|  |
| --- |
| **SIA-GAN: ECG Signal과 STFT를 통해**  **이미지화 된 ECG를 이용한 GAN 기반의**  **multimodal 심부전 탐지 알고리즘**  **요 약**  현재 병원에서는 심장질환 환자의 심전도 데이터를 장시간 측정하고, 결과를 심장내과 전문의가 직접 확인 후 부정맥을 진단한다. 따라서 진단에 많은 시간이 소요되며, 굉장히 피로한 작업이다. 이러한 문제점을 해결하고자 자동으로 심부전을 탐지하는 딥러닝 알고리즘 연구가 활발히 진행되고 있지만, 기존 연구는 대부분 signal 데이터만을 이용했기 때문에 주파수 등의 다른 유용한 정보들을 학습할 수 없다. 따라서 우리는 STFT 변환을 통해 심전도 시그널을 이미지화 시킨 데이터와 signal 데이터를 함께 활용해 서로에게 긍정적인 영향을 줄 수 있도록 학습한다. 또한 의료 데이터에서는 정상 데이터와 비정상 데이터 사이에 개수 불균형이 존재하기 때문에 GAN 기반의 CNN 모델을 구성함으로써 정상데이터 만으로도 학습을 할 수 있도록 설계한다. 더불어 Transformer의 attention 메커니즘을 활용함으로써 signal과 STFT 이미지 사이에서 feature fusion이 잘 될 수 있도록 한다. 실험결과 우리는 AUC 0.9588의 높은 정확도를 달성하였다. |

**1. 서론**

**1.1. 연구배경**

현재 심부전 데이터 분석에는 심장내과 전문의가 24시간 동안 측정된 심전도 데이터를 모두 확인하는 과정이 필요하다. 이렇게 장시간 동안 측정된 데이터를 분석하는 것은 상당히 부담이 되는 작업이다.

이를 해소하기 위해 많은 연구들이 진행되고 있지만, 아직 문제점이 존재한다. 대표적으로 심부전 환자의 ECG 데이터가 부족하기 때문에 비정상 데이터가 충분치 않다는 점에 있다. 따라서 본 연구에서는 GAN을 활용하여 정상 데이터만 학습시켜도 비정상을 탐지할 수 있도록 한다.

또한 기존의 연구들에서는 ECG signal만을 사용해 비정상을 탐지했기 때문에 2차원 데이터에서 얻을 수 있는 색상, 주파수 등의 정보를 활용할 수 없었다. 따라서 우리는 지난 연구에서 STFT 변환을 통해 ECG signal을 2D로 변환해 비정상을 탐지하는 B2I-GAN (Beat To Image GAN) 모델을 고안했고, [표1]과 같이 frequency 정보가 anomaly detection에 긍정적인 효과를 미친다는 것을 밝힌 바 있다.

[표 1] B2I-GAN (our model) 및 다른 모델 성능

|  |  |  |
| --- | --- | --- |
| 모델 | AUC | AP |
| AE | 0.8944 | 0.8415 |
| AnoGAN | 0.8642 | 0.8035 |
| Ganomaly | 0.9083 | 0.8701 |
| BeatGAN | 0.9447 | 0.9143 |
| ***B2I-GAN*** | ***0.9460*** | ***0.9058*** |

하지만 STFT로 변환된 이미지만을 사용하면 사람이 변환된 이미지를 해석하기 난해하고, encoder와 decoder에 모두 2D cnn을 사용하기 때문에 학습에 오랜 시간이 걸린다는 한계점이 있었다. 따라서 본 연구에서는 STFT 변환을 통해 signal 정보와 frequency 정보를 함께 활용하여 각 모달리티의 단점을 보완하고 장점을 극대화 시키고자 한다.

최근 transformer를 활용해 모달리티를 fusion하는 연구가 많이 진행되었고 효과적이라는 것이 밝혀졌기 때문에 우리는 transformer를 활용해 feature를 fusion하는 다양한 연구를 진행하였다. 그리고 이를 토대로 실험을 진행한 결과 우리 SIA-GAN 모델은 epoch=100에 대해 AUC 0.9588라는 우수한 정확도를 가짐을 확인할 수 있었다.

**1.2. 연구목표**

첫째, 우리는 signal과 STFT 이미지를 동시에 사용함으로써 부정맥을 자동으로 진단하는 딥 러닝 모델을 구축하여 높은 정확도를 가진 부정맥 진단 시스템을 만들고자 한다.

둘째, 우리는 transformer를 적절히 사용함으로써 각 모달리티의 feature가 잘 섞일 수 있는 방안을 탐구하여 성능이 뛰어나고 효율적인 multimodal transformer를 제안한다.

**2. 관련 연구**

**2.1. Disease Detection**

ECG데이터와 딥 러닝 모델을 이용해 질병 분석한 대표적인 연구는 cardiac arrhythmias 감지[4]와 ECG 데이터의 ST변화를 모니터링하는 연구가 있다[5]. 또한 convolutional neural network를 통해 MI (myocardial infarction), CHF (congestive heart failure) 검출을 자동화[6][7]하는 연구도 진행되었다.

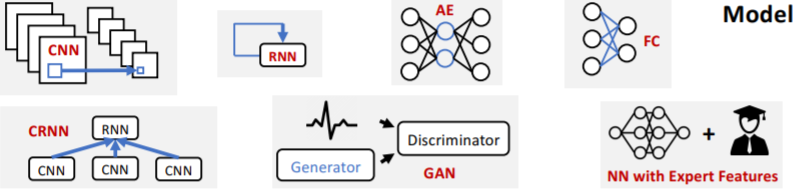
**2.2. Localization and Annotation**

Localization과 annotation은 심장병 전문의에게 매우 중요한 연구이다. 질병을 진단하는데 이러한 정보들을 사용하기 때문이다. 이와 관련된 연구로는AE, RNN을 사용하여 12-lead ECG data로부터 심실 급속증의 출구를 자동으로 localization한 연구[8][9]가 있다. 또한 MITBIH arrhythmia dataset을 이용해 QRS 복합체에 주석을 다는 연구와[10], PhysioNet의 QT database를 이용해 ECG 파장에 주석을 다는 색다른 방법을 탐구(P,T wave annotation)하는 연구도 진행된 바 있다[11].

