# TadGAN: Time Series Anomaly Detection using Genrative Adversarial Networks


Alexander Geiger*
*MIT*
Cambridge, USA
geigera@mit.edu

Dongyu Liu*
*MIT*
Cambridge, USA
dongyu@mit.edu

Sarah Alnegheimish
*MIT*
Cambridge, USA
smish@mit.edu

Alfredo Cuesta-Infante
*Universidad Rey Juan Carlos*
Madrid, Spain
alfredo.cuesta@urjc.es

Kalyan Veeramachaneni
*MIT*
Cambridge, USA
kalyanv@mit.edu



arXiv 2020




2021.06.23

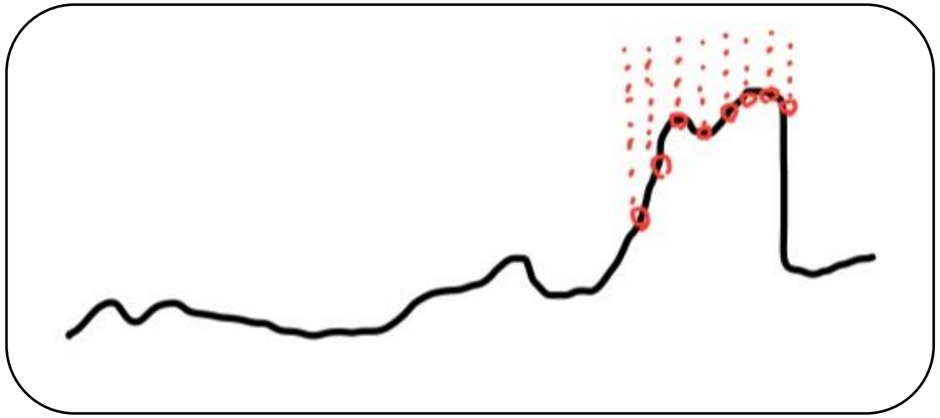
최영제

# Index

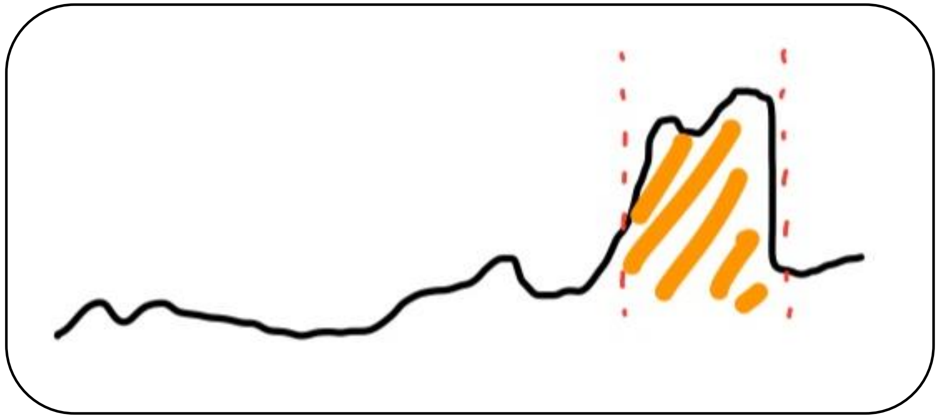
# Introduction

## Two types of anomalies

- Point anomaly



- Collective anomaly



- TadGAN aims to find collective anomalies

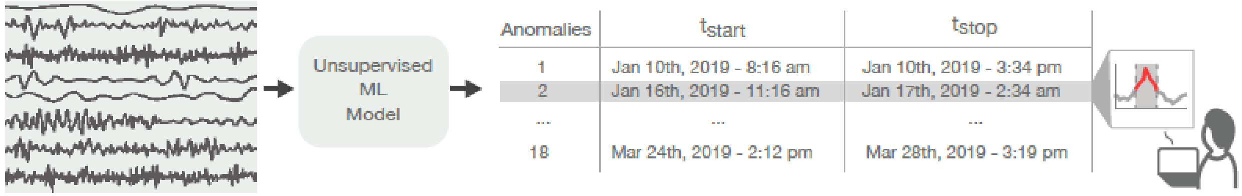



Fig. 1. An illustration of time series anomaly detection using unsupervised learning. Given a multivariate time series, the goal is to find out a set of anomalous time segments that have unusual values and do not follow the expected temporal patterns.

