

시계열 생성적 적대 신경망을 이용한 비행체 궤적 합성 데이터 생성 및 비행체 궤적 예측에서의 활용에 관한 연구

A Study on Synthetic Flight Vehicle Trajectory Data Generation Using Time-series Generative Adversarial Network and Its Application to Trajectory Prediction of Flight Vehicles

박인희*, 이창진*, 정찬호**★

In Hee Park*, Chang Jin Lee*, Chanho Jung**★

Abstract

In order to perform tasks such as design, control, optimization, and prediction of flight vehicle trajectories based on machine learning techniques including deep learning, a certain amount of flight vehicle trajectory data is required. However, there are cases in which it is difficult to secure more than a certain amount of flight vehicle trajectory data for various reasons. In such cases, synthetic data generation could be one way to make machine learning possible. In this paper, to explore this possibility, we generated and evaluated synthetic flight vehicle trajectory data using time-series generative adversarial neural network. In addition, various ablation studies (comparative experiments) were performed to explore the possibility of using synthetic data in the aircraft trajectory prediction task. The experimental results presented in this paper are expected to be of practical help to researchers who want to conduct research on the possibility of using synthetic data in the generation of synthetic flight vehicle trajectory data and the work related to flight vehicle trajectories.

요약

딥러닝을 포함한 머신러닝 기법을 기반으로 비행체의 궤적 설계, 제어, 최적화, 예측 등의 작업을 수행하기 위해서는 일정한 양 이상의 비행체 궤적 데이터를 필요로 한다. 그러나 다양한 이유(예를 들어 비행체 궤적 데이터셋 구축에 필요한 비용, 시간, 인력 등)로 일정한 양 이상의 비행체 궤적 데이터를 확보하기 어려운 경우가 존재한다. 이러한 경우 합성 데이터 생성이 머신러닝을 가능하게 하는 방법 중 하나가 될 수 있다. 본 논문에서는 이와 같은 가능성을 탐구하기 위하여 시계열 생성적 적대 신경망을 이용하여 비행체 궤적 합성 데이터를 생성하고 평가하였다. 또한 비행체의 상태를 인식하기 위한 비행체 궤적 예측 작업에서 합성 데이터의 활용 가능성을 탐구하기 위하여 다양한 ablation study(비교 실험)를 수행하였다. 본 논문에서 제시된 생성 평가 및 비교 실험 결과는 비행체 궤적 합성 데이터 생성 및 비행체 궤적 관련 작업에서 합성 데이터의 활용 가능성에 대한 연구를 수행하고자 하는 연구자들에게 실질적인 도움이 될 것으로 예상된다.

Key words : Generative adversarial network, Synthetic data generation, Trajectory data, Flight vehicles, Trajectory prediction

* The 5th R&D Institute - 1st Directorate, Agency for Defense Development, Korea. Senior Engineer.

** Dept. of Electrical Engineering, Hanbat National University, Korea. Associate Professor.

★ Corresponding author

E-mail : peterjung@hanbat.ac.kr, Tel : +82-42-821-1095

Manuscript received Nov. 13, 2021; revised Nov. 30, 2021; accepted Dec. 1, 2021.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

I. 서론

생성적 적대 신경망(generative adversarial network) [1]은 준지도(semi-supervised) 및 비지도(unsupervised) 학습을 위한 신경망 중 하나로써, 합성 데이터(synthetic data) 생성을 통한 지도 학습(판별) 방법을 이용한다. 생성적 적대 신경망은 한 개의 신경망으로 구성된 기존의 신경망과는 다르게 두 개의 신경망으로 구성된다[2]: 1) 생성(generative) 및 2) 판별(discriminative) 신경망. 생성 신경망은 랜덤 노이즈로 구성된 벡터를 입력받아 실제 데이터(real data)와 비슷한 특성을 가지는 합성 데이터를 생성한다. 판별 신경망은 실제 데이터와 합성 데이터를 분류하는 것을 목표로 한다. 생성적 적대 신경망은 객체 검출, 텍스트 합성, 초해상화 등과 같은 이미지 및 비디오 데이터를 다루는 분야에서 활발하게 이용되어 왔다. 또한 자연어 처리, 음성 및 오디오 합성, 음악 생성 등과 같은 시계열(time-series) 데이터를 다루는 분야에서도 그 유용성이 검증되었다. 학습 데이터의 양이 충분하지 않아 학습이 어려운 경우 실제 데이터와 유사한 합성 데이터를 생성하여 학습 데이터를 보완함으로써 학습 성능을 극대화하는 것도 생성적 적대 신경망의 활용 방법 중 하나이다.

