

## 실시간 재식별 기술을 통한 반려동물 움직임 분석 서비스

윤 경 윤<sup>1</sup> · 설 재 형<sup>1</sup> · 조 영 준<sup>2\*</sup><sup>1</sup>전남대학교 IoT인공지능융합학과 학사과정<sup>2\*</sup>전남대학교 소프트웨어공학과 교수

## Pet Movement Analysis Service using Real-time Re-identification

Kyung-Yoon Yoon<sup>1</sup> · Jae-Hyung Seol<sup>1</sup> · Yeong-Jun Cho<sup>2\*</sup><sup>1</sup>Undergraduate Student, Department of IoT Artificial Intelligence Convergence, Chonnam National University<sup>2\*</sup>Assistant Professor, Department of Software Engineering, Chonnam National University

### [요 약]

저출산·고령화에 따라 반려동물을 양육하는 가정이 늘고 있다. 반려동물을 집에 홀로 두는 시간이 길어지며, 모니터링 CCTV 및 카메라의 수요가 증가하고 있다. 그러나 기존의 펫테크 기기는 가격이 비싸고 기능도 단순하다. 사용자가 일과 중 계속해서 영상을 살펴보아야 하고, 주어진 영상 이외에는 반려동물에 대한 구체적인 정보를 얻기 어렵다. 이에 본 연구는 인공지능을 이용하여 반려동물의 움직임을 추적하고 분석하는 웹 서비스를 제안한다. 높은 정확도의 다중 객체 추적 성능을 갖추고 있어, 여러 마리 반려동물에 대한 개별적인 분석 결과를 사용자에게 제공한다. 기술 구현을 위해 객체 분류기, 검출기, 추적기, 분석 및 시각화 알고리즘을 적용한다. 제안하는 서비스를 실험한 결과, 여러 마리의 반려동물이 자유자재로 움직이는 상황에서 다중 객체 추적과 결과 분석이 효과적으로 수행되는 것을 확인했다. 본 연구는 기존의 펫테크 기기가 가진 기술적 한계를 극복하여 다양한 분야에 적용할 수 있게 했다. 가정과 사업장 및 기관에서 동물 모니터링 용도로 사용할 수 있으며, 학습 데이터를 조정하여 새로운 분야에 적용할 인공지능 어플리케이션을 만들 수 있다.

### [Abstract]

The number of families raising companion animals is increasing. The time pets are left alone at home is getting longer, and the demand for monitoring CCTVs and cameras is increasing. However, existing pet tech devices are expensive and have simple functions. The user has to continuously monitor the video during the day, and it is difficult to obtain specific information about the companion animal other than the given video. Therefore, this study proposes a web service that tracks and analyzes the movement of companion animals using artificial intelligence. High-accuracy multi-object tracking performance is achieved, providing users with individual analysis results for multiple companion animals. To implement the technology, object classifier, detector, tracker, analysis, and visualization algorithms are applied to the web service. The analysis was performed effectively in a situation where several companion animals were freely moving. This study overcomes the technical limitations of existing pet tech devices and can be used for animal monitoring at home and in institutions. The technology can be applied to new fields by adjusting the learning data.

**색인어** : 객체 검출, 객체 추적, 객체 재식별, 반려동물 움직임, 펫테크 기기**Keyword** : Object detection, Object tracking, Object re-identification, Pet movement, Pet-tech device<http://dx.doi.org/10.9728/dcs.2023.24.3.531>

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 13 December 2022; Revised 10 January 2023

Accepted 27 February 2023

**\*Corresponding Author; Yeong-Jun Cho**

Tel: +82-62-530-3432

E-mail: yj.cho@chonnam.ac.kr

## 1. 서론

반려동물을 양육하는 가정이 증가함에 따라, 관련 산업의 규모가 커지고 있다. 2020년 전 세계 펫테크 사업 규모는 55억 달러(약 6조8,000억 원)에 이른다[1]. 저출산·고령화와 1인 가구 수 증가가 진행되며 이 수치는 점차 커질 것으로 보인다. 반려동물에 대한 사람들의 관심이 늘면서 반려동물 맞춤형 기술의 필요성이 커지고 있다.

구체적인 펫테크 기기 사용 현황[2]에 대해 살펴보면 그림 1과 같다. 가장 많이 이용한 펫테크 기기는 자동 급식/급수기이며, 두 번째로 많이 이용한 펫테크는 모니터링 CCTV 및 카메라이다. 사용자 중 30.3%가 모니터링 CCTV 및 카메라를 사용했다고 답했다. 가정에서 효과적인 반려동물 양육을 위해 다양한 펫테크 기기를 사용한다. 이 중 모니터링 CCTV는 컴퓨터 비전 기술과 접목하여 반려동물 양육에 대한 편의를 증대시킬 수 있다.

기존의 상용화된 모니터링 CCTV 및 카메라는 영상 확인 및 저장의 단순한 기능을 가지고 있다. 사용자가 계속해서 촬영된 영상을 확인해야 하며 통계적 기법을 적용하지 못하는 한계가 있다. 인공지능을 적용한 홈 CCTV 어플리케이션[3]이 존재하나, 한 마리에 대해서만 분석이 가능하여 여러 마리의 반려동물을 양육하는 가정에서는 사용할 수 없다. 본 연구에서 이를 극복하기 위해 “반려동물의 움직임을 추적하여 분석하는 인공지능 웹 어플리케이션”을 제안한다. 반려동물 영상을 입력으로 받아 인공지능 모델로 분석한 후 사용자에게 필요한 정보를 제공한다. 컴퓨터비전 기술을 웹 어플리케이션 형태로 구현하여 모바일 기기와 컴퓨터에서 편리하게 활용할 수 있다.

웹 서비스는 다음의 과정을 수행한다. 각 반려동물 영상이 입력으로 들어오면 객체 분류기를 이용하여 학습한다. 이후 분석할 CCTV 영상을 이용하여 객체 검출기와 객체 추적기, 객체 분류기를 순차적으로 동작한다. 인공지능의 결과물로 반려동물 움직임 정보가 출력되면 분석 및 시각화 알고리즘을 적용한다. 최종적으로 반려동물이 이동한 시간과 특정 구역에 머무른 시간 및 횟수, 움직임을 시각화한 이미지를 사용자에게 전달한다.

제안한 기술을 통해 얻을 수 있는 기대효과는 다음과 같다. 첫 번째, 외출 시 반려동물에 대한 걱정을 덜 수 있다. 통계 결과[2]에 따르면 응답자가 펫테크 기기를 구매하는 주된 이유는 반려동물 홀로 두고 외출할 때 걱정을 덜기 위해서이다. 제안한 기술은 사용자에게 반려동물 영상과 분석 결과를 제공하여 효과적인 모니터링을 수행하게 한다. 두 번째, 비싼 펫테크 기기를 구매하지 않아도 된다. 웹 어플리케이션 서비스 이기에 사용자는 카메라가 달린 웹 기기 이외 별도의 펫테크 기기를 구매하지 않아도 된다. 세 번째, 여러 마리의 반려동물을 효과적으로 돌볼 수 있다. 다중 객체 추적 기술을 이용하여 각 반려동물의 움직임을 추적하고 분석하기 때문에 사용자가 많은 시간을 들이지 않아도 된다.

