

유통 서비스 혁신을 위한 인공지능 이미지 인식 기술 개발 및 적용*

김 형 진(제1저자)

명지대학교 미래융합경영대학 미래융합경영학과(교수)

권 보 람(공동저자)

서울대학교 경영대학 경영학과(교수)

한 진 영(공동저자)

중앙대학교 교양대학(교수)

손 영 우(공동저자)

고려대학교 일반대학원 첨단기술비즈니스학과(교수)

이 지 은(교신저자)

한양사이버대학교 경영정보·AI비즈니스학과(교수)

Development and Application of Artificial Intelligence Image Recognition Technology for Innovation in Retail Services

Kim, Hyungjin(First Author)

College of Future Convergence, Myongji University(Professor)

Kwon, Boram(Co Author)

College of Business Administration, Seoul National University(Professor)

* 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2021R1G1A1094469).

Han, Jinyoung(Co Author)

College of General Education, Chung-Ang University(Professor)

Sohn, Youngwoo(Co Author)

Department of Future Science and Technology Business, Graduate School, Korea
University(Professor)

Lee, Ji-Eun(Corresponding Author)

Department of MIS & AI Business, Hanyang Cyber University(Professor)

Abstract

In demographic shifts, changes in consumer patterns, and heightened uncertainty, traditional retail enterprises are experiencing a sense of crisis. Nevertheless, innovative business models and digital transformations have propelled some retail enterprises to achieve remarkable growth and success by maximizing customer convenience and enhancing operational efficiency through diverse technologies. This study aims to develop partial image recognition technology applicable to the situation of unmanned convenience store kiosks, grounded in the Retail Tech paradigm. To address limitations in previous research concerning the technical and empirical approaches to partial image matching, this study proposes deep learning-based partial image matching technology utilizing semantic matching and cross-verification to improve the accuracy of image recognition technology. The developed algorithm achieves a accuracy of over 95% in various visual scenarios, contributing significantly to the advancement of image processing and recognition technologies. In addition, the developed technology, designed to consider most noise situations in the application process of image recognition technology, is applicable to various situations, including the logistics/distribution field where this study was conducted.

Keywords : Object Recognition, Partial Image Matching, Potential Similarity Measurement, Retail Service Innovation, Case Study

접수일(2023년 11월 24일), 수정일(2023년 12월 14일), 게재확정일(2023년 12월 16일)

I. 서 론

코로나 19로 야기된 팬데믹(Pandemic)과 디지털 전환은 인류의 삶과 기업 경영에 큰 변화를 불러왔다. 일, 학업, 영화관람 등 일상생활의 상당 부분이 온라인상에서 이뤄지고, 플랫폼의 영향력은 커지고 있으며, 공급 과잉 시대가 되면서 자신에게 꼭 맞는 제품이나 서비스를 정확하게 추천해주는 ‘큐레이션 서비스’나 나만을 위한 ‘커스텀 서비스’를 선호하는 소비자가 늘고 있다. 키오스크나 무인 매장을 이용하는 것이 일반화되고 대형 할인점에서는 무인 계산대가 확대되고 있다.

오래 전부터 온라인 시장이 급속도로 성장해오고 있었지만, 팬데믹을 계기로 온라인 서비스의 영향력은 더욱 커졌고 오프라인 기반의 서비스도 상당 부분 온라인으로 전환되고 있으며 엔데믹(Endemic) 상황에서도 셀프서비스와 비대면 방식을 선호하는 고객이 늘고 있다.

인구구조 변화와 소비 패턴의 변화, 불확실성의 증가로 서비스 산업의 경영환경 역시 큰 변화를 겪고 있으며, 기업들은 위기감을 느끼고 있다(박정수 2021). 하지만 그 가운데서도 혁신적인 비즈니스 모델과 디지털 전환으로 큰 폭의 성과와 괄목할만한 성장을 기록하는 기업들이 있다. 이들 기업은 고객의 편의성을 극대화하고 점포 및 플랫폼의 운영 효율성을 높이기 위해 다양한 기술을 적극 활용한다는 공통점이 있다.

서비스 경영환경은 AI, 빅데이터, 로봇 등 신기술의 도입과 함께 서비스 혁신을 요구하고 있다(박정수 2021). 기술혁신은 기술 자체의 변화에 그치지 않고 생산·유통·소비 방식을 바꾸고 사회의 모습을 변화시켜 인류의 생활방식 전체에 변화를 가져온다(이향숙, 이찬구, 2021). 생성형 AI의 등장으로 일반인들에게 더욱 가까워진 인공지능(AI)과 빅데이터, VR, AR, MR, XR과 사물인터넷(IoT) 기술은 혁신을 촉진하고 있다. 이러한 변화를 일컬어 ‘유통 4.0’으로 지칭하고 있는데, 4차 산업혁명 기반 기술이 유통산업에 활용되어 유통서비스의 초연결화, 초지능, 초실감을 가능하게 하는 현상으로 국내외 기업들은 유통 서비스를 리테일테크(Retail Tech) 중심으로 발전시키고 노력하고 있다(정우진, 김정연, 2023). ‘유통 4.0’ 시대를 맞아 유통 산업에는 O2O 서비스, 옴니채널화, 유통과 물류의 융합 등 온·오프라인과, 산업, 업태 간 경계가 붕괴되고, 유통산업 내 가치창출의 원천이 상품·서비스 거래에서 생산/소비에 대한

