

# AI 音樂生成與實時評分平台：基於 Transformer-WGAN 與神經網路訓練互動介面

## AI Music Generation and Real-Time Scoring Platform: Based on Transformer-WGAN with Neural Network Training Interactive Interface

指導教授：陳朝鈞

專題成員：陳曄翰

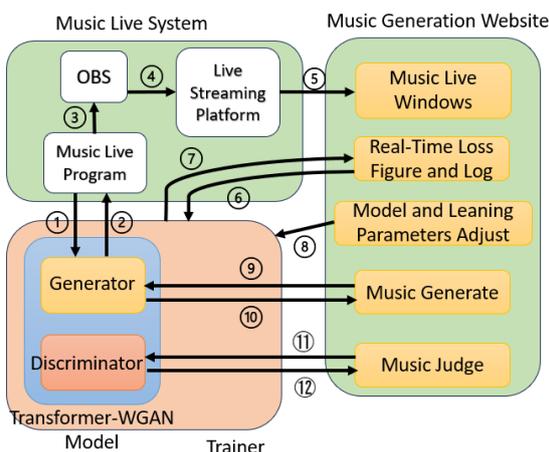
開發工具：Python, Html, CSS, JavaScript, C#, PyTorch, Django, OBS Studio

測試環境：Ubuntu, Windows 10, Google Chrome

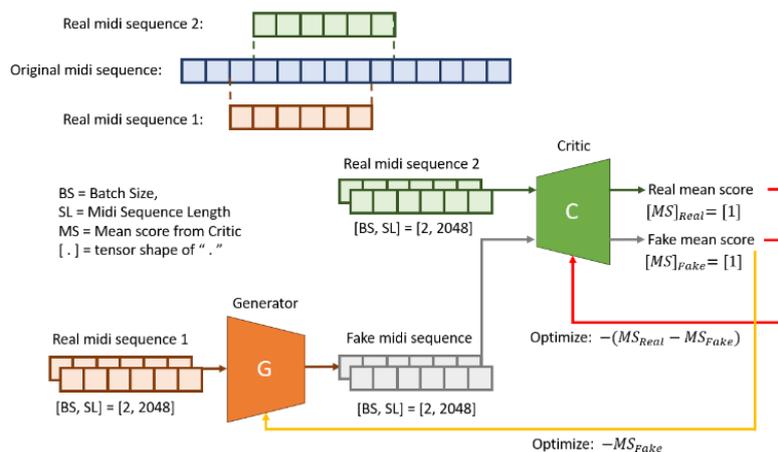
### 一、簡介：

本專題目的是開發一種利用深度學習技術的創新音樂生成方法。這一方法不僅使用 AI 進行音樂創作，也藉由 AI 即時評分來客觀評估音樂品質。由於目前學術界缺乏一套公認的音樂品質量化標準，因此我們使用生成對抗網路（WGAN）中的判別器來評估生成音樂的品質、提供實時的反饋。專題的核心內容包含以下幾個部分：

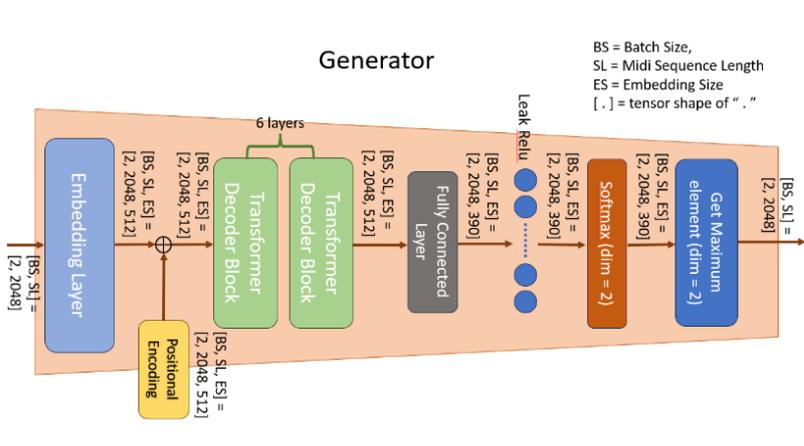
1. **音樂生成與即時評分**：利用 AI 基於 Transformer 之 WGAN 中的生成器進行音樂生成，並使用判別器對指定音樂進行即時評分。
2. **網站展示與互動平台**：架設網站用以展示生成的 MIDI 文件，也可以透過將 YouTube Live 嵌入網站來實時監看音樂生成結果。
3. **神經網路訓練互動介面**：開發一個互動式用戶界面，讓開發者可以直接在網站上調整神經網路的層級結構和參數（如學習率、批量大小、訓練周期等），並即時查看訓練過程中的損失圖和訓練日誌。以下為系統架構圖：



