

선행연구 분석을 통한 그래픽 디자인에서의 공학적 분석방법론 적용가능성에 관한 연구

이미지 인식 기술을 중심으로

A Study on the Applicability of the Engineering Analysis in the field
of Graphic Design through Preceding Research

Focusing on Image Recognition

주 저 자 : 홍석주 (Hong, Suk Joo)

한국과학기술원 문화기술대학원

sukjoo.hong@kaist.ac.kr

공동저자 : 이지현 (Lee, Ji Hyun)

한국과학기술원 문화기술대학원

Abstract

With an advance development at the area of computer graphics, image recognition technology has been improved dramatically. Computer vision is making rapid progress due to the improvements of computer hardware, quality & quantity of image data and Machine-Deep Learning algorithms. Now images can be segmented into semantic units and derived meanings from features. Graphic designers can make use of image processing techniques in their working area. With this research, the significance and limitation of preceding researches in the field of graphic design are analyzed. Also, image recognition technology in the field of computer vision is introduced, for the sake of the quantification method in engineering basis. The applicable proposition is offered for graphic design methodology and this research will work as a phased reference to the future quantitative experiments.

Keyword

Graphic Design, Survey, Quantitative Analysis, Computer Vision, Image Recognition

요약

컴퓨터 성능의 향상과 이미지 데이터의 양적 증가와 질적 향상, 머신러닝·딥러닝 분야의 발전으로, 사진 및 평면 이미지를 정보 단위로 분할하고 각각의 의미의 확보가 가능하게 되었다. 이는 해당 기술을 그래픽 디자인에 적용한다면 향후 디자이너가 작업을 진행함에 있어 유용한 도구로 활용될 수 있음을 뜻한다. 이에 본 연구는 파악된 그래픽 디자인 선행연구의 동향을 바탕으로 그 의의와 한계점을 분석하고, 그래픽 디자인을 공학적 방법으로 정량화할 수 있는 컴퓨터 비전(Computer Vision) 분야의 이미지 인식(Image Recognition) 기술을 소개하고자 한다. 최종적으로 본 연구는 해당 기술의 적용이 가능한 디자인 분석방법을 제안하고자 하며, 이는 향후 그래픽 디자인을 대상으로 실험할 정량적 분석에 앞서, 공학적 연구의 단계적 접근으로써의 의의를 지닌다.

목차

1. 서론

- 1-1. 연구 배경 및 목적
- 1-2. 연구 방법 및 범위

2. 이미지 인식 기술 관련 그래픽 디자인 연구 동향

- 2-1. 조형 요소 연구
- 2-2. 레이아웃 구성 원리 연구

3. 공학적 연구 방법

- 3-1. 이미지 인식(Image Recognition) 기술
- 3-2. 활용 방안
- 3-3. 의의 및 한계

4. 결론

참고문헌

1. 서론

1-1. 연구의 배경 및 목적

인터넷에서 구할 수 있는 이미지양의 증가와 데이터의 질적인 향상, 이미지 처리에 효율적인 컴퓨터 GPU의 보급, 수학적·통계적 알고리즘 연구들의 축적으로 이미지 인식 기술의 정확도는 사람보다 앞서고 있다. 2015년 ILSVRC(The ImageNet Large Scale Visual Recognition Challenge)에서 우승한 ResNet은 이미지 인식 오류율이 3.57%로 인간의 오류율 5%보다 낮은 수치를 보였다.¹⁾ 이미지 인식 기술은 사진에서 사물과 글자의 위치와 범위를 확보하고 그것의 의미 정보를 추출한다. 이 기술은 학문적 연구뿐만 아니라 얼굴 인식²⁾, 사진 스타일 변경³⁾, 주제별 사진 분류⁴⁾, 간판 글자 번역과 스캔 문서 변환⁵⁾ 등과 같이 실생활에서 사용하는 어플리케이션에도 적용되고 있다. 이와 같이 이미지 인식 기술을 사용해 사진 이미지를 활용하는 응용 연구는 활발하지만, 그래픽 디자인을 대상으로 하는 연구가 부재했다. 또한 시각적 표현의 예술성, 주관성, 해석의 다양성 등의 특징들은 그래픽 디자인 연구가 정성적 분석 방법에 제한되는 결과를 가져왔다. 이에 조형요소의 구성 원리나 시각적 표현 방법들이 디자인 연구자와 디자이너의 이론적이고 경험적인 분석 방법을 토대로 연구되어 왔다.

