

# 2023 K-ium 의료 인공지능 경진대회

---

Team VITAL **채윤병**  
ybchae@biosoft.kaist.ac.kr



# Contents

1. 팀 소개
2. 데이터 전처리 및 데이터 증강
3. 모델링

## VITAL team

- Team leader



채윤병 Yoonbyung Chae

Major: Biomedical engineering  
ybchae at biosoft.kaist.ac.kr



- Team members



한성주 Sungjoo Han

Major: Biology and Computer science  
sjhan at biosoft.kaist.ac.kr



김세현 Sehyun Kim

Major: Electrical engineering  
shkim at biosoft.kaist.ac.kr



정다연 Dayeon Jeong

Major: Bio and brain engineering  
dyjeong at biosoft.kaist.ac.kr



김서진 Seojin Kim

Major: Bio and brain engineering  
sjkim at biosoft.kaist.ac.kr

- 프로젝트 목적

뇌혈관조영술 영상 사진을 활용한 뇌동맥류 여부와 위치정보를 예측하는 인공지능 알고리즘 개발

- 이전 관련 연구

Jaehyuk Heo et al., 'Prediction of Intracranial Aneurysm Risk using Machine Learning'. Nature(2020)

➡ 머신러닝이용(XGB), AUROC : 0.765

- 주요과정

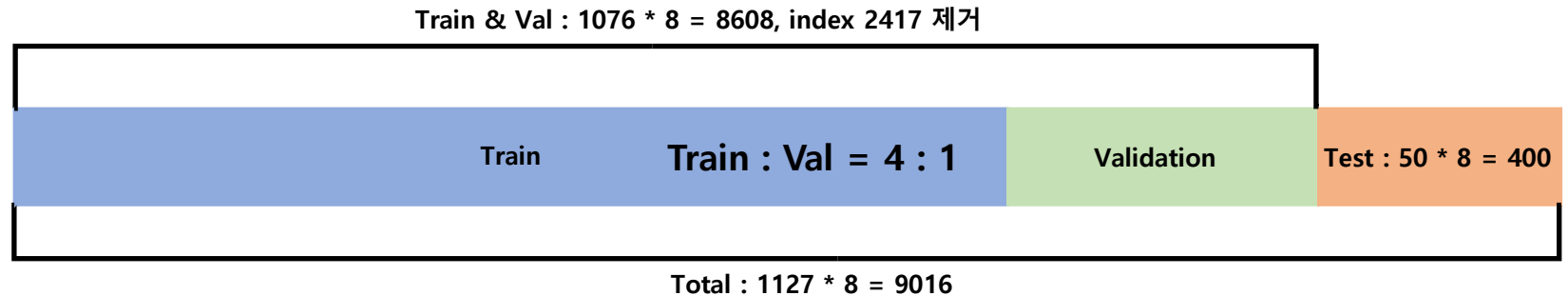
전처리 및 증강 – 지역별 이미지 자르기, 지역별 이미지 정답 값 설정, 데이터 증강

모델링 – Pretrained model 사용(CNN 기반 모델)

## 2. 데이터 전처리 및 데이터 증강

### ■ 데이터 셋

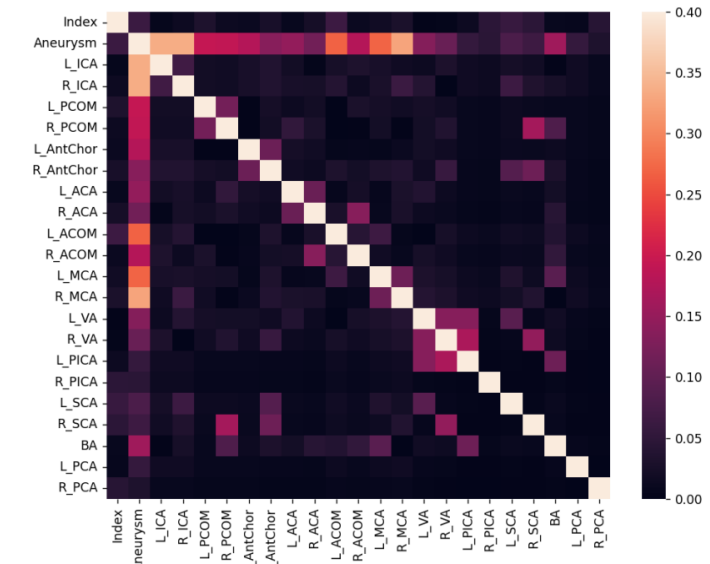
- ✓ 개인당 8개 이미지  
(LI, LV, RI, RV, 2장씩)  
1127명, 9016개 뇌영상
- ✓ Label : 뇌동맥류 여부 1열  
: 위치정보 21열



