

효과적인 의료 영상 처리를 위한 그룹 압축 자극 블록 기반 다중 헤드 네트워크에 관한 연구

A Study on Group SE Block based Multi-Head Neural Network
for Effective Medical Image Processing

Sanghyuck Lee

tkdgur658@gmail.com

12th Oct, 2022

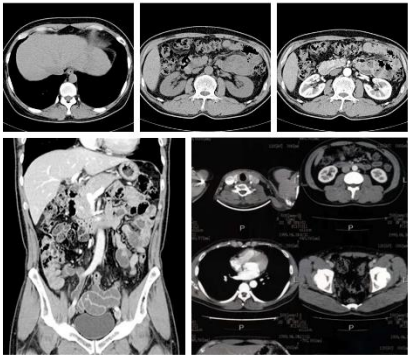


중앙대학교
CHUNG-ANG UNIVERSITY

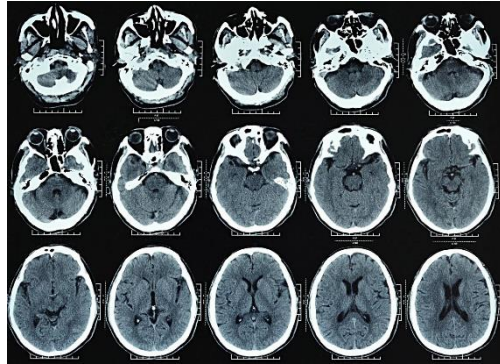


서론(의료 영상 처리란?)

- **의료 영상 처리** : 인체 내부에 존재하는 의학적으로 유용한 생체 정보들을 영상화한 자료들로부터, 진단이나 치료에 필요한 임상 정보를 추출하고 처리하는 과정
- **의료 영상 종류** : CT*, MRI**, 초음파, X-ray, 내시경, 안저, Dermoscopic, 유방조영술 등



