**Summary for CV articles**

1. YOLO 기반의 광학 음악 인식 기술 및 가상현실 콘텐츠 제작 방법
2. 문제 인식 및 배경 지식

OMR은 사진 형태로 있는 악보를 디지털화 시키는 것을 의미한다. OMR을 활용하여 악보 정보를 컴퓨터에 불러들이고 이를 토대로한 오디오를 재생하는 등 다양한 방면에서 활용된다. 최근에는 합성곱 신경망이 OMR에 적용되어 YOLO, SSD, Fast R-CNN 등 다양한 기술이 사용된고 있다. VR은 4차 산업혁명의 핵심 기술로 주목받고 있다. VR는 타 플랫폼에 비해 렌더링에 더 높은 사양의 하드웨어 자원이 필요하다는 단점이 있다. 이러한 단점은 VR 게임의 컨텐츠 부족으로 이어진다. 따라서 이 논문에서는 신경망을 통해 악보를 추출하고 이를 VR 게임의 기반으로 사용할 수 있는지 연구한다.

1. YOLO 기반의 악보 인식

사람이 악보 위의 음표를 해석하기 위해서는 악보 종류 분류, 보표 크기 설정, 오선 좌표 검출, Hough Transform을 이용한 특정 음악 객체 검출 및 분류, 보표 크기 수정, 보표 내외 음표 개체 분류, 보표 크기 재수정, Beam, Flag와 음표 객체 연결, 온쉼표와 2분 쉼표 분류, 가상의 선 생성, 출력의 과정을 거친다. 음표, 오선 등 악보에서 사용되는 대부분의 객체를 검출하는 데는 YOLOv5가 사용된다.

1. VR 콘텐츠 생성 – 악보 검출 데이터 추출 및 처리

위와 같은 과정을 거쳐 얻은 악보 검출 데이터를 채보라고 한다. 채보에는 음정, 박자 등 곡의 전반적인 요소가 들어있다. 채보는 VR 게임 콘텐츠에서 노트를 생성하는 데 사용된다. 우선 음표의 박자 길이를 실수 형태로 변환하기 위해 채보에서 BPM을 추출하고, 각 노트를 채점하기 위해 전체 음표 수를 추출한다. 두 데이터를 채보 파일에서 분리하면 노래의 음정과 박자에 대한 부분이 남는다. 음정과 박자는 1차 분해, 2차 분해를 거쳐 노트를 생성하는 데 쓰이는 정보로 가공된다.

1. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift
2. Background Knowledge and Problem

Training DNN is complicated by the fact that the distribution of each layer’s inputs changes during training, as the parameters of the previous layers change. This slows down the training by requiring lower learning rate and careful parameter initialization, and makes it notoriously hard to train models with saturating nonlinearities. This phenomenon is called internal covariate shift, and there were lots of theories to compensate it.

1. Whitening

We can make inputs whitened, linearly transformed to have zero means and unit variances, and decorrelated. We could consider whitening activations at every training step or at some interval. However, if these modifications are interspersed with the optimization steps, then the gradient descent step may attempt to update the parameters in a way that requires the normalization to be updated, which reduces the effect of the gradient step. Consequently, he combination of the update to bias and subsequent change in normalization led to no change in the output of the layer nor the loss. The other problem is that whitening the layer inputs is expensive, as it requires computing the covariance matrix.

1. Batch Normalization

We can normalize each scalar feature independently by making it have the mean of zero and the variance of 1. Since simply normalizing each layer may change what the layer can represent, we introduce a pair of parameters which scale and shift the normalized value, also learnable. We would use the mini batches in stochastic gradient. The statistics for normalization can fully participate in the gradient backpropagation because each mini-batch produces estimates of the mean and variance of each activation