3.9	-15.7 ± 3.7	<0.001†	<0.001†	<0.001†
VF PSD (decibels)	0.51 ± 0.21	6.81 ± 2.97	9.09 ± 5.17	<0.001†	<0.001†	<0.001†

Values are mean ± standard deviation.
D, diopters; CCT, central corneal thickness; VF, visual field;
MD, mean deviation; PSD, pattern standard deviation.
† student t test., ‡ chi-square test

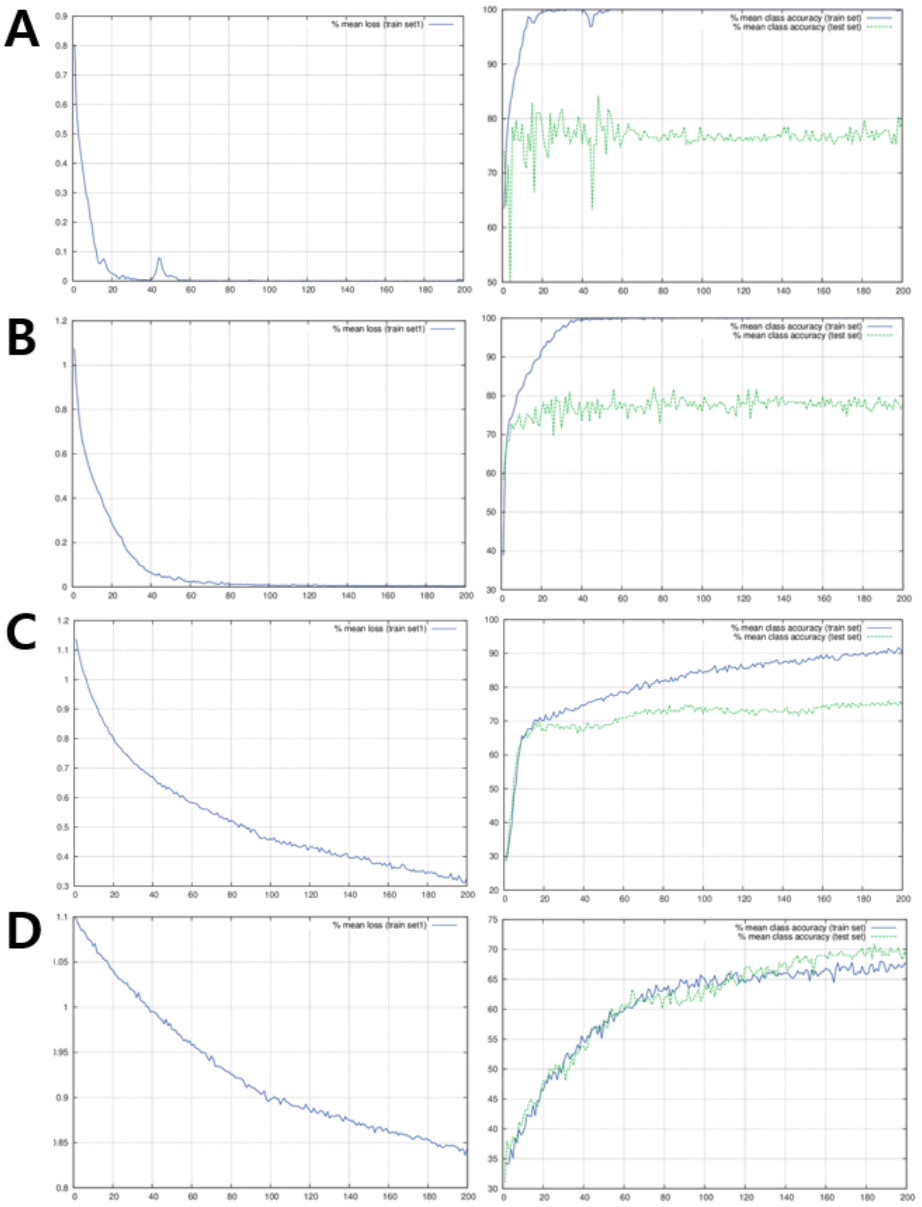

Figure 4.2. Graph of training loss and accuracy according to epoch at various learning rates. (A) learning rate = 0.1, (B) learning rate = 0.01, (C) learning rate = 0.001, (D) learning rate = 0.0001

The accuracy of training and testing according to epoch while the learning rate was varied from 0.01 to 0.001, momentum was changed 0.3, 0.5 and 0.7 was analyzed. When LeNet is trained, the accuracy gradually increased and reached almost 100%, but the tested accuracy converged to a constant value from the moment when it went up with the first time. The ROC curve using the test results when convergence was achieved. The AUROC and accuracy for glaucoma subjects and normal subjects were calculated as shown in Table 4.2 by plotting ROC curves for learning rate, momentum and trained data respectively. Also, the AUROC and accuracy for glaucoma subjects of early stage and normal subjects were calculated as shown in Table 4.3 by plotting ROC curves respectively. The accuracy classifying glaucoma subjects and normal subjects when trained data was original data and learning rate was 0.001 and momentum was 0.9 was highest to 86.68% and AUROC was highest to 0.9387 when trained data was original data and learning rate was 0.001 and momentum was 0.5. And the accuracy and AUROC dividing glaucoma subjects of early stage and normal subjects were highest to 81.18%, 0.8811 when trained data was original data and learning rate was 0.001 and momentum was 0.9. Totally,

when LeNet is trained by original data and at 0.001 of learning rate and at 0.9 of momentum had the best performance.

In transfer learning with VGGNet, the learning rate was changed from 0.001 to 0.0001 because VGGNet was trained at 0.01 of learning rate to ImageNet and set to the most appropriate learning rate and evaluated according to trained data and changed layers. When trained with VGGNet like LeNet, the accuracy gradually increased to nearly 100%, but the tested accuracy converged to a constant value from the moment when it went up with the first time. The accuracy and AUROC distinguishing between glaucoma subjects and normal subjects were calculated according to trained data, learning rate, changed layer in Table 4.5. Also, the accuracy and AUROC for glaucoma subjects of early stage and normal subjects were calculated for each in Table 4.5. The accuracy and AUROC splitting glaucoma subjects and normal subjects was highest to 87.69%, 0.9353 when trained data was original data and changed layers were convolutional layers and fully connected layers and learning rate was 0.0001. And the accuracy and AUROC classifying glaucoma subjects of early stage and normal subjects were highest to 83.39%, 0.8880 when trained data was original data and changed layers were

convolutional layers and fully connected layers and learning rate is 0.0001. Overall performance was best when VGGNet was trained by original data and at 0.0001 of learning rate and convolutional layers and fully connected layers were changed.