본 연구의 목적은 그래픽 디자인 분야의 선행연구들을 살펴보고, 컴퓨터 비전 분야의 기술을 이용해 정량적으로 분석할 수 있는 연구 방법론을 제시함에 있다. 선행연구들이 그래픽 디자인을 분석하는 관점과 이론을 도출하는 방법을 정리하고, 그것들의 의의와 한계점을 고찰한다. 그래픽 디자인에서도 공학적 연구가 가능함

에 따라, 관련된 컴퓨터 비전 분야의 이미지 인식 기술들을 소개한다. 이를 바탕으로 해당 기술을 적용한 그래픽 디자인 분석 방법을 제안하여 향후 연구에 대한 기초적인 토대를 제공한다.

1-2. 연구의 방법 및 범위

연구 방법은 문헌조사를 통해 그래픽 디자인의 이론적 개념과 방법론적 원칙을 정리하고 사례를 분석한 선행연구를 바탕으로 진행했다. 디자인 요소들을 분류하고 각각의 특징을 서술하며, 표현된 조형 원리를 시각적 효과로 도출한 논문들을 조형 요소 연구로 분류했다. 타이포그래피와 이미지가 표현하는 의미와 내용을 중심으로 선행연구들의 의의와 한계점을 지적했다. 다음으로 조형 요소들의 구성으로 레이아웃의 형태가 표현하는 시각적 효과를 분석한 논문들을 레이아웃 구성 원리 연구로 분류했다. 레이아웃의 분류 기준과 표현 특징, 검증 방법에 대한 의의와 한계점을 지적했다.

연구 영역은 평면으로 제작되어 출판물로 유통되거나 화면으로 표시되는 그래픽 디자인으로 한정했다. 특정한 목적을 위해 예외적인 크기로 제작되거나 시간에 따라 달라지는 동적인 그래픽 디자인은 선행연구들에서 고려하지 않았다. 이미지 인식 기술 부분에서는 분석하고자 하는 데이터의 종류에 따라 얼굴 검출(Face Detection), 지형 검출(Landmark Detection), 웹 검출(Web Detection), 로고 검출(Logo Detection) 등 분야가 다양하지만, 그래픽 디자인 조형요소와 관련된 이미지와 글자 2 가지로 분류했다.

2. 이미지 인식 기술 관련 그래픽 디자인 연구 동향

2-1. 조형 요소 연구

그래픽 디자인에서 조형 요소를 중심으로 한 연구들은 타이포그래피나 이미지의 형태가 표현하는 시각적인 의미를 분석했다. 선행연구들은 공통적으로 그래픽 디자인의 기본적인 조형 요소로 타이포그래피와 이미지를 제시했다. 이미지 인식 기술에서 글자와 이미지의 내용과 의미를 확보할 수 있음을 바탕으로 선행연구들을 살펴보았다.

먼저 남미현·백진경(2004)은 책표지 디자인에서 조형 요소들의 속성이 선호도에 미치는 영향을 분석했다. 속성의 기준은 타이포그래피에서 글자의 서체와 자간,

- 1) Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition, Computer Vision and Pattern Recognition, 2015, pp.772-777.
- 2) Paul Viola, Michael J. Jones, Robust Real-Time Face Detection, International Journal of Computer Vision, 2004, pp.137-154.
- 3) Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, Image Style Transfer Using Convolutional Neural Networks, Computer Vision and Pattern Recognition, 2016, pp.2414-2421.
- 4) Ali Borji, Ming-Ming Cheng, Qibin Hou, Huaizu Jiang, Jia Li, Salient Object Detection: A Survey, Pattern Analysis and Machine Intelligence, 2017, pp.11-13.
- 5) Qixiang Ye, David Doermann, Text Detection and Recognition in Imagery: A Survey, Pattern Analysis and Machine Intelligence, 2015, pp.1480-1482.

