

다중모달 기반 생육·환경 데이터 융합을 통한 수직농장 작물 성장 예측 모델 개발

(Development of a Crop Growth Prediction Model for Vertical Farms through Multimodal-Based Fusion of Growth and Environmental Data)

김현준*, 이명훈**

(Kim Hyun Jun*, Lee Meong Hun**)

요약

본 연구의 목적은 수직농장에서 수집한 영상 및 환경 데이터를 통합하여 작물의 생육 단계와 성장률을 정밀하게 예측할 수 있는 다중모달 융합 모델을 개발·평가하는 것이다. 이를 위해 수직농장 실증지에서 일정 기간 동안 RGB 영상과 CO₂ 농도, 온도, 상대습도, 조도 등 주요 환경 데이터를 동기화하여 수집하고, 전처리와 지표 산출을 거쳐 다중모달 학습용 데이터셋을 구축하였다. 영상 데이터로부터 엽면적, 엽색도, 형상지표, proxy VI 등의 특징을 추출하고, 환경 센서 데이터의 시계열 변화를 LSTM으로 학습한 뒤 CNN 기반 영상 특징과 병합하여 생육 단계 분류 및 성장률 예측을 동시에 수행하도록 모델을 설계하였다. 다중모달 융합 모델은 단일모달(CNN 또는 LSTM) 대비 생육 단계 분류 정확도와 성장률 예측 오차에서 모두 우수한 성능을 보였다(정확도 91.3%, RMSE 0.028). 혼동행렬 분석에서도 각 생육 단계별 높은 분류율이 확인되었으며, 실제값과 예측값 비교에서도 일관성 있는 예측이 나타났다. 또한, 입력 변수별 중요도 분석을 통해 엽면적과 proxy VI가 주요 단서로 작용하고 CO₂ 농도와 광량이 예측력 향상에 기여함을 확인하였다. 이 연구는 수직농장 환경에서 다중모달 데이터를 활용해 작물의 생육 상태를 보다 정밀하게 진단·예측할 수 있는 모델링 프레임워크를 제시함으로써, 향후 스마트 수직농장의 데이터 기반 의사결정과 제어 기술 고도화에 기여할 수 있는 가능성을 보이고 있다.

■ 중심어 : 다중모달 융합 모델 ; 수직농장 ; 생육 단계 ; 성장률 예측 ; 데이터 통합

Abstract

The purpose of this study is to develop and evaluate a multi-modal fusion model that integrates image and environmental data collected from a vertical farm to accurately predict crop growth stages and growth rates. RGB images and key environmental parameters, including CO₂ concentration, temperature, relative humidity, and light intensity, were synchronously collected over a defined period at a vertical farm test site, and a multi-modal training dataset was constructed through preprocessing and feature extraction. Leaf area, greenness, shape indices, and proxy VI were extracted from the image data, while the temporal variations of the environmental sensor data were modeled using a long short-term memory (LSTM) network. These were then merged with convolutional neural network (CNN) -based image features to perform simultaneous growth stage classification and growth rate prediction.

The multi-modal fusion model outperformed single-modal models (CNN or LSTM alone) in both growth stage classification accuracy and growth rate prediction error (accuracy 91.3%, RMSE 0.028). Confusion matrix analysis also showed high classification rates for each growth stage, and the comparison between predicted and actual growth rates confirmed consistent performance. Feature importance analysis revealed that leaf area and proxy VI were the most influential variables, while CO₂ concentration and light intensity contributed to improved predictive power. This study presents a modeling framework for accurately diagnosing and predicting crop growth status in vertical farming environments using multi-modal data, demonstrating its potential to enhance data-driven decision-making and control technologies in smart vertical farms.

■ keywords : Multi-modal Fusion Model ; Vertical Farm ; Growth Stage ; Growth Rate Prediction ; Data Integration

1. 서론

최근 인구 증가와 도시화, 기후변화로 인해

안정적이고 지속 가능한 식량 생산의 필요성이 커지며, 이에 대한 대안으로 밀폐형 환경제어가 가능한 수직농장이 주목받고 있다. 수직농장은 도심이나

* 정회원, 국립순천대학교 스마트농업전공

** 정회원, 국립순천대학교 융합바이오시스템기계공학과

이 논문은 2025학년도 국립순천대학교 글로벌대학 사업비의 지원을 받아 연구되었음

접수일자 : 2025년 10월 15일

수정일자 : 1차 2025년 11월 05일, 2차 2025년 11월 17일

게재확정일 : 2025년 11월 21일

교신저자 : 이명훈 e-mail : leemh777@snu.ac.kr

농촌의 유희시설을 활용해 설치할 수 있어 공간 활용도와 연중 생산성의 보장이 높다는 장점을 가지고 있다. 그러나, 현존하는 수직농장은 대부분 설정값 기반의 단순 자동화 운영에 머물러 있어 작물의 생육 반응을 실시간으로 반영하지 못하고, 이로 인해 광합성과 무관한 에너지 및 이산화탄소(CO₂) 투입이 발생하여 운영비 상승과 생산성 저하의 원인으로 작용하고 있다[1,2].