**2.3. Denoising**

ECG 신호를 수집하는 과정에서는 많은 양의 노이즈가 발생하게 된다. 이는 질병 진단의 정확도를 떨어뜨려 부정적인 결과를 초래하며, 특히 원격 진료 환경에서는 더욱 부정적인 영향을 미친다. 이러한 노이즈를 제거하기 위해서 Encoder-decoder CNN을 적용한 연구와[12] GAN의 생성자와 판별자 사이의 minimax game을 통해 ECG 노이즈 분포 지식을 축적하는 연구가 진행되었다[13].

**2.4. 딥러닝 모델**



[그림 1] ECG 데이터 분석을 위한 딥러닝 모델 연구 분야

ECG 데이터 분석에 특화된 딥러닝 모델연구도 많이 존재한다. 대표적으로 ECG classification을 위해 1D CNN과 2D CNN을 사용한 연구가 있으며[11], Vanishing gradient를 처리하기 위해 GRU/LSTM, bidirectional-LSTM을 사용한 연구도 존재한다. DeepHeart는 CRNN 프레임워크를 사용해 심혈관 위험 예측을 하고[12], DAE[13], SAE[14][15], CAE[16] 스택을 이용해 ECG denosing하는 연구도 있다[13].

**2.5.** **GAN 기반의 Anomaly Detection**

최근 GAN에 대한 관심이 높아지면서 우리 모델과 같이 GAN을 활용한 비정상 탐지 연구도 제안되었다. 대표적으로GAN을 활용해 이미지에서 비정상을 감지하는 AnoGAN [1] 모델, encoder-decoder- encoder구조로 잠재표현을 비교해 이미지에서 비정상을 감지하는 Ganomaly [2] 모델, auto-encoder와 GAN을 결합해 1차원 ECG 신호에 대해 심부전을 탐지하는 BeatGAN [3] 모델이 존재한다.

**2.6. STFT 기반의 ECG 데이터 분석 연구**

STFT를 활용해 ECG 데이터를 분석하려는 시도도 존재했다. 대표적으로 ECG 신호에 STFT를 적용해 이미지로 변환한 후 특징을 추출함으로써 심방 세동을 검출하는 알고리즘을 고안한 연구와[20]. 부정맥 환자의 STFT 스펙트로그램 이미지를 2D-CNN의 입력으로 활용하여 심전도를 정상과 비정상으로 분류하는 연구가 존재한다[21]

**3. 프로젝트 진행 과정**

**3.1. 1D->2D 변환을 위한 이미지화 알고리즘 선정**

프로젝트 진행을 위해서는 데이터를 2차원으로 변환할 수 있는 다양한 이미지화 알고리즘 가운데 하나를 선택해야 한다. 이러한 역할을 하는 이미지화 알고리즘은 대표적으로STFT, GAFs, 2D-FFT가 있다. 각각의 알고리즘이 나타내는 의미는 아래와 같다.

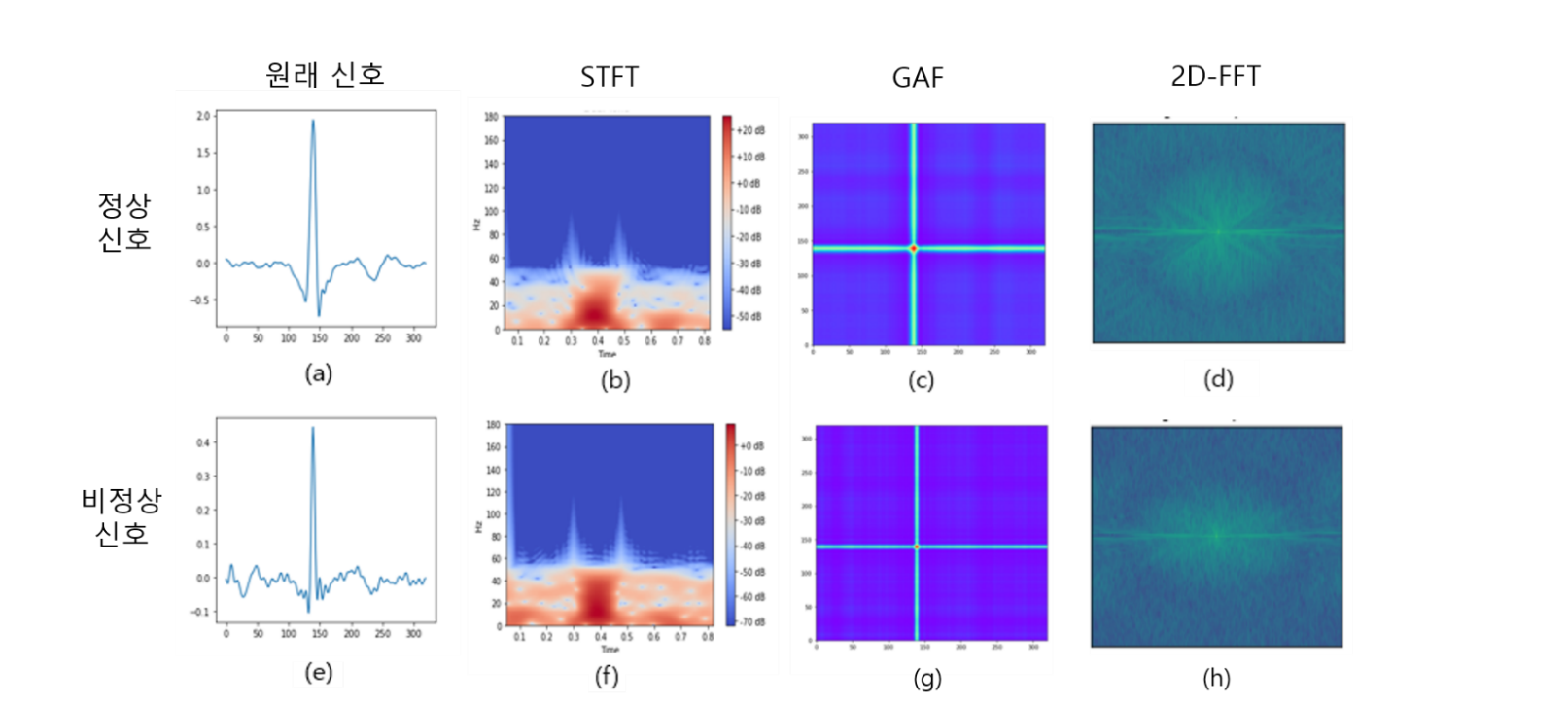
- STFT: Short-Time Fourier Transform의 약자로, 신호를 윈도우 길이에 따라 분리시켜 각 구간에 대해 푸리에 변환을 진행하는 방법이다. 그 결과 시간에 따른 주파수 변화를 나타내는 spectrogram을 생성하게 된다.

