

안진 검출을 위한 딥러닝 기반의 동공 추적 알고리즘

이예린¹, 이세나¹, 류지승¹, 서경덕¹, 왕형준², 서영준^{2,3*}, 양세정^{1*}

연세대학교 의공학과¹

연세대학교 원주의과대학 이비인후과학교실²

연세대학교 원주의과대학 청각재활연구소³

Pupil tracking algorithm for nystagmus detection with deep learning

Yerin Lee¹, Sena Lee¹, Jiseung Ryu¹, Kyungdeok Seo¹, Hyeong Jun Wang², Young Joon Seo^{2,3*}, Sejung Yang^{1*}

Department of Biomedical Engineering, Yonsei University, Korea¹

Department of Otorhinolaryngology²,

Research Institute of Hearing Enhancement³,

Yonsei University Wonju College of Medicine, Korea

yerin001@hotmail.com, *okas2000@hanmail.net, *syang@yonsei.ac.kr

Abstract

Nystagmus is repetitive, involuntary movement of eyes. Various vestibular diseases can be diagnosed by nystagmus induced by positional tests. In recent years, video-oculography has become a common screening method for nystagmus due to its non-invasive feature, but classification of nystagmus pattern requires professional knowledge. Therefore, there is a need for automatic detection of nystagmus to diagnose disease in a faster and accurate manner. The aim of this study is to obtain precise trajectory of pupil to determine nystagmus pattern. Pupil tracking algorithm consists of deep-learning model for pupil segmentation and center tracking algorithm to determine the center of the pupil. In this paper, the performance of U-Net is compared with previous deep-learning based pupil tracking model DeepEye, and the center tracking algorithm is evaluated with signal analysis. U-Net outperformed DeepEye with dice coefficient of 93.5% and the center tracking algorithm effectively reduced noises associated with tracking the pupil.

1. 연구 배경

전정기능검사 중 비디오안구운동검사는 전정신경기능의 이상이나 어지럼증의 원인을 파악하기 위하여 시행하는 검사로, 비자발적이고 주기적인 안구운동인 안진을 관찰하여 이루어진다[1]. 이 검사는 타 측정법과는 달리 비침습적으로 이루어지기 때문에 최근 널리 사용되고 있다. 비디오안구운동검사는 환자의 자세를 변경하여 자극을 일으킨 후 발생하는 안진을 고글에 장착된 적외선 카메라로 기록된 동영상을 통하여 의사가 판단하는 방식으로 이루어진다. 이러한 검사는 전문적인 지식을 요구하며 특히 안진의 유형을 구분하는 작업은 쉽지 않다. 따라서 검사를 보다 빠르고 정확하게 시행할 수 있는 자동화된 방법을 개발할 필요가 있다. 본 연구에서는 동공을 추적하는 알고리즘을 설계하여 비디오안구운동검사의 자동화를 위한 기반을 마련한다.

비디오안구운동검사는 비침습적이고 설계가 간단하다는 장점이 있지만 눈의 깜박임이나 검사 시 움직임으로 발생한 동잡음으로 인하여 동공 추적의 정확도가 낮아지는 경향이 있다. 안진에 관련한 기존의 연구에서는 동공의 중심을 찾기 위해 영상을 이진화한 후 수학적 방법인 허프 변환을 적용하는데 그쳤다[2-4]. 이러한 방법은 눈꺼풀로 인해 중심의 위치가 왜곡되는 등의 한계점이 있다. 이를 해결하기 위해 본 연구에서는 딥러닝을 도입하였고, 딥러닝 모델의 한계점을 보완하기 위하여 위치 추적 알고리즘을 적용했다. 본 연구에서는 동공을 딥러닝으로 분할하는 기존 연구인 DeepEye 모델과 새로운 데이터셋으로 훈련된 U-Net을 비교하고[5], 신호 분석을 통해 위치 추적 알고리즘의 효용성을 평가했다.

2. 연구 방법

실험에 사용한 데이터셋은 딥러닝의 훈련 시 사용하는 데이터셋과 알고리즘의 평가 시 사용하는 데이터셋으로 구성되어 있다. 딥러닝의 훈련 시 사용하는 데이터셋은 길이가 약 10초인 165개의 비디오안구운동검사 적외선 비디오에서 선별한 프레임으로 구성되어 있다. 알고리즘의 평가 시 사용하는 데이터셋은 길이가 약 2분에서 10분 사이의 길이인 8개의 적외선 비디오로 구성되어 있다. 모든 데이터는 연세대학교 원주 세브란스 병원에서 획득했다. 본 연구는 연세대학교 원주의과대학 생명윤리심의위원회의 승인(No. CR319082)과 모든 실험 대상자로부터 서면동의를 받은 후 진행하였다.