그동안 LED 광원 효율 개선, 냉난방 시스템 최적화, CO₂ 공급 장치 개선 등 개별 설비 중심의 연구가 이루어져 왔으며, 일부 연구에서는 IoT 센서를 활용한 환경 모니터링과 자동제어 시스템을 구축하여 에너지 효율을 높이고자 한 바 있다. 그러나, 이러한 접근은 조명·온도·CO₂ 등 핵심 생육 환경 요소를 독립적으로 운용하거나 시간 기반으로 제어하는 수준에 머물러 있어, 작물 생육 단계별 요구량을 반영하지 못하고 있다. 국내외 기술 현황을 살펴보면 스펙트럼 센서를 이용한 LED 제어, DAC(Direct Air Capture) 기반 CO₂ 순환 시스템 등 일부 고도화 사례가 존재하지만, 여전히 생육 반응 중심의 통합 제어 수준에는 미치지 못하고 있다[3,17].

작물의 생육 반응을 정량적으로 파악하고, 이를 환경제어에 반영하기 위해서는 단일 센서 정보만으로는 한계가 있다. 영상 기반 생육 지표(엽면적·엽색·형상)와 CO₂·온도·조도·습도 등 환경 센서값을 동시에 수집·분석하는 다중모달 접근은 보다 정밀한 생육 진단 예측을 가능하게 한다[4]. 특히 시간·위치가 일치하는 데이터 셋을 기반으로 합성곱신경망(CNN)과 장단기기억망(LSTM) 등 최신 딥러닝 기법을 적용하면, 생육 단계 분류와 생장률 예측, 광합성 반응 추정 등 다중 출력 모델 구현이 가능하다.

본 연구는 수직농장 내에서 작물의 영상 데이터와 환경 센서 데이터를 통합적으로 수집·분석하여 생육 단계와 생장률을 동시에 예측할 수 있는 다중모달 기반 생육·환경 데이터 융합 모델을 개발하는 것을 목표로 한다. 본 논문은 다음과 같이 구성된다. 2장에서는 연구에 사용된 데이터 수집 방법과 전처리 과정, 그리고 다중모달 융합 모델의 설계 및 학습 방법을 설명한다. 3장에서는 제안한 모델을 수직농장 실증 데이터에 적용한 결과와 성능 평가

지표를 제시한다. 4장에서는 결과를 바탕으로 다중모달 융합이 생육·환경 예측 성능 향상에 미치는 효과와 한계점을 서술한다. 끝으로 5장에서는 본 연구의 결론과 향후 연구방향을 제시하며 마친다.

II. 관련 연구

최근 수직농장 자동화 분야에서는 영상·환경 데이터를 융합하여 작물 생육을 정량적으로 진단하고 예측하려는 시도가 활발히 이루어지고 있다. 이는 기존의 설정값 기반 단순 제어 방식에서 벗어나, 실제 생육 반응을 반영한 데이터 중심 제어로 전환하려는 흐름과 맞닿아 있다. Lee et al. (2024)은 Transformer 기반의 다중모달 구조를 활용하여 상추의 엽면적과 생장률을 동시 예측하였으며, 기존 CNN 기반 모델 대비 유의미한 예측 성능 향상을 보고하였다[5]. 이러한 연구는 고정형 영상 데이터를 정량화하여 생육 정보를 추출하는 방식의 한계를 넘어서, 시계열적 생육 변화를 반영하는 특징을 가진다는 점에서 의의가 있다.