<복부 CT 영상>



<두부 MRI 영상>



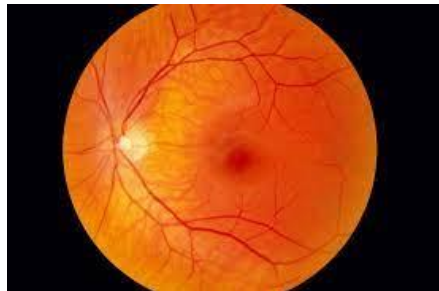
<신장 초음파 영상>



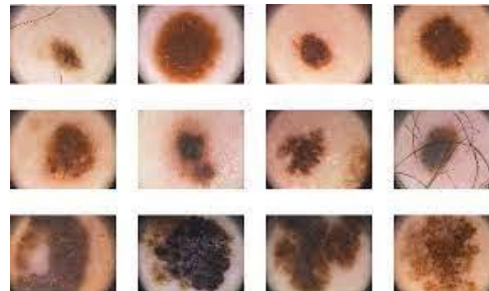
<흉부 X-ray 영상>



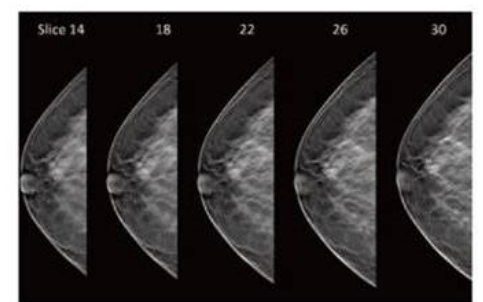
<대장 내시경 영상>



<안저 영상>



<Dermoscopic 영상>

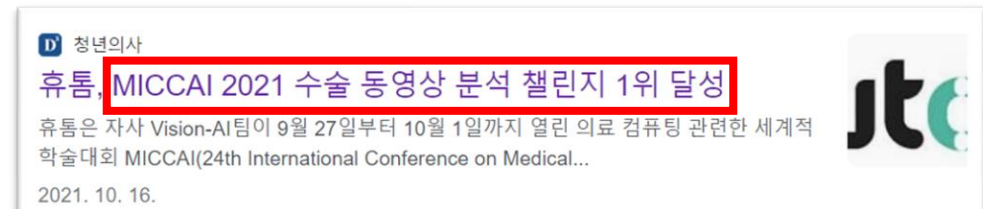
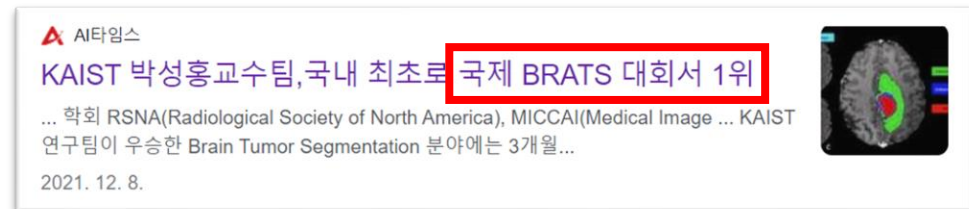
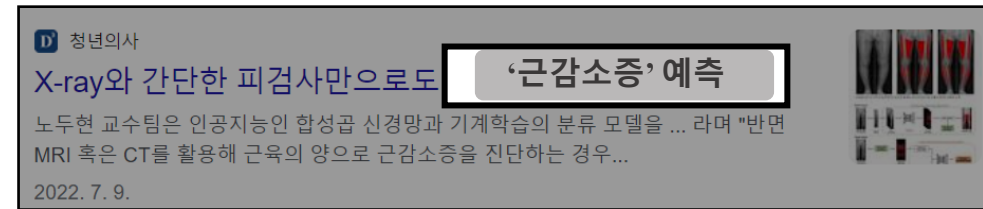
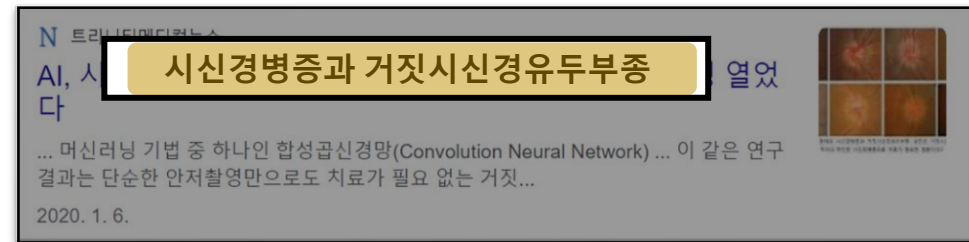
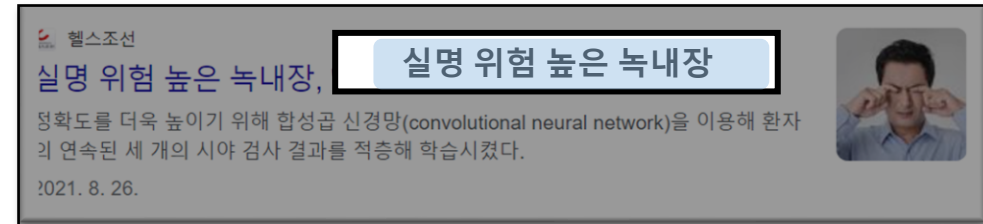
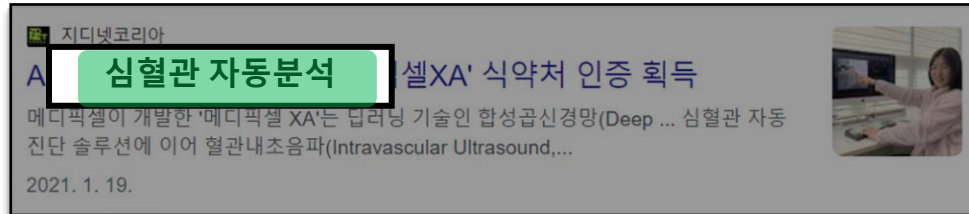
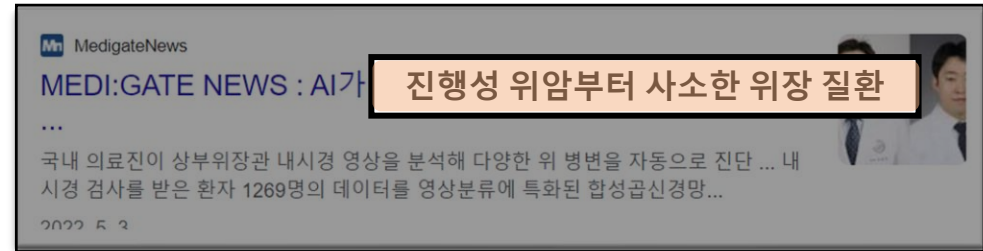
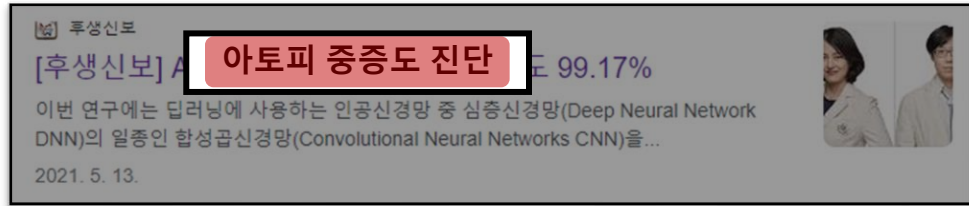


<유방조영술 영상>

* 컴퓨터 단층 촬영(Computed Tomography, CT), **자기공명영상(Magnetic Resonance Imaging, MRI)

서론(CNN과 의료 영상 처리)

- 합성곱 신경망(Convolutional Neural Network, CNN)의 눈부신 발전
- 영상의학과, 조직학, 피부과 등 다양한 의료 영상에서의 유망성



서론(의료용 CNN 연구 가속화)

- 의료용 CNN은 관련 전문가들에게 상당히 장려됨 [1].
- 2017년, 구글(google)사 AI 모델의 89% 정확도로 인간 전문가를 뛰어넘는 성능으로 암 탐지 이후, 의료용 CNN 발전은 가속

Google Research

Philosophy

Research Areas

Publications

People

Tools & Downloads

Outreach

Careers

Blog

BLOG ›

Assisting Pathologists in Detecting Cancer with Deep Learning

FRIDAY, MARCH 03, 2017

Posted by Martin Stumpe, Technical Lead, and Lily Peng, Product Manager

A pathologist's report after reviewing a patient's biological tissue samples is often the gold standard in the diagnosis of many diseases. For cancer in particular, a pathologist's diagnosis has a profound impact on a patient's therapy. The reviewing of pathology slides is a very complex task, requiring years of training to gain the expertise and experience to do well.

⋮

In fact, the prediction heatmaps produced by the algorithm had improved so much that the localization score (FROC) for the algorithm reached 89%, which significantly exceeded the score of 73% for a pathologist with no time constraint². We were not the only ones to see promising results, as other groups were getting scores as high as 81% with the same dataset. Even more exciting for us was that our model generalized very well, even to images that were acquired from a different hospital using different scanners. For full details, see our paper "[Detecting Cancer Metastases on Gigapixel Pathology Images](#)".

