

# 5G 환경에서 전자기장 예측을 위한 ViT 기반 멀티모달 물리 데이터 결합 연구

김도연\*, 김성신\*\*  
승실대학교 지능형 반도체학과\*  
승실대학교 AI 융합학부\*\*

## Multimodal Physical Data Fusion using ViT for 5G Electromagnetic Field Estimation

Doeon Kim\*, Seongsin Kim\*\*  
Department of Intelligent Semiconductors, Soongsil University\*  
School of AI Convergence, Soongsil University \*\*  
E-mail : \*ilsin@soongsil.ac.kr, \*\*kss0222@ssu.ac.kr

### Abstract

5G 이동통신망의 최적 설계를 위해서는 정확한 전자기장(EMF) 강도 예측이 필수적이다. 딥러닝 기반 예측 모델은 기존 레이트레이싱 방법 대비 계산 효율성이 뛰어나지만, 단순히 건물 형상 정보(building map)만을 입력으로 사용할 경우 송신기의 복잡한 물리적 특성을 충분히 반영하지 못한다는 한계가 있다. 본 연구에서는 순수 Vision Transformer(ViT)에 기반하여 건물 이미지와 여러 물리 정보를 결합하는 두 가지 멀티모달 퓨전(multimodal fusion) 기법을 제안하고 비교 분석한다. 첫 번째는 물리 데이터를 임베딩하여 토큰 형태로 주입하는 토큰 융합(Token Fusion) 방식이며, 두 번째는 물리 데이터를 2 차원 공간으로 확장하여 입력 채널로 구성하는 2 차원 공간 매핑(2D Spatial Mapping)방식이다. 실험 결과, 두 방식 모두 기존 최고 성능 모델인 HRFormer 보다 우수한 예측 정확도를 보였다. 특히 2D 공간 매핑 방식은 HRFormer 대비 평균 오차(MAE)를 60% 감소시켰고, 토큰 융합 방식 대비 MAE 를 50% 감소시키며, 전파 해석 분야에서 물리적 데이터를 결합할 때 공간적 정합성(spatial alignment)을 유지하는 것이 문맥적 정보 결합보다 더욱 효과적임을 시사한다.

### I. 서론

5G 통신 기술의 상용화로 인해 Massive MIMO

본 논문은 정보통신기획평가원(과제: 2022-0-00986), 한국연구재단 G-LAMP 사업(과제: RS-2025-25441317)의 지원하에 수행되었음.  
정회원: 승실대학교 AI 융합학부, 김성신\*\*

(Multiple-Input Multiple-Output) 기술과 고밀도 소형 셀(small cell) 배치가 가속화되면서, 도심 지역 내 전자파 분포는 더욱 복잡해지고 있다. 이에 따라 인체에 대한 전자파 노출 안전성을 평가하고 관리하는 것이 필수적인 과제로 대두되고 있다.

전통적인 EMF 평가 방식 중 측정 기반 접근법 [1],[2]은 정확도가 높으나 많은 인력과 시간이 소요되며, 모든 지역을 포괄하기 어렵다는 한계가 존재한다. 시뮬레이션 기반 방식[3]은 정밀하지만 계산 비용이 매우 높다는 문제가 있다. 이에 최근에는 딥러닝을 활용한 EMF 예측 연구가 활발히 진행되고 있다. 초기에는 CNN 기반 모델[4]이 주로 사용되었으나, 최근 연구[5]에서는 ViT 계열 모델들이 전역적 문맥을 학습하여 CNN 보다 우수한 성능을 보임을 입증하였다. 특히 [5]에서는 HRFormer 가 가장 우수한 정확도를 달성한 바 있다.

