

딥러닝 모델을 통한 혀 움직임 추출 방법 개발

김근호

한국 한의학 연구원

rkim70@kiom.re.kr

Development of a Deep Learning-Based Method for Tongue Motion Extraction

Kim Keun Ho

Korea Institute of Oriental Medicine

요약

혀는 다양한 신경과 근육으로 구성되어 있으며, 이들은 뇌 및 신경계와 긴밀하게 연결되어 있다. 따라서 혀의 움직임이 비정상적이거나 제한되는 경우, 뇌졸중이나 신경계 질환, 근육 이상을 시사할 수 있다. 한편, optical flow는 각 픽셀의 움직임을 벡터 형태로 표현하므로, 모든 픽셀에 대해 이동 벡터를 지정해야 한다. 이로 인해 수백만 개의 픽셀에 대한 수작업 레이블링이 필요하게 되며, 이는 막대한 시간과 노동력을 요구하는 비효율적인 작업이 된다. 이러한 한계는 실제 데이터 활용에 장애로 작용하기 때문에, 이를 보완하기 위해 기존 딥러닝 모델의 성능 평가가 중요하다. 본 연구에서는 PWC-Net, FlowNet2, GMA 세 가지 모델을 비교 분석하였다. 실험 결과, 딥러닝 기반 모델들은 특히 외곽선 추적에서 높은 성능을 보였으며, 그중 GMA 모델은 복잡한 구조를 가진 혀의 움직임 추적에 탁월한 성과를 나타냈다. 이에 따라, GMA 모델을 활용하여 U-Net으로 분할한 영역 내외부의 혀 움직임을 정밀하게 추적하는 방법을 제안하였다.

I. 서론

혀는 여러 신경과 근육으로 구성되어 있어 뇌와 신경계와 밀접하게 연결되어 있다. 따라서 혀의 움직임이 비정상적이거나 제한될 경우, 이는 뇌졸중, 신경 질환, 또는 근육 이상을 시사할 수 있다. 또한 혀의 움직임은 구강 건강과 전신 건강 상태를 반영한다. 움직임이 느리거나 제한되면 구강 내 염증이나 감염, 혀 자체 질환을 의심할 수 있으며, 혀의 떨림이나 비정상적인 부종은 파킨슨병, 갑상선 기능 저하증, 아밀로이드증과 같은 질환의 초기 징후일 수 있다 [1,2]. 혀는 음식 섭취와 소화 과정의 첫 단계를 담당하며, 이상이 있을 경우 삼킴 장애나 식도 문제로 이어질 수 있다. 또한 혀의 움직임은 발음과 언어 발달에도 중요한 영향을 미치며, 이상은 신경학적 또는 발달적 문제와 관련될 수 있다. 이에 따라 혀 움직임을 정확히 도출하는 것이 중요하다. 본 연구에서는 설진기를 이용해 혀를 촬영하고, 다양한 딥러닝 모델을 적용하여 혀 움직임을 분석하고 모델 간 성능을 비교하였다.

을 비교하여 분석하였다.

