

음성인식과 안면인식을 활용한 NUI LMS

구성모*, 안인균*, 이지훈*, 문호**
서울과학기술대학교 ITM학과*
한전 KDN**

gu0112@naver.com, ingyun01@naver.com, enschs21@gmail.com,
maskgold@naver.com

NUI LMS using Webcam & Mic
(Natural User Interface Learning Method System)

Seong-mo Gu*, In-kun Ahn*, Ji-hoon Lee*, Ho Moon**
Dept. of ITM, Seoul National University of Science and Technology*
KEPCO Knowledge Data & Network Co.,Ltd.**
요 약

최근 코로나 관련 온라인 강의가 늘어남에 따라 적은 도구를 이용한 온라인과 오프라인 강의의 장점은 부각, 단점은 보완한 새로운 LMS가 필요함. 웹캠과 마이크를 이용하여 수강자의 수강태도를 파악 후, 수강자의 수업태도를 향상시키는 시스템임.

1. 서론

최근 코로나 사태와 관련하여 온라인 강의의 활용성이 증가. 하지만 온라인 강의는 오프라인 강의와는 다르게 몰입도가 낮고 타인과의 동시학습을 통한 분위기 조성 불가 및 비교를 통한 학습정도가 높지 않음. 온라인과 오프라인 강의의 장점은 강조하고 단점은 보완한 새로운 LMS의 필요성 대두. 본 시스템은 웹캠과 마이크를 활용하여 수강자의 행동 데이터를 얻어 수강자의 학습 태도를 분석하는 시스템임.

2. 작품의 개발 배경 및 필요성

1) 기존 LMS 분석 연구의 복잡함과 높은 수준의 장비 요구
- 기존의 LMS 태도 분석 연구는 뇌파 감지, 토비 아이 트래커 등 고가의 장비들을 필요로 하며 많은 전문성 장비가 요구됨, 국내 및 해외에서 학습자 태도 분석에 대한 연구는 진행되고 있지만 실용화되지 않음.
- 본 프로젝트는 기존의 노트북 기능(마이크, 웹캠)만으로 수강자의 태도 데이터를 수집 및 분석.

2) 오프라인 학습과 유사한 환경 조성
- 학생들의 집중력 강화와 대리 출석 방지를 위한 얼굴인식 기능을 활용.
- 눈 비율 정도, 감정, 표정 등을 이용한 집중도 흐름 파악 집중도 검사 등을 이용해 데이터 수집 후 모델 설립. 이를 이용해 수강태도를 분석하여 점수부여 및 온라인 강의 효과 향상.
- 강의를 듣지 않고도 들은 것처럼 조작 및 강의를 실행시키고 영상플레이 시간동안 다른 행동하는 경우가 빈번함. 이는 곧 수강자의 지식습득 정도와 연관되기 때문에 강제로라도 집중도를 높여줄 시스템이 필요함.

3. 작품의 특징 및 장점

1) 적은 장비로 최대한의 효과
- 웹캠과 마이크를 이용하여 일정수준 수강자의 행동을 제한하고, 집중도를 시각적으로 나타내 주어 온라인 강의임에도 오프라인 강의와 같은 효과를 보여줄 수 있는 웹 시스템임.

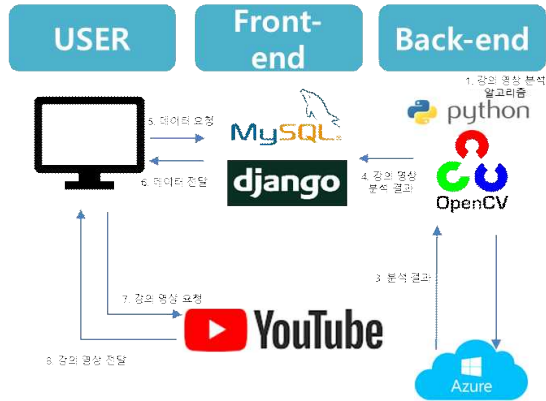
2) 오프라인 학습과 유사한 환경 조성
- 등록된 수강자 안면과 일치하지 않거나 수강자의 안면이 인식되지 않으면 강의 재생 불가.
- 수강자의 실시간 눈의 비율을 계산하여 수강자의 졸음 여부를 판단.
- 수강자의 얼굴을 분석하여, 감정과 목과 얼굴의 각도, 미소, 졸음 정도를 파악하여, 수강자에게 그래프로 시각화 하여, 강의 중 수강자의 학습 태도를 보여줌. 이는 수강자뿐만 아니라 강사에게도 좋은 지표가 될 것이라 예상함.
- 수강자의 현재 점수를 직접적으로 알 수 있게 초록(정상), 노랑(경고1), 주황(경고2), 빨강(경고3)으로 나타내고 현재 태도점수를 직관적으로 알 수 있게 함.

