

## 딥러닝에 의한 항공사진 구름 분류 및 탐지 비교 실험 Comparative Experiment of Cloud Classification and Detection of Aerial Image by Deep Learning

송준영<sup>1)</sup> · 원태연<sup>2)</sup> · 조수민<sup>3)</sup> · 어양담<sup>4)</sup> · 박소영<sup>5)</sup> · 신상호<sup>6)</sup> · 박진수<sup>7)</sup> · 김창재<sup>8)</sup>

Song, Junyoung · Won, Taeyeon · Jo, Su Min · Eo, Yang Dam

Park, So young · Shin, Sang ho · Park, Jin Sue · Kim, Changjae

### Abstract

As the amount of construction for aerial photography increases, the need for automation of quality inspection is emerging. In this study, an experiment was performed to classify or detect clouds in aerial photos using deep learning techniques. Also, classification and detection were performed by including satellite images in the learning data. As algorithms used in the experiment, GoogLeNet, VGG16, Faster R-CNN and YOLOv3 were applied and the results were compared. In addition, considering the practical limitations of securing erroneous images including clouds in aerial images, we also analyzed whether additional learning of satellite images affects classification and detection accuracy in comparison a training dataset that only contains aerial images. As results, the GoogLeNet and YOLOv3 algorithms showed relatively superior accuracy in cloud classification and detection of aerial images, respectively. GoogLeNet showed producer's accuracy of 83.8% for cloud and YOLOv3 showed producer's accuracy of 84.0% for cloud. And, the addition of satellite image learning data showed that it can be applied as an alternative when there is a lack of aerial image data.

Keywords : Deep Learning, Classification, Detection, GoogLeNet, VGG16, Faster R-CNN, YOLOv3

### 초 록

항공사진 촬영량이 증가함에 따라 품질검사 자동화의 필요성이 대두되고 있다. 본 연구에서는 딥러닝 기법으로 항공사진 내 구름을 분류 또는 탐지하는 실험을 수행하였고, 또한 위성영상을 학습자료에 포함시켜 분류 및 탐지를 수행하였다. 실험에 사용한 알고리즘으로는 GoogLeNet, VGG16, Faster R-CNN과 YOLOv3를 적용하여 결과를 비교하였다. 또한 구름이 포함된 오류영상 확보의 현실적 제한을 고려하여 항공영상만 존재하는 학습 데이터셋에서 위성영상을 활용한 추가학습이 분류 및 탐지정확도에 영향을 미치는지도 분석하였다. 실험결과, 항공사진의 구름 분류와 탐지에서 각각 GoogLeNet과 YOLOv3 알고리즘이 상대적으로 우월한 정확도를 나타냈고, GoogLeNet은 구름에 대한 생산자정확도 83.8% 그리고 YOLOv3는 구름에 대한 생산자정확도 84.0%를 보여주었다. 또한, 위성영상 학습자료 추가가 항공사진 자료의 부족 시 대안으로 적용가능 함을 보여주었다.

핵심어 : 딥러닝, 분류, 탐지, GoogLeNet, VGG16, Faster R-CNN, YOLOv3

Received 2021. 11. 14, Revised 2021. 12. 02, Accepted 2021. 12. 18

1) Member, Dept. of Advanced Technology Fusion, M.S. student, Konkuk University, Seoul, Korea (E-mail: songjy95@konkuk.ac.kr)

2) Dept. of Advanced Technology Fusion, M.S. student, Konkuk University, Seoul, Korea (E-mail: teadone@konkuk.ac.kr)

3) Dept. of Advanced Technology Fusion, M.S. student, Konkuk University, Seoul, Korea (E-mail: jsm302@konkuk.ac.kr)

4) Corresponding Author, Member, Dept. of Civil and Environmental Engineering, Professor, Konkuk University, Seoul, Korea (E-mail: eoandrew@konkuk.ac.kr)

5) Geographic Information Division, National Geographic Information Institute, Senior Officer, Ministry of Land, Infrastructure and Transport, Suwon, Korea (E-mail: soyoung331@korea.kr)

6) Geographic Information Division, National Geographic Information Institute, Deputy Director, Ministry of Land, Infrastructure and Transport, Suwon, Korea (E-mail: ssh1263@korea.kr)

7) Member, Project Development Division, Assistant Manager, ALLforLAND.Co.Ltd, Seoul, Korea (E-mail: mofeste@all4land.com)

8) Member, Dept. of Civil and Environmental Engineering, Professor, Myongji University, Yongin, Korea (E-mail: cjkim@mju.ac.kr)

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.

## 1. 서론

우리나라는 디지털트윈국토의 기반데이터 확보를 위해 2021년부터 항공영상 해상도를 다양화하고 전국에 대한 촬영주기 단축을 추진하고 있다. 이에 따라 항공사진과 정사영상 구축물량이 기존보다 대폭 증가하고 향후에도 지속적으로 그 양이 증가할 전망이다. 국토지리정보원에서는 항공사진 데이터 품질에 관한 표준을 제정하여 항공사진 데이터 품질 측정방법을 규정하고, 생산자와 사용자가 항공사진 공유와 활용을 증진하도록 하였고(NGII, 2015), 앞으로는 촬영된 항공사진의 대규모 물량에 대한 품질관리가 필요함에 따라 현행 육안검사 방식에서 시스템 기반의 자동검사방식으로 전환함으로써 안정적·체계적으로 품질을 관리하고 업무 생산성을 높일 필요가 있다. 항공사진 품질관리 분야 연구로는 Cho *et al.* (2002)이 필름 방식의 항공사진 촬영당시 시험노출과 농도측정 등의 정량적 방법을 제안하여 당시의 작업방법과 품질검사방법의 개선방안을 제안한 바 있다. Zhang *et al.* (2014)는 사진모자이크에 필요한 단일 사진 품질 평가, 사진 매칭을 기반으로 하는 사진시리즈 품질 평가 등의 품질 분석 시스템을 제안하였다. 이외에 최근 들어서 UAV(Unmanned Aerial Vehicle) 촬영영상에 대한 품질 평가 연구들이 발표된 바 있으나, 항공사진의 정성적인 기준, 즉 구름, 구름 그림자, 빛 반사 등에 대한 자동화된 추출과 기준에 대한 연구는 거의 이루어지지 않은 실정이다(Zhang *et al.* 2016; Lee and Sung, 2019).

