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

[UCI]I804:48009-200000366841

의학석사 학위논문

인공신경망을 통한 녹내장성 시야손상 진행의 예측에 관한 연구

Prediction of glaucomatous visual field progression using artificial neural network

울 산 대 학 교 대 학 원

의 학 과

손 길 환

인공신경망을 통한 녹내장성 시야손상 진행의 예측에 관한 연구

지 도 교 수 성 경 림

이 논문을 의학석사학위 논문으로 제출함

2021년 2월

울 산 대 학 교 대 학 원

의 학 과

손 길 환

손길환의 의학석사학위 논문을 인준함

심사위원 이 주 용 인

심사위원 성 경 림 인

심사위원 사 호 석 인

울 산 대 학 교 대 학 원

2021 년 2 월

국문요약

목적: 본 연구에서는 인공신경망을 이용하여 녹내장성 시야 진행을 예측할 수 있는지 알아보고자 하였다.

연구 대상: 1998 년부터 2018 년까지 서울아산병원 녹내장 클리닉에서 정기적으로 시야 검사를 시행한 3,313 명의 5,224 안을 대상으로 하였다.

연구 방법: 연속된 3 회의 검사를 기본 검사로 하여 이후 첫 검사 후 3 년 이내에 mean deviation (MD), pattern standard deviation (PSD) 과 visual field index (VFI), 세 가지 global indices 중 두 가지 이상이 기준치(MD; 1.46dB, PSD; 0.97dB, VFI; 3.69%) 이상 악화되었으며, 3 년 이상, 최대 6 년 동안 호전이 없는 경우 시야 진행군으로 분류하였다. 3 회의 기본 시야 검사를 이용해 global indices 를 이용한 선형회귀모형, pointwise linear regression (PLR) 모형, 가우시안 혼합 모델을 통해 변환한 PLR 모형으로 기본검사 후 3 년째 MD, PSD 와 VFI 를 계산하여 예측된 시야진행여부를 판별하였다. 인공신경망 모형은 두 단계로 설계되었다: 첫 번째 단계에서는 각각 기본검사 후 1 년, 2 년과 3 년째 시야를 예측하는 세개의 인공신경망을 훈련시켰고, 두 번째 단계에서는 기본검사와 첫 번째 단계에서 예측된 시야를 이용하여 시야진행여부를 예측하였다. 시야진행 예측에 있어서 선형모형과 인공신경망모형의 성능을 비교하기 위해 민감도, 특이도 및 진단오즈비를 이용하였다.

결과: 5,224 안 중 327 안(5.92%)이 진행군으로 분류되었다. 진행군은 비진행군에 비해 유의하게 연령이 높았으며, 최초 MD, PSD 및 VFI 가 유의하게 나빴다 (모두 $p < 0.001$). 대상안 중 75%(3,918 안)는 트레이닝군으로, 나머지 25%(1,306 안)는 테스트군으로 배정하였다. 세 가지 선형회귀모형 중에서는 변환 PLR 모형의 성능이 가장 우수하여, 민감도 0.21, 특이도 0.80, 진단오즈비 1.67 을 보여주었다. 인공신경망 모형의 경우 선형 회귀 모형보다 훨씬 좋은 성능을 보여주었으며 area under the receiver operating characteristic curve (AUC) 0.86 으로 특이도 0.90 으로 계산 시 민감도 0.54 및 진단 오즈비 10.57 을 보여주었다.

결론: 선형회귀모형에 비해 인공신경망 모형이 녹내장성 시야 진행의 예측에 더 유용할 수 있음을 보여주었다.

차례

서론 1

방법 2

대상 2

시야 진행군과 비진행군의 분류 2

Global indices 의 계산 3

가우시안 혼합모델을 이용한 시야 감도역치의 변환 3

인공신경망 모델 4

선형회귀 모델 5

통계적 분석 5

결과 5

고찰 8

부록 11

참고문헌 12

서론

녹내장은 2013년 40세에서 80세까지의 세계 인구 중 약 6천4백만명이 앓고 있는 것으로 추정되었으며, 이 중 약 6백 9십만명이 중등도 이상의 원거리 시력장애를 가지고 있거나 혹은 실명한 것으로 추산된다.[1,2] 녹내장은 시신경의 손상이 진행하며 이에 따른 시기능의 손실이 동반되는 질환인데 전형적인 임상 양상은 주변부 시야 손상이 먼저 나타난 후 말기에 중심 시야 손상이 나타나는 것으로 되어 있다. 이와 같은 기능적 손실은 유의하게 호전되지 않는 것으로 알려져 있다.[3] 따라서 녹내장 치료의 일차적인 목표는 녹내장의 진행을 예방하는 것이며 치료적 결정에 있어서 녹내장성 시야 손상의 진행을 미리 예측할 수 있다면 많은 도움이 될 것으로 생각된다.

전통적으로 녹내장성 시야 손상의 추이를 분석하고 변화를 예측하기 위해 두 가지 선형 모형이 사용되어 왔다. 시야 검사의 mean deviation (MD)이나 visual field index (VFI) 등의 global index를 이용한 선형회귀모형과 각 시야지점별로 선형회귀분석(pointwise linear regression; PLR)을 시행하는 것이다.[4–6] 두 가지 모형 중에서는 PLR이 녹내장성 시야 손상의 진행 여부에 대한 임상적 판단에 더 유용하다는 연구결과가 있으며, 비선형회귀모형들(이차회귀모델, 지수회귀모델, 로지스틱 회귀모델, M-추정량 로버스트 회귀모델)의 경우 선형회귀분석에 비해 유의한 우수성이 입증되지 못하여 널리 사용되지 않는다.[7,8] 또한 전통적인 회귀분석의 경우 신뢰성 있는 예측을 위해서는 비교적 많은 수의 검사가 필요하며 한 연구에 따르면 정확한 예측 결과를 보장하기 위해서는 적어도 10개 이상의 시야 검사가 필요하다고 한다.[8]