### ■ EDA

L_ICA	R_ICA	L_ACA	R_ACA	L_VA	R_VA	L_PCA
18.75	18.548	4.032	2.218	3.226	2.218	0.403
L_PCOM	R_PCOM	L_ACOM	R_ACOM	L_PICA	R_PICA	R_PCA
7.056	6.25	12.298	6.048	0.403	0.403	0.202
L_AntChor	R_AntChor	L_MCA	R_MCA	L_SCA	R_SCA	BA
5.847	3.427	12.903	18.75	1.210	0.605	4.435

뇌동맥류가 있는 환자 샘플에서의 위치 정보 통계(%)



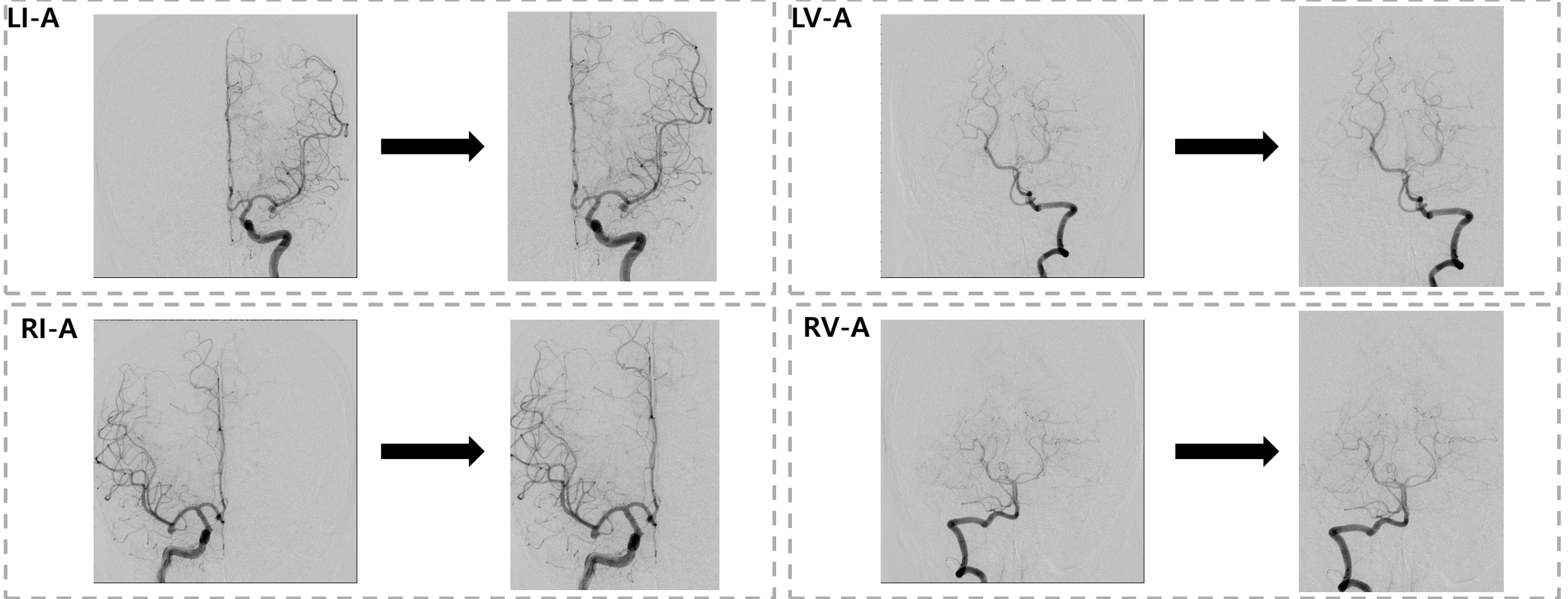
변수들 간 상관관계

- 2,000 번대 index 노이즈 제거

## 2. 데이터 전처리 및 데이터 증강

### ■ 지역별 뇌영상 자르기

지역별로 혈관이 있는 부분이 잘 나타나도록 부위별로 규칙을 정해 뇌영상 전처리 실행

