창의융합프로젝트 Project 3

2021. 05. 14.

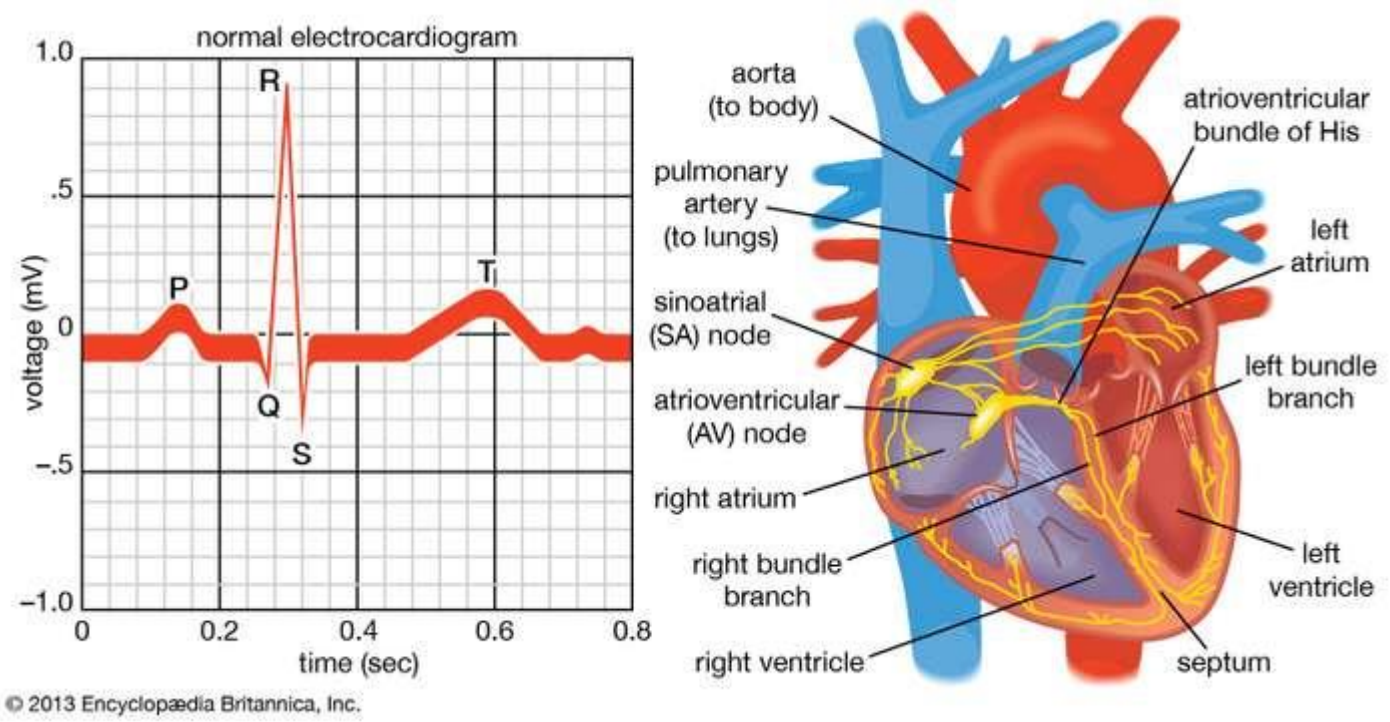
서장원
Seoul National University
Graduate School of Convergence Science and Technology
Applied Data Science Lab.

ADS Lab.

Electrocardiogram

Representation of the electrical activity of the heart

Electrocardiogram (ECG, 심전도)

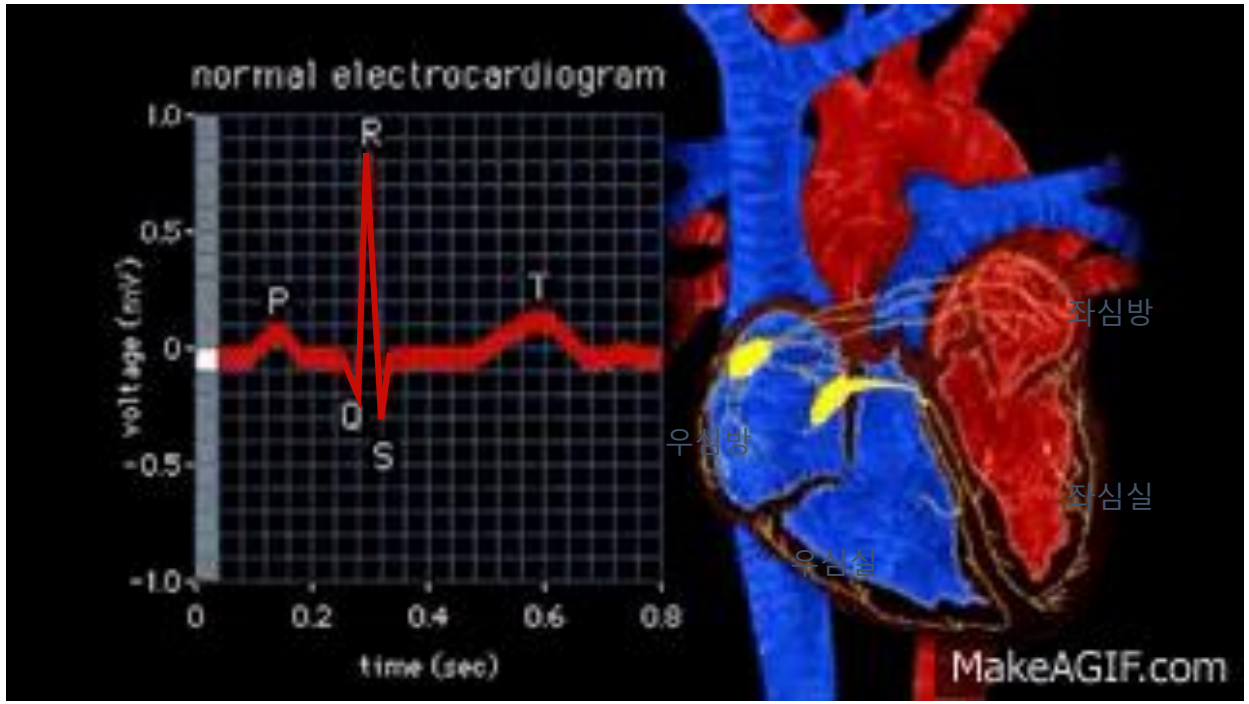


https://www.britannica.com/science/electrocardiography

http://www.internistsassoc.org/electrocardiography.html

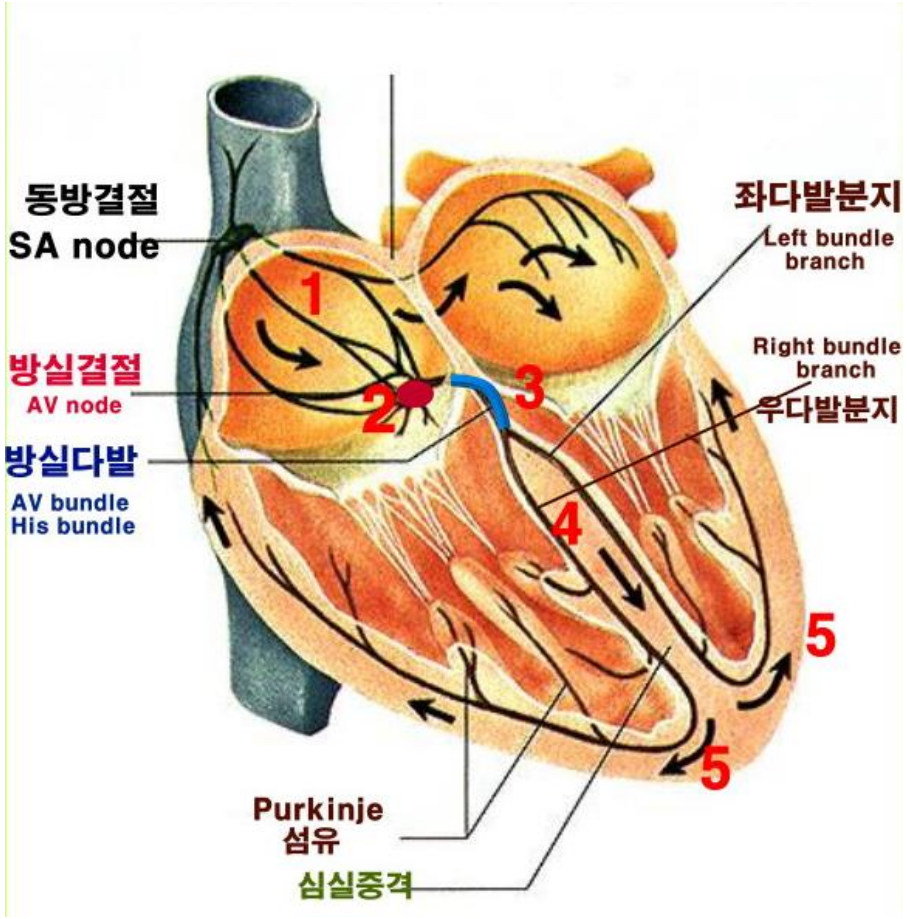
- 몸에 전극을 붙여 전위를 측정함으로써 심장의 활동전류를 기록.