- GAFs: Gramian Angular Field의 약자로 각도에 대한 시간 상관 행렬을 만드는 알고리즘이다. 그 결과 각 시점 간의 시간적인 상관 관계를 극좌표를 기반으로 표현할 수 있게 된다.

- 2D-FFT: 2D Fast Fourier Transform의 약자로 x축, y축으로 FFT를 수행하고 스케일링한 뒤 magnitude를 추출함으로써 입력 데이터를 주파수 영역으로 변환하는 알고리즘이다.

우리는 이 세 가지의 알고리즘을 python 코드로 작성하여 이미지화 결과를 비교하고, 연구에 사용할 최적의 이미지화 알고리즘을 선정하였다.

**3.2. 이미지화 알고리즘 선정**

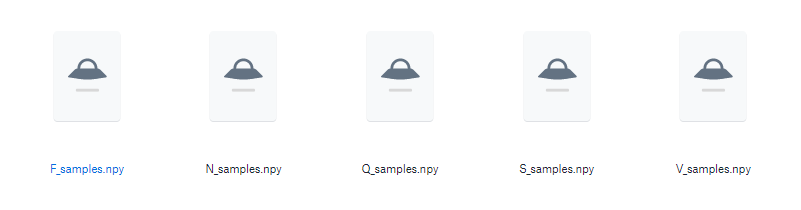


[그림 2] 원본 데이터와 2D 데이터로 변환한 결과 비교

[그림 2]는 정상인의 ECG 데이터와 심부전을 가진 환자의 ECG 데이터에 대해 각각의 이미지화 알고리즘을 적용한 결과이다. 이 결과에 대한 분석은 아래와 같다.

2D-FFT의 경우 코드 상에서 ECG 신호를 이미지로 저장한 뒤 그것을 불러온 뒤에 2D-FFT를 수행해야 하는 번거로움을 가지며, (d)와 (h)에서 보이는 것처럼 정상과 비정상 데이터에 대해 큰 차이를 보이지 못했다. GAFs의 경우 같은 라벨을 가진 데이터 내에서도 제 각각의 패턴을 보였다. 또 [그림 2]에서 보이는 것과 같이 (g)와 같은 비정상 데이터도 정상 데이터인 (c)와 비슷한 형상을 띠기 때문에 학습이 어려울 것으로 판단된다. 무엇보다 이 두가지 알고리즘은 사람이 인식하기에 난해한 모습을 보인다.

반면 STFT의 경우 비정상 데이터 (f)는 정상 데이터 (b)에 비해 양 사이드의 주파수 분포가 넓게 퍼져 있어 그 차이가 두드러지게 나타난다. 정상, 비정상 데이터 대부분에서 이렇게 사람이 보더라도 어느정도 차이점이 있다고 느낄 만큼의 패턴 구분이 있었다. 또한 STFT를 이용해 심부전 종류를 분류하는 한 연구[21]에서 STFT의 효과를 입증하기도 했다. 또한 x축이 시간, y축이 주파수이기 때문에 사람이 인식하기에도 친숙하다는 장점이 있다. 따라서 우리는 STFT 변환 알고리즘을 이용해 이미지화를 진행하기로 결정했다.

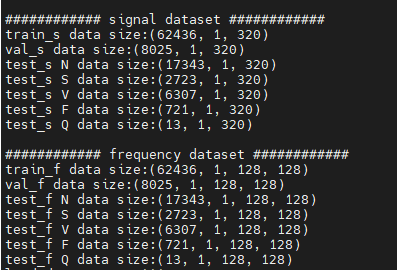
**3.3. Dataset**

[그림 3] signal 데이터

텍스트, 실내이(가) 표시된 사진

자동 생성된 설명

[그림 4] freq 데이터



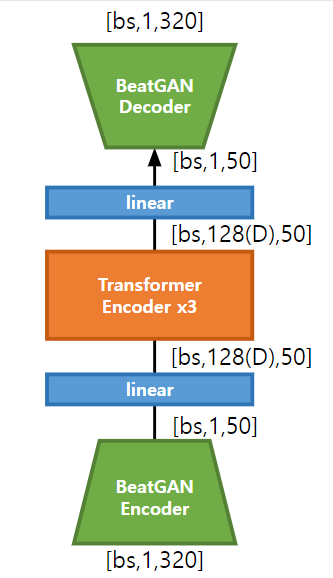
[그림 5] dataset

우리 연구의 기반 dataset은 MIT-BIH Arrhythmia Database로 선정하였다. 그 이유는 다양한 종류의 심부전증을 나타내는 데이터를 포함하고 있으며 ECG 딥러닝 분야의 많은 연구에서 사용되어 왔기 때문에 신뢰성이 높기 때문이다.

