



# SSD 알고리즘 기반 MI-FL을 적용한 회전 불변의 다중 객체 검출 시스템 구현

박수빈\*, 임혜연\*, 강대성\*\*

## Implementation of Rotating Invariant Multi Object Detection System Applying MI-FL Based on SSD Algorithm

Su-Bin Park\*, Hye-Youn Lim\*, and Dae-Seong Kang\*\*

이 논문은 2017 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2017R1D1A1B04030870)

### 요 약

최근 CNN을 기반으로 한 객체 검출 기술의 연구가 활발하다. 객체 검출 기술은 자율주행차, 지능형 영상분석 등에서 중요한 기술로 사용된다. 본 논문에서는 CNN 기반의 객체 검출기 중 하나인 SSD(Single Shot Multibox Detector)에 MI-FL(Moment Invariant-Feature Layer)을 적용하여 회전 변형에 강인한 객체 검출 시스템을 제안한다. 먼저 VGG 네트워크를 기반으로 입력 이미지의 특징을 추출한다. 그 후 총 6개의 특징 계층 (Feature layer)을 적용하여 객체의 위치 정보와 종류를 예측해 경계 박스들을 생성한다. 그 후 NMS 알고리즘을 이용해 가장 객체일 확률이 높은 경계 박스를 얻는다. 하나의 객체 경계 박스가 정해지면 MI-FL을 이용해 해당 영역의 불변 모멘트 특징을 추출하여 미리 저장하고 학습한다. 이후 검출 과정에서 미리 저장해둔 불변 모멘트 특징 정보를 이용해 검출함으로써 회전된 이미지에 대해 기존 방법보다 더 강인한 검출이 가능하다. 기존의 SSD와 MI-FL을 적용한 SSD의 비교를 통해 약 4~5%의 성능 향상을 확인하였다.

### Abstract

Recently, object detection technology based on CNN has been actively studied. Object detection technology is used as an important technology in autonomous vehicles, intelligent image analysis, and so on. In this paper, we propose a rotation change robust object detection system by applying MI-FL (Moment Invariant-Feature Layer) to SSD (Single Shot Multibox Detector) which is one of CNN-based object detectors. First, the features of the input image are extracted based on the VGG network. Then, a total of six feature layers are applied to generate bounding boxes by predicting the location and type of object. We then use the NMS algorithm to get the bounding box that is the most likely object. Once an object bounding box has been determined, the invariant moment feature of the corresponding region is extracted using MI-FL, and stored and learned in advance. In the detection process, it is possible to detect the rotated image more robust than the conventional method by using the previously stored moment invariant feature information. The performance improvement of about 4 ~ 5% was confirmed by comparing SSD with existing SSD and MI-FL.

### Keywords

object detection, CNN, SSD, moment invariant

\* 동아대학교 전자공학과  
- ORCID<sup>1</sup>: <https://orcid.org/0000-0001-6346-8960>  
- ORCID<sup>2</sup>: <https://orcid.org/0000-0002-5189-461x>  
\*\* 동아대학교 전자공학과 교수(교신저자)  
- ORCID: <https://orcid.org/0000-0003-0186-2430>

• Received: Jan. 28, 2019, Revised: May 01, 2019, Accepted: Apr. 04, 2019  
• Corresponding Author: Dae-Seong Kang  
Dept. of Dong-A University, 37 NaKdong-Daero 550, beon-gil saha-gu,  
Busan, Korea,  
Tel.: +82-51-200-7710, Email: [dskang@dau.ac.kr](mailto:dskang@dau.ac.kr)

## 1. 서 론

최근 인공지능망의 한 종류인 CNN(Convolutional Neural Network)을 이용한 영상처리 연구가 활발하다. 그중 객체 검출 기술은 지능형 영상분석 시스템, 자율 주행 자동차 등 많은 실생활 분야에 적용되고 있다. 각종 영상으로부터 객체를 검출하고 인식하는 기술은 인간의 안전과 편의를 위해 매우 중요한 기술이다.

