

텐서플로우를 이용한 미디어아트 제작 연구
-마젠타API의 응용을 중심으로-

Study of Making Media Art using Tensorflow
- Focused on applying Magenta API -

주저자

노 승 관 No, Seung-kwan

한양대학교 디자인대학 영상디자인학과 교수 | Professor of Hanyang University
skno@hanyang.ac.kr

투고일	2019.09.10	심사일	2019.10.24	게재확정일	2019.10.28
-----	------------	-----	------------	-------	------------

이 논문은 2017년 한양대학교 교내연구비 지원으로 연구되었음(HY-2017-G)

www.kci.go.kr

목 차

1. 서론
 - 1.1. 연구목적
 - 1.2. 연구방법 및 범위
 2. 이론적 고찰
 - 2.1. 인공지능과 머신러닝의 개념과 발전
 - 2.2. 머신러닝 프레임워크와 텐서플로우
 - 2.3. 기존 연구 고찰
 3. 텐서플로우 프레임워크와 마젠타 API
 - 3.1. 텐서플로우 마젠타 뮤직 모델
 - 3.2. 텐서플로우 마젠타 이미지 모델
 4. 마젠타 활용 미디어아트 제작
 - 4.1. 제작 개요 및 구성
 - 4.2. 사운드와 이미지 모듈 별 구현
 5. 결론
- 참고문헌

Keyword

머신러닝, 텐서플로우, 미디어아트
Machine Learning, Tensorflow, Media Art

Abstract

Machine learning is the core technology of 4th industrial revolution and brought big changes in art and design creations. Media art is a form of art practice which converge cutting edge art and design and possibilities of creating art with machine learning is active in this field. This paper investigate the practical possibilities of machine learning in art and design through analysing the creative process of interactive web art work FVTM:From Vera to Magenta(2019). The paper surveyed the development of artificial intelligence and machine learning technology and compared the characteristics of four major machine learning frameworks. Main study illustrates the characteristic of Google's Tensorflow, the chosen framework for project and analyse two components of Magenta API which was developed based on Tensorflow especially for artists and designers; it's sound models and image models. In creation stage, the paper illustrates evolution of open source remixing process for FVTM and explain the two main elements of the work; generation of sound applying pre-trained VAE model of Magenta API and creation and mapping process of visual elements which utilize the process of color value overlapping. Significance of FVTM work and possible of future work using new model build, train and mapping. By presenting the process of creating generative art work using machine learning API, the paper aim to stimulate more active collaboration between machine learning and art of various disciplines.

논문요약

머신러닝은 4차 산업혁명의 핵심 기술로 디자인과 예술 창작에도 큰 변화를 가져오고 있다. 미디어아트는 첨단 예술과 테크놀로지를 결합시킨 아트 분야로 머신러닝을 활용한 창작의 가능성의 실험이 활발한 분야이다. 본 논문은 머신러닝을 활용한 미디어아트 작품 FVTM:From Vera to Magenta(2019)의 제작 과정을 분석을 통해 머신러닝의 아트와 디자인의 실제적 활용 가능성을 탐구하였다. 먼저 인공지능과 머신러닝의 기술의 발전과정을 살펴보고, 4가지 주요 머신러닝 프레

임워크들의 특징을 비교하였다. 본 연구의 프레임워크로 선택된 구글 텐서플로우의 특징을 밝히고, 아티스트를 위해 개발된 마젠타 API의 사운드 모델과 이미지 모델의 구성을 분석하였다. 작품 제작 단계에서 FTVA의 오픈소스를 활용한 리믹싱의 진화과정을 밝히고, 작품 구현의 두 가지 축을 이루는 마젠타 API의 사전 훈련된 VAE 모델을 적용한 사운드 생성과 색상값의 중첩 과정을 통한 시각 요소의 생성 및 매핑과정을 설명하였다. 완성된 FVTM 작품의 의의와 새로운 모델 훈련 및 적용을 통한 후속 연구의 가능성을 논의하였다. 본 논문은 머신러닝 API를 활용한 실제 생성형 작품제작을 통하여 머신러닝과 아트와 디자인의 협동 창작을 촉진하기를 기대한다.

1. 서론

1.1. 연구목적

인공지능은 4차 산업혁명의 핵심 요소로, 생활의 모든 측면에 근본적인 변화를 가져오고 있다. 2000년대 초반까지 컴퓨터 공학의 한 분야로 인공지능은 주식 매매 등 제한된 분야에서 활용되었지만, 인공지능, 특히 머신러닝 기법은 2012년 이미지넷이미지인식대회¹⁾에서 머신러닝의 한 종류인 딥러닝 기법이 적용된 알렉스넷이 84%의 정확도로 우승한 이후 학계의 큰 주목을 받기 시작했다. 딥러닝 프로그램은 오픈소스로 공개되어 활발한 연구를 촉진했고, 2016년 구글의 딥러닝으로 훈련된 인공지능 바둑 프로그램인 알파고(AlphaGo)와 이세돌 기사의 대국 생중계를 통해 딥러닝은 크게 대중적 주목을 받게 되었다. 딥러닝은 자율주행 자동차, 스마트폰 개인비서, 질병진단 등 광범위한 영역에 적용되며 인공지능의 실용화를 앞당기고 있다.