이 dataset에는 비트 단위로 나누어진 ECG 시그널이 F, N, Q, S, V의 총 5개 클래스로 분류되어 있다. 여기서 N은 normal beat, V는 VEBs 환자의 beat, S는 SVE 환자의 beat, Q는 unknown beats, F는 fusion beats를 나타낸다. 각 클래스별 데이터의 개수는 N데이터 86717개, S데이터 3026개, V데이터 802개, Q데이터 15개로 구성돼 있다. 우리는 이 dataset을 signal dataset으로 활용했다.

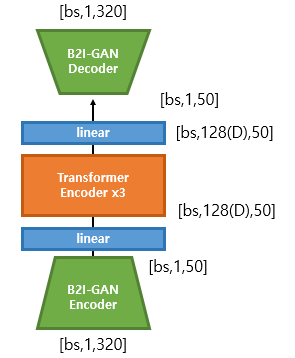
그리고 freq dataset을 만들기 위해 이 dataset에 존재하는 모든 beat에 대해 STFT 변환을 거쳐 그 결과를 [그림 4]와 같이 numpy배열로 저장해 학습을 진행하였다. 데이터 개수는 위와 동일하며, 한 stft 이미지 당 1x128x128의 크기를 갖는다.

**3.4. Transformer 추가 성능 테스트**



[그림 6] BeatGAN + Transformer 모델

Transformer를 사용하는 것이 효과적인 지 확인하기 위해 기존 GAN 기반 심부전 탐지 모델 중 AUC가 가장 높았던 BeatGAN의 latent에 Transformer를 추가해 self-attention을 진행하여 성능을 확인했다. 그 결과 우리가 BeatGAN을 직접 돌려봤을 때 AUC는 0.9404였는데, Transformer를 추가하니 AUC가 0.9449로 향상했다. BeatGAN 논문에서 나온 AUC는 0.9446이었는데, Transformer를 추가한 결과는 이 보다 더 높은 성능을 보임을 볼 수 있다.

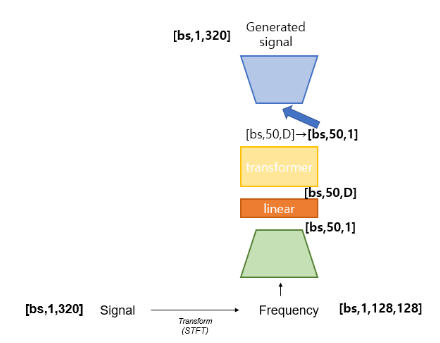


[그림 7] B2I-GAN + Transformer 추가

마찬가지로, B2I-GAN의 latent에도 transformer를 추가해 성능을 확인했다. 원래 B2I-GAN의 AUC는 0.9460이었는데, transformer를 추가하니 0.9517로 상당한 성능향상이 있었다.

이렇게 transformer를 추가하는 것이 합당하다는 것을 실험을 통해 밝혔다. 하지만 B2I-GAN의 모델을 그대로 사용하면 generator에 2D encoder와 2D decoder를 사용하기 때문에 학습시간이 오래걸리고 STFT 이미지를 생성하기 때문에 signal에 비해 분석이 어렵다는 단점이 여전히 존재하게 된다.

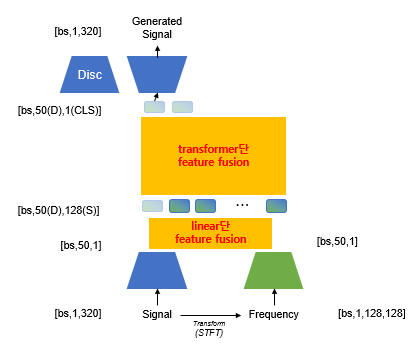
그래서 학습시간을 단축하고 분석에 용이하도록 encoder는 2d 그대로, decoder는 signal decoder와 같은 구조의 1d decoder를 사용해 freq에서 signal을 사용하면 어떨까? 라는 아이디어가 생각나 실험을 진행했다.



[그림 8] freq -> transformer -> signal

실험 결과, AUC는 0.9269로 정확도는 다소 낮아졌으나 학습시간이 절반 가량으로 줄어들고, signal을 생성하기 때문에 분석에 용이하다는 큰 장점이 생겼기 때문에 이 모델과 BeatGAN을 토대로 feature를 잘 fusion한다면 좋은 결과를 낼 수 있을 것이라고 판단했기 때문에 이 모델을 활용하기로 결정했다.

**3.5. 모델**



[그림 9] 기본 모델

기본적인 모델 구조는 위와 같다. ECG signal는 generator의 1D encoder에 넣어 [bs,50,1] 크기의 feature를 추출한다. 또한 ECG signal을 STFT로 변환한 이미지를 generator의 2D encoder를 통과시켜 똑같이 [bs,50,1] 크기의 feature를 추출한다. 그리고 이 과정에서 signal과 freq의 feature 사이의 distribution이 매우 다를 수 있기 때문에, linear layer를 통해 먼저 feature를 fusion 시킴으로써 distribution을 matching한다.

그리고 그 결과를 transformer에 넣어 attention을 진행함으로써 두 모달리티의 feature가 더 잘 fusion될 수 있도록 한다. 그리고 CLS token을 얻어 signal을 generation한다. 그리고 discriminator에서 real signa(1)인지 generator를 통해 만들어진 signal(0)인 지 판별한다.

Train 단계에서는 정상 ECG 데이터로만 학습을 하기 때문에 이 과정을 통해서 generator와 discriminator가 서로 경쟁적으로 학습하면서 모델이 정상 ECG의 특징을 학습하게 되고, test 단계에서는 모델을 통과해 만들어진 정상 ecg에 가까운 fake signal과 input signal 사이의 차이를 통해 이상을 탐지한다.

