

들어가기

컴퓨터 생성 예술의 문제는 향후 수년에서 늦어도 수십 년 내 분명 우리가 주목하게 될 문제가 될 것이다. 거의 모든 예술 형식이 그 영향권에 들어가게 될 것이기 때문이다. 인공지능기반 창작어플리케이션들의 최근 발전상을 보면 매우 긍정적인 미래를 그리게 된다. 오늘날 인공지능 기반 생성 시스템은 실제와 같은 이미지와 영상의 구현은 물론 수준 높은 글쓰기와 작곡까지 가능하다. 그런데 여기서 한 가지 의문이 제기된다. 과연 기계가 만든 결과물이 예술작품이 그러하듯 우리의 정서적인 반응을 이끌어낼 수 있을까? 필자는 음악과 컴퓨터 과학을 전공하고 현재 예술가이자 오스트리아 린츠에 위치한 아르스 일렉트로니카의 퓨처랩 연구원으로 활동하고 있다. 따라서 이 글은 음악을 중심으로 인공지능과 창의성의 관계, 작곡에 활용되는 인공지능의 창의성, 인공지능 생성 시스템들의 매력과 더불어 그 한계와 문제점들을 짚어보고 나아가 인공지능으로만 만들어진 음악의 예술적 가치와 인공지능기술이 실제 예술가들과 아마추어 예술가들에게 줄 혜택에 대해 논의하는 글이 될 것이다.

인공적 창의성의 숨은 기술

현재 창의적, 예술적 작업에 활용되는 인공지능 시스템의 근본적인 문제점과 한계를 감안하더라도, 최근의 발전된 기술들은 앞으로 나올 시스템의 모습과 그 가능성을 엿볼 수 있는 강력한 예술적 잠재력을 보여준다. 이에 그 기술적 한계에도 불구하고

예술가들은 이러한 시스템들에 내재한 잠재력을 이용하여 대단히 흥미로운 작품과 프로젝트들을 쏟아내고 있는데 그 바탕이 되는 인공지능기술이 딥 러닝이다. 딥 러닝은 인공지능망을 이용하여 주어진 임무수행에 필요한 데이터의 유의미한 특징들을 학습하는 확률적 인공지능기술이다. 기본적으로는 입력된 데이터에서 주어진 문제 해결을 위한 데이터 도출 시스템을 구축하는 것이 목적이다.

하지만 바로 이 입력에서 출력까지 이어지는 과정이 매우 복잡해질 수 있으므로 인공지능망은 반드시 사례를 들어 학습시켜야 한다. 인공지능망모델은 주어진 임무수행에 필요한 데이터의 모든 특징들을 독자적으로 익혀야 한다. 학습된 모델에게는 종종 입력된 데이터와 유사한 데이터를 생성하는 임무가 주어진다. 예를 들어, 자기회귀모델은 텍스트와 음악 생성에, 생성적 적대 신경망(GAN)은 다양한 이미지 생성에 자주 활용된다. 여기엔 "유사함"이 "동일함"을 의미하지 않음을 명시하는 작업이 필요하다. 대부분의 경우 신경망모델은 학습 데이터에 들어있지 않은 데이터를 생성해야한다. 예를 들어, 인간의 얼굴을 학습한 신경망모델이 있다고 하자. 그 모델은 얼굴로 인식될 수 있는 이미지를 생성하되 학습데이터에 들어있는 이미지와 동일하지 않은 이미지를 생성해야 한다. 이런 종류의 신경망모델은 딥 러닝 기술을 창의적 작업에 적용하려는 학계의 높은 관심 덕분에 불과 몇 년 만에 엄청난 진보를 거듭했다. 그런데 왜 학계에서 유독 이러한 모델들에 관심을 가지는 것일까?

