

# A coupled-GAN architecture to fuse MRI and PET image features for multi-stage classification of Alzheimer's disease

Information Fusion (14.8, Q1)

2025.05.13.

딥러닝 (김대영 교수님)

순천향대학교 소프트웨어융합학과

박사과정 20247447 길태연



# Outline



1. Introduction

2. Literature review

3. Methodology

4. Experiment results

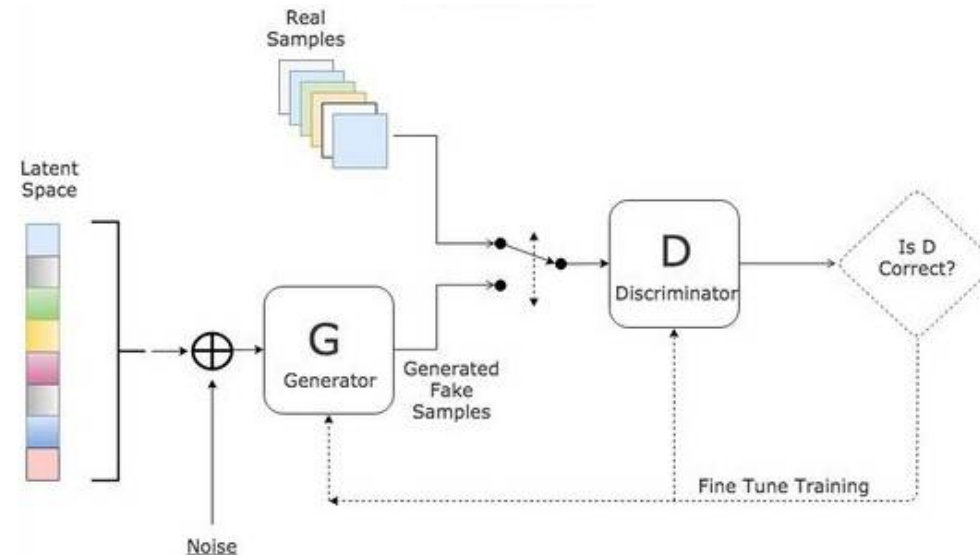
5. Conclusion

# 1. Introduction

- An adversarial learning-based fusion of MRI and PET scans has been proposed for Alzheimer's disease(AD) diagnosis.  
→ 알츠하이머의 진단을 위해 MRI, PET을 융합하는 적대적 학습 기반 방법을 제안한다.
- MRI scans contain structural atrophies, and PET scans consist of metabolic information.  
→ MRI에는 뇌의 구조적 위축을 확인할 수 있고, PET에는 대사 정보가 포함되어 있다.
- We aim to extract structural as well as metabolic information for effective diagnosis of AD at an early stage.  
→ 본 연구에서는 알츠하이머의 조기 진단을 위해 구조적 정보 뿐만 아니라 대사 정보도 추출하는 것을 목표로 한다.
- MRI and PET images are fused using adversarial learning-based coupled-GAN(cGAN) which uses dual convolutional auto encoders and discriminators to fuse both MRI and PET images.  
→ MRI와 PET 영상은 적대적 학습 기반 cGAN을 사용하여 융합된다.

## 2. Literature review - Generative Adversarial Network(GAN)

- 기존 생성모델은 이미지와 같은 고차원 데이터의 분포를 학습하는 데 어려움이 있음. → **GAN** 제안
- **Generator - 생성자**
  - 최대한 실제처럼 보이는 데이터를 생성
  - 판별자를 속이려고 시도
  - 학습이 진행됨에 따라 점점 더 실제와 같은 데이터 생성 가능
- **Discriminator - 판별자**
  - 실제 데이터와 만들어진 가짜 데이터를 구별
  - 학습이 진행됨에 따라 실제와 가짜를 더 잘 구별
- 두 모델이 서로 경쟁하면서 생성자가 판별자를 속이기 위해 점점 더 진짜 같은 데이터를 만들게 됨.