# 1 Introduction

## Background

- 기존 GAN은 generator와 discriminator의 균형을 유지하기 어려우며(학습이 어려움), model collapse 현상이 발생할 수 있음
- 이를 개선한 모델: Wasserstein GAN (WGAN) / TadGAN은 이를 사용(Wasserstein loss)

GAN ⇨ WGAN

**GAN**

- $\mathcal{L} = \mathbb{E}_{x\sim\mathbb{P}_X}[log\mathcal{C}_x(x)] + \mathbb{E}_{z\sim\mathbb{P}_Z}[\log(1 - \mathcal{C}_x(\mathcal{G}(z))]$
- Generator: 판별 결과를 1에 가깝도록 학습
- Discriminator 최대화 = KL Divergence 최소화
- KL Divergence가 적합하지 않은 상황이 존재

**WGAN**

- $\mathcal{L} = \mathbb{E}_{x\sim\mathbb{P}_X}[\mathcal{C}_x(x)] - \mathbb{E}_{z\sim\mathbb{P}_Z}[\mathcal{C}_x(\mathcal{G}(z))]$
- Generator: 데이터 분포와의 Wassertein distance(두 분포간의 차이)를 줄이는 방향으로 학습
- Critic은 Wassertein distance를 구하기 위한 Lipschtz 함수를 근사

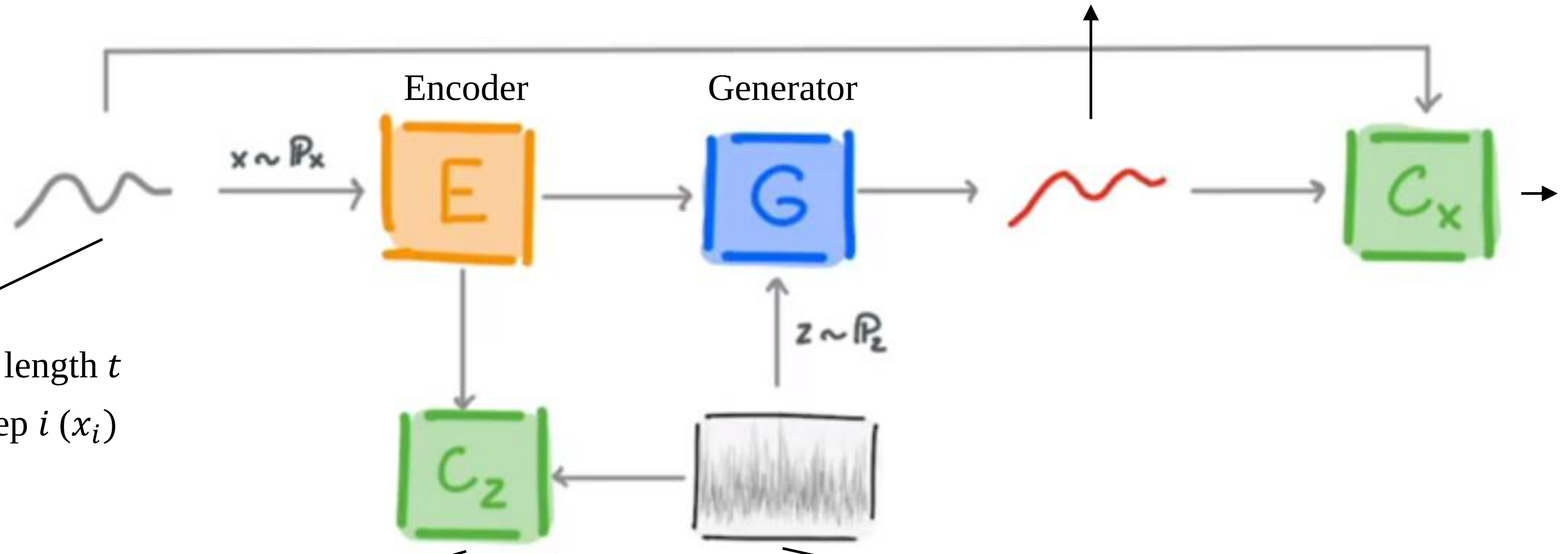
참고:https://kionkim.github.io/2018/06/01/WGAN_1/

## Model architecture of TadGAN

- TadGAN has encoder, generator and two critic networks

Reconstructed sequence

$\mathcal{G}(\mathcal{E}(x_i)) \approx \hat{x}_i$




Critic x

: distinguish $x_i$ from $\mathcal{G}(z)$

Time sequence of length $t$

starting at time step $i$ ($x_i$)

Critic z

: distinguish $z \sim \mathbb{P}_z$ from $\mathcal{E}(x_i)$

Noise (standard multivariate normal dist.)

