

기계학습 모델 기반의 사진으로부터 감성 분석에 관한 연구

이 나 혁¹ · 이 태 민^{2*}

¹중앙대학교 소프트웨어대학 소프트웨어학부 학사과정

^{2*}중앙대학교 다빈치SW교육원 특임교수

The Study for Emotion Analysis from Images depend on Machine Learning Models

Nahyuk Lee¹ · Taemin Lee^{2*}

¹Bachelor's Course, Computer Science & Engineering, Chung-Ang University, Seoul 06974, Korea

^{2*}Special Affair Professor, Davinci SoftWare Institute, Chung-Ang University, Seoul 06974, Korea

[요 약]

인간은 일반적으로 시각 의존적인 동물이다. 눈으로 들어오는 정보들이 인간에게 가장 많은 영향을 미친다. 인터넷의 발달을 통해서 단순한 경험이 아닌 온라인을 통해서 더 많은 이미지들에 쉽게 접근할 수 있게 되었다. 주어진 이미지들이 사용자들에게 어떤 감성을 줄 수 있는지를 예측해둔다면, 사용자들이 이미지를 검색하거나 분류할 때 조금 더 도움이 될 수 있다. 기계학습의 발전에 따라서 콘텐츠에서 감성을 예측하는 연구들이 많아졌지만, 기계학습 모델중 어떤 것이 감성 예측에 효과적인지에 대한 분석은 미비하였다. 본 연구는 시각적 정보인, 이미지들의 색상정보를 기반으로 감성을 예측해내는 연구를 진행한다. 네 가지의 대표적 감성을 분류하고 주어진 이미지들을 네 가지의 감성 중 하나로 크롤링한다. 수집된 데이터를 기준으로 다양한 기계학습을 통해서, 각 모델 중에 가장 어울리는 모델을 찾고, 각각의 기계학습 모델의 결과를 분석한다. 이를 통해 최종적으로 색채를 기반으로 이미지를 분석하는 것이 타당한지에 대해서 정리한다.

[Abstract]

Humans are generally visually dependent animals. Information coming into the eye affects humans the most. Through the development of the Internet, more images are easily accessible through online rather than just experience. Predicting what sensitivity a given image can give users, it can be a little more helpful when users search or categorize images. While more studies have been conducted to predict sensitivity in content as machine learning advances, analysis of which of the machine learning models is effective in predicting emotions has been insufficient. This study is conducted to predict sensitivity based on color information of images, which is visual information. Categorize the 4 representative emotions and crawl the images into one of the four emotions. At this time, through various machine learning, we find the model that best suits each model, and analyze the results of each machine learning model. This finally summarizes whether it is reasonable to analyze images based on color.

색인어 : 기계학습, 감성 컴퓨팅, 이미지 분석, 색상 분석, 데이터 크롤링

Key word : Machine Learning, Affective Computing, Image Analysis, Color Analysis, Data Crawling

<http://dx.doi.org/10.9728/dcs.2021.22.8.1295>



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 16 July 2021; Revised 30 July 2021

Accepted 13 August 2021

*Corresponding Author; Taemin Lee

Tel: [REDACTED]

E-mail: kevinlee@cglab.cau.ac.kr

I. 서론

인간은 시각적 자극을 통해서 들어오는 정보에 많은 영향을 받는다. 눈으로 들어오는 정보들이 인간의 감정에 많은 영향을 미치게 된다는 뜻이다. 이전에는 이런 감정적 영역은 인간 고유의 영역으로서 객관적인 분석이 어렵다는 평가가 많았다. 기계가 발전함에 따라서, 많은 정보를 기반으로 분석하게 되는 기계 학습, 인공지능 등이 발전함에 따라서, 이런 인간의 고유 영역도 컴퓨터를 통해서 분석이 가능해지기 시작하였다. 수많은 정보를 기반으로 감성을 분석한다면 분석의 타당성도 올라갈 수 있게 되었다.

일반적으로 눈으로 들어오는 정보가 가장 많은 만큼 인간의 감성에 영향을 미치는 가장 큰 요인 중 하나는 바로 시각적 자극일 수밖에 없다. 우리는 본 연구에서 이 시각적 정보들을 기반으로 감성을 분석하는 연구를 제안한다. 시각의 대표적 요소라고 할 수 있는 색채를 기반으로 주어진 사진 이미지에서 느껴지는 감성을 컴퓨터로 분석하고자 한다.

본 연구에서는 특정 감정에 따른 색채 데이터를 수집하여 학습을 시키고, 감정을 예측 분석하는 기계학습 모델을 구축하고자 한다. 보편적으로 사용하는 구글 검색 엔진을 통해서 이미지들을 수집한다. 본 연구에서 분류한 감정 키워드를 기반으로 이미지들을 4가지로 검색하여 수집한다. 수집된 이미지들을 기반으로 색상 스펙트럼을 추출하여 모델에 학습시킨 뒤, 감정을 예측하는 모델을 구축한다. 다양한 기계학습 모델을 통해서 색채 기반의 감정 모델들을 구성하고 각각의 모델들을 비교 분석한다. 이것을 기반으로 감정예측이 실제와 유사성이 있는지, 검증한다. 최종적으로, 시각적 정보와 감정과의 상관관계를 비교하여 예측된 감정의 타당성에 관해서 확인한다.