Table 4.2. Accuracy and area under receiver operating characteristic curve (AUROC) of LeNet trained on various hyper-parameter distinguishing between normal subjects and glaucoma subjects.

Hyper-parameter		Accuracy (%)		AUROC	
		Original Data	Augmented Data	Original Data	Augmented Data
Learning rate	Momentum				
0.001	0.5	85.79	84.77	0.9387	0.9133
0.001	0.7	85.15	85.91	0.9297	0.9292
0.001	0.9	86.68	83.88	0.9377	0.9076
0.003	0.5	83.50	86.42	0.9093	0.9206
0.003	0.7	85.41	85.28	0.9178	0.9161
0.003	0.9	85.66	82.87	0.9237	0.9009
0.005	0.5	85.15	85.53	0.9166	0.9139
0.005	0.7	84.77	84.26	0.9208	0.9010
0.005	0.9	84.52	84.01	0.9100	0.9157
0.007	0.5	84.39	84.39	0.9134	0.9059
0.007	0.7	83.50	83.63	0.9102	0.9034
0.007	0.9	81.47	83.88	0.8922	0.8993
0.01	0.5	83.76	84.90	0.9035	0.9044
0.01	0.7	84.26	83.88	0.9050	0.9053
0.01	0.9	86.04	85.28	0.9264	0.8976

AUROC = area under receiver operating characteristic curve

Table 4.3. Accuracy and area under receiver operating characteristic curve (AUROC) of LeNet trained on various hyper-parameter distinguishing between normal subjects and glaucoma subjects of early stage.

Hyper-parameter		Accuracy (%)		AUROC	
		Original Data	Augmented Data	Original Data	Augmented Data
Learning rate	Momentum				
0.001	0.5	79.70	77.68	0.8806	0.8528
0.001	0.7	78.60	78.97	0.8632	0.8712
0.001	0.9	81.18	77.31	0.8811	0.8350
0.003	0.5	76.20	79.34	0.8362	0.8533
0.003	0.7	78.78	79.34	0.8505	0.8587
0.003	0.9	79.52	75.65	0.8586	0.8294
0.005	0.5	78.23	80.26	0.8521	0.8514
0.005	0.7	77.49	77.68	0.8587	0.8395
0.005	0.9	77.31	76.38	0.8427	0.8492
0.007	0.5	77.12	77.31	0.8426	0.8367
0.007	0.7	77.12	76.01	0.8373	0.8339
0.007	0.9	74.35	77.49	0.8110	0.8374
0.01	0.5	76.20	77.31	0.8321	0.8373
0.01	0.7	77.31	77.49	0.8429	0.8359
0.01	0.9	80.07	77.12	0.8710	0.8347

AUROC = area under receiver operating characteristic curve

Table 4.4. Accuracy and area under receiver operating characteristic curve (AUROC) of VGGNet trained on various hyper-parameter distinguishing between normal subjects and glaucoma subjects.

	Accuracy (%)		AUROC	
Hyper-parameter	Original Data	Augmented Data	Original Data	Augmented Data
Convolutional layers & Fully connected layer				
Learning rate - 0.0001	87.69	86.80	0.9353	0.9312
Learning rate - 0.0003	86.55	86.80	0.9294	0.9292
Learning rate - 0.0005	86.29	86.68	0.9290	0.9280
Learning rate - 0.0007	86.68	86.80	0.9282	0.9284
Learning rate - 0.001	86.80	87.56	0.9293	0.9297
Fully Connected layers				
Learning rate - 0.0001	85.91	86.17	0.9306	0.9273
Learning rate - 0.0003	86.55	84.90	0.9294	0.9205
Learning rate - 0.0005	86.17	85.03	0.9242	0.9180
Learning rate - 0.0007	84.77	84.52	0.9201	0.9184
Learning rate - 0.001	84.77	84.39	0.9192	0.9183

AUROC = area under receiver operating characteristic curve

Table 4.5. Accuracy and area under receiver operating characteristic curve (AUROC) of VGGNet trained on various hyper-parameter distinguishing between normal subjects and glaucoma subjects of early stage.

	Accuracy (%)		AUROC	
Hyper-parameter	Original Data	Augmented Data	Original Data	Augmented Data
Convolutional layers & Fully connected layer				
Learning rate - 0.0001	83.39	82.29	0.8880	0.8829
Learning rate - 0.0003	81.92	81.92	0.8794	0.8804
Learning rate - 0.0005	81.18	82.29	0.8808	0.8791
Learning rate - 0.0007	82.29	82.29	0.8787	0.8781
Learning rate - 0.001	82.29	83.21	0.8809	0.8823
Fully Connected layers				
Learning rate - 0.0001	81.37	80.81	0.8770	0.8726
Learning rate - 0.0003	81.55	78.78	0.8746	0.8593
Learning rate - 0.0005	81.00	79.15	0.8676	0.8555
Learning rate - 0.0007	78.60	78.23	0.8594	0.8560
Learning rate - 0.001	78.60	78.86	0.8564	0.8550

AUROC = area under receiver operating characteristic curve

In DFFN, features included features extracted by test data from first and second fully connected layer when LeNet and VGGNet were trained to most well performance respectively. The accuracy and AUROC were calculated for DFFN that fused features extracted from Chapter 2, Chapter 3 and features extracted by deep learning in Table 4.6. Table 4.6 also shows results that were trained by features in previous studies. In Figure 4.3, ROC curves about all analysis algorithm were plotted. The accuracy and AUROC were 82.49%, 0.9053 about features extracted in previous studies and 85.79%, 0.9240 about features extracted in Chapter 2, Chapter 3 distinguishing between normal subjects and glaucoma subjects. And the accuracy and AUROC were 75.46%, 0.8241 about features extracted in previous studies and 80.26%, 0.8682 about features extracted in Chapter 2, Chapter 3 distinguishing between normal subjects and glaucoma subjects of early stage.

