

## Fantastic AI Sinawi: 심층신경망 기반 국악 선율 작곡을 활용한 시나위 창작

박 한 나<sup>1</sup> · 한 단 비 내 린<sup>2</sup> · 오 채 령<sup>3</sup> · 정 다 샘<sup>4\*</sup>

<sup>1</sup>서강대학교 컴퓨터공학과 학사과정 <sup>2</sup>서강대학교 아트&테크놀로지학과 석사과정

<sup>3</sup>서강대학교 아트&테크놀로지학과 학사 <sup>4</sup>서강대학교 아트&테크놀로지학과 조교수

## Fantastic AI Sinawi: Composing Korean Traditional Music Using Deep Neural Networks

Hannah Park<sup>1</sup> · Danbinaerin Han<sup>2</sup> · Chaeryeong Oh<sup>3</sup> · Dasaem Jeong<sup>4\*</sup>

<sup>1</sup>Bachelor's Course, Department of Computer Science & Engineering, Sogang University, Seoul 04107, Korea

<sup>2</sup>Master's Course, Department of Art & Technology, Sogang University, Seoul 04107, Korea

<sup>3</sup>Bachelor, Department of Art & Technology, Sogang University, Seoul 04107, Korea

<sup>4</sup>Professor, Department of Art & Technology, Sogang University, Seoul 04107, Korea

### [요 약]

본 연구에서는 심층신경망을 이용하여 새로운 국악 선율을 생성하고 이를 활용해 인간과 인공지능 간의 합주 시나위를 완성하는 과정을 소개한다. 국악 선율 생성을 위해 시나위 악보 데이터셋을 수집하고 편집하였으며 이를 이용해 순환신경망 기반의 선율 생성 모델을 학습하였다. 학습 데이터의 부족함을 극복하기 위해 다른 단선율 데이터셋을 이용한 사전학습을 적용하였다. 학습된 모델이 생성한 선율들을 활용하여 하나의 시나위 악곡으로 구성하였고 가상악기를 활용하여 각 악기에 적합한 주법을 적용할 수 있도록 선율을 편집하였다. 이렇게 구성된 반주 파트에 해금 연주자가 선율을 모방하는 즉흥 연주를 함께하여 Fantastic AI Sinawi를 완성하였다. 작품의 시각적인 몰입을 돕기 위해 해금 연주 오디오를 입력으로 받는 실시간 인터랙티브 시각화를 함께 제안하였다.

### [Abstract]

This study presents a method for generating new Korean traditional music melodies using deep neural networks. Furthermore, it explores their use in collaboration between humans and artificial intelligence. To create these melodies, we collected and edited a dataset of Sinawi scores, which were then used to train a melody generation model based on the recurrent neural network. Because of limited training data, we used pre-training techniques using additional monophonic melody datasets. We composed a single Sinawi piece using 20 melodies generated by the trained model. We used virtual instruments to synthesize melodies while editing each phrase with proper transposition and articulation for each instrument. The generated accompaniment part was performed with a Haegeum player who improvised to mimic the melody, thereby completing Fantastic AI Sinawi. We also proposed a real-time interactive visualization that responds to the audio of the performance.

**색인어** : 국악, 딥러닝, 시나위, 음악 시각화, 음악 인공지능, 음악 정보 검색

**Keyword** : Korean traditional music, Deep learning, Sinawi, Music visualization, Music AI, Music information retrieval

<http://dx.doi.org/10.9728/dcs.2023.24.4.785>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Received** 23 February 2023; **Revised** 20 March 2023

**Accepted** 22 March 2023

**\*Corresponding Author; Dasaem Jeong**

**Tel:** +82-2-705-8381

**E-mail:** [dasaemj@sogang.ac.kr](mailto:dasaemj@sogang.ac.kr)



평가하는 AI Music Generation Challenge가 진행되기도 하였다[22]. 이 대회의 경우 2020년 1회 대회에서는 아일랜드 전통 음악 중 Double Jig를 생성하는 것을 주제로 하였고 [22], 이듬해 2회 대회에서는 스칸디나비아의 전통 춤곡인 Slängpolska를[23], 2022년 3회 대회에서는 아일랜드의 Reel 음악을 생성하는 것이 주제였다[24]. 이외에 MG-VAE[25]에서는 중국 민요의 특정 음악 스타일을 포착하여 조작할 수 있는 방식으로 새로운 곡을 생성하는 연구를 진행했다.

우리나라의 전통 음악에 대해서 심층신경망을 활용한 연구 사례가 존재한다. 배준은 날리리아, 진도 아리랑, 강강술래와 같은 대표적인 국악곡 35개를 데이터셋으로 사용해 국악 멜로디를 생성하는 연구를 진행했으며 LSTM, Music Transformer, 그리고 Self Attention 모델을 선택해 비교하였다[26]. 이동진 외 2명은 정악 중 수연장지곡의 해금 선율을 예로 들어 국악의 기하학적인 구조를 시각화했으며, 음표들의 나열을 직접 학습하는 대신 곡 내부에서 발견한 기하학적 구조의 패턴을 인공신경망에 학습해 음악을 생성하는 방법을 제안하였다[27]. 이러한 접근법의 경우 학습 악곡에 포함된 특징적인 패턴을 직접적으로 활용한다는 장점이 있지만 음표의 나열을 중단형으로 학습하는 방법에 비해 생성된 출력의 자유도가 떨어지고 다른 종류의 음악을 사전학습 데이터로 활용하기 어렵다는 단점이 있다.

최근 심층신경망 기반의 음악 생성의 연구에서 데이터셋이 한정된 전통 음악이나 특수한 장르의 음악을 다루는 경우 큰 규모의 데이터셋에 모델을 사전 학습시킨 후 작은 규모의 데이터셋에 적용하는 방식을 사용하는 예시가 많다. 이러한 방법은 아일랜드 전통 음악으로 사전 학습한 후 스웨덴 전통 음악 생성을 학습한 연구[11]나 대규모 팝 음악 데이터셋에 사전 학습하고 게임 음악 데이터셋에 미세조정 하는 방식으로 적용되었다[28]. 본 연구에서는 이러한 사례를 따라 사전 학습과 미세조정을 활용하였다.

### III. 시나위

시나위는 무속음악에서 유래하여 발달한 기악음악으로서, 한국의 민속악을 대표하는 장르라고 할 수 있다. 시나위가 즉흥성을 기반으로 한반도 전 지역에서 다양한 토리와 음계로 구성되는 만큼 시나위라는 장르에 대해서 심도 있게 다루기 위해서는 한국음악의 선법과 음계에 대한 종합적인 이해가 선행되어야 함이 분명하다. 하지만 본 연구에서는 특정 선법과 장르를 인공지능이 모방하여 창작할 수 있는지를 살펴보는 것에 목적이 있기에 구현하고자 하는 장르와 그 특징의 범위를 간소화할 필요성이 있었다. 이러한 이유로 이 연구에서는 육자배기토리(남도계면조) 음계 중심의 허튼가락으로 연주되는 남도시나위를 학습 대상으로 설정하였다.