본 논문의 구성은 다음과 같다. 2장에서는 감성을 분석하는 연구들을 논한다. 기계학습을 통해 혹은 딥러닝을 통해서 감성을 분석하는 연구들을 정리하고 우리의 연구와의 차이점을 서술한다. 3장에서는 우리가 이미지에서 색상정보들을 추출하는 방법과 기계학습 모델을 제작하는 과정에 관해서 서술한다. 4장에서는 기계학습 모델을 통해서 나온 결과들을 서술하고, 그에 대해 분석한다. 5장에서는 마지막으로 결과들을 기반으로 우리의 연구에 대한 결론 및 향후 연구에 관해서 서술한다.

II. 관련연구

Chambers & Moulton(1978)은 인간의 생리적 반응 현상과 더불어 여러 가지 사고, 판단 작용과 감정, 심리상의 반응을 색상이 일으키는 역할을 할 수 있다고 하였다[1]. 색채가 인간의 인지 과정에서 심리 정서적으로 영향이 높으므로, 색채와 감정 간의 연구는 활발히 이루어지고 있다. [2]는 색채가 지니는 특성과 효과를 알아보고, 색채 심리의 실질적인 활용 방안에 대해서 모색하였다. “색”이 모든 사물의 고유한 시각적 특성으로 빛의

반사에 의해 이루어진다면, “색채”는 대비와 조화 그리고 동화를 통해 다양한 심리적인 반응을 이끌어내는 심리적 현상이라고 분석하였다. 이 외에도 [3], [4]와 같이 통계 분석적으로 색채에 따라 인간이 보편적으로 가지는 감정 및 느낌에 관해 기술하고, [5]와 같이 여러 가지 색상조합을 통해서도 정서를 유추해내는 연구들도 진행되었다. 이렇게 색상과 감정 간의 관계가 연구되면서, 이를 통해서 실생활에서 유용하게 사용되는 연구들도 존재한다. [6]은 색채와 광고 간을 분석한 연구로서, 색채가 인간의 마음의 상태와 인지 능력에 영향을 미칠 수 있다고 서술한다. 이를 통해서 색채를 통해서 고객의 관심을 끌고, 의사결정에 영향을 주고, 브랜드와 상품 수준을 전달하는 커뮤니케이션 역할이 가능하다고 하였다.

사진을 이용한 감정예측 연구들은 다양하게 존재한다 [7-11]. 이 연구들은 사진에서 색상정보 또는 간단한 물리적 정보들을 추출하여, 기본적인 분류법이나 기계학습을 이용하여 감성을 추출한다. 가장 대표적인 기계학습법인 선형 회귀(Linear Regression)를 이용한 분류법을 주로 사용한다. 감정이란 복잡한 정보를 추출하기 위해서 간단한 기계학습법을 이용한다면 오버샘플링(over-sampling)과 같은 문제점이 생길 수 있으므로 감정 모델에 어울리는 분석법이라고 할 수 없다. 이외에도 [12]과 같이 비디오에서 감정을 추출하기 위한 연구도 진행되었다. 뮤직비디오 제작을 위해, 음악과 비디오를 각각 분석하여 감정을 바탕으로 매칭시켰다. 이때, 영상을 스틸샷으로 쪼개서 각각의 감정을 추출하여 그림 이미지의 A.V(Arousal-Valence) 값을 계산하였다. 사진에서 도메인을 더 축소하여, 그림에서의 감정을 추출하는 연구들도 진행되었다. [13]의 연구는 lab 색상 값과 Scale Evaluant Feature Transform을 이용하여 기계학습을 통해 그림에서 감정을 추출하였다. 이미지에서 직접적으로 얻을 수 있는 물리적 특징들을 이용해서 감정을 예측해내었고, 기존의 사진 분석 연구에서 도메인을 감정과 밀접하게 연결되어있는 그림으로 축소를 분석하였다는 특징이 있다. [14]의 연구는 그림의 대표적인 색채 조합을 통해 감정 형용사를 추출하였다. 감정 형용사를 러셀의 Arousal-Valence 모델에 매핑해서 다른 감정 연구에도 쉽게 적용할 수 있도록 하였다. [15]와 [16] 역시 감정예측에 예술성을 사용하였지만 사용된 것들이 모두 색과 관련된 요소들이었다.

다양한 색상을 이용하여 사진들 혹은 그림들을 분석해내는 연구들이 많이 진행되었지만, 이 연구들은 모두 간단한 기계학습 모델들을 이용하여 감정예측 모델을 제작하였다. 사용되는 특징들도 다양하게 사용되었지만 모두 시각적 정보, 특히 색에 관련된 정보들이 대부분이었다. 우리의 연구는 기존의 연구들 처럼 색상을 기반으로 기계학습 모델을 제작하지만, 단순 색 정보가 아닌 색상의 분포를 뜻하는 스펙트럼을 이용한다. 또한, 단순한 기계학습 모델인 선형 회귀분석이 아닌, 다양한 분석 모델을 이용하여 여러 가지 예측모델을 제작하고 각각을 비교 분석하여, 색상과 감정예측에 대한 연관성을 분석하고 감정예측에 있어서 색상의 중요성에 대해서 분석한다.

