방학 프로젝트 최종 발표

Kaggle ISIC 2024 Skin Cancer Detection with 3D-TBP

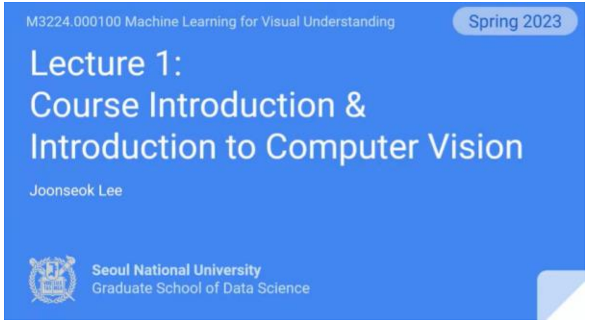
고급 유정민 팀

EWHA EURON

커리큘럼

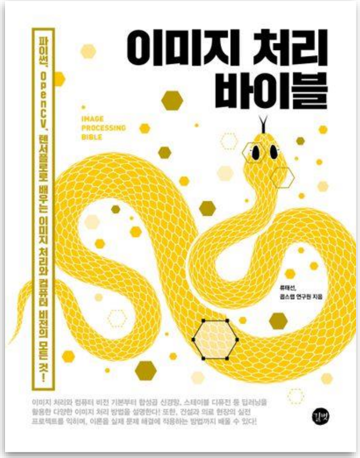
7월 - 스터디(1, 2, 3, 4, 5주차)

- 시각적 이해를 위한 머신러닝 강의 일주일에 2개
- 이미지 처리를 위한 바이블 일주일에 1~2장

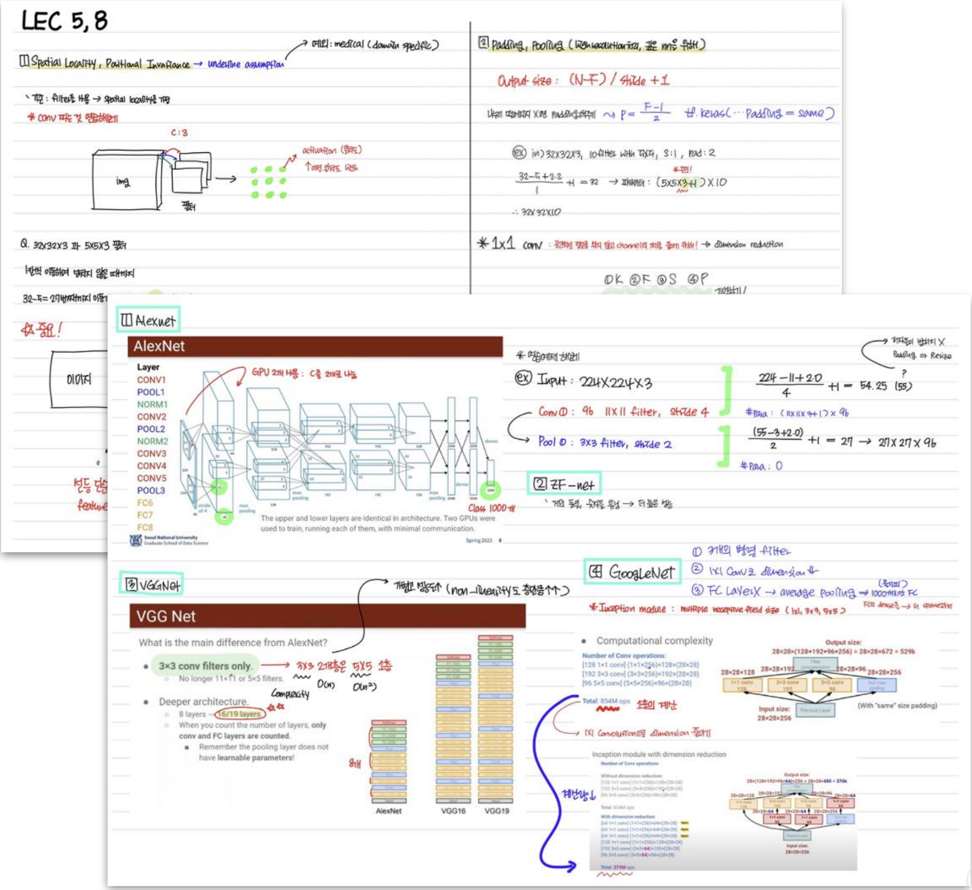


8월 - 대회 준비(6, 7, 8주차)

