

글로벌 AIGC를 활용한 작곡 기술 발전 현황과 창작 플랫폼에 관한 사례연구

이 빙 옥¹ · 최 중 원^{2*}

¹중앙대학교 첨단영상대학원 영상학과 박사과정

²중앙대학교 첨단영상대학원 교수

A Case Study on the Development of Composition Technology and a Creative Platform in a Global AIGC Background

Bing-Yu Li¹ · Jongwon Choi^{2*}

¹Ph.D. Course, Graduate School of Advanced Imaging Science and Film, Chung-Ang University, Seoul 06974, Korea

²Professor, Graduate School of Advanced Imaging Science and Film, Chung-Ang University, Seoul 06974, Korea

[요 약]

본 연구는 AIGC를 활용한 인공지능 작곡 플랫폼의 사례를 분석함으로써 인공지능 작곡 기술의 발전 추세와 문제점을 파악하는 것을 목적으로 하였다. 사례분석을 통해 AI 작곡 기술의 발전 경향은 요약하면 다음과 같다. 확장된 창작 공간과 협업 측면에서 AI 작곡 플랫폼은 사용자가 정의한 감정, 스타일, 기간 등을 기반으로 음악을 생성하고, 사용자의 실시간 피드백을 반영하여 음악을 조정하는 기능을 제공한다. 고도의 맞춤화와 개인화 측면에서 AI 작곡 기술은 사용자의 행동, 환경, 심지어 생체 신호까지 분석하여 개인의 감정 상태나 상황에 맞는 음악을 실시간으로 생성할 수 있다. 사례분석을 통해 드러난 AI 작곡 기술의 발전 문제점 중 하나는 '창의성'을 표현할 수 있는 기술적 구현이 부족하다는 점이다. 인공지능 작곡 플랫폼을 통한 작곡 및 콘텐츠 제작, 그리고 문화예술교육 등의 융합화 현황을 조명함으로써, 향후 AI 작곡의 활용 가능성에 대한 기초자료를 제공하고자 한다.

[Abstract]

The purpose of this study was to identify the development trends and problems of artificial intelligence composition technology by analyzing cases of artificial intelligence composition platforms using AIGC. Through a case analysis, the development trends of artificial intelligence (AI) composition technology can be summarized as follows. In terms of an expanded creative space and collaboration, the AI composition platform provides the abilities to create music based on user-defined emotions, styles, and periods as well as to adjust the music by reflecting the user's real-time feedback. In terms of a high level of customization and personalization, AI composition technology can analyze the user's behavior, environment, and biological signals to generate music tailored to the individual's emotional state or situation in real time. One of the problems with the development of AI composition technology revealed through case analysis is the lack of technical implementation that can express "creativity". By highlighting the current status of the convergence of composition, content production, and cultural arts education through the artificial intelligence composition platform, we aim to provide basic data on the possibility of using AI composition in the future.

색인어 : AIGC, 음악 인공지능, 작곡 기술, 창작 플랫폼, 사례분석

Keyword : AIGC, Music AI, Composition Technology, Creative Platform, Case Analysis

<http://dx.doi.org/10.9728/dcs.2024.25.7.1677>



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 08 May 2024; **Revised** 28 May 2024

Accepted 19 June 2024

***Corresponding Author; Jongwon Choi**

E-mail: choijw@cau.ac.kr

1. 서론

2016년 1월, 스위스 다보스에서 열린 세계경제포럼에서는 4차 산업혁명의 시작을 공식적으로 선언하였다. 이 혁명은 메타버스, 인공지능, 빅데이터, NFT(Non-Fungible Token), 가상 및 증강현실 등의 기술 발전을 특징으로 하며, 이 시점으로부터 디지털로의 전환은 가속화되고 있다[1]. 이러한 기술적 진보는 음악 산업을 포함한 다양한 분야에서의 패러다임 변화를 촉진하고 있다. 음악 산업에서는 기획, 창작, 제작 방식뿐만 아니라 향유 및 소비 방식에도 혁신이 이루어지고 있다. 특히 한국에서는 인공지능 분야가 정부의 정책 지원과 대기업의 진출, 새로운 관련 기업의 등장 및 산업 수요의 증가에 힘입어 급속도로 성장하고 있다[2]. 또한 이러한 기술은 음악, 영상, 이미지 등의 분야에서 융합 콘텐츠 산업의 발전과 지속적인 연구를 촉진하고 있다.

2023년 초에 이루어진 GPT 4.0의 출시는 인공지능에 대한 대중의 관심과 주목도를 크게 증가시켰다. 이 플랫폼은 자연어 처리 기술을 통해 사용자의 질문에 대답하는 대화 형식의 접근 방식을 가능하게 하며, 이를 통해 사용자들은 AI가 제공할 수 있는 개인화된 경험, 효율성 향상 및 복잡한 문제 해결 능력에 더 큰 관심을 가지게 되었다[3]. 이러한 추세는 AI 기술의 발전을 가속화하는 중요한 요인으로 작용하고 있다. 또한, 기술의 발전은 인공지능 기술을 적용한 제품과 응용 융합의 범위를 확대시키고 있다. 특히, 동영상 공유 플랫폼을 중심으로 한 1인 미디어의 확산, OTT의 대중화로 인한 구독형 콘텐츠 및 자유 이용 허가 표시 저작물의 수요 증가는 음악, 영상, 디자인 등 창작에 필요한 다양한 구독형 미디어 콘텐츠 플랫폼의 시장 수요 증가를 확인할 수 있게 해준다[4].

음악 분야에서도 인공지능 작곡 소프트웨어가 화성학을 비롯한 이론 데이터를 기반으로 작곡 활동을 반복하고 평가하면서 스스로 진화하고 있다[5]. 이미 인공지능이 작곡한 음악을 인간 가수가 녹음하여 시장에 소개한 사례도 존재한다. 예를 들어, 2018년에 유튜브 스타이자 가수인 Taryn Southern은 인공지능을 사용하여 작곡된 최초의 앨범인 "I AM AI"를 발표하였다. 이 앨범은 AI 음악 제작 도구인 Amper Music을 사용하여 제작되었으며, 인공지능이 작곡한 음악에 Southern의 가사와 보컬이 결합된 결과물이었다. 또한, 사운드 스타트업 Endel은 AI를 활용하여 맞춤형 사운드 스케이프를 생성하고, 이 기술을 사용하여 여러 앨범을 출시하였다. 이들 앨범은 인간의 생체 리듬과 환경에 최적화된 음악을 제공하는 것을 목표로 하고 있다. 이러한 예시는 인공지능 기술이 음악 산업 내에서 어떻게 창의적인 활용이 이루어지고 있는지를 보여준다.

예술의 본질은 시대의 흐름과 사회의 변화에 따라 변모해간다. 이에 따라, 4차 산업혁명이 사회에 깊숙이 자리 잡은 미래에서는 인간과 기술이 밀접하게 결합되어 인간이 그 변화에 익숙해진 시대에서 인공지능에 의한 예술작품이 사회적

합의에 의해 예술성을 갖는 하나의 작품으로 인정받을 가능성이 있다[6],[7]. 반면, 기술의 발달로 지친 인간의 감성을 회복하려는 대중은 인간과 예술의 본질적 철학에 다시 접근할 수 있다. 이러한 접근을 통해 예술은 영원히 인류의 고유한 영역으로 남을 수 있다. 인공지능은 기계에 의한 사고 및 행동을 대표하며 창작의 영역에서 더욱 활발하게 작업될 것으로 예측된다. 따라서 4차 산업혁명이 도래한 현재, 새로운 사회 변화에 따른 인공지능의 예술 창작 가능 영역을 탐구하고 이에 대해 적극적으로 논의해야 할 시점이다.

본 연구는 AIGC를 활용한 인공지능 작곡 플랫폼의 사례를 분석함으로써 인공지능 작곡 기술의 발전 추세와 문제점을 파악하는 것을 목적으로 한다. 제2장에서는 선행연구를 토대로 AIGC에 대한 개요를 설명하고 인공지능을 활용한 작곡 과정을 검토한다. 제3장에서는 상용화된 인공지능 작곡 플랫폼의 사례를 조사하여, 해당 기술의 상용화 현황과 창작 범위를 분석한다. 이어지는 제4장에서는 사례 분석을 통해 인공지능 작곡 기술의 발전 경향과 그에 따른 문제점을 심층적으로 분석하고자 한다.

