

Development of a Diagnostic Model for Glaucoma

Jihyung Lee¹, Sudong Lee¹, Baekhee Lee¹, Jaheon Kang³, Chi-Hyuck Jun¹, and Heecheon You^{1,2}

¹Department of Industrial Engineering and Management Engineering, POSTECH, Pohang, 790-784

²Department of Creative IT Excellence Engineering, POSTECH, Pohang, 790-784

³Kyung Hee University Hospital at Gangdong, College of Medicine, Kyung Hee University, Seoul, 134-727

ABSTRACT

Objective: The aim of this study was to develop a glaucoma diagnostic model applying the statistical classification method by various testing results for glaucoma. **Background:** Statistical diagnostic models applied with structural (e.g., fundus examination) and functional (e.g., visual field testing) testing results have been developed to diagnose glaucoma more accurately. The existing studies with regards to the glaucoma diagnostic model considered statistical techniques by various input data such as intraocular pressure and cup-to-disc ratio; however, constant exertions for improving the performance of a diagnostic model are necessary through applying unvalidated methods and input data. **Method:** Glaucoma testing results of 145 eyes (glaucomatous eyes = 74, normal eyes = 71) were used as input data. Five classification techniques (linear and Gaussian support vector machines [L-SVM and G-SVM], linear discriminant analysis [LDA], binary logistic regression [BLR], and classification and regression tree [CART]) were applied to develop the model. Used as input, the performance (accuracy, sensitivity, specificity and AUC) of the developed models were compared with among the developed models. **Results:** The PD/G-SVM model which applied with G-SVM with pattern deviation (PD) was determined as an optimal model (accuracy = 0.87; sensitivity = 0.81; specificity = 0.92; and AUC = 0.95). **Conclusion:** The developed PD/G-SVM model can contribute to making a diagnosis of glaucoma more precisely. **Application:** The developed model would be expected to apply to develop a comprehensive diagnostic system for glaucoma.

Keywords: Glaucoma, diagnostic model, classification, support vector machine, pattern deviation

1. Introduction

녹내장(glaucoma)은 시신경 손상이 진행되어 실명에 이르는 안과 질환으로 구조적 검사(structural test) 및 기능적 검사(functional test) 결과를 바탕으로 진단되고 있으나 보다 정확한 진단을 위해서는 진단모형이 필요하다. 녹내장은 안압 증가, 시신경 유두 함몰, 시야 장애 등의 증상이 다양하게 나타나기 때문에 구조적 및 기능적 검사 결과를 통해 진단되고 있다(Park et al., 2005). 구조적 검사는 시신경을 포함한 안구의 내, 외부 형태와 구조를 파악하는 검사로 시신경 유두 검사, 전방각경검사, 망막신경섬유층 두께검사 등이 있다(Caprioli, 1988). 기능적 검사는 시신경 부위별 빛에 대한 민감도, 시표의 움직임 등 시신경의 성능을 측정하는 검사로 시야검사와 시신경

유발전위검사(visual evoked potential) 등이 있다(Caprioli, 1992; Towel et al., 1983). 안과 의료진에 의한 녹내장 진단은 의료진 주관에 개입될 수 있으나 진단모형에 의한 진단은 정량적인 data와 통계적 분류기법을 사용하고 있어 객관적으로 녹내장을 진단할 수 있다. 따라서, 보다 정확한 녹내장 진단을 보조하기 위해서는 통계적 분류기법을 활용한 녹내장 진단모형의 개발이 필요하다.

녹내장 진단모형은 녹내장 유발 인자를 다양한 기술로 측정할 수 있게 됨에 따라 활발하게 연구되고 있으나 진단모형의 성능 향상을 위해서는 핵심 인자 탐색과 통계적 분류기법 개발에 대한 지속적인 연구가 필요하다. Chan et al.(2002)은 시야검사 결과 중 threshold sensitivity (dB)를 mixture of Gaussian(MOG)기법에 적용하여 area under the ROC curve(AUC) = 0.923의 우수한 성능의 녹내장 진단모형을 개발하였다. Townsend et al.(2008)은 시신경 유두

형태 분석 장비인 Heidelberg Retina Tomograph III(HRT, Heidelberg Engineering GmbH, Germany)에서 측정된 이미지 분석 결과(예: cup-to-disc area ratio, vertical cup-to-disc ratio 등)를 recursive partitioning and regression trees(RPART) 기법에 적용하여 AUC = 0.899인 녹내장 진단모형을 개발하였다. 한편, 보다 성능이 높은 진단모형 개발을 위해서는 기존에 사용되지 않은 변수들(예: pattern deviation)과 다양한 통계적 기법이 적용된 진단모형 연구가 필요하다.

