

딥러닝 기반 손 그림 심리분석 알고리즘 연구

장 정 인¹ · 김 병 규^{2*} · 김 영 나³ · 박 용 배⁴

¹숙명여자대학교 IT공학과 석사과정 ^{2*}숙명여자대학교 IT공학과 교수

³(주)지영커뮤니티 대표이사 ⁴(주)지영커뮤니티 연구소장

A Study on Deep learning-based Hand Drawing Psychological Analysis Algorithm

Jung-In Jang¹ · Byung-Gyu Kim^{2*} · Young-Nah Kim³ · Young-Bae Park⁴

¹MS Student, Department of IT Engineering, Sookmyung Women's University, Seoul, Korea

^{2*}Professor, Department of IT Engineering, Sookmyung Women's University, Seoul, Korea

³CEO GYoungcommunity Co., Ltd., Seoul Rep. of Korea

⁴CTO, GYoungcommunity Co., Ltd., Seoul Rep. of Korea

[요 약]

간단한 아이들의 낙서부터 세계적인 미술가들의 그림들까지, 인간의 심리는 그림 속에 녹아들었다. 그림은 그들이 가진 환경에 대한 태도, 개념, 경험, 욕망을 표현하며 감정과 성격을 확인할 수 있게 한다. 이것은 그림 심리 치료의 매개체가 될 수 있으며 '예술 치료'라 부른다. 본 논문에서는 손 그림을 계층적으로 분석하고 특정한 심리 상태를 분석할 수 있는 효과적인 방법을 제시한다. 먼저, 손으로 그려진 분석용 그림이 입력되면 객체 검출 기법을 통해 전체 객체를 구성하고 있는 세부적인 부분 객체들을 추출한다. 이 추출된 부분 객체들은 개별 속성 분류를 위해 2차원 CNN에 입력으로 주어진다. 최종적인 분석 결과는 분류기를 통해 분석된 패턴 속성을 기반으로 생성한다. 제시한 알고리즘은 다양한 그림 심리 검사 중 '나무 그림 검사'와 '고양이 그림 검사'에 적용되었으며 분석 정확도는 각각 93%와 96%를 획득하였다.

[Abstract]

From simple scribbles of children to paintings by world-class artists, human psychology is melted into drawings. They express the concept, experience, desire, and attitude to the environment they have acquired, and can identify the person's tendency and personalities. This can be used as a medium for psychotherapy, called "art therapy." We devise an efficient scheme to analyze the hand drawings based on a hierarchical structure, which allows us to produce certain results. First, we extract some separated parts which is consisted of whole objects by object detection. With the extracted parts, we perform the classification task based on 2D CNN to define key attribute. For the given image, the analyzed report is constructed based on the key attributes. The proposed algorithm is applied to "tree test" and "cat test" among many drawing psychological tests, and achieves an analysis accuracy of about 93% and 96%, respectively.

색인어 : 그림 심리 검사, 객체 분류, 딥러닝, 객체 검출, 계층적 구조

Keyword : Art Therapy, Classification, Deep learning, Object detection, Hierarchical structure

<http://dx.doi.org/10.9728/dcs.2021.22.9.1501>



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 22 July August **Revised** 07 September 2021

Accepted 23 September 2021

***Corresponding Author, Byung-Gyu Kim**

Tel: +82-2-2077-7293

E-mail: bg.kim@sookmyung.ac.kr

I. 서론

미술은 조형적인 방법으로 사람의 생각을 표현하는 예술의 한 분야로 자신의 감정이나 사상을 전달하는 수단이 되기도 한다. 이러한 미술 활동은 단순히 타인에게 시각적인 즐거움을 주는 요소가 되기도 하지만 사람의 심리를 진단하고 교육, 치료 등 다양한 분야에서 활용할 수 있는 수단이 되기도 한다. 미술 치료를 한마디로 정의할 수는 없지만 ‘미술 치료(Art Therapy)’라는 말은 1940년대 말 영국의 예술가 Adrian Hill이 처음 사용함으로써 그 역사가 시작되었다고 할 수 있다. 그는 폐결핵 환자들을 돌봤는데, 이때 그들이 그림을 그리는 행위가 마음을 편안하게 해주고 불행한 마음을 바로잡게 하는 치료적 효과를 주었다고 주장하였다. 이와 동시대의 미국의 Margaret Naumberg는 그림이 환자와 치료사 간 소통의 매개체가 될 수 있으며 이를 심리적인 이론과 관련지을 수 있다고 주장하였다. 실제로, 미술 치료는 이미지를 만드는 과정과 결과물 그리고 치료 관계의 제공을 모두 포함한다[1].

그림을 통한 심리 검사는 대표적인 미술 치료의 방법으로서 성인뿐만 아니라 아동들에게 많이 적용되는 방법이다[2]. 이는 언어를 매개로 하지 않기 때문에 검사 대상자의 방어와 왜곡에 따른 정확한 진단의 어려움을 덜어주고 복잡한 도구가 필요하지 않아 접근성이 좋다. 보통 검사자가 그림을 그리면 그 그림을 치료사가 해석하는 방향으로 진행이 된다. 그러나 이는 그림의 해석에 대해 주관적인 견해가 개입되거나 치료사의 경험에 따라 결과가 달라질 수 있고, 검사에서 판단까지 많은 시간이 소요될 수 있다. 그러므로 검사의 과정 중 일부에 딥러닝 기법을 도입하면 그림 심리 검사의 접근성, 분석의 객관성과 서비스 속도를 개선시킬 수 있을 것이다.