그러나 기존 딥러닝 기반 연구들은 두 가지 한계를 지닌다. 첫째, 물리적 특성이 고려되지 않은 모달리티 퓨전 방식이다. 기존 모델들은 안테나 방사 패턴이나 전력, 기울기 등의 수치 정보를 MLP 를 통해 1 차원 벡터로 변환하여 모델에 입력한다. 이 과정에서 데이터 고유의 공간적 정보가 소실될 수 있다. 둘째, HRFormer 와 같은 계층적 모델의 다운 샘플링과 업 샘플링을 반

복하는 과정에서 전자파 신호의 국소적인 고주파 값들이 평활화되며(smoothing) 신호 정보가 소실되는 현상이 관찰되었다.

본 논문에서는 이러한 문제를 해결하기 위해 순수 Vision Transformer 를 백본(backbone)으로 채택하여 과도한 평활화를 완화하고, 미세한 픽셀 단위 예측에 유리하게 하였다. 또한 안테나의 물리적 정보를 모델에 효과적으로 주입하기 위해 두 가지 새로운 멀티모달 융합 방법을 제안한다. i) 2D 공간 매핑(2D Spatial Mapping): 물리 데이터를 2 차원 공간으로 확장하여 이미지와 결합해 입력 채널로 구성한다. ii) 토큰 융합(Token Fusion): 물리 데이터를 독립적인 토큰으로 변환하여 이미지 토큰과 상호작용한다. 실험을 통해 위 제안 방법의 성능이 기존 HRFormer 대비 정성적, 정량적으로 우수함을 입증하였으며, 전파 예측 문제에 최적화된 멀티모달 데이터 퓨전 전략을 제시하고자 한다.

## II. 데이터셋

입력 데이터는 U-Net 으로 EMF 예측을 연구한 [6]의 데이터셋을 사용하였다. OpenStreetMap 과 USGS 고도 데이터를 활용하여 100 개 미국 대학 캠퍼스의 2D 환경 모델을 생성한 후, 이 모델에 다양한 5G 기지국 구성(안테나 방향, 전송 전력 등)과 WinProp(FEKO) 광선 추적 시뮬레이터를 활용하여 EMF 분포를 계산하여 생성된 데이터셋이다. 초기 시뮬레이션을 통해 생성된 약 700 개의 샘플 데이터 증강 과정을 거쳐 총 약 3,000 개의 훈련 샘플로 구성된다.

| Category | Data                   |
|----------|------------------------|
| Image    | 빌딩 맵                   |
| Pattern  | 수직/수평 방사 패턴            |
| Numeric  | 방위각                    |
|          | SSB 전력                 |
|          | 안테나 기울기                |
|          | 안테나 좌표<br>(경도, 위도, 고도) |

표 1. 입력 데이터 요약

데이터셋 구성은 표 1 과 같다. 빌딩 맵은 500x500 크기의 2 차원 건물 평면도이고, 방사 패턴은 수직, 수평 각도별 이득(gain) 값을 나타내며 1x360 크기를 가진다.

Numeric 데이터인 방위각, 전력, 기울기, 좌표 데이터는 모두 단일 실수 값들이며 1x6 크기의 하나의 벡터로 처리된다.

이러한 물리적 특성이 포함된 특수 데이터는 III 절에서 토큰 융합 방식, 2D 공간 매핑 방식을 적용하여 각각의 물리적 특성을 보존하며 모델에 입력되어 EMF의 특성을 효과적으로 학습 가능하도록 한다.

## III. 제안 방법

본 논문은 빌딩 맵 이미지와 안테나의 물리적 정보를 입력받아 EMF 분포를 예측하는 것을 목표로 한다. 백본 네트워크로는 ViT-Base 아키텍처를 사용하여 입력의 고해상도 특징을 유지하고자 하였고, 물리 기반 데이터를 모델에 효과적으로 주입하기 위한 다음의 두 가지 모달리티 융합 방법을 제안한다.

### 3.1 2D 공간 매핑 방법(2D Spatial Mapping)

2D 공간 매핑은 일종의 Early-Fusion 방식으로, 모든 물리 데이터를 입력 빌딩 맵과 동일한 500x500 해상도의 2 차원 맵으로 변환하여 채널 축으로 결합(Concat)하는 방법이다. 이를 통해 물리 정보가 실제 공간상에 어떻게 분포하는지를 모델에게 직관적으로 제공한다.