본 연구는 다양한 녹내장 유관 변수들을 활용하여 진단모형 개발하고 성능을 분석하였다. 녹내장 검사 결과에서 추출된 다양한 변수들의 조합을 통계적 분류기법에 적용하여 녹내장 진단모형을 개발하였다. 개발된 녹내장 진단모형들의 성능 비교 분석을 통해 최적 녹내장 진단모형을 파악하였다.

2. Materials and Methods

2.1. Data

본 연구의 진단모형 개발에 사용된 data에는 세 가지 시야검사 결과(visual field raw test result, VFR)와 두 가지 global index(GI), 그리고 여덟 가지 clinical index(CI)가 포함되었다. 첫째, 세 가지 VFR은 각각이 다변량 변수로 구성된 threshold sensitivity(TS; $TS_1 \sim TS_{54}$), total deviation(TD; $TD_1 \sim TD_{52}$), 그리고 pattern deviation(PD; $PD_1 \sim PD_{52}$)이다. TS는 시야영역에 grid 형태로 mapping된 test point(24-2 threshold test의 경우 54개)마다 algorithm에 의해 밝기가 결정된 자극이 제시되었을 때, 환자가 반응한 가장 sensitive한 자극의 수치를 나타내며 0 ~ 50 dB 값으로 표현된다. TD는 피검사자의 반응으로 측정된 각 test point별 TS값과 동일한 연령의 정상안 median value의 차이를 나타낸다. PD는 국소적인 시야결손 위치와 정도를 파악하기 위해 TD value 중 맹점 주위 두 개의 test point 값을 제외하고 7번째로 높은 값을 각 tested point에서 뺀 값이다(Figure 1 참조). TD와 PD의 맹점 주변 두 개의 test point 값은 다변량 변수에서도 제외되고 있기 때문에, TS의 다변량 변수에서도 맹점 주변의 두 개의 test point 값은 data에서 제외시켰다. 둘째, 두 가지 GI에는 mean deviation(MD)와 pattern standard deviation(PSD)가 사용되었다. MD는 전반적인 시야 손상 정도를 나타내며 TD 다변량 변수 값들의 평균으로 산출된다. PSD는 전반적인 시야손상 정도 대비 국소적인 시야 손상 정도를 나타내며, 피검사자의 hill of vision과 MD의 표준편차로 산출된다. 마지막으로, 여덟 가지 CI에는 나이(age), 성별(gender), visual acuity(VA), intraocular pressure(IOP), central corneal thickness (CCT), vertical, horizontal,

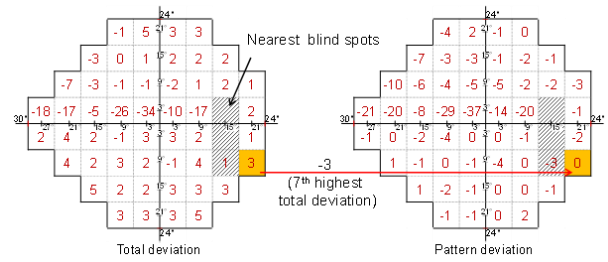


Figure 1. General height adjustment

and average cup-to-disc ratio(VCDR, HCDR, 그리고 ACDR)가 사용되었다. 상기 자료는 5-fold cross validation 기법을 적용하여 진단모형 개발을 위한 learning set과 성능 평가를 위한 testing set을 각각 80%($n = 116$)와 20%($n = 29$)로 나누어 사용되었다. 145개 sample은 녹내장안 74개, 정상안 71개로 구성되었다.