이어지는 장을 통해 구체적인 연구 과정을 제시한다. 먼저

제2장 배경 연구에서 반려동물 양육 현황 및 선행 기술에 대해 논의한다. 이후 제3장 제안 기술 및 서비스에서 사용된 컴퓨터비전 기술에 대해 논의한다. 구체적으로 객체 검출, 객체 추적, 분석 및 시각화 알고리즘, 사용자 인터페이스에 대해 다룬 후, 제4장 실험 결과에서 연구를 통한 성능 향상을 살펴본다. 제5장 결론으로 마친다.

### Experience with Pet Tech

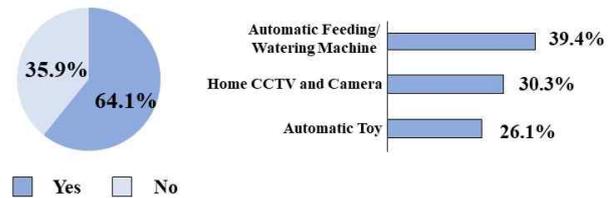
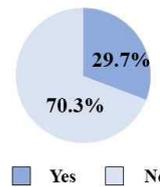


그림 1. (통계) 펫테크 기기 사용  
Fig. 1. (Statistics) Pet tech device use

### Pet raising family



### Pet Species

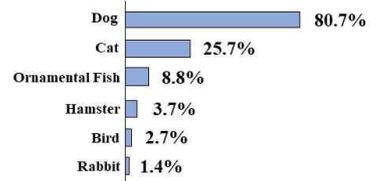
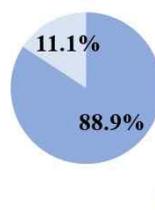


그림 2. (통계) 반려동물 양육 가정 및 반려동물 종류  
Fig. 2. (Statistics) Pet raising family and pet species

### Consider pet as family



### Leave pet alone at home

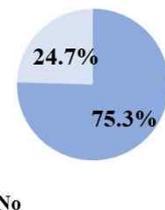


그림 3. (통계) 반려동물을 가족으로 생각하는지 여부; 반려동물을 집에 홀로 둔 적이 있는지 여부

Fig. 3. (Statistics) Consider pet as family; Leave pet alone at home

## II. 배경 연구

### 2-1 반려동물 양육 현황 및 실태

반려동물(companion animal) 용어는 “사람이 정서적으로 의지하고자 가까이 두고 기르는 동물”[4]의 의미를 지니고 있다. 이 용어는 동물학자 콘라트 로렌츠(Konrad Lorenz)가 1983년 열린 ‘사람과 애완동물의 관계(The human-pet

relationship)’를 주제로 한 국제심포지엄에서 처음 제안했다. 국내에서는 2007년에 동물보호법이 개정되며 반려동물이라는 용어를 공식적으로 사용하기 시작했다[5]. 이처럼 단순히 귀여워하고 예뻐하는 ‘애완’의 의미에서 가족이자 친구로 의지하는 ‘반려’의 의미로 동물 양육에 대한 인식이 변화하고 있다.

구체적으로 반려동물 양육 실태[2]를 살펴보면 그림 2와 같다. 국내 총 604만 가구의 1,448만 명이 반려동물을 양육하는 것으로 조사됐다. 이는 전체 가구 중 29.7%에 해당하는 수치로, 10가구 중에서 3가구는 반려동물을 양육하고 있음을 보여준다. 반려동물의 종류별 양육 비율을 살펴보면 개 80.7%, 고양이 25.7%, 판상어 8.8%, 햄스터 3.7%, 새 2.7%, 토끼 1.4%이다. 반려동물 종류 중에서는 양육하고 교감하기 좋은 개와 고양이를 많이 키우는 것으로 보인다.

최근에는 ‘펫팸족’과 같은 용어도 생겨나 사용되고 있다. “반려동물(Pet)을 가족(Family)처럼 여기며 살아간다”[6]라는 의미이다. 반려동물을 가족과 같이 보호해야 할 대상으로 여기고 있다. 그러나 이러한 인식의 변화와 달리 반려동물을 양육하는 것에는 여전히 어려움이 따른다. 통계[2]에 따르면 반려동물 양육 가정 중 75.3%는 반려동물을 집에 혼자 두는 경우가 있으며, 그 시간은 하루 평균 5시간 40분이라고 답했다(그림 3). 1인 가구의 경우에는 반려동물을 집에 두는 시간이 2시간가량 더 길었다. 반려동물 양육 가정의 구성원은 사회적 활동을 위해 필연적으로 집을 비우는 일이 생긴다. 재택근무가 증가하는 추세지만 여전히 반려동물을 집에 홀로 두는 상황이 많아 대책이 필요하다.

한편, 반려동물을 집에 혼자 두는 상황을 해결하기 위해 기술이 혼합된 펫테크 기기를 사용하는 사람이 생겨나고 있다. 대표적으로 사용되는 펫테크 기기에는 홈 CCTV 및 카메라, 자동 급식기/자동 급수기, 반려동물용 자동 장난감 등이 있다. 조사[1]에 따르면 2020년의 전 세계 펫테크 산업 규모는 55억 달러(약 6조8,000억 원)에 이르렀다고 한다. 반려동물에 기술을 더하는 펫테크 사업의 규모가 커짐에 따라 펫테크 기기에 차별화된 기술을 적용하는 것이 점차 중요해지고 있다.

## 2-2 인공지능 및 영상분석 기술

본 연구에서는 다양한 인공지능 및 컴퓨터비전 기술을 활용한다. 인공지능이란 인간의 지능이 가지는 추리, 적응, 논증 따위의 기능을 갖춘 컴퓨터 시스템을 의미한다[4]. 컴퓨터비전은 인공지능의 한 응용 분야로 시각기관의 역할을 하는 기술이다. 이미지나 동영상 자동 분석하여 필요한 정보를 추출한다.

합성곱 신경망(CNN)[7]은 심층 신경망의 종류 중 하나이며, 필터 학습 기법을 이용해 이미지를 효과적으로 분석한다. 2차원 영상을 처리하기 적합하여 객체 탐지, 객체 분류 등 다양한 컴퓨터 비전 분야에서 연구가 진행되고 있다. 한 개 이상의 합성곱 계층(Convolution layer)과 풀링 계층(Pooling layer), 완전 연결 계층(fully connected layer)으로 구성되어 있다[8],[9]. 합성곱 신경망 알고리즘 중 하나인 VGGNet[10]은 합성곱 계층의 크기를

3×3으로 고정하여 파라미터 개수를 줄이면서 신경망(Networks)의 계층(Layer)을 깊게 쌓았다. 이전 CNN 모델보다 학습의 효율성을 향상하여 다양한 인공지능 모델에 활용되고 있다.

