

# 동적 관심구간 선정법을 활용한 디퓨전 기반 PPG-to-ECG 변환 방법론

추창욱 · 김성범<sup>†</sup>

고려대학교 산업경영공학과

## Diffusion-Based PPG-to-ECG Translation with Dynamic ROI Selection

Changwook Chu · Seoung Bum Kim

Department of Industrial and Management Engineering, Korea University

Electrocardiography (ECG) is one of the simplest and fastest methods for examining the cardiovascular system. Despite its widespread use in assessing cardiac health, collecting ECG data can be a challenge, often requiring the attachment of multiple sensors to the body or the use of specialized equipment. To overcome these limitations, many research has been conducted to find alternative approaches. One promising avenue is the use of photoplethysmography (PPG) data, which has a high correlation with ECG data and can be easily collected in real-time through wearable devices. Consequently, there has been growing interest in using PPG data to generate ECG data for real-time cardiac analysis. In this study, we propose a diffusion-based PPG-to-ECG translation with dynamic region of interest (ROI) selection, using the diffusion model success in various domains, including image, video, audio, and time-series data. Our proposed methodology addresses a key limitation of traditional diffusion models, which struggle to accurately generate the QRS complex which is an essential component of ECG data. We enhance the model performance by selectively adding noise to the ROI, focusing on the most critical segments. Instead of relying on hyperparameters for ROI selection, we develop a dynamic ROI selection algorithm. Furthermore, instead of using a standard U-Net architecture commonly used in image processing, we propose using a diffusion model suitable for time-series data, DiffWave, enabling more precise and detailed extraction of signal characteristics. To validate the performance of our model, we use a benchmark comprising five bio-signal datasets. Experimental results show that, compared with the baseline RDDM, our method reduced the average RMSE from 0.533 to 0.212 and FD from 129.88 to 14.28 across five benchmark datasets.

**Keywords:** Diffusion Model, Cardiovascular Health, Dynamic ROI Selection, Time-Series Data, ECG Data Generation

### 1. 서론

심혈관 질환은 심장과 혈관 등 순환계통에 발생하는 질환을 통칭하며, 전 세계적으로 연간 약 1,790만 명의 사망을 초래하는 주요 사망 원인 중 하나이다(Kauffman, 2020). 특히 심방세동은 성인에서 가장 흔한 리듬 장애로, 조기 발견은 뇌졸중의 1차 및

2차 예방에 필수적이다(Olier *et al.*, 2021). 심혈관 질환을 탐지하기 위해 가장 널리 사용되는 생체 신호는 심전도 (electrocardiography, ECG)로, QRS 복합체, P파, T파 등 심장의 탈분극 및 재분극 과정을 반영하는 다양한 진단 정보를 제공한다 (Hannun *et al.*, 2019). 그러나 ECG 측정을 위해서는 12리드 홀터 모니터와 같은 전문 장비가 필요하며, 다수의 전극을 부착

<sup>†</sup> 연락저자 : 김성범 교수, 02841 서울특별시 성북구 안암로 145 고려대학교 산업경영공학부, Tel: 02-3290-3397, Fax: 02-929-5888, E-mail : sbkim1@korea.ac.kr

2025년 5월 7일 접수; 2025년 9월 2일; 2025년 10월 24일 수정본 접수; 2025년 10월 24일 게재 확정.

해야 하는 번거로움과 장시간 연속 측정 시 신호 품질 저하 등의 문제가 존재한다. 최근 스마트워치 등 소비자용 웨어러블이 상용화되었으나, 손가락 접촉 방식으로 30초 내외의 짧은 시간 동안만 측정할 수 있어 실질적인 연속 모니터링에는 한계가 있다(Bayoumy *et al.*, 2021). 이러한 제약을 극복하기 위한 대안으로, LED와 광검출기를 이용해 피부 표면의 혈류 변화를 비침습적으로 감지하는 광용적맥파(photoplethysmography, PPG)가 주목받고 있다. PPG는 스마트워치와 같은 웨어러블에서 실시간으로 손쉽게 수집할 수 있다는 장점이 있지만, 피부 접촉 상태, 움직임, 주변 광원 간섭 등에 민감하여 ECG에 비해 높은 노이즈를 포함한다(Lee *et al.*, 2020). 그럼에도 불구하고 PPG와 ECG는 동일한 심장 박동 메커니즘을 반영하기 때문에 높은 상관성을 가지며, 이를 기반으로 실시간 PPG 신호로부터 ECG 신호를 생성하려는 다양한 연구가 활발히 진행되고 있다(Lan, 2023; Sarkar and Etemad, 2021; Vo *et al.*, 2024).

기존 연구들은 주로 시간-주파수 영역에서 수작업으로 추출한 특징을 기반으로 RR, PR, QRS, QT 간격을 예측하는 머신러닝 방식을 사용했으나, 복잡한 특징 설계에 대한 의존도가 높아 실제 환경에서의 일반화 성능에 한계가 있었다(Banerjee *et al.*, 2014; Islam *et al.*, 2017). 최근에는 이러한 한계를 극복하기 위해 트랜스포머(transformer), 디퓨전 모델(diffusion model) 등 딥러닝 기반 기법이 제안되었다(Sohl-Dickstein *et al.*, 2015). 특히, 디퓨전 모델은 이미지, 오디오, 비디오 등 다양한 분야에서 고품질의 생성 성능을 입증하였으며, 이는 생체 신호 생성 분야에서도 정밀하고 신뢰도 높은 시그널 생성을 가능하게 할 잠재력을 보여주고 있다. 그러나 전통적인 디퓨전 기법인 denoising diffusion probabilistic models(DDPM) 방식은 시그널 전체 구간에 균등하게 노이즈를 추가하고 제거하는 방식으로 동작하기 때문에, ECG의 핵심 구조인 QRS 복합체와 같은 미세한 세부 정보를 복원하는 데 한계가 있다(Ho *et al.*, 2020). QRS 복합체는 Q파, R파, S파로 구성된 파형으로, 심실의 탈분극 과정이 가장 뚜렷하게 나타나는 구간이다. 이를 보완하기 위해 일부 연구에서는 관심 구간(region of interest, ROI)을 설정하여 QRS 구간에 집중적으로 노이즈를 부과하고 제거하는 학습 방식을 도입했으나, ROI 구간을 하이퍼파라미터로 수동 설정해야 하고, 데이터 간 QRS 길이에 차이가 존재하여 일반화에 어려움이 있었다(Shome *et al.*, 2024). 또한, 기존 디퓨전 모델의 백본으로 사용된 U-Net은 원래 바이오메디컬 이미지 세그멘테이션을 위해 설계된 구조로, 1차원 시계열 신호의 시간적 연속성과 주기성을 효과적으로 반영하지 못할 수 있다는 구조적 한계를 가진다(Ronneberger *et al.*, 2015).

본 연구에서는 기존 디퓨전 기반 PPG-to-ECG 변환 방법론의 한계를 극복하기 위해, 동적 ROI선정 알고리즘을 적용한 새로운 디퓨전 기반 변환 프레임워크를 제안한다. 먼저, ECG의 핵심 정보가 포함 되어있는 QRS 구간을 ROI로 활용하기 위해 P, Q, R, S, T 피크들을 전부 활용하는 알고리즘을 개발하였다. 제안 알고리즘은 별도의 하이퍼파라미터 설정 없이 각

데이터셋의 특성에 맞게 ROI를 동적으로 선정할 수 있어, 다양한 환경에 대한 사용 가능하다. 또한, 최근 시계열 디퓨전 모델에 대한 연구가 활발히 이루어지고 있으며, 이러한 모델은 시계열 데이터의 연속성과 주기성을 효과적으로 반영할 수 있다는 점에서 시그널 생성에 유리한 특성을 가진다. 이에 본 연구는 오디오 생성용으로 개발된 DiffWave 구조를 기반으로, PPG와 ECG의 시그널 특성을 정교하게 반영할 수 있도록 구조를 확장하였다(Kong *et al.*, 2021). 구체적으로, PPG와 ECG 각각을 처리하는 독립적인 잔차 블록(residual block)을 설계하고, 이들 블록의 최종 특징 맵(feature map) 간에 다중 헤드 크로스 어텐션(multi-head cross-attention)을 수행함으로써 두 시그널 간의 연관성을 효과적으로 학습할 수 있도록 하였다. 아울러, 시간 정보를 모델에 반영하기 위해 학습 가능한 위치 인코딩(learnable positional encoding)을 추가하여 시간 조건을 복합적으로 반영하였다.

본 연구에서는 PPG와 ECG 쌍을 포함하는 다섯 개의 공개 데이터셋을 Shome *et al.*(2024)의 방식에 준하여 하나의 통합 벤치마크로 구성하였으며, 이를 기반으로 학습을 수행한 후, 각 개별 데이터셋에 대해 정량적 및 정성적 성능을 비교·분석하였다. 실험 결과, 제안한 방법론은 모든 벤치마크와 디퓨전 샘플링 단계에서 정량적 지표와 정성적 평가 모두 기존 방법에 비해 우수한 성능을 나타내었다. 본 논문의 기여점은 다음과 같이 요약할 수 있다.

- ECG 데이터의 모든 피크점들을 활용하여 ROI 영역을 동적으로 선정할 수 있는 알고리즘을 제안하였다. ROI 구간의 길이를 하이퍼파라미터로 수동 설정해야 하는 한계가 있었던 반면, 제안된 알고리즘은 데이터의 구조적 특성을 기반으로 ROI를 자동으로 결정하며, 해당 영역 내에서 QRS 복합체를 정밀하게 탐지할 수 있다는 장점을 지닌다.
- 시계열 신호 처리를 위해 설계된 디퓨전 모델 구조를 활용하여, PPG와 ECG 각각의 특성을 세부적으로 반영할 수 있는 프레임워크를 설계하였다. 본 연구에서는 디퓨전 기반 시계열 생성 모델인 DiffWave를 백본으로 채택하여 양방향 팽창 합성곱(bi-dilated convolution) 기반의 잔차 블록을 구성함으로써 심박 간 주기성과 같은 장기 의존 패턴을 효과적으로 학습하고, ECG 파형의 전역적 리듬과 국소적 형태를 동시에 재현할 수 있도록 하였다. 특히, 디퓨전 모델의 점진적 복원 과정을 통해 QRS 복합체와 같은 고주파 파형 요소를 정밀하게 복원할 수 있으며, PPG 조건 정보는 별도의 조건 네트워크를 통해 추출된 후 각 잔차 블록의 최종 특징 맵에 다중 헤드 크로스 어텐션 방식으로 통합되어, 심박 리듬과 같은 전역 시간 정보가 ECG 생성에 직접 반영되도록 설계하였다. 이를 통해 주요 생리학적 파형의 표현력을 강화하여 정밀한 ECG 시그널을 생성할 수 있었으며, 정량적, 정성적 평가에서 일관된 우수한 성능을 보여주었다.