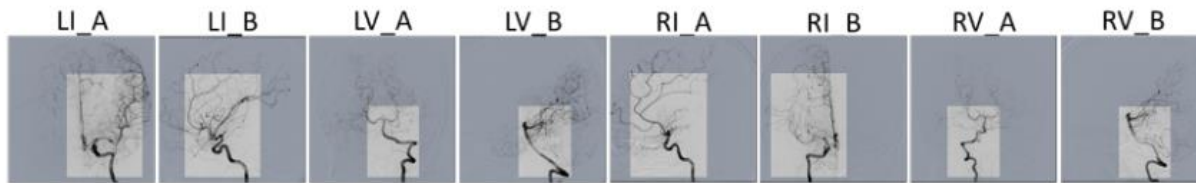




## 2. 데이터 전처리 및 데이터 증강

### ■ 지역별 뇌영상 정답 값 설정

a. 이미지 자르기



b. 라벨 설정하기

L_ICA	L_PCOM
1	0
L_ACA	L_ACOM
0	0
L_Antchor	L_MCA
0	0

라벨 : 1 (뇌동맥류 0)

L_VA	L_PICA
0	0
L_SCA	L_PCA
0	0
BA	
0	

라벨 : 0 (뇌동맥류 X)

R_ICA	R_PCOM
0	0
R_ACA	R_ACOM
0	0
R_Antchor	R_MCA
0	0

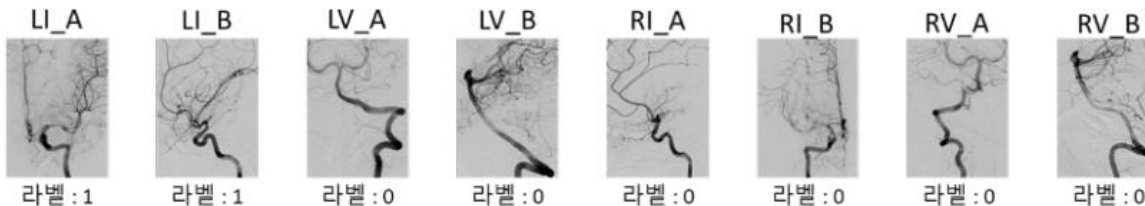
라벨 : 0 (뇌동맥류 X)

R_VA	R_PICA
0	0
R_SCA	R_PCA
0	0
BA	
0	

라벨 : 0 (뇌동맥류 X)