본 논문에서는 아동의 그림을 받아 분석하는 단계를 딥러닝 기술을 통해 해결하고자 한다. 딥러닝을 통해 실제 객체가 포함된 이미지에서 객체를 찾거나 분류하는 작업은 지난 몇 년간 계속 연구되고 있으며 현재 이미지 분류 같은 경우, 사람보다 좋은 성능을 보여주고 있다. 딥러닝이 사진을 처리하는 것처럼 아동의 그림을 데이터로 하여 훈련한다면 그림 심리 검사의 과정을 일부 자동화시키며 기존의 아날로그적 방식보다 빠르고 정확한 분석이 가능할 것이다.

본 논문의 구성은 다음과 같다. 2장에서는 제안한 알고리즘의 필요성을 중심으로 관련된 연구를 통해 적용 가능한 기술들을 살펴본다. 3장에서는 제안하는 알고리즘의 설계와 구현에 대하여 설명하고 4장에서는 이에 대한 성능을 평가하며 결과에 대한 예시를 보여준다. 5장에서는 결론을 맺으며 추후 연구 방향을 제시한다.

II. 관련 연구

2-1 그림 심리 검사

나무 그림 검사는 ‘Baum Test’ 혹은 ‘Koch test’라고도 불리며 개인의 성격과 감정 상태를 평가하기 위해 사용하는 투영검사법이다[3]. 이 검사는 1952년 스위스 심리학자 겸 상담사 Charles Koch에 의해 고안되었으며 처음부터 누구나 쉽게 사용할 수 있도록 설계되었기 때문에 주로 어린이들에게 유용하게 쓰인다[4]. 2015년 연구[5]에 따르면, 나무 그림 검사는 인지장애와 치매를 성공적으로 진단하는 것으로 증명이 되었는데 인지장애가 있는 환자가 그린 나무는 그 정도에 일반 사람이 그린 나무와의 차이점을 보였다. 예를 들어 알츠하이머 환자가 그린 나무는 전체적인 크기가 작으며 비교적 기울어지지 않았다. 검사의 순서는 다음과 같다. 우선 검사자에게 A4 크기의 종이를 제공하고 마음대로 나무를 그리도록 한다. 제한된 시간 없이 자유롭게 나무 한 그루를 그린 후 제출하면 그 모양새를 분석하여 해석한 결과를 낼 수 있다. 해석은 나무의 크기와 그림의 밀도와 같은 전반적인 특성은 물론 줄기의 패턴과 과일, 가지의 배열과 같은 구체적인 특징들에 따라 달라진다[6].

고양이 그림 검사는 뇌의 오른쪽 또는 왼쪽 반구와 연관이 있는 그림 패턴을 관찰하게 한다. 그림에 색상 사용 빈도나 고양이의 움직임 및 특이한 특징이 있는 그림은 우뇌를 더 많이 사용했다 할 수 있고 정적이고 단순한 경우 좌뇌를 더 많이 사용했다는 것으로 해석한다. 이 검사는 검사자 뇌의 기여 수준을 알 수 있게 해 후에 인지 과정과 성격을 파악하는 것에 도움을 준다. 본 논문에서는 나무와 고양이 두 그림에 대한 심리분석 알고리즘을 각각 제시한다.

2-2 이미지 처리

컴퓨터 비전은 ‘컴퓨터 과학’의 한 분야로 디지털 이미지에서 특징을 추출하고 활용하는 것이 그 핵심이라 할 수 있다. 실제 사물은 다양한 형질로 이루어져 있지만, 컴퓨터에서의 이미지는 픽셀이라는 작은 단위로 이루어져 있다. 픽셀값은 0부터 255 사이의 값을 가지고 있으며 색을 표현할 수 있는데 여러 픽셀이 모이면, 실제의 사물을 표현할 수 있게 된다. 사람의 눈이 사물을 보고 이해하는 것처럼 컴퓨터 이미지를 이해하기 위해서도 객체에서 유의미한 정보를 얻어야 한다. 전통적으로 주성분 분석 같은 통계적인 방법[7]과 지역적 특징을 찾아 유사한 곳을 매칭시키는 방법 등이 있다[8]. 최근 몇 년간 컴퓨터 비전에서 떠오른 방법은 딥러닝이다. 이는 여러 층의 신경망을 쌓은 후 입력을 받아 계층별로 특징을 추출한 후 객체에 대한 정보를 추론할 수 있게 하는 방법이다. 딥러닝이라는 개념은 1980년대에 이미 등장했지만 2012년 ImageNet Large Scale Visual Recognition Challenge (ILSVRC)에서 AlexNet이 우승 이후 기하급수적으로 발전하게 된다[9]. 현재 딥러닝에서 사용되는 네트워크는 초기보다 몇 배는 더 깊어지고 복잡한 구조를 갖게 되었다. 다양한 연구를 통해 훈련방법도 더 정교해졌고, 컴퓨터가 이미지를 이해하는 정도가 인간의 수준보다 더 높아졌으며 얼굴 인식, 이미

지 화질 개선 등 다양한 분야에서 사용되고 있다[10, 11, 12].

이미지 분류(Image Classification)란 주어진 여러 장의 이미지들을 클래스로 나누어 학습시킨 후 후에 새로운 이미지를 다시 입력으로 했을 때 어떤 클래스에 속하는지를 파악하는 문제이다. 1998년 Yann Lecun의 팀이 합성곱 신경망(Convolutional Neural Network)을 이용해 성공적으로 손글씨를 분류하였고 [13], 이후 합성곱 신경망은 더 깊게 쌓거나 [9, 14], 앞단 정보를 유지하는 등 [15] 다양한 방향으로 발전하였다. 본 논문에서는 비교적 간단한 합성곱 신경망을 쌓아 분류에 사용하고 이후 tensorflow2에서 제공하는 학습된 모델을 불러와서 전이학습을 해 그 성능을 비교해본다.