행간, 크기, 색상으로 하고 사진 및 일러스트레이션에서는 이미지의 스타일, 글자와의 비율로 정했다. 하지만 각각의 속성에 대한 정의가 추상적 개념에 의존하고 있어 불분명하고, 실험조사에서 형용사 평가 방법을 사용해 선호도 검증이 명증하지 않다.

김응철·김후성(2016)은 디자인의 표현요소의 변화를 연구하기 위해, 조형 요소들의 중요도에 따라 그래픽 디자인을 사진, 일러스트, 타이포그래피, 하이브리드 네 가지 유형으로 분류했다. 작업물에 여러 조형 요소가 속한 경우 가장 주목성이 큰 조형 요소를 기준으로 하고, 분류가 어려운 경우 하이브리드로 정했다. 본 연구자들이 지적했듯이 조형 요소의 분류 기준이 모호함에 따라, 표현 방법의 경향의 변화를 정량적으로 연구하는데 있어서 한계점이 존재했다.

전원화·이철영(2014)은 그래픽 디자인의 조형요소의 변화가 소비자의 평가에 미치는 영향을 알아보기 위해, 조형요소를 로고타입, 바탕색상, 일러스트, 레이아웃 네 가지로 설정했다. 실험을 위해 로고타입에서는 타입페이스의 속성에 따라, 일러스트에서는 이미지의 모습의 변화를 기준으로 실험 조사를 실시했다. 하지만 조사에 사용된 디자인 예시에서 각각의 조형요소의 변화 요인이 하나의 경우밖에 보여주지 못했다.

송민정·최은유(2013)는 책표지 디자인의 분야별 시각적 특징을 분석하기 위해, 시각적 요소를 타이포그래피, 이미지, 색채로 분류했다. 타이포그래피는 명조체, 고딕체, 장식체, 캘리그래피로, 이미지는 일러스트레이션, 사진, 컴퓨터그래픽으로, 색채는 형용사 이미지 스케일로 세분화했다. 하지만 타이포그래피와 이미지의 세분화 기준은 명확하나, 특징과 효과는 추상적이고 주관적이다.

2-2. 레이아웃 구성 원리 연구

레이아웃 구성 원리를 중심으로 한 연구들은 레이아웃의 형태가 만들어내는 시각적 표현 효과를 분석했다. 선행연구들은 템플릿 안에서 복수의 타이포그래피와 이미지들의 배치를 레이아웃의 개념으로 정의한다. 이에 이미지 인식 기술에서 글자와 이미지의 위치를 확보하고 차지하는 크기만큼 영역을 표시해주는 경계 상자(Bounding Box)의 적용 가능성을 확인하기 위해 선행연구들을 조사했다.

최선주(2004)는 웹디자인에서 조형 요소의 활용방법을 확인하기 위해, 레이아웃 유형을 점·선·면 단일요소 레이아웃, 가로·세로 그리드 레이아웃, 곡선·사선 레

이아웃 등으로 분류하고 사례를 통해 표현 효과를 분석했다. 하지만 제시하는 레이아웃 유형들의 분류 기준이 모호하고, 각각의 표현 효과가 검증되지 않았다.

홍영일(2010)은 레이아웃을 기준으로 응용할 수 있는 가이드라인을 제시하기 위해, 레이아웃 구성 원리를 통일, 변화, 균형, 율동, 강조로 분류했다. 나아가 통일은 조화와 질서, 변화는 비례와 변형, 균형은 대칭과 비대칭, 율동은 반복과 연속, 점미, 방사, 교대로, 강조는 대비와 배경으로 세분화 했다. 구성 원리의 시각적 특징을 사례를 통해 설명하지만, 조형 요소들의 배치에 대한 설명이 부족하다.