기존의 객체 검출 기술들은 목적에 따라 차이는 있으나 다양한 특징 추출 기법을 통해 이루어졌다. 주로 영상 내 픽셀 값들을 이용하여 객체의 특징을 찾아내며 대표적으로 Haar-like[1], HOG(Histogram of Oriented Gradient)[2] 방법들이 있다. 그러나 위 방법들은 객체의 특징 형태를 직접 설계하기 때문에 성능에 한계가 있다.

최근에는 대부분의 객체 검출에 CNN이 사용되고 있다. CNN은 합성곱 필터(Convolution filter)를 이용해 객체의 의미 있는 특징들을 스스로 추출하고 학습함으로써 기존의 특징 추출 알고리즘들에 비해 월등한 성능을 보였다[3]. 대표적인 CNN 구조로는 AlexNet[4], VGG-Net[5], GoogLeNet[6], ResNet[7] 등이 있다.

한편 CNN 학습 기반의 객체 검출 방법에는 R-CNN(Region-CNN)[8] 방법이 주로 사용되고 있는데 R-CNN이란 입력 영상으로부터 객체가 있을 만한 후보 영역들을 먼저 찾은 후 해당 영역에 있는 객체를 인식하는 2가지의 단계로 이루어진 객체 검출 방법이다. 현재 객체 검출 기술은 영상 내부 객체의 위치와 종류를 한 번에 검출하는 단계로 발전하였으며 대표적인 R-CNN 방식의 검출 기술에는 Faster R-CNN[9], YOLO(You Only Look Once)[10], SSD(Single Shot Multibox Detector)[11] 등이 있다. 위 검출 기술들은 검출 성능이 뛰어나지만 영상 내 객체의 각도 변화나 회전이 생길 경우 성능이 떨어지는 경우가 있다. 실제 카메라 환경에서는 카메라의 각도나 위치에 따른 회전과 각종 변화가 불가피하기 때문에 검출 성능이 저하된다.

본 논문에서는 위에 언급한 검출 기술 중 하나인 SSD에 MI-FL(Moment Invariant-Feature Layer)를 적

용한 회전 및 변화에 좀 더 강인한 객체 검출 시스템을 제안한다.

학습 및 검출에 사용할 데이터 셋은 객체 검출 연구의 성능 지표로서 사용되는 Pascal VOC 데이터 셋을 사용하였다[12]. 특징 추출에 사용될 CNN 구조는 VGG-16 네트워크를 사용하였다. 특징 추출 과정은 ImageNet 데이터셋으로 사전에 학습된 VGG-16 모델의 가중치를 이용해 새로운 데이터셋에 대해 재학습하는 미세 조정(Fine-tune) 기법을 사용하였다. 특징 추출 후 6개의 특징 계층(Feature layer)을 통해 객체 후보 영역과 영역 내 객체의 클래스를 예측한다. 또한 학습 및 검출 과정에서 MI-FL을 적용하여 객체의 불변 모멘트 특징 값을 추출하여 검출에 사용함으로써 회전이나 각도 변화가 일어난 객체의 검출 성능을 향상시킬 수 있다. 기존의 SSD와 YOLO 그리고 MI-FL을 추가한 SSD 검출기의 비교 실험을 통해 제안하는 방법의 성능을 확인하였다.

## II. 관련 이론

### 2.1 VGG Network

VGG 네트워크는 구조적으로 간단하게 구성되어 있으면서도 좋은 성능을 가진 CNN 구조이다. 이해하기 쉽고 변형이 쉬워 현재 대부분 CNN 기반 연구들의 기본 구조로 사용되고 있다. 기본적인 CNN 구조와 같이 합성곱 계층(Convolution layer), 풀링 계층(Pooling Layer), 그리고 전결합 계층(Fully Connected Layer)으로 구성되어 있다. 구조에 사용된 합성곱 필터의 크기는 모두 3x3의 크기로 고정되어 있어 타 구조들에 비해 간단하게 구현할 수 있다. 합성곱 계층의 개수에 따라 여러 종류가 있는데 주로 사용되는 구조는 총 16개의 층을 가진 VGG-16과 19개의 층을 가진 VGG-19이다. 그림 1은 VGG-16의 구조를 나타낸 그림이다.