인공지능의 발달에 따라 인간 고유의 영역으로 여겨지던 창작 분야에도 인공지능의 적용이 가능할 것인지도 본격적으로 논의 및 연구가 시작되었다. 하지만 현재까지 인공지능 기술, 특히 머신러닝 관련 기술의 진입장벽이 높은 상태이며 실제 아티스트, 디자이너들이 머신러닝을 창작 작업에 활용할 수 있는 범위는 매우 제한적이다. 본 연구는 인공지능, 특히 머신러닝 기술을 활용한 예술 창작의 활성화를 위해

머신러닝 제작 환경들을 고찰하고, 딥러닝 모델 구현을 위한 라이브러리 분석, 텐서플로우의 응용을 통한 작품의 구현을 통해 머신러닝을 디자인과 미디어아트에 실제적으로 적용시킬 수 있는 가능성을 검증 해보는 것을 목적으로 한다.

1.2. 연구방법 및 범위

본 연구는 방법은 인공지능의 구현법 중 하나인 머신러닝 모델의 미디어아트적 활용에 초점을 맞출 것이다. 본 논문의 범위로 머신러닝 모델의 수학적 구현방법과 알고리즘은 포함되지 않으며, 인공지능과 예술 디자인에 대한 결합에 대한 이전 연구들에 대한 고찰을 바탕으로, 사전에 훈련된 모델들이 로우레벨 머신러닝 프레임워크 구글 텐서플로우(Tensorflow), 응용 모델 API 텐서플로우 마젠타(Magenta)의 구성과 자바스크립트버전인 Magenta.js에서 제공하는 모델들을 분석하고, 이를 기반으로 생성형 웹 미디어 작품의 제작과정을 제시하는 것을 연구의 방법으로 한다.

제작과정의 고찰을 통해 머신러닝의 생태계에서 아티스트와 디자이너들이 어떻게 실제적인 개발과 창작의 과정에 참여할 수 있을지 필요한 도구와 제작의 단계를 모색할 것이다.

2. 이론적 고찰

2.1. 인공지능과 머신러닝의 개념과 발전

컴퓨터가 단순히 외부의 입력을 받아 처리한 결과를 출력하는 기계 이상의 용도를 가질 수 있음을 최초로 주장한 학자는 앨런 튜링(Alan Turing)이다. 그는 ‘계산 기계와 지능(1950)’이라는 논문에서 컴퓨터가 수학적 계산 뿐 아니라 논리회로를 통해 인간과 유사한 사고의 과정을 구현하는데 사용할 수 있음을 최초로 주장하였다.

‘인공지능(Artificial Intelligence)’이란 용어는 1956년 ‘인간 지능의 모든 측면을 그대로 재현하는’ 기계를 만들기 위해 학계 전문가들이 한 자리에 모였던 다트머스회의에서 수학자 존 맥카시(John McCarthy)가 제안하였다.²⁾ 이 용어는 곧 대중적으로 큰 반향을 일으켜 인간처

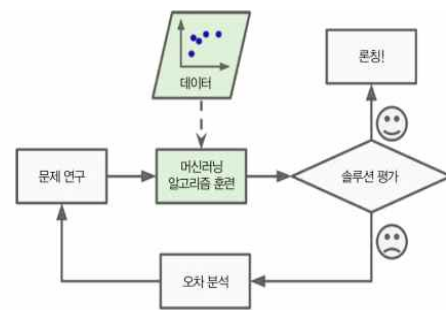
1) 2010년 시작된 이미지인식 경진대회. 대용량의 이미지데이터를 제시하고 이미지 인식 알고리즘의 성능을 평가하는 대회

2) Jerry Kaplan. (2015). 신동숙 역. 인간은 필요없다 - 인공지능 시대의 부와 노동의 미래. 한스미디어. p.37.