Search blog by keyword or author



Archive



Labels



관련 연구

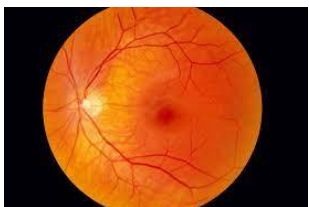
1) 단일 2차원 이미지 입력 CNN



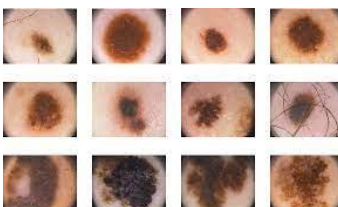
<내시경 영상>



<초음파 영상>

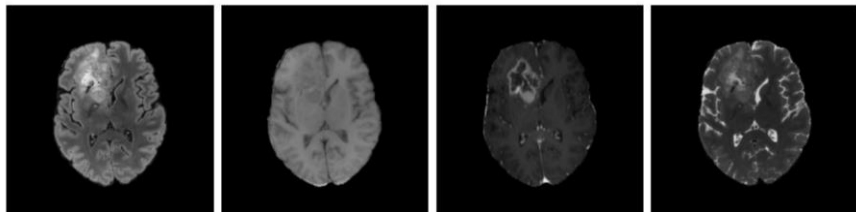


<안저 영상>

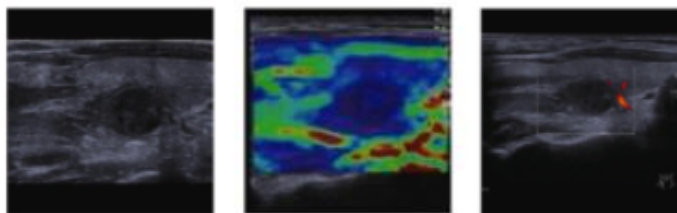


<Dermoscopic 영상>

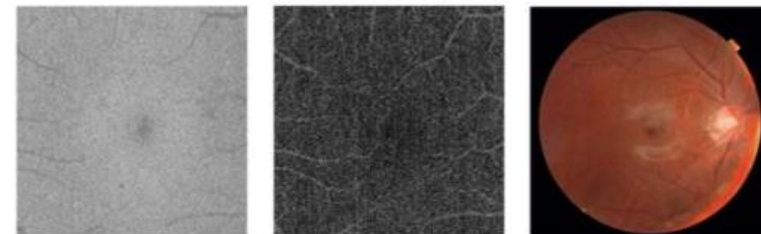
3) 다중 2차원 이미지 입력 CNN



<MRI 영상의 다중 모달리티>

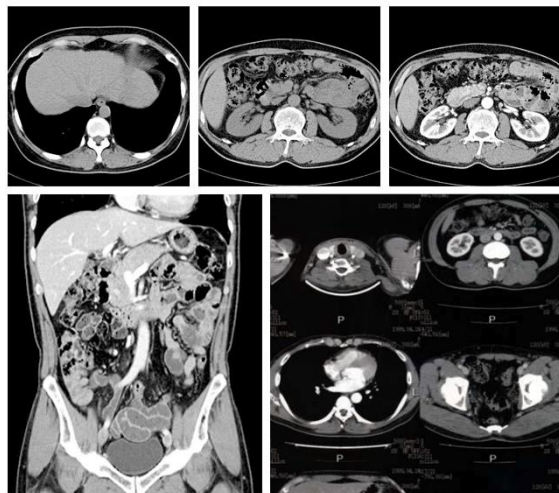


<초음파 영상의 다중 모달리티>

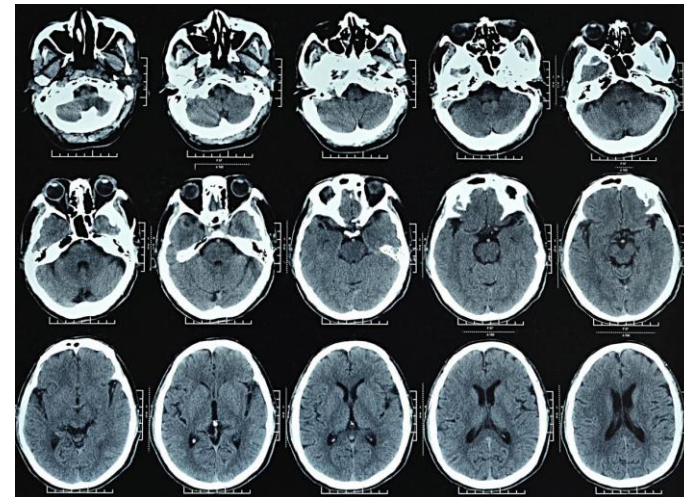


<안저 영상의 다중 모달리티>

2) 3차원 이미지 입력 CNN



<CT 영상>

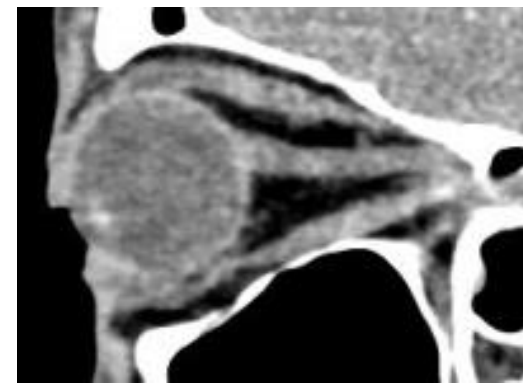
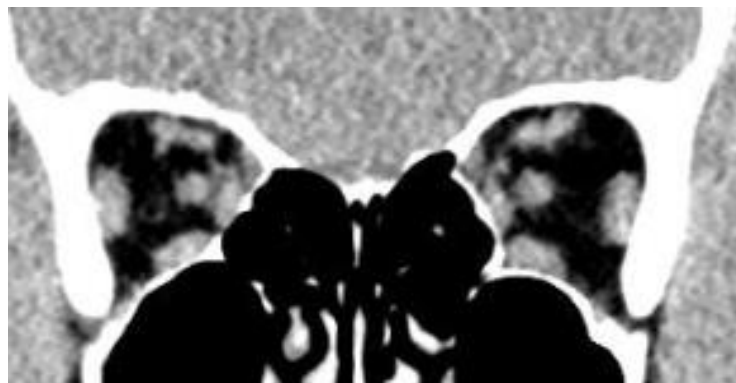


<MRI 영상>

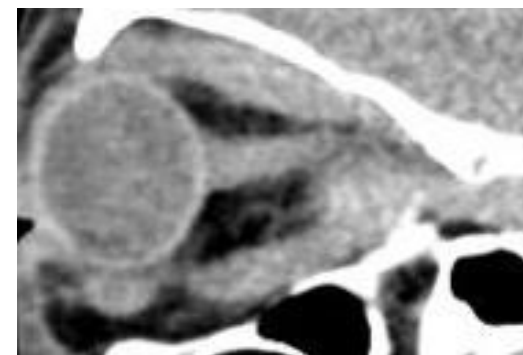
기존 방식들의 한계점

- 기존 방식들의 한계 : 안와내 다수의 근육들, 상안검과 같은 다양한 해부학적 구조물을 다양한 뷰에서 관찰하여 진단하여야 하는 갑상선 안병증(Thyroid-associated ophthalmopathy, TAO)과 같은 질병의 경우, 기존 CNN들은 구조적인 한계로 인해 효과적으로 정보를 수용하기 어려움.