그림과 같이 총 5개의 합성곱 구역(Convolution block)이 있고 각 구역은 2~3개의 합성곱 계층과 1개의 최대 풀링 계층(Max Pooling Layer)으로 구성되어 있다. 총 13개의 합성곱 계층을 거쳐 추출된

특징 맵을 평탄한 계층(Flatten Layer)에서 1차원 벡터 형태로 변환하고 이후 2개의 전결합 계층을 거쳐 학습과 분류를 수행한다.

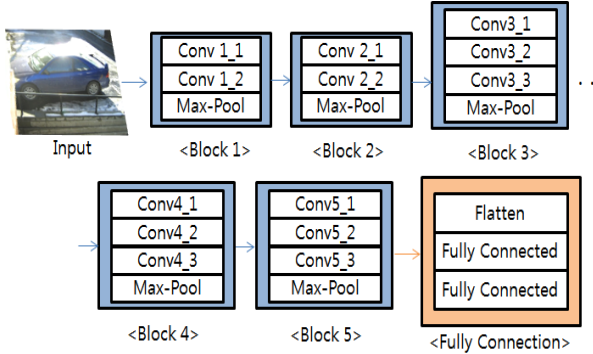


그림 1. VGG-16의 구조  
Fig. 1. Structure of VGG-16

## 2.2 SSD

SSD는 한 번의 네트워크 진행으로 다중 객체를 검출할 수 있는 검출기이다. 기존에 있던 R-CNN 방법의 경우 RPN(Region Proposal Network)이라고 불리는 객체 후보 영역을 추출하기 위한 네트워크와 CNN 기반의 객체 분류를 위한 네트워크를 따로 설계하여 검출이 2단계에 걸쳐 이루어진다. 때문에 성능은 뛰어났으나 속도가 다소 느린 단점이 있었다. SSD는 2단계의 검출 과정을 1번의 네트워크로 줄여 성능 대비 속도를 크게 개선하였다[9][11].

SSD에서는 기반 네트워크로 VGG 네트워크를 사용하였으며 기반 네트워크 뒤에 특징 계층이라 불리는 여러 크기의 특징 맵들을 이용하여 각각 다른

크기의 객체들을 검출한다. 그림 2는 SSD의 객체 검출 과정을 나타낸 그림이다.

그림과 같이 VGG 네트워크의 마지막 합성곱 계층인 Conv 5\_3 layer를 거친 38x38 크기의 최종 특징 맵에 여러 크기의 특징 계층들을 적용하여 객체를 검출한다. 특징 계층의 특징 맵들은 미리 정해진 크기의 셀(Cell)로 분할되어 있고 각 셀에는 k개의 정해진 개수의 경계 박스들이 있다. 이 경계 박스들은 다양한 모양과 비율을 가지고 있다. 한 칸의 셀에서 k개의 경계 박스를 생성하고 각 경계박스마다 객체의 존재 유무와 객체의 종류를 분류한다. 그 후 NMS(Non-Maximum Suppression) 알고리즘을 통해 가장 객체에 가까운 하나의 경계 박스를 제외하고 모두 제거되면서 검출이 이루어진다. 그림 3은 특징 계층의 경계 박스 검출의 예를 나타낸다.

그림 3과 같이 8x8 크기의 셀을 가진 특징 맵에서는 비교적 작은 객체를 검출하고 4x4 크기의 셀을 가진 특징 맵에서는 큰 객체를 검출하는 것을 알 수 있다[11].