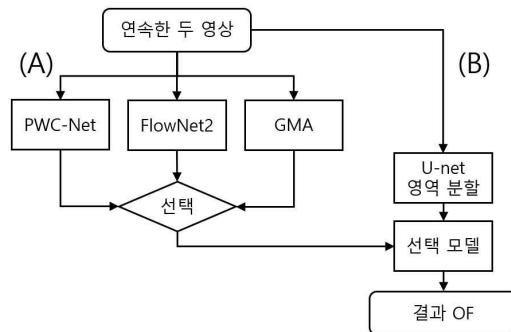


그림 1. 움직임 추출 방법 개발 과정

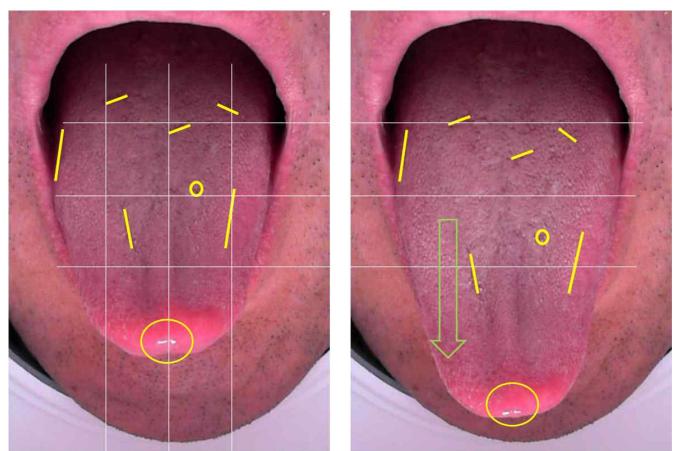


그림 2. 혀 내부 특징의 움직임

II. 본론

1. 레이블링의 한계

Optical flow (OF)는 각 픽셀의 움직임을 나타내기 위해 모든 픽셀에 이동 벡터를 지정해야 하며, 이 과정은 매우 시간 소모적이고 노동 집약적이다. 특히, 1픽셀 미만의 미세한 움직임은 인간의 눈으로 감지하기 어려워 레이블링 오류가 발생할 가능성이 높다. 현실 영상에서는 복잡한 배경과 다양한 물체들이 동시에 움직이는데, 예를 들어 바람에 흔들리는 나뭇잎이나 인체 기관인 혀처럼 부분마다 다른 움직임을 보이는 경우, 정확한 레이블링이 어렵다. 또한, 카메라의 이동이나 회전에 따른 시점 변화와 기하학적 왜곡은 픽셀 이동을 복잡하게 만들어 레이블링을 더욱 어렵게 한다. 정확한 optical flow ground truth를 얻기 위해서는 복잡한 시뮬레이션이나 고가의 특수 장비가 필요하지만, 이는 비용과 적용 범위에 한계가 있다. 이에 본 연구에서는 PWC-Net, FlowNet2, GMA 세 가지 딥러닝 모델

2. PWC-Net

PWC-Net(Pyramid, Warping, and Cost volume Network) [3]은 피라미드 구조와 워핑 기법을 활용해 프레임 간의 작은 지역적 움직임을 효과적으로 처리하는 딥러닝 모델이다. 다중 해상도의 피쳐 피라미드를 통해 낮은 해상도부터 높은 해상도로 점진적으로 움직임을 추정한다. 이러한 구조 덕분에 지역적 움직임에는 강하지만, 자연물처럼 전역적인 큰 변화를 포함하는 경우에는 성능이 떨어질 수 있다. 또한, PWC-Net은 메모리 효율성과 빠른 연산 속도를 바탕으로 실시간 응용에 적합하지만, 복잡한 전역적 움직임을 처리하는 데는 한계가 있다.

3. FlowNet2

FlowNet2 [4]는 여러 서브 네트워크로 구성되어 다양한 종류의 움직임을 효과적으로 처리할 수 있도록 설계된 딥러닝 모델이다. 빠른 움직임, 느린 움직임, 다양한 규모의 움직임을 모두 다룰 수 있으며, 일반적인 상황에서는 우수한 성능을 보인다. 그러나 자연물처럼 복잡한 전역적 움직임을 처리하는 데 있어서는 GMA(Global Motion Aggregation)보다 한계가 있을 수 있다. FlowNet2는 잔차 학습을 통해 세밀한 움직임을 예측하지만, 지역적 움직임 처리에 초점을 맞추고 있어 전역적인 복잡성에는 취약할 수 있다. 따라서 FlowNet2는 다양한 움직임에 강점을 가지고지만, 복잡한 자연물 움직임을 다루기 위해서는 추가적인 개선이 필요하다.

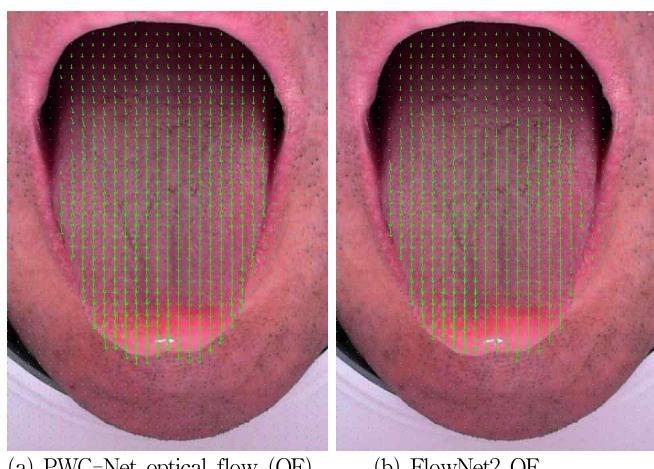
4. GMA

GMA [5]는 글로벌 컨텍스트를 활용해 이미지 전체의 움직임을 분석하는 딥러닝 모델로, 복잡하고 비정형적인 자연물의 움직임을 효과적으로 포착할 수 있다. 나무의 흔들림이나 물결처럼 복잡한 패턴 분석에 강점을 가지며, self-attention 메커니즘을 통해 각 픽셀 간의 관계를 학습하여 복잡한 움직임을 정확히 예측한다. 또한, 픽셀 간 관계를 동적으로 조정하여 움직임을 집계하는 방식을 통해 예측이 어려운 복잡한 장면에서도 뛰어난 성능을 보인다. 결과적으로, GMA는 전역적 움직임을 포착하고 복잡한 자연물 패턴을 이해하는 데 매우 효과적이다.