본 논문의 구성은 다음과 같다. 제2장에서는 기존의 ECG 생성 모델 및 PPG-to-ECG 변환과 관련된 선행 연구들을 소개한다. 제3장에서는 본 연구에서 제안하는 동적 ROI 선정 알고리즘과 PPG-to-ECG 디퓨전 모델의 구조 및 학습 과정에 대해 상세히 설명한다. 제4장에서는 데이터셋 구축 및 전처리 과정, 평가 지표, 실험 결과를 통해 제안한 방법론의 정량적·정성적 성능을 분석한다. 마지막으로 제5장에서는 본 연구의 결론을 제시한다.

## 2. 관련 연구

### 2.1 ECG 생성 모델

ECG 시그널을 인공적으로 생성하려는 시도는 생체 신호 분석 분야에서 데이터 증강, 프라이버시 보호 등의 필요성에 의해 지속적으로 연구되어왔다. 초기 연구에서는 McSharry *et al.*(2003)의 3차원 동적계 모델이나 Quiroz-Juarez *et al.*(2018)의 이중 발진기 기반 모델처럼, 심장 전기생리학의 메커니즘을 수식화하여 파형을 합성하는 해석 기반 접근법이 주로 활용되었다. 그러나 이러한 해석 기반 ECG 합성 모델은 심장 전기생리를 미리 정의된 미분방정식으로 기술하므로, 모델 파라미터가 정상 성인의 평균적 심전도를 기준으로 조정되는 경우가 많아 연령, 성별, 질환 상태에 따른 형태 변이를 포괄적으로 반영하기 어렵다. 특히, 드물게 나타나는 부정맥이나 병리적 패턴을 재현할 때 일반화 성능이 급격히 저하되는 문제가 발생한다. 또한, 심박 변동성과 같은 정보를 원하는 형태로 얻기 위해서는 파형의 높이, 폭, 위치 등 다양한 파라미터를 반복적으로 조정해야 한다. 그러나 이러한 파라미터는 비선형적으로 상호 의존하므로 최적화 과정이 복잡하고 대규모 데이터 증강에는 비효율적이다.

최근에는 대규모 생체 신호 데이터셋과 딥러닝 기법의 발전에 힘입어, variational autoencoder(VAE), generative adversarial network(GAN), 디퓨전 모델 등 확률적 생성 기법을 활용한 연구가 활발히 진행되고 있다. 이러한 방법들은 다양한 심장 리듬, 잡음, 개인 간 생리적 차이를 데이터 기반으로 학습함으로써, ECG 생성의 현실성과 다양성을 크게 향상시켰다. 예를 들어, Zhu *et al.*(2019)은 양방향 long short-term memory-convolutional neural network(BiLSTM-CNN) 기반 GAN 구조를 제안하여 단일 리드 ECG 시그널을 합성하였다. 이 모델은 2층의 양방향 LSTM을 생성기로 사용하여 노이즈 시퀀스를 ECG 파형으로 변환하고, 1차원 CNN 기반 판별기를 통해 생성 신호의 진위를 판별함으로써 학습의 안정성을 높였다. 해당 모델은 비교적 짧은 학습 시간으로도 기존 합성 모델 대비 높은 형태학적 유사도와 우수한 데이터 증강 효과를 보였다. 그러나 단일 리드 ECG만 생성이 가능하며, PPG 신호나 개인별 생리적 특성을 조건으로 반영하지 못한다는 점에서 여전히 한계가 존

재한다.

Xia *et al.*(2023)은 ECG 데이터 부족 문제를 해결하기 위해 조건부 생성 프레임워크를 도입하였다. 제안된 방법은 두 갈래로 구성된다. 첫째, short-time Fourier transform(STFT)기반 스펙트로그램(spectrogram)을 라벨과 함께 학습하는 conditional Wasserstein GAN(CWGAN)은 2차원 스펙트럼을 합성한 뒤 역방향 STFT를 통해 ECG를 복원한다. 둘째, 조건부 VAE는 실제 심박 시퀀스와 클래스 라벨을 인코더에 입력하고, 잠재 공간에서 샘플링 된 벡터를 디코더로 복원하여 직접 ECG를 생성한다. 실험 결과, 해당 모델은 기존 GAN 대비 절반 수준의 파라미터 수로도 빠른 수렴을 보였다. 그러나 단일 리드, 고정된 시퀀스 길이, 낮은 수준의 노이즈 환경을 전제로 설계되었기 때문에, 실제 PPG 시그널을 조건으로 활용하거나 다양한 환경에서 ECG를 생성하는 데에는 한계가 존재한다.

Adib *et al.*(2023)은 ECG 데이터 생성을 위해 딥러닝 기반 확률적 생성 모델인 DDPM을 활용하는 새로운 접근 방식을 제안하였다. 해당 연구에서는 MIT-BIH Arrhythmia 데이터셋의 normal sinus beat을 개별 심장박동 단위로 추출하여 2차원 이미지로 변환한 뒤, 2D DDPM으로 학습 및 합성하였다. 생성된 합성 심장박동의 품질은 정상 박동과 이상 박동을 구분하는 이진 분류기를 생성된 합성 데이터셋에 적용하여 확인하였다. 실험 결과, precision-recall AUC에서 0.95를 달성하며 실제 ECG와 높은 유사성을 보여주었으나, 1D 시그널을 2D로 변환하고 다시 복원하는 과정에서 정보 손실이 발생하여 생성된 ECG의 세밀한 파형 형태가 원본에 비해 부정확해지는 문제가 있었다. 이처럼 기존 연구들은 다양한 생성 기법을 통해 ECG 데이터를 합성하려는 시도를 이어왔으나, 실제 수집이 용이한 생체 신호인 PPG를 조건으로 활용하여 신뢰성 높은 ECG를 생성하는 데에는 한계가 있었다. 특히, 실제 환경에서는 ECG 시그널을 직접 수집하기 어려운 경우가 많고, 개인별 생리적 특성이나 노이즈가 존재하는 신호로부터 높은 품질의 ECG를 생성하는 것이 필수적임에도 불구하고, 이러한 조건부 생성에 초점을 맞춘 연구는 상대적으로 부족하다. 이에 따라, 본 연구는 PPG 시그널을 조건으로 활용하면서도 개인별 심박 특성과 환경적 노이즈를 반영하여 다양한 형태의 ECG를 생성할 수 있는 새로운 접근 방식을 제안한다.

### 2.2 PPG-to-ECG 변환 모델

최근에는 PPG 시그널을 활용하여 ECG를 생성하는 새로운 패러다임이 주목받고 있다. PPG는 혈류에 따른 광학적 변화를 측정하는 비침습적 시그널로, 수집이 간편하고 심혈관계 상태를 간접적으로 반영할 수 있다는 장점이 있다. 이러한 특성으로 인해, PPG를 기반으로 ECG를 복원하거나 생성하려는 다양한 연구가 진행되고 있으며, 이는 ECG의 직접적인 수집이 어려운 환경에서 대안적 접근법으로서 유용할 뿐만 아니라, 개인 맞춤형 헬스케어 구현 측면에서도 중요한 가능성을 제시한다.

Sarkar and Etemad(2021)은 PPG 시그널을 활용하여 ECG를 생성하는 CardioGAN을 제안하였다. CardioGAN은 GAN 구조를 기반으로 생성기에 시퀀스 기반 어텐션 메커니즘을 적용하여 ECG 시그널 내 세밀한 시간 의존성을 학습하도록 설계하였다. 생성자는 가우시안 분포(Gaussian distribution)에서 샘플링한 노이즈 시퀀스를 입력으로 받아 양방향 LSTM 기반의 어텐션 블록들을 통과시키며 ECG 시그널을 생성한다. 어텐션 메커니즘을 통해 P파, QRS 복합체, T파 등 주요 ECG 요소들의 구조적 특징과 시간적 관계를 효과적으로 복원할 수 있도록 설계되었다. 판별자는 ID CNN 기반 구조를 채택하여 생성된 시그널의 진위를 평가하며, Wasserstein loss와 gradient penalty를 적용해 학습 안정성을 강화하였다. 실험 결과, CardioGAN은 기존 양방향 LSTM-GAN 대비 다양한 정량적 평가 지표에서 우수한 성능을 보였다. 하지만 CardioGAN은 GAN 특유의 훈련 불안정성으로 인해 생성자와 판별자의 학습 균형 조절이 어려워 모드 붕괴(mode collapse)와 같은 문제가 발생하였고, 이는 생성된 ECG 시그널의 다양성과 품질을 저해하여 실제 적용 가능성에 한계를 초래하였다.

Lan(2023)은 PPG 시그널로부터 ECG를 생성하기 위해 트랜스포머 기반의 모델을 제안하였다. 제안된 모델은 시계열 생체 시그널에 특화된 구조로 PPG 시그널을 입력 받아 ECG 파형을 복원하는 sequence-to-sequence 학습을 수행한다. 해당 모델은 기존 CNN이나 RNN 기반 접근법들과 달리, 트랜스포머의 셀프 어텐션 메커니즘을 활용하여 긴 시계열 내 글로벌 컨텍스트를 효과적으로 학습하고, 중요한 파형 구간에 집중할 수 있도록 설계되었다. 특히 입력 시퀀스를 다양한 크기의 패치로 나누고 이들을 교차 연결함으로써, 생체 시그널의 지역적 패턴과 전체적 구조를 동시에 포착하는 데 성공하였다. 그러나 단순히 PPG와 ECG 간 전역적 매핑을 학습하는 데 중점을 두기 때문에, ECG 시그널 내에서 진단적으로 중요한 관심 영역이나 미세한 국소적 특성을 별도로 강조하거나 세밀하게 제어하는 구조적 메커니즘은 부족하다는 한계가 존재한다.

