

객체 탐지 및 호모그래피 추정을 이용한 안저영상 자동 조정체계 시스템 연구

인상규*, 범정현**, 추현승***

*성균관대학교 인공지능학과

**성균관대학교 소프트웨어플랫폼학과

***성균관대학교 컴퓨터공학과

e-mail : *insg1993@skku.edu, **bumjh@skku.edu, ***choo@skku.edu

A Study on Automatic Alignment System based on Object Detection and Homography Estimation

Sanggyu In*, Junghyun Beom**, Hyunseung Choo***

*Dept. of Artificial Intelligence, Sungkyunkwan University

**Dept. of Software Platform, Sungkyunkwan University

***Dept. of Computer Science and Engineering, Sungkyunkwan University

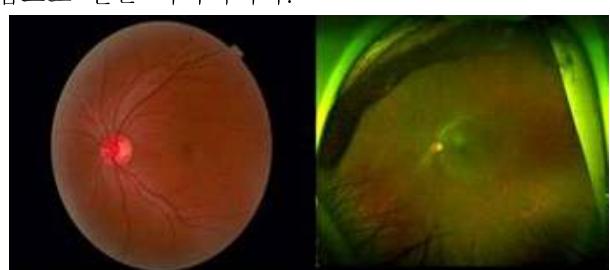
요약

본 시스템은 같은 환자로부터 촬영한 기존 안저영상과 초광각 안저영상을 Paired Dataset으로 지니고 있으며, 영상의 크기 및 해상도를 똑같이 맞추고, 황반부와 신경유두 및 혈관의 위치를 미세 조정하는 과정을 자동화하는 것을 목표로 하고 있다. 이 과정은 황반부를 중심으로 하여 영상을 잘라내어 이미지의 크기를 맞추는 과정(Scaling)과, 황반부를 중심으로 잘라낸 한 쌍의 영상을 포개었을 때 황반부, 신경 유두, 혈관 등의 위치가 동일하도록 미세조정하는 과정(Warping)이 있다. Scaling Stage에선 기존 안저영상과 초광각 안저영상의 촬영범위가 현저하게 차이나기 때문에, 황반변성 부위를 잘 나타내도록 사전에 잘라낼 필요가 있으며, 이를 신경유두의 Object Detection을 활용할 예정이다. Warping Stage에선 동일한 위치에 같은 황반변성 정보가 내포되어야 하므로 규격조정 및 위치조정 과정이 필수적이며, 이후 안저영상 내의 특징들을 매칭하는 작업을 하기 위해 회전, 회절, 변환 작업 등이 이루어지며, 이는 Homography Estimation을 통하여 이미지 변환 matrix를 구하는 방법으로 진행된다. 자동조정된 안저영상 데이터는 추후에 GAN을 이용한 안저영상 생성모델을 위한 학습데이터로 이용할 예정이며, 현재로선 2500쌍의 데이터를 대상으로 실험을 진행중이지만, 최종적으로 3만 쌍의 안저영상 데이터를 목표로 하고 있다.

1. 서론

안구는 시각을 담당하는 시신경이 분포한 중요한 기관으로, 시각세포의 질환, 대표적으로 황반변성을 정확하게 진단되어야 할 필요가 있으며, 이를 위해 크게 두 가지 방법으로 안저영상을 촬영하게 된다. 한 가지는 산동과정을 통해 환자의 동공을 확장한 후 눈에 직접 가시광선을 비추어 촬영하는 방식인 기존 안저영상(Standard fundus photography) 촬영 방법이 있다. 해당 방식은 동공확장의 영향으로 인하여 안저영상을 촬영한 후 환자에게 몇 시간 동안 앞을 보기 힘들다는 부담을 주며, 촬영 각도인 시야각 또한 45도 정도로 망막 전체의 15% 부분만 촬영 가능하다는 단점이 있다. 하지만 가시광선으로 촬영하기에 색상 그대로 안저영상을 획득할 수 있으며, 선명하게 촬영 가능하여 황반변성 부위를 보다 명확하게 알아낼 수 있다. 다른 한 가지는 최근 몇 년간 개발된 초광각 안저영상(Ultrawide field fundus photography) 촬영방법이 있다. 기존 안저영상의 촬

