

EfficientNet-B1 기반 오로라 이미지의 분류 및 시각화 분석

2148052 여정인

1. 연구 배경	I	4. 딥러닝 접근 방법	4
<ul style="list-style-type: none">- 오로라 직접 관측 경험- 극지연구소 인턴십- 우주-대기-딥러닝 접점 탐구		<ul style="list-style-type: none">- 다양한 딥러닝 모델 시도- EfficientNet-B1 선정 이유- 딥러닝 전처리 과정	
2. 연구 동기 및 목표	2	5. 결과 및 분석	5
<ul style="list-style-type: none">- 오로라 특성의 정량화- 태양풍 변수와의 연결- 최소한의 지도학습		<ul style="list-style-type: none">- 분류 정확도 분석- GradCAM을 통한 시각화- 시간대별 성능 평가	
3. 데이터 선정	3	6. 결론 및 향후 계획	6
<ul style="list-style-type: none">- AuroraMAX 데이터셋- 이미지 전처리 과정- 데이터 균형 확보 전략		<ul style="list-style-type: none">- 연구의 주요 발견- 한계점 및 개선 방향- 향후 연구 계획 제시	

연구 배경



오로라 직접 관측 경험

- 캐나다 옐로나이프에서 2회 관측
- 실제 오로라 현상의 복잡성 인식
(규칙을 찾고 싶어짐)



극지연구소 인턴십

- 1주일 간의 딥러닝 전처리 실무 경험...
- 오로라 데이터의 복잡한 특성 이해
(해결하고 싶어짐)

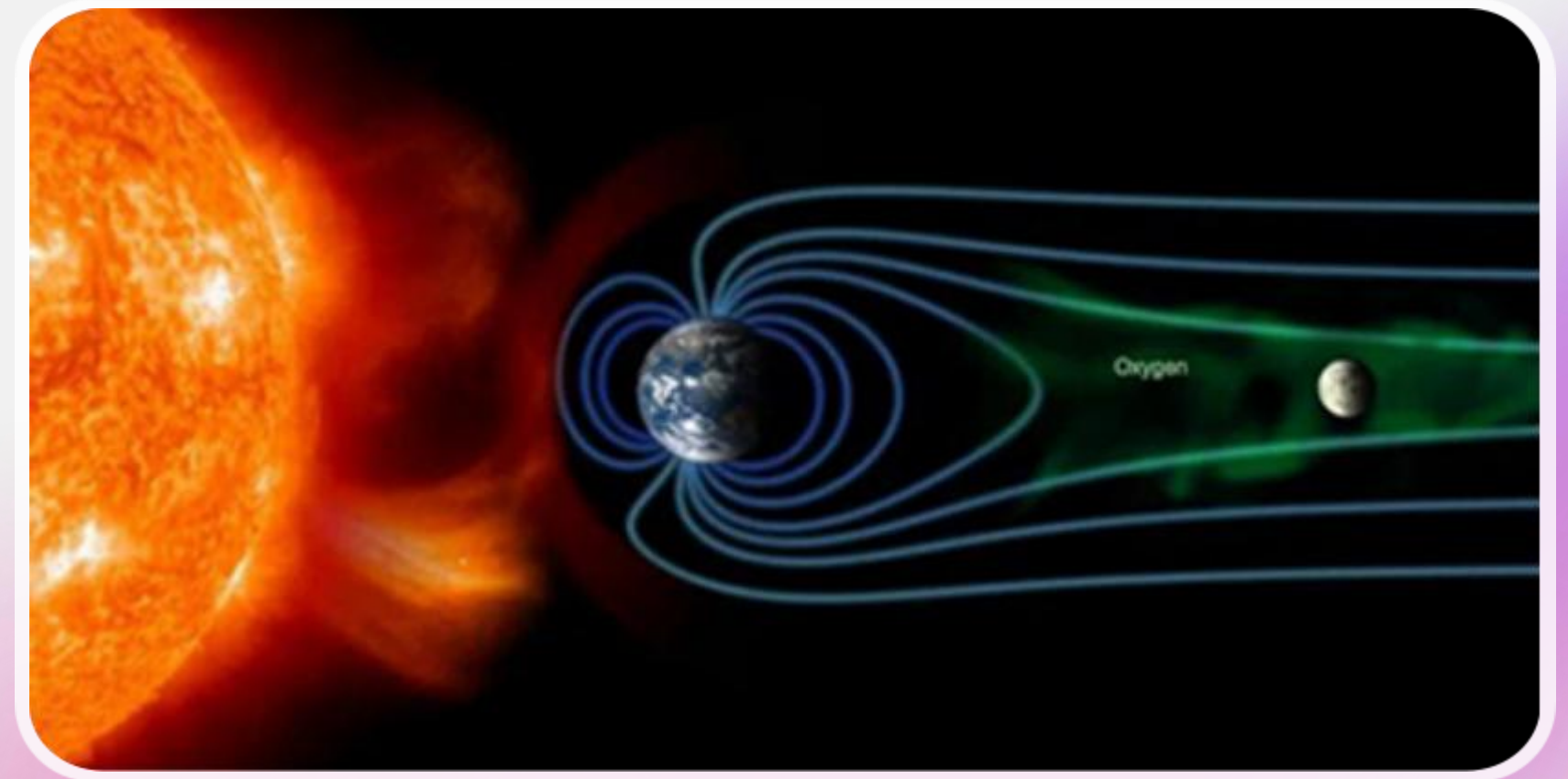


우주-대기-딥러닝 접점

- 우주 기상과 지구 대기 현상 연계
- 딥러닝의 이미지 분석 능력 활용

연구 동기 및 목표

오로라 현상은 태양-지구 상호작용의 가시적 증거로, 우주 기상 연구에 중요한 역할을 한다. 본 연구는 딥러닝을 활용하여 오로라 이미지를 자동으로 분석하고, 이를 통해 태양풍 변수와의 관계를 규명하고자 했었다.



I

최종 목표

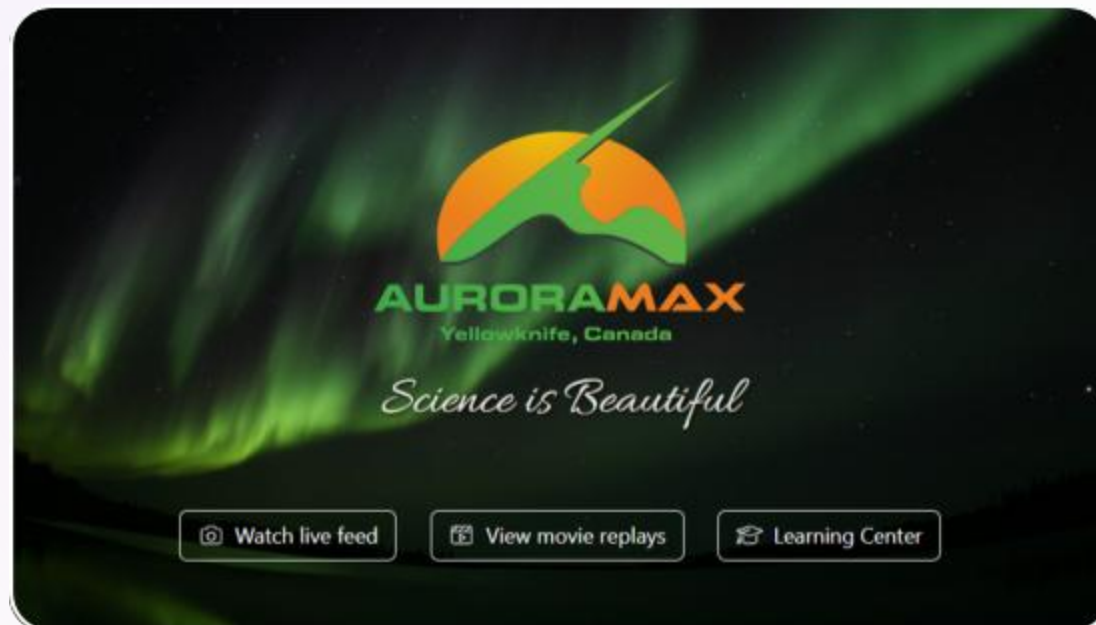
- 오로라의 형태, 밝기, 곡률 정량화
- 태양풍 변수(B_z , K_p , V_{sw} 등)와 연결
 - 우주 기상 예측 모델 개선에 기여
- 자동화된 오로라 분석 시스템 구축

2

현재 단계

- 오로라 존재 여부 이진 분류 (Yes/No)
- EfficientNet-B1 모델 적용 및 최적화
 - GradCAM을 통한 모델 해석
- 오로라 분할 방법 찾는 중

