

CNN 기반 딥러닝을 이용한 패션 아이템 분류 및 결합

김희숙¹

¹광주과학기술원 대우교수

Classification and Combination of Fashion Items Using CNN-Based Deep Learning

Hye-Suk Kim¹

¹Lecture Professor, Gwangju Institute of Science and Technology, Gwangju Cheomdangwagi-ro 123, Korea

[요 약]

본 논문에서는 패션 아이템을 상황에 따라 적절하게 매칭할 수 있는 인공지능 시스템을 구현하기 위해서 fashion-MNIST를 기반으로 10가지 패션 아이템으로 레이블링한 후 합성곱 신경망 기반의 딥러닝 모델을 설계하였다. 설계된 딥러닝 모델을 기반으로 45,000개 패션 이미지를 훈련 데이터셋으로 사용하고, 15,000개 패션 이미지를 검증 데이터셋으로 사용하여 총 15에폭으로 딥러닝을 실행하였다. 딥러닝 실행 결과, 패션 아이템 분류의 정확도 평가에서 훈련 데이터 학습 정확도 96%, 검증 데이터 학습 정확도 94%가 출력되었다. 본 논문에서는 구현된 인공지능 패션 아이템 분류 시스템을 기반으로 7가지 유형의 패션 매칭을 제공할 수 있는 데이터셋을 구축하였으며, 구축된 데이터셋은 향후 개인의 다양한 패션 욕구를 충족시켜줄 수 있는 인공지능 패션 매칭 서비스의 기반이 될 것이라 기대한다.

[Abstract]

In this paper, in order to implement an artificial intelligence system that can appropriately match fashion items according to the situation, a deep learning model based on convolutional neural network was designed after labeling with 10 fashion items based on fashion-MNIST. Based on the designed deep learning model, 45,000 fashion images were used as a training dataset, and 15,000 fashion images were used as a verification dataset, and deep learning was performed with a total of 15 epochs. As a result of the deep learning execution, the training data learning accuracy 96% and the verification data learning accuracy 94% were output in the accuracy evaluation of fashion item classification. In this paper, we constructed a dataset that can provide 7 types of fashion matching based on the implemented artificial intelligence fashion item classification system. The established dataset is expected to become the basis of an artificial intelligence fashion matching service that can satisfy various fashion needs of individuals in the future.

색인어 : 인공지능, 딥러닝, 머신러닝, 합성곱 신경망, 생성적 적대 신경망, 패션

Key word : Artificial Intelligence, Deep Learning, Machine Learning, Convolutional Neural Network, Generative Adversarial Network, Fashion

<http://dx.doi.org/10.9728/dcs.2021.22.3.475>



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 08 January 2021; **Revised** 27 January 2021
Accepted 05 February 2021

***Corresponding Author; Hye-Suk Kim**

Tel: +82-62-530-0147

E-mail: idishskim@naver.com

I. 서론

이미지 인식 기술이 발전함에 따라 패션에도 인공지능(AI; artificial intelligence)을 이용한 다양한 서비스가 도입되고 있다. 영국의 패션 브랜드 버버리는 인공지능과 빅데이터 기술을 이용하여 마케팅과 선호도 조사, 모조품 구별 등에 활용하고 있다. 전자상거래 쇼핑몰 아마존은 인공지능 기술을 이용하여 옷을 디자인하고 온라인으로 옷을 입어보고 구매할 수 있는 서비스를 소개하였다. 뿐만 아니라 패션 기업들은 소셜 네트워크 서비스(SNS; social network service)에 업로드되고 있는 인플루언서(influencer)들의 패션 이미지를 분석하여 패션 트렌드 파악에 사용하고 있다. 인공지능 기술이 도입된 패션 쇼핑몰에서는 소비자 행동과 니즈를 반영하여 개인에게 맞는 패션 제품 추천 서비스를 제공하고, 판매 제품의 전반적인 수요 예측과 마케팅 및 전략적 의사 결정에 도움을 받고 있다. 따라서 인공지능 기술이 도입된 패션 쇼핑몰을 이용하는 소비자는 패션 쇼핑몰에서 구매하려고 하는 제품과 비슷한 제품을 추천받거나 매칭(matching)할 수 있는 제품을 소개받을 수 있게 되었다[1]. 이처럼 패션 산업에 활용되고 있는 인공지능 기술은 빅데이터를 기반으로 하는 데이터 과학(data science), 머신 러닝(machine learning), 딥 러닝(deep learning)과 같은 기술이 발전하였기에 가능한 일이 되었다[2].

현재 다양한 산업 분야에 활용되고 있는 인공지능 기술은 매우 빠르게 발전하고 있으며 LeNet-5 개발에 사용된 얀 르쿤(Yann LeCun)의 합성곱 신경망(CNN; convolutional neural network)[3]은 이미지 분류는 물론 자율주행 자동차, 의료 영상 분석 등에 유용하게 응용되고 있다. 손글씨 숫자 이미지 데이터에 해당하는 MNIST(modified national institute of standards and technology database)[3]를 벤치마킹하여 다양한 이미지 데이터셋이 등장하고 있으며 패션 분야에서는 잘란도(zalando)에서 패션 데이터셋 fashion-MNIST[4]를 2017년도에 발표했다. 최근 이를 기반으로 CNN 알고리즘을 활용하여 패션 아이템들을 분류할 수 있도록 딥러닝을 설계하고 실행할 수 있는 연구가 소개되었다[5], [6], [7].

