주제

* 음원에서 특정 악기(피아노)의 음을 분리해 내어, 악보로 만들어 출력한다.
  + 원활한 학습을 위해, 보컬은 없는 음원을 사용하며, 비교적 간단한 음원(적은 악기가 사용된) 음원부터 시작한다. 딥러닝 학습의 결과가 좋다면, 추후 복잡한 음원이나 보컬이 포함된 음원에 대해서도 학습을 시도해볼 수 있다.
* 자료 획득 -> 전처리 -> 딥러닝 -> 출력의 과정을 거친다.
  + 자료 획득: 실제 음원을 이용할 수도 있고, 가상악기 프로그램을 통해 음원을 조합하여 사용할 수도 있다.
  + 전처리: mel-spectrogram, MFCC등 음원처리에 적절한 형태로 데이터를 가공한다.
  + 딥러닝: pyTorch와 같은 툴을 이용하여 딥러닝을 구현한다. 전처리한 데이터의 형태에 따라 CNN(합성곱)과 Affine(완전연결) 신경망을 적절히 이용한다.
  + 출력: 출력 결과를 악보로 가공한다. 악보 생성 프로그램을 이용할 수도 있고, 간단한 형태로 직접 악보를 생성할 수도 있을 것이다.

유의 사항 – 주제

* 음성/음악에 대한 딥러닝은 연구가 부족하여 자료가 부족하다. 특히 음악이 더 그렇다.  
  대부분의 자료는 이미지와 같은 정적인 자료에 대한 딥러닝이며, 음원/음성처럼 동적인 자료에 대한 내용은 부족하다.
  + 비디오+음원을 이용하여 음원을 악기별로 분류한 딥러닝이 존재하는데, 해당 논문을 읽어보면 도움을 얻을 수도 있다.  
    (http://sound-of-pixels.csail.mit.edu/)  
    (<https://arxiv.org/abs/1804.03160>)
  + 음성인식, 음성분류와 같은 분야의 딥러닝은 음악과는 다르다고 한다. 그러나 음원 자체를 다루는 방법(mel-spectrogram, MFCC)들은 도움이 될 수 있다. 자료가 부족하면 음성분야에서 참고해보자.
  + 보컬과 악기소리를 분류하는 딥러닝에서 힌트를 얻을 수도 있다.  
    ((<http://danetapi.com/chimera>)

유의사항 – 자료획득

* 연구가 부족하므로 학부수준에서 고성능의 딥러닝을 구현하긴 어려울 수 있다.
  + 보다 학습하기 쉬운 자료들 위주로 딥러닝을 먼저 구현한다. 성능이 좋다면 더 복잡한 자료(더 많은 악기, 보컬 포함)에 대해서도 학습을 시도할 수 있다.
* 자료 획득의 어려움이 있을 경우
  + 가상악기 프로그램을 이용하여, 음원을 직접 만들어 낼 수도 있다. 간단한 음원을 구현하는데 유용하다.  
     다만, 실제 음원의 경우 사람의 오차나 잡음 등 여러 변수가 존재하는데, 가상악기 프로그램은 너무 깨끗한 결과이므로, 이를 통해 학습할 시 실제 결과는 좋지 않을 수 있다.  
     직접 만든 음원의 사용 비율을 제한하거나, 아래 나오는 노이즈를 추가하는 방법을 쓰는 것도 괜찮을 듯하다.
  + 노이즈를 추가하여 비슷하지만 다른 데이터를 다량 확보할 수 있다. 가상악기 프로그램으로 음원을 만들었을 때도 사용하면 유용할 듯 하다.   
    (<https://engineering.linecorp.com/ko/blog/voice-waveform-arbitrary-signal-to-noise-ratio-python/>)

유의사항 – 전처리

* 일반적으로 우리가 아는 음악 그래프는, 음압(amplitude)에 대한 그래프이며, 이를 통해 딥러닝을 구현하기엔 어렵다.
  + 음성분야에서는 mel-spectrogram, MFCC등으로 자료를 변환하여 사용한다. 이러한 자료는 FFT를 이용하여 만들어지는데, FFT는 시간축이 없으므로 음악을 잘라서 변환해야한다. 일반적으로는 1프레임(20-40ms)단위로 잘라서 이용한다고 알려져 있다. 자세한 내용은 자료 추가바람   
    (<https://brightwon.tistory.com/m/11>)