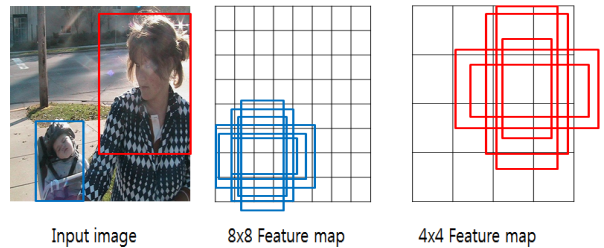


그림 3. Feature layer의 경계 박스 검출  
Fig. 3. Detection bounding box of feature layer

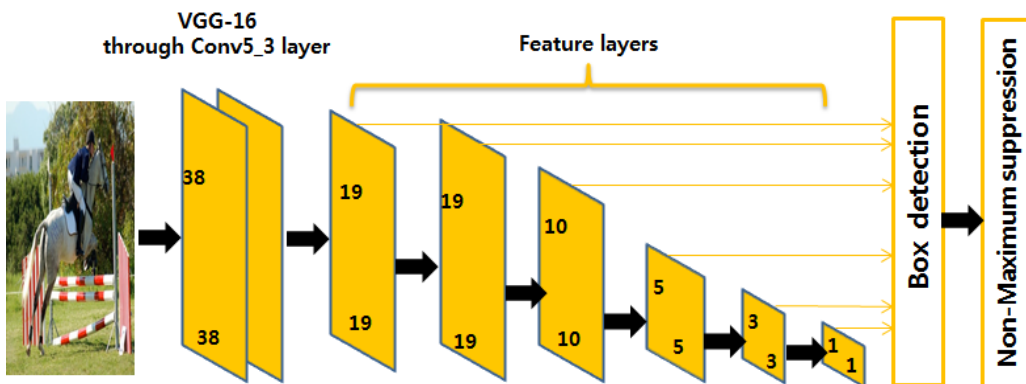


그림 2. SSD의 객체 검출 과정  
Fig. 2. Object detection process of SSD

### 2.3 Hu 불변 모멘트

Hu의 불변 모멘트는 객체의 모양을 특징으로 사용하는 방법 중 하나로 영상 내 객체의 전체 픽셀 값과 위치값을 이용해 고유한 값을 획득해 다른 객체들과 구별할 수 있다. 따라서 불변 모멘트 특징은 영상의 회전, 이동, 크기변화 등 각종 변화에 불변한 특징 값을 표현할 수 있다.

x축과 y축을 가진 2차원 영상을 함수  $f(x,y)$ 로 나타낼 때 영상의 기하학적 모멘트는 식 (1)과 같이 정의된다.  $(p+q)$ 를 모멘트의 차수라고 부르며  $M, N$ 은 픽셀의 가로축과 세로축의 크기이다. 기하학적 모멘트는 영상의 이동 변환 시 그 값이 변하는 단점이 있다. 따라서 식 (2)와 같이 영상의 무게 중심을 고려하여 계산한 중심 모멘트를 사용한다. Hu는 모멘트 차수가 3차 이하인 중심 모멘트 값들을 조합하여 식 (3)과 같이 크기에 불변한 특성을 기반으로 한 총 7개의 불변 모멘트 특징을 만들었다.

$$m_{pq} = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} x^p y^q f(x,y) \quad (1)$$

$$\mu_{pq} = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \quad (2)$$

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00} \left(1 + \frac{i+j}{2}\right)} \quad (3)$$

표 1. Hu 불변 모멘트 특징

Table 1. Hu moment invariant features

$I_1$	$\eta_{20} + \eta_{02}$
$I_2$	$(\eta_{20} - \eta_{02})^2 + (2\eta_{11})^2$
$I_3$	$(\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$
$I_4$	$(\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$
$I_5$	$(\eta_{30} - 3\eta_{12})^2(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]$
$I_6$	$(\eta_{20} - \eta_{02})[(\eta_{03} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$
$I_7$	$(3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]$

이는 크기를 정규화한 중심 모멘트의 비선형 조합들로 구성되어 있으며 영상의 회전, 크기 및 위치 변화에 불변한 특징 값을 가진다. 표 1은 7개의 불변 모멘트 특징이다[13][14].