영법과는 다르게 무산동으로 촬영이 가능하여 환자에게 부담을 덜어주며, 시야각 또한 200도 가량 넓어 망막 전체의 80%를 촬영 할 수 있다는 장점이 있다. 하지만 가시광선을 사용한 촬영이 아닌 합성 이미지 이기에 녹색계통의 색상 변환이 주로 발견되며, 황반부 이미지의 왜곡이나 인공물(Artifacts)의 생성으로 인하여 황반부나 주변부 병변들의 감별진단이 비교적 어렵다. 그림 1은 동일 인물의 안 구를 기존 안저영상(좌)과 초광각 안저영상(우)의 촬영방법으로 얻은 이미지이다.



(그림 1) 기존 안저영상과 초광각 안저영상

기존 안저영상은 황반변성 부위를 정확하게 찾아낼 수 있다는 장점이 있으며, 초광각 안저영상의 경우 간편한 촬영방법으로 넓은 시야각의 안저영상을 얻어낼 수 있다는 장점이 있다. 위 장점을 통합하기 위해, 간편하게 촬영 가능한 초광각 안저영상만을 촬영한 후, 생산적 적대 신경망(Generative Adversarial Network, GAN)[1]을, 그중에서 Pix2Pix[2]를 참조하여 황반변성 부위가 잘 반영된 기존 안저상을 생성해내는 ‘안저영상 생성모델’을 최종 목표로 잡고 있으며, 내부에 학습데이터 생성을 위한 자동 조정 체계(Auto Alignment System, AAS)

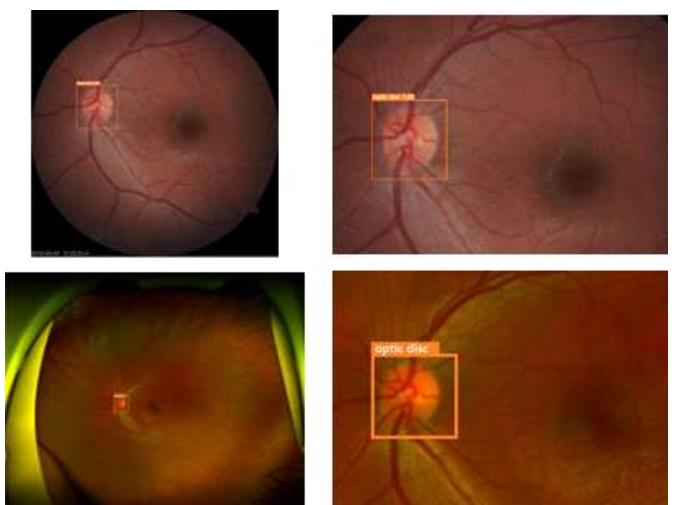
안저영상 생성모델을 학습하는데 있어서 동일인물의 기존 안저영상과 초광각 안저상을 Paired Dataset으로 하여 딥러닝을 수행할 필요가 있다. 위 과정에서 기존 안저영상과 초광각 안저영상간의 이미지 크기 및 해상도에서 차이가 있어, 원본을 그대로 학습하게 되면 좋은 결과를 얻을 수 없다. 이러한 이유로 딥러닝 학습 이전에 학습데이터를 사전에 조정해야 할 필요가 있으며, 수많은 데이터를 자동으로 처리하기 위해 고안한 연구가 본 논문의 주제인 안저영상 대상 자동 조정(Auto Alignment for Fundus Photography)라 할 수 있다. 이번 논문은 AAS의 두 번째 단계인 ‘Warping Stage’에 관하여 초점을 두고자 한다.