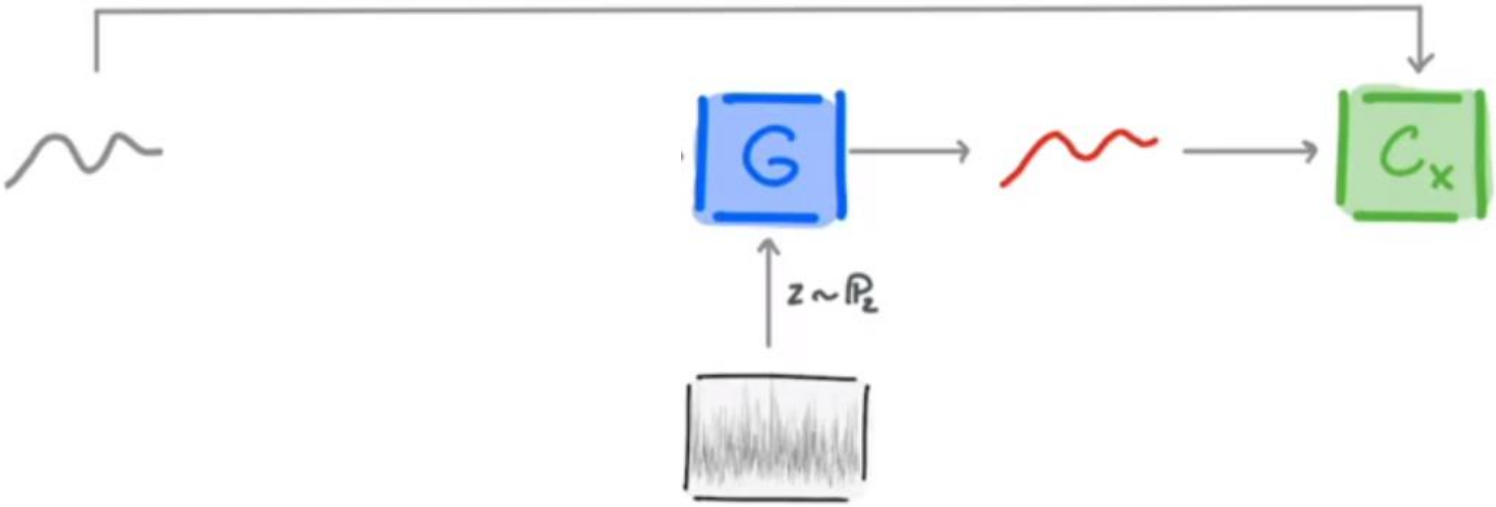
$z \sim \mathbb{P}_z = \mathcal{N}(0,1)$

출처 : https://www.youtube.com/watch?v=jIDj2dhU99k

# Methods

## Adversarial loss using Wasserstein−1 distance ($\mathcal{G}$ & $\mathcal{C}_x$)

- Generator ($\mathcal{G}$): Standard multivariate normal distribution을 따르는 z를 토대로 real과 유사한 sequence를 생성토록 학습($\mathcal{G}(z)\rightarrow X$)
- Critic x ($\mathcal{C}_x$): Real sequence ($X$)와 generated sequence ($\mathcal{G}(z)$)를 구분하도록 학습




$$V_X(\mathcal{C}_x, \mathcal{G}) = \mathbb{E}_{x\sim\mathbb{P}_X}[\mathcal{C}_x(x)] - \mathbb{E}_{z\sim\mathbb{P}_Z}[\mathcal{C}_x(\mathcal{G}(z))]$$

# Methods

## Adversarial loss using Wasserstein−1 distance ($\mathcal{E}$ & $\mathcal{C}_z$)

- Encoder ($\mathcal{E}$): Time sequence ($x$)를 토대로 noise ($z$)와 유사한 latent vector를 만들도록 학습($\mathcal{E}(x)$→z)
- Critic z ($\mathcal{C}_z$): Sampling된 z와 $\mathcal{E}(x)$를 구분하도록 학습


$$V_z(\mathcal{C}_z, \mathcal{E}) = \mathbb{E}_{z\sim\mathbb{P}_Z}[\mathcal{C}_z(z)] - \mathbb{E}_{x\sim\mathbb{P}_X}[\mathcal{C}_z(\mathcal{E}(x))]$$