코디

Anders

창의적 모델에 대한 연구자들의 관심은 새로운, 유사 데이터의 생성이 입력 데이터의 본질적 속성에 대한 기초적인 이해가 있어야만 가능하다는 가정에 근거한다. 말하자면 사람의 얼굴 이미지를 생성할 수 있는 모델은 인간 얼굴의 모든 핵심적인 특징들과 인간의 얼굴을 얼굴처럼 보이게 만드는 요소들, 또한 인간 얼굴의 다양성과 차이점들까지 (즉, 성, 나이, 안색, 눈의 색깔 등) 모두 섭렵한 모델이다. 더욱이 그 모델은 눈의 색깔이 사람마다 조금씩 다르지만 개인의 눈 색깔은 양쪽이 동일하다는 점까지 학습해야 한다.

일반적으로 딥 러닝 시스템은 엄청난 양의 데이터를 학습할 수 있다. 종종 임무를 수행하기 위해 인간이 미처 인지하지 못한 데이터의 특징들까지 학습한다. 이 시스템들은 인간 관찰자가 결코 주목해본 적이 없는 관계들과 종속성을 파악할 수 있으며 따라서 창의적 임무 전반에 새로운 영감으로 작용할 수 있다. 특히 예술가들에게 인공지능은 새로운 예술작품의 가능성을 발견하고 실험할 수 있는 창의적 분야로 인식될 가능성이 크다.

인공지능기반 음악 생성 프로그램의 숨은 기술

이번 섹션은 인공지능 시스템의 기술적 배경에 중점을 두고 이러한 시스템들이 어떤 학습과정을 거쳐 음악을 만들어내는 데 활용되는지 알아볼 것이다. 기술적인 관점에서 여러 연구 프로젝트들은 인공지능기반 자연어 처리(NLP) 기술이 음악 생성에 성공적으로 활용될 수 있다는 걸 보여주었다. 이러한 시스템들을 이용하기 위해서는 우선 음악 데이터를 연속되는 단어들로 표현할 수 있는 방법이 필요하다. 여기에 여러 가능성들이 있다. 예를 들어, 음의 높낮이는 고유"명사들"로 표현될 수 있으며, 음의 길이와 높임은 기호들로 표현될 수 있다. 이러한 특수한 단어들의 연결로 약보나 미디 파일같은 음악 데이터를 어떤 정보의 손실도 없이 나타낼 수 있다. 음악 데이터를 단어들의 순차적 연결로 표현하는 것은 음악가들에게 전혀 직관적이지 않아 보일 수 있다. 하지만 최근 연구에 따르면 인공지능 시스템은 이러한 표현들을 매우 효과적으로 사용하고 있음을 알 수 있다.

크리스틴 맥리비 페인이 2019년 오픈에이아이 [OpenAI]에서 만든 뮤즈넷(MuseNet)<sup>2</sup>은 그 가장 좋은 예라 할 수 있다. 뮤즈넷은 GPT-2<sup>3</sup>, GPT-3<sup>4</sup> or Bert<sup>5</sup> 같은 가장 강력한 자연어 처리 모델들 중 몇몇과 유사한 기술을 배경 기술로 사용하며 언어와 비슷한 음악 데이터의 순차적 표현들을 학습한다. 여기에는

트랜스포머<sup>6</sup>라는 기술이 사용되는데, 트랜스포머 구조를 가진 인공지능망은 곡의 특정 부분에 주목하는 방법을 배운다. 이는 시퀀스 출력을 생성하는 데 중요하다. 이러한 주의(attention) 기제를 사용하는 것은 음악을 생성하는 데 결정적인 중요성을 갖는데 이는 악보의 음표들이 (즉, 언어에서 단어들처럼) 연주되는 곡의 맥락에 크게 의존하기 때문이다. 자연어에서 모든 단어가 그것이 쓰이는 맥락안에서만 이해될 수 있듯이, 음악에서 모든 음의 높낮이는 그 맥락안에서만 의미를 갖는다.