데이터 선정 이유



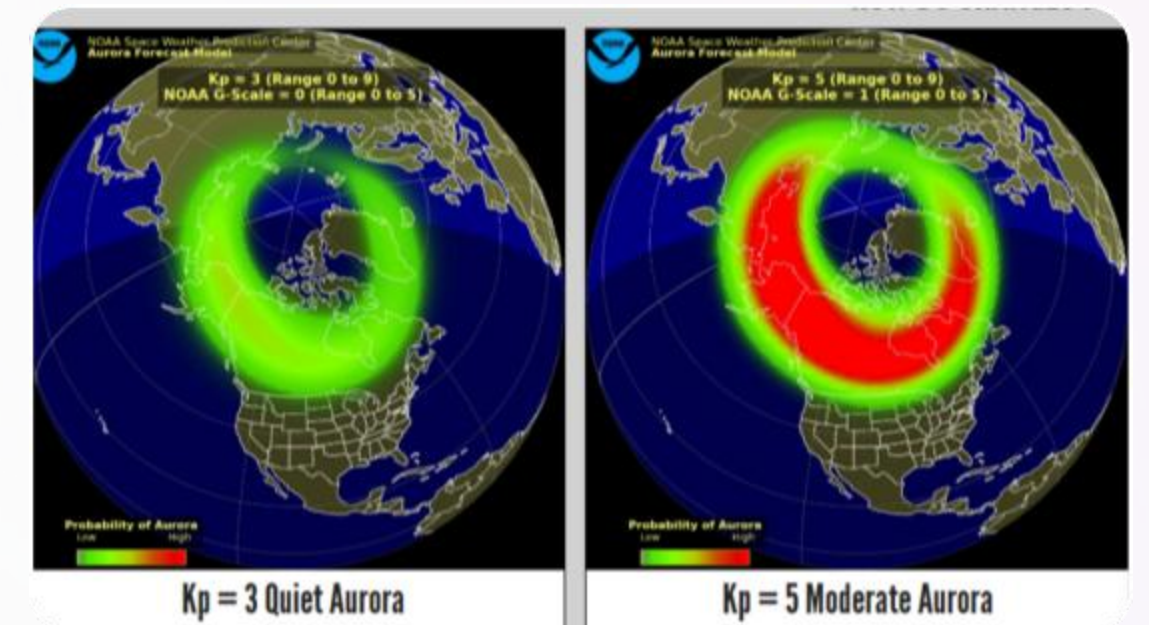
AuroraMAX 프로젝트

- 캐나다 우주국 제공 공식 데이터
- 고품질 전천 카메라 영상 제공
- 장기간 연속 관측 데이터 확보
- 신뢰성 있는 연구 기반 마련



데이터 접근 유연성

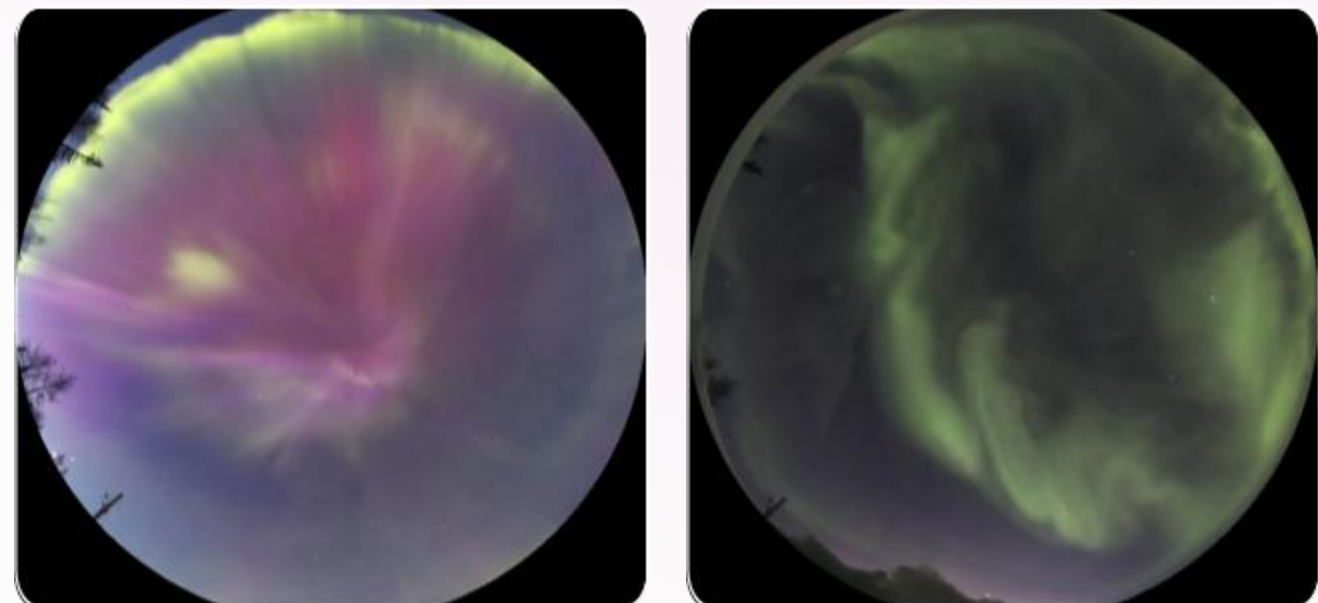
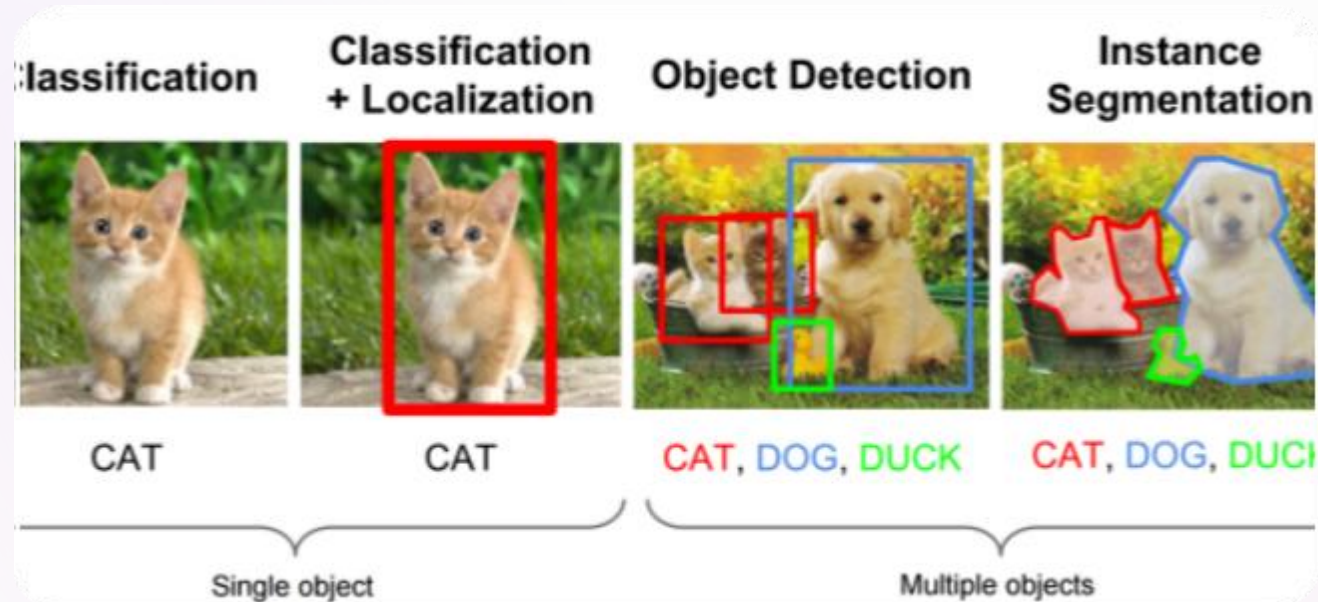
- 대중과학 목적으로 홈페이지 제공
- 원하는 인터벌로 프레임 추출 가능
- 광해나 지형지물 적은 편



데이터 균형 확보

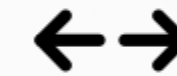
- 옐로나이프는 오로라 오벌 내 위치
- Yes/No 샘플 균형 확보 용이
- 모델 학습의 편향 방지

오로라 딥러닝의 어려움 (1)