# Methods

## Cycle consistency loss ($\mathcal{E}$ & $\mathcal{G}$)

- 앞에서 제시한 adversarial losses (i.e. Wasserstein losses) 만으론 sequence의 올바른 reconstruction을 보장할 수 없음
- 따라서 CycleGAN에서 제안된 cycle consistency loss를 추가하여 $x \rightarrow \mathcal{E}(x) \rightarrow \mathcal{G}(\mathcal{E}(x_i)) \approx \hat{x}_i$ 관계를 학습하고자 함
- Cycle consistency loss는 forward, backward loss가 동시 적용되나 TadGAN의 실험 결과 backward loss 적용 시 성능이 나오지 않아 forward consistency loss만 적용(=일반적인 L2-reconstruction error와 동일)


Forward consistency loss: $V_{L2}(\mathcal{E},\mathcal{G}) = \mathbb{E}_{x\sim\mathbb{P}_X}[\|\mathrm{x} - \mathcal{G}(\mathcal{E}(\mathrm{x}))\|_2]$

Backward consistency loss: $V_{L2}(\mathcal{E},\mathcal{G}) = \mathbb{E}_{z\sim\mathbb{P}_Z}[\|z - \mathcal{E}(\mathcal{G}(z))\|_2]$

## Full objective of TadGAN

- TadGAN의 최종 objective function은 다음과 같음

Wasserstein loss ($\mathcal{C}_x$, $\mathcal{C}_z$, $\mathcal{E}$, $\mathcal{G}$)    Cycle consistency loss ($\mathcal{E}$, $\mathcal{G}$)

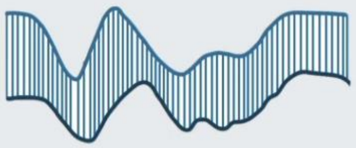
$$\min_{\{\mathcal{E},\mathcal{G}\}} \max_{\{\mathcal{C}_x,\mathcal{C}_z\}} \underline{V_X(\mathcal{C}_x,\mathcal{G}) + V_z(\mathcal{C}_z,\mathcal{E})} + \underline{V_{L2}(\mathcal{E},\mathcal{G})}$$

- Encoder ($\mathcal{E}$): Input sequence의 latent vector를 noise z와 유사하게함과 동시에 generator가 original input sequence를 잘 복원할 수 있도록 mapping
- Generator ($\mathcal{G}$): Noise z가 들어올 때와 latent vector가 들어올 때 모두 original input sequence를 잘 복원하도록 학습
- Critic x ($\mathcal{C}_x$): Real sequence ($X$)와 generated sequence ($\mathcal{G}(z)$)를 구분하도록 학습
- Critic z ($\mathcal{C}_z$): Sampling된 z와 $\mathcal{E}(x)$를 구분하도록 학습

## Anomaly detection using TadGAN

- Estimating anomaly scores using reconstruction errors

**Pointwise Difference**



$$|x^t - \hat{x}^t|$$

**Area Difference**

$$\frac{1}{2*l}\left|\int_{t-l}^{t+l} x^t - \hat{x}^t \, dx\right|$$

**Dynamic Time Warping**

$$\min_{W}\left[\frac{1}{K}\sqrt{\sum_{k=1}^{K} w_k}\right]$$

- Estimating anomaly scores with critic outputs
  - Using the outputs of $C_x$ as an anomaly measure (how real - larger value / how fake - smaller value)

# Methods

## Anomaly detection using TadGAN

- Combining both scores

[ Convex combination ]

$$a(x) = \alpha Z_{RE}(x) + (1-\alpha) Z_{\mathcal{C}_x}(x)$$

[ Multiplication ]

$$a(x) = \alpha Z_{RE}(x) \odot Z_{\mathcal{C}_x}(x)$$

- Finding anomalous sequences with locally adaptive thresholding
  - Sliding window를 사용, 각 window의 평균 대비 4 standard deviations를 넘기면 anomaly로 판단함
  - Window가 K개라면 anomaly sequence $\{a^i_{seq}, i = 1,2,\ldots,K\}$가 생성, 각각의 seq는 $a^i_{seq} = (a_{start(i)}, \ldots a_{end(i)})$의 형태