Shome *et al.*(2024)은 PPG 시그널을 조건으로 받아 ECG를 생성하는 새로운 접근으로 region-disentangled diffusion model(RDDM)을 제안하였다. 기존 DDPM 기반 모델들이 전 구간에 균일하게 노이즈를 추가함으로써 ECG의 복잡한 국소적 구조를 충분히 학습하지 못하는 문제를 지적하고, 이를 해결하기 위해 관심 영역에 선택적으로 노이즈를 추가하는 새로운 포워드 프로세스(forward process)를 설계하였다. 이후 리버스 프로세스(reverse process)에서 ROI와 비 ROI 구간을 분리하여 독립적으로 복원하는 구조를 적용하여 ECG의 전역적 구조와 세밀한 지역적 패턴을 모두 복원하도록 설계하였다. 하지만 RDDM은 이미지 작업에 특화된 U-Net을 백본 모델로 사용하여, 시계열 데이터의 정보를 풍부하게 반영하지 못한다는 한계점이 있다. 또한, ROI 선택 과정에서 고정된 하이퍼파라미터를 기반으로 영역을 지정하기 때문에, 사전에 도메인 전문가의 개입이 필요하며, 시그널마다 상이한 ROI 길이를 유연하

게 반영하지 못한다는 구조적 제약이 존재한다.

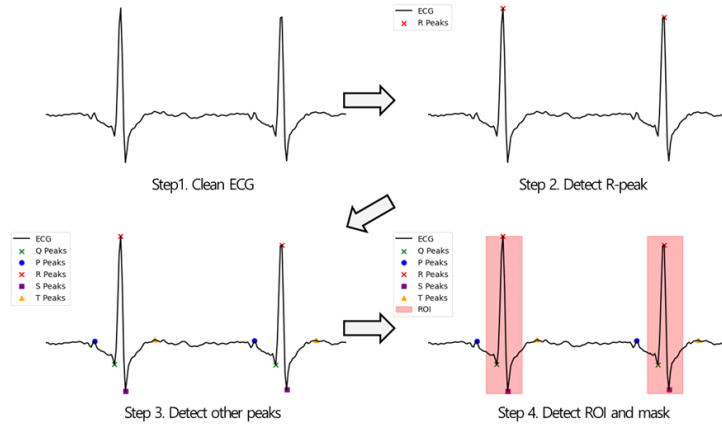
위에 설명한 바와 같이 기존 PPG-to-ECG 변환 연구들은 각각 의미 있는 성과를 거두었으나, 훈련의 불안정성, 모드 붕괴와 같은 생성 모델 특유의 문제 또는 진단적으로 중요한 관심 구간 및 미세한 파형 복원을 정밀하게 조정할 수 있는 구조적 메커니즘의 부재라는 한계가 존재한다. 이러한 문제를 해결하기 위해, 본 연구에서는 시계열 데이터 특성에 최적화된 모델 구조를 기반으로, PPG 시그널을 조건으로 받아 ECG를 생성하되, 관심 영역을 동적으로 식별하고 해당 영역에 집중적으로 노이즈를 부여하는 방식을 제안한다. 이를 통해 시계열 시그널의 전역적 구조와 국소적 패턴을 동시에 정밀하게 복원함으로써, 개인별 생리적 특성과 다양한 노이즈 조건을 효과적으로 반영할 수 있는 고품질 ECG 생성 모델을 구현하는 것을 목표로 한다.

### 3. 제안 방법론

#### 3.1 동적 ROI 선정 알고리즘

본 연구에서 제안하는 동적 ROI 선정 알고리즘은 ECG 시그널의 생리학적 파형 구조를 충실히 반영하면서도 데이터마다 상이한 QRS 복합체 폭과 노이즈 특성에 자동으로 적응할 수 있도록 설계되었다. <Figure 1>은 동적 ROI 선정 과정을 보여주고 있다. 먼저, 128Hz의 샘플링율로 수집된 ECG 시그널에 대해 NeuroKit2의 Pan-Tompkins 전처리 함수를 사용하여 필터링을 적용하였다(Makowski *et al.*, 2021; Pan and Tompkins, 1985). 이 전처리 과정은 심방 탈분극을 나타내는 P파, 심실 탈분극 초기 음성 전위를 나타내는 Q파, 심실 탈분극 주전위를 의미하는 R파, 심실 탈분극 후 음성 전위를 나타내는 S파, 심실 재분극을 의미하는 T파 등 ECG의 주요 파형 요소들이 손실 없이 유지될 수 있도록 하기 위함이다.

이후 전처리 된 시그널로부터 R파 후보를 탐지하기 위해, SciPy 모듈을 활용하여 피크 탐색을 수행하였다. 이때 최소 피크 간격은 30샘플(약 234ms)로 제한하였고, 피크 높이 기준은 시그널 평균에 0.5 표준편차를 더한 값 이상으로 설정하여 잡음성 스파이크를 효과적으로 배제하였다(Virtanen *et al.*, 2020). 또한 인접한 R파 간 거리가 50샘플(약 390 ms) 이하로 연속 검출되는 경우, 안정 시 성인의 전형적 R-R 간격(약 600~1000 ms)에 비추어 T파 과대나 잡음에 따른 오검출 가능성이 높으므로, 더 큰 진폭 하나만 남기도록 추가 필터링을 적용하였다. 이를 통해 중복 검출을 방지하고 안정적인 R파 리스트를 확보하였다. 확정된 각 R파를 기준으로, R파 이전 30샘플 구간에서 최솟값을 Q파로, 해당 Q파 이전 30샘플 구간에서 최댓값을 P파로, R파 이후 30샘플 구간에서 최솟값을 S파로, S파 이후 30샘플 구간에서 최댓값을 T파로 정의하여 PQRST 파형의 좌표를 탐색하였다. 이때 30샘플(약 234 ms) 창은 전형적 QRS 폭(약 70~110 ms)을 충분히 포함하면서 인접 박동의 간섭을 최소



**Figure 1.** Process of selecting dynamic ROI. The first step involves preprocessing the ECG signal to obtain a clean waveform. R-peaks are then detected using the SciPy module, followed by detection of the remaining P, Q, S, and T peaks using the same method. When all five characteristic peaks are successfully identified, the midpoint between the P and Q waves is designated as the start of the ROI, and the midpoint between the S and T waves is set as the end. The ROI is then masked with a value of 1.

화하기 위한 크기이다. 만약 Q파, P파, S파 또는 T파 중 하나라도 탐지되지 않을 경우, QRS 복합체의 최소 폭으로 알려진 약 15샘플(약 117ms)을 기준 거리로 삼아 해당 피크의 전후 구간을 일정 거리만큼 이동시켜 임시 경계를 설정하였다.

P, Q, R, S, T 파형이 모두 탐색된 후에는, P-Q 구간과 S-T 구간을 연결하는 두 지점의 중간 지점을 각각 ROI의 좌우 경계로 설정함으로써, QRS 복합체를 포함한 주요 탈분극 및 재분극 활동이 영역 내에 균형 있게 포함되도록 하였다. 이렇게 계산된 경계는 시그널 전체 길이 내에서 자동으로 조정되며, 과도한 확장이나 축소 없이 각 파형의 생리학적 의미를 충실히 반영한다. 마지막으로 경계 구간에 대응하는 이진 마스크를 생성하고 해당 구간을 1로 설정한 뒤, 이를 디퓨전 모델의 손실함수에 해당 마스크를 가중치로 적용함으로써, 모델이 QRS 복합체 영역을 집중적으로 학습하도록 유도하였다. 제안 알고리즘은 별도의 하이퍼파라미터 조정 없이도 다양한 ECG 데이터셋에서 일관된 ROI 탐지 성능을 보이며, 특히 실제 착용 환경에서 흔히 발생하는 움직임 기반 노이즈에 대해서도 높은 강건성을 유지하는 것으로 확인되었다.

### 3.2 PPG-to-ECG 디퓨전 모델

디퓨전 생성 모델은 원시 시그널에 점진적으로 노이즈를 추가하는 포워드 프로세스와 추가된 노이즈를 예측하고 제거하는 리버스 프로세스로 구성된다. 포워드 프로세스는 깨끗한 시그널  $x_0$ 에 점차적으로 가우시안 노이즈를 주입하며  $x_1, x_2, \dots, x_T$ 를 생성하며, 이는 다음과 같은 마르코프 연쇄 형태로 정의된다:

$$q(x_T | x_0) := \prod_{t=1}^T q(x_t | x_{t-1}), \quad (1)$$

where  $q(x_t | x_{t-1}) := N(x_t; \sqrt{1 - B_t} x_{t-1}, B_t I)$

식 (1)에서  $x_0 \sim q(x_0)$ 은 노이즈가 포함되지 않은 원 시그널을 나타내며,  $B_t \in (0, 1)$ 은 각 시점  $t$ 에서 시그널에 주입되는 노이즈 강도를 제어하는 분산 스케줄을 의미한다. 일반적으로  $B_t$ 는  $t$ 가 증가할수록 점진적으로 커지도록 설계되며, 이를 통해 초기에는 소량의 노이즈를, 후반에는 보다 강한 노이즈를 추가하도록 구성된다. 수식의 단순화를 위해  $\alpha_t = 1 - B_t$  및  $\bar{\alpha} = \prod_{s=1}^t \alpha_s$ 를 정의하면,  $x_0$ 에서  $x_t$ 를 샘플링할 때의 분포를 식 (2)와 같이 나타낼 수 있다:

$$q(x_T | x_0) := N(x_T; \sqrt{\bar{\alpha}} x_0, (1 - \bar{\alpha}) I) \quad (2)$$

그러나 앞서 정의한 포워드 과정을 단순히 역방향으로 적용하여  $x_t$ 로부터  $x_{t-1}$ 의 조건부 확률  $q(x_{t-1} | x_t)$ 는 closed-form으로 계산이 불가능하다. 따라서 DDPM에서는 evidence lower bound(ELBO)를 극대화 하는 방식으로 리버스 프로세스를 근사하며, 식 (3)과 같이 파라미터화 된 가우시안 전이 분포로 정의된다:

$$p_\theta(x_{t-1} | x_t) = N(x_{t-1}; \mu_\theta(x_t, t), \sigma_\theta^2 I) \quad (3)$$

이때 평균  $\mu_\theta(x_t, t)$ 는 포워드 과정에서 주입된 노이즈를 제거하고 시그널을 복원하기 위해 다음과 같이 정의된다:

$$\mu_\theta(x_t, t) = \frac{1}{\sqrt{\alpha_t}} \left( x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right) \quad (4)$$

여기서  $\epsilon$ 은 포워드 과정에서 원시 시그널  $x_0$ 에 추가된 실제 가우시안 노이즈 샘플이며,  $\epsilon_\theta(x_t, t)$ 은 리버스 네트워크가  $x_t$ 와 시점  $t$ 를 입력으로 받아 예측한 노이즈 성분을 의미한다. 학습 목표는 실제 노이즈  $\epsilon$ 과 예측된 노이즈  $\epsilon_\theta$  사이의 평균

제공 오차를 최소화하는 것으로, 이는 ELBO 기반 손실 함수로 구체화된다.

