

# 인공지능 기반 혈당 데이터 예측 및 데이터 무결성 보장 연구

이태강\*

한국항공대학교

## Predicting Blood Glucose Data and Ensuring Data Integrity Based on Artificial Intelligence

Tae Kang Lee\*

Korea Aerospace University

E-mail : tkllll@naver.com

### 요 약

최근 5년간 당뇨병으로 진료받은 환자가 322만 명으로 27.7% 증가하였으며 여전히 손가락 채혈을 통해 혈당을 확인하므로 연속적인 혈당 측정과 혈당 피크 확인이 어렵고 고통스러워한다. 이를 해결하기 위해 14일 간 측정된 혈당 데이터를 기반으로 인공지능 기술을 사용하여 3개월간의 혈당 예측 데이터를 당뇨 환자들에게 제공해준다.

### ABSTRACT

Over the past five years, the number of patients treated for diabetes has increased by 27.7% to 3.22 million, and since blood sugar is still checked through finger blood collection, continuous blood glucose measurement and blood sugar peak confirmation are difficult and painful. To solve this problem, based on blood sugar data measured for 14 days, three months of blood sugar prediction data are provided to diabetics using artificial intelligence technology.

### 키워드

SinGAN (Single Generative Adversarial Network), ADA (Adaptive Differentiable Augmentation), RNN (Recurrent Neural network), LSTM(Long Short Term Memory), GRU (Gated Recurrent Unit)

## I. 서 론

최근 5년간 당뇨병으로 진료받은 환자가 322만 명으로 27.7% 증가하였으며 여전히 손가락 채혈을 통해 혈당을 확인하므로 연속적인 혈당 측정과 혈당 피크 확인이 어렵고 고통스러워한다. 이를 해결하기 위해 14일 간 측정된 혈당 데이터를 기반으로 인공지능 기술을 사용하여 3개월간의 혈당 예측 데이터를 당뇨 환자들에게 제공해준다. 본 연구에서는 의료 데이터 부족 문제에 대한 해결책으로 SinGAN 모델을 구성하였다. 먼저 사용자로부터 2주 간 측정된 혈당 데이터를 수집한다. 인공지능을 통해 혈당 예측 시 빅데이터 생성이 필요하므로 수집된 혈당 데이터를 기반으로 SinGAN 모델을 통해 합성 혈당 데이터를 생성하여 데이터 부

족에 대한 문제점을 해결한다. 다음으로, 합성 혈당 데이터를 기반으로 RNN, LSTM, GRU 모델을 통해 3개월간의 미래 혈당 예측 데이터를 생성하여 데이터 예측에 대한 문제를 해결한다. 추가로 사용자가 제공한 원본 혈당 데이터를 블록체인에 저장하여 데이터에 대한 무결성 보장을 연구한다. 최종적으로 사용자는 앱을 통해 미래 혈당 예측 데이터를 기반으로 식단 및 운동을 기록하여 식단 및 운동 관리 서비스를 제공할 수 있게 한다.

## II. 관련 연구

SinGAN은 Single Image를 입력으로 Generator와 Discriminator 네트워크를 통해 원본 데이터를 학습시켜 데이터를 증강시키는 모델이다. 원본 데이터를 학습시켜 나온 합성 데이터들이 증강된 데이터

\* corresponding author

들인데 일본 데이터와 얼마나 유사한지 연구하는 모델이다.

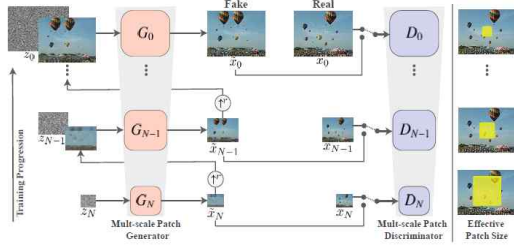


그림 1. SinGAN 모델

ADA는 Training GAN with Limited Data가 목표이다. Generator로 데이터를 증강할 수 있을 뿐만 아니라 Discriminator로도 데이터를 증강할 수 있게끔 해주는 모델이다. GAN을 학습시키기 위해서는 학습 데이터가 최소 100,000개에서 1,000,000개가 필요한데 ADA에서는 100개의 데이터만으로도 일본 데이터와 유사한 데이터를 추출할 수 있도록 연구하는 모델이다.