### III. 제안하는 방법

그림 4는 본 논문에서 제안하는 네트워크 구조의 순서도를 나타낸다. 먼저 입력 영상은 300×300의 크기로 변환한다. 학습 과정에서는 먼저 ImageNet 데이터셋을 기반으로 미리 학습된 VGG-16 구조의 마지막 합성곱 계층까지의 가중치를 이용하여 입력 영상의 특징을 추출한다. 생성된 특징 맵에 다시 Pascal VOC 데이터셋으로 미세 조정한다. 미세 조정 기법은 미리 학습된 가중치를 이용해 새로운 데이터를 재학습하는 방법으로서 학습 시간을 크게 줄일 수 있다. 미세 조정 과정에서 6개의 특징 계층들을 통해 객체 후보 영역과 그 안의 객체의 클래스를 예측한다. 그 후 예측된 모든 객체 후보 영역들 중 가장 객체일 확률이 높은 영역들을 NMS 알고리즘을 통해 얻어 낸다. 생성된 하나의 객체 영역에 대해 MI-FL을 적용하여 불변 모멘트 특징 값을 구하여 특징 정보들을 미리 저장해두고 학습을 마친다.

검출 과정에서 6개의 특징 계층들을 통해 각 후보 영역 내 객체의 클래스를 예측할 때 미리 저장해둔 불변 모멘트 특징 정보와의 매칭을 통해 검출하면 회전 변형된 입력 영상에서도 보다 정확하게 객체를 검출할 수 있다.

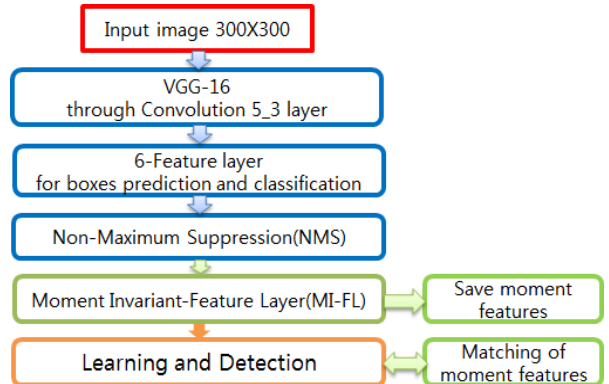


그림 4. 제안하는 구조의 순서도  
Fig. 4. Flowchart of proposed structure

## IV. 실험 결과

### 4.1 Fine-tune VGG-16

미세 조정 기법은 최근 CNN 연구에서 주로 사용되는 학습 방법이다. 많은 양의 데이터에 대해 미리 학습된 CNN 구조의 가중치를 그대로 사용하여 또 다른 데이터를 학습하는 방법이다. ImageNet 데이터셋에 대해 학습된 가중치를 주로 사용하며, 많은 CNN 구조들의 개발자들이 제공하기 때문에 손쉽게 사용할 수 있다. 본 논문에서 사용된 SSD 역시 VGG-16 구조의 마지막 계층인 Convolution 5\_3 계층까지의 가중치를 사용하여 특징 맵을 생성한다.

### 4.2 6-Feature layer

SSD의 핵심은 기반 네트워크인 VGG-16 네트워크 이후 6개의 특징 계층들이다. 기존 VGG-16의 전 결합 계층 2개를 포함하여 점점 줄어드는 크기를 가진 6개의 계층들을 통해 여러 크기의 객체들을 검출하게 된다. 특징 계층을 통해 생성되는 특징 맵을 다중 스케일(Multi-Scale) 특징 맵이라고 하며 SSD는 이 특징 맵들을 다른 크기들로 만들어 크기가 큰 맵에서는 작은 물체를 검출하고 작은 맵에서는 큰 물체를 검출한다.