본 논문에서는 “비행체 궤적 데이터”의 양이 부족한 경우 “생성적 적대 신경망의 학습 및 활용”에 대해 탐구하였다. 학습 성능 평가를 위해 시계열 생성적 적대 신경망을 이용하여 생성한 비행체 궤적 합성 데이터의 품질을 측정하였으며, 활용 가능성 탐구를 위해 비행체 궤적 예측 작업을 수행하였다. 특히 효율적인 활용 가능성 탐구를 위해 “합성 데이터 기반 학습 및 실제 데이터 기반 테스트”(train on synthetic and test on real, TSTR) 프레임워크를 이용하였다.

II. 생성적 적대 신경망을 이용한 비행체 궤적 합성 데이터 생성 및 품질 평가

시계열 생성적 적대 신경망의 궁극적인 목적은 실제 시계열 데이터와의 구분이 매우 어려운 합성 시계열 데이터를 생성하는 것이다. 그림 1은 생성적 적대 신경망을 이용한 비행체 궤적 합성 데이터 생성의 도식을 보여준다. 효율적인 합성 시계열 데이터 생성을 위해 다양한 프레임워크가 제안되어

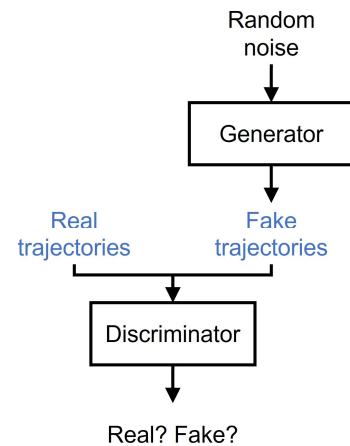


Fig. 1. Synthetic trajectory data generation using GAN. 그림 1. GAN을 이용한 비행체 궤적 합성 데이터 생성

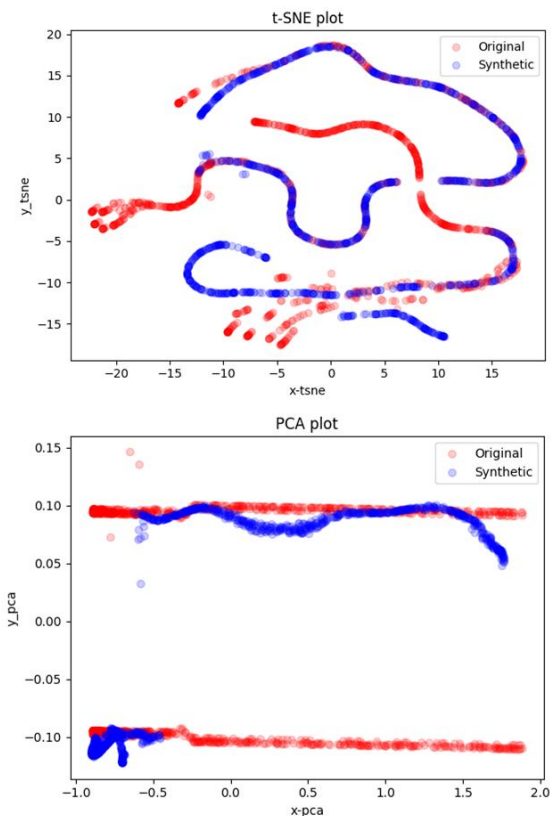


Fig. 2. Quality assessment of synthetic trajectory data based on visualization using t-SNE and PCA.

그림 2. t-SNE 및 PCA를 이용한 시각화 기반 합성 데이터 품질 측정(빨강: 실제, 파랑: 합성)

왔다[3-5]. 본 논문에서는 비지도 GAN 프레임워크의 적응성과 자동회귀 모델을 이용한 지도 학습을 통해 제공되는 제어 능력을 결합한 시계열 생성적 적대 신경망을 이용한다[5]. 또한 합성 데이터 생성을 위해 우리의 최근 이전 연구[6]에서 이용된 총 18개의 텔레메트리 기반 비행체 이동 궤적 데이터를 이용하였다. 합성 시계열 데이터의 품질 측정을 위해

[5]에서와 같이 시각화 및 판별 스코어(discriminative score)를 이용하였다. 그림 2는 t-SNE 및 PCA를 이용한 시각화 기반 품질 측정 결과를 보여준다. 그림 2에서 보는 바와 같이 비행체 궤적 실제 데이터와 합성데이터가 2차원 공간상에서 상당히 겹치는 것을 볼 수 있다. 겹치는 정도가 클수록 합성 데이터의 품질이 높다고 판단한다[5]. 표 1은 [5]에서 이용된 다양한 벤치마크 데이터셋 및 비행체 궤적 데이터셋의 판별 스코어를 보여준다. 판별 스코어는 다음과 같이 정의한다[5].

