

기계학습 기반의 파이썬 모듈을 이용한 밀양아리랑우주천문대 전천 영상의 운량 모니터링 프로그램 개발

임 구^{1,2,*} · 김도형^{1,2,3,*} · 김동현³ · 박근홍^{4,5}

¹부산대학교 지구과학과, 46241, 부산광역시 금정구 부산대학로 63번길 2
²부산대학교 미래지구환경연구소, 46241, 부산광역시 금정구 부산대학로 63번길 2
³부산대학교 지구과학교육과, 46241, 부산광역시 금정구 부산대학로 63번길 2
⁴밀양아리랑우주천문대, 50420, 경상남도 밀양시 밀양대공원로 86
⁵밀양시 시설공단, 경상남도 밀양시 시청로 26

Development of the Cloud Monitoring Program using Machine Learning-based Python Module from the MAAO All-sky Camera Images

Gu Lim^{1,2,*}, Dohyeong Kim^{1,2,3,*}, Donghyun Kim³, and Keun-Hong Park^{4,5}

¹Department of Earth Sciences, Pusan National University, Busan 46241, Korea
²Institute for Future Earth (IFE), Pusan National University, Busan 46241, Korea
³Department of Earth Science Education, Pusan National University, Busan 46241, Korea
⁴Miryang Arirang Astronomical Observatory, Miryang 50420, Korea
⁵Miryang City Facilities Management Corporation, Miryang 50420, Korea

Abstract: Cloud coverage is a key factor in determining whether to proceed with observations. In the past, human judgment played an important role in weather evaluation for observations. However, the development of remote and robotic observation has diminished the role of human judgment. Moreover, it is not easy to evaluate weather conditions automatically because of the diverse cloud shapes and their rapid movement. In this paper, we present the development of a cloud monitoring program by applying a machine learning-based Python module “cloudbright” on all-sky camera images obtained at Miryang Arirang Astronomical Observatory (MAAO). The machine learning model was built by training 39,996 subregions divided from 1,212 images with altitude/azimuth angles and extracting 16 feature spaces. For our training model, the F1-score from the validation samples was 0.97, indicating good performance in identifying clouds in the all-sky image. As a result, this program calculates “Cloudiness” as the ratio of the number of total subregions to the number of subregions predicted to be covered by clouds. In the robotic observation, we set a policy that allows the telescope system to halt the observation when the “Cloudiness” exceeds 0.6 during the last 30 minutes. Following this policy, we found that there were no improper halts in the telescope system due to incorrect program decisions. We expect that robotic observation with the 0.7 m telescope at MAAO can be successfully operated using the cloud monitoring program.

Keywords: Machine learning, All-sky camera, Observational astronomy

*Corresponding authors: dohyeong@pusan.ac.kr;
lim9gu@gmail.com
Tel: +82-51-510-2705

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.

요약: 운량은 천체 관측을 지속하는 데에 중요한 요소 중 하나이다. 과거에는 관측자가 날씨를 직접 판단할 수밖에 없었으나, 원격 및 자동 관측 시스템의 개발로 관측자의 역할이 상대적으로 줄어들었다. 또한 구름의 다양한 형태와 빠른 이동 때문에 자동으로 운량을 판단하는 것은 쉽지 않다. 이 연구에서는 기계학습 기반의 파이썬 모듈인 “cloudynight”을 밀양아리랑우주천문대의 전천 영상에 적용하여 운량을 모니터링하는 프로그램을 개발하였다. 전천 영상을 하위 영역으로 나누어 각 39,996개 영역의 16개의 특징을 학습하여 기계학습 모델을 생성하였다. 검증 표본에서 얻은 F1 점수는 0.97로, 기계학습 모델이 우수한 성능을 가짐을 보여준다. 운량(“Cloudiness”)은 전체 하위 영역 개수 중 구름으로 식별된 하위 영역 개수의 비율로 계산하며, 운량이 지난 30분 동안 0.6을 초과할 때 관측을 중단하도록 자동 관측 프로그램 규칙을 정하였다. 이 규칙을 따를 때, 기계학습 모델이 운량을 오관하여 관측에 영향을 미치는 경우는 거의 발생하지 않았다. 본 기계학습 모델을 통하여, 밀양아리랑우주천문대 0.7 m 망원경의 성공적인 자동 관측을 기대한다.

주요어: 기계학습, 전천 카메라, 관측천문학

서 론

천문학 연구에서 관측은 천체의 물리량을 얻을 수 있는 가장 중요한 실험 활동 가운데 하나다. 관측 자료를 얻기 위하여 다양한 파장 대역을 관측할 수 있는 다양한 종류의 망원경들을 활발히 이용하고 있다. 그 중 광학 망원경이 널리 이용되는데, 천체에서 나오는 빛을 관측하여 항성 및 은하의 밝기와 온도, 나이, 그리고 금속 함량(metallicity) 등 다양한 물리량들을 얻을 수 있다.

기본적으로 집광력이 좋은 큰 망원경일수록 어둡고 멀리 있는 천체의 빛을 많이 모을 수 있다는 장점이 있다. 따라서 구경이 수 cm 크기인 초기 망원경에서 점차 발전하여, 현재는 8 m급 또는 그 이상의 대형 망원경들이 천문학 연구에 활발히 이용되고 있으며, 가까운 미래에는 25-30 m급의 차세대 대형 망원경들도 완성될 예정이다.

그러나 최근에는 어두운 천체를 관측하는 데 적합한 대형 망원경뿐만 아니라 변광 관측에 적합한 소형 망원경의 중요성이 크게 대두되고 있다. 특히 식 현상 외계 행성(Wang et al., 2017; Alsubai et al., 2018), 활동성 은하핵의 변광(Walsh et al., 2009; Kim et al., 2019), 초신성과 같은 돌발 천체 연구(Im et al., 2015; de Jaeger et al., 2019; Cai et al., 2022; Deckers et al., 2022; Lim et al., 2023), 그리고 중력과 광학 검출 및 탐색 연구(Abbott et al., 2017; Paek et al., 2024) 등이 활발해짐에 따라, 넓은 영역을 자주 관측하거나 돌발 천체가 나타난 지역을 빠르게 후속 관측할 수 있는 중소형 망원경들이 점차 활발히 이용되는 추세이다(Djorgovski et al., 2011; Itoh et al., 2013; Shappee et al., 2014; Im et al., 2021).