객체 검출(Object Detection)은 분류 문제에서 특정 클래스의 개체가 이미지 안 어느 위치에 있는지 위치 정보를 나타내주는 Localization 과정이 추가된 것이다. 이때 보통 경계 상자(Bounding box)를 그려 찾은 개체를 표시해준다. 이러한 객체 검출 모델은 추론 속도와 탐지 정확도에 따라 유형을 분류할 수 있다. YOLO[16], SSD[17], Retinanet[18] 등은 One-stage Detector로 분류와 지역화를 동시에 수행하기 때문에, 그 속도가 빠르지만, 상대적으로 저조한 성능을 가진다. 그에 비해 Fast R-CNN[19], Faster R-CNN[20], Mask R-CNN[21] 등의 모델은 Two-stage Detector로 검출 과정이 순차적으로 진행되어 상대적으로 느리지만, 정확도가 높다는 장점이 있다.

2-3 컴퓨터를 통한 손 그림 심리 검사

기존 그림 심리 검사는 보통 종이와 펜을 가지고 그림을 그린 후 사람이 그림을 분석하여 결과를 내고 후에 공간을 따로 두어 보관을 하는 방식으로 진행되었다. Jeon et al.은 그림 심리 진단과 컴퓨터를 접목해 컴퓨터 미술 치료를 위한 컴퓨터 심리 진단 장치를 제시하였다[22]. 아동이 터치 모니터에 그림을 그리면 색의 분포, 선의 특성, 그림을 그린 시간 등의 정보를 얻어 심리 진단을 사람 없이 할 수 있도록 하였다. 하지만 이때 선에 대한 정보는 필압을 통해 얻었기 때문에 특별히 센서가 붙은 펜이 필요했다. 이에 반해 본 논문이 제시하는 방법은 이러한 부가적인 펜과 무관하게 그림으로부터 직접 특성을 얻은 후 이를 해석에 활용하는 것이다.

또한 Park et al.은 딥러닝 사물 인식 알고리즘을 활용하여 아동이 그린 그림으로부터 사물 검출이 가능하다는 것을 확인하였고 감성 분석 분야에서의 활용 가능성을 진행하였다 [2]. 하지만 인식 후 이를 활용하여 심리에 대한 진단을 내리는 부분은 구현하지 않았다. 본 연구는 이미지로부터 사물을 검출한 후 해석을 내는 부분까지의 전체 과정을 모듈로 제시하고자 한다.

Ⅲ. 제안된 심리분석 방법

기본 논문에서 제안하는 딥러닝 기반 그림 심리분석 알고리즘은 어린이가 그린 그림으로부터 객체를 인식하고 각 개체를 어떻게 해석할 수 있는지 판단하는 단계를 거친다. 전체 시스템의 구조는 Fig. 1과 같으며 Faster R-CNN을 이용한 객체 인식 모듈과 합성곱 신경망을 이용한 클래스 분류 모듈로 나뉠 수 있다. 먼저 아동이 그린 그림을 컴퓨터 이미지로 만들고 입력으로 집어넣어 객체 검출 모듈을 통해 객체를 찾아 해당 영역을 자른다. 이후 찾은 객체에 대한 특성을 알기 위해 클래스 분류 모듈에 넣어서 어떠한 속성을 가졌는지 파악하고 그에 따른 심리 진단 결과를 얻게 된다.

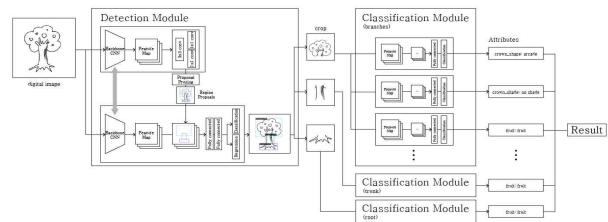


그림 1. 시스템 구조도
Fig. 1. System structure

3-1 객체 검출 모듈

클래스가 잘 분류되기 위해서는 공통으로 가지는 특징이 많아야 한다. 따라서 이미지를 통째로 넣고 세세한 특징을 찾는 것보단 큰 객체를 입력해 분리하고 분리된 곳만의 특징을 찾는 일이 더 쉬울 것이다. 객체 검출 모듈에는 Faster R-CNN[20]을 시킨 딥러닝 모델로 이전 두 모델이 선택적 탐색이라는 기법을 사용해 관심 영역(Region of Interest, ROI)을 생성했다면 Faster-RCNN의 경우 영역 제안 네트워크(Region on Proposal Network, RPN)을 사용하는 방식으로 바뀌었다. 모델은 두 부분으로 구성되는데 앞부분의 RPN은 입력으로 이미지를 받아 합성곱 신경망을 사용해 관심 영역을 추출하고 사후 처리를 거쳐 RoI 리스트를 출력한다. 뒷부분은 분류의 역할을 하며 특징 맵과 RoI를 받아 최종 결과를 낸다. 이 모델은 굉장히 빠른 속도를 가지게 되는데 이는 그림 심리 검사 속도의 단축이라는 목표에 부합한다. Fig 2.는 모델의 구조를 보여준다.