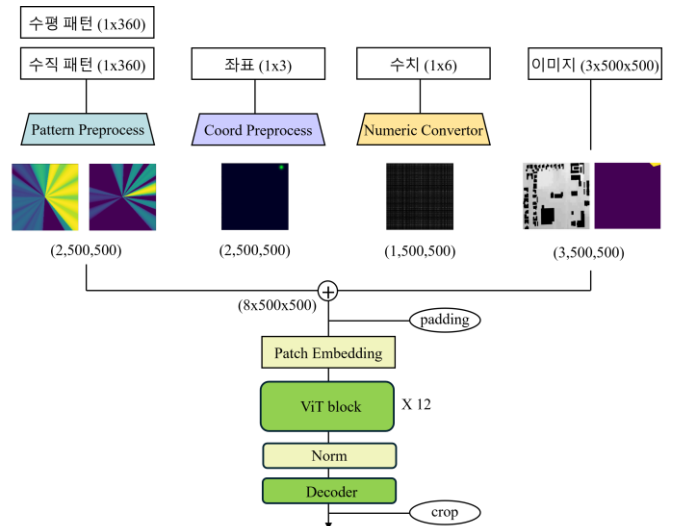


그림 1. 2D 공간 매핑 방법 구조도

좌표 데이터는 그림 1 과 같이 경도, 위도 값으로 1 채널 가우시안 히트맵을 생성하고, 고도 값으로 모든 픽셀 값이 동일한 평면 채널을 생성하

여, 최종적으로 2x500x500 크기의 평면으로 변환된다. 수직/수평 방사 패턴은 2 차원 공간으로 투영된 후 실제 방위각에 맞춰 회전시킴으로써 전파의 각도별 세기와 지향성을 시각적으로 반영한다. 수치 데이터들은 [5]의 MLP 기반 Numeric convertor 를 통해 500x500 크기의 맵으로 확장한다. 최종적으로 생성된 모든 2D 맵들은 건물 맵과 방사 이미지와 함께 채널 축으로 결합되어 ViT 의 입력으로 주입되며, 이를 통해 모델은 첫 번째 레이어부터 이미지 정보와 물리 정보가 상호작용하는 Early-Fusion 학습을 수행하게 된다.

### 3.2 토큰 융합 방법 (Token Fusion)

토큰 융합 방법은 일종의 Mid-Fusion 방식으로, 각 입력 모달리티를 독립적인 토큰으로 변환하여 self-attention 메커니즘 내에서 상호작용하도록 설계되었다. 이미지 모달리티는 빌딩 맵, 방사 이미지로, 패턴 모달리티는 안테나의 수직/수평 방사 패턴, 수치 모달리티는 표 1 의 Numeric 데이터들로 정의한다.