객체 탐지는 이미지 내의 객체 클래스를 판별하고, 위치 및 크기를 예측하는 기술이다. 객체 탐지 기술의 한 종류인 YOLO(You Only Look Once)[11]는 객체 탐지 상자(Bounding Box)를 그려 위치, 크기, 클래스, 정확도를 표시한다. 기존의 R-CNN은 객체 검출에 분류 방식을 사용하여 속도가 느리고 최적화가 어려운 문제가 있었다. YOLO는 객체 검출에 회귀 방식을 사용하여 객체 탐지 상자의 위치와 크기, 클래스, 정확도를 구하는 과정을 단순화했다. 자동차 자율주행, 이미지 검색, 글자 탐지, 얼굴인식 등 다양한 분야에 활용되고 있다.

객체 추적은 입력으로 들어온 영상에서 프레임 간 객체의 위치 변화를 이용하여 움직임을 추적하는 기술이다[12]. 기존의 객체 추적 기술은 객체끼리 겹치면 추적 ID가 변경되고 미탐(False Negative)이 발생하는 문제가 있다. DeepSort[13]는 객체의 추적이 잠시 끊어지더라도 객체의 외형 정보를 기반으로 추적을 지속하여 이 문제를 해결한다. 그러나 DeepSort는 장시간 추적이 끊기면 추적 ID 정보를 삭제하여 동일 객체가 영상에 다시 등장했을 때 새로운 추적 ID를 부여하는 한계가 있다. 본 연구에서 움직임을 분석하는 객체인 반려동물은 매우 활동적인 특성이 있다. 영상 안팎으로 자유롭게 드나들기 때문에 객체 추적 기술만 사용하면 추적이 장시간 끊겨 추적 ID가 변경되기 쉽다. 이를 해결하기 위해 객체 재식별[14] 기술을 추가로 활용하여 재등장한 객체에 대해 기존 추적 ID를 부여하게 했다.

## 2-3 관련 연구 및 서비스

시중에서 많이 거래되고 있는 반려동물 CCTV는 대체로 단순 실시간 영상 송출 기능만 제공한다. 일부 제품은 기술이 추가되어 대상의 움직임을 감지하여 영상을 저장하거나 사용자의 조작에 의해 화각을 바꾸는 기능을 가진 것이 있다. 그러나 여전히 반려동물의 상태를 수시로 확인해야 한다는 한계점이 있다. 반려동물이 집에 혼자 있을 것을 걱정하며 반려동물 CCTV를 구매하는 가정이 많은데, 정작 위급 상황에는 영상을 보지 못하는 경우가 생긴다. 여러 마리 반려동물을 양육할 경우에는 개별의 상태를 확인하는데 많은 시간이 소요되어, 정확한 반려동물의 상태를 판단하기 더 어려워진다.

인공지능을 이용한 반려동물 CCTV에는 객체 탐지 기술을 사용하여 동물의 움직임을 분석하고 통계 내는 어플리케이션이 존재한다[3]. 그러나 탐지 모델의 정확도가 완전하지 않아 작거나 멀리 떨어져 있는 객체에 대해 잘 검출해내지 못한다. 객체 탐지가 끊기면 반려동물의 움직임 정보를 얻을 수 없어 부정확한 통계 결과가 나오는 문제가 있다. 또한 해당 어플리케이션은 여러 마리의 반려동물을 구분해내지 못한다. 객체 탐지 기술은 반려동물이라는 하나의 클래스에 대해 탐지할 수 있지만, 여러 마리 반려동물을 양육하는 가정에서 각 객체를 구분하지는 못하는 한계점이 존재한다.

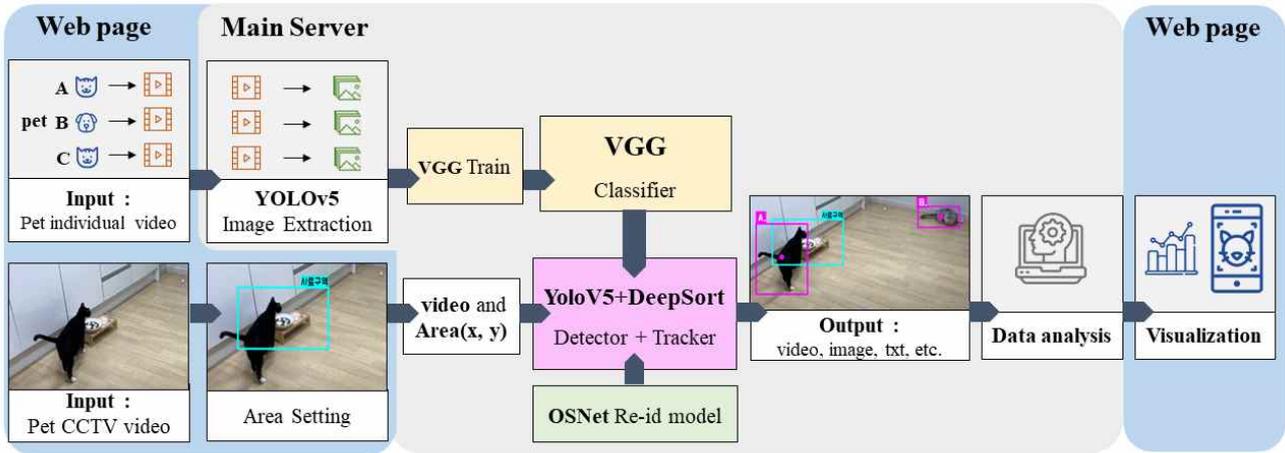


그림 4. 반려동물의 움직임을 추적 및 분석하는 인공지능 웹 서비스  
 Fig. 4. AI web service that tracks and analyzes the movement of pet

이 밖에도 반려동물의 행동을 분석하거나[15], 움직임을 모니터링하는 선행 연구가 존재한다[16]. 그러나 이 연구는 객체 탐지 기술만을 이용하므로, 활동량이 많은 반려동물에게 적용할 경우 정확도가 확연히 떨어지며 영상의 중간에 추적되지 않는 부분이 발생하는 한계가 있다.

시한다. 세 번째, 사용자에게 유용한 분석을 제공하는 것이다. 반려동물의 CCTV 영상을 분석하여 반려동물의 정보를 사용자에게 전달한다. 자체적인 분석 알고리즘을 개발하여 반려동물의 이동 시간, 특정 구역 방문 횟수 및 시간, 이동 경로 시각화 이미지를 생성한다.

### III. 제안 기술 및 서비스

#### 3-1 제안 서비스의 전체 구조 및 목표

본 연구는 반려동물의 움직임을 추적하여 분석 및 시각화한다(그림 4). 웹 페이지를 통해 사용자 입력을 받고 결과를 출력한다. 사용자 입력 정보는 각 반려동물의 30초 정도의 영상과 분석할 CCTV 영상이다. 입력된 정보에 대해 메인 서버에서 인공지능 모델과 분석 및 시각화 알고리즘을 동작하여 웹 페이지로 결과를 보낸다. 웹 서버로 Node.js와 Flask 두 가지를 사용했다. 웹 페이지 관련된 기능은 프론트엔드와 동일한 언어인 Node.js 웹 프레임워크를 사용하여 개발 생산성을 높였다. 인공지능 알고리즘이 수행되는 부분은 Python 코드를 실행하기에 용이한 Flask 웹 프레임워크를 사용하였다. 두 서버 간 API 방식으로 통신하게 했다. 최종적인 출력은 영상에서 반려동물이 움직인 시간, 특정 장소에 머무른 횟수와 시간, 시각화 이미지이다.