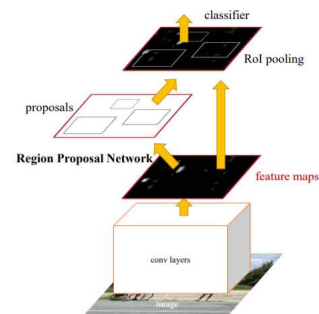


그림 2. Faster R-CNN 구조
Fig. 2. Faster R-CNN Architecture

객체 인식 모듈을 통해 나무 그림 심리 검사의 경우 흰 배경에 나무 한 그루가 그려진 이미지에서 Branches(관), Trunk(기둥), Root(뿌리)를 찾을 수 있고, 고양이 그림 심리 검사의 경우 다양한 그림이 그려진 이미지 한 장에서 Cat(고양이), Head(머리), Body(몸통)를 찾을 수 있다. 객체 단위로 검출을 완료한 뒤, 추출된 각 객체의 위치 정보를 가져와 이 영역만을 잘라내 새로운 이미지 파일을 생성하게 되고 이때 생성된 이미지 파일은 두 번째 단계인 클래스 분류에 쓰이게 된다.

3-2 클래스 분류 모듈

이전 모듈에서 객체 이미지가 추출되어 잘리면, 잘린 이미지를 다시 분류하기 위한 모듈이 필요하다. 본 연구에서는 모듈에 사용되는 13개의 계층으로 구성된 순차적인 구조의 합성곱 신경망을 만들었다. 합성곱 계층과 최대-풀링 계층이 번갈아가며 나오며 활성화 함수로 ReLU[24]를 사용한다. 잘린 이미지는 224x224의 해상도로 변환되며 3개의 색 채널을 그대로 합성곱 신경망의 입력으로 쓴다. 커널의 크기는 3x3이며 계층의 깊이가 깊어짐에 따라 특징 맵의 깊이는 늘어난다. 마지막으로 세 개의 밀집 계층으로 구성되며 softmax 함수가 적용된다. Fig. 3은 위 설명한 합성곱 신경망을 보여주며 conv와 max-pool, dense는 각각 합성곱 계층, 최대 풀링 계층, 밀집 계층이며, n은 특징 맵의 수, k는 커널의 크기를 나타낸다.

한편 딥러닝으로 분석하기 힘든 클래스는 일반적인 영상처리 기법을 이용하기도 하였다. 우선 나무 그림에서는 그림의 밀도를 파악하기 위해 사진 가장자리 20% 속에 있는 픽셀값이 255가 아닌 픽셀의 수를 세었다. 그 수가 일정 수준을 넘어가면 그림의 밀도가 높은 것으로 하고 그렇지 않다면 그림의 밀도가 낮은 것으로 판단하였다.

고양이 그림에서는 그림이 얼마나 화려한가를 판단하기 위한 colorfulness를 판단하였는데 [25]의 방법을 따랐다. 원저 논문에서는 화려함의 기준을 세우기 위해 20명의 비전문가 참가자들이 1~7등급의 화려함으로 84개의 이미지 세트를 평가해 달라고 요청했다. 그 후 일련의 실험적인 계산을 통해, 이미지 색공간과 앞의 결과가 상응하는 수식 (1)을 모델링 하였다.

$$\begin{aligned}
 rg &= R - G & (1) \\
 yb &= \frac{1}{2}(R + G) - B \\
 \sigma_{rgyb} &= \sqrt{\sigma_{rg}^2 + \sigma_{yg}^2} \\
 \mu_{rgyb} &= \sqrt{\mu_{rg}^2 + \mu_{yg}^2} \\
 C &= \sigma_{rgyb} + 0.3 * \mu_{rgyb}
 \end{aligned}$$

먼저 이미지의 RGB채널을 나눈 후 R과 G를 더한 rg와 R과 G를 더한 후 B를 뺀 yb 값을 구한다. 그 다음 각각 표준편차와 평균값을 구하고 표준편차는 표준편차끼리, 평균값은 평균값끼리 더해 나온 값들을 더하여 최종값(C)을 계산한다. 본 논문에서는 화려한 정도를 두 단계로 나누었는데 이는 고양이 데이

터의 경우 두 단계 분류로도 충분히 분석 가능하기 때문이다.

Table 1에서 클래스와 서브 클래스를 볼 수 있는데 합성곱 신경망은 클래스별로 학습되고 입력 이미지를 이 신경망 모델을 통해 분류하였을 때 나오는 결과가 서브 클래스가 된다. 그러므로 전체 모듈은 여러 개의 합성곱 신경망 모델들로 구성되었다고 할 수 있다.

한편 딥러닝으로 분석하기 힘든 클래스는 일반적인 영상처리 기법을 이용하기도 하였다. 우선 나무 그림에서는 그림의 밀도를 파악하기 위해 사진 가장자리 20% 속에 있는 픽셀값이 255가 아닌 픽셀의 수를 세었다. 그 수가 일정 수준을 넘어가면 그림의 밀도가 높은 것으로 하고 그렇지 않다면 그림의 밀도가 낮은 것으로 판단하였다. 고양이 그림에서는 그림이 얼마나 화려한가를 판단하기 위해 색의 사용 빈도수를 세었다.