**3.6. LOSS**

**Generator Loss:**

1. **Signal Loss**
2. **Freq Loss**

: adversarial regularization의 효과를 조정하는 weighting parameter

: signal data

: signal과 매치되는 frequency data

: signal을 discriminator에 넣은 값

: frequency로 생성된 fake signal을 discriminator에 넣은 값

Signal loss의 경우 좌측항은 real signal과 real signal을 generator에 넣어서 얻은 fake signal 사이의 차이를 의미한다. Real 과 fake 사이의 차이가 적으면 fake signal을 진짜에 가깝게 잘 생성했다는 뜻이므로 이 값은 작을수록 좋다.

그리고 우측항은 real signal을 판별기가 판별한 값과 fake signal을 판별기가 판별한 값의 차이인데, 이것도 마찬가지로 real 과 fake의 판별값의 차이가 적으면 fake 이미지가 real과 비슷하다는 것을 의미하므로 우측항도 작을수록 좋다.

따라서 전체 식이 작아지는 방향으로 학습을 진행하면 좋은 generator를 얻을 수 있다.

Freq loss의 경우에도 마찬가지인데, 다른 점은 fake signal 자리에다가 real signal을 stft로 변환한 후에 generator에 입력하여 만들어진 fake signal을 사용한다는 것이다.

**Discriminator Loss:**

1. **Signal Loss**
2. **Freq Loss**

: i 번 째 signal data

: i 번 째 signal data를 입력으로 generation된 signal data

: i 번 째 signal data에 대응하는 frequency data

: i 번 째 signal data에 대응하는 frequency data를 입력으로 generation된 signal data

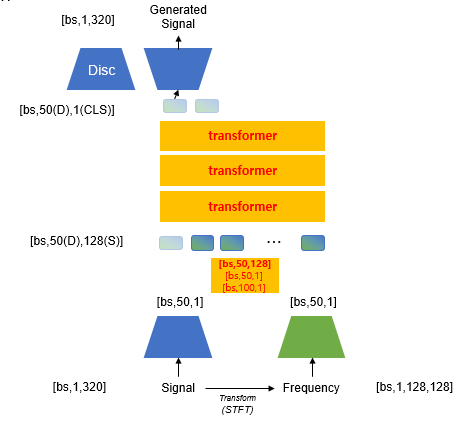
Discriminator loss는 다음과 같다. Signal loss에서 좌측 항은 real signal을 discriminator에 넣은 값을 의미하는데 real signal이면 1으로 판별하고 fake signal이면 0으로 판별해야 좋은 discriminator이므로 이 값은 높을수록 좋다. 우측 항의 는 fake signal을 discriminator에 넣은 값이므로 0으로 판별해야 하니까 작을수록 좋다. 반대로 는 클수록 좋기 때문에 우측항은 높은 값을 가질수록 좋다. 결국 전체 식이 높아지는 방향으로 학습하면 좋은 discriminator를 얻을 수 있다.

**3.7. 실험**

**3.7.1 transformer 구성**

우리는 Transformer 구조를 결정하기 위한 실험을 진행했다.

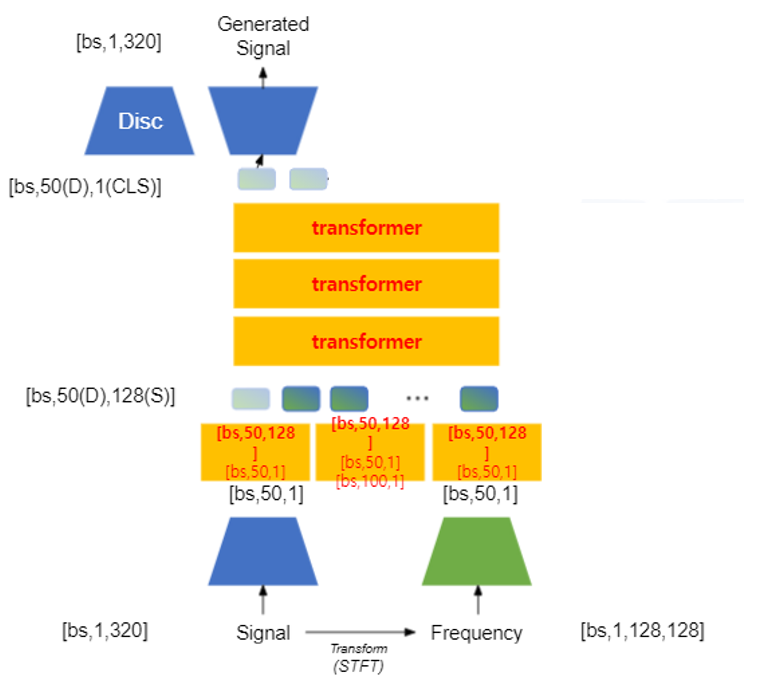
**3.7.1.1 SELF-ATTENTION**



[그림 10] self-attention

우선 임의로 Linear fusion을 진행한 후 그 결과를 transformer에 입력해 self-attention을 진행하였다. 그 결과 AUC는 0.9529로 많은 정확도 상승이 있었다. 이는 BeatGAN의 AUC인 0.9404(\*0.9446)과 B2I-GAN의 AUC인 0.9460보다 훨씬 높은 결과임을 확인할 수 있다. BeatGAN의 괄호친 AUC는 BeatGAN 논문에 나온 AUC를 뜻한다.