오늘날 일반적으로 전문 연주자들이 활발히 연주하는 남도 시나위는 삼현육각이라고 구분되는 피리, 대금, 해금, 장구, 징을 비롯하여 가야금, 거문고, 아쟁 등의 악기로 구성되는 남도계면조의 기악합주곡을 의미한다. 연주자들은 장단의 초반부에 구성되는 보편적인 선율 이후로는 고정된 가락이 없이 유동적, 즉흥적으로 연주자들 간의 호흡을 맞추어 자유분방하게 연주한다. 다른 전통기악곡에서 피리가 가락을 주도하는 경우가 많은 것과 달리 시나위에서는 아쟁과 대금을 가락을 주도하는 경향이 크며, 악기달은 각기 다른 선율과 음색을 가지면서도 다른 악기 선율의 흐름을 예민하게 주시하며 그 안에서 조화를 이루어낸다. 기악합주곡으로 연주될 때 연주자가 자신의 기량을 뽐내는 독주 부분이 들어가는 것도 시나위의 눈에 띄는 형식 중 하나이다.

#### 3-1 시나위의 장단 및 음계

국악에서 장단은 “일정한 길이의 리듬(rhythmic pattern)”, “박자(meter), 빠르기(tempo), 강약(accent), 한배(pattern)”로 그 개념을 구체화할 수 있다[29]. 이보형의 장단분류이론[30]에 따르면, 시나위에서 사용되는 ‘굿거리’와 ‘자진모리’는 우리나라에서 장단에서 그 박자 구조가 3소박 4박자(12/8박자)로 되어 있는 경우에 해당한다.

시나위는 장구 연주자가 굿거리 한 장단을 처춤으로써 기악연주자들에게 템포를 알려주며 시작한다. 이후 징은 각 장단의 첫박에 1회 연주하는 것이 일반적이며 장구는 각 장단의 기본 및 변형 리듬을 연주한다. 이때 변형 리듬은 잔가락 및 헤미올라(hemiola, 2:3) 리듬을 활용하여 구성한다.

대부분의 시나위 선율은 대개 육자배기토리 음계로 이루어져 있다. 이보형은 경기도 남부에서 충청도와 전라도, 경상도 서부를 경계로 하여 공통적으로 가진 음악적 특징을 ‘육자배기토리’로 명명하고 ‘시나위권’의 음악을 규정하였다[31]. 이는 대표적인 남도민요인 ‘육자배기’의 이름을 따서 붙인 명칭으로 <미-라-(시)-도>의 단순형, <미-솔-라-(시)-도-레>의 5음 복합형 변형태 음악으로 나눌 수 있다.

### IV. 데이터셋 수집

딥러닝 기반 언어 모델을 이용하여 새로운 선율을 생성하기 위해서는 선율 정보가 음고와 길이 및 박자 정보로 정리되어 있는 심볼릭 음악(symbolic music) 데이터가 학습 시 필요하다. 기존에 공개되어 활용되는 국악 악보 데이터셋은 악보를 디지털 스캔한 래스터 이미지(raster image) 형태로 존재하는데 이 경우 광학악보인식(OMR; optical music recognition)을 사용하지 않고서는 이미지상에서 음표 정보를 기호화하여 처리할 수 없다. 광학문자인식과 달리 광학악보인식의 정확도는 자동으로 데이터셋을 처리할 수 있을 만

**표 1.** 수집한 사전학습 데이터셋  
**Table 1.** Collected pre-training dataset

Dataset	Format	Number of Songs
Essen Folk Song	ABC, SM(EsAC)	9,034
O'Neil's Dance Music of Ireland	ABC	750
ORCHESET	MIDI	64
Total		9,848

큼 정확도가 높지 않기 때문에 저자가 직접 사보 프로그램을 활용하여 시나위를 포함한 여러 국악 악보를 서양 오선보 형태의 기호형 데이터로 변환하였다.

하지만 국악 악보의 경우 수집할 수 있는 모든 예시를 데이터셋으로 사용하여도 신경망을 충분하게 학습하기에 그 크기가 부족할 수 있다. 전통 음악을 생성하는 신경망을 다룬 이전 연구 사례에서는 이런 문제를 해결하기 위하여 규모가 큰 다른 음악 데이터셋으로 신경망을 사전 학습(pre-training)한 뒤, 작은 규모의 데이터셋으로 사전 학습된 모델을 미세조정(fine-tuning) 하는 방법을 활용하였다. 스코틀랜드 전통 음악을 모델링하는 BinaryMuseGAN[32]의 경우 공개된 MIDI (musical instrument digital interface) 데이터셋 중 가장 규모가 큰 Lakh MIDI 데이터셋[33]을 사전학습에 활용하여 신경망이 기본적인 음악 어법을 학습하도록 하고 미세 조정을 통해 신경망이 스코틀랜드 전통 음악 스타일을 생성하게 하였다.

#### 4-1 사전학습 데이터셋

본 연구는 단선울 생성을 목적으로 하기에 반주 트랙 등이 없고 단선울만으로 구성된 데이터셋을 사전학습에 활용하였다. 수집한 데이터셋은 표 1과 같다.

Essen Folk[34]은 유럽과 중국의 전통 음악을 ABC와 EsAC (essen associative code) 형식으로 나타낸 데이터셋이다. O'Neill's Dance Music of Ireland[35]은 아일랜드 춤곡으로 이루어진 Frank Nordberg의 필사본으로 구성된 데이터셋이다. 그리고 ORCHESET[36],[37]은 베토벤부터 20세기 작곡가까지의 관현악곡 오디오와 이에 대응하는 단선울 MIDI 레이블로 구성된 데이터셋으로서 본 연구에서는 이중 단선울 데이터만을 사용하였다.

#### 4-2 시나위 데이터셋

시나위 데이터셋은 주로 2010년 서울 국립국악원에서 편찬한 『한국음악, 한국음악, 제38집 : 시나위』의 남도시나위 각 파트별 악보를 채보하여 제작하였다. 해당 악보는 한국음악선집 제32집으로 출판된 음원을 채보하여 발간한 오선악보로 해당 악보를 채보한 음원의 연주자는 피리 한세현, 대금 원

장현, 해금 김정림, 가야금 박준호, 거문고 한민택, 아쟁 김영길, 장구 조용복, 징 황광엽으로 구성되었다. 이들 모두 당시 국립국악원 민속악단 단원으로 활동하고 있던 연주자들로 남도시나위의 성음을 잘 이해하여 시나위 가락을 창작할 수 있는 공신력 있는 연주단으로 구성되었다고 볼 수 있다. 해당 자료에서는 대금, 피리, 해금, 아쟁, 가야금, 거문고, 장구, 징이 전부 나타나는 총보와, 악기별 파트보로 구성되어 있고 두 악보는 같은 내용을 담고 있다. 해당 악보는 기록을 목적으로 한 것이기 때문에 실음으로 채보를 하였다고 기술되어 있으나 (청: C, Cm) 실제로 음원과 비교하여 악보를 확인해보면 음원의 청(중심음) 높이가 2도가량 높은 것을 확인할 수 있었다.

남도시나위 가락 데이터셋을 위한 가락을 선별할 때, 고려해야 하는 점은 다음과 같았다. 1) 학습 데이터셋이 각 장단의 특징을 살릴 수 있는가? 2) 해당 가락이 남도시나위의 음계를 잘 반영하는 가락인가? 3) 가락에 특정 악기의 개성이 과하게 드러나거나 해당 가락을 다른 악기로 연주로 하기에 불가능하지는 않은가?

첫째, 시나위는 굿거리와 자진모리 두 장단으로 구성되어 있다. 국립국악원 악보에서는 두 장단 모두 12/8박자로 같으나 굿거리는 점4분음표 기준 45-50, 자진모리는 점4분음표 기준 75-96의 템포로 기록되어 있다. 두 장단의 템포와 장단 형이 다른만큼, 악기 선율에서 나타나는 가락의 밀집도나 리듬형이 상이하다. 본 연구에서는 두 장단의 선율적 특성을 모델이 학습을 할 수 있도록 하기 위해서 굿거리와 자진모리의 데이터셋을 각각 따로 제작하였다.