## 2. Literature review - Generative Adversarial Network(GAN)의 활용

- **Conditional GAN**

- GAN에 조건 정보를 추가하여 특정 조건에 맞는 이미지를 생성할 수 있도록 변형한 모델

- **Pix2Pix**

- Conditional GAN과 구조 동일, Image to Image를 위한 모델

- A style 이미지를 입력받아 B style 이미지 생성

- paired training data 필요

- **CycleGAN**

- A style > B style > A style

- 스타일 변환과 이미지 복원이 가능한 모델

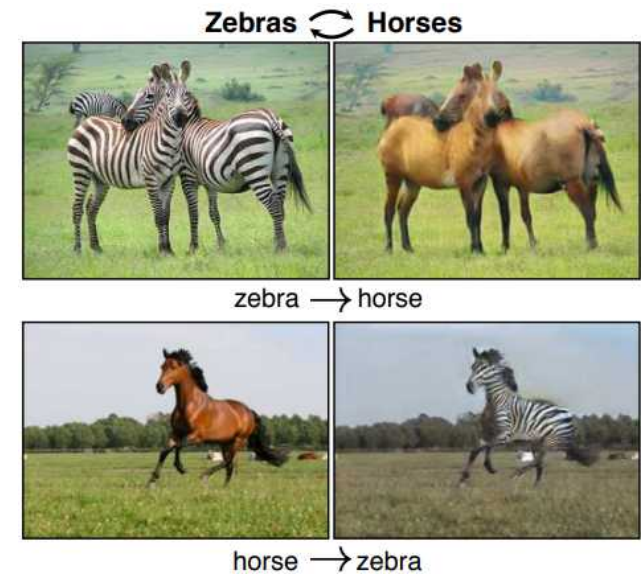
- unpaired training data 대상

➔ 데이터를 통한 데이터 부족 문제 완화

➔ 합성을 통한 데이터 다양성 확보



▲ Pix2Pix의 예시



▲ CycleGAN의 예시

\* Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

\* Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.

### 3. Methodology - Pre-processing

- For a comprehensive automated diagnosis of AD, it is necessary to consider both the structural and metabolic information from the MRI and PET images. → AD의 자동 진단을 위해 MRI와 PET 영상의 구조적 정보와 대사 정보를 모두 고려해야한다.
- Certain pre-processing has been done to make both images suitable for the fusion.  
→ MRI와 PET 영상의 촬영 환경 및 특성이 서로 다르므로 pre-processing을 수행한다.
- After pre-processing, the adversarial network is used for image fusion and classification.  
→ pre-processing 이후 이미지 융합 및 분류를 수행한다.

- Pre-processing 단계**

- 1) Realignment**

- To eliminate motion artifacts.

- 2) Normalization**

- To standardize the intensity values in both scans.

- 3) Registration**

- To ensure that the two modalities fit together accurately.