럼 사고할 수 있는 컴퓨터에 대한 성급한 기대와 함께 그 기계의 오용이 가져올 우려를 함께 가져왔으나 '사전에 프로그래밍 된 기능만을 수행'하는 것이 아닌 스스로 의사결정이 가능한 컴퓨터의 개발을 위해 해결해야 할 기술적 문제가 너무 많았고, 초기 결과물들의 성과가 기대에 크게 못 미쳐, 범용적 인공지능 개발에 대한 관심과 지원도 줄어들었다. 인공지능의 이론적 발전은 연역적 방식과 귀납적 방식의 두 가지 갈래로 이루어졌는데 귀납적 방식은 프랑크 로젠블라트(Frank Rosenblatt)(1958)가 제안한 인간의 뇌신경망 구조를 모방한 퍼셉트론(Perceptron)의 방식이 기원이며, 연역적 방식으로는 지식 데이터베이스를 통해 인간과 유사한 수준의 컴퓨터를 실제적으로 개발하는 전문가 시스템(Expert System)의 방식에서 기원을 찾아볼 수 있다. 1970년대 다양한 전문가 시스템이 등장했지만 입력해야 하는 자료의 양이 너무 방대해서, 신경망 시스템이 다시 주목을 받게 되었고, 1990년대 들어 컴퓨터의 대중화, 인터넷을 통한 데이터의 빠른 생성과 호환, 프로세서 속도의 발달이 본격적으로 시작되어 이전까지 이론적으로만 연구되었던 신경망 연구의 실용화를 위한 기틀이 다져졌다.³⁾ 2000년대 들어 인터넷 네트워크 이용자수와 생성자료 수가 폭발적으로 증가하는 빅데이터(Big Data)시대가 시작되면서 기존의 귀납적 처리 방법에서 사용한 것 보다 훨씬 많은 데이터를 이용해 신경망을 구성할 수 있게 되었다.

2006년 제프리 힌튼은 기존의 인공신경망에 여러 개의 은닉층(Hidden layer)⁴⁾들을 더한 심층신경망(Deep Neural Network)을 제안하여 2012년 이미지넷대회에서 인간의 수행도를 능가하는 정확도를 보였다. 이때부터 일반적인 사고체계를 모방하는 인공지능 대신 특정 과제(task)를 효율적으로 처리할 수 있는 머신러닝(Machine Learning)이 주목받게 되었다. 머신러닝은 '인공지능' 연구의 하위 분야로, 방대한 데이터집합을 바탕으로 주어진 특정 과제를 해

결할 수 있는 최적의 함수를 모델의 형태로 구현하는 방법이다. 머신러닝의 급격한 성능향상을 가능케 된 이유들은 CPU 기술의 발전과 병렬컴퓨팅 기법의 도입으로 컴퓨터의 처리속도가 빨라졌고, 수학적 모형의 발전을 통해 처리 알고리즘이 향상되었으며, 인터넷을 통해 실시간으로 많은 양의 데이터가 생성되기 때문에 들 수 있다. 머신러닝은 입력 - 처리 - 출력의 함수로 기술될 수 있는데 처리과정이 사전에 작성된 알고리즘을 바탕으로 하는 것이 아닌 다층 레이어의 학습과정을 통해 생성, 수정된다는 점이 기존 프로그래밍 방식과 근본적인 차이점이다. 머신러닝은 '인간이 사전에 설계한 알고리즘대로 실행되지 않고 데이터를 바탕으로 스스로 최적화 학습하는 과정을 통해 입력에 따른 출력값을 조절하는 프로그램'이다.



[Fig. 1] 머신러닝의 훈련 및 데이터 처리과정⁵⁾

[Fig. 1]에 표시된 것처럼 머신러닝의 과정은 입력 데이터를 바탕으로 학습할 피쳐(feature)⁶⁾들을 찾아내고, 이를 바탕으로 레이어 연산을 거치면서 출력 값을 산출해낸다. 이 과정을 학습(learning)이라 하는데 대표적 학습법은 지도학습, 비지도학습, 강화학습이 있다. 입력데이터의 수가 증가할수록 훈련으로 예측되는 출력값의 정확도가 향상되는데 여러 단계의 심층(deep)레이어에서 입력 값이 어떻게 출력되는 지는 데이터에 따라 계속 조정되면서 최적값을 찾게 된다. 훈련을 통해 생성된 함수를 모델(model)이라 한다.

3) Michael Negnevitsky(2013). 김용혁 역. 인공지능 개론. 한빛아카데미. pp.37-42.

4) 신경망을 구성하는 여러 계층의 노드 중, 외부에서 직접 접근할 수 없는 계층의 노드들을 지칭하는 용어

5) Aurlien Gron(2018). 박해선 역. 핸즈온 머신러닝 - 사이킷런과 텐서플로를 활용한 머신러닝, 딥러닝 실무. 한빛미디어. p.36.

6) 관찰대상에서 발견된 개별적이고 측정가능한 경험적 속성

2.2. 머신러닝 프레임워크와 텐서플로우

인공신경망을 활용한 연구를 촉진시킨 배경에는 효율적인 프레임워크의 기여가 중요한 역할을 하였다. [table 1]에 현재 많은 개발자들이 참여 중인 4개의 주요 프레임워크들과 특징들을 제시하였다. 머신러닝 중 다층레이어로 구성된 딥러닝(Deep Learning)을 구현한 프레임워크들은 다양한 모델을 효율적으로 제작할 수 있는 라이브러리와 인터페이스를 제공하여 개발자들이 다양한 목적에 맞는 모델 및 머신러닝 프로그램을 개발 할 수 있는 환경을 지원한다.⁷⁾ 또한 이들 다수가 오픈소스로 개발되고 유지되어 전 세계의 개발자, 디자이너, 아티스트들이 활용할 수 있다. 머신러닝의 성능 향상을 위해서 효율적 프레임워크의 역할이 가장 중요하기 때문에 아마존, 마이크로소프트 등 주요 IT 기업들도 자체적으로 머신러닝 프레임워크를 개발, 오픈소스로 공개하고 있다.

