

장기 시계열 예측 모델을 활용한 가상현실에서의 사이버멀미 선제적 예측*

최윤선⁰¹, 정다영², 김보관¹, 한경식^{1†}

¹한양대학교 데이터사이언스학과

²한양대학교 인공지능학과

yoonseonchoi@hanyang.ac.kr, dayoungjeong@hanyang.ac.kr,

bogoankim@hanyang.ac.kr, kyungsikhan@hanyang.ac.kr

Early Prediction of Cybersickness in Virtual Reality Using a Long-term Time Series Forecasting Model

Yoonseon Choi⁰¹, Dayoung Jeong², Bogoan Kim¹, Kyungsik Han^{1†}

¹Department of Data Science, Hanyang University

²Department of Artificial Intelligence, Hanyang University

요약

사이버멀미는 가상현실(Virtual Reality, VR)에서 사용자의 몰입도를 방해하는 주요한 요소 중 하나이다. 최근 연구들은 인공지능을 활용하여 사이버멀미 예측 모델을 제안해왔으나, 이러한 모델들은 대부분 사용자가 사이버멀미를 경험한 후에야 적용이 가능하여 불편을 겪기 전의 선제적 대응에는 제한적이다. 이를 해결하기 위해, 본 논문에서는 장기 시계열 예측(Long-term Time Series Forecasting, LTSF) 모델을 활용한 사이버멀미 선제적 예측 모델을 제안한다. 제안하는 모델은 45명의 참여자로부터 수집된 VR 표준 장비인 Head Mounted Display (HMD) 센서 데이터를 기반으로, LTSF 모델 중 하나인 PatchTST를 활용하여 구축되었다. 실험 결과, MAE는 0.14, RMSE는 0.67로 baseline 모델인 DLinear 대비 높은 성능을 보였다. 이는 VR 표준 장비인 HMD 센서 데이터만으로도 효과적인 사이버멀미 예측이 가능하며, LTSF 모델을 활용한 선제적 예측이 가능하여 모델 개발의 발전 가능성을 시사한다.

1. 서론

가상현실(Virtual Reality, VR) 기술이 교육, 의료, 엔터테인먼트 등 다양한 분야에서 점차 중요한 도구로 자리 잡으면서, 그 활용도는 계속해서 증가하고 있다[1]. 하지만 VR 사용 중 발생하는 사이버멀미 현상은 사용자에게 눈의 피로, 메스꺼움, 현기증 등 신체적 불편을 초래하여 몰입감을 방해하고, VR의 사용성을 저해하는 주요한 요인 중 하나이다.

최근 사이버멀미 연구들은 VR 사용자 경험을 증진하기 위해 뇌전도, 비디오, 센서 데이터 등을 활용하여 사이버멀미 예측 모델을 제안하고 있다[2]. 특히, 뇌전도 데이터는 사용자의 인지적 활동을 포착하는 데 효과적이지만, Head Mounted Display (HMD) 이외에 추가적인 측정 장치의 침습적 착용은 사용자 불편함과 VR 사용자 경험 감소를 초래할 수 있다. 이로부터 수집된 데이터는 사용자가 편안함을 느끼는 환경에서 수집된 데이터 대비 신뢰도 및 범용성이 부족할 수 있다.

대부분의 사이버멀미 예측 모델은 현재 시점의 point-wise 센서 데이터를 입력으로 사용하여 단일 시점의 사이버멀미를

예측하며, 높은 성능을 보여왔다[3]. 그러나 이러한 모델을 VR 콘텐츠에 탑재할 경우, 사용자가 사이버멀미를 체감하였을 때를 기점으로 완화 조치가 가능하기에 시야각 조절, blurring 등과 같은 기존 사이버멀미 완화 기술의 선제적 적용이 어려울 수 있다.

트랜스포머 기반의 장기 시계열 예측(Long-term Time Series Forecasting, LTSF) 모델은 시계열 데이터의 특성(e.g., 데이터의 증감 경향성)을 학습하여 높은 성능을 보이며[4], 그 중, PatchTST [5]는 다변량 데이터를 분해하여 단변량 데이터로 처리함으로써 높은 시·공간적 계산 효율성을 보인다는 점에서, 사이버멀미와 연관성이 높은 다양한 변수(e.g., 눈/머리 움직임)를 포함하고 있는 VR 멀티모달 센서 데이터 학습에 효율적이다.