천체 관측에서 가장 중요한 요소는 바로 날씨이다. 구름이 없고 수증기가 적은 깨끗한 날씨에 관측을 하면 대기 소광의 영향이 거의 없는 좋은 자료를 얻을 수 있지만, 구름이 많거나 비가 온다면 관측을 진행할 수조차 없기 때문이다. 일반적으로 관측자는 관측 중 날씨를 직접 판단하고 관측 진행 여부를 결정하지만, 중소형 망원경의 경우 원격 또는 자동 관측 화가 널리 이루어져 있기에, 관측자가 항상 직접 날씨를 판단하기는 매우 어렵다.

특히, 구름이 천체를 가리면 천체의 빛이 망원경에 도달하기 전에 흡수 혹은 산란되므로 매우 밝은 천체만 관측 가능하거나 아예 관측이 불가능하다. 따라서 관측을 성공적으로 수행하기 위해서는 운량을 판단하는 것이 매우 중요하다. 그러나, 구름의 모양은 매우 다양하며 바람에 의해 빠르게 이동하기 때문에 운량을 정량화하기가 어렵다. 전통적으로는 관측자가 경험을 토대로 운량을 판단해 왔는데, 원격 및 자동 관측에서는 관측자가 운량을 판단해서 관측을 진행하기 어렵다. 따라서 기계학습(machine learning) 모델을 적용한 자동화된 운량 판단 프로그램 개발이 원격 및 자동 관측 시스템 구축을 이루는데 있어 필수적이다.

Cloudynight (Mommert 2020)는 기계학습 모델을 이용하여 전천 관측 영상에서 다양한 특징(feature)으로부터 구름을 식별하도록 해주는 파이썬(Python) 기반의 모듈(module)¹⁾이다. cloudynight는 처리 속도가 빠르고 컴퓨터 메모리의 사용이 작아 대용량 데이터의 다양한 특징을 학습하기에 유리한 기계학습 모델로 알려진 lightGBM (Ke 2017) 모델을 사용한다. Mommert (2020)은 Lowell 천문대의 Discovery Channel

망원경으로부터 얻은 1,000개의 학습 데이터를 이용해 기계 학습 모델을 구축했고, 이 모델은 95%의 높은 정확도(accuracy)를 보여주었다.

밀양아리랑우주천문대에는 지름이 0.7 m이고 초점비가 6.5인 수정 Dall-Kirkham 형식의 반사 망원경이 설치되어 있으며, 최근 자동 관측 시스템(Lim et al. in prep.)이 구축되었다. 밀양아리랑우주천문대의 자동 관측 시스템을 통해 앞으로 초신성 폭발, 식현상 외계행성 등 각종 시계열 관측 연구를 활발히 할 수 있으리라 기대한다. 본 연구에서는 자동 관측을 성공적으로 수행하기 위하여, 밀양아리랑우주천문대의 전천 카메라(all-sky camera) 영상에 cloudynight 모듈을 적용하여 자동으로 운량을 계산하는 프로그램을 구축했다. 뿐만 아니라, 이러한 운량 모니터링 프로그램을 활용한다면, 국내의 많은 소형 망원경들에 자동 관측 시스템 구축이 성공적으로 이루어지고 이를 통해 더욱 다양한 시계열 관측 연구가 활발히 이루어질 수 있으리라 기대한다.

전천 카메라 장비

밀양아리랑우주천문대에서 이용하는 전천 카메라 장비의 모습은 Fig. 1과 같으며, 제원은 다음과 같다. 전천 카메라 광학계는 ALCOR SYSTEM사의 OMEA 6M을 사용하며, 초점거리가 1.8 mm (f/3)인 어안렌즈를 사용한다. 검출기는 ZWO사의 ASI178MM CMOS 카메라를 사용한다. 저장 장치의 용량 자원을 아끼고, 많은 점광원을 검출하기 위하여 2×2 비닝(binning)을 하여 전천 이미지를 얻는다. 이렇게 얻어진 전천 이

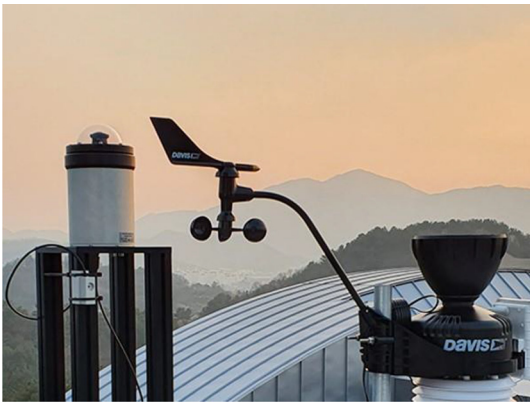


Fig. 1. All-sky camera equipment (a dome on the left) used in the MAAO (Image credit: METASPACE Inc.).

미지는 1,548×1,040 픽셀 배열(pixel array)로 구성되어 있으며, 픽셀 당 9.2 각분의 픽셀 스케일(pixel scale)을 가진다.

전천 카메라 장비는 자동날씨관측시스템(Automated Weather System) 컴퓨터에 연결되어 있다. 영상은 전천 카메라 장비의 전용 소프트웨어인 skywatcher를 통하여 1분에 한 번씩 촬영되어 압축된 fits.bz2 형식으로 컴퓨터에 저장되나, 운량 모니터 프로그램은 3분에 한 장씩만 저장된 전천 영상을 읽어 들인다. 노출 시간은 광량에 따라 자동으로 조절되며, 최소 32 마이크로초에서 최대 35초까지로 설정되어 있다.