현재 우리나라는 항공사진 성과를 촬영원본보다 모자이크 색깔 조정 및 보정 등을 거쳐서 제공한다. 따라서 항공사진 품질 검사 시 사진 내 화소값을 직접 이용하기보다는 상대적인 화소값들의 분포와 객체 형상, 크기, 질감 등에 많이 의존하게 된다. 최근 들어 딥러닝 기반의 영상추출분야가 급격히 확대되어 의료영상의 판독, 하수관 결함탐지, 콘크리트 균열 탐지 등 촬영된 영상들을 학습시켜 목적으로 하는 판독요소를 자동 추출하는 연구가 활발히 진행 중이다(Song *et al.*, 2019; Kim *et al.* : 2019). 따라서 항공사진의 특징적 요소들을 학습시켜서 정성적 품질 평가대상을 자동 추출하는 방안을 고려할 수 있다.

본 연구에서는 촬영된 항공사진의 구름 존재에 대한 학습 기반의 품질검사 자동화를 위해 구름에 대한 객체 탐지, 분류 기법을 적용하여 정확도를 계산하였고, 학습에 필요한 학습 자료 제작방법을 구분하여 결과를 산출하였으며 그 결과를 비교 분석하였다. 항공영상 데이터를 Arcmap 및 MATLAB 프로그램을 사용하여 분류와 객체 탐지 알고리즘에서 요구

하는 학습 데이터세트 형식에 맞게 가공하였고, 실험에 사용한 알고리즘으로는 GoogLeNet, VGG16, Faster R-CNN 그리고 YOLOv3를 적용하였다. 분류 및 객체 탐지에 따른 탐지 결과는 오차행렬 형태로 구축하여 평가하였으며 구름에 대한 생산자정확도에 초점을 맞춰 결과를 비교 분석하였다. 딥러닝을 사용하여 항공영상 품질검사 자동화를 이루게 되면 시간과 비용적인 절감이 가능할 것으로 기대된다.

## 2. 학습기반 항공사진 구름 탐지 기술

원격탐사에서 구름탐지는 대부분 구름의 분광특성을 이용하여 구름 추출에 적합한 파장대역인 단파적외 밴드나 costal/aerosol, cirrus밴드와 같이 특정 밴드로부터 임계값을 적용하여 구름을 탐지하는 것으로 (Sun *et al.*, 2017), Landsat 8호 영상의 경우 QA(Quality Assessment)밴드를 이용하거나(Lee *et al.*, 2019) 추가적인 밴드를 활용하며 threshold Otsu 알고리즘을 이용하여 두껍고 얇은 구름지역을 탐지하는 연구도 진행되었다(Kim *et al.*, 2014). 또한, 구름탐지와 직접적인 관련 있는 밴드가 없는 위성영상에 대한 구름탐지 연구도 진행되었고, 구름탐지에 효과적인 밴드를 가진 위성영상으로 사전학습하여 전이학습을 하면 구름탐지에 효과적임을 입증한 바 있다(Seong *et al.*, 2020). 그런데 위성영상을 대상으로 한 기존 연구들은 영상이 가진 원화소값이나 분광반사율을 기반으로 연구가 진행되었다. 사용자에게 제공되는 최종 산출물인 항공사진의 경우 원화소값을 유지하지 않고, 특히 정사영상의 경우도 모자이크 등의 작업을 거치면서 분광특성을 활용하기는 어렵다. 따라서 본 연구에서는 위와 같은 환경을 고려하여 원화소값을 고려하지 않은 상태에서 딥러닝 기술을 적용하기로 한다. 일반적으로 영상에 적용하는 딥러닝 기법으로는 분류, 객체 탐지 그리고 의미론적 분할 등을 들 수 있다. 이 중 의미론적 분할의 경우 구름 학습 데이터세트 제작에 대한 별도의 연구가 필요할 것으로 판단되어 분류와 객체 탐지뿐만 아니라 실험을 구성하였다. 분류와 객체 탐지 각각에 대해 2가지 알고리즘을 적용하여 결과를 나타냄으로써 일관성 있는 결과 도출이 되는지 확인하였다.

### 2.1 딥러닝 영상 분류

딥러닝에 의한 영상분류는 다양한 그룹에 해당하는 이미지들을 그룹별로 나눠 이를 학습시키고 이렇게 학습시킨 모델에 신규 영상을 입력하여 특정 그룹으로 분류하는 기법을 말한다. 본 연구의 대상이 되는 구름 영상분류의 경우, 구름

영상과 비구름영상을 나눠서 학습을 하도록 한다. 분류 알고리즘에는 ResNet(He *et al.*, 2015), AlexNet(Krizhevsky *et al.*, 2012), ZFNet(Zeiler *et al.*, 2014), VGGNet(Simonyan *et al.*, 2015), GoogLeNet(Szegedy *et al.*, 2015) 등이 있으며 본 연구에서는 AlexNet, ZFNet에 비해 오분류를 크게 줄였다고 평가받는 VGGNet, GoogLeNet을 실험에 적용하였다.

VGGNet은 2014년 이미지넷 인식 대회(ILVRC)에서 2등을 차지한 알고리즘으로, CNN(Convolutional Neural Network) 모델의 깊이가 모델의 성능에 중요한 역할을 한다는 것을 보여주었다. VGGNet은 크게 4가지로 계층의 수로 구분하며 VGG11, VGG13, VGG16, VGG19가 존재한다. 이 VGGNet들은 모두 3x3 합성곱 필터를 여러 번 사용하여 기존 존재하던 CNN인 AlexNet보다 높은 성능을 보여주었다(Simonyan and Zisserman, 2015). GoogLeNet은 2014년 이미지넷 이미지 인식 대회(ILSVRC)에서 VGG19를 이기고 우승을 차지한 알고리즘으로, VGG19 보다 많은 22개의 계층으로 구성되었다.

GoogLeNet은 다른 CNN 신경망들과는 다르게 1x1 합성곱 필터를 주로 사용하여 연산량을 줄임과 동시에 중요한 특징을 추출하며, 신경망 내부에 Inception 모듈을 사용하여 다양한 크기의 합성곱 필터를 이용한 특징들을 추출하는 신경망이다(Szegedy *et al.*, 2015). 이미지넷 인식 대회에서는 GoogLeNet이 우승을 차지하였지만, 다른 연구에서는 VGGNet이 더욱 좋은 성능을 보여주었다(Simonyan and Zisserman, 2015).