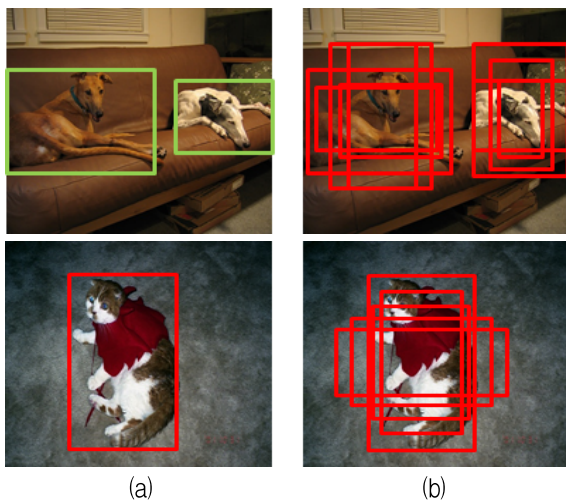


그림 5. Feature layer의 예측 경계 박스  
Fig. 5. Predictive bounding box of feature layer

각 특징 맵의 셀들은 k개의 경계 상자를 가지며 경계 상자들은 다양한 모양과 비율을 가지고 있다. 각 특징 계층들은 공통적으로  $3 \times 3 \times p$  크기의 필터를 사용한다. p는 채널의 수를 뜻한다.

본 논문에서는 각 특징 맵 하나의 셀에 4개 혹은 6개의 경계 상자를 가지고 있다. 그림 5는 VOC 데이터셋의 실제 라벨링 된 객체의 경계 상자 위치와 특징 계층들을 통해 예측된 경계 상자들을 시각화한 그림이다. (a)는 라벨링된 실제 객체의 위치 정보를 경계 상자로 나타내고 (b)는 예측된 경계 상자들을 나타낸다.

### 4.3 NMS

특징 계층을 통해 객체의 후보 영역들을 추출한 뒤 보다 정확한 하나의 후보 영역을 찾기 위해 NMS 알고리즘을 이용한다. NMS는 주어진 경계 상자들 중 IoU(Intersection of Union) 값이 가장 큰 상자를 남기고 나머지는 제거하는 방법이다. IoU란 2개의 경계 상자 중 겹치는 부분의 비율을 나타낸다. NMS의 과정은 크게 3단계로 나뉘는데 먼저 생성된 경계 상자 중 가장 클래스 점수가 높은 영역을 찾는다. 이후 그 영역과 나머지 영역들의 IoU를 계산하고 일정 임계값 이상인 영역은 제거하고 이하인 영역은 유지한다. 임계값 이상인 영역의 경우 똑같은 객체를 예측하고 있다고 판단하기 때문에 제거 된다. 그 후 남은 영역들에 대하여 하나의 영역이 남을 때까지 반복하면 가장 객체에 가까운 하나의 경계 상자를 찾을 수 있다.

### 4.4 MI-FL

본 논문에서 사용된 Hu 불변 모멘트 특징은 7개의 특징 중 크기, 회전, 위치에 불변인 특징인  $I_5$ ,  $I_6$ ,  $I_7$  3개의 특징만 사용하였다. 나머지 특징들은 값이 작기 때문에 약간의 차이에도 오인식 할 가능성이 높아 제외하였다.

NMS 알고리즘을 통해 객체 후보 영역이 정해지면 경계 박스의 좌표 값 정보를 가져와 해당 영역을 Grayscale로 변환한 뒤 불변 모멘트 특징을 계산

하여 라벨링 된 객체의 클래스와 함께 저장하여 검출 과정에서 사용한다.