지식과 정보로 전환되는 등 기술혁신에 직접 영향을 받는 큰 변화가 이루어지고 있다. 또한, 국경 간 장벽이 무너짐에 따라 산업 내에서 글로벌 기업의 영향력이 커지고 있다(이정섭 외, 2021). 기술을 활용해 쇼핑의 즐거움과 편의성, 접근성을 높이는 것은 기업의 경쟁력을 강화하는 근원이 될 것이다.

본 연구는 리테일테크(Retail Tech)를 바탕으로 비대면 서비스의 확산과 서비스의 디지털화가 강화되고 있는 가운데 무인 편의점 키오스크의 상황에 적용할 수 있는 이미지 인식 기술을 개발하고자 한다. 특히, 본 연구는 전체 이미지를 활용하기 어려운 현실을 고려하여, 부분 이미지 매칭(Partial Image Matching) 기술에 초점을 맞춘다. 즉, 이미지 인식 기술을 활용하여 다양한 시각적 노이즈 노출 상황에서 판별 정확도가 높은 딥러닝 기반 부분 이미지 매칭 기술을 개발함으로써 이미지 인식 기술의 발전만 아니라 소비자의 편의성과 기업의 효율성에 실질적으로 기여하고자 한다. 또한, 본 연구를 통해 개발된 기술은 이미지의 부분 정보를 활용하여 이미지 분할 없이 회전 및 배경 불변성을 제공함에 따라 다양한 시나리오에 활용할 수 있을 것이다. 이러한 연구 결과는 서비스 경영 분야에서의 기술 기반 혁신과 효율성 향상에 기여할 것으로 기대된다.

따라서 본 연구는 기반 부분 이미지 매칭 기술을 활용하여 이미지 인식 기술이 현실에서 활용되는 과정에서 직면하는 여러 오류를 개선하는 데에 목적이 있다. 이를 위해, 무인 편의점 키오스크 개발과정에 부분 이미지 매칭 기술을 적용하여 실질적 활용 가능성을 검증하였으며, 이를 기반으로 유통 서비스 혁신을 위한 이미지 인식 기술의 범용적 활용 가능성과 접근 전략을 제시하고자 한다.

II. 이론적 배경

2.1 유통 서비스의 기술적 혁신

유통 분야는 리테일테크를 통해 서비스 혁신활동이 활발하게 일어나고 있다(이제영, 김정호, 2021). 리테일테크는 리테일(Retail)과 기술(Technology)을 결합한 신조어로, 점포 운영의 효율성과 소비자의 사용 편의성을 높이는데 기여하는 모든 것을 포괄한다. 리테일테크의 대표적인 예로 무인점포를 들 수 있다. 해외에서는 아마존 고(Amazon go)의 'Just Walk Out'이 대표적이다. 2016년 시애틀에 오픈한 아마존 고는 고객이 매장에서 물건을 고른 후 매장을 나서면 고객의 신용카드로 물건값이 자동 결제되는 방식으로 운영된다. 즉, 계산대에서 직원이 바코드를 찍어 계산하는 방식을 자동화한 것으로, 점포 운영을 위해 필수적이었던 계산 인력과 계산대를 매장에서 없앤

혁신적인 모델이다(이성훈, 이동우, 2018). 소비자가 매장에 들어가서 원하는 상품을 고른 후 계산대를 거치지 않고 곧바로 매장을 나갈 수 있게 한다는 의미에서 ‘Just Walk Out’ 모델이라 부른다. ‘Just Walk Out’을 구현하기 위해 아마존은 수십 대의 카메라를 매장에 장착하고 카메라 센서와 컴퓨터 비전, 딥러닝 기술을 활용해 매장 내 고객들의 움직임을 실시간으로 파악하고 있다.

아마존 고로 대표되는 셀프 체크아웃의 진화는 소비자와 기업에게 이점을 제공한다. 소비자로서는 대기시간 감소와 사용 편의성이 장점이며, 소매업체의 경우 대기 줄을 줄여 빠른 서비스를 제공함으로써 고객의 만족과 신뢰를 강화하여 수익을 강화할 수 있다(Polacco & Backes, 2018). 이처럼 아마존 고는 기술혁신을 통한 소매업의 미래를 보여주었지만 현재 정체기에 빠져 있는 것으로 보인다. 그러나 기술 불확실성을 제거할 경우 미래의 가장 일반적인 소매업의 형태로 자리 잡을 것으로 예상된다.