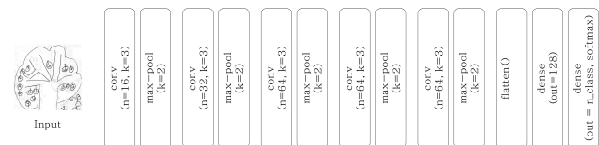


그림 3. 분류를 위한 합성곱 신경망
Fig. 3. Convolution Neural Network for classification

3-3 해석 모듈

딥러닝을 통해 그림에 대한 정보를 얻으면 이를 토대로 해석을 하여 결과를 검사자에게 보여줘야 한다. 현재는 미리 클래스에 대한 해석을 몇 개의 단어로 저장한 뒤 새로운 입력 이미지의 분류 클래스와 비교한 후 해당하는 단어를 가져와 조사와 합쳐 출력하는 식으로 결과를 볼 수 있다. 지금은 단어 끼우기 형식으로 결과가 출력되나 후에 자연어 처리 기술을 적용한다면 사람이 상담 결과를 말해주는 것과 매우 유사한 수준으로 발전할 것이다. 두 개의 그림 심리 검사에서 찾을 수 있는 클래스와 이에 대한 해석은 Table 1, 2에서 볼 수 있다.

IV. 실험 결과

4-1 개발 환경

본 연구는 Ubuntu 18.04.4 LTS에서 Intel Core i7-2600K CPU 3.40GHz와 GeForce GTX 1070 GPU를 기반으로 진행되었다. 객체 검출 모듈은 Faster R-CNN 모델을 불러와 Detectron2 0.4버전을 사용해 그림 데이터를 이용해 재학습시켰다. 클래스 분류 모듈의 경우 Tensorflow2 2.3.2버전과 Keras 2.4.0버전을 사용해 합성곱 신경망을 만들어 학습한 것을 기반으로 한다.

4-2 데이터 수집 및 레이블링

표 1. 나무 그림 클래스와 의미

Table 1. Class and meaning for tree drawing

class	sub_class	form of drawing	analysis
crown_shape	branches	consists of branches	-
	crown_arcade	arcade-shaped branches	sensitivity to form, courtesy, duty-oriented.
	crown_ball	ball-shaped branches	customary, popular, lack of sense of composition, sympathetic to the crowd, daydreaming, bluffing, empty-hearted, blustering, lack of energy.
crown_shade	no_shade	Shaded or unshaded branches	-
	shade	Shaded or shaded	sensitivity to form, courtesy, duty-oriented.
trunk_shape	trunk_base	the base of a broad trunk	containment, learning difficulties, slow understanding, containment thinking, delay in development.
	straight	a straight, parallel trunk	normative, accurate, strong, stubborn, cold, lifeless
trunk_wave	no_wave	a non-curling trunk	-
	wave	a wave-like trunk	vital, adaptable.
trunk_lines	lines	scattered lines	sensitivity, strong empathy, loss of personality, lack of a clear sense of territory (me/you, self/object)
	no_lines	non-scattered lines	-
trunk_shade	full_shade	a whole shady trunk	-
	left_shade	a shaded trunk on the left	extrovert tendencies, sensitivity, inhibition, pernicious.
	right_shade	a shaded trunk on the right	willing to adapt, capable of contacting
	no_shade	an unshaded trunk	-
trunk_tilt	trunk_tilt_left	a tree leaning to the left	defensive attitude, restraint, oppression, suppression of emotions, obsession with the past, challenging.
	trunk_tilt_right	a tree leaning to the right	concentration, sensitivity, easily influenced, seductive, love of things.
	no_tilt	an uninclined tree	-
trunk_pattern	no_pattern	an unpatterned bark	-
	trunk_round	round bark pattern	ability to prepare for contact, willing to adapt
	trunk_scratch	scratched or scaly bark	angry, vulnerable, sensitive, stubborn, observant.
low_branch	low_branch	a branch at the bottom	partial developmental inhibition, unreliable, secret-spreading, behavioral like a child.
	no_branch	no branches under trunk	-
fruit	fix	fruit on branches	visual talent, ability to observe
	no_fruit	No fruit	-
cut_branch	cut_branch	a branch that looks like a saw cut.	controlled tendency, resistance, willingness to live.
	pipe_branch	branch or stem in pipe form	attractions to the unknown, pioneering types, the variety of interests that arise from candid reality and unclear goals.
	no_cut_branch	no cut branches	-

표 2. 고양이 그림 클래스와 의미

Table 2. Class and meaning for cat drawing

class	sub_class	form of drawing	analysis
concept	conceptual	colorful	right brain type
	no_conceptual	not colorful	left brain type
movement	dynamic	dynamic cat appearance: a cat with a different angle between legs, showing changes such as bending or bending legs.	right brain type
	static	a cat with legs that stretch straight up and down without showing any changes such as bending or bending.	left brain type



그림 4. 분석용 데이터셋 예시: (a) 나무 그림, (b) 고양이 그림
Fig. 4. Examples of datasets: (a) Tree drawing, (b) Cat drawing

분류 학습 시 각 데이터 세트에 대해 20 Epoch가 반복되었고 학습률은 0.001로 고정했으며 최적화 알고리즘은 RMSprop을 사용했다. 고양이 그림 심리 검사의 경우 색상의 쓰임 정도를 알아야 하는 부분이 있는데 이는 OpenCV 4.5.2 버전을 사용하여 판단하였다.

본 논문의 실험에 쓰인 데이터는 실제 아동이 그린 그림을 스캔한 이미지와 google-image-download 라이브러리를 사용해 구글에서 크롤링한 선화, 클립아트 등 다양한 유형의 이미지로 구성되며 총개수는 나무 그림 학습용 1,946장, 평가용 159장, 고양이 그림 학습용 492장, 평가용 61장이다. Fig. 4에 나무 그림 데이터셋과 고양이 그림 데이터셋을 나타내었다.