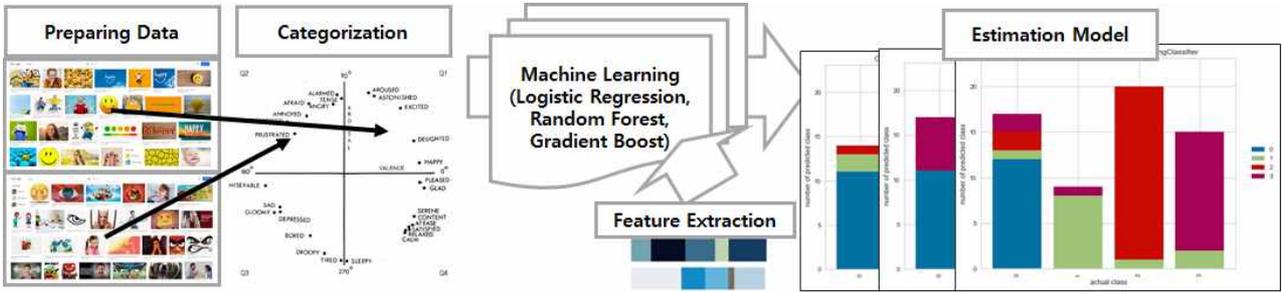


그림 1. 시스템 개요도 (이미지들을 크롤링한 후, 수집된 이미지들을 기반으로 색상정보를 추출하고 이를 토대로 다양한 기계학습 모델을 제작하여 결과를 비교한다.)

Fig. 1. A system overview(After crawling the images, color information is extracted based on the collected images and various machine learning models are created based on them to compare the results.)

III. 본론

본 연구는 이미지에 대해 색채 정보를 추출한 뒤, 이를 특징으로 학습하는 기계학습 모델을 통해 감성 레이블을 예측할 수 있는 분류 모델을 개발하였다. [그림 1]은 우리의 전체 연구 흐름을 도식화한 것이다. 우리의 연구는 다음의 과정으로 진행된다. 첫 번째, 데이터 수집(Preparing Data) 단계에서는 가장 큰 규모의 검색 엔진으로 잘 알려진 구글(Google)에 특정 감성 키워드를 검색하여 나타난 이미지를 크롤링한다. 이렇게 얻은 이미지들은 각각의 검색 키워드 감성으로 레이블링 된 것으로 간주한다. 두 번째 K-평균 군집화(K-mean Clustering)에서는 이미지 내 모든 픽셀의 색상을 군집화함으로써 이미지를 대표할 수 있는 색상 클러스터를 추출한다. 이는 이후에 색상과 관련된 특징을 계산하는 데 사용된다. 세 번째 특징 추출(Feature Extraction)에서는 RGB 색 공간에서의 각각의 색상 값, 그리고 색 조화도와 명도, 채도의 조화도를 계산한다. 이렇게 계산된 특징은 보다 더 나은 학습 결과를 위해 이후 정규화(Normalization) 과정을 거친다. 마지막으로 기계학습 자동화 기법(Automated Machine Learning)을 이용하여 최적의 기계학습 알고리즘을 찾아내고 그에 대한 하이퍼 파라미터 튜닝을 진행한다.

3-1 실험 환경(Experience Environment)

실험은 AMD Ryzen 5 3600 6-Core 프로세서, 16GB RAM, Nvidia GTX 1660 SUPER 6GB VGA 환경에서 진행되었다. 실험을 위한 코드는 Python으로 작성하였으며, 3.6.10 버전의 인터프리터(Interpreter)를 사용하였다. 데이터 수집을 위한 크롤링은 Selenium[16]과 Chrome 웹 브라우저를 사용하였으며, 기계학습에 대한 구현은 Scikit-learn, Pycaret[17], Pytorch 등을 이용하였다.

3-2 감성 키워드 선정(Selecting Emotion Keywords)

앞서 관련 연구에서 확인할 수 있었던 색채와 감성의 연관성을 데이터를 통해 입증하기 위해 분명하게 대조될 수 있는

‘Happy(행복)’, ‘Calm(안정)’, ‘Angry(화남)’, ‘Depress(우울)’의 총 4가지 감성에 대한 키워드를 선정하였다. 실제로 우리가 연구에 활용하기 위해 선정한 4개의 감성 키워드의 위치를 [그림 2]의 Russell AV 감성 모델 4분면[1]에서 찾으면 모두 분명히 대조될 수 있는 위치에 존재한다는 것을 확인할 수 있다.

3-3 데이터 수집(Preparing Data)

사용자 조사(User Survey) 방식을 통해 이미지를 보고 인간이 느끼는 감성이 어떠한지 조사했던 기존 연구와 다르게 우리는 크롤링(Crawling)을 통해 이미지 데이터와 감성 레이블을 획득하였다. 이 방식은 기존 연구와 비교하여 다음의 장점을 가진다. 첫째, 학습을 수행할 이미지 데이터를 구함에 있어 객관성을 가질 수 있다. 크롤링을 통해 감성 키워드 검색 결과의 상단 이미지셋을 얻기 때문에 데이터 선택에 있어 개입이 들어가지 않게 된다.