## 2.2 딥러닝 객체 탐지

객체 탐지는 영상 내에 특정객체를 영역으로 지정하여 이를 학습시키고, 지정된 영역과 지정되지 않은 영역간의 특징 차이를 학습하여 신규 영상을 제시했을 때 기존에 학습시킨 영역이 영상 내 어디에 존재하는지를 나타내도록 한다. 분류 기법과 달리 객체 탐지의 경우 객체 영역을 일일이 지정하여 학습자료를 제작하도록 한다. Fig. 1(a)에서는 항공사진에서 구름 영역을 지정한 그림을 나타냈고, Fig. 1(b)는 테스트영상에서 결과로 나온 탐지영역을 나타낸다. 탐지 알고리즘으로는 R-CNN(Girshick *et al.*, 2014), MR-CNN(Liu *et al.*, 2019), Mask R-CNN(He *et al.*, 2017), RetinaNet(Lin *et al.*, 2017), SSD(Liu *et al.*, 2016), YOLO(Redmon *et al.*, 2016), Faster R-CNN(Ren *et al.*, 2015) 등이 있고, 본 연구에서 적용한 실험에는 YOLO 알고리즘 중 YOLOv3(Redmon *et al.*, 2018)와 Faster R-CNN을 사용하였다.

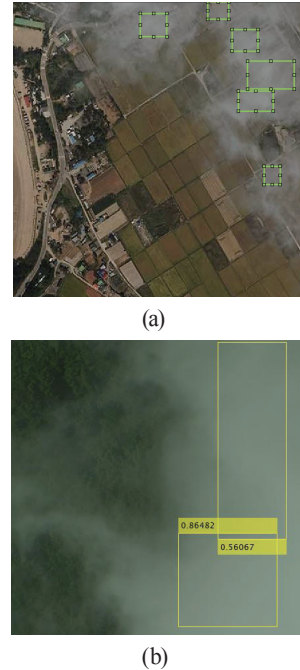


Fig. 1. Object detection in experimental images

(a) Object detection data labeling example,

(b) Examples of detection results

Faster R-CNN은 2015년 발표된 객체 탐지 알고리즘으로, 기존 R-CNN과 다른 방식인 Region Proposal 생성을 RPN이라는 새로운 방식으로 대체하여 RoI를 찾고, 찾은 RoI를 객체 분류 하는 과정으로 이루어져있다. RPN은 Fully convolutional network 형태로, Fast R-CNN과 합성곱 계층을 공유하여 기존 R-CNN이 지니고 있는 RoI(Region of Interest)마다 CNN 연산을 하여 발생하는 속도저하와 학습에서 발생하는 파이프라인과 같은 한계점을 극복하여 기존에 비해 실시간 객체 탐지가 원활하도록 개선한 알고리즘이다(Ren *et al.*, 2015).

YOLO는 You Only Look Once의 줄임말로 2016년 새로운 접근방식으로 제안된 객체 탐지 알고리즘이다. 기존 Fast R-CNN과 같은 객체 탐지에 비하여 월등한 예측속도를 지니고 있으며, 기존 객체 탐지에서 발생하는 복잡한 연산 과정을 이미지 전체가 하나의 신경망으로 한 번의 예측이 가능하도록 구현하여 1초에 약 40프레임 가량을 처리할 수 있다(Redmon *et al.*, 2016). 이런 우수한 처리속도, 정확도 때문에 YOLO는 꾸준히 개선되어왔고, 현재 Python 코드로 YOLO 버전5까지 개발되었으며, 기존에 비해 더욱 빠른 처리속도, 높은 정확도로 많은 어플리케이션에 응용되고 있다.

### 3. 실험 결과 분석

#### 3.1 실험자료 및 방법

실험에 사용된 항공사진은 국토지리정보원의 협조로 제공되었고, 구름 영상의 경우에는 구름이 포함된 12cm 해상도의 다양한 사이즈의 원본 영상을 활용하였다. 프로그램에 입력될 영상 크기를 일정하게 하기 위해 다양한 해상도를 가진 원본 영상을 448×448 크기로 split을 진행하였으며, 그 중 구름 영상만을 추출하여 총 8,021장의 구름 학습 자료를 구축하였다. 추가적으로 사용한 구름 위성영상 자료는 NWPU\_RESICS45(Northwestern Polytechnical University-Remote sensing Image Scene Classification) (<https://www.tensorflow.org/datasets/catalog/resisc45>)와 Sentinel 2A MSIL1C(<https://zenodo.org/record/4172871#.YYnmTGBByUk>)를 사용하였다. NWPU\_RESICS45는 Northwestern Polytechnical University에서 만든 자료이며, 20cm급부터 30m급까지 다양한 공간 해상도를 가지는 256×256 크기의 영상 31,500장으로 구성되어 있고 그 중 구름에 대한 영상은 700장으로 구성되어있다. Sentinel 2A MSIL1C은 10m급 해상도를 가지는 512×512 크기의 구름 영상 513장으로 이루어져 있다.

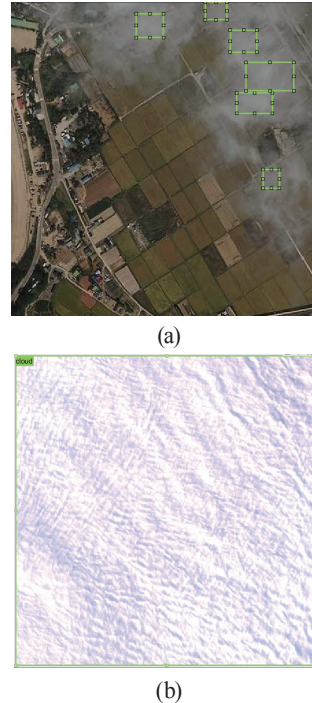
항공영상으로 구축한 구름 학습 자료에서는 바다와 같은 큰 면적의 수역은 존재하지 않았기 때문에 위성영상자료에서도 수역이 포함된 영상은 제외하였다. 결과적으로 NWPU\_RESICS45에서는 410장의 영상을 획득하였으며, Sentinel 2A MSIL1C에서는 154장의 영상을 구름에 대한 추가 학습 자료로써 사용하였다.