Electrocardiogram (ECG, 심전도)



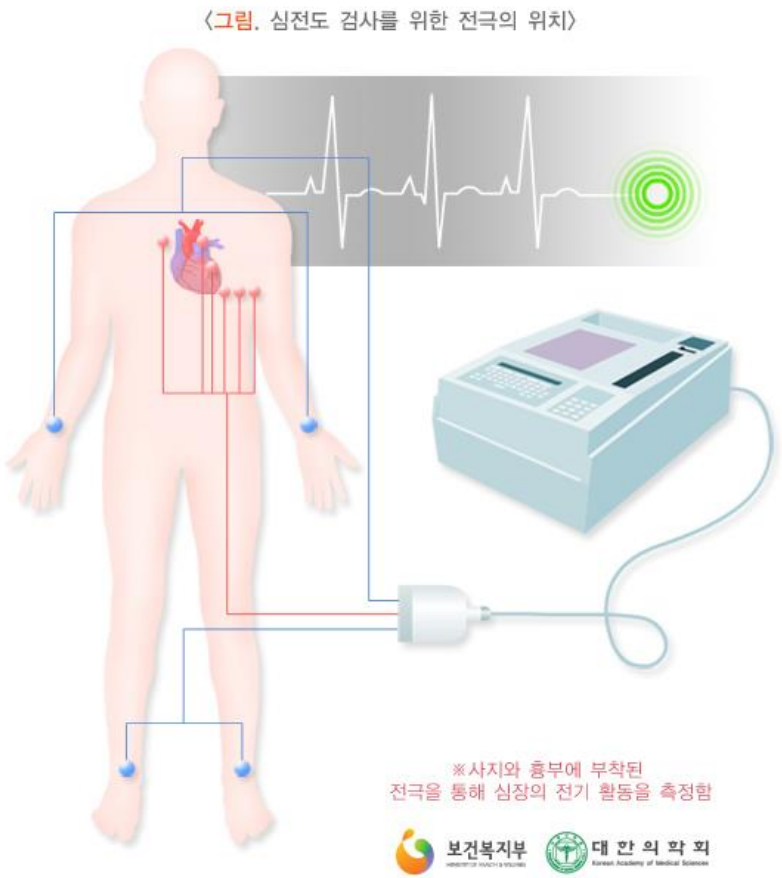
https://makeagif.com/i/rGaGQ9

Electrocardiogram (ECG, 심전도)



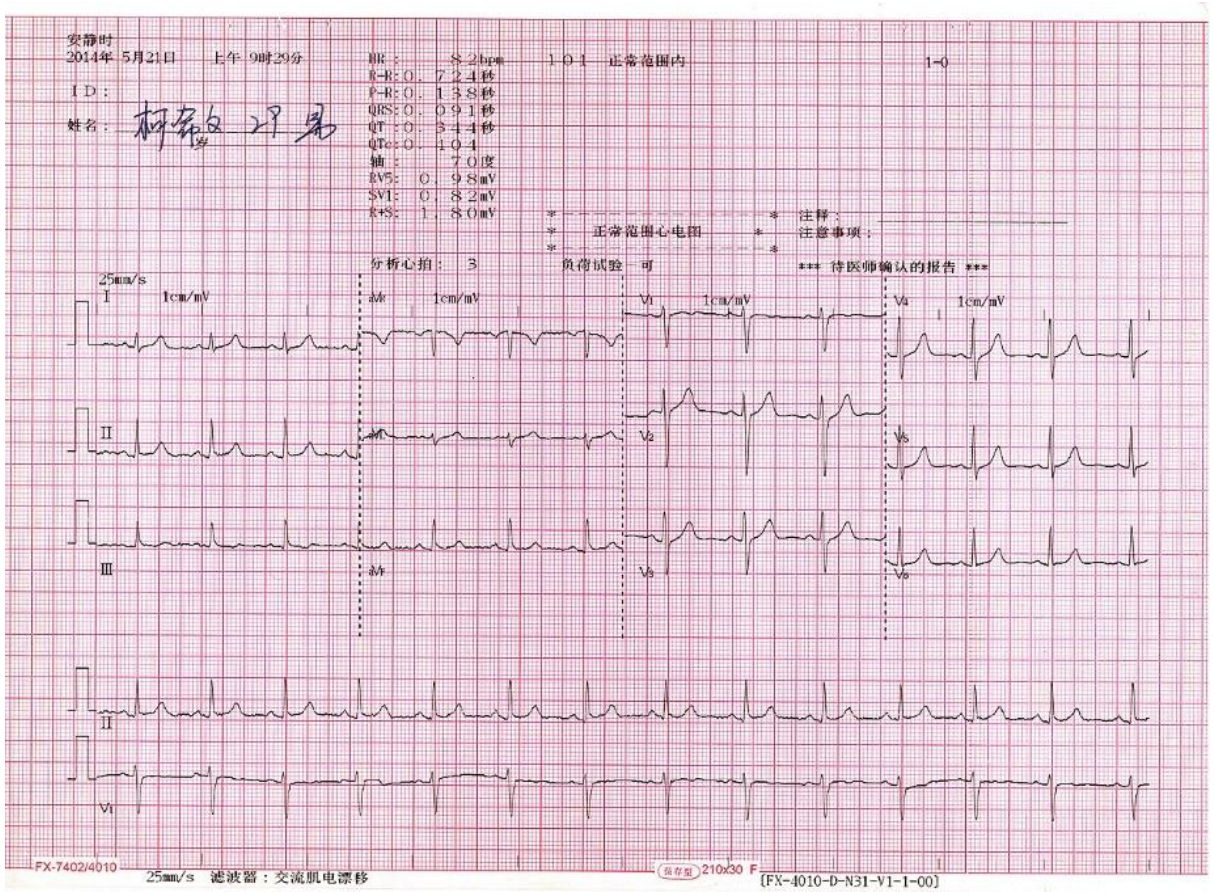
https://heartguide.tistory.com/entry/심장의-생리-심장전도계-Conduction-system-부정맥-Arrhythmia-%E2%91%A1

ECG Measurement



http://health.cdc.go.kr/

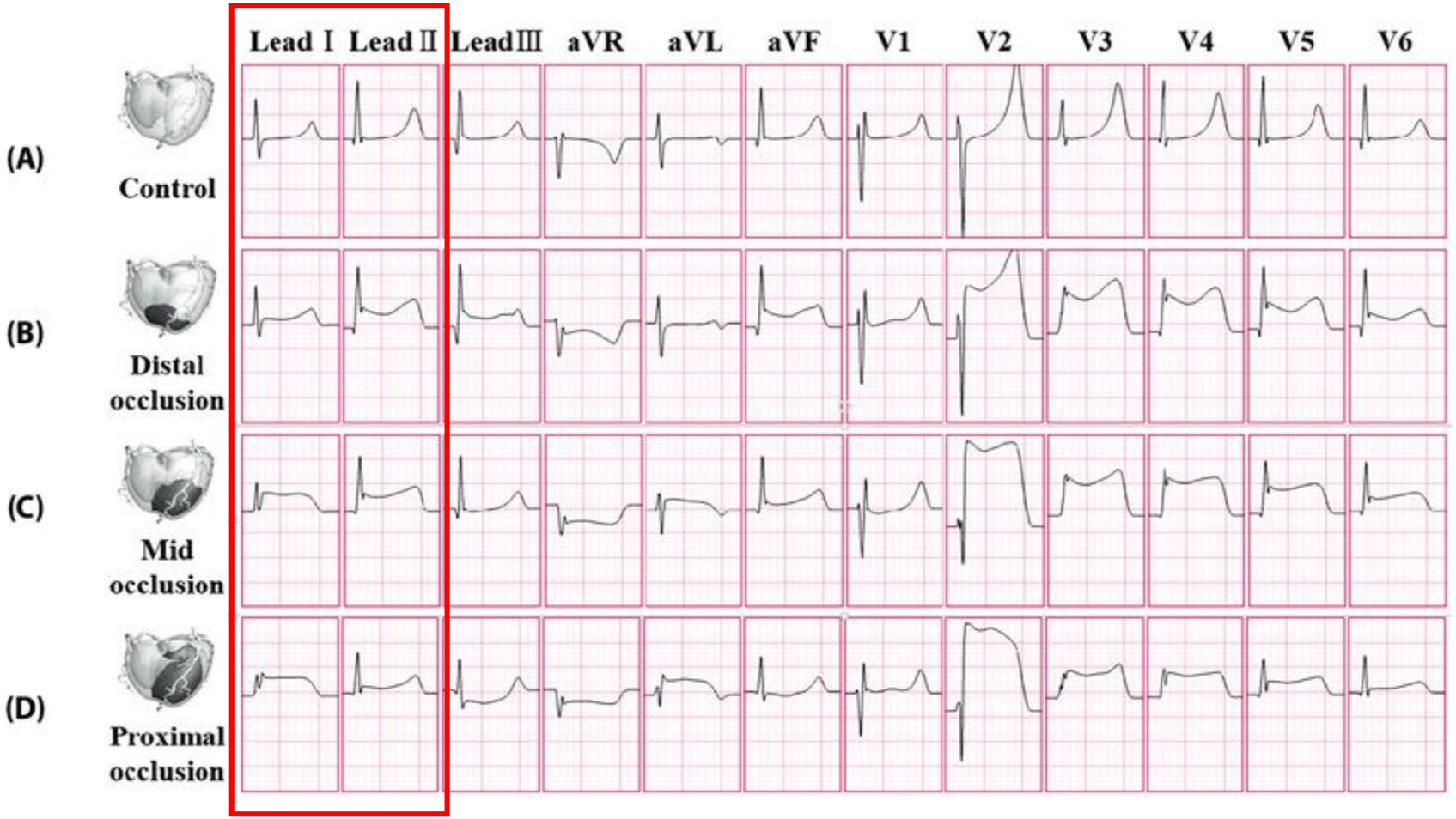
- Using 10 electrodes,
- Record a snapshot of 12 different views(leads) of the electrical activity of a heart.