해당 연구의 목표는 다음과 같다. 첫 번째, 반려동물의 움직임을 높은 정확도로 추적하는 것이다. 이를 위해 객체 추적 모델과 함께 동작하는 객체 탐지 모델의 성능을 향상한다. 두 번째, 여러 마리 반려동물에 대해 효과적인 추적 및 분석을 할 수 있게 하는 것이다. DeepSort 객체 추적 모델은 특정 시간 이상 추적이 끊기면 객체에 대한 정보를 삭제하고 추적 ID를 새롭게 부여한다. 본 연구를 통해 이에 대한 해결책을 제

#### 3-2 반려동물 자동 검출

본 연구는 객체 검출기를 이용하여 클래스를 탐지한 후, 객체 추적기를 이용하여 각 객체의 움직임을 추적한다. 객체 추적을 효과적으로 수행하기 위해 뛰어난 탐지 성능과 수행 속도를 가진 YOLOv5(이하 YOLO)[17] 객체 검출기를 사용한다.

YOLO가 공개한 모델은 MS coco 데이터셋[18]으로 학습한 것이다. 해당 데이터셋은 개와 고양이의 정면 모습에 대한 이미지가 많다. 기존 모델로 Test를 수행했을 때, 반려동물의 뒷모습과 작게 웅크린 모습에 대해 정확하게 탐지해내지 못하는 문제가 발생했다. 이를 개선하기 위해 탐지 임계값을 낮추어 수행해보았으나, 여전히 반려동물을 탐지하지 못하는 미탐(False Negative)과 다른 객체를 반려동물이라고 탐지하는 오탐(False Positive)이 발생했다. 이러한 문제를 해결하기 위해 웅크린 자세에 대한 이미지와 다각도 촬영 이미지를 추가 수집했다. 학습 데이터를 늘려 탐지 모델의 정확도를 개선하였다.

기존 YOLO는 개와 고양이의 특징이 비슷한데, 클래스가 따로 나뉘어 있어 정확한 분류가 수행되지 않았다. 이에 따라 탐지 임계값을 넘기지 못하여 YOLO가 객체를 탐지하지 못하는 문제가 발생했다. 본 연구에서는 개와 고양이 클래스를 하나의 반려동물 클래스로 통합하여 탐지 정확도가 떨어지는 것을 방지한다. 개와 고양이 클래스를 구분하여 받아들이는 문제에서 반려동물 클래스를 받아들이는 간단한 문제로 변환했다.

### 3-3 다중 반려동물 추적 및 재식별

객체 검출기를 이용한 후, 객체 추적기로 반려동물의 움직임을 추적한다. 본 연구는 OSNet[14]을 재추적 모델로 사용하여 ID Switch 문제와 미탐 발생을 개선한 DeepSort[19] 객체 추적기를 사용했다. 객체 검출기 YOLO와 객체 추적기 DeepSort를 함께 사용하여[20] 반려동물 클래스에 속하는 객체 움직임 정보를 얻는다.

객체 추적기에서 재추적 모델을 사용하여 ID Switch와 미탐 발생을 해결했지만, 여전히 한계가 존재한다. DeepSort는 장시간 추적이 끊기면 객체 ID 정보를 삭제한다. 즉, 반려동물이 카메라 화각을 벗어났다가 해당 장소에 재등장할 경우 기존 반려동물에 대한 결과가 아닌 새로운 반려동물로 결과를 낸다.

이를 해결하기 위해 객체 분류기를 사용했다. 본 연구는 간단한 구조와 우수한 분류 성능을 가진 VGG[10]를 사용했다. 각 반려동물의 짧은 영상을 입력으로 넣어 VGG를 학습한 후 DeepSort의 추적이 시작되는 부분에 삽입하여, VGG에 의해 분류된 반려동물 이름이 추적할 객체에 할당되게 한다.

### 3-4 분석 및 시각화 알고리즘

객체 추적기의 결과로 나온 값은 분석 알고리즘을 이용하여 사용자에게 제공하기 적합한 형태로 가공한다. 사용자에게 제공할 정보는 다음과 같다. 반려동물이 이동한 시간, 반려동물이 특정 위치에 머무른 시간과 횟수, 반려동물이 움직임을 시각화한 이미지이다. 객체 추적기의 출력으로 나온 텍스트 파일을 분석하고 시각화하여 웹 페이지에 전달한다.

반려동물의 이동 시간, 방문 시간 및 횟수를 측정하기 위해 IoU(Intersection over Union)를 이용했다(수식 1). IoU는 객체 검출에서 정확도를 측정할 때 사용하며, 실제 객체 탐지 상자와 예측한 객체 탐지 상자 두 영역이 얼마나 겹쳐 있는지 나타낸다. 교집합 영역 넓이를 합집합 영역 넓이로 나누어 계산한다. 0~1 사이의 값을 가지며, 1에 가까운 값일수록 두 영역이 많이 겹침을 의미한다.

$$B_{gt} = \text{groundtruth}, B_p = \text{prediction},$$

$$IoU = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{area}(B_{gt} \cap B_p)}{\text{area}(B_{gt} \cup B_p)}, \quad (1)$$

반려동물의 이동시간을 측정하는 방식은 다음과 같다. 1초에 해당하는 프레임마다 반려동물 객체 탐지 상자의 IoU를 계산하여 값이 작으면 움직인 것으로 간주한다. 이때 기준이 되는 IoU는 여러 차례 반복하여 실험한 후에 가장 적합한 결과를 내는 0.7을 사용했다. 움직이는 경우에만 이동 시간을 1초씩 증가한다.

반려동물의 방문 시간과 횟수를 측정하는 방식은 다음과

같다. 사용자가 웹에서 특정 구역을 드래그 방식으로 선정하여 좌측상단과 우측하단 위치 좌표 정보(수식 2)를 넘겨준다.  $j$ 는 사용자가 입력한 각각의 특정 구역 인덱스를 의미하며,  $M$ 은 사용자가 입력한 전체 특정 구역 수를 나타낸다.

$$a^j = [x_1^j, y_1^j, x_2^j, y_2^j], j = 1, 2, 3, \dots, M \quad (2)$$

프레임별로 특정 구역의 넓이와 반려동물 객체 탐지 상자 넓이 간 IoU를 계산하여 일정 시간 이상 큰 IoU 값을 유지하면 방문한 것으로 보았다. 지속 시간을 도입한 이유는 객체가 단순히 특정 구역을 스치는 경우를 방문으로 보지 않기 위함이다. 이때 기준이 되는 지속 시간은 1초, IoU는 0.2를 사용했다. 이는 여러 차례 반복하여 실험한 후에 가장 적합한 결과를 내는 값을 적용한 것이다.