실험을 수행하기 위해 사용한 구름 영상이 아닌 데이터는 항공 구름 영상과 마찬가지로 국토지리정보원으로부터 12cm/25cm 해상도 영상을 제공받아 구름 영상과 동일한 JPG 형식으로 구축하였다. 구름이 없는 영상도 구름 영상과 동일하게 448×448 크기로 split을 수행하였다.

객체 탐지를 수행하기 위한 구름 클래스의 데이터라벨링은 실험에 사용한 MATLAB 2021a 버전의 기능에 적합한 학습 자료 형식에 맞추기 위해 MATLAB 내부의 'imageLabeler' 도구를 활용하여 수행하였다. 학습 자료는 영상의 경로와 영상 내 데이터라벨링을 수행한 위치가 절대좌표의 형태로 표현되며 MATLAB 작업공간에 테이블 형태로 저장된다. Fig. 2(a)는 항공영상에서 구름 객체의 데이터라벨링을 진행한 예시이며, Fig. 2(b)는 Sentinel 2A MSIL1C에서 구름 객체의 데이터라벨링을 진행한 예시이다.

항공영상에서의 구름 데이터라벨링은 라벨링 영역 전체를

구름이라고 판단할 수 있게 구름의 외곽지역이 아닌 구름의 중심부에 라벨링을 수행하였다. 인공위성 영상에서는 항공영상보다 크고 다양한 구름이 하나의 영상에 존재하기 때문에 영상 전체를 구름이라고 판단할 수 있는 영상들에 대해 라벨링을 수행하였다. Table 1은 객체 탐지 학습 자료 구축을 위해 사용한 영상의 개수와 영상 내에서 구름 객체를 라벨링 한 개수에 대한 표이다.



**Fig. 2. Cloud object labeling in the image (a) Cloud object labeling example in aerial image, (b) Cloud object labeling example in Sentinel 2A MSIL1C image**

**Table 1. Number of images used for object detection and number of object labeling**

Type of images	No. of Images	No. of Labeling
Aerial Images	366	1,209
Sentinel 2A MSIL1C	55	55

분류와 객체 탐지 기법 모두 동일한 테스트 자료를 사용하여 실험을 수행하였다. 테스트 자료는 원본영상을 448×448 크기로 split을 수행하여 총 2,912장의 영상으로 구성하였다. 테스트 자료의 영상별 클래스는 육안 검사를 통해 부여하였



으며 Fig. 3(a)는 테스트 자료에서 구름 영상의 예시이고, Fig. 3(b)는 테스트 자료에서 구름이 존재하지 않는 구름이 없는 영상의 예시이다. 테스트 자료 내에서 구름 클래스로 분류된 영상은 519장이며, 구름이 없는 영상 클래스로 분류된 영상은 1,673장이다. 테스트 자료에 대한 분류 및 객체 탐지 예측 결과에 대해 오차행렬 형식으로 결과를 정리하고 그 중에서 전체정확도와 구름에 대한 생산자정확도를 구하는 영상 단위의 정확도 분석을 수행하였다.

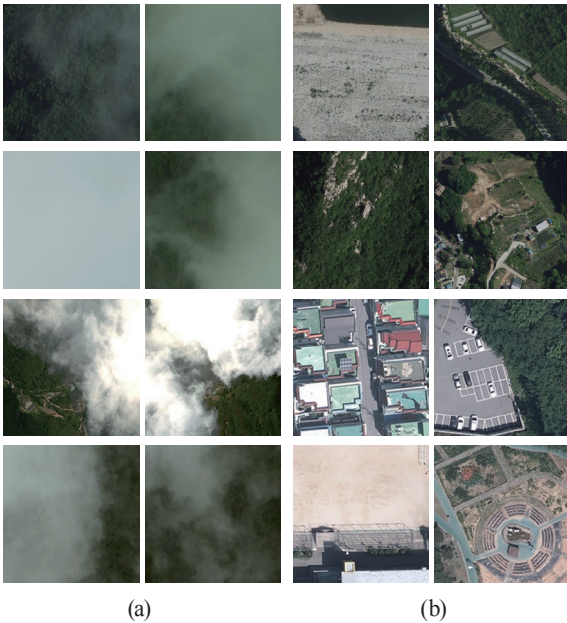


Fig. 3. Test data(examples of image by class) (a) An example of an image classified by cloud class, (b) An example of an image classified by non-cloud class

### 3.2 결과 및 분석

실험은 항공사진 내 구름의 존재여부를 판단하는 분류와 영상 내 구름 객체 탐지를 수행하는 2가지 기법에 대하여 각각 실험하였다. 각각의 기법별로 2가지의 알고리즘을 적용하였으며 결과의 일관성과 우수성을 살펴보았다. 또한, 현실적으로 구름을 포함하는 영상의 확보가 어려운 상황에서 위성 영상의 활용성을 실험을 통해 분석하였다.

#### 3.2.1 구름 분류와 탐지 실험 결과

분류 실험은 GoogLeNet 알고리즘과 VGG16 알고리즘을 사용하여 수행하였으며 두 가지 알고리즘 모두 epoch 10, batchsize 10 환경 하에서 실험하였다. 실험결과 GoogLeNet

은 전체정확도 56.7% 이고 구름에 대한 생산자정확도가 80%로 계산되었고, VGG16의 경우 전체정확도 75.0%, 구름에 대한 생산자정확도가 58.2%로 나타났다. 두 알고리즘이 전체정확도와 생산자정확도가 각각 상반되게 나왔으나, 오류 영상을 누락하지 않는다는 측면에서 생산자정확도가 높은 것이 현실적이라 판단한다.