2.2. Classifiers

진단모형은 최적화 기반의 linear와 Gaussian support vector machine (L-SVM와 G-SVM), 통계 기반의 linear discriminant analysis(LDA), binary logistic regression(BLR), 그리고 트리(tree)기반의 classification and regression tree(CART)의 다섯 가지 분류기법을 적용하여 개발되었다. L-SVM과 G-SVM의 개발에는 package LibSVM(Chang and Lin, 2011)이 사용되었다. 각 모형의 성능을 최적화하기 위해 L-SVM과 G-SVM은 simulation을 통해 도출된 parameter인 $c = .0313$ 와 $r = .008$ 가 적용되었고, LDA와 BLR은 녹내장안이 정상안으로 오분류될 때의 비용을 정상안이 녹내장안으로 오분류될 때 비용보다 5배 높도록 설정하였다.

2.3. Glaucoma diagnostic model

녹내장 진단 모형은 세 가지 VFR, 두 가지 GI, 그리고 여덟 가지 CI를 조합한 data set을 통하여 L-SVM, G-SVM, LDA, BLR, 그리고 CART의 다섯 가지 통계적 분류기법을 적용하여 개발되었다. 진단모형 개발에 사용된 learning set은 5-fold cross validation 기법을 적용하여 각 data set에서 80% data($n = 116$)를 무작위로 추출하여 구성되었다. 개발된 진단모형의 명칭은 입력된 변수와 사용된 분류기법의 명칭을 ‘/’를 기준으로 앞과 뒤에 표기하였다. VFR, GI, 그리고 CI 항목에 포함된 변수가 모두 사용될 경우에는 각각의 항목명으로 표기하고 각 변수만 사용되었을 경우에는 변수명을 ‘&’으로 연결하여 표기하였다. 예를 들어, VFR에서 PD가 입력변수로 사용되고 G-SVM이 분류기법으로 사용되었을 경우 진단 모형의 명칭은 PD/G-SVM으로 표기하였다.

2.4. Performance evaluation

개발된 진단모형들의 성능은 sensitivity, specificity, accuracy, 그리고 area under of receiver operating characteristic curve(AUC)로 평가되었다. sensitivity, specificity, accuracy, 그리고 AUC는 총 100회의 5-fold cross validation이 수행되는 동안 매 회 무작위로 구성되는 learning set과 training set을 사용하여 얻은 결과의 평균이 사용되었다.

3. Results

본 연구에서는 PD/G-SVM의 성능이 가장 우수한 것(accuracy = 0.87; sensitivity = 0.81; specificity = 0.92; AUC = 0.95)으로 나타났다(Table 1 참조). VFR중 다변량 변수 PD가 입력변수로 사용된 진단모형의 성능이 나머지 다변량 변수

TS와 TD가 입력변수로 사용된 진단모형의 성능에 비해 우수한 것(AUC: PD/L-SVM = 0.88; TS/L-SVM = 0.82; TD/L-SVM = 0.80)으로 평가되었다. VFR에 속하는 다변량 변수가 입력 변수로 사용된 진단모형들 중에서는 G-SVM기법이 적용된 진단모형의 성능이 다른 두 가지 기법 L-SVM과 LDA가 적용된 진단모형의 성능에 비해 우수한 것(AUC: TS/G-SVM = 0.92; TS/L-SVM = 0.82; TS/LDA = 0.46)으로 나타났다. 일변량 변수들로 구성된 CI 또는 GI&CI가 입력 변수로 사용된 진단모형들 중에서는 BLR기법이 적용된 진단 모형의 성능이 우수한 것(AUC: CI/BLR = 0.76; GI&CI/BLR = 0.92)으로 나타났다. 다변량 변수와 일변량 변수가 복합적으로 구성된 PD&GI&CI가 입력변수로 사용된 진단모형들 중에서도 G-SVM기법이 적용된 진단모형의 성능이 우수한 것(AUC: PD&GI&CI/G-SVM = 0.91)으로 평가되었다.