딥러닝에서 정답을 의미하는 라벨은 4명의 연구원이

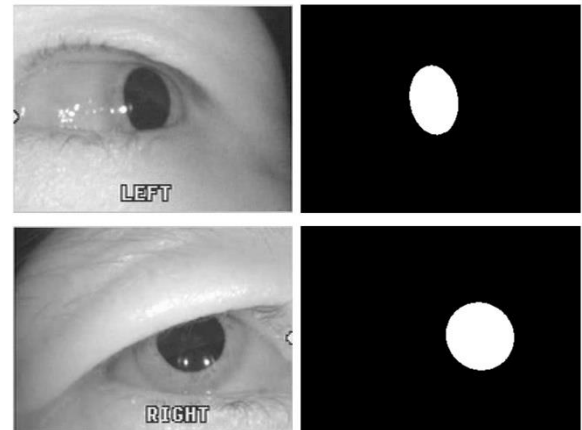


그림 1. 획득한 데이터와 라벨의 예시

자체적으로 구축한 프로그램을 사용하여 생성했다. 프로그램은 사용자가 동공의 가장자리에 최소 10개의 점을 지정하면 RANSAC 알고리즘을 통해 자동적으로 타원을 근사하고 화면에 표시하는 방식으로 작동한다[6]. 사용자가 생성된 타원이 동공에 일치한다고 확인하면 프로그램이 자동으로 타원의 내부를 동공이라고 표현하는 라벨을 생성한다. 해당 프로그램을 통해 총 7,834의 프레임을 획득하였고, 교차검증을 위해 이를 7:1:2의 비율로 나누어 각각 딥러닝 모델의 training, validation, test에 사용했다. 획득한 데이터의 예시는 그림 1에 표시했다.

동공 분할에 사용한 딥러닝 모델은 U-Net으로[7], 섬세한 형태를 예측하는 데에 우수한 성능을 보이기 때문에 의료영상의 분할에 널리 사용하는 구조다. 동공 영상은 회색조이므로 MRI나 CT와 같은 의료영상과 유사하고 동공의 정확한 중심을 찾고자 하므로 U-Net이 가장 적합하다고 판단했다. 모델에 입력하는 영상은 256x256 픽셀로 조정했으며 평균 0.5와 표준편차 0.5로 정규화했다. 훈련 시에는 입력 영상에 -10° 에서 10° 사이의 기울기와 0.8에서 1.2 비율의 크기 조정, 0.5 비율의 위치 조정을 임의로 적용했다. Adam optimizer에 0.001의 학습률을 적용했으며 batch size는 8, epoch는 100을 사용하고 validation의 loss가 가장 낮은 epoch의 모델을 최종 test에 사용했다.

딥러닝 모델의 결과로 동공을 추적하기 위해서는 동공의 위치 추적 알고리즘이 필요하다. 모델의 출력에서 동공이 완전히 추정되지 않은 경우 중심의 위치가 왜곡될 수 있기 때문에 타원 근사 알고리즘을 통해 중심의 위치를 찾았다. 또한 안진 영상의 가쪽에 생기는 그림자는 동공의 명도와 유사하여 딥러닝 모델이 혼동하는 경우가 있다. 이는 동공에 비해 면적과 위치가 확연히 다르다는 특성을 가지고 있어 프로그램상으로 이전에 판별한 동공의 면적과 위치를 저장하여 비교하는 과정을 통해 제거했다. 딥러닝 모델이 동공을 찾지 못한 프레임과 알고리즘 상으로 제거된 위치는 선형 보간법으로 값을 추정했다.

3. 연구 결과

딥러닝 모델의 분할 성능을 확인하기 위하여 같은 데이터를 DeepEye와 U-Net 모델에 입력했다. 실험 결과 U-Net의 Dice coefficient는 93.5%, DeepEye는 7.7%로 U-Net이 더 높은 결과를 보였다. 두 모델의 결과에 대한 예시는 라벨을 오버레이하여 그림 2에 표시했다. 이는 통상적인 시선 추적 데이터와 안진 검사 데이터의 도메인이 다르며 U-Net이 현재 데이터셋에 더욱 적합하게 훈련되었기에 일어난 차이로 보인다.