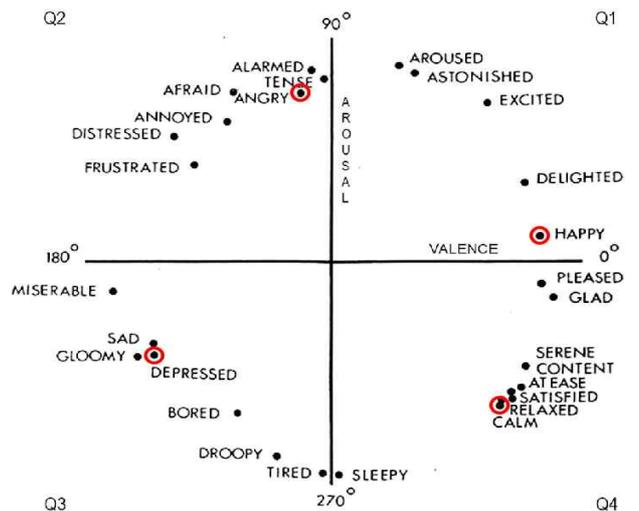


그림 2. Russell의 AV 감성 모델 사분면(빨간 동그라미가 우리가 본 논문에서 사용하는 대표 감성 형용사들이다.)

Fig. 2. Russell's Arousal-Valence emotional coordinate(Red Circles are the representative emotional adjectives we use in this paper.)

둘째, 이미지에 대한 감성의 트렌드(Trend)를 반영할 수 있다. 검색 결과는 특정 감성에 대하여 인터넷 사용자들이 업로드한 이미지를 나타내기 때문이다. 셋째, 기존 방식과 비교하였을 때 데이터 수집에 필요한 비용이 획기적으로 감소된다. 크롤링을 통해 자동화된 방식으로 데이터를 수집할 수 있게 된다.

우리는 다음의 과정을 거쳐 크롤링을 진행하였다. 먼저 감성에 대한 이미지 데이터를 획득하기 위해 가장 큰 규모의 검색 엔진으로 잘 알려진 구글(Google)에 특정 감성 키워드를 검색한다. [그림 3]과 같이 특정 감성 키워드를 검색하면 실제 사진, 회화, 삽화와 같이 다양한 종류의 이미지가 나타남을 확인할 수 있다. 단순 풍경, 회화를 넘어 모든 종류의 이미지 데이터를 다루기 때문에 감성 분류에 대한 정확도는 떨어질 수 있으나, 인간의 모든 감성 범위를 포함한다는 데 큰 의미가 있으므로 별도의 이미지 선별 과정은 거치지 않았다.

우리는 앞서 실험 환경에서 언급하였듯 Chrome 브라우저에 대한 자동 조작을 통해 이미지 데이터를 크롤링하였다. 자동 조작을 위한 라이브러리로 Selenium을 활용하였다. ‘Happy’, ‘Calm’, ‘Angry’, ‘Depress’의 4가지 검색어를 검색한 후 이미지 검색 결과 상위 250개의 이미지를 크롤링하여 총 1,000장의 이미지 데이터셋을 구축하였다.



그림 3. 감성 키워드에 대한 이미지 검색 결과 예시 (위에서부터 happy, angry 키워드 검색)

Fig 3. Image search results example for emotional keyword (up : happy, down : angry)

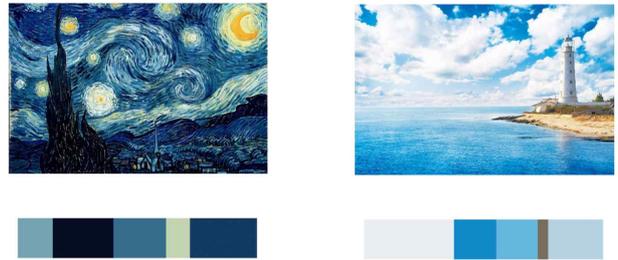


그림 4. K-평균 군집화를 통해 구한 색상 스펙트럼 예시
Fig 4. Color spectrum examples obtained through K-means clustering

3-4 K-평균 군집화를 통한 색상 스펙트럼 추출 (Extracting Color Spectrum using K-Means Clustering)

단일 이미지는 수많은 픽셀로 이루어져 있기 때문에, 모든 픽셀에 대한 색상 값을 개별로 추출하고 분석하는 것은 많은 연산을 필요로 한다. 일반적인 이미지들은 그들이 가지고 있는 주 색상들이 존재하기 때문에, 많은 시간을 사용하여 모든 색상을 분석하더라도 분석된 색상이 주 색상이 아닌 경우 결과에 미치는 영향이 미미하다. 따라서 본 연구에서는 이미지를 구성하고 있는 주 색상들을 군집화하여 색상 스펙트럼을 추출하고 이를 분석에 활용하였다.

군집화에는 비지도 학습 방식의 K-평균 군집화 알고리즘을 활용하였다. 연산 속도가 지나치게 느려지지 않을 정도를 유지하면서도 단일 이미지의 색상정보를 잘 대표할 수 있을 정도의 그룹을 구성하는 것을 목표로 하였다. [그림 4]에서 볼 수 있듯, 우리는 5개의 그룹으로 군집화된 색상 스펙트럼 데이터를 활용하여 이미지를 분석하고 향후 진행될 기계학습 과정에 활용하였다.