Table 1. Quality assessment of synthetic trajectory data based on discriminative score.

표 1. 판별 스코어 기반 합성 데이터 품질 측정

Time-series dataset	Score
Sines	0.011
Stocks	0.102
Energy	0.236
Events	0.161
Trajectories (ours)	0.103

$$D = |0.5 - A| \quad (1)$$

식 (1)에서 A 는 판별 신경망의 분류 정확도를 나타낸다. 판별 스코어가 작을수록 합성 데이터셋의 품질이 높다고 판단한다[5]. 실제 데이터셋과 합성 데이터셋의 분류가 어려울수록 판별 스코어는 작은 값을 가지기 때문이다. 비행체 궤적 데이터셋의 경우 Sines보다 크며 Energy보다 작은 판별 스코어를 가지며, Stocks와 비슷하게 작은 판별 스코어를 가진다. 즉, 주가 실제 데이터셋을 이용하여 생성한 주가 합성 데이터셋의 품질과 비행체 궤적 실제 데이터셋을 이용하여 생성한 비행체 궤적 합성 데이터셋의 품질이 거의 비슷하다. 그림 2 및 표 1에서 종합적으로 보는 바와 같이 비행체 궤적 합성 데이터셋의 품질이 상당히 높음을 볼 수 있다.

III. 실험 결과: 비행체 궤적 예측에서의 활용 결과

본 논문에서는 우리의 최근 이전 연구[6]에서 이용했던 2차원 장단기 메모리를 이용하여 비행체의 궤적 예측 작업을 수행하였다. 또한 [6]에서와 같이 손실 함수로써 MSE(mean squared error)를 이용하였다. 총 100 epoch에 걸쳐 학습을 진행하였으며 초기 학습률은 0.1로 설정하였다. 안정적인 학습을

위해 학습률을 50 epoch에서 0.01로 감소시켰다.

Table 2. Prediction error w.r.t. input and output.

표 2. 입출력 관계 변화에 따른 비행체 궤적 예측 오차 (O : 비행체 궤적 실제 데이터 기반 학습, S : 비행체 궤적 합성 데이터 기반 학습)

Method	Avg. error	
Baseline(absolute value-based)	O	0.1356
	S	0.1402
Difference between absolute values-based[7]	O	0.0437
	S	0.0429
Acceleration value-based[8]	O	0.1396
	S	0.1394
Relative value-based[6]	O	0.0053
	S	0.0057

Table 3. Ablation study.

표 3. 비교 실험 결과(O : 비행체 궤적 실제 데이터 기반 학습, S : 비행체 궤적 합성 데이터 기반 학습)

3.1. [모듈 변화에 따른 비행체 궤적 예측 오차]

Method	Avg. error	
RNN-based	O	0.0178
	S	0.0177
LSTM-based[9]	O	0.0053
	S	0.0057
GRU-based[10]	O	0.0082
	S	0.0085

3.2. [상대값 범위 변화에 따른 비행체 궤적 예측 오차]

Method	Avg. error	
[0, 1]	O	0.0053
	S	0.0057
[0, 100]	O	0.0054
	S	0.0059
[0, 10000]	O	0.0056
	S	0.0060

3.3. [위도/경도 동시 학습 여부에 따른 위도 예측 오차]

Method	Avg. error	
training only with latitude	O	0.0078
	S	0.0080
training with latitude and longitude	O	0.0076
	S	0.0079

3.4. [위도/경도 동시 학습 여부에 따른 경도 예측 오차]

Method	Avg. error	
training only with longitude	O	0.0028
	S	0.0033
training with latitude and longitude	O	0.0031
	S	0.0037

3.5. [리샘플링 여부에 따른 비행체 궤적 예측 오차]

Method	Avg. error	
without resampling	O	0.0053
	S	0.0057
with resampling	O	0.0054
	S	0.0057