- Mitigating false positive
  - Anomaly를 판단하는 기준이 각 window 단위기 때문에 false alarm이 빈번히 발생, 이를 줄이고자 다음과 같은 pruning을 거침
  - For each anomalous sequence, use the maximum anomaly score to represent it $\{a^i_{max}, i = 1,2,\ldots,K\}$
  - Maximum anomaly score를 내림차순 정렬 후 decrease percent $(p^i = (a^{i-1}_{max} - a^i_{max})/a^{i-1}_{max})$를 구함
  - 만약 decrease percent가 threshold (default 0.1)를 넘지 못하면 i부터 K까지의 sequence를 normal로 변경

# Experiments and Results

## Experimental setup

- 11 datasets as follow
  - NASA spacecraft telemetry data – mars science laboratory (MSL), soil moisture active passive (SMAP)
  - Yahoo S5 (A1-A4)
  - Numenta anomaly benchmark (NAB) – Art, AdEx, AWS, Traf, Tweets
- Baseline models
  - Autoregressive integrated moving average (ARIMA)
  - Hierarchial temporal memory (HTM)
  - Long short term memory (LSTM)
  - Autoencoder (AE)
  - MAD-GAN
  - Microsoft Azure anomaly detector (MS Azure)
  - Amazon DeepAR

* Prediction-based

* Reconstruction-based

* Commercial tool

# Experiments and Results

## Results

- F1-scores of baseline models

| | NASA | | Yahoo S5 | | | | NAB | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Baseline** | MSL | SMAP | A1 | A2 | A3 | A4 | Art | AdEx | AWS | Traf | Tweets | Mean±SD |
| TadGAN | **0.623** | **0.704** | **0.8** | 0.867 | 0.685 | 0.6 | **0.8** | **0.8** | 0.644 | 0.486 | **0.609** | **0.700±0.123** |
| (P) LSTM | 0.46 | 0.69 | 0.744 | **0.98** | 0.772 | 0.645 | 0.375 | 0.538 | 0.474 | **0.634** | 0.543 | 0.623±0.163 |
| (P) Arima | 0.492 | 0.42 | 0.726 | 0.836 | **0.815** | **0.703** | 0.353 | 0.583 | 0.518 | 0.571 | 0.567 | 0.599±0.148 |
| (C) DeepAR | 0.583 | 0.453 | 0.532 | 0.929 | 0.467 | 0.454 | 0.545 | 0.615 | 0.39 | 0.6 | 0.542 | 0.555±0.130 |
| (R) LSTM AE | 0.507 | 0.672 | 0.608 | 0.871 | 0.248 | 0.163 | 0.545 | 0.571 | **0.764** | 0.552 | 0.542 | 0.549±0.193 |
| (P) HTM | 0.412 | 0.557 | 0.588 | 0.662 | 0.325 | 0.287 | 0.455 | 0.519 | 0.571 | 0.474 | 0.526 | 0.489±0.108 |
| (R) Dense AE | 0.507 | 0.7 | 0.472 | 0.294 | 0.074 | 0.09 | 0.444 | 0.267 | 0.64 | 0.333 | 0.057 | 0.353±0.212 |
| (R) MAD-GAN | 0.111 | 0.128 | 0.37 | 0.439 | 0.589 | 0.464 | 0.324 | 0.297 | 0.273 | 0.412 | 0.444 | 0.35±0.137 |
| (C) MS Azure | 0.218 | 0.118 | 0.352 | 0.612 | 0.257 | 0.204 | 0.125 | 0.066 | 0.173 | 0.166 | 0.118 | 0.219±0.145 |

TABLE IV

F1-SCORES OF BASELINE MODELS USING WINDOW-BASED RULES. COLOR ENCODES THE PERFORMANCE OF THE F1 SCORE. ONE IS EVENLY DIVIDED INTO 10 BINS, WITH EACH BIN ASSOCIATED WITH ONE COLOR. FROM DARK RED TO DARK BLUE, F1 SCORE INCREASES FROM 0 TO 1.