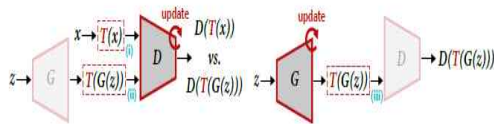


그림 2. ADA 모델

RNN은 입력과 출력을 시퀀스 단위로 처리하는 모델이다. 여기서 시퀀스는 연관된 연속의 데이터를 의미하며, 시계열 데이터에 적합한 신경망 모델이라 자연어 처리, 순서를 가지는 정보, 연속적인 시간 간격으로 배치된 데이터 등을 연구하는 모델이다.

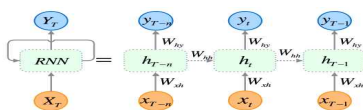


그림 3. RNN 모델

LSTM은 RNN 모델의 문제점을 개선하기 위해 나온 모델이다. 일반적인 RNN의 경우 짧은 시퀀스를 처리할 경우 유리하며, 거리가 멀어지는 경우 학습 능력이 현저하게 저하된다. 데이터의 시퀀스가 길어지면 장기 의존성 학습에 어려움이 있기 때문인데 가중치를 업데이트 하는 과정에서 1보다 작은 값들이 계속 곱해지면 기울기가 사라지는 기울기 소실 문제가 발생하기 때문에 장기 의존성을 해결하기 위해서 셀 상태라는 구조를 만들었다. 데이터를 계산하는 위치에 입력, 망각, 출력 게이트 3가지가 추가되어 각 상태 값을 메모리 공간 셀에

저장하고, 데이터를 접하는 게이트 부분을 조정하여 불필요한 연산, 오차 등을 줄여 장기 의존성 문제를 해결할 수 있도록 연구하는 모델이다.

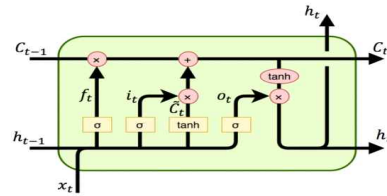


그림 4. LSTM 모델

GRU는 RNN 모델의 문제점을 개선하기 위해 나온 모델이다. LSTM의 변형 모델로 구조가 크게 다르지 않지만 update gate와 reset gate만 있고 output gate가 없는 구조이기 때문에 LSTM 보다 매개변수가 적으며 소규모 데이터 셋일 때는 더 나은 성능을 보여주고 보통 대위법 음악 모델링 및 음성 신호를 연구하는 모델이다.

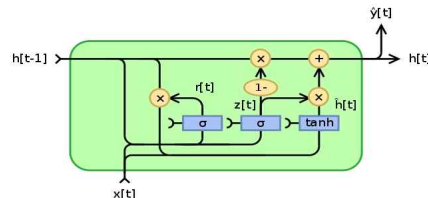


그림 5. GRU 모델

### III. 모델 설계

일본 데이터를 기반으로 SinGAN과 ADA 모델에 대입하여 학습시킨다. 학습시켜서 나온 증강된 합성 데이터를 기반으로 RNN과 LSTM 모델에 대입하여 학습시킨다. 학습시켜서 나온 3개월간의 미래 예측 데이터를 생성하여 데이터를 예측한다.

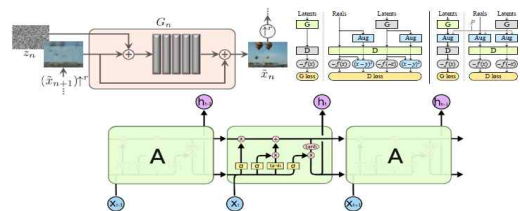


그림 6. SinGAN ADA LSTM 모델 합성

### IV. 제안 방법

사용자로부터 CGM(Continuous Glucose Monitoring) Sensor를 통해 14일 간 측정된 혈당 데이터 기반으로 SinGAN과 ADA 모델을 통해 데이터를 증강시켜 혈당 합성 데이터를 생성하고 생성된 합성