최근 기계학습 기술의 발전으로 녹내장의 기능적, 구조적 평가 및 예측을 위해 다양한 기계학습 기법들이 적용되고 있다.[9] 대표적으로 시야에 관련해서는 최근 Wen 등이 0.5년 간격으로 최소 0.75년에서 최대 5.5년 간격의 시야 검사 쌍들을 만든 다음 딥러닝을 이용해 시야손상의 진행을 예측하는 모형을 만들었으며, 기존의 선형회귀 모형들과 비교하여 딥러닝 모형에서 pointwise mean absolute error (PMAE)가 유의하게 낮음을 보고하였다.[10]

본 논문에서는 딥러닝 기법을 이용하여 연속된 3개 시야 검사 결과만을 가지고 이후 약 3년간 발생할 시야 결손의 진행을 예측할 수 있는지를 검증하고자 하였다.

방법

대상

본 연구에서는 1998년 5월부터 2018년 5월까지 서울아산병원 녹내장클리닉을 방문한 환자 중 초진 시 30세 이상이며, 3년 이상 정기적으로 험프리 시야검사계(Humphrey Field Analyzer; Carl Zeiss Meditec, Dublin, CA)의 Swedish Interactive Threshold Algorithm과 standard 24-2 프로그램을 이용한 시야검사를 시행하였으며, 신뢰성 있는 시야검사(가양성 15% 미만, 가음성 15%미만, 시고정소실 20% 미만) 중 3년 이상의 기간동안 시야검사를 7회 이상 시행한 환자 4,959명의 8,521안을 일차적으로 선별하였다. 대상안의 선별과 녹내장성 시야손상의 진행여부를 판단하기 위해 절단값 (cut-off value)을 정할 필요가 있었기에, 다음과 같은 방법으로 구하였다: 1) 우선 전체 84,908개의 시야 검사 중 각 안 별로 첫 세 검사를 제외하였고, 2) 각 안 별로 이후 3년간 시행된 시야 검사를 선별한 다음, 3) 각 안 별로 MD, PSD와 VFI의 분산을 구한 다음 그 평균의 제곱근을 절단값으로 삼았다(MD 1.46dB, PSD 0.97dB, VFI 3.69%). 과도한 단기시야변동을 배제하기 위하여 각 안 별로 이전 시야 검사 대비 MD와 VFI의 향상폭이 상기 절단값보다 작은 연속적인 3개 시야 검사 중 시간적으로 가장 빠른 3개 검사를 기본 검사로 삼았다. 또한 정기적이고 지속적인 경과 관찰을 보장하기 위해 기본 검사 이후 첫 검사로부터 3년 이내 시야 검사를 시행하였고(평균 2.74±1.38회), 3년 이후 6년 이내에도 시야 검사를 시행한(평균 2.23±1.30회) 경우만을 선별하였으며 그 결과 3,313명의 5,224안이 연구대상이 되었다. 이 중 25%를(1,306안) 무작위로 추출하여 테스트군으로 배정하고 나머지 75%(3,918안)은 트레이닝군으로 배정하였다.

시야 진행군과 비진행군의 분류

시야진행을 판단하기 위해 기본 검사 후 첫 번째 검사로부터 3년 이내에 MD, PSD와 VFI 중 두 가지 이상이 기본 검사의 평균치보다 절단값 이상 악화되었으며 (MD 기준으로 1.46dB 이상, PSD는 0.97dB이상, VFI는 3.69% 이상), 악화된 이후 6년까지의 경과 관찰 기간 동안 단 한번도 해당 지표들이 기준치보다 개선되지 않은 경우 시야진행군으로 분류하였다. 기본 검사 후 첫번째 검사로부터 6년 이상 경과한 시점의 시야검사는 연구대상에서 제외하였다.

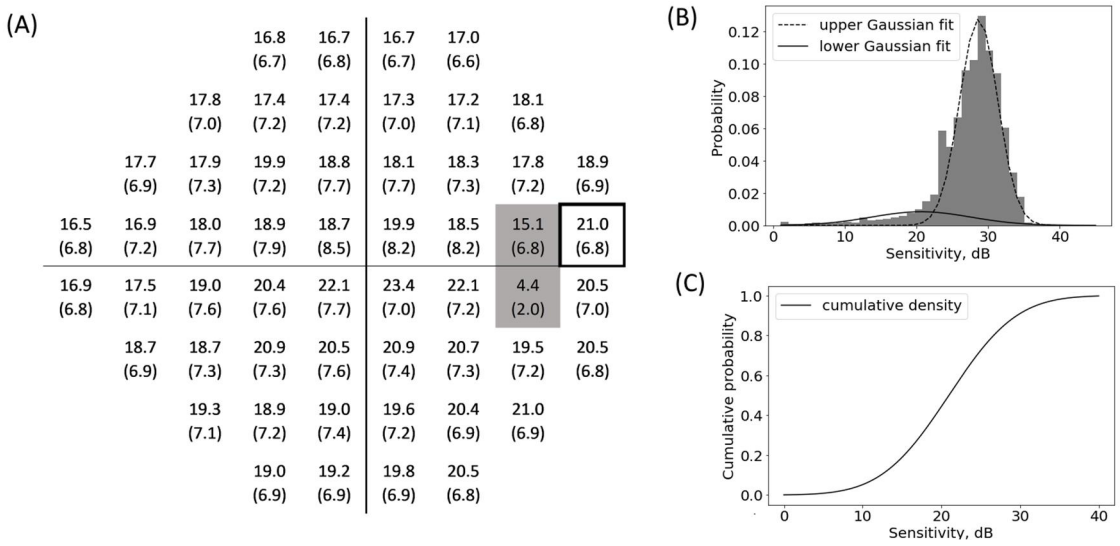
Global indices 의 계산

PLR을 이용해 예측된 시야의 시야결손 진행여부를 판별하기 위해서는 MD, VFI와 PSD를 산출할 필요가 있어 2003년 5월부터 2013년 5월까지 서울아산병원 녹내장클리닉을 방문한 환자 중 30세부터 75세까지 1세 단위로 정상안 (교정시력 0.8 이상, 시신경유두 및 망막신경섬유층의 형태학적 이상이 없으며, 시야 검사에서 patter deviation (PD) plot에서 비정상점이 없으며, MD > -0.5이고 VFI 100%인 안)을 각 50안씩 선정하였으며, 75세 이상 정상안 150안을 따로 선정하였다. 이후 상기 모집된 정상군을 이용하여 기존에 기술된 방식대로 MD, pattern standard deviation (PSD)와 VFI를 계산하였다.[11–13]