두 번째, 산조합주나 산조 등의 악곡과 연주자가 자신의 기량을 보여주기 위한 독주 부분에서는 변조 및 변청 등 변화가 일어날 수 있으나 시나위에서는 그 조의 변화가 크지 않은 편이다. 또한 해당 악보자료에서는 합주 부분과 대부분의 독주 선율에서 남도시나위 음계(육자배기토리)로 연주하는 것을 확인하였다. 따라서 가야금을 비롯한 변청-변조 부분을 제외하고는 대부분의 선율을 데이터셋으로 활용할 수 있었다.

마지막으로 가야금 및 거문고는 발현악기의 연주 특성상 옥타브 음을 활용하는 경우를 많이 확인할 수 있는데, 이러한 가락은 해당 악기의 가락에서만 자주 출현한다. 다른 관악기(대금, 피리, 해금) 이러한 가락을 연주하기 어려울 뿐만 아니라 연주 시 자연스럽게 않게 느껴질 수 있기 때문에 대금, 피리, 해금, 아쟁의 파트보는 해당 가락 그대로 사용하였으며 가야금과 거문고의 가락에서 옥타브가 처음으로 도약하는 경우 해당 처음을 한 옥타브를 올려 악보를 수정하여 사용하였다.

위에 언급한 기준에 따라 데이터셋에 쓸 가락을 선별하고, 선별된 가락들을 모델 학습에 용이하게 만들기 위해 2차로 수정하였다.

우선, 앞꾸밈음 및 뒷꾸밈음은 심볼릭 표현에서 길이 값을 가지고 있지 않기 때문에 데이터셋 학습 후 생성 과정에서 오류를 일으킬 수 있음을 확인하였다. 이 문제를 해결하고자 꾸밈음으로 되어 있는 가락 중에 필수적이지 않은 잔가락은 제거하였고, 음악적인 맥락에서 중요한 가락으로 판단되는 경우

꾸밈음이 붙은 본음의 길이를 줄여 꾸밈음이 그만큼의 길이를 가지도록 편집하였다.

이 작업 과정 중에서, 꺾는 음, 즉 Eb-D음과 F-Eb과 같은 꺾는 목에서 첫 음을 꾸밈음으로 처리한 것을 눈에 띄게 많이 확인할 수 있었다. 꺾는 목을 처리할 때 가상악기의 주법을 활용하여 후처리할 계획을 하였다. D는 육자배기토리에서 대부분 Eb으로 꺾어서 연주하기 때문에 Eb-D 꺾는 음을 전부 D로 수정하였고, F-Eb으로 나타나는 꺾는 목의 경우 Eb-D 꺾는 목과 달리 항상 나오지 않기 때문에 붓점 리듬형으로 악보를 수정하였다.

또한 악기의 음역대를 고려하여 대금의 경우 기준음이 C5, 피리와 해금은 C4, 아쟁은 C3으로 구성되어 있음을 확인할 수 있었다. 이는 악기마다 남도시나위 음계의 중심음이라는 점에서 같은 역할을 하지만, 데이터셋을 학습하는 데 있어서 C5, C4, C3는 독립적인 음고값으로서 여겨져 각각의 특징을 개별적으로 학습해야 한다는 문제가 있다. 학습할 데이터셋의 양이 많지 않기 때문에 학습해야 하는 음고 값의 개수를 줄여 성능을 높이고자 대금은 전체 가락을 한 옥타브 내리고, 아쟁은 한 옥타브 올려 기준 음을 C4로 통일시켜주었다.

생성된 가락이 음악적인 맥락을 파악하여 가락을 종지할 수 있게끔 선율의 음악적 흐름이 마무리되는 부분을 기준으로 악구(phrase)를 분할하였다. 굵거리는 총 40개의 악구, 자진모리는 35개의 악구로 나누어 학습을 진행하였다.

## V. 선율 작곡 모델 구조 및 학습

### 5-1 모델 구조

단선율 정보의 경우 각 음표를 자연어의 단어와 같이 취급할 수 있어 자연어 생성 모델의 하나인 언어 모델의 구조를 차용할 수 있다. 언어 모델은 지금까지 등장한 토큰(token)을 기반으로 다음에 등장할 토큰의 확률을 계산하는 모델이다. 음악에 적용할 경우, 지금까지 등장한 음표들을 사전 조건으로 주었을 때, 미리 정의된 어휘에 포함된 각각의 음표에 대해 다음에 등장할 수 있는 확률을 계산한다.

2018년 이후로 GPT[38]와 같이 transformer 구조를 활용하는 것이 언어 모델 설계에서 주류가 되었으나 음악의 경우 비교적 한정적인 데이터셋의 크기 때문에 LSTM (long short-term Memory)이나 GRU (gated recurrent units)와 같은 순환신경망(RNN; recurrent neural network)을 활용하는 예시가 많다. 본 연구에서도 데이터셋의 크기를 고려하여 순환신경망의 일종인 GRU를 사용하였다. 그리고 그림 1과 같이 학습 파이프라인을 구성했다.

심볼릭 음악을 위한 라이브러리인 ‘MusPy’[39]를 사용해 MIDI, ABC, JSON 등 다양한 형식으로 되어있는 선율 파일을 읽고 각 선율을 음표의 열로 표현하여 데이터셋에 추가하

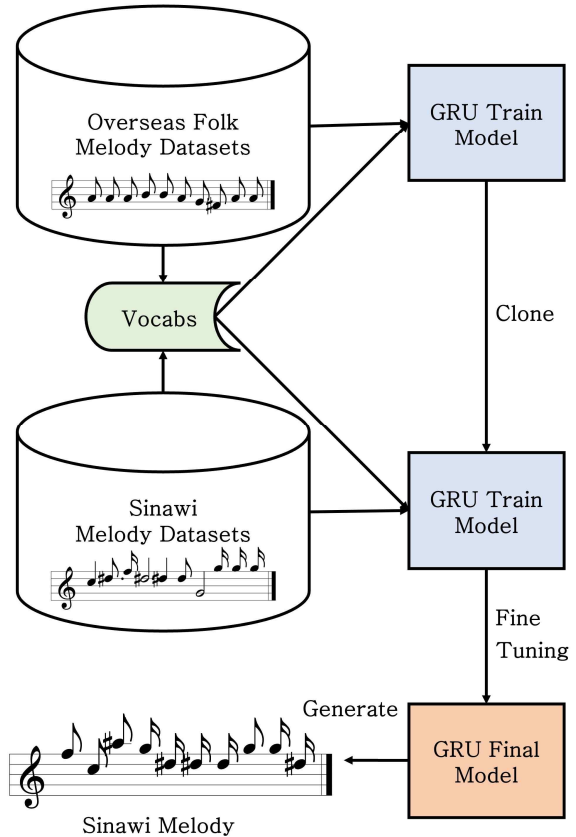


그림 1. 학습 파이프라인 구조도  
Fig. 1. Diagram of training pipeline

였다. 이때 각 음표는 음고와 길이의 조합으로 구성된다. 주어진 데이터셋으로 언어 모델을 학습하기 위해서는 데이터를 다시 토큰의 열로 표현해야 한다. 이때 각 토큰은 음표나 쉽표 하나를 나타내게 되며 이외에 특수 토큰으로 시작 지점('start')과 끝 지점('end')을 나타내는 것으로 설계하였다.