3-5 이미지 특징 추출(Extracting Image Feature)

사람이 사진 혹은 그림을 인식할 때 영향을 받는 여러 시각적 요소들이 존재한다. 사진을 찍거나, 그림을 그릴 때 이런 요소들을 기반으로 제작가가 자신의 의도를 담아 작품에 표현한다. 가장 직관적으로 표현할 수 있는 정보로는 색상 정보가 있고, 색의 배치, 조화 혹은 대비 또한 이미지를 접하는 관찰자의 감성에 영향을 끼칠 수 있다. 이를 기계학습에 적용시키기 위해서는 감성을 계산하기 위한 여러 이미지의 특징을 추출해야 한다.

1) RGB 모델을 활용한 색상 정보 계산

K-평균 군집화를 통해 군집화된 5개의 색상 정보와 그 비율을 통해 이미지의 RGB 색상 정보를 판단한다. 평균 R, G, B 값의 경우 아래의 수식을 통해 계산할 수 있다.

2) 유클리디언 거리를 활용한 색 분포(분산)의 계산

RGB 색상 모델은 두 색 사이의 차이를 유클리디언 (Euclidean) 거리를 이용하여 구할 수 있다. 단일 이미지 내에 존재하는 여러 가지 색상이 어떻게 분포하였는지에 대하여 수

식 (1), (2)를 통해 계산할 수 있다. 수식 (1)은 각 군집의 RGB 분산 값 D_n 을 구하기 위한 것으로, 군집 각각의 RGB 3채널 값 (R_n, G_n, B_n) 과 전체 RGB 3채널의 평균 값 $(\bar{R}, \bar{G}, \bar{B})$ 의 차를 제공한다. 이후 수식 (2)를 통해 각 군집에 대해 구해진 분산을 평균 내어 해당 이미지 내 k 개 군집의 RGB 색상 분포(분산) D_v 을 계산하였다. 앞서 언급하였듯, 우리는 군집화를 통해 5개의 색상 정보와 비율을 획득하고 이를 이용하였다.

$$D_n = \sqrt{(R_n - \bar{R})^2 + (G_n - \bar{G})^2 + (B_n - \bar{B})^2} \quad (1)$$

$$D_v = \frac{\sum_{1 \leq n \leq k} (D_n - \bar{D})^2}{k}, k=5 \quad (2)$$

3) RGB 모델을 활용한 색상 정보 계산

색은 색상 명도 채도의 속성으로 구성되며 이 중 명도와 채도를 동시에 고려한 것을 톤(Tone)이라고 말한다. 배색 관계에 의하여 색채의 조화가 나타나게 되는데 이때, 인접한 색과의 관계에 따라 배색 조화 혹은 배색 대비가 일어나게 된다. 이 또한 이미지를 관찰하는 사람의 감정에 영향을 끼치게 되기 때문에 우리는 해당 이미지의 명도와 채도가 이루는 분포에 대한 조화도를 계산하였다. 수식 (3)은 명도(Value)에 대한 계산 식이다. RGB 채널의 분포를 계산할 때와 마찬가지로 각 군집에 대한 명도 V_n 와 이미지 평균 명도 \bar{V} 를 이용하여 k 개 군집 전체의 분포를 계산하였다.

$$V_v = \frac{\sum_{1 \leq n \leq k} (V_n - \bar{V})^2}{k}, k=5 \quad (3)$$

채도(Saturation)에 대한 분석은 수식 (4)를 활용하였으며, 색상 정보, 명도와 마찬가지로 각 군집에 대한 채도 S_n 와 이미지 평균 채도 \bar{S} 를 이용하여 k 개 군집 전체의 분포를 계산하였다.

$$S_v = \frac{\sum_{1 \leq n \leq k} (S_n - \bar{S})^2}{k}, k=5 \quad (4)$$

3-6 표준 정규화(Standard Normalization)

전처리 과정을 통해 얻은 여러 특징들에 대한 정규화를 통해 기계학습 알고리즘 적용 시의 학습 성능을 향상시킬 수 있도록 하였다. [표 1]에서 확인할 수 있듯, 각 피처의 최대 및 최소 범위가 제각각이기 때문에 정규화 과정을 통해 특정 특징이 학습에 지나치게 관여하지 않도록 조절하였다.

다만, 우리는 일반적인 정규화 과정에서 평균과 분산 값을 활용하는 것과 다르게 중간값(Median)과 사분위값(IQR: InterQuartile Range)를 활용하였다. 우리가 사용하는 데이터셋의 경우 검색 엔진 기반 크롤링 데이터이기 때문에 이상치

(Outlier)의 영향을 최소화할 수 있는 정규화를 필요로 하였고, 따라서 수식 (5)를 통해 계산하였다. 수식 (5)는 RobustScaler로 불리는 표준 정규화를 설명하는데, 이 스케일러는 기준값과 중앙값의 편차를 사분위값의 3분위와 1분위 값의 차로 나누어 계산하여 데이터를 스케일링한다.

$$X_{scale} = \frac{x_i - x_{med}}{x_{75} - x_{25}} \quad (5)$$

RobustScaler를 통해 여러 특징은 같은 규모(Scale)를 갖게 되지만 평균 대신 중앙값을 사용함을 통해 극단적인 이상치의 영향을 줄일 수 있도록 하였다.