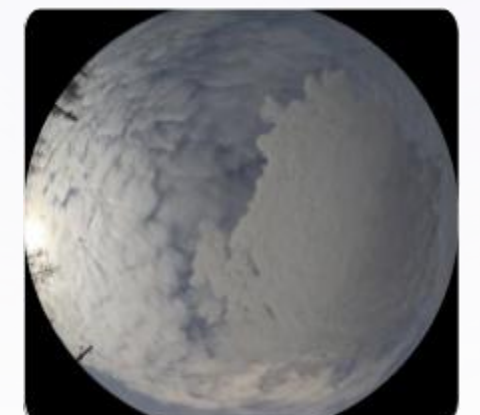
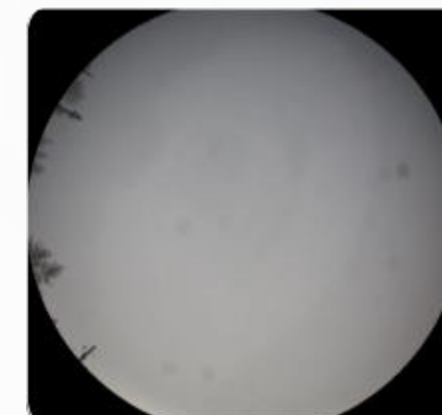
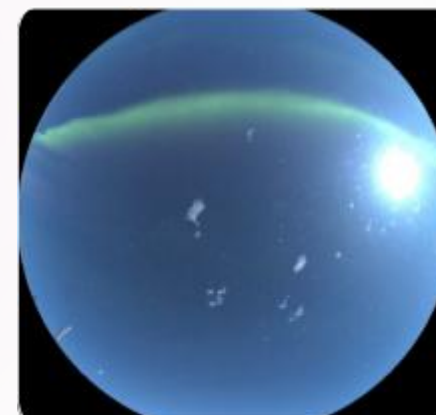
일반 이미지

- 모양이 일정함
 - 윤곽이 뚜렷함
 - 불투명함
- ImageNET 학습 데이터 풍부
- 특징 추출이 용이함
 - 분류 정확도가 높음

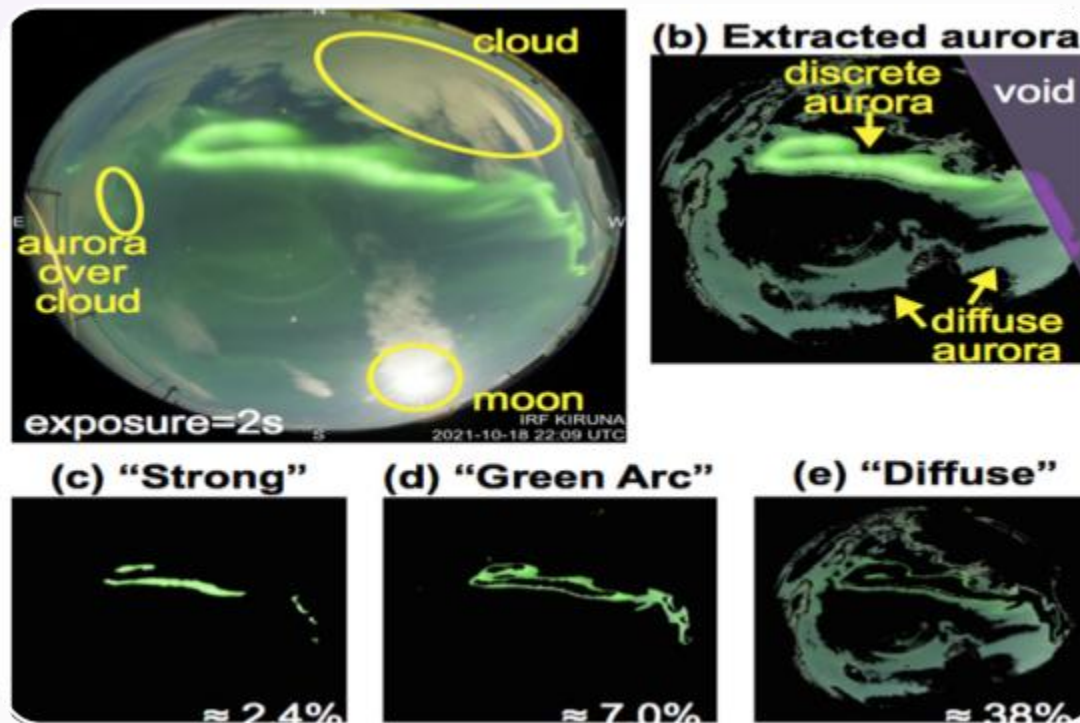


오로라 이미지

- 모양이 불규칙적
 - 경계가 흐릿하고 투명함
 - 배경이 매우 다양 (구름, 달, 눈)
- ImageNET 정답셋 부족
- 특징 추출이 어려움
 - 분류 정확도 향상에 한계



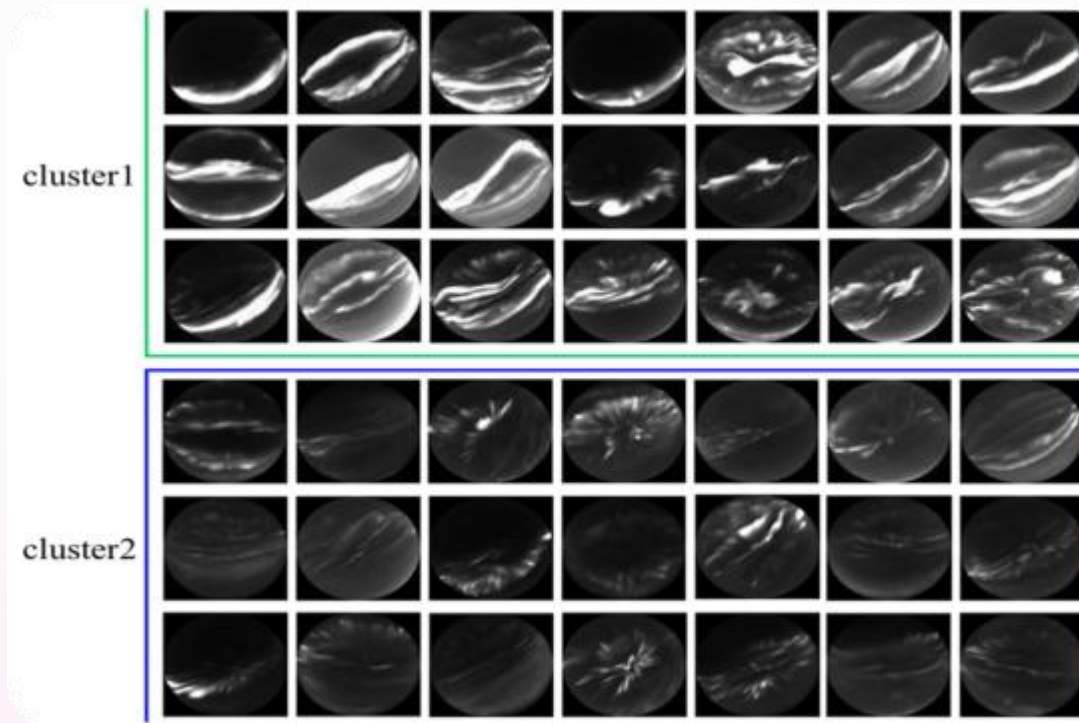
오로라 딥러닝의 어려움 (2)



Yamauchi, M., & Brändström, U. (2023). Auroral alert version 1.0: two-step automatic detection of sudden aurora intensification from all-sky JPEG images. Geoscientific Instrumentation, Methods and Data Systems, 12(1), 71-90.

비지도 RGB 임계값 분할

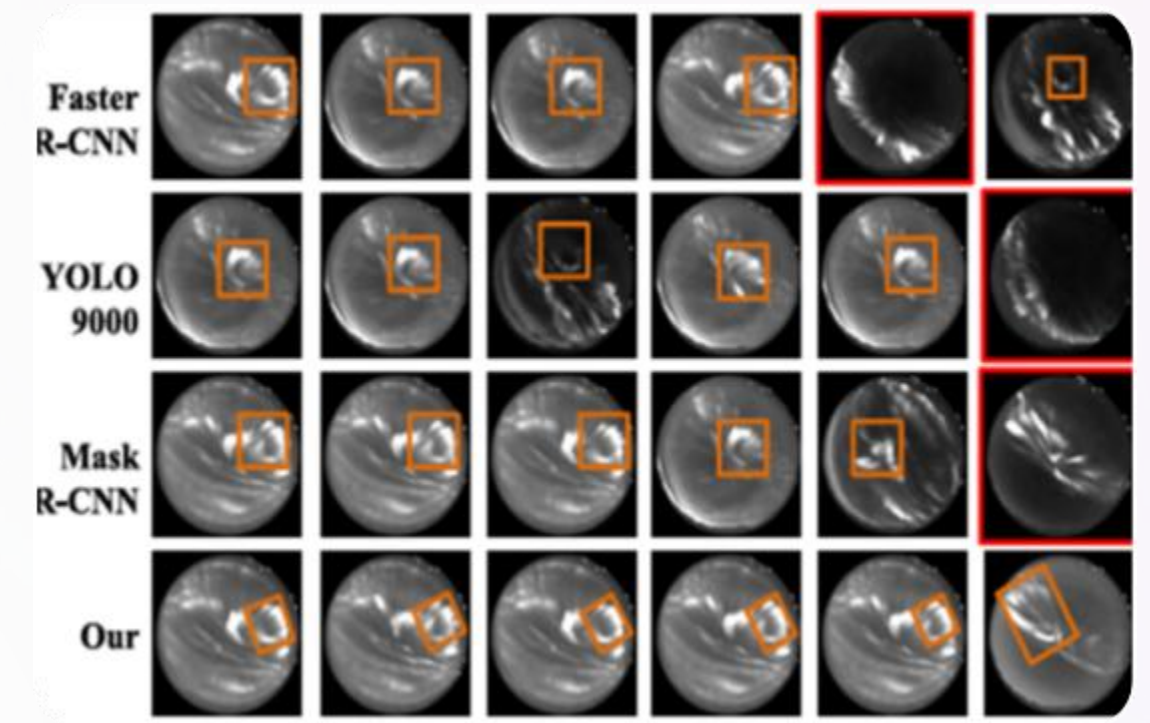
- 알고리즘 단순화, 사전 학습 불필요
- 투명하여 배경하늘 영향 큼



Yang, Q., Liu, C., & Liang, J. (2021). Unsupervised automatic classification of all-sky auroral images using deep clustering technology. Earth Science Informatics, 14(3), 1327-1337.