학습 데이터 준비

학습 데이터 수집

기계학습 모델을 구축하기 위하여 확보된 전천 영상들을 수집하여 학습 데이터를 준비하였다. 2023년 11월 6일부터 2023년 12월 6일까지, 그리고 2024년 1월 24일부터 2024년 2월 29일까지 얻어진 영상은 모두 145,562장이다. 이 중 임의로 2,000장의 영상을 추출하여 썸네일로 만들었으며, 이를 다음의 기준으로 안시 분류(visual inspection)하였다. 구름의 형태는 매우 다양하고, 고도, 방위각, 그리고 촬영 시각에 따라 빛(태양과 달빛, 광해)의 영향이 각기 다르므로, 구름의 유무를 식별하는 것은 사용자의 주관에 상당히 개입할 수밖에 없다. 연구의 목적이 돛 슬릿의 개방 가능 여부를 결정하고자 하는 것임을 고려하여, 점광원이 보이지 않는 불투명한 영역이 적은 영상을 구름의 영향이 적은 영상(Not cloudy)으로, 나머지를 구름의 영향이 큰 영상(Cloudy)으로 분류하였다. Fig 2는 안시 분류된 전천 영상의 예시를 보여주며, 상단의 두 영상은 Not cloudy, 하단의 두 영상은 Cloudy로 분류된 것이다. 오른쪽 위의 영상은 얇은 권운(thin cirrus)을 포함하여 완전히 맑지는 않지만, 돛 슬릿을 개방하는 데 문제가 없을 것으로 판단하여 Not cloudy로 분류하였다. 그 결과 Not cloudy 영상은 512개로 분류되었다. Mommert (2020)에서 훈련 데이터 개수가 1,000개가 될 경우에 95%의 정확도(accuracy)를 보여주며, 그 이상 훈련 데이터 개수가 초과하더라도 정확도가 뚜렷하게 증가하지 않는다는 것을 보여주었다. 그럼에도 불구하고 이보다는 더 여유를 두어 700개의 Cloudy 영상을 추가하여 1,212개 영상을 선택하였다. 이 과정에서 일부 분류가 모호한

영상들이 not cloudy 혹은 cloudy쪽으로 재분류가 되면서 Not cloudy 523개 (42.7%), Cloudy 689개 (57.3%)로 최종적으로 결정되었다. 따라서 본 연구에서도 충분한 정확도의 기계학습 모델을 구축하는 데에 큰 문제가 없을 것으로 판단하였다.

전처리 정보 입력 및 전처리

Cloudynight는 전천 영상을 읽어들이며 마름질(trimming)하고, 마스크(masking)하여 모델 학습에 필요한 형태로 영상을 전처리한다. 먼저 하나의 전천 영상에 대하여, 천정을 포함한 다양한 고도와 방위각에 따른 33개의 하위영역(subregion)으로 분할한다. 그런 다음 각 하위 영역의 특징을 추출한다. 각 특징은 태양의 고도, 달의 고도, 달의 위상, 하위 영역의 위치와, 각 하위영역에서의 점광원 밀도, 배경하늘값의 단순평균(mean), 중간값(median), 표준편차(standard deviation)이다. 여기에 3분 전과 15분 전 영상의 점광원 밀도, 배경하늘값의 단순평균, 중간값, 표준편차까지 포함하여 총 16개 특징이 학습에 사용된다. 3분 전과 15분 전 영상의 특징이 사용되는 이유는 시간에 따른 구름의 이동을 반영하고자 추출한다. 자세한 설명은 Mommert (2020)을 참조하라.

다음으로 천문대 위치 정보, 검출기의 픽셀 배열, 파일 경로, 그리고 전천 영상을 cloudynight에 입력한다. 천문대의 위도(북위 35도 30분 08.6초), 경도(동경 128도 45분 40.0초), 그리고 해발고도(해발 95 m) 정보는 태양과 달의 고도를 계산하는 데 사용된다. 전천 영상의 점광원 밀도를 계산하는 측광 프로그램이 cloudynight에 포함되어 있으며, 측광 파라미터의 설정은 cloudynight의 설정을 따랐다. Fig. 3은 밀양 아리랑우주천문대의 전천 영상을 처리하는 과정을 보여준다. 첫번째로 원자료(Raw image)에서 돔 구조물과 영상 전천 카메라 가장자리를 마름질하여 제거한다. 두번째로 배경하늘값이 비슷한 100장의 이미지를 중간값으로 합성(stacking)한 후, 픽셀값이 11,000 ADU 이상인 영역을 마스크(masking)하여 최종 전처리된 영상을 얻는다. 마스크를 하는 이유는 측광 과정에서 점광원 밀도를 계산할 때 전천 카메라 가장 자리에 남아있는 신호가 없는 영역을 배제하여 올바른 배경하늘값과 잡음을 측정할 수 있기 때문이다. 특히 2023년 11월 6일부터 12월 6일까지의 영상은 비닝이 되어 있지 않았기 때문에, 따로 비닝 처리하여 다른 영상 자료의 차원과 맞추는 작업을 수행하였다.

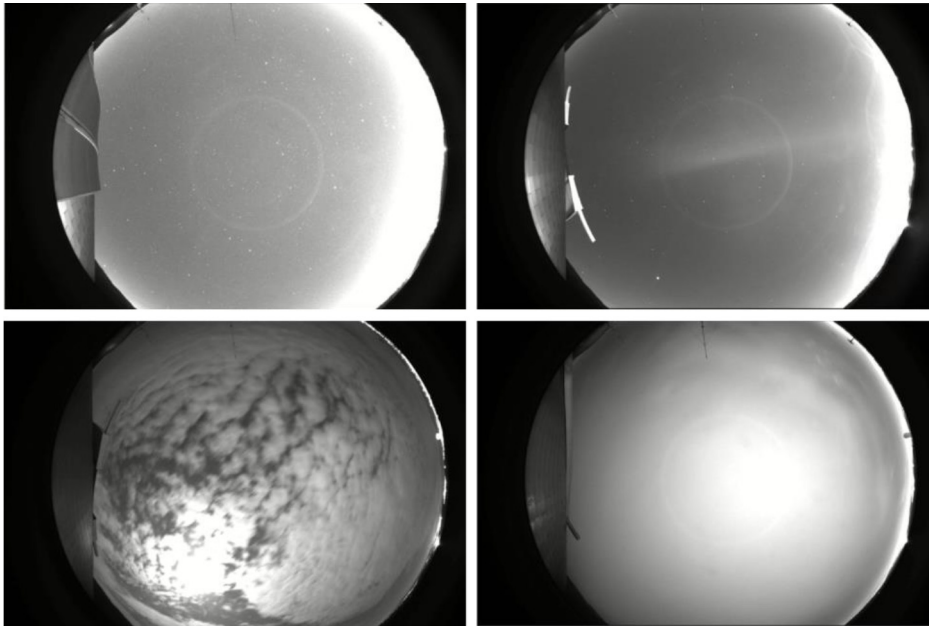


Fig. 2. Examples of the all-sky images visually classified as “Not cloudy” (top) and “Cloudy” (bottom) in the training data preparation. The images classified as “Not cloudy” indicate sky conditions suitable for opening dome-slit, while those classified as “Cloudy” do not.

