

딥러닝을 활용한 시각적 드럼 가이드 구현

오승엽¹, 김나연, 이유림, 이수빈, 장진이², 이기영³

¹인천대학교 정보통신공학과

²인천대학교 소비자 행동학과

³인천대학교 정보통신공학과

osy9907@naver.com¹, kny010901@inu.ac.kr, yurim0412@inu.ac.kr, ssubinn@inu.ac.kr,
528794@naver.com², kylee@inu.ac.kr³

Implementation of the Visual Drum Guide Using Deep Learning

Seung-Yeop Oh¹, Na-Yeon Kim, Yu-Rim Lee, Su-Bin Lee, Jin-Lee Jang²,
Ki Young Lee³

¹Dept. of Info and Telecom Eng. Incheon National University

²Dept. of Consumer & Child studies. Incheon National University

요 약

드럼 연습 및 연주를 쉽게 배우고 즐기기 위한 시각적 가이드를 제공한다. CNN과 OpenCV를 활용하여 악보의 음표 이미지를 처리하고, 데이터 수집, 전처리, 모델 학습, 서버와 안드로이드 애플리케이션 연동 등 다양한 과정으로 구현된다. '드럼라이트'는 드럼 연주의 접근성을 높이고, 다른 악기와 음악 학습 분야에도 확장 가능한 가능성을 제시한다.

1. 서론

드럼은 피아노, 기타보다 수요가 적다. 근거로 네이버 데이터랩 쇼핑 분야 트렌드 비교 분석 결과를 확인 가능하다. 2023년 07월 기준 최대 클릭 횟수를 100회로 한정했을 때 '디지털피아노'는 100회, '어쿠스틱 기타'는 92회, '전자드럼'은 16회의 비율로 확인되었다(기기별, 성별, 연령 전체 조회 시). 드럼 강습 학원, 드럼 제작 업체는 드러머들이 많아져 드럼 시장의 매출 증가를 원할 것이다. 이를 위해서 대중들이 한 번쯤 시도해보고자 하는 악기로 자리 잡아야 한다. 새로운 콘텐츠를 추가해 드럼의 즐거움을 더할 수 있다면 대중들이 주목할 것이다. 악보를 기반으로 드럼을 연타하는 타이밍에 맞춰 드럼 표면 위에 시각적 효과를 더하고자 한다. 본 논문은 드럼을 쉽게 즐기기 위한 서비스 '드럼라이트'를 제공하기 위해 OpenCV와 CNN을 활용한 방안을 서술한다.

2. 본론

2.1 CNN(Convolution Neural Network)

딥러닝은 네트워크의 구조와 계층 구성에 따라 종류가 다양하다. 악보의 음표 이미지를 처리하기 위해 CNN을 선택했다. CNN은 컴퓨터 비전 분야에서

이미지 처리 및 인식에 사용되는 딥러닝 아키텍처이다. CNN의 합성곱 레이어(Convolutional Layers)는 합성곱 레이어들이 순차적으로 쌓인 것이며 이를 통해 이미지의 작은 부분에서 패턴과 특징을 감지할 수 있다. 그리고 풀링 레이어(Pooling Layers)를 사용해 특징 맵의 크기를 줄이고 특징을 요약해 계산량을 감소시킨다. 또한 앞의 두 레이어를 계층적으로 쌓아 추상적인 정보를 표현해 다양한 이미지 처리 작업에 적응하고 복잡한 특징을 학습할 수 있다[1].

2.2 OpenCV(Open Source Computer Vision Library)

컴퓨터 비전과 이미지 처리 작업을 위한 오픈 소스 라이브러리로, 이미지, 비디오 데이터를 처리하고 분석한다. OpenCV를 활용해 이미지 데이터의 전처리, 특징 추출 등 다양한 작업을 수행가능 하며, 딥러닝 모델과도 통합 가능하다. 본 서비스는 딥러닝 시 사용되는 드럼의 음표 이미지를 악보로부터 추출 및 딥러닝 학습을 위한 이미지 전처리에 사용된다[2].

3. 시각적 드럼 가이드 구현

3.1 데이터 수집

악보 사이트 뮤즈 스코어(musescore.com)를 통해 드럼 악보 및 곡에 대한 정보와 유튜브(YouTube)를

통해 음악의 MR 및 mp3 데이터를 수집한다.

3.2. 데이터 전처리

전처리에 사용된 환경으로는 python과 OpenCV를 사용했다. 어플리케이션의 화면 상단에 악보를 표시하기 위해 보표별로 이미지를 잘라 저장한다. 또한 드럼은 cymbal과 tom에 따라 음표의 위치가 지정되어있기 때문에 보표 이미지에서 음표만 잘라내고 매칭된 번호에 맞게 음표들을 나눠 저장한다[3].