비지도 VGG 특징 추출 및 군집화

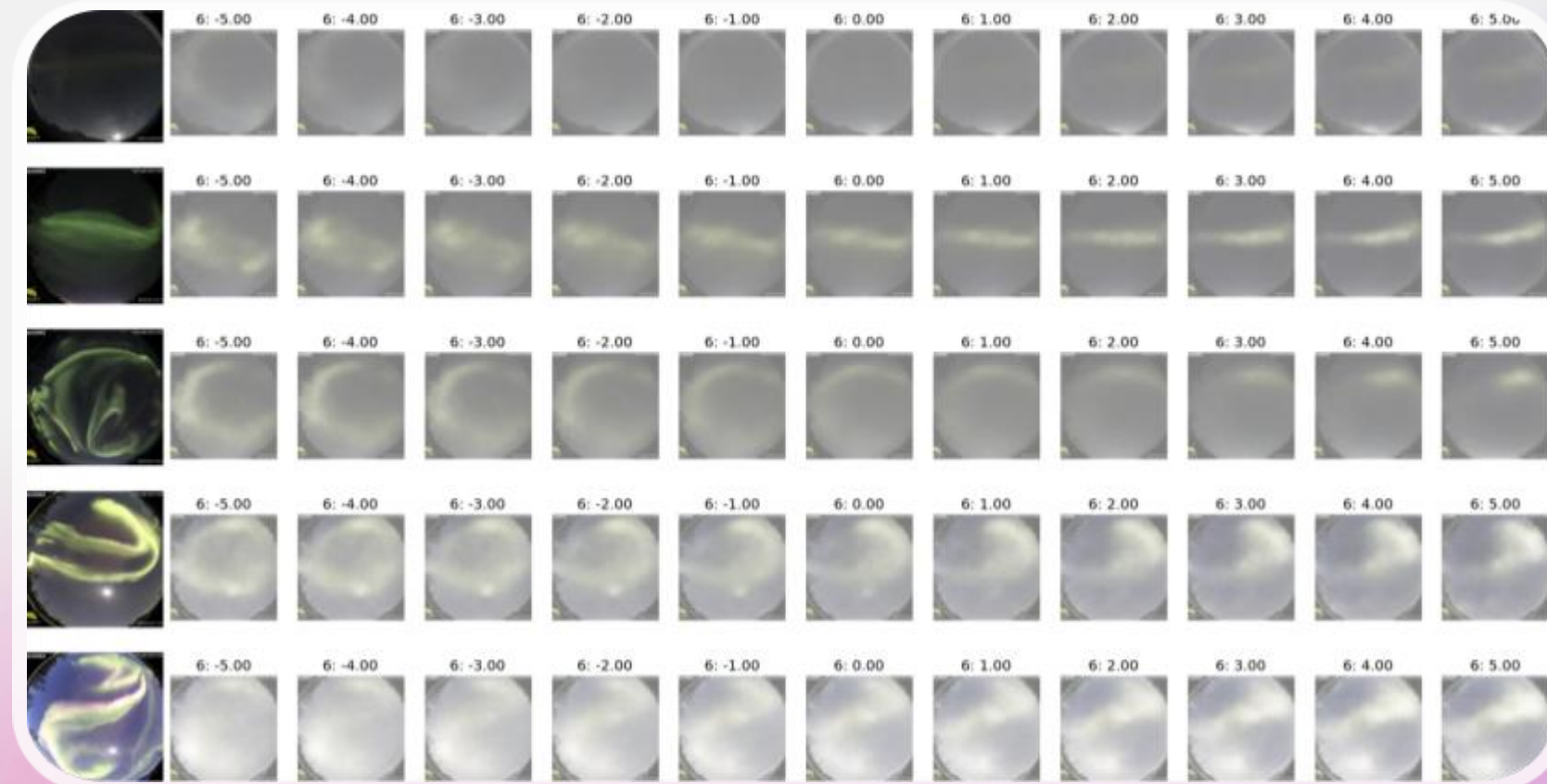
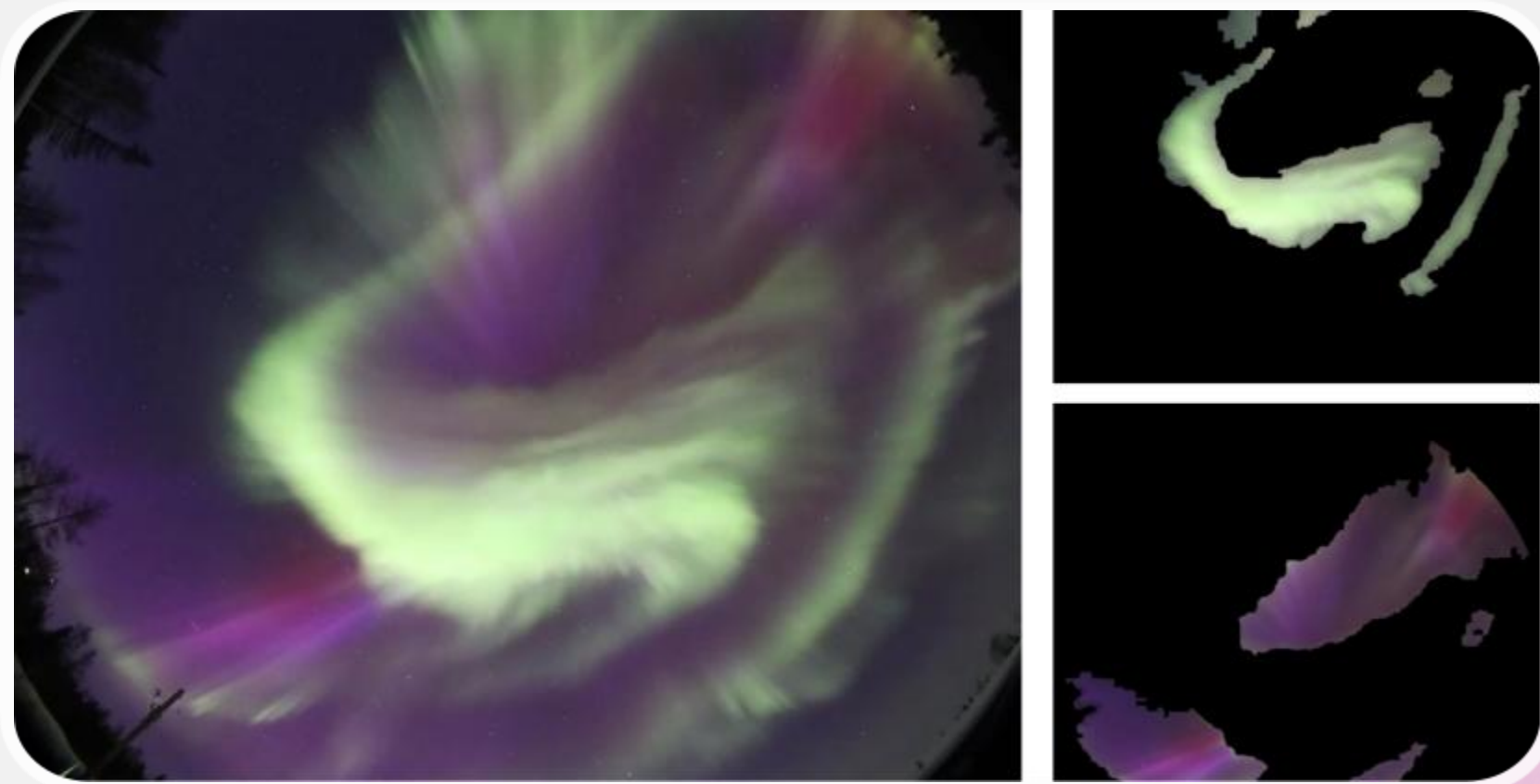
- 수동 라벨링 불필요
- 블랙박스라 클러스터 의미 해석 주관



Yang, X., Wang, N., Song, B., & Gao, X. (2019). Aurora image search with saliency deep features. IEEE Access, 7, 65996-66006.