국내에서는 이마트24가 아마존 고와 유사한 형태의 무인 편의점을 2021년도에 도입하였다. ‘이마트24 스마트 코엑스점’도 고객이 물건을 가지고 나가면 자동으로 결제되는 방식으로 운영되었는데 여기에 인공지능, 컴퓨터 비전, 음성 인식, 클라우드 POS(Point-of-Purchase) 등의 리테일테크가 적용되었다. 고객이 매장 앞 키오스크를 통해 신용·체크카드로 인증을 받은 뒤 출입 QR코드를 받아 입장하고 매장에서 원하는 상품을 들고 나오면 자동으로 결제되는 방식으로, 이러한 방식을 ‘저스트 픽앤고(Just Pick&Go)’라고 칭했다. 이남경과 박현준(2021)은 이마트24 무인 매장의 차별적 강점이 고객의 이용의도에 영향을 미치는지를 실증한 결과, 편리성, 신속성, 신뢰성, 오락성, 심미성이 고객이 느끼는 효익임을 확인했다. 현재는 100% 무인 매장보다는 주간은 직원이 매장을 관리하고 심야에는 셀프계산 시스템으로 운영하는 하이브리드 형태가 확대되고 있다. 이처럼 무인 편의점 운영을 위해서는 다양한 혁신기술을 필요로 하며, 특히 객체를 정확하게 인식하는 컴퓨터 비전 기술이 요구된다.

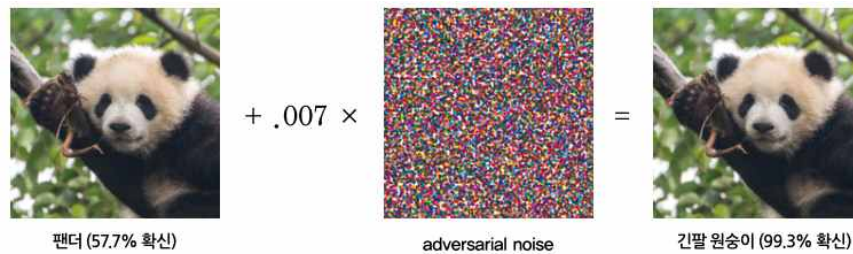
2.2 인공지능 이미지 인식 기술의 적용

컴퓨터 비전(Computer Vision) 기술 중 하나인 인공지능 이미지 인식 기술은 이미지 내 특정 사물을 분류(Classification), 검출(Detection), 픽셀단위로 식별하여 분할(Segmentation)하는 기술로(Bao et al., 2019; Liu et al., 2018; Saritha et al., 2019; Zellers et al., 2019), ILSVRC(ImageNet Large Scale Visual Recognition Challenge)의 기록을 살펴보면 이미 사람의 인식률(94.90%)을 추월하여 최근 98.7%를 상회하는 수준으로 발전하여 다양한 영역에 적용되어 왔다(송재민 외, 2020; 이주열, 2020). 딥러닝의 발전으로 인공지능 이미지 인식 수준은 정확도 측면에서 사람의

수준을 뛰어넘기도 하였지만, 제조, 물류/유통, 의료, 자율주행, 안전 등 실제 산업 영역에 적용되면서부터 한계를 보이고 있다(이주열, 2020).

실제 상황에서는 전체 이미지가 활용되기 어려운 경우가 많아, 부분 이미지 매칭기술을 위한 정확도 제고가 필요하나 이에 대한 기술적, 실증적 접근이 부재한 것이 현실이다. 이런 상황에서, 사람이 판단하기에는 문제가 없으나 인공지능망의 판정을 교란시키는 적대적 예제(Adversarial examples)의 문제가 발생하기도 하며(Eykholt et al., 2018; Goodfellow et al., 2015; Thys et al., 2019), 블러링(Blurring), 포깅(Fogging)등 현실에서 흔하게 발생할 수 있는 노이즈에 대한 대응 문제(Hendrycks & Dietterich 2019), 측광학적 편차(Photometric variation), 기하학적 편차(Geometric variation)문제(Kim et al., 2017) 등이 발생하여 실제 상황에서는 기술의 정확도가 떨어져 활용이 어려운 상황이 지속되고 있다.

<그림 1> 적대적 예제(Adversarial Examples)의 사례



(출처: Goodfellow et al. 2015)

2.2.1 부분 이미지 매칭을 위한 잠재적 유사성 측정의 필요성

이러한 한계에서 실무적 적용을 위한 여러 시도가 있었으며, 대표적인 사례가 아마존의 계산대 없는 무인 점포 아마존 고이다. 아마존 고의 고객은 매장에 들어가서 자신이 원하는 제품을 장바구니에 담으면 자동으로 고객의 앱의 가상 장바구니에 해당 제품을 담고 쇼핑이 끝나 체크아웃 레인을 통과하면 자동으로 결제가 가능하다. 아마존 고는 고객을 식별하고 쇼핑 동선에 맞춰 이들을 추적하기 위한 목적과 선택한 제품을 정확하게 식별하기 위한 이미지 인식 기술을 활용하고 있으며, 결제가 연동되어 있어 사실상 100% 수준의 정확도를 보이는 이미지 인식 기술이 필요한 상황이다.