II. 선행연구

2-1 글로벌 AIGC 환경의 이해

Artificial Intelligence Generated Content (AIGC)는 인공지능이 생성한 콘텐츠를 의미한다. AIGC는 텍스트, 이미지, 음악, 비디오 등 다양한 미디어 형태의 콘텐츠를 포함하며, 인공지능 기술을 활용해 인간의 창조성 또는 데이터에 기반한 새로운 콘텐츠를 자동으로 생성한다[8]. AIGC 기술은 기계학습(ML), 자연어 처리(NLP), 컴퓨터 비전 등 여러 인공지능 분야의 발전을 바탕으로 빠르게 성장하고 있다. 이 기술들은 개인화된 콘텐츠 제작, 대량 데이터 분석을 통한 창의적 아이디어 제공, 사용자 경험의 향상을 목적으로 다양하게 활용될 수 있다[9].

AIGC 창작의 계층적 과정은 그림 1과 같다[10]. 기본층(Basic Layer)은 시스템의 기반 인프라와 자원을 제공하는 것이다. 이 층에는 클라우드-엣지 컴퓨팅을 통한 강력한 계산 능력이 포함되어 있으며, 이를 통해 데이터 소스에 가까운 곳에서 데이터 처리를 할 수 있어 지연 시간과 대역폭 사용을 줄일 수 있다. 대규모 데이터는 규범에 맞는 다양한 모달 데이터로 구성되어 있으며, 데이터 주석과 필터링을 통해 시스템에 입력되는 정보의 품질과 관련성을 보장한다. 또한, 감독 학습 및 비감독 학습 알고리즘, TensorFlow 및 Pytorch와 같은 여러 AI 프레임워크, 변환기와 확산 모델과 같은 생성 AI 모델 등 다양한 알고리즘이 포함되어 있다. 이러한 알고리즘을 지원하기 위해 AI 칩과 대용량 저장장치와 같은 필수 하드웨어 지원이 포함되어 있다.

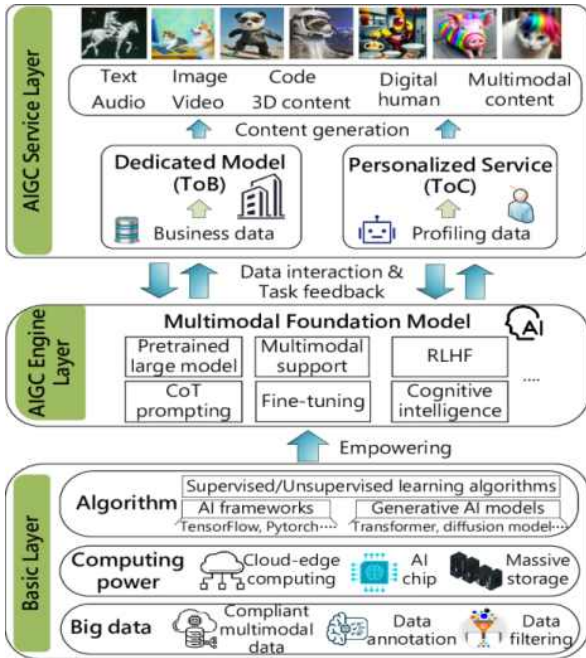


그림 1. AIGC 창작과정
Fig. 1. AIGC creation process

중간층은 AIGC 엔진층(Engine Layer)으로서, 프레임워크의 핵심부분이며 다양한 모델 작업 수행 능력을 제공한다. 이 층은 사과의 연쇄(CoT) 프롬프팅 방식을 활용한 사전 훈련된 대규모 모델을 포함하고 있으며, 특정 작업에 적합하게 미세 조정이 가능하다. 다양한 모델 지원을 통해 시스템은 텍스트, 이미지, 오디오 등 다양한 형태의 콘텐츠를 처리하고 생성할 수 있다. 추가적으로, 인간 피드백을 통해 학습하는 강화 학습(RLHF) 및 인간 인지과정을 모방하여 AI의 의사결정 및 상호작용 품질을 개선하는 인지 지능 기술도 포함되어 있다. 이 층의 핵심 기능은 "역량 부여"로, 최상위 층의 서비스를 위한 기술적 및 데이터 지원을 제공한다.

최상위 층은 AIGC 서비스층(Service Layer)으로, 사용자와 직접 소통하며 맞춤형 콘텐츠 생성 서비스를 제공한다. 이 층은 기업 고객을 위한 전용 모델(Dedicated Model ToB)과 소비자 고객을 위한 개인화된 서비스(Personalized Service ToC)로 구분된다. 기업용 서비스는 비즈니스 데이터를 기반으로 전문 콘텐츠를 생성하는 반면, 소비자용 서비스는 사용자 개인 데이터를 바탕으로 콘텐츠를 맞춤 설정한다. 서비스 층은 텍스트, 이미지, 오디오, 비디오, 코드, 3D 콘텐츠, 디지털 인간, 멀티모달 콘텐츠 등 다양한 형식의 콘텐츠 생성을 지원한다. 사용자 데이터와 작업 피드백의 상호작용을 통해 서비스 층은 지속적으로 개선될 수 있다.

2-2 인공지능 작곡 기술의 이해

인공지능 작곡 기술은 컴퓨터 알고리즘을 활용하여 음악을 자동으로 생성하는 기술이다. 이 과정은 데이터 수집, 기계학

표 1. 인공지능 작곡의 과정
Table 1. Process of AI composition

Step	Description
Data Collection	Collect musical data necessary for training the AI model. This data can come in formats like MIDI files or audio files and includes various musical elements such as melody, harmony, rhythm, and dynamics. Data may be preprocessed for optimization.
Model Training	Train the AI model using the collected data, often employing deep learning techniques. Algorithms like Recurrent Neural Networks (RNN), Variational Autoencoders (VAE), and Generative Adversarial Networks (GAN) are commonly used to learn musical patterns and structures.
Music Generation	Use the trained model to create new music. The AI synthesizes original melodies and harmonies based on its learned knowledge and combines them to produce complete musical compositions. Adjustments can be made to customize the music to specific requirements such as tempo and key.
Post-Processing	Review and refine the AI-generated music. Human musicians may intervene to enhance emotional depth or correct aspects that the AI may have overlooked, ensuring the final product is polished and suitable for its intended audience.

습, 음악생성, 후처리의 4 단계로 나누어진다(표 1).

첫 번째 단계인 데이터 수집(Data Collection)에서는 기계 학습 알고리즘이 필요로 하는 음악적 데이터를 수집한다. 이러한 데이터는 MIDI 혹은 오디오 파일의 형태로, 다양한 음악 장르 및 스타일에 따른 멜로디, 화음, 리듬, 동기와 같은 음악적 요소들을 포함한다. 데이터는 종종 전처리 과정을 거친다. 전처리 과정에는 AI가 생성한 음악을 오디오 파일로 추출한 후, 배경 노이즈 제거를 통해 신호 대 잡음비를 개선하고, 불균형 정규화를 통해 음량의 일관성을 유지하며, 샘플링 레이트 조절을 통해 모든 오디오 파일의 일관된 분석을 가능하게 한다. 이어서 파일 형식 변환을 통해 분석에 용이한 형태로 바꾸고, 악기 트랙 분리를 통해 멜로디, 화음, 리듬 등의 요소를 개별적으로 분석할 수 있게 한다. 마지막으로, 시간축 정렬을 통해 템포와 박자에 맞게 데이터를 정리하여 학습에 적합하도록 조정한다.

기계학습(Model Training) 단계에서는 수집된 데이터를 기반으로 인공지능 모델이 훈련을 진행한다. 이 단계에서는 주로 딥러닝 기법이 적용되며, 순환신경망(RNN), 변형 오토인코더(VAE), 생성적 적대 신경망(GAN) 등의 다양한 알고리즘을 활용한다[11]. 이러한 모델들은 음악 데이터의 구조적 및 시계열적 특성을 학습함으로써 음악적 패턴을 인식하고 이를 이해하게 된다.

세 번째 단계인 음악 생성(Music Generation) 단계에서는 훈련된 모델을 활용하여 실제 음악을 생성한다. 이 과정에서 인공지능은 학습된 지식을 바탕으로 창의적인 멜로디나 화음을 구성하고, 이를 통합하여 완성도 높은 음악 작품을 제작한다. 생성된 음악은 사용자의 요구에 따라 템포, 조성, 계리

등 다양한 파라미터를 조정하여 맞출 수 있다.

마지막 후처리(Post-Processing) 단계에서는 인공지능에 의해 생성된 음악을 작곡가가 검토하고 수정하거나 향상시키는 작업을 진행한다. 이 단계에서 음악가들은 작품에 더욱 깊은 감성을 부여하거나, 인공지능이 간과한 부분을 보완함으로써 최종적으로 소비자에게 제공될 음악의 품질을 향상시킨다. 인공지능 작곡 기술은 그 활용 범위가 매우 넓어 영화, 광고, 게임 음악 제작뿐만 아니라 개인 음악 창작자들에게도 도구로서의 가능성을 제공한다. 이러한 기술은 전통적인 음악 창작 방식에 혁신을 가져오고 있어, 음악 산업에 새로운 전망을 제시하고 있다.