뿐만 아니라, 음악은 많은 구간의 반복으로 이루어지는데 이렇게 반복되는 구간은 앞선 내용을 알고 있어야 인지될 수 있다. 반복되는 구간들은 시간상 거리가 가깝게 나오도록 하고 매우 띄엄띄엄 등장하기도 한다. 이 구간들은 해당 구간들이 이후 반복될 때 감상자들이 기억할 수 있는 주제부와 구조들을 가지고 있다. 이와 유사하게, 앞선 내용을 기억하는 것은 자연어 처리에 매우 중요한 부분이다. 이야기 초반에 주요 등장인물들이 소개되고 나면 나중엔, 그, 그녀, 그의, 그녀의 등과 같은 대명사로 지칭된다. 인공지능 시스템은 이 모든 대명사들이 누구를 지칭하는지, 심지어 지칭하는 인물이 모호한 경우에도 파악할 수 있어야 한다.

뮤즈넷과 같은 모델들을 자기회귀모델 (autoregressive models)이라고 부른다. 자기회귀모델은 주어진 단어들 다음에 나올 단어들을 예측할 수 있게 훈련된다. 이렇게 이전의 단어들로부터 다음 단어를 생성할 수 있으면 학습 데이터가 나온 출처 분야에 대한 기초적인 이해가 전제되어야 한다. 언어의 경우, 이것은 단어들 간의 관계와 의미 그리고 문법을 아우른다. 따라서 데이터를 "내면화"하는 방법을 습득한 모델은 [엄밀히 말해 데이터의 비선형적 내적 표상을 학습한 모델은], 텍스트 생성 이외의 임무에 활용될 수 있다. 즉, 텍스트 요약 (주어진 텍스트를 요약할 수 있다), 정서 분석 (문장에서 느껴지는 분위기, 예를 들어 긍정적인지 부정적인지를 알아낼 수 있다), 질의응답 (주어진 질문에 대한 답을 찾아낸다)을 비롯한 많은 임무들을 수행할 수 있다. 한 마디로, 텍스트 생성 시스템은 언어의 내적 작용을 파악하는 데 활용될 수 있다는 것이다. 마찬가지로 데이터 표상을 습득한 모델들은 그와 관련된 임무를 수행하는 데 활용될 수 있다. 음악의 경우, 이 모델은 음표들로 표현된 이전의 내용을 기반으로 다음 음을 생성하도록 훈련된다. 제시되는 구간의 길이는 모델에 따라 달라진다. 모델의 메모리나 종류에 따라 수백 개에서 수천 개에 이르는 요소들로 구성된 구간이 제시되기도 한다.

2 크리스틴 페인, "뮤즈넷[MuseNet]", 오픈에이아이, 2019년 4월 25일, <http://openai.com/blog/musenet>

3 알렉 레드포드 외, "언어 모델은 자율적 멀티태스킹 학습자다 [Language Models are Unsupervised Multitask Learners]", 2019. url: [https://cdn.openai.com/better-language-models/language\\_models\\_are\\_unsupervised\\_multitask\\_learners.pdf](https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf)

4 탐 B 브라운 외, "언어 모델은 최소의 정보로 학습이 가능한 모델이다 [Language Models are Few-Shot Learners]", 2020. url: <https://arxiv.org/abs/2005.14165>

5 야쿰 데블린 외, "버트(BERT): 언어 학습을 위한 심층 양방향 트랜스포머의 사전 훈련 [BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding]", 2018. url: <https://arxiv.org/abs/1810.04805>

6 애쉬리 왓슨니 외, "필요한 것은 어텐션이다", 2017 <https://arxiv.org/abs/1706.03762>

1 이안 굿펠로우 외, "생성적 적대 신경망", 2014. 출처: <https://arxiv.org/abs/1406.2661>



이 섹션에서, 우리는 음악과 자연어가 엄연히 다른 영역임에도 인공지능 시스템이 도전할 과제에 있어서는 별 다른 점이 없다는 것을 보았다. 인간이 음악을 만들 때 앞서 작업한 부분을 바탕으로 다음 음표를 적어 나가듯, 인공지능도 앞선 내용을 바탕으로 그 다음의 음을 생성한다.