사람들은 인공지능 기능이 탑재된 쇼핑몰에서 제공되는 다양한 서비스를 이용하여 자신의 개성을 표현하고 장소와 목적에 맞는 패션을 완성할 수 있는 패션 아이템(items)들을 구매하고 있다. 하지만 구매한 패션 아이템들을 활용할 장소 및 목적에 변화가 생길 경우, 구매한 패션 아이템들을 다시 적절하게 매칭하기 위해 시간과 노력을 소비해야 하는 문제는 여전히 존재하고 있다. 이러한 문제를 해결하기 위해 감각적인 색상 배색으로 손쉽게 의상을 코디네이션할 수 있도록 보색 대비의 스마트 의상 코디네이션에 관한 서비스가 소개되기도 하였다[8].

본 논문에서는 패션 아이템을 환경에 따라 적절하게 매칭할 수 있는 인공지능 서비스를 설계하기 위해서 fashion-MNIST를 기반으로 티셔츠/탑(T-shirt/top), 바지(Trouser), 풀오버(Pullover), 드레스(Dress), 코트(Coat), 샌들(Sandal), 셔츠(Shirt), 스니커즈(Sneaker), 가방(Bag), 앵클 부츠(Ankle boot)에 해당하

는 10가지 패션 아이템으로 구분하여 레이블링한 후 딥러닝으로 학습을 실행할 수 있는 시스템을 구축한다. 10가지 종류 중 하나로 분류가 가능해진 패션 아이템들을 7가지 패션 매칭 요소로 결합할 수 있는 데이터셋을 구성하여 향후 인공지능 패션 아이템 매칭 서비스를 위한 프레임워크를 설계한다. 본 연구에서 제안된 7가지 패션 매칭 요소는 티셔츠/탑(T-shirt/top)와 바지(Trouser) 매칭, 풀오버(Pullover)와 바지(Trouser) 매칭, 코트(Coat)와 바지(Trouser) 매칭, 셔츠(Shirt)와 바지(Trouser) 매칭, 드레스(Dress)와 샌들(Sandal) 매칭, 바지(Trouser)와 스니커즈(Sneaker) 매칭, 바지(Trouser)와 앵클 부츠(Ankle boot) 매칭으로 설계된다.

본 논문의 구성은 다음과 같다. 2장에서는 패션 분야에 활용되고 있는 딥러닝 알고리즘을 살펴보고 3장에서는 10가지 패션 아이템으로 구분하여 레이블링한 후 딥러닝으로 학습을 실행할 수 있는 시스템을 구축하기 위해서 패션 데이터셋 fashion-MNIST를 분석하고, 분석된 fashion-MNIST를 기반으로 7가지 패션 매칭 요소로 결합할 수 있는 데이터셋을 설계한다. 4장에서는 fashion-MNIST를 기반으로 10가지 패션 아이템으로 분류할 수 있는 CNN 패션 딥러닝 구현 모델을 설계하여 실행하고 테스트 및 평가한다. 이어서 5장에서는 결론 및 향후 연구 방향에 관해 서술한다.

II. 관련 연구

패션 디자인 분야는 디자이너의 역량 및 인지도에 따라서 소비자들의 반응에 영향을 줄 수 있는 분야이다. 패션 디자이너의 고유 영역이었던 디자인 분야에 생성적 적대 신경망(GAN; generative adversarial network)[9] 및 NST(neural style transfer)[10] 등의 딥러닝 기반의 생성 모델이 나오면서 훈련 데이터에 존재하지 않았던 새로운 패션 이미지를 자연스럽게 생성하는 일이 가능해졌다. 2-1절에서는 생성적 적대 신경망 GAN을 이용한 패션 디자인 원리에 대해 살펴보고 2-2절에서는 본 연구에서 활용하고자 하는 합성곱 신경망 CNN에 관해 설명한다.

2-1 GAN을 이용한 패션 디자인 원리

생성적 적대 신경망은 서로 대립하는 두 개의 신경망에 의해 생성되는 알고리즘으로 GAN이라고 축약되어 사용되고 있으며 이안 굿펠로우(Ian Goodfellow)에 의해 제안되었다[9]. ‘Adversarial’이란 단어의 의미처럼 ‘적대적’이라는 의미를 포함하고 있으며 대립하는 적대적 관계에 해당하는 두 부분의 신경망으로 구성된다. 즉, 이미지를 생성하는 부분에 해당하는 생성자(generator)와 생성된 이미지를 평가하는 판별자(discriminator)가 있다. 이안 굿펠로우는 GAN의 생성자와 판별자의 역할을 지폐위조범과 경찰로 비유하였다. 지폐위조범에 해당하는 생성자는 경찰을 최대한 열심히 속여서 위조지폐를 만들려고 하고, 판별자에 해당하는 경찰은 위조된 지폐와 실제 지폐를 감별해내는 역할을 하는 것으로 비유하여 설명하였다.