In DFFN, the accuracy and AUROC were 87.56%, 0.9418 when features included features extracted by the first fully connected layer of LeNet and 85.41%, 0.9320 when features included features extracted by the second fully connected layer of LeNet discriminating between glaucoma subjects and normal

subjects. In addition, the accuracy and AUROC were 82.66%, 0.8925 when features included features extracted by the first fully connected layer of LeNet and 79.15%, 0.8708 when features included features extracted by the second fully connected layer of LeNet discriminating between glaucoma subjects and normal subjects. And the accuracy and AUROC were 86.68%, 0.9435 when features included features extracted by the first fully connected layer of VGGNet and 88.45%, 0.9555 when features included features extracted by the second fully connected layer of VGGNet discriminating between glaucoma subjects and normal subjects. In addition, the accuracy and AUROC were 83.39%, 0.9117 when features included features extracted by the first fully connected layer of VGGNet and 84.50%, 0.9204 when features included features extracted by the second fully connected layer of VGGNet discriminating between glaucoma subjects and normal subjects.

Table 4.6. Accuracy and area under receiver operating characteristic curve (AUROC) of various machine learning algorithm distinguishing between normal subjects and glaucoma subjects and discriminating between normal subjects and glaucoma subjects of early stage.

	Normal vs Glaucoma		Normal vs Early Glaucoma	
	Accuracy (%)	AUROC	Accuracy (%)	AUROC
Previous research	82.49	0.9053	75.46	0.8241
Chapter 2,3	85.79	0.9240	80.26	0.8682
Deep learning (LeNet)	86.68	0.9377	81.18	0.8811
Deep learning (VGGNet)	87.69	0.9353	83.39	0.8880
DFFN (fc1 of LeNet)	87.56	0.9418	82.66	0.8925
DFFN (fc2 of LeNet)	85.41	0.9320	79.15	0.8708
DFFN (fc1 of VGGNet)	86.68	0.9435	83.39	0.9117
DFFN (fc2 of VGGNet)	88.45	0.9555	84.50	0.9204

AUROC = area under receiver operating characteristic curve, DFFN = deep feature fusion network, fc = fully connected layer

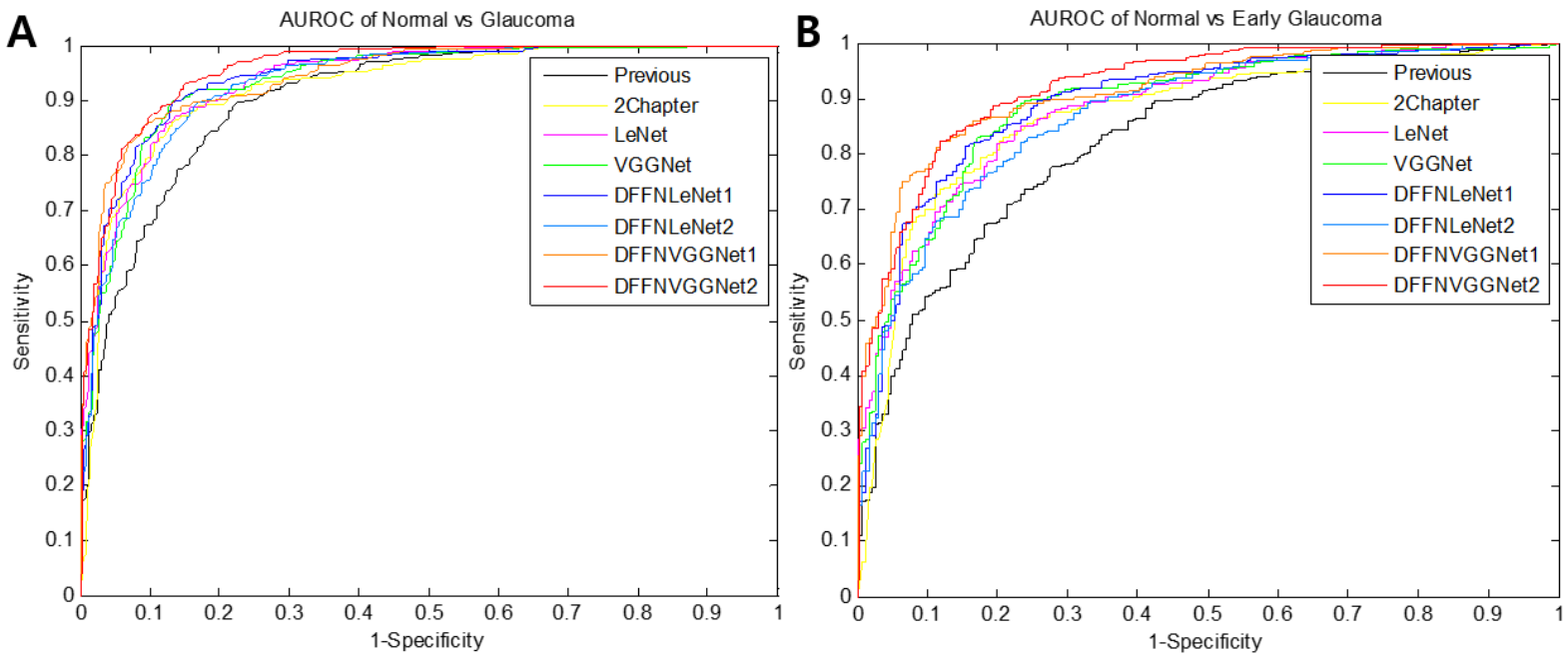

Figure 4.3. Receiver operating characteristic curve (AUROC) of various machine learning algorithm (A) distinguishing between normal subjects and glaucoma subjects and (B) discriminating between normal subjects and glaucoma subjects of early stage.

4.4. DISCUSSION

In this study, the computer was directly learned and analyzed by deep learning and DFFN was developed by fusing and analyzing the features extracted from deep learning and other chapters for accurate glaucoma diagnosis. The machine learning algorithms which were trained by features in previous researches and by DFFN and deep learning were evaluated by the accuracy and AUROC. DFFN didn't only classify into glaucoma subjects and normal subjects most accurately than any other machine learning algorithm, but also classified into glaucoma subjects of early stage and normal subjects most accurately than any other machine learning algorithm.