https://en.wikipedia.org/wiki/Electrocardiography

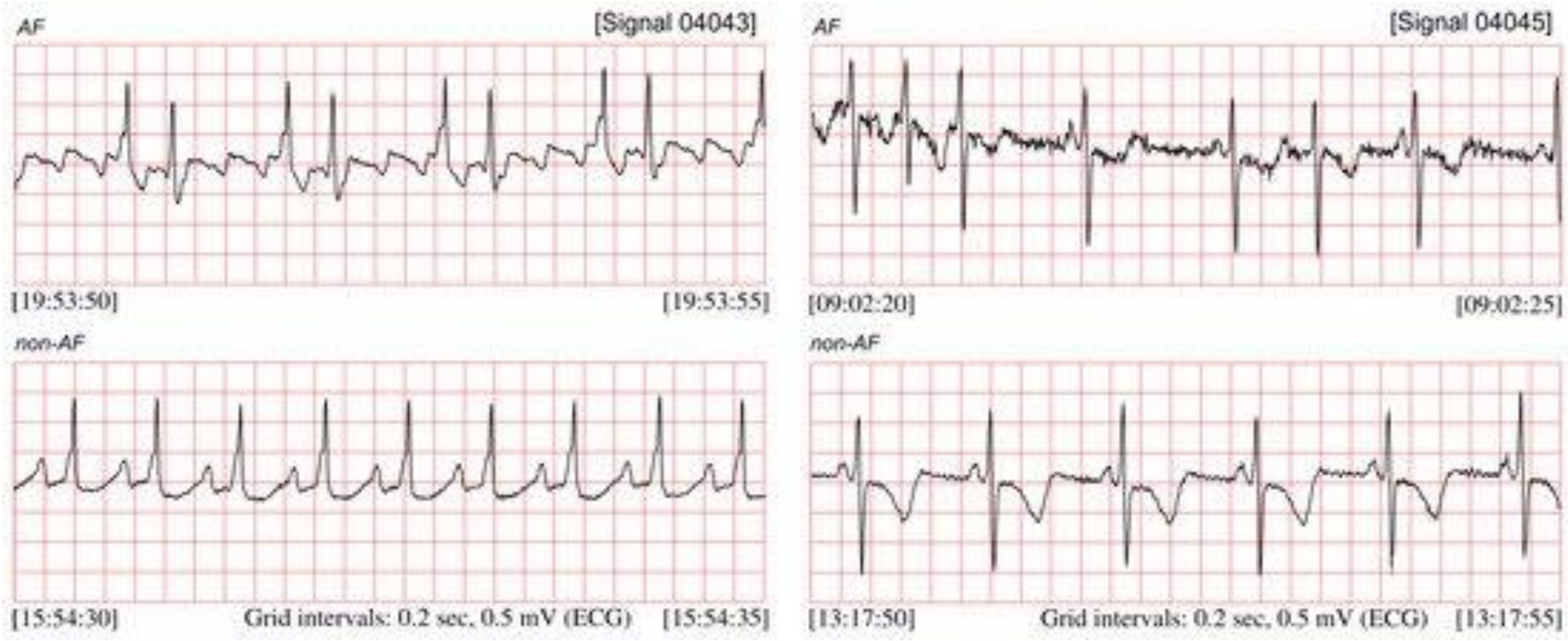
12-lead ECG

We will use ECGs of lead I, lead II in this challenge



https://www.researchgate.net/figure/12-lead-ECG-waveform-in-10-minutes-after-coronary-occlusion-A-control-model-without_fig4_330107119

Detecting heart disease using ECG



From *Detection of Atrial Fibrillation Episodes in Long-Term Heart Rhythm Signals Using a Support Vector Machine*

Atrial Fibrillation(AF): 심방 세동 – 심방이 무질서하게 매우 빠르고 미세하게 떨리면서 불규칙한 맥박을 형성하는 부정맥 질환의 일종

Project 3

Classification of cardiac abnormalities from two-lead ECGs

Project 3

- Competition due: 6/4 (금) 오후 12:59 (수업 시간 1시간 전까지)
- 최종 발표 자료 및 코드 제출 due: 6/10 (목) 오후 11:59
- Submission 제한: 1인당 하루 최대 3회

Project 3

- 개인 과제 → 팀 과제

- 개인 과제: “Baseline Example – PyTorch”의 hyperparameter를 바꿔가며 public leaderboard 기준 0.403이 넘는 submission 만들어내기

#	Team Name	Notebook	Team Members	Score	Entries	Last
	baseline_example_CNN.csv			0.39853		

- 5/21(금) 수업시간 전까지 간단한 자료 / 코드 제출
- 이후 team merger, 4인 1팀으로 project 3 진행

Classification from two-lead ECGs

- ECGs of lead I, lead II are provided
- Multi-label classification: A sample is labeled with one or more labels
- 12 labels used in this competition
  - 0: 1st degree av block
  - 1: Atrial fibrillation
  - 2: Left anterior fascicular block
  - 3: Left axis deviation
  - 4: Premature atrial contraction
  - 5: Q wave abnormal
  - 6: Sinus arrhythmia
  - 7: Sinus bradycardia
  - 8: Sinus rhythm
  - 9: Sinus tachycardia
  - 10: T wave abnormal
  - 11: T wave inversion

Classification from two-lead ECGs

파일 구성

`train.zip`과 `test.zip`을 다운로드 받으실 수 있습니다. 각 폴더 안에는 `{id}.txt` 파일과 `{id}.npy` 파일이 쌍으로 존재하며, `{id}.txt` 파일엔 해당 sample 환자의 나이, 성별, 심전도 판독(label) 정보가, `{id}.npy` 파일엔 해당 sample의 심전도 recording 정보가 저장되어 있습니다.

`train` 폴더 안에는 19212개의 sample(id: 0~19211)이, `test` 폴더 안에는 7389개의 sample(id:0~7388)이 들어있습니다. `train` 폴더 안의 sample과 `test` 폴더 안의 sample은 id가 같더라도 서로 다른 sample입니다.