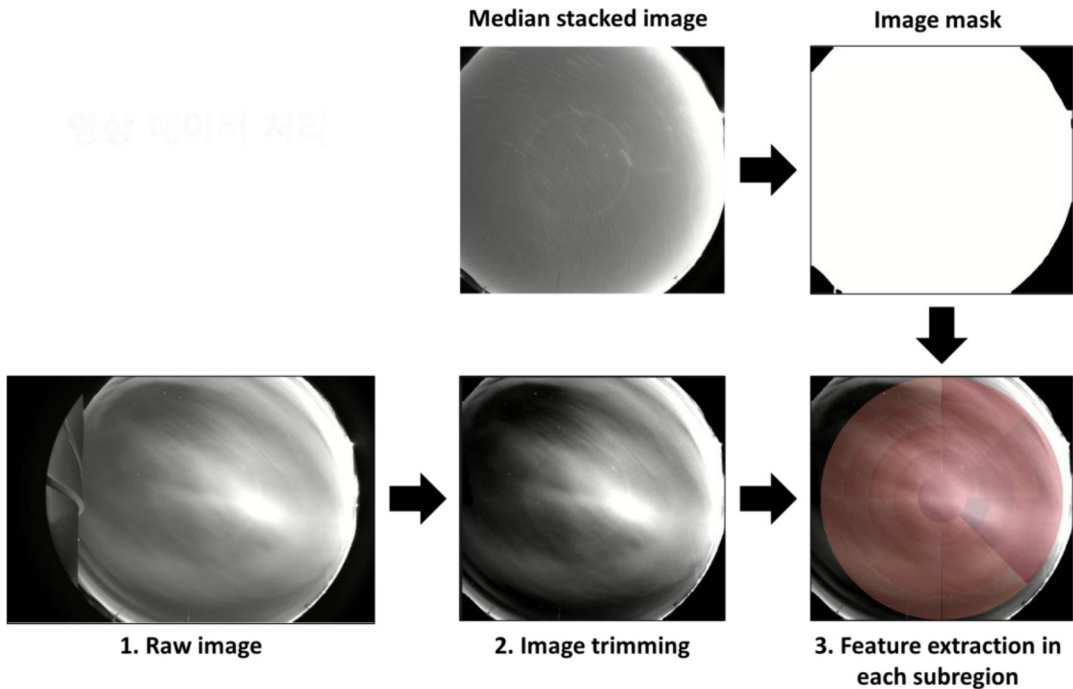


Fig. 3. Data analysis scheme for the MAAO all-sky images for the machine learning model training. The subregions are marked as reddish areas in the image for feature extraction. Here, the intensity of color represents the median value of the sky background in each subregion.

하위 영역 내 구름의 분류

처리된 하나의 영상에 대하여 각 하위영역 내 구름의 유무를 판별하는 작업을 수행하였다. cloudy night 은 전처리된 영상을 업로드하여 각 하위영역 내 구름 유무를 사용자가 분류할 수 있도록 장고(Django) 기반의 웹앱(Webapp)을 자체적으로 제공한다. 이를 이용하여, 39,996개의 하위 영역에 대해 구름의 유무를 판단하여 분류하였다. 전체 하위 영역에 대해 49.3%인 19,738개의 하위 영역이 구름인 영역으로, 나머지 50.7%인 20,258개의 하위 영역이 구름이 아닌 영역으로 분류되었다.

연구 방법

데이터 학습 (Data training)

데이터 학습에 앞서 추출된 영상의 모든 특징 매개변수들은 웹앱에 포함된 PostgreSQL 기반 데이터베이스에 저장된다. 이를 내려받은 후 lightGBM 모델로 맞추기(fitting)하여 기계학습 모델을 생성하였다. 학습에서 사용한 초매개변수(Hyperparameter)는 의사

결정나무 최대 깊이(Maximum depth), 의사결정나무 개수(Number of estimators), 학습률(Learning rate), 의사결정나무 말단노드 개수(Number of leaves), 말단노드 최소샘플 개수(Minimum child samples), L1 정규화 값(α), L2 정규화 값(λ)이다. 이들은 Table 1에 정리되어 있으며, Mommert (2020)에서 제시한 최적의 초매개변수를 사용하였다. 이 연구에서는 Mommert (2020)에서와 마찬가지로 5겹 교차검증(5-fold cross validation)을 통하여 학습 데이터 중 10%인 4,000개를 검증 표본(validation sample)로 구분한 뒤 학습시켰다.

Table 1. Hyperparameters used in this work (Same as Mommert 2020)

Hyperparameter	value
Maximum depth	5
Number of estimators	500
Learning rate	0.25
Number of leaves	30
Minimum child samples	100
α	10
λ	100

연구 결과 및 논의

학습 모델의 성능 평가

데이터 학습의 결과로 생성된 학습 모델의 성능을 평가하면 아래와 같다. 먼저 분류결과표(confusion matrix)를 통하여 실제값과 예측값이 서로 얼마나 일치하는지를 비교할 수 있다. 하위 영역에 대하여 실제 구름이 있는 영역을 구름이 있다고 예측한 경우가 True Positive (TP; 97.2%), 구름이 없다고 예측한 경우가 False Negative (FN; 2.8%)가 되며, 실제 구름이 없는 영역을 구름이 있다고 예측한 경우가 False Positive (FP; 2.8%), 구름이 없다고 예측한 경우가 True Negative (TN; 97.2%)가 된다. 이를 정리한 것이 Table 2이다.

Table 2. The confusion matrices for the data training in this work. The values in the parenthesis are the percentage of samples

	Actual Positive	Actual Negative
Predicted Positive	TP=34,991 (0.972)	FP=1,005 (0.028)
Predicted Negative	FN=995 (0.028)	TN=35,001 (0.972)

Table 2의 분류결과표를 이용하여, 재현율(Recall)과 정밀도(Precision)를 모두 고려한 F1 점수(F1 score)를 구할 수 있는데, 이러한 F1 점수는 기계학습 모델의 성능평가에 널리 이용되고 있다. 각 지표들에 대한 정의는 아래와 같다.