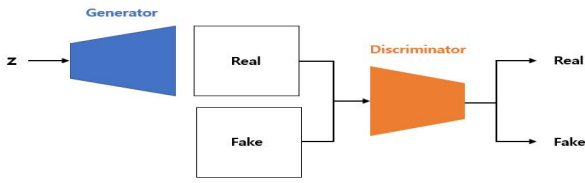


그림 1. GAN 흐름도
Fig. 1. Flowchart of GAN

GAN은 위조지폐범에 해당하는 생성자와 경찰에 해당하는 판별자를 경쟁적으로 학습시킨다. 생성자의 목적은 그럴듯한 가짜(fake) 데이터를 만들어서 판별자를 속이는 것이며, 판별자의 목적은 생성자가 만든 가짜 데이터와 실제 데이터를 구분하는 것이다. 이처럼 생성자와 판별자 모두를 함께 학습시키면서 실제와 위조된 상태를 구분할 수 없는 위조 상태를 만들어내는 생성자를 얻는 것이 목표이다.

아래의 그림 1은 생성자가 z를 받아서 실제 데이터와 비슷한 데이터를 만들어내도록 학습하고, 판별자는 실제 데이터와 생성자가 생성한 가짜 데이터를 판별하도록 학습하는 것을 표현하고 있다.

딥러닝 구현을 위해서는 생성하고자 하는 목표에 해당하는 샘플 데이터가 필요하며 생성자는 샘플 데이터의 분포를 보고 그럴듯한 새로운 데이터를 생성한다. 판별자는 생성자가 생성한 데이터가 실제 데이터인지 생성된 가짜 데이터인지를 판별하도록 한다. 이처럼 생성자는 속이는 모델에 해당하고 판별자는 속지 않기 위한 모델로 비유할 수 있으며 이 둘은 한 쌍이 되어서 정교한 학습이 이루어진다. 결국, 마지막에는 생성자의 성능이 올라가서 판별자가 0.5로 예측 할 수밖에 없게 하는 것이 가장 이상적인 상태가 된다.

위 그림 2에서 초록색 실선은 생성자가 만들어내는 가짜 데이터의 분포이고, 파란색 점선은 판별자가 실제 데이터와 가짜 데이터를 판별하는 선에 해당한다. 그리고 검정색 점들은 실제 데이터들의 분포를 나타내고 있다[9]. GAN 알고리즘은 식 (1)과 같이 간단하게 정리할 수 있다.

식 (1)에서 D는 전체적인 식의 값을 크게 만드는 방향으로 파라미터를 학습시킨다. 그러므로 D(x)는 크게 만들어야 좋고, D(G(z))는 작게 만들어야 좋다. 또한, G(z)는 D(G(z))를 최소화하려고 하므로 최대화되는 D(x)의 x와 같아지려고 노력한다. 위의 식 (1)을 이용한 GAN 딥러닝으로 새로운 이미지를 생성할 수 있게 되었다.

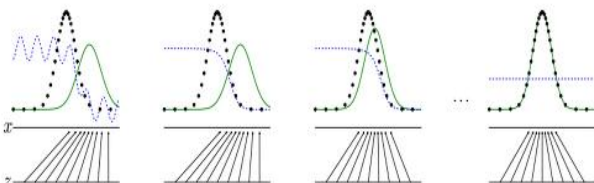


그림 2. GAN으로 생성된 데이터 분포 및 판별 시각화
Fig. 2. Visualization of distribution and discrimination of data generated by GAN

$$\min_G \max_D V(D, G) = \tag{1}$$

$$E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$



그림 3. 딥러닝으로 채색된 패션 디자인
Fig. 3. Fashion design colored by deep learning

스케치에 채색하거나, 흑백 사진을 컬러로 만들거나, 낮 사진을 밤 사진으로 바꾸고 싶을 때처럼 상황에 따라 이미지를 처음부터 생성하기보다는 이미 존재하는 이미지를 변형하고 싶은 경우도 있다. cGAN(conditional GAN)은 이를 가능하게 해주는 모델이다. GAN의 생성자가 랜덤 벡터를 입력으로 받는 것에 비해 cGAN의 생성자는 변형할 이미지를 입력으로 받는다. 그 뒤 생성자는 입력 이미지에 맞는 변형된 이미지를 출력한다.

가령 스케치 사진을 받은 생성자는 그 스케치에 맞는 색상을 채색한 후 채색된 이미지를 출력하는 것이다. 판별자는 스케치와 채색된 이미지를 확인하고 채색된 이미지가 스케치에 어울리는지 판별한다. 판별자를 속이기 위해서 생성자는 실제와 같은 채색을 하면서 스케치에 적합한 이미지를 만들어야 한다. cGAN의 핵심은 주어진 이미지를 새로운 이미지로 변형하는 수많은 문제를 하나의 간단한 네트워크 구조로 해결했다는 점이다. 이렇게 한 영역의 이미지를 다른 영역의 이미지로 변형할 때 cGAN이 유용하게 쓰일 수 있으며 패션 디자인을 완성할 때도 사용할 수 있게 되었다[11].

그림 3과 같이 cGAN을 이용하여 패션 디자인 아이템에 해당하는 스케치 이미지를 기반으로 자동으로 에지를 추출하여 채색을 입히는 딥러닝을 수행하여 인간 디자이너의 디자인 결과물인지, 인공지능 디자이너의 결과물인지 구별하기 힘든 패션 디자인을 완성할 수 있다[12].

