

메타버스 환경에서의 딥 러닝 기반 알고리즘을 활용한 상표권 탐지 시스템

이지은[○], 이형수^{*}, 신용태^{**}

[○]숭실대학교 컴퓨터학과,

^{*}한국폴리텍대학교 사이버보안과,

^{**}숭실대학교 컴퓨터학부

e-mail: lhsgsse10@soongsil.ac.kr[○], hslee01@kopo.ac.kr^{*}, shin@ssu.ac.kr^{**}

Trademark detection system using deep learning-based algorithms in a metaverse environment

Ji-Eun Lee[○], Hyung-Su Lee^{*}, Yong-Tae Shin^{**}

[○]Dept. of Computer, Soongsil University,

^{*}Dept. of Cyber Security, Korea Polytechnic University,

^{**}Dept. of Computer Science, Soongsil University

● 요약 ●

코로나 19(Covide-19)이후 가상과 현실이 융·복합 되어 사회·경제·문화활동과 가치 창출이 가능한 메타버스가 차세대 핵심산업으로 부상하고 있다. 이에 자사 보유 기술, IP(Intellectual Property) 등을 활용하여 메타버스 플랫폼을 구축하고자 하는 기업들이 증가하여 지식재산권을 둔 법적 이슈들이 새롭게 나타나고 있다. 따라서 본 논문에서는 상표권 침해를 보호하기 위하여 딥 러닝 기반 객체 탐지모델인 YOLOv5 모델을 활용한 메타버스 환경에서의 상표권 탐지 시스템을 제안한다.

키워드: 객체 탐지(object detection), IP(Intellectual Property), 메타버스(metaverse), 인공지능(AI)

I. Introduction

코로나 19(Covid-19) 이후 다가올 미래의 글로벌 신성장산업으로 주목받고 있는 ‘메타버스(metaverse)’는 초월을 의미하는 ‘메타(meta)’와 세계 또는 우주를 의미하는 ‘유니버스(universe)’의 합성어로 가상과 현실이 융·복합되어 사회·경제·문화 활동과 가치 창출이 가능한 디지털 세계이다[1]. 현재 네이버제트(NAVER Z)가 출시한 ‘제페토(ZEPETO)’는 2022년 말 기준 글로벌 누적 가입자가 4억 명을 넘어섰다[2]. 이렇듯, 메타버스를 활용한 행사들도 많아지고 있다. 대표적으로 방탄소년단(BTS)이 팬 커뮤니티 플랫폼 ‘위버스(Weverse)’에서 개최한 가상 콘서트에는 270만 명이 넘는 이용자가 몰리기도 하였으며, 제페토에서 방탄소년단의 공연 관련 한정 판매 굿즈와 상품을 디지털로 제작해 판매하는 특별 이벤트도 실시하였다. 이처럼 메타버스가 차세대 핵심산업으로 부상하면서 자사 보유 기술, IP(Intellectual Property) 등을 활용하여 메타버스 플랫폼을 구축하고자 하는 기업들도 늘어나고 있다. 이러한 사회변화로 메타버스 환경에서 다양한 법적 문제가 떠오르고 있다. 그 중 산업재산권에 해당하는 상표권 침해 사례로 프랑스 명품 브랜드 에르메스(Hermès)는 디지털 예술가인 메이슨 로스차일드를 상대로 상표권 침해 소송을 제기하였다

[3]. 그 이유는 메이슨 로스차일드가 메타버스의 ‘meta’와 에르메스 자사의 핵심 상품인 버킨백의 ‘birkin’을 합성한 ‘메타버킨즈(MetaBirkins)’라는 명칭으로 도메인을 등록하고 버킨백의 걸면을 모피 질감으로 처리한 메타버킨즈라는 NFT(Non-Fungible Token)를 만들어 메타버스 공간에서 판매하였다는 것이었다. 이와 같은 사례는 메타버스 환경에서 상표권 침해가 발생하고 있고 앞으로도 지속적으로 발생할 수 있음을 알 수 있다.

따라서 본 논문에서는 메타버스 환경에서 상표권을 보호하고자 딥 러닝 기반 객체 탐지 알고리즘인 YOLOv5 모델을 활용한 상표권 탐지 시스템을 제안한다.