**3.7.1.2. QKV 조정**

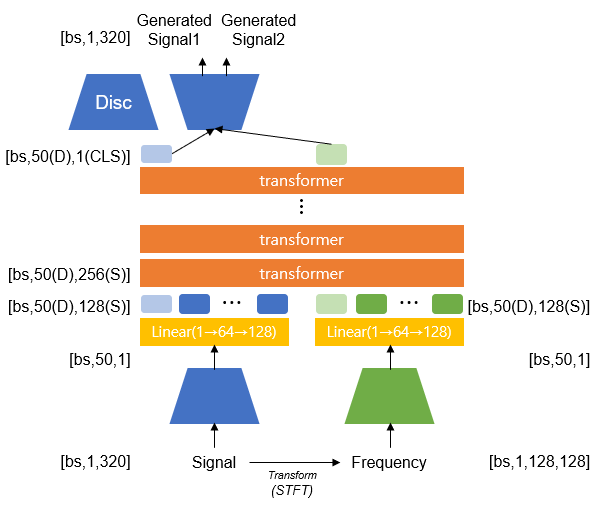


|  |  |  |  |
| --- | --- | --- | --- |
| Layer | Q | K | V |
| **1** | **SF** | **SF** | **SF** |
| **2** | **S** | **SF** | **SF** |
| **3** | **F** | **SF** | **SF** |
| **4** | **SF** | **SF** | **SF** |

[그림 10] Transformer의 QKV 변경

두번째로 시도한 방법은, transformer에서 self-attention이 아닌 [그림 10]처럼 transformer의 각 레이어 마다 쿼리, 키, 밸류를 다르게 해attention을 진행하는 방법이다. 결과는 0.9417의 AUC로, self-attention 방법보다 낮은 성능을 보였다.

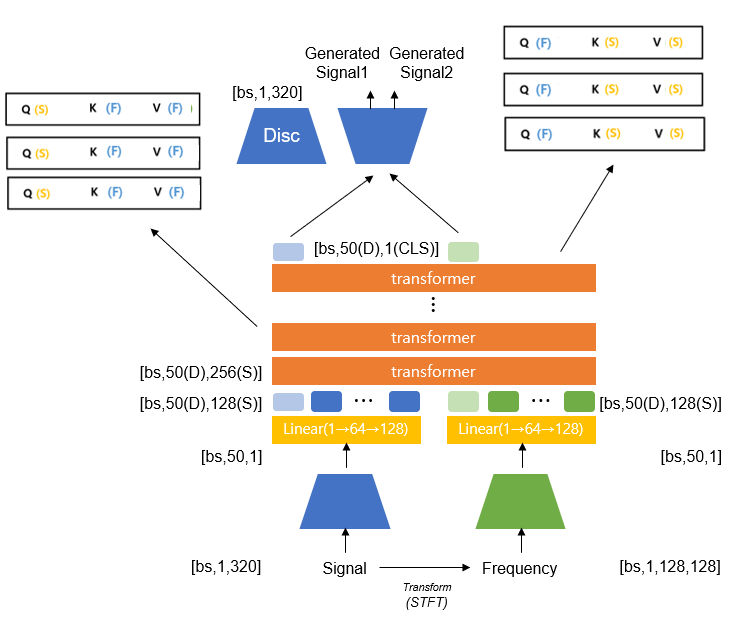
**3.7.1.3. 반 씩 나눠서 attention**



[그림 11] 반 씩 나눠서 attention

보통 multimodal에서 feature fusion을 할 때는 transformer를 두개 이상 사용하는데, 그렇게 되면 파라미터가 많아지고 학습시간이 오래 걸린다. 이를 해결하기 위한 방법으로 우리는 한 transformer에서 구간을 둘로 나눠서 왼쪽 반은 signal을 attention하고 오른쪽 반은 freq를 attention 하는 방법을 고안해 실험했다. AUC는 0.9256으로 성능향상을 이뤄내진 못했다.

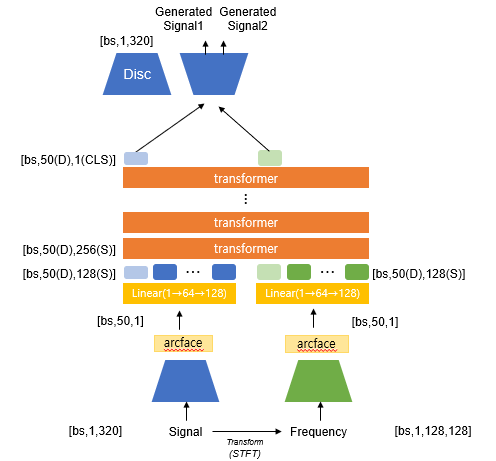
**3.7.1.3. 반 씩 나눠서 QKV 변경하여 attention**



[그림 12] 반 씩 나눠서 QKV 변경해서 attention

[그림 11]에서 self-attention이 아닌, 왼쪽 반은 signal을 쿼리로, freq를 키, 밸류로 설정해 signal에 freq 정보를 추가하는 느낌으로 attention을 진행하고, 오른쪽 반은 freq를 쿼리로, signal을 키,밸류로 설정해 freq에 signal 정보를 추가하는 느낌으로 attention을 진행했다. 그 결과 AUC는 0.9376으로 좋은 결과를 보이진 못했다.

**3.7.1.5. arcface loss 적용**



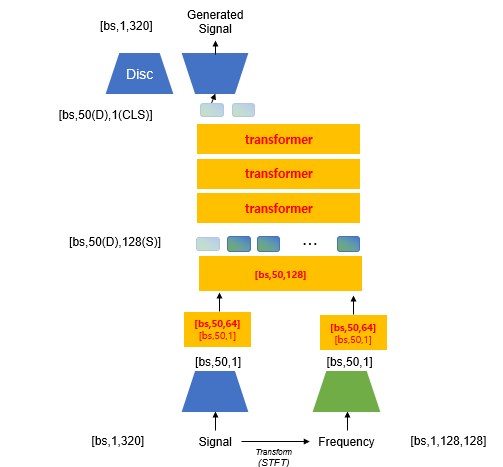
[그림 12] arcface loss 적용

Latent에 arcface를 적용해 freq와 signal의 latent 사이의 거리를 벌려 서로 다른 형태의 signal을 generation한다면 서로에게 더 의미있는 정보를 전달해줄 수 있지 않을까 라는 아이디어에 기반하여 latent에 arcface loss를 적용하였다. 이 실험의 결과 AUC는 0.9219에 그쳐 아쉽게도 성공적인 결과를 내지는 못하였다.