✓ 학습데이터의 위치정보  
정답 값을 이용해 지역별  
로 Aneurysm 여부에 대한  
정답 값 설정

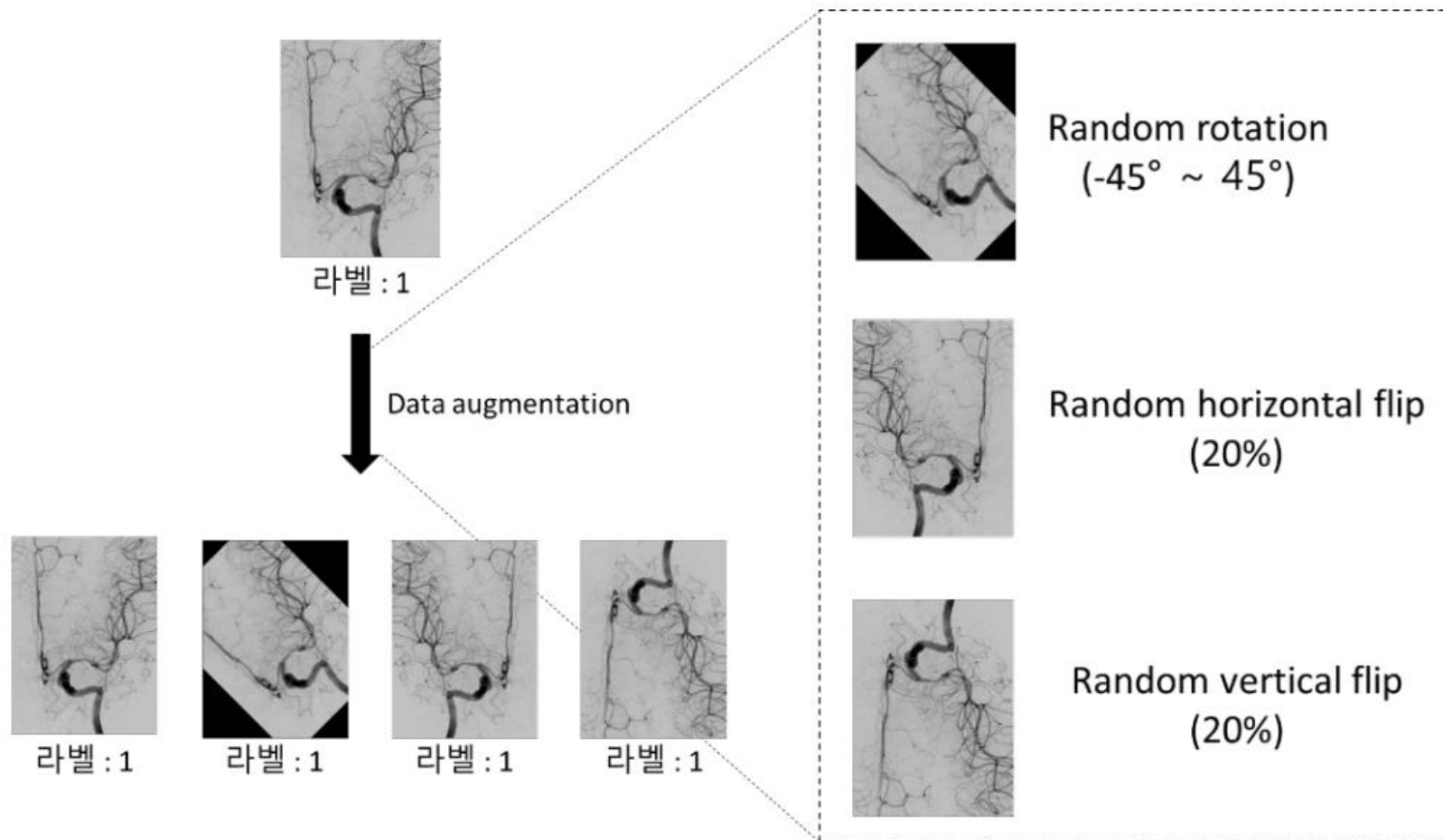
c. 이미지마다  
라벨 부여하기



✓ 같은 사람의 뇌 영상이 부  
위에 따라 다른 정답 값을  
가질 수 있도록 설계

## 2. 데이터 전처리 및 데이터 증강

### ■ 데이터 증강



✓ 뇌 영상을 무작위적으로  
 $-45^\circ \sim 45^\circ$  기울이기