Table 1. A comparison of performance of classifiers by their accuracy, sensitivity, specificity, and ROC areas by input variables

Data category	Input			Process (model)	Output	Performance			
	variable	n(eye)	# of variables			Accuracy	Sensitivity	Specificity	AUC
Visual field raw test result (VFR)	TS (TS ₁ , TS ₂ , ..., TS ₅₂)	145	52	L-SVM	Normal /glaucoma	0.77	0.68	0.87	0.82
				G-SVM		0.85	0.80	0.90	0.92
				LDA*		0.65	0.64	0.68	0.46
	TD (TD ₁ , TD ₂ , ..., TD ₅₂)			L-SVM		0.76	0.68	0.84	0.80
				G-SVM		0.84	0.79	0.90	0.92
				LDA*		0.64	0.62	0.65	0.67
	PD (PD ₁ , PD ₂ , ..., PD ₅₂)			L-SVM		0.84	0.76	0.91	0.88
				G-SVM		0.87	0.81	0.92	0.95
				LDA*		0.66	0.63	0.69	0.71
Clinical index (CI)	Age, gender, VA, IOP, CCT, VCDR	145	6	L-SVM	0.62	0.64	0.62	0.71	
				G-SVM	0.56	0.57	0.58	0.63	
				LDA*	0.58	0.97	0.19	0.51	
				BLR*	0.67	0.68	0.67	0.76	
				CART	0.68	0.67	0.69	0.73	
Global index (GI) & CI	MD, PSD, age, gender, VA, IOP, CCT, VCDR	145	8	L-SVM	0.84	0.73	0.95	0.91	
				G-SVM	0.74	0.74	0.75	0.81	
				LDA*	0.61	0.94	0.28	0.53	
				BLR*	0.84	0.78	0.90	0.92	
				CART	0.86	0.85	0.87	0.88	
PD & GI & CI	PD (PD ₁ , PD ₂ , ..., PD ₅₂), MD, PSD, age, gender, VA, IOP, CCT, VCDR	145	60	L-SVM	0.83	0.74	0.92	0.87	
				G-SVM	0.84	0.89	0.79	0.91	
				LDA*	0.67	0.62	0.73	0.49	
				BLR*	0.67	0.60	0.74	0.71	
				CART	0.85	0.85	0.85	0.88	

*Cost ratio, C(normal|glaucoma) : C(glaucoma|normal) = 1 : 5

Table 2. A comparison of performance of classifiers

Variable	Study (year)	Process (model)	Performance			
			Accuracy	Sensitivity	Specificity	AUC
TS	Chan et al.(2002)	L-SVM	-	0.66	0.85	0.89
	The present study		0.77	0.68	0.87	0.82
	Chan et al.(2002)	G-SVM	-	0.78	0.88	0.91
	The present study		0.85	0.80	0.90	0.92
	Chan et al.(2002)	LDA	-	0.58	0.75	0.82
	The present study		0.65	0.64	0.68	0.46
PD	The present study	L-SVM	0.84	0.76	0.91	0.88
		G-SVM	0.87	0.81	0.92	0.95
		LDA	0.66	0.63	0.69	0.71

4. Discussion

본 연구는 기존 연구에서 사용되지 않았던 녹내장 유관 다변량 변수인 PD를 진단모형 개발에 적용함으로써, 녹내장안을 보다 정확하게 분류할 수 있는 것으로 파악하였다. 본 연구에서 가장 우수한 진단모형으로 분석된 PD/G-SVM의 성능(AUC = 0.95)은 Chan et al. (2002)의 연구에서 TS를 사용하여 개발된 TS/G-SVM모형의 성능(AUC = 0.91)보다 약 4% 우수한 것으로 분석되었다(Table 2 참조). PD는 시야검사결과에서 피검자사와 동일한 연령의 평균 hill of vision과의 차이를 나타내는 값으로 부분적인 녹내장성 시신경 손상 부위를 정량적으로 파악하는데 중요한 인자로 알려져 있다(Anderson and Patella, 1999). 또한, 각막 또는 수정체가 뿌옇게 흐려지는 증상(media opacities)에 의해 다변량 변수 값이 전체적으로 감소되기 쉬운 TD와 달리 PD는 산출과정에서 TD의 다변량 변수 값이 전체적으로 감소되는 것을 제외 시켰기 때문에 그 영향이 적은 것으로 추정된다(Katz, 2000). 따라서, PD와 같은 인자들이 진단모형 개발에 활용된다면 진단모형의 성능이 보다 향상될 것으로 기대된다.

