



## 안개 제거가 객체 검출 성능에 미치는 영향 분석

### An Empirical Study of the Effect of Single Image Dehazing on Object Detection

정치윤\*\*<sup>ID</sup>, 하동휘\*\*\*<sup>ID</sup>, 문경덕\*<sup>ID</sup>, 김무섭\*<sup>ID</sup>

Chi Yoon Jeong, Donghee Ha, KyeongDeok Moon, and Mooseop Kim<sup>†</sup>

\*한국전자통신연구원 인공지능연구소 책임연구원, \*\*과학기술연합대학원대학교 인공지능 전공 부교수, \*\*\*과학기술연합대학원대학교 인공지능 전공 박사과정

\*Principal Researcher, Artificial Intelligence Research Lab., ETRI

\*\*Associate Professor, Artificial Intelligence Major, University of Science and Technology

\*\*\*Ph.D Course, Artificial Intelligence Major, University of Science and Technology

#### 요약

안개 제거 기술은 안개 제거 영상의 품질 향상에 초점을 맞추고 있어 모델 복잡도가 비약적으로 증가하고 있다. 하지만 안개 제거 기술은 다른 태스크의 전처리 과정으로 활용되어 안개 제거 성능과 태스크의 성능을 모두 고려하여 모델 복잡도의 균형을 이루는 것이 중요하다. 이를 위해, 본 논문에서는 안개 제거 기술의 성능과 객체 검출 성능과의 관계를 분석하는 연구를 수행하였다. 먼저 최신 안개 제거 모델을 사용하여 다양한 성능을 가지는 9개 안개 제거 모델을 생성하였다. 그리고 공개 데이터셋에 적용하여 안개 제거 영상을 생성한 후, 6개의 객체 검출 모델을 적용하여 검출 성능을 측정하였다. 실험 결과 객체 검출 성능은 안개 제거 성능에 비례하여 증가하지 않고, 객체 검출 성능에 영향을 주는 안개 제거 성능의 최소 및 최대 임계치가 존재함을 확인하였다. 이를 활용하면 안개 제거 성능과 객체 검출 성능 향상에 균형을 맞춘 효율적인 안개 제거 모델을 생성할 수 있을 것으로 기대된다.

**키워드** : 안개 제거, 객체 검출, 특징 융합 네트워크, 성능 비교 분석, 딥러닝

#### Abstract

Existing methods for image dehazing are focused on increasing the quality of restored images, so they require huge parameters. Considering that the main purpose of image dehazing is a preprocessing step for various tasks, it is important to balance the model complexity by considering the performance of image dehazing and that of the main tasks. Therefore, in this paper, we analyze the effect of image dehazing on object detection, a typical task of computer vision. We use nine dehazing models having different performances and the six object detection methods for analyzing the relation between the performance of dehazing and its influence on object detection. Experimental results showed that the performance of the object detection methods does not increase in proportion to the performance of the image dehazing and it converges when the performance of the image dehazing exceeds a certain level. By using the results, it is expected that an efficient dehazing model can be developed by considering the performance improvement in the main task and the model complexity.

**Key Words** : Single Image Dehazing, Object Detection, Feature Fusion Attention Network, Performance Analysis, Deep Learning

Received: Jan. 11, 2022  
Revised : Feb. 12, 2022  
Accepted: Feb. 23, 2022  
<sup>†</sup>Corresponding author  
(gomskim@etri.re.kr)

이 논문은 한국전자통신연구원 연구용 영비지원사업(기본사업)으로 수행한 인간중심의 자율지능시스템 원천기술연구 과제(22ZS1200)의 논문입니다.



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 noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. 서론

안개는 대기에 존재하는 스모그, 먼지, 입자들로 인해서 발생하는 대기 현상으로, 빛을 흡수하거나 산란시켜 영상의 품질을 열화시킨다[1]. 안개로

인한 영상의 품질 저하는 객체 인식, 객체 검출, 의미 영역 분할 등 다양한 컴퓨터 비전 기술들의 성능 저하를 유발한다[2]. 따라서, 안개 영상으로부터 안개를 분리하여 깨끗한 영상을 생성하는 안개 제거 기술에 관한 연구가 진행되었다[1-13]. 특히 주변 상황인식을 위해 영상 정보를 활용하는 자율주행 자동차, 드론 등 무인 이동체의 안정적인 동작을 위해 안개 제거 기술에 관한 관심이 급격히 증가하고 있다[14-16].

안개 제거 기술은 사전 지식 기반 방법과 딥러닝 기반 방법으로 구분할 수 있다[3]. 사전 지식 기반 방법들은 안개 영상들로부터 추출한 통계적인 특성 정보를 활용하여 안개 영역을 제거하였다. 대표적으로 안개가 존재하지 않는 영상의 일정 영역에서 밝기가 0의 값을 가지는 컬러 채널이 존재한다는 가정을 사용하는 다크 채널 프라이어(Dark Channel Prior : DCP) 방법[4]이 있다. 이 방법은 영상의 일정 영역에서 가장 낮은 값을 가지는 컬러 채널을 다크 채널로 정의하고, 이를 활용하여 안개 제거에 필요한 정보를 계산하였다. 하지만, 이런 가정은 대기의 빛과 유사한 색상을 가지는 객체가 존재할 때 성립되지 않는다[5]. 이처럼 사전 지식 기반 방법들은 모든 경우에 적용될 수 없으므로 안개 제거 성능 향상에 한계가 있다.

사전 지식 기반 방법의 문제점을 해결하기 위하여 최근 딥러닝을 사용한 안개 제거 기술들이 연구되고 있다. 딥러닝 기반 안개 제거 기술은 안개 영상으로부터 종단간(End-to-end) 학습을 통하여 안개가 존재하지 않는 깨끗한 영상을 복원하는 방법으로, 합성곱 기반 방법[11,12], 인코더-디코더 기반 방법[1,2,8-10], 적대적 생성 신경망(Generative Adversarial Network) 기반 방법[3,13] 등으로 구분된다. 이와 같은 딥러닝 기반 방법들은 기존의 사전 지식 기반 방법들에 비해서 높은 성능을 보여주었지만, 안개 제거 성능 향상에 초점을 두고 연구가 진행되어 안개 제거 모델의 복잡도가 점점 더 증가하고 있다.