{id}.txt

Ex) `train/1.txt`

`#Age: 76.0 -> 환자의 나이`

`#Sex: F -> 환자의 성별 (M: 남성, F: 여성)`

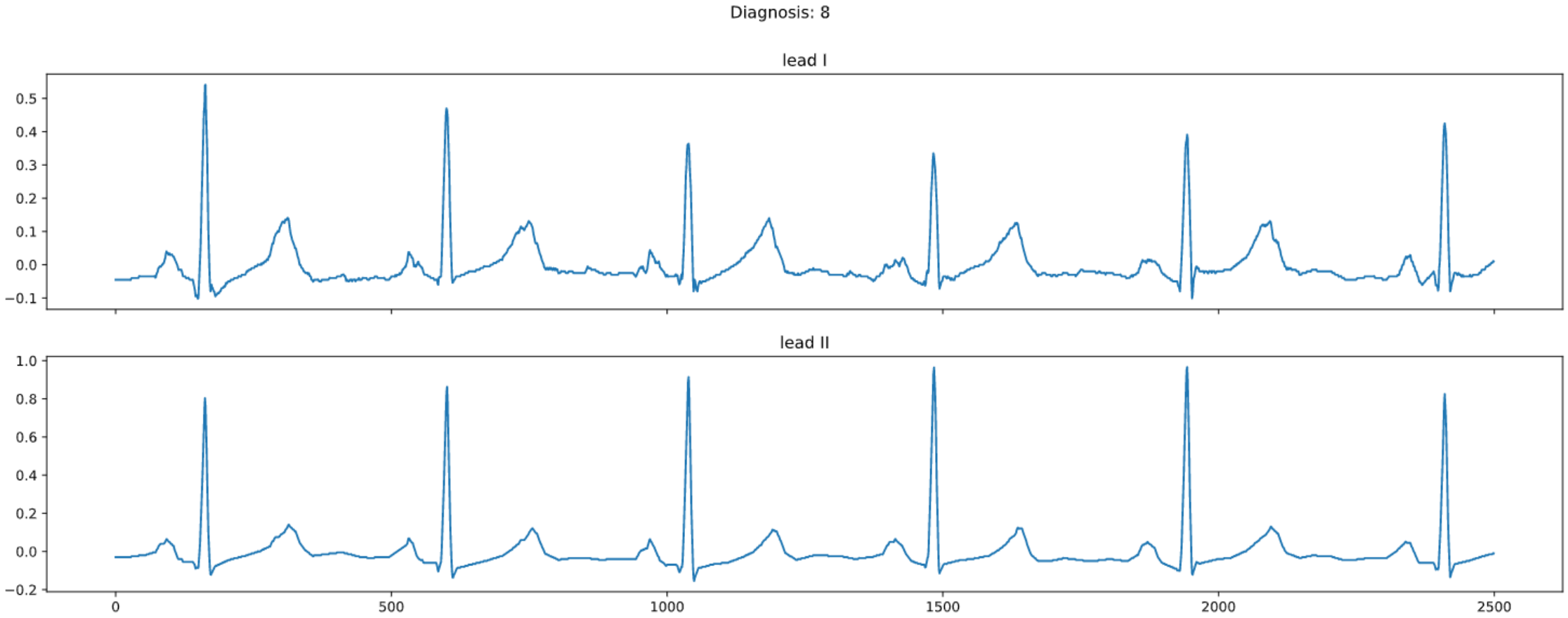
`#Dx: 2 3 8 -> label 정보. test 폴더 안의 {id}.txt 파일에는 label 정보가 없습니다.`

{id}.npy

- shape가 (2, 5000)인 numpy array
- dtype=float32
- 1st row: lead I ECG, 2nd row: lead II ECG
- Sampling rate: 500Hz
- Duration: 10 seconds