#### 4.5 검출 결과

실험의 개발환경은 Python 기반의 Tensorflow를 사용하였고, GPU는 Geforce GTX980 Ti를 사용하였다. 데이터셋은 Pascal VOC 2007을 사용했다. VOC 2007은 객체 검출 분야의 성능 분석을 위해 주로 사용하는 데이터셋으로 배경을 포함해 총 21개의 클래스를 가지고 있고 학습 데이터에는 영상 내 모든 객체의 위치 정보와 클래스 정보가 라벨링 되어 있다. 총 이미지의 개수는 5,011개이며 이미지 내 총 객체의 수는 12,608개가 있다[12].

실험은 원본 테스트 이미지를 시작으로 이미지를 임의로 15도씩 회전시켜 15도, 30도 2번의 단계로 회전된 이미지에 대해서도 검출 하였다.

그림 6은 원본 이미지에 대한 검출 결과이다. 그림 7과 그림 8은 영상에 회전 변형이 일어난 후 검출이 되지 않은 경우와 제안하는 방법을 이용하여 검출에 성공하는 결과를 나타낸다. 각각 15도, 30도의 회전 변형이 일어난 경우에 대해 나타 내었다. 그림 7(a)는 기존의 SSD를 이용해 검출한 결과이고, 7(b)는 제안하는 구조를 이용해 검출한 결과이다. 그림 8(a)와 그림 8(b)도 동일하다.

표 2는 전체 실험 결과의 검출 성능을 판단하는 mAP(mean Average Precision)에 대한 표이다. mAP는 객체 검출 분야의 대표 성능 평가 지표로 객체 검출기가 검출한 정보들과 라벨링 된 실제 정보가 일치하는 비율을 나타낸다[12].

실험 결과의 비교에는 SSD와 YOLOv3[15]의 검출 결과와 MI-FL을 적용한 SSD를 비교하였다.

표 2. 실험 결과 mAP 비교

Table 2 Experimental results comparison of mAP

Methods	mAP(%)		
	Original	15 °	30 °
SSD	81.6	77.4	68.9
YOLOv3	82.3	77.9	70.2
MI-FL+SSD	81.6	80.4	73.7



그림 6. 원본 이미지의 검출 결과  
Fig. 6. Detection result of original image

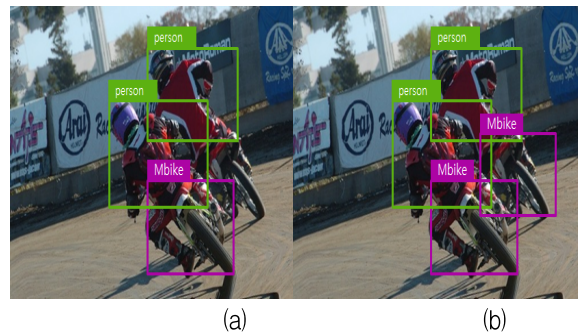


그림 7. 15도 회전 변형된 영상의 검출 결과  
Fig. 7. Detection result of 15 degree rotated image

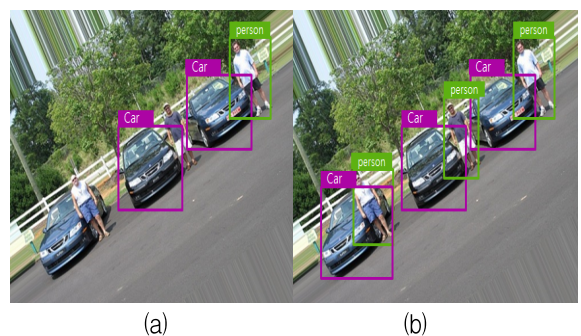


그림 8. 30도 회전 변형된 영상의 검출 결과  
Fig. 8. Detection result of 30 degree rotated image

## V. 결 론

본 논문에서는 회전 변형이 일어난 영상에서도 보다 정확하게 객체를 검출하기 위해 MI-FL을 적용

한 SSD 검출기를 제안하였다. 제안하는 구조를 통해 실험한 결과 회전된 이미지에 대해 기존의 SSD와 YOLO보다 더 좋은 성능의 검출 결과를 얻을 수 있었다. 각종 변화에 불변한 특징을 가지는 Hu 불변 모멘트 특징들을 검출에 이용함으로써 회전된 객체의 경우도 추가적으로 검출할 수 있었다.