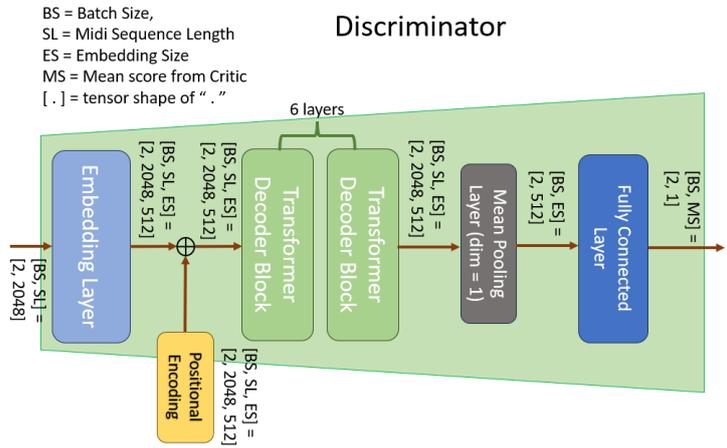
圖一、系統架構圖。



圖二、WGAN 架構圖。



圖三、生成器架構圖。



圖四、判別器架構圖。

## 二、測試結果：

上傳 mid file 並評分

選擇檔案 未選擇任何檔案

提交評分

音樂真實度評分: 90.1%

### 模型調整區

Add New Layer

**Layer 1**

embedding layer

Embedding Dimension: 1

Number of Embeddings: 512

Delete Layer

**Layer 2**

fully-connected layer

Input Dimension: 50

Output Dimension: 100

Delete Layer

**Layer 3**

transformer

Number of Heads: 8

Model Dimension: 512

Feedforward Dimension: 1024

Dropout: 0.1

Delete Layer

Apply

### 生成的PyTorch程式碼

```
import torch
import torch.nn as nn

model = nn.Sequential(
    nn.Embedding(num_embeddings=512, embedding_dim=1),
    nn.Linear(50, 100),
    nn.Transformer(d_model=512, nhead=8, dim_feedforward=1024, dropout=0.1),
)
```

### Midi 檔案展示區

0:00 / 0:29

0:00 / 0:33

0:00 / 2:12

0:00 / 0:28

0:00 / 1:53

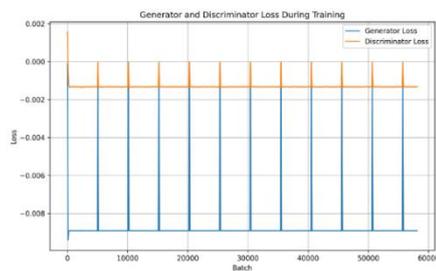
0:00 / 1:47

### 生成音樂直播區

ai音樂生成

到以下平台觀看: YouTube

### 訓練實時監控



### Learning Log

[Epoch 1/1000]	[Batch 100/5066]	[D loss: 0.932387]	[G loss: 0.163629]
[Epoch 1/1000]	[Batch 200/5066]	[D loss: 0.641731]	[G loss: 0.533460]
[Epoch 1/1000]	[Batch 300/5066]	[D loss: 0.404271]	[G loss: 0.711196]
[Epoch 1/1000]	[Batch 400/5066]	[D loss: 0.600578]	[G loss: 0.644716]
[Epoch 1/1000]	[Batch 500/5066]	[D loss: 0.371243]	[G loss: 0.719220]
[Epoch 1/1000]	[Batch 600/5066]	[D loss: 0.387324]	[G loss: 0.512547]
[Epoch 1/1000]	[Batch 700/5066]	[D loss: 0.373870]	[G loss: 1.180949]
[Epoch 1/1000]	[Batch 800/5066]	[D loss: 0.412254]	[G loss: 0.747303]
[Epoch 1/1000]	[Batch 900/5066]	[D loss: 0.254941]	[G loss: 1.090486]
[Epoch 1/1000]	[Batch 1000/5066]	[D loss: 0.309474]	[G loss: 1.114308]
[Epoch 1/1000]	[Batch 1100/5066]	[D loss: 0.266445]	[G loss: 1.220899]
[Epoch 1/1000]	[Batch 1200/5066]	[D loss: 0.179959]	[G loss: 1.304349]
[Epoch 1/1000]	[Batch 1300/5066]	[D loss: 0.197302]	[G loss: 1.130355]
[Epoch 1/1000]	[Batch 1400/5066]	[D loss: 0.224983]	[G loss: 1.361877]
[Epoch 1/1000]	[Batch 1500/5066]	[D loss: 0.239294]	[G loss: 1.793091]
[Epoch 1/1000]	[Batch 1600/5066]	[D loss: 0.162379]	[G loss: 1.344646]