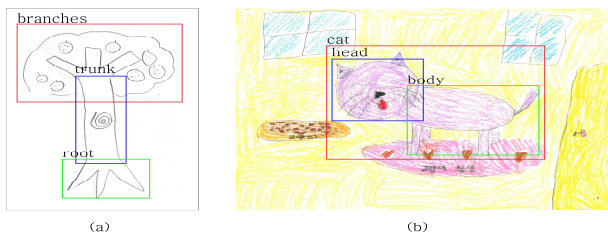


그림 5. 레이블링된 데이터셋 예시: (a) 나무 그림, (b) 고양이 그림
Fig. 5. Example of labeled data: (a) labeled tree drawing, (b) labeled cat drawing

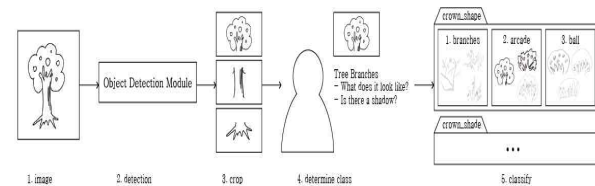


그림 6. 클래스에 따른 데이터셋 구축 과정
Fig. 6. Process of creating class-specific dataset

이미지를 학습시키기 전 데이터에 레이블링 작업이 필요한데 이는 이미지 데이터 속에 찾고자 하는 객체의 위치와 종류를 지정하기 위함이다. 파이썬 이미지 레이블링 툴인 LabelImg를 사용해 경계 상자를 그리고 레이블을 붙여 PascalVOC 형식으로 저장한 뒤 다시 json파일로 변환하였다. Fig. 5는 레이블링 예시이다.

클래스 분류 모델을 학습시키기 위해서는 그림 전체 이미지가 아닌 클래스별 특징을 보여줄 수 있는 이미지가 필요하다. 이를 위해 나무 그림 데이터 전체를 객체 검출 모듈을 적용해 객체의 영역을 잘라내 새로운 이미지들을 만들었다. 잘린 이미지는 사람이 직접 보고 판단하여 어떤 클래스와 서브 클래스에 해당하는지 분류해 서브 클래스 데이터셋에 추가한다. Fig. 6 나무 그림에서 Branch가 검출된 후 세부 클래스로 분류하는 과정을 보여준다.

4-3 학습 및 검증 결과

객체 인식에서는 성능 지표로 재현율(recall)을 사용했다. 재현율이란 실제 True인 것 중에서 모델이 True라고 예측한 비율로 대상 물체를 얼마나 잘 검출하는지 나타낸다. Precision은 positive 예측값으로 모든 positive 예측 중 True의 비율을 계산한 것이다. 그러므로 precision이 높으면 positive 예측 중 정답을 positive라고 예측한 비율이 높다. Table 3은 심리 검사별 AP를 보여준다. 이는 validation 데이터셋으로 측정된 값으로 학습에 사용하지 않았고 오로지 평가에만 사용하였으며 이를 (주지영커뮤니티 전문가의 분석을 토대로 정확도를 측정하였다.

표 3. 객체 검출 평균 정밀도(IoU=0.50)

Table 3. Average Precision (AP) of object detection (IoU=0.50)

Drawing Test	AP
Tree Drawing Test	0.934
Cat Drawing Test	0.856

표 4. 나무 그림 분류 손실값과 정확도

Table 4. Classification loss and accuracy for Tree drawing

Name	Total Loss	Accuracy
crown_shape	0.0163	0.9966
trunk_wave	0.0606	0.9835
trunk_tilt	0.0426	0.9827
fruit	0.5184	0.7509
crown_shade	0.5531	0.7490
trunk_lines	0.0739	0.9812
trunk_pattern	0.0486	0.9906
cut_branch	0.0529	0.9847
trunk_shape	0.0767	0.9688
trunk_shade	0.0265	0.9948
low_branch	0.1442	0.9437

표 5. 고양이 그림 분류 손실값과 정확도

Table 5. Classification loss and accuracy for Cat drawing

Name	Total Loss	Accuracy (%)
concept	0.1119	95.89
movement	0.0821	96.19

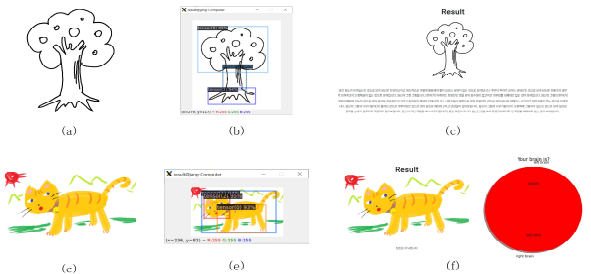


그림 7. (a) 입력 나무 그림, (b) 가지, 줄기, 뿌리에 대한 영역 상자, (c) 나무 그림 테스트 결과 (d) 입력 고양이 그림, (e) 고양이, 머리, 몸통에 대한 영역 상자, (f) 고양이 그림 테스트

Fig. 7. (a) Input tree drawing, (b) Bounding box for branches, trunk, root, (c) Tree drawing test result, (d) Input cat drawing, (e) Bounding box for cat, head, body (f) Cat drawing test result

클래스 분류에서는 손실 값(loss)와 정확도(accuracy)를 측정하는데 각 심리 검사 그림에 대한 개별 성능은 다음 Table 4, 5와 같다. Table 4에서 fruit과 crown_shade 부분이 정확도가 다소 낮은 경향을 보이는 것은 fruit은 과일 부분이 정보가 명확하지 않아 학습이 어려운 부분이 있는 것으로 보이며 crown_shade 단순 검은색으로 표현되는 shade 정보가 학습정보로는 충분하지 않아서 나타난 결과로 해석할 수 있다. 나머지 속성 식별 결과들은 최대 99.6%까지로 매우 높은 것을 볼 수 있다.