G-SVM 분류기법은 TS, TD, 그리고 PD와 같은 다변량 변수를 입력변수로 하는 진단모형에 적용되었을 때, 녹내장 분류성능이 우수한 것을 파악하였다. Table 1과 같이 입력변수로 다변량 변수가 사용된 진단모형들 중에서는 G-SVM기법이 적용된 진단모형의 성능(AUC: TS/G-SVM = 0.92; TD/G-SVM = 0.92; PD/G-SVM = 0.95; PD&GI&CI/G-SVM = 0.91)이 가장 우수한 것으로 평가되었다. 그러나 일변량 변수로 구성된 CI 또는 GI&CI가 입력변수로 사용된 진단모형들 중에서 G-SVM기법이 사용된 진단모형의 성능(AUC: CI/G-SVM = 0.63; GI&CI/G-SVM = 0.81)은 다른 분류기법이 사용된 진단모형의 성능(AUC: CI/L-SVM = 0.71;

CI/BLR = 0.76; CI/CART = 0.73; GI&CI/L-SVM = 0.91; GI&CI/BLR = 0.92; GI&CI/CART = 0.88)에 비해 낮게 평가되었다. 그 이유는 G-SVM이 두 개의 범주를 분류할 때 사용하는 hyperplane이 비선형함수이기 때문에 시야검사 결과와 같이 특정한 pattern의 형태로 나타나는 다변량 변수를 잘 분류하였기 때문으로 사료된다(Sangeetha & Kalpana, 2011).

가장 성능이 우수한 것으로 평가된 PD/G-SVM에 의해 오분류된 13 samples은 녹내장 전문의료진에 의해서도 정확한 진단이 어려운 samples로 파악되었다. 본 연구에서 개발된 PD/G-SVM을 사용하여 145개 sample을 분류한 결과를 Table 3과 같이 confusion matrix로 나타내었다. 오분류된 13 samples는 실제범주가 녹내장안에 속하지만 PD/G-SVM에 의해 추정된 범주에서는 정상안으로 분류된 sample(N/G)이 10개, 실제범주가 정상안에 속하지만 PD/G-SVM에 의해 추정된 범주에서는 녹내장안으로 분류된 sample(G/N)이 3개인 것으로 분석되었다. 두 가지 오분류가 발생한 이유는 N/G sample 10개의 경우 시야검사 결과에서는 정상안으로 파악되었으나 시신경 유두 형태와 같은 다른 인자들에 대한 분석에서는 녹내장안의 특성을 나타내고 있었고, G/N sample 3개의 경우 시야검사 결과에서는 녹내장안으로 파악되었으나 IOP, VCD 등의 다른 인자들의 분석 결과에서는 정상안의 특성을 나타냈기 때문으로 사료된다.

본 연구의 결과는 정확한 녹내장 진단뿐만 아니라 녹내장 핵심 인자 추출 및 녹내장 검사 절차 개선에 활용될 수 있다. 본 연구에서 성능이 우수하게 평가된 진단모형에 사용된 입력변수(예: PD)들은 녹내장안의 특성을

Table 3. Confusion matrix by PD/G-SVM model

	Predicted class	Actual class		Total
		G	N	
G	G	64	3	67
	N	10	68	78
Total		74	71	145

※G, glaucomatous eye; N, normal eye

정량적 정보의 형태로 담고 있기 때문에 녹내장 진단을 위한 핵심 인자로 활용될 수 있다. 또한 추출된 핵심 인자를 결과로 취득할 수 있는 검사는 피검사자에 우선적으로 수행되어 녹내장을 조기에 선별하는데 활용될 수 있으므로 녹내장 검사 절차에서 우선적으로 수행되도록 절차를 개선하는데 활용될 수 있다. 그러나 본 연구는 진단모형 개발에 안과 전문의료기관에서 수행되고 있는 검사결과 모두를 인자로 사용하지 않았기 때문에, 추후 연구에서는 본 연구에서 사용되지 않은 인자를 포함하여 진단모형을 개발하고, 핵심인자를 추출하여 기존 녹내장 검사 절차 개선을 위한 연구가 필요하다.