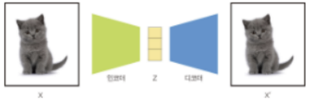
- 이전 대회 수상자들 코드 및 모델 분석
- 데이터 전처리
- 후보 모델 수행
- 최종 모델 선정



스터디 결과물



오토 인코더



- **인코딩, z = f(x):** 이미지 표현 파악 후 작은 사이즈로 압축
  - 인코더: 이미지 차원을 축소하고 잠재 공간으로 표현
- **디코딩, x' = g(z):** 압축된 정보를 원본 이미지의 사이즈와 비슷하게 재구성
  - 디코더: 잠재 공간 표현을 원래의 데이터 공간으로 복원하여 유사한 이미지 생성하도록 학습됨
- 재생성된 이미지(x')가 원본 이미지(x)와 차이나지 않도록 모델을 학습시킴

생성적 적대 신경망

- 정답이 없는
- **생성자:** 실
  - 랜덤 노
- **판별자:** 생
  - 이진 분
- 판별자가

```
import num
import ter
from tens
```

```python
import cv2
import matplotlib.pyplot as plt

!wget https://raw.githubusercontent.com/Lilcob/test_colab/main/three%20young%20man.jpg

# 이미지 로드
image_path = "/content/three young man.jpg"
image = cv2.imread(image_path)
gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

# 하르 캐스케이드 로드
face_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + "haarcascade_frontalface_default.xml")

# 얼굴 탐지
faces = face_cascade.detectMultiScale(gray, scaleFactor=1.1, minNeighbors=5, minSize=(30, 30))
print(faces)

# 탐지된 얼굴에 사각형 그리기
for (x, y, w, h) in faces:
    cv2.rectangle(image, (x, y), (x+w, y+h), (255, 0, 0), 2)

# 결과 표시
plt.imshow(cv2.cvtColor(image, cv2.COLOR_BGR2RGB))
plt.axis('off')  # 축 정보 숨기기
plt.show()
```

대회 소개

[주제] 피부암 탐지 (이진 분류) 대회 종료일 : 9/6

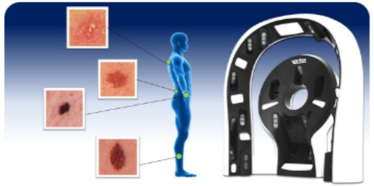
크롭된 피부 이미지를 입력으로 받아 피부암 여부를 판단하는 CV 모델 개발

INTERNATIONAL SKIN IMAGING COLLABORATION (ISIC) · RESEARCH CODE COMPETITION · 18 DAYS TO GO

Submit Prediction ···

ISIC 2024 - Skin Cancer Detection with 3D-TBP

Identify cancers among skin lesions cropped from 3D total body photographs



Overview Data Code Models Discussion Leaderboard Rules Team Submissions

Overview

In this competition, you'll develop image-based algorithms to identify histologically confirmed skin cancer cases with single-lesion crops from 3D total body photos (TBP). The image quality resembles close-up smartphone photos, which are regularly submitted for telehealth purposes. Your binary classification algorithm could be used in settings without access to specialized care and improve triage for early skin cancer detection.