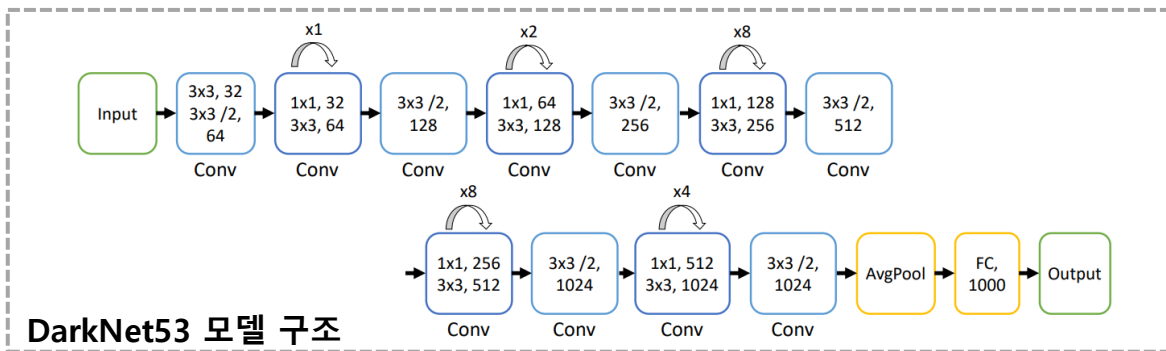
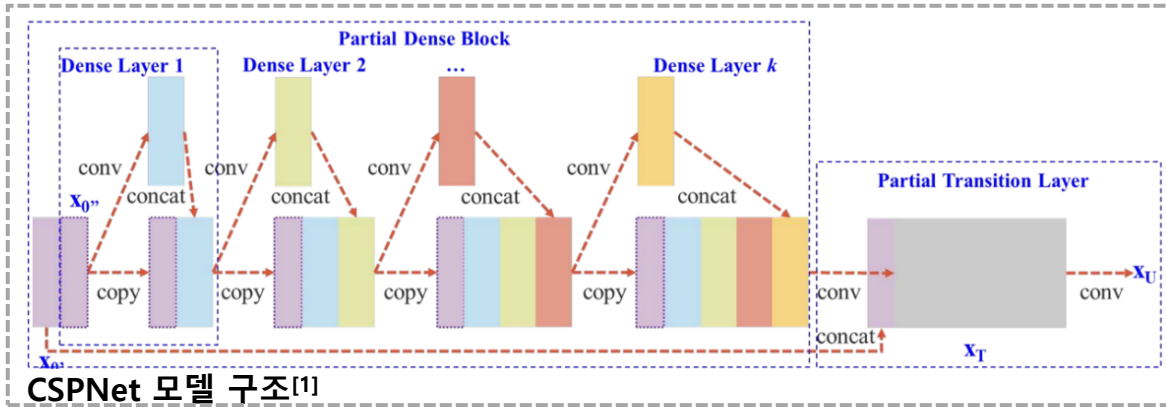
✓ 확률적으로 수평, 수직  
으로 뒤집기



# 3. 모델링

## ■ 뇌동맥류 여부 예측

사전학습 모델 : cs3darknet(CSPNet + DarkNet53)

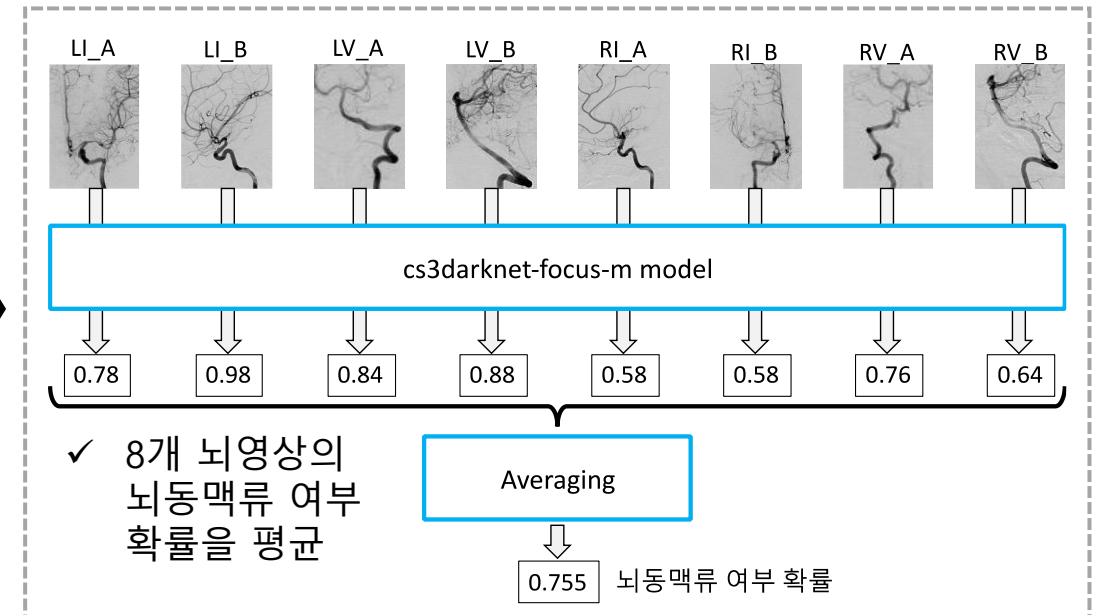


## ✓ Hyperparameter

Optimizer : Adam (Learning rate : 2e-4)