가우시안 혼합모델을 이용한 시야 감도역치의 변환

인공신경망 계산의 효율을 향상시키기 위해 데시벨로 측정된 시야 감도 역치를 0과 1사이의 값으로 변환하였으며 이를 위해 트레이닝군과 테스트군을 포함한 84,908개의 시야 검사의 각 54개 점에 대해 1dB 이상인 감도역치분포를 scikit-learn의 Bayesian Gaussian Mixture() 함수를 이용하여 두 개의 가우시안 분포로 근사하였다. 그 다음, 평균값이 더 낮은 가우시안 분포의 누적분포함수를 이용하여 감도 역치의 데시벨 값을 변환하였다 (그림 1).[14–17]

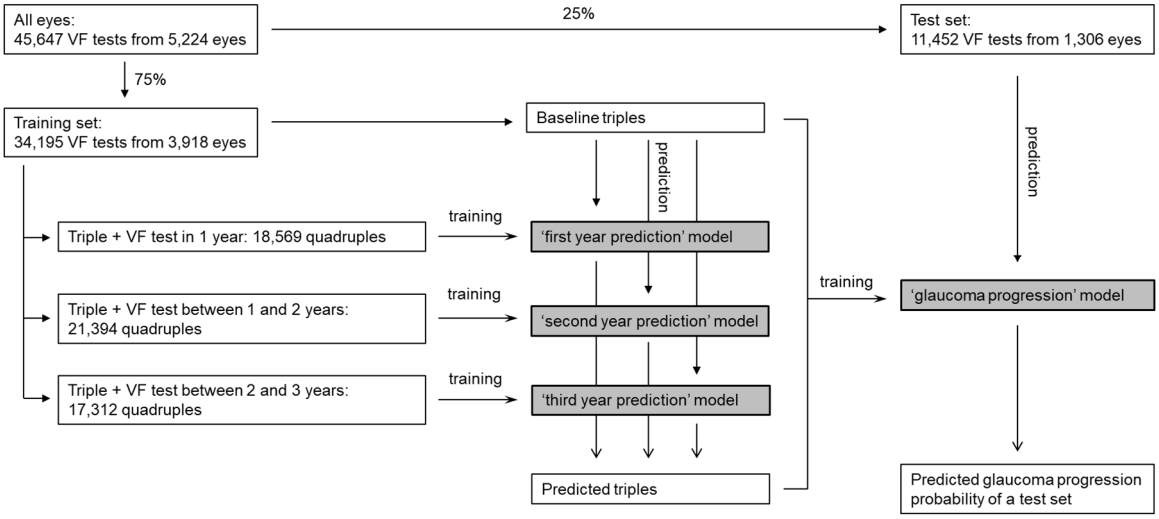
그림 1. 변환에 사용된 각 점 별 정규 분포의 평균과 표준편차 (A), 상측의 최비측 점 (진한 테두리로 표시)의 히스토그램 및 가우시안 혼합 모델을 이용한 근사 (B), 근사된 두 가우시안 분포 중 평균값이 낮은 분포의 누적분포함수 (C).



인공신경망 모델

우선 분석을 위한 준비단계로 트레이닝군과 테스트군의 모든 시야검사를 가우시안 혼합모델을 이용해 변환한 다음 8x9 크기의 2차원텐서로 만들었다. 인공신경망 모델은 두 단계로 설계되었다. 첫 번째 단계에서는 기본 검사 3개를 8x9x3의 3차원 텐서로 적층한 다음, 기본 검사 이후 365일까지, 366일에서 730일까지, 731일에서 1095일까지의 세 기간별로 시행된 시야 검사 1개와 가능한 모든 짝을 만들어 세 개의 동일한 구조의 인공신경망을 훈련시켰다. 두 번째 단계에서는 첫 번째 단계에서 훈련된 세 가지 인공신경망을 이용하여 기본 검사 3개로 이루어진 3차원 텐서를 각 시간대에 대한 인공신경망의 예측에 해당하는 세 개의 시야 검사로 변환하였으며 이를 순서대로 적층하여 두 번째 8x9x3의 3차원 텐서를 생성하였다. 이렇게 생성된 기본 검사와 예측 검사의 두 8x9x3 텐서를 입력으로 하여 시야 검사 여부에 대해 인공신경망을 훈련하였다. 두 가지의 인공신경망 훈련에는 트레이닝군만 이용하였으며 손실 함수로는 두 모델 모두 binary crossentropy를 이용하였다. 첫 번째 단계의 3개 신경망은 최적화 알고리즘으로 Adam 법을 이용해 각각 300에폭만큼 훈련하였다. 두 번째 단계의 신경망은 최적화 알고리즘으로 모멘텀 확률적 경사하강법을 이용하였으며 과적합을 피하기 위해 검증손실값이 연속 2회 이상 증가한 경우 훈련을 중단하였다 (그림 2 및 부록 1).

그림 2. 인공신경망을 이용한 녹내장성 시야손상 예측모형의 모식도.