# Experiments and Results

## Results

- vs ARIMA

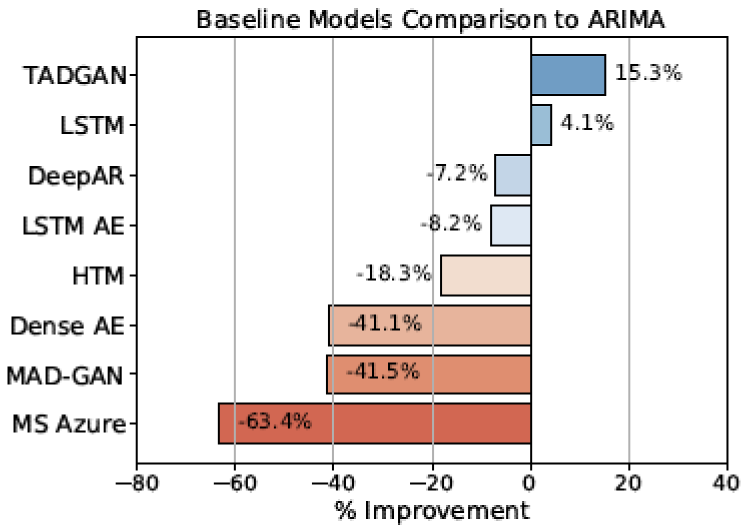



Fig. 3. Comparing average F1-Scores of baseline models across all datasets to ARIMA. The x-axis represents the percentage of improvement over the ARIMA score by each one of the baseline models.

# Experiments and Results

## Results

- Ablation study
  - Point-wise difference / Area difference / DTW
  - Convex combination / Multiplication

| | NASA | | Yahoo S5 | | | | NAB | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Variation** | MSL | SMAP | A1 | A2 | A3 | A4 | Art | AdEx | AWS | Traf | Tweets | Mean+SD |
| Critic | 0.393 | 0.672 | 0.285 | 0.118 | 0.008 | 0.024 | 0.625 | 0 | 0.35 | 0.167 | 0.548 | 0.290±0.237 |
| Point | 0.585 | 0.588 | 0.674 | 0.758 | 0.628 | **0.6** | 0.588 | 0.611 | 0.551 | 0.383 | 0.571 | 0.594±0.086 |
| Area | 0.525 | 0.655 | 0.681 | 0.82 | 0.567 | 0.523 | 0.625 | 0.645 | 0.59 | 0.435 | 0.559 | 0.602±0.096 |
| DTW | 0.514 | 0.581 | 0.697 | 0.794 | 0.613 | 0.547 | 0.714 | 0.69 | 0.633 | 0.455 | 0.559 | 0.618±0.095 |
| Critic×Point | 0.619 | 0.675 | 0.703 | 0.75 | **0.685** | 0.536 | 0.588 | 0.579 | 0.576 | 0.4 | 0.59 | 0.609±0.091 |
| Critic+Point | 0.529 | 0.653 | **0.8** | 0.78 | 0.571 | 0.44 | 0.625 | 0.595 | **0.644** | 0.439 | 0.592 | 0.606±0.111 |
| Critic×Area | 0.578 | **0.704** | 0.719 | **0.867** | 0.587 | 0.46 | **0.8** | 0.6 | 0.6 | 0.4 | 0.571 | 0.625±0.131 |
| Critic+Area | 0.493 | 0.692 | 0.789 | 0.847 | 0.483 | 0.367 | 0.75 | 0.75 | 0.607 | 0.474 | 0.6 | 0.623±0.148 |
| Critic×DTW | **0.623** | 0.68 | 0.667 | 0.82 | 0.631 | 0.497 | 0.667 | 0.667 | 0.61 | 0.455 | 0.605 | **0.629±0.091** |
| Critic+DTW | 0.462 | 0.658 | 0.735 | 0.857 | 0.523 | 0.388 | 0.667 | **0.8** | 0.632 | **0.486** | **0.609** | 0.620±0.139 |
| Mean | 0.532 | 0.655 | 0.675 | 0.741 | 0.529 | 0.438 | 0.664 | 0.593 | 0.579 | 0.409 | 0.580 | |
| SD | 0.068 | 0.039 | 0.137 | 0.211 | 0.182 | 0.154 | 0.067 | 0.209 | 0.081 | 0.087 | 0.02 | |

TABLE V
F1-SCORES OF ALL THE VARIATIONS OF OUR MODEL.

# Q&A