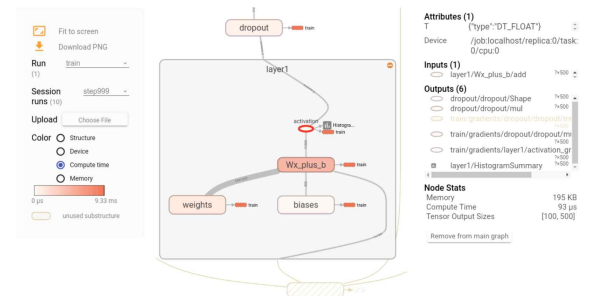
[Table 1] 주요 머신러닝 프레임워크⁸⁾

프레임워크	개발사	특징	적용분야
Tensor flow	구글	<ul style="list-style-type: none"> 하위레벨 API 제공 구글서비스로 적용되어 실제 활용력 검증됨 다양한 언어지원 	<ul style="list-style-type: none"> 언어탐지 얼굴인식 오브젝트 구분 소리인식, 합성 비디오 분석
PyTorch	페이스북	<ul style="list-style-type: none"> 파이썬 Torch 기반 개발 실행중 계산 메모리생성 효율적 메모리 관리 	<ul style="list-style-type: none"> 페이스북 기반 머신러닝 서비스
Keras	Francis Collet	<ul style="list-style-type: none"> 상위레벨 API 다양한 로우레벨 플랫폼 빠른 실행을 위한 UI 	<ul style="list-style-type: none"> CoreML 응용앱 Android 응용앱
Caffe	Yangqing Jia	<ul style="list-style-type: none"> 다양한 언어 인터페이스 이미지 처리와 학습에 빠른처리속도 	<ul style="list-style-type: none"> 이미지 인식 스타일 인식

현재 가장 많은 사용자가 이용하는 텐서플로우는 구글브레인팀에서 2011년 개발한 디스트벨리프(DistBelief)를 기원으로 한다. 구글은 디스트벨리프를 자사의 검색, 음성검색, 광고, 포토, 맵 등의 서비스의 분석과 제품 개발에 적용한 후, 이를 발전시켜, 2015년 텐서플로우

(Tensorflow)라는 이름으로 오픈소스로 공개하였다. 텐서플로우는 맥, 윈도우, 리눅스 등 다양한 OS를 지원하며, 구글의 다양한 서비스를 통해 다년간에 걸친 실제 검증이 이루어진 높은 신뢰성으로 현재 가장 높은 관심과 활용을 받는 라이브러리가 되었다.⁹⁾

텐서플로우는 C++ 로 제작된 로우레벨 프레임워크로, 파이썬 언어를 통해 모델을 생성시키고 훈련시키는 방식을 기본으로 하고 있지만



[Fig 2]텐서보드를 통한 GUI 노드 구성¹⁰⁾

구글브레인팀은 텐서플로우는 머신러닝의 대중화에 기여할 수 있게 하기 위하여 [Fig.2]와 같은 텐서보드라는 GUI를 함께 제공하였다. 텐서보드를 통해 각 텐서의 흐름도를 시각적으로 확인할 수 있고 노드를 직관적으로 구성할 수 있다. 2017년, 웹 환경에서 텐서플로우의 활용을 위해 자바스크립트API가 추가되어 웹을 활용하여 텐서플로우 라이브러리를 직접 사용하는 것이 가능해졌다. tensorflow.js를 통해 아티스트와 디자이너를 포함한 머신러닝 비전 공자들도 기존에 트레이닝된 모델을 사용하거나, 자체 데이터를 이용하여 머신러닝 모델을 목적에 맞게 최적화시키는 것이 용이해졌고, 자바스크립트를 활용하여 새로운 모델을 빌드하고 학습시키는 것도 가능해져 다양한 협업작업들이 등장하기 시작했다.¹¹⁾

2.3. 기존 연구 고찰

인공 지능과 머신 러닝에 대한 아트와 디자인에 관련된 이전 연구는 크게 세 가지 범주로 구분해 볼 수 있다.

9) 조만석.(2017). 인공지능 오픈소스 라이브러리 텐서플로우(tensorflow)와 인공지능 응용 소프트웨어 개발. 한국통신학회지(정보와통신), 34(10), 57.