Competition Host

International Skin Imaging Collaboration (ISIC)

Prizes & Awards

$80,000

Awards Points & Medals

대회 전략 분석

2020 ISIC 대회 우승자 코드, 논문 분석

- 2018, 2019, 2020 데이터셋 활용
- Image augmentation 적극 활용: Transpose, Flip, Rotate, Random Brightness 등 적용
- 앙상블 기법: 하나의 CNN이 아니라 여러 개의 모델을 앙상블하여 성능을 올림
- meta data: Linear하게 입력

4 Augmentations and training setup

In small to medium sized datasets, image augmentation is important to prevent overfit. We used in our pipeline the following augmentations from the popular and powerful Pytorch augmentation library Albumentations [1] : Transpose, Flip, Rotate, RandomBrightness, RandomContrast, MotionBlur, MedianBlur, GaussianBlur, GaussNoise, OpticalDistortion, GridDistortion, ElasticTransform, CLAHE, HueSaturationValue, ShiftScaleRotate, Cutout[2]. Figure 2 is an illustration of the before-after of these augmentations.

For training schedule, we used cosine annealing with one warm up epoch [4]. The total number of epochs is 15 for most models. The initial learning rate of the cosine cycle is tuned for each model, which ranges from $1e-4$ to $3e-4$. The learning rate in the warm up epoch is always one tenth of the initial learning rate of the cosine cycle. Batch size is 64 for all models. All training were done on NVIDIA Tesla V100 GPUs in mixed precision. Up to 8 GPUs were used in parallel.

5 Ensemble

In order to make the ensemble stronger, we trained a variety of diverse models. 18 models were selected to comprise the final ensemble as listed in Table 2. The model diversity comes from multiple angles.

- Backbones: EfficientNet B3, B4, B5, B6, B7, SE-ResNeXt-101 [3], ResNeSt-101 [6]
- Different targets: Model 9 has only 4 classes
- 5 folds are split differently among models

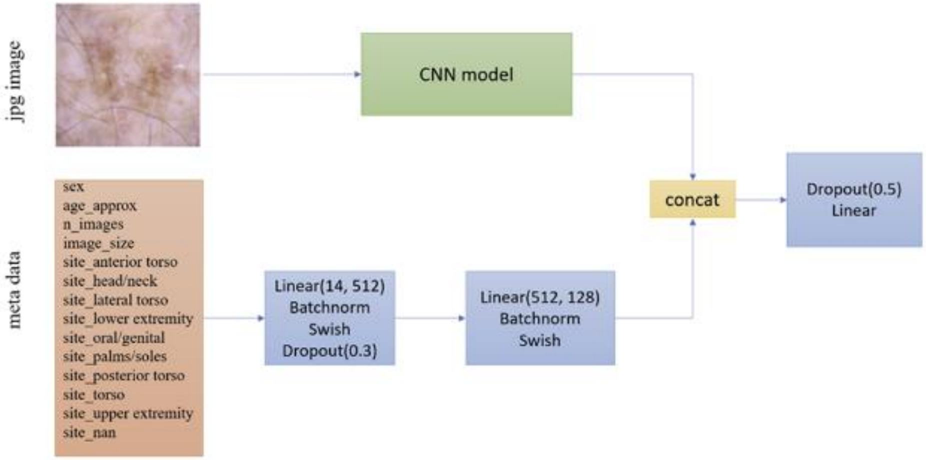
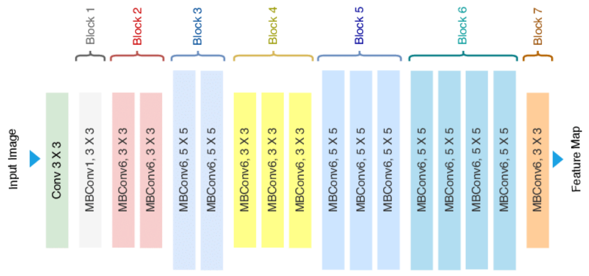


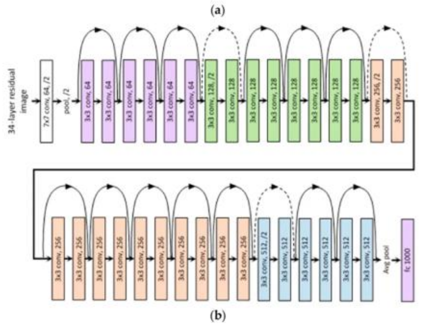
Figure 1: **Model architecture of metadata models.**