Batch size : 32, Epoch : 40

Early stopping : 13 epochs

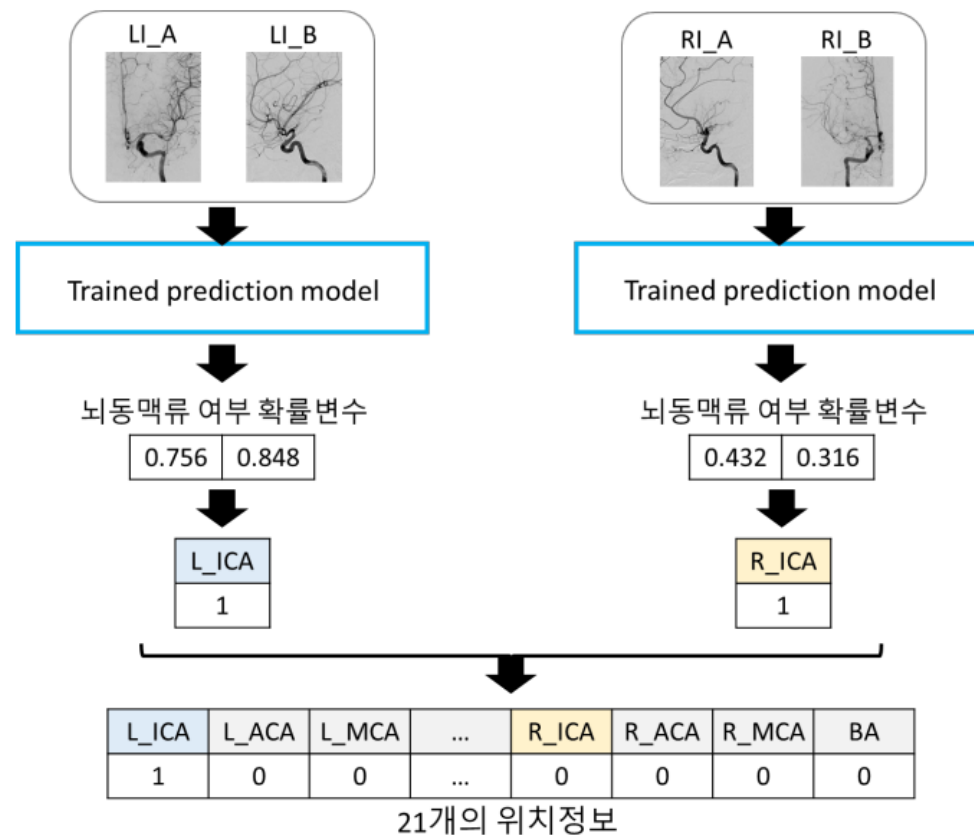


## ■ 뇌동맥류 위치 예측

L_ICA	R_ICA	L_ACA	R_ACA	L_VA	R_VA	L_PCA
18.75	18.548	4.032	2.218	3.226	2.218	0.403
L_PCOM	R_PCOM	L_ACOM	R_ACOM	L_PICA	R_PICA	R_PCA
7.056	6.25	12.298	6.048	0.403	0.403	0.202
L_AntChor	R_AntChor	L_MCA	R_MCA	L_SCA	R_SCA	BA
5.847	3.427	12.903	18.75	1.210	0.605	4.435

뇌동맥류가 있는 환자 샘플에서의 위치 정보 통계(%)

- ✓ 데이터의 수가 적고 위치 정보에 대한 값이 부족하므로 모든 위치 정보에 대한 예측을 할 수 없다고 판단
- ✓ 위치 정보 예측은 통계량이 가장 높은 L\_ICA와 R\_ICA에 대해서 진행
- ✓ 뇌동맥류 여부 예측 모델에서 LI와 RI에 대한 확률값만을 이용하여 예측



#### ■ 최종 성능

	AUROC(뇌동맥류 여부)	Accuracy(뇌동맥류 위치)
임시 Test set(index 1001~1052)	0.889	0.96
Test set	0.741	0.957

#### ■ 아쉬운 점

- 적은 데이터
- 앙상블 모델 도입
- 위치 정보 예측 고도화

**THANK YOU ☺**