정상



환자



Axial 뷰

Coronal 뷰

Sagittal 뷰

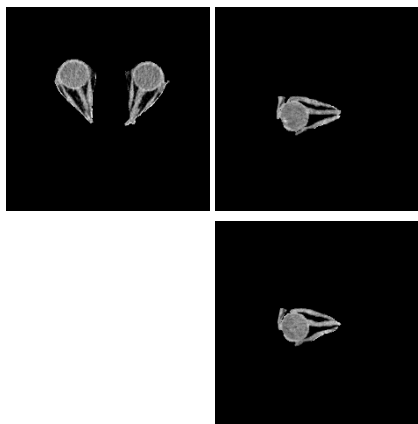
기존 방식들의 한계점



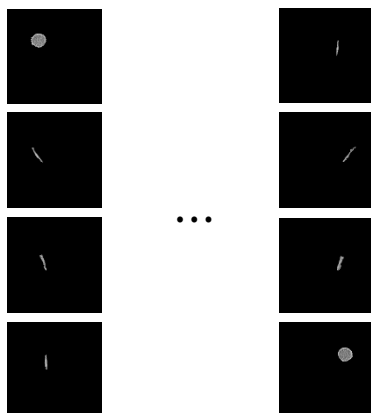
1) 단일 2차원 이미지



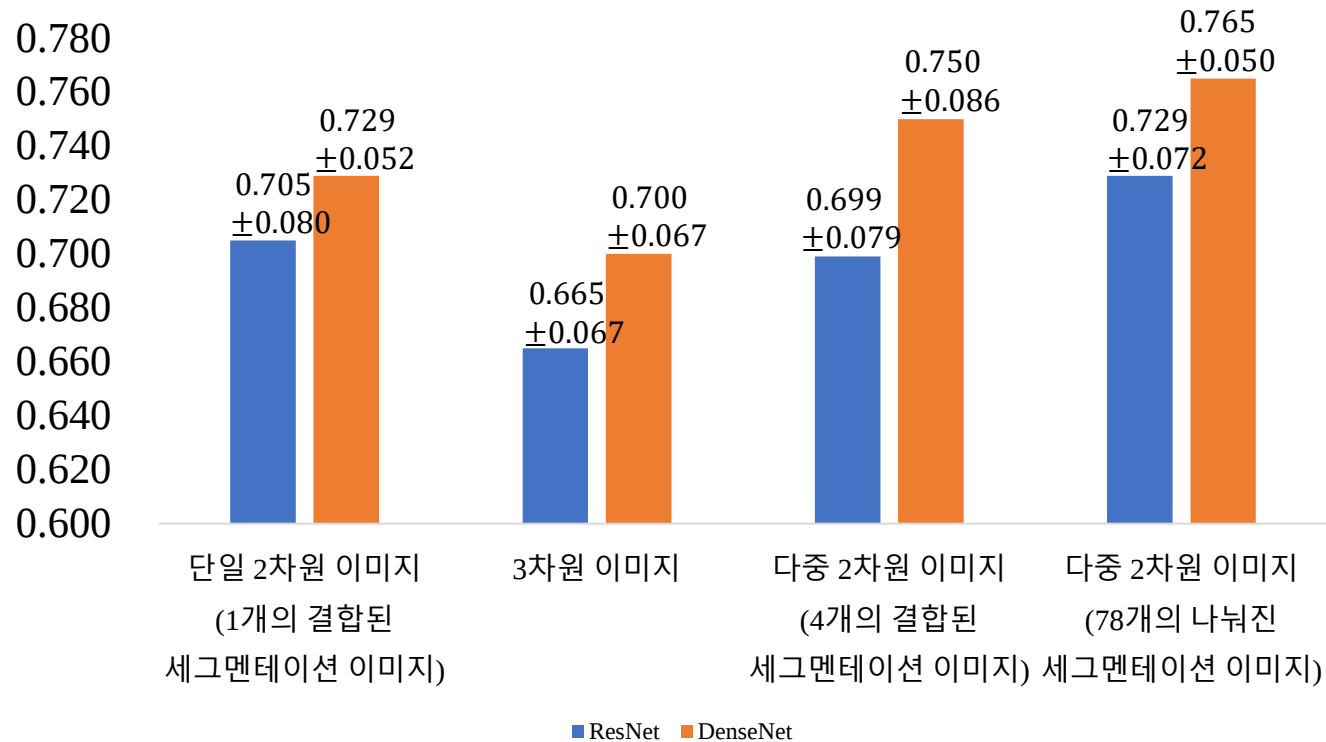
2) 3차원 이미지



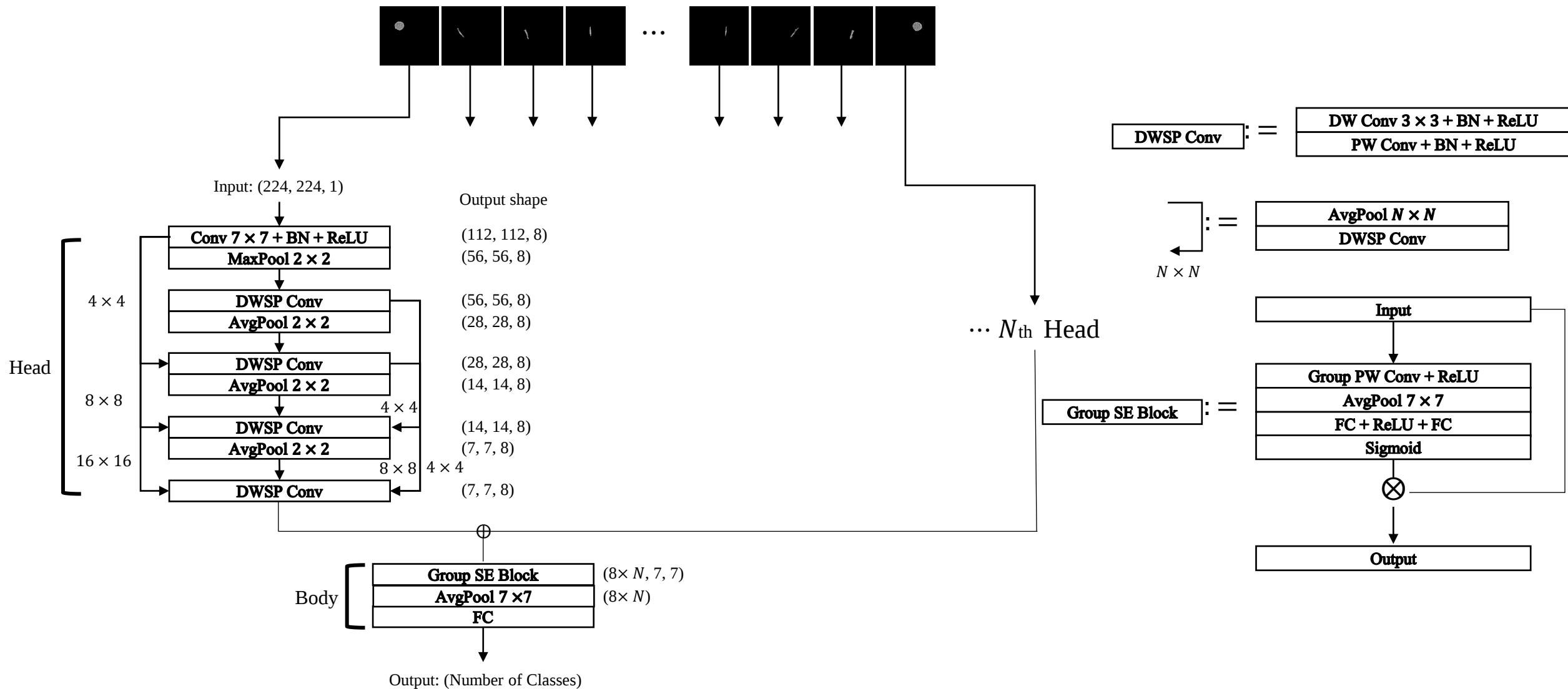
3) 다중 2차원 이미지
(4개의 결합된
세그멘테이션 이미지)