2. 관련 연구

GANs를 기반으로 한 안저영상 생성모델의 딥러닝을 진행하기 위해 기존안저영상과 초광각 안저영상의 Paired Dataset이 필요하며, 두 영상을 비슷하게 정합하기 위해 AAS 개념을 도입한다. AAS는 크게 두 단계로, 두 이미지의 크기와 위치를 비슷하게 맞추는 Scaling Stage와 특징점들을 비교하여 미세조정을 하는 Warping Stage로 나뉜다. Scaling Stage에서는 Object Detection 기법을 이용하고자 Faster R-CNN[3] 및 YOLOv3[4]를 활용하고자 하며, Warping Stage에서는 두 영상의 Feature Matching을 통해 Homography Matrix를 산출하고 이를 기준으로 정합과정을 진행하고자 한다. 이때 Homography Matrix를 구하는 방식은 총 3가지로, SIFT나 SURF와 같은 Traditional 한 접근법이 있으며, Supervised[7]/Unsupervised[8] 딥러닝을 기반으로 한 Homography Matrix Estimation 방법이 있다.

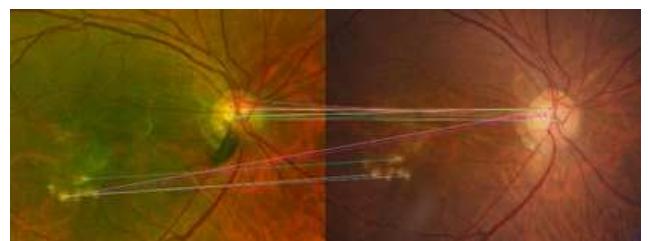
3. Object Detection 기능을 활용한 Scaling Stage

AAS는 크게 두 단계로 이루어져 있다. 그중 첫 번째 단계인 ‘Scaling Stage’에서는 Object Detection 기술을 활용하여 신경유두를 검출하고, 검출된 Bounding Box의 정보를 이용하여 두 영상을 잘라내고 크기와 위치를 대략적으로 정합하는 과정을 하고 있다. 본래 목적인 황반부위를 직접 찾기에는 성능이 낮아, 비교적 특징이 두드러지는 신경유두를 먼저 검출하고, 해당 안저영상이 좌안(O.S.)인지 우안(O.D.)인지의 해당 영상의 Tag를 확인하여 신경유두를 기준으로 황반부위가 위치한 쪽을 기준으로 넓게 영상을 잘라낸다.



(그림 2) 신경유두 검출결과를 통한 Scaling Stage과정

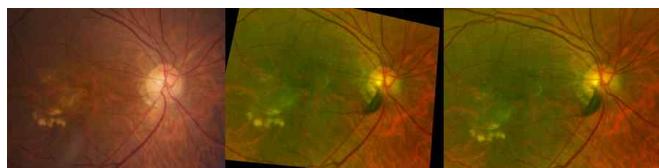
그림 2는 YOLOv3를 이용하여 신경유두가 검출된 결과 및 이를 기준으로 영상을 잘라내는 Scaling 과정을 나타내고 있다. 상단이 기존 안저영상, 하단이 초광각 안저영상이며, 좌측이 영상 원본, 우측이 Scaled된 이미지이다. 신경유두 뿐만 아니라 황반변성부위나 Artifact를 신경유두 Object로 인식하여 검출하는 경우가 있으나, Objectness Score를 비교하여 신경유두 Object가 무엇인지 판별이 대부분 가능하다.