본 논문은 HMD에서 수집 가능한 센서 데이터를 기반으로 PatchTST를 활용한 사이버멀미 선제적으로 예측하는 모델을 제안한다. 실험 결과, 본 모델은 baseline 모델인 DLinear[6] 대비 높은 성능(MAE=0.14, RMSE=0.67)을 달성하였다. 이는 HMD 센서 데이터만으로도 사이버멀미를 예측할 수 있음을 입증하고, LTSF 모델을 사이버멀미 도메인에 적용해 선제적 예측이 가능하다는 연구적 의의를 제시한다. 더 나아가 사이버멀미 선제적 예측 모델을 VR 콘텐츠에 탑재함으로써 사용자에게 높은 사용자 경험 제공 가능성을 기대할 수 있다.

* 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 한국연구재단 바이오·의료기술개발사업의 지원을 받아 수행된 연구임(No. NRF-2021M3A9E4080780).

† 교신저자

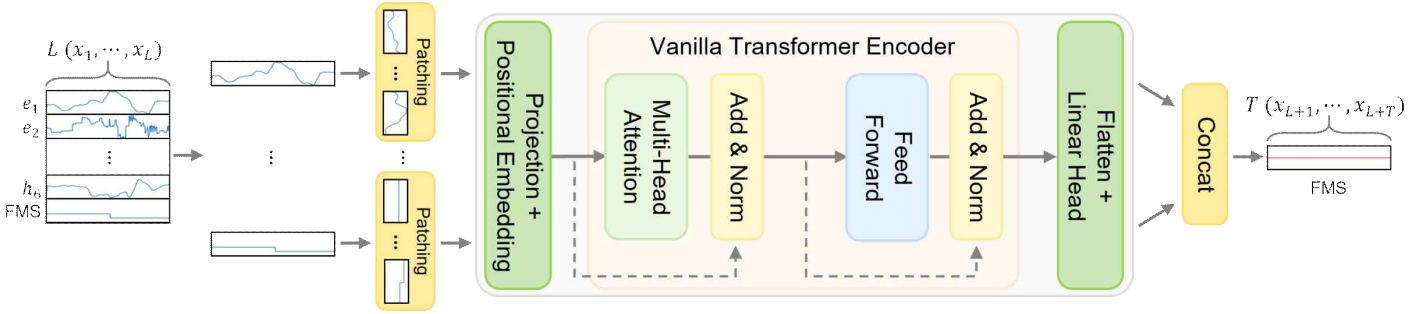


그림 1. PatchTST 전체 구조

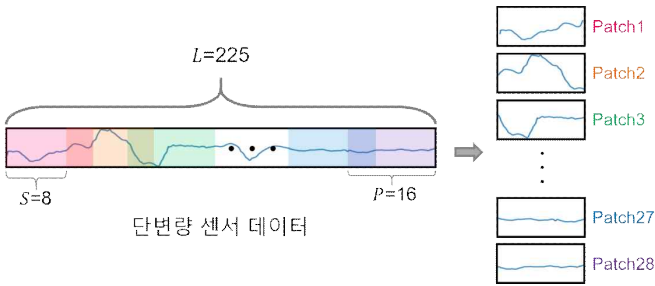


그림 2. Patching

2. 사이버멀미 예측 모델링

2.1. 데이터 수집

우리는 데이터 수집을 위해 HTC VIVE Pro Eye HMD와 선행연구[7]를 통해 선정된 45초 길이의 사이버멀미 유발 영상 20개를 사용자 스터디에 활용하였다. 스터디에는 총 45명의 참여자(평균연령=24.4세; 남자: 24명, 여자: 21명)가 모집되었으며, 각 참여자에게 영상 시청이 끝날 때마다 멀미 정도를 정량적으로 판단할 수 있는 Fast Motion Sickness Scale (FMS) 점수로 응답할 것을 요청했다. FMS 점수는 0부터 20까지로 설정되어 있으며, 0은 ‘멀미가 없음,’ 20은 ‘멀미가 심함’을 나타낸다.

또한, HMD를 통해 2개의 모달리티 센서 데이터를 수집하였다. 센서 데이터는 눈 움직임에 대한 23개의 신호(e.g., 양쪽 눈의 시선 방향, 시선의 중심)와 머리 움직임에 대한 6개의 신호(e.g., 머리 위치 및 회전)로 구성되어 있으며, 이들 신호는 초당 25Hz씩 측정되었다.