데이터셋에 포함된 모든 음고와 길이 중 고유한 값을 어휘(vocabulary)로 만들어 각 음고와 길이마다 고유한 순번을 가지도록 만들었다. 이를 통해 각각의 음고와 길이를 사전 상의 순번으로 치환하여 표현하여 신경망의 입력으로 활용한다. 본 작품에서는 사전학습 데이터셋과 시나위 데이터셋 공통의 어휘를 가지고 있어야 사전학습 데이터셋으로 학습한 모델을 시나위 데이터셋으로 미세 조정하여 최종 모델을 학습할 수 있다. 따라서 데이터셋을 전처리할 때 공통의 어휘를 만드는 과정을 포함했다.

음고와 길이는 각각 독립된 원-핫 벡터(one-hot vector)와 같은 형태로 표현되며 각각의 벡터는 임베딩 레이어를 통해 밀집 벡터(dense vector)로 변환된다. 이때 임베딩 레이어의 가중치 값은 모두 학습 가능한 파라미터이다. 임베딩 레이어를 활용하여 각 음고와 길이를 40차원의 벡터로 변환하였으며 이를 이어 붙여 80차원의 임베딩 벡터를 순환신경망 레이어의 입력으로 활용했다.

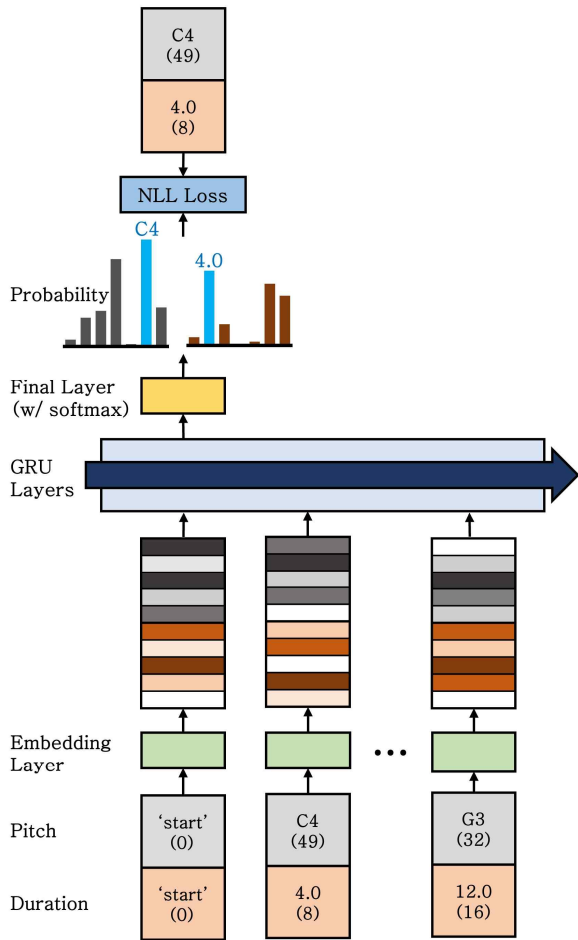


그림 2. 모델 구조도  
 Fig. 2. Diagram of model architecture

순환신경망은 입력으로 들어온 벡터  $x_t$ 와 이전 단계의 은닉 상태(hidden state)  $h_{t-1}$ 를 이용하여 현재 단계의 은닉 상태  $h_t$ 를 계산한다. 이 계산 과정에서 GRU는 리셋 게이트와 업데이트 게이트를 활용하여 각 단계에서 은닉 상태를 얼마나 보존하거나 업데이트할 것인지를 결정한다[40]. 본 연구에서는 이러한 GRU 레이어를 5개를 쌓아올린 구조를 활용하였다. 이렇게 생성된 각각의 단계에 대한 은닉 상태 출력  $h_t$ 를 선형 레이어에 입력으로 주어 어휘에 존재하는 모든 ‘단어’에 대한 등장 확률 로짓(logit)을 계산하게 하고 여기에 소프트맥스(softmax) 함수를 활용하여 각 단어의 등장 확률이 0과 1 사이의 값을 가지게 하며 어휘에 존재하는 모든 단어의 등장 확률의 합이 1이 되게 만든다. 본 연구에서 어휘의 단어는 어휘에 대한 소프트맥스와 길이 어휘에 대한 소프트맥스 출력값을 별도로 계산한다. 이 과정을 도식화한 것은 그림 2와 같다.

5-2 학습

전체 데이터셋을 9:1의 비율로 훈련 데이터셋과 검증 데이

터셋으로 나누어서 진행했다. 일반적으로 인공지능경망 모델을 학습할 때 훈련, 검증, 시험 데이터셋으로 나누지만 본 연구의 목적이 시험 데이터셋에 대해 펄플렉서티(perplexity)와 같은 성능을 평가하는 것이 아니라 창작에 활용할 수 있는 선율을 생성할 수 있도록 모델을 학습시키는 것이기 때문에 최대한 많은 데이터를 훈련에 사용하고 별도의 시험 데이터셋을 활용하지는 않았다.

모델 학습을 위한 손실 함수로는 분류 모델 및 언어 모델 학습에 널리 활용되는 NLL(negative log-likelihood)을 사용하였다. 이전에 등장한 음표들이 주어졌을 때 모델이 예측한 다음에 올 음표의 음고와 길이의 확률 분포에서, 실제 다음에 오는 음고와 길이에 해당하는 확률값의 음의 로그 값에 해당한다. 모델이  $n$ 번째 음표로 예측한 확률 분포 중  $i$ 번째 음고의 확률을  $\hat{y}_n^{(i)}$ 라고 표기하고 실제  $n$ 번째 음표의 음고를  $y_n$ 이라고 할 때,  $N$ 개의 입력 음표에 대한 NLL 손실값  $L(\mathbf{y}, \hat{\mathbf{y}})$ 를 수식으로 표현하면 아래와 같다.

$$L(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N} \sum_{n=1}^N -\log \hat{y}_n^{(y_n)} \tag{1}$$

위와 같은 NLL 손실 함수를 음고와 길이 각각에 대해 계산하여 합산한 값을 최종 손실 값으로 활용하여 이를 최소화하는 방향으로 학습하였다.

Adam optimizer[41]를 이용해 모델의 가중치를 학습했으며, 초기 학습률은 0.001을 사용했다. 훈련 데이터에 포함된 모든 곡을 한 번씩 학습에 활용하는 것을 에포크(epoch)라고 할 때, 매 세 번의 학습 에포크마다 검증 데이터셋에 대한 손실값(validation loss)을 측정하여 모델의 학습 상황을 관찰하였다. 훈련 중 두 번 연속 (6 에포크에 해당)으로 검증 손실값을 역대 최솟값 보다 낮게 개선하는 데 실패한다면 학습률에 0.3을 곱해 학습률을 낮춰주는 방식을 적용하였다. Early stopping을 이용해 검증 손실값이 9 에포크에 동안 줄어 들지 않는다면 학습을 멈추도록 진행했다. 사전 학습에서는 early stopping에 의해 에포크가 65일 때 학습을 멈췄다.

사전 학습한 모델을 바탕으로 장단별 시나위 멜로디 데이터셋을 이용해 미세 조정을 진행했다. 특정 장단의 시나위를 생성하는 모델을 만들기 위해 해당 시나위 장단 데이터셋으로 미세 조정하였다. 미세 조정에 사용된 요소와 파라미터들은 사전학습을 진행했을 때와 동일하다. 미세 조정에서는 굵거리는 84 에포크, 자진모리는 93 에포크가 지나고 학습을 중단했다.