### 매력

오늘날 인공지능 기반 작곡 시스템은 불과 일 이 년 전에는 상상도 못했던 수준의 작곡 능력을 보여주고 있다. 비전문가들에게는 인공지능이 작곡한 음악과 사람이 작곡한 음악 사이의 차이가 거의 느껴지지 않을 정도다. 인공지능의 이러한 작곡 능력은 '말리의 미완성 프로젝트'라는 주제하에 열린 2019년 아리스 일렉트로니카 페스티벌을 통해 소개된 바 있다. 당시 우리는 출시된 지 얼마 되지 않은 뮤즈넷을 이용하여 인공지능 시스템이 구스타프 말리의 마지막 교향곡 (그의 미완성된 교향곡)의 테마를 바탕으로 곡을 전개시킬 수 있다는 사실을 보여주었다. 사실 이 교향곡의 테마는 매우 독특하다. 조성이 뚜렷하게 느껴지지 않는 단조로운 선율이 아무런 반주 없이 시작되기 때문이다. 그럼에도 뮤즈넷은 그 테마에 새로운 악상을 결합하여 곡을 전개하며 새로운 교향악단 (린츠 브루크너 오케스트라)의 연주로 공연되었다. 더불어 말리의 원곡도 연주되었다. 이를 통해 우리가 의도한 것은 인공지능과 클래식 음악에서의 창의성에 관한 논의의 기초를 마련하는 것이었다. 대부분의 관객들은 인공지능이 작곡한 음악을 자연스럽고, 서정적인 연주처럼 감상했고 소리만으로 [이것은 오케스트라 단원들이 곡에 모든 열정을 쏟아 연주했다는 사실과도 관계가 있다] 그 곡이 인공지능에 의해 생성된 곡이라는 사실을 눈치채지 못했다. 물론 말리 전문가들은 거창히 테마를 풀어나가는 방식과 인공지능 시스템이 테마를 전개시키는 방식의 차이를 인지할 수 있었지만 말이다.



마커스 포슈너가 미완성 말리 교향곡 프로젝트에서 린츠 브루크너 오케스트라를 지휘하고 있다. 아리스 일렉트로니카 페스티벌, 2019. 사진: 유르겐 그룬발트

그러나 이 공연을 비롯하여 이와 비슷한 프로젝트들은 앞으로 인공지능 생성 음악의 예술적 의미에 대한 논의를 심화시키게 될 것이다. 순수하게 기술적인 관점에서 인공지능

시스템이 자연스럽고 서정적이며 거장들이 작곡한 음악처럼 우리의 정서적 반응을 이끌어 낼 수 있는 음악을 만들어 낸다는 사실은 물론 멋진 일이다. 인공지능 시스템이 작곡한 몇몇 곡들은 전문 음악 이론의 관점에서 보더라도 매력적이라고 할 수 밖에 없는 복잡한 하모니와 리듬을 가지고 있다. 그럼에도 우리는 여전히 다음의 질문들을 제기할 수 있다. 인공지능이 생성한 음악을 예술이라고 부를 수 있을까? 인공지능의 개입 없이 예술적 의미를 창출하는 것이 가능한가? 그 모든 매력적인 측면에도 불구하고 철저히 인공지능으로만 생성한 음악의 예술적 의미에 대한 이 질문은 여전히 답을 찾지 못하고 있다.

다음 섹션에서 우리는 인공지능기반 작곡 시스템의 근본적인 한계점들에 대해 논의할 것이다. 이 섹션은 두 부분으로 구성된다. 전반부에서는 현재 작곡에 활용되고 있는 인공지능 시스템의 한계점들을 논할 것이다. 앞으로 극복될 수 있는 인공지능경망모델의 기술적 본질과 관련된 이슈들이 여기에서 다루어질 것이다. 그 다음으로는 현재 기술을 비롯한 인공지능 기술의 한계점들에 대해 다룰 것이다. 인공지능이 생성한 예술과 인공지능이 인간의 예술과 문화에 어떤 도움이 될 수 있는지에 대한 논의에 일반적으로 등장하는 한계점들이다.