It is important to set hyper-parameters before training, when LeNet was trained by randomly distributed weights. So, learning rate and momentum were used variously and set to best trained range through the loss. When the learning rate is too high, the loss converged to nearly zero quickly but features were selected locally and tested data were over-fitted. On the contrary, when the learning rate is too low, the loss converged slowly but bigger than zero and features were not trained

accurately. However, when the learning rate is appropriate, the loss value converged to nearly zero slowly which indicates that the learning rate is from 0.01 to 0.001. And there was no significant differences to the loss about momentum because the direction of the loss was similar to the overall trained direction.

The three group were trained close to 100% when was trained by LeNet, but tested data could not be divided accurately. Because features trained by LeNet were found to global features for the first time and to local features at end. In other words, local features among the 400 features extracted by LeNet affected the over-fitted results. The accuracy for normal subjects and glaucoma subjects of early stage was more inaccurate because features extracted by LeNet were selected to classify between normal subjects and glaucoma subjects of mild or severe stage. In order to enhance accuracy, LeNet should be trained more accurately about between normal group and glaucoma subjects of early stage. The accuracy according to the momentum was not changed as significantly as according to the learning rate because direction of the loss was similar to the overall trained direction. But the momentum is slightly valid for features trained and shows the differences in accuracy. When

trained according to the trained data, there were slight differences in the accuracy and AUROC trained by various hyper-parameters.

The loss and accuracy according to epoch when network was learned by trained VGGNet was similar to when network was learned by LeNet from randomly distributed weights. But the accuracy and AUROC of network trained by VGGNet was better than that of network trained by LeNet as a whole, which shows that it is better to train with transfer learning. However, the accuracy and AUROC of two learning methods were not significantly different, which shows that features trained by transfer learning were found for global features at the beginning and some of features found for local features at the end like when trained by LeNet. Among the 4608 features extracted by VGGNet, specific features existing in the only trained data affected the accuracy of tested data. The accuracy for normal subjects and glaucoma subjects of early stage was more inaccurate because features extracted by VGGNet were selected to classify between normal subjects and glaucoma subjects of mild or severe stage like LeNet. When the network was trained by VGGNet according to trained data, the accuracy and AUROC

didn't have much differences between original data and augmented data. That is because the weights of VGGNet were already affected by Gaussian noise when trained by ImageNet. And glaucoma subjects and normal subjects were classified similarly according to changed layers, but glaucoma subjects of early stage and normal subjects were more divided at network changed by the convolutional layers and fully connected layers. These results shows that the features classifying between glaucoma subjects of early stage and normal subjects was slightly different from the features trained by ImageNet in the convolutional layers and pooling layers.

The accuracy and AUROC of network trained by DFFN were highest than those of network trained by deep learning and features extracted by other chapters, which is because DFFN had more different features analyzing glaucoma subjects and normal subjects(67). And that DFFN had more complicated features extracted by using the information between thickness map pixels than that extracted simply in the previous researches is the reason why the accuracy and AUROC of DFFN were better than that of the previous researches(65). But features extracted by VGGNet classified more accurately than that extracted by

LeNet, which is because the features extracted by LeNet had more specific features to over-fit the thickness map data.

There was difference in accuracy about trained data and tested data because features were extracted effectively for only train data by deep learning. Since the train data was not many enough to be trained accurately, the thickness map data should increase in order to enhance glaucoma diagnosis accuracy. In addition, when data are collected at hospitals and other facilities, data are more likely to be unbalanced because of the greater number of glaucoma patients of mild of severe stage rather than normal or glaucoma patients of early stage. Therefore, this problem can be solved by the augmentation adding the Gaussian noise and flipping or rotation and if the amount of data increased evenly, the accuracy can be improved(77). Also, in order to enhance the accuracy, pre-trained data needs to be more similar to the thickness map than ImageNet when transfer learning method is used and the thickness map data need to be trained with deeper neural network like GoogLeNet or ResNet(78, 79). Since the glaucoma diagnosis is confirmed by several tests, it is possible to diagnose more accurate glaucoma if DFFN can fuse

features extracted by the thickness map data as well as the visual field test data, IOP, demographic data.

Glaucoma can be diagnosed accurately through DFFN fusing the features extracted by deep learning in this study. We developed DFFN that can be analyzed by using various test results and demographic data and made a stepping stone for future research. This will allow us to accurately diagnose glaucoma as well as early glaucoma using OCT.

CHAPTER 5

Thesis Summary and Future Work

5.1. THESIS SUMMARY AND CONTRIBUTION

We developed the deep feature fusion network (DFFN) using the retinal nerve fiber layer (RNFL) optical coherence tomography (OCT) thickness map for accurate glaucoma diagnosis. Since DFFN is network fusing features trained by deep learning and hand-crafted features extracted manually, studies extracting features which affects glaucoma diagnosis are conducted. We extracted features using the area and depth of the defects on the thickness map and deviation map. And we developed deep learning algorithm which allowed only computer to directly extract features related to glaucoma. We distinguished between normal subjects and glaucoma subjects of early stage as well as between normal subjects and glaucoma subjects using DFFN fusing features extracted by method in Chapter 2, 3 and deep learning and evaluated DFFN using the accuracy and the area under receiver operating characteristic curve (AUROC).

Chapter 2 calculated localized RNFL defects on the ganglion cell inner plexiform layer (GCIPL) and RNFL deviation maps and extracted features as severity indices according to probability levels. All of the severity indices calculated by image

processing had significantly difference between glaucomatous eyes and normal eyes. We found that the AUROCs for detection of localized RNFL defects were larger on the pRNFL deviation map than on the GCIPL deviation map, but the differences were not statistically significant. In the detection of glaucomatous eyes with localized RNFL defects, indices calculated in the macular GCIPL thickness map deviation map showed a level of diagnostic performance comparable to that of indices calculated in the RNFL thickness deviation map.

Chapter 3 segmented RNFL defects on the deviation map according to inclusion criteria and extracted feature as RNFL defect depth percentage index (RDPI) on the RNFL thickness deviation map using the proportion of the RNFL defect depth. The RDPI, a new parameter using RNFL thickness deviation map, can be a useful adjunct tool for objective quantification of RNFL defect depth. This parameter has an advantage over RNFL thickness in discriminating between mild and moderate RNFL defects, not in discriminating between moderate and severe defects.