5. U-net 영역 분할 및 선택 모델 적용

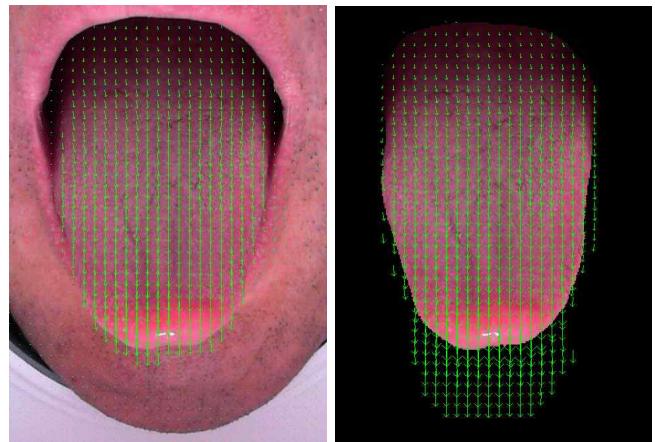
그림 1은 움직임 추출 방법 개발 과정을 보여 주고 있다. 왼쪽의 (A) 과정에서는 세 가지 모델 중 선택을 하는 과정을 보여주고 있고, (B) 과정에서는 딥러닝 방법인 U-net 방법을 이용하여 혀 영역을 분할하는 방법을 보여 주고 있다. 움직임의 정확도를 향상시키기 위해 혀 영역을 분할하여 선택된 모델을 이용하여 OF를 구하는 과정을 보여 주고 있다.

III. 결론



(a) PWC-Net optical flow (OF)

(b) FlowNet2 OF



(c) GMA OF

(d) 영역 분할된 영상의 GMA OF

그림 3. 세 개의 DL(Deep learning) model들의 optical flow

그림 2는 왼쪽에 첫 번째 영상, 오른쪽에 두 번째 영상을 배치하여 두 영상 간 특징 이동을 시각적으로 나타낸 것이다. 그림 2(a)에서 (b)로 시간 흐름에 따라 혀가 아래로 펼쳐지며 이동하는 모습이 확인되며, 혀의 특징들이 하향 이동하는 것을 볼 수 있다. 그림 3은 다양한 optical flow 기법의 결과를 비교한 것으로, (a)는 PWC-Net, (b)는 FlowNet2, (c)는 GMA 모델의 결과를 각각 보여준다. 특히, 그림 3(c)에서 GMA 모델은 혀 끝부분의 돌출 방향을 보다 정확하게 표현하여 우수한 성능을 보였다.

세 딥러닝 모델 모두 전체적인 혀 움직임 추적에는 양호한 결과를 보였으며, GMA 모델은 복잡한 구조를 가진 혀 움직임을 추적하는 데 특히 뛰어난 성능을 나타냈다. 그래서 GMA 모델이 선택되었다. (그림 1의 (A) 과정) 다음은 연속한 두 영상을 U-net으로 영역 분할하여 선택된 GMA 모델을 이용하여 OF를 구한 결과를 그림 3(d)에서 보여 주고 있다. (그림 1의 (B) 과정)

이러한 연구 결과는 단순한 색상을 가지면서도 복잡한 움직임을 보이는 객체, 예를 들어 혀와 같은 대상에 대해 새로운 적용 가능성을 제시하며, 향후 진단 자동화 기술 개발에도 기여할 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

This study was supported by Korea Institute of Oriental Medicine. (Project no.: KSN2511012)

참 고 문 현

- [1] Hirose, M. et. al. "Visualization of Parkinsonian tongue tremor on ultrasonography," *JAMA Neurol.*, pp. 1081-1082, 2022.
- [2] Je, J. T. and Lee S. K. "A clinical report of acupuncture on tongue involuntary movement," *The Journal of Internal Korean Medicine*, pp. 331-334, 2010.
- [3] D. Sun, X. Yang, M.-Y. Liu, J.n Kautz, "PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume", *CVPR*, pp. 8934-8943, 2018.
- [4] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, T. Brox, "FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks", *CVPR*, pp. 1647-1655, 2017.
- [5] S. Jiang, D. Campbell, Y. Lu, H. Li, R. Hartley, "Learning to Estimate Hidden Motions with Global Motion Aggregation", pp. 9752-9761, 2021.