2.2. 데이터 전처리

LTSF 모델을 사이버멀미 선제적 예측 모델로 활용하기 위해 수집된 데이터를 모델 구조에 맞게 전처리하였다. 첫 번째로, HMD를 착용하고 벗을 때의 데이터와 중복되는 Hz 값을 제거하여, 영상별 데이터의 길이를 1125Hz (45초*25Hz)로 균일화하였다. 다음으로, 데이터 간 시간의 흐름을 나타내기 위해, 1부터 1125까지 Hz가 표기된 열을 추가하였다. 기존의 LTSF 모델은 연속된 값을 예측하는 구조이지만, FMS는 0에서 20 사이의 값을 갖는 등간척도 데이터이기 때문에, 45초마다 측정된 FMS 값을 1125Hz 길이로 평탄화하였다.

2.3. VR 멀티모달 센서 데이터 기반 PatchTST

우리는 사이버멀미 선제적 예측을 위해 PatchTST 모델을 채택하였다. 그림 1과 같이, PatchTST는 L 길이의 데이터($x_1, \dots, x_L$, x_t 는 t 시점의 모든 신호 값)를 입력받아 T 길이의 미래 데이터($x_{L+1}, \dots, x_{L+T}$)를 예측한다. 이때, L 길이의 다변량 데이터를 단변량 데이터로 분해하여 처리한다. 그림 2는 각 단변량 시계열 데이터를 일정 길이의 연속된 patch로 쪼개는 patching 기법을 보여주며, 하나의 patch를 입력 토큰으로 사용한다. 각 patch는 트랜스포머의 latent space 차원으로 변환되고, 시간적 순서 정보를 보존하기 위해 linear projection 및 positional embedding을 수행한다. 트랜스포머 인코더 및 flatten layer를 거쳐 도출된 각 단변량 데이터는 합쳐진 후 최종 예측값이 도출된다.

우리는 VR 환경에서 수집된 데이터를 PatchTST에 적용하기 위해 모델을 튜닝하였다. VR 멀티모달 센서 데이터는 참여자와 영상별로 시계열 데이터가 구분되어 있다. 따라서 입력 데이터가 연속적으로 모델에 들어갈 때, 한 영상의 마지막 시점 데이터와 다음 영상의 첫 시점 데이터가 함께 입력으로 사용되지 않도록 모델을 개발하였다.

3. 실험 및 결과

3.1. PatchTST 실험 세팅

본 연구에서는 29개 변수(눈 움직임 23개, 머리 움직임 6개)의 센서 데이터와 평탄화한 FMS 데이터를 학습 데이터로 사용하였다. 영상별 1125Hz (45초*25Hz) 길이의 데이터에서, $L=225$ Hz (9초*25Hz) 길이의 데이터를 입력 데이터로 사용하여, 다음 $T=75$ Hz (3초*25Hz) 길이의 FMS score를 예측하였다. 그림 2와 같이, 입력 데이터의 patch 길이 P 는 16, slide 길이 S 는 8로 설정하였다. patch의 개수 N 은 $N = \left\lfloor \frac{L-P}{S} \right\rfloor + 2$ 로 정의되며, 이에 따라 28개의 patch를 입력 토큰으로 설정하였다.

PatchTST에서의 hyperparameter는 batch size=64, epoch=5, learning rate=0.0001, optimizer=Adam으로 설정하였다. 또한, 회귀 모델의 예측 성능 평가를 위해 MAE와 RMSE 평가지표를 사용하였다.

표 1. LTSF 모델별 성능

모델	MAE	RMSE
DLinear	0.25	0.68
PatchTST	0.14	0.67

3.2. LTSF 모델별 성능 비교

DLinear는 단순한 선형 모델임에도 불구하고, 트랜스포머 기반의 모델보다 높은 성능을 보였기에, 비교 모델로 사용하였다. 표 1과 같이, 사이버멀미 예측 모델인 PatchTST는 MAE=0.14, RMSE=0.67 성능을 보였으며, DLinear보다 높은 성능을 보였음을 확인하였다. 이러한 결과는 patching을 적용함으로써 연속된 시계열 데이터의 지역적 정보를 보존하여 모델의 정확도를 높였으며, 다변량 데이터를 단변량 데이터로 처리하는 PatchTST가 다수의 변수를 포함한 VR 멀티모달 센서 데이터에 효율적으로 활용될 수 있음을 의미한다. 또한, 사이버멀미 도메인에서도 트랜스포머 기반의 LTSF 모델을 활용하여 선제적 예측이 가능함을 시사한다.