안개 제거 기술은 객체 검출, 객체 인식 등 다양한 컴퓨터 비전 태스크의 성능 향상을 위한 전처리 과정으로 활용되기 때문에 주 태스크의 성능에 영향을 주지 않아야 한다. 딥러닝 기반 안개 제거 방법 중 높은 성능을 나타내는 FFANet[11], DuRN[12]의 경우 딥러닝 모델의 복잡도를 나타내는 학습 파라미터의 수가 각각 4.5M, 9M이다. 객체 검출에 사용되는 EfficientDet-D0 모델[17]의 학습 파라미터의 수인 3.9M 과 비교할 때, 안개 제거 방법들의 모델 복잡도가 주 태스크에 사용되는 객체 검출모델보다 더 높은 것을 확인할 수 있다. 따라서, 안개 제거 기술은 안개 제거 성능 향상도 중요하지만 다른 태스크의 전처리 과정으로 활용되는 점을 고려할 때 안개 제거 성능과 모델 복잡

도의 균형을 이루는 것이 중요하다. 하지만 현재까지 안개 제거 성능 향상에 관한 연구가 주를 이루고 있으며, 안개 제거 기술이 활용되는 태스크와 연계하여 성능을 비교 분석한 연구는 수행되지 않고 있다.

따라서 본 논문에서는 안개 제거 기술의 성능과 이를 활용하는 태스크 성능 간의 관계를 분석하는 연구를 수행하였다. 먼저 안개 제거 모델의 학습 파라미터를 조정하여 다양한 안개 제거 성능을 가지는 모델을 생성하였다. 그리고 객체 검출 분야의 대표적인 검출 알고리즘들과 안개가 존재하는 객체 검출 데이터셋을 사용하여 안개 제거 성능에 따른 객체 검출 모델의 성능 변화 정도를 분석하였다. 실험 결과 객체 검출 성능은 안개 제거 성능에 비례하여 증가하지 않고, 일정 수준 이상이면 객체 검출 성능이 수렴된다는 것을 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서는 안개 제거 기술과 관련된 연구 동향을 소개하고, 3장에서는 안개 제거 성능이 객체 검출 성능에 미치는 영향을 분석하기 위한 제안 방법을 기술한다. 4장에서는 실험 방법 및 결과를 제시하고, 5장에서는 결론 및 향후 연구 방향에 관해 논한다.

## 2. 안개 제거 기술 관련 연구

안개 제거 기술은 안개 현상을 물리적으로 모델링하는 식 1에 기반하고 있다[18].

$$I(x) = J(x)t(x) + A(1-t(x)) \quad (1)$$

위의 식에서  $x$ 는 영상에서의 좌표를 의미하며,  $I(x)$ 는 카메라로 획득되는 영상,  $J(x)$ 는 안개가 없는 깨끗한 영상,  $A$ 는 대기광(Atmospheric light),  $t(x)$ 는 거리에 따라 변화하는 전달량(Transmission map)을 의미한다. 기존의 사전 지식 기반 안개 제거 모델들은 전문가가 정의한 정보를 기반으로 영상 정보를 분석하고 전달량과 대기광을 추정함으로써 안개가 없는 영상을 복원하였다. 안개 제거를 위한 사전 지식은 모든 경우를 고려할 수 없으므로 성능 향상에 제한이 있으며, 이를 해결하기 위해서 딥러닝 기반의 안개 제거 기술에 관한 연구가 진행되었다.

초기 딥러닝 기반 안개 제거 기술들[5-7]은 기존의 사전 지식 기반 방법과 같이 딥러닝 네트워크를 사용하여 전달량을 추정한 후, 영상 정보로부터 추론된 대기광 정보와 통합하여 계산함으로써 안개를 제거하였다. DehazeNet[7]은 다수의 합성곱 계층을 사용하여 전달량을 추정하는 모델을 제안하였다. MSCNN[5]은 전역적인 전달량을 추정하는 모델을 사용하여 개략적인 전달량을 추정한 후,

표 1. 다양한 안개 제거 방법 비교  
Table 1. Comparison of various dehazing methods

Method	Class	Network	Multi-scale	Attention mechanism	PSNR (dB)	Feature
MSCNN[5]	ASM-based	CNN	✓		17.57	Multiple neural networks
AODNet[6]		CNN			18.80	Light-weight model
DehazeNet[7]		CNN			-	Nonlinear activation function
GFN[8]	End-to-End	Encoder-decoder	✓		24.11	Gated fusion
GCANet[9]		Encoder-decoder	✓		28.13	Smoothed dilated convolution
PFFNet[1]		Encoder-decoder	✓		29.22	Progressive feature fusion
GDN[10]		Encoder-decoder		✓	31.51	Multiple interconnected paths
Dehaze-GAN[3]		GAN			-	Cycle GAN
DuRN[12]		CNN		✓	31.92	Dual residual connection
MSBDN[2]		Encoder-decoder	✓		33.79	Dense feature fusion
FFANet[11]		CNN		✓	34.98	Channel attention and pixel attention

미세조정을 수행하는 모델을 사용하여 지역적으로 전달량을 최적화하였다. AODNet[6]은 식 1을 재구성하여 대기광과 전달량으로 구성된 특징맵을 정의한 후, 단순한 구조의 딥러닝 모델을 사용하여 이를 계산한 후 안개 제거 영상을 복원하였다. 하지만, 이와 같은 방법들은 각 단계에서의 에러가 다른 단계에 영향을 줌으로써 최종 에러가 증폭되는 문제점이 있다.

따라서 최근에는 전달량과 대기광을 각각 추정하지 않고 입력 영상으로부터 안개 제거 영상을 직접 생성할 수 있는 종단간 학습 모델들에 관한 연구들이 주를 이루고 있다. 종단간 학습을 통하여 안개를 제거하는 방법은 사용되는 네트워크 모델에 따라서 합성곱 기반 방법[11,12], 인코더-디코더 기반 방법[1,2,8-10], 적대적 생성 신경망 기반 방법[3,13] 등으로 구분된다.

FFANet[11]은 합성곱 기반 안개 제거 모델의 대표적인 방법이다. FFANet은 합성곱 계층과 잔차 연결(Residual connection) 및 어텐션(Attention) 계층으로 구성된 기본 블록을 적층하여 합성곱 그룹을 생성하고, 다수의 합성곱 그룹을 사용하여 입력 영상으로부터 안개 제거 영상을 생성하였다. FFANet은 픽셀 기반의 어텐션과 채널 기반의 어텐션을 모두 활용하고, 블록 내부의 지역적인 잔차 연결 및 그룹들을 연결하는 전역적인 잔차 연결을 사용함으로써 안개 제거 성능을 향상시켰다. DuRN[12]은 네트워크의 각 계층에서 처리되는 정보들이 유연하게 전달될 수 있도록 이중 잔차 연결(Dual residual connection) 방법을 제안하고, 이를 활용하여 다양한 영상 품질 저하(블러링, 안개, 비 등)를 완화할 수 있는 네트워크 모델을 제안하였다.