### 한계

인공지능기반 음악 생성 시스템은 주어진 음의 나열에 기초하여 다음에 올 음을 생성한다. 하지만 다음에 올 음과 그 전의 음들이 잘 어울린다고 해서 생성된 음들이 창의적 가치와 고차원적 의미를 가진다고 보기에 우리가 있다.

그런데 우리는 인공지능 시스템이 작곡한 음악의 수준을 어떻게 판단할 수 있을까? 인공지능이 생성한 결과물의 "참신함"이 하나의 기준이 될 수 있다. 음악 생성에 사용되는 심층인공신경망은 많은 데이터를 학습하여 그와 유사한 데이터를 생성한다. 자연어 생성에 이용되는 신경망 모델들 역시 비슷한 목적을 가지고 있다. 하지만, 언어 모델의 경우, 결과 데이터의 참신성은 음악 생성 모델에서처럼 중요성을 갖지 않는다. 예를 들어, 어떤 영어 문장을 독일어로 번역하는데 학습한 데이터와 조금 다른 방식으로 했다면 참신한 것으로 본다. 하지만 "참신함"은 음악을 생성하는 데 있어서는 매우 중요한 문제다. 어떤 음악이 익숙한 멜로디들을 포함하고 있을 경우 사람들은 그것을 쉽게 알아차린다. 그 멜로디들이 변주 되거나 다른 조성과 박자로 연주되어도 마찬가지다. 따라서 출력된 멜로디가 학습 데이터에 들어있는 선율들을 (음악 생성 모델 학습 데이터는 이미 많이 알려진 곡들로 구성된다 - 역주) 상기시키지 않게 하는 게 중요하다. 한편, 뮤즈넷과 같은 현대 음악 생성 시스템은 학습 데이터와 다른 새로운 데이터를 생성할 수 있는 것처럼 보인다. 더욱이 임의적 전략으로 멜로디를 생성하기 때문에 그럴 가능성이 (최소한 긴 구간에서는) 현저히 떨어진다.

"참신함"에 더하여 인공지능 생성 음악의 수준은 "내용"에

기초해 평가될 수 있다. 완벽하게 인공지능이 생성한 음악 내용의 예술적 수준을 객관적으로 측정하는 것은 불가능하다. 하지만 우리는 여전히 인간 음악의 이론적 기준을 적용해 출력된 음악의 수준을 판단해 볼 수 있다. 그 기준들은 하모니의 복잡성, 리듬의 복잡성 등 많은 음악적 측면들을 고려한다. 하지만 이러한 음악 이론들도 우리에게 출력된 음악의 예술적 가치에 대해 그 무엇도 말해주지 못한다. 클래식 음악의 가장 탁월한 순간들은 의외로 매우 단순한 구조를 갖춘 경우가 많기 때문이다.

참신함과 내용이 주는 감흥 외에 "독창성"또한 중요한 요소다. 현재 딥 러닝 기반 컴퓨터 모델들은 학습 데이터의 통계적 패턴을 익히는데 이런 방법은 사람들이 음악을 들을 때 어디에서 그 곡의 "참신함"과 "독특함"을 발견하는 지 찾기 어렵게 한다. 인공지능기반 작곡 시스템들은 특별한 구조보다 평범한 구조를 학습하고 생성해내려는 성향을 지닌다. 이러한 시스템들이 출력하는 곡은 학습 데이터에서 흔하게 발견할 수 있는 하모니, 멜로디, 리듬 구조를 가지고 있다. 하지만, 음악에서 [또한 예술 전반에서] "독창성"은 위대한 거장들을 일반인들과 구별시키는 것이다. 특수한 단어보다 평범한 단어를 선호하는 것은 자연어 처리 분야의 이미 잘 알려진 문제이기도 하다. 최근의 연구들을 보면 예술과 창작 분야를 넘어 인간이 기계보다 특수한 단어를 더 사용하려는 성향이 있다는 사실을 알 수 있다.<sup>7</sup> 조성 음악 (tonal music)은 한 악절의 결말이나 불협화음의 해소 등 쉽게 예측 가능한 순간들을 포함하고 있지만 동시에 특별하고, 색다른, 예측불가능한 순간들을 가지고 있기에 우리의 귀를 사로잡는다. 이것이 음악에 유일성을 부여되는 순간들이다.