3) 기존 기술 활용과 차별성
- 얼굴 인식 기술은 Azure Face API, 음성인식은 Google Speech Recognition API, 졸음 인식은 깃허브 오픈 소스에서 알고리즘을 가져와 사용 중에 있음 [1].
- 기본적인 웹 구성은 장고(Django)를 이용하여 제작. Database는 MySQL을 이용.

4. 작품 내용

1) 서비스 구성도 (그림 1)
- Back-end
1. 강의 영상 분석 알고리즘: 영상 분석 알고리즘으로 동영상 분석 진행
2. 이미지(동영상): Azure Face API사용 위해 이미지 전달
3. 분석 결과: 이미지를 분석한 결과 값 전달
4. 강의 영상 분석 결과: 이미지를 분석한 결과 값인 임계 값, 표정, 감정 등을 저장

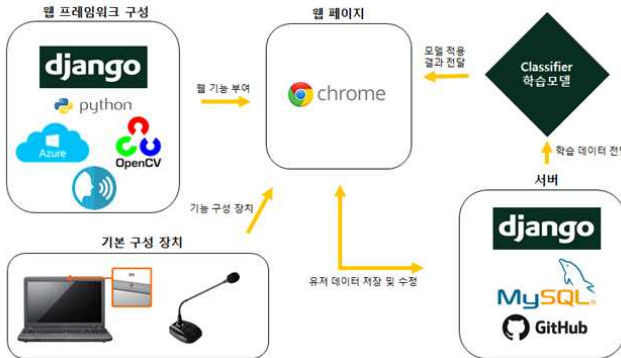
- Front-end
- 5. 데이터 요청: 회원 정보, 홈페이지 정보, 강사 정보, 강의 정보 등 데이터 요청
- 6. 데이터 전달: 요청 데이터 전달
- 7. 강의 영상 요청: 강의 영상 요청
- 8. 강의 영상 전달: 요청받은 영상 데이터 전달



(그림 1) 서비스 구성도

2) 시스템 구성도 (그림 2)

- Back-end



(그림 2) 시스템 구성도

3) 유즈케이스

유즈케이스 ID	유즈케이스	하위 유즈케이스	액터	대상
Uc-01-001	유저정보관리	회원가입 기능	학생	WEB
Uc-01-002	유저정보관리	로그인 기능	학생	WEB
Uc-01-003	유저정보관리	학습 기록 조회	학생	WEB
Uc-01-004	유저정보관리	학습 결과 다운 로드	학생	WEB
Uc-02-001	학습설정	테스트 창 진입	학생	WEB
Uc-02-002	학습설정	마이크 테스트	학생	WEB
Uc-02-003	학습설정	웹 캠 테스트	학생	WEB
Uc-02-004	학습설정	눈 일계 값 계산	학생	WEB
Uc-02-005	학습설정	얼굴 사진 촬영, 등록	학생	WEB
Uc-02-006	학습설정	테스트 결과	학생	WEB
Uc-03-001	강의 목록	동영상 학습 목록	학생	WEB
Uc-03-002	강의 목록	동영상 학습 선택	학생	WEB
Uc-04-001	동영상 학습	얼굴인식 확인	학생	SERVER
Uc-04-002	동영상 학습	동영상 플레이	학생	WEB SERVER
Uc-04-003	동영상 학습	졸음 인식 및 방지 경고음	학생	SERVER
Uc-04-004	동영상 학습	졸음 상태 신호 등 표시	학생	WEB
Uc-04-005	동영상 학습	눈 크기 실시간 그래프	학생	SERVER