10) <https://www.tensorflow.org/tensorboard/r1/graphs>

11) <https://www.tensorflow.org/js?hl=ko>

7) <https://www.analyticsvidhya.com/blog/2019/03/deep-learning-frameworks-comparison/>

8) <https://towardsdatascience.com/deep-learning-framework-power-scores-2018-23607ddf297a> 을 바탕으로 재구성

먼저, 인공지능의 등장에 따른 예술 작품 제작의 미학적 의미와 가치에 대한 고찰을 다룬 연구이다. 김희선(2017)은 인공지능시대에 작품의 구현을 위한 기술적 노력은 기계의 몫으로 변환됨에 따라 예술작품의 생산-소비의 간극이 줄어들 것이며, 동시에 예술의 가치가 기술의 완성도가 아닌 창작자의 아이디어와 의도가 중요한 방향으로 발전될 것으로 보았다.¹²⁾ 두 번째는, 사례 분석연구로 인공지능과 머신러닝 기법을 활용하여 구현된 작품들에 대한 분석 연구들인데 최효승(2017)등은 창작 사례들을 분석하여 인공지능 예술의 특징을 창의성, 유희성, 편리성, 가변성의 네 가지 특성으로 구분했다. 가변성의 경우 작품 상 큰 차이를 보이지 않아 유의미한 특성으로 보기 힘들다.¹³⁾ 조서현(2019)등은 인공지능을 활용한 예술작품들을 창의성과 행위의 주체자가 누구인가에 따라 간접적 표현측면과 직접적 표현측면으로 구분하고 각 사례들의 특징과 한계에 대해 언급했다.¹⁴⁾ 세 번째는, 머신러닝을 활용하여 작품을 제작하는 과정에 대한 연구로 송윤섭(2016)등은 비전문가들을 대상으로 인공지능 프로그래밍에 대한 연구를 진행하였는데 실험대상이 이미 Python 등의 프로그래밍 언어를 구사할 줄 안다는 점에서 완전한 비전문가로 볼 수는 없다. 아직 머신러닝을 아티스트 혹은 비전문가의 시각에서 실제로 접목시킨 연구 사례는 많지 않다.

3. 텐서플로우 프레임워크와 마젠타 API

2017년 구글브레인팀에서 텐서플로우 기반으로 개발한 크리에이터를 위한 마젠타(Magenta)API를 공개하였다. 마젠타는 이전까지 다양한 머신러닝 모델의 개발과정에서 개별적인 결과물로 제공되던 주요 머신러닝 모델들을 오픈소스 라이브러리 형태로 제공함으로써, 아티스트들이 좀 더 용이하게 다양한 머신러닝을 이용한 창작 결과물을 도출하고, 이를 응용

12) 김희선 (2017). 4차 산업혁명과 인공지능시대 전통예술의 미래전망과 과제. 한국예술연구, 16. pp.18-19.

13) 최효승, 손영미 (2017). 인공지능과 예술창작 활동의 융복합 사례분석 및 특성 연구. 한국과학예술포럼, 2017. 3. pp.296-297.

14) 조서현, 이필하(2019). 딥 러닝을 활용한 인공지능의 예술표현 사례 연구. 조형디자인연구, 22(1), p.206.

할 수 있는 가능성을 제공하였다. 마젠타 API의 구성은 예술창작 분야 중 사운드와 이미지 처리 부분에 활용할 수 있는 모델들로 구성되어 있으며 현재 파이썬과 자바스크립트 라이브러리로 제공된다.

3.1. 텐서플로우 마젠타 뮤직 모델

마젠타의 사운드API는 사운드 아트 및 음악 창작을 위한 사전 훈련된 다양한 모델을 제공하는데 [Table. 2]에 각 모델이 어떤 기능을 하는지 비교하였다.

[Table 2]. 마젠타 뮤직 API 의 주요 모델 구성

이름	딥러닝 모델	데모 앱	기능
Onset sAndFrames	Piano Transcription Model	piano Scribe	오디오신호를 MIDI 파일 형식으로 변경해 줌. 피아노 솔로곡에 가장 좋은 성능
Music RNN	LSTM-Based Language Model	Neural Drum Machine	마젠타의 LSTM - 기반의 다양한 RNN 모델 구현
Music VAE	VAE(Variational Auto Encoder) Model	Endless Trios	마젠타의 변환 자동인코딩 모델 구현. 멜로디, 드럼 등
Piano Genie	VQ-VAE model	Piano Genie	8 버튼 입력을 실시간으로 88 건반 피아노로 매핑함

마젠타 API를 바탕으로 함께 제공되는 데모 앱을 통하여 다양한 모델의 실제 구현과정을 분석하고, 이를 응용한 딥러닝 뮤직 앱 창작이 가능하며 자바스크립트 버전을 통해 웹 앱의 형태로도 창작물을 제작할 수 있다.

마젠타 스튜디오는 한 단계 더 나아가 커맨드 기반 프로그래밍의 단계를 없애고, 뮤지션들에게 친숙한 DAW(Digital Audio Workstation)인 에이블톤 라이브(Ableton Live)의 플러그인 형태로 제작되어 [Fig. 3] 과 같이 아티스트들이 친숙한 직관적인 에이블톤 환경에서 머신러닝을 직접 구현할수록 더 편리한 인터페이스로 구성되어있다.