Acknowledgements

The present research was supported by ministry of knowledge economy & National IT industry promotion agency for IT consilience program (C1515-1121-0003).

References

- Anderson, D.R., and Patella, V. M. Automated static perimetry, 2nd ed. St. Louis: C.V. Mosby, Co., 1999.
- Caprioli Joseph. Correlation of structure and function in glaucoma. Quantitative measurements of disc and field. *Ophthalmology*, 95(6), 723-727, 1988.
- Caprioli Joseph. Discrimination between normal and glaucomatous eyes. *Investigative Ophthalmology & Visual Science*, 33(1), 153-159, 1992.
- Chan, K., Lee, T. W., Sample, P. A., Goldbaum, M. H., Weinreb, R. N., and Sejnowski, T. J. Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Transactions on Biomedical Engineering*, 49(9), 963-974, 2002.
- Chang, C. C., and Lin, C. J. LIBSVM: a library for support vector machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2011.
- Park, K. H., Medical treatment of glaucoma. *Pharmacotherapeutics*, 1, 189-196, 2005.
- Sangeetha, R., and Kalpana, B. Performance evaluation of Kernels in multiclass support vector machines. *International Journal of Soft Computing and Engineering*, 1(5), 138-145, 2011.
- Towel, V. L., Moskowitz, A., Samuel, S., and Schwartz, B. The Visual Evoked Potential in Glaucoma and Ocular Hypertension: Effects of Check Size, Field Size, and Stimulation Rate. *Investigative Ophthalmology & Visual Science*, 24(2), 175-183, 1983.
- Townsend, K. A., Wollstein, G., Danks, D., Sung, K. R., Ishikawa, H., Kagemann, L., Gabriele, M. L., and Schuman, J. S. Heidelberg Retina Tomograph 3 machine learning classifiers for glaucoma detection. *British Journal of Ophthalmology*, 92(6), 814-818, 2008.

Author listings

Jihyung Lee: iwoneye@postech.ac.kr

Highest degree: M.S., Department of Industrial & Management Engineering, POSTECH

Position title: Ph.D. student, Department of Industrial & Management Engineering, POSTECH

Areas of interest: Ergonomic Product Design & Development, Digital Human Modeling & Simulation, Ergonomic Interface System for the Disabled Person, Ophthalmic Medical Device

Sudong Lee: dongdong@postech.ac.kr

Highest degree: B.S., Department of Industrial Engineering, Hanyang University

Position title: Ph.D. student, Department of Industrial & Management Engineering, POSTECH

Areas of interest: Data Mining in Healthcare & Medical Diagnosis, Probabilistic Graphical Model

Baekhee Lee: x200won@postech.ac.kr

Highest degree: M.S., Department of Industrial & Management Engineering, POSTECH

Position title: Ph.D. student, Department of Industrial & Management Engineering, POSTECH

Areas of interest: Ergonomic Product Design & Development, Digital Human Modeling & Simulation, Vehicle Ergonomics

Jaheon Kang: kjh0614@khu.ac.kr

Highest degree: Ph.D., Medical college of Kyung Hee University

Position title: Department of Ophthalmology, Kyung Hee University Hospital at Gangdong

Areas of interest: glaucoma

Chi-hyuck Jun: chjun@postech.ac.kr

Highest degree: Ph.D., Department of Industrial Eng. & Operations Research, University of California

Position title: Professor, Department of Industrial & Management Engineering, POSTECH

Areas of interest: Applied Statistics, Informatics, Probability Modeling

Heecheon You: hcyou@postech.ac.kr

Highest degree: Ph.D., Industrial Engineering, Pennsylvania State University

Position title: Associate Professor, Department of Industrial & Management Engineering, Department of Creative IT Excellence Engineering, POSTECH

Areas of interest: Ergonomic product design & development, User interface design & evaluation, Digital human modeling & simulation, Human performance & workload assessment, Work-related musculoskeletal disorders (WMSDs) prevention, Usability testing