본 연구에서는 이러한 디퓨전 모델의 강점을 유지하면서도, ECG에서 가장 중요한 QRS 복합체에 노이즈를 선택적으로 주입하고 제거할 수 있도록 설계된 동적 ROI 기반 디퓨전 모델을 제안한다. 먼저, 앞서 설명한 동적 ROI 선정 알고리즘을 통해 QRS 복합체가 포함된 관심 영역을 탐색하고, 해당 영역을 1로 마스킹하는  $m \in \{0,1\}^L$ 을 생성한다. 학습 단계에서는 각 배치 샘플마다 무작위로 디퓨전 스텝  $t$ 를 선택하고, 해당 시점에 대응하는 누적 스케일 계수  $\sqrt{\alpha_t}$ 와 노이즈 세기 계수  $\sqrt{1-\alpha_t}$ 를 참조한다. 이후, 전역 가우시안 노이즈  $\epsilon \sim N(0,1)$ 를 활용하여 ROI 영역에만 노이즈를 부여한  $\epsilon_{ROI} = m \odot \epsilon$ 와, 시그널 전 구간에 일괄적으로 노이즈를 부여한  $\epsilon_{full} = \epsilon$ 을 각각 생성한다. 여기서,  $\odot$ 는 곱연산을 의미하며, 이는 두 벡터 또는 행렬의 원소별 곱을 나타낸다. 이 두 형태의 노이즈를 기반으로 다음과 같이 시그널을 생성한다:

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{1-\alpha_t} \epsilon_{ROI} \quad (5)$$

$$x_t^{unmasked} = \sqrt{\alpha_t} x_0 + \sqrt{1-\alpha_t} \epsilon_{full} \quad (6)$$

식 (5)의  $x_t$ 는 노이즈 예측 네트워크( $\epsilon_\theta$ )에 입력되어 ROI 내 노이즈 성분인  $\epsilon_{ROI}$ 를 복원하도록 학습되며, 식 (6)의  $x_t^{unmasked}$ 는 시그널 전체를 복원하는 네트워크( $f_\phi$ )에 입력된다. 각각의 손실 함수는 다음과 같이 정의된다:

$$L_{ROI} = E[\|\epsilon_{ROI} - \epsilon_\theta(x_t, t, c_{ppg})\|^2] \quad (7)$$

$$L_{region} = E[\|x_t - f_\phi(x_t^{unmasked}, t, c_{ppg})\|^2] \quad (8)$$

여기서  $c_{ppg}$ 는 조건으로 활용되는 PPG 시그널을 의미한다. 전체 손실은 두 항에 중요도 기반 가중치를 부여한 다음과 같은 형태로 정의된다:

$$L = \lambda_1 L_{ROI} + \lambda_2 L_{region} \quad (9)$$

본 연구에서는  $\lambda_1$ 을 100,  $\lambda_2$ 를 1로 설정하여 ROI 영역을 생성하는데 집중하였다.

샘플링 단계에서는 학습된 노이즈 예측 네트워크와 전체 시그널 복원 네트워크를 순차적으로 호출하여 초기 가우시안 노이즈  $x_T \sim N(0, I)$ 로부터  $x_0$ 을 복원한다. 각 리버스 스텝  $t = T, \dots, 1$ 에서 먼저 전체 시그널 복원 네트워크  $f_\phi$ 를 통해 임시 보정값  $\tilde{x}_t$ 을 다음과 같이 산출한다:

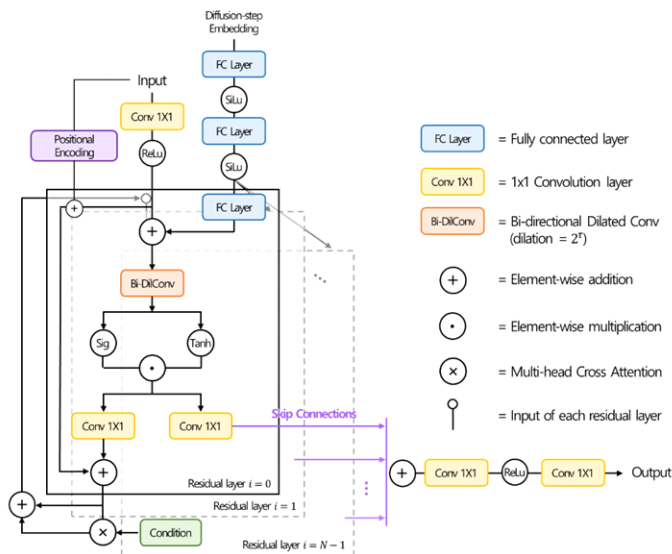
$$\tilde{x}_t = f_\phi(x_t^{unmasked}, t, c_{ppg}) \quad (10)$$

이후, 노이즈 예측 네트워크로부터 예측된 ROI 내 노이즈  $\hat{\epsilon}$ 을 제거하여 최종적으로 다음 식을 적용한다:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left( \tilde{x}_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}} \hat{\epsilon} \right) + \sqrt{B_t} z, \quad z \sim N(0, I) \quad (11)$$

이와 같은 교차 업데이트를 통해 노이즈가 집중된 관심 영역은 점차 정제되고, 전역 시그널 구조와 국소 ROI 정보가 균형 있게 통합된 고품질 ECG 시그널  $x_0$ 이 최종적으로 생성된다.

본 연구에서 사용하는 PPG-to-ECG 변환 모델은 시계열 시그널 합성에 최적화된 DiffWave 아키텍처를 백본으로 채택하였다. DiffWave는 비순차적(non-autoregressive) 양방향 팽창



**Figure 2.** Overall architecture of the proposed model. The model retains the DiffWave backbone while incorporating a learnable positional encoding at the input stage. To better capture time-series characteristics, multi-head cross-attention is applied to the final feature map of each residual block, conditioning on the PPG signal.

합성곱(bi-dilated convolution) 구조를 통해, 적은 디퓨전 스텝 만으로도 긴 시퀀스를 병렬 생성할 수 있는 장점을 지닌다. 특히, 각 잔차 블록 내에서 지수적으로 확장되는 팽창률(dilation)을 적용함으로써 모델의 수용 영역(receptive field)을 학습 샘플이 거치는 전체 디퓨전 스텝 수와 샘플 수를 곱한 만큼 확대하였으며, 이를 통해 ECG와 같은 고해상도 바이오 시그널의 미세 파형부터 전역 리듬 구조까지 균형 있게 포착할 수 있다. <Figure 2>는 제안한 모델의 전반적인 개요를 나타내며, 특히 기존 DiffWave 구조에 다음 세 가지 주요 개선점을 적용하였다. 첫째, 입력 투영(input projection) 이후 각 레이어에 학습 가능한 위치 인코딩을 추가하여, 시그널의 위상 변화와 주기성을 효과적으로 모델링할 수 있게 하였다. 이를 통해 ECG 데이터의 P, Q, R, S, T 파형의 시점 별 특성을 보다 정교하게 반영할 수 있다. 둘째, PPG 조건 정보를 효과적으로 활용하기 위해 별도의 조건 DiffWave 네트워크를 구축하였다. 이 네트워크는 PPG 시그널을 동일한 잔차 블록 구조로 처리하며, 초기 투영 후 학습 가능한 위치 인코딩을 추가하여 PPG의 시계열 패턴을 특징 맵 형태로 추출한다. 셋째, 메인 DiffWave 네트워크의 각 잔차 블록 뒤에 다중 헤드 크로스 어텐션 모듈을 배치하여, ECG 생성 과정에서 PPG 조건 특징과 중간 표현 간의 정보를 효과적으로 융합하였다. 이러한 구조를 통해 제안된 모델은 PPG가 제공하는 전역 심박 리듬 정보를 활용하여, ROI 내 미세한 노이즈 제거에 집중하면서도 전역과 국소의 관점을 모두 반영한 ECG 시그널을 생성할 수 있다. 특히, PPG-to-ECG 전환 과정에서 전파되는 전역 맥락 정보를 정밀하게 반영함으로써, QRS 복합체와 같은 생리학적으로 중요한 파형 요소의 디테일을 놓치지 않는 고품질 ECG 생성을 가능하게 하였다.

## 4. 실험 결과

### 4.1 데이터셋 및 전처리

본 연구에서는 제안 방법론의 성능을 검증하기 위해 다섯 개의 공개된 바이오 시그널 도메인 데이터셋을 활용하였다. 첫 번째 데이터셋은 a multimodal dataset for wearable stress and affect detection(WESAD)로, 총 15명의 피험자로부터 수집된 700Hz 샘플링의 ECG 데이터와 64Hz로 샘플링 된 PPG의 24시간 분량으로 구성되어 있으며, 스트레스 및 정서 상태에 대한 주석이 포함되어 있다(Schmidt *et al.*, 2018). 두 번째 데이터셋은 MIMIC-AFib으로 Johnson *et al.*(2016)이 공개한 MIMIC-III의 일부로부터 구성되었다(Bashar *et al.*, 2019). 해당 데이터셋은 부정맥을 경험한 19명의 환자와 중증 성인 16명으로부터 수집된 데이터이다. ECG와 PPG 시그널로 구성되며, 두 신호 모두 125Hz의 샘플링률(sampling rate)로 기록되었다. 세 번째로 활용된 데이터셋은 PPG dataset for motion compensation and heart rate estimation in daily life activities(DALIA)로 15명의 피

험자로부터 걷기, 운전 등 일상생활 활동을 수행하는 동안 수집된 PPG와 ECG 데이터로 구성 되어있다(Reiss *et al.*, 2019). 각 피험자별 35시간 분량의 데이터가 기록되었으며, PPG는 64Hz, ECG는 700Hz의 샘플링률로 수집되었다. 네 번째 데이터셋은 capnography based dataset(CAPNO)로 의료 감독하에 총 42명의 피험자로부터 약 5.6시간 분량의 ECG와 PPG 시그널로 구성되어 있다(Karlen, 2021). ECG와 PPG 모두 동일한 샘플링률인 300Hz로 수집되었다. 마지막으로 사용한 Beth Israel Deaconess medical center(BIDMC) 데이터셋은 중환자실에 입원한 53명의 환자로부터 수집된 약 7시간 분량의 ECG 및 PPG 시그널로 구성되며, 두 신호 모두 125Hz로 샘플링 되었다(Pimentel *et al.*, 2016).

