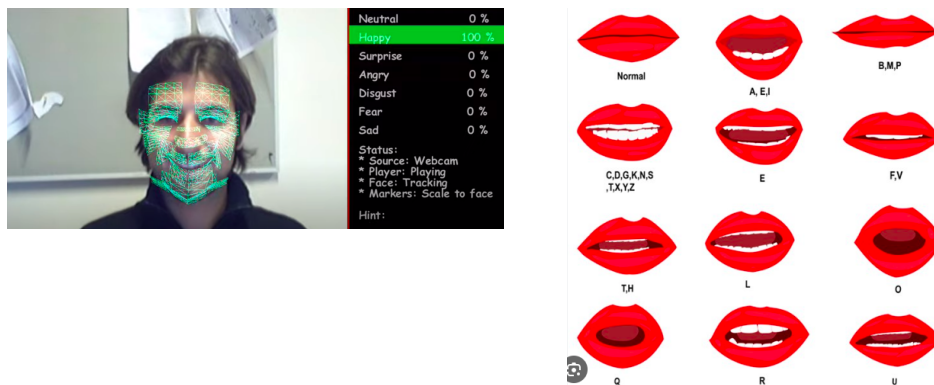


10주차 연구노트

Emotion Detection & Lip Reading

20191342 김기현, 20211428 안태현

1. 연구 주제 : Emotion Detection & Lip Reading



2. 선정 이유: 청각 장애인들은 소통 시에 상대가 전하고자 하는 뉘앙스를 잘 파악하지 못하는 경우가 많습니다. 이에 따라, 소통하는 상대의 입모양과 감정을 AI가 파악하여 상대가 말하고자 하는 뉘앙스를 사용자에게 전달하는 어플리케이션을 만들고자 합니다.

3. 기존 연구

1. DeepMind의 LipNet: Google DeepMind에서 개발한 LipNet은 2016년에 소개되었습니다. 이 모델은 비디오에서 입 모양을 분석하여 전체 문장을 인식하는 최초의 엔드-투-엔드 학습 시스템이었으며, 기존의 단어 수준의 인식 방법보다 훨씬 정확한 결과를 보여주었습니다. LipNet은 수화 데이터베이스를 사용하여 훈련되었고, 그 정확도는 인간의 리더보다도 높았습니다.

2. Oxford University의 Lip Reading in the Wild (LRW) 데이터셋: Oxford의 연구자들은 'Lip Reading in the Wild'라는 대규모 데이터셋을 구축하였습니다. 이 데이터셋은 자연스러운 환경에서 촬영된 TV 방송의 비디오 클립을 포함하며, 입술 읽기 모델의 학습 및 평가에 널리 사용됩니다.

3. Visual Speech Recognition (VSR) 시스템: 여러 대학과 연구 기관에서는 시각적 음성 인식(VSR)을 위한 다양한 시스템을 개발하고 있습니다. 이 시스템들은 주로 딥러닝 기술을 사용하여 입술의 움직임만으로 말하는 사람의 말을 인식하려고 합니다.

4. 어떤 데이터를 사용 : 입모양 데이터, 표정 데이터

- Lip reading image dataset - kaggle
- 립리딩 데이터셋 - AiHub
- 한국인 감정인식을 위한 복합 영상

5. 기대 효과 :

1. **비언어적 신호의 이해** : 표정 감지 기능은 사용자가 상대방의 감정 상태를 이해하는데 도움을 줄 수 있습니다. 이는 특히 청각 장애인에게 중요한데, 소리를 통한 감정 전달이 어려우므로 시각적 신호를 통해 감정을 인식할 수 있습니다.
2. **의사소통의 정확성 및 효율성 증가** : 입술 읽기 기능은 청각 장애인이 다른 사람의 말을 더 잘 이해할 수 있게 돕습니다. 자동화된 립리딩은 인간의 능력을 넘어서는 정확도를 제공할 수 있으며, 이는 의사소통의 오해를 줄이고 더 원활한 대화를 가능하게 합니다.
3. **사회적 상호작용의 촉진** : 표정 인식과 입술 읽기 기능이 통합된 어플리케이션은 사용자가 사회적 상황에서 더 적극적으로 참여하도록 합니다. 이 기능들은 사람들이 서로의 의도와 감정을 더 잘 파악하게 하며, 더 깊고 의미 있는 인간 관계를 쌓는 데 기여할 것입니다.
4. **교육 및 학습 도구로서의 활용** : 교육적 환경에서도 유용하게 사용될 수 있습니다. 예를 들어, 청각 장애 학생들은 교실에서 선생님의 말을 립리딩을 통해 더 잘 이해할 수 있으며, 표정을 통해 선생님의 강조나 감정을 인식할 수 있습니다.
5. **개인 맞춤형 서비스 제공** : 사용자의 반응을 기반으로 어플리케이션의 반응을 조정함으로써, 개인 맞춤형 서비스를 제공할 수 있습니다. 예를 들어, 어떤 사용자가 특정 표정이나 입모양을 더 잘 이해하는 경우, 앱이 이를 인지하고 그에 맞게 조정할 수 있습니다.