R-CNN 특징 추출 및 관심 영역 추출

- 100만장 수동 라벨링 필요
- 시각적 해석은 일치함



I

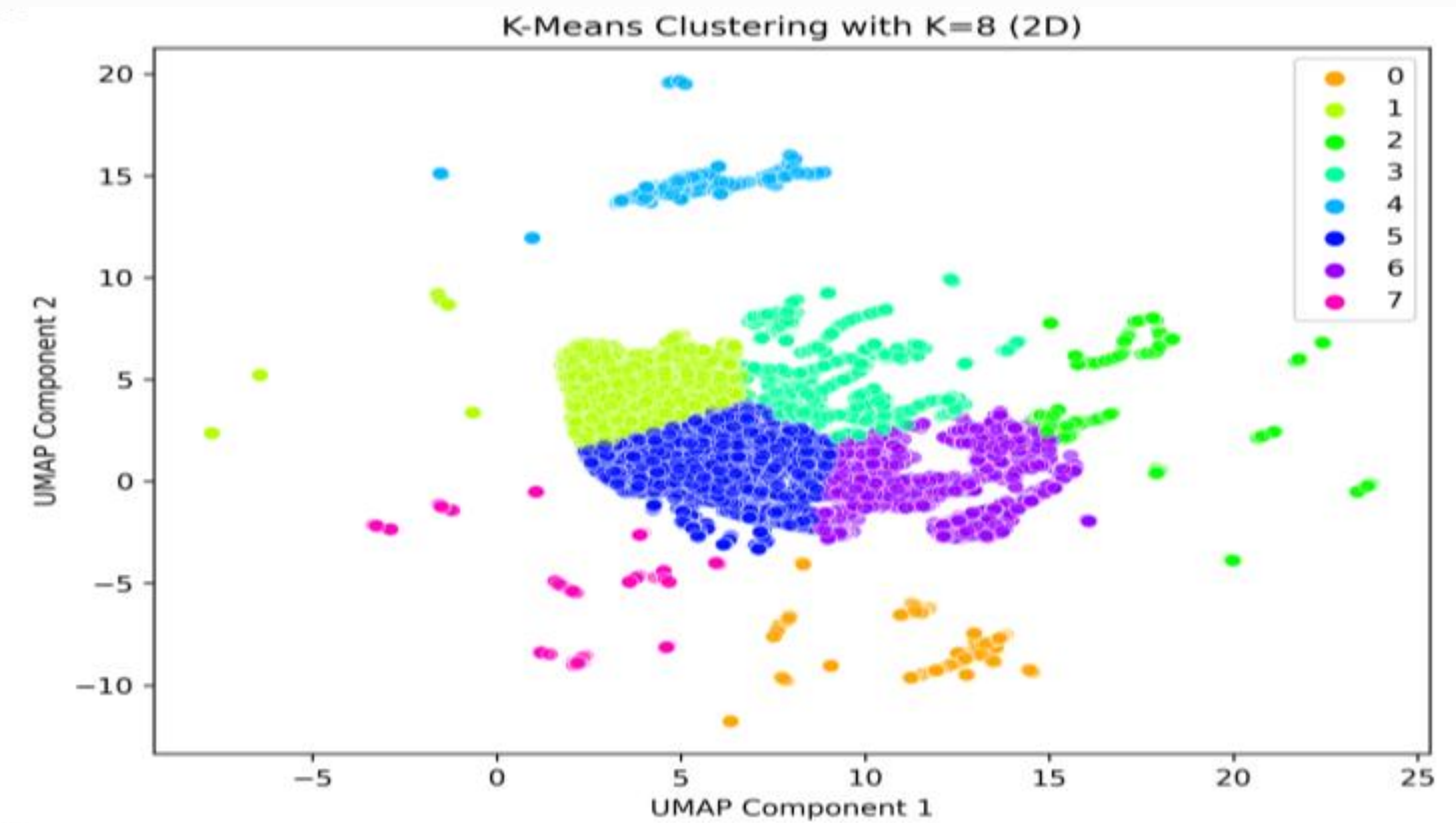
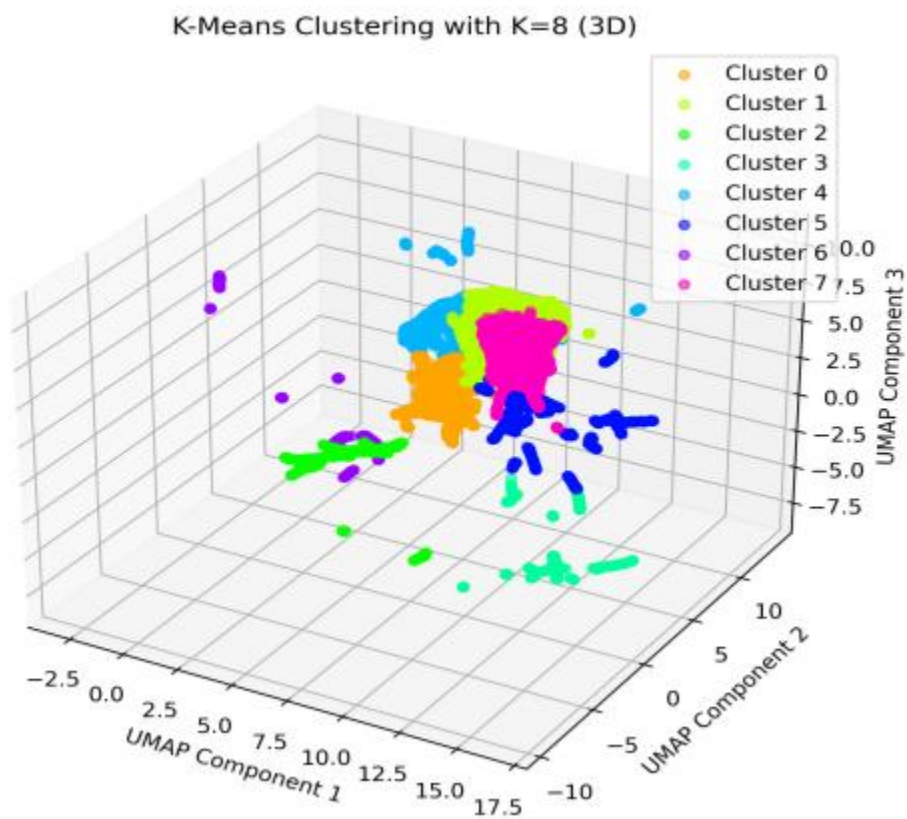
RGB/HSV 마스크

- RGB 및 HSV, YgCrCb 색공간 실험
초록: $(g==v) \ \& \ (s>=25) \ \& \ (v-yg>1)$
핑크: $((h<=10)|(h>=120)) \ \& \ (cr>=cb) \ \& \ (s>50) \ \& \ (r>30))|((b==v) \ \& \ (cr>130))$
- 구름, 달 등 배경하늘에 따른 색감 차이 심함
- 2024년 이후 카메라 변경 문제

2

β -TCVAE

- 오로라 관련 잠재 특징 벡터 추출 시도
- 오로라 위치, 색감, 형태, 밝기 불규칙성 문제
- 블랙박스: UMAP, T-SNE, DBSCAN 모두 사용해봤으나 뚜렷한 군집을 찾지 못함



3

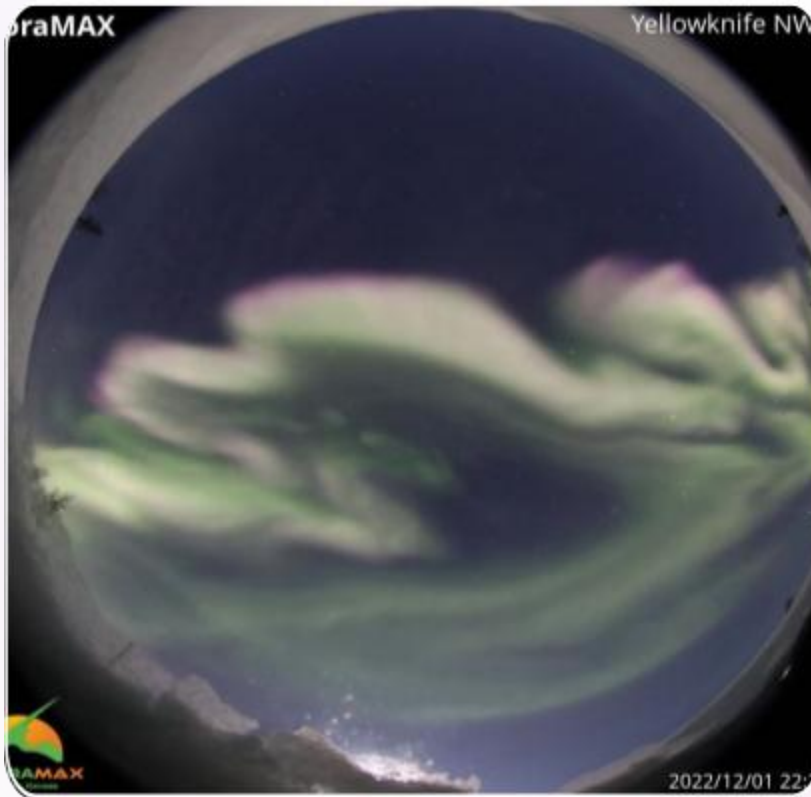
SimCLR + DBSCAN

- 대조 학습 기반 특징 추출
- 데이터 증강 기법 활용
- 클러스터링으로 그룹화 시도
- 전처리 없이 학습 불가 (loss 감소 X)

