

멀티모달 데이터 융합 기반 작물 병해 실시간 예측 시스템 개발

임대환*, 황현호**

Multimodal Fusion Framework for Real-Time Crop Disease Forecasting

Lim DaiHwan*, Hwang HyunHo**

요약

본 연구는 작물 병해 예측의 정확도와 실시간성을 극대화하기 위해 멀티모달 데이터를 통합한 혁신적인 예측 프레임워크를 제안한다. 기존의 유전체 및 환경 데이터에 더하여, 드론 기반 이미지와 IoT 센서 데이터를 통합 분석함으로써 질병 발생의 시공간적 패턴을 정밀하게 추적한다. CNN-LSTM 기반 하이브리드 모델에 Transformer 기반 Feature Fusion Layer를 결합하여, 시계열적 질병 진행과 영상 기반 병징 변화, 환경적 트리거를 통합적으로 반영하였다. 본 시스템은 현장 실증을 통해 기존 모델 대비 예측 정확도와 조기 경고 성능이 향상되었음을 입증하였으며, 지속 가능한 병해 관리와 스마트팜 운영에 실질적인 기여가 가능함을 확인하였다.

Abstract

This study proposes an innovative prediction framework that maximizes the accuracy and real-time performance of crop disease forecasting through the integration of multimodal data. In addition to conventional genomic and environmental data, the model incorporates and analyzes drone-based imagery and IoT sensor data to precisely track the spatiotemporal patterns of disease outbreaks. A hybrid model combining CNN and LSTM is enhanced with a Transformer-based feature fusion layer to holistically capture time-series disease progression, visual symptom development, and environmental triggers. Field validation demonstrated that the proposed system significantly outperforms existing models in prediction accuracy and early warning capabilities, offering practical contributions to sustainable disease management and smart farming operations.

Keywords

multimodal data, crop disease prediction, real-time forecasting, deep learning, smart farming

1. 서론

기후 변화, 병원체 다양성, 환경 스트레스 등으로 인해 작물 병해 발생 양상은 점점 복잡해지고 있다. 이에 따라 정밀한 예측 기술의 수요가 급증하고 있으며, 특히 실시간으로 작물 병해를 감지하고 조기

대응할 수 있는 기술의 필요성이 대두되고 있다. 기존 연구는 오믹스와 기상 데이터를 활용한 정적인 분석이 중심이었으며, 실제 병징이나 환경 변화에 대한 실시간 반영은 미흡하였다[1].

본 연구는 작물 병해 예측의 기술적 어려움을 오믹스·환경 데이터 활용, 드론 및 고정형 카메라 기

* 호서대학교 스마트팜AI데이터센터 부교수
- ORCID: <https://orcid.org/0000-0002-0749-7983>
** 한양사이버대학교 국방융합기술학과 조교수(교신저자)
- ORCID: <https://orcid.org/0009-0001-4065-0943>

• Received: Jul. 31, 2025, Revised: Sep. 30, 2025, Accepted: Oct. 03, 2025
• Corresponding Author: Hwang HyunHo
Hanyang Cyber University, 220 Wangsimni-ro, Seongdong-gu, Seoul, Korea
Tel.: +82-2-2290-0303, Email: dhmk85h3@naver.com

반 영상 기반 병해 진단 기술, IoT 센서를 활용한 생육 환경 분석, 멀티모달 융합을 위한 딥러닝 프레임워크 적용하여 한계를 극복하고자 한다. 특히, CNN-LSTM 모델에 Transformer 기반 데이터 융합층을 도입함으로써 시공간적 예측의 정밀도를 향상시키고, 실제 농업 현장에 적용 가능한 실시간 경보 시스템을 구현하였다.

II. 시스템 구조 및 설계

본 연구에서 제안하는 멀티모달 기반 작물 병해 예측 시스템의 유전체, 환경, 시각, 센서 데이터를 통합하여 시공간적으로 정밀한 병해 예측을 가능하게 한다. 시스템은 세 가지 핵심 구성요소는 그림 1과 같이 데이터 수집 및 정제, 멀티모달 특성 추출 및 통합, 병해 예측 모델 및 실시간 경보 시스템이다[2].