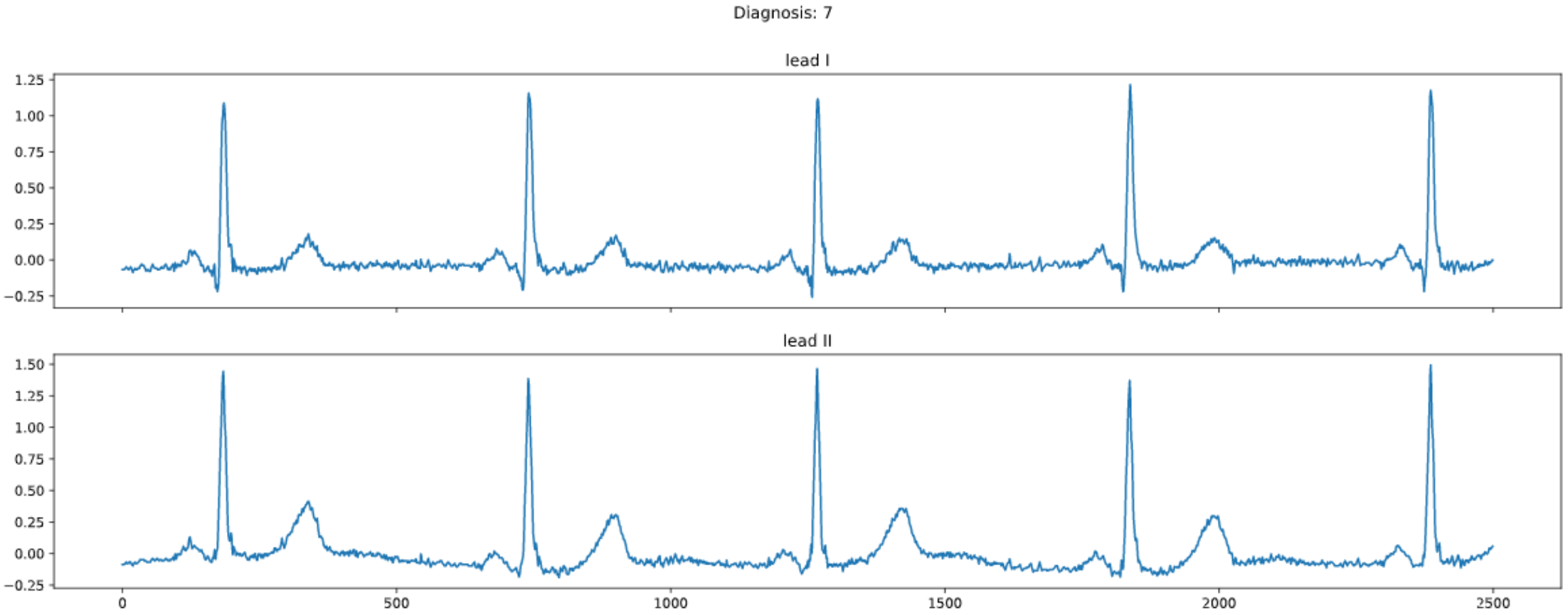
Ex) train/0.npy

- 8. Sinus rhythm



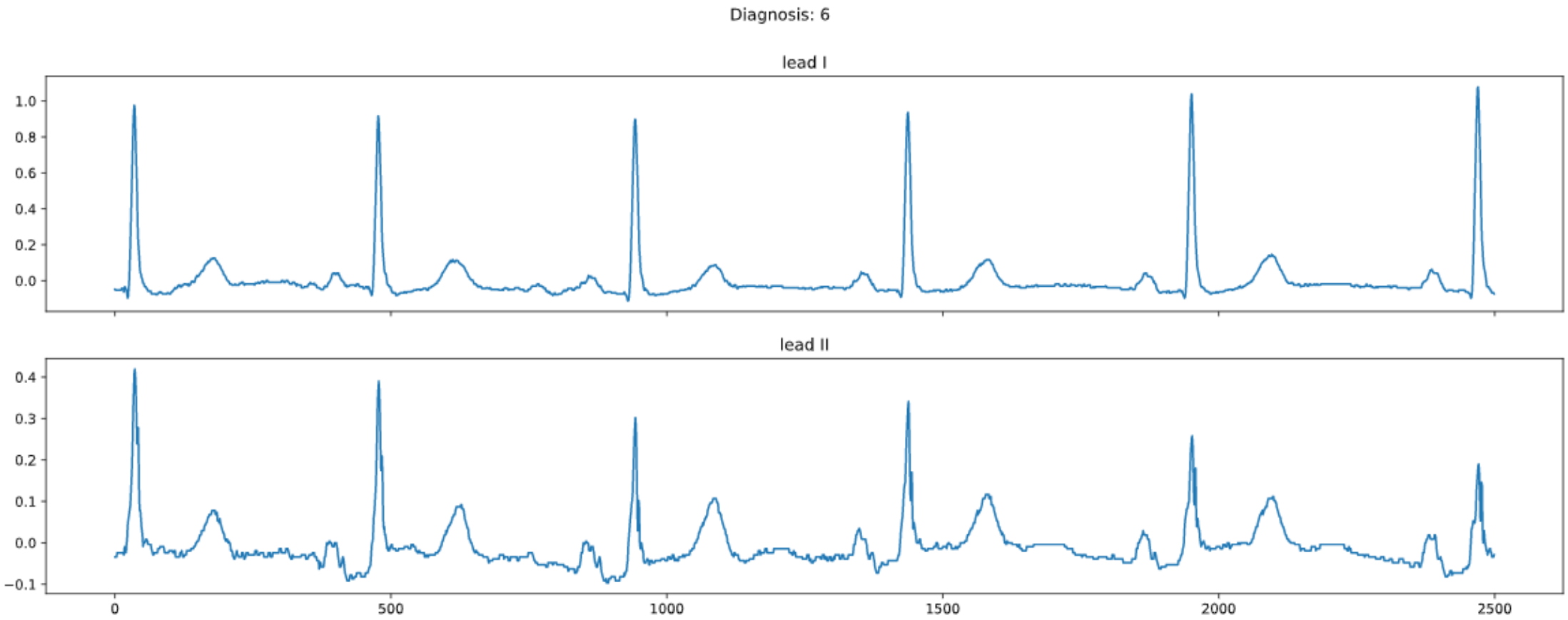
Ex) train/21.npy

- 7. Sinus bradycardia



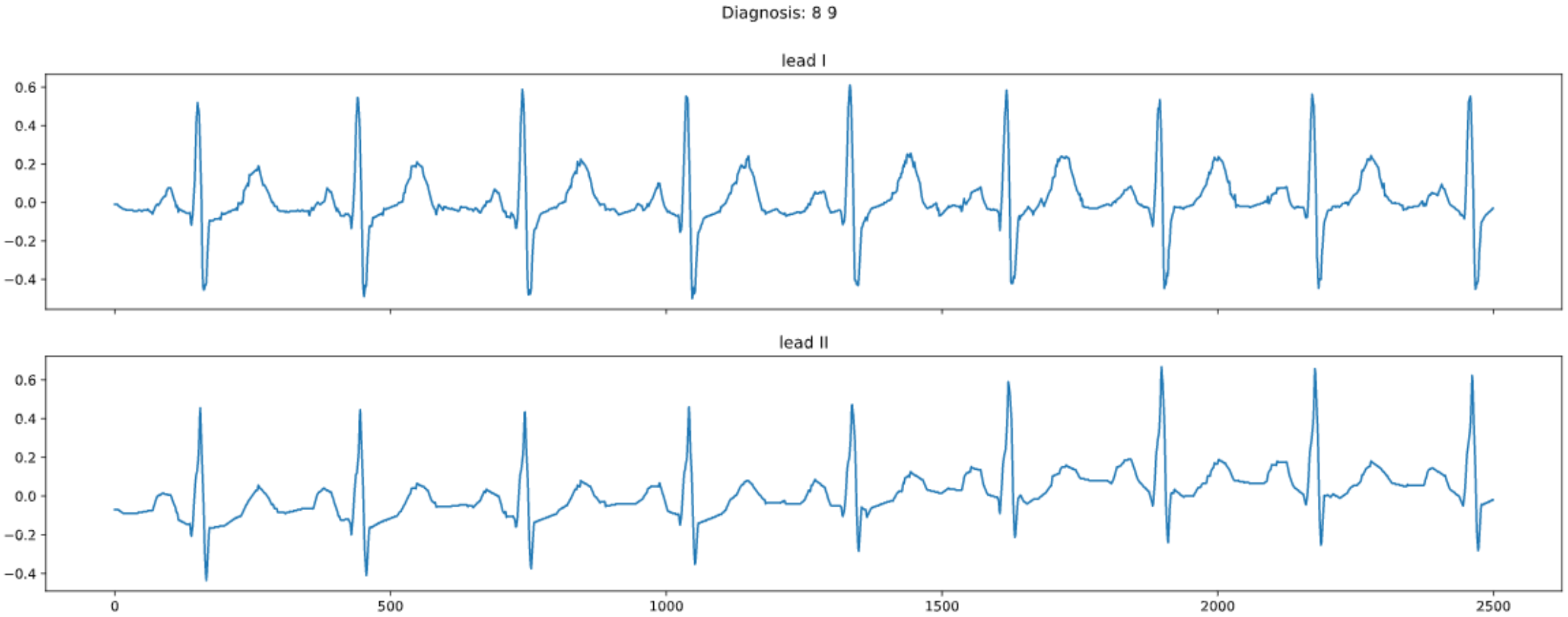
Ex) train/35.npy

- 6. Sinus arrhythmia



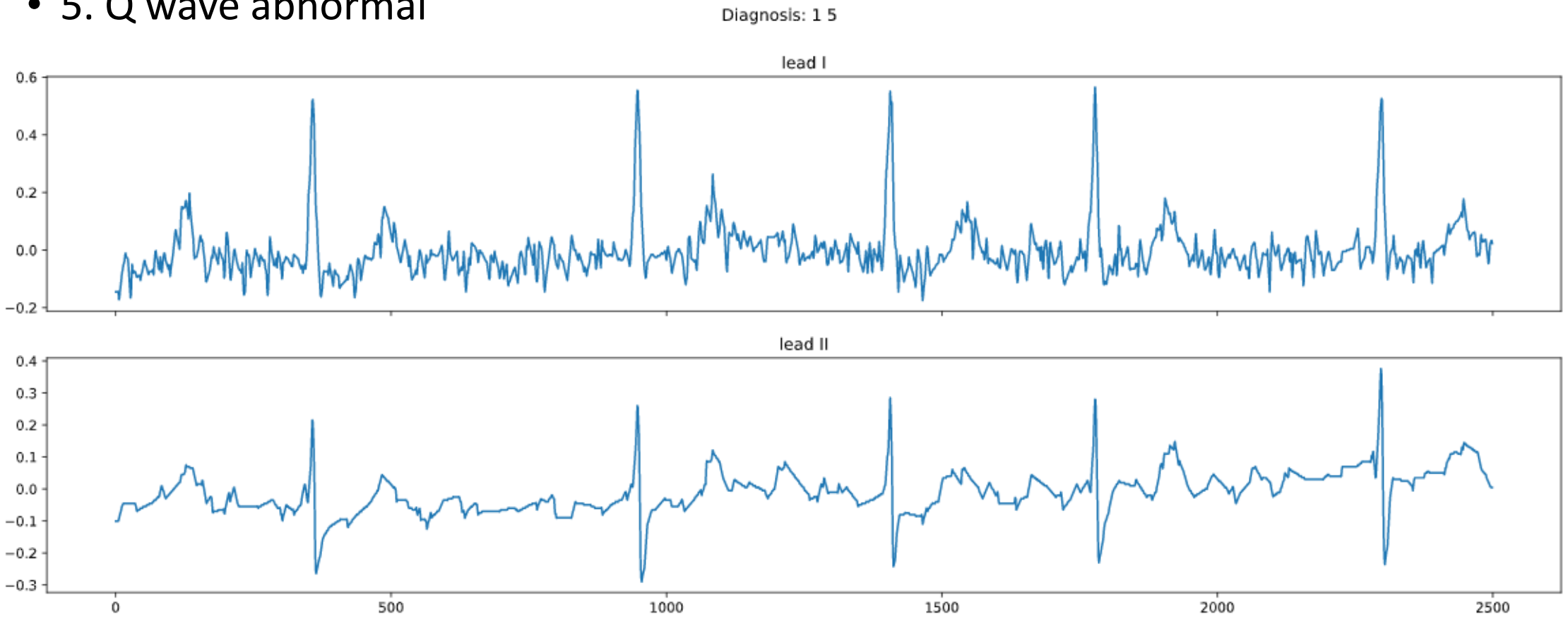
Ex) train/41.npy

- 8. Sinus rhythm
- 9. Sinus tachycardia



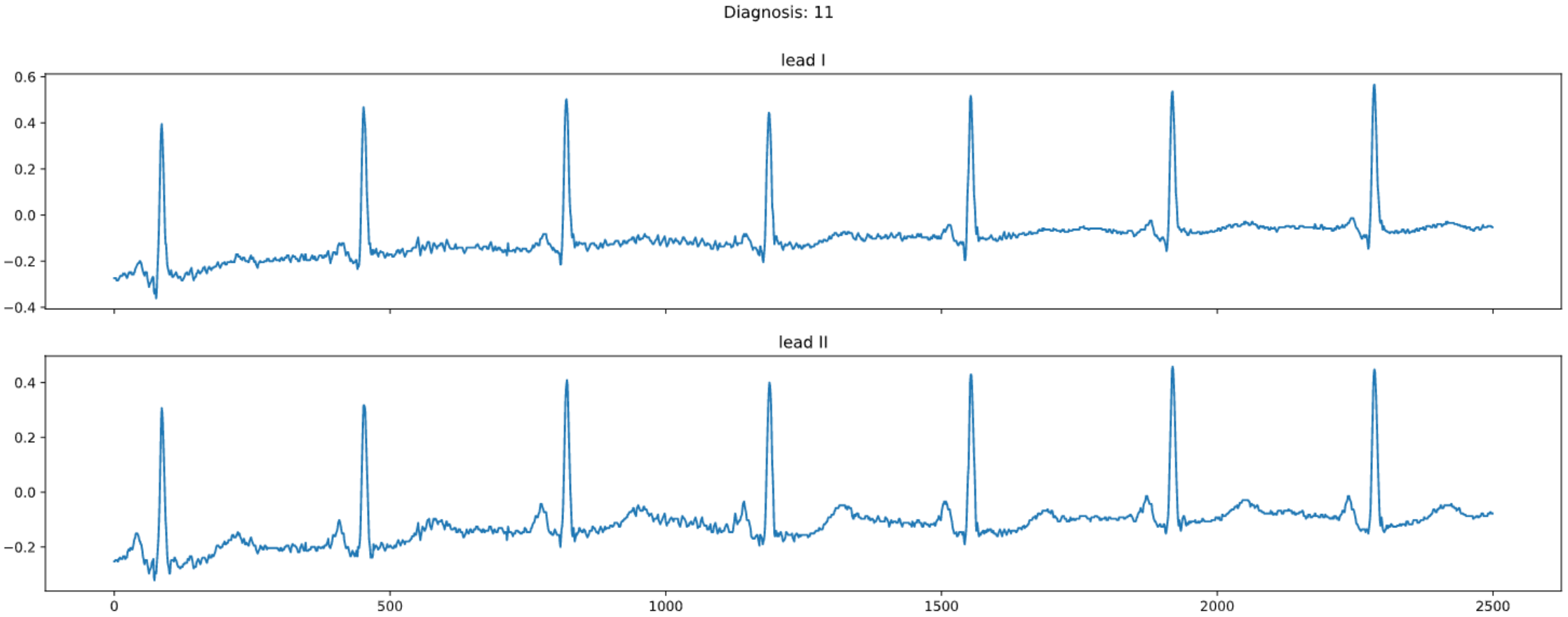
Ex) train/15.npy

- 1. Atrial Fibrillation