Table 5는 고양이 그림에 대한 실험 결과를 보여주고 있다. 기본적으로 속성 클래스 수가 많지 않기 때문에 검출 및 분류를 통한 속성 식별 결과가 96% 정도를 보여주었으며 이를 실제 상용 서비스로 고려할 때 상당히 고무적인 결과로 판단된다.

실제 그림을 입력으로 해 모듈을 테스트해본 결과는 Fig 7과 같다. 입력 이미지는 모델이 이전에 학습한 적 없는 이미지를 사용했고 나무 그림의 경우 branches, trunk, root를 각각 객체 검출 정확도 99%, 96%, 94%로 찾았으며 고양이 그림은 cat, head, body를 각각 99%, 96%, 94%의 정확도로 검출하였다.

나무 그림 심리 검사는 클래스별 해당하는 분석 내용이 줄 글로 출력된다(Fig. 7(e)). 나무 그림에서 branch는 모양이 구름 모양인데 이에 해당하는 클래스는 crown_arcade로, 형태에 대한 감수성이 있고 예의가 바르며 의무감이 있는 사람이 그렸다고 해석할 수 있다. 과일이 달린 것을 인식했기 때문에 fruit 클래스가 해당하고 이는 시각적 재능과 관찰력이 있다고 할 수 있다. 나무의 줄기 부분은 straight하며 scratch 무늬가 있으므로, 그린 이가 규범적이고 고집이 세며 냉정한 면이 있다는 결과가 나온다.

고양이 그림 심리 검사는 성향에 대해서만 파이 그래프로 나타낸다(Fig. 7(f)). 그림은 색이 많이 사용된 편이라 colorful하다 할 수 있으며 다리 동적으로 그려졌기 때문에 이에 대한 평가로 그린 이가 우뇌형이라 할 수 있다.

V. 결 론

본 논문에서는 딥러닝 알고리즘을 통해 손 그림 이미지 데이터로부터 찾고자 하는 요소를 인식하고 분류하는 것이 가능함을 확인하고, 이를 통해 아동의 그림을 분석하고 해석하는 과정의 자동화가 가능함을 보여주었다. 객체 인식 모듈은 Faster R-CNN 모델을 백본으로 재학습 시켜 만들었고 추출된 개별 객체를 통해 클래스 분류 모듈은 간단한 합성곱 신경망 통해 만들었다. 최종 분류 결과는 나무 그림의 경우 그 정확도가 평균 94%, 고양이 그림의 경우 95% 이상의 높은 수치를 보여주었다.

본 연구에서는 나무와 고양이 그림의 많은 특성 중 인식을 위한 데이터가 충분한 특성들을 중심으로 학습하고 분석을 진행하였다. 따라서 심리 진단에 필요한 모든 특성에 대한 분석 및 평가는 이루어지지 않았다. 그러므로 추후 연구되지 않은 요소들에 대한 추가적인 분석이 필요하고 새로운 분석을 토대로 더 일반화된 검사 결과를 도출해 내야 할 것이다.

감사의 글

This research was supported by Seoul R&BD Program (CT200009, Seoul Campus Town Technology R&D Project-AI-based Psychological Analysis Algorithm Using Image Recognition).

참고문헌

- [1] D. Edwards. Art Therapy, *SAGE Pub.*, ch. 2, pp. 1-16, 2014.
- [2] J. Park, S. Shin, J. Kim, K. Park, S. Lee, M. Jeon and S. Kim, "Preliminary Research of HTP Sentiment Analysis Automation on Children's Drawings," in *Proceedings of the HCI Society of Korea*, pp. 867-871, February 2019.
- [3] K. Koch, "Der Baumtest: Der Baumzeichen-versuch als psychodiagnostisches Hilfsmittel" *Bern Stuttgart Huber*, 1957
- [4] Karl Koch's Tree Testing: What A Drawing Can Reveal About Your Emotions [Internet]. Available: <https://themindsjournal.com/karl-kochs-tree-testing/>.
- [5] M. S. Maserati, C. Maticena, L. Sambati, F. Oppi, R. Poda, M. D. Matteis and R. Gallassi, "The Tree-Drawing Test (Koch's Baum test): a Useful Aid to Diagnose Cognitive Impairment." in *Behavioural neurology*, June 2015. <https://doi.org/10.1155/2015/534681>
- [6] W. Baek, Establishing Interpretation System and Improving