본 연구에서는 서로 다른 샘플링 주파수와 수집환경을 지닌 다섯 개의 바이오 시그널 도메인의 데이터셋인 WESAD, DALIA, CAPNO, BIMC, MIMIC-AFib을 통합하여 학습용 데이터를 구축하였다. 우선, 시간 축 해상도의 일관성을 확보하기 위해 모든 시그널을 128Hz로 재샘플링(resampling)하였다. 다음으로 ECG 시그널에는 0.5Hz 이상의 저주파 잡음을 제거하기 위해 컷오프 주파수 0.5Hz 하이패스 버터워스 필터(Butterworth filter)를 적용하였다. 또한 PPG 시그널에는 0.5Hz에서 8Hz 사이의 주파수 성분만 통과 시키는 밴드패스 버터워스 필터를 적용하여 호흡 및 움직임에 의한 저주파 잡음과 고주파 잡음을 동시에 제거하였다. ECG 및 PPG 시그널에 적용된 버터워스 필터는 바이오 시그널 처리에서 널리 사용되며, 주파수 응답 특성이 평탄하여 대역 내 진폭 왜곡을 최소화하는 장점이 있다(Chatterjee *et al.*, 2018; Liu *et al.*, 2016). 필터링 이후 각 피험자별 시그널 진폭과 분포 차이를 보정하기 위해 z-score 표준화를 수행하였으며, 이렇게 정규화 된 시그널을 다시 [-1, 1] 범위로 min-max scaling하여 모델 입력의 수치 안정성을 확보하였다. 전처리된 시그널은 4초 길이의 윈도우로 세분화하였으며, 시계열의 연속성을 유지하기 위해 각 윈도우는 10% 중첩되도록 구성하였다. 최종적으로 전체 데이터의 80%는 학습용, 20%는 테스트용으로 분할하여 모델 학습과 평가에 활용하였으며, 학습용 데이터와 테스트 데이터에 동일한 환자에 동시에 포함되지 않도록 구성하였다. 이와 같은 일련의 전처리 과정을 통해 서로 이질적인 데이터셋 간 특성 차이를 최소화하고, 학습 과정에서 안정적이고 일관된 입력 신호 특성을 확보할 수 있도록 하였다.

### 4.2 실험 세팅

본 연구에서 제안하는 방법론은 세 가지 주요 구성 요소로 이루어진다. 첫째, 시그널 생성의 핵심인 메인 모델은 DDPM 기반 노이즈 예측기와 ROI 재구축기로 구성된 두 개의 서브 모듈을 한 쌍으로 구성하여 생성 과정에서 시그널의 전역 구조와 관심 구간 정보를 동시에 학습할 수 있도록 설계되었다. 둘째, 조건 추출 모델은 PPG로부터 다중 레이어 특징 맵을 추

출하는 별도의 네트워크를 두개 사용하여, 시간 및 주파수 차원의 풍부한 정보가 디퓨전 과정의 각 단계에 조건으로 제공 되도록 설계되었다. 마지막으로 메인 모델의 각 잔차 블록에서 생성된 특징 맵과 PPG 특징 추출 네트워크를 통해 얻은 잔차 블록 별 특징 맵 간에 다중 헤드 크로스 어텐션을 적용함으로써, PPG 패턴이 효과적으로 반영된 정밀한 ECG 시그널을 생성할 수 있도록 하였다.

모델 학습은 단일 GPU 환경에서 수행되었으며, 내부적으로 병렬 처리를 지원하는 구조로 구현되어 다중 GPU 환경에서도 확장 가능하도록 설계되었다. 모든 실험은 동일한 조건에서의 재현성을 보장하기 위해 난수 시드를 고정된 상태로 수행하였다. 디퓨전 단계 수는 10, 25, 50, 100으로 설정하였고, 배치 사이즈는 64로 고정하였다. 최대 500 에포크(epoch)로 학습하되, 100 에포크 동안 성능이 개선되지 않으면 학습을 중단하도록 하여 과적합을 방지하였다. 손실 함수는 ROI 영역에 해당하는 항에 가중치  $\lambda_1 = 100$ , 전체 시그널 복원 항에  $\lambda_2 = 1$ 을 적용하여, ROI 복원에 중점을 두도록 설계하였다. 학습률 스케줄링은 초기 단계에서 선형적으로 증가시키고 이후에는 코사인 감소 방식을 적용하여 점진적으로 감소시키는 전략을 선택하였다. 비교 대상 방법인 RDDM은 오픈 소스로 제공된 공식 구현체를 활용하였으며, 본 연구에서 사용한 동일한 데이터셋을 기반으로 재학습을 수행하여 비교 분석하였다.

### 4.3 평가지표

본 연구는 생성된 ECG가 실제 ECG의 형태, 분포, 진폭을 얼마나 정확히 재현하는지를 다각적으로 검증하기 위해 평균 제곱근 오차(root mean square error, RMSE)와 프레셰 거리(Fréchet distance, FD)를 평가 지표로 사용하였다(Alt, 1995; Berndt, 1994). RMSE는 생성된 ECG인  $ECG_g$ 와 정답 ECG인  $ECG_o$ 의 진폭 차이를 전 구간에서 걸쳐 제공하고 평균한 후, 제곱근을 취한 값으로 정의된다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (ECG_g^i - ECG_o^i)^2} \quad (12)$$

여기서  $N$ 은 샘플 수를 의미하며, RMSE 값이 작을수록 두 파형의 평균 진폭이 전역적으로 일치함을 나타낸다. 그러나 RMSE는 시간 축의 위상 차이나 파형의 전역적 형태를 충분히 반영하지 못한다는 한계가 있다. 이에 따라, 본 연구에서는 추가적인 평가 지표로 FD를 활용하여 파형의 형태적 및 시간적 유사성을 평가하였다. FD는 생성된 ECG 파형과 실제 ECG 파형을 각각 하나의 다각형 곡선으로 간주하고, 두 곡선이 서로 얼마나 유사한 형태로 움직이는지를 다음 식으로 평가한다.

$$FD = \inf_{\alpha, B} \max_{t \in [0,1]} \|ECG_g(\alpha(t)) - ECG_o(B(t))\| \quad (13)$$

여기서  $\alpha$ 와  $B$ 는 두 곡선 간의 재매개화 함수이며, FD 값이 작을수록 두 곡선의 형태적 유사성이 높음을 의미한다. 본 연구는 RMSE와 FD를 함께 활용하여 생성된 ECG 파형의 진폭 일치성, 형태 유사성, 시간 정렬성을 종합적으로 평가함으로써, 생성 품질과 신뢰성을 다각도로 검증하였다.

### 4.4 동적 ROI 선정 알고리즘 효과 검증

<Table 1>은 동적 ROI 선정 알고리즘의 유효성을 평가하기 위해 수행한 비교 실험의 정량적 결과이다. 본 실험에서는 제안 알고리즘과 기존 RDDM의 ROI 선정 방식 간 성능 차이를 확인하고자, DALIA와 WESAD 두 데이터셋을 대상으로 동일한 실험 조건 하에 네 가지 디퓨전 샘플링 단계를 설정하고 RMSE 및 FD를 평가지표로 활용하였다. RDDM의 ROI 선정 방식은 탐색된 R파를 기준으로 일정한 길이의 구간을 고정적으로 선택하는 방식이며, 본 실험에서는 RDDM 원문에서 사용한 길이 값인 32를 그대로 적용하였다. DALIA 데이터셋 기준으로, 10 스텝에서 동적 ROI 모델은 RMSE 0.223, FD 15.77

**Table 1.** Comparison of ROI selection methods across 4 different diffusion sampling steps(10, 25, 50, 100) in terms of RMSE and FD scores. The best performance values among methods are highlighted in bold. Results are reported as mean and standard deviation over three independent runs. Statistical significance between RDDM and RDDM + dynamic ROI is evaluated using paired  $t$ -tests, where \* indicates  $p$ -value < 0.1 and \*\* indicates  $p$ -value < 0.05.

Sampling step		10		25		50		100	
Dataset	Method	RMSE	FD	RMSE	FD	RMSE	FD	RMSE	FD
DALIA	RDDM	0.316 (0.027)	25.21 (3.48)	0.318 (0.014)	31.62 (5.77)	0.533 (0.050)	139.55 (30.63)	0.668 (0.256)	243.61 (165.12)
	RDDM + dynamic ROI	<b>0.223**</b> <b>(0.027)</b>	<b>15.77*</b> <b>(6.92)</b>	<b>0.243**</b> <b>(0.018)</b>	<b>8.91**</b> <b>(0.85)</b>	<b>0.299**</b> <b>(0.024)</b>	<b>26.11**</b> <b>(10.36)</b>	<b>0.366*</b> <b>(0.105)</b>	<b>58.23*</b> <b>(48.69)</b>
WESAD	RDDM	0.316 (0.026)	25.80 (3.77)	0.317 (0.014)	32.18 (5.62)	0.529 (0.048)	138.14 (29.22)	0.670 (0.261)	246.27 (168.36)
	RDDM + dynamic ROI	<b>0.219*</b> <b>(0.023)</b>	<b>15.37**</b> <b>(6.12)</b>	<b>0.243**</b> <b>(0.019)</b>	<b>9.357**</b> <b>(0.901)</b>	<b>0.298**</b> <b>(0.025)</b>	<b>26.54**</b> <b>(10.08)</b>	<b>0.367*</b> <b>(0.109)</b>	<b>59.94*</b> <b>(50.73)</b>