- 5. Q wave abnormal



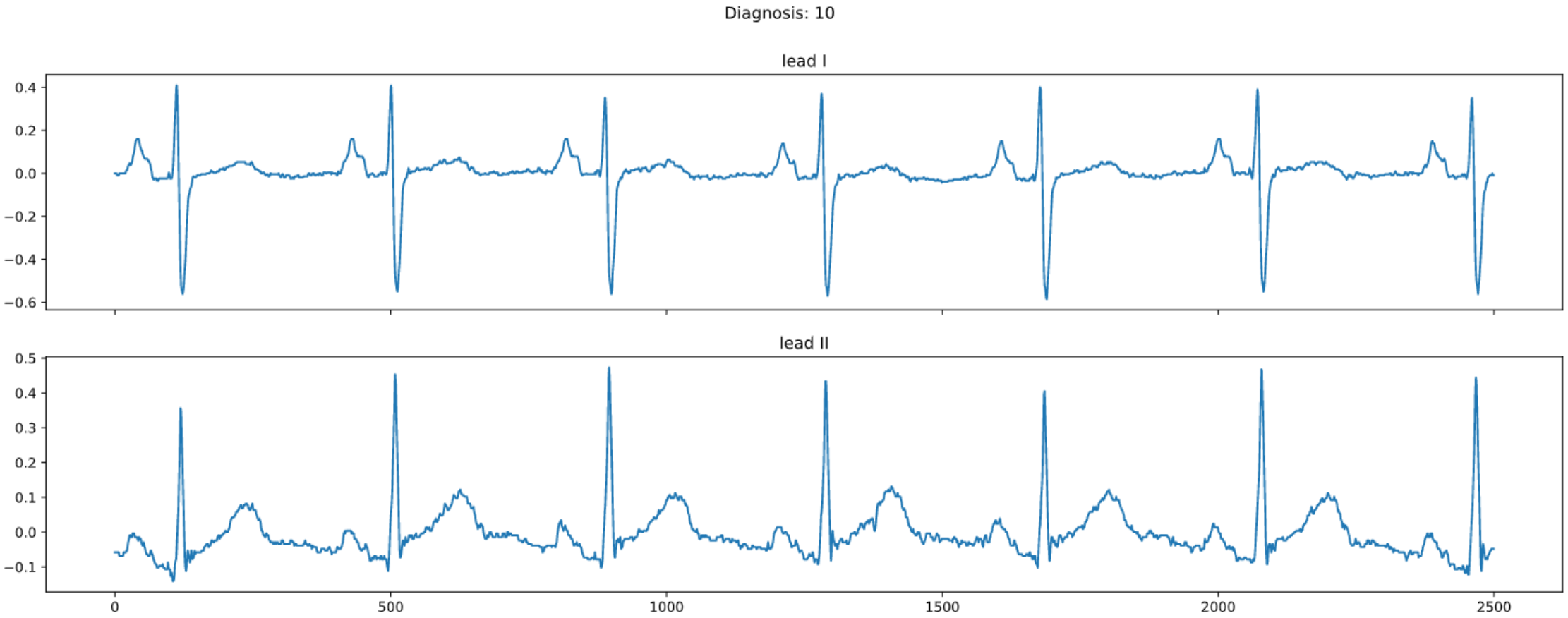
Ex) train/3.npy

- 11. T wave inversion



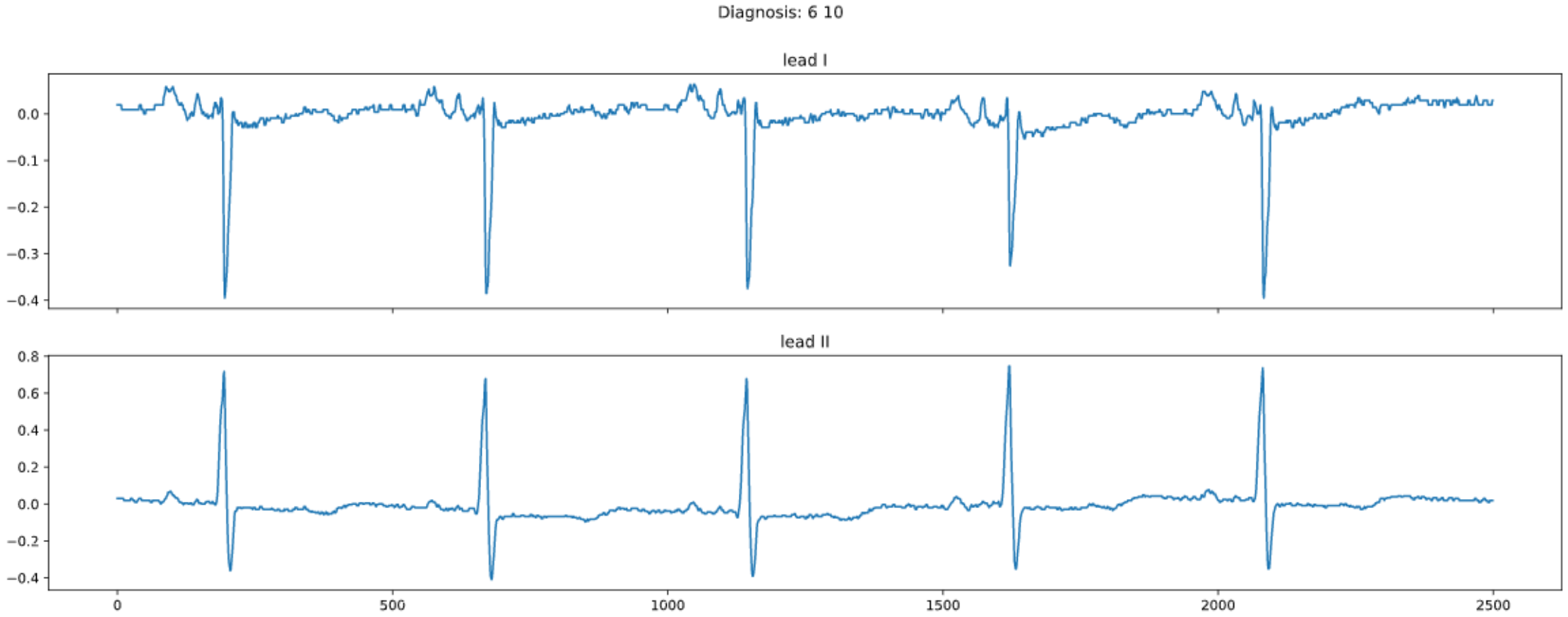
Ex) train/4.npy

- 10. T wave abnormal



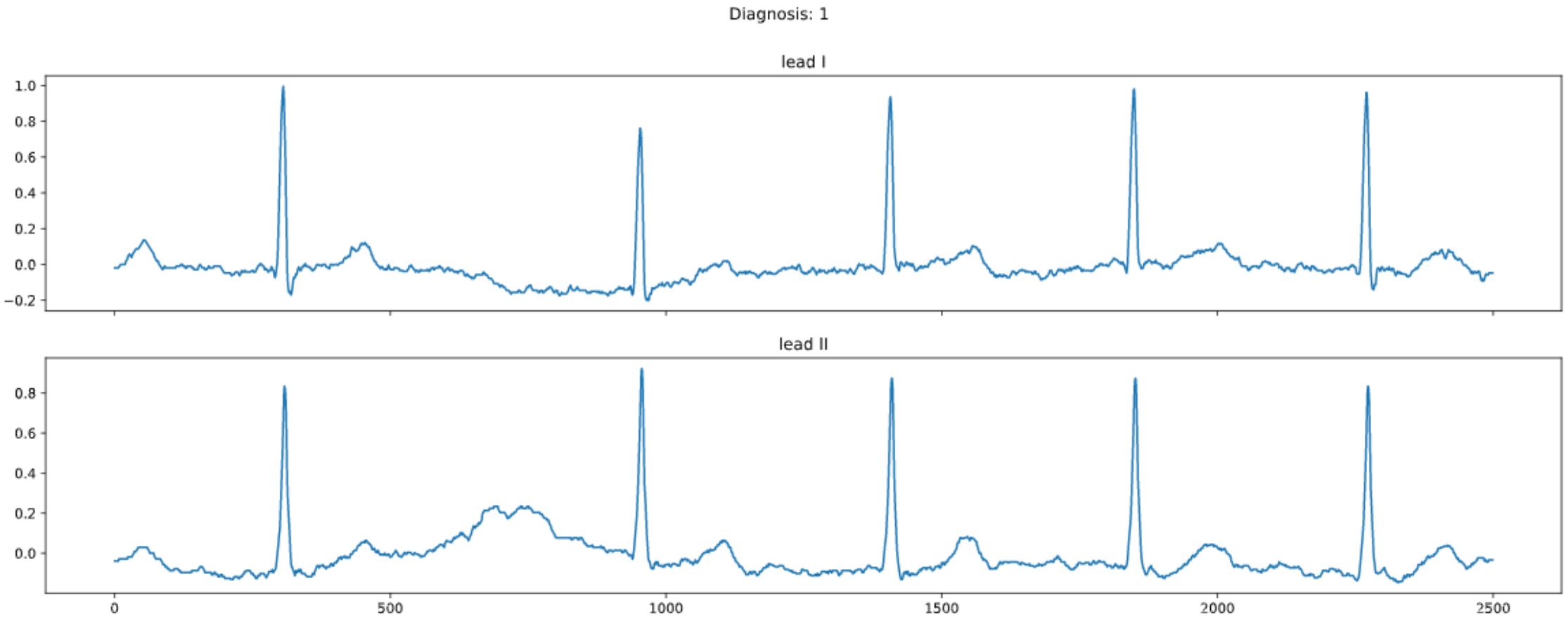
Ex) train/22.npy

- 6. Sinus arrhythmia
- 10. T wave abnormal



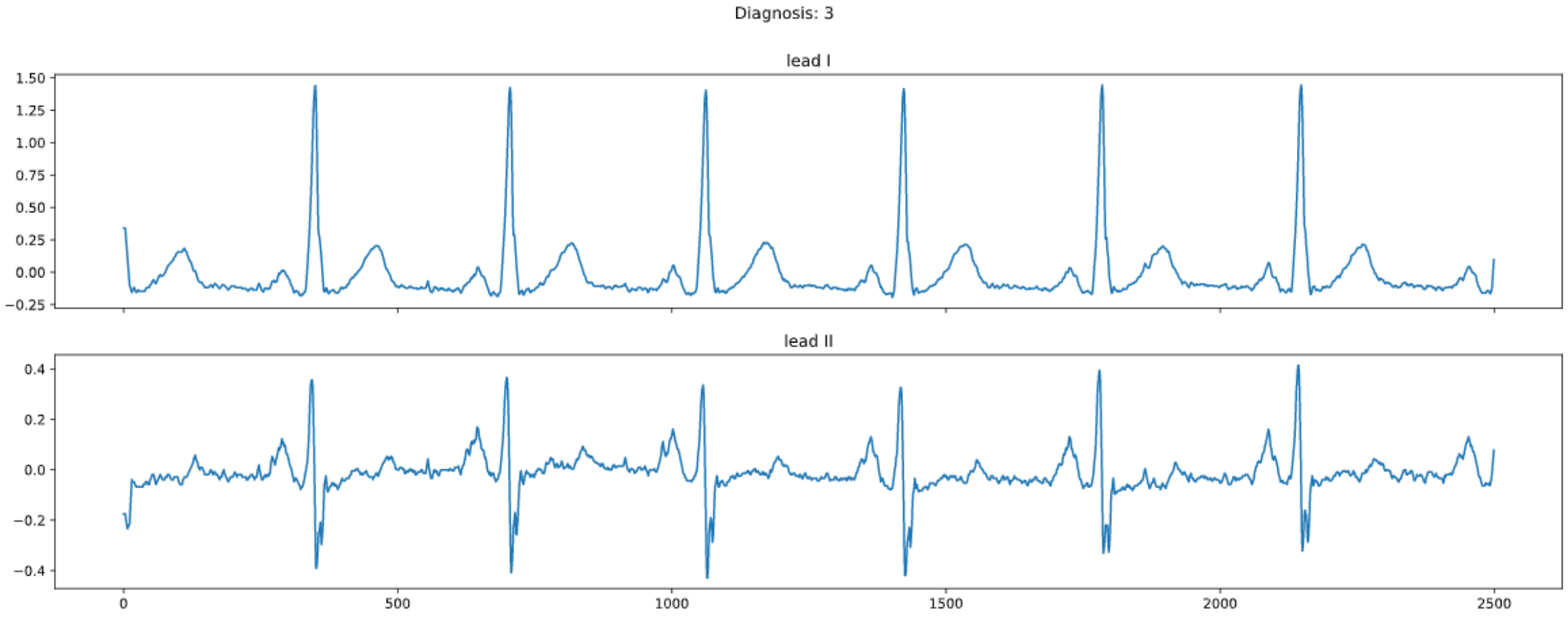
Ex) train/27.npy

- 1. Atrial Fibrillation



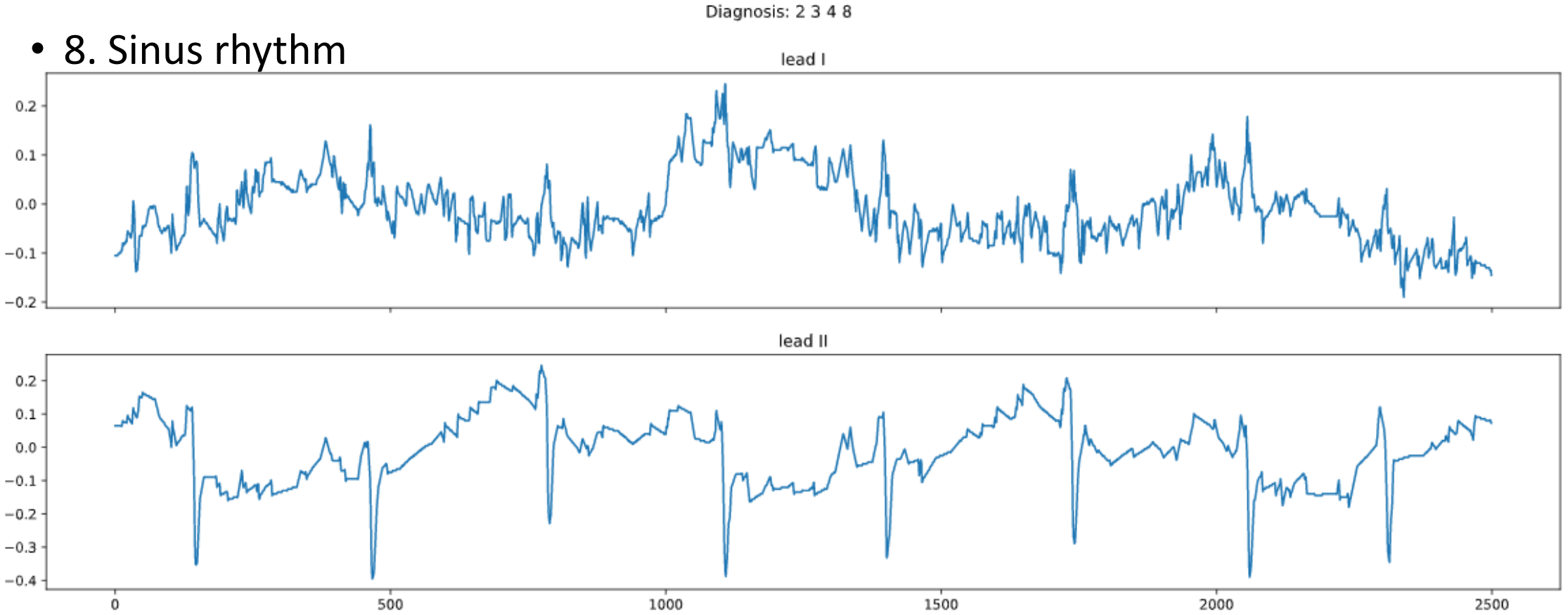
Ex) train/32.npy