반려동물 움직임 시각화의 목표는 반려동물의 이동 경로와 많이 머무른 장소를 한눈에 보여주는 것이다. 시각화를 위한 라이브러리로 Matplotlib을 사용하여 구현했다. 분석 영상 크기와 동일한 0으로 채워진 벡터를 생성한 후, 프레임별로 반려동물의 위치를 구하여 해당 위치 벡터값을 1씩 증가한다. 이때 반려동물의 위치를 나타내기 위한 객체 탐지 상자의 중심점은 다음의 수식 3으로 구한다.  $x, y$ 는 객체 탐지 상자의 좌측 상단 좌표이며,  $w$ 와  $h$ 는 너비와 높이이다.

$$c = [x_i + \frac{1}{2}w_i, y_i + \frac{1}{2}h_i], i = 1, 2, 3, \dots, N, \quad (3)$$

$i$ 는 반려동물 객체 이름에 매칭된 인덱스를 의미하며,  $N$ 은 영상에 등장하는 전체 반려동물의 수를 의미한다. 영상의 길이에 상관없이 균형 있는 이미지를 보여주기 위해 최대 최소 정규화(MinMaxScaler)를 진행한다. 정규화식은 다음과 같다.

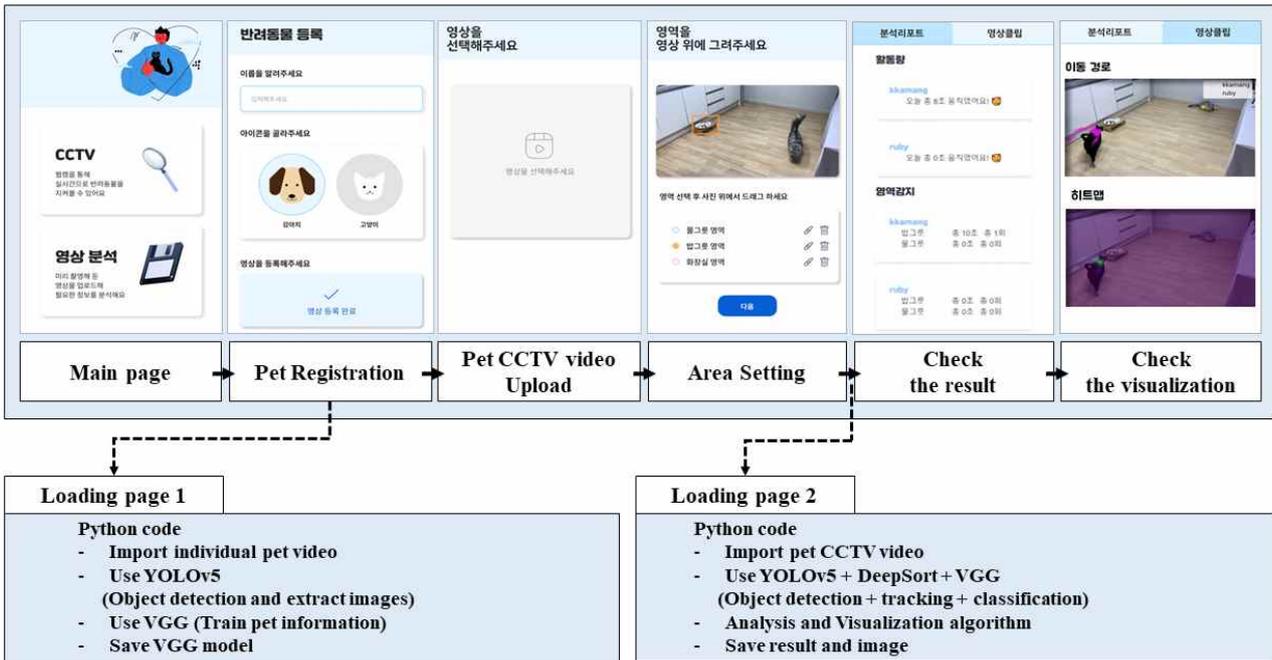
$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}, \quad (4)$$

최종적으로 그림 6과 같은 json 파일과 이미지가 생성되어 웹으로 넘겨진다.

### 3-5 사용자 인터페이스

본 연구에서는 사용자의 편의를 고려하여 웹 어플리케이션으로 기술을 구현한다. 웹 프론트엔드 개발은 JavaScript 기반의 라이브러리인 React를 이용하였다. React는 선언적인 컴포넌트 기반의 언어이며 코드 작성이 간결하고 유지보수에 용이하다. 웹 백엔드는 React와 동일한 언어인 Node.js를 사용하였다. 인공지능 모델을 실행하는 서버로는 Pytorch 환경을 구축한 Flask를 사용했다. Pytorch는 딥러닝 모델을 구현하는 데 사용하는 라이브러리이다. Node.js와 Flask는 API

(a) Web Page



(b) Pet Registration Algorithm

(c) Pet Analysis Algorithm

그림 5. 웹 페이지 구성 및 동작 알고리즘  
Fig. 5. Web page and Algorithm

"Pet Name": Move Time	<pre>                 "Pet Name": {                 "Area1": [                 Visit Time,                 Number of Visit                 ],                 "Area2": [                 Visit Time,                 Number of Visit                 ]                 }             </pre>	
Move.json	Visit.json	Visualization Image

그림 6. 분석 및 시각화 알고리즘 결과  
Fig. 6. Analysis and visualization algorithm results

를 생성하여 요청과 응답을 주고받는다. 웹 페이지의 구성은 다음과 같다(그림 5-a). 가장 먼저 사용하는 키우는 반려동물 전체에 대한 이름과 영상을 등록한다. 이후 영상 분석 페이지로 이동하여 분석할 CCTV 영상을 등록하고, 날짜와 특정 구역 선택을 완료한다. 이때 특정 구역은 입력 영상에서 드래그 방식으로 영역을 선택한다. 개수에 상관없이 자유롭게 특정 구역을 추가하고 삭제할 수 있으며, 장소 이름을 수정할 수 있다. 이후 로딩 페이지로 넘어가 분석을 진행한다. 분석을 완료하면 결과 페이지를 통해 반려동물의 움직임에 대한 통계 정보와 이미지, 영상을 제공한다. 마이페이지 영상 등록 시 로딩 페이지에서 수행하는 부분은 다음과 같다(그림 5-b). 반려동물 영상을 불러와 YOLO로 객체를 탐지하고 이미지로 추출한다. 추출한 이미지로 VGG의 학습을 진행하고 완성된 모델을 서버에 저장한다. 영상 분석 로딩 페이지에서 진행되는 내용은 다음과 같다(그림 5-c).

반려동물 CCTV 영상을 YOLO, DeepSort, VGG가 통합된 모델에 넣어 실행한다. 나온 출력 결과물을 분석 알고리즘에 적용하여 반려동물 이동시간, 특정 구역 방문 횟수 및 시간, 움직임 시각화한 이미지를 출력으로 얻는다.