선형회귀 모델

인공신경망모델과 비교를 위하여 세 가지의 선형회귀모형을 구축하여 기본 검사 후 3년(1095일)이 경과한 시점에의 예상치를 구한 후 시야 진행군과 비진행군의 분류에 사용된 것과 같은 기준으로 (기본검사 대비 MD:1.46 dB, PSD: 0.97dB, VFI 3.69% 중 두 가지 이상 악화) 시야 진행과 비진행의 예상치를 구하였다. 첫 번째 모형은 MD, PSD와 VFI에 대하여 각각 선형회귀분석을 시행하였으며, 두 번째 모형은 시야의 54개 점에 대해 각각 선형회귀를 시행하는 PLR을 사용하였으며, 예측된 감도 역치의 범위는 0dB 이상, 50dB미만으로 제한하였다. 세 번째 모형은 상기 기술한 가우시안 혼합 모델을 이용하여 각 점을 누적분포함수로 변환한 다음 PLR을 시행하고 역 누적 분포 함수(percent point function)를 이용하여 다시 데시벨로 변환하였으며 PLR과 마찬가지로 예측된 감도 역치의 범위는 0dB 이상, 50dB미만으로 제한하였다.

통계적 분석

인공신경망 분석에는 Python (version 3.7, https://www.python.org/)환경에서 구동되는 Tensorflow (version 2.3, https://www.tensorflow.org/)와 Keras(version 2.4, https://keras.io/)를 이용하였다. 기타 통계적 분석에는 SAS 9.4 (SAS Institute Inc., Cary, NC)를 이용하였으며 범주형 자료의 비교에는 카이제곱검정을, 연속형 변수의 비교에는 독립 t-검정을 적용하였다.

결과

연구에 포함된 3,313명의 평균연령은 56.6세였으며 5,224안의 기본검사 3회의 평균 MD는 -3.99 (±5.48)dB, VFI는 88.7 (±16.9)%였으며 기본 검사의 첫 검사로부터 평균 7.25 (±2.89)년 동안 평균 8.74 (±2.72)번의 검사를 시행하였으며 트레이닝군과 테스트군에서 유의한 통계척 차이는 없었다 (표1). 경과 관찰 시작 시 연령은 시야 진행군에서 59.6세로 비진행군의 56.3세보다 더 높았으며 ($p < 0.001$), 시야 진행군에서는 3회 기본검사의 평균 MD (-7.04dB vs -3.76dB, $p < 0.001$), PSD (7.50dB vs 4.54dB, $p < 0.001$)와 VFI (79.8% vs 89.3%, $p < 0.001$)가 모두 비진행군에 비해 유의하게 나빴다 (표2).

표1. 트레이닝군과 테스트군의 기본 특성

	Train group	Test group	p-value
Patients, count	3375	1249	
Eyes, count	3918	1306	
Laterality (right eye), count	1949 (50.3%)	645 (49.4%)	0.823
Age, years	56.5 (11.7)	56.6 (11.8)	0.707
No. of HVF tests, count	8.73 (2.83)	8.77 (3.05)	0.668
Follow-up time, years	7.24 (2.77)	7.27 (2.54)	0.669
Initial MD, dB*	-3.96 (5.41)	-4.09 (5.67)	0.478
Initial PSD, dB*	4.69 (4.05)	4.82 (4.17)	0.332
Initial VFI, %*	88.8 (16.7)	88.3 (17.4)	0.319

**Arithmetic mean of three baseline tests*

HVF: Humphrey visual field, MD: mean deviation, PSD: pattern standard deviation, VFI: visual field index

Numbers in the parentheses are standard deviations if not denoted otherwise.

표2. 시야 진행군과 비진행군의 기본 특성

	Not progressed	Progressed	p-value
Eyes, count	4897	327	
Laterality (right eye), count	2419 (49.4%)	175 (46.5%)	0.149
Age, years	56.3 (11.6)	59.6 (12.4)	<0.001*
No. of HVF tests, count	8.73 (2.86)	8.88(3.28)	0.423
Follow-up time, years	7.27 (2.72)	6.82 (2.57)	0.002*
Initial MD, dB	-3.76 (5.44)	-7.04 (5.19)	<0.001*
Initial PSD, dB	4.54 (4.04)	7.50 (3.75)	<0.001*
Initial VFI, %	89.3 (16.7)	79.8 (16.5)	<0.001*

**statistically significant at $p < 0.05$ level **Arithmetic mean of three baseline tests*

HVF: Humphrey visual field, MD: mean deviation, PSD: pattern standard deviation, VFI: visual field index

Numbers in the parentheses are standard deviations if not denoted otherwise.

인공신경망을 이용한 시야결손 진행의 예측결과 area under the receiver operating characteristic curve (AUC)값은 0.86으로 나타났으며 특이도를 0.90으로 계산할 경우 민감도는 0.54로 나타났다 (그림 3). 선형회귀모형의 경우, 세 모형 모두 음성예측도는 0.93으로 나타났으나 민감도는 최고 0.21에 불과하였으며(변환 PLR 모형) 특이도 역시 인공신경망모형보다 낮았다 (변환 PLR 모형에서 0.8). 선형회귀모형의 경우 AUC를 그릴 수 없으므로 진단오즈비로 비교할 경우 인공신경망의 10.57에 비해 가장 우수한 변환 PLR

모형에서 1.06으로 선형회귀모형에 비해 인공신경망모형이 약 10배가량 우수한 것으로 나타났다. 실제 분류와 예측된 분류가 일치하는 정도를 나타내는 Cohen's kappa값은 인공신경망에서 0.31이었으며 선형모형에서는 0.01 미만으로 이 역시 인공신경망이 훨씬 우수하였다.

그림 3. 테스트군에서 인공신경망을 이용한 녹내장 진행 예측의 receiver operating characteristic curve 및 특이도 0.90에서 오분류표.