인공지능 기술의 발전은 문화예술 분야에서 창작 데이터를 처리하여 콘텐츠를 생성하는 방식에 혁신을 가져왔다. 특히 MIDI 데이터의 추출과정은 오디오 소스로부터 MIDI 정보를 분리하거나 음악 연주를 분석하여 MIDI 데이터를 생성하는 것을 포함한다. 이 과정은 선택된 오디오 파일이나 라이브 음악에서 음표의 높이, 지속 시간, 강도 등을 정확하게 식별하고 분석하는 것으로 시작된다. 이후 특수 개발된 소프트웨어를 사용하여 이러한 정보를 MIDI 메시지로 변환하고 오류를 수정하여 데이터를 정제함으로써 최종 MIDI 파일의 정확도를 향상시킨다. 추출된 MIDI 데이터는 추가적인 편집과 최적화 과정을 거쳐 음악 제작 소프트웨어에서 사용될 수 있으며, 가상 악기와 통합을 통해 더욱 다양하고 세밀한 음악적 표현이 가능하게 된다. 이와 같은 음악 및 사운드 데이터 처리와 생성을 위한 대표적인 모델로는 마젠타 스튜디오(Magenta Studio), 뮤즈넷(MuseNet), 엔신스(NSynth) 등이 있으며, 음성인식 및 합성을 위해서는 스크립블링 스피치(Scribbling Speech), 명 트랜스레이터(Thing Translator) 등이 사용되고 있다. 이러한 모델들은 사운드 및 MIDI 데이터를 처리하고 학습 데이터를 기반으로 딥러닝 기술을 활용하여 지속적으로 기술을 발전시키고 있다.

기계학습을 활용한 음악생성에 대한 복잡한 신경망 아키텍처를 설명하는 그림 2는 세 가지 주요 부분으로 구성되어 있다[12]. 이 아키텍처에는 생성기(Generator), 컨트롤 네트워크(Control Network), 그리고 확률 네트워크(Probability Network)가 포함되어 있으며, 각각의 구성 요소는 음악 작곡 과정에서 중요한 역할을 수행한다. 생성기 부분은 LSTM(Long Short-Term Memory) 블록을 포함하여, 음악 작곡에 있어 필수적인 시퀀스 학습과 기억 능력을 제공한다. 이 LSTM 블록은 'Actor'로 명명되어, 장기 의존성을 지닌 데이터인 음악에 필요한 패턴을 학습하는 데 중요한 역할을 한다. 'Actor'는 'Score'라는 평가 메커니즘을 통해 특정 음악 규칙에 기반한 평가를 받고, 이를 통해 자신의 생성 과정을 조정하게 된다. 컨트롤 네트워크는 'Music Rules' 데이터베이스를 통해 생성된 음악이 특정 음악적 기준을 만족하는지 평가하는 역할을 한다. 이 네트워크에서 'Critic'이라 불리는 부분은 생성된 음악의 음악성을 평가하고, 'Score' 형태로 생성기에게 보상이나 피드백을 제공한다. 이러한 점수는 생성기가

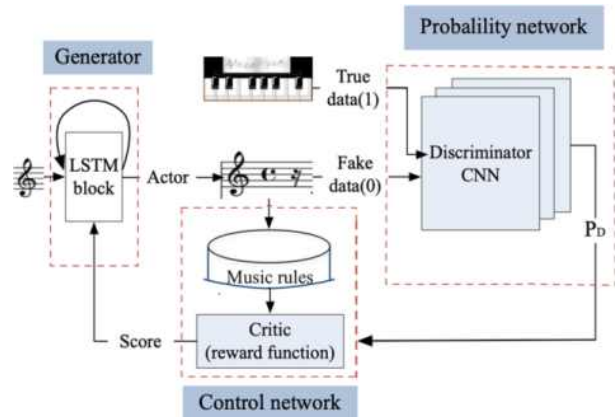


그림 2. GAN의 시스템 다이어그램
Fig. 2. System diagram of GAN

더욱 음악적인, 즉 인간 작곡가가 만들었을 법한 음악을 생성할 수 있도록 유도하는 기능을 수행한다. 이 아키텍처는 음악 창작 과정에서 인공지능이 어떻게 인간의 창의성을 보완하고 확장할 수 있는지를 보여주는 중요한 예시이다.

확률 네트워크는 복잡한 CNN(Convolutional Neural Network)을 활용하며, 이 내부에 'Discriminator'라 불리는 구성 요소를 포함하고 있다. 이 구성 요소는 생성된 음악(가짜 데이터)과 실제 음악(진짜 데이터)을 구분하는 기능을 수행한다. 이 판별자는 생성된 음악이 실제 음악에 얼마나 가까운지를 확률 P_D 로 나타내며, 생성기가 만든 음악이 실제 음악과 구별되지 않도록 하는 학습 신호를 제공한다. 이 아키텍처는 생성적 적대 신경망(GAN, Generative Adversarial Network)의 원리를 따른다. 생성기는 실제 음악과 구별하기 어려운 음악을 생성하려 하고, 판별자는 생성기가 만든 음악이 창작한 음악인지 AI 음악인지를 판별하려 한다. 이 두 네트워크는 서로 경쟁하면서 생성기는 점점 더 나은 음악을 만들어내는 능력을 개발하게 된다. 컨트롤 네트워크는 이러한 경쟁 과정을 보완하며, 음악적 규칙을 바탕으로 생성기가 음악적 가치가 있는 결과물을 만들도록 유도하는 역할을 한다. 이러한 과정은 GAN이 음악 창작의 다양한 모델에서 주목받는 이유를 설명해준다. GAN은 생성자와 판별자가 경쟁적으로 학습하는 구조를 통해 매우 현실감 있고 고품질의 데이터를 생성할 수 있으며, 이러한 상호 작용은 생성자의 성능을 지속적으로 개선시킨다. 이러한 우수성 때문에 GAN은 음악 생성 분야에서 다른 모델들보다 더 많이 설명되고 연구되는 경향이 있다.

III. 인공지능 작곡 플랫폼

최근 대기업들은 인공지능 기술을 활용한 창작 플랫폼을 출시하면서 대중의 관심을 이끌었다. 구글, 아마존, 바이트댄스와 같은 기업들이 이러한 플랫폼을 개발하여 인공지능의

작곡 능력을 공개하였다. 이와 동시에, 앰퍼뮤직(Amper music), AIVA, 주크 덱(Juke Deck) 등의 스타트업들은 인공지능 작곡 기술을 상용화하기 위한 다양한 사업 계획을 추진하고 있다. 또한, 텐센트와 바이도 같은 기업들은 인수 및 협력을 통해 이 기술의 연구 및 개발을 지속하며 상용화 전략을 모색하고 있다. 이러한 움직임은 인공지능 기술의 발전과 그 가능성을 널리 알리는 데 기여한다.

표 2와 같이 인공지능을 활용한 작곡 플랫폼은 세 가지 주요 유형으로 구분될 수 있다. 첫 번째 유형인 자동 작곡(Automated Composition)은 인간의 최소한의 개입을 통해 AI가 독립적으로 음악을 생성하는 형태로, 대규모 데이터셋에서 학습된 패턴과 복잡한 알고리즘을 기반으로 고품질의 작품을 만들어낸다. 두 번째 유형인 협력적 작곡(Collaborative Composition)에서는 AI가 초기 아이디어나 음악적 요소를 제공하고, 인간 아티스트가 이를 바탕으로 창의적인 결정을 내리며 작품을 완성한다. 이 과정에서 AI는 창작의 동반자 역할을 수행한다. 마지막으로 적응적 작곡(Adaptive Composition)은 사용자의 행동, 상황, 또는 환경에 따라 음악이 실시간으로 변화하는 방식으로, 주로 비디오 게임, 인터랙티브 미디어, 개인화된 음악 경험 등에 활용된다. 이 세 가지 유형은 각각 다양한 응용 분야에서 인공지능의 가능성을 보여주며, 음악 창작의 새로운 지평을 열고 있다.