김병주(2015)는 반응형 웹에서 레이아웃 유형별 사례에 대한 사용자 경험 디자인을 연구하기 위해, 레이아웃의 유형을 유동형, 하단 이동형, 자리 이동형, 미세 수정형, 부분 화면형으로 분류했다. 반응형 웹의 특성상 조형요소의 위치와 크기의 속성은 명확하지만, 본 연구자가 결론에서 지적하듯이 경험요소의 특징 검증방법이 리커트 척도 하나에 그쳐 추가적인 실험조사가 필요하다.

고건·김보연(2015)은 표지디자인이 구매 결정에 미치는 영향을 분석하기 위해, 구성 요소를 타이포그래피, 이미지 등으로 분류했다. 나아가 레이아웃을 일러스트레이션, 사진과 같은 이미지, 타이포그래피, 색으로 이루어져 주목성, 가독성, 조형성, 창조성을 표현하는 조형 계획이라 정의했다. 레이아웃의 조건으로 타이포그래피와 이미지 등을 제시하지만, 조형 요소들이 어떠한 구성을 통해 레이아웃의 형태를 갖추게 되는지에 대한 설명이 부족했다.

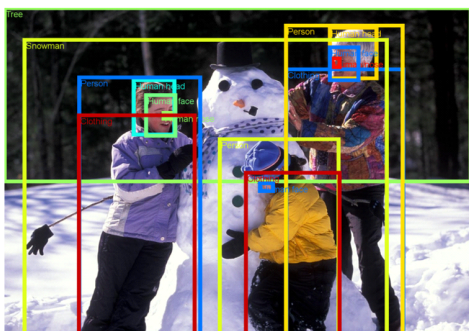
3. 공학적 연구 방법

3-1. 이미지 인식(Image Recognition) 기술

이미지 데이터의 양적인 증가와 질적인 향상, 병렬 계산에 효율적인 GPU의 보급, 수학적·통계적 알고리즘 연구들의 축적으로 머신러닝과 딥러닝의 발전이 이루어졌다. 특히 이미지 처리 관련 연구들은 컴퓨터 비전 분야에서 이미지 인식 기술에 새로운 방법론을 제시했다.⁶⁾ 이미지 인식 기술은 사진에서 배경으로부터 사물

6) Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems 25, 2012, pp.1-9.

의 영역을 확보하고 해당 이미지의 의미를 파악하는 것을 의미한다. 검출하고자 하는 이미지의 종류에 따라 객체 검출과 글자 검출로 분류할 수 있다. 객체 검출은 얼굴, 고양이, 의자 등의 사물을 확보하는 것을 목표로 한다. [그림 1]은 Google의 OpenImages dataset⁷⁾을 활용해 V2 data visualizer를 사진에 적용한 모습이다. 글자 검출은 그래픽적인 모양의 글자를 확보하고 해당 이미지를 글자로 변환하는 것을 목표로 한다. [그림 2]는 Google Cloud Vision API에서 Optical Character Recognition을 간편에 적용한 모습이다.



[그림 1] 객체 검출 기술을 사진에 적용한 모습



[그림 2] 글자 검출 기술을 간편에 적용한 모습

이미지 인식 기술을 사용하기 위해서는 직접 모델을

만드는 방법과 기업 및 연구 기관에서 제공하는 모델을 사용하는 방법이 있다. 전자는 해당 연구에 적합한 데이터를 확보하고, 머신러닝과 딥러닝 알고리즘을 통해 이미지 분석 프로그램을 제작하는 방법이다. 때문에 실험에 최적화된 모델을 사용할 수 있게 됨으로써 유의미한 결과를 얻어낼 수 있는 확률이 높아진다. 하지만 이 경우 이미지를 수집하는 것뿐만 아니라 모델에 학습이 가능하도록 데이터 마이닝 과정을 수행해야 한다. 또한 학습 과정에서 많은 컴퓨팅 자원이 필요하기 때문에 고성능의 장비가 필요하다.