- 3. Left axis deviation



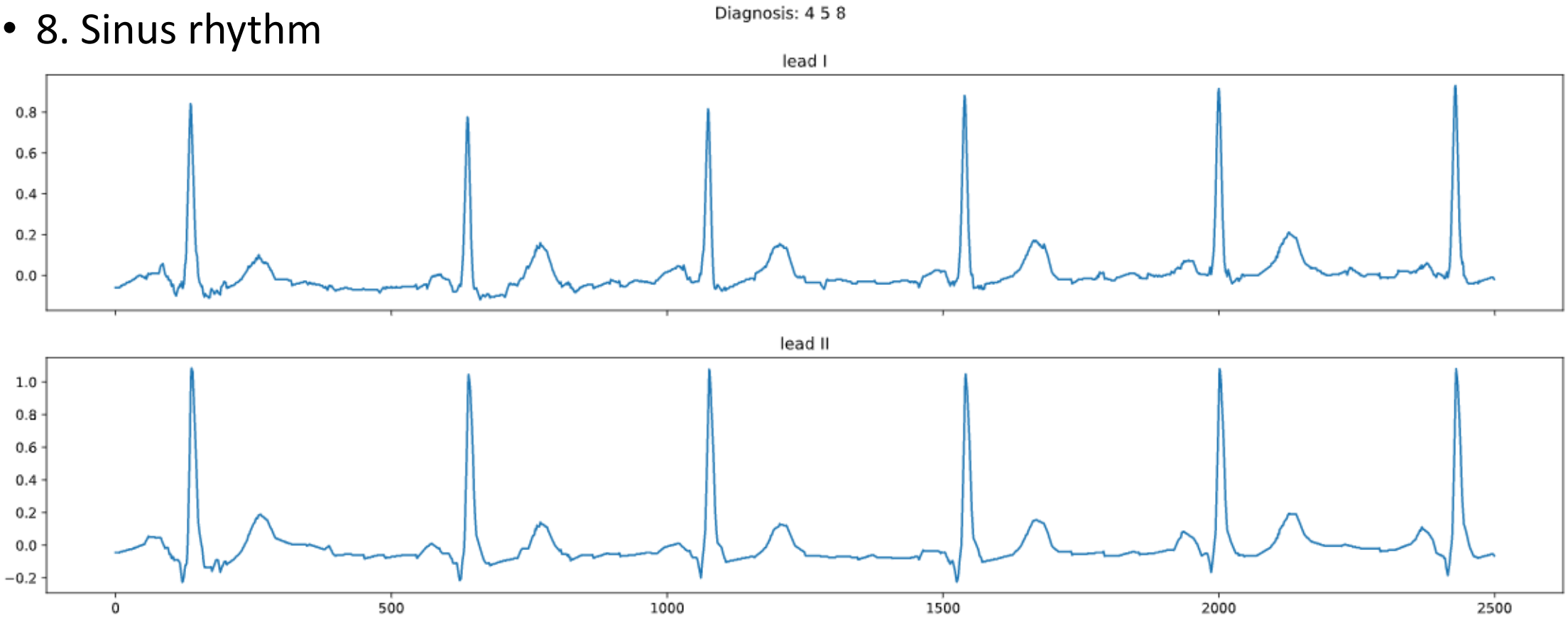
Ex) train/38.npy

- 2. Left anterior fascicular block
- 3. Left axis deviation
- 4. Premature atrial contraction
- 8. Sinus rhythm



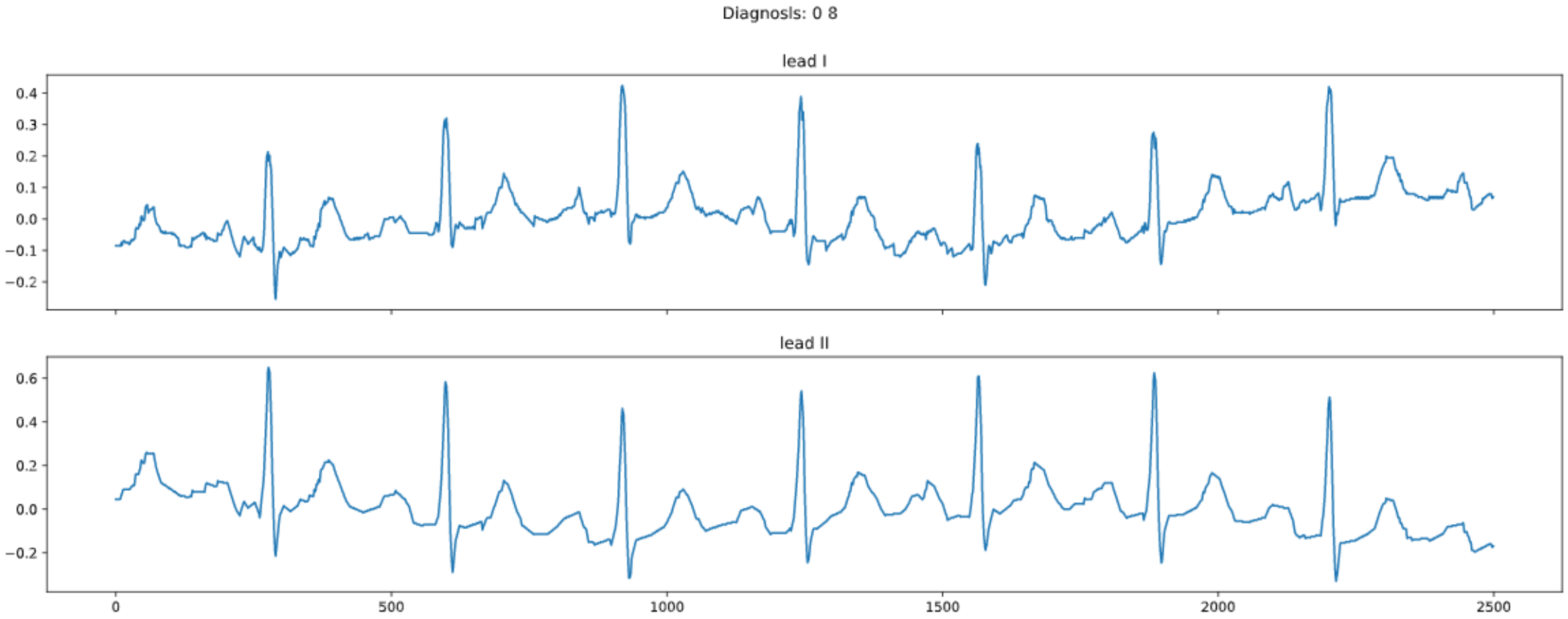
Ex) train/59.npy

- 4. Premature atrial contraction
- 5. Q wave abnormal
- 8. Sinus rhythm



Ex) train/72.npy

- 0. 1st degree av block
- 8. Sinus rhythm

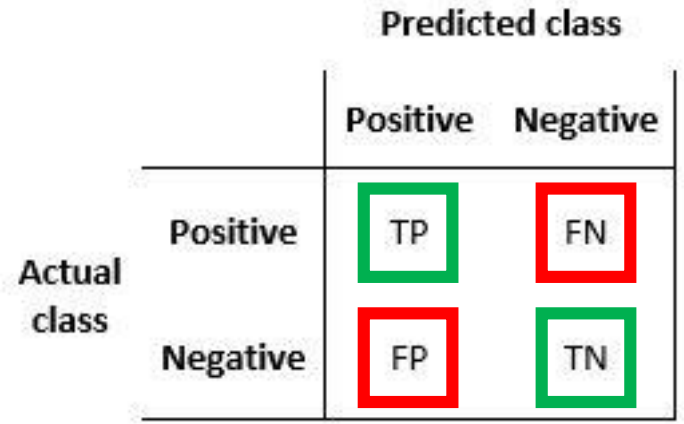


Evaluation metric

- Macro F-score
- 각 class별로 계산된 F-score(F1-score)를 평균한 점수

(Submission example)

Id	Predicted labels
69	7 8
70	2 3 8
71	8
72	8
73	1 3 8
74	2 8
75	8
76	0
77	1



https://commons.wikimedia.org/wiki/File:Binary_confusion_matrix.jpg

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

$$F_1 = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{\text{tp}}{\text{tp} + \frac{1}{2}(\text{fp} + \text{fn})}.$$

Thank you!

ADS Lab.