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{F1score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

기계학습 모델을 이용해 검증 표본에 적용한 F1 점수는 0.97로, 학습 모델이 매우 좋은 성능을 보여주고 있음을 알 수 있다. 이는 다른 천문대에서 얻은 전천 영상에서도 cloudynt 모듈이 잘 적용될 수 있음을 보여준다. 또한 전처리 과정에서 돔 구조물이 제거되고, 하위 영역에 관측돔, 나무, 건물이 포함되지 않았으므로 밀양아리랑우주천문대 전천영상의 특

수성이 학습 모델 형성에 있어 큰 영향을 미치지 않음을 보여준다.

운량의 계산과 모니터링 프로그램

자동 관측 시 돔의 개폐 여부를 자동으로 판단하도록 하기 위하여, 운량을 정량화할 필요가 있다. 영상 하나에 대해 기계학습 모델 맞추기로 얻어지는 결과는 각각의 하위영역에 대한 구름의 유무에 대한 디서너리 자료형으로 주어지며, 각 요소는 구름이 존재하는 경우 1로, 그렇지 않은 경우 0이 된다. 본 연구에서 운량(Cloudiness)은 간단하게 아래와 같이 정의하였다.

$$\text{운량(Cloudiness)} = \frac{\text{구름으로 식별된 하위 영역의 개수}}{\text{전체 하위 영역의 개수}}$$

이때, 본 연구에서 사용한 전체 하위 영역 개수는 33개이다. 예를 들어, 16개의 하위 영역이 구름으로 예측되었다면, 운량은 0.48로 계산된다. 이러한 정의를 이용하여 3분마다 촬영된 영상을 읽고, 전처리 후 학습된 모델과 영상을 적합하여 운량을 출력해 주는 모니터링 프로그램을 개발하였다. 영상이 확보되는 시간은 태양 고도가 -6도 이하일 때부터로 설정되어 있기 때문에, 낮에는 운량이 계산되지 않는다. 하위영역 별 구름의 유무와 운량은 JavaScript Object Notation (JSON) 형식의 데이터로 저장되어, 밀양아리랑우주천문대 0.7 m 망원경을 구동하는 컴퓨터에 실시간 공유된다. 돔 슬릿은 30분간 연속으로 운량이 0.6 이하로 측정되면 개방이 가능한 것으로 판단하게 하였다. Fig. 4는 새로 얻은 전천 영상들에 대하여 학습 모델이 어떻게 운량을 계산하였는지를 보여준다. 왼쪽 그림은 운량이 0으로 측정된 영상으로, 구름이 없는 날에는 달빛이 있더라도 운량이 잘 계산되는 것을 볼 수 있다. 오른쪽 그림은 운량이 0.64로 측정된 영상으로, 약간 구름이 적은 영역이 확인되며 자동관측을 정지하는 기준을 만족한다. 이러한 자동관측 운영 기준이 적절한지 그리고 학습된 기계학습 모델이 가지는 한계는 토의에서 다루었다.

토 의

밀양아리랑우주천문대 자동 관측 시스템에서의 운량 판단 기준 평가

이렇게 구축한 기계학습 모델의 운량 판단 정확도

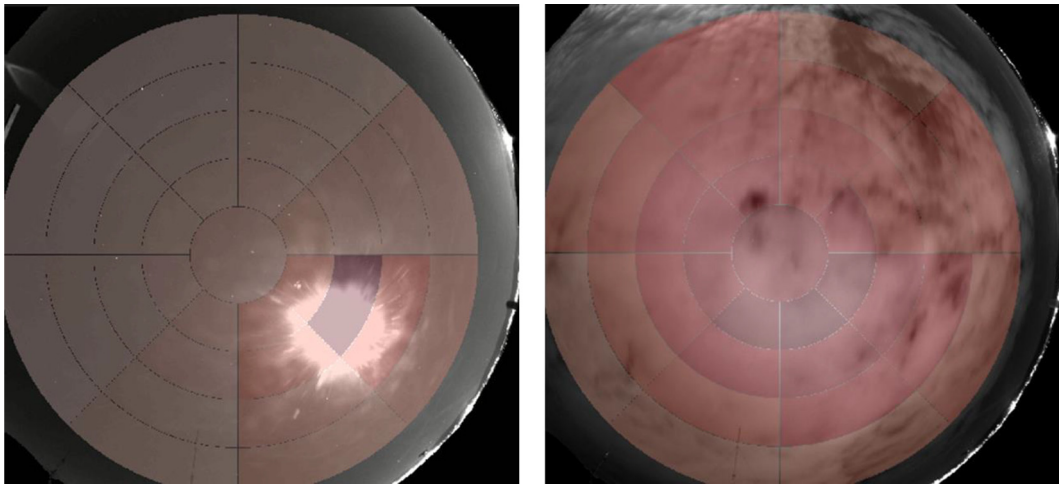


Fig. 4. Examples of reduced all-sky cam images with Cloudiness calculation. Each figure is obtained on (left) 00:00:08, 20th March 2024 with Cloudiness of 0, and (right) on 21:20:44, 26th March 2024 with Cloudiness of 0.64. The intensity of color is also the median value of the sky background.

를 고려하였을 때 밀양아리랑우주천문대의 자동 관측 기준이 얼마나 적절한지 조사하였다. 운량 모니터링 프로그램은 전천 영상을 3분에 1장씩 읽어 들이고, 이를 바탕으로 자동 관측 진행 여부를 결정하는데 다음과 같은 기준들을 가진다. 첫째, 지난 30분간 연속으로 운량이 0.6 이하라고 측정되면 날씨가 맑은 것으로 판단해 돐을 열고 자동 관측을 수행할 수 있다. 둘째, 지난 30분간 한번이라도 운량이 0.6을 초과하면 구름이 낀 것으로 판단해 자동관측을 시작할 수 없으며, 진행되고 있던 자동 관측 역시 강제 중지된다.