인코더-디코더 구조는 연산 시간을 줄일 수 있으므로 안개 제거를 위하여 이를 활용하는 방법들이 많이 제안되었다[1,2,8-10]. GFN[8]은 입력 안개 영상이 인코더-디코더를 통과하여 생성된 신뢰맵(Certainty map)과 영상 보정이 적용된 영상들로부터 생성된 특징 맵을 선택적으로 활용함으로써 안개 제거 성능을 향상시켰다. 또한 연산량을 줄이면서 합성곱의 수용 영역(Receptive field)을 확장시킬 수 있는 팽창 합성곱(Dilated convolution)을 사용하였다. GCANet[9]은 팽창 합성곱으로 인하여 발생하는 아티팩트(Artifact)를 제거하기 위하여 매끄러운 합성곱 연산(Smoothed dilated convolution)을 제안함으로써 기존 연구의 단점을 보완하였다. 단일 경로를 갖는 인코더-디코더 구조가 아닌 입력 영상에서 출력 영상까지 다양한 경로를 가질 수 있는 그리드 구조를 활용하여 안개 제거 성능을 향상하려는 방법이 제안되었다[10]. PFFNet[1]은 인코더-디코더 기반의 U-Net 구조에서 인코더의 각 계층과 대응되는 디코더 계층을 직접 연결하는 잔차 연결을 사용함으로써 각 계층의 정보가 안개 영상 복원 시 잘 활용될 수 있도록 하였다. MSBDN[2]은 영상 노이즈 제거에 효과적인 부스팅 알고리즘(Boosting algorithm)과 이전 계층의 정보를 효과적으로 전달할 수 있는 고밀도 특징 융합(Dense Feature Fusion)을 사용하여 안개 제거 성능을 향상시켰다.

적대적 생성 신경망을 활용한 안개 제거 방법[3,13]도 제안되었다. 제안 방법[13]은 입력 영상을 RGB의 각 채널로 분리한 후, 각 영상 채널 정보를 분석하여 안개가 제거된 영상을 생성하는 생성자와 실제 안개 영상과 합성된 안개 영상을 구분하

는 판별자 모델을 사용하였다. Cycle-GAN을 안개 제거에 적용한 방법[3]도 제안되지만, 안개 영상의 구조적인 정보는 잘 복원하는 데 반하여 색상 정보를 잘 복원하지 못하는 단점이 존재하였다.

현재 연구되는 딥러닝 기반 안개 제거 방법을 정리하면 대기 산란 모델(ASM : Atmospheric scattering model) 기반 방법과 종단간 학습 방법으로 구분되며, 사용하는 네트워크 구조 및 어텐션 메커니즘의 적용 여부에 따라 표 1과 같이 구분될 수 있다. 표 1에서 안개 제거 성능은 RESIDE 데이터셋[19]의 실내 및 실외 영상을 시험하여 계산된 PSNR(Peak Signal to Noise Ratio)을 사용하여 표시하였다[2]. 표 1을 살펴보면 종단간 학습 기반의 합성곱 네트워크를 사용하는 방법들이 좋은 성능을 보여주고 있으며, 어텐션 메커니즘을 적용하는 것이 안개 제거 성능 향상에 도움이 됨을 알 수 있다.

기존의 안개 제거 방법들은 주로 결과 영상의 품질 향상에 초점을 맞추고 있지만, 안개 제거 기술이 컴퓨터 비전 태스크의 전처리 과정으로 활용되기 위해서는 안개 제거 기술이 주 태스크의 성능에 미치는 영향에 대한 분석이 필요하다. 이와 관련하여 안개 제거 방법의 적용 여부에 따른 객체 검출 성능 변화를 제시하는 연구가 있었다[6]. 또한 기존 연구[19]에서는 안개 제거 방법에 따른 객체 검출 성능 변화를 분석하였다. 하지만 해당 연구는 사전 지식 기반의 안개 제거 방법과 초기 딥러닝 기반 안개 제거 방법을 사용하여 안개 제거 성능 자체가 낮고, 객체 검출 방법의 학습 데이터가 달라서 객관적인 성능 비교의 어려움이 있었다.

### 3. 객체 검출에서 안개 제거 영향 분석 방법

본 논문에서는 안개 제거 기술의 성능과 이를 활용하는 객체 검출 성능과의 관계를 분석하기 위한 연구를 수행하였다. 안개 제거 기술이 객체 검출 성능에 미치는 영향을 분석하기 위한 분석과정의 흐름도는 그림 1과 같다. 먼저 안개 제거 모델의 성능 평가를 위해 사용되는 합성 데이터셋을 사용하여 서로 다른 성능을 가지는 다수의 안개 제거 모델들을 생성하였다. 그리고, 안개가 포함된 객체 검출 데이터셋에서 안개 제거를 수행하지 않은 경우와 다양한 성능의 안개 제거 모델을 전처리 과정으로 적용한 경우의 객체 검출 알고리즘 성능을 측정하였다. 마지막으로 안개 제거 모델의 성능 변화에 따른 객체 검출 알고리즘의 성능 변화를 분석하여 안개 제거 기술의 성능과 이를 활용하는 객체 검출 성능과의 관계를 분석하였다.

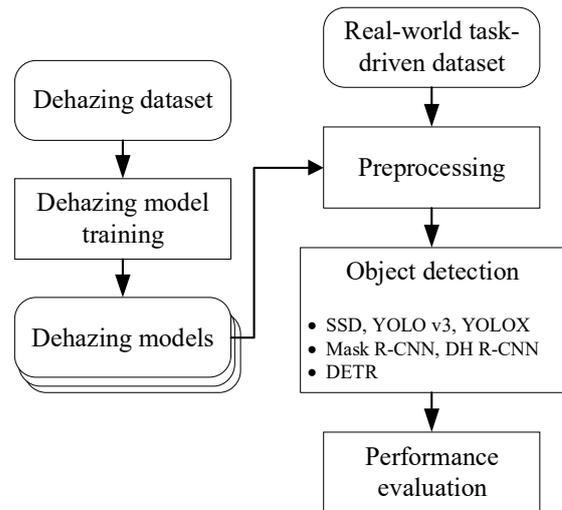


그림 1 안개 제거와 객체 검출 성능 관계 분석 흐름도  
Fig. 1. Flowchart for analyzing performance relation between image dehazing and object detection

안개 제거를 위한 네트워크 모델로는 공개 데이터셋에서 가장 좋은 성능을 보여주고 있는 FFANet[11]을 사용하였다. FFANet은 기본 합성곱 블록을 적층하여 그룹을 생성하고, 다수의 합성곱 그룹을 사용하여 네트워크를 구성하기 때문에 합성곱 블록의 수와 합성곱 그룹 수를 조절하여 다양한 성능을 갖는 안개 제거 모델을 생성할 수 있는 장점이 있다. 본 논문에서 사용한 FFANet의 기본 구조는 그림 2 (a)와 같이 전처리 과정 및 후처리 과정과 3개의 합성곱 그룹으로 구성된다. FFANet에서는 안개 영상이 입력으로 주어지면 전처리 과정을 수행한 후 3개의 합성곱 그룹을 순차적으로 통과하게 된다. 각 합성곱 그룹을 통과한 특징맵들은 하나로 합쳐지게 되며, 이후 특징맵의 채널별 가중치를 부여하는 채널 어텐션 계층과 공간정보에 가중치를 부여하는 픽셀 어텐션 계층을 통과하게 된다. 마지막으로 후처리 과정을 수행하여 안개가 제거된 영상을 생성하게 된다.