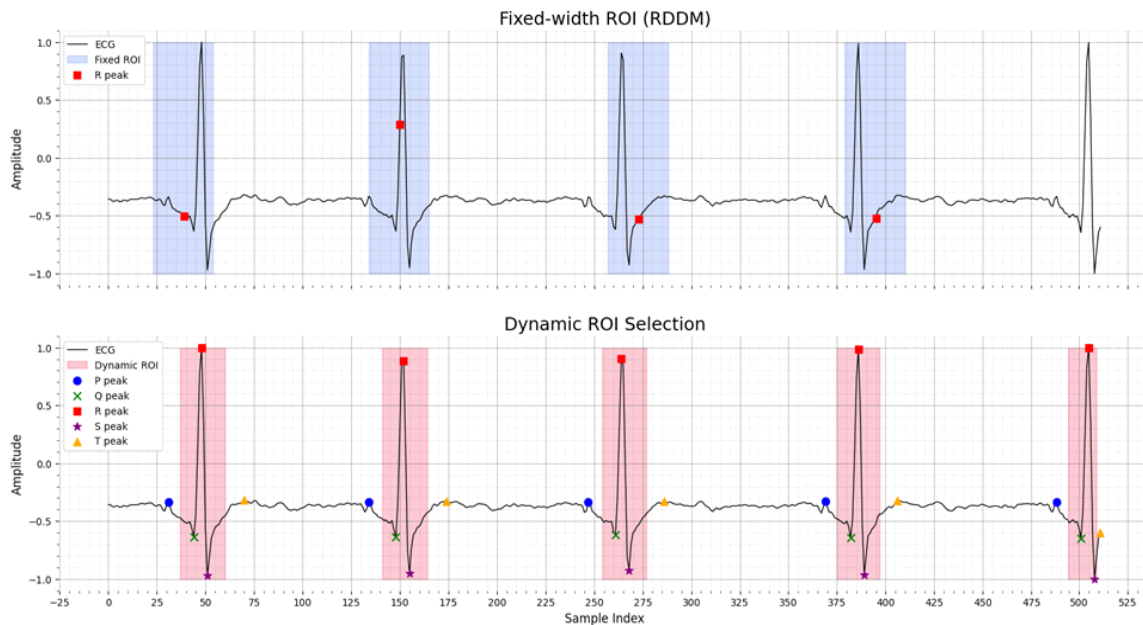
을 기록하여 RDDM 대비 각각 29.4%, 37.5% 개선되었다. 25 스텝 구간에서도 RMSE와 FD는 각각 23.6%, 72.0% 감소하며 가장 두드러진 개선을 보였다. 이후 50 스텝과 100 스텝에서도 RMSE는 각각 5.0%, 2.9% 줄었고, FD는 14.4%, 70.7% 줄어들어 전반적으로 동적 ROI의 효용성이 유지되었다. WESAD 데이터셋에서도 유사한 경향이 나타났다. 10 스텝에서 RMSE는 0.219, FD는 15.37로 각각 30.7%, 40.5% 개선되었고, 25 스텝에서도 RMSE 0.243, FD 9.36으로 성능 향상이 뚜렷했다. 50 스텝과 100 스텝에서도 RMSE는 각각 5.9%, 3.3% 감소했으며, FD는 21.7%, 75.8% 줄어들어 동적 ROI 선정 알고리즘의 안정적인 효과가 입증되었다. 두 데이터셋 전체 스텝 평균 기준, 동적 ROI는 DALIA에서 RMSE가 0.445에서 0.424로 약 4.7%, FD가 114.2에서 61.7로 약 46.0% 감소하였다. WESAD에서는 RMSE가 0.446에서 0.426으로 약 4.5%, FD가 123.2에서 42.2로 약 65.7% 감소하였다. 이는 정량적 오차와 파형 유사성 모두에서 유의미한 향상임을 보여준다. 특히 DALIA와 WESAD 모두에서 가장 큰 개선 효과는 25 스텝에서 관측되었으며, 이는 동적 ROI가 QRS 복합체 뿐만 아니라 P파와 T파를 포함한 다양한 형상 요소를 균형 있게 포착함으로써 파형 복원력을 극대화한 결과로 해석된다.

<Figure 3>은 제안한 동적 ROI 선정 알고리즘의 효과를 정성적으로 시각화한 결과이다. 상단 그래프는 기존 RDDM의 ROI 지정 방식을 보여주며, 모든 심박 주기에 대해 일정한 길이의 ROI를 일률적으로 적용하였다. 이로 인해 R파 주변의 급격한 기울기만 부분적으로 포착되고, P, Q, S, T파처럼 저진폭이거나 위치 변동성이 큰 세부 특징은 ROI 외부로 누락되는

현상이 발생하였다. 그 결과, ROI 내부에서는 고주파 성분이 과도하게 학습되었고, ROI 외부에는 노이즈가 잔존하여 재현된 파형의 정밀도가 저하되는 한계가 있었다. 또한, RDDM에 적용된 R 피크 탐지 모듈의 성능이 낮아 R 피크 자체를 안정적으로 식별하지 못하는 문제도 나타났다. 반면, 하단 그래프는 제안한 동적 ROI 선정 알고리즘의 적용 결과로, 각 심박 주기에 따라 ROI의 위치와 폭을 동적으로 조정함으로써 P, Q, R, S, T파가 모두 포함되도록 하였다. 특히 R파를 중심으로 전후 구간이 균형 있게 포함되어 수축기와 이완기 전반의 전위 변화를 포괄할 수 있었으며, 심박 간격이 변화하는 상황에서도 ROI가 자동으로 재배치되어 심박 변이도와 같은 시간적 특징이 왜곡 없이 반영되었다. 결과적으로, 제안 동적 ROI 선정 기법은 기존 RDDM 대비 ECG의 주요 구성 요소인 QRS 복합체를 보다 정밀하게 포착함으로써 파형 복원력을 향상시켰다. 이는 <Table 1>에서 확인된 FD 감소와도 직접적으로 연관되며, 정량적 지표 외에도 정성적 측면에서의 향상 효과를 입증하였다. 또한, 두 데이터셋의 모든 샘플링 스텝에 대해 RDDM과 제안한 동적 ROI 모델 간 성능 차이를 paired *t*-test로 검증한 결과, *p*-value가 0.05 미만으로 나타나 귀무 가설을 기각하였다. 이는 RMSE와 FD에서의 개선이 통계적으로 유의미함을 보여준다.

#### 4.5 PPG-to-ECG 디퓨전 모델 효과 검증

<Table 2>는 제안한 모델과 기존 방법인 RDDM 간의 정량적 비교 결과를 제시한다. 제안 모델은 동적 ROI 선정 기법,



**Figure 3.** Qualitative comparison between the fixed-width ROI selection used in RDDM and the proposed dynamic ROI selection. Selected ROIs are highlighted with blue (RDDM) boxes and red (proposed) boxes. The characteristic peaks used to define each ROI are also indicated.

DiffWave 기반 백본, 그리고 PPG 조건 기반 크로스 어텐션으로 구성되어 있으며, 실험은 다섯 개의 공개 데이터셋 (DALIA, WESAD, BIDMC, CAPNO, MIMIC)을 대상으로 수행되었다. 모든 실험은 샘플링 스텝을 10, 25, 50, 100으로 수행되었으며, 평가지표로는 진폭 오차를 나타내는 RMSE와 파형의 형상 유사성을 측정하는 FD를 사용하였다. 샘플링 스텝이 가장 짧은 10스텝에서 제안 모델은 모든 데이터셋에서 RMSE와 FD가 크게 감소되었다. 예를 들어, DALIA에서는 RMSE가 RDDM의 0.440에서 0.209로 감소하였고, FD는 86.15에서 15.53로 줄었다. 유사한 개선은 WESAD, BIDMC, CAPNO, MIMIC에서도 반복되었으며, 다섯 개 데이터셋의 평균 RMSE는 약 0.441에서 0.211으로 절반 수준으로 줄었고, 평균 FD는 약 88.29에서 14.21로 크게 개선되었다. 특히 잡음이 많은 BIDMC에서도 FD가 93.66에서 16.00로 감소하여, 제안 모델의 강건한 노이즈 대응 능력을 확인할 수 있었다. 샘플링 스텝을 25, 50, 100으로 증가시켜도 제안 모델의 성능은 일관되게 유지되었다. RMSE는 0.192에서 0.233 사이의 좁은 범위 내에서 안정적으로 유지되었으며, FD 역시 12.01에서 17.00 사이로 낮은 값을 기록하였다. 반면, RDDM은 스텝 수가 증가할수록 RMSE가 약 0.478에서 0.643, FD가 70 이상으로 높아지는 경향을 보였다. 다섯 개 데이터셋 전반에서 평균적으로 제안 모

델은 RMSE와 FD 모두에서 뚜렷한 성능 우위를 보였고, 특히 FD의 감소 폭이 RMSE보다 훨씬 크다는 점은, 단순한 진폭 일치 뿐 아니라 QRS 복합체와 P파, T파 등 세부 파형 구조까지 정밀하게 재현했음을 시사한다. 또한, 모든 데이터셋의 모든 샘플링 스텝 조건에서 RDDM과 제안한 모델 간 성능 차이를 paired *t*-test를 통해 검증하였다. 그 결과, 대부분의 실험결과 *p*-value가 0.05 미만으로 나타나 제안 모델의 성능 향상이 통계적으로 강한 유의성을 보여주었다. 또한, RDDM의 FD 지표는 표준편차가 20 이상으로 과도하게 산출되는 경우가 빈번하게 나타난 반면, 제안한 방법론의 FD 표준편차는 모든 경우에서 3에서 8 사이로 안정적으로 관찰되었다. 이는 제안한 모델이 기존 방법론에 비해 결과의 변동성이 현저히 낮으며, 보다 일관된 성능을 제공함을 시사한다.

<Figure 4>는 샘플링 스텝 10 조건에서 BIDMC 데이터셋의 한 사례를 통해 제안 모델과 기존 RDDM 간의 정성적 차이를 시각적으로 비교한 결과이다. 첫 번째 그래프는 입력 조건으로 사용된 PPG 신호를, 두 번째는 실제 ECG를, 세 번째는 RDDM이 생성한 ECG를, 마지막 네 번째는 제안 모델이 생성한 ECG를 각각 나타낸다. RDDM의 결과는 전반적으로 진폭 변동이 크고 QRS 복합체의 형상이 불명확했으며, 파형 전반에 왜곡이 분포되어 있는 것 확인할 수 있었다. 반면, 제안 모

**Table 2.** Quantitative comparison between RDDM and the proposed method across 4 different diffusion sampling steps (10, 25, 50, 100) in terms of RMSE and FD scores. The best performance values among methods are highlighted in bold. Results are reported as mean and standard deviation over three independent runs. Statistical significance between RDDM and proposed method is evaluated using paired *t*-tests, where \* indicates *p*-value < 0.1 and \*\* indicates *p*-value < 0.05.