표 2. 인공지능 창작의 3가지 유형
Table 2. 3 Types of AI composition

Type	Description	Platform
Automated Composition	AI generates music independently with minimal human intervention.	AIVA Technologies, Amper Music, Jukedeck
Collaborative Composition	AI and humans work together to create music.	Watson Beat, Magenta, AWS Deep Composer
Adaptive Composition	AI dynamically changes music based on user behavior, situation, or environment.	Endel, Mubert

3-1 자동 작곡 플랫폼

1) AIVA

AIVA[13]는 다양한 장르와 악기를 활용하여 음악 작곡이 가능한 창조 도구로서, 룩셈부르크와 영국에 기반을 둔 스타트업 AIVA 테크놀로지에 의해 2016년 개발되었다. 이 인공지능 시스템은 프랑스 및 룩셈부르크의 작곡가 권리 협회인 SACEM에 의해 인증된 첫 번째 AI 작곡가로, 고도의 딥러닝 모델과 순환 신경망을 통해 방대한 음악 데이터를 분석하여 멜로디, 하모니, 리듬 등 음악적 요소를 학습한다. AIVA는 클래식 음악 데이터와 베토벤, 바흐, 모차르트 등 유명 작곡가들의 악보를 바탕으로 훈련되었으며, 이를 통해 클래식 앨범을 발매하고 2017년에는 프랑스 아비뇽 국립 오케스트라와 협

력하여 AI가 작곡한 음악을 연주하는 등 주목을 받았다. 더불어, 2018년에는 세계적인 반도체 회사인 엔비디아와 협업하여 GPU 및 딥 뉴럴 네트워크를 활용, 록 그룹 퀸의 음악 스타일을 분석하며 큰 관심을 이끌었다. AIVA의 사용 용이성과 유연성 덕분에 음악 산업에서 혁신적 도구로 평가받고 있으며, AI와 창의성이 결합된 새로운 음악 창작의 가능성을 모색하고 있다. AIVA는 전문 음악가뿐만 아니라 음악을 취미로 즐기는 사람들에게도 매력적인 옵션으로 자리 잡고 있으며, 기술 발전과 함께 지속적으로 진화할 것으로 기대된다.

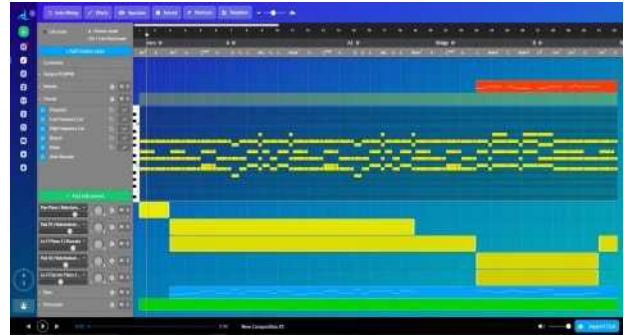


그림 3. AIVA의 인터페이스[13]
Fig. 3. Interface of AIVA[13]

그림 3에서는 AIVA의 인터페이스가 소개되며, 이는 디지털 오디오 워크스테이션(DAW)과 유사하게 사용자가 인공지능이 생성한 음악 데이터를 편집할 수 있는 기능을 제공한다. 이 인터페이스를 통해 사용자는 자신의 음악 창작 활동에 필요한 데이터를 수집하고 조작할 수 있다. 또한, AIVA는 음악 데이터의 세부 조정이 가능하며, 옥타브, 샘플링, 트랙별 볼륨, 좌우 팬 값 등을 조절할 수 있는 기능을 포함한다. AIVA는 다양한 프리셋 스타일을 제공하여 사용자가 자신의 멜로디나 화성 아이디어를 바탕으로 음악을 생성할 수 있도록 한다. 제공되는 스타일에는 모던 시네마틱(Morden Cinematic), 오케스트라(Ambient Orchestral), 20세기 시네마틱(20th Century Cinematic), 뱃노래(Sea Shanty), 신스웨이브(Synthwave), 탱고(Tango), 사이버펑크(Cyberpunk), 차이니즈(Chinese) 등이 포함된다. 사용자는 조성(단조 또는 장조), 박자(Time Signature)와 같은 음악적 요소를 선택할 수 있으며, 심포닉 오케스트라, 현악과 관악(Strings & Brass), 피아노와 현악(Piano & Strings) 등의 편성을 선택하여 곡의 길이를 30초 단위로 설정할 수 있다. 이러한 기능을 통해 AIVA가 생성한 음악 데이터는 다운로드 받아 사용자가 자신의 DAW를 이용하여 추가적인 창작 활동을 할 수 있도록 하며, 이는 인공지능과의 협업을 통한 창의적 과정을 가능하게 한다.

2) Amper Music

Amper[14]는 2014년에 설립된 인공지능 작곡을 전문으로 하는 스타트업으로, 개발자 API 및 인공지능 기반 작곡 플

랫폼인 Amper Music을 개발하여 제공하고 있다. 이 플랫폼은 세계 최초로 기업형 콘텐츠 크리에이터들을 대상으로 한 것으로, 인공지능이 생성한 음악 콘텐츠를 영상, 팟캐스트 및 기타 다양한 매체와의 융합을 통해 활용할 수 있는 점을 특징으로 한다. 특히, Amper는 독자적으로 개발한 수백만 개의 샘플과 가상 악기를 이용하여 고품질의 창작 데이터를 오디오 신호로 추출하는 기술을 보유하고 있다. 또한, 사용자는 Amper의 워크 프로세스를 통해 생성된 원곡을 다양한 음악 스타일로 변환하여 추출할 수 있는 기능을 활용할 수 있다.

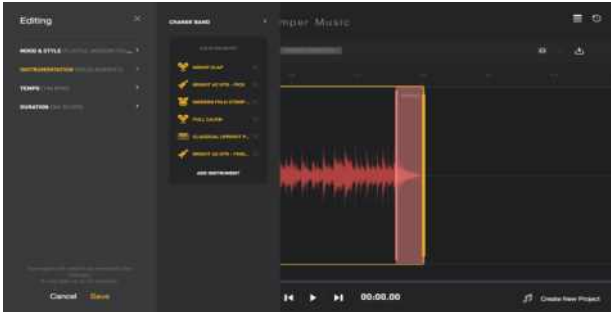


그림 4. Amper Music의 인터페이스[14]
Fig. 4. Interface of Amper Music[14]

위의 그림 4에서는 Amper Music의 편집기능을 소개하고 있으며, 이 기능을 통해 사용자는 인공지능이 창작한 음악을 효과적으로 편집할 수 있다. Amper Music은 다양한 스타일과 장르의 음악을 제공하며, 사용자는 대시보드(Dashboard)를 통해 원하는 음악 스타일을 미리 청취하고 자신의 프로젝트에 적합하게 활용할 수 있다. 사용자는 필요에 따라 음악의 길이를 조절하거나, 특정 영상에 맞는 길이의 음악을 생성할 수 있는 기능을 이용할 수 있다. 또한, 다큐멘터리(Documentary), 시네마틱(Cinematic)와 같은 영상 스타일 또는 클래식(Classical), 포크(Folk), 힙합(Hip hop), 재즈(Jazz) 등 다양한 음악 장르를 선택할 수 있다. 이후 사용자는 세부 장르, 분위기 설정을 통해 자신이 원하는 음악 스타일을 더욱 세밀하게 구현할 수 있으며, 참고 음악을 활용하여 구체적인 작곡 스타일을 결정할 수 있다. Amper Music은 AI로 작곡된 원본 음원 파일을 사용자가 직접 편집할 수 있는 기능도 제공한다. 사용자는 자신이 보유한 음원 파일과 혼합하거나, 새로운 가상 악기 음향을 추가하고 BPM(Beat Per Minute)이나 조표를 변경하여 개성 있는 곡을 조합할 수 있다. 이러한 기능들은 사용자가 자신의 창의적 비전을 실현할 수 있게 하며, 음악 창작과정에서의 융통성을 높인다.

3) Jukedeck

Jukedeck[15]은 2012년 영국에서 설립된 인공지능 기반 음악 작곡 서비스 스타트업으로 작곡가, 프로듀서, 엔지니어 등으로 구성된 개발 팀이 머신러닝 기술을 활용하여 개발하였다. Jukedeck은 13개의 음악 장르에 걸쳐 1만 곡 이상의

학습 데이터를 기반으로 음악 콘텐츠를 생성한다. 핵심 기술에는 작곡 엔진과 프로덕션 엔진이 포함되어 있다. 작곡 엔진에서는 딥러닝과 신경망 기술을 사용하여 화성 구조와 같은 음악적 요소들을 학습하고 이를 기반으로 음악 데이터를 완성한다. 이어서 그림 5와 같이 프로덕션 엔진에서는 연주 형태와 악기 편성을 선택하고, 이퀄라이저 및 컴프레서와 같은 다양한 이펙터를 사용하여 최종 음악 콘텐츠를 생성한다. 이 과정을 통해 몇 분 내에 특정 스타일과 감정을 반영한 독창적인 음악 작품을 창작할 수 있다. 또한, Jukedeck은 사용자에게 생성된 음악의 저작권을 부여하고, 추가적인 저작권 비용 없이 다양한 프로젝트에 자유롭게 사용할 수 있는 권한을 제공한다. 이 플랫폼은 창의적이고 상업적인 목적으로 맞춤형 음악을 생성하는 혁신적인 AI 음악 생성 도구로서, 사용자의 아이디어, 이야기 및 감정 등 다양한 요소를 음악으로 변환하는데 도움을 주는 역할을 한다.