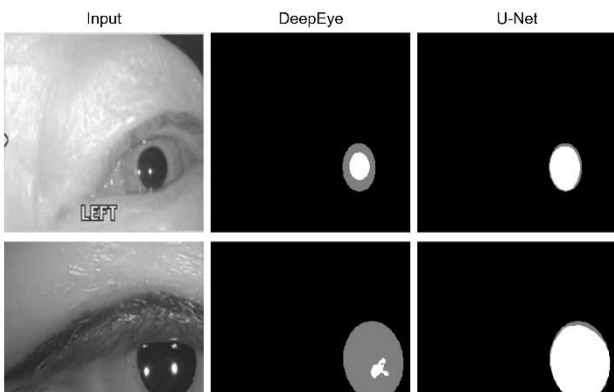


그림 3. 입력에 대한 DeepEye와 U-Net의 출력 비교

다음으로 위치 추적 알고리즘을 평가하기 위해서 원 신호와 알고리즘을 적용한 후의 신호를 비교했다. 원 신호는 딥러닝 모델의 예측 결과 중 가장 큰 면적을 가지는 영역을 선택하고 그 중심을 동공의 중앙이라고 정의하여 산출했다. 두 신호의 예시 영상은 그림 3에 표시했다. 분석 결과 원 신호에서 발생하는 노이즈가 위치 추적 알고리즘의 결과 신호에서는 제거되는 모습을 확인할 수 있었다. 이는 위치 추적 알고리즘이 정상적으로 작동하며 노이즈를 제거하는데 효과적임을 의미한다.

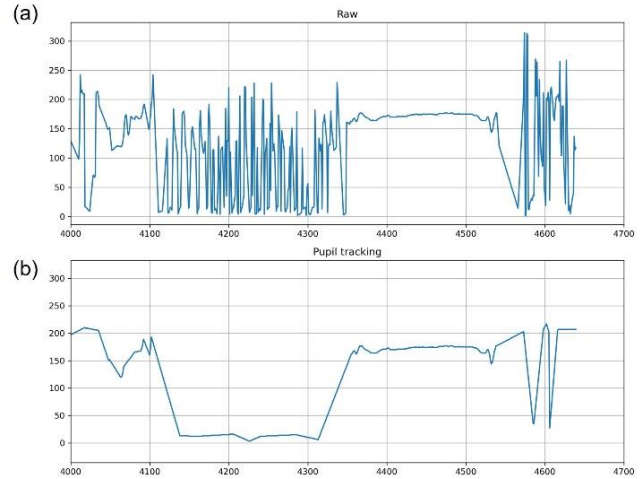


그림 2. 원 신호(a)와 위치 추적 알고리즘의 결과(b) 비교

안진 검사 데이터에서는 통상적인 시선 추적 알고리즘의 성능이 떨어지므로 데이터를 추가하여 학습한 딥러닝 모델이 필수적임을 확인할 수 있었다. 안진 검사는 평상시와는 달리 암전 상황에서 측정하기 때문에 동공이 확대되어 발생하는 차이로 보인다. 위치 추적 알고리즘은 기존의 동공에 대한 정보를 활용하여 딥러닝 모델에서 나오는 출력이 비정상적일 경우 위치를 혼동하는 현상을 예방하였다. 본 연구에서 개발한 동공 추적 알고리즘은 비디오안구운동검사 영상으로부터 동공의 위치를 데이터화하고, 더 나아가 전정신경기능 이상의 진단을 자동화하는 데에 활용할 수 있을 것으로 기대된다.

4. Acknowledgements

이 연구는 National Research Foundation of Korea (NRF) 과제의 지원을 받아 수행하였음. (NRF-2019R1F1A1058971).

5.참고 문헌

- [1] R. W. Hertle, "Nystagmus in infancy and childhood: characteristics and evidence for treatment," *American Orthoptic Journal*, vol. 60, no. 1, pp. 48-58, 2010.
- [2] W. Zhang *et al.*, "Deep learning based torsional nystagmus detection for dizziness and vertigo diagnosis," *Biomedical Signal Processing and Control*, vol. 68, p. 102616, 2021.
- [3] S. Reinhardt, J. Schmidt, M. Leuschel, C. Schüle, and J. Schipper, "VertiGo—a pilot project in nystagmus detection via webcam," *Current Directions in Biomedical Engineering*, vol. 6, no. 1, 2020.
- [4] E.-C. Lim *et al.*, "Developing a diagnostic decision support system for benign paroxysmal positional vertigo using a deep-learning model," *Journal of clinical medicine*, vol. 8, no. 5, p. 633, 2019.

- [5] F. J. Vera-Olmos, E. Pardo, H. Melero, and N. Malpica, "DeepEye: Deep convolutional network for pupil detection in real environments," *Integrated Computer-Aided Engineering*, vol. 26, no. 1, pp. 85-95, 2019.
- [6] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381-395, 1981.
- [7] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015: Springer, pp. 234-241.