2-2 CNN을 이용한 패션 아이템 분류

CNN 알고리즘은 1998년 LeNet-5를 발표하면서 필터링 기법을 인공신경망에 적용함으로써 이미지를 더욱 효과적으로 처리하기 위해 소개되었다[3].

CNN은 입력된 이미지로부터 이미지의 고유한 특징 맵(feature map)을 만들어 신경망에 연결하여 이미지가 어떤 레이블에 속하는지를 분류한다.

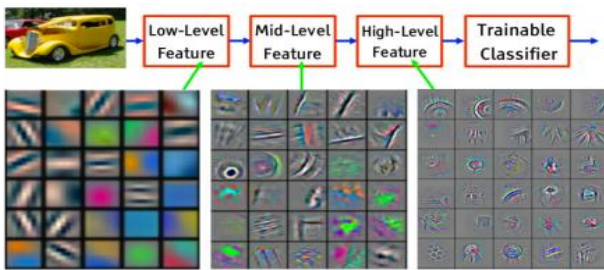


그림 4. CNN의 특징 맵 시각화
 Fig. 4. Visualization of CNN feature maps

즉, CNN은 특징 추출 신경망과 분류 신경망을 직렬로 연결한 구조로 구성된다.

이미지의 특징 맵을 얻기 위해서 수평 에지 필터와 합성곱 연산을 하면 수평 에지가 강조된 특징 맵을 얻게 되고, 수직 에지 필터와 합성곱 연산을 하면 수직 에지가 강조된 특징 맵을 얻게 된다. 특징을 추출하는 필터 행렬의 값은 학습을 통해 결정되고, 합성곱을 통해 얻은 특징 맵은 활성화 함수(activation function)를 거쳐 출력된다. 여기서 서브샘플링(subsampling) 또는 풀링(pooling)의 과정을 통해 활성화된 특징 맵들의 크기를 줄일 수 있다.

이 특징 맵들에 다시 합성곱, 활성화 함수, 서브샘플링을 수행해서 지역적인 특징 맵으로부터 점차적으로 전역적인 특징 맵을 만들어간다. 이 과정을 여러 번 반복하여 얻어진 최종 특징 맵은 FC Layer (fully connected layer) 신경망에 입력되어 레이블링 된 상태에 맞는 분류를 할 수 있게 된다.

그림 4는 CNN 네트워크 안에서 합성곱, 활성화 함수, 서브샘플링 과정을 반복함으로 저차원적인 특징 맵으로 시작해서 고차원적인 특징을 도출해나간 후 최종 특징 맵으로 훈련 가능한 분류기를 만들 수 있음을 보여준다[13].

사람의 시각 인지 과정을 모방해서 신경망에 추가한 CNN은 컴퓨터 비전 분야의 대표적인 딥러닝 알고리즘으로 정착되었고, CNN의 구조를 변경하여 R-CNN[14], Fast R-CNN[15] 등의 방식으로 CNN 기반의 딥러닝 알고리즘은 발전하고 있다.

III. 패션 아이템 매칭을 위한 데이터셋 설계

3-1 패션 데이터셋 분석

본 연구에서는 패션 데이터셋 fashion-MNIST[4]를 활용하여 10가지 종류로 패션 아이템을 분류할 수 있도록 딥러닝 모델을 설계하고 실행한다. 본 연구에서 사용되는 fashion-MNIST는 60,000개의 훈련 데이터셋과 10,000개의 테스트 데이터셋으로 총 70,000개의 데이터셋 이미지로 구성되어 있다. 제공되는 이미지는 너비와 높이가 28x28 (총 784) 픽셀로 구성되어 있으며 0~255 사이의 밝기에 해당하는 값이 들어있다. 아래의 그림 5는 fashion-MNIST의 생성 과정을 보여준다.

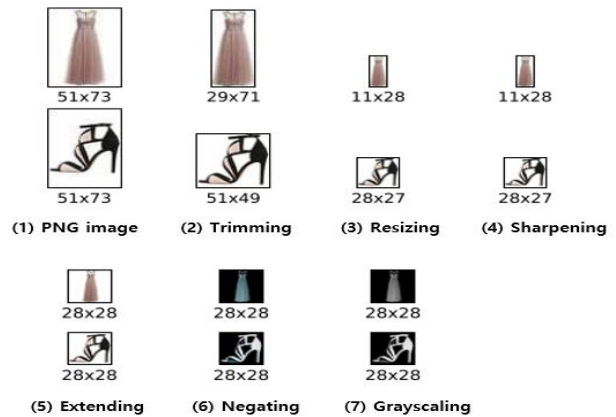


그림 5. fashion-MNIST의 생성 과정
 Fig. 5. The creation process of fashion-MNIST

위의 그림 5에서 fashion-MNIST의 생성 과정은 다음과 같이 7 단계로 구성된다.