Park et al. (2023)은 영상 정보와 IoT 환경센서를 결합한 Edge-AI 기반 제어 구조를 제안하여, 실시간 분석을 통한 온도 및 광량 조절의 효율성을 제시하였다[6]. 이 연구는 영상·환경 데이터의 결합이 단순 예측 모델 수준을 넘어 실제 제어 구조에 적용될 수 있음을 보여주었다는 점에서 의미가 있다. 더 나아가 Wang et al. (2024)은 Vision Transformer (ViT)과 Temporal Transformer (TSFM)를 결합한 모델을 제안하여, 작물 생육 예측에서 장기적 추세 파악의 중요성을 강조하였다. 그러나, 본 연구는 실증 기반 다중모달 데이터에 대한 통합형 생육 단계·생장률 동시 예측 모델링 프레임워크를 제시함으로써 기존 Transformer 기반 접근과 차별화된 실용성과 범용성을 제공하며, RGB 기반 생육지표 + 환경 센서 시계열을 실증 기반으로 동기화한 다중모달 데이터셋을 구축한 점, 그리고 CNN-LSTM 결합 구조에 환경 변수 상호작용(Quadratic term) 및 Feature-wise Linear Modulation(FiLM)을 통합하여 실제 수직농장

조건에서의 예측 성능을 검증한 점에 고유한 기여가 있다[7]. 이러한 연구들은 다중모달 접근이 높은 일반화 성능을 제공할 수 있음을 시사하지만, 실질적으로는 실증 데이터 확보의 한계로 인해 여전히 시뮬레이션 또는 제한적 환경에서의 검증이 대부분이라는 공통된 한계를 갖는다.

한편, 환경 데이터 기반 제어 기술 중 스마트 관수 분야에서도 다양한 연구가 축적되어 왔다. Bwambale et al. (2023)은 모델 기반 관수 제어 전략을 적용하여 작물 재배 단계별 물 사용 효율 변화를 분석하였으며, 환경 변수에 따른 제어 정책 변화가 실질적인 효율 향상에 기여할 수 있음을 제시하였다[8]. 이러한 연구는 환경 반응 기반 제어 모델이 단순 임계값 방식보다 우수한 제어 성능을 제공할 수 있음을 뒷받침한다.

Upendra Roy et al. (2024)은 IoT 환경센서와 머신러닝 모델을 결합하여 관수 여부를 자동 판단하는 모델을 제시하였으며, 센서 기반 수분 상태 예측이 자동 관수 시스템 구성의 핵심 요소임을 확인하였다[9]. Ndunagu et al. (2022)은 무선 센서 네트워크(WSN) 기반의 관수 시스템을 구축하여 토양수분과 기상 데이터를 실시간으로 수집·송신하는 구조를 검증하였고, 이는 스마트팜 전반에서 데이터 기반 제어의 필수 인프라 구성을 설명하는 중요한 근거가 된다[15].

또한 Allam et al. (2022)은 머신러닝 알고리즘을 통해 관수에 필요한 수분 수준을 분류하는 모델을 제안하였으며, 환경 변수의 패턴 분석이 관수 의사결정 모델의 정밀성을 향상할 수 있음을 보고하였다[11]. Hamouda et al. (2024)은 토양수분 기반 관수를 위한 센서 네트워크를 설계·검증함으로써, 실증 환경에서의 자동 관수 구조가 안정적으로 작동함을 제시하였다[12]. 이러한 스마트 관수 관련 연구들은 환경 데이터 기반 의사결정 모델이 실제 농업 현장에서 활용 가능성을 확인한 실증적 연구들로, 환경 데이터 분석과 제어 알고리즘 구성의 중요성을 강조한다.

이처럼 영상 기반 생육 분석 연구와 IoT·환경

데이터를 활용한 제어 기술 연구는 각각의 영역에서 상당한 발전을 이루어 왔으나, 두 데이터 유형을 동시에 활용하여 작물 생육 단계와 생장률을 통합적으로 예측하는 다중모달 접근은 상대적으로 제한적으로 보고되고 있다. 기존 연구 상당수는 영상 또는 환경 데이터 중 하나에 집중하거나, 시뮬레이션 기반의 평가에 머무르는 경향을 보이며, 실제 수직농장 실증 데이터가 동기화된 형태로 확보된 경우는 매우 드물다.

따라서 실증 기반 다중모달 데이터를 활용하여 작물 생육 상태를 종합적으로 예측하는 연구는 학술적으로도 그 필요성이 높다고 할 수 있다. 본 연구는 수직농장 실증지에서 수집된 영상·환경 데이터를 시공간적으로 동기화하여 다중모달 융합 모델을 설계하고, 생육 단계와 생장률을 동시에 예측할 수 있는 구조를 제시하였다. 이는 기존 연구의 한계로 지적된 실증 기반 검증의 부족 문제를 해소하는 동시에, 영상·환경 데이터 융합의 효과성을 실험적으로 확인했다는 점에서 학술적·실무적 의의를 갖는다.