Sampling step		10		25		50		100	
Dataset	Method	RMSE	FD	RMSE	FD	RMSE	FD	RMSE	FD
DALIA	RDDM	0.440 (0.037)	86.15 (17.64)	0.514 (0.049)	129.09 (26.78)	0.558 (0.083)	154.46 (44.40)	0.628 (0.095)	175.05 (33.78)
	Proposed method	<b>0.209**</b> <b>(0.011)</b>	<b>15.53**</b> <b>(3.93)</b>	<b>0.218**</b> <b>(0.016)</b>	<b>13.89**</b> <b>(5.87)</b>	<b>0.218**</b> <b>(0.018)</b>	<b>15.33**</b> <b>(4.55)</b>	<b>0.202**</b> <b>(0.006)</b>	<b>13.78**</b> <b>(4.72)</b>
WESAD	RDDM	0.433 (0.040)	84.19 (18.44)	0.507 (0.052)	126.07 (28.00)	0.557 (0.084)	154.43 (44.82)	0.624 (0.096)	153.11 (39.83)
	Proposed method	<b>0.203**</b> <b>(0.009)</b>	<b>12.78**</b> <b>(4.04)</b>	<b>0.203**</b> <b>(0.013)</b>	<b>13.10**</b> <b>(4.34)</b>	<b>0.208**</b> <b>(0.019)</b>	<b>13.57**</b> <b>(5.15)</b>	<b>0.192**</b> <b>(0.006)</b>	<b>14.02**</b> <b>(2.37)</b>
BIDMC	RDDM	0.457 (0.023)	93.66 (11.64)	0.528 (0.046)	136.40 (25.35)	0.565 (0.088)	126.72 (26.38)	0.643 (0.099)	175.83 (30.20)
	Proposed method	<b>0.224**</b> <b>(0.014)</b>	<b>16.00**</b> <b>(7.20)</b>	<b>0.233**</b> <b>(0.011)</b>	<b>16.54**</b> <b>(7.02)</b>	<b>0.229**</b> <b>(0.013)</b>	<b>17.10**</b> <b>(7.38)</b>	<b>0.227**</b> <b>(0.006)</b>	<b>15.24**</b> <b>(6.74)</b>
CAPNO	RDDM	0.473 (0.031)	102.88 (16.22)	0.537 (0.042)	141.97 (24.45)	0.571 (0.089)	156.61 (39.35)	0.651 (0.102)	194.60 (49.87)
	Proposed method	<b>0.201**</b> <b>(0.010)</b>	<b>13.00**</b> <b>(3.53)</b>	<b>0.201**</b> <b>(0.012)</b>	<b>12.38**</b> <b>(4.41)</b>	<b>0.205**</b> <b>(0.020)</b>	<b>13.53**</b> <b>(4.43)</b>	<b>0.197**</b> <b>(0.010)</b>	<b>12.01**</b> <b>(3.33)</b>
MIMIC	RDDM	0.404 (0.115)	74.57 (43.95)	0.478 (0.123)	97.17 (39.69)	0.517 (0.179)	105.97 (57.79)	0.584 (0.208)	128.69 (83.67)
	Proposed method	<b>0.217**</b> <b>(0.008)</b>	<b>13.76*</b> <b>(5.90)</b>	<b>0.219**</b> <b>(0.012)</b>	<b>14.92**</b> <b>(4.87)</b>	<b>0.219**</b> <b>(0.017)</b>	<b>14.97*</b> <b>(6.26)</b>	<b>0.208**</b> <b>(0.002)</b>	<b>14.06*</b> <b>(4.56)</b>



**Figure 4.** Qualitative comparison between the conditioning PPG signal (first plot), the corresponding ground-truth ECG (second plot), the ECG generated with RDDM (third plot), and the ECG generated by the proposed diffusion model (bottom plot) on the BIDMC dataset. The proposed method more accurately reproduces the sharp QRS complexes and overall morphological patterns of the original ECG, while the RDDM output exhibits noticeable amplitude fluctuations and waveform distortions.

델의 생성 결과는 실제 ECG와 유사한 QRS 첨도, S파의 급강하, T파 회복 구간 등의 세부 구조를 효과적으로 재현하였다. 진폭의 안정성과 배경 잡음 억제로 인해 파형은 전반적으로 매끄럽게 이어지며, 심박 주기 간 간격 또한 실제 신호와 일치하였다. 이러한 결과는 동적 ROI가 QRS 복합체를 중심으로 노이즈를 선택적으로 제어하고, PPG 조건 정보를 크로스 어텐션을 통해 융합함으로써 전역적 구조와 국소적 디테일을 동시에 보정했기 때문이다. 또한 백본 모델로 시계열 데이터에 특화된 DiffWave 구조를 채택함으로써, 이미지 기반 U-Net과 달리 ECG와 같은 연속적 신호의 시간적 특성을 보다 효과적으로 반영하였음을 시사한다.

#### 4.6 제안 방법론과 트랜스포머 및 GAN기반 방법론 비교 실험

<Table 3>은 제안 방법과 디퓨전 기반이 아닌 일반적 방법론(트랜스포머 기반, GAN 기반)의 성능을 다섯 개의 공개 데이터셋에서 비교한 결과를 보여준다. 트랜스포머 기반 방법은 Chiu *et al.*(2020)을 재현하여 공개 코드의 기본 하이퍼 파라미터를 사용하되, 학습 에포크를 제안 방법과 동일하게 500으로 설정하였다. GAN 기반 방법은 Sarkar *et al.*(2021)의 CardioGAN에서 보고된 수치를 인용하여 동일 데이터셋 기준으로 비교하였다. 제안 방법론은 샘플링 스텝 10의 결과를 사용하였으며,

**Table 3.** Quantitative comparison of the proposed method against transformer and GAN-based(CardioGAN) baselines on 5 datasets. Metrics are RMSE and FD and the best value for each dataset and metric is highlighted in bold.

Dataset	DALIA		WESAD		BIDMC		CAPNO		MIMIC	
	RMSE	FD	RMSE	FD	RMSE	FD	RMSE	FD	RMSE	FD
Transformer	0.34	<b>10.6</b>	0.29	<b>10.13</b>	0.50	42.71	0.38	18.82	0.66	18.75
CardioGAN	0.42	27.51	0.37	29.15	0.63	154.64	0.38	31.10	0.54	99.03
Proposed method	<b>0.21</b>	11.52	<b>0.193</b>	11.17	<b>0.22</b>	<b>15.82</b>	<b>0.19</b>	<b>10.8</b>	<b>0.21</b>	<b>12.48</b>

평가 지표로는 RMSE와 FD를 채택하였다. 전반적으로 제안 방법은 모든 데이터셋에서 가장 낮은 RMSE를 기록하였다. 평균 RMSE는 0.205로, 트랜스포머의 0.434와 GAN의 0.468 대비 각각 52.9%, 56.3% 낮았다. 평균 FD 또한 제안 방법론이 12.36으로, 트랜스포머의 20.20과 GAN의 68.29 보다 낮았으며, 감소율은 각각 38.8%, 81.9%였다. 이는 동적 ROI 선정으로 QRS 복합체를 집중적으로 학습하고, 시계열 데이터에 특화된 DiffWave 백본과 크로스 어텐션을 통해 전역 심박 주기 구조와 국소적 파형 디테일을 동시에 포착했기 때문으로 해석된다. 데이터셋별로 살펴보면, 비교적 안정적인 신호 환경의 DALIA와 WESAD에서는 트랜스포머가 FD에서 근소한 우위를 보였으나, 제안 방법의 RMSE가 현저히 낮았다. 이는 트랜스포머의 시퀀스 정렬 및 보정 특성이 매끈한 파형 형상에는 유리하기 때문으로 보인다. 하지만 잡음과 왜곡이 상대적으로 큰 BIDMC, CAPNO, MIMIC에서는 트랜스포머 기반 방법론의 성능이 저하된 반면, 제안 방법론은 FD와 RMSE 모두에서 뚜렷한 우위를 보이며 안정적인 성능을 유지하였다.

## 5. 결론

본 연구는 장기간 연속적인 심전도 계측이 어려운 환경에서 심혈관 이상을 조기에 탐지하기 위한 방안으로, PPG 시그널만을 이용해 고품질 ECG를 생성하는 동적 ROI 기반 디퓨전 모델을 제안하였다. 제안 방법론은 다음과 같은 세 가지 핵심 구성으로 이루어진다. 첫째, P, Q, R, S, T 파형의 위치 정보를 기반으로 QRS 복합체를 완전히 포함하는 동적 ROI 선정 알고리즘을 설계하였다. 본 알고리즘은 별도의 하이퍼파라미터 조정 없이, 시그널별·개인별로 상이한 심박 구조에 자동으로 적용할 수 있다. 둘째, 시계열 신호에 특화된 DiffWave 아키텍처를 백본 모델로 채택함으로써, ECG 파형의 전역적 리듬과 국소적 형태를 동시에 학습할 수 있도록 하였다. 셋째, PPG 조건 정보를 잔차 블록의 최종 특징 맵에 다중 헤드 크로스 어텐션 형태로 통합함으로써, 심박 간 리듬과 같은 전역적 시간 정보를 반영하는 동시에 QRS 복합체와 같은 핵심 구간의 표현력을 강화하였다. 또한 학습 전략 측면에서는 ROI 영역에는 노이즈 손실 가중치를 100으로, 전역 손실 가중치는 1로 설정해 모델이 핵심 구간 복원에 집중하도록 유도하였다.