비지도 학습의 한계

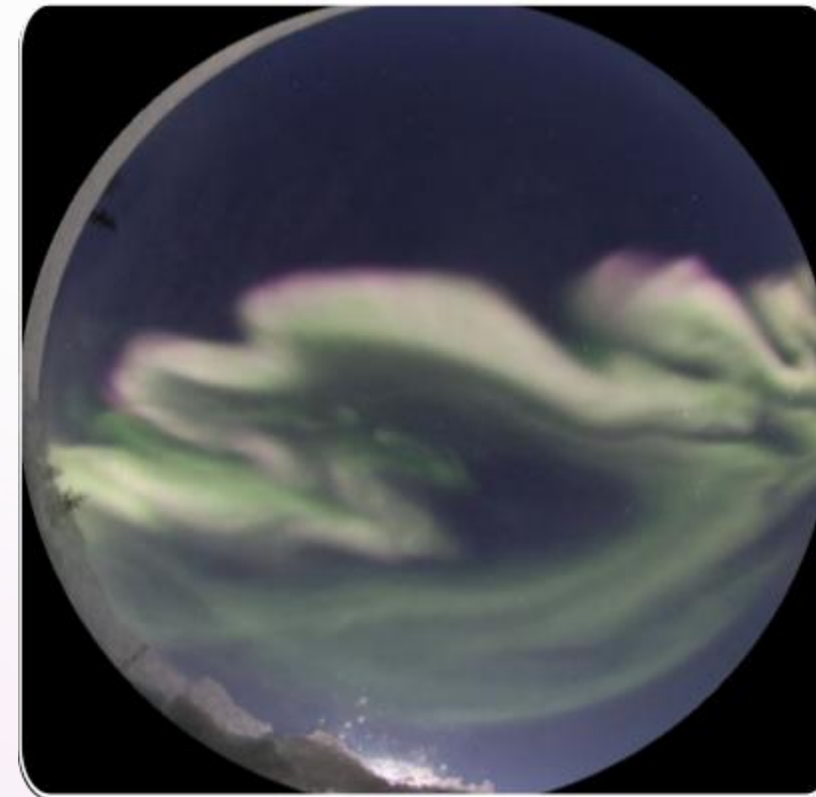
- 오로라의 높은 변동성
- 배경과 오로라 구분 어려움
- 시간적 연속성 반영 부족
- 물리적 특성 고려 필요성 확인

전처리 과정 요약



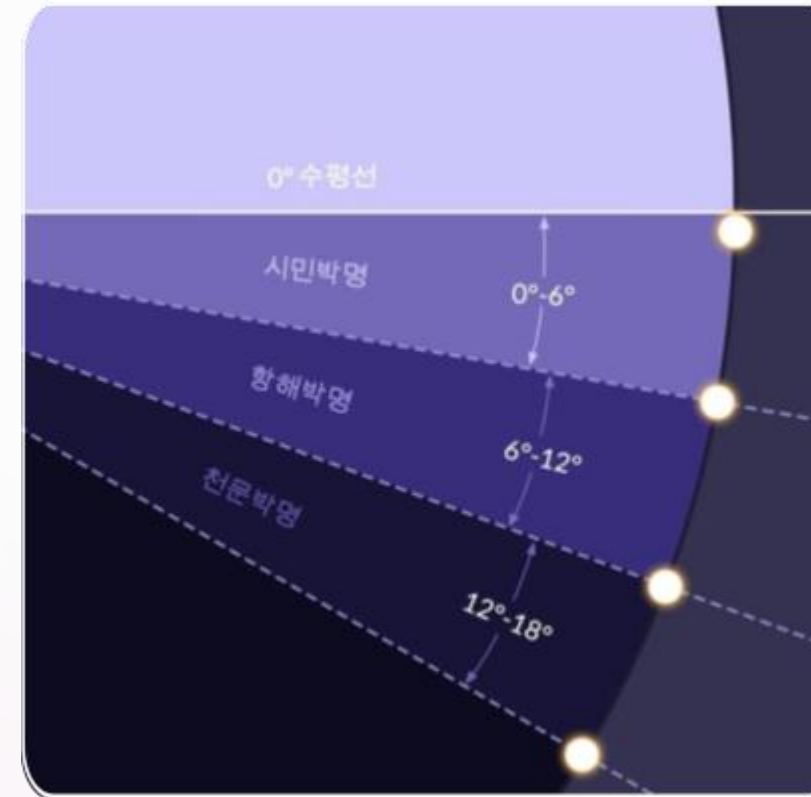
영상 프레임 추출

- 331일치 영상 사용
- 1분 간격으로 추출
- 시간은 텍스트 추출



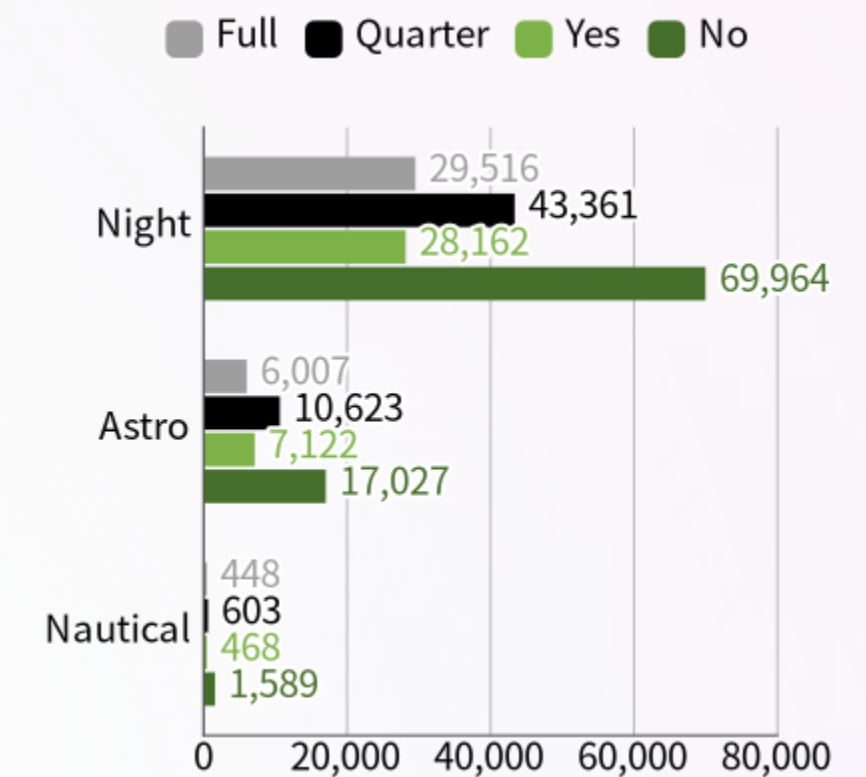
원형 마스크

- 원형 crop 적용
- 불필요한 지평선 제거



데이터 필터링

- 태양 고도 > -18°
- 보름달 고도 > 5°
- Quarter 달 고도 > 10°



수동 라벨링

- 날짜별 20장씩 총 4563장 처리
 - Train: 1024장 * 2
 - Validation: 326장 * 2

EfficientNet-B1 선택 이유

Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.