- 1 단계: 입력된 이미지를 PNG 이미지로 변환한다.
- 2 단계: 이미지의 가장자리를 트리밍(trimming) 처리한다.
- 3 단계: 이미지 픽셀을 서브샘플링(subsampling)하여 이미지에서 가장 긴 가장자리 크기를 28픽셀로 조정한다.
- 4 단계: 가우시안 연산자를 사용하여 이미지의 윤곽선이 선명하게 보이도록 한다.
- 5 단계: 가장 짧은 가장자리를 28픽셀로 확장하고 이미지를 중앙에 배치한다.
- 6 단계: 색상을 보색으로 반전한다.
- 7 단계: 이미지를 8비트 그레이스케일 픽셀로 변환한다.

3-2 딥러닝으로 분류된 패션 아이템 매칭을 위한 데이터셋 설계

패션 데이터셋 fashion-MNIST를 대상으로 아래의 그림 6과 같이 티셔츠/탑(T-shirt/top), 바지(Trouser), 풀오버(Pullover), 드레스(Dress), 코트(Coat), 샌들(Sandal), 셔츠(Shirt), 스니커즈(Sneaker), 가방(Bag), 앵클 부츠(Ankle boot)에 해당하는 총 10개의 패션 아이템으로 분류할 수 있도록 레이블링한다.

Label	Description	Examples
0	T-Shirt/Top	
1	Trouser	
2	Pullover	
3	Dress	
4	Coat	
5	Sandals	
6	Shirt	
7	Sneaker	
8	Bag	
9	Ankle boots	

그림 6. Fashion-MNIST의 레이블링 시각화
 Fig. 6. Visualization of labeling in Fashion-MNIST

본 연구에서는 패션 데이터셋 fashion-MNIST를 기반으로 딥러닝을 수행한 후 분류된 10가지 종류의 패션 아이템을 기반으로 일상에서 활용할 수 있는 패션 매칭 서비스를 완성하기 위해서 다음과 같이 A ~ G까지 7가지 유형의 패션 아이템 매칭을 위한 데이터셋을 설계한다.

- A : 티셔츠/탑(T-shirt/top)와 바지(Trouser) 매칭
- B : 풀오버(Pullover)와 바지(Trouser) 매칭
- C : 코트(Coat)와 바지(Trouser) 매칭
- D : 셔츠(Shirt)와 바지(Trouser) 매칭
- E : 드레스(Dress)와 샌들(Sandal) 매칭
- F : 바지(Trouser)와 스니커즈(Sneaker) 매칭
- G : 바지(Trouser)와 앵클 부츠(Ankle boot) 매칭

IV. 패션 아이템 매칭 데이터셋 구축

4-1 CNN 기반의 패션 아이템 분류를 위한 딥러닝 모델

본 연구에서는 패션 이미지 데이터셋 fashion-MNIST를 기반으로 10개의 패션 아이템으로 레이블링하여 분류하는 딥러닝 모델을 tensorflow 2.0기반의 koras를 이용하여 설계하여 실행하였다. fashion-MNIST의 이미지 데이터 값이 0 부터 255 사이의 정수이므로 0 부터 1 사이의 범위로 정규화(normalization)하여 신경망에 입력하였다. CNN을 이용한 딥러닝 과정은 그림 7 과 같이 특징 추출 신경망과 분류 신경망을 직렬로 연결한 형태로 구성하였다.

그림 7에서 특징 추출 신경망의 경우, 컨볼루션 레이어(Convolution Layer)와 풀링 레이어(Pooling Layer)를 반복시켜 이미지의 특징에 해당하는 에지를 추출할 수 있도록 하고, 분류 신경망의 첫 번째 계층인 플래튼 레이어(Flatten Layer)에서는 28 x 28 픽셀의 2차원 배열 이미지를 784 픽셀의 1차원 배열로 변환한다. 그리고 fashion-MNIST에서 10개의 패션 아이템 중 하나로 인식시키는 학습을 위해서 텐스 레이어(Dense Layer)와 과적합 방지를 위한 드롭아웃 레이어(Dropout Layer)를 이용하여 딥러닝 모델을 구성한다.