객체 탐지 실험에서의 알고리즘은 Faster R-CNN과 YOLOv3를 사용하였다. 각각의 알고리즘은 시스템적 환경 및 코드의 구성에 따라 Faster R-CNN은 epoch 20, batchsize 3, default featureExtractionNetwork 옵션인 Resnet50과 default featureLayer 옵션인 activation\_40\_relu에서 실행되었으며, YOLOv3의 경우에는 epoch 20, batchsize 10, default featureExtractionNetwork 옵션인 squeezeNet과 default featureLayer 옵션인 fire5-concat 환경에서 실험을 수행하였다. epoch는 학습에 사용한 데이터셋 전체에 대해 한 번의 학습을 완료한 상태이며, batchsize는 1회 학습에 사용하는 영상의 개수를 의미한다. 실험 컴퓨터의 성능 내에서 Faster R-CNN을 구동시킬 수 있는 batchsize 최대치인 3으로 동일하게 YOLOv3의 batchsize를 지정할 경우 구름에 대한 생산자정확도가 47.5%로 매우 낮게 측정되었다. 또한 YOLOv3의 경우에는 Faster R-CNN보다 더 높은 batchsize 환경을 구성하여도 알고리즘의 정상적인 구동이 가능하였기 때문에 상대적으로 더 높은 구름에 대한 생산자정확도를 도출한 batchsize 10에 맞추어 YOLOv3에 대한 실험을 수행하였다. featureExtractionNetwork는 알고리즘에서 특징을 추출한 뒤 학습하기 위해 사용되는 네트워크의 종류를 의미하며, featureLayer는 특징을 추출하기 위해 사용되는 네트워크의 종류를 의미한다. Faster R-CNN의 경우 일반적인 객체 탐지 확률 기준에서는 정확도가 매우 낮게 도출되었고, 객체의 threshold를 낮출수록 구름에 대한 생산자정확도와 전체정확도가 증가하는 경향을 보여주었다. 객체의 threshold란 학습된 객체 탐지 알고리즘을 사용하여 영상 내 객체를 탐지할 때 threshold로 지정한 값 이상의 확률을 가진 영역을 객체로써 탐지하라는 뜻이다. 본 실험에서는 구름 객체에 대한 threshold를 0.1로 설정하였을 때 전체정확도 83.6%, 구름에 대한 생산자정확도 60.1%를 보여주었다. YOLOv3의 경우에는 구름에 대한 객체 탐지 확률의 threshold를 0.5로 설정하였을 때 전체정확도 89.0%에 구름에 대한 생산자정확도 84.0%를 보여주었다.

결과적으로 MATLAB 환경 하에서 구름 항공영상 분류 실험의 경우 GoogLeNet이 높은 구름에 대한 생산자정확도 분류 결과를 보여주었으며, 항공영상 내 구름 객체 탐지 실

험의 경우에는 YOLOv3가 높은 구름에 대한 생산자정확도를 보여주었다.

### 3.2.2 위성영상 학습자료를 활용한 구름 분류 추가 실험 결과

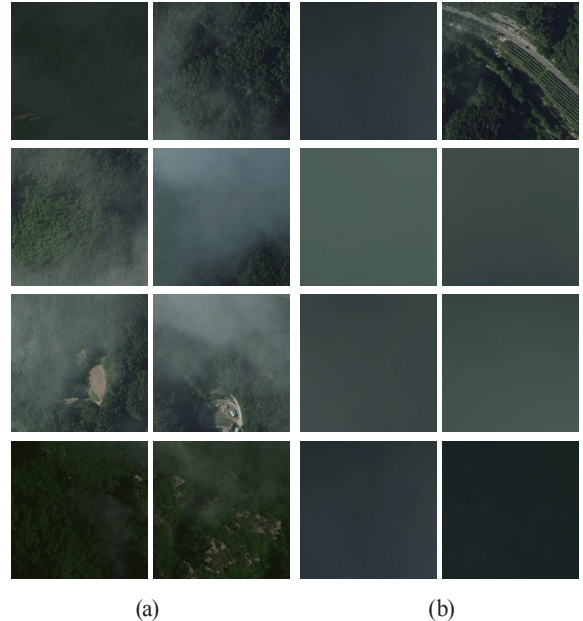
구름이 포함된 항공사진 학습자료가 부족한 경우를 고려해서 인공위성 구름영상을 추가로 포함하여 학습하는 것이 구름에 대한 생산자정확도에 영향을 미치는지를 확인하기 위해 Table 2와 같이 다양한 실험을 수행하였다. NWPU\_RESICS45 구름 학습 자료와 Sentinel 2A MSILIC 구름 영상으로 구성된 구름 학습 자료를 각각 추가한 경우, NWPU\_RESICS45 구름 학습 자료를 추가했을 때는 구름 항목에 대한 생산자정확도가 오히려 하락했지만 Sentinel 2A MSILIC 학습 자료의 경우에는 NWPU\_RESICS45 구름 학습 자료보다 영상의 개수가 더 적었음에도 불구하고 구름 항목에 대한 생산자정확도가 더 높아진 결과를 보여주었다. 여기서 구름에 대한 생산자정확도란 학습된 알고리즘이 테스트 데이터셋의 영상들 중 구름이 존재하는 영상으로부터 구름이 존재한다고 탐지해낸 경우를 백분율로 나타낸 값이다. 즉, 구름에 대한 생산자정확도가 높을수록 입력 영상으로부터 구름을 잘 탐지해낸다고 볼 수 있다.

**Table 2. Classification experiment result with satellite image learning data added**

case	cloud learning images	algorithm	overall accuracy (%)	cloud producer's accuracy (%)
1	Aerial images	GoogLeNet	56.7	80
2	Aerial images +①	GoogLeNet	60.4	73.0
3	Aerial images +②	GoogLeNet	43.0	83.8
4	Aerial images +①+②	GoogLeNet	69.7	68.0
5	Aerial images	VGG16	75	58.2
6	Aerial images +①	VGG16	75.4	49.5
7	Aerial images +②	VGG16	65.1	52.8
8	Aerial images+①+②	VGG16	74.8	63.2

\* ① : NWPU\_RESICS45, ② : Sentinel 2A MSILIC

Fig. 4는 Table 2의 2번 실험에 대한 결과로써, 구름이 존재하는 영상을 구름 영상으로 분류해낸 영상과 구름이 존재하지만 구름 영상으로 분류해내지 못한 영상에 대한 예시이다.