비행체 궤적 예측 작업에서 합성 데이터의 활용 가능성을 탐구하기 위하여 다양한 ablation study를 수행하였다. 구체적으로 우리의 최근 이전 연구 [6]에서 효율성이 검증되었던 비행체 궤적 예측 작업의 정량적 평가 방법을 따랐다. 특히 현실성 있는 활용 가능성 탐구를 위해 [5]에서와 같이 “합성 데이터 기반 학습 및 실제 데이터 기반 테스트” 프레임워크를 이용하였다. 표 2는 입출력 관계 변화에 따른 비행체 궤적 예측 오차를 보여준다. 표 2에서 보는 바와 같이 비행체 궤적 실제 데이터 기반 학습 결과와 비행체 궤적 합성 데이터 기반 학습 결과가 거의 비슷함을 볼 수 있다. 즉, 비행체 궤적 예측 작업에서 시계열 생성적 적대 신경망을 이용하여 생성한 비행체 궤적 합성 데이터의 활용 가능성이 매우 높음을 알 수 있다. 표 3은 다양한 ablation study 결과를 보여준다. 표 3에서 보는 바와 같이 순환 신경망 모듈 변화, 상대값 범위 변화, 위도/경도 동시 학습 여부, 리샘플링 여부에 따라 보여주었던 비행체 궤적 실제 데이터 기반 학습 결과 양상과 비행체 궤적 합성 데이터 기반 학습 결과 양상이 거의 비슷함을 볼 수 있다. 이를 통해 비행체 궤적 합성 데이터가 비행체 궤적 예측 작업에서 매우 유용하게 이용될 수 있음을 알 수 있다.

IV. 결론

본 논문에서는 비행체 궤적 데이터의 양이 부족한 경우 생성적 적대 신경망의 학습 및 활용에 대해 탐구하였다. 본 논문에서 제시된 1) 생성 데이터 품질 평가 결과 및 2) 비행체 궤적 예측 작업에서의 활용 가능성 탐구를 위한 비교 실험 결과는 비행체 궤적 합성 데이터 생성 및 비행체 궤적 관련 작업에서 합성 데이터의 활용 가능성에 대해 연구·개발하는 연구자 및 개발자들에게 큰 도움을 제공할 것으로 판단된다.

References

- [1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Neural Information Processing Systems*, pp.2672-2680, 2014. DOI: 10.3156/JSOFT.29.5_177_2
- [2] Y. J. Jo, K. M. Bae, and J. Y. Park, “Research Trends of Generative Adversarial Networks and Image Generation and Translation,” *Electronics and Telecommunications Trends*, vol.35, no.4, pp. 91-102, 2020. DOI: 10.22648/ETRI.2020.J.350409
- [3] O. Mogren, “CC-RNN-GAN: Continuous recurrent neural networks with adversarial training,” *arXiv preprint arXiv:1611.09904*, 2016.
- [4] C. Esteban, S. L. Hyland, and G. Rätsch, “Real-valued (Medical) Time Series Generation with Recurrent Conditional GANs,” *arXiv preprint arXiv:1706.02633*, 2017.
- [5] J. Yoon, D. Jarrett, and M. van der Schaar, “Time-series Generative Adversarial Networks,” *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, vol.32, 2019.
- [6] C. J. Lee, I. H. Park, and C. Jung, “Objective Evaluation of Recurrent Neural Network Based Techniques for Trajectory Prediction of Flight Vehicles,” *Journal of IKEEE*, vol.25, no.3, pp.540-543, 2021. DOI: 10.7471/ikeee.2021.25.3.540 <https://doi.org/10.7471/ikeee.2021.25.3.540>
- [7] J. Kim, C. Jung, D. Kang, and C. J. Lee, “A New Vessel Path Prediction Method using Long Short-term Memory,” *The Transactions of the Korean Institute of Electrical Engineers*, vol.69, no.7, pp.1131-1134, 2020. DOI: 10.5370/KIEE.2020.69.7.1131
- [8] J. Kim, C. Jung, D. Kang, and C. J. Lee, “A New Vessel Path Prediction Method Based on Anticipation of Acceleration of Vessel,” *Journal of Institute of Korean Electrical and Electronics Engineers*, vol.24, no.4, pp.1176-1179, 2020. DOI: 10.7471/ikeee.2020.24.4.1176
- [9] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol.9, no.8, pp.1735-1780, 1997. DOI: 10.1162/neco.1997.9.8.1735
- [10] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.