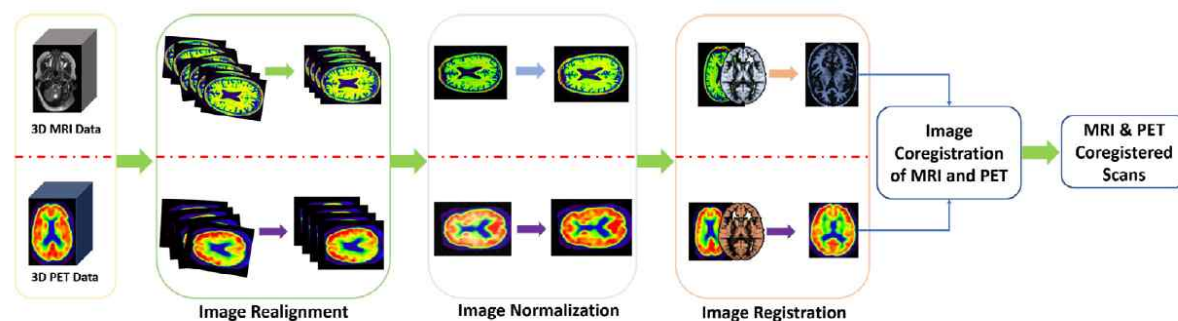
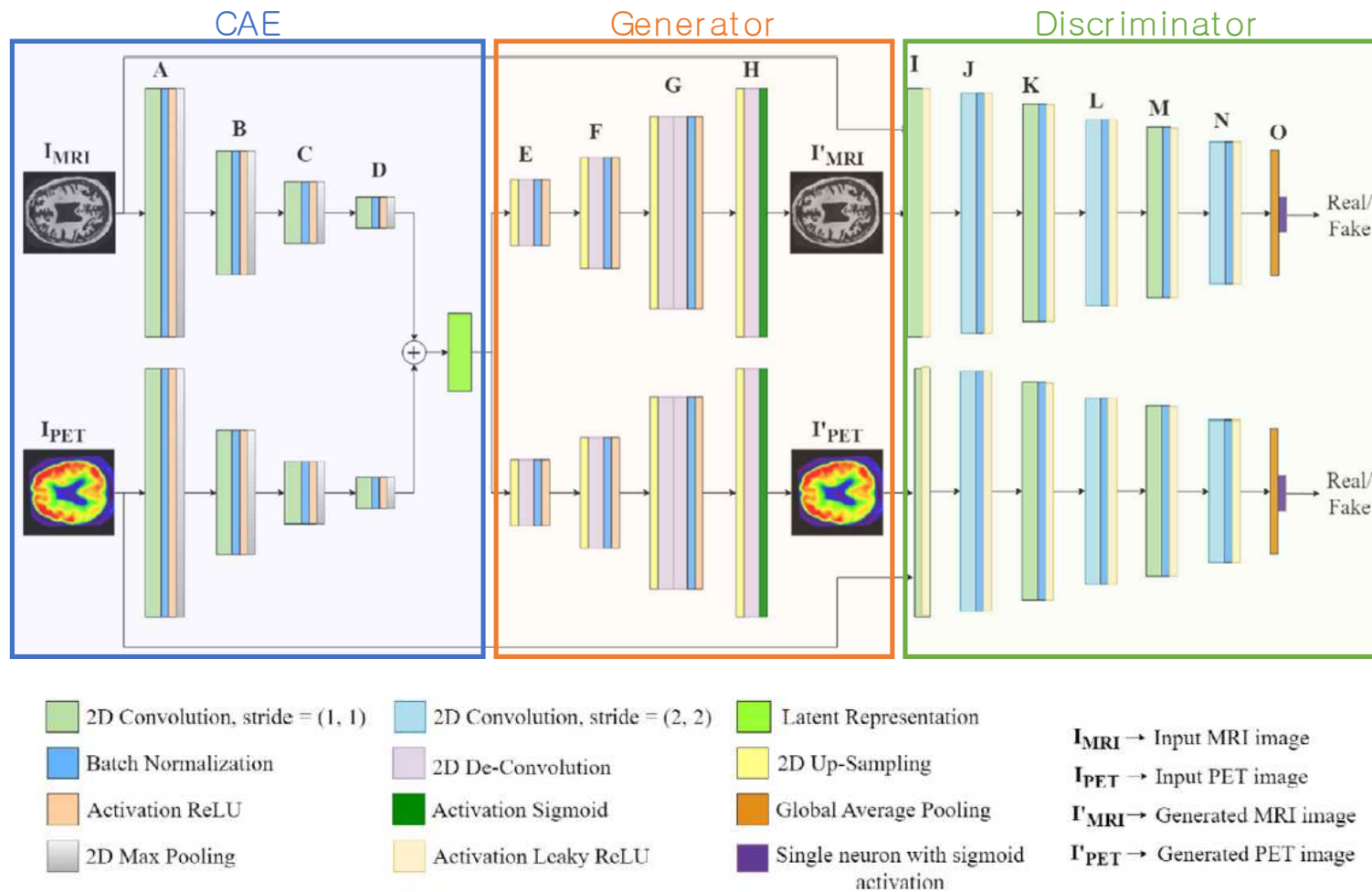


Fig. 1. Preprocessing pipeline of MRI and PET scans.