날씨 상황을 구름이 있을 때와 구름이 없을 때 크게 둘로 나누어 이러한 자동 관측 기준들이 적절한지 조사해 보았다. 첫째, 실제로 구름이 가득 찬 상황임에도 불구하고 30분간 연속으로 운량이 0.6 이하라고 오판하고 돐을 열 확률을 계산해 보았다. 이를 위해 몬테카를로 시뮬레이션(Monte Carlo simulation)을 100만 회 수행했다. 해당 몬테카를로 시뮬레이션에서 가상의 전천 영상을 33개의 영역으로 나눈 뒤 각 영역의 구름 판별 여부를 학습모델 운량 판단 성능(Table 2)을 참고해 독립적으로 할당하였다. 시뮬레이션을 100만 회 수행하였음에도 불구하고 30분간 연속으로 운량이 0.6 이하라고 판단되어 돐이 열리는 경우는 없었으며, 30분간 연속으로 운량이 0.95 이하로 판단하는 확률도 0.0004%밖에 되지 않았다. 해당

시뮬레이션은 비가 오는 경우에 돐이 열리는 상황을 막기 위해 전천을 구름으로 모두 뒤덮은 경우를 상정했기에, 비가 오지 않으나 부분적으로 구름으로 뒤덮인 하늘에서 돐이 열리는 확률은 더 높을 것으로 예상된다. 둘째, 실제로 완전히 맑은 하늘임에도 불구하고 30분간 한번이라도 운량이 0.6 이상인 것으로 오판하고 돐을 닫을 확률을 계산해 보았다. 역시 동일한 몬테카를로 시뮬레이션을 수행해 해당 확률을 계산하였으며, 시뮬레이션을 100만 회 수행하여도 오판으로 인하여 돐을 닫을 경우는 확인할 수 없었다. 그뿐만 아니라, 30분간 한번이라도 운량이 0.2 이상으로 측정될 확률도 약 0.03%밖에 되지 않았다. 다만, 이 결과는 하늘이 완전 맑을 때를 상정하고 얻은 결과이기에 구름이 부분적으로 끼어있는 날씨 아래 운량 오판으로 인한 관측 중지 확률은 더 높다. 그러나, 자동관측 활동은 주로 날씨가 완전히 맑은 날 이루어질 것이므로 이러한 운량 오판 가능성은 자동 관측 운영에 영향을 크게 미치지 않을 것이다. 결론적으로 이 연구의 기계 학습 모델 성능이 Table 2와 같다면, 본 연구의 운량 측정 프로그램을 이용하여 원격 및 자동 관측을 성공적으로 수행할 수 있으리라 생각된다.

기계학습 모델을 이용한 운량 판단의 한계

본 연구에서 기계학습 모델을 만들기 위하여 약

1,200장의 전천 영상을 학습 데이터로 사용하였다. 이러한 1,200장의 전천 영상은 굉장히 다양한 상황의 전천 영상들을 포함하고 있지만, 특수한 몇몇 경우의 전천 영상에서는 사용자의 주관에 따라 운량을 잘못 식별한 결과가 기계학습 모델에 반영되었을 가능성이 있다. 첫째, 일몰 직후와 일출 직전의 전천 영상에서는 태양과 가까운 방향의 구름 유무를 제대로 분류하기 어려웠다. 일몰 직후 또는 일출 직전이라 할지라도 여전히 하늘이 꽤 밝은 상태이므로, 태양과 가까운 방향의 배경 하늘값이 매우 높다. 따라서 해당 지역에서는 사용자가 구름을 판별하는데 있어 어려움을 겪었고, 이러한 결과가 기계학습 모델에 반영되었을 가능성이 있다. 둘째, 달의 위상이 증가할수록, 달의 위치와 겹치는 하위 영역의 배경 값이 매우 높기 때문에 구름의 존재를 식별하기 쉽지 않다. 본 연구에서는 보수적인 기준을 적용하여, 구름이 부분적으로 있으면 달 근처에는 구름이 있는 것으로 분류했기에, 이에 대한 편향이 기계학습 모델에 반영되어 있을 가능성이 있다.

그뿐만 아니라, 학습에 사용한 전천 영상들은 매우 특이한 상황의 전천영상들까지 모두 담고 있지는 못하다. 특히 전천 카메라 위에 이물질이 올라간 경우, 생성된 기계학습 모델은 이러한 상황에서의 운량을 적절히 판단하지 못하는 경향이 있다. Fig. 5와 같이, 작은 나뭇가지로 추정되는 물질이 전천 카메라를 가리게 되면, 완전히 구름 낀 날씨임에도 불구하고 생성된 기계학습 모델은 모든 영역에서 구름이 끼지 않은 날씨로 오판하였다. 구름 낀 날씨를 구름이 끼지 않은 날씨로 오판하는 것은, 비가 온다든지 돛을 반드시 닫아야 하는 상황에서 돛을 열어버릴 가능성이 있기에, 밀양아리랑우주천문대의 자동 관측 운영에 있어 매우 위험한 오판 가능성이라 할 수 있다.

하지만, 전천 카메라 위에 이물질이 올라간 상태의 전천 영상을 학습 데이터에 포함하여 기계학습 모델을 생성할 수는 없었다. 왜냐하면, 큰 이물질이 전천 카메라에 올라간 경우에는 전천 이미지를 통해 구름의 정보를 얻을 수 없기 때문이다. 그뿐만 아니라 전천 카메라 위에 올라간 이물질의 크기, 모양, 위치, 개수 등이 모두 다 다를 수 있기에, 이물질이 올라간 모든 경우를 포함해 운량 판단 프로그램을 구축할 수 없었기 때문이다. 따라서, 우리는 두 가지의 해결 방안을 모색해 볼 수 있다. 첫째, 자동 관측 직전 전천 카메라에 이물질이 떨어졌는지를 일일이 확인한다.



Fig. 5. An all-sky image with small materials, though to be branches, on its camera. Even though it was a cloudy night, the all-sky image is misclassified as “Not Cloudy” via our machine learning model.