- Validity of HTP (House-Tree-Person) Test, Ph.D. Samyook University, 2019.
- [7] M. Turk and A. Pentland, "Eigenfaces for Recognition." *Journal of cognitive neuroscience*, Vol. 3, No. 1, pp. 71-86, January 1991.
<https://doi.org/10.1162/jocn.1991.3.1.71>
- [8] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, November 2004.
<https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [9] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, pp. 1097-1105, January 2012. <https://doi.org/10.1145/3065386>
- [10] D. Jeong, B. Kim and S. Dong, "Deep Joint Spatiotemporal Network for Efficient Facial Expression Recognition," *Sensors*, Vol. 2020, No. 20, pp. 1963, March 2020.
<https://doi.org/10.3390/s20071936>
- [11] J. Kim, B. Kim, P.P. Roy and D. Jeong, "Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure" in *Proceedings of the Institute of Electrical and Electronics Engineers*, Vol. 7, No. -, pp. 41273-41285, March 2019.
<https://doi.org/10.1109/ACCESS.2019.2907327>
- [12] Y. Choi, Y. Lee and B. Kim, "Wavelet Attention Embedding Networks for Video Super-Resolution," in *Proceedings of IEEE International Conference on Pattern Recognition*, pp. 7314-7320, January 2021.
<https://doi.org/10.1109/ICPR48806.2021.9412623>
- [13] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," in *Proceedings of the Institute of Electrical and Electronics Engineers*, Vol. 86, No. 11, pp. 2278-2324, November 1998.
<https://doi.org/10.1109/5.726791>
- [14] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," <https://arxiv.org/abs/1409.1556>, September 2014.
- [15] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the Institute of Electrical and Electronics Engineers*, pp. 770-778, June 2016.
<https://doi.org/10.1109/CVPR.2016.90>
- [16] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection." in *Proceedings of the Institute of Electrical and Electronics Engineers Conference on Computer Vision and Pattern Recognition*, pp. 770-788, June 2016.
<https://doi.org/10.1109/CVPR.2016.91>
- [17] W. Liu, D. Anguelov, D. Erhan. C. Szegedy, S. Reed, C. Fu and A. C. Berg, "SSD: Single Shot Multibox Detector." in *European Conference on Computer Vision*, pp. 21-37, September 2016.
https://doi.org/10.1007/978-3-319-46448-0_2
- [18] T. Lin, P. Goyal, R. Girshick, K. He and P. Dollar, "Focal Loss for Dense Object Detection," in *Proceedings of the Institute of Electrical and Electronics Engineers International Conference on Computer Vision*, pp. 2980-2988, October 2017.
<https://doi.org/10.1109/ICCV.2017.324>
- [19] R. Girshick, "Fast R-CNN," in *Proceedings of the Institute of Electrical and Electronics Engineers International Conference on Computer Vision*, pp. 1440-1448, February 2015. <https://doi.org/10.1109/ICCV.2015.169>
- [20] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Advances in neural information processing systems*, Vol.39, No. 6, pp. 91-99, June 2015.
<https://doi.org/10.1109/TPAMI.2016.2577031>
- [21] K. He, G. Gkioxari, P. Dollar and R. Girshick, "Mask R-CNN." in *Proceedings of the Institute of Electrical and Electronics Engineers International Conference on Computer Vision*, pp. 2961-2969, December 2017.
<https://doi.org/10.1109/ICCV.2017.322>
- [22] I. Jeon, S. Kim, K. Kim, E. Lim and Y. Kim, "Computerized Psychological Test System for Art Therapy." *Journal of Korean Institute of Information Scientists and Engineers: Computing Practices and Letters*, Vol. 16, No. 2, pp. 135-143, 2010.
- [23] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in *Proceedings of the Institute of Electrical and Electronics Engineers Conference on Computer Vision and Pattern Recognition*, pp. 580-587, June 2014. <https://doi.org/10.1109/CVPR.2014.81>
- [24] V. Nair and G. E. Hinton. "Rectified Linear Units Improve Restricted Boltzmann Machines," *International Conference on Machine Learning*, June 2010.
<https://doi.org/10.5555/3104322.3104425>
- [25] D. Hasler and S.E. Suesstrunk, "Measuring Colorfulness in Natural Images." *International Society for Optics and Photonics*, Vol. 5007, pp. 87-95, June 2003.
<https://doi.org/10.1117/12.477378>



장정인(Jung-In Jang)

1921 년 : 숙명여자대학교 IT공학과

2016년~2021년: 숙명여자대학교 IT공학과 공학사

2021년~현 재: 숙명여자대학교 IT공학과 석사과정

※관심분야: 딥러닝(Deep Learning), 초해상도 이미지 처리 (Image Super-resolution)



김병규(Byung-Gyu Kim)

1996 년 : 부산대학교 전기공학과 공학사

1998 년 : 한국과학기술원(KAIST) 전기 및 전자공학 석사

2004 년 : 한국과학기술원(KAIST) 전기 및 전자공학 박사

2004년~2008년: 한국전자통신연구원 선임연구원

2009년~2015년: 선문대학교 컴퓨터공학과 부교수

2016년~2018년: 숙명여자대학교 IT공학과 부교수

2019년~현 재: 숙명여자대학교 IT공학과 교수

2018년~현 재: Journal of Multimedia Information System 편집위원장 및 한국멀티미디어학회 부회장

※관심분야: 비디오 및 영상 신호처리, 비디오 압축이론, 지능 시스템, 딥러닝 기반 인식 알고리즘 등



김영나(Young-Nah Kim)

1986 년 : 이화여자대학교 서양학과 미술학사

2000 년 : 이화여자대학교 교육대학원 미술교육 석사

2003년~2016년: 명지전문대 커뮤니케이션디자인 “관찰과 표현” 강사

2011년~2016년: 서울교육대학교 평생교육원 “우뇌로 그림 그리기” 강사

2011년~2015년: 우뇌개발미술원격평생교육원 “우뇌개발 미술교육” 강사

2008년~현 재: 아트샘미술교습소 대표

2018년~현 재: (주)지영커뮤니티 대표이사

※관심분야: 우뇌로 그림 그리기, 그림을 통한 심리 진단, 미술교육을 통한 심리 치료 등



박용배(Young-Bae Park)

1987년 : 숭실대학교 법학과 법학사

1989년~2003년: (수)신홍정밀 전산실

2004년~2007년: 한국IBM (BAS/400 프로젝트)

2008년~2011년: (주)신홍산업 전산실

2012년~2015년: (주)엡디스 연구실 이사(Bus Information System 개발)

2019년~현 재: (주)지영커뮤니티 AI응용연구팀 이사

※관심분야: 딥러닝 기반 이미지 인식, Big Data를 활용한 제조현장 검사 및 자동화 등