Uc-04-006	동영상 학습	음성 명령어 영상 제어	학생	WEB
Uc-04-007	동영상 학습	동영상 러닝시간 기록	학생	SERVER
Uc-04-008	동영상 학습	음성 집중도 검사	학생	SERVER
Uc-04-009	동영상 학습	집중도 검사 음성대답	학생	WEB
Uc-04-010	동영상 학습	강의 다시 보기	학생	WEB
Uc-05-001	학습 평가	수강 태도 분석	학생	SERVER
Uc-05-002	학습 평가	분석 결과 전달	학생	SERVER

<표1> 유즈케이스

3) 주요 적용 기술

○ Google Speech Recognition API(Annyang API)

[1]

- API가 애플리케이션의 마이크로부터 스트리밍되거나 사전 녹음된 오디오 파일로부터 전송되는 (인라인 또는 Cloud Storage 사용) 오디오 입력을 처리할 때 음성 인식 결과를 실시간으로 수신.
- 강의 재생 시 수강자의 음성명령어의 인식을 위한 STT(Speech To Text)를 위한 API
- 수강자 마이크 정상작동 확인, 음성 명령어를 통한 영상제어 등을 위한 API

○ Azure Face API [2]

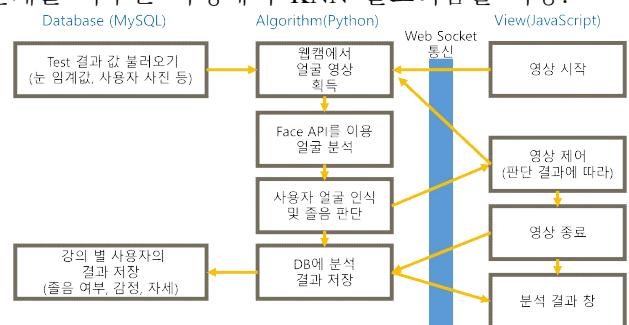
- 이미지(동영상)에서 얼굴과 특성을 인지, 별도의 DB가 필요 없고 사람의 다양한 표정을 감지한다. 얼굴감지여부, 수강자일치여부와 웃음, 화남, 공포, 중립, 슬픔 등의 감정 등을 제공.
- 이미지(동영상)에서 얼굴과 특성을 인지, 별도의 DB가 필요 없고 사람의 다양한 표정을 감지함.

○ Web Socket을 이용한 비동기 통신

- 동영상 강의 중 실시간 졸음 분석은 Backend에서 진행하고, 영상 제어 및 음성 인식은 Frontend 쪽에서 동시에 진행됨. Web Socket 통신을 통하여 비동기적으로 서로 정보를 실시간으로 교환함.

○ Drowsiness driving detection system with OpenCV & KNN 알고리즘 [3]

- 얼굴 및 안구 검출을 하기 위해 Azure Face API를 이용하여 검출
- 졸음 상태를 감지하기 위해선 Eye Aspect Ratio라는 개념을 사용.
- 마지막으로 수강자의 졸음 위험 수준을 세 단계로 나눠 단계별로 차등 알람이 울리게 하였고, 단계를 나누는 과정에서 KNN 알고리즘을 사용.