5-3 생성

‘start’ 토큰에서부터 시작해 매 음표 임베딩을 순환신경망에 입력하고 이를 통해 계산된 다음 음표의 확률 분포를 기반으로 음고와 길이를 가중치 샘플링하여 음표를 생성한다. 생

성된 음표는 다시 재귀적으로 입력되어 그다음 음표를 순차적으로 생성하게 된다. 사용자가 지정한 음표 길이에 도달하거나 'end' 토큰을 생성하였을 때 선율 생성을 종료한다. 생성된 선율은 MIDI 형태로 변환하여 그림 3과 같이 악보로 출력할 수 있다.



그림 3. (a), (b) 사전 학습으로 만들어진 선율 (c), (d) 미세조정을 통해 생성한 시나위 굿거리 선율 (e), (f) 미세조정을 통해 생성한 시나위 자진모리 선율

Fig. 3. (a), (b) Melodies generated by model after pre-training (c), (d) Sinawi melodies in goodgeori jangdan generated by model after fine-tuning, (e), (f) Sinawi melodies in jajinmori jangdan generated by model after fine-tuning

## VI. 선율 작곡 모델 구조 및 학습

인공지능 모델의 작곡을 통해 인간 연주자와 기계의 합주를 선보이는 작품을 만들기 위해서 기계가 맡은 기계가 맡은 파트의 연주를 자연스러운 연주 오디오로 합성해야 한다. 이를 위해 현재 활용 가능한 국악기 가상악기가 어떤 종류가 있는지 조사하였고 가상악기의 표현력을 충분히 잘 활용할 수 있는 방식으로 선율을 편집하여 작품을 구성하였다.

### 6-1 가상악기 선택 및 시김새 구현

생성된 MIDI를 국악 가상악기로 연주하기 위해 'OX 가상악기 (한양대학교)', '국악 (Gugak)(서울대학교)', '삼현육각 (전북대학교)'의 현존하는 세 가지의 가상악기를 비교했을 때, 이 중 '삼현육각'이 주법 및 피치 벤더(pitch bender)를 자유롭게 조작하여 가장 실제 연주에 가까운 소리를 구현해 낼 수 있음을 확인하였다.

'삼현육각'은 실제 연주를 샘플링하여 제작하였으며 키 스위치(Key Switch) 및 벨로시티(Velocity)를 조절하여 선택할 수 있어 국악기 특유의 시김새 주법을 적용하여 연주 소리를 합성할 수 있다. 사실적인 국악기 표현을 재현하기 위하여 악기의 특성과 음계의 특성에 맞게 선율의 각 음표에 어울리는 주법을 규칙기반으로 적용하였다. 대금의 경우 C5(본청한 옥타 위의 음)보다 높은 음의 경우 청소리(E2, Cheong Sori)를 적용하여 대금이 고음을 낼 때 나는 청소리를 구현하였으며 떠는 목(G3, G4)은 굽게 떠는 청(D#2, Slow Vibrato) 주법을 적용하여 육자배기토리의 떠는 목 시김새(굽은 요성)를 구현하였다. 해당 사항에 없는 음의 경우 일반 주법(C2, Relaxed Long)을 적용하였다. 또한 꺾는 청(D4)에서는 피치 벤더를 이용해 음표의 시작과 동시에 원음보다 반음가량 높은 음에서 원음으로 음고가 변하도록 하여 육자배기토리의 꺾는 목 시김새를 구현하였다. 가야금의 경우 떠는 청(G3, G4)은 빠른 비브라토(D#1, Fast Vibrato)를 적용하여 떠는 목 시김새(굽은 요성)를, 꺾는 청(D3, D4)은 벨로시티를 90아래로 조절 후 E1(Bend Dn) 주법을 적용하여 꺾는 목 시김새를 구현하였다.

### 6-2 작곡 과정

작품의 구성은 다음과 같이 작곡하였다.

- 도입가락 1장단: 장구
- 굿거리 46장단: 합주 20장단, 대금 독주 8장단, 합주 4장단, 가야금 독주 8장단, 합주 6장단
- 자진모리 45장단: 합주 18장단, 해금 독주 7장단, 합주 20장단

학습한 모델을 통해 생성한 150개의 선율 중에서 굿거리와 자진모리 중 각각 20개의 악구를 선별하여 작곡에 활용하였다. 무대용 기악합주곡으로 시나위를 연주할 때 합주 중간에 악기마다 연주자의 기량을 보여줄 수 있는 독주를 구성하는 것에 착안하여 굿거리에서는 차례대로 대금과 가야금 독주 파트(8마디)를, 자진모리에서는 연주자가 자유롭게 연주할 수 있도록 타악기 반주만 연주되는 파트(7마디)를 구성하였다. 앞서 언급한 대로 발현악기의 경우에는 옥타브 관계의 장식음을 자주 활용하므로 가야금 선율의 경우 길이가 긴 음표가 있을 때 한 옥타브 아래 음표를 거쳐 연주하는 식으로 선율을 수정하였다. 장구와 징은 Native Instrument 사의사물놀이(samulnori) 가상악기를 사용하여 합성하였다. 징은 각 장단의 첫 박(강박)을 연주하도록 하였다. 굿거리 및 자진모리의 첫 장구 장단은 도입 장단을 사용하였으며, 도입 장단 이후에는 굿거리 12박 중 앞의 6박자는 5종의 장단형, 뒤의 6박자는 4종의 장단형을 무작위로 조합하여 굿거리는 20개, 자진모리는 4개의 자진모리 장단형 중 무작위로 샘플링하여 총 92개의 장단 음원을 제작하였다.

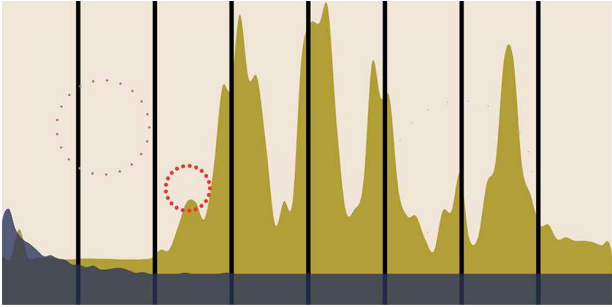


그림 4. 굿거리에서의 1차 시각작업 캡처  
 Fig. 4. Screenshot during goodgeori(굿거리) from the first visualization design

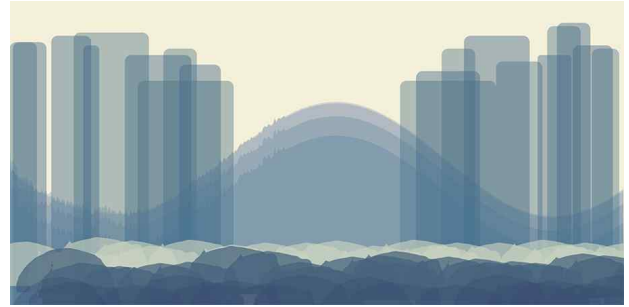


그림 6. 굿거리에서의 2차 시각작업 캡처  
 Fig. 6. Screenshot during goodgeori from the second visualization design

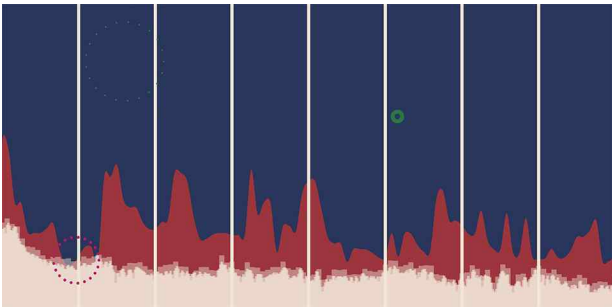


그림 5. 자진모리에서의 1차 시각작업 캡처  
 Fig. 5. Screenshot during jajinmori(자진모리) from the first visualization design



그림 7. 자진모리에서의 2차 시각작업 캡처  
 Fig. 7. Screenshot during jajinmori from the second visualization design

### 6-3 인터랙티브 시각화

본 작품에 시각적인 효과를 가미하기 위해 연주자의 해금 소리와 반주의 장구 소리에 맞추어 연주자의 뒤에서 움직이는 배경 형태의 인터랙티브 시각화를 적용하였다. 시각화의 스타일은 인공지능과 사람의 연주가 어우러져 만들어진 현대적인 한국화라는 콘셉트로 디자인하였다.