실험 결과 크기가 큰 객체들은 회전이 일어나면 영상 내부 픽셀 값들의 변화가 매우 커져서 다소 효과가 없었으나, 주로 작은 크기의 객체들의 경우 회전이 일어나도 객체 영역 내부의 픽셀 값들의 변화가 적어 효과적이었다.

실험 결과 전체 검출 성능을 판단하는 mAP는 기존의 SSD를 사용했을 때와 비교하여 15도 회전 변형의 경우 약 3% 향상되었고, 30도 회전 변형의 경우 약 5%가량 향상되었다. SSD보다 약간 성능이 좋은 YOLOv3의 경우는 원본 이미지는 YOLOv3가 약 1% 성능이 좋았으나 회전된 이미지에 대해서는 제안하는 구조가 약 3%의 성능 향상을 보였다.

추후 연구에서는 실시간 영상에서의 적용 가능성에 대해 연구를 진행할 계획이다.

## References

- [1] P. Viola, and M. J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", Proc. IEEE CVPR, Vol. 1, pp. 511-518, Dec. 2001.
- [2] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection", IEEE Computer Society Conference on CVPR, pp. 886-893, June. 2005.
- [3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition", Proc. IEEE, pp. 2278-2323, Nov. 1998.
- [4] K. Alex, S. Ilya, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", NIPS, Vol. 1, pp. 1097-1105, Dec. 2012.
- [5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", In ICLR, May. 2015.
- [6] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions", Proc. IEEE, CVPR, pp. 1-9, June. 2015.
- [7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", Proc. IEEE, CVPR, pp. 770-778, June. 2016.
- [8] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", Proc. IEEE, CVPR, pp. 580-587, June. 2014.
- [9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks", NIPS, Dec. 2015.
- [10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection", Proc. IEEE, CVPR, pp. 779-788, June. 2016.
- [11] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "SSD: Single shot multibox detector", ECCV, pp 21-37, Oct. 2016.
- [12] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes (VOC) Challenge", IJCV, Vol. 88, No.2, pp. 303-338, June. 2010.
- [13] M. K. Hu, "Visual pattern recognition by moment invariants", IRE Trans, Vol. 8, pp. 179-187, Feb. 1962.
- [14] G.W. Kim, J.H. Park, H.Y. Lim, and Dae-Seong Kang, "A System Development on ASL Numbers Recognition Using Hierarchical Moment Invariants Based on Edge Image", JKIIIT, Vol. 13, No. 4, pp. 37-43, Apr. 2015.
- [15] J. Redmon, A. Farhadi, "Yolov3: An incremental improvement", arXiv preprint arXiv:1804.02767, Apr. 2018.

저자소개

박 수 빈 (Su-Bin Park)



2015년 8월 : 동아대학교  
전자공학과 (공학사)  
2015년 8월 ~ 현재 : 동아대학교  
전자공학과 석사과정  
관심분야 : 컴퓨터비전, AI

임 혜 연 (Hye-Youn Lim)



2013년 : 동아대학교  
전자공학과(공학박사)  
2013년 ~ 현재 : 동아대학교  
전자공학과 조교수  
관심분야 : 영상처리, 패턴인식

강 대 성 (Dae-Seong Kang)



1994년 5월 : Texas A&M 대학교  
전자공학과(공학박사)  
1995년 ~ 현재 : 동아대학교  
전자공학과 교수  
관심분야 : 영상처리, 패턴인식