대회 전략 고안

Baseline Model

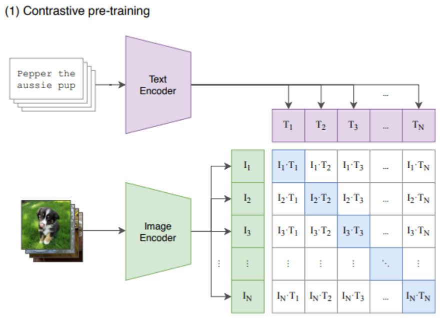
1) ResNet50
2) EfficientNet
3) Inception-v3

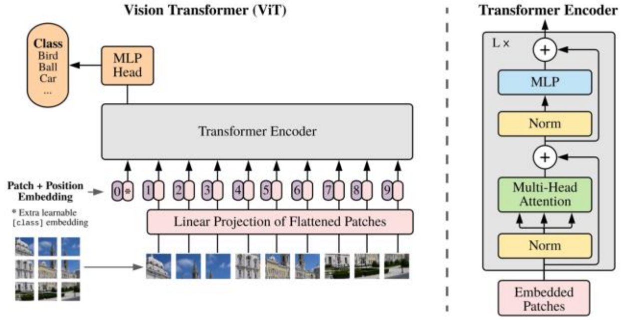




Pre-trained Vision Encoder 활용

1) Google ViT(based-sized model, vit-base-patch16-224)
2) CLIP





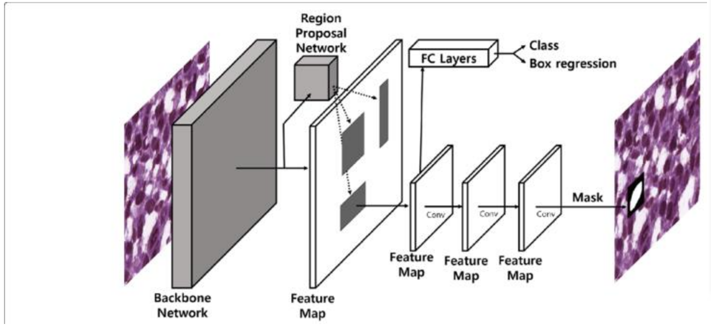
추가 전략 고안

Segmentation

- R-CNN을 활용하여 피부암 부분만 segmentation한 후 feature extract

⇒ Focus, Noise Reduction, Improved Feature Extraction 기대 가능

- U-Net, Attention U-Net, Mask R-NN, SAM



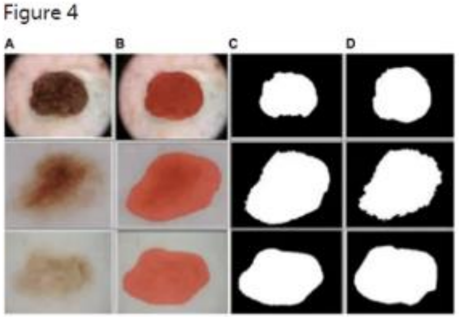
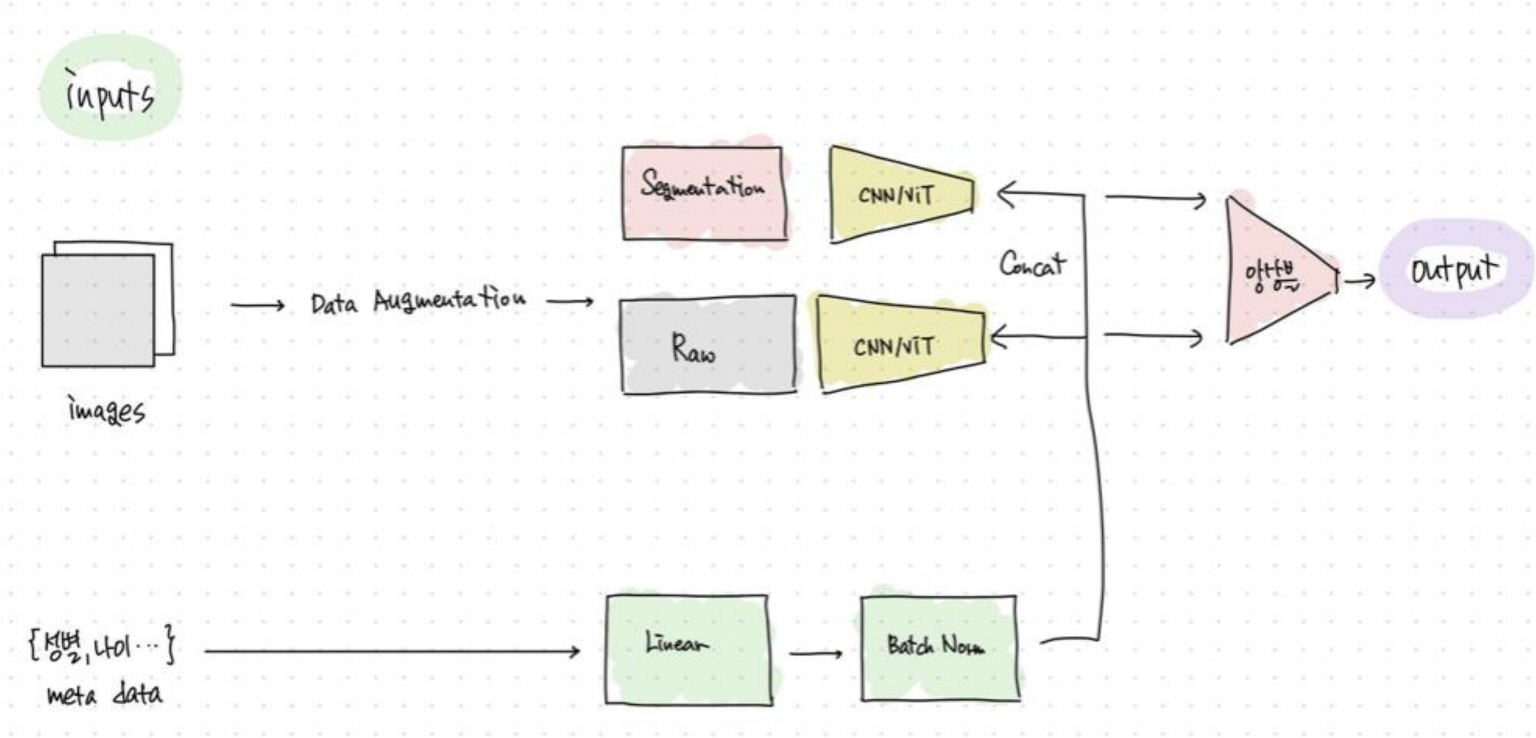