FFANet의 합성곱 그룹은 그림 2(b)와 같이 N개 합성곱 블록과 한 개의 합성곱 계층 및 잔차 연결로 구성된다. 각 합성곱 블록은 다수의 합성곱 계층과 채널 어텐션 및 픽셀 어텐션 계층으로 구성된다. 합성곱 그룹에서는 특징들이 다수의 네트워크 계층을 통과하는 동안 잘 전달될 수 있도록 합성곱 블록 내부 및 합성곱 블록 간의 다양한 잔차 연결을 사용하였다.

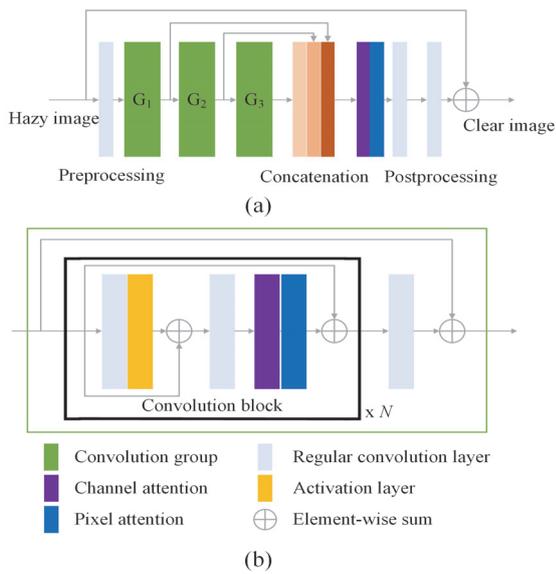


그림 2 (a) FFANet 구조도 (b) 합성곱 그룹 구조  
 Fig. 2. (a) FFANet architecture (b) Convolution group structure

객체 검출 알고리즘은 객체가 존재할 가능성이 있는 영역을 탐색한 후, 이를 기반으로 객체의 영역과 범주를 분류하는 이 단계 검출 방법과 사전 정의된 영역에서 물체의 크기와 위치 정보를 사용하여 해당 영역에 객체의 존재 여부와 영역 정보를 판단하는 단일 단계 검출 방법으로 구분된다. 또한 최근에는 트랜스포머(Transformer) 기반의 객체 검출 방법이 활발히 연구되고 있다. 따라서, 안개 제거 모델의 성능 변화가 객체 검출에 미치는 영향을 분석하기 위해서는 다양한 객체 검출 알고리즘의 적용이 필요하다. 본 논문에서는 전체 6가지의 객체 검출 알고리즘을 사용하였다. 먼저 단일 단계 검출 방법으로 SSD(Single Shot Detector)[20], YOLO(You Only Look Once)v3[21], YOLOX[22]를, 이 단계 검출 방법으로는 Mask R-CNN [23], DH(Double-Head) R-CNN[24]을 선정하였으며, 마지막으로 트랜스포머 기반 방법인 DETR(DEtection TRansformer)[25]을 사용하였다.

SSD[20]는 이미지를 분류하는 백본 네트워크를 통하여 특징맵을 생성한 후, 추가 네트워크를 사용하여 다양한 크기의 객체를 검출하는 방법이다. 추가 네트워크에서는 다양한 크기의 합성곱 연산을 수행하여 특징 맵의 채널별로 다양한 중첩비를 가지는 바운딩 박스의 객체별 확률을 계산하고, 이를 통합하여 최댓값을 가지는 바운딩 박스만을 추출함으로써 객체를 검출하게 된다. SSD는 멀티스케일 특징 맵을 통하여 작은 객체를 검출할 수 있고, 객체 검출을 위하여 특징값을 리샘플링하지 않으며 후보 영역 추출과정을 제거함으로써 객체 검출 속도를 향상하였다.

YOLOv3[21]는 기존 YOLOv2 모델을 변형하여 객체 탐지 성능을 향상한 방법이다. YOLOv3는 객체 검출에 사용되는 특징을 추출하기 위하여 Darknet-53으로 알려진 새로운 백본 네트워크를 사용하였다. Darknet-53은 53개의 합성곱 계층으로 구성되어 있으며, 각각의 합성곱 계층은 3x3 커널 또는 1x1 커널만으로 구성된다. 또한, Darknet-53은 잔차 연결을 사용하여 네트워크의 깊이가 깊어 지더라도 특징값들이 잘 전달될 수 있도록 하였다. 기존 모델은 객체 검출을 위한 바운딩 박스와 참값(Ground truth)의 차이를 제곱 오차(Squared error)로 모델링하여 학습하였지만, 이 모델에서는 바운딩 박스의 좌표와 크기를 직접 예측 하여 참값과의 차이를 L1 손실을 사용하여 학습하였다. 또한 세 개의 특징맵 스케일에서 바운딩 박스를 예측하여 통합 분석하고, 바운딩 박스로부터 객체 범주를 추정할 때 소프트맥스를 사용하지 않고 로지스틱 분류기(Logistic classifiers)를 사용하여 멀티 레이블 예측이 가능하게 함으로써 객체 검출 성능을 향상시켰다.

YOLOX[22]는 YOLOv3 모델을 기반으로 다른 검출 알고리즘에서 많이 활용되는 기법들을 적용하여 성능을 개선한 방법이다. YOLOv3는 객체 검출과 검출 영역 정합을 위하여 단일 헤드를 사용하였지만, YOLOX는 객체 검출과 검출 영역 정합을 분리하여 각기 다른 헤드를 사용하였다. YOLOX는 YOLOv3와 같은 Darknet-53을 백본 네트워크로 사용하였으며, 검출 성능을 향상하기 위해서 모자이크(Mosaic)과 믹스업(Mixup) 등의 데이터 증강 방법을 적용하였다. 기존 YOLO 기반의 객체 검출 알고리즘들은 학습 전 클러스터링을 통해서 생성된 앵커 박스를 사용하여 바운딩 박스를 예측하였지만, 이는 도메인에 종속적이며 일반화에 어려움이 있다. 따라서 YOLOX에서는 앵커 프리 매커니즘을 사용하였으며, 좀 더 진보된 레이블 할당 방법을 사용하여 성능을 개선하였다.