Chapter 4 extracted features using deep learning which allows computer to distinguish between normal subjects and glaucoma patients on RNFL thickness maps. RNFL Thickness map was trained in two ways for deep learning: one is to be trained with LeNet from randomly distributed weights and the other is to be transfer learned with VGGNet from weights pre-trained by ImageNet. When training RNFL thickness maps with deep learning, we conducted data augmentation for deep neural network to be trained accurately and selected the most appropriate hyper-parameters through the loss and accuracy according to epoch. We developed DFFN fusing features extracted by Chapter 2, 3 and deep learning and evaluated performances with previous studies. DFFN fusing features including features extracted by VGGNet discriminated between glaucoma subjects of early stage and normal subjects as well as between glaucoma subjects and normal subjects most accurately than other previous studies.

5.2. FUTURE WORK

DFFN developed in this study can classify more accurately if meaningful features are more accurately extracted manually and by deep learning. Anatomical features like area, angle, and correlation with GCIPL thickness map extracted meaningfully will be analyzed more accurately on the RNFL thickness map. Hand crafted features extracted on the red-free fundus images and visual field tests by using other image processing will be extracted more correctly. And it can be important to use demographic data as meaningful features affecting glaucoma.

The amount of RNFL thickness map data to be trained must be large enough to be trained with deep neural network more accurately. It is necessary for doctors to get data from the hospital or other facilities and classify data into glaucoma. Also, it is important to augment image data such as various noises addition, image flipping, image rotation and hue, saturation, and lightness (HSL) color space transformation. If amount data are big enough to be trained with deep learning, deeper network like ResNet or GoogLeNet can be trained more correctly. And if deep

neural network is trained with image data similar to RNFL thickness map, better deep neural architecture will be build.

DFFN which is cascaded or trained parallel with deep learning and machine learning can be greatly improved. DFFN can be studied in distinguishing between glaucoma patients of early stage and normal subjects and analyzing progression of glaucoma. And if glaucoma is classified by using DFFN accurately, the study analyzing factor of glaucoma on RNFL thickness map can be developed by using features of DFFN.

BIBLIOGRAPHY

1. Weinreb RN, Khaw PT. Primary open-angle glaucoma. The Lancet. 2004;363(9422):1711-20.
2. Foster PJ, Johnson GJ. Glaucoma in China: how big is the problem? British Journal of Ophthalmology. 2001;85(11):1277-82.
3. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. British journal of ophthalmology. 2006;90(3):262-7.
4. Sakai H, Morine-Shinjyo S, Shinzato M, Nakamura Y, Sakai M, Sawaguchi S. Uveal effusion in primary angle-closure glaucoma. Ophthalmology. 2005;112(3):413-9.
5. Hollands H, Johnson D, Hollands S, Simel DL, Jinapriya D, Sharma S. Do Findings on Routine Examination Identify Patients at Risk for Primary Open-Angle Glaucoma?: The Rational Clinical Examination Systematic Review. Jama. 2013;309(19):2035-42.
6. Urban RC, Dreyer EB. Corticosteroid-induced glaucoma. International ophthalmology clinics. 1993;33(2):135-40.
7. Dielemans I, Vingerling JR, Algra D, Hofman A, Grobbee DE, de Jong PT. Primary open-angle glaucoma, intraocular pressure, and systemic blood pressure in the general elderly population: the Rotterdam Study. Ophthalmology. 1995;102(1):54-60.
8. Hayashi K, Hayashi H, Nakao F, Hayashi F. Changes in anterior chamber angle width and depth after intraocular lens implantation in eyes with glaucoma. Ophthalmology. 2000;107(4):698-703.
9. Kwon YH, Kim C-s, Zimmerman MB, Alward WL, Hayreh SS. Rate of visual field loss and long-term visual outcome in primary open-angle glaucoma. American journal of ophthalmology. 2001;132(1):47-56.
10. Quigley HA, Katz J, Derick RJ, Gilbert D, Sommer A. An evaluation of optic disc and nerve fiber layer examinations in monitoring progression of early glaucoma damage. Ophthalmology. 1992;99(1):19-28.
11. Quigley HA, Addicks EM, Green WR, Maumenee A. Optic nerve damage in human glaucoma. Archives of Ophthalmology. 1981;99(4):635-49.

12. Medeiros FA, Vizzeri G, Zangwill LM, Alencar LM, Sample PA, Weinreb RN. Comparison of retinal nerve fiber layer and optic disc imaging for diagnosing glaucoma in patients suspected of having the disease. Ophthalmology. 2008;115(8):1340-6.
13. Medeiros FA, Zangwill LM, Bowd C, Weinreb RN. Comparison of the GDx VCC Scanning Laser Polarimeter, HRT II ConfocalScanning Laser Ophthalmoscope, and Stratus OCT Optical Coherence Tomographfor the Detection of Glaucoma. Archives of Ophthalmology. 2004;122(6):827-37.
14. Chauhan BC, O'Leary N, AlMobarak FA, Reis AS, Yang H, Sharpe GP, et al. Enhanced detection of open-angle glaucoma with an anatomically accurate optical coherence tomography-derived neuroretinal rim parameter. Ophthalmology. 2013;120(3):535-43.
15. Strouthidis NG, Gardiner SK, Sinapis C, Burgoyne CF, Garway-Heath DF. The spatial pattern of neuroretinal rim loss in ocular hypertension. Investigative ophthalmology & visual science. 2009;50(8):3737-42.
16. Drexler W, Morgner U, Ghanta RK, Kärtner FX, Schuman JS, Fujimoto JG. Ultrahigh-resolution ophthalmic optical coherence tomography. Nature medicine. 2001;7(4):502-7.
17. Drexler W, Sattmann H, Hermann B, Ko TH, Stur M, Unterhuber A, et al. Enhanced visualization of macular pathology with the use of ultrahigh-resolution optical coherence tomography. Archives of ophthalmology. 2003;121(5):695-706.
18. Mohamed S, Lee GK, Rao SK, Wong AL, Cheng AC, Li EY, et al. Repeatability and reproducibility of pachymetric mapping with Visante anterior segment-Optical coherence tomography. Investigative ophthalmology & visual science. 2007;48(12):5499-504.
19. Schuman JS. Spectral domain optical coherence tomography for glaucoma (an AOS thesis). Transactions of the American Ophthalmological Society. 2008;106:426.
20. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, et al. Optical coherence tomography. Science (New York, NY). 1991;254(5035):1178.
21. Wojtkowski M, Srinivasan VJ, Ko TH, Fujimoto JG, Kowalczyk A, Duker JS. Ultrahigh-resolution, high-speed, Fourier domain optical coherence tomography and methods for dispersion compensation. Optics express. 2004;12(11):2404-22.