그림 1) 원본 악보의 일부분

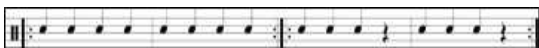


그림 2) 보표별로 자른 이미지

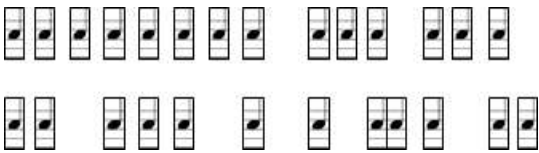


그림 3) 전처리 완료된 음표들의 개별 이미지

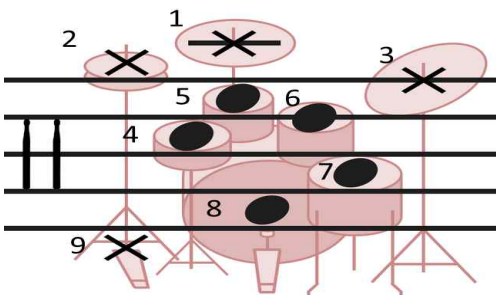


그림 4) 음표와 cymbal, tom 번호 매칭

3.3 데이터 학습 및 결과 분석

전처리된 음표 데이터를 CNN을 이용해 딥러닝을 진행했다. 사용된 이미지의 개수는 음표별로 약 3000개씩 총 27000개 정도 사용됐다. 음표를 총 9가지로 분류하기 위해 class_mode를 categorical로 설정하고 모델을 구축했다.

```
Image: D:/note_deep_learning/test#3-1.jpg, Predicted Class: 3
1/1 [=====] - 0s 25ms/step
Image: D:/note_deep_learning/test#4-1.jpg, Predicted Class: 4
1/1 [=====] - 0s 25ms/step
Image: D:/note_deep_learning/test#5-1.jpg, Predicted Class: 5
```

그림 5) CNN으로 구축한 모델로 예측한 결과

3.4 알고리즘 구현

구축된 모델과 vscode, python을 이용해 서버에서 악보 이미지를 받아오면 전처리하여 보표, 음표 이

미지를 저장한다. 모델을 통해 음표들을 예측해 txt 파일로 생성하여 서버로 보낸다.

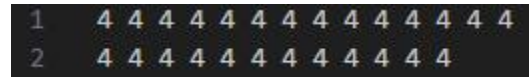


그림 6) 예측 후 생성된 txt파일

3.5 서버와 안드로이드 연동

사용자의 접근성을 위해 안드로이드 스튜디오를 활용하여 애플리케이션을 구현한다. 그 후, AWS 서버에서 동작하도록 안드로이드와 연동한다.

4. 결론 및 향후 연구 방향

본 서비스는 드럼 연주에 시각적인 도움을 더해 누구나 드럼을 재밌고 쉽게 접할 수 있도록 접근성을 높여 드럼 대중화에 목적을 둔다. 빔프로젝터를 통해 나타나는 움직이는 노드는 음악과의 조화를 통해 연주자에게 신선한 경험을 제공할 것이다. 이러한 시각적 효과를 통해 눈으로 보이지 않는 연주의 표현력을 나타내 드럼 연주에 흥미를 가져다준다.

악보 분석 알고리즘을 응용한다면 드럼뿐 아니라 피아노와 같은 다른 악기에도 적용할 수 있다. 또한 음악 학원과의 연계를 시도할 수도 있다. 본 서비스의 일일 과제 기능에 선생님이 연습량을 작성하면 모바일을 통해 실시간으로 수강생의 과제 수행 상황을 확인할 수 있다. 녹음 기능도 추가해 수행한 과제를 비대면으로도 실시간 피드백이 가능하다.

라이트 아트를 추가하여 드럼 공연에도 활용될 수 있다. 드럼 무대에 본 서비스를 활용한 라이트 아트를 추가하면 관객들에게 드럼 연주를 통한 시각적 즐거움을 더해줄 수 있다.

※ 본 프로젝트는 과학기술정보통신부 정보통신창의 인재양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.

참고문헌

- [1] 허이륜. "합성곱 신경망(CNN)기반 이미지 처리 시스템." 국내박사학위논문 배재대학교, 2018. 대전
- [2] 문지수,김민지,임영규,and 공기석. "영상 처리와 딥러닝을 이용한 악보 코드 변환 프로그램." 한국인터넷방송통신학회 논문지 21.1 (2021): 69-77.
- [3] 제승완, 이덕우. (2022). 악보의 음표 인식을 위한 형태학 기반의 오선 제거. 한국산학기술학회 논문지, 23(5), 334-340, 10.5762/KAIS.2022.23.5.334