표 2. 사이버멀미 예측 모델별 성능

모델	모달리티	MAE	RMSE
DeepTCN [8]	눈 움직임 머리 움직임 생체신호	-	0.49
Kundu et al. [9]	LSTM 눈 움직임	0.21	0.36
	GRU 머리 움직임	0.24	0.41
	MLP 생체신호	0.40	0.57
PatchTST (본 연구)	눈 움직임 머리 움직임	0.14	0.67

3.3. 사이버멀미 예측 모델 성능 비교

PatchTST의 patching 기법의 효과성을 확인하기 위해, point-wise를 입력 데이터로 사용한 기존 모델과 성능을 비교하였다. 표 2와 같이, PatchTST는 Kundu et al.[9]의 LSTM보다 MAE가 0.07만큼 감소하여 높은 성능을 보였다. 그러나 RMSE는 0.31만큼의 성능 감소가 있었음을 확인했다. Patching된 센서 데이터는 표준화된 전처리 방식(e.g., 중복값 제거, smoothing)으로는 다소 처리하기 어려운 다수의 이상치를 포함할 수 있으며, RMSE는 이상치에 비교적 민감하기 때문에 성능에 영향을 미쳤을 것으로 예상된다. 또한, 선행연구에서는 모델 학습에 있어 HMD 뿐만 아니라 추가 생리 신호 수집 장비에서도 모달리티를 수집하여 함께 학습을 진행하였다. 이는 사이버멀미와 상대적으로 더 직결되는 features를 학습하여 모델을 구축하였다는 점에서 성능에 긍정적 요인을 주었을 수도 있다. 이러한 점에서, HMD 센서만으로 기존 모델들 대비 상회하는 MAE 값을 보여준 제안된 모델은 사이버멀미 도메인에서 HMD 센서 데이터를 기반으로 patching 기법을 적용하는 방법론의 활용 가능성이 충분함을 시사한다. 또한, 향후 계획된 대규모 사용자 스터디를 통해 추가적으로 데이터를 수집하고, 이를 기반으로 전처리 방법 및 모델 구조를 보완하여 RMSE 또한 상회하는 결과를 기대한다.

4. 결론 및 향후 연구

본 논문에서는 VR 멀티모달 센서 데이터를 기반으로 LTSF 모델을 활용하여 사이버멀미를 선제적으로 예측하는 모델을 제안하였다. 모델의 성능 평가 결과는 사이버멀미 도메인에서도 PatchTST를 사용하여 선제적으로 사이버멀미를 예측할 수 있음을 보여주었다. 또한, 비침습적 기기인 HMD만을 사용하여 수집된 데이터만으로도 사이버멀미의 효과적인 예측이 가능함을 보였다.

본 연구에서 제안한 모델을 VR 콘텐츠에 탑재하였을 때, 모델의 예측값을 기반으로 사이버멀미의 선제적인 완화 조치가 가능하며, 이는 VR 사용자 경험을 높일 수 있음을 시사한다.

참 고 문 헌

- [1] Kim, B., et al. V-DAT (Virtual Reality Data Analysis Tool): Supporting Self-Awareness for Autistic People from Multimodal VR Sensor Data. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (pp. 1-13). 2023.
- [2] Jeong, D., et al. Cybersickness analysis with eeg using deep learning algorithms. In 2019 IEEE conference on virtual reality and 3D user interfaces (VR) (pp. 827-835). IEEE. 2019.
- [3] Jeong, D., et al.. MAC: multimodal, attention-based cybersickness prediction modeling in virtual reality. Virtual Reality, 27(3), 2315-2330. 2023.
- [4] Wu, N., et al. Deep transformer models for time series forecasting: The influenza prevalence case. arXiv preprint arXiv:2001.08317. 2020.
- [5] Nie, Y., et al. A time series is worth 64 words: Long-term forecasting with transformers. arXiv preprint arXiv:2211.14730. 2022
- [6] Zeng, A., et al. Are transformers effective for time series forecasting?. In Proceedings of the AAAI conference on artificial intelligence (Vol. 37, No. 9, pp. 11121-11128). 2023.
- [7] Shahid Anwar, M., et al. Evaluating the factors affecting QoE of 360-degree videos and cybersickness levels predictions in virtual reality. Electronics, 9(9), 1530. 2020.
- [8] Islam, R., et al. Towards forecasting the onset of cybersickness by fusing physiological, head-tracking and eye-tracking with multimodal deep fusion network. In 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 121-130). IEEE. 2022.
- [9] Kundu, R. K., et al. LiteVR: Interpretable and Lightweight Cybersickness Detection using Explainable AI. In 2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR) (pp. 609-619). IEEE. 2023.