22. Lim H, Mujat M, Kerbage C, Lee E, Chen Y, Chen TC, et al. High-speed imaging of human retina in vivo with swept-source optical coherence tomography. Optics express. 2006;14(26):12902-8.
23. Wojtkowski M, Srinivasan V, Fujimoto JG, Ko T, Schuman JS, Kowalczyk A, et al. Three-dimensional retinal imaging with high-speed ultrahigh-resolution optical coherence tomography. Ophthalmology. 2005;112(10):1734-46.
24. Schuman JS, Hee MR, Arya AV, Pedut-Kloizman T, Puliafito CA, Fujimoto JG, et al. Optical coherence tomography: a new tool for glaucoma diagnosis. Current opinion in ophthalmology. 1995;6(2):89-95.
25. Takayama K, Hangai M, Durbin M, Nakano N, Morooka S, Akagi T, et al. A Novel Method to Detect Local Ganglion Cell Loss in Early Glaucoma Using Spectral-Domain Optical Coherence TomographyDetection of Local RGC Loss in Glaucoma. Investigative ophthalmology & visual science. 2012;53(11):6904-13.
26. Hwang YH, Kim YY, Kim HK, Sohn YH. Ability of cirrus high-definition spectral-domain optical coherence tomography clock-hour, deviation, and thickness maps in detecting photographic retinal nerve fiber layer abnormalities. Ophthalmology. 2013;120(7):1380-7.
27. Jeoung JW, Park KH. Comparison of Cirrus OCT and Stratus OCT on the ability to detect localized retinal nerve fiber layer defects in preperimetric glaucoma. Investigative ophthalmology & visual science. 2010;51(2):938-45.
28. Kim NR, Lee ES, Seong GJ, Choi EH, Hong S, Kim CY. Spectral-domain optical coherence tomography for detection of localized retinal nerve fiber layer defects in patients with open-angle glaucoma. Archives of ophthalmology. 2010;128(9):1121-8.
29. Leung CK, Lam S, Weinreb RN, Liu S, Ye C, Liu L, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: analysis of the retinal nerve fiber layer map for glaucoma detection. Ophthalmology. 2010;117(9):1684-91.
30. Mwanza J-C, Durbin MK, Budenz DL, Girkin CA, Leung CK, Liebmann JM, et al. Profile and predictors of normal ganglion cell–inner plexiform layer thickness measured with frequency-domain optical coherence tomography. Investigative ophthalmology & visual science. 2011;52(11):7872-9.

31. Mwanza J-C, Oakley JD, Budenz DL, Chang RT, O'Rese JK, Feuer WJ. Macular ganglion cell–inner plexiform layer: automated detection and thickness reproducibility with spectral domain–optical coherence tomography in glaucoma. Investigative ophthalmology & visual science. 2011;52(11):8323-9.
32. Kotowski J, Folio LS, Wollstein G, Ishikawa H, Ling Y, Bilonick RA, et al. Glaucoma discrimination of segmented cirrus spectral domain optical coherence tomography (SD-OCT) macular scans. British Journal of Ophthalmology. 2012;96(11):1420-5.
33. Mwanza J-C, Durbin MK, Budenz DL, Sayyad FE, Chang RT, Neelakantan A, et al. Glaucoma diagnostic accuracy of ganglion cell–inner plexiform layer thickness: comparison with nerve fiber layer and optic nerve head. Ophthalmology. 2012;119(6):1151-8.
34. Airaksinen PJ, Nieminen H. Retinal nerve fiber layer photography in glaucoma. Ophthalmology. 1985;92(7):877-9.
35. Hoyt WF, Frisen L, Newman NM. Fundoscopy of nerve fiber layer defects in glaucoma. Investigative ophthalmology & visual science. 1973;12(11):814-29.
36. Savini G, Carbonelli M, Barboni P. Spectral-domain optical coherence tomography for the diagnosis and follow-up of glaucoma. Current opinion in ophthalmology. 2011;22(2):115-23.
37. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988:837-45.
38. Sommer A, Miller NR, Pollack I, Maumenee AE, George T. The nerve fiber layer in the diagnosis of glaucoma. Archives of ophthalmology. 1977;95(12):2149-56.
39. Tuulonen A, Lehtola J, Airaksinen PJ. Nerve fiber layer defects with normal visual fields: do normal optic disc and normal visual field indicate absence of glaucomatous abnormality? Ophthalmology. 1993;100(5):587-98.
40. Leung CK-s, Cheung CY-l, Weinreb RN, Qiu Q, Liu S, Li H, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study. Ophthalmology. 2009;116(7):1257-63. e2.
41. Zhao L, Wang Y, Chen CX, Xu L, Jonas JB. Retinal nerve fibre layer thickness measured by Spectralis spectral-domain optical coherence tomography: The Beijing Eye Study. Acta