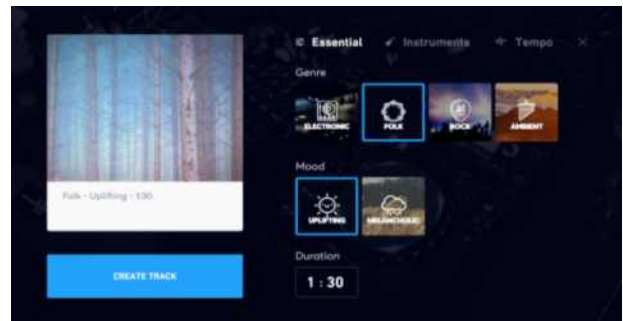


그림 5. Jukedeck의 인터페이스[15]
Fig. 5. Interface of Jukedeck[15]

3-2 협력적 작곡 플랫폼

1) Watson Beat

IBM Watson Beat은 IBM의 인공지능 플랫폼인 Watson을 기반으로 개발된 음악 창작 도구이다. 이 도구는 사용자가 지정한 감정, 스타일, 기간 등의 다양한 변수를 입력받아, 이를 기반으로 독특한 음악을 자동 생성할 수 있는 기능을 제공한다. Watson Beat은 인공지능의 강력한 학습 능력을 활용하여 음악 창작 과정의 각 단계에 깊이 관여할 수 있게 하며, 이를 통해 최종 음악 작품이 사용자의 창의적 및 감정적 의도를 정확히 반영할 수 있도록 한다.

Watson Beat은 사용자가 생성할 음악에 대한 구체적인 요구 사항을 입력할 수 있는 상세한 대화 시스템을 제공한다. 사용자는 음악이 전달해야 할 감정(예: 즐거움, 슬픔, 신비로움), 음악 스타일(예: 록, 재즈, 전자음악), 리듬의 빠르기, 볼륨의 세기 등을 지정할 수 있다. 이러한 입력은 단순한 선택을 넘어서 음악의 특정 특성, 예를 들어 조화의 복잡성이나 멜로디의 유창성 등을 더 구체적으로 설명할 수 있다. 음악 창작 과정에서 Watson Beat은 실시간 피드백을 제공하여 사용자가 AI가 생성한 음악이 자신들의 기대에 부합하는지 즉

시 듣고 평가할 수 있게 한다. 이러한 실시간성은 사용자가 필요한 경우 즉시 조정 파라미터를 수정하여 AI가 음악 방향을 수정하도록 지시할 수 있게 해 준다.

기본 음악 속성 외에도 Watson Beat은 사용자가 악기 선택, 음표 배치, 음높이 변화 등을 깊이 있게 맞춤 설정할 수 있도록 한다. 사용자는 특정 스타일을 강조하기 위해 특정 악기 파트를 선택하거나, 비디오나 기타 미디어 콘텐츠의 필요에 더 잘 맞도록 음악의 특정 부분을 조정할 수 있다. 이와 같은 고도의 상호작용 창작 프로세스를 통해 IBM Watson Beat은 단순한 자동 음악 생성 도구를 넘어 창의적 협업 플랫폼으로 기능하며, 사용자의 상상과 요구에 더욱 근접한 음악 작품을 창출할 수 있다. 이는 전통적인 음악 제작 과정을 혁신적으로 변화시키며 음악 창작을 더욱 민주화하고 개인화하는 경험을 제공한다.

2) Magenta

Google Magenta[16]는 창작 과정에서 기계 학습의 역할을 탐구하는 오픈 소스 연구 프로젝트이다. 음악과 미술 생성에 기계 학습을 적용하는 것을 주된 목적으로 하며, 이를 위해 새로운 딥러닝과 강화 학습 알고리즘을 개발한다. Magenta Studio는 이 프로젝트의 결과물 중 하나로, Ableton Live와 원활하게 통합되도록 업그레이드된 음악 창작 도구 모음이다. 사용자들은 이 도구를 사용하여 AI가 만들어낸 음악적 아이디어를 바탕으로 자신만의 음악을 만들 수 있다. Magenta Studio는 Ableton Live 플러그인 형식과 독립 실행형 애플리케이션 형식 두 가지로 제공되며, 두 형태 모두 무료로 이용할 수 있다.

Magenta가 사용하는 기술 중 하나인 MusicVAE는 반복적인 변형 오토인코더로, 양방향 LSTM 셀로 이루어진 순환 신경망(RNN)과 표준 인코더 및 계층적 디코더를 포함한다. 이 기술은 MIDI 파일로부터 단선 멜로디 라인이나 폴리포닉 드럼 패턴을 생성하는 데 사용된다. 그림 6과 같이 Megent studio에서 제공되는 기능이다. 'Continue' 기능은 사용자가 시작한 멜로디를 AI가 자동으로 이어서 완성하게 해주고, 'Generate' 기능은 사용자의 입력을 바탕으로 새로운 멜로디를 생성한다. 또한, 'Interpolate' 기능은 두 개의 서로 다른 멜로디를 AI가 분석하여 중간 멜로디를 생성하며, 이는 전혀 새

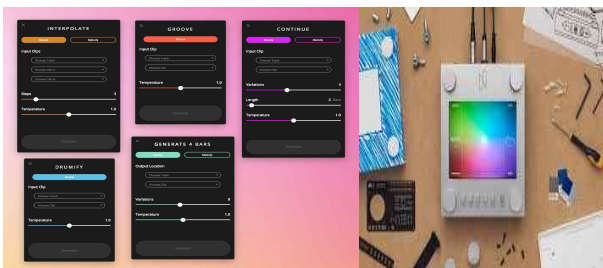


그림 6. Megent studio의 인터페이스와 NSynth Super[16]
Fig. 6. Interface of Megent studio and NSynth Super[16]

로운 창작물을 탐색하는 데 매우 유용하다. 이처럼, Magenta는 완전한 음악 제작 솔루션이 아니라, 창작 또는 음악 제작 과정에서 사용할 수 있는 도구로서의 역할을 한다. 또한, Megent Studio는 딥러닝 데이터를 활용하여 인공지능 모델인 NSynth(Neural Synthesizer)를 학습시켜 음악 데이터를 생성한다. NSynth는 딥러닝 기반의 신경 오디오 합성 모델로, 기존 음향에서는 구현할 수 없었던 새로운 소리를 창출한다. 특히, NSynth Super는 물리적 인터페이스를 통해 음악가가 직접 조작하며 소리를 탐색할 수 있는 기능을 제공한다.

3) AWS Deep Composer

AWS DeepComposer[17]는 Amazon Web Services가 제공하는 기계학습 및 생성적 인공지능 기술을 활용하여 사용자가 멜로디를 입력하고, 사전에 훈련된 AI 모델을 통해 새로운 음악 작품을 생성할 수 있도록 지원하는 서비스이다. 이 서비스는 생성 AI, 즉 생성적 인공지능 기술로 개발되었으며, GAN(생성적 적대 신경망)을 사용하여 학습 데이터를 생성하고 최적화한다. 이는 두 개의 신경망이 결합하여 콘텐츠를 창출하는 방식으로, 사용자의 창작 아이디어 및 창작 분야에 관련된 기본 매개변수들을 입력받아 인공지능이 콘텐츠를 창작해 나가는 과정을 가능하게 한다.

AWS DeepComposer는 뮤직 스튜디오와의 연동을 통해 사전 학습된 모델들을 모니터링하고 창작 활동에 적용할 수 있는 기능을 제공한다. 사용자는 콘솔을 통해 모델의 하이퍼파라미터 조정, 학습 속도 설정 등을 최적화할 수 있다. 또한, 아마존의 클라우드 기반 기계 학습 플랫폼인 Amazon SageMaker를 활용하여 머신러닝 모델을 생성하고, 맞춤형 GAN 아키텍처를 구축할 수 있다. 이 서비스는 사용자가 원하는 DAW(Digital Audio Workstation)과 연동되며, 인공지능을 활용한 작곡 활동을 지원한다. 건반을 통해 입력된 간단한 멜로디에 대해, AWS의 음악 관련 학습 데이터를 기반으로 나머지 악기구성을 인공지능이 자동으로 완성해주며, 음악의 템포 설정, 레코딩, 자동 아르페지오 생성 및 코드 생성기 등을 이용하여 창작활동을 할 수 있다.