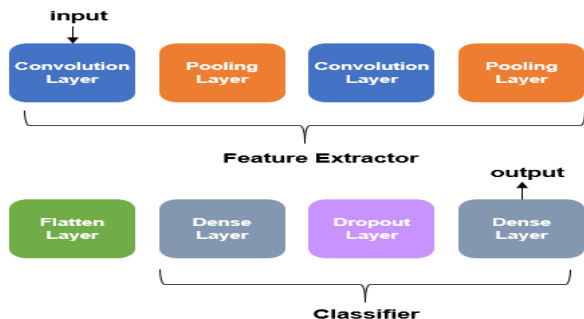


그림 7. CNN을 이용한 딥러닝 과정 흐름도
Fig. 7. Flowchart of deep learning process using CNN

본 연구에서는 딥러닝 실행 시 과적합을 방지하기 위해서 구성된 딥러닝 모델을 이용하여 fashion-MNIST 훈련 데이터셋 60,000개 중에서 75%에 해당하는 45,000개를 훈련(training) 데이터셋으로, 나머지 25%에 해당하는 15,000개를 검증(validation) 데이터셋으로 재구성하여 총 15회의 에폭(epoch)으로 딥러닝 모델을 학습시켰으며 소스 코드는 아래와 같다.

```
# 특징 추출을 위한 모델링
cnn_model = Sequential()
cnn_model.add(Conv2D(filters=32, kernel_size=(3, 3),
                    input_shape=(28,28,1), activation='relu', padding='same'))
cnn_model.add(BatchNormalization())
cnn_model.add(MaxPooling2D(pool_size=(2, 2)))
cnn_model.add(Dropout(0.2))

cnn_model.add(Conv2D(filters=64, kernel_size=(3, 3),
                    input_shape=(28,28,1), activation='relu', padding='same'))
cnn_model.add(BatchNormalization())
cnn_model.add(MaxPooling2D(pool_size=(2, 2)))
cnn_model.add(Dropout(0.2))

# 분류를 위한 모델링
cnn_model.add(Flatten())
cnn_model.add(Dense(units=128, activation='relu'))
cnn_model.add(Dropout(0.2))
cnn_model.add(Dense(units=10, activation='softmax'))

# 딥러닝 모델 컴파일
METRICS = [ 'accuracy',
            keras.metrics.Precision(name='precision'),
            keras.metrics.Recall(name='recall')]
cnn_model.compile(loss='sparse_categorical_crossentropy',
                  optimizer='adam', metrics=METRICS)

# 딥러닝 모델 학습
epochs = 15
batch_size = 32
history = cnn_model.fit(
    X_train, y_train,
    epochs=epochs,
    verbose=1,
    validation_data=(X_validate, y_validate))
```

4-2 CNN 기반의 패션 아이템 분류를 위한 딥러닝 실행 결과

본 연구에서 설계한 CNN을 이용하여 10가지 패션 아이템 분류를 위한 딥러닝 모델을 실행한 결과 그림 8 과 같은 결과가 출력되었다.

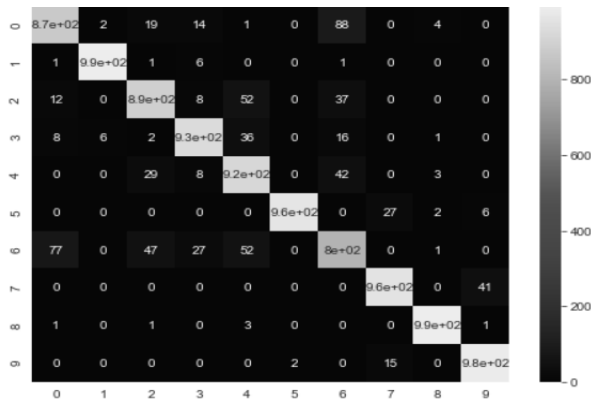


그림 8. 딥러닝 실행 후 예측 상황
 Fig. 8. Prediction situation after running deep learning

그림 8의 x축은 지도 학습(supervised learning)에서 레이블링으로 사용된 값을 의미하고, y축은 딥러닝 수행 결과 예측된 값을 의미한다. 그림 8에서 0으로 레이블링된 값을 6으로 잘못 예측한 경우는 77회, 2로 레이블링된 값을 6으로 잘못 예측한 경우는 47회에 해당함을 의미한다. 대각선 패턴은 레이블링된 값과 예측된 값이 일치한 횟수를 의미한다.

본 연구에서 설계한 CNN을 이용하여 10가지 패션 아이템 분류를 위한 딥러닝 모델을 실행한 결과, 표 1의 검증 데이터셋의 정확도(val_accuracy)를 통해서 알 수 있듯이 과적합을 방지하면서 적절하게 학습이 수행되었음이 확인되었다. 15번째 에폭(epoch)을 실행한 결과 훈련 데이터셋은 손실(loss) 0.1165, 정확도(accuracy) 0.9572를 산출하였고, 검증 데이터셋은 손실(loss) 0.2227, 정확도(accuracy) 0.9364로 나타났다.

표 1. 훈련 데이터셋과 검증 데이터셋을 기반으로 CNN을 이용한 딥러닝 실행 후 손실 및 정확도

Table 1. Loss and accuracy after deep learning execution using CNN based on the training dataset and validation dataset

Epoch	loss	accuracy	val_loss	val_accuracy
1	0.4840	0.8281	0.3383	0.8816
2	0.3231	0.8810	0.2882	0.9004
3	0.2844	0.8960	0.3990	0.8683
4	0.2547	0.9058	0.2510	0.9089
5	0.2297	0.9147	0.2535	0.9096
6	0.2129	0.9216	0.2385	0.9184
7	0.1968	0.9270	0.2353	0.9249
8	0.1830	0.9320	0.2276	0.9258
9	0.1664	0.9383	0.2420	0.9249
10	0.1546	0.9422	0.2271	0.9287
11	0.1484	0.9429	0.2177	0.9335
12	0.1388	0.9489	0.2256	0.9300
13	0.1330	0.9502	0.2362	0.9320
14	0.1229	0.9532	0.2361	0.9264
15	0.1165	0.9572	0.2227	0.9364

위의 표 1을 대상으로 손실 및 정확도를 그래프로 시각화하면 그림 9와 같다.