#### 1) 1차 작업

조선시대 궁궐에 배치된 일월오봉도[42]는 어좌의 뒤에 놓여 배경을 이루는 병풍이다. 일월오봉도가 산과 물이 있는 병풍이라는 데에서 착안하여, 그림 4 및 그림 5와 같이 1차 시각화 작업을 진행하였다.

해금 소리를 마이크를 통해 실시간으로 받아, 시간에 따른 음압 변화는 황토색 산으로, 스펙트로그램은 남색 물로 표현하였다. 검은 세로선으로 병풍의 형태를 모사하였으며, 컴퓨터가 내는 장구의 ‘덩’ 소리에 맞춰 불꽃이 터지는 효과를 주었다.곡의 중반에 장단이 바뀌며 빠른 템포로 넘어감에 따라, 색상 변경을 통해 더 도시적이고 강렬한 느낌을 주었다. 또한, 스펙트로그램을 연속적이지 않은 디지털 신호의 그래프와 같이 표현하여 현대적인 느낌을 가미하였다.

#### 2) 2차 작업

두 번째 작업에서는 장단의 변화에 따른 컬러 팔레트 변경, 스펙트로그램을 통한 산 표현은 갖고 가되, 작업의 디테일에 변화를 주었다. 정선의 ‘산월도’와 같은 조선 시대 회화 작품을 보면 먹의 농담과 선의 반복을 통해 산과 바위를 표현한 것을 볼 수 있다. 정홍래의 ‘옥해창용’에서는 비슷한 형태의 반복을 통해 물을 표현했다. 이 점에서 착안하여 시각화 작업에 변화를 주었다.

그림 6과 그림 7처럼 산의 경우, 산 모양의 기본 형태를 잡은 후 이 안에서 스펙트로그램의 변화를 나타내도록 하여 좀 더 산의 모양을 부각하고, 다수의 레이어를 통해 산의 농담을 표현하고자 하였다. 산의 양옆에 배치된 바위 또한 같은 방식으로 농담을 표현하였다. 직사각형 형태의 기본 도형을 활용함으로써 바위의 조형에 현대적인 느낌을 가미하였다. 물의 경우, 그림 8의 (a)와 (b)에서 제시된 SVG (scalable vector graphics) 그림을 키프레임으로 장단에 맞추어 바뀌게 하여 일렁이는 물의 모습을 표현하였다.

해당 움직임을 여러 개 배치하고 같이 움직이게 하여, 하나의 거대한 물길처럼 보이게 하였다.

장구의 ‘덩’ 소리에 맞추어서는 물이 튀기는 모습을 표현했다. ‘덩’ MIDI 신호에 맞추어 그림 8의 (c)처럼 4가지의 SVG 그림 중 하나가 무작위로 나타났다 사라지는 형식으로 물이 튀는 모습을 표현하였다.



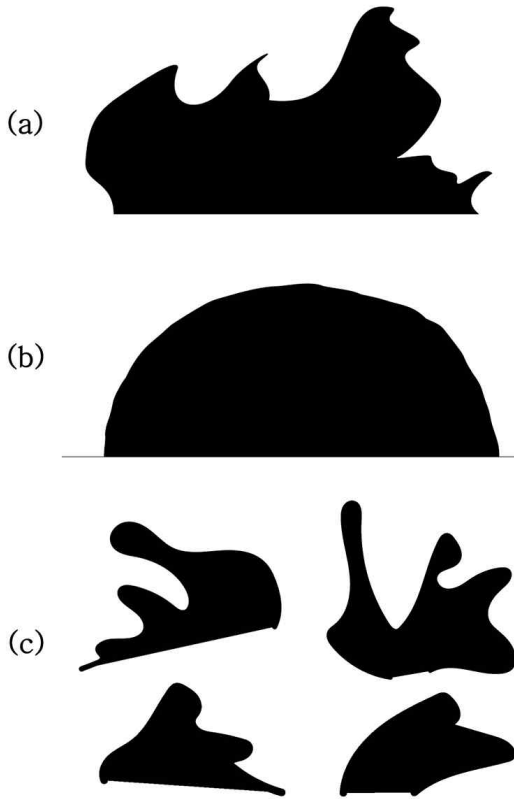


그림 8. (a) 변화하는 물결의 첫 키프레임 (b) 변화하는 물결의 마지막 키프레임 (c) 물 튀김을 표현한 4가지 SVG 그림  
 Fig. 8. (a) The first keyframe of the changing wave (b) The last keyframe of the changing wave (c) Four SVG images that represent water splashing

#### 6-4 실연 과정

1차, 2차 녹화는 서강대학교 하비에르관 이멀시브홀에서 이루어졌다. 그림 9와 같이 시각화 영상은 4K 빔프로젝터를 통해 10,290mm x 5,788mm 스크린에서 사영되고, 연주자는 스크린 앞에 앉아 연주한다. 실연을 위한 설치에는 전면 스크린과 빔프로젝터 외에도 PC 1대와 마이크 1대, 스피커가 사용되었다. PC에서는 해금 연주와 컴퓨터의 장구 MIDI 신호를 인풋으로 받아, 음악과 이에 맞는 시각화 영상을 각각 스피커와 스크린에 내보낸다. 이를 위해 마이크는 해금에 가까이 두고, 음악 시퀀스를 재생할 DAW (digital audio workstation)를 컴퓨터 내부에서 실행한다. 이 중 장구 MIDI 시퀀스를 시각화 프로그램에서 인풋으로 받을 수 있도록 가상 MIDI 버스를 이용했다. 1차 작업을 이용한 작품 연주 영상(<https://youtu.be/d9kmgF1R94M>)과 2차 작업을 이용한 작품 연주 영상(<https://youtu.be/MJNGXpRdKpg>)을 각각 그림 10의 모습으로 기록하였다.