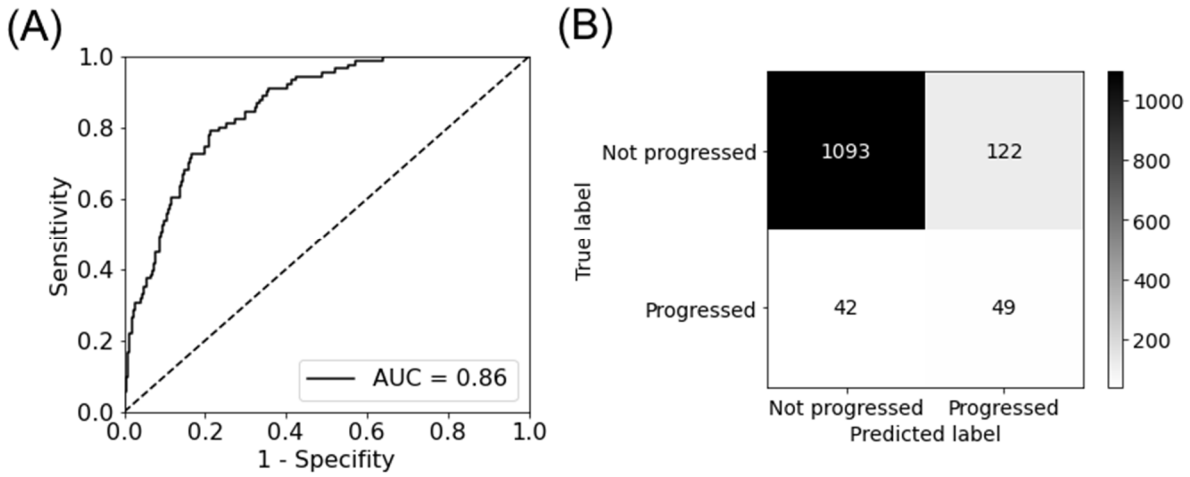


표 3. 테스트군에서 인공신경망 및 선형회귀모형을 이용한 녹내장 진행의 예측 정확도.

	Sensitivity	Specificity	PPV	NPV	DOR	Kappa*
Neural network model	0.54	0.90	0.29	0.96	10.57	0.31
Indices linear model	0.19	0.75	0.05	0.93	0.70	-0.03
Pointwise linear model	0.03	0.96	0.06	0.93	0.74	-0.01
Transformed pointwise linear model**	0.21	0.80	0.07	0.93	1.06	0.01

** Cohen's kappa*

***Visual field thresholds have been transformed using Gaussian mixture model before PLR and inverse-transformed after regression.*

PPV = positive predictive value, NPV = negative predictive value, DOR = diagnostic odds ratio

고찰

본 연구에서는 인공신경망을 이용하여 신뢰성 있는 단 3개의 시야검사 결과만으로도 향후 3년간 시야진행을 AUC 0.86 수준으로 예측할 수 있음을 보였다. 특이도를 0.90으로 계산할 경우, 본 연구에 사용된 인공신경망 모형이 시야결손이 진행하지 않을 것으로 예측한 경우 실제 약 96%에서 진행하지 않았으며, 진행할 것으로 예측한 경우에는 약 29%에서 실제로 진행한 것으로 나타났다.

시야진행여부를 결과값으로 삼았기 때문에 시야진행여부를 판단하는 방법이 중요하다. 녹내장 전문의가 진행여부를 판독하는 것이 가장 이상적이겠으나, 본 논문과 같이 자료의 양이 많을 경우 현실적으로 다수의 전문의가 모두 판독하기에 어려움이 있으며, 선행연구에 따르면 녹내장 전문의간의 판독 일치율도 median kappa = 0.32 수준으로 완벽하지 못하다.[7] 따라서 본 연구에서는 MD, PSD와 VFI를 기준으로 녹내장 진행여부를 판단하였다. 아직까지 global indices에서 어느 정도의 변화가 유의미한 시야진행으로 볼 수 있는지에 대한 표준적인 수치는 없으며, Early Manifest Glaucoma Trial의 경우 시야진행이 없다고 판단된 치료군에서 median MD 진행속도는 -0.03dB/month로 36개월로 환산할 경우 약 -1.08dB/3yrs 가 된다.[18] 다수의 선행연구들을 종합하여 분석한 De Moraes et al. 에 따르면 FDA에서 공인된 녹내장 진행의 기준인 Glaucoma Change Probability (GCP) criteria 에 해당되는 MD의 감소속도는 −0.5 dB/yr 이상이라고 하였으며 이는 본 연구에서 사용한 MD -1.46dB/3yrs 와 근사한 수치이다.[19] 녹내장 환자에서 MD, PSD와 VFI의 수치 및 변화가 상당한 차이가 있음이 알려져 있으며 본 논문에서는 한가지 index에 전적으로 의존하기보다는 세가지 index 중 두가지 이상이 악화된 경우에만 녹내장 진행으로 판단하였다.[20–22]

선형회귀모형의 경우 3개의 검사만으로는 좋은 결과를 얻기 어려웠으며 가우시안 혼합모델 변환을 적용한 모형이 다소 향상된 예측력을 보였다. 가우시안 혼합모델을 이용한 기법들은 CT, MRI, 병리 슬라이드 등 의학분야에서 광범위하게하게 활용되고 있으며, 안과학적으로도 안저사진의 정합 (registration)이나 전안부안구광학단층촬영을 이용한 원발폐쇄각의증의 분류 등에 활용되었다.[23–27] 시야검사에서는 각 점이 정상역치보다 감소했는지 여부를 판별하는 것이 중요하며, 기존에 널리 사용되고 있는 MD, PD, PSD와 VFI 모두 동 연령대에서 각 점의 정상적인 감도역치의 분포에 근거하여 비교하여 산출되고 있다. 본 연구에서는 데시벨 단위를 0에서 1로 변환 시 기본적인 정규화 방식보다 비정상

감도역치를 잘 반영하는 변환방식을 찾고자 하였다. 널리 사용되는 로지스틱 함수를 포함한 시그모이드형 함수를 이용할 경우 비정상 감도역치를 잘 반영하는 매개 변수를 계산할 필요가 있다. 본 연구에서는 연구대상 시야검사의 감도역치가 정상역치 부근에 밀도있게 분포하고 있음에 착안하여 정상역치에 인접한 다수의 점들을 가우시안분포로 근사하고 나머지 정상감도보다 감소한 점들을 대표할 수 있는 제2의 분포를 근사하고자 하였다.