후자의 방법은 해당 기관에서 의도한 목적에 부합하는 데이터셋과 알고리즘을 사용하기 때문에 실험 이미지의 종류에 따라 성능과 결과의 차이를 보이게 된다. 하지만 대부분의 경우 90% 이상의 정확도를 검증 단계에서 보여주기 때문에 모델을 연구에 적절한 형태로 수정하는 방법을 취할 수 있다. 대표적인 프로그램으로 Google Cloud Vision⁸⁾과 IBM Watson Visual Recognition⁹⁾, Amazon Rekognition¹⁰⁾, Microsoft Computer Vision API¹¹⁾, Clarif.ai¹²⁾, CloudSight¹³⁾ 등과 같은 API(Application Programming Interface)가 있다.

3-2. 활용 방안

그래픽 디자인이 타이포그래피와 이미지를 기초로 조형 요소로 사용한다는 점에서 이미지 인식 기술을 연구에 적용할 수 있다. 우선 그래픽 디자인 작업물을 이미지 인식 모델로 분석하면 글자와 사물의 위치와 크기를 알 수 있다.¹⁴⁾ 이는 템플릿 안에서 글자와 이미지가 개별적으로 차지하는 절대적인 크기와 비율의 정보를 분석할 수 있음을 의미한다. 나아가 디자인 요소들의 위치 정보를 확보하게 되면 그래픽 디자인의 레이아웃 또한 분석이 가능하다. 이에 Bylinskii 외 8명(2017)은 그래픽 디자인에서 이미지 인식 기술의 하나

7) Krasin I., Duerig T., Alldrin N., Ferrari V., Abu-El-Haija S., Kuznetsova A., Rom H., Uijlings J., Popov S., Veit A., Belongie S., Gomes V., Gupta A., Sun C., Chechik G., Cai D., Feng Z., Narayanan D., Murphy K., OpenImages: A public dataset for large-scale multi-label and multi-class image classification, 2017, github.com/openimages.

8) cloud.google.com/vision

9) ibm.com/watson/services/visual-recognition

10) aws.amazon.com/rekognition

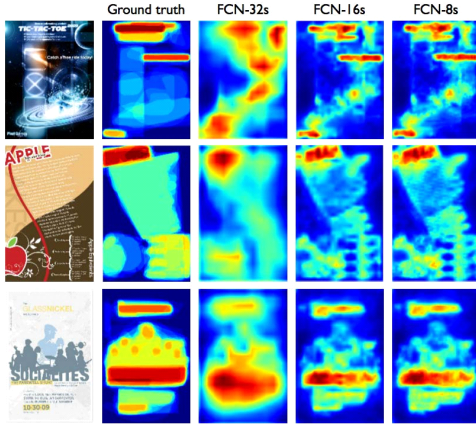
11) azure.microsoft.com/services/cognitive-services/computer-vision

12) clarifai.com

13) cloudsight.ai

14) Dim P. Papadopoulos, Jasper R. R. Uijlings, Frank Keller, Vittorio Ferrari, We don't need no bounding-boxes: Training object class detectors using only human verification, Computer Vision and Pattern Recognition, 2016, pp.854-863.

인 돌출 감지(Saliency Detection) 기술을 통해 조형 요소의 속성에 따라 레이아웃의 변화를 분석하는 접근을 시도해보았다. [그림 3]은 그래픽 디자인의 유형과 분석 모델의 종류에 따라 돌출 정도의 차이를 보여준다.



[그림 3] 그래픽 디자인 요소의 돌출 정도 비교

또한 도형으로 존재하는 글자가 가진 의미 정보를 확보할 수 있고, 사물의 의미 정보를 이름의 형식으로 레이블링할 수 있다. 이를 통해 그래픽 디자인에서 사용된 조형 요소들의 의미 정보들을 확보할 수 있다. 동시에 해당 작업물의 콘텐츠를 고려하게 되면, 디자인 의도가 시각적으로 표현되는 조건을 살펴볼 수 있다. 이에 타이포그래피에서 타이페이스의 속성을 이미지의 상태에서 분석하고 새로운 폰트를 생성해내는 연구가 있었다.¹⁵⁾ 나아가 그래픽 디자인 세부 분야나 장르를 기준으로 분류하게 되면, 특정 그래픽 디자인이 사용하는 관습적 조형 원리를 발견할 수 있다.