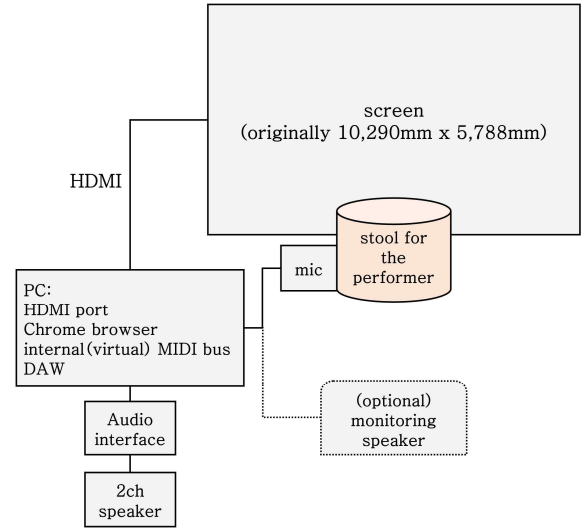


그림 9. 프로젝트 실연을 위한 테크니컬 라이더  
 Fig. 9. The technical rider for performance

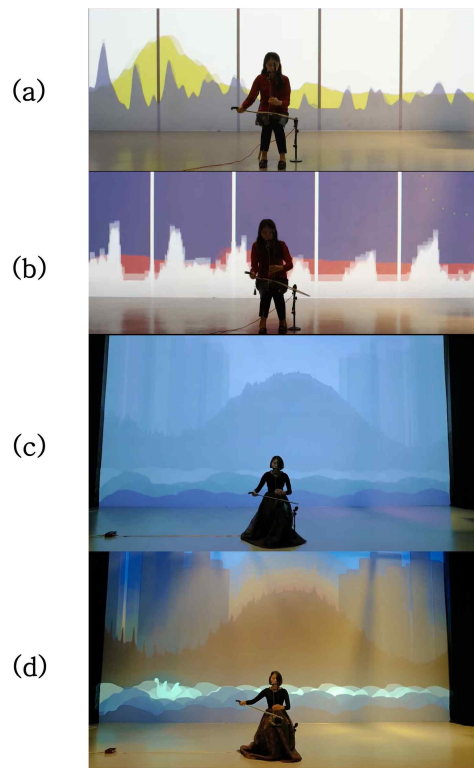


그림 10. (a) 1차 굿거리, (b) 1차 자진모리, (c) 2차 굿거리, (d) 2차 자진모리 시연 캡처. 1차, 2차의 구분은 인터랙티브 시각화 영상의 버전에 따라 나뉜다. 연주 영상의 링크는 6-4 실연 과정에 소개되어있다.

Fig. 10. Screenshots of 1st performance with goodgeori (a) and jajinmori (b), and 2nd performance with goodgeori (c) and jajinmori (d). The difference between the two performances is the version of the interactive visualization video. The link for the video is introduced in section 6-4.

또한 2022년 12월에 인도 벵갈루루에서 개최된 음악 정보 학 분야의 저명한 국제학회인 ISMIR (International Society for Music Information Retrieval Conference)의 Music Program에 선정되어 인도 과학원(IISc; Indian Institute of Science)의 J N Tata Auditorium에서 본 작품의 실제 라이브 연주를 선보였다[43].

## VII. 결론

본 논문에서는 인공지능 모델을 활용하여 새로운 국악 선율을 작곡하고 이를 이용해 인간 연주자의 즉흥연주와 인공 본 연구의 의의는 전통적인 음악적 특성을 존중하면서도 새로운 창작 방법을 탐색하고자 함에 있다. 이러한 탐색은 인간 연주자와 인공지능과의 공존을 도모해 전통 예술과 현대 기술의 융합에 대한 새로운 가능성을 보여준다. 또한 시나위뿐만 아니라 다른 국악 곡에서의 창작에서도 참고할 수 있는 중요한 자료가 될 것으로 기대한다.

이번 연구에서는 시나위를 대상으로 딥러닝을 활용한 작곡 시스템을 구현했으나, 향후에는 다양한 국악 곡에 대한 작곡 시스템으로 확장이 가능하다. 또한 본 작업에서는 데이터셋 전처리와 멜로디에 대한 규칙 기반 방법을 사용하여 적합한 연주 기법에 따라 시김새와 농현을 구현했지만, 앞으로 선율 생성 모델이 시김새와 농현을 직접 모델링하는 방향으로의 추가적인 연구가 필요하다. 또한 딥러닝 기반 언어 모델을 이용하여 다양한 모델을 학습하기 위해서는 국악의 특징을 충분히 고려한 심볼릭 데이터셋의 구축이 필요하다는 것을 확인하였다. 현재 공개된 국악 관련 심볼릭 데이터셋이 없는 시점에 국악 전공자들과 관련 연구자들이 함께 국악 데이터셋의 형태와 기준을 고민하고 합의하여 정립하는 것은 향후 딥러닝을 활용한 국악 연구에 커다란 숙제이자 도약을 위한 발판일 것이다.

## 감사의 글

본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. NRF-2022R1F1A1 074566).

## 참고문헌

[1] M. Klingemann, Memories of Passersby I [Internet]. Available: <https://underdestruction.com/2018/12/29/memories-of-passersby-i/>

[2] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, ... and I. Sutskever, "Zero-Shot Text-to-Image Generation," in *Proceedings of the 38th International Conference on Machine Learning*, Vol. 139, pp. 8821-8831,

July 2021. <https://doi.org/10.48550/arXiv.2102.12092>

[3] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," in *Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans: LA, pp. 10674-10685, 2022.

[4] C. J. Carr and Z. Zukowski, "Generating Albums with SampleRNN to Imitate Metal, Rock, and Punk Bands," arXiv preprint, November 2018. <https://doi.org/10.48550/arXiv.1811.06633>

[5] DADABOTS. RELENTLESS DOPPELGÄNGER \m/ \m/ \m/ \m/ \m/ \m/ \m/ \m/ \m/ \m/ \m/ \m/ [Internet]. Available: <https://www.youtube.com/live/MwtVtPKx3RA/>

[6] OpenAI. MuseNet [Internet]. Available: <https://openai.com/blog/musenet/>

[7] ARS ELECTRONICA. Mahler-Unfinished [Internet]. Available: <https://ars.electronica.art/futurelab/en/projects-mahler-unfinished/>

[8] C. Huang, H. Koops, E. Newton-Rex, M. Dinculescu, and C. Cai, "AI Song Contest: Human-AI Co-Creation in Songwriting," arXiv preprint, October 2020. <https://doi.org/10.48550/arXiv.2010.05388>

[9] STHLM. Evigt Förlopad (Forever Lost) [Internet]. Available: <https://www.aisongcontest.com/participants/sthlm-2021>

[10] Maskinspelmanslaget. It's all about me [Internet]. Available: <https://www.aisongcontest.com/participants-2022/maskinspelmanslaget>

[11] B. Sturm and L. Casini, "Tradformer: A Transformer Model of Traditional Music," in *Proceedings of International Joint Conference on Artificial Intelligence*, Vienna, pp. 4915-4920, 2022. <https://doi.org/10.24963/ijcai.2022/681>

[12] Korean Convergence Arts Company (EMOTIONWAVE ARTS). Gugak Mechatronica [Internet]. Available: <https://arts.emotionwave.com/km/>

[13] The Korea Creative Content Agency. Music Meets AI [Internet]. Available: <https://edu.kocca.kr/edu/archiveUser/contentsDeptList.do?menuNo=500261&taskSeq=8>

[14] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, ... and K. Kavukcuoglu, "WaveNet: A Generative Model for Raw Audio," in *Proceedings of 9th ISCA Speech Synthesis Workshop*, Sunnyvale: CA, pp. 125-125, 2016. <https://doi.org/10.48550/arXiv.1609.03499>

[15] B. Sturm and O. Ben-Tal, Folk the Algorithms: (Mis)Applying Artificial Intelligence to Folk Music, in *Handbook of Artificial Intelligence for Music*. Cham: Springer International Publishing, pp. 423-454, 2021.