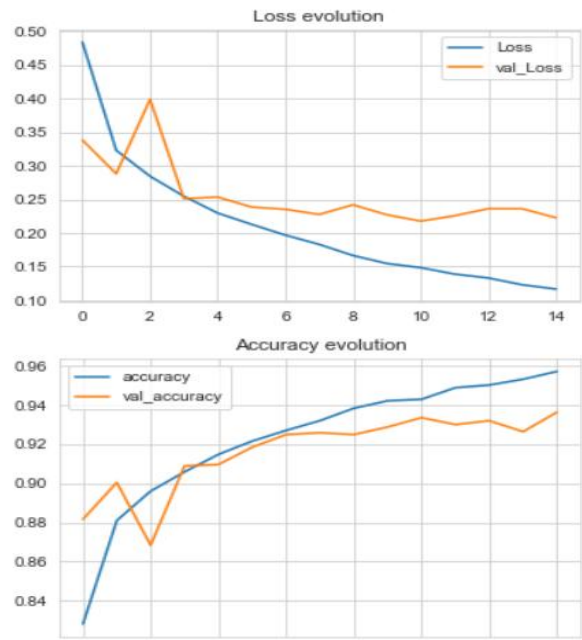


그림 9. 딥러닝 실행 후 손실 및 정확도
 Fig. 9. Loss and accuracy after executing deep learning

기존 연구 [16]에서 제시한 학습률 최적화 알고리즘을 사용할 경우 92.87% 정확도를 출력했던 것에 비하여 본 연구에서 설계한 특징 추출 및 분류를 위한 CNN 딥러닝 모델을 기반으로 딥러닝을 수행한 결과 2.85%가 향상된 95.72%까지 정확도를 획득할 수 있었다.

4-3 패션 아이템 매칭 데이터셋 구축

본 연구에서는 딥러닝 기반의 CNN을 이용하여 훈련 데이터셋 학습 정확도 0.9572, 검증 데이터셋 학습 정확도 0.9364를 가진 패션 아이템을 10가지로 분류하여 인식할 수 있는 시스템을 구현하였다. 구현된 패션 아이템 분류를 위한 딥러닝 시스템을 기반으로 학습에 사용하지 않은 fashion-MNIST의 테스트 데이터셋 10,000개를 사용하여 랜덤으로 데이터를 선택하여 3-2절에서 설계한 7가지 유형의 패션 아이템 매칭을 위한 데이터셋 구축을 그림 10과 같이 구축하였다.

패션 매칭 데이터셋 구축은 python에서 제공하는 리스트(list) 형식의 데이터 타입을 이용하였다. 패션 매칭을 위해서 A 유형의 경우, 티셔츠/탑(T-shirt/top)를 선택하면 매칭된 카테고리에 해당하는 바지(Trouser) 리스트에 등록된 이미지들이 보이도록 구축되었다. 반면에 바지(Trouser)를 선택하면 티셔츠/탑(T-shirt/top) 리스트에 등록된 이미지가 보이도록 구축되었다. G 유형의 경우, 바지(Trouser)를 선택하면 앵클 부츠(Ankle boot) 리스트에 등록된 이미지가 보이도록 구축되었고, 앵클 부츠(Ankle boot)를 선택하면 바지(Trouser) 리스트에 등록된 이미지가 보이도록 구축되었다.

V. 결 론

본 연구에서는 fashion-MNIST를 이용하여 티셔츠/탑(T-shirt/top), 바지(Trouser), 풀오버(Pullover), 드레스(Dress), 코트(Coat), 샌들(Sandal), 셔츠(Shirt), 스니커즈(Sneaker), 가방(Bag), 앵글 부츠(Ankle boot)에 해당하는 10가지 패션 아이템을 레이블로 셋팅하여 CNN을 이용한 딥러닝 모델을 설계하고 지도 학습을 실행하였다. 그 결과 패션 이미지를 입력하면 10가지 종류 중 하나의 패션 아이템으로 자동 분류되는 인공지능 시스템이 구축되었다. 이렇게 구축된 인공지능 시스템을 통해서 10가지 종류 중 하나로 분류가 가능해진 패션 아이템들을 7가지 패션 매칭 요소로 결합할 수 있는 데이터셋을 구성하여 향후 인공지능 패션 아이템 매칭 서비스 제공을 위한 기반을 마련하였다. 본 연구에서 제안된 7가지 패션 매칭 요소는 티셔츠/탑(T-shirt/top)와 바지(Trouser) 매칭, 풀오버(Pullover)와 바지(Trouser) 매칭, 코트(Coat)와 바지(Trouser) 매칭, 셔츠(Shirt)와 바지(Trouser) 매칭, 드레스(Dress)와 샌들(Sandal) 매칭, 바지(Trouser)와 스니커즈(Sneaker) 매칭, 바지(Trouser)와 앵글 부츠(Ankle boot) 매칭으로 구성되었다.

현재 본 연구는 패션 매칭을 위한 데이터셋을 10개의 카테고리 분류하고, 분류된 결과에 따라 다시 상의와 하의로 결합할 수 있도록 7가지 패션 매칭 요소로 재분류하는 것으로 마무리되었다. 향후 장소와 목적에 맞는 다양한 패션 매칭 서비스를 제공하기 위해서 색상, 계절, 재질, 트렌드, 브랜드명 등의 여러 가지 요소들을 포함하여 개선된 패션 매칭 딥러닝을 수행하여 개인의 다양한 패션 욕구를 충족시켜줄 수 있는 인공지능 패션 매칭 서비스가 될 수 있도록 연구를 확장할 계획이다.