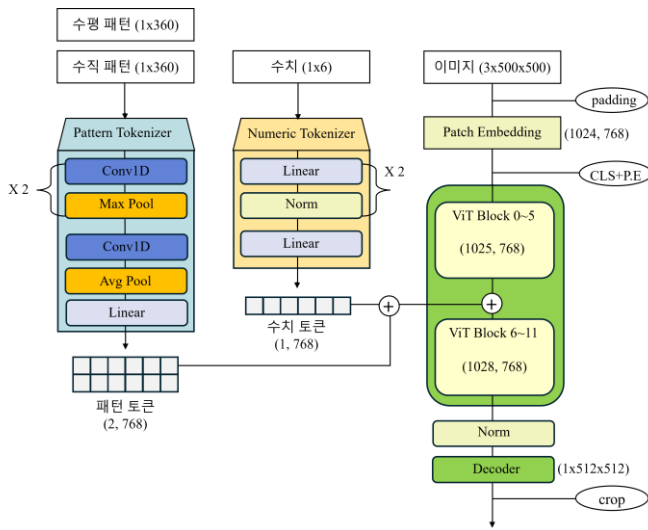


그림 2. 토큰 융합 방법 구조도

이러한 이중 데이터들은 그림 2 와 같이 각각의 토큰으로 변환된다. 패턴 모달리티의 경우 Pattern Tokenizer 를 통해 1D Convolution 으로 360° 연속 데이터의 특징을 효과적으로 추출 후 Max Pooling 을 통해 길이를 점진적으로 줄인 다음, Average Pooling 연산을 거쳐 1x768 크기의 특징 벡터로 압축된다. 수치 모달리티는 Numeric Tokenizer 의 Linear 연산을 통해 수치 데이

터 간 관계를 학습한 후 1x768 크기의 단일 토큰으로 변환된다. 모달리티 융합은 ViT 의 전체 12 개의 인코더 블록 중 후반부에서 진행된다. 초기 6 개 블록에서는 이미지 패치 토큰만을 입력 받아 self-attention 으로 공간적 특징을 우선적으로 학습하고, 이후 후반 6 개 블록에서 앞서 인코딩된 패턴 토큰과 수치 토큰을 이미지 토큰에 주입함으로써 모델이 이미지의 공간 정보와 안테나의 물리적 정보를 심층적으로 교차 학습하게 한다.

## IV. 실험

### 4.1 실험 환경

실험은 NVIDIA GeForce RTX 4090, 3090 GPU 에서 진행되었고, 초기 학습률 0.0001 의 Adam 옵티마이저를 사용하였다. 배치 크기는 4 로 설정하였고, 학습은 200 에포크 동안 진행되었다. ViT-Base 는 ImageNet-1K 데이터셋에서 사전 학습된 가중치를 사용하였으며, 자세한 파라미터 설정은 표 2 에 제시되어 있다.

| Parameter        | Value |
|------------------|-------|
| Patch size       | 16    |
| Hidden dimension | 768   |
| Depth            | 12    |
| Number of heads  | 12    |
| MLP dimension    | 3072  |
| Head dimension   | 64    |
| Dropout rate     | 0.0   |

표 2. ViT Backbone 하이퍼 파라미터

### 4.2 실험 결과

기존 연구[5]에서 가장 우수한 성능을 보였던 HRFormer 를 베이스라인으로 선정하여, 본 논문에서 제안한 두 가지 방법들과 비교하였다. 평가 지표로는 MAE 와 이미지의 품질 검증 지표인 PSNR, SSIM 을 사용하였다. 평가는 전파가 도달하지 않는 건물 내부를 제외한 실외 영역에 대해서만 수행된다.

| Method                   | MAE (↓) | PSNR (dB, ↑) | SSIM (↑) |
|--------------------------|---------|--------------|----------|
| HRFormer (baseline)      | 0.05    | 18.08        | 0.85     |
| ViT (Token Fusion)       | 0.04    | 25.46        | 0.93     |
| ViT (2D Spatial Mapping) | 0.02    | 28.33        | 0.94     |

표 3. 베이스라인과 제안 방법의 정량 평가 비교

두 방법 모두 단일 MLP 변환을 사용하여 데이터를 인코딩한 베이스라인 대비 월등한 성능 향상을 보였다. 토큰 융합 방식의 경우 MAE 20%, PSNR 41%, SSIM 9%의 성능 향상이 있었고, 2D 공간 매핑 방식은 MAE 60%, PSNR 57%, SSIM 11% 향상으로 물리 정보를 토큰이나 이미지 형태로 변환하여 모델에 주입한 제안 방법의 우수성을 입증하였다.