하지만, 실제 상황에서 발생할 수밖에 없는 다양한 노이즈(e.g., 고객이 제품을 집을 때 동일 제품이라 하더라도 선반 내의 배치 위치, 제품을 집는 각도 및 포장의 구겨짐 등이 발생하는 기하학적 편차, 매장 내의 조도 등에 따른 측광학적 편차, 제품과 고객이 아닌 배경의 패턴을 잡아 인식하려는 적대적 예제, 냉장/냉동고에서의 온도 차로 인한 포깅 등에 대응하기 위해 엄청난 양의 카메라를 활용함과 동시에 선반 내의

무게 센서 설치(Sensor fusion), 위치별로 미리 정해진 제품의 배치 등 실질적으로는 이미지 인식 기술을 보조역할로만 활용하고 있다.

이미지 인식 기술로만 이를 해결할 수 있다고 하더라도 수많은 카메라가 필요하고 이로 인한 계산량 증가에 따른 비용문제가 발생할 뿐만 아니라, 거래(Transaction) 시점에 발생하는 이미지 데이터 간의 동기화(Synchronization) 문제로 인한 높은 오류 발생 가능성이 존재한다. 따라서, 인공지능 이미지 인식 기술이 실제 상황에서 온전하게 활용되기 위해서는 부분 이미지 매칭을 위한 잠재적 유사성 측정(Potential similarity measuring)이 가능해야 하며 이를 통한 실제 상황에서의 판별함수 정확도를 높일 수 있다.

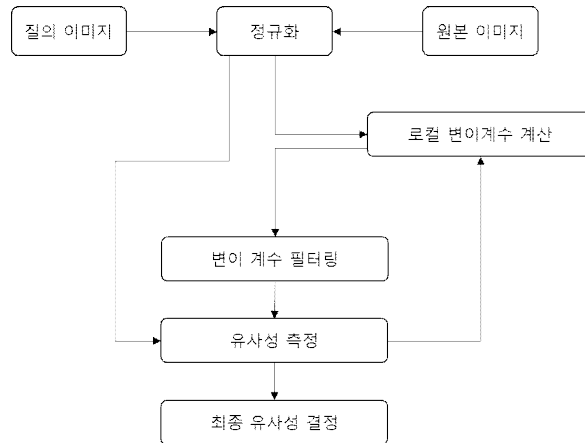
III. 기술개발 프레임워크 및 사례적용

3.1 부분 이미지 매칭을 활용한 기술개발 프레임워크

이미지의 부분 정보를 활용하여 가장 유사도가 높은 제품을 찾아내기 위해 질의 이미지(Query image)를 바탕으로 저장된 이미지들의 전체 이미지 간의 유사도를 계산하여 가장 유사도가 높은 이미지를 제시하는 방식을 활용하였다(유채곤 외, 1999; 전민철 외, 2010).

우선, 질의 이미지와 원본 이미지의 농도치를 정규화한다(Step 1). 그리고 원본 이미지를 스캔하면서 위치별로 변이 계수를 사용하여 질의 이미지와 비교하여 질의 이미지의 변이 계수와의 차이가 임계치 내에 드는 부분에 대해서만 유사성 측정을 시도하기 위한 변이계수 필터링을 시도한다(Step 2). 마지막으로, 질의 이미지와 원본 이미지 간의 유사도 분석을 한다(Step 3). 변이 계수 필터링을 통과한 위치에 대해서는 원형 윈도우의 중심과 해당 픽셀과의 거리에 비례하여 픽셀의 농도치를 변환해 주는 작업을 통해, 농도치 히스토그램에 농도치 정보와 더불어 윈도우 내의 픽셀의 위치 정보가 포함되도록 한다. 이미지 유사성 계산을 위한 프레임워크는 <그림 2>에 도식화하였다.

<그림 2> 이미지 유사성 계산 프레임워크



3.1.1 시맨틱 정합(Semantic Correspondence) 기술의 활용

고객이 구매를 원하는 제품에 대한 이미지와 실제 학습 데이터에 존재하는 이미지 간의 정합을 활용하여 가장 유사한 영상을 검색(Retrieval)하는 기술을 적용하기 위해, 제품의 부분 혹은 전체에 대한 정합도를 구하고자 한다. 제품 이미지 인식 시, 고객이 제품을 집는 과정에서 발생할 수 있는 포장의 구겨짐, 위치에 따른 상대적 크기 차이, 등이 발생할 수 있어, 크기 및 변형에 강인한 특징 표현자를 인공지능으로 학습하여 활용하고자 한다. Kim et al.(2020)과 Sucontphunt(2020)의 연구에서 VGG-16 딥러닝 모델을 기반으로 측광학적, 기하학적 편차에 강인한 특징 표현자를 추출한 방식을 활용하여 이미지 판별을 수행하였다.