웹 서비스 디자인은 Figma 도구를 사용하였다. 모든 웹 페이지에 탭 바를 생성하여 각 페이지를 편리하게 이동할 수 있게 했다. 사용자 입력은 데이터 특성을 고려하여 날짜 선택 도구, 드롭다운, 폼 등의 인터페이스 요소를 사용했다. 분석 결과 페이지는 활동량 폼과 특정 구역 폼으로 나누어 반려동물별 정보를 확인하도록 구성했다. 끝부분에 이미지를 배치하여 사용자가 원하는 정보를 빠르게 파악하도록 했다.

#### IV. 실험 결과

##### 4-1 반려동물 검출기 개선

YOLO가 제공한 모델은 성능의 개선이 필요하다. 개와 고양이는 비슷한 생김새 특징을 가지고 있는데 클래스가 따로 나뉘어 있어서 탐지 성능이 떨어진다. 또한 학습한 데이터로 사용한 MS coco 데이터셋[18]은 개와 고양이 클래스에 대해 정면 이미지가 많아서 반려동물의 뒷모습이나 웅크린 모습에 대해 정확하게 학습하고 탐지하지 못한다. 반려동물 클래스를 찾아내는 객체 검출기의 성능이 떨어지면 추적 모델

에도 영향을 끼쳐 성능이 떨어진다. 사용자에게 부정확한 반려동물 통계 정보를 전달하게 되는 문제가 생긴다.

이를 해결하기 위해 YOLO의 개와 고양이 클래스를 통합하여 반려동물이라는 하나의 클래스를 생성했다. 또한 추가적인 데이터 수집을 진행하여 학습 이미지로 사용했다. MS coco 데이터셋[18]에서 개와 고양이 이미지 8,000장을 사용했으며, AIHub ‘반려동물 구분을 위한 동물 영상’ 데이터셋[21], kaggle ‘개와 고양이 분류’ 데이터셋[22], 다양한 자제의 반려동물 크롤링 이미지 10,000장을 추가로 수집하여 사용했다. 개와 고양이 이미지 18,000장을 이용하여 반려동물의 단일 클래스로 학습을 진행했다. 학습(Train) : 검증(Validation) = 8 : 2의 비율을 이용했으며, 여러 차례 학습 파라미터 조정을 통해 최적의 모델을 학습했다.

학습을 마친 검출기 성능을 평가하기 위해 Test 데이터를 수집했다. 학습에 사용하지 않은 개와 고양이 이미지 900장과 이와 상관없는 Noise 이미지 100장을 이용하여 성능 평가를 진행했다. 검출기 성능을 비교하는 평가 지표로는 재현율(Recall)과 정밀도(Precision)를 사용했다. 재현율은 실제 값이 긍정(Positive)인 것 중 모델이 긍정이라고 예측한 것의 비율을 의미하며, 정밀도는 모델이 긍정이라고 예측한 것 중 실제로 긍정인 것의 비율을 의미한다.

실험은 다음과 같은 방법으로 진행한다. 기존 YOLO 객체 검출기(Baseline)와 데이터를 수집하여 새롭게 학습한 YOLO 객체 검출기(Improved)의 탐지 임계값을 동일하게 설정하여 객체 탐지를 수행한다. 최종적으로 개와 고양이 반려동물 이미지 900장 중 탐지한 데이터 수(True Positive)와 탐지하지 못한 데이터 수(False Negative), Noise 이미지 100장 중 탐지한 데이터 수(False Positive)와 탐지하지 못한 데이터 수(True Negative)를 구해 재현율과 정밀도를 계산한다.

1,000장의 이미지를 활용하여 반려동물 검출기 성능을 평가한 결과, 그림 7과 같이 개선한 모델(Improved)에서 높은 정밀도와 재현율을 확인할 수 있다(표 1). 기존 모델(Baseline)보다 향상된 결과이다. 실제 어플리케이션에서 객체 탐지 임계값은 0.5로 설정하였으며, 이는 원하는 값으로 세부 조정할 수 있다.

#### 4-2 다중 반려동물 추적기 개선 및 분석 알고리즘

DeepSort는 동일한 객체의 추적이 장시간 끊겼다가 재시작될 경우 새로운 객체로 인식하여 다른 ID 값을 부여한다. 본 연구에서는 객체 분류기 VGG를 사용하여 이러한 문제를 극복했다. 웹 서비스 시작 부분의 등록 페이지에서 반려동물 이름과 영상을 입력하여 최초 1회 VGG를 학습한다. DeepSort의 추적이 시작되는 부분에서 학습한 VGG를 사용하여 추적이 진행될 때마다 동일 이름을 반려동물 객체에 부여한다. 기존의 객체 추적 모델은 다중 객체 추적이 불가능하였으나, 제안하는 연구를 통해 반려동물 수와 영상 출입과 관계없이 효과적인 추적을 수행한다.

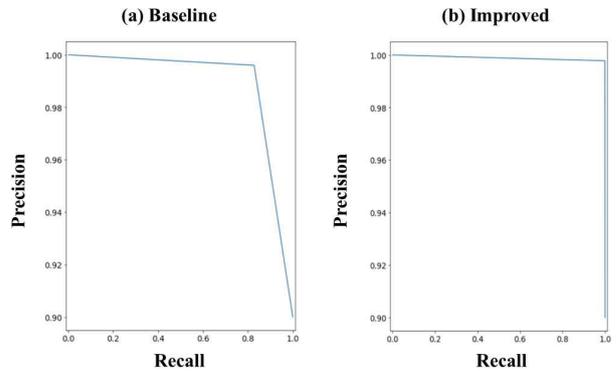


그림 7. 정밀도-재현율 그래프  
Fig. 7. Precision-Recall Curve

표 1. 개선 전후 YOLO v5 모델의 성능 비교

Table 1. Comparison to YOLO v5 models

Confidence Threshold	Pet Detection Model	Recall	Precision
0.2	YOLO v5 (baseline)	90.6%	99.5%
	YOLO v5 (improved)	100%	97.8%
0.3	YOLO v5 (baseline)	88.5%	99.6%
	YOLO v5 (improved)	100%	98.3%
0.4	YOLO v5 (baseline)	86.2%	99.6%
	YOLO v5 (improved)	99.9%	98.9%
0.5	YOLO v5 (baseline)	82.8%	99.6%
	YOLO v5 (improved)	99.9%	99.8%
0.6	YOLO v5 (baseline)	76.1%	99.7%
	YOLO v5 (improved)	99.9%	100%
0.7	YOLO v5 (baseline)	68.3%	99.8%
	YOLO v5 (improved)	99.4%	100%
0.8	YOLO v5 (baseline)	52.3%	99.8%
	YOLO v5 (improved)	95.1%	100%

객체 분류기 VGG의 학습 방식은 다음과 같다. 각 반려동물의 영상을 입력받으면, YOLO를 이용하여 반려동물 클래스에 해당하는 객체를 탐지하고 잘라내어 이미지로 저장한다(그림 8). 잘라낸 이미지를 이용하여 ImageNet 데이터셋[23]으로 사전 학습된 VGG의 전이 학습을 수행한 후, DeepSort에 삽입하여 객체 분류기로 사용한다. 반려동물 정보가 변경될 경우 등록 페이지에서 반려동물 객체 정보를 추가하거나 삭제할 수 있으며, 이때 VGG를 다시 학습하여 DeepSort에 반영한다.