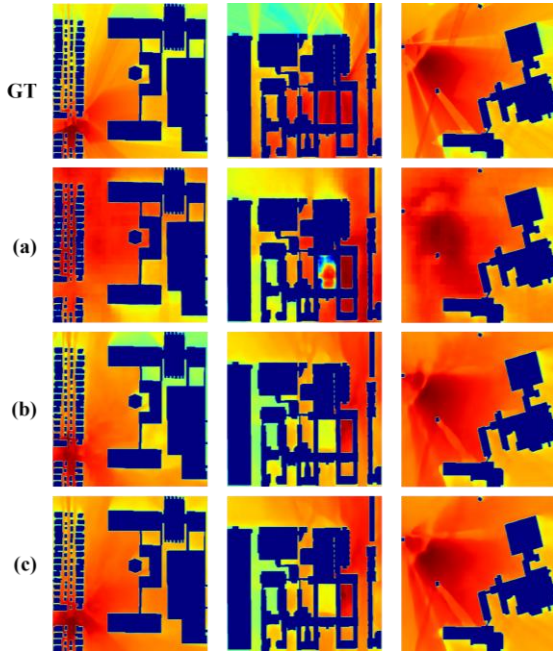


그림 3. 방법별 EMF 예측 결과 시각화.

(a): baseline, (b): Token Fusion, (c): 2D Spatial Mapping

실제 EMF 분포 예측 결과를 시각화 후 정성적으로 비교했을 때에도 베이스라인 대비 두드러진 예측 정확도 향상을 확인할 수 있었다. 그림 3의 첫 번째 열에서 (a)는 강한 신호가 평활화되어 예측 결과가 전체적으로 붉은색을 띠는 반면, (b), (c)에서는 전파 세기의 차이를 잘 포착한 것을 확인할 수 있다. 두 번째 열에서는 (a), (b), (c) 모두 전파 세기 경향성을 잘 포착하였으나, (a)는 건물 중간 빈 부분에 대한 예측 품질이 떨어졌다. 세 번째 열에서는 (b), (c)는 전파의 선형성을 잘 예측한 반면, (a)에서는 이러한 선형성이 아예 드러나지 않음을 확인하였다.

## V. 결론

본 논문에서는 ViT 기반의 EMF 예측 성능을 극대화하기 위해 이중 물리 데이터를 결합하는 최적의 방법을 탐구하였다. 실험 결과, 물리 데이터를 활용한 두 가지 제안 기법인 토큰 융합, 2D 공간 매핑 방법 모두 모든 평가 지표에서 베이스라인 모델인 HRFormer를 능가하였으며, 특히 물리 정보를 2차원 공간으로 확장하여 결합하는 2D 공간 매핑 기법이 가장 우수한 성능을 보였다. 이는 5G 기지국 설계 시 물리적 파라미터와 공간 정보를 통합적으로 고려하는 것이 필수적임을 시사하며, 제안된 방법은 향후 다양한 EMF 예측 및 네트워크 분석 모델의 기반 기술로 활용될 수 있을 것이다.

## 참고문헌

- [1] Kapetanakis, Theodoros N., et al. "Assessment of radiofrequency exposure in the vicinity of school environments in Crete Island, South Greece." *Applied Sciences* 12.9 (2022): 4701.
- [2] Kopacz, Thomas, et al. "Effective SSB beam radiation pattern for RF-EMF maximum exposure assessment to 5G base stations using massive MIMO antennas." *2021 15th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2021.
- [3] Lodato, Francesca, et al. "Ray Tracing Tools Assessment for the Evaluation of EMF Levels Generated by 5G NR Systems: An Overview." *2024 IEEE International Symposium on Measurements & Networking (M&N)*. IEEE, 2024.
- [4] Anusha, Veluru Sai, G. K. Nithya, and Sethuraman N. Rao. "A comprehensive survey of electromagnetic propagation models." *2017 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2017.
- [5] D. Kim et al., "Estimation of Electromagnetic Field Strength: Experiments Using Vision Transformers," in *IEEE Access*, doi: 10.1109/ACCESS.2025.3631972.
- [6] Park, Dongryul, et al. "5G Base Station Electromagnetic Field Strength Estimation Method in Complex Hotspot Area using Deep Learning." *2024 IEEE International Symposium on Electromagnetic Compatibility, Signal & Power Integrity (EMC+ SIPI)*. IEEE, 2024.