3-7 기계학습(Machine Learning)

표준 정규화 전처리가 이루어진 각 이미지에 대한 특징들은 0, 1, 2, 3으로 레이블링 되었으며, 각각은 Angry(화남), Calm(안정), Depress(우울), Happy(행복)을 의미한다. AutoML은 머신러닝 자동화 기법으로, 현존하는 많은 기계학습 알고리즘을 자동으로 적용 및 비교하여 최적의 모델을 선택하는데 도움이 되기 위해 개발되었다[18]. 우리는 이 AutoML을 활용하여 감성 분류 성능이 가장 좋은 모델을 찾은 후 하이퍼 파라미터 최적화 과정을 거쳤다. 평가 지표로는 AUC를 활용하여 각각의 감성 Label에 대한 분류 성능을 측정하고자 하였다.

IV. 결과

널리 알려진 기계학습 알고리즘을 통해 감성 분류 모델을 만든 후 그 성능을 AUC 평가 지표를 통해 비교 분석하였다. AUC는 오차행렬의 FP(False Positive)와 TP(True Positive)를 두 개의 축으로 하여 그린 그래프의 ROC 곡선 아래 부분의 넓이를 의미한다. 이진 분류의 성능을 측정할 때 널리 쓰이는 성능 지표 중 하나로, 우리는 각 감성이 나머지 감성들과 비교해 얼마나 구별되는지에 대해 파악하고자 AUC를 사용하였다. 총 학습 데이터의 70%를 Train Set, 나머지 30%를 Test Set으로 하여 10회의 교차 검증(K-Fold Cross Validation)을 거쳤다.

[표 2]는 기계학습의 여러 알고리즘을 AutoML 기법을 활용하여 데이터를 학습하고 평가한 결과이다. 각 모델은 [표 3]에서 간략하게 설명되어 있다. 약한 분류기 모델인 Decision Tree를 제외하면 반복적 데이터 샘플링을 통해 모델을 학습시키는 Bagging 계열, 순차적 가중치 갱신 방식을 이용하는 Boosting 계열의 모델에서 대부분 약 86~87% AUC의 비슷한 성능을 나타내는 것을 확인할 수 있다. 우리는 보편적으로 데이터 분석에 많이 사용되는 Logistic Regression, Random Forest, Gradient Boost 모델을 선택하여 세 분류기에 대한 자세한 분석을 실시하였다.

표 1. 기계학습을 위해 추출된 이미지의 특징 데이터 예시

Table 1. Example of extracted feature data for machine learning

Image	Color Spectrum	R_avg	G_avg	B_avg	Label
		Color_var	V_var	S_var	
		34.8860043	1.57057268	1.81187042	Angry(0)
		5125.9221	15481.1014	0.47699126	
		30.1372226	42.7017917	36.7001137	Calm(1)
		10712.9634	24681.8625	0.33652975	
		24.7386989	24.7313521	25.0412616	Depress(2)
		14798.6729	14906.9441	0.00072891	
		41.7242885	34.3936207	20.7583408	Happy(3)
		8839.32621	26805.3865	0.19867737	

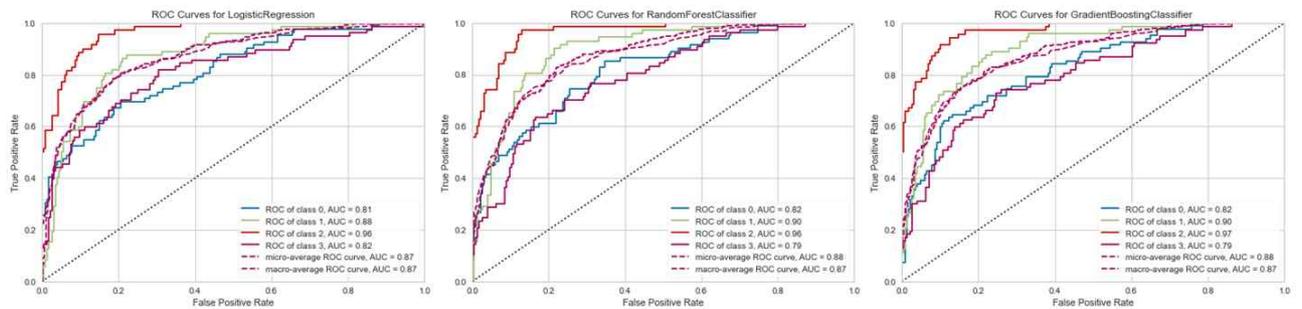


그림 5. 분류기 학습에 대한 ROC Curve (왼쪽부터 Logistic Regression, Random Forest, Gradient Boost 알고리즘을 활용한 모델의 그래프이며, Label 0, 1, 2, 3은 각각 Angry, Calm, Depress, Happy를 의미함.)