Mask R-CNN[23]은 이 단계 객체 검출 방법인 Faster R-CNN 방법을 확장한 방법이다. Faster R-CNN 방법은 백본 네트워크를 통하여 추출된 특징 맵에서 객체 후보 영역을 찾는 단계와 특징 맵에서 RoIPool 방법[26]을 사용하여 객체 후보 영역의 특징을 추출한 후 객체 분류와 검출 영역 정합을 수행하는 단계로 구성된다. 이때, 각 단계에서 사용되는 특징맵이 같다는 특징이 있다. Mask R-CNN은 Faster R-CNN과 객체 후보 영역을 찾는 단계는 동일하며, 객체 분류와 검출 영역 정합을 수행 단계에서 객체의 마스크를 예측하는 브랜치를 추가하여 객체 검출 성능을 향상했다. 객체 마스크를 예측하는 브랜치는 모든 객체에 대한 마스크를 생성하며, 객체 분류 결과에 종속되지 않고 독립적으로 동작한다. 기존 RoIPool 방법은 특징맵

에서 특징을 가져올 때, 객체 후보 영역과 특징맵의 크기가 일치하지 않아서 생기는 양자화 에러(Quantization error)가 발생하였다. 이를 해결하기 위해 Mask R-CNN 방법에서는 이중 선형 보간(Bilinear interpolation)을 통하여 정확한 특징값이 계산될 수 있도록 하는 ROIAlign 방법을 사용하였다.

기존 이 단계 객체 검출 방법들은 객체 검출 네트워크에서 한 개의 전 연결 헤드(Fully connected head) 또는 합성곱 헤드(Convolution head)를 통하여 객체 분류와 객체 영역 정합을 모두 수행하였다. 하지만, 헤드의 특성에 따라서 더 적합한 태스크가 달라질 수 있다. DH R-CNN 방법[24]은 전 연결 헤드와 합성곱 헤드가 어떤 태스크에 더 적합한지 분석하였으며, 실험 결과 전 연결 헤드와 합성곱 헤드가 각각 객체 분류 및 객체 영역 집합에 더 적합함을 확인하였다. 이런 특성을 활용하여 DH R-CNN 방법은 Mask R-CNN 방법을 기반으로 두 개의 헤드를 각기 다른 태스크에 적용하여 객체 검출 성능을 개선하였다.

최근 기계번역, 음성인식 등에서 셀프 어텐션 기반의 트랜스포머가 높은 성능을 보여주면서, 객체 검출 분야에서도 트랜스포머를 기반으로 하는 객체 검출 방법들이 제안되고 있다. DETR 방법[25]은 인코더와 디코더 구조를 갖는 트랜스포머 네트워크를 사용하여 객체 영역 집합을 생성한 후, 이를 참값과 비교하여 이분 매칭(Bipartite matching) 손실을 계산함으로써 객체 검출 네트워크를 학습하였다. DETR 방법은 기존 객체 검출 방법이 전처리 또는 후처리 과정에서 사용하는 도메인 지식에 의존하지 않기 때문에 단순한 구조를 가지며, 기존의 최적화가 잘 수행된 Faster R-CNN과 유사한 성능을 보여주었다.

#### 4. 실험 결과

안개 제거 모델의 학습을 위해서는 영상 합성 기반의 공개 데이터셋인 RESIDE-β 데이터셋[19]을 사용하였다. RESIDE-β 데이터셋은 실외용 학습 데이터(Outdoor Training Set : OTS) 72,135장으로 구성되어 있다. 학습된 모델의 성능 평가를 위해서는 RESIDE:V0 데이터셋 중 SOTS(Synthetic Objective Testing Set)의 실외 영상 500장을 사용하였다. 안개 제거 모델의 네트워크 학습 및 성능 평가를 위해서 사용된 데이터셋의 샘플 영상들은 그림 3과 같다.

본 논문의 실험에 사용한 FFANet 네트워크는 3개의 합성곱 그룹으로 구성되며, 각 그룹은 19개의 합성곱 블록으로 구성된다. 네트워크 학습을 위한 배치 크기는 4로 설정하였으며, 학습 비율은 0.0001로 설정한 후 코사인 어닐링(Cosine Annealing) 방법

을 적용하고 ADAM 최적화 방법을 사용하였다. 입력 영상을 240 x 240 크기로 변환하여 모델 학습에 사용하였으며, 안개 제거 모델의 성능 평가를 위한 평가 척도로는 아래 식 2와 같이 정의되는 PSNR을 사용하였다.

$$PSNR = 10 \log_{10} \frac{MAX^2}{MSE} \quad (2)$$

위의 식에서 MAX는 입력 영상의 픽셀이 가질 수 있는 최대값을 의미하며, MSE(Mean squared error)는 안개가 없는 영상과 안개 제거 모델을 통하여 생성된 영상의 평균 제곱 에러를 의미한다. 본 논문에서는 FFANet을 학습하는 과정에서 다양한 안개 제거 성능을 가지도록 24.7에서 32.0까지 9개의 PSNR 값을 갖는 안개 제거 모델을 생성하여 객체 검출 알고리즘 성능과의 관계를 분석하기 위해 사용하였다.

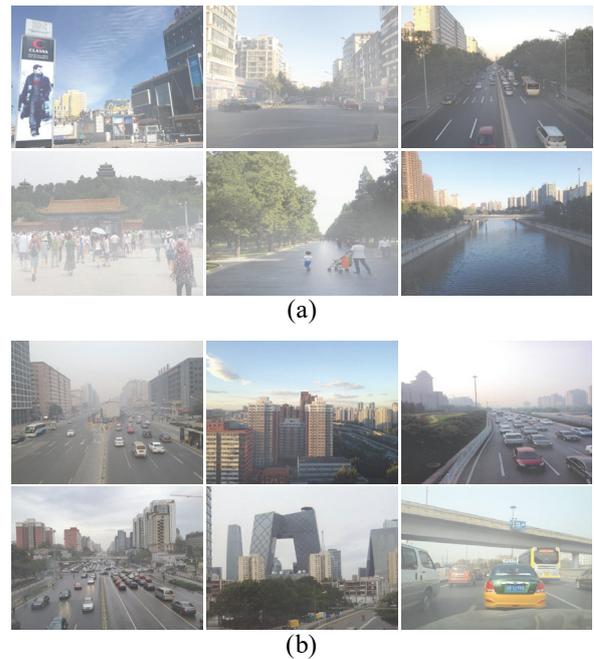


그림 3. (a) RESIDE-β 실외 학습 데이터 샘플 영상  
(b) SOTS 실외 평가 데이터 샘플 영상  
Fig. 3. (a) Sample images of OTS on RESIDE-β dataset (b) Sample images of OTS on SOTS

FFANet은 PyTorch 라이브러리를 사용하여 구현되었으며, 인텔 i9-9980XE 프로세서와 RTX A6000 그래픽카드가 탑재된 윈도우 10 64bit 환경에서 연산 시간을 측정하였다. 총 4,322장의 영상(평균 해상도 682 x 481)을 사용하여 처리 시간을 측정하였을 때, FFANet은 영상 한 장을 처리하는

데 평균 0.371초가 소요되었으며, 텐서의 입출력을 모두 고려하는 경우 영상 한 장을 처리하는 데 평균 0.387초가 소요되었다. 본 논문의 실험에서 사용한 9개 모델의 경우 모두 동일한 네트워크 구조를 가지기 때문에 처리 시간을 측정하였을 때 동일한 처리 시간을 보여주었다.