독창성과 유일성은 단순히 덜 자연스러운 전개를 가진 구조를 사용한다고 해서 생겨나지 않는다. 그것은 데이터와 인간 예술의 본질 모두에 대한 깊은 이해를 요구한다. 나아가 특별한 구조는 그에 맞는 강렬함과 적절한 순간에 드러날 때라야 그 빛을 발한다. "독창성"에 대한 어떤 공식적인 정의는 없다. 그러나 기술적으로 미래의 신경망 모델들이 현재 모델들보다 이러한 특별한 기술들을 더 잘 습득하게 되리라는 것은 충분히 예상 가능한 일이다.

컨셉트를 생각하는 것 또한 현재 인공지능기반 생성 시스템 관련 이슈 중 하나다. 컨셉트는 예술 작품의 아이디어와 작가의 의도를 매력적으로 요약하여 제시한다. 특히 음악에는 많은 컨셉트 구조들이 있다. 예술작품의 컨셉트는 매우 복잡하며 작품의 숨은 의도를 표상하기도 한다. 거의 모든 음악에서 찾을 수 있으며 감상자들이 쉽게 알아차릴 수 있는 기본적 음악 컨셉트들은 다음과 같다. 우선 곡의 주제부가 그렇다. 즉, 전체적인 악상을 보여주는 컨셉트이다. 클래식 음악에서 특정 테마와 그 변주는 여러 악장으로

확장된다. 특정 반복을 또한 일종의 컨셉트다. 해당 악절들이 한 곡이 연주되는 동안 번갈아 반복적으로 등장한다. 이러한 컨셉트들은 주의(attention) 기제를 이용하여 통계적 패턴을 학습하거나 인공지능 연구에서 아이디어를 얻어 적용시킴으로써 어느 정도까지 재생산될 수 있다. 그러나 이렇게 학습될 수 있는 컨셉트들은 매우 한정적이며 인간 작곡자의 컨셉트 작업에 견줄 수 없다. 예를 들어, 곡 전체에 등장하는 주제부와, 계속해서 변주되며 규칙적으로 등장하는 주제부를 갖게 하는 것이 불가능하다. 이러한 컨셉트를 학습할 수 없다는 점은 인공지능 기반 작곡의 근본적인 한계점들 중 하나다.

마지막으로 가장 중요한 질문은 (현재 기술 뿐만 아니라 미래 기술과도 관련된) 의미의 문제이다. 모든 기술적인 문제들이 해결되고 "작곡" 버튼만 누르면 가장 수준 높은 음악을 만들 수 있는 시스템이 마침내 만들어졌다고 가정해보자. 그렇게 생성된 음악은 어떤 의미를 갖는가? 의미 창출의 의도가 없는 기술이 의미를 만들어낸다는 것이 가능한 일일까?

의미를 만들어내기 위해서는 의도가 필요하다. 의도는 기계가 제공할 수 있는 것이 아니다. 인간이 그린 최초의 동굴벽화에도 의미를 만들고 그리려는 의도가 들어있다. 머리를 스치는 상상을 영원히 고정시켜 다른 사람에게 전달할 수 있는 방법이 어느 날 갑자기 가능해진 것이다. 만들어진 지 수천년이 지났는데도 우리는 거기에 그려진 장면을 해석할 수 있다. 아무도 그 그림의 진짜 의도가 무엇인지 모르며 누구도 어떤 예술작품의 의도가 무엇인지 말하지 못한다. 우리는 단지 작품을 해석하여 그것을 작가의 의도로 생각한다. 그럼에도 우리는 의도가 없는 창작은 없다고 말할 수 있다. 물론, 이 글을 읽는 독자들 중 현대음악의 의도<sup>8</sup>는 의도되지 않은 작품을 "창작하는" 것이라고 주장하는 사람들도 있을 것이다. 그러나 이런 경우에도 새로운 개념을 의도적으로 창작한 것이다.