**Fig. 4. Classification results of case 2 in Table 2, (a) An image in which clouds exist is classified as a cloud image, (b) An image in which clouds exist but cannot be classified as a cloud image.**

또한, 구름에 대한 학습데이터로 NWPU\_RESICS45 구름 학습 자료와 Sentinel 2A MSILIC 학습 자료를 추가했을 때 경우에 따라 구름에 대한 생산자정확도의 향상을 보여주었지만 사용한 알고리즘에 따라 효과적인 인공위성 영상 학습 자료가 다르다고 할 수 있었다. GoogLeNet의 경우 위성영상을 추가하기 전 가장 좋은 구름 항목에 대한 생산자정확도를 보여준 실험은 구름과 구름이 없는 영상 모두 고해상도의 영상을 사용하여 학습을 수행한 경우였으며, 이 때 구름에 대한 생산자정확도는 80.0%에 달했다. 인공위성 학습 자료를 추가하여 실험할 때에도 Sentinel 2A MSILIC 구름 학습 자료를 추가하여 실험한 경우에 구름에 대한 생산자정확도가 증가한 경향을 보여주었다. 이 경우 실험에 사용한 Sentinel 2A MSILIC 구름 학습 자료가 NWPU\_RESICS45 구름 학습 자료보다 상대적으로 적었음에도 불구하고 구름 탐지를 더 잘 해낸 것을 볼 수 있다. VGG16의 경우에는 NWPU\_RESICS45 구름 학습 자료와 Sentinel 2A MSILIC 구름 학

습 자료 모두를 추가했을 때 가장 높은 구름에 대한 생산자정확도를 보여주었다.

실험결과 GoogLeNet 알고리즘이 VGG16에 비해 구름분류에서 구름에 대한 생산자정확도가 좋았던 이유는 분류하고자 하는 객체의 특성 때문으로 여겨진다. GoogLeNet 알고리즘의 경우 영상에서 특징을 추출할 때 1×1 크기의 필터, 3×3 크기의 필터, 5×5 크기의 필터 등 다양한 크기의 필터를 사용하여 특징을 추출한다. 반면, VGG16의 경우에는 3×3 크기의 필터만을 사용하여 네트워크를 깊게 구성하기 때문에 구름과 같은 부정형의 객체의 특징추출에 한계가 있을 수 있다. 따라서 이러한 부정형의 객체에 대하여 특징을 추출하고 학습하며 분류할 때는 상대적으로 다양하고 작은 사이즈의 필터를 사용하여 특징을 추출해 내는 GoogLeNet이 VGG16에 비해 유리하며 이는 외부 데이터를 사용하여 추가적인 학습 자료를 구축할 때에도 유효하다고 볼 수 있다. 따라서 부정형 객체에 대한 특징 추출을 진행할 때에는 그 경계를 명확하게 하기 위한 방편으로 높은 해상도를 가진 영상을 사용하거나 다양한 크기의 필터를 사용하여 특징을 추출해내는 것이 유리할 것이다.

구름의 추가가 구름에 대한 생산자정확도를 대체적으로 높여주긴 했지만 전체정확도의 측면에서는 오히려 좋지 않은 경향을 보여준 케이스 또한 많았다. 따라서 본 실험에서의 결과로써는 위성영상의 추가가 좋은 영향을 미쳤다고는 볼 수 없으나 전체정확도를 낮추더라도 탐지하지 못하는 구름 영상이 없어야 하는 품질검사에 적용하기 위한 모델을 구축할 때 현실적인 자료 확보가 어렵다면 위성영상을 사용한 구름 학습자료가 대안이 될 수 있을 것이다. 또한, 구름은 자연 내에서 농도, 크기, 모양, 위치가 매우 다양하다. 구름의 농도에 의해 영상에서 구름이 가지는 시각적 특성이 다를 수 있고, 구름의 크기에 따라 영상에서 구름이 차지하는 비율이 달라질 수 있다. 구름의 모양과 위치에 따라 영상에서 구름이 존재하는 영역 외의 토지피복이 학습에 영향을 줄 수도 있다. 이러한 경향은 더 넓은 지역을 한 번에 촬영하는 위성영상에서 더 두드러진다. 하지만 본 실험에서는 자연에 존재하는 구름의 이러한 특징을 고려하지 않고 구름이 존재하는 영상을 모두 하나의 구름 영상으로 분류하여 디렉토리를 구축하였다. 이로 인해 구름 영상으로부터 추출되는 특징이 정형화되지 않았을 것으로 사료된다.

### 3.2.3 위성영상 학습 자료를 활용한 구름 객체 탐지 추가 실험 결과

객체 탐지 실험에 대해서도 인공위성영상을 활용하여 구축한 구름 객체의 라벨링 학습 자료를 추가하여 실험을 수

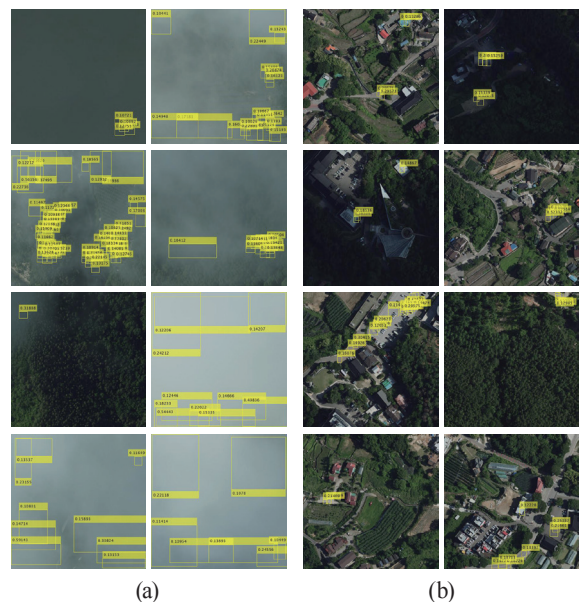
행하였다. Table 3은 위성영상을 추가하여 수행한 객체 탐지 실험결과를 정리한 표이다. 객체 탐지에 추가한 위성영상은 분류 실험에서 데이터를 추가함으로써 성능의 개선을 보여준 Sentinel 2A MSILIC 위성영상에 대해서만 데이터 라벨링을 수행하여 데이터의 범용성을 확인하는 실험을 수행하였다.

**Table 3. Detection experiment result with satellite image learning data added**

case	cloud training images	algorithm	overall accuracy(%)	cloud producer's accuracy(%)
1	Aerial images	Faster R-CNN	83.6	60.1
2	Aerial images+②	Faster R-CNN	65.0	80.9
3	Aerial images	YOLOv3	89	84
4	Aerial images+②	YOLOv3	86.9	67.1

\* ② : Sentinel 2A MSILIC

Fig. 5는 영상내 구름을 정상적으로 탐지해낸 경우와 구름이 존재하지 않는 영상에서 구름을 오탐지한 경우에 대한 그림이다.