안개가 포함된 영상에서의 객체 검출 성능을 평가하기 위해 RESIDE- $\beta$  데이터셋의 객체 검출 성능 평가 데이터셋(Real-world Task-driven Testing Set : RTTS)을 사용하였으며, 해당 데이터셋은 총 4,322장 영상으로 구성되어 있다. RTTS 영상은 자동차, 자전거, 모터바이크, 사람, 버스 등의 5개 범주의 객체를 포함하고 있으며, 평균 난이도를 가지는 29,599개의 객체 레이블과 난이도가 높은 11,064개의 객체 레이블로 구성된다. 객체 레이블의 난이도 및 범주에 따른 분포는 표 2와 같다. 그림 4는 RTTS의 샘플 영상을 나타내며, 각 영상에서 바운딩 박스의 색상은 객체 범주를 의미한다.

표 2. RTTS 객체 범주별 구성 정보  
Table 2. Detailed classes information of RTTS

Category	Normal	Difficult	Total
Person	7,950	3,416	11,366
Bicycle	534	164	698
Car	18,415	6,902	25,317
Bus	1,838	752	2,590
Motorbike	862	370	1,232
Total	29,599	11,604	41,203

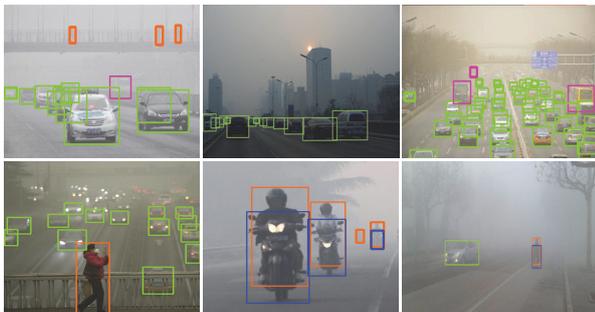


그림 4. RTTS 샘플 영상  
Fig. 4. Sample images of RTTS

본 논문에서 사용한 6가지 객체 검출 방법은 객체 검출 알고리즘의 성능 비교를 위해 사용되는 MMDetection[27] 라이브러리를 사용하여 구현하였다. MMDetection은 다양한 객체 검출 알고리즘을 PyTorch로 구현하여 제공하고 있으며, 알고리즘 간 객관적인 성능 비교를 위해 다양한 데이터셋에 대한 사전 학습된 모델도 제공하고 있다. 본 논문에서 사용한 객체 검출 방법들은 COCO 데이터셋[28]을 사용하여 사전 학습된 모델을 사용하였으며, 객체 검출모델들의 상세한 정보는 표 3과 같다. 표 3는 객체 검출모델들이 사용한 백본 네트워크와 COCO 검증 데이터셋에 대한 객체 검출 성능을 나타낸다. COCO 데이터셋은 80개의 객체 범주로 구성되어 있어, 제안 방법에서는 RTTS 영상이 포함하는 5개 객체 범주를 검출하도록 하였다. 객체 검출 방법의 성능을 평가하는 기준은 IoU(Intersection over Union)의 기준 값을 0.5로 설정했을 때 mAP(mean Average Precision)를 측정하였다.

표 3. 객체 검출 방법 상세 정보  
Table 3. Detailed information of object detection methods

Method	Backbone	mAP
SSD	VGG16	25.5
YOLOv3	Darknet-53	30.9
YOLOX	YOLOX-S	40.5
Mask R-CNN	R-50-FPN	38.2
DH R-CNN	R-50-FPN	40.0
DETR	R-50	40.1

안개 제거 기술의 성능과 객체 검출 성능의 관계 분석을 위해 FFANet을 학습하여 생성한 각기 다른 성능을 가지는 9개 안개 제거 모델을 RTTS 영상에 적용하여 안개 제거 영상을 생성하였다. 그리고 안개 제거 영상들에 대해 6개 객체 검출 모델을 적용하여 객체 검출 성능을 측정하였다. 마지막으로, 안개 제거 기술을 적용하지 않은 원본 영상에 대한 객체 검출 성능을 기준 성능으로 안개 제거 성능이 객체 검출 성능에 미치는 영향을 분석하였다.

표 4. 안개 제거성능에 따른 객체 검출 결과  
Table 4. Results of the object detection according to the performance of the dehazing method

Method	Baseline	PSNR 24.68 dB	PSNR 25.74 dB	PSNR 26.76 dB	PSNR 27.64 dB	PSNR 28.18 dB	PSNR 29.64 dB	PSNR 30.93 dB	PSNR 31.57 dB	PSNR 32.01 dB
SSD	32.87	34.64	34.64	35.39	34.95	<b>35.44</b>	35.26	35.16	34.57	34.95
YOLOv3	41.77	41.46	41.46	42.02	41.93	42.22	42.13	<b>42.54</b>	42.02	42.45
YOLOX	53.27	53.84	53.84	54.68	54.58	54.67	<b>54.88</b>	54.62	54.78	54.54
Mask R-CNN	56.07	55.14	55.14	56.49	56.46	56.79	56.90	<b>57.63</b>	57.22	56.89
DH R-CNN	55.12	54.37	54.37	55.97	55.96	55.73	55.87	<b>56.00</b>	55.89	55.91
DETR	56.53	56.18	56.18	56.37	56.38	56.95	57.23	<b>57.80</b>	57.26	57.29

표 5. 안개 제거성능에 따른 객체 검출 결과 (높은 난이도 포함)

Table 5. Results of the object detection according to the performance of the dehazing method (including difficult annotations)

Method	Baseline	PSNR 24.68 dB	PSNR 25.74 dB	PSNR 26.76 dB	PSNR 27.64 dB	PSNR 28.18 dB	PSNR 29.64 dB	PSNR 30.93 dB	PSNR 31.57 dB	PSNR 32.01 dB
SSD	24.51	26.25	25.76	26.38	26.03	<b>26.41</b>	26.23	26.19	25.74	26.06
YOLOv3	32.60	31.80	32.50	32.90	32.76	33.09	33.04	33.27	32.92	<b>33.28</b>
YOLOX	41.54	42.50	42.11	43.00	42.78	42.95	<b>43.07</b>	42.74	42.86	42.73
Mask R-CNN	45.73	45.81	45.11	46.58	46.49	46.64	46.66	<b>47.27</b>	46.83	46.70
DH R-CNN	44.99	45.48	44.61	<b>46.25</b>	46.13	45.98	45.97	45.96	45.80	45.96
DETR	46.8	47.10	46.52	46.99	47.01	47.44	47.59	<b>48.00</b>	47.60	47.75

안개 제거성능에 따른 객체 검출 방법의 성능 실험 결과는 표 4, 표 5와 같다. 표 4는 RTTS 영상 중 평균 난이도를 갖는 객체 레이블만을 검출한 경우를 나타내며, 표 5는 높은 난이도를 갖는 객체 레이블을 포함하여 검출한 객체 검출 성능(mAP)을 나타낸다.