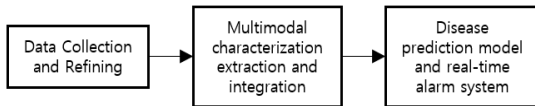


그림 1. 시스템 구조 및 설계
Fig. 1. System structure and design

2.1 데이터 수집 및 정제

그림 2와 같이 연구에 사용된 데이터는 총 세 가지 유형으로 구성되며, 각각의 수집 경로와 전처리 방식은 다음과 같다. 유전체 및 표현형 데이터는 AI-Hub 농업 플랫폼에서 제공되는 주요 작물의 유전자 서열 정보와 주요 표현형 형질 데이터를 활용하였다. 이는 질병 저항성과의 상관관계를 규명하기 위한 입력으로 사용된다[3]. 기상 및 환경 데이터는 NOAA 및 KMA의 공공 API를 통해 수집된 일일 평균 기온, 상대 습도, 강수량 데이터를 이용하였으며, 시간 동기화를 통해 질병 발생 시점과 정렬하였다. 영상 데이터는 드론 및 고정형 RGB 카메라를 통해 촬영된 작물의 고해상도 이미지와 열화상 데이터를 수집하였다. 병정의 조기 발생 탐지를 위해 5초 간격의 프레임으로 수집하였으며, YOLO 기반 사전 필터링으로 이상 프레임을 제거하였다. IoT 센서 데이터는 토양 수분, 온도, 조도, pH, CO₂ 농도 등의 생육환경 데이터를 실시간으로 수집하였으며, 10분

간격의 시계열 데이터로 구성되었다. 결측치 보간은 KNN 및 Cubic Spline을 병행하여 처리하였다.

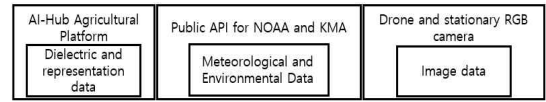


그림 2. 데이터 수집
Fig. 2. Data collection

모든 데이터는 타임스탬프 기준으로 동기화되며, Min-Max 정규화 및 노이즈 필터링(예: Z-score 이상치 제거)을 거쳐 딥러닝 학습에 적합한 형태로 전처리된다[4][5].

2.2 멀티모달 특성 추출 및 통합

영상 데이터는 ResNet50 기반의 CNN(Convolutional Neural Network)를 활용하여 공간적 병징 패턴을 추출하며, 그림 3과 같이 최종 Layer에서 512차원의 벡터로 임베딩되고 시계열 센서 및 기상 데이터는 LSTM(Long Short-Term Memory) 네트워크를 통해 시간 축에 따른 변화 패턴을 학습하며, 병해 유발 트리거와의 상호작용을 모델링한다.

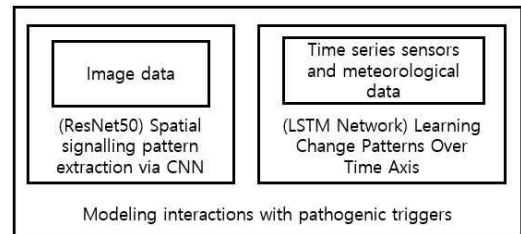


그림 3. 병해 유발 트리거와의 상호작용 모델링
Fig. 3. Modeling the interaction with disease-inducing triggers

오믹스 데이터의 경우는 그림 4와 같이 VAE(Variational Autoencoder)를 통해 차원 축소 및 노이즈 제거된 유전자 표현을 추출하며, 고차원 유전정보를 병해 민감성 지표로 반영한다[6].

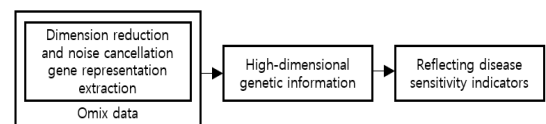


그림 4. 오믹스 데이터 특성 추출 및 통합
Fig. 4. Extraction and integration of omics data features

다음 단계로 그림 5는 Self-Attention 기반의 Transformer Encoder를 구성하여 각 모달 간의 상호 작용 정보를 학습하고, 질병 예측에 유의미한 통합 표현을 생성하고 마지막으로 Cross-modal attention weight를 활용하여 각 모달의 기여도를 동적으로 반영하였다[7][8].