3.1.2 측광학적/기하학적 편차 문제 해결을 위한 교차검증

Kim et al.(2016)의 연구를 통해 DASC (Dense Adaptive Self-Correlation) 기술을 통해 멀티 모달(Multi-modal), 멀티 스펙트럴 조밀 대응(Multi-spectral dense correspondence) 등으로 측광학적 편차 문제의 해결이 가능할 수 있음을 판단하고 이를 활용하였다. 또한, 범위(Scale)와 회전(Rotation)과 같은 기하학적 문제를 해결하기 위해 Kim et al.(2016, 2020)이 제안한 GI-DASC(Geometry-Invariant DASC)를 활용하였다. 본 방식은 사용자가 서비스를 사용하는 상황요소에 관계없는 제품 인식이 가능하도록 한다. 또한, 존재하는 기하학적 변형을 예측할 수 있는 최적화 기술을 개발하고자 하는데, Kim et al.(2017, 2018)이 제안한 DCTM(Discrete-Continuous Transformation Matching) 프레임워크를 바탕으로 교차검증을 수행하

였다.

3.2 사례적용: 무인 편의점 키오스크

본 연구의 목표는 실제 상황에서 활용 가능한 수준으로 고도화 된(e.g., 다양한 시각적 노이즈 노출 상황에서의 판별 정확도 95% 이상) 딥러닝 기반 부분 이미지 매칭 기술의 개발에 있다. 복잡도가 높은 무인 키오스크(Kiosk) 거래에서의 결과물을 통해, 다른 상황에서도 높은 정확도로 적용이 가능한 알고리즘을 개발하고자 한다. 무인 키오스크에서의 쇼핑 상황에서는(1) 고객이 제품을 집었을 때 제품 이미지의 일부가 가려질 수 있고(Partial Image Matching), (2) 밝기차이 등과 같은 주변 환경적 요소가 인식률에 영향을 미칠 수 있으며(Photometric variation), (3) 판매되는 제품의 특성상 포장이 구겨질 수 있고(Geometric variation), (4) 제품과 고객이 아닌 배경의 패턴이 인식될 수 있고(Adversarial examples), (5) 냉장/냉동 온도 차로 인한 포깅(Fogging) 등의 문제로 인식 오류가 발생 가능성이 다양하게 존재하여, 여러 상황에 적용하기 위한 알고리즘 개발의 최적화된 환경으로 볼 수 있다. 부분 이미지 매칭 기법은 이미지의 부분 정보를 사용하며, 회전, 배경에 불변적이고, 이미지 분할이 필요 없어 다양한 상황에서의 활용 가능성이 높다. 정량적으로는 본 연구에서 설정한 성능 지표를 상회하는 결과를 목표로 하며, 동일항목을 다양한 지표로 측정하여 판별함수 정확도를 높이하고자 한다.

3.2.1 소비자 실제 구매 환경과 동일한 실험 환경 구축

모바일 앱 구동 ▶ 잠금 해제 & 쇼케이스 열기 ▶ 구매희망 제품선택 ▶ 쇼케이스 닫기 ▶ 자동결제의 일련의 과정이 자연스럽게 이어지는 방식의 스마트 키오스크 환경을 설정하였다. 사용자가 물건을 선택하는 행동(Take-out action)을 찾기 위해 스마트 키오스크 내 4개 모서리에 카메라를 설치(카메라의 각도 조정($\pm 15^\circ$)이 가능하게) 설치하였다. 카메라 4대의 시야각의 합이 고객이 제품을 꺼낼 수 있는 범위를 포함할 수 있도록 설정하였다.

3.2.2 공개 데이터를 활용한 기본 알고리즘 개발

실험상황을 만들기 전에 기본적인 알고리즘 개발을 위해 오픈되어 있는 Grocery data, CUB data, Airplane data, Stanford data 등의 공개 데이터(Public data)를 바탕으로 architecture: VGG16, VGG19, ResNET34, ResNET51, ResNET101, ResNET152, densenet121, densenet161 등의 알고리즘을 활용하였다. 실제 기술 적용 상황에서 발생할 수 있는 유사한 객체들 간의 편차로 인해 생길 수 있는 적대적

예제의 발생을 최소화하기 위해 공개 데이터 내의 유사하지만 각기 다른 9개의 상품 분류를 활용하였다. 포함된 상품은 주스(juice), 우유(milk), 귀리우유(oat-milk), 귀리요거트(oatghurt), 사워크림(sour-cream), 발효유(soured-milk), 두유(soy-milk), 두유요거트(soyoghurt), 요거트(yoghurt) 등이다.

선정한 공개 데이터(Public data)의 왼쪽 끝부터 10 픽셀(Pixel)씩 수직, 수평으로 움직여서 잘라내기(Cropping)를 진행하였다. 사진별로 약 170개의 분절화된 224 x 224 크기의 이미지를 저장하여 알고리즘 개발에 활용하였으며, 공개 데이터를 활용한 인공지능 알고리즘 개발의 중간 결과는 <표 1>과 같다.