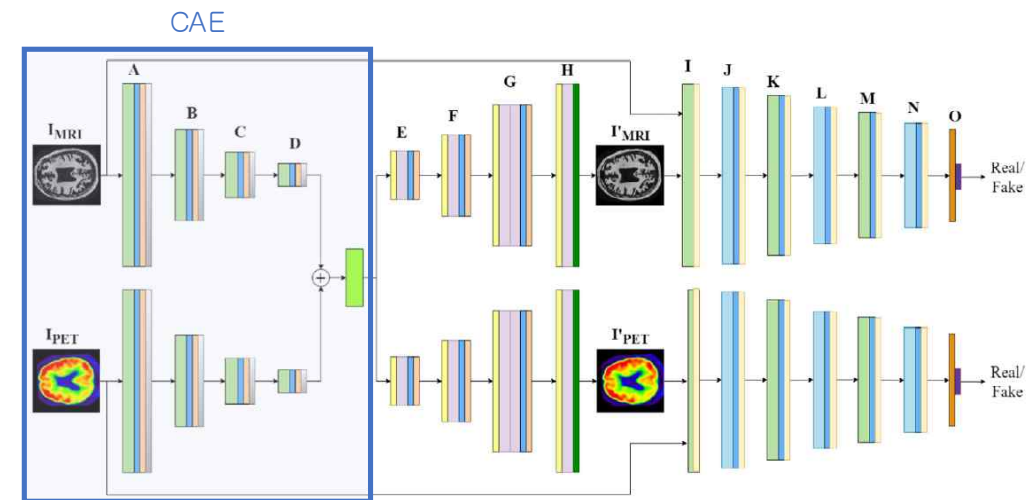
### 3. Methodology - coupled-GAN (cGAN)

- Proposed cGAN architecture



### 3. Methodology - Convolution Auto-Encoders(CAE)

- To fuse the mutually exclusive features from PET and MRI images, we propose to couple two Convolution Auto-Encoders (CAE) at their latent space. → PET과 MRI 영상의 상호 배타적인 특징들을 융합하기 위해, 두 개의 CAE를 잠재 공간에서 결합하는 것을 제안한다.  
(Auto-Encoder: 입력 데이터를 비선형적으로 압축된 잠재 벡터로 인코딩하며, 대부분의 정보를 유지함)
- We propose to encode the MRI and PET image separately into two separate latent vectors.
- These vectors are then added to form a fused latent representation.  
→ PET과 MRI 영상을 각각 두 개의 잠재 벡터로 인코딩한 후 더하여 융합된 잠재 표현을 형성한다.
- The fused latent space should logically contain the information required to reconstruct the MRI as well as the PET image separately. → 융합된 잠재 공간은 각 영상을 재구성하는 데 필요한 정보를 포함한다. (다른 영상 유형에만 포함된 정보도 포함됨)





### 3. Methodology - Generator & Discriminator

#### • Generator

- This generator model will yield two outputs of same shape as the inputs.
  - 생성기 모델은 입력과 동일한 형태의 두 출력을 생성한다. (출력에는 서로 다른 영상 특성도 포함된 상태임)

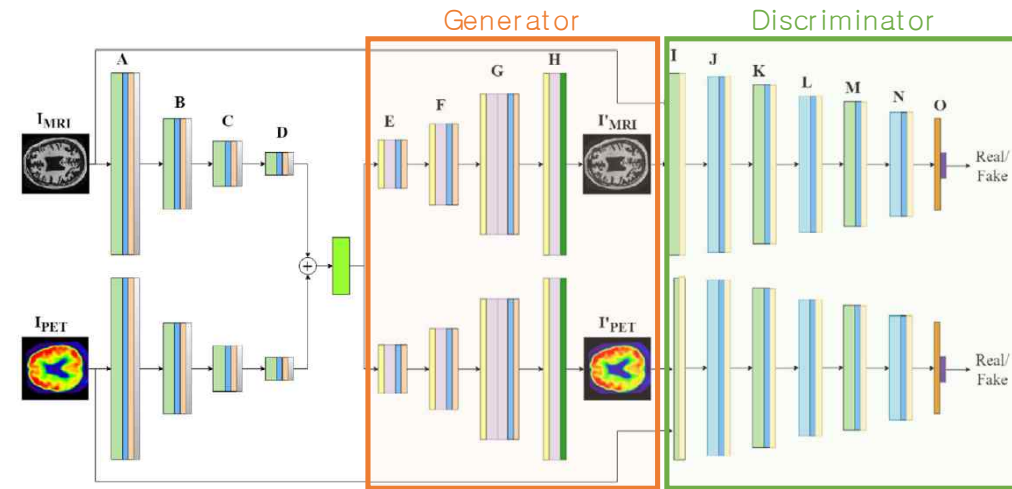
$$Cost_G = \frac{\alpha}{N} \sum_{i=1}^N (\{I_{MRI}\}_i - \{I'_{MRI}\}_i)^2 + (\{I_{PET}\}_i - \{I'_{PET}\}_i)^2 + \frac{(1-\alpha)}{N} \sum_{i=1}^N -\{\log(p_i^1) + \log(p_i^2)\} \quad (1)$$