II. Related work

본 장에서는 학습시킬 데이터셋이 주어졌을 때, 객체를 탐지해줄 YOLO 알고리즘과 동작 원리, PyTorch 그리고 본 논문에서 활용한 YOLOv5 구조를 살펴본다.

1. YOLO(You Only Look Once)

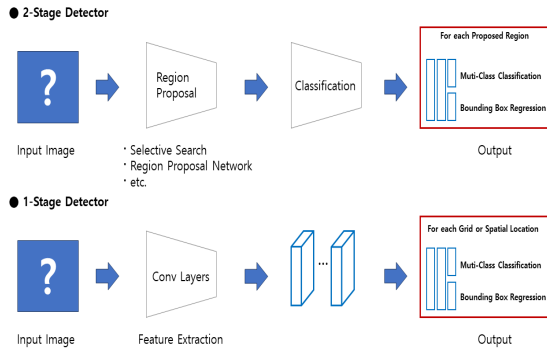


Fig. 1. Difference between 1-stage detector and 2-stage detector

YOLO[4]는 1-stage detector 방식의 객체 탐지 딥 러닝 기법으로 매우 빠른 속도와 추론 과정을 통한 실시간 물체 인식 기술이다[5]. 또한 Fast R-CNN[6]과 같은 기존의 실시간 객체 탐지 모델은 [Fig. 1]과 같이 2-stage detector 방식으로 Region Proposal과 Classification을 순차적으로 처리했다면 YOLO는 이 단계를 동시에 간략화하여 처리한다. YOLO는 이러한 접근법을 통하여 전체 이미지를 한 번에 처리하여 각 그리드 셀에서 객체의 존재 여부와 클래스를 예측함으로써, 실시간 처리를 가능하게 하며 훨씬 빠른 속도를 보여주기 때문에 기존 객체 탐지 모델들보다 뛰어난 성능을 보인다.

1.1 CNN(Convolutional Neural Network)

YOLO는 CNN을 기반으로 한 객체 탐지 모델이다. CNN은 이미지의 공간적 특징을 학습하는 데 효과적인 딥 러닝 구조로 합성곱 레이어, 활성화 함수, 풀링 레이어 등을 통해 복잡한 이미지 데이터에서 중요한 정보를 추출한다. YOLO는 이러한 CNN의 특징을 활용하여 각 이미지 내의 객체를 정확하게 탐지하고 분류한다.

1.2 동작 원리

YOLO의 동작 원리는 다음과 같다. 1) 입력 이미지를 $S \times S$ 와 같은 동일한 크기의 그리드(grid)로 나눈다. 2) 각 그리드에 대해 그리드 중앙을 중심으로 미리 정의된 형태(predefined shape)로 지정된 경계박스의 개수를 예측하고 이를 기반으로 신뢰도(confidence)를 계산한다. 3) 높은 신뢰도를 가진 위치를 선택해 해당 구역의 객체를 인식한다[7]. 이와 같은 YOLO는 기존의 복잡한 객체 검출 과정과 달리 후보 영역을 추출하지 않는 단일 신경망 구조를 가지고 있다.

2. PyTorch

PyTorch는 2016년에 Facebook의 AI 연구팀이 개발한 딥 러닝 구현을 위한 파이썬 기반의 오픈 소스 머신 러닝 라이브러리이다[8]. 또한 PyTorch는 Tensorflow와 다르게 Torch 기반 딥 러닝 프레임워크이다. 따라서 동적 그래프를 사용하여 동작 도중 그래프를 정의할

수 있는 장점이 있다. 기존 YOLO의 모델 중 v1, 2, 3, 4까지의 모델은 Tensorflow 기반의 모델이며 DARKNET을 이용하여 학습을 진행하였다[9]. 본 논문에서 활용할 YOLOv5[10]는 PyTorch를 기반으로 구현된 모델이다. YOLOv5 구조는 신경망의 크기에 따라 N(Nano), S(Small), M(Medium), L(Large), X(XLarge) 5가지로 분류할 수 있다. 신경망의 크기가 커질수록 더 많은 가중치로 정확도는 높아지지만 속도는 느려지는 단점이 있다.