<표 1> 인공지능 알고리즘 개발 중간 결과

	CUB Dataset	Stanford Dataset	Airplane Dataset
Resnet 101	58.9%	64.3%	67.8%
Resnet 101+STN	84%	76.2%	85.21%
Resnet 101+deformable(5c)	—	—	65.14%
Resnet101+deformable(5b)	61.9%	—	62.23%

본 연구에서는 적대적 예제를 최소화 할 수 있는 여러 객체들 중 실제 적용 가능성이 높은 우유 용기를 대상으로 AI 알고리즘을 도출했다. 이는 탐색적 연구 성격의 초기 모델로 85~88% 사이의 정확도를 확보할 수 있었다. 다만, 공개 데이터 사용의 한계(e.g., 사진의 각도 및 제약식을 사례에 맞게 조절하지 못하는 상황)를 가지고 있는 만큼, 실제 데이터 확보 및 알고리즘 개발을 통해 한계를 극복한다면 실제 상업적 활용이 가능한 알고리즘 개발이 가능할 것으로 기대되었다. 특히, 공개 데이터가 과도하게 크거나 작게 확대/축소되거나, 모호하거나, 목표물의 중앙에 초점이 맞춰지지 않은 등의 문제가 있는 것에 주목하여, 학습용 데이터를 새롭게 촬영하고 어노테이션(Annotation)하는 과정에서 주의를 기울이고자 하였으며, 이를 통해 95% 이상의 정확도를 가진 알고리즘 개발이 가능할 것으로 예상하였다.

3.2.3 실제 데이터 확보 및 성능 평가

개발된 기본 알고리즘을 고도화하기 위해 존재하는 공개 데이터 외에 자체적으로 선별한 총 50개 제품에 대한 이미지 촬영 및 어노테이션(Annotation) 작업을 수행하였다. 이미지 어노테이션을 위해 CVAT(Computer Vision Annotation Tool)을 별도로

설치하였으며, 약 8,000장의 이미지를 수집하였다. 데이터 확보 후, 이미지 데이터는 6:2:2 비율로 학습용 데이터(Training set), 검증용 데이터(Validation set), 평가용(Test set)으로 나눠 알고리즘 고도화를 수행했다. 인공지능 기반 이미지 인식 모형의 분류성능을 평가하기 위해 판별의 정밀도(Precision)와 정확도(Accuracy)를 측정하였으며, 판별 속도를 측정하여 실제 서비스 제공 상황에서의 적용 가능성을 검증하였다.

정밀도는 모델이 참(True)이라고 분류한 것 중에 실제 참인 것의 비율로 긍정 정답률(Positive Predictive Value)라고도 불린다. 정밀도는 $TP/(TP+FP)$ 로 계산되며, 값이 클수록 정확하게 예측함을 의미한다(<표 2>).

<표 2> 정밀도 측정을 위한 분류표

		실제 정답	
		True	False
분류결과	True	True Positive(TP)	False Positive(FP)
	False	False Negative(FN)	True Negative(TN)

True Positive(TP): 실제 True인 정답을 True라고 예측(정답)

False Positive(FP): 실제 False인 정답을 True라고 예측(오답)

False Negative(FN): 실제 True인 정답을 False라고 예측(오답)

True Negative(TN): 실제 False인 정답을 False라고 예측(정답)

정확도는 Top-1과 Top-5으로 측정하였으며 Top-1 정확도는 이미지를 분류하는 과정에서 가장 확률 높은 클래스를 예측했을 때 이것이 정답이었다면 이러한 경우가 전체 예측 수에 비해 어느정도 존재하는지를 의미한다. Top-5 정확도는 가장 가까운 클래스를 다섯 개 뽑았을 때 그 안에 정답이 존재한 경우의 비율을 의미한다.

정밀도는 총 12번의 주기(epoch)를 수행 후 평균 99.5%를 상회하여 높은 분류 정밀도를 보였다. 정확도는 Top-1일 때 평균 0.954, Top-5일 때 0.984로 나타났다(<표 3>). 일반적으로 Top-5 분별 정확도가 0.950이 넘으면 분류능력이 상당히 좋다고 판단할 수 있다(Singh et al., 2022). 판별 속도는 초당 약 9.0장(1024*384)의 이미지 인식 처리 성능을 보여 실제 서비스 운영 상황에서 활용이 가능한 수준임을 확인하였다.

<표 3> 이미지 판별 정확도(Top-1, Top-5)

성능 평가 시험 데이터	이미지 수	Top-1	Top-5
test_230920_1	244	0.947	0.971
test_230920_2	80	1.000	1.000
test_230920_3	80	1.000	1.000
test_230920_4	54	0.944	0.981
test_230920_5	102	0.922	0.980
test_230920_6	40	0.975	1.000
test_230920_7	97	0.887	0.959
평균		0.954	0.984

IV. 결론과 시사점

4.1 연구결과 요약 및 시사점

본 연구는 스마트 키오스크 쇼핑 상황에서 활용 가능한 딥러닝 기반 부분 이미지 매칭 기술을 개발했다. 수집된 데이터를 기반으로 본 연구에서 개발된 기술은 다양한 시각적 노이즈 상황에서 높은 판별 정확도(95%)를 제공하며, 높은 분별 정확도 (Top-1 평균: 0.954, Top-5 평균: 0.984)와 빠른 판별 속도(초당 약 9.0장)를 보여 준다.

본 연구는 아마존 고와 이마트24의 무인 점포 도입 사례와 같이 유통 분야의 서비스 혁신이 빠르게 증가하는 가운데, 소비자에게 편의점을 제공하고 기업에는 효율성을 증대시킬 수 있는 스마트 키오스크라는 중요한 트렌드를 검토했다는 점에서 의미가 있다.