이를 위해서는 매일 관측 직전 사람이 직접 카메라에 이물질이 떨어졌는지를 전천 이미지를 통해 판단하고, 관측을 수행해야 한다. 가장 확실한 방법이긴 하지만 자동 관측의 의미가 크게 퇴색될 수 있다는 단점이 있다. 둘째, 전천 카메라에 이물질이 떨어졌는지를 판단하는 기계 학습 모델을 새롭게 개발해, 카메라에 이물질이 떨어졌다면 자동 관측을 정지시킨다. 이를 위해서는 이물질의 크기, 모양, 위치, 개수 등이 다 다른 다양한 상황을 감안한 추가 전천 이미지들을 충분히 얻고, 이들을 이용한 새로운 기계 학습 모델을 추가로 개발하는 것이 필요하다.

요약 및 결론

밀양아리랑우주천문대 0.7 m 망원경의 자동 관측 (Lim et al. in prep.)을 원활하게 수행하기 위하여, 주기적으로 운량을 계산하는 모니터링 프로그램을 개발하였다. 이 프로그램은 기계학습 기반의 파이썬 모듈인 “cloudynight” (Mommert 2020)을 밀양아리랑우주천문대의 전천 영상에 적용해 학습 모델을 구축하고 운량(Cloudiness)을 계산한다. 전처리된 전천 영상을 학습 데이터로 사용했으며, 전천 영상을 고도와 방위각에 따라 33개의 하위 영역으로 분할한 뒤 학습하였다. 전체 39,996개의 하위영역 데이터 가운데 10%의 하위 영역을 검증 표본으로 사용해 5겹 교차검증을 수행한 결과, 생성된 기계 학습 모델의 F1 점수는 0.97이었다. 이러한 결과는 밀양아리랑우주천문대의 전천 영상에서도 학습 모델이 좋은 성능을 보

여준다는 것을 시사하며, 다른 천문대의 전천 영상에서도 이를 적용할 수 있다는 가능성을 보여준다. 밀양아리랑우주천문대의 자동관측 시스템은 3분에 한 번씩 전천 영상을 읽어 들이고, 학습 모델을 이용해 전천 영상의 운량을 측정한다. 이렇게 측정된 운량이 30분간 연속으로 0.6 이하면 돔이 열리고 자동 관측이 시작되며, 지난 30분간 한번이라도 운량이 0.6 초과라면 돔이 닫히고 자동관측이 정지된다. 본 연구의 학습 모델 성능을 고려해 보았을 때 자동 관측 수행 시 운량 오관으로 인해 돔이 잘못 열리거나 닫힐 확률은 거의 존재하지 않았다. 다만, 이러한 학습 모델을 구축하는 데 이용된 학습 데이터의 한계로 인해 태양과 달빛 또는 전천카메라의 이물질이 운량 측정에 있어 영향을 미칠 가능성을 가지고 있다. 학습 모델은 차후 특수한 상황의 영상을 포함하여 재학습을 진행하는 등 추가적인 보완 작업을 통하여 이러한 한계들을 극복할 수 있으리라 기대한다. 그리고 본 연구의 운량 모니터링 프로그램 활용을 통하여 국내의 많은 소형 망원경에 자동 관측 시스템이 구축되고 더욱 다양한 시계열 관측 연구가 활발히 이루어 지기를 바란다.

감사의 글

이 과제는 부산대학교 기본연구지원사업(2년)에 의하여 연구되었음.

References

- Abbott, B. P., R. Abbott, T. D. Abbott, et al. 2017, *The Astrophysical Journal*, 848, L12.
- Alsubai, K., Z. I. Tsvetanov, D. W. Latham, A. Bieryla, G. A. Esquerdo, D. Mislis, S. Pyrzas, E. Foxell, J. McCormac, C. Baranec, N. P. E. Vilchez, R. West, A. Esamdin, Z. Dang, H. M. Dalee, A. A. Al-Rajih, and A. K. Al-Harbi, 2018, *The Astronomical Journal*, 155, 52.
- Cai, Y.-Z., A. Pastorello, M. Fraser, X.-F. Wang, A. V. Filippenko, A. Reguitti, K. C. Patra, V. P. Goranskij, E. A. Barsukova, T. G. Brink, N. Elias-Rosa, H. F. Stevance, W. Zheng, Y. Yang, K. E. Atapin, S. Benetti, T. J. L. de Boer, S. Bose, J. Burke, R. Byrne, E. Cappellaro, K. C. Chambers, W.-L. Chen, N. Emami, H. Gao, D. Hiramatsu, D. A. Howell, M. E. Huber, E. Kankare, P. L. Kelly, R. Kotak, T. Kravtsov, V. Y. Lander, Z.-T. Li, C.-C. Lin, P. Lundqvist, E. A. Magnier, E. A. Malygin, N. A. Maslennikova, K. Matilainen, P. A. Mazzali, C. McCully, J. Mo, S. Moran, M. Newsome, D. V. Oparin, E. Padilla Gonzalez, T. M. Reynolds, N. I. Shatsky, S. J. Smartt, K. W. Smith, M. D. Stritzinger, A. M. Tatarnikov, G. Terreran, R. I. Uklein, G. Valerin, P. J. Vallely, O. V. Vozyakova, R. Wainscoat, S.-Y. Yan, J.-J. Zhang, T.-M. Zhang, S. G. Zheltoukhov, R. Dastidar, M. Fulton, L. Galbany, A. Gangopadhyay, H.-W. Ge, C. P. Gutiérrez, H. Lin, K. Misra, Z.-W. Ou, I. Salmaso, L. Tartaglia, L. Xiao, and X.-H. Zhang, 2022, *Astronomy and Astrophysics*, 667, A4.
- Deckers, M., Maguire, K., Magee, M. R., Dimitriadis, G., Smith, M., Sainz de Murieta, A., Miller, A. A., Goobar, A., Nordin, J., Rigault, M., Bellm, E. C., Coughlin, M., Laher, R. R., Shupe, D. L., Graham, M., Kasliwal, M., & Walters, R., 2022, *Monthly Notices of the Royal Astronomical Society*, 512, 1317-1340.
- de Jaeger, T., W. Zheng, B. E. Stahl, A. V. Filippenko, T. G. Brink, A. Bigley, K. Blanchard, P. K. Blanchard, J. Bradley, S. K. Cargill, C. Casper, S. B. Cenko, S. Channa, B. Y. Choi, K. I. Clubb, B. E. Cobb, D. Cohen, M. de Kouchkovsky, M. Ellison, E. Falcon, O. D. Fox, K. Fuller, M. Ganeshalingam, C. Gould, M. L. Graham, G. Halevi, K. T. Hayakawa, J. Hestenes, M. P. Hyland, B. Jeffers, N. Joubert, M. T. Kandrashoff, P. L. Kelly, H. Kim, M. Kim, S. Kumar, E. J. Leonard, G. Z. Li, T. B. Lowe, P. Lu, M. Mason, K. J. McAllister, J. C. Mauerhan, M. Modjaz, J. Molloy, D. A. Perley, K. Pina, D. Poznanski, T. W. Ross, I. Shivvers, J. M. Silverman, C. Soler, S. Stegman, S. Taylor, K. Tang, A. Wilkins, X. Wang, X. Wang, H. Yuk, S. Yunus, and K. D. Zhang, 2019, *Monthly Notices of the Royal Astronomical Society*, 490, 2799-2821.
- Djorgovski, S. G., Drake, A. J., Mahabal, A. A., Graham, M. J., Donalek, C., Williams, R., Beshore, E. C., Larson, S. M., Prieto, J., Catelan, M., Christensen, E., & McNaught, R. H., 2011, arXiv e-prints, arXiv: 1102.5004.
- Im, M., C. Choi, and K. Kim, 2015, *Journal of Korean Astronomical Society*, 48, 207-212.
- Im, M., Y. Kim, C.-U. Lee, H.-W. Lee, S. Pak, H. Shim, H.-I. Sung, W. Kang, T. Kim, J.-E. Heo, T. C. Hinse, M. Ishiguro, G. Lim, C. T. K. Ly, G. S. H. Paek, J. Seo, J.- na . Yoon, J.-H. Woo, H. Ahn, H. Cho, C. Choi, J. Han, S. Hwang, T.-G. Ji, S.-K. J. Lee, S. Lee, S. Lee, C. Kim, D. Kim, J. Kim, S. Kim, M. Jeong, B. Park, I. Paek, D. Kim, and C. Park, 2021, *Journal of Korean Astronomical Society*, 54, 89-102.
- Itoh, R., Y. Fukazawa, Y. T. Tanaka, Y. Abe, H. Akitaya, A. Arai, M. Hayashi, T. Hori, M. Isogai, H. Izumiura, K. S. Kawabata, N. Kawai, D. Kuroda, R. Miyanoshta, Y. Moritani, T. Morokuma, T. Nagayama, J. Nakamoto,