객체 분류기 VGG의 성능을 측정하는 실험을 진행했다. 성능 측정 시 사용한 데이터는 서로 다른 고양이 두 마리의 영상을 각각 YOLO로 탐지하고 잘라낸 이미지이다. 한 마리당 850장의 이미지를 사용하여 총 1,700장을 학습에 사용하였다. 이때 학습(Train) : 검증(Validation) = 8 : 2의 비율로 학습을 진행했다. 분류기 성능을 평가하기 위해 한 마리당 60장의 이미지를 사용하여 총 120장의 이미지로 Test를 수행했다. 표 2의 결과를 보면 VGG가 높은 재현율과 정밀도를 보

이며 효과적으로 기능하고 있음을 알 수 있다.

YOLO+DeepSort 결과로 객체 탐지 상자가 그려진 영상과 추적 결과 텍스트(.txt)파일이 저장된다. 객체 추적 결과는 다음의 6개의 값으로 이루어져 있다.

$$[f_i, i, x_i, y_i, w_i, h_i] \text{ where } i = 1, 2, 3, \dots, N, \quad (5)$$

$f$ 는 탐지된 프레임 번호  $i$ 는 반려동물 객체에 매칭된 인덱스,  $x$ 와  $y$ 는 각각 객체의 좌측 픽셀 좌표와 상단 픽셀 좌표,  $w$ 와  $h$ 는 객체의 너비와 높이를 나타낸다.  $N$ 은 영상에 등장하는 전체 반려동물의 수를 의미한다. 이때, 탐지 임계값 0.75 이상에 해당하는 값만 결과로 인정한다.

DeepSort의 수행이 끝난 후에는 추적 결과로 출력된 프레임별 객체 탐지 상자 정보를 분석하여 사용자에게 필요한 형태로 제공한다. 분석 알고리즘을 이용하여 반려동물이 이동한 시간, 반려동물이 특정 위치에 머무른 시간과 횟수 정보가 담긴 json 파일을 생성하고, 시각화 알고리즘을 이용하여 반려동물의 움직임을 시각화한 이미지(그림 9)를 생성한다. 완성된 json 파일과 이미지는 웹 페이지를 통해 사용자에게 전달한다.

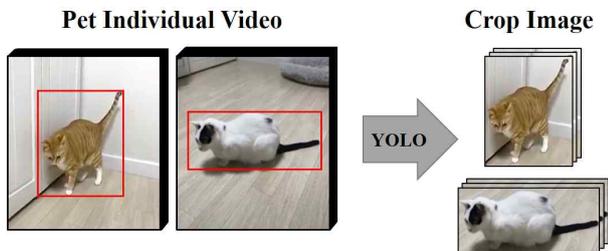


그림 8. 각 반려동물의 짧은 영상을 프레임별 잘린 이미지로 저장  
 Fig. 8. Save a short video of each pet as a frame-by-frame cropped image

표 2. 객체 분류기 VGG 성능 평가

Table 2. Object Classifier VGG Performance Evaluation

	Precision	Recall	F1-score
cat 1	94%	100%	97%
cat 2	100%	93%	97%

### 4-3 반려동물 움직임 분석

본 연구에서 제안한 인공지능 웹 서비스는 객체 분류기 VGG, 객체 추적기 DeepSort, 객체 검출기 YOLO를 사용한다. 반려동물 영상을 입력으로 받아 컴퓨터비전 기술로 이를 분석한다. 최종적으로 반려동물이 이동한 시간, 반려동물이 특정 위치에 머무른 횟수 및 시간 정보, 반려동물의 움직임을 시각화한 이미지가 웹 페이지를 통해 사용자에게 제공된다.

본 연구의 인공지능 모델 성능 측정에 사용한 5개의 테스트 케이스는 다음과 같다. 첫 번째 시나리오는 한 마리의 반려동물이 특정 구역에 계속 머무르는 경우이다. 두 번째 시나

표 3. 실제 측정 결과와 모델의 측정 결과 비교

Table 3. Compare actual measurements with AI model measurements

Scenario 1.

	Actual Measurements	AI Measurements
Move Time	1 s	1 s
Stay Time	32.554 s	31.733 s
Number of Visits	1	1

Scenario 2.

	Actual Measurements	AI Measurements
pet 1	Move Time	8 s
	Stay Time(1)	9.689 s
	Number of Visits(1)	1
pet2	Move Time	0 s
	Stay Time(2)	12.878 s
	Number of Visits(2)	1

Scenario 3.

	Actual Measurements	AI Measurements
pet 1	Move Time	4.3 s
	Stay Time(1)	7.09 s
	Number of Visits(1)	2
pet2	Move Time	2 s
	Stay Time(1)	0 s
	Number of Visits(1)	0

Scenario 4.

	Actual Measurements	AI Measurements
pet 1	Move Time	1.4 s
	Stay Time(2)	0 s
	Number of Visits(2)	0
pet2	Move Time	7 s
	Stay Time(2)	13.349 s
	Number of Visits(2)	2

Scenario 5.

	Actual Measurements	AI Measurements
pet 1	Move Time	7.365 s
	Stay Time(1)	5.337 s
	Number of Visits(1)	1
	Stay Time(2)	2.861 s
	Number of Visits(2)	1
pet2	Move Time	21 s
	Stay Time(1)	17.05 s
	Number of Visits(1)	2
	Stay Time(2)	7.624 s
	Number of Visits(2)	1

\* ( ) denotes an index of the area

리오는 두 마리의 반려동물 중 한 마리는 특정 구역1에 들어간 후에 일정 시간 후 나오며, 다른 한 마리는 특정 구역2에 계속 머무는 경우이다. 세 번째 시나리오는 두 마리의 반려동물이 영상에 등장할 때, 그중 한 마리가 특정 구역 1에 들어간 후 일정 시간이 지나고 재방문하는 경우이다. 네 번째 시나리오는 두 마리의 반려동물이 영상에 등장할 때, 그중 한 마리가 특정 구역 2에 들어간 후 일정 시간이 지나고 재방문하는 경우이다. 다섯 번째 시나리오는 두 마리의 반려동물이 모두 특정 구역 1과 2를 방문하는 경우이다.

각 시나리오에 따라 인공지능 모델의 성능을 측정해보면 표 3과 같다. 실제 측정 결과와 모델의 측정 결과가 거의 일치하는 것을 볼 수 있다. 전체 테스트 케이스에서 반려동물이 이동하거나 특정 위치에 머무른 시간은 0.5초 미만의 오차가 발생하였으며, 방문한 횟수는 정확히 일치한다. 본 연구가 충분히 가치 있는 결과를 도출하였음을 알 수 있다.