또한, 본 연구는 실제 상황에서 전체 이미지 활용이 어렵고, 부분 이미지 매칭 기술에 대한 기술적 및 실증적 접근이 부족하다는 이전 연구의 한계를 보충한다. 이러한 한계를 극복하기 위해, 본 연구는 이미지의 부분 정보를 활용하여 가장 유사한 제품을 찾기 위한 새로운 방법을 제안한다. 구체적으로, 시멘틱 정합 기술과 교차 검증을 활용하여 이미지 인식 기술의 정확성을 향상시킴으로써 적대적 예제, 노이즈, 기하학적 변형 등 정확한 이미지 판별을 어렵게 만드는 다양한 상황에 적용할 수 있는 부분 이미지 매칭 기술을 개발 했다. 개발된 기술은 다양한 시각적 노이즈 노출 상황에서 95% 이상의 판별 정확도를 달성함으로써, 이미지 처리 및 인식 기술의 발전에 기여한다. 이는 실제 서비스 운영 중에 발생할 수 있는 다양한 환경에서도 안정적이고 정확한 이미지 인식을 가능하게 한다.

뿐만 아니라, 연구를 통해 개발된 기술은 이미지 인식 기술의 적용 과정에서 발생할 수 있는 대부분의 노이즈 상황을 고려하여 개발된 알고리즘이므로 이는 범용적으

로 활용한 기술이다. 즉, 본 연구가 수행된 물류/유통 분야만 아니라 다양한 서비스 분야에서의 응용이 가능하다. 예를 들면, 생산 과정에서 식물의 질병을 탐지하거나 생산된 과일의 등급 자동 분류, 원유 재고량 확인만 아니라 안전인식을 통한 고객 관리에도 개발된 기술은 활용 가능하다. 또한, 이러한 기술은 더 효율적인 서비스 운영을 가능하게 하고 기업 비용의 절감과 빠른 서비스 제공에도 기여할 수 있다.

본 연구는 스마트 키오스크에서의 구매 상황을 가정하였지만, 다양한 노이즈 상황을 모두 고려할 수 있어 범용 인공지능(Artificial General Intelligence) 분야 연구에 기여할 수 있다. 또한, 부분 이미지 매칭을 통해 딥러닝 기술의 경량화에 기여할 수 있을 것이다.

인간의 눈으로 판별해야 하는 대부분의 작업과 산업에서 활용이 가능한 기술적 접근이며 범죄 예방 및 사후처리 등에 있어서도 활용할 수 있어, 경제적, 산업적 파급효과가 높을 뿐만 아니라 사회문제 해결 측면에서도 기여할 수 있는 부분이 상당한 기술적용 사례이다.

4.2 한계점과 미래의 연구방향

본 연구는 스마트 키오스크 상황에서의 부분 이미지 매칭 기술에 중점을 두고 있다. 그러나 이는 단일 사례 연구로 한정되어 있어 다양한 상황에서의 시나리오를 설계하고 검증함으로써 일반화하는 것이 필요하다. 구체적으로, 제품의 특성, 환경적 요인, 고객 행동 등이 고려될 수 있다.

또한, 알고리즘의 개발 과정에서 정해진 데이터 셋을 활용했다. 이러한 한계를 극복하기 위해서는 단일 데이터 셋에서 벗어나 여러 데이터 셋을 활용하고 현실적인 환경에서 발생할 수 있는 다양한 상황을 반영한 데이터 수집이 중요하다. 이를 통해 알고리즘의 강건성과 일반화 가능성을 향상시킬 수 있을 것으로 기대된다. 초기 모델은 공개 데이터를 사용하여 개발되었으나, 추가적으로 여러 데이터 셋을 활용함으로써 알고리즘을 보다 견고하게 발전시킬 수 있을 것이다.