- [16] I. Goodfellow, “NIPS 2016 Tutorial: Generative Adversarial Networks,” arXiv preprint, December 2016. <https://doi.org/10.48550/arXiv.1701.00160>
- [17] G. Hadjeres, F. Pachet, and F. Nielsen, “DeepBach: a Steerable Model for Bach Chorales Generation,” in *Proceedings of the 34th International Conference on Machine Learning, PMLR*, Sydney, Australia, August 2017, pp. 1362-1371. <https://doi.org/10.48550/arXiv.1612.01010>
- [18] H. W. Dong, W. Y. Hsiao, L. C. Yang, and Y. H. Yang, “MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment,” in *Proceedings of the AAAI Conference on Artificial Intelligence, Virtual Event*, Vol. 32, No. 1, April 2018. <https://doi.org/10.1609/aaai.v32i1.11312>
- [19] L. C. Yang, S. Y. Chou, and Y. H. Yang, “MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation,” arXiv preprint, July 2017. <https://doi.org/10.48550/arXiv.1703.10847>
- [20] N. Trieu and R. Keller, “JazzGAN: Improvising with Generative Adversarial Networks,” in *Proceedings of the 6th MUME Workshop*, Salamanca, Spain, June 2018. <https://doi.org/10.5281/zenodo.4285166>
- [21] B. L. Sturm, J. F. Santos, O. Ben-Tal, and I. Korshunova, “Music transcription modeling and composition using deep learning,” in *1st Conference on Computer Simulation of Musical Creativity*, Huddersfield, UK, June 2016. <https://doi.org/10.48550/arXiv.1604.08723>
- [22] B. L. Sturm and H. Maruri-Aguilar, “The Ai Music Generation Challenge 2020: Double Jigs in the Style of O’Neill’s 1001,” *Journal of Creative Music Systems*, Vol. 5, pp. 1-26, December 2021. <https://doi.org/10.5920/jcms.950>
- [23] B. L. Sturm. The ai music generation challenge 2021 [Internet]. Available: <https://github.com/boblsturm/aimusic-generationchallenge2021>
- [24] B. L. Sturm. The ai music generation challenge 2022 [Internet]. Available: <https://github.com/boblsturm/aimusic-generationchallenge2022>
- [25] J. Luo, X. Yang, S. Ji, and J. Li, “MG-VAE: Deep Chinese folk songs generation with specific regional styles,” in *Proceedings of the 7th Conference on Sound and Music Technology (CSMT)*, Springer, pp. 93-106, 2020. [https://doi.org/10.1007/978-981-15-2756-2\\_8](https://doi.org/10.1007/978-981-15-2756-2_8)
- [26] J. Bae, “Korean Traditional Music Melody Generator using Artificial Intelligence,” *Journal of the Korea Institute of Information and Communication Engineering*, Vol. 25, No. 7, pp. 869-876, July 2021. <https://doi.org/10.6109/jkiice.2021.25.7.869>
- [27] D. Lee, M. L. Tran, and J. H. Jung, “Geometric Structure of Korean Music and AI Music Composition,” *Tongyang U’mak*, Vol. 51, pp. 63-98, June 2022. <https://doi.org/10.33452/AMRI.2022.51.63>
- [28] C. Donahue, H. H. Mao, Y. Li, G. Cottrell, and J. McAuley, “LakhNES: Improving multi-instrumental music generation with cross-domain pre-training,” in *Proceedings of International Society for Music Information Retrieval Conference (ISMIR)*, Delft, Netherlands, November 2019. <https://doi.org/10.48550/arXiv.1907.04868>
- [29] B. H. Lee, “A Study on the Rhythmic Pattern in Changdan of Korean Folk Music,” *Journal of the Asian Music Research*, Vol. 16, pp. 39-70, December 1994.
- [30] B. H. Lee, “Types of Rhythmic Patterns,” *Studies in Korean Music*, Vol. 23, pp. 26-131, December 1995.
- [31] B. H. Lee, “The shamanistic music of sinawi area,” *Korean Cultural Anthropology*, Vol. 4, pp. 79-86, December 1971.
- [32] F. Marchetti, C. Wilson, C. Powell, E. Minisci, and A. Riccardi, “Convolutional Generative Adversarial Network, via Transfer Learning, for Traditional Scottish Music Generation,” in *Proceedings of Artificial Intelligence in Music, Sound, Art and Design: 10th International Conference, EvoMUSART 2021, Virtual Event*, April 2021, pp. 187-202. [https://doi.org/10.1007/978-3-030-72914-1\\_13](https://doi.org/10.1007/978-3-030-72914-1_13)
- [33] C. Raffel, “Learning-Based Methods for Comparing Sequences, with Applications to Audio-to-MIDI Alignment and Matching,” Ph.D. dissertation, Columbia University, New York, NY, July 2016. <https://doi.org/10.7916/D8N58MHV>
- [34] Essen Folk. Essen Folk Song Database [Internet]. Available: <https://ifdo.ca/~seymour/runabc/esac/esacdatabase.html>
- [35] O’neill’s Dance Music of Ireland. [Internet]. Available: <http://john-chambers.us/~jc/music/book/ONEills/>
- [36] J. J. Bosch, R. Marxer, and E. Gómez, “Evaluation and combination of pitch estimation methods for melody extraction in symphonic classical music,” *Journal of New Music Research*, Vol. 45, No. 2, pp. 101-117, May 2016. <https://doi.org/10.1080/09298215.2016.1182191>
- [37] Orchset: a dataset for melody extraction in symphonic music recordings description [Internet]. Available: <https://www.upf.edu/web/mtg/orchset>
- [38] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving Language Understanding by Generative Pre-Training,” 2018.

[39] H. W. Dong, K. Chen, J. McAuley, and T. Berg-Kirkpatrick, "MusPy: A Toolkit for Symbolic Music Generation," in *Proceedings of International Society for Music Information Retrieval Conference (ISMIR)*, Virtual Event, October 2020.  
<https://doi.org/10.48550/arXiv.2008.01951>

[40] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the Properties of Neural Machine Translation: Encoder-Decoder Approaches," in *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation. Association for Computational Linguistics*, Doha, Qatar, pp. 103-111, October 2014.  
<https://doi.org/10.3115/v1/W14-4012>

[41] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Proceedings of 3rd International Conference on Learning Representations, ICLR 2015*, San Diego, CA, May 2015.  
<https://doi.org/10.48550/arXiv.1412.6980>

[42] Irworobongdo. Encyclopedia of Korean Culture [Internet]. Available: <http://encykorea.aks.ac.kr/>

[43] ISMIR 2022. Accepted Music Submissions [Internet]. Available: <https://ismir2022.ismir.net/program/music>



**오채령 (Chaeryeong Oh)**

2023년 : 서강대학교  
아트&테크놀로지학과  
(예술공학사)  
서강대학교 컴퓨터공학과  
(공학사)

※ 관심분야 : 음악 인공지능(Music AI), Creative Coding, Generative Music



**정다샘 (Dasaem Jeong)**

2015년 : KAIST 문화기술대학원  
(공학석사)  
2020년 : KAIST 문화기술대학원  
(공학박사)

2020년~2021년: SK 텔레콤  
2021년~현 재: 서강대학교 아트&테크놀로지학과 조교수  
※ 관심분야 : 딥러닝 (Deep Learning), 음악 인공지능 (Music AI), 음악 정보 검색 (Music Information Retrieval)



**박한나(Hannah Park)**

2018년~현 재 : 서강대학교  
컴퓨터공학과,  
아트&테크놀로지학  
과 학사과정

※ 관심분야 : 음악 정보 검색(Music Information Retrieval), 음악 인공지능(Music AI), Data Art, Creative Coding



**한단비내린(Danbinaerin Han)**

2021년 : 서울대학교 음악대학 (음악학사)

2022년~현 재 : 서강대학교 아트&테크놀로지학과 석사과정  
※ 관심분야 : 음악 정보 검색(Music Information Retrieval), 한국음악(Korean Traditional Music), 음악 인공지능(Music AI)