→ 정합 손실 (MSE 방식)  
 → 분류 손실 (CCE 방식)

#### • Discriminator

- Two separate critic networks, one for each of the generated output images, criticize the output pair from the generator.
  - 생성기의 출력 이미지는 각각의 판별기를 통해 평가된다.

$$Cost_{C_1, C_2} = \frac{1}{N} \sum_{i=1}^N -\{y_i^1 \times \log(C_1(\{I_{MRI}\}_i)) + (1 - y_i^1) \times \log(1 - C_1(\{I'_{MRI}\}_i))\} + \frac{1}{N} \sum_{i=1}^N -\{y_i^2 \times \log(C_2(\{I_{PET}\}_i)) + (1 - y_i^2) \times \log(1 - C_2(\{I'_{PET}\}_i))\} \quad (3)$$



### 3. Methodology - Classifier

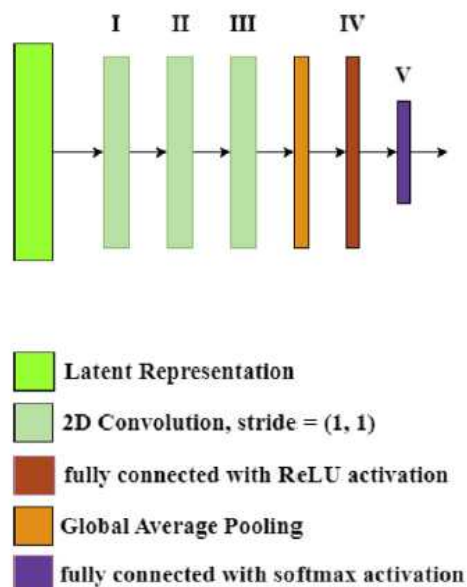


Fig. 3. Proposed CNN architecture for classifier.