Fig 5. ROC Curve for classifier learning(from left to right is a graph of models utilizing the Logic Regression, Random Forest, and Gradient Boost algorithms, with Label 0, 1, 2 and 3 representing Angry, Calm, Depress, and Happy, respectively)

표 2. AutoML을 통한 기계학습 알고리즘 적합도 평가

Table 2. Evaluation of suitability of machine learning algorithms with AutoML

Algorithm	Accuracy	AUC
Logistic Regression	0.6700	0.8701
LDA(Linear Discriminant Analysis)	0.6685	0.8696
QDA(Quadratic Discriminata Analysis)	0.6626	0.8694
Decision Tree	0.5571	0.7042
Random Forest	0.6580	0.8602
Gradient Boost	0.6385	0.8609
Light Gradient Boost	0.6294	0.8559
AdaBoost	0.5647	0.7675

[그림 5]는 학습이 완료된 모델의 분류 결과에 따라 ROC(Receiver Operating Characteristic) 곡선을 그린 그래프이다. 실선의 경우 4개의 레이블에 대한 각각의 ROC 곡선을 표시한 것이며 파랑, 녹색, 빨강, 보라 순으로 각각 Angry(0), Calm(1), Depress(2), Happy(3)을 의미한다. 두 점선의 경우 각

ROC Curve에 대해 micro 평균과 macro 평균을 낸 결과이다. 세 가지 분류기에서 모두 Depress 감성에 대한 AUC가 가장 높고 Happy 감성에 대한 AUC가 가장 낮은 것을 확인할 수 있다.

표 3. 각 AutoML에서 사용된 모델에 대한 특징

Table 3. Characteristic of models used in each AutoML(seg.= segmentation, dep. = dependant)

Algorithm	Characteristic of models
Logistic Regression	Use logistic functions to estimate causality between dep. and indep. variables
LDA(Linear Discriminant Analysis)	Learn probability distribution to create decision boundaries based on bayes theory
QDA(Quadratic Discriminata Analysis)	Similar to LDA, but ach class has its own covariance matrix
Decision Tree	Classify data according to a set of classification rules constructed with tree
Random Forest	Construct mulptle decision tress using multiple sampling of training data
Gradient Boost	Use the boosting algorithm with gradient descent optimization
Light Gradient Boost	Continuous seg. of leaf nodes with maximum loss, not balanced tree seg.
AdaBoost	Gather the weighted weak classifiers to create a strong classifier

[표 4]은 학습이 완료된 모델의 분류 결과에 따라 오차 행렬을 나타낸 표이다. [그림 6]는 세 개의 분류기 중 가장 AUC 평가 지표 기준 가장 분류 성능이 높았던 Random Forest 모델의 특징 중요도(Feature Importance)를 그래프로 도식화한 것이다. 추출한 6개의 종류의 특징이 모두 고르게 분류에 이용되었다는 것을 확인할 수 있으며, 특히 일반적인 RGB 색상보다 이미지 내 채도의 분포, 색상의 분포가 상대적으로 중요하게 쓰였음을 확인할 수 있다. 이로써 우리는 일반적으로 인간이 이미지를 통해 감성을 느낄 때 색상 톤이 중요하게 쓰인다는 특성이 기계학습 알고리즘을 통한 예측에도 반영된다는 것을 확인할 수 있다.

표 4. 분류기 학습에 대한 오차행렬 (lr은 Logistic Regression, rf는 Random Forest, gb는 Gradient Boost, avg는 세 분류기 평균 값을 의미한다.)

Table 4. Error matrix for classifier learning (lr : Logistic Regression, rf : Random Forest, gb : Gradient Boost, avg is average of 3 classifiers)

True class	Predicted class							
	lr	rf	lr	rf	lr	rf	lr	rf
	gb	avg	gb	avg	gb	avg	gb	avg
0	45	38	11	12	11	11	15	21
	47	43.3	9	10.6	10	10.6	16	17.3
1	7	3	53	58	8	5	4	6
	7	5.6	54	55	5	6	6	5.3
2	3	2	4	8	63	60	0	0
	1	2	9	7	60	61	0	0
3	17	15	15	15	5	7	40	40
	21	17.6	14	14.6	5	5.6	37	39
	0		1		2		3	

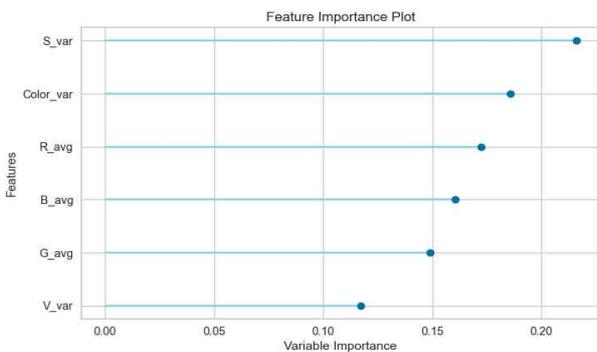


그림 6. Random Forest 분류기에 대한 특징 중요도(Random Forest 모델의 분류 시 모델이 결과 판단에 사용하는 특징들에 대한 중요도를 나타낸 그래프이다.)

Fig. 6. Feature importance for Random Forest classifiers (a graph of the importance of features that the model uses to judge results when classifying Random Forest models).

V. 결 론

인공지능 기술이 대두되며 기계학습 알고리즘을 활용한 감성 컴퓨팅 연구가 활발하게 진행되고 있다. 최근 딥 러닝(Deep Learning) 기술이 발달하며 분류 문제는 굉장한 발전을 이루고 있다. 하지만 딥 러닝 기술은 인공 신경망의 수많은 은닉층으로 구성되어 인간이 모델을 이해할 수 없는 블랙 박스(Black Box)를 이루고 있다. 이에 우리는 설명 가능한 기계 학습 모델을 만들기 위해 이미지 전처리 과정 후 특징을 추출하여 이를 통해 감성을 추출하였다. 이는 딥 러닝 모델에 비해 성능이 안 좋다는 한계를 가지지만, 현재까지 연구된 인간의 감성 인지 분야와 기계학습을 통한 감성 추론의 비교 분석을 용이하다.