4) 다중 2차원 이미지
(78개의 나뉜
세그멘테이션 이미지)

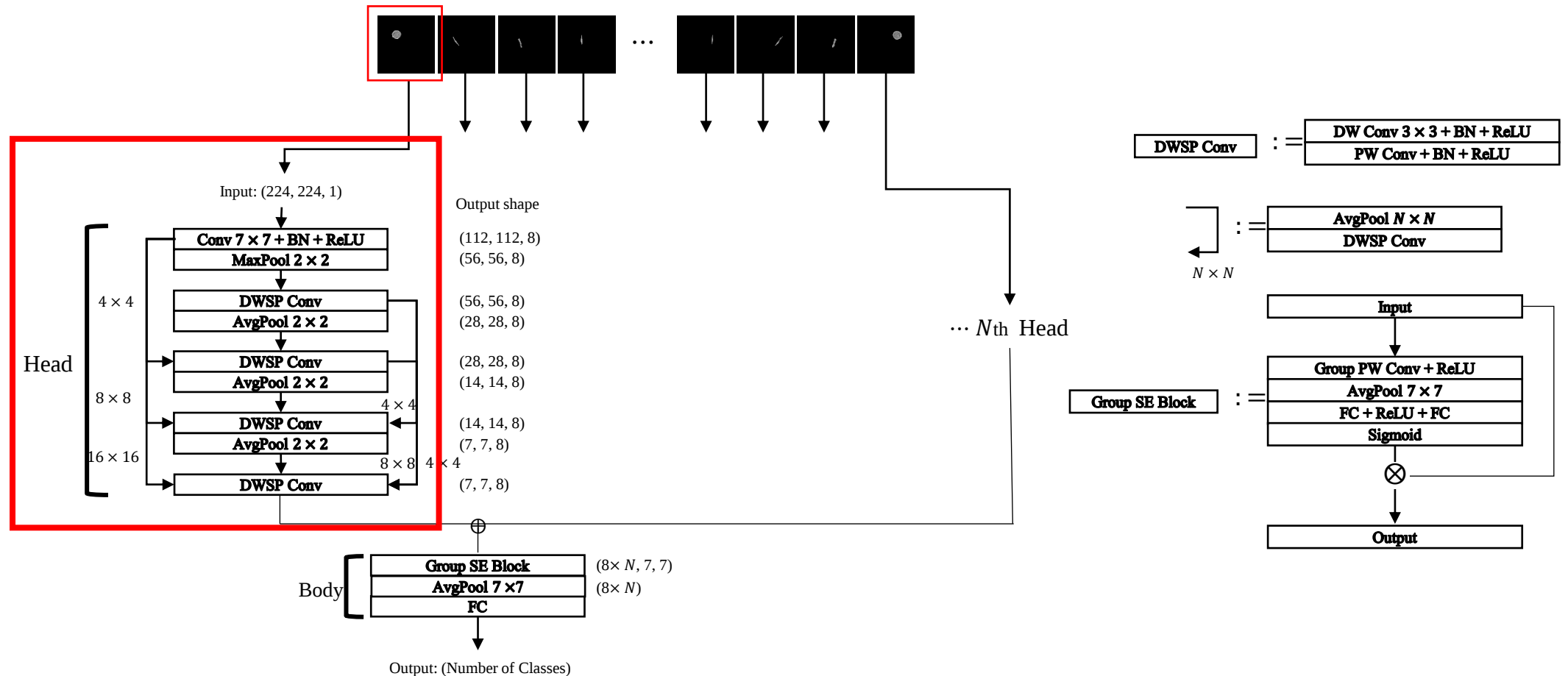


제안 모델



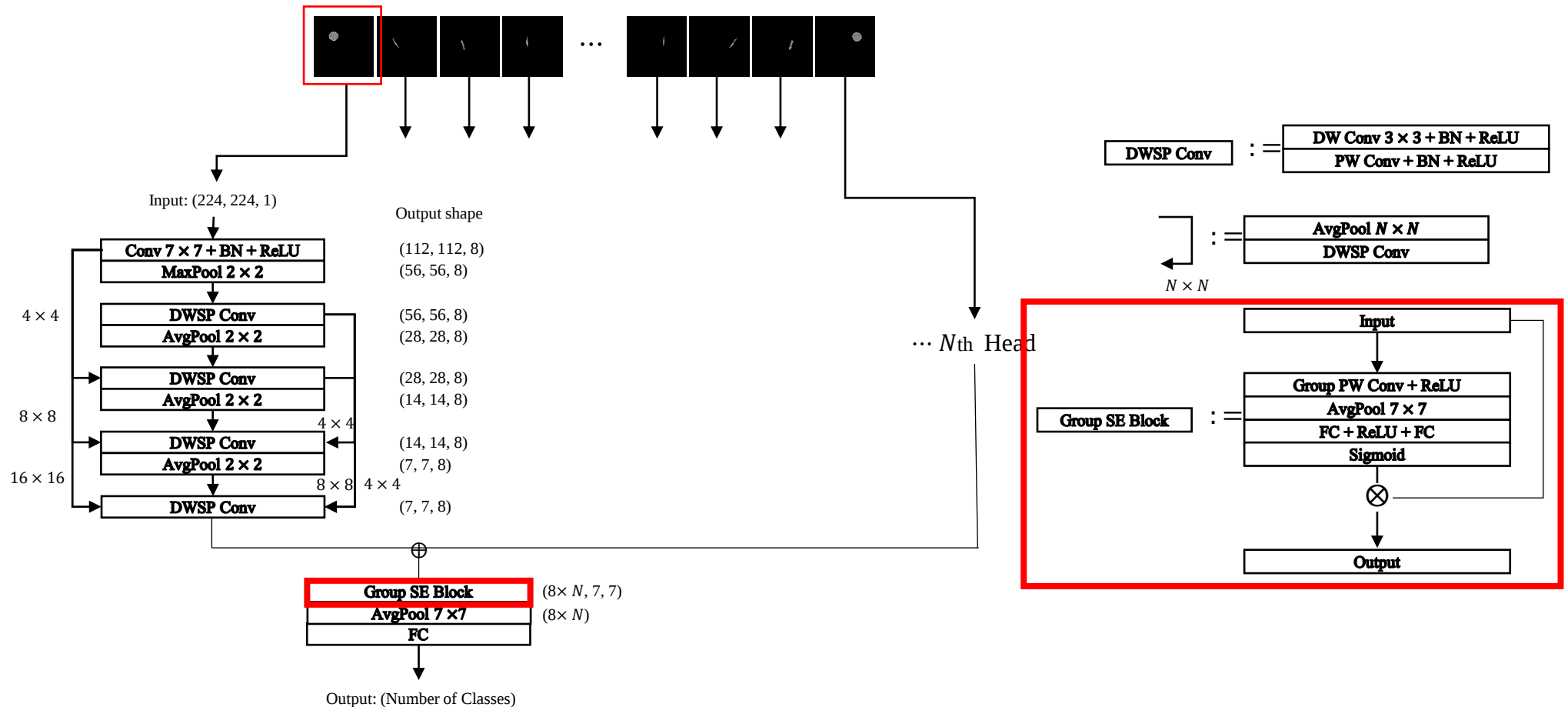
제안 모델(효율적인 다중 헤드 설계)

- 기존 다중 2차원 모델 연구의 일부 접근방식처럼 **다중 헤드 설계**를 통해 각 특징 이미지의 고유한 공간적 특징을 학습할 수 있는 모델을 설계
- 깊이별 분리 합성곱 연산(DWSP Conv 연산)** [2] 을 주요하게 사용하여 전체적인 모델 효율성 증대
- 서로 다른 차원 간의 **스킵 커넥션** 기술을 적용하여 각 계층들이 추가적인 감독을 받도록 설계하여 중간 계층들을 강화 [3, 4]



제안 모델(그룹 압축 자극 블록)

- 압축 자극 블록(squeeze-and-excitation block, SE block) [5] 에 영감을 받아 각 특징 이미지별로 가중치를 재조정하는 그룹 압축 자극 블록을 설계
- 그룹 점별 합성곱 연산을 통해 같은 헤드에서 출력된 특징맵들이 하나의 어텐션을 할당받을 수 있도록 1개 채널로 압축함.



| 실험 설계

➤ 데이터셋

- 클래스 수: 2 (active vs inactive)
- 샘플 수: 144 actives vs 924 inactives
- 입력: (224, 224) 크기의 78개 특징 이미지

➤ 교차 검증

- 50 반복 실험
- 무작위 층화표집(stratified random sampling)
- 데이터 분할: (Train : Validation : Test) = (0.6 : 0.2 : 0.2)

➤ 초매개변수 설정

Model	Batch Size	Epochs	Loss Function	Optimizer	LR	Weight Decay
CSPDenseNet-121 [6], EfficientNet-V2-S [7]	32	30	Focal Loss [8]	AdamW [9]	1e-3	1e-4

➤ 평가 지표

- AUROC, AUPRC, 정확도, F1 점수, 민감도, 특이도, 정밀도

➤ 통계 검정

- 프리드만 검정 [10], 본페르니-돈 검정 [11]