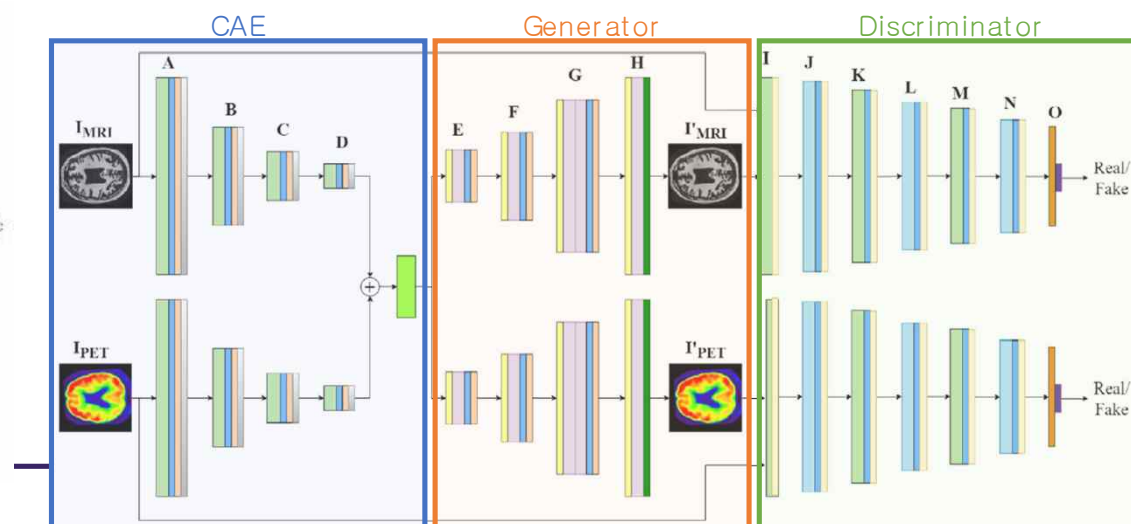
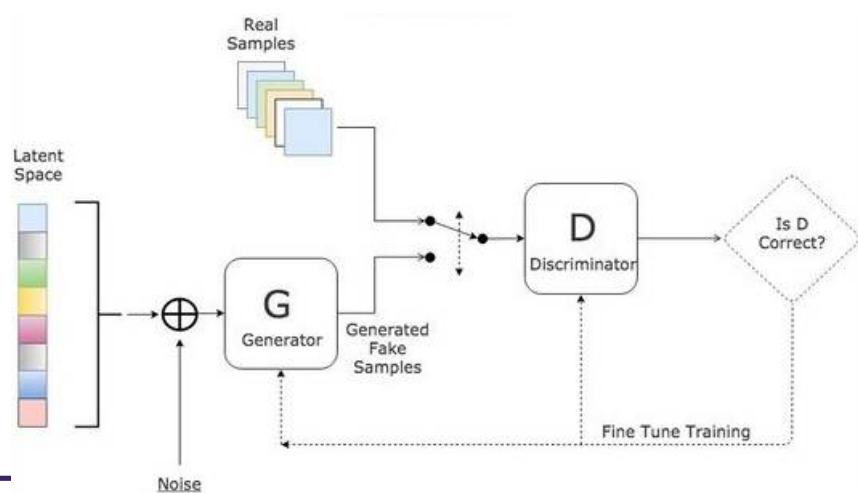
- Based on the latent representation from the above CGAN network, the classification is done.
  - cGAN 네트워크의 잠재 표현을 기반으로 분류가 수행된다.  
(알츠하이머, 경도 인지 장애, 주관적 기억력 저하, 인지 정상)
- For classification a separate convolution neural network (CNN) is built.
  - 분류를 위해 별도의 CNN을 구축하였다.
- For training the classifier the same training data-set is used as for training the CGAN model.
  - 훈련 데이터로는 cGAN 훈련에 사용된 데이터세트와 동일한 것을 사용하였다.
- The cost function used for training this network is categorical cross entropy:
  - CNN의 훈련을 위해 사용한 cost 함수는 범주형 교차 엔트로피이다.

$$Cost_{classifier} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^3 -\{y_i^j \times \log(Class_j(L_i))\} \quad (4)$$

## + 기본 GAN과의 차이점

### • GAN vs cGAN

	GAN	cGAN
입력	단일 모달(ex. Image)	이중 모달(MRI+PET)
모델 구조	단일 Generator, Discriminator	2개의 Generator, Discriminator Fusion module + Classifier
학습 목적	Real/Fake 구분 학습	각 모달리티의 특징 표현을 융합 가능하도록 학습
최종 목표	실제같은 이미지 생성	알츠하이머의 4단계 분류



## 4. Experiment results

- **ANDI Dataset**

- Includes both males and females with a follow-up during the last 18 months.
- Age range of 55 to 90 years.
- One hundred subjects have been chosen from each group—AD, MCI, SMC, and CN to conduct the experiments.
  - ➔ 55~90세 남녀에 대한 18개월 동안의 추적 관찰 데이터
  - ➔ 실험을 위해 알츠하이머(AD), 경도 인지 장애(MCI), 주관적 기억력 저하(SMC), 인지 정상(CN) 총 4개 그룹에서 각 100명의 피험자 선정