이미지 인식 기술을 통해 분석한 디자인 원리는 출판 기반 그래픽 디자인을 제작 과정에 발생하는 결정에 정량적인 근거가 된다. 나아가 이는 본 논문에서 연구 범위로 한정된 정적인 그래픽 디자인에 머물지 않는다. 이는 이미지 인식 기술이 래스터 이미지의 픽셀들을 특징적인 요소에 따라 단위별 패턴으로 추출하기 때문에, 벡터 기반의 동적인 그래픽 디자인 분야인 웹 디자인에도 응용이 가능하다.

3-3. 의의 및 한계

15) erikbern.com

본 연구 방법의 의의는 그래픽 디자인을 컴퓨터 기술을 사용하여 디자인의 관점에서 연구할 수 있는 가능성을 제시한데 있다. 그래픽 디자인이 가진 시각적 표현의 예술성과 디자이너에 따른 주관성, 해석의 다양성은 정성적 분석 연구에 도움이 되지만, 정량적 분석 연구에는 활용 불가능한 정보로 취급된다. 하지만 컴퓨터를 사용해 디자인 요소를 이미지 자체로 확보할 수 있다면, 정량적 분석 방법에 기초한 정성적 연구가 가능하다. 이를 바탕으로, 그래픽 디자인은 글자와 이미지를 활용해 시각적 결과물을 만들어내기 때문에, 디자이너가 만들어낸, 차용한, 변형된 디자인 요소들이 혼재한다. 각각의 부분들은 위의 기술들을 활용해 그래픽 디자인 요소들의 위치, 크기, 종류, 의미 등을 정량적으로 확보할 수 있다. 개별 그래픽 디자인을 시각적 특징과 의미를 기준으로 분석할 수 있고, 웹 크롤링 기법을 통해 대량의 작업물을 자동적으로 일괄 처리하여 결과를 통계적 분석을 진행할 수 있다. 나아가 그래픽 디자인을 대상으로 하는 연구가 컴퓨터를 통해 대량의 사례들을 분석할 수 있다. 디자인 요소의 종류와 배치, 사용 빈도를 통해 그리드와 레이아웃을 파악하고, 시대적·장소적 경향과 분야별·장르별 특징 등까지 통계적으로 분석이 가능하다.

하지만 이미지 인식 기술 모델들은 사진 이미지를 학습 데이터로 사용하기 때문에 과도한 변형이 일어난 이미지나 콜라주 이미지 등을 검출하는데 낮은 정확도를 보인다. 또한 모델은 이미지를 의미적으로 분류하기 위한 이름 카테고리가 있는데, 사진에 많이 등장하는 대상들을 중심으로 구성했기 때문에 일러스트나 그래픽 도형 등의 이미지를 분류하지 못하는 경우가 있다. 이미지 인식 기술에서 글자 검출은 다양한 타이페이스의 글자 모양을 인식하는 기술에 한정되어, 글자의 디자인적 요소를 곧바로 도출하기는 힘들다.

4. 결론

본 논문은 그래픽 디자인 연구의 동향을 분석 대상의 범주에 따라 분류하고, 의의와 한계점을 중심으로 정리했다. 이에 정량적 분석방법이 부재함을 확인하고, 그래픽 디자인에 적용이 가능한 컴퓨터 비전 분야의 이미지 인식 기술을 소개했다. 이는 그래픽 디자인을 공학적 관점에서 연구할 수 있는 분석방법을 제시함에 의의가 있다. 나아가 선행연구의 이론들을 보충하거나 재해석할 수 있는 기술적 토대가 된다.

후속 연구에서는 이미지 인식 기술을 사용해 책 표지, 앨범 커버, 영화 포스터를 대상으로 레이아웃의 유형을 분석할 것이다. 이 연구는 템플릿의 비율에 따른 레이아웃의 변화 양상을 통계적으로 도출할 수 있을 것이다. 동시에 실험 과정에서 그래픽 디자인 이미지에 생성된 속성값은 컴퓨터 비전 분야에 활용 가능한 데이터셋이 될 것이다.