실험 결과

- 제안 모델이 두 개의 비교 모델보다 7개 지표에서 통계적 우위에 있음.

표 1: 실험 결과

모델 이름	AUROC	AUPRC	정확도	F1 점수	민감도	특이도	정밀도
제안 모델	0.806 ± 0.053 (1.23)*	0.498 ± 0.104 (1.37)*	0.730 ± 0.056 (1.34)*	0.429 ± 0.076 (1.34)*	0.731 ± 0.056 (1.37)*	0.729 ± 0.057 (1.36)*	0.306 ± 0.070 (1.34)*
CSPDenseNet	0.751 ± 0.070 (2.13)	0.459 ± 0.107 (1.96)	0.686 ± 0.061 (2.09)	0.378 ± 0.062 (2.09)	0.688 ± 0.060 (2.09)	0.686 ± 0.061 (2.08)	0.262 ± 0.051 (2.09)
EfficientNet-V2	0.725 ± 0.073 (2.64)	0.390 ± 0.101 (2.67)	0.658 ± 0.061 (2.57)	0.347 ± 0.059 (2.57)	0.658 ± 0.058 (2.54)	0.658 ± 0.061 (2.56)	0.237 ± 0.048 (2.57)

- 강조체는 각 지표별로 가장 높은 성능 값을 나타냄.

- 괄호는 평균 순위를 나타냄.

- *표기는 해당 모델이 다른 모델과 비교하여 통계적으로 우위에 있음을 나타냄.

실험 결과

- 제안 모델은 두 비교 모델보다 매우 낮은 공간 및 계산 효율성을 보임.

표 2: 복잡도

모델 이름	매개변수 수 (백만)	FLOPS (십억)
제안 모델	0.18	0.24
CSPDenseNet	117.76	6.27
EfficientNet-V2	5.52	2.70

토의(그룹 압축 자극 블록에 대한 절제 연구)

- 제안 모델의 성능에 대한 그룹 압축 자극 블록의 영향을 검증하기 위해 그룹 압축 자극 블록을, 가장 널리 사용되는 어텐션 모듈인 CBAM [12], SE [5] 블록을 사용할 경우, 어텐션 블록을 사용하지 않을 경우를 각각 비교함.
- 제안 모델이 7개 지표에서 모두 가장 높은 평균 성능을 보이고, AUROC 측면에서 통계적으로 우위에 있음.