제안 모델의 효용성을 검증하기 위해 WESAD, DALIA, BIDMC, CAPNO, MIMIC 다섯 개의 공개 데이터셋을 통합한 학습 데이터를 구축하고, 디퓨전 샘플링 스텝을 10, 25, 50, 100으로 다양화하여 실험을 수행하였다. 성능 평가는 진폭 오차와 파형 형상 유사성을 동시에 고려하기 위해 RMSE와 FD를 지표로 사용하였다. 실험 결과, 샘플링 스텝 10 조건에서 제안 모델은 RDDM 대비 평균 RMSE를 0.441에서 0.211로 52%, FD를 88.29에서 13.67으로 85% 감소시켰다. 스텝 수가 증가한 25, 50, 100 조건에서도 우수한 성능을 유지하였으며, 전체 평

균 기준으로 RMSE는 0.531에서 0.211로 약 60%, FD는 129.88에서 14.42로 약 89% 개선되어 두 지표 모두에서 뚜렷한 성능 향상을 입증하였다. 특히 CAPNO와 BIDMC처럼 움직임 노이즈가 빈번하고 광학 간섭이 있는 임상 환경에서도 FD가 70 이상에서 15 이하로 감소해 모델의 견고성을 입증하였다. 정성적 비교에서도 제안 모델은 R 파 침도, S 파 급강하, T 파 회복 구간을 실제 ECG와 거의 동일한 형태로 복원해 고정폭 ROI 기반 RDDM에서 발생하던 파형 왜곡 문제를 해소하였다. 결론적으로, 본 연구는 동적 ROI 선정 기법, 시계열 데이터에 특화된 디퓨전 모델, 그리고 PPG 조건 기반 크로스 어텐션의 결합을 통해, 웨어러블 환경에서 수집된 PPG 신호만으로도 고신뢰도의 ECG 생성을 달성할 수 있음을 정량적·정성적 분석을 통해 입증하였다. 제안된 방법은 센서 부착 부담을 최소화하면서 심혈관 이상 탐지의 정확도를 향상시키며, 향후 모바일 헬스케어, 원격 생체 신호 모니터링, 스마트워치 기반 조기 경보 시스템 등 다양한 응용 분야에서 새로운 진단 및 유지보수 패러다임을 제시할 수 있을 것으로 기대된다.

그럼에도 본 연구는 몇 가지 한계점이 남아 있다. 첫째, 생성된 ECG의 진폭이 원 시그널과 완벽하게 정렬되지 않아, 절대 전위 해석이 필요한 임상 관독 과정에서 오류가 발생할 수 있다. 둘째, 원본 PPG에 심한 움직임 노이즈나 조도 변화가 포함될 경우, PPG와 ECG의 상관 구조가 깨져 QRS 복합체 위치가 불안정하게 예측되고 파형 전체가 흐려지는 현상이 나타났다. 이러한 한계를 보완하기 위해, 후속 연구에서는 시그널 전처리 단계에 스펙트럼 소거 기반의 노이즈 억제 기법을 도입하고, 다채널 PPG 융합 및 ROI 경계 보정을 위한 전위 정렬 모듈을 추가할 예정이다. 아울러 생성된 ECG의 임상적 유효성을 보다 실질적으로 평가하기 위해 부정맥 분류 등 다운스트림 태스크와의 연계를 추진하고, 다양한 환자군 및 센서 부착 위치에 대한 교차 검증을 수행함으로써 제안 방법의 일반화 가능성을 확장해 나갈 계획이다.

## 참고문헌

- Adib, E., Fernandez, A. S., Afghah, F., and Prevost, J. J. (2023), Synthetic ECG signal generation using probabilistic diffusion models, *IEEE Access*, **11**, 75818-75828.
- Alt, H. and Godau, M. (1995), Computing the Fréchet distance between two polygonal curves, *International Journal of Computational Geometry & Applications*, **5**(1-2), 75-91.
- Banerjee, R., Sinha, A., Choudhury, A. D., and Visvanathan, A. (2014), PhotoECG: Photoplethysmography to estimate ECG parameters, In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4404-4408.
- Bashar, S. K., Ding, E., Walkey, A. J., McManus, D. D., and Chon, K. H. (2019), Noise detection in electrocardiogram signals for intensive care unit patients, *IEEE Access*, **7**, 88357-88368.
- Bayoumy, K., Gaber, M., Elshafeey, A., Mhaimeed, O., Dineen, E. H., Marvel, F. A., ... and Elshazly, M. B. (2021), Smart wearable

- devices in cardiovascular care: where we are and how to move forward, *Nature Reviews Cardiology*, **18**(8), 581-599.
- Chatterjee, A. and Roy, U. K. (2018), PPG based heart rate algorithm improvement with Butterworth IIR Filter and Savitzky-Golay FIR Filter, In *2018 2nd International Conference on Electronics, Materials Engineering & Nano-technology (IEMENTech)*, 1-6.
- Chiu, H. Y., Shuai, H. H., and Chao, P. C. P. (2020), Reconstructing QRS complex from PPG by transformed attentional neural networks, *IEEE Sensors Journal*, **20**(20), 12374-12383.
- Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., and Ng, A. Y. (2019), Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, *Nature Medicine*, **25**(1), 65-69.
- Ho, J., Jain, A., and Abbeel, P. (2020), Denoising diffusion probabilistic models, *Advances in Neural Information Processing Systems*, **33**, 6840-6851.
- Islam, M. T., Zabir, I., Ahamed, S. T., Yasar, M. T., Shahnaz, C., and Fattah, S. A. (2017), A time-frequency domain approach of heart rate estimation from photoplethysmographic (PPG) signal, *Biomedical Signal Processing and Control*, **36**, 146-154.
- Johnson, A. E., Pollard, T. J., Shen, L., Lehman, L. W. H., Feng, M., Ghassemi, M., ... and Mark, R. G. (2016), MIMIC-III, a freely accessible critical care database, *Scientific Data*, **3**(1), 1-9.
- Karlen, W. (2021), CapnoBase IEEE TBME respiratory rate benchmark, *Sch. Portal Dataverse*.
- Kauffman, D. (2020), Cardiovascular disease burden, deaths are rising around the world, *American College of Cardiology*, **9**.
- Kong, Z., Ping, W., Huang, J., Zhao, K., and Catanzaro, B. (2021), DiffWave: A Versatile Diffusion Model for Audio Synthesis. In *International Conference on Learning Representations*.
- Lan, E. (2023), Performer: A novel PPG-to-ECG reconstruction transformer for a digital biomarker of cardiovascular disease detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 1991-1999.
- Lee, J., Kim, M., Park, H. K., and Kim, I. Y. (2020), Motion artifact reduction in wearable photoplethysmography based on multi-channel sensors with multiple wavelengths, *Sensors*, **20**(5), 1493.
- Liu, M., Hao, H., Xiong, P., Lin, F., Hou, Z., and Liu, X. (2018), Constructing a guided filter by exploiting the butterworth filter for ECG signal enhancement, *Journal of Medical and Biological Engineering*, **38**, 980-992.
- Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., ... and Chen, S. A. (2021), NeuroKit2: A Python toolbox for neurophysiological signal processing, *Behavior research methods*, 1-8.
- McSharry, P. E., Clifford, G. D., Tarassenko, L., and Smith, L. A. (2003), A dynamical model for generating synthetic electrocardiogram signals, *IEEE Transactions on Biomedical Engineering*, **50**(3), 289-294.
- Olier, I., Ortega-Martorell, S., Pieroni, M., and Lip, G. Y. (2021), How machine learning is impacting research in atrial fibrillation: Implications for risk prediction and future management, *Cardiovascular Research*, **117**(7), 1700-1717.
- Pimentel, M. A., Johnson, A. E., Charlton, P. H., Birrenkott, D., Watkinson, P. J., Tarassenko, L., and Clifton, D. A. (2016), Toward a robust estimation of respiratory rate from pulse oximeters, *IEEE Transactions on Biomedical Engineering*, **64**(8), 1914-1923.
- Quiroz-Juarez, M. A., Jiménez-Ramírez, O., Vazquez-Medina, R., Ryzhii, E., Ryzhii, M., and Aragon, J. L. (2018), Cardiac conduction model for generating 12 lead ECG signals with realistic heart rate dynamics, *IEEE Transactions on Nanobioscience*, **17**(4), 525-532.
- Reiss, A., Indlekofer, I., Schmidt, P., and Van Laerhoven, K. (2019), Deep PPG: Large-scale heart rate estimation with convolutional neural networks, *Sensors*, **19**(14), 3079.
- Ronneberger, O., Fischer, P., and Brox, T. (2015), U-Net: Convolutional Networks for Biomedical Image Segmentation, In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234-241.
- Sarkar, P. and Etemad, A. (2021), Cardiogan: Attentive generative adversarial network with dual discriminators for synthesis of ecg from ppg. In *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**(1), 488-496.
- Schmidt, P., Reiss, A., Duerichen, R., Marberger, C., and Van Laerhoven, K. (2018), Introducing wesad, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, 400-408.
- Shome, D., Sarkar, P., and Etemad, A. (2024), Region-disentangled diffusion model for high-fidelity ppg-to-ecg translation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, **38**(13), 15009-15019.
- Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., and Ganguli, S. (2015), Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning*, 2256-2265.
- Pan, J. and Tompkins, W. J. (1985), A real-time QRS detection algorithm, *IEEE Transactions on Biomedical Engineering*, **32**(3), 230-236.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... and van Mulbregt, P. (2020), Fundamental algorithms for scientific computing in python and SciPy 1.0 contributors, *SciPy 1.0. Nat.Methods*, **17**, 261-272.
- Vo, K., El-Khamy, M., and Choi, Y. (2024), PPG-to-ECG Signal Translation for Continuous Atrial Fibrillation Detection via Attention-based Deep State-Space Modeling. In *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1-7.
- Xia, Y., Wang, W., and Wang, K. (2023), ECG signal generation based on conditional generative models, *Biomedical Signal Processing and Control*, **82**, 104587.
- Zhu, F., Ye, F., Fu, Y., Liu, Q., and Shen, B. (2019), Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network, *Scientific Reports*, **9**(1), 6734.

## 저자소개

**추창욱:** 고려대학교 산업경영공학부에서 2023년 학사 학위를 취득하고, 고려대학교 산업경영공학과에서 석사 학위를 취득하였다. 연구 분야는 Diffusion Model, Generative Model이다.

**김성범:** 고려대학교 산업경영공학부 교수로 2009년부터 재직하고 있으며, 인공지능공학연구소 소장, 기업산학연협력센터 센터장, 한국데이터마케팅학회 회장을 역임했다. 미국

University of Texas at Arlington 산업공학과에서 교수를 역임하였으며, 한양대학교 산업공학과에서 학사학위를 미국 Georgia Institute of Technology에서 산업시스템공학 석사 및 박사학위

를 취득하였다. 인공지능, 머신러닝, 최적화 방법론을 개발하고 이를 다양한 공학, 자연과학, 사회과학 분야에 응용하는 연구를 수행하고 있다.