- C. Nakata, Y. Oasa, T. Ohshima, T. Ohsugi, S.-. ichiro . Okumura, Y. Saito, Y. Saito, M. Sasada, K. Sekiguchi, Y. Takagi, J. Takahashi, Y. Takahashi, K. Takaki, M. Uemura, I. Ueno, S. Urakawa, M. Watanabe, M. Yamanaka, Y. Yonekura, and M. Yoshida, 2013, *The Astrophysical Journal*, 768, L24.
- Ke, G., Meng, Q., Finley, T., Wang T., Chen W., Ma W., Ye Q., Liu T.-Y., 2017, in *Advances in Neural Information Processing Systems* 30, ed. I. Guyon et al. (San Diego, CA: NeurIPS), 3149.
- Kim, J., M. Im, C. Choi, and S. Hwang, 2019, *The Astrophysical Journal*, 884, 103.
- Lim, G., M. Im, G. S. H. Paek, S.-C. Yoon, C. Choi, S. Kim, J. C. Wheeler, B. P. Thomas, J. Vinkó, D. Kim, J. Seo, W. Kang, T. Kim, H.-I. Sung, Y. Kim, J.-N. Yoon, H. Kim, J. Kim, H. Bae, S. Ehgamberdiev, O. Burhonov, and D. Mirzaqulov, 2023, *The Astrophysical Journal*, 949, 33.
- Mommert, M., 2020, *The Astronomical Journal*, 159, 178.
- Shappee, B. J., J. L. Prieto, D. Grupe, C. S. Kochanek, K. Z. Stanek, G. De Rosa, S. Mathur, Y. Zu, B. M. Peterson, R. W. Pogge, S. Komossa, M. Im, J. Jencson, T. W.-S. Holoiien, U. Basu, J. F. Beacom, D. M. Szczygieł, J. Brimacombe, S. Adams, A. Campillay, C. Choi, C. Contreras, M. Dietrich, M. Dubberley, M. Elphick, S. Foale, M. Giustini, C. Gonzalez, E. Hawkins, D. A. Howell, E. Y. Hsiao, M. Koss, K. M. Leighly, N. Morrell, D. Mudd, D. Mullins, J. M. Nugent, J. Parrent, M. M. Phillips, G. Pojmanski, W. Rosing, R. Ross, D. Sand, D. M. Terndrup, S. Valenti, Z. Walker, and Y. Yoon, 2014, *The Astrophysical Journal*, 788, 48.
- Paek, G. S. H., M. Im, J. Kim, G. Lim, B. Park, C. Choi, S. Kim, C. Barbieri, O. S. Salafia, I. Paek, S. Shin, J. Seo, H. M. Lee, C.-U. Lee, S.-L. Kim, and H.-I. Sung, 2024, *The Astrophysical Journal*, 960, 113.
- Walsh, J. L., Minezaki, T., Bentz, M. C., Barth, A. J., Baliber, N., Li, W., Stern, D., Bennert, V. N., Brown, T. M., Canalizo, G., Filippenko, A. V., Gates, E. L., Greene, J. E., Malkan, M. A., Sakata, Y., Street, R. A., Treu, T., Woo, J.-H., & Yoshii, Y., 2009, *The Astrophysical Journal Supplement Series*, 185, 156-170.
- Wang, Y.-H., S. Wang, H.-G. Liu, T. C. Hinse, G. Laughlin, D.-H. Wu, X. Zhang, X. Zhou, Z. Wu, J.-L. Zhou, R. A. Wittenmyer, J. Eastman, H. Zhang, Y. Hori, N. Narita, Y. Chen, J. Ma, X. Peng, T.-M. Zhang, H. Zou, J.-D. Nie, and Z.-M. Zhou, 2017, *The Astronomical Journal*, 154, 49.

Manuscript received: April 11, 2024

Revised manuscript received: April 15, 2024

Manuscript accepted: April 15, 2024