AWS DeepComposer는 그림 7에 제시된 바와 같이 뮤직 스튜디오 인터페이스를 통해 USB를 활용하여 PC와 연결되

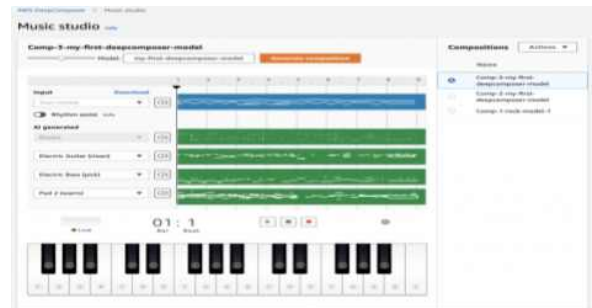


그림 7. AWS DeepComposer의 인터페이스[17]
Fig. 7. Interface of AWS DeepComposer[17]

며, 이를 통해 건반과 연동하여 사용이 가능하다. 사용자는 기존에 생성된 MIDI 데이터를 입력하거나 자신의 연주를 AI와 연계하여 작곡하는 작업을 수행할 수 있다. 멜로디나 화성을 입력하면, 이를 기반으로 AI가 멜로디, 화성, 편곡 등의 음악적 요소를 조합하여 사용자가 요구하는 창작물을 생성한다. 또한 사용자는 록, 팝, 재즈, 클래식 등 선호하는 음악 스타일 및 장르에 맞춰 학습 데이터를 활용하여 음악 창작 모델을 학습시킬 수 있다. 이 과정은 음악 연주나 작곡에 대한 사전 지식이 없는 사용자도 자신의 멜로디 아이디어를 선호하는 스타일로 음악 신호를 생성할 수 있게 해준다.

3-3 적응적 작곡

1) Endel

Endel 그림 8은 2018년에 설립되었으며, 독일 베를린에 본사를 두고 있다. 사용자의 상황과 상태에 맞추어 실시간으로 개인화된 사운드 환경을 제공하며, 집중력 증진, 이완, 수면 개선 등을 돕기 위해 설계되었다. Endel의 핵심특징 중 하나는 사용자의 상황과 요구에 맞게 실시간으로 개인화된 음악을 생성할 수 있는 알고리즘이다. 이 알고리즘은 사용자의 위치, 시간, 날씨 상태, 생리 데이터(예: 심박수, 활동 수준) 등 다양한 정보를 수집한다.



그림 8. Endel의 인터페이스
Fig. 8. Interface of Endel

또한, Endel은 음악이 감정, 스트레스 수준, 생리적 반응에 어떻게 영향을 미치는지를 연구하는 음악 심리학 및 음악 치료 분야의 연구에 기반하여 설계되었다. 특정 멜로디, 리듬, 화음을 사용하여 특정 감정 반응을 유발하거나 억제할 수 있다. 예를 들어, 이완 및 수면을 촉진하기 위해 느린 리듬과 부드러운 멜로디를 사용하거나, 경계성과 에너지를 높이기 위해 빠른 리듬과 복잡한 화음을 사용할 수 있다. Endel의 음악은 사전 녹음된 트랙을 재생하는 것이 아니라 알고리즘을 통해 실시간으로 생성된다. 이는 음악이 사용자의 즉각적인 피드백에 따라 조정될 수 있다. 예를 들어, 사용자의 심박수 데이터가 이완되지 않았다는 것을 보여주면 알고리즘이 음악의 일부 요소(예: 리듬을 낮추거나 음조를 조정하는 등)를 변경하여 더욱 효과적으로 이완을 돕도록 조정할 수 있다. 사용자와의 장기적인 상호작용을 통해 지속적으로 최적화되고 학습하

는 기계 학습 모델을 포함한다. 시간이 지남에 따라, 이 알고리즘은 특정 사용자에게 어떤 유형의 음악이 가장 적합한지를 더 정확하게 이해하고, 음악 체험을 더욱 효과적으로 개인화할 수 있다.

2) Mubert

Mubert[18]는 사용자의 환경, 활동 유형 또는 감정 상태에 따라 음악 콘텐츠를 자동으로 조정할 수 있는 인공지능 기반으로 한 적응형 음악 생성 플랫폼이다. 이 플랫폼은 특히 개인 사용자에게 맞춤형 음악 경험을 제공하는 데 초점을 맞추고 있으며, 소매점, 체육관 또는 온라인 방송과 같은 상업적 환경에서 사용하기에 적합한 배경 음악을 제공한다. Mubert의 핵심 기능 중 하나는 실시간으로 연속적인 음악 스트림을 생성할 수 있는 능력이다. 이는 미리 녹음된 수천 개의 음악 조각을 알고리즘으로 조합하고 수정하여 이음새 없이 연결되는 음악을 만든다. 또한, Mubert는 사용자의 환경 변화와 감정 상태를 감지하고 분석하여 현재 장면이나 감정 요구에 맞게 음악 스타일과 리듬을 자동으로 조정할 수 있는 능력을 갖추고 있다. 예를 들어 요가나 명상 시에는 이완된 음악을, 활동적인 체육 환경에서는 리듬감 있는 음악을 생성할 수 있다. 플랫폼은 전자부터 클래식까지 다양한 음악 스타일을 제공하여, 다양한 사용자의 음악 취향과 적용 장면의 요구를 만족시킬 수 있다. 이러한 기능은 저작권 문제 없이 제공되며, 사용자는 어떤 상업 환경에서든 자유롭게 음악을 사용할 수 있으며, 저작권 문제에 대해 걱정할 필요가 없다.

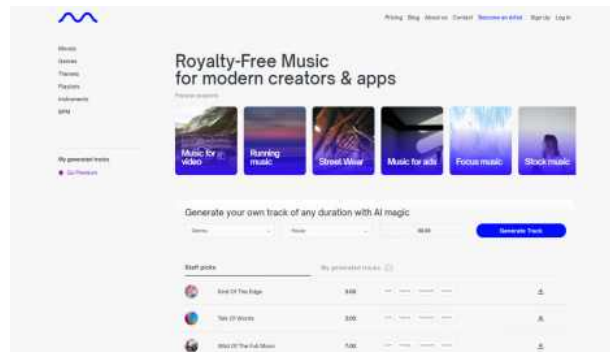


그림 9. Mubert의 인터페이스[18]
Fig. 9. Interface of Mubert[18]

Mubert는 고급 음악 이론과 인공지능 기술을 결합하여, 기계 학습과 방대한 음악 샘플 데이터베이스를 사용하여 음악 생성 알고리즘을 지속적으로 최적화한다. 이 기술은 음악의 창작 방식을 변화시키고, 음악의 적용 분야를 확장하여 사용자 경험과 브랜드 상호 작용을 강화하는 새로운 도구로서의 역할을 한다. Mubert는 이처럼 혁신적인 적응형 음악 생성 플랫폼으로서 음악 창작과 소비를 보다 지능적이고 개인화되며 다양화할 수 있는 방법을 제공함으로써, 모든 사용자가 자신만의 독특한 음악적 경험을 할 수 있도록 돕는다.

IV. AI 작곡 기술의 발전 경향 및 문제점

4-1 기술의 발전 경향

1) 확장된 창작 공간과 협업

AI 작곡 기술은 인간 작곡가와 AI의 협업을 가능하게 하여 새로운 창의적 공간을 제공하고 있다. 예를 들어, IBM의 Watson Beat는 사용자가 정의한 감정, 스타일, 기간 등을 바탕으로 음악을 생성하며, 사용자의 피드백을 실시간으로 반영해 음악을 조정할 수 있는 기능을 갖추고 있다.

Google의 Magenta 프로젝트는 이러한 협업 개념을 한층 발전시켜 사용자가 AI와 직접 상호작용하면서 음악을 생성할 수 있는 툴셋을 제공한다. Magenta Studio는 딥러닝 데이터를 활용하여 AI 모델인 NSynth를 학습시켜 음악 데이터를 생성하며, NSynth Super[19]는 물리적 인터페이스를 통해 음악가가 직접 조작하며 소리를 탐색할 수 있는 기능을 제공한다[9].



그림 10. 오픈 소스 하드웨어인 NSynth Super[19]
Fig. 10. Open source hardware NSynth Super[19]



그림 11. AWS DeepComposer 연주용 키보드[20]
Fig. 11. AWS DeepComposer keyboard[20]

그림 11과 같이 AWS DeepComposer 연주용 키보드 [20]는 뮤직 스튜디오 인터페이스와 USB를 활용해 PC와 연결된다. 사용자는 기존에 생성된 MIDI 데이터를 입력하거나 자신의 연주를 AI와 연계하여 작곡하는 작업을 수행할 수 있으며, 멜로디나 화성을 입력하면 AI가 이를 기반으로 멜로디, 화성, 편곡 등 음악적 요소를 조합해 사용자의 요구에 맞는

창작물을 생성한다. 이처럼 실제 물리 환경과 AI 환경의 결합은 AI와 작곡가 간 협업 공간의 다양화 경향을 보여준다.