표 3: 그룹 압축 자극 블록에 대한 절제 연구

모델 이름	AUROC	AUPRC	정확도	F1 점수	민감도	특이도	정밀도
제안 모델	0.806 ± 0.053 (1.95)*	0.498 ± 0.104 (2.23)	0.730 ± 0.056 (2.13)	0.429 ± 0.076 (2.11)	0.731 ± 0.056 (2.1)	0.729 ± 0.057 (2.13)	0.306 ± 0.070 (2.11)
CBAM	0.785 ± 0.050 (2.75)	0.473 ± 0.094 (2.66)	0.717 ± 0.050 (2.75)	0.399 ± 0.079 (2.79)	0.700 ± 0.110 (2.68)	0.720 ± 0.061 (2.76)	0.280 ± 0.061 (2.79)
SE	0.793 ± 0.048 (2.57)	0.482 ± 0.097 (2.59)	0.718 ± 0.044 (2.55)	0.411 ± 0.056 (2.53)	0.715 ± 0.044 (2.62)	0.718 ± 0.044 (2.54)	0.289 ± 0.050 (2.53)
어텐션 사용 하지 않음	0.788 ± 0.044 (2.73)	0.475 ± 0.086 (2.52)	0.718 ± 0.041 (2.57)	0.411 ± 0.053 (2.57)	0.716 ± 0.044 (2.6)	0.719 ± 0.041 (2.57)	0.289 ± 0.046 (2.57)

- 강조체는 각 지표별로 가장 높은 성능 값을 나타냄.

- 괄호는 평균 순위를 나타냄.

- *표기는 해당 모델이 다른 모델과 비교하여 통계적으로 우위에 있음을 나타냄.

토의(매개변수 수에 대한 절제 연구)

- 기계학습 모델은 매개변수 수가 과도하게 많으나, 훈련 샘플 수가 적을 경우, 훈련 데이터에 과도하게 적합되어 낮은 성능을 보일 수 있음.
- 이에 대해 CSPDenseNet의 매개변수 수를 대폭 낮춰, 제안 모델의 구조적 우수성을 검증함.
- CSPDenseNet은 매개변수를 낮추어 전체적인 성능이 상승하였지만, 여전히 제안 모델이 통계적으로 7개 지표에서 모두 가장 뛰어난 성능을 보임.

표 4: 매개변수 수에 대한 절제 연구

모델 이름	매개변수 수	AUROC	AUPRC	정확도	F1 점수	민감도	특이도	정밀도
제안 모델	0.179	0.806 ± 0.053 (1.27)*	0.498 ± 0.104 (1.68)*	0.730 ± 0.056 (1.43)*	0.429 ± 0.076 (1.43)*	0.731 ± 0.056 (1.48)*	0.729 ± 0.057 (1.44)*	0.306 ± 0.070 (1.43)*
CSPDenseNet	5.515	0.751 ± 0.070 (2.45)	0.459 ± 0.107 (2.19)	0.686 ± 0.061 (2.35)	0.378 ± 0.062 (2.35)	0.688 ± 0.060 (2.33)	0.686 ± 0.061 (2.34)	0.262 ± 0.051 (2.35)
CSPDenseNet (축소버전)	0.184	0.771 ± 0.053 (2.28)	0.478 ± 0.097 (2.13)	0.698 ± 0.047 (2.22)	0.388 ± 0.054 (2.22)	0.698 ± 0.047 (2.19)	0.698 ± 0.047 (2.22)	0.270 ± 0.046 (2.22)

- 강조체는 각 지표별로 가장 높은 성능 값을 나타냄.

- 괄호는 평균 순위를 나타냄.

- *표기는 해당 모델이 다른 모델과 비교하여 통계적으로 우위에 있음을 나타냄.

토의(어텐션 블록을 이용한 중요 특징 이미지 분석)

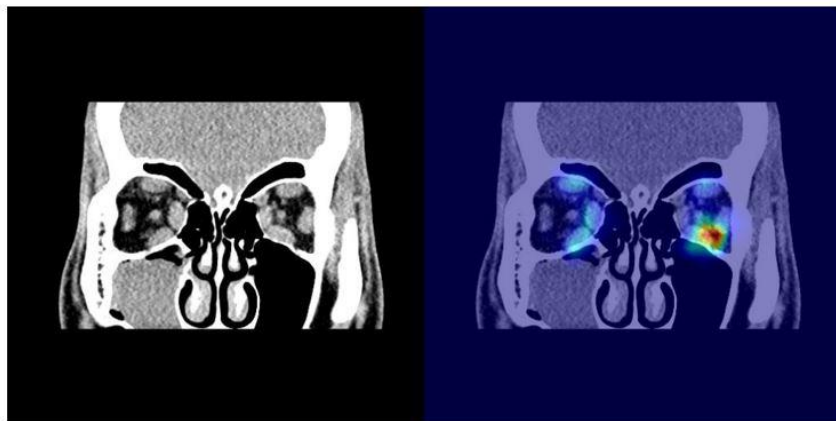
- 하직근, 상직근의 중요도가 높음.
- Coronal 2 슬라이스가 전체적으로 중요도가 높음.
- 눈은 하위 순위에 대거 포함되었음.