[Fig. 3] 마젠타 스튜디오 에이블톤 라이브 플러그인 15)

3.2. 텐서플로우 마젠타 이미지 모델

텐서플로우 마젠타의 이미지API는 드로잉을 생성해내는 스타일 트랜스퍼와 스케치RNN을 주요 모델로 제공한다. [Table. 3]에 각 모델의 특징을 비교하였다.

[Table. 3]. 마젠타 이미지 API 의 주요 모델 구성

이름	딥러닝 모델	데모 앱	기능
Image Stylization	Fast Arbitrary Image Stylization	Arbitrary Style Transfer	오디오신호를 MIDI 파일 형식으로 변경해 줌. 피아노 솔로곡에 가장 좋은 성능
Sketch RNN	LSTM-Based Language Model	Teach Machine to Draw	스트로크 기반의 벡터 드로잉을 바탕으로 완성 이미지를 예측

스타일 트랜스퍼는 Ghiasi(2017)등이 제시한 임의의 회화스타일로 만들어진 이미지를 입력 받아 사전에 학습된 다양한 회화 스타일로 출력해주는 모델을 텐서플로우로 구현한 것이다.

스케치RNN은 David Ha(2017)등이 제안한 순환신경망(Recurrent Neural Network)모델을 바탕으로 구현한 것으로, 획(stroke)기반의 벡터이미지에서 각 획의 위치와 펜의 상태를 입력데이터로 받아 어떤 오브젝트를 그리는지를 예측하도록 훈련된 모델이다. 현재 사전 훈련된 1014개의 오브젝트를 구분할 수 있으며 데모 앱을 통하여 실제 구현 방법을 실험할 수 있다. 이미지 모델들은 직관적으로 확인할 수 있는 결과물들을 산출하기 때문에 머신러닝 기법을 아트에 활용한 사례로 가장 먼저 연구된 분야이며 다양한 연구사례들을 찾아볼 수 있다.

4. 마젠타 활용 미디어아트 제작

4.1. 제작의도 및 구성

미디어아트와 머신러닝의 발전에 공통된 동력으로 창작의 코드를 공유하는 오픈소스(Open Source)의 철학을 들 수 있다. 머신러닝의 경우 구글, 페이스북, 아마존 등 거대 기업들도 자사의 머신러닝 프레임워크를 오픈소스로 공개하여 누구나 기술에 접근하고 활용할 수 있다. 본 연구의 주제가 된 머신러닝 미디어아트 작품 “FVTM : From Vera to

Magenta(2019)”도 오픈 소스를 적극적으로 활용하였고 작품의 변환 및 발전에 리믹스 작업 방식을 적용하였다.

작품의 기본 아이디어는 텐서플로우 마젠타의 사운드모듈을 이용하여 임의의 피치(pitch)를 생성한 후 머신러닝 모델을 통해 멜로디와 리듬을 생성하고 이를 시각적인 움직임에 매핑시키는 것이었다. 기본 제작 구조는 오픈소스의 리믹싱 형태로 진행되었는데 각 반복(iteration)단계에서 이전 단계의 변수의 변형과 새로운 변수의 추가로 작품은 새로운 형태로 진화되었다. 기원은 컴퓨터아트와 생성형아트(Generative Art)의 선구자 베라 몰나르(Vera Molnar, 1924)의 작품 Un Deux Trois로, 캔버스를 세 개의 구역으로 나누고 각 영역 별 선분의 복잡도를 알고리즘을 이용하여 세 단계의 선으로 구성된 추상적 평면 회화 작품이었다.¹⁶⁾ 디지털 진화의 첫 번째 단계는 팀홀만(Tim Holman)이 베라의 작품을 자바스크립트코드를 주 언어로 캔버스오브젝트를 이용하여 코딩한 것으로 웹아트 작품으로 발전시키고 제작과정을 오픈소스로 공개한 것이다. 베라의 컨셉이 변수화되어 표현된 것이 주요 변화이다.¹⁷⁾ 팀의 작품이 오픈소스로 공개되었기 때문에 다양한 응용작업이 창작되었고, 코드펜(codepen.io)을 통해 사운드에 반응하는 리믹스작업으로 발전되었다.¹⁸⁾ 두 번째 진화로, 구글의 마젠타 프로젝트 개발팀의 일원인 모니카 딘클루스쿠(Monica Dinculescu)가 팀의 작품을 마젠타 뮤직 API의 세 가지 악기 모델에 반응하도록 매핑(Mapping)시켜 마젠타를 통해서 생성된 음악에 반응하는 머신러닝 모델이 생성하는 작품 “Hearing Vera molnar”로 발전시켰다.¹⁹⁾

본 연구는 세 번째 진화단계에 해당하며, 모니카의 작품을 기반으로 텐서플로우 마젠타를 이용하여 사운드 생성 모델을 다른 사전 훈련된 멜로디와 악기세트로 변경하고, 이미지와 움직임의 생성을 연구자의 기존 인터랙티브 미디어 퍼포먼스 작품 ‘MNRP:My Name is Rain