2) 고도의 맞춤형과 개인화

AI 작곡 기술은 사용자의 실시간 피드백을 기반으로 적응하는 능력을 강화하여 개발되고 있다. 이 기술은 사용자의 행동, 환경, 심지어 생체 신호까지 분석하여, 개인의 감정 상태나 상황에 맞는 음악을 실시간으로 생성할 수 있다. 예를 들어, Endel은 사용자의 위치, 시간, 날씨, 심박수 등의 데이터를 수집하여 이를 기반으로 음악을 조정하고, 사용자의 일상 패턴, 기분 변화, 활동 수준에 따라 이완, 집중력 증진, 수면 개선 등 다양한 필요를 충족시키는 맞춤형 음악을 제공한다. 음악 심리학의 원리를 바탕으로 특정 주파수와 리듬이 인간의 심리 상태에 미치는 영향을 이해하고 이를 응용하여 사용자에게 긍정적인 영향을 미치는 사운드스케이프를 창조한다. Mubert는 사용자의 활동 유형, 환경, 감정 상태에 따라 음악 콘텐츠를 자동으로 조정하는 기능을 갖추고 있으며, 미리 녹음된 음악 조각을 알고리즘으로 조합하여 각 사용자의 상황에 맞게 음악을 매끄럽게 조정한다.

사용자의 환경 변화와 감정 상태를 감지하여 필요에 따라 음악 스타일과 리듬을 자동으로 조정하는 AI 작곡 기술은 다양한 상황에서 음악적 경험을 최적화한다. 예를 들어, 요가나 명상 시에는 이완된 음악을, 활동적인 운동 환경에서는 리듬 감 있는 음악을 자동으로 생성할 수 있다. 이러한 기술은 사용자의 변화하는 필요에 신속하게 대응하여, 각 상황에 적합한 음악을 제공함으로써 보다 풍부한 음악적 경험을 가능하게 한다. 이처럼, AI는 단순한 자동화 도구를 넘어, 인간의 감성과 창의성을 확장하고 개인화된 사용자 경험을 제공하는 파트너로서 음악 창작의 미래를 형성하고 있다.

3-4 인공지능 작곡 플랫폼에서 제공하는 공통적 기능과 문제점

표 3에서 제시된 바와 같이, 인공지능을 이용한 작곡에서 음악적 아이디어의 표현은 뮤직 프로덕션 과정에서 스타일, 장르, 템포 설정 외에 형식이나 음악적 표현과 같은 세부적인 내용을 설정하는 데 한계가 있으며, 멜로디 및 편성에서도 특정 멜로디에 대한 학습과 편성 구성 외에 ‘창의성’을 표현할 수 있는 기술적 구현이 부족함을 확인할 수 있었다. 이는 초기 MIDI 시퀀서가 등장했을 때 제기되었던 문제점과 유사하며, 이를 기술적으로 음악적으로 발전시키기 위해서는 많은 연구와 기술적 분석이 요구된다. 이로 인해 많은 음악 관련 전문가들과 비평가들이 인공지능이 작곡한 콘텐츠의 품질에 대해 의구심을 가지고 있다는 사실이 명확하다. 문화예술 분야에서는 창작이라는 ‘독창성’을 필요로 하며, 단순한 수학적 계산을 넘어 작품의 수요자들에게 감성적 요소를 전달하여 공감을 얻는 것이 중요하다.

특히 인공지능 작곡 기반의 플랫폼 서비스는 확률형 모델을 기반으로 하여, 학습 데이터로부터 음악적 결과물을 생성

표 3. 인공지능 작곡 플랫폼 공통적 기능

Table 3. AI composition platform common functions

Platform functions	Description
Input music creativity	Input your own melody as a music signal or select a preset
The choice of style and rhythm	Choose a music style or genre and provide rhythm options
Choose music length	Select/automatically generate the desired music length
Create melody	Automatically generate melodies that match the characteristics of the genre
Arrangement	The arrangement and automatic arrangement of instruments such as brass, strings, piano, and percussion
Output of creative works	Direct or indirect use/resolution of copyright disputes in creative works

하고 있다. 이 과정에서 벨로시티(Velocity), 서스테인(Sustain), 익스프레션(Expression) 등의 컨트롤 체인지 값에 대한 표현력 부재와 음악적 오류가 발생할 수 있다. 이로 인해 전문적인 콘텐츠 제작에는 한계가 있으며, 주로 창작 활동의 협업 수단으로 활용되는 제약이 따른다. 선행연구들에서도 이러한 인공지능 작곡의 기술적 한계를 명확히 지적하고 있다. 예를 들어, Fernández, Vico[21]는 예술적 창의성을 형식적이고 명확한 정의 및 연산으로만 표현할 수 없기 때문에 시스템적으로 창작물을 평가하기 어렵다고 하였다. 또한 Ariza[22]는 인공지능 시스템의 알고리즘 구성이 창의적이고 혁신적인 예술 작품을 창작하기보다는 모방에 효과적이라고 언급하였다.

인공지능 작곡 기술의 발전을 위해 현재 전문 음악인과의 연계를 통한 학습 모델의 개선이 필수적이다. 기존의 인공지능 작곡 학습 모델은 다양한 음악적 장르, 스타일, 리듬 정보를 바탕으로 새로운 창작물을 생성하는 프로세스를 갖추고 있으나, 창작 과정 자체가 최종 결과물에 미치는 영향을 고려할 때 이러한 창작 과정을 보다 효과적으로 학습할 수 있는 모델의 개발이 요구된다. 이를 위해 인공지능 작곡 모델 개발자와 음악 콘텐츠 제작 전문가 간의 지속적인 연구 및 협업이 중요하다.

V. 결 론

본 연구는 AIGC를 활용한 인공지능 작곡 플랫폼의 사례를 분석함으로써 인공지능 작곡 기술의 발전 추세와 문제점을 파악하는 것을 목적으로 하였다. 사례분석에서 인공지능을 활용한 작곡 플랫폼은 세 가지 주요 유형으로 구분되었다. 첫 번째 유형은 자동작곡으로, 인간의 최소한의 개입으로 AI가 독립적으로 음악을 생성하는 형태이다. AIVA Technologies, Amper Music, Jukedeck 등 플랫폼을 사례분석 대상으로 분석하였다.

두 번째 유형인 협력적 작곡에서는 AI가 초기 아이디어나 음악적 요소를 제공하며, 인간 아티스트가 이를 바탕으로 창의적인 결정을 내려 작품을 완성한다. 이 유형에는 Watson Beat, Magenta, AWS Deep Composer 등이 포함된다. 마지막 유형은 적응적 작곡으로, 사용자의 행동, 상황 또는 환경에 따라 음악이 실시간으로 변화하는 방식이며, Endel, Mubert 등 플랫폼을 사례분석 대상으로 분석하였다.

사례분석을 통해 AI 작곡 기술의 발전 경향과 문제점은 요약하면 다음과 같다. 첫째, 확장된 창작 공간과 협업을 들 수 있다. AI 작곡 플랫폼은 사용자가 정의한 감정, 스타일, 기간 등을 기반으로 음악을 생성하고, 사용자의 실시간 피드백을 반영하여 음악을 조정하는 기능을 제공한다. 예를 들어, Magenta Studio와 AWS DeepComposer의 사례에서 사용자는 기존에 생성된 MIDI 데이터를 입력하거나 자신의 물리적 도구를 통해 연주하면서 AI는 이를 기반으로 멜로디, 화성, 편곡 등 음악적 요소를 조합하여 사용자의 요구에 부합하는 창작물을 생성한다. 이러한 실제 물리 환경과 AI 환경의 결합은 AI와 작곡가 간의 협업 공간을 확장하는 산업 발전 경향을 보여준다. 둘째, 고도의 맞춤화와 개인화가 이루어지고 있다. AI 작곡 기술은 사용자의 실시간 피드백을 토대로 적용하는 능력을 강화하여 개발되었다. 이 기술은 사용자의 행동, 환경, 심지어 생체 신호까지 분석하여 개인의 감정 상태나 상황에 맞는 음악을 실시간으로 생성할 수 있다. 이러한 기술은 사용자의 변화하는 요구에 신속하게 대응하여, 각 상황에 적합한 음악을 제공함으로써 보다 풍부한 음악적 경험을 제공한다. 셋째, 사례분석을 통해 드러난 AI 작곡 기술의 발전 문제점 중 하나는 '창의성'을 표현할 수 있는 기술적 구현이 부족하다는 점이다. 인공지능 작곡 플랫폼은 주로 확률형 모델을 기반으로 하여 학습 데이터로부터 음악적 결과물을 생성하고 있다. 이 과정에서 벨로시티, 서스테인, 익스프레션 등의 컨트롤 체인지 값에 대한 표현력 부재로 음악적 오류가 발생할 수 있다. 이러한 문제는 전문적인 콘텐츠 제작의 한계를 드러내며, 대체로 창작 활동의 협업 수단으로서의 활용에 제약을 받는다.