표 3: 그룹 압축 자극 블록에 의해 산출된 어텐션 상위 10 특징 이미지

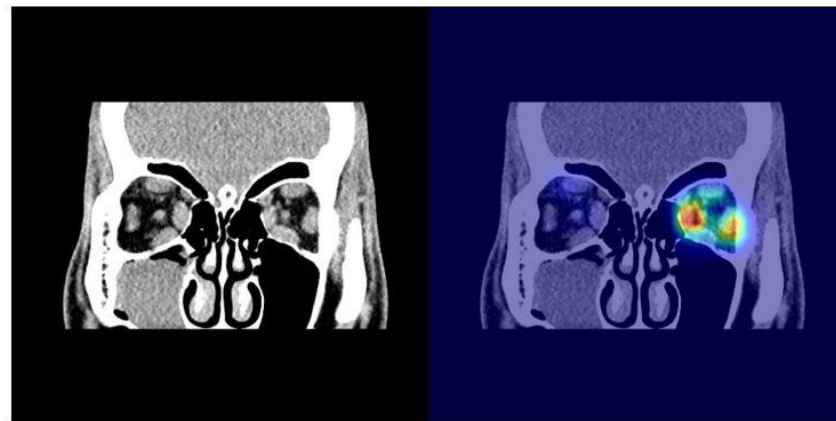
순위	어텐션 점수	슬라이스 ID	좌우	구조물
1	0.619 ± 0.107	Coronal 2	좌	하직근
2	0.616 ± 0.090	Coronal 2	우	상직근
3	0.614 ± 0.099	Coronal 2	좌	상직근
4	0.586 ± 0.092	Coronal 2	우	하직근
5	0.585 ± 0.077	Sagittal 1	우	하직근
6	0.585 ± 0.088	Sagittal 2	좌	하직근
7	0.585 ± 0.074	Sagittal 1	우	상직근
8	0.581 ± 0.091	Coronal 2	우	내직근
9	0.573 ± 0.070	Sagittal 2	좌	상직근
10	0.569 ± 0.074	Axial 2	우	내직근

토크의(어텐션 블록을 이용한 중요 특징 이미지 분석)

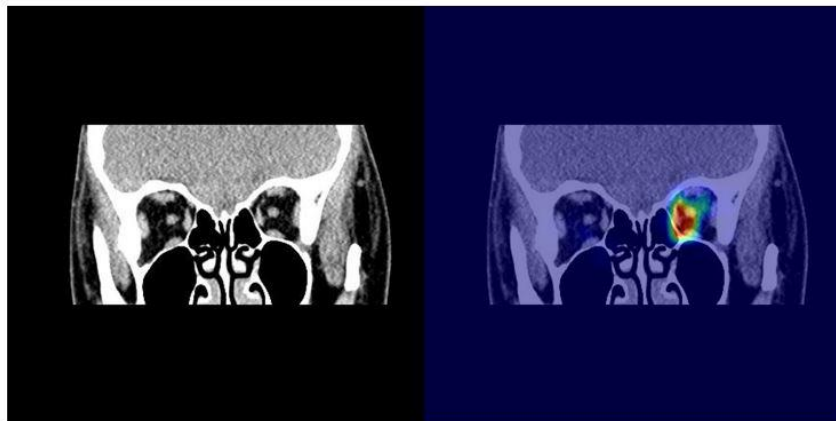
➤ Grad-CAM [13]



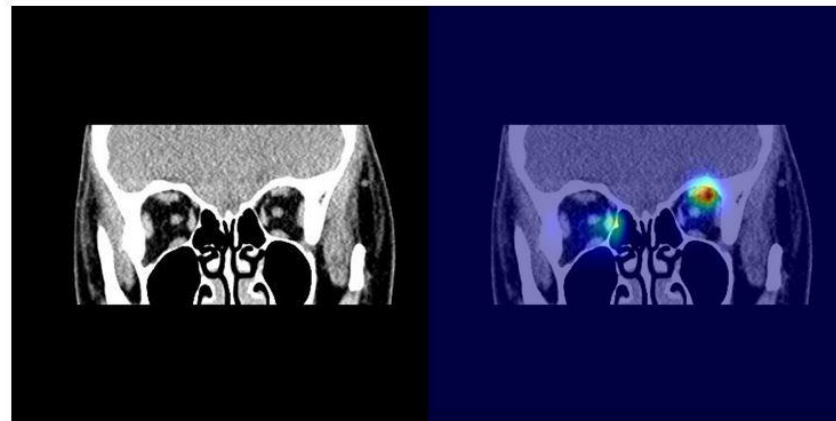
스텝 계층 tp 케이스



첫번째 DWSP Conv 계층 tp 케이스



스텝 계층 tn 케이스



첫번째 DWSP Conv 계층 tn 케이스

결론

- 본 연구에서는 기존 연구에서는 없었던 많은 수의 특징 이미지를 입력하는 방식이 의료 영상 처리를 위해 효과적일 수 있음을 확인하였다.
- 제안 모델로써 다중 특징이미지를 독립적으로 처리하는 효율적 다중 헤드를 설계했고, 그룹 압축 자극 블록을 통해 중요한 특징이미지 및 헤드를 강하게 학습하였다.
- 제안 모델은 TAO 데이터셋에서 2개의 최신 CNN 모델보다 더 나은 성능과 높은 파라미터 및 계산 효율성을 보였다.

기여

In International Conferences

- Multi Head Network for Effective Multiple Image Processing in the Medical Field, *ICCE-Asia*, 2022, **1st author**
- An Efficient Neural Network based on Early Compression of Sparse CT Slice Images, *PlatCon*, 2021, **Co-author**
- An Adaptation of Bayesian Theorem for Passenger Destination Estimation, *IEEE ICOIN*, 2021, **1st author**

In Domestic Conferences

- Interpretable Medical Diagnostic Pipeline using Multi-head Design, *IEIE*, 2022, **1st author**

참고 문헌

- [1] Sarvamangala, D. R., and Raghavendra V. Kulkarni. "Convolutional neural networks in medical image understanding: a survey." *Evolutionary intelligence* (2021): 1-22.
- [2] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).
- [3] Huang, Gao, et al. "Densely connected convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [4] Huang, Huimin, et al. "Unet 3+: A full-scale connected unet for medical image segmentation." *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020.
- [5] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
- [6] Wang, Chien-Yao, et al. "CSPNet: A new backbone that can enhance learning capability of CNN." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2020.
- [7] Tan, Mingxing, and Quoc Le. "Efficientnetv2: Smaller models and faster training." *International Conference on Machine Learning*. PMLR, 2021.
- [8] Lin, Tsung-Yi, et al. "Focal loss for dense object detection." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [9] Loshchilov, Ilya, and Frank Hutter. "Decoupled weight decay regularization." *arXiv preprint arXiv:1711.05101* (2017).
- [10] Demšar, Janez. "Statistical comparisons of classifiers over multiple data sets." *The Journal of Machine learning research* 7 (2006): 1-30.
- [11] Dunn, Olive Jean. "Multiple comparisons among means." *Journal of the American statistical association* 56.293 (1961): 52-64.
- [12] Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." *Proceedings of the European conference on computer vision (ECCV)*. 2018.
- [13] Selvaraju, Ramprasaath R., et al. "Grad-cam: Visual explanations from deep networks via gradient-based localization." *Proceedings of the IEEE international conference on computer vision*. 2017.

Thank you