15) <https://magenta.tensorflow.org/studio/>

16) <https://www.ara.na/block/2110218>

17) <https://generativeartistry.com/tutorials/un-deux-trois/>

18) <https://codepen.io/Rumyra/full/qMaEmM>

19) <https://glitch.com/~hearing-vera-molnar>

Performance(노승관, 2015)²⁰에서 구현한 시각 현상적 움직임의 시각요소들을 변환하였다. MNRP 원작에서 사운드는 미리 준비된 사운드 샘플들을 실시간으로 호출(trigger)하는 방식으로 구현하였는데 머신러닝을 이용하여 자동적으로 새로운 멜로디를 생성하고 시각적 결과물과 매핑 시키는 함수의 구현은 본 연구에서 가장 중요한 부분이었다. [Fig. 4]에 전체적인 작품의 구상 및 발전과정의 흐름도를 보였다. 기본피치를 먼저 생성하고 이를 멜로디 및 리듬으로 변환한 후 비주얼매핑을 한 결과를 렌더링하는 구성이다.



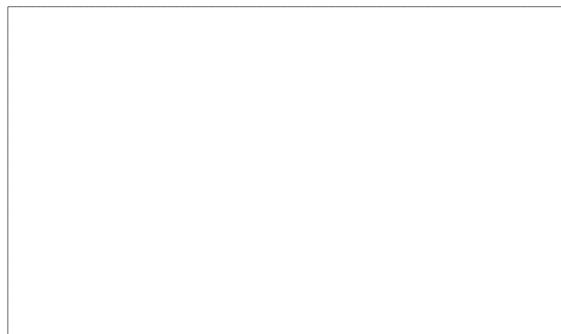
[Fig. 4] FVTA 작품 구현 기본 플로우

4.2. 사운드와 이미지 모듈 구현

“From Vera to Magenta” 작품의 실제 제작 과정은 전체 프레임워크의 설계, 사운드 모듈의 구현, 시각 모듈 구현의 단계로 구성되어 있다.

1) 사운드 모듈 구현과정

사운드의 생성은 훈련된 VAE 모델 사용하여 구현하였다. 멜로디는 연속되는 시간 정보로 이루어져 있기 때문에 기존의 공간적 신경망으로는 시간에 따른 정보의 순차적 상호관계가 중요한 멜로디나 애니메이션을 구현하기 어렵다.



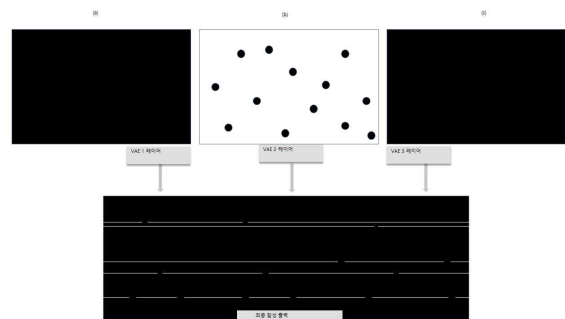
[Fig. 5] 사운드 모듈의 RNN 전처리 Conductor²⁰

이를 보완하기 위해 [Fig. 5]에 표시된 것처럼

럼 전처리코드를 RNN디코더에 직접 연결하기 전에 지휘자역할을 하는 conductorRNN레이어를 거치도록 하여 좀 더 자연스럽게 악기의 음색에 따른 멜로디 마디를 구성할 수 있다. 이 설계를 이용하여 훈련된 다양한 악기별 16마디의 멜로디 모델이 제공된다. 마젠타라이브러리에서는 현재 드럼루프, 멜로디, 트리오, 멀티트랙 VAE 모델을 제공하며, 본 연구의 작품에는 trio_4bar_lokl_small_q1 멀티트랙 모델을 사용하여 사운드를 생성시켰다. 이 멀티트랙 모델은 멜로디, 베이스, 드럼의 멀티트랙 구성으로 되어있고, conductor LSTM에 의해 3개의 코드 벡터를 배치하고 위치시킨다. 시작하는 노트가 주어지면 훈련된 신경망 모델이 멜로디를 생성해낸다. 3개의 사운드 값은 각각 다른 시각적 요소의 렌더링을 위한 변수로 전달된다.

2) 비주얼 구현과정

화면 비주얼의 구성은 자바스크립트의 캔버스 객체를 통해 구현되었다. 백색 배경에 세 가지 다른 두께의 시각적 요소가 나타나는 이전 진화단계 작품들과 달리, MNR의 제작 컨셉인 시각적으로 같은 색상값의 중첩기법을 사용하였다. 검정 바탕 위에 검정 원 요소들과 흰색 선 요소들을 교차 배치하여 중첩에 의해 비가 내리는 듯한 새로운 애니메이션의 시각을 경험하도록 하는 것이 “From Vera to Magenta” 작품의 시각적 구현이며, 그 구성은 [Fig. 6]에 제시하였다. 검정 바탕, 수직 움직임을 하는 검정 블랍, 흰색 수평선이 중첩되며 멜로디 단계에서 전달 받은 변수에 의해 중첩선의 위치가 변하는 구조이다.