- **Implementation details**

- From each class of both MRI and PET images, we have taken 800 image pairs for training purpose and the rest 200 image pairs for testing purpose.
  - ➔ 1,000장의 이미지 데이터 쌍 중 800 쌍은 학습, 200 쌍은 테스트 데이터로 사용하였다.
- We have trained the proposed cGAN for 70 iterations, and the classifier has been trained for 50 iterations.
  - ➔ cGAN은 70번, 분류기는 50번 학습하였다. (조기 종료 ➔ cGAN 59번, 분류기 17번)

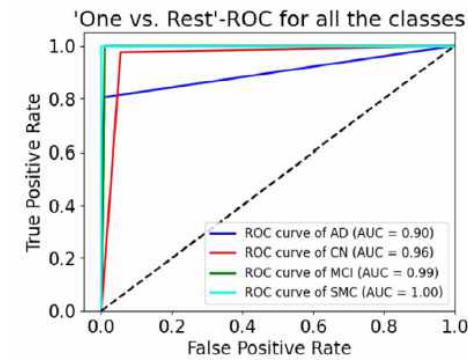
## 4. Experiment results - Performance results

### ◆ Performance of the proposed method

**Table 1**

Performance parameters achieved by the proposed method.

Class	Accuracy	Sensitivity	Specificity	Precision	F1-score
AD	80.5%	81%	99.19%	97%	88%
CN	97.5%	97%	93.5%	86%	91%
MCI	100%	100%	92.67%	97%	99%
SMC	100%	100%	92.67%	100%	100%
Overall	<b>94.5%</b>	<b>95%</b>	<b>94.49%</b>	<b>95%</b>	<b>94%</b>



(a) ROC Curve

	Multimodality			
AD	161	33	6	0
CN	5	195	0	0
MCI	0	0	200	0
SMC	0	0	0	200

(b) Confusion Matrix

**Fig. 4.** Performance of the proposed method.

## 4. Experiment results - Performance results

### ◆ Comparison of single and multi-modality

**Table 2**  
Comparison of single and multi-modality

Modality	Accuracy	Sensitivity	Specificity	Precision	F1-score
MRI	81.88%	82%	81.87%	82.5%	82%
PET	74.13%	74%	74.5%	74.5%	74%
Fused	94.5%	95%	94.49%	95%	94%

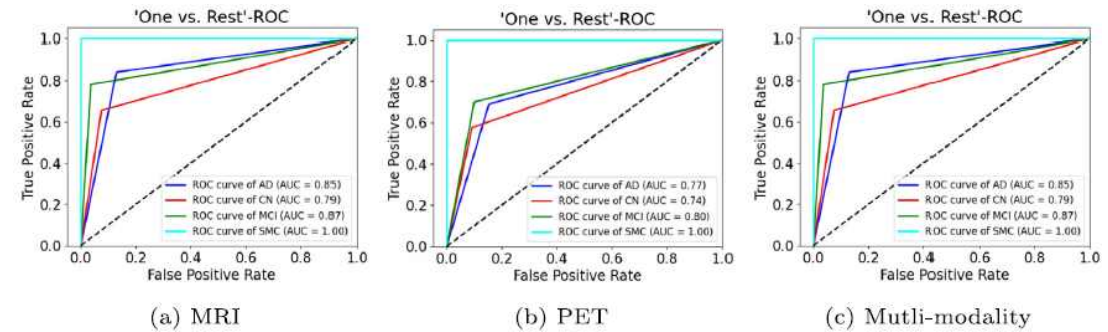


Fig. 5. ROC curve for Single and Multi Modality.

MRI					PET					Multimodality				
AD	168	25	7	0	AD	138	35	27	0	AD	161	33	6	0
CN	54	131	15	0	CN	52	115	33	0	CN	5	195	0	0
MCI	24	20	156	0	MCI	40	20	140	0	MCI	0	0	200	0
SMC	0	0	0	200	SMC	0	0	0	200	SMC	0	0	0	200

(a) MRI

(b) PET

(c) Mutli-modality

Fig. 6. Confusion matrix for Single and Multi Modality.