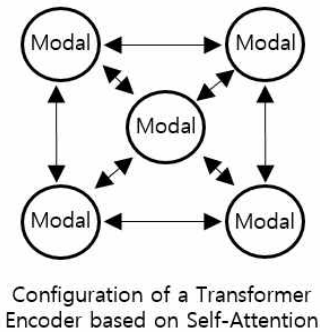


그림 5. Self-Attention 기반의 Transformer Encoder 구성
Fig. 5. Configuration of a Transformer Encoder based on Self-Attention

2.3 병해 예측 모델 및 실시간 경보 시스템

통합된 멀티모달 표현은 Fully Connected Layer 및 Softmax 출력층을 통해 병해 발생 확률을 추정하며, 출력값은 다음과 같다. 먼저 Disease Risk Score: 0~1 사이의 연속값으로 표현된 병해 발생 가능성도, Alert Level: Green(정상), Yellow(주의), Red(고위험) 등급, Disease Type Prediction: 주요 병해 유형(예: 도열병, 탄저병 등)의 분류 결과로 나타난다. 실시간 경보 시스템은 해당 출력값을 기반으로 사용자 맞춤형 경고 메시지를 전송하며, 대시보드에는 병징 진행 타임라인, 환경 조건 변화 그래프, 권장 방제 시점 등이 시각화된다. 이 시스템은 클라우드 기반 아키텍처로 구축되며, 데이터 동기화는 MQTT(Message Queue Telemetry Transport) 프로토콜을 기반으로 이루어진다. 또한, 주기적 학습 재훈련 및 외부 실측 데이터 반영을 통해 모델 성능을 지속적으로 향상시킨다.

그림 6에서 본 시스템은 드론·센서·환경 데이터를 포함한 멀티모달 데이터를 수집·전처리하여 병해 예측 모델에 입력하고, 성능평가를 통해 유효성을 검증하였다. 특히, 유전체·환경·영상·IoT 데이터를

를 통합하고 Transformer 기반 융합을 적용함으로써 시공간적 병해 패턴을 정밀하게 포착하고 실시간 경보 기능을 제공하는 점에서 기존 연구와 차별성을 갖는다.

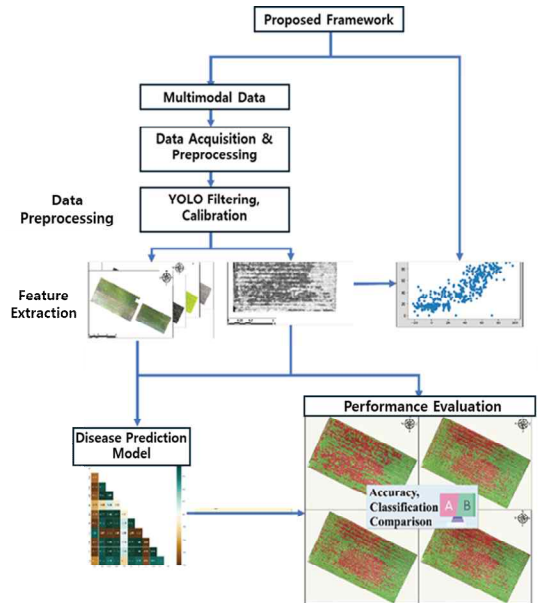


그림 6. 제안된 멀티모달 기반 병해 예측 시스템 구조
Fig. 6. System architecture of the proposed multimodal crop disease prediction framework

III. 결론

본 논문은 멀티모달 데이터를 융합한 작물 병해 예측 프레임워크를 제안하였다. 제안된 모델은 유전체·환경·영상·IoT 센서 데이터를 통합하고 Transformer 기반 Cross-modal Attention Layer를 적용함으로써 기존 모델에 비해 성능이 크게 향상되었다. 표 1의 결과에서 확인할 수 있듯이, LSTM 단일 모델(Baseline 1)은 Accuracy 0.81, F1-score 0.74, 조기경보 리드타임 1.4일, CNN 단일 모델(Baseline 2)은 Accuracy 0.79, F1-score 0.72, 리드타임 1.2일에 머물렀다. CNN과 LSTM을 결합한 하이브리드 모델(Baseline 3)은 Accuracy 0.86, F1-score 0.82, 리드타임 1.8일로 개선되었으나, 여전히 초기 병징 탐지에는 한계가 있었다. 반면, 제안 모델은 Accuracy 0.94, Precision 0.91, Recall 0.90, F1-score 0.91을 달성하였고, 리드타임도 3.6일로 향상되어 기존 모델 대비

모든 평가지표에서 유의미한 성능 향상을 입증하였다. 특히 병해 발생 최소 3일 전에 조기 경보를 발신할 수 있어 방제 전략 수립에 실질적인 도움이 될 수 있음을 확인하였다. Baseline 1과 2는 단일 작물 중심, Baseline 3은 병해 심각도 분류에 초점을 두는 반면, 본 연구의 모델은 멀티모달 데이터 통합과 실시간 경보 기능에서 뚜렷한 차별성을 가진다.