ophthalmologica. 2014;92(1):e35-e41.
42. Sung M-S, Yoon J-H, Park S-W. Diagnostic validity of macular ganglion cell-inner plexiform layer thickness deviation map algorithm using cirrus HD-OCT in preperimetric and early glaucoma. Journal of glaucoma. 2014;23(8):e144-e51.
43. Curcio CA, Allen KA. Topography of ganglion cells in human retina. Journal of comparative Neurology. 1990;300(1):5-25.
44. Leite MT, Zangwill LM, Weinreb RN, Rao HL, Alencar LM, Sample PA, et al. Effect of disease severity on the performance of Cirrus spectral-domain OCT for glaucoma diagnosis. Investigative ophthalmology & visual science. 2010;51(8):4104-9.
45. Quigley HA, Reacher M, Katz J, Strahlman E, Gilbert D, Scott R. Quantitative grading of nerve fiber layer photographs. Ophthalmology. 1993;100(12):1800-7.
46. Tuulonen A, Airaksinen PJ. Initial glaucomatous optic disk and retinal nerve fiber layer abnormalities and their progression. American journal of ophthalmology. 1991;111(4):485-90.
47. Suh M, Kim D, Kim Y, Kim T, Park K. Patterns of progression of localized retinal nerve fibre layer defect on red-free fundus photographs in normal-tension glaucoma. Eye. 2010;24(5):857-63.
48. Jeoung JW, Kim SH, Park KH, Kim T-W, Kim DM. Quantitative assessment of diffuse retinal nerve fiber layer atrophy using optical coherence tomography: diffuse atrophy imaging study. Ophthalmology. 2010;117(10):1946-52.
49. Jeoung JW, Kim SH, Park KH, Kim T-W, Kim DM. Diagnostic accuracy of OCT with a normative database to detect diffuse retinal nerve fiber layer atrophy: diffuse atrophy imaging study. Investigative ophthalmology & visual science. 2011;52(9):6074-80.
50. Leung CK-S, Yu M, Weinreb RN, Lai G, Xu G, Lam DS-C. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: patterns of retinal nerve fiber layer progression. Ophthalmology. 2012;119(9):1858-66.
51. Leung CK, Choi N, Weinreb RN, Liu S, Ye C, Liu L, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: pattern of RNFL defects in glaucoma. Ophthalmology. 2010;117(12):2337-44.
52. Ye C, To E, Weinreb RN, Yu M, Liu S, Lam DS, et al.

Comparison of retinal nerve fiber layer imaging by spectral domain optical coherence tomography and scanning laser ophthalmoscopy. Ophthalmology. 2011;118(11):2196-202.
53. Roh KH, Jeoung JW, Park KH, Yoo BW, Kim DM. Long-term reproducibility of Cirrus HD optical coherence tomography deviation map in clinically stable glaucomatous eyes. Ophthalmology. 2013;120(5):969-77.
54. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide: Cambridge University Press; 2011.
55. Nassif N, Cense B, Park BH, Yun SH, Chen TC, Bouma BE, et al. In vivo human retinal imaging by ultrahigh-speed spectral domain optical coherence tomography. Optics letters. 2004;29(5):480-2.
56. Choma MA, Sarunic MV, Yang C, Izatt JA. Sensitivity advantage of swept source and Fourier domain optical coherence tomography. Optics express. 2003;11(18):2183-9.
57. Kang SH, Park KH, Kim JM, Seo JM, Kim DM. Korean normative database for time domain optical coherence tomography to detect localized retinal nerve fiber layer defects (preliminary study). Japanese journal of ophthalmology. 2010;54(2):144.
58. Quigley HA, Addicks EM. Quantitative studies of retinal nerve fiber layer defects. Archives of Ophthalmology. 1982;100(5):807-14.
59. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. Progress in retinal and eye research. 2007;26(6):688-710.
60. Hood DC, Anderson SC, Wall M, Kardon RH. Structure versus function in glaucoma: an application of a linear model. Investigative ophthalmology & visual science. 2007;48(8):3662-8.
61. Hood DC, Anderson S, Rouleau J, Wenick AS, Grover LK, Behrens MM, et al. Retinal nerve fiber structure versus visual field function in patients with ischemic optic neuropathy: a test of a linear model. Ophthalmology. 2008;115(5):904-10.
62. Niessen AG, LANGERHORST CT, BOSSUYT PM. Grading of retinal nerve fiber layer with a photographic reference set. American journal of ophthalmology. 1995;120(5):577-86.
63. Burgansky-Eliash Z, Wollstein G, Chu T, Ramsey JD, Glymour C, Noecker RJ, et al. Optical coherence tomography machine learning classifiers for glaucoma detection: a

preliminary study. Investigative ophthalmology & visual science. 2005;46(11):4147-52.
64. Huang M-L, Chen H-Y. Development and comparison of automated classifiers for glaucoma diagnosis using Stratus optical coherence tomography. Investigative ophthalmology & visual science. 2005;46(11):4121-9.
65. Bizios D, Heijl A, Hougaard JL, Bengtsson B. Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by Stratus OCT. Acta ophthalmologica. 2010;88(1):44-52.
66. Sun Y, Wang X, Tang X, editors. Deep convolutional network cascade for facial point detection. Proceedings of the IEEE conference on computer vision and pattern recognition; 2013.
67. Suggu SP, Goutham KN, Chinnakotla MK, Shrivastava M. Deep feature fusion network for answer quality prediction in community question answering. arXiv preprint arXiv:160607103. 2016.
68. Le Callet P, Viard-Gaudin C, Barba D. A convolutional neural network approach for objective video quality assessment. IEEE Transactions on Neural Networks. 2006;17(5):1316-27.
69. Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE transactions on medical imaging. 2016;35(5):1285-98.
70. Bottou L, Cortes C, Denker JS, Drucker H, Guyon I, Jackel LD, et al., editors. Comparison of classifier methods: a case study in handwritten digit recognition. Pattern Recognition, 1994 Vol 2-Conference B: Computer Vision & Image Processing, Proceedings of the 12th IAPR International Conference on; 1994: IEEE.
71. Nair V, Hinton GE, editors. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th international conference on machine learning (ICML-10); 2010.
72. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.
73. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L, editors. Imagenet: A large-scale hierarchical image database. Computer Vision and Pattern Recognition, 2009 CVPR 2009 IEEE

Conference on; 2009: IEEE.
74. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large scale visual recognition challenge. International Journal of Computer Vision. 2015;115(3):211-52.
75. White H. Learning in artificial neural networks: A statistical perspective. Neural computation. 1989;1(4):425-64.
76. Scholkopf B, Sung K-K, Burges CJ, Girosi F, Niyogi P, Poggio T, et al. Comparing support vector machines with Gaussian kernels to radial basis function classifiers. IEEE transactions on Signal Processing. 1997;45(11):2758-65.
77. Koppers S, Haarburger C, Merhof D, editors. Diffusion MRI Signal Augmentation: From Single Shell to Multi Shell with Deep Learning. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2016: Springer.
78. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al., editors. Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition; 2015.
79. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.