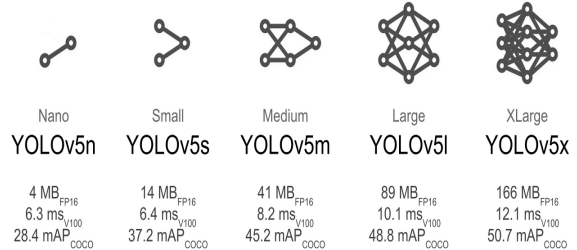


Fig. 2. Model structure differences in YOLOv5

III. Proposed Method

메타버스 환경에서는 실시간 처리가 중요하기 때문에 실시간 객체 탐지에 적합한 1-stage object detection model인 YOLOv5[10]을 활용하여 [Fig. 3]와 같은 상표권 탐지 시스템을 제안한다.

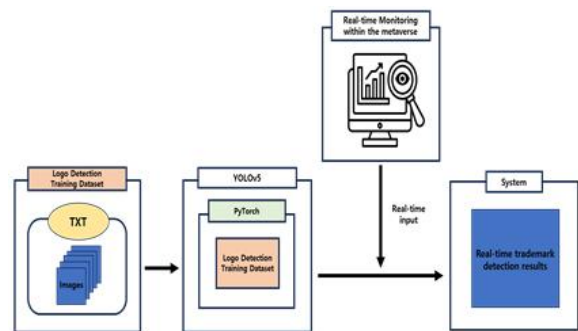


Fig. 3. Proposed system structure

제안하는 Object detection model인 YOLOv5는 PyTorch를 이용하여 학습시키기 위한 데이터셋이 필요하다. 이 데이터셋은 웹에서 크롤링을 수행하여 이미지를 저장하고, 저장된 이미지 속 각 상표권에 해당하는 로고와 텍스트 로고를 바운딩 박스(bounding box) 좌표와 라벨링 정보 등을 포함하여 TXT 파일로 저장한다. 이때 라벨링된 정보를 담고 있는 TXT 파일 형식은 다음 [Table 1]과 같은 정보를 포함한다.

Table 1. Information contained in text files

항목	설명
class index	객체의 클래스를 나타내는 정수 인덱스이다.
x_center	객체의 바운딩 박스 중심의 x좌표이며, 이 값은 이미지의 전체 너비에 대한 비율로 표현된다.
y_center	객체의 바운딩 박스 중심의 y좌표이며, 이 값은 이미지의 전체 높이에 대한 비율로 표현된다.
width	객체의 바운딩 박스 너비이며, 이 값은 이미지의 전체 너비에 대한 비율로 표현된다.
height	객체의 바운딩 박스 높이이며, 이 값은 이미지의 전체 높이에 대한 비율로 표현된다.

IV. EXPERIMENTAL RESULTS

상표를 탐지하기 위해 본 연구에서 사용된 학습 데이터셋은 Roboflow의 Logo detection Dataset으로 총 6301장의 사진으로 구성되어있다. 이 데이터셋은 각 사진에 포함된 로고에 대한 바운딩 박스 좌표와 라벨링 정보 등을 포함하고 있으며 제안하는 시스템을 구현하기 위해 실시간으로 객체 탐지가 가능한 YOLOv5를 PyTorch 이용하여 학습을 진행하였다. 학습을 위해 전체 6301장의 이미지 중 80%를 학습 데이터로 사용하였고 20%를 평가 데이터로 사용하였다. 또한 파라미터 값으로 학습할 이미지 크기는 416*416, batch size는 16, epochs는 50으로 지정하여 적용하였다. 실험결과는 다음 [Fig. 3]과 같다.

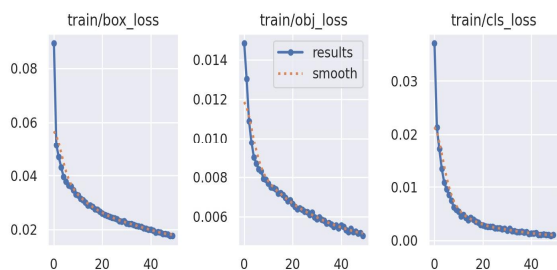


Fig. 3. Loss of the training set output as a result of object detection model learning

[Fig 3]는 YOLOv5가 학습 과정을 거친 후 출력되는 훈련 데이터 세트의 loss를 보여준다. 첫 번째 그래프인 Box loss의 경우, IoU에 대한 loss를 뜻한다. IoU의 식은 다음과 같다.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (Eq. 1)$$

subject to:

Area of Overlap = 모델이 예측한 bounding box와 ground truth bounding box가 겹치는 교집합 영역

Area of Union = 모델이 예측한 bounding box와 ground truth bounding box를 포함한 합집합 영역

따라서 IoU가 높을수록 주어진 정답과 모델이 예측한 객체의 범위의 차이가 작다는 것을 의미한다. 두 번째 그래프인 Obj loss는 Objectness Loss로 객체 인식에 대한 loss 값을 의미한다. 이 경우 학습된 모델이 사진 내의 객체를 인식하는 비율을 의미하며, 사진 내의 모든 객체를 인식하였는지, 인식된 객체가 원하는 객체였는지를 나타낸다. 세 번째 그래프인 Cls loss는 클래스 분류에 대한 loss이다. 모델이 인식한 객체에 대해 클래스 분류를 제대로 진행하였는지를 알 수 있는 loss 값으로, 이 모델에서는 상표를 잘 분류하였는지를 의미한다.

학습 결과 최종 모델의 유효성 지표(mAP)는 [Fig. 4]와 같이 0.872로 나왔으며 [Fig. 5]는 학습 결과물을 보여준다.

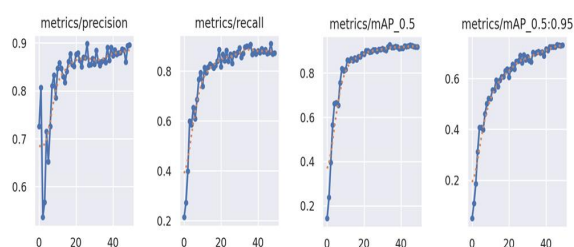


Fig. 4. Object detection model training results



Fig. 5. Results of trademark detection system

V. Conclusions

본 논문에서는 딥 러닝 기반 알고리즘인 YOLOv5를 활용하여 메타버스 환경에서 상표를 탐지함으로써 상표권을 보호할 수 있는 시스템을 제안하였다. 하지만 상표의 각도에 따른 왜곡(Distortion) 또는 상표의 일부가 다른 객체에 의해 가려졌을 때(Partial Occlusion)와 같은 기하학적 변형에서는 상표를 잘 인식할 수 없는 문제점이

발생하였다. 따라서 향후 연구에서는 훈련 데이터에 다양한 각도와 크기가 변형된 이미지를 포함시켜 다양한 기하학적 변형에서도 높은 정확성을 보일 수 있도록 본 연구의 한계점을 개선할 것이다.

REFERENCES

- [1] Joon-Sung. Bang “The Era of Metaverse is Coming.” National Assembly Future Institute, Future Column. October 6, 2021 <http://www.nafi.re.kr/new/contribution.do>
- [2] Eun-Young. Jeong “Naver’s ZEPETO, Rising as a Hotspot for Generation Z” Korea Finance, Telecommunication & Internet. (February 13, 2023) http://www.fntimes.com/html/view.php?ud=202302110904341191dd55077bc2_18
- [3] Shin & Kim’s Metaverse-NFT Team. “NFT Metaverse and Trademark Rights”, Law Firm Sejong, Newsletter, June 21, 2022 <http://www.shinkim.com/kor/media/newsletter>
- [4] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A., You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, May 2016
- [5] Seong-Woo. Kang, Hoon. Song, Young-Hak. Shin, June, 2022. “Deep learning-based real-time colorectal polyp detection study Using YOLOv5” Proceedings of The Institute of Electronics and Information Engineers summer conference
- [6] Girshick, R. Fast r-cnn. In Proceedings of the IEE international conference on computer vision, pp. 1440-1448, 2015
- [7] Yong-Hwan. Lee and Young-seop. Kim, “Comparison of CNN and YOLO for Object Detection.” Journal of the Semiconductor and Display Technology, vol.19, no.1, pp. 85-92, 2020
- [8] UR, [AI] What is PyTorch?, UR Dev Log, 2019 <https://blog.naver.com/os2dr/221565409684>
- [9] Hyoung-Seok. Jeon, Dong-Hae. Yeom and Young-Hoon. Joo, “Video-based Intelligent Unmanned Fire Surveillance System,” pp. 516-521, 2010
- [10] GitHub, ultralytics/yolov5, <https://github.com/ultralytics/yolov5.git>