표 1. 실험 결과

Table 1. Experimental results

Model name	Accuracy	Precision	Recall	F1 Score	Early warning lead time
Baseline 1 (LSTM only)	0.81	0.77	0.72	0.74	1.4d
Baseline 2 (CNN only)	0.79	0.75	0.7	0.72	1.2d
Baseline 3 (CNN+LSTM)	0.86	0.83	0.81	0.82	1.8d
Proposed model	0.94	0.91	0.9	0.91	3.6d

성능 향상의 원인은 세 가지로 요약된다. 첫째, 멀티모달 데이터 통합으로 단일 모달 대비 평균 7~10%의 정확도 향상을 확보하였다. 둘째, Transformer 기반 융합층이 CNN-LSTM만 사용하는 경우보다 F1-score를 약 9% 높이며, 초기 병징 탐지 민감도를 개선하였다. 셋째, Ablation 실험 결과 영상(35%), 기상·센서(40%), 오믹스(25%) 데이터가 상호보완적으로 기여하여 최종 성능을 높이는 데 중요한 역할을 하였다. 다만 특정 지역·작물에 대한 일반화 성능은 아직 제한적이며, 흐린 날씨나 해 질 녘 등 영상 품질 저하 상황에서 성능이 감소하는 한계가 있다. 또한 현 시점에서는 고성능 GPU 환경에서만 안정적인 학습이 가능하므로, 향후에는 다지역 학습과 데이터 증강, 조도 보정 기법 적용, Edge 환경 최적화 및 경량화 연구가 필요하다.

종합적으로, 제안된 프레임워크는 단일 모달 모델이나 기존 하이브리드 모델보다 우수한 성능을 입증하였으며, 실시간 병해 예측과 조기 경보를 통해 스마트팜 현장에서 실질적으로 활용될 수 있는 가능성을 보여주었다. 이는 향후 지속 가능한 농업과 스마트농업 운영을 위한 중요한 기여가 될 것으로 기대된다.

References

- [1] H. Choudhury and A. Kumar, "Climate change and plant diseases: Understanding the complex relationship for sustainable agriculture", *Frontiers in Sustainable Food Systems*, Vol. 7, pp. 1152031, Mar. 2023. <https://doi.org/10.1111/ppa.14033>
- [2] H. Noh, G. Jang, and J. Lee, "A Comparative Study on the Performance of RNN-based Deep Learning Models for Tomato Yield and LAI Prediction", *KIIT*, Vol. 23, No. 1, Jan. 2025. <https://doi.org/10.14801/jkiit.2025.23.1.21>.
- [3] S. McCouch, et al., "Agriculture: Feeding the future", *Nature*, Vol. 499, No. 7456, pp. 23-24, Jul. 2013. <https://doi.org/10.1038/499023a>.
- [4] R. K. Varshney, et al., "Designing future crops: Genomics-assisted breeding comes of age", *Trends in Plant Science*, Vol. 26, No. 6, pp. 631-649, Jun. 2021. <https://doi.org/10.1016/j.tplants.2021.03.010>.
- [5] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey", *Computers and Electronics in Agriculture*, Vol. 147, pp. 70-90, Apr. 2018. <https://doi.org/10.1016/j.compag.2018.02.016>.
- [6] A.-K. Mahlein, "Plant disease detection by imaging sensors—Parallels and specific demands for precision agriculture and plant phenotyping", *Plant Disease*, Vol. 100, No. 2, pp. 241-251, Feb. 2016. <https://doi.org/10.1094/pdis-03-15-0340-fe>.
- [7] J.-H. Bae, J.-H. Lee, G.-H. Yu, G. J. Kwon, and J.-Y. Kim, "A Study about Learning Graph Representation on Farmhouse Apple Quality Images with Graph Transformer", *Smart Media Journal*, Vol. 12, No. 1, pp. 9-16, Feb. 2023. <https://doi.org/10.30693/smj.2023.12.1.9>.
- [8] C. Hwang, C. P. Yoon, and S. W. Lee, "Attention-based word correlation analysis system for big data analysis", *KIICE*, Vol. 27, No. 1, pp. 41-46, Jan. 2023. <https://doi.org/10.6109/jkiice.2023.27.1.41>.