**Fig. 5. Object detection results of case 1 in Table 3, (a) Images in which clouds are detected in images with clouds, (b) Images in which clouds are detected in images without clouds**



MATLAB 환경 하에서 Faster R-CNN의 경우 학습 시에 라벨링 형태로 지정한 객체를 제외한 나머지 영역을 background 클래스로 분류하여 학습을 진행한다. 이로 인해 탐지하려고 하는 객체를 영상에서 놓치거나 객체에 대한 영역을 정확하게 지정하지 않을 경우 탐지하려는 객체임에도 불구하고 background 클래스로 학습되게 된다. 이러한 학습 특징은 결과적으로 구름과 같은 부정형객체를 탐지할 때 객체의 특징을 희석시키고 오탐지 가능성을 높이는 원인이 될 수 있다. Faster R-CNN의 경우 위성영상에 대해 라벨링을 수행한 학습 자료를 추가하였을 때 threshold 0.5에서는 여전히 낮은 구름에 대한 생산자정확도를 보여주었지만 threshold를 0.1로 낮추었을 때는 기존보다 더 높은 80.9%의 구름에 대한 생산자정확도를 보여주었다. YOLOv3의 경우에는 Sentinel 2A MSILIC 영상에 대해 수행한 데이터라벨링을 추가했을 때 오히려 구름에 대한 생산자정확도가 낮아지는 경향을 보여주었다. 구름의 경우 다양한 모양을 가지고 있는 부정형 객체이기 때문에 본 실험에서 수행한 사각형 형태의 라벨링은 적절하지 않다고 볼 수 있다. 정확한 구름 탐지를 위해 영역 내가 모두 구름일 경우에 대해서 라벨링을 진행하였으므로 상대적으로 지표의 형상과 섞이는 외곽지역의 경우 경계가 매우 복잡하여 구름의 다양한 형상에 대한 충분한 학습이 이루어지지 않은 것으로 보인다. 또한 추가한 위성영상에서의 데이터라벨링도 항공영상과 동일하게 사각형 형태의 라벨링을 수행하였다. 따라서 객체 라벨링에 따른 케이스의 다양화가 개선되지 않은 상태에서 위성영상에서의 구름에 대한 특징만이 기존의 항공영상에서의 구름에 대한 특징과 섞이며 정확도에서의 개선이 이루어지지 않은 것으로 사료된다. 그럼에도 불구하고 YOLOv3의 경우 GoogLeNet의 특징과 유사하게 입력된 이미지에 대해 서로 다른 3가지 크기의 필터를 사용하여 특징을 추출하며, Faster R-CNN과 다르게 라벨링을 수행한 영역에 대해서만 학습을 진행함으로써 구름과 같은 부정형 객체에 대해 학습하는 데 Faster R-CNN 기법보다 상대적으로 유리하다는 것을 확인할 수 있었다.

#### 4. 결론

본 논문에서는 항공영상에서 구름이 있는 영상과 구름이 없는 영상을 학습하여 구름 분류와 탐지를 실험하였고, 인공위성영상을 활용하여 구축한 구름 학습 자료를 추가함으로써 구름 분류와 탐지를 실험하였다. 실험결과, 항공영상만을 사용하여 학습 자료를 구축한 뒤 실험을 수행했을 때 GoogLeNet 알고리즘에서 80%의 구름에 대한 생산자정확도

를 나타내어 구름이라는 부정형 객체가 포함된 영상에 대한 자동화 검사의 적용 가능성을 보여주었다.

인공위성영상에 대한 구름 학습 자료를 추가하였을 때는 구름에 대한 생산자정확도 대해서 VGG16보다 GoogLeNet이 더 좋은 결과 보여주었다. 이와 같은 결과는 VGG16과 GoogLeNet 알고리즘의 특성 추출 단계에서의 차이점이 원인으로 보이고, 항공영상에 인공위성영상을 학습 자료로 추가했을 때 구름에 대한 생산자정확도가 증가하는 경향을 보여주었다. 객체 탐지의 경우에는 GoogLeNet과 유사하게 다양한 크기의 필터로 특징을 추출하는 YOLOv3 알고리즘이 구름에 대한 높은 생산자정확도를 보여주었다. 하지만 YOLOv3의 경우에는 GoogLeNet과 다르게 인공위성영상에 대한 학습 자료를 추가하였을 때 전체정확도와 구름에 대한 생산자정확도 모두 낮아지는 경향을 보여주었는데 이는 항공영상에서의 구름과 인공위성 영상에서의 구름 객체 간 형상적 차이 때문인 것으로 보인다.

객체 탐지의 경우 인공위성 영상을 기반으로 추가적으로 라벨링을 진행하여 학습을 진행하였을 때 전반적인 정확도의 개선이 이루어지지 않았다. 이는 데이터라벨링의 단계에서 항공영상에서의 구름 형상과 인공위성 영상에서의 구름 형상 간의 차이로 발생하는 데이터라벨링의 어려움을 해결하여 일관된 데이터의 경향성을 구축할 수 있는 데이터라벨링에 대한 방법론을 고안해내야 해결할 수 있을 것으로 보인다.

구름은 그 농도에 따라 다양한 시각적 강도를 가지는 객체이다. 또한 모양이 불특정하여 다양한 형상을 가지는 객체이기도 하다. 본 실험에서는 구름의 농도, 영상 내에서 구름이 존재하는 영역 비율, 영상 내 구름이 존재하지 않는 영역의 토지피복 특성 등을 고려하지 않아서 이로 인한 정확도 영향을 추후 연구할 필요가 있다 또한 객체 탐지 학습 자료를 구축할 때에도 사각형 형태의 라벨링을 사용함으로써 구름 객체의 정확한 영역 지정에 한계가 존재하였다. 따라서 추후 구름에 대한 학습 데이터셋을 구축하기 위해서는 영상 내 존재하는 구름의 농도, 구름이 존재하는 영역의 비율, 구름이 존재하지 않는 영역의 토지피복에 대한 특성 등을 고려하여 다양한 경우의 구름 영상 디렉토리로 구축하여 상세한 클래스로 나누어 실험하고 정확도에 미치는 영향요소를 분석하고 이러한 실험과정을 통해 학습자료를 제작하는 작업자가 표본적인 학습자료 제작과정 수립연구가 진행되어야 할 것으로 사료한다.