[Fig. 6] 중첩을 통한 이미지 생성 과정

20) <https://magenta.tensorflow.org/music-vae>

교차되는 반대 레이어의 상호과정으로 최종적으로 표출되는 비주얼은 소리에 따라 예측 할 수 없는 새로운 형태를 창출해낸다.

3) 최종 결과

최종 구현작품은 웹에서 생성형(generative) 미디어 설치물로 구현되었다. 머신러닝으로 구현된 사운드 엔진은 기존 생성형 알고리즘의 정형성을 벗어나 실행 시 마다 다른 멜로디와 리듬을 만들어내며 작품이 가지는 우연성에 의한 시각적 구성을 표현한다. 현재 구현은 사전 훈련된 모델을 바탕으로 사운드폰트를 사용하여 출력되지만, 차후 마이크 입력 등을 통해 샘플 자체를 실시간 분석 사용하는 형태로 확장시켜 인터랙티브 미디어 설치 형태로 발전시킬 계획이다.

5. 결론

본 연구에서는 머신러닝 프레임워크의 발전을 고찰하고, 구글의 오픈소스 프레임워크 텐서플로우를 기반으로 제작된 마젠타 API를 이용하여 기존의 미디어아트 작품을 머신러닝으로 구동되는 생성형아트 작품으로 변화시킨 작품 "From Vera to Magenta"의 제작과정을 통해 머신러닝의 실제적 창작 가능성을 실증적으로 제시하였다. 본 연구는 오픈소스의 리믹싱 방식으로 진행하여 사운드, 애니메이션 등 제작 단계 별 머신러닝의 적용을 실험하였다. 머신러닝은 가장 활발한 연구가 이루어지는 분야로 제작을 위한 도구와 환경도 빠른 속도로 기술적 요구사항들이 창작에 장애가 되지 않도록 진화하고 있다. 아직까지 머신러닝을 아티스트나 디자이너가 창작에 적용하기 위해서는 자바스크립트나 파이썬 등 기존의 프로그래밍언어를 구사할 수 있어야 하며, 머신러닝 모델의 작동을 이해하기 위해 수학적 구현에 대한 저수준레벨의 지식이 필요하다. 포토샵과 같은 직관적인 작업이 가능한 고수준 머신러닝 저작 도구는 없다. 하지만 마젠타 스튜디오와 같은 DAW에서 그대로 불러서 사용할 수 있는 플러그인 형태의 머신러닝 모델이나, runwayML²¹⁾ 같은 시각적 인터페이스 환경이 지속적으로 등

장하고 있어 머신러닝과 아트, 디자인 간의 간격은 빠른 속도로 좁혀지고 있다. 인공지능의 발전에 있어 특정 과제를 효율적으로 수행할 수 있는 머신러닝 모델 개발이라는 공학자의 역할, 인공지능을 활용한 창작 예술 작품이 가지는 미학적 가치를 논하는 철학, 미학자의 역할이 모두 중요하다. 머신러닝의 적용을 통한 구체적인 예술적 결과물의 수가 증가할수록 인공지능이 창작에 어떤 영향을 주게 될지 구체적인 사례를 통한 학습을 통해 향후의 방향에 대한 예측도 가능해질 것이다.

참고문헌

- 김희선 (2017). 4차 산업혁명과 인공지능시대 전통예술의 미래전망과 과제. 한국예술연구, 16, 5-26.
- 송운섭, 권영미(2016). 비전문가의 인공지능 분야 진입 실험. 한국멀티미디어학회지, 20(3), 53-61.
- 조만석 (2017). 인공지능 오픈소스 라이브러리 텐서플로우와 인공지능 응용 소프트웨어 개발. 한국통신학회, 9, 55-63.
- 조서현, 이필하.(2019). 딥 러닝을 활용한 인공지능의 예술표현 사례 연구. 조형디자인연구, 22(1), 193-211.
- 최효승, 손영미 (2017). 인공지능과 예술창작 활동의 융복합 사례분석 및 특성 연구. 한국과학예술포럼, 2017. 3, 289-299.
- 황원지연, 탕위안.(2017). 예제로 풀어보는 텐서플로우 원리와 응용. 선문각.
- Aurlien Gron(2018). 박해선 역. 헨즈온 머신러닝 - 사이킷런과 텐서플로를 활용한 머신러닝, 딥러닝 실무. 한빛미디어.
- Jerry Kaplan. (2015). 신동숙 역. 인간은 필요없다 - 인공지능 시대의 부와 노동의 미래. 한스미디어.
- Michael Negnevitsky(2013). 김용혁 역. 인공지능 개론. 한빛아카데미.

21) <https://runwayml.com/>