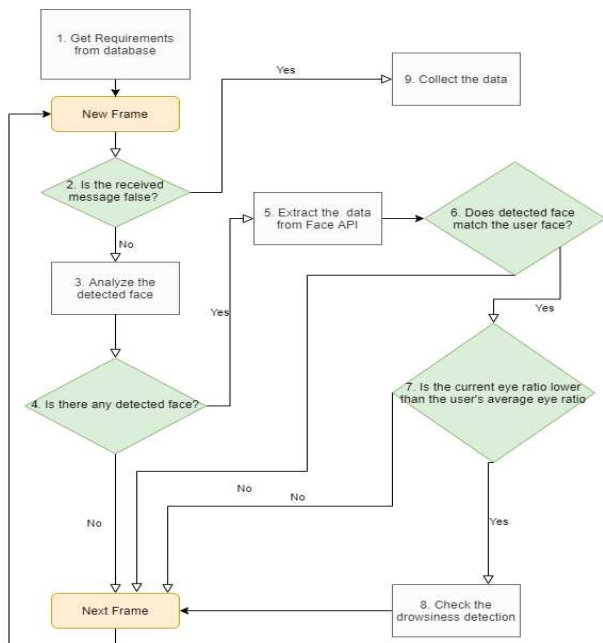
(그림 3) 강의 알고리즘 명세서

• 강의 중 얼굴 인식 알고리즘 시나리오

(그림4)

- ① 테스트 화면에서 저장한 유저 정보들을 받아온다. 등록된 유저의 face id와 눈의 일계 값(유저의 평균 눈 비율)을 받아온다.
- ② 웹 소켓을 이용하여 강의가 종료되면, false라는 메시지를 받게 된다. False 값을 받으면 반복문이 종료된다.
- ③ Azure Face API 를 이용하여 현재 프레임

- 의 얼굴을 분석한다.
- ④ 현재 프레임에 얼굴이 인식 되지 않으면, 바로 다음 프레임으로 넘어가며, 웹 소켓을 통해 강의가 정지된다.
 - ⑤ 얼굴이 인식 되어 있을 때, API에서 필요한 데이터를 추출한다. .ex) 얼굴 고유 id, 감정, 얼굴 포즈, 미소, 눈의 위치를 이용한 눈 비율
 - ⑥ 얼굴 고유 id 와 등록된 유저 얼굴의 고유 id 를 비교하여 등록된 유저가 보고 있는지 아닌지를 판단한다.
 - ⑦ 현재 눈의 비율이 등록된 눈의 임계값 보다 낮으면, 졸음이 시작 되었다고 판단한다
 - ⑧ 위의 조건 하에서, 위의 조건이 유지 되 있는 시간을 기반으로 졸음 정도를 판단한다.
 - ⑨ 강의 도중에 추출해 낸 모든 데이터를 강의가 끝난 후에 DB에 저장한다.



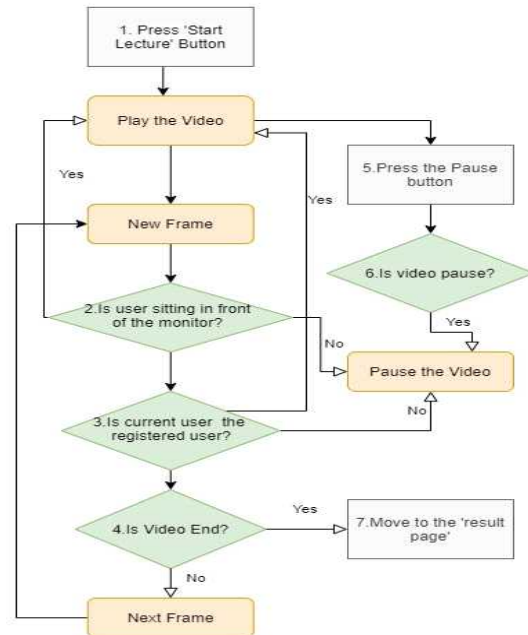
(그림 4) 강의 중 얼굴 인식 알고리즘

- 강의 중 영상 제어알고리즘 시나리오 (그림 5)

- ① 'Start Lecture' 버튼을 누르면 강의가 강의 영상이 시작되며, 파이썬 함수 'loadvideoplayer'가 실행 된다
- ② 화면 앞에 본인 얼굴이 보이지 않으면 웹 소켓을 통해 파이썬 함수에서 얼굴인식을 하여, 메시지를 보내 얼굴이 안 잡히면 영상을 일시정지, 잡히면 재생한다
- ③ 등록된 사용자와 현재 사용자의 얼굴이 일치 하지 않으면, 파이썬 함수에서 얼굴인식을 하여, 동영상을 일시 정지한다. 본인 일 경우 영상이 계속 재생된다.
- ④ 등록된 비디오가 끝나게 되면, 파이썬 함수가 종료되며, 집중도 결과 창으로 이동하게 된다.
- ⑤ 강의 재생 중 사용자가 'Pause' 버튼을 누르면 강의가 멈추게 된다.
- ⑥ 강의가 재생 중인지 아닌지 파악한다.