본 연구에서는 시야자료를 convolutional neural network (CNN)을 이용하여 통상적인 red, green, blue 의 3 채널 이미처럼 분석하였다. 시계열자료의 활용을 위해 1 단계 신경망모형에서 연속되는 3 개 시야와 그 후 각각 1 년, 2 년 , 3 년안에 시행된 시야검사와 가능한 모든 짝을 만들어 3 개의 인공신경망을 훈련시키는 일종의 sliding window 방법을 사용하였다.[28] 하지만 2 단계 인공신경망에서는 각 안별로 기본검사 3 개만을 이용하였으며 이러한 방법으로는 각 안의 누적된 모든 시야검사를 활용할 수 없어 임상적으로 중요한 정보를 누락시키는 결과로 이어질 수 있다. 본 연구에서 사용한 CNN 모형의 또다른 문제로, 각 시야검사간의 시간적 간격을 반영기 어렵다. 이를 보완하기 위해 상기한 바와 같이 1 단계 모형에서는 시야검사들을 1 년 단위로 분할하였으며, 시야진행의 기준도 3 년 이내 진행으로 제한하였다. 하지만 실제 임상에서 시야검사가 꼭 정기적으로 이루어지는 것이 아니며 시야손상의 진행속도 역시 매우 중요한 고려사항이므로 시간정보를 활용하지 못하는 것은 중요한 제한점이라 할 수 있다.

이러한 한계점을 극복하기 위해 활용할 수 있는 분석기법 중에 대표적인 것으로 순차 분석과 시계열 분석이 있다. 순차적 자료의 대표적인 예는 자연어 처리이며, 과거부터 인공지능 분야에서 적극적으로 연구되어왔다. 특히 인공신경망에서는 순차적으로 입력되는 자료에서 은닉 상태(hidden state)를 통해 과거 자료의 특성을 기억하고 새로 입력된 값을 적용해 갱신하는 recurrent neural network (RNN)의 개념이 1982 년에 이미 제안되었으며, 이를 토대로 발전시킨 long short term memory networks (LSTM)이나 2014 년에 조경현 박사가 제안한 gated recurrent units (GRU) 등이 널리 활용되고 있다.[29–31] RNN 계통 모형의 경우 순차적으로 자료를 입력해야하는 단점이 있어 병렬화에 어려움이 있으며 최근에는 자연어처리에서 self-attention 만으로 구성된 네트워크가 더욱 좋은 성능을 보이고 있다.[32] 인공신경망을 이용한 시계열 분석 역시 상기한 CNN 과 RNN 을 포함하여 다양한 기법들이 연구되고 있으나 대부분의 시계열 모형은 시간간격이 일정하며 시간적으로 오래 축적된 자료를 대상으로 한다.[33] 이에 비해, 임상적 자료는 다양한 형식의 자료가 섞여있으며,

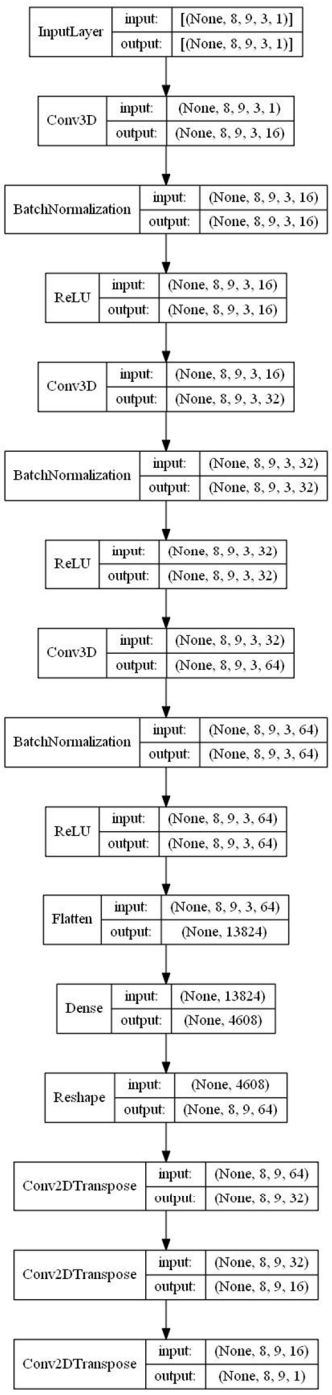
잡음이 많으며 (noisy), 기간이 불규칙하고, 많은 자료의 확보가 어려우며, 각 개인별 차이가 크며 시간적으로는 많은 자료가 축적되기 어려운 특성이 있으며 이에 따라 임상적 자료의 특성에 보다 적합한 인공신경망을 개발하기 위한 연구들이 이루어지고 있다.[34–36]

Wen et al. 은 한 시점의 시야검사로 향후 5 년까지의 시야를 예측하는 연구를 진행하였으며 선형모형과 비교하여 인공신경망 모형이 유의하게 우수함을 보고하였다.[10] 본 연구는 상기 연구와 CNN 을 활용하여 각각 다른 시간대의 시야검사들을 짝을 지어 분석하였다는 점에서 유사성이 있으나 뚜렷하게 발전된 몇 가지 특징이 있다. 우선, 입력으로 한 개의 시야검사 대신 연속적이면서 서로 다른 시점의 3 개의 시야 검사를 사용함으로써 시간에 따른 시야결손의 변화를 보다 잘 반영하고 있으며, 2 단계 신경망에서는 시야결손 진행 예측의 정확도를 평가하므로서 임상적 유용성을 높였다. 인공신경망을 비교하는 기준인 선형모형의 경우 상기 연구에서는 한 시점의 검사만으로는 회귀분석이 불가능하기 때문에 상기 연구에서는 모든 안의 시야 진행속도를 정규분포로 가정하여 일률적으로 적용하였기에 임상에서 활용되는 선형회귀분석과는 차이가 있으며 시야결손의 양상에 따른 시야진행속도의 차이를 반영하기 어려운 한계가 있으나, 본 연구에서는 세 시간대의 시야검사를 이용해 선형회귀분석을 시행하였기에 보다 임상적 관련성을 높였다.