#### 감사의 글

본 연구는 2021년 국토지리정보원 “항공영상 품질검사 자



동화체계 연구”사업의 지원을 받아 수행된 연구임.

## References

- Cho, W., Lee, S.H., Choi, S.S., Hwang, H.D., and Lee, H.J. (2003), Quality Improvement on Aerial Photographs, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 21, No. 3, pp. 277-287. (in Korean with English abstract)
- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014), Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 24-27 June, Columbus, Ohio, pp. 580-587.
- He, K., Gkioxari, G., Dollar, P., and Girshick, R. (2017), Mask R-CNN, *IEEE International Conference on Computer Vision (ICCV)*, 22-29 October, Venice, Italy, pp. 2961-2969.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015), Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, *IEEE International Conference on Computer Vision (ICCV)*, 13-16 December, Santiago, Chile, pp. 1026-1034.
- Kim, B.H., Kim, Y.H., Han, Y.K., Choi, W.S., and Kim, Y.I. (2014), Fully Automated Generation of Cloud-free Imagery Using Landsat-8, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 32, No. 2, pp. 133-142. (in Korean with English abstract)
- Kim, B., Kim, G., Jin, S., and Cho, S. (2019), A Comparative Study on Performance of Deep Learning Models for Vision-based Concrete Crack Detection according to Model Types, *Journal of the Korean Society of Safety*, Vol. 34, No. 6, pp. 50-57. (in Korean with English abstract)
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012), ImageNet classification with deep convolutional neural networks, *In Proceedings of NIPS*, pp. 1106-1114.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998), Gradient-based learning applied to document recognition, *IEEE*, vol. 86, no. 11, pp. 2278-2324.
- Lee, M.H., Cheon, E.J., and Eo, Y.D. (2019), Cloud Detection and Restoration of Landsat-8 using STARFM, *Korean Journal of Remote Sensing*, Vol. 35, No. 5-2, pp. 861-871. (in Korean with English abstract)
- Lee, J.O., and Sung, S.M. (2019), Quality Evaluation of UAV Images Using Resolution Target, *Journal of the Korean Association of Geographic Information Studies*, Vol. 22, No. 1, pp. 103-113. (in Korean with English abstract)
- Lin, T.Y., Goyal, P., Girshick, R., He, K., and Dollar, P. 2017, Focal Loss for Dense Object Detection, *IEEE International Conference on Computer Vision (ICCV)*, 2017, 22-29 October, Venice, Italy, pp. 2980-2988.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., and Berg, A. C. 2016, SSD: Single Shot MultiBox Detector, In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.), *Computer Vision – ECCV 2016*, Springer, Cham. pp. 21-37.
- Liu, Z., Du, J., Tian, F., and Wen, J. (2019), MR-CNN: A Multi-Scale Region-Based Convolutional Neural Network for Small Traffic Sign Recognition, *IEEE Access*, vol. 7, pp. 57120-57128.
- Mittal, A., Moorthy, A.K., and Bovik, A.C. (2012), No-Reference Image Quality Assessment in the Spatial Domain, *IEEE TRANSACTIONS ON IMAGE PROCESSING*, Vol. 21, No. 12, pp. 4695-4708.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016), You Only Look Once: Unified, Real-Time Object Detection, *29th IEEE Conference on Computer Vision and Pattern Recognition*, 27-30 June, Las Vegas, Nevada, USA, pp. 779-788.
- Redmon, J., and Farhadi, A. (2018), YOLOv3: An Incremental Improvement, *arXiv*, <https://arxiv.org/abs/1804.02767> (last date accessed: 17 December 2021).
- Ren, S., He, K., Girshick, R., and Sun, J. (2015), Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 6, pp. 1137-1149.
- Seong, S.K., Choi, S.K., and Choi, J.W., Cloud Detection of PlanetScope Imagery Based on Deeplab-V3+ by Using Transfer Learning, *Journal of Korean Society for Geospatial Information Science*, Vol. 28, No. 4, pp. 25-32. (in Korean with English abstract)
- Simonyan, K. and Zisserman, A. (2015), Very Deep

- Convolutional Networks For Large-Scale Image Recognition, *Computer Vision and Pattern Recognition*, Cornell University, Ithaca, NY, USA.
- National Geographic Information Institute (2015), *Data Quality for Aerial Photograph*, NGII-STD.2015-7/2020, National Geographic Information Institute, Suwon, Korea.
- Song, H.J., Lee, E.B., Jo, H.J., Park, S.Y., Kim, S.Y., Kim, H.J., and Hong, J.W. (2020), Evaluation of Classification and Accuracy in Chest X-ray Images using Deep Learning with Convolution Neural Network, *Journal of the Korean Society of Radiology*, Vol. 14, No. 1, pp. 39-44. (in Korean with English abstract)
- Sun, L., Mi, X., Wei, J., Wang, J., Tian, X., Yu, H., and Gan, P. (2017), A cloud detection algorithm generating method for remote sensing data at visible to short-wave infrared wavelengths, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 124, pp. 70-88.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015), Going Deeper with Convolutions, *2015 IEEE Conference on Computer Vision and Pattern Recognition*, 7-12 June, Boston, MA, USA, pp. 1-9.
- Zhang, H., Li, B., Zhang, J., and Xu, F. (2014), Aerial Image Series Quality Assessment, *35th International Symposium on Remote Sensing of Environment-2013*, 22-26 April, Beijing, China, Vol. 17, pp. 012183.
- Zhang, G.F., Wang, L., Zheng, Z.N., Chen, Y.F., Zhou, Z.B., and Zhao K. (2016), No-reference aerial image quality assessment based on natural scene statistics and color correlation blur metric, *13th International Conference on Transmission & Distribution Construction, Operation & Live-Line Maintenance*, 21 August, Columbus, OH, USA, pp. 1-4.
- Zeiler, M.D., and Fergus, R. (2014), Visualizing and understanding convolutional networks, In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.), *ECCV 2014*, Springer, Cham, vol. 8689, pp. 818-833.