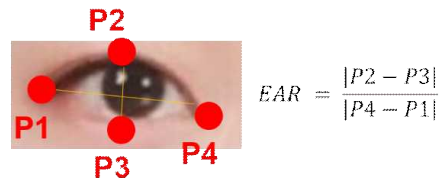
- ⑦ 강의가 종료 될 시 '집중도 결과' 창으로 이동한다

(그림 5) 강의 영상 제어 알고리즘



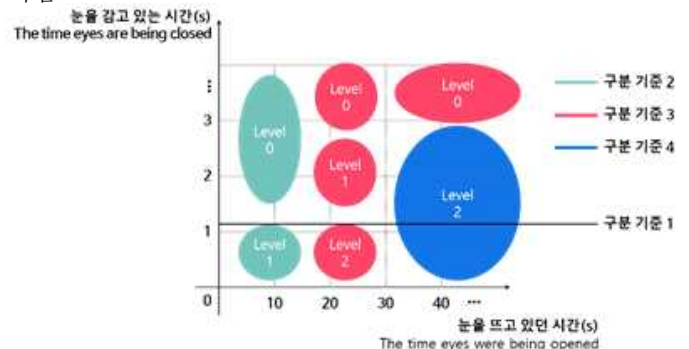
(그림 5) 강의 영상 제어 알고리즘

- 2016년 Tereza Soukupova & Jan ' Cech에 의해 제시된 Eyes Aspect Ratio(이하 EAR) 방식을 사용. EAR은 검출된 안구에 네 개의 (x, y) 좌표를 이용하여 계산. (그림 6)



(그림 6) 눈 임계 값 계산 식

- 위와 같은 식으로 졸음 수준을 세 단계로 분리함.



(그림 7) 졸음 수준 단계 분리 기준

- 구분 기준 1. 첫 번째 알람은 (약0.9초경) 졸음 단계 1~2 사이에서 울림.
- 구분 기준 2. 15초 이하로 졸음을 반복하고 있다면 졸음단계는 1부터 시작하고 연속으로 올리는 다음 알람은 0까지 올라감.

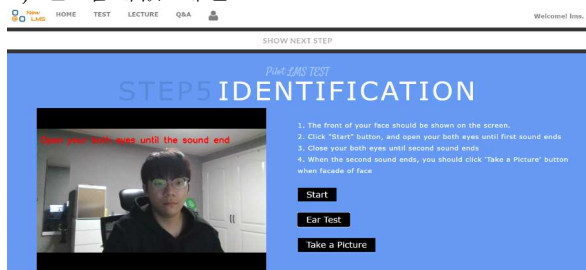
- 구분 기준 3. 15초 초과 30초 미만 간격의 졸음 운전 시 첫 번째 알람은 level2, 두 번째 알람은 level1, 세 번째 알람은 level 0이 울림.
- 구분 기준 4. 30초 이상 졸음운전을 하지 않으면 2단계로 설정.