또한 인공지능이 작곡한 음악을 고르고 공개하는 것 또한 의식적인 창작 행위다. 우리가 인공지능기반 작곡 소프트웨어를 쓸 때, 우리는 분명 의도를 가지고 "버튼을 누르며" 결과를 저장한다. 우리는 또한 그 결과를 공개하고 다른 사람들과 공유하는데 이 또한 의도가 들어가 있는 행위다. 이는 선택된 음악에 그 음악만의 인간적 의미를 부여한다. 물론, 예술적 관점에서 의도는 인간의 개입 없이 작곡하는 시스템을 만드는 일 자체일 수도 있다. 그러나 이 경우 시스템을 만든 사람이 의도의 원천이 될 것이다.

### 협력적 인공지능

지금까지 우리는 새로운 인공지능기반 작곡 기술의 매력과

7 연구 예, 세바스티안 게르만 외, "GLTR: 생성 텍스트의 통계적 검출과 시각화", 2019. <https://arxiv.org/abs/1906.04043>  
8 우리는 존 케이지의 음악에서 이와 유사한 의도를 찾을 수 있다. 예술가들이 임의성을 이용하여 음악적 결정을 무언이나 공연에 맡기는 음악들에서도 그러한 의도를 찾을 수 있다.

그 한계 및 문제점들을 짚어봤다. 인공지능기반 음악 창작 시스템이 어떻게 주어진 내용을 바탕으로 다음에 올 음을 만들어가는지 살펴봤다. 또한 의도를 가지지 않는 인공지능모델이 학습 기간 동안 학습된 통계적 패턴을 바탕으로 만든 결과물의 예술적 의미에 대해 논의했다.

의미를 만들려면 의미를 만들려는 의도가 필요하다. 이 점이 그러한 시스템에 인간의 협력이 중요한 이유다. 그러나 현재 작곡에 이용되는 창조적 인공지능 시스템의 제한된 상호작용 가능성을 보면 알 수 있듯이 인공지능은 아직 협력자라고 하기 보다 단순한 도구로서 더 많이 이용되고 있다. 일반적으로 창의적 인공지능 시스템과 인간 사용자들 간 상호작용과 협력은 최근 몇 년 동안 그리 많은 주목을 받지 못했다.

그렇다면 인간 예술가들과 인공지능 시스템들과의 협력은 어떻게 이루어질까? 협력이란 상호 동등한 위치에서 상대방과 함께 일하는 것을 의미한다. 훈련된 인공지능경망은 학습 과정에서 배운 음악에 대한 자기만의 기대와 가정을 가지고 있다. 이러한 가정들을 바탕으로 다음 음의 개연성을 계산한다. 인간 예술가는 인공지능과는 다른 가정들과 기대들을 가지고 있다. 연구자와 예술가들이 우리의 목표는 전문 예술가, 아마추어 예술가들 모두가 사용할 수 있는 환경을 조성함으로써 이 두 세계를 통합시키는 것이다. 현재 창의적 임무를 수행하는 인공지능 시스템의 잠재력은 인간과의 협력을 통해서만 활용될 수 있다.

#### 리체르카레: 인공지능 기반 음악 동반자

필자를 포함한 아리스 일렉트로니카 퓨처랩의 연구원들은 창의적 인공지능 시스템과 인간 예술가들이 협력할 수 있는 부분을 연구하고 있다. 당연히게도 창의적 임무는 창의적인 협력 방법을 필요로 한다. 협력에 최적화된 환경을 찾기 위해 우리는 인간과 인공지능이 좀 더 직관적으로 협력할 수 있는 방법을 제공하는 새로운 인터랙티브 인공지능 기반 작곡 시스템, "리체르카레"를 개발하고 있다. 리체르카레는 바로크와 르네상스 시대의 음악 형식을 지칭하며 "탐구하다(search out)"를 뜻하는 이탈리아어 단어에서 유래했다. 작곡자들은 테마나 악상을 실현하고 순열과 변주 가능성, 잠재적 하모니 같은 성질들을 발견하는 작품들에 이 용어를 사용했다.