## 4. Experiment results - Performance results (Cont.)

### ◆ Comparison of different cGAN architecture

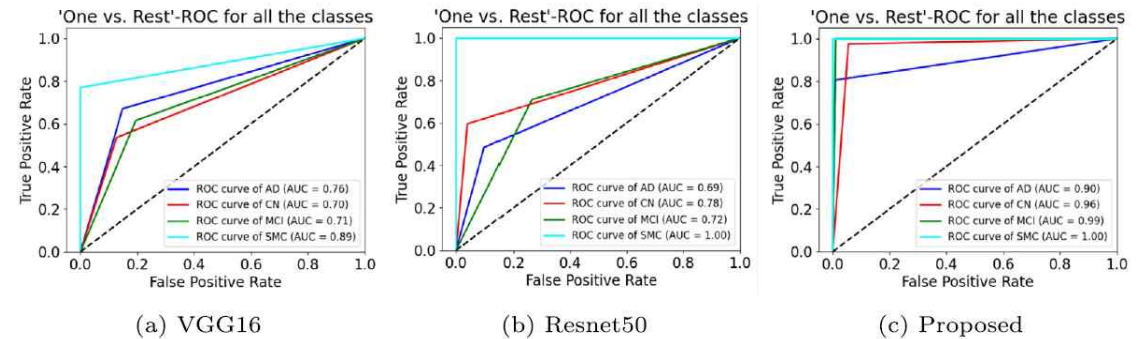
**Table 3**  
Comparison of different CGAN architecture.

Architecture	Accuracy	Sensitivity	Specificity	Precision	F1-score
VGG16	64.75%	65%	64.75%	67%	66%
Resnet50	69.75%	70%	62%	73%	70%
<b>Proposed</b>	<b>94.5%</b>	<b>95%</b>	<b>94.49%</b>	<b>95%</b>	<b>94%</b>

### ◆ Comparison of time complexity

**Table 4**  
Comparison of time complexity of different fusion techniques.

Architecture	fps	Test time/ image (s)	Accuracy
VGG16	84.33	0.0199	64.75%
Resnet50	97.23	0.0103	69.75%
<b>Proposed</b>	<b>137.79</b>	<b>0.0073</b>	<b>94.5%</b>



**Fig. 7.** ROC curve for different architecture.

<b>VGG16</b>					<b>Resnet50</b>					<b>Multimodality</b>				
AD	134	14	52	0	AD	97	11	92	0	AD	161	33	6	0
CN	32	107	61	0	CN	14	119	67	0	CN	5	195	0	0
MCI	43	34	123	0	MCI	45	13	142	0	MCI	0	0	200	0
SMC	14	28	4	154	SMC	0	0	0	200	SMC	0	0	0	200

(a) VGG16

(b) Resnet50

(c) Proposed

**Fig. 8.** Confusion matrix for different architectures.



## 5. Conclusion

- AD includes structural, functional, metabolite, and chemical changes in the brain of the affected person.  
→ 알츠하이머는 뇌의 구조적, 기능적, 대사적, 화학적 변화를 포함한다.
- Diagnosis accuracy will increase if both the structural as well as metabolic alterations in the brain are considered.  
→ 따라서, 뇌의 구조적 변화와 대사적 변화를 모두 고려할 경우 진단의 정확도를 향상시킬 수 있다.
- The idea of this work is to fuse the MRI and PET images such that their feature set complement each other for classification of different stages of AD. → 본 연구는 MRI와 PET 이미지 융합을 통해 서로의 특성을 보완하여 여러 단계의 알츠하이머를 분류한다.
- We demonstrated that the extracted fused feature set enables more accurate classification of AD stages than using features from only MRI or PET images.  
→ 추출 및 융합된 특징을 사용하면 각 이미지의 단일 특징만을 사용하는 것보다 더 정확하게 분류할 수 있다는 것을 확인하였다.
- In the future the cost function of the CGAN can be extended to accommodate constraints on the latent space representation so that the proposed approach can be used for classification of data with larger number of classes.  
→ 향후 cGAN의 비용 함수를 확장을 통해 제안 방법을 더 많은 클래스 분류에 활용할 수 있을 것이다.  
→ 나아가, 분할, 기능적 MRI, 확산 텐서 영상 등과 같은 또 다른 모달리티를 추가한다면 더욱 정밀한 알츠하이머 진단이 가능할 것이다.



감사합니다