5. 작품의 기대효과 및 활용분야

- 1) 작품의 기대효과
 - 기존의 LMS와 차별화된 시스템
 - 노트북 내장 웹캠과 마이크를 이용하여 지속적인 피드백으로 실제 오프라인 수업과 비슷한 학습 환경 조성
 - 자칫 느슨할 수 있는 온라인 수업을 마이크 및 웹캠을 통한 수강자 집중도 파악으로 반강제적 집중도 향상
 - 사용자 태도 점수를 최종적으로 부여하여 타인과의 비교를 통한 참여 유도, 출석 태도 점수에 반영하여 적극적인 수업 태도 유도 가능
- 2) 작품의 활용분야
 - 교육 시설에 도입, 웹캠과 마이크만 있으면 어디서든 활용 가능
 - 다양한 교육 시설에 도입함으로써 학생들의 수업 집중도 향상과 집중도의 수준을 확인할 수 있음.
 - 집중도의 수준 평가를 통하여, 강사 역시 학생의 수업 태도 및 집중도의 정도를 파악할 수 있음.
 - 초, 중, 고등학교 및 학원과 같은 교육 시설
 - 초, 중, 고등학교 및 학원과 같은 교육 시설에서 수업에 디지털 기술 접목 가능.
 - 기존 온라인 강의와 달리 기술 접목을 통하여, 보다 학생들의 집중도 향상 및 평가를 통하여 강의 청취율의 질을 향상시킬 수 있다고 보임.
 - 학습 관리 기법 업그레이드(학습집중 시간, 수업 충실도, 개인별 집중 시간관리)
 - 실시간 비 대면 온라인 강의 평가(수업출석분) 정량적 관리 가능.
 - 개인별 온라인 강의 특징을 반영한 맞춤형 강의 패턴 개발 ⇒ 학습 결과의 정량화로 수업 참여(예:출석)에 대한 객관적 평가 가능
 - 위와 같은 기법을 이용하여 특허 신청도 가능할 것으로 사료됨
 - 수강자가 아닌 강사 평가 가능
 - 사용자의 표정 분석을 활용하여 사용자가 영상을 시청하며 어떤 표정을 짓는지 확인할 수 있음. 강사가 수강자의 태도 점수를 판단할 수도 있지만, 그와 반대로 수강자가 강사에게 점수를 부여할 수 있음. 타 강사와 비교하여 집중도가 떨어지는 영상이 많은 강사라면 강사에게 경각심을 줄 수 있음.

6. 주요 화면 설계서

1) 눈 임계값 계산



사용자의 눈의 임계값 등록과 본인 얼굴 사진을 Database에 등록한다.

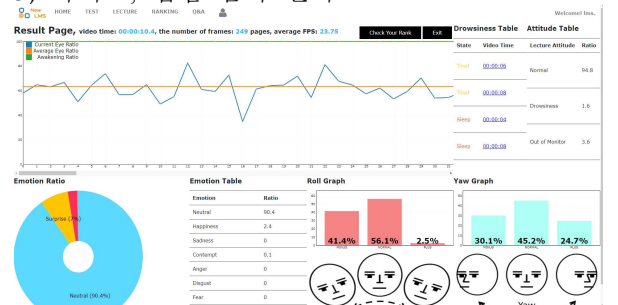
2) 실시간 강의 화면



얼굴 인식 확인, 동영상 재생, 졸음 인식 및 방지 경고음, 음성 명령어 영상제어, 동영상 러닝시간 기록.

동영상 강의를 진행 되면서, 본인이 강의를 수강하지 않거나 화면에서 사라지면 강의는 멈추며, 강의가 진행 되는 도중, 사용자의 졸음을 인식하여 환기 해주는 기능이다.

3) 마이크, 웹캠 검사 결과



선택한 강의 도중 수강자의 얼굴분석 데이터와 음성인식 데이터를 이용하여 해당자가 어떻게 들었는지를 그래프와 표로 보여준다. 수업 집중도에 따른 점수 부여 및 졸음 인식한 구간에 영상을 다시 볼 수 있다.

본 논문은 과학기술정보통신부 정보통신창의인재양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트의 결과물입니다.

참고문헌

- [1] <https://github.com/TalAter/annyang>
- [2] <https://azure.microsoft.com/ko-kr/services/cognitive-services/face/>
- [3] <https://github.com/woorimlee/drowsiness-detection>
- [4] Tereza Soukupova and Jan Cech "Real-Time Eye Blink Detection using Facial Landmarks" 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia, February 3-5, 2016