FIGURE 4. Sample segmented dermoscopic image: **(A)** pre-processed images, **(B)** output images of mask-RCNN model, **(C)** binary images of mask-RCNN model, and **(D)** ground-truth images.

최종 구상 모델


inputs
Segmentation
CNN/ViT
Concat
앙상블
output
images
Data Augmentation
Raw
CNN/ViT
{성별, 나이 ...}
meta data
Linear
Batch Norm

진행 상황

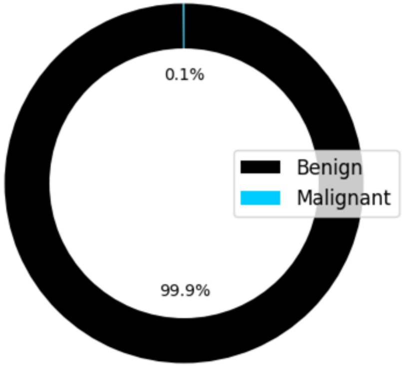
Pre-trained ResNet, EfficientNet으로 1차 baseline 성능 확인 중

⇒ Epoch 1, GPU 1개 1.5시간

```
Epoch [2/10], Loss: 0.005410635624700039
Checkpoint saved at epoch 2
```

Issue: 심한 class imbalance

⇒ 이전 2017, 2018, 2019, 2020 데이터셋 활용, augementation, synthesis 필요



400666 393

THANK YOU