Abstract in Korean

국문 초록

광간섭 단층 촬영술의 발전에 따라 시신경 두께에 대한 분석을 통해 녹내장의 비침습적 진단이 가능해졌다. 그러나 녹내장의 특성 상 조기진단을 통한 적절한 안압의 유지가 필수적임에 따라 녹내장을 조기에 정확하고 객관적으로 분석하기 위한 컴퓨터 보조진단 시스템의 개발이 요구되고 있다. 본 논문은 조기 녹내장을 정확하게 진단할 수 있는 컴퓨터 보조 시스템을 구현하는 방법으로 기존 기계학습에서 사용되어온 특징기반 분류법과 최근 관심이 고조되고 있는 심층신경망(deep neural network)에서 심층학습(deep learning)에 의해 생성된 특징을 통합한 “특징 융합 심층신경망”을 제안하고 실제 정상인과 환자 영상을 이용한 성능 평가를 통해 임상적 유효성을 검증하였다.

특징 융합 심층신경망은 영상처리(image processing)와 심층학습을 통해서 추출된 이질적인 특징들을 통합하여 구성한 심층신경망이다. 녹내장과 관련된 시신경 결함 부분의 넓이와 깊이 특징들은 전통적인 영상처리 기법을 통해 추출하였고, 심층학습을 통해서는 녹내장과 정상인의 구별과 관련된 특징들을 신경망의

중간단(middle layer) 출력에서 추출하였다. 이렇게 추출한 특징들을 통합하여 특징 융합 심층신경망을 구성하였다.

영상처리 기반 특징의 첫 번째로 시신경 결함 부분의 넓이에 관련된 특징을 추출하기 위해서 시신경 유두 주변의 망막 신경 섬유층과 황반 주변의 신경세포절 속얼기층을 촬영한 두께 이미지와 두께 편차 이미지를 이용해서 분석하였다. 69 명의 녹내장 환자들과 79 명의 정상인에 대해서 세 가지 기준에 의해서 각각의 두께 편차 이미지에서 시신경의 결함 부분을 찾아냈고 그 영역의 넓이를 계산했다. 넓이를 통해서 계산된 심각 지표들의 성능을 ROC 곡선 아래면적(AUC)을 통해서 평가하였다. 모든 심각 지표들에 대해서 녹내장 환자들과 정상인들 간에 두드러진 차이를 보였고($p < 0.0001$), 녹내장 환자들과 정상인들을 정확히 구분하였다(AUC = 0.91~0.95). 이를 통해서 시신경 결함 부분의 넓이 특징들이 녹내장 진단의 객관적인 지표로 이용될 수 있음을 알 수 있었다.

두 번째 영상처리 기반 특징으로 시신경 결함 부분의 깊이에 관련된 특징을 추출하기 위해서 시신경 유두 주변의 망막 신경 섬유층을 촬영한 두께 이미지와 두께 편차 이미지를 분석하였다. 두께 편차 이미지로 분석된 결함 부분에 대해서 정상인의 시신경 두께와 분석된 부분의 시신경 두께의 비를

이용하여 깊이에 관련된 지표를 개발하였다. 이 지표를 이용하여 108 명의 초기 녹내장과 96 명의 중기 녹내장, 그리고 111 명의 말기 녹내장 환자들에 대해서 분석하였고 지표의 성능을 ROC 곡선 아래면적을 통해서 평가하였다. 이 때, 지표에 대해서 각 집단들 간의 두드러진 차이가 있었고($p < 0.001$), 초기 녹내장과 중기 녹내장 환자들을 잘 구분할 수 있을 뿐만 아니라(AUC = 0.98), 중기 녹내장 환자들과 말기 녹내장 환자들도 잘 구분할 수 있었다(AUC = 0.97). 이를 통해 시신경 결함 부분의 깊이 특징이 녹내장의 정도를 구분하는 데 의미 있는 지표임을 알 수 있었다.

시신경 두께 영상을 심층학습에 적용하기 위해서 두 가지 방법을 이용하였는데 한 가지 방법은 LeNet 신경망에 무작위로 분포되어 있는 가중치를 이용하여 심층학습하는 방법이고 다른 한 가지 방법은 VGGNet 신경망을 다른 방대한 이미지 데이터에 미리 학습을 한 가중치를 이용하여 심층학습하는 방법이다. 316 명의 정상인과 226 명의 초기 녹내장, 그리고 246 명의 중말기 녹내장 환자들에 대해서 두 가지 방법으로 학습하였고 각각에 대해서 ROC 곡선 아래면적으로 성능을 평가하였다. LeNet 과 VGGNet 으로 학습한 신경망 모두 정상인들과 녹내장 환자들을 잘 구분할 수 있을 뿐만 아니라(AUC = 0.94, 0.94), 초기 녹내장 환자들도 정확히 구분하였다(AUC = 0.88, 0.89). 이를 통해서 심층학습을

통해서 분석한 두 방법 모두 녹내장에 관련된 특징들을 잘 추출함을 알 수 있었다.

최종적으로 영상처리를 통해서 추출한 특징들과 심층학습으로 추출했던 특징들을 통합하여 특징 융합 심층신경망을 개발하였고 이를 ROC 곡선 아래면적으로 기존 연구들과 성능을 비교하였다. VGGNet 신경망으로 추출한 특징으로 융합한 특징 융합 심층 신경망이 다른 신경망들보다 정상인들과 녹내장 환자들을 정확히 구분할 뿐만 아니라(AUC = 0.96), 초기 녹내장 환자들도 정확히 구분하였다(AUC = 0.92). 이는 기존 연구 결과(AUC = 0.91, 0.82)에 비해서 뛰어난 성능을 보였으며 특히, 초기 녹내장과 정상인을 구분하는 데 뛰어난 성능을 보였다.

이러한 결과를 통해 본 논문에서 제안한 특징 융합 심층신경망이 기존 방법에 비해 녹내장의 진단 및 초기 진단에 높은 정확도를 제공함을 확인하였다. 추후 인구통계학 정보들과 녹내장에 관련된 여러 검사들의 특징들을 특징 융합 심층신경망에 추가 적용한다면 정확도를 더욱 향상시킬 수 있을 것으로 예상된다. 본 논문에서 제안한 특징 융합 심층신경망은 녹내장 조기 진단 뿐만 아니라 녹내장의 진행 정도를 분석하는 데에도 적용할 수 있을 것으로 기대된다.

핵심어 : 특징 융합 심층신경망, 녹내장, 광간섭 단층 영상, 망막 신경 두께 영상, 심층학습

학 번 : 2010-23352