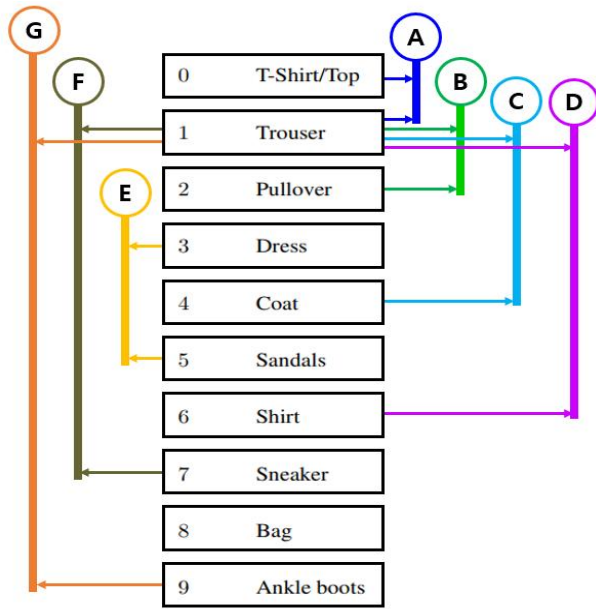


그림 10. 패션 아이템 결합을 위한 데이터셋 프레임워크
 Fig. 10. Dataset framework for combining fashion items

그림 11은 테스트 데이터셋 10,000개를 이용하여 패션 아이템 10가지 종류로 분류한 결과 중 일부를 보여주고 있다. 실제 클래스(True Class)는 테스트 데이터셋에서 10가지 종류로 레이블링 된 패션 아이템 번호를 의미하고, 예측 클래스(Prediction Class)는 딥러닝으로 학습된 시스템에서 예측한 패션 아이템 번호를 의미한다.



그림 11. 10 가지 패션 아이템으로 분류된 이미지
 Fig. 11. Classified image with 10 fashion item types

참고문헌

- [1] H. M. Gang, "A Study on the Trend of Knit Industrial Fashion Using Big Data and Artificial Intelligence," in Proceeding of the 8th Korean Society Of Knit Design, pp. 41-42, 2020.
- [2] H. J. Lee, S. J. Lee, M. H. Choi, S. Y. Kim, and H. Y. Moon, "Fashion Image Searching Website based on Deep Learning Image Classification," The Journal of practical engineering education, Vol. 11, No. 2, pp. 175-180, November 2019.
- [3] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in Proceeding of the IEEE, pp. 2278-2324, 1998
- [4] H. Xiao, K. Rasul and R. Vollgraf, Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms, Zalando of Berlin, Zalando Research, 2017.
- [5] K. Mohammed, A. Ahmed, and M. Hadeer, "Classification of Garments from Fashion MNIST Dataset Using CNN LeNet-5 Architecture," in Proceedings of the International

Conference on Innovative Trends in Communication and Computer Engineering, pp. 238-243, 2020.

- [6] I. Cristian, and V. E. Neagoie, "A Deep CNN Approach with Transfer Learning for Image Recognition," in Proceedings of the 11th International Conference on Electronics, Computers and Artificial Intelligence, pp. 1-6, 2019.
- [7] N. Karthika, B. Janet, and H. Shukla, "A Novel Deep Neural Network Model for Image Classification," The Journal of Engineering and Advanced Technology, Vol. 8, No. 6, pp. 3241-3249, August 2019.
- [8] H. S. Kim, and H. D. Kim, "Coordination of Smart Costume based on Complementary Colors using Image Segmentation," The Journal of Digital Contents Society, Vol. 19, No. 8, pp. 1453-18462, August 2018.
- [9] I. J. Goodfellow, P. A. Jean, M. Mirza, B. Xu, W. F. David, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in Proceedings of the 27th International Conference on Neural Information Processing Systems, pp. 2672-2680, 2014.
- [10] L. A. Gatys, A. S. Ecker, and M. Bethge, A Neural Algorithm of Artistic Style, Journal of Vision, 2015.
- [11] L. Gaopeng, S. Zhao, and Z. Yun, "ISAR Target Recognition Using Pix2pix Network Derived from cGAN," in Proceedings of the International Radar Conference, pp. 1-4, September 2019.
- [12] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1125-1134, 2017.
- [13] B. B. Traore, K. F. Bernard, and F. Tangara, "Deep convolution neural network for image recognition," The Journal of Ecological Informatics, Vol. 48, pp. 257-268, October 2018.
- [14] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1-21, 2014.
- [15] R. Girshick, "Fast R-CNN" in Proceedings of the IEEE international conference on computer vision. pp. 1440-1448, 2015.
- [16] S. Y. Shin, G. Wang, and K. S. Shin, "Fashion Clothing Image Classification Based on Improved CNN" in Proceedings of the The Korea Institute of Information and Communication Engineering. pp. 83-84, 2020.



김희숙(Hye-Suk Kim)

1999년 : 전남대학교 대학원
전산통계학과 (이학석사)

2009년 : 전남대학교 대학원
전산학과 (이학박사)

2018년~ 현재 : GIST(Gwangju Institute of Science and Technology) Lecture Professor

※관심분야 : 인공지능, 딥러닝, 머신러닝, 데이터과학, 영상처리, 웹콘텐츠, UX, UI 등