유의사항 – 딥러닝

* 어떤 툴을 사용해야 할까?
  + 현재 많이 사용하는 딥러닝 툴은 pyTorch, tensorflow등이 있다. 혹시나 음원/음악 처리에 특화된 딥러닝이 있다면 추가바람
* 동적인 자료의 딥러닝에 대한 이해가 필요하다. 많이 알려진 이미지는 정적인 자료이지만, 음악은 동적인 자료라고 볼 수 있다.
  + 음성인식/분류 등의 딥러닝에서 힌트를 얻을 수 있다.
* 악기(피아노)에 대한 이해가 필요할까?
  + 피아노를 치면 얼마동안 울리는지, 그렇다면 정답을 레이블 할 때는 그 시간만큼 레이블을 해야 하는지, 사람이 연주한 음악은 오차가 존재하는데 어떻게 반영할 것인지, 기준은 어떻게 정하는지 등등  
    이 부분은 하이퍼파라미터라고 볼 수 있다. 즉 학습률과 같이 인간이 설정해주어야 하는 영역이라는 것. 따라서 이론과 여러 번의 실험을 통해 이를 정하면 될 듯 하다.
* 어떤 방법으로 딥러닝을 할 것인가?
  + 파형을 쪼개어 변환하고(mel-spectogram, MFCC), 해당 자료를 가지고 딥러닝.  
    출력은 음계+시간의 형태.
  + 파형에서 피아노의 음원만 분리하고, 해당 음원에서 다시 음계를 분리  
    사실상 두번의 딥러닝을 해야하는 것이라고도 볼 수 있다. 물론 피아노의 음원을 성공적으로 분리할 수 있다면, 두번째 음계분리는 다소 쉽기 때문에 괜찮긴하다. 문제는 음원이 복잡 해질수록, 여러 음원이 섞이기 때문에 피아노의 음원만 분리하는 것은 쉽지 않을 듯하다.

유의사항 – 출력

* 악보를 만들어 출력하려면 어떤 프로그램을 사용해야 할까?
  + 추가바람, 다만 우선순위가 낮다고 생각됨

유의사항 – 기타

* CUDA코어를 사용할 수 있도록 NVIDIA 그래픽 카드가 장착된 컴퓨터가 필요하다
  + 대여할 수 있는가?
  + 구매해야 한다면 어느 수준의 그래픽카드가 필요한가?
  + 지원금을 받을 수 있나?  
    ([한양대학교 융합전자공학부 (hanyang.ac.kr)](http://electronic.hanyang.ac.kr/kor/education/board.php?m=v&idx=4885&pNo=1&code=grd_notice)) (인당 30만원)  
    단, 발표회 필수 참여인듯
* 졸업작품 진행을 어떻게 할 것인가?
  + 딥러닝은 시간이 많이 필요하기 때문에, 몰아서 하기보다는 딥러닝을 꾸준하게 돌리면서 하는게 좋을 것 같다.
  + 졸업작품 최종 보고서, 연구노트, 제안서는 어떻게 작성?
  + 현 시점에서는 배정 승인서/제안서를 제출해야 한다.  
    작년 내용을 보니, 2월 즈음에 승인서 양식이 공지로 올라올 듯하다.
  + 자세한 내용은 capstone-design-설명회.pdf 참고
* 그래서 전체적인 진행은?
  + 2-4월달: 이론적인 부분을 끝내기  
    이론 + 간단한 딥러닝 모델
  + 5-6월달: 딥러닝 수행 및 1차적인 완성작 만들기
  + 6월 말: 중간보고서 제출
  + 7-9월달: 딥러닝 성능 높이기 및 다듬기
  + 10월말: 1차 심사 ------ 사실상 최종제출이라고 봐야할 듯 -----
  + 11월: 상위작 및 하위작 심사
  + 12월 초: 최종 제출물 제출

피아노 음계 여러 개

Densenet 기초로 하는중