그림 9. 반려동물 움직임 시각화 예시

Fig. 9. Pet movement visualization

## V. 결론

반려동물 양육 가정이 증가함에 따라 이들의 걱정과 부담을 덜어줄 기술의 필요성이 커지고 있다. 본 연구는 반려동물의 움직임을 추적 및 분석하여 사용자에게 제공한다. 다중 객체 추적 기술을 적용하여 여러 마리의 반려동물을 효과적으로 분석할 수 있으며, 서비스 형태를 웹 어플리케이션으로 제작하여 사용자가 편리하게 사용할 수 있다. 배경 연구를 통해 정확한 다중 객체 추적 및 분석이 불가능한 현 기술의 한계점을 제시했다. 이를 해결하기 위해 컴퓨터 비전 기술을 응용하여 반려동물 객체 추적 모델의 성능을 향상하는 방안을 연구했으며, 반려동물 움직임 정보를 분석하는 알고리즘을 개발했다. 일련의 과정을 웹 어플리케이션 형태로 제공하여 사용자가 편리하게 사용하도록 했다.

본 연구에서 제안한 인공지능 모델은 다양한 방식으로 활용할 수 있다. 반려동물 홈 CCTV와 같은 팻테크 기기에 적용하여 반려동물 양육 가정에서 사용할 수 있으며, 동물 관련 기관에서 모니터링 용도로 사용할 수 있다. 또한 인공지능

모델의 학습 데이터를 조정하여 새로운 분야에 적용할 수 있다. 반려동물 산업이 점차 확장됨에 따라 본 연구가 제안한 인공지능 모델이 효과적으로 활용될 것을 기대한다.

## 감사의 글

이 논문은 전남대학교 학술연구비(과제번호: 2021-2264) 지원에 의하여 연구되었음

## 참고문헌

- [1] Chosun Biz. Petech develops in the direction of controlling devices by pets themselves [Internet]. Available: <https://biz.chosun.com/industry/company/2022/04/16/H3O3N273W5FJPLGWBRNWFBJUTE/>
- [2] KB Financial Group Management Research Institute. 2021 Pet Report [Internet]. Available: <https://www.kbfg.com/kbresearch/report/reportView.do?reportId=2000160>
- [3] Petpeotalk. Dogibogi Application [Internet]. Available: <https://dogibogi.co.kr/>
- [4] National Institute of Korean Language Standard Dictionary [Internet]. Available: <https://stdict.korean.go.kr/main/main.do>
- [5] Korean Animal Welfare Association. What's wrong with the word "pet"? [Internet]. Available: <https://www.animals.or.kr/campaign/friend/53083>
- [6] Hanwha Journal. The era of 10 million pets! Pet Lifestyle Guide [Internet]. Available: <https://m.hanwha.co.kr/media/discover/view.do?seq=2945>
- [7] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," *Neural Computation*, Vol. 1, No. 4. pp. 541-551, December 1989. <https://doi.org/10.1162/neco.1989.1.4.541>
- [8] Architectural Institute of Korea. Architectural Glossary, Convolutional Neural Network [Internet]. Available: <https://terms.naver.com/entry.naver?docId=6046497&cid=67350&categoryId=67350>
- [9] Telecommunications Technology Association. IT Glossary, Convolutional Neural Network [Internet]. Available: <https://terms.naver.com/entry.naver?docId=3686122&cid=42346&categoryId=42346>
- [10] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv, 2014.

<https://doi.org/10.48550/arXiv.1409.1556>

- [11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," arXiv, 2015. <https://doi.org/10.48550/arXiv.1506.02640>
- [12] Scienceall. Science Encyclopedia, object tracking [Internet]. Available: <https://www.scienceall.com/>
- [13] N. Wojke, A. Bewley, and D. Paulus, "Simple Online and Realtime Tracking with a Deep Association Metric." arXiv, 2017. <https://doi.org/10.48550/arXiv.1703.07402>
- [14] K. Zhou, Y. Yang, A. Cavallaro, and T. Xiang, "Omni-Scale Feature Learning for Person Re-Identification." arXiv, 2019. <https://doi.org/10.48550/arXiv.1905.00953>
- [15] Y. H. Jo, H. J. Lee, and Y. H. Kim, "Implementation of a Classification System for Dog Behaviors using YOLI-based Object Detection and a Node.js Server," *The Journal of Korea Institute of Convergence Signal Processing*, Vol. 21, No. 1, pp. 29-37, March 2020. <https://doi.org/10.23087/jkicps.2020.21.1.005>
- [16] M.-H. Lee, J.-Y. Kang, and S.-J. Lim, "Design of YOLO-based Removable System for Pet Monitoring," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 24, no. 1, pp. 22-27, January 2020. <https://doi.org/10.6109/jkiice.2020.24.1.22>
- [17] Github. Gleen jocher, yolov5 [Internet]. Available: <https://github.com/ultralytics/yolov5>.
- [18] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, ... and Piotr Dollár, "Microsoft COCO: Common Objects in Context," arXiv, 2014. <https://doi.org/10.48550/arXiv.1405.0312>
- [19] Github. Nicolai Wojke, Deep SORT [Internet]. Available: [https://github.com/nwojke/deep\\_sort](https://github.com/nwojke/deep_sort)
- [20] Github. Mikel Brostrom, Yolov5 Deepsort OSNet [Internet]. Available: [https://github.com/mikel-brostrom/Yolov5\\_DeepSort\\_OSNet](https://github.com/mikel-brostrom/Yolov5_DeepSort_OSNet)
- [21] AI Hub. Animal video for classification of companion animals [Internet]. Available: <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=59>
- [22] Kaggle. Dogs vs. Cats [Internet]. Available: <https://www.kaggle.com/competitions/dogs-vs-cats/data>
- [23] Stanford Vision Lab, ImageNet [Internet]. Available: <https://www.image-net.org/>

### 윤경윤(Kyung-Yoon Yoon)



2019년~2022년: 전남대학교  
AI융합대학  
IoT인공지능융합전공  
학사과정

※ 관심분야 : 컴퓨터 비전(Computer Vision), 인공지능(Artificial Intelligence) 등

### 설재형(Jae-Hyung Seol)



2015년~2022년: 전남대학교 AI융합대학  
IoT인공지능융합전공  
학사과정

※ 관심분야 : 컴퓨터 비전(Computer Vision), 인공지능(Artificial Intelligence) 등

### 조영준(Yeong-Jun Cho)



2014년 : 광주과학기술원 정보통신공학  
(공학석사)  
2018년 : 광주과학기술원  
전기전자컴퓨터공학  
(공학박사-컴퓨터비전)

2018년~2020년: 현대모비스 데이터사이언스팀 책임  
2020년~현재: 전남대학교 소프트웨어공학과 조교수  
※ 관심분야 : 컴퓨터비전(Computer Vision), 딥러닝(Deep Learning), 스마트팩토리 등