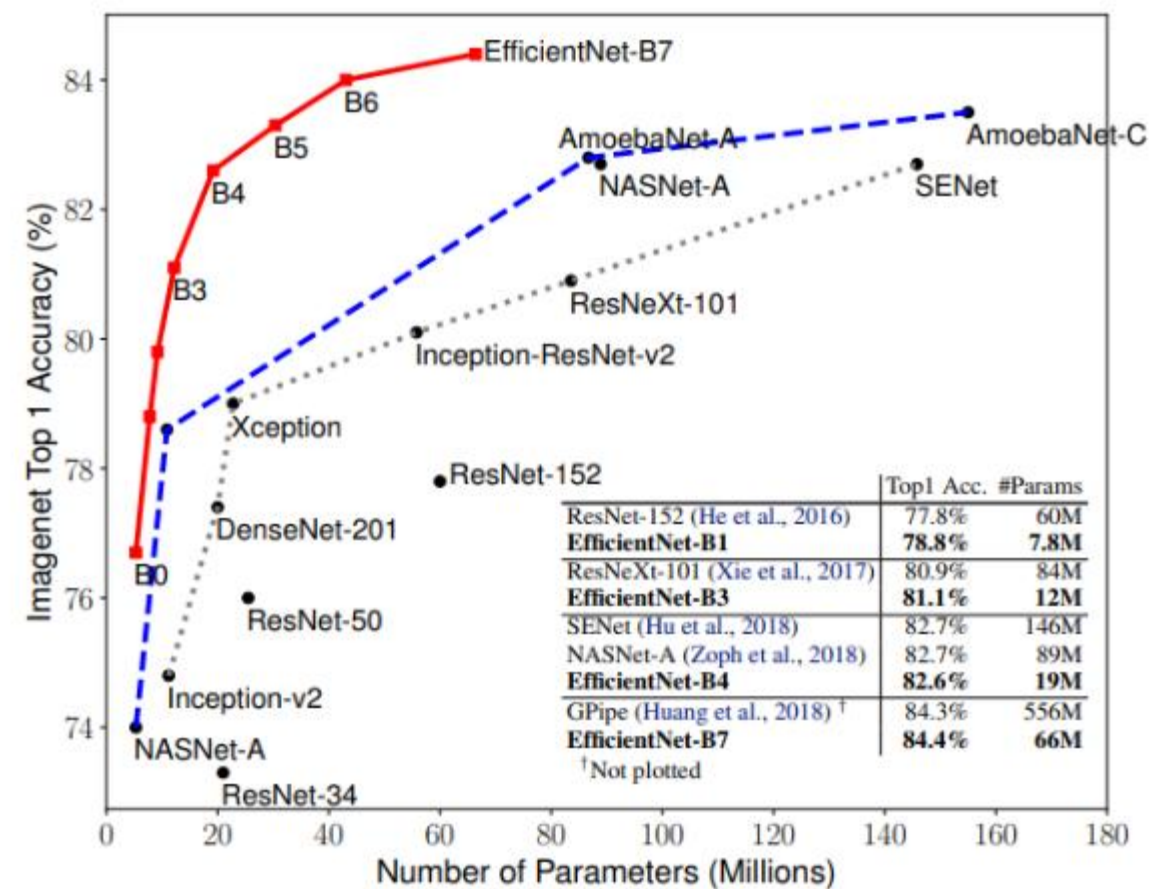
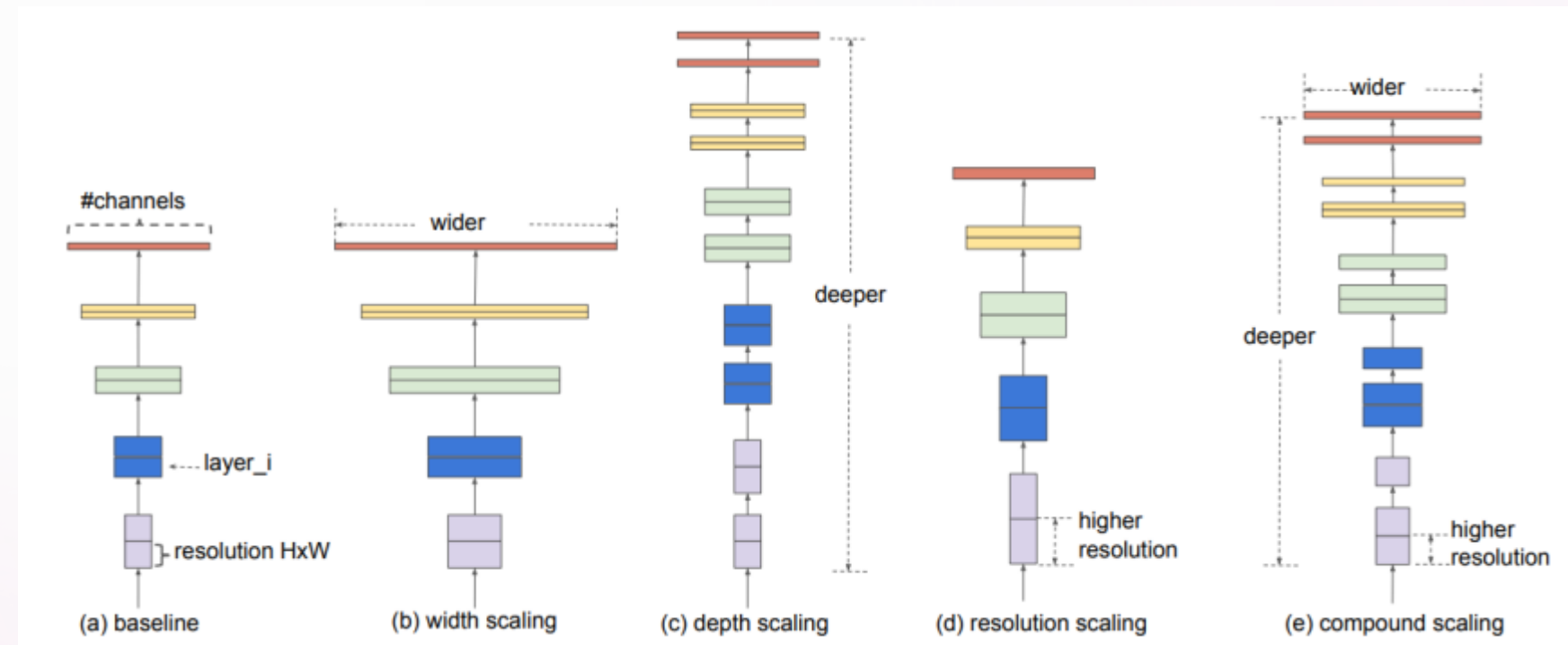
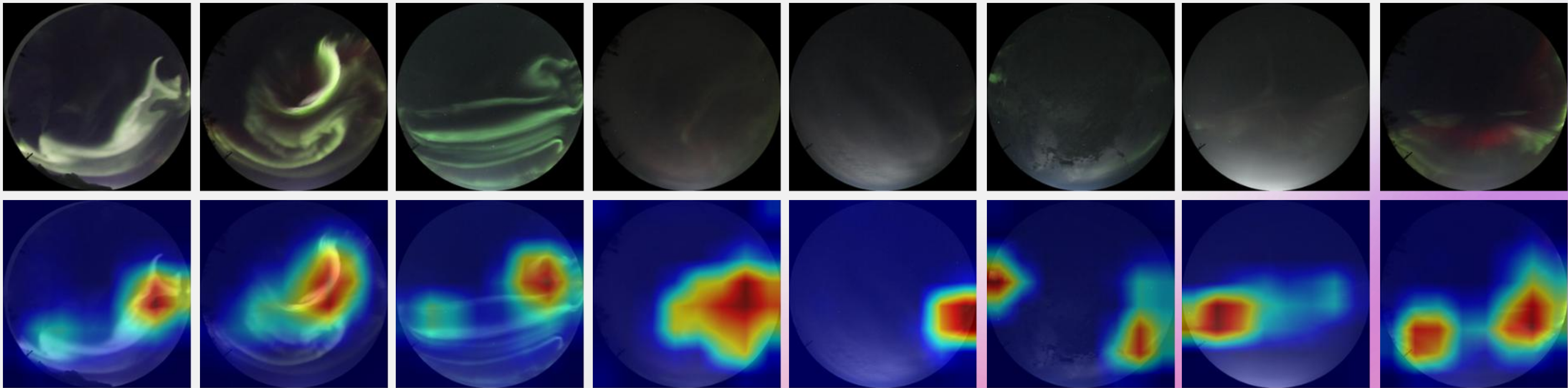


Figure 1. Model Size vs. ImageNet Accuracy. All numbers are for single-crop, single-model. Our EfficientNets significantly outperform other ConvNets. In particular, EfficientNet-B7 achieves new state-of-the-art 84.4% top-1 accuracy but being 8.4x smaller and 6.1x faster than GPipe. EfficientNet-B1 is 7.6x smaller and 5.7x faster than ResNet-152. Details are in Table 2 and 4.