(그림 3) Feature Matching 결과화면(Akaze)

4. Homography Matrix을 적용한 Warping Stage

AAS의 두 번째 단계인 ‘Warping Stage’에서는 황반부, 신경 유두, 혈관 등의 위치가 동일하도록 정합하는 과정이 진행되는데 이를 Homography Matrix를 이용하여 해결해보고자 한다. Traditional한 Feature Matching 방법으로는 대표적으로 SIFT, SURF가 있으며, 그 외에도 ORB, Akaze 방법 등등이 있다. 그림 3의 경우 Traditional Feature Matching의 결과화면 예시 중 하나로, Akaze 기법을 이용하여 얻은 결과화면이다. 기존 안저영상과 초광각 안저영상의 매칭되는 특이점을 연결한 결과가 다음과 같이 나타나며, 이 결과값을 이용하여 Homography Matrix를 산출하고, Matrix 연산을 통하여 그림 4와 같이 Warping 작업이 진행된다.



(그림 4) Homography Matrix를 이용한 정합과정

좌측영상이 기존안저영상, 우측 영상이 초광각 안저영상 Paired Data를 Scaling 과정까지 마친 영상이며, Warping Stage를 거쳐 우측의 초광각 안저영상을 좌측의 기존안저영상에 맞게 Homography 추정을 통해 회전 및 회절되어 정합된 결과를 가운데 영상에서 나타내고 있다.

5. 결론 및 향후계획

Scaling Stage에 집중했던 예전 결과에 비하여, 혈관, 신경유두, 황반변성부위의 위치의 미세조정 과정을 Warping Stage에서 다룰 수 있었다. 다만, Traditional Feature Matching 방법으로는 안저영상에서 Feature가 제대로 Matching되지 않는 영상들이 40% 이상 존재한다. 이를 보다 많은 안저영상 대상으로 Homography Matrix를 산출하기 위해서는 Deep Homography Estimation, 즉 딥러닝 기법을 적용할 필요성이 있다. 기존 안저영상과 초광각 안저영상은 다른 domain의 영상이며 별도의 Homography Matrix ground Truth가 존재하지 않기 때문에 딥러닝을 활용하게 된다면 Unsupervised Learning기법을 적용해야 하며, 충분한 성능향상 가능성을 기대 가능하다.

ACKNOWLEDGEMENT

본 연구는 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원(No.2019-0-00421, 인공지능대학원지원(성균관대학교)), 과학기술정보통신부 및 정보통신기획평가원의 지역지능화혁신인재양성(Grand ICT 연구센터) 사업(IITP-2021-2015-0-00742), 과학기술정보통신부 및 정보통신기획평가원의 글로벌핵심인재양성지원사업(2019-0-01579)의 지원을 받아 수행된 연구임

참고문현

- [1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, and Y. Bengio, "Generative adversarial nets." *Advances in neural information processing systems*, pp. 2672–2680, 2014.
- [2] P. Isola, J. Y. Zhu, T. Zhou and A. A. Efros, "Image-to-image translation with conditional adversarial networks", *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.
- [3] Ren, S., He, K., Girshick, R., and Sun, J., "Faster r-cnn: Towards real-time object detection with region proposal networks", In *Advances in neural information processing systems*, pp. 91–99, 2015.
- [4] Redmon, J., & Farhadi, A, "Yolov3: An incremental improvement". *arXiv preprint arXiv:1804.02767*, 2018.
- [5] Ng, P. C., & Henikoff, S. "SIFT: Predicting amino acid changes that affect protein function." *Nucleic acids research*, 31(13), pp. 3812–3814, 2003.
- [6] Bay, H., Tuytelaars, T., & Van Gool, L, "Surf: Speeded up robust features." In *European conference on computer vision*, pp. 404–417, 2006.
- [7] DeTone, D., Malisiewicz, T., and Rabinovich, A. "Deep image homography estimation." *arXiv preprint arXiv:1606.03798*, 2016.
- [8] Nguyen, T., Chen, S. W., Shivakumar, S. S., Taylor, C. J., and Kumar, V. "Unsupervised deep homography: A fast and robust homography estimation model." *IEEE Robotics and Automation Letters*, 3(3), 2346–2353, 2018