"리체르카레"시스템은 그와 유사한 음악 형식을 추구한다. 이 시스템은 인간 예술가와 인공지능 작곡 시스템의 직관적 인터페이스를 구축하는 것을 목표로 하는데 이 인터페이스를 통해 인간과 인공지능은 협력하며 악상의 (사용자가 시스템에 주었거나 시스템 자체에서 시작된) 잠재성을 발견할 수 있다.

작곡 외에도 리체르카레 같은 인공지능 모델은 우리가 한 번도 고려해본 적 없는 음악 데이터의 새로운 측면들을 보여줌으로써 음악 자체의 내적 작용에 대한 통찰을 제공할 수 있다. 이로써 우리는 음악의 본질에 대해, 우리를 깊게 감동시키는 음악의

힘에 대해 지금보다 더 잘 이해할 수 있게 될 것이다. 텍스트나 음악 생성 모델을 훈련시킨다는 것은 결국에는 텍스트나 음악을 생성하는 것이라기 보다 데이터의 통계적 표상을 학습시키는 것임을 기억하자. 이 표상은 생성보다 임무의 수행에 활용될 수 있다. 예를 들어, 리체르카레의 내적 작용을 이해하기 위해 우리는 음악에 대한 그 시스템의 내적 표상을 이용해 "기계는 어떻게 음악을 보는가"라는 프로젝트를 진행한 바 있다.

이 프로젝트에서 우리는 작곡 대신 기존의 음악을 모델에 입력하고 음악이 연주되는 동안 각 신경망 층에서 신경세포들이 작용하는 모습을 관찰했다. 곡이 연주되는 동안 인공지능세포들은 특정 순간의 연주에 반응했다. 또한 특정 구간에 활성화되는 인공지능세포들의 유사성이 "유사도 매트릭스"상에 드러났다. 다시 말해, 이 프로젝트는 곡의 유사한 부분들 (즉, 반복되는 구간들) 그 부분이 비슷한 신경세포들을 활성화시키 모습을 통해 보여주었다.

그 결과 곡의 복잡한 구조를 시각화하는 새로운 방법을 발견할 수 있었다. 이로써 다시 한 번 우리는 음악이 우리의 귀를 즐겁게 할 뿐만 아니라 곡 전반에 걸쳐 아름다운 숨은 구조들을 지니고 있다는 사실을 알 수 있었으며 인공지능 기술에 의해 그것이 시각화 될 수 있다는 사실 또한 알게 되었다.



요한 세바스티안 바흐의 골드베르크 변주곡 11번을 시각화 한 이미지. 시각화를 통해 두 개의 리체르카레 시스템 신경망 층이 활성화된 모습을 볼 수 있다.

인간 예술가와 기계가 협력할 수 있는 최고의 방법을 찾기 위해 우리는 리체르카레의 인터페이스를 연구 도구로 이용해 다양한 상호작용 및 협력 방식을 비교 평가해 보았다. 이상적인 협력은 사용자들이 예술적 요소를 입력하여 얻은 결과를 자기화시키는 능력과 함께 인공적 창의성의 이점을 취할 수 있는 방식으로 이루어져야 한다. 우리는 그러한 상호작용과 협력 방식이 인공지능을 창의적 임무 수행에 성공적으로 활용하는 데 매우 중요한 역할을 하게 되리라 생각한다. 오늘날 인공지능기반 작곡 시스템들의 한계점에도 불구하고 그 시스템들이 현재 가진 잠재력은 전문 예술가들과 아마추어 예술가가 시스템과 좀 더 나은 방식으로 협력하고 상호작용함으로써 그들의 창의력을 확장하고 작품을 창작하는 새로운 방법을 찾아내는 데 훨씬 더 많은 혜택을 주게 될 것이다. 음악은 매우 인간적이고 직관적인 예술이다. 작곡은 모두에게 직관적으로 다가갈 수 있는 방식으로 이루어져야 한다. 그리고 인공지능 기술은 작곡을 좀 더 직관적이고 누구나 할 수 있는 것으로 만들 수 있다.