혈당 데이터를 기반으로 LSTM 모델을 통해 3개월 간 미래 예측 혈당 데이터 생성한다. 사용자가 제공한 원본 혈당 데이터는 블록체인에 저장하여 데이터에 대한 무결성 보장 및 데이터 제공에 대한 보상을 토큰으로 제공한다. 추후 사용자는 앱을 통해 미래 혈당 예측 데이터를 기반으로 식단 및 운동을 기록하여 식단 및 운동 관리 서비스를 제공할 수 있게 한다.



그림 7. 모델 전체 흐름도

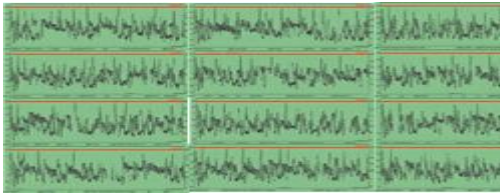


그림 8 데이터 증강된 혈당 합성 데이터

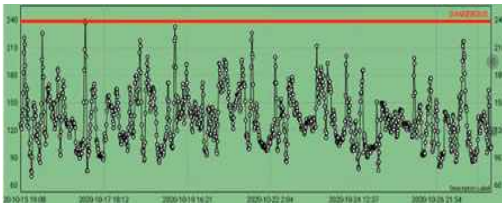


그림 9. 14일 간 측정된 혈당 데이터

```

contract data {
    struct Data {
        uint256 carbs;
        uint256 protein;
        uint256 fat;
    }
    Data[] data;
    function setData(uint256 carbs, uint256 protein, uint256 fat) external {
        data.push(Data(carbs, protein, fat));
    }
}

contract data {
    struct Data {
        uint256 carbs;
        uint256 protein;
        uint256 fat;
    }
    Data[] data;
    function setData(uint256 carbs, uint256 protein, uint256 fat) external {
        data.push(Data(carbs, protein, fat));
    }
}

contract data {
    struct Data {
        uint256 carbs;
        uint256 protein;
        uint256 fat;
    }
    Data[] data;
    function setData(uint256 carbs, uint256 protein, uint256 fat) external {
        data.push(Data(carbs, protein, fat));
    }
}
    
```

그림 10. 데이터 블록체인 저장 및 토큰 보상

### V. 결 론

본 논문에서는 혈당 데이터 증강을 위해 SinGAN과 ADA 모델을 사용했으며 혈당 예측을 위해 LSTM 모델을 사용하였다. 또한 원본 데이터에 대한 무결성 보장과 토큰 보상을 위해 Smart Contract를

사용하여 데이터를 블록체인 네트워크에 저장하여 무결성을 보장하고 데이터 제공에 대한 토큰을 보상하도록 구현하였다. 그러나 데이터 예측에 대한 정확도가 부족하며 예측 데이터 기반으로 식단 및 운동 관리를 통해 당뇨 예방에 대한 연구가 필요하다.

### References

- [1] Shaham, Tamar Rott, Tali Dekel, and Tomer Michaeli, "Singan: Learning a generative model from a single natural image," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019.
- [2] Karras, Tero, et al, "Training generative adversarial networks with limited data," *Advances in Neural Information Processing Systems* 33, 2020.
- [3] Zhao, Shengyu, et al, "Differentiable augmentation for data-efficient gan training," *Advances in Neural Information Processing Systems* 33, 2020.
- [4] Wang, Qin, et al, "Non-fungible token (NFT): Overview, evaluation, opportunities and challenges," *arXiv preprint arXiv:2105.07447*, 2021.
- [5] tamarott (2019) SinGAN[Source code]. Available : <https://github.com/tamarott/SinGAN>.
- [6] NVlabs (2020) stylegan2-ada[Source code]. Available : <https://github.com/NVlabs/stylegan2-ada>.
- [7] mit-han-lab (2020) data-efficient-gans[Source code]. Available : <https://github.com/mit-han-lab/data-efficient-gans>.