참고문헌

- 박정수. (2021). I. 서비스산업의 미래 모습과 전략 로드맵. *월간 KIET 산업경제*, 269, 36-46.
- 송재민, 이새봄, 박아름 (2020). 이미지 인식 기술의 산업 적용 동향 연구. *한국콘텐츠학회논문지*, 20(7), 86-96.
- 유채곤, 이성환, 최영수, 김진용, & 황치정. (1999). 부분 영상 매칭을 위한 잠재적 유사성 측정 기법. *한국정보과학회 학술발표논문집*, 26(2Ⅱ), 491-493.
- 이남경, 박현준(2021), 4차 산업혁명 시대 무인편의점의 성공요인에 대한 제언: 이마트24 셀프 스토어 사례 중심으로, *한국산업정보학회논문지*, 26(2), 73-94
- 이성훈, 이동우(2018), 무인 판매점 변화를 리드하는 ICT 기술에 대한 연구, *융합정보논문지*, 8(4), 109-114.
- 이정섭, 정진섭, 고초균(2021), 4차 산업혁명 시대, 첨단 유통기업의 사례 : 디지털 전환을 중심으로, *경영컨설팅연구*, 21(1), 335-352.
- 이제영, 김정호(2021), 유통서비스 기업의 혁신 결정요인이 혁신활동 수행과 매출 성과에 미치는 영향 : 혁신목적, 저해요인, 규제효과를 중심으로, *서비스경영학회지*, 22(2), 28-54.
- 이주열 (2020). 인공지능 이미지 인식 기술 동향. *TTA저널*, 187, 44-51
- 이향숙, 이찬구(2021), 기술혁신 개념의 새로운 탐색, *기술혁신학회지*, 24(4):777-798.
- 정우진, 김정연 (2023), 유통 4.0 관련기술의 성장성 및 수익성 기여 효과분석, *한국전자거래학회지*, 28(3):73-83.
- Bao, Y., Tang, Z., Li, H., & Zhang, Y. (2019). Computer vision and deep learning-based data anomaly detection method for structural health monitoring. *Structural Health Monitoring*, 18(2), 401-421.
- Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., ... & Song, D. (2018). Robust physical-world attacks on deep learning visual classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1625-1634).
- Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*.
- Hendrycks, D., & Dietterich, T. (2019). Benchmarking neural network robustness to common corruptions and perturbations. *arXiv preprint arXiv:1903.12261*.

- Jeon, M. C., Eo, Y. D., Han, D. Y., Kang, N. G., & Pyeon, M. W. (2010). Experiment for 3D coregistration between scanned point clouds of building using intensity and distance images. *Korean Journal of Remote Sensing*, 26(1), 39–45.
- Kim, S., Min, D., Ham, B., Do, M. N., & Sohn, K. (2016). DASC: Robust dense descriptor for multi-modal and multi-spectral correspondence estimation. *IEEE transactions on pattern analysis and machine intelligence*, 39(9), 1712–1729.
- Kim, S., Min, D., Lin, S., & Sohn, K. (2020). Dense cross-modal correspondence estimation with the deep self-correlation descriptor. *IEEE transactions on pattern analysis and machine intelligence*, 43(7), 2345–2359.
- Liu, X., Deng, Z., & Yang, Y. (2019). Recent progress in semantic image segmentation. *Artificial Intelligence Review*, 52(2), 1089–1106.
- Polacco, A., & Backes, K. (2018). The amazon go concept: Implications, applications, and sustainability. *Journal of Business and Management*, 24(1), 79–92.
- Saritha, R. R., Paul, V., & Kumar, P. G. (2019). Content based image retrieval using deep learning process. *Cluster Computing*, 22(2), 4187–4200.
- Singh, A., Saini, K., Nagar, V., Aseri, V., Sankhla, M. S., Pandit, P. P., & Chopade, R. L. (2022). Artificial intelligence in edge devices. In *Advances in Computers* (Vol. 127, pp. 437–484). Elsevier.
- Sucontphunt, T. (2021). Geometric Surface Image Prediction for Image Recognition Enhancement. In *Smart Computing and Communication: 5th International Conference, SmartCom 2020, Paris, France, December 29–31, 2020, Proceedings 5* (pp. 259–268). Springer International Publishing.
- Thys, S., Van Ranst, W., & Goedemé, T. (2019). Fooling automated surveillance cameras: adversarial patches to attack person detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*.
- Zellers, R., Bisk, Y., Farhadi, A., & Choi, Y. (2019). From recognition to cognition: Visual commonsense reasoning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 6720–6731).

*** 저자소개 ***

· **김 형 진(kimhj@mju.ac.kr)**

한국과학기술원(KAIST)에서 경영공학 박사학위를 취득하였으며, 현재 명지대학교 미래융합경영학과 조교수로 재직 중이다. 주요관심 분야는 빅데이터, 인공지능, CRM·디지털마케팅, 뉴미디어전략, 데이터리터러시, 정보보안관리 등이다.

· **권 보 람(brkwon@snu.ac.kr)**

서울대학교에서 경영학 박사학위를 취득하였다. 현재 서울대학교 경영대학 연구부교수로 재직 중이며, 주요 관심분야는 IT전략, 온라인 플랫폼, 미디어 콘텐츠 등이다.

· **한 진 영(han1618@cau.ac.kr)**

고려대학교에서 경영학 박사학위를 취득하였고 현재 중앙대학교 교양대학 부교수로 재직 중이다. 주요 강의 및 연구 분야는 정보보안관리, 지식관리, 데이터리터러시, 웹3.0 등이다

· **손 영 우(이메일주소)**

고려대학교에서 경영학 박사학위를 취득하였으며, 현재 고려대학교 대학원 첨단기술비즈니스학과 부교수로 재직 중이다. 주요 강의 및 연구분야는 첨단기술비즈니스 인사이트, 비즈니스 모델, 기술 스타트업, Open Innovation, Business Dynamics 이다.

· **이 지 은(scully1215@hycu.ac.kr)**

한양대학교 정보기술경영학과에서 공학박사 학위를 취득하고 현재 한양사이버대학교 경영정보·AI비즈니스학과에 교수로 재직 중이다. 주요 관심분야는 인공지능, 기술경영, 에듀테크이다.