본 연구는 국내에서 발표된 논문들을 대상으로 문헌 조사를 진행하여 국외 그래픽 디자인 논문을 추가적으로 탐구할 필요가 있다. 또한 이미지 인식 기술은 컴퓨터 비전 뿐만 아니라 컴퓨터 과학, 인지 과학, 수학, 통계학 등 여러 분야와 함께 발전한다. 이에 타 학문에 대한 종합적인 이해를 바탕으로 한 연구가 필요하다.

참고문헌

- 고건, 김보연, 베스트셀러 표지디자인과 소비자의 구매결정에 관한 연구 -온라인 서점을 중심으로-, 커뮤니케이션 디자인학연구, 2015, Vol.53.
- 김병주, 사용자 경험디자인 요소에 따른 반응형 웹 GUI 디자인 분석 -남녀대학생을 중심으로-, 한국디자인문화학회지, 2015, Vol.21, No.2.
- 김응철, 김후성, 한국 대중 음악 음반 커버 디자인의 표현 요소 트렌드 변화 연구 -2000년대 이후 발매된 음반의 음반 커버를 중심으로-, 한국디자인문화학회, 2016, Vol.22, No.2.
- 남미현, 백진경, 책표지 디자인에서 시각적 요소가 독자 선호도에 미치는 영향 -문학도서를 중심으로-, Archives of Design Research, 2004, Vol.55.
- 송민정, 최은유, 분야별 책표지디자인의 시각적 요소 분석을 통한 시각적 특징, 한국디자인문화학회지, 2013, Vol.19, No.1.
- 전윤희, 이철영, 디자인 중요성 인식수준이 패키지의 조형요소와 이미지의 관계에 미치는 매개적 효과, 디지털디자인학연구, 2014, Vol.14, No.4.
- 최선주, 그래픽 요소를 활용하는 레이아웃 원칙에 대한 연구 -웹디자인을 중심으로-, 한국콘텐츠학회 종합학술대회 논문집, 2004, Vol.2, No.2.
- 홍영일, 레이아웃 구성원리를 적용한 그래픽디자인의 효용성 증대에 관한 연구, 디지털디자인학연구 2010, Vol.10, No.4.
- Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems 25, 2012.
- Ali Borji, Ming-Ming Cheng, Qibin Hou, Huaizu Jiang, Jia Li, Salient Object Detection: A Survey, Pattern Analysis and Machine Intelligence, 2017, NIPS'12.
- Dim P. Papadopoulos, Jasper R. R. Uijlings, Frank Keller, Vittorio Ferrari, We don't need no bounding-boxes: Training object class detectors using only human verification, Computer Vision and Pattern Recognition, 2016,
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition, Computer Vision and Pattern Recognition, 2015.
- Krasin I., Duerig T., Alldrin N., Ferrari V., Abu-El-Hajja S., Kuznetsova A., Rom H., Uijlings J., Popov S., Veit A., Belongie S., Gomes V., Gupta A., Sun C., Chechik G., Cai D., Feng Z., Narayanan D., Murphy K., OpenImages: A public dataset for large-scale multi-label and multi-class image classification, github.com/openimages, 2017.
- Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, Image Style Transfer Using Convolutional Neural Networks, Computer Vision and Pattern Recognition, 2016.
- Paul Viola, Michael J. Jones, Robust Real-Time Face Detection, International Journal of Computer Vision, 2004.
- Qixiang Ye, David Doermann, Text Detection and Recognition in Imagery: A Survey, Pattern Analysis and Machine Intelligence, 2015.
- Zoya Bylinskii, Nam Wook Kim, Peter O'Donovan, Sami Alsheikh, Spandan Madan, Hanspeter Pfister, Fredo Durand, Bryan

Russell, Aaron Hertzmann, Learning Visual Importance for Graphic Designs and Data Visualizations, Computer Vision and Pattern Recognition, 2017.

18. aws.amazon.com/rekognition

19. azure.microsoft.com/services/cognitive-services/computer-vision

20. clarifai.com

21. cloud.google.com/vision

22. cloudsight.ai

23. erikbern.com

24. ibm.com/watson/services/visual-recognition