표 4를 살펴보면 안개 제거 모델의 성능이 낮은 경우 안개 제거를 수행하지 않는 경우보다 객체 검출 성능이 저하되는 경우가 발생하는 것을 알 수 있다. 특히 이러한 현상들은 이 단계 검출 방법들(Mask R-CNN, DH R-CNN)에서 더 현저하게 나타남을 알 수 있다. 하지만 안개 제거 모델의 성능(PSNR)이 28.18 이상일 때 모든 객체 검출 알고리즘에서 객체 검출 성능이 향상되는 것을 확인할 수 있다. 또한, 안개 제거 모델의 PSNR이 30.93인 경우 대부분의 객체 검출 방법에서 가장 높은 성능을 보여주었다. 반면에 안개 제거 모델의 PSNR이 32.01로 증가하더라도 객체 검출 성능이 지속해서 증가하지 않음을 확인할 수 있다.

표 5의 실험 결과는 검출이 어려운 객체를 포함하고 있어서 안개 제거를 수행하지 않은 기준 성능이 평균 난이도의 객체만을 검출한 표 4의 결과에 비해서 낮음을 알 수 있다. 하지만 안개 제거 모델의 성능과 객체 검출 성능의 변화는 표 4과 비슷한 결과를 볼 수 있다. 표 5의 실험 결과를 살펴보면 안개 제거 모델의 PSNR이 26.76 이상이면 모든 객체 검출 알고리즘의 성능이 향상되고, 안개 제거 모델의 성능이 계속 증가하더라도 객체 검출 성능이 지속해서 증가하지 않는다는 것을 확인할 수 있다.

기존 연구 결과[19]에서는 안개 제거성능과 객체 검출 성능과의 연관성을 찾기 어려웠다. 이는 기존 연구에서 사용한 안개 제거 모델의 PSNR이 최소 15.72, 최대 21.14인 방법들을 사용하여 객체 검출 성능 향상의 일관성이 나타나지 않았기 때문이다. 하지만 본 논문에서는 현재 가장 좋은 성능을 보여주는 안개 제거 모델을 사용하여 실험함으로써, 객체 검출의 전처리 과정으로 적용하는 안개 제거 모델의 성능이 최소 28.18 정도(Lower Bound)를

만족하면 객체 검출 알고리즘에서의 성능 향상이 가능함을 확인하였다. 또한 안개 제거 모델의 성능이 32 이상일 때 객체 검출 알고리즘의 성능이 더 증가하지 않는다는 점에서 최고 임계치(Upper Bound)를 확인할 수 있었다.

### 5. 결론 및 향후 연구

본 논문에서는 안개 제거 기술의 성능이 안개 제거 기술을 전처리 과정으로 사용하는 객체 검출 성능에 미치는 영향을 분석하는 연구를 수행하였다. 이를 위하여 최신 안개 제거 모델인 FFANet을 학습하여 각기 다른 9개의 안개 제거성능을 갖는 안개 제거 모델을 생성하였다. 그리고 RTTS 영상에 적용하여 안개 제거 영상을 생성하고 6개 객체 검출모델을 적용함으로써, 각기 다른 안개 제거 성능이 다양한 객체 검출모델의 성능에 미치는 영향을 분석하였다. 실험 결과 안개 제거 모델의 성능이 낮은 경우 객체 검출 성능이 저하되는 경우가 발생하는 것을 확인하였다. 하지만 안개 제거 성능(PSNR)이 28 이상이면 모든 객체 검출 알고리즘에서 성능 향상이 가능함을 확인하였다. 또한, 객체 검출 성능은 안개 제거성능에 비례하여 증가하지 않고 안개 제거 성능(PSNR)이 32 이상이면 객체 검출 성능이 수렴하는 것을 확인하였다.

현재 안개 제거 기술에 관한 연구는 안개 제거 성능 향상에 초점을 맞추고 있어 모델의 복잡도가 비약적으로 증가하고 있다. 하지만 실험 결과와 같이 안개 제거 모델의 성능이 일정 수준 이상일 경우 객체 검출 성능 향상이 수렴되어, 안개 제거 모델의 성능과 복잡도의 균형을 맞추는 것이 필요하다. 따라서, 본 논문의 실험으로 도출한 안개 제거 성능의 최저, 최고 임계치를 기반으로 안개 제거 모델의 복잡도(계층 수, 채널 수 등)를 조절함으로써 처리 속도와 모델 복잡도를 고려하여 효율적으로 안개를 제거할 수 있는 모델에 관한 연구를 수행할 예정이다.

본 논문에서는 안개 제거 모델의 성능 지표로 레퍼런스 기반의 대표적인 성능 지표인 PSNR을

사용하였지만, 향후 연구에서는 비 레퍼런스 기반 성능 지표를 포함하여 다양한 성능 지표와의 관계를 분석할 예정이다. 또한 본 논문에서는 동일한 네트워크 구조를 사용하여 성능이 다른 모델을 생성하였기에 연산 시간을 고려하지 않았지만, 향후 연구에서는 다양한 네트워크 구조를 가지는 모델들을 활용하여 객체 검출을 비롯 객체 인식, 장면 분할 등 다른 태스크에서도 안개 제거 모델의 연산 시간, 모델 복잡도 등이 주 태스크의 성능에 미치는 영향을 분석하는 연구를 수행할 예정이다.

## Conflict of Interest

저자는 본 논문에 관련된 어떠한 잠재적인 이해 상충도 없음을 선언한다.