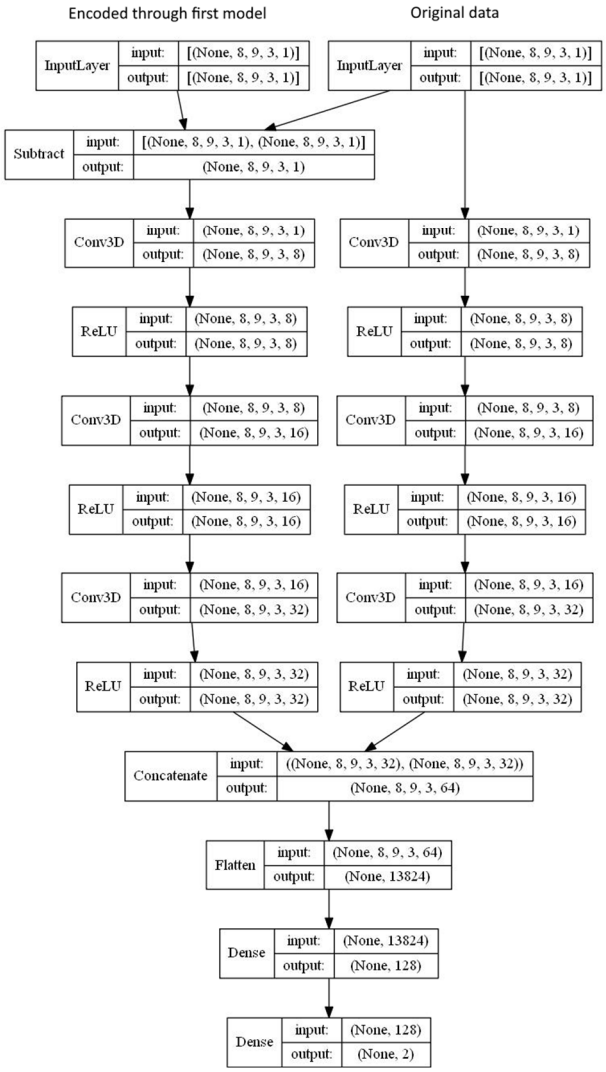
본 연구에서는 5,224 안을 대상으로 하였으나 통상적으로 딥러닝에 사용되는 자료에 비하면 양적으로 많은 부족함이 있으며, 트레이닝군과 테스트군의 손실값이 테스트군 기준으로 AUC 0.86 인근에서 더 이상 동행하지 못하고 발산되는 현상이 발생하였다. 또한 본 연구에서는 시야검사만을 이용하였지만 실제 임상에서는 시야검사 이외에도 안압, 세극등검사, 치료에 대한 순응도, 안저사진과 안구광학단층촬영 등의 다양한 소견들을 종합해서 시야진행 가능성을 판단하게 된다. 보다 다면적이고 종합적인 자료의 수집과 보다 발전된 인공신경망모형의 개발을 통하여 본 연구보다 향상된 예측결과를 얻을 수 있을 것으로 기대된다.

부록 1. 본 연구에서 사용된 인공신경망의 모식도: 1 단계 모델 (A)와 2 단계 모델 (B)

(A)



(B)



참고문헌

1. Tham YC, Li X, Wong TY, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis. *Ophthalmology*. 2014;121(11):2081-2090.

2. WHO. *World Report on Vision*. Vol 214.; 2019.

3. Quigley HA. Glaucoma. *Lancet*. 2011;377(9774):1367-1377.

4. Bengtsson B. Prediction of Glaucomatous Visual Field Loss by Extrapolation of Linear Trends. *Arch Ophthalmol*. 2009;127(12):1610.

5. Wilkins MR, Fitzke FW, Khaw PT. Pointwise linear progression criteria and the detection of visual field change in a glaucoma trial. *Eye*. 2006;20(1):98-106.

6. Shon K, Wollstein G, Schuman JS, et al. Prediction of glaucomatous visual field progression: Pointwise analysis. *Curr Eye Res*. 2014;39(7):705-710.

7. Viswanathan AC, Crabb DP, McNaught AI, et al. Interobserver agreement on visual field progression in glaucoma: A comparison of methods. *Br J Ophthalmol*. 2003;87(6):726-730.

8. Taketani Y, Murata H, Fujino Y, et al. How Many Visual Fields Are Required to Precisely Predict Future Test Results in Glaucoma Patients When Using Different Trend Analyses? *Investig Opthalmology Vis Sci*. 2015;56(6):4076.

9. Devalla SK, Liang Z, Pham TH, et al. Glaucoma management in the era of artificial intelligence. *Br J Ophthalmol*. 2020;104(3):301-311.

10. Wen JC, Lee CS, Keane PA, et al. Forecasting future humphrey visual fields using deep learning. *PLoS One*. 2019;14(4):1-14.

11. Anderson DR, Patella VM. *Automated Static Perimetry.* 2nd ed. St. Louis: Mosby; 1999.

12. Nordmann JP, Mesbah M, Berdeaux G. Scoring of visual field measured through Humphrey perimetry: Principal component varimax rotation followed by validated

cluster analysis. *Investig Ophthalmol Vis Sci*. 2005;46(9):3169-3176.