Compound Scaling

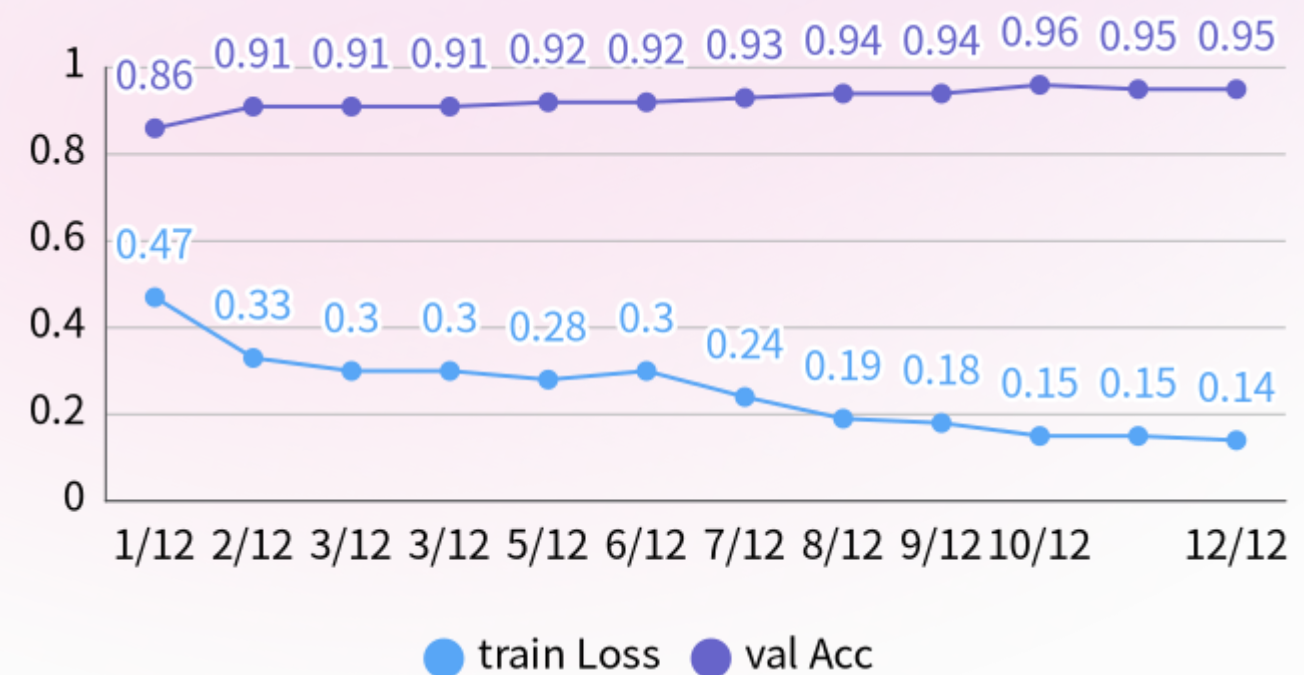
- 깊이, 폭, 해상도를 동시에 **균형 있게 확장**
- 적은 파라미터로 더 높은 정확도와 더 적은 계산량 달성
- RTX 2070 GPU로 2시간이면 처리 가능

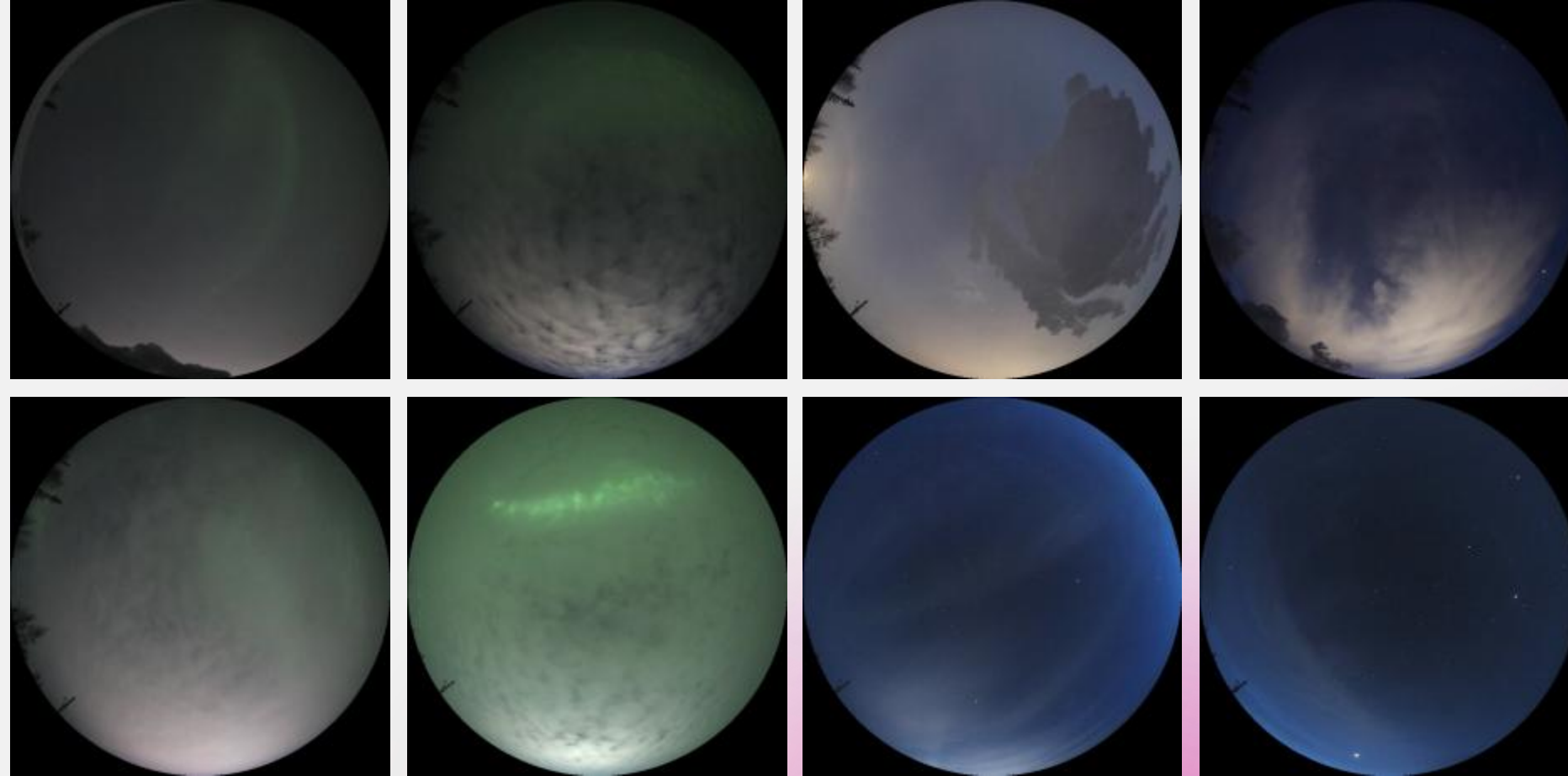


I

전체 분류 정확도

- 다양한 오로라 형태와 밝기에 대해 높은 정확도의 이진 분류 성능 확보
- 구름 유무 및 형태와 관계없이 안정적인 식별 가능
- 특히 Night에서 오로라 유무를 효과적으로 구분함





2

오류 분석

- 밝은 부분이 없을 경우 No→Yes 오류 발생
 - Astro twilight 시점에서 오류율 증가
- 얇은 구름이 오로라처럼 보이는 경우 혼동 발생
- 배경과 오로라 사이 급격한 밝기 변화 경계에 주로 반응

오류 분석

- 시간 정보 활용 고려 중 (10분 전 오로라 유무 반영)
 - 물리적 특성(곡선, RGB 등) 반영 고려 중
- Segmentation 활용 방안 검토 중

결론 요약

딥러닝 분류의 가능성

- 약 1천장만 지도하여 일정 수준 성과 달성
- EfficientNet-B1 모델의 효과적인 이진 분류
- 오로라 유무 이진 분류에서 높은 정확도 (0.96)
- 딥러닝의 오로라 연구 적용 가능성 확인

현재 한계점

- 복잡한 오로라 형태 완전 이해 부족
- 물리적 특성 반영 미흡

향후 연구 방향

- 물리 기반 특징 추출 방법 개발
- 세분화된 오로라 형태 분류 체계 구축
- 시공간적 변화 고려한 모델 설계
- 다중 스펙트럼 데이터 통합 분석

향후 계획

오로라 현상과 태양-지구 상호작용에 대한 이해를 높이고, 우주 기상 예측 능력을 향상시키는 것이 궁극적인 목표

I

오로라 지수 정량화

- 밝기, 면적, 곡률 기반 지수 개발
- 딥러닝 모델을 통한 자동 측정
- 시간에 따른 오로라 변화 추적
- 물리적 의미를 갖는 특징 추출

2

태양풍 상관관계 분석

- Bz, Kp, IMF 등 태양풍 변수 활용
- 시간지연 효과 고려한 상관관계 분석
 - 머신러닝 기반 예측 모델 개발
- 우주 기상 예보 정확도 향상 목표

참고문헌

- Tan & Le (2019). EfficientNet: Rethinking Model Scaling for CNNs. ICML.
- Chen et al. (2020). A Simple Framework for Contrastive Learning of Visual Representations. ICML.
- Jackel et al. (2021). Auroral Image Classification with Machine Learning. Space Weather.
- Zhang et al. (2020). Deep Learning for Night Sky Imaging. JGR Space Physics.
 - NOAA SWPC CME/Flare Data Archives
 - NASA OMNIWeb Database (Solar Wind Parameters)
 - AuroraMAX Live Canadian Space Agency
 - EfficientNet Wikipedia page