우리는 transformer 실험 결과에서 linear fusion 결과를 바로 self-attention하는 방법이 가장 좋은 성능을 내는 것을 확인했기 때문에 transformer 구조는 이 방법으로 결정하고 linear fusion 방식에 대한 실험을 이어서 했다.

**3.7.2 linear fusion**

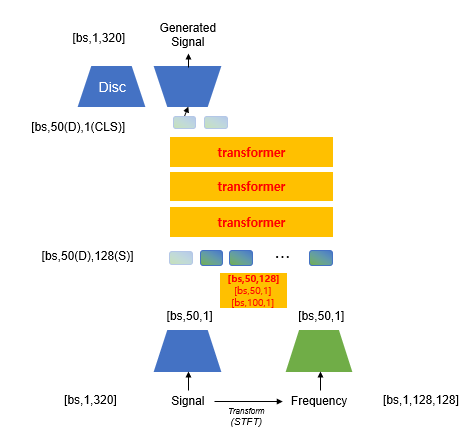
**3.7.2.1. sequence를 늘린 후에 sequence를 concat**



[그림 14] sequence를 늘리고 sequence를 concat

Signal과 freq encoder를 통과해서 나온 feature에서 linear를 통해 seqence를 64로 증가시키고 각 feature를 sequence기준으로 concat 하는 방법으로 [bs,50,128]의 형태를 만들어 transformer를 통과시키고 CLS token을 얻어 signal을 generation 했다. 이 방법의 AUC는 0.9399로 다소 아쉬운 성능을 보였다.

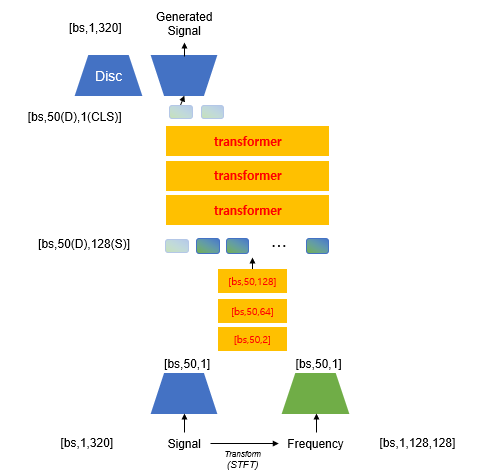
**3.7.2.2. dimension을 concat한 후에 sequence를 증가**



[그림 15] dimension을 concat한 후 sequence를 증가

[그림 14]에서는 sequence만을 이용했는데, dimension도 함께 이용한다면 성능 향상이 있지 않을까 라는 가설을 세우고 실험을 진행했다. 각 feature를 dimension기준으로 concat해 100으로 늘린 뒤 다시 dimension을 50으로 줄인 후에 sequence를 128로 늘려 [그림 14]처럼 [bs, 50, 128]의 형태를 만들어 transformer에 입력으로 주었다. 이 결과는 0.9529로 beatgan과 b2i-gan의 AUC를 훨씬 상회하는 결과로, 성공적인 결과를 얻었다.

**3.7.2.3. sequence를 concat한 후 sequence 증가**



[그림 16] sequence를 concat한 후 sequence 증가

[그림 15]의 구조에서는 dimension과 sequence를 모두 변경하기 때문에 왜곡이 클 수 있다는 문제점이 있다. 그래서 sequnce만을 늘리되 [그림 14]와 달리 먼저 sequence를 concat한 후에 linear를 통해 sequence를 128로 늘려 [bs, 50, 128]로 만들어 transformer에 입력한다. 이 방법의 AUC는 0.9588로 SOTA성능을 달성하여 이 모델을 최종 모델로 결정하고, 이름을 SIA-GAN (signal and image attention GAN)으로 지어줬다.

**4. 결과**

**4.1. 실험 결과 비교**

[표 2] AUC 비교

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | |  | 모델 | | AUC |
| UNIMODAL | | | BeatGAN (signal -> signal) | | 0.9404 |
| BeatGAN + TF (signal -> tf -> signal) | | 0.9449 |
| B2I-GAN ( freq-> freq) | | 0.9460 |
| B2I-GAN + TF ( freq -> tf -> freq) | | 0.9517 |
| Freq -> tf -> Signal | | 0.9269 |
| MULTI  MODAL | Transformer  Fusion | | Linear fusion 적용 | S+f self attention | 0.9529 |
| QKV 변경 | 0.9357 |
| CLS 두개 | Half attention | 0.9376 |
| Half attention + QKV | 0.9376 |
| Arcface loss | 0.9219 |
| Linear fusion | | [bs,50,1], [bs,50,1] -> [bs,100,1] -> [bs,50,128] | | 0.9529 |
| [bs,50,1], [bs,50,1] -> [bs,50,128] -> [bs,50,256] | | 0.9399 |
| **[bs,50,1], [bs,50,1] -> [bs,50,2] -> [bs,50,64] -> [bs,50,128] (최종 모델)** | | **0.9588** |

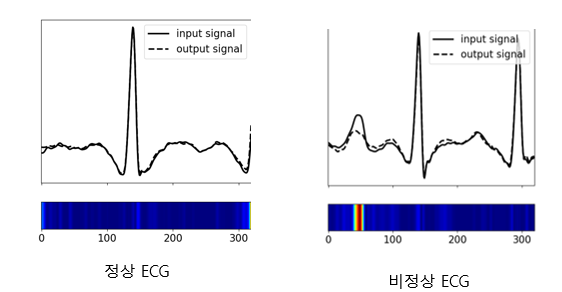
앞선 실험의 결과를 비교하자면 [표 2]와 같다. SIA-GAN의 최종 모델을 보면 unimodal의 어떤 모델과 비교해도 더 높은 AUC를 달성하며, feature fusion이 성공적으로 이루어졌음을 알 수 있다.