13. Bengtsson B, Heijl A. A Visual Field Index for Calculation of Glaucoma Rate of Progression. *Am J Ophthalmol*. 2008;145(2):343-353.

14. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res*. 2011;12(85):2825-2830. http://jmlr.org/papers/v12/pedregosa11a.html.

15. Bishop C. *Pattern Recognition and Machine Learning*. New York: Springer; 2006.

16. Attias H. A Variational Bayesian Framework for Graphical Models. In: *In Advances in Neural Information Processing Systems 12*. MIT Press; 2000:209-215.

17. Blei D, Jordan M. Variational inference for Dirichlet process mixtures. *Bayesian Anal*. 2004;1(1).

18. Heijl A, Leske MC, Bengtsson B, et al. Reduction of intraocular pressure and glaucoma progression. *Evidence-Based Eye Care*. 2003;4(3):137-139.

19. De Moraes CG, Liebmann JM, Levin LA. Detection and measurement of clinically meaningful visual field progression in clinical trials for glaucoma. *Prog Retin Eye Res*. 2017;56(1):107-147.

20. Rao HL, Senthil S, Choudhari NS, et al. Behavior of visual field index in advanced glaucoma. *Investig Ophthalmol Vis Sci*. 2013;54(1):307-312.

21. Valente C, D'Alessandro E, Iester M. Classification and Statistical Trend Analysis in Detecting Glaucomatous Visual Field Progression. *J Ophthalmol*. 2019;2019.

22. Gardiner SK, Demirel S. Detecting Change Using Standard Global Perimetric Indices in Glaucoma. *Am J Ophthalmol*. 2017;176:148-156.

23. Nishio M, Tanaka Y. Heterogeneity in pulmonary emphysema: Analysis of CT attenuation using Gaussian mixture model. Fehrenbach H, ed. *PLoS One*. 2018;13(2):e0192892.

24. Yantao Song, Zexuan Ji, Quansen Sun. An extension Gaussian mixture model for

brain MRI segmentation. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE; 2014:4711-4714.

25. Ghazvinian Zanjani F, Zinger S, de With PHN. Deep Convolutional Gaussian Mixture Model for Stain-Color Normalization of Histopathological Images. In: ; 2018:274-282.

26. Liu C, Ma J, Ma Y, et al. Retinal image registration via feature-guided Gaussian mixture model. *J Opt Soc Am A*. 2016;33(7):1267.

27. Nongpiur ME, Gong T, Lee HK, et al. Subgrouping of primary angle-closure suspects based on anterior segment optical coherence tomography parameters. *Ophthalmology*. 2013;120(12):2525-2531.

28. Sadouk L. CNN Approaches for Time Series Classification. In: *Time Series Analysis - Data, Methods, and Applications*. IntechOpen; 2019.

29. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci*. 1982;79(8):2554-2558.

30. Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Comput*. 1997;9(8):1735-1780.

31. Chung J, Gulcehre C, Cho K, et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. December 2014. http://arxiv.org/abs/1412.3555.

32. Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need. *CoRR*. 2017;abs/1706.0. http://arxiv.org/abs/1706.03762.

33. Ismail Fawaz H, Forestier G, Weber J, et al. Deep learning for time series classification: a review. *Data Min Knowl Discov*. 2019;33(4):917-963.

34. Baytas IM, Xiao C, Zhang X, et al. Patient subtyping via time-aware LSTM networks. *Proc ACM SIGKDD Int Conf Knowl Discov Data Min*. 2017;Part F129685:65-74.

35. Che Z, Purushotham S, Cho K, et al. Recurrent Neural Networks for Multivariate Time Series with Missing Values. *Sci Rep*. 2018;8(1):1-12.

36. Ruan T, Lei L, Zhou Y, et al. Representation learning for clinical time series prediction tasks in electronic health records. *BMC Med Inform Decis Mak*. 2019;19(Suppl 8):1-14.

Abstract

Purpose: To assess predictability of glaucomatous visual field (VF) progression using artificial neural network.

Participants: A total of 5,224 eyes of 3,313 glaucoma patients who underwent regular VF examination at the glaucoma clinic of Asan Medical Center between 1998 and 2018.

Methods: Eyes which had shown significant deterioration in at least two criteria based on global indices (mean deviation MD worse than 1.46dB, pattern standard deviation (PSD) 0.97dB, and visual field index (VFI) < -3.69%) in three years from the first test after consecutive 3 baseline VF tests and the deterioration was kept after 3 years up to 6 years were classified as progression group. Three linear models, linear regression based on global indices, pointwise linear regression (PLR) and PLR after transformation through Gaussian mixed model were fitted with the three baseline tests. Predicted MD, PSD and VFI at 3 years after baseline tests from linear models were calculated to determine predicted progression. Neural network was constructed in two stages: first stage consisting of three networks predicting visual field at one, two and three years after baseline tests respectively, and the second stage predicting visual field progression using three baseline tests and predicted three tests from the first stage. Linear and neural network models were compared using sensitivity, specificity and diagnostic odds ratio (DOR).

Results: Among 5224 eyes, 327 eyes were classified as progression group (5.92%). Progression group had significantly higher age, worse initial MD, PSD and VFI than non-progression group ($p < .001$ for all). Training set was constructed with 75% of eyes (3,918 eyes) and remaining 25% of eyes (1,306 eyes) were used as test set. Of three linear models, transformed PLR has performed the best with sensitivity of 0.21, specificity of 0.80 and DOR of 1.06. The neural network model was significantly better with area under the receiver operating characteristic curve (AUC) of 0.86 and sensitivity of 0.54 at specificity 0.90 resulting in DOR of 10.57.

Conclusion: The neural network model has shown significantly better performance than linear models for the prediction of the VF progression.