결론적으로, 현재 인공지능을 이용한 작곡이 인간을 완벽하게 대체할 수는 없다. 그러나 과거 음악 녹음 방식의 변화, MIDI를 활용한 작곡, 가상 악기를 이용한 편곡과 같이 기술의 발전과 사회적 패러다임 변화에 따라 문화예술 전반, 특히 음악 분야에서 과도기적인 문제들을 해결하며 자리를 잡았던 다양한 기술들처럼, 인공지능을 활용한 콘텐츠 생성과 음악 창작 또한 지속적인 연구와 논의, 신기술 개발을 통해 발전하고 활용되어 나갈 것으로 기대된다. 이 과정에서 관련 기업, 학계, 업계 종사자들의 지속적이고 다각도의 연구가 필수적이며, 이를 통해 인공지능 작곡 기술은 향후 인간에게 효율적이고 유용한 기술로 자리매김할 것이다.

본 연구의 한계점은 연구자가 주관적으로 AIGC를 활용한 인공지능 작곡 플랫폼의 사례를 분석했다는 점으로, 객관적인 자료나 관찰된 통계 자료를 통해 검증하지 않았다는 것이다.

향후연구에서는 다양한 인공지능 작곡 플랫폼 간의 사용자 객관적 자료와 관찰된 통계 자료를 수집하여, 구조방정식이나 기타 통계적 방법을 활용함으로써 문제점으로 제시된 부분을 보강할 것이다.

참고문헌

- [1] H.-J. Yoo, "A Case Study on Artificial Intelligence's Music Creation: Focusing on Google's AI Magenta Project," *Journal of Next-Generation Convergence Technology Association*, Vol. 6, No. 9, pp. 1737-1745, September 2022. <http://dx.doi.org/10.33097/JNCTA.2022.06.09.1737>
- [2] B. K. Park, "Analysis of Domestic and International Research Trends Related to Artificial Intelligence in Music Field," *Research in Music Pedagogy*, Vol. 25, No. 1, pp. 1-12, February 2024.
- [3] H.-J. Kim, "A Study on the Intentions of ChatGPT Users Using the Extended UTAUT Model," *Journal of Digital Contents Society*, Vol. 24, No. 7, pp. 1465-1473, July 2023. <http://dx.doi.org/10.9728/dcs.2023.24.7.1465>
- [4] Y.-Y. Kim and Y.-H. Nam, "Fairy Tale Reconstruction Service Using Conversational AI," *Journal of Digital Contents Society*, Vol. 24, No. 9, pp. 1955-1964, September 2023. <http://dx.doi.org/10.9728/dcs.2023.24.9.1955>
- [5] K. H. Kang, "A Study on the Challenges and Legislative Response of Artificial Intelligence (AI) to Cultural Arts Law," *Journal of Legislative Studies*, Vol. 21, No. 1, pp. 5-41, February 2024. <https://doi.org/10.31536/jols.2024.21.1.001>
- [6] J. Park, "A Research into Potential for the Development of Convergence Art of Dancing in the Era of the Fourth Industrial Revolution -Focusing on Artificial Intelligence-," *The Korean Journal of Dance Studies*, Vol. 68, No. 1, pp. 37-52, March 2018. <https://doi.org/10.16877/kjds.68.1.2018.03.37>
- [7] K.-H. Ko, "AI, Human, and the Possibility of Change in the Dance Paradigm," *The Journal of Korean Dance*, Vol. 39, No. 1, pp. 1-25, March 2021. <https://doi.org/10.15726/jkd.2021.39.1.001>
- [8] D. Guo, H. Chen, R. Wu, and Y. Wang, "AIGC Challenges and Opportunities Related to Public Safety: A Case Study of ChatGPT," *Journal of Safety Science and Resilience*, Vol. 4, No. 4, pp. 329-339, December 2023. <https://doi.org/10.1016/j.jnlssr.2023.08.001>
- [9] K. Cheng, P. Neisch, and T. Cui, "From Concept to Space: A New Perspective on AIGC-Involved Attribute Translation," *Digital Creativity*, Vol. 34, No. 3, pp. 211-229, 2023. <https://doi.org/10.1080/14626268.2023.2248103>
- [10] Y. Wang, Y. Pan, M. Yan, Z. Su, and T. H. Luan, "A Survey on ChatGPT: AI-Generated Contents, Challenges, and Solutions," *IEEE Open Journal of the Computer Society*, Vol. 4, pp. 280-302, August 2023. <https://doi.org/10.1109/OJCS.2023.3300321>
- [11] O. Ben-Tal, M. T. Harris, and B. L. T. Sturm, "How Music AI is Useful: Engagements with Composers, Performers and Audiences," *Leonardo*, Vol. 54, No. 5, pp. 510-516, October 2021. https://doi.org/10.1162/leon_a_01959
- [12] C. Jin, Y. Tie, Y. Bai, X. Lv, and S. Liu, "A Style-Specific Music Composition Neural Network," *Neural Processing Letters*, Vol. 52, pp. 1893-1912, December 2020. <https://doi.org/10.1007/s11063-020-10241-8>
- [13] Sound Tech Insider. AIVA Review - Everything You Need To Know - AI Composition [Internet]. Available: <https://soundtechinsider.com/aiva-review-everything-you-need-to-know-ai-composition/>.
- [14] The Korea Bizwire. Amper Music Launches First AI Music Composition Platform for Enterprise Content Creators [Internet]. Available: <http://koreabizwire.com/amper-music-launches-first-ai-music-composition-platform-for-enterprise-content-creators/131411/>.
- [15] Music Ally. Jukedeck Hopes Artificial Intelligence Can 'Democratise Music' [Internet]. Available: <https://musically.com/2017/08/09/jukedeck-artificial-intelligence-music/>.
- [16] Ableton. Magenta Studio: Free AI tools for Ableton Live [Internet]. Available: <https://www.ableton.com/en/blog/magenta-studio-free-ai-tools-ableton-live/>.
- [17] Amazon Web Services. AWS DeepComposer - Now Generally Available with New Features [Internet]. Available: <https://aws.amazon.com/cn/blogs/aws/aws-deep-composer-now-generally-available-with-new-features/>.
- [18] PopularAITools.ai. Mubert: Revolutionizing Music Creation with AI? [Internet]. Available: <https://www.popularaitools.ai/tools/mubert>.
- [19] Synth Anatomy. NSYNTH SUPER is a New Open Source Machine Learning Synthesizer Supported by Google [Internet]. Available: <https://synthanatomy.com/2018/03/n-synth-super-new-open-source-machine-learning-synthesize-r-supported-google.html>.
- [20] Synthtopia. Amazon Intros AWS DeepComposer, a New Platform for Artificial Intelligence & Music Composition [Internet]. Available: <https://www.synthtopia.com/content/2019/12/02/amazon-intros-aws-deepcomposer-a-new-platform-for-artificial-intelligence-music-composition/>.
- [21] J. D. Fernández and F. Vico, "AI Methods in Algorithmic Composition: A Comprehensive Survey," *Journal of Artificial Intelligence Research*, Vol. 48, pp. 513-582, 2013.

<https://doi.org/10.1613/jair.3908>

[22] C. Ariza, "The Interrogator as Critic: The Turing Test and the Evaluation of Generative Music Systems," *Computer Music Journal*, Vol. 33, No. 2, pp. 48-70, 2009.

이빙옥(Bing-Yu Li)



2019년 : 중국 호남사범대학교
(예술학 학사)

2021년 : 중앙대학교(영상학 제작석사)

2021년~현 재: 중앙대학교 첨단영상대학원
영상제작전공 박사과정

※관심분야 : 디지털 미디어아트(Digital media art),
음악시각화(music visualization)

최종원(Jongwon Choi)



2008년 : KAIST 전기전자공학과
(공학사)

2012년 : KAIST 전기전자공학과
(공학석사)

2018년 : 서울대학교 전기정보공학과
(공학박사 - 컴퓨터비전)

2018년~2020년: 삼성SDS 연구소

2020년~현 재: 중앙대학교 첨단영상대학원 영상학과 교수

※관심분야 : 컴퓨터비전(Computer Vision),
인공지능(Artificial Intelligence),
딥러닝(Deep Learning) 등