**4.2. 모델 비교**

|  |  |
| --- | --- |
| **모델** | **AUC** |
| AE | 0.8944 |
| AnoGAN | 0.8642 |
| GANnomaly | 0.9083 |
| BeatGAN | 0.9404 (\*0.9447) |
| B2I-GAN | 0.9460 |
| ***SIA-GAN*** | ***0.9588*** |

기존의 GAN기반 unimodal 연구들 과도 AUC를 비교해 보아도 높은 AUC를 보이면서 성공적으로 연구를 끝마쳤다.

**4.3. signal 생성 결과**



정상 ECG를 SIA-GAN에 입력하면 점선(fake signal)이 실선(input signal)과 거의 비슷하게 나타나는 것을 볼 수 있다. 반대로 비정상 ECG를 입력하면 우측과 같이 점선이 정상 ecg의 형태로 generation되기 때문에 비정상이 일어난 부분이 아래 heatmap에 붉은색으로 표시되는 것을 볼 수 있다. 따라서 비정상 탐지를 성공적으로 마쳤다고 할 수 있다.

**5. 결론 및 향후 연구 방향**

기존 GAN기반 심부전 탐지 연구들 에서는 unimodal인 signal만을 이용했는데, 우리 모델은 signal과 함께 freq 정보도 함께 사용해서 정확도를 높였다. 또한 multimodal에서 feature를 효율적으로 fusion하는 새로운 방식을 제시하면서 심부적을 효과적으로 탐지하는데 기여했다. 또한 추후 transformer fusion 방식을 조금만 더 발전시킨다면 더욱 성능향상이 있을 것으로 기대된다.

**6. 참고문헌**

[1] Thomas Schlegl, Philipp Seeböck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, Georg Langs 2017. Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery

[2] Samet Akcay, Amir Atapour-Abarghouei, Toby P. Breckon 2018. GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training

[3] Bin Ahou, Shenghua Liu, Bryan Hooi, Xueqi Cheng, Jing Ye., 2019c. BeatGAN: Anomalous Rhythm Detection using Adversarially Generated Time. IJCAI-19.

[4] Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P., Ng, A.Y., 2019. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature medicine 25, 65.

[5] Park, Y., Yun, I.D., Kang, S.H., 2019. Preprocessing method for performance enhancement in cnn-based stemi detection from 12-lead ecg. IEEE Access 7, 99964–99977.

[6] Liu, W., Zhang, M., Zhang, Y., Liao, Y., Huang, Q., Chang, S., Wang, H., He, J., 2017. Real-time multilead convolutional neural network for myocardial infarction detection. IEEE journal of biomedical and health informatics 22, 1434–1444.

[7] Wang, L., Zhou, W., Chang, Q., Chen, J., Zhou, X., 2019c. Deep ensemble detection of congestive heart failure using short-term rr intervals. IEEE Access.

[8] Gyawali, P.K., Chen, S., Liu, H., Horacek, B.M., Sapp, J.L., Wang, L., 2017. Automatic coordinate prediction of the exit of ventricular tachycardia from 12-lead electrocardiogram, in: 2017 Computing in Cardiology (CinC), IEEE. pp. 1–4.

[9] Gyawali, P.K., Horacek, B.M., Sapp, J.L., Wang, L., 2019. Sequential factorized autoencoder for localizing the origin of ventricular activation from 12-lead electrocardiograms. IEEE Transactions on Biomedical Engineering .

[10] Lee, J.S., Lee, S.J., Choi, M., Seo, M., Kim, S.W., 2019b. Qrs detection method based on fully convolutional networks for capacitive electrocardiogram. Expert Systems with Applications 134, 66–78.

[11] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E., 2000. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation 101, e215–e220.

[12] Fotiadou, E., Konopczyński, T., Hesser, J.W., Vullings, R., 2020. End-to-end trained cnn encoder-decoder network for fetal ecg signal denoising. Physiological Measurement .

[13] Wang, J., Li, R., Li, R., Li, K., Zeng, H., Xie, G., Liu, L., 2019b. Adversarial de-noising of electrocardiogram. Neurocomputing 349, 212–224.

[14] Li, Y., Zhang, Y., Zhao, L., Zhang, Y., Liu, C., Zhang, L., Zhang, L., Li, Z., Wang, B., Ng, E., et al., 2018c. Combining convolutional neural network and distance distribution matrix for identification of congestive heart failure. IEEE Access 6, 39734–39744.

[15] Ballinger, B., Hsieh, J., Singh, A., Sohoni, N., Wang, J., Tison, G.H., Marcus, G.M., Sanchez, J.M., Maguire, C., Olgin, J.E., et al., 2018. Deepheart: semi-supervised sequence learning for cardiovascular risk prediction, in: Thirty-Second AAAI Conference on Artificial Intelligence.

[16] Xia, Y., Zhang, H., Xu, L., Gao, Z., Zhang, H., Liu, H., Li, S., 2018b. An automatic cardiac arrhythmia classification system with wearable electrocardiogram. IEEE Access 6, 16529–16538.

[17] Yin, Z., Zhao, M., Wang, Y., Yang, J., Zhang, J., 2017. Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. Computer methods and programs in biomedicine 140, 93–110

[18] Zhang, J., Lin, F., Xiong, P., Du, H., Zhang, H., Liu, M., Hou, Z., Liub, X., 2019a. Automated detection and localization of myocardial infarction with staked sparse autoencoder and treebagger. IEEE Access .

[19] Jingshan Huang, BINQIANG CHEN, , BIN YAO, WANGPENG HE. 2019. ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network. IEEE Access 7, 92871-92880

[20] Sara Ross-Howe, Kimia Lab, Waterloo. 2019. Atrial Fibrillation Detection Using Deep Features and Convolutional Networks

[21] Jingshan Huang, BINQIANG CHEN, , BIN YAO, WANGPENG HE. 2019. ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network. IEEE Access 7, 92871-92880