## References

- [1] Kangfu Mei, Aiwen Jiang, Juncheng Li, and Mingwen Wang, "Progressive Feature Fusion Network for Realistic Image Dehazing," *Proceedings of the Asian Conference on Computer Vision*, Perth, Australia, pp. 203-215, Dec. 2018.
- [2] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, and Xinyi Zhang, "Multi-Scale Boosted Dehazing Network with Dense Feature Fusion," *Proceeding of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, pp. 2157-2167, June 2020.
- [3] Chenghuan Wang, Zhijun Meng, Ronglei Xie, and Xiaoi Jiang, "A Single Image Dehazing Algorithm Based on Cycle-GAN," *Proceedings of the 2019 International Conference on Robotics, Intelligent Control and Artificial Intelligence*, Shanghai, China, pp. 247-251, Sep. 2019.
- [4] Kaiming He, Jian Sun, and Xiaoou Tang, "Single Image Haze Removal Using Dark Channel Prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341-2353, Dec. 2011.
- [5] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang, "Single Image Dehazing via Multi-scale Convolutional Neural Networks," *Proceedings of the European Conference on Computer Vision*, Amsterdam, Netherlands, pp. 154-169, Sep. 2016.
- [6] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng, "AOD-Net: All-in-One Dehazing Network," *Proceedings of IEEE International Conference on Computer Vision*, Venice, Italy, pp. 4780-4788, Oct. 2017.
- [7] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao, "DehazeNet: An End-to-End System for Single Image Haze Removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187-5198, 2016.
- [8] Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, and Ming-Hsuan Yang, "Gated fusion network for single image dehazing," *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 3253-3261, June 2018.
- [9] Dongdong Chen, Mingming He, Qingnan Fan, Jing Liao, Liheng Zhang, Dongdong Hou, Lu Yuan, and Gang Hua, "Gated context aggregation network for image dehazing and deraining," *Proceedings of IEEE Winter Conference on Applications of Computer Vision*, Waikoloa Village, HI, USA, pp. 1375-1383, Jan. 2019.
- [10] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen, "GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing," *Proceedings of IEEE International Conference on Computer Vision*, Seoul, Korea (South), pp. 7314-7323, Oct. 2019.
- [11] Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia, "FFA-Net: Feature Fusion Attention Network for Single Image Dehazing," *Proceedings of the AAAI Conference on Artificial Intelligence*, New York, USA, vol. 34, no. 7, pp. 11908-11915, Feb. 2020.
- [12] Xing Liu, Masanori Suganuma, Zhun Sun, and Takayuki Okatani, "Dual Residual Networks Leveraging the Potential of Paired Operations for Image Restoration," *Proceeding of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 7000-7009, June 2019.
- [13] Patricia L. Suarez, Angel D. Sappa, Boris X. Vintimilla, and Riad I. Hammoud, "Deep Learning Based Single Image Dehazing," *Proceeding of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Salt Lake City, UT, USA, pp. 1250-1257, June 2018.
- [14] Yufeng Fan, Yongfeng Cao, and Xiuzhang Yang, "Unmanned Aerial Vehicle Remote Sensing Image Dehazing via Global Parameters," *Proceedings of IEEE International Conference on Artificial Intelligence and Information Systems*, Dalian, China, pp. 349-354, Mar. 2020.
- [15] Sidharth Gautam, Tapan Kumar Gandhi, and B.K. Panigrahi, "A Model-based Dehazing Scheme for Unmanned Aerial Vehicle System using Radiance Boundary Constraint and Graph Model," *Journal of Visual Communication and Image Representation*, vol. 74, pp. 1-11, Jan. 2021.
- [16] Seok Bong Yoo and Mikyong Han, "Temporal Matching Prior Network for Vehicle License Plate Detection and Recognition in Videos," *ETRI Journal*, vol. 42, no. 3, pp. 411-419, June 2020.
- [17] Mingxing Tan, Ruoming Pang, and Quoc V. Le, "EfficientDet: Scalable and Efficient Object Detection," *Proceeding of the IEEE/CVF Conference*

on *Computer Vision and Pattern Recognition*, Seattle, WA, USA, pp. 10781-10790, June 2020.

[18] A. Cantor, "Optics of the atmosphere-Scattering by molecules and particles," *IEEE Journal of Quantum Electronics*, vol. 14, no. 9, pp. 698-699, Sep. 1978.

[19] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang, "Benchmarking Single-Image Dehazing and Beyond," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 492-505, 2019.

[20] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg, "SSD: Single Shot MultiBox Detector," *Proceedings of European Conference on Computer Vision*, Amsterdam, Netherlands, pp. 21-37, Sep. 2016.

[21] Joseph Redmon and Ali Farhadi, "YOLOv3: An Incremental Improvement," *arXiv:1804.02767 [cs.CV]*, Apr. 2018.

[22] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun, "YOLOX: Exceeding YOLO Series in 2021," *arXiv:2107.08430[cs.CV]*, Aug. 2021.

[23] Kaiming He Georgia Gkioxari Piotr Dollar Ross Girshick, "Mask R-CNN," *Proceedings of IEEE International Conference on Computer Vision*, Venice, Italy, pp. 2961-2969, Oct. 2017.

[24] Yue We, Yinpeng Chen, Lu Yuan, Zicheng Liu, Lijuan Wang, Hongzhi Li, and Yun Fu, "Rethinking Classification and Localization for Object Detection," *Proceeding of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, pp. 10186-10195, June 2020.

[25] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko, "End-to-End Object Detection with Transformers," *Proceedings of European Conference on Computer Vision*, Glasgow, UK, pp. 213-229, Aug. 2020.

[26] Ross Girshick, "Fast R-CNN," *Proceedings of IEEE International Conference on Computer Vision*, Washington DC, USA, pp. 1440-1448, Dec. 2015.

[27] Kai Chen et al., "MMDetection: Open MMLab Detection Toolbox and Benchmark," *arXiv:1906.07155 [cs.CV]*, June 2019.

[28] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollar, and C Lawrence Zitnick, "Microsoft coco: Common objects in context," *Proceedings of European Conference on Computer Vision*, Zurich, Switzerland, pp. 740-755, Sep. 2014.

**저 자 소 개**



**정치윤(Chi Yoon Jeong)**

2002년: POSTECH 전자전기공학부 학사  
 2004년: POSTECH 전자전기공학부 석사  
 2018년: KAIST 전산학부 박사  
 2004년~현재: 한국전자통신연구원  
 책임연구원  
 2021년~현재: UST 인공지능 전공 부교수

관심분야 : Computer vision, Pattern recognition, Machine learning, Data analytics

ORCID Number : 0000-0001-7089-2516

E-mail : iamready@etri.re.kr



**하동휘(Donghee Ha)**

2018년: 충남대학교 컴퓨터공학 학사  
 2020년: 충남대학교 컴퓨터공학 석사  
 2021년~현재: UST 인공지능 전공 박사  
 과정

관심분야 : Continual learning, On-device deep learning, Embedded system

ORCID Number : 0000-0002-8806-7108

E-mail : dhha@etri.re.kr



**문경덕(KyeongDeok Moon)**

1990년: 한양대학교 전산학 학사  
 1992년: 한양대학교 전산학 석사  
 2005년: 한국과학기술원 정보공학부  
 박사  
 1992년~현재: 한국전자통신연구원  
 책임연구원

관심분야 : Artificial intelligence, Human augmentation technology, Autonomous ship

ORCID Number : 0000-0002-7221-8303

E-mail : kdmoon@etri.re.kr



**김무섭(Mooseop Kim)**

1998년: 경북대학교 전자공학과 석사  
 2008년: 충남대학교 컴퓨터공학과 박사  
 1999년~현재: 한국전자통신연구원  
 책임연구원  
 1998년~1999년: LG종합기술원 연구원

관심분야 : Wearable computing, Activity recognition, Sensory substitution

ORCID Number : 0000-0003-4914-0584

E-mail : gomskim@etri.re.kr