우리가 제시한 모델은 색상에 주목한다. 이미지에는 감성을 일으킬 수 있는 수많은 요소가 존재한다. 이를테면 이미지 내에 존재하는 키워드, 물체 등이 그 예가 될 수 있다. 그렇기에 우리의 감성 분류 결과가 100%에 수렴할 정도로 좋은 모델을 제시할 수는 없지만, 색상만으로도 일정 수준 이상의 결과를 제시할 수 있음에 색상과 감성 사이의 상관관계에 대해 설명을 할 수 있다. 특히, 보편적으로 많이 쓰이는 기계학습 모델들을 활용하여 특징 중요도를 제시함으로써 설명할 수 있는 인공지능을 구현했다고 말할 수 있다. 우리 연구는 위 서술한 바의 의의를 지니며 향후 기계학습 기반의 색상과 감성의 상관관계 연구에 이바지할 수 있을 것으로 예상된다.

참고문헌

- [1] H.S. Shin, K.A. Shin, I.G. Shin, "The Study on Receiver's Memory and Attitude on the Stores by Signboard Layout Style and Color Harmony", Korean Academy of OOH advertising, 6(4), 127-149, 2009.
- [2] B. H. Seo, and K.M. Lee, "A Study on the Reliability and Usefulness of Color Psychology for Clinical Practices", Journal of Arts Psychotherapy, 9(22), 139-160, 2013
- [3] S.K. Langer, "Problems of Art", Art History, 1957
- [4] N. Kolenda, "The psychology of choice", Psychology & Marketing, 2016.
- [5] S. Kobayashi, "The aim and method of the color image scale," Color Res. Appl., 6(2), 3534-3539, 1981
- [6] Patil. Daivata (2012), "Coloring Consumer's Psychology Using Different Shades the Role of Perception of Colors by Consumers in Consumer Decision Making Process: a Micro Study of Select Departmental Stores in Mumbai City, Journal of Business & Retail Management Research, 7(1), 60-73, 2012.
- [7] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying aesthetics in photographic images using a computational

approach,” Computer Vision –ECCV 2006, 288–301, 2006

[8] J. Machajdik A. Hanbury, “Affective image classification using features inspired by psychology and art theory,” Int. Conf. Multimedia, 92–93, 2010.

[9] D. Joshi, R. Datta, E. Fadorovskaya, Q. -T. Luong, J. Z. Wang, J. Li, and J. Luo, “Aesthetics and emotions in images,” Signal Processing Magazine, IEEE, 28(5), 94–115, 2011.

[10] S. Zhao, Y. Gau, X. Jiang, H. Yao, T. -S. Cua, and X. Sun, “Exploring principles-of-art features for image emotion recognition,” Proceedings of the ACM International Conference on Multimedia, 47–56, 2014.

[11] J.-K. Chang and S.-T. Ryou, "Real-time emotion retrieval scheme in video with image sequence features", Journal of Real-Time Image Processing, 9(3), 541-547, 2013.

[12] K. Shin, and I. Lee, “Music synchronization with video using emotion similarity”, Big Data and Smart Computing, 47-50, 2017.

[13] V. Yanulevskaya, J. C. Van Germert, K. Roth, A. K. Herbold, N. Sebe, and J. M. Geusebroek, “Emotional valence categorization using holistic image features,” In IEEE International Conference on Image Processing, 101–104, 2008.

[14] T. Lee, H. Lim, D. Kim, S. Hwang, and K. Yoon, “System for matching paintings with music based on emotions,” In Technical Brief of the Siggraph Asia. Article 31, 1–4, 2016.

[15] D. Kang, H. Shim and K. Yoon, “A method for extracting emotion using colors comprises the painting image,” Multimedia Tools and Applications, 77(4), 4985–5002, 2018.

[16] SeleniumHQ Browser Automation [Online]. Available: <https://www.selenium.dev>

[17] M. Ali, “PyCaret: An open source, low-code machine learning library in Python”. [Online]. Available: <https://www.pycaret.org>, 2020

[18] T. Nagarajah and G. Poravi, "A Review on Automated Machine Learning (AutoML) Systems," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), 1-6, doi: 10.1109/I2CT45611.2019.90338 10, 2019.



이나혁(Nahyuk Lee)

2019년~현재 : 중앙대학교 소프트웨어학부 학사과정

※ 관심분야 : 컴퓨터비전(Computer Vision), 기계학습(Machine Learning)



이태민(Taemin Lee)

2011년 : 중앙대학교 컴퓨터공학과 (공학사)

2013년 : 중앙대학교 일반대학원 (공학석사-컴퓨터그래픽스)

2019년 : 중앙대학교 일반대학원 (공학박사-컴퓨터그래픽스)

2019년~현재 : 중앙대학교 다빈치 SW교육원 특임교수
※ 관심분야 : 비사실적 렌더링(Non-Photorealistic Rendering), 색상 이론(Color Theory), 감정 컴퓨팅(emotional computing), 인공지능(Artificial Intelligence)