III. 시스템 구성 및 모델링

1. 시스템 구조 및 설계 원리

본 연구는 본 대학에서 공동연구를 수행 중인 수직농장 기업의 실증지를 대상으로 진행하였으며, 작물 대상은 상추로 선정하였다.

실증지는 순천시 소재의 기존 유희시설을 활용하여 조성된 층고 2.5 m, 3단 구조의 소형 수직농장으로, LED 조명, CO₂ 주입기, 냉·난방 장치, 환경 센서, 영상 촬영 장치 등을 기본 구성으로 갖추고 있다.

각 층에는 환경 센서와 RGB 카메라가 설치되어 있어 생육 단계별 환경 반응을 세밀하게 모니터링 할 수 있으며, CO₂ 는 공통 공급라인을 사용하되 층별 센서 데이터를 통해 농도 차이를 확인하고 필요 시 밸브 조절을 통해 주입 패턴을 차등 운용하였다.

이를 통해 상업용 농장의 일반적인 구조를 유지하면서도 연구 목적에 맞는 데이터 확보가 가능하도록 하였다. 실험 기간 동안 각 층의 환경 조건은 LED 광량 200~350 $\mu\text{mol}/\text{m}^2\cdot\text{s}$, CO₂ 농도 400~900

ppm, 온도 22~26°C, 상대습도 60~70% 범위 내에서 조절하였다. 영상 데이터는 6시간 간격으로 해상도 1920×1080의 RGB 카메라로 촬영하였으며, 초기·중기·수확기 각 생육 단계별로 120장씩 총 360장의 영상을 확보하였다. 환경 센서 데이터는 동일 기간 동안 CO₂ 농도, 온도, 조도, 상대습도를 분 단위로 수집하였으며, 영상 촬영 시각과 동일한 타임스탬프를 부여하여 시공간적으로 동기화된 형태의 다중모달 데이터셋을 구축하였다.



그림 1. 수직농장 실증 모듈 구성

LED 조명은 작물의 광합성에 필요한 기본 파장을 제공하는 상용 LED를 사용하였으며, 층별로 광량을 조절할 수 있게 설치하였다. 냉·난방 장치는 설정값을 변경해 다양한 재배 환경을 구현할 수 있도록 하였고, 실험 기간 동안 각 설비의 운전 조건과 에너지 사용량을 병행 기록하였다.

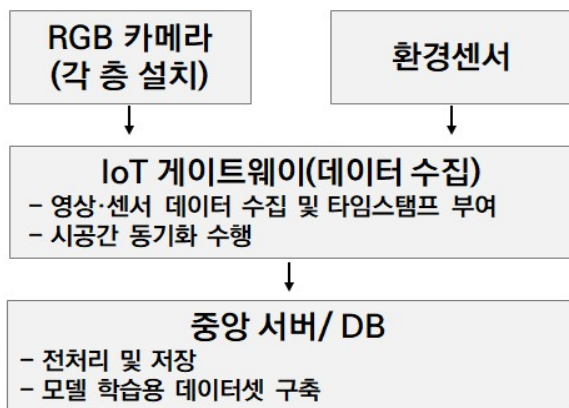


그림 2. 영상·환경 데이터 수집 및 동기화 흐름도

영상 데이터는 RGB 카메라를 이용하여 일정 주기로 촬영하였고, 주요 생육 단계(초기 - 중기 - 수확기)별 엽면적·엽색·형상 변화를 확인하였다. 필요 시 간단한 필터를 적용하여 조명·노출 조건을

맞춘 후 분석에 활용하였다. 환경 센서값은 CO₂ 농도, 온도, 조도, 상대습도를 중심으로 수집하였으며, 영상 촬영 시각과 동일한 시간 정보를 부여해 자료를 정리하였다. 이를 통해 영상과 환경 데이터가 시공간적으로 일치된 형태로 확보되었다.

2. 데이터 수집 및 전처리

본 연구에서는 수직농장 실증지에서 수집된 영상 데이터와 환경 센서 데이터를 사용하였다. 영상은 각 재배층 상부에 설치된 RGB 카메라로 일정 간격(6시간) 촬영하였으며, 촬영 시각·노출·조명 조건을 일정하게 유지하였다. 촬영된 영상은 분석 단위별로 크롭(crop)하고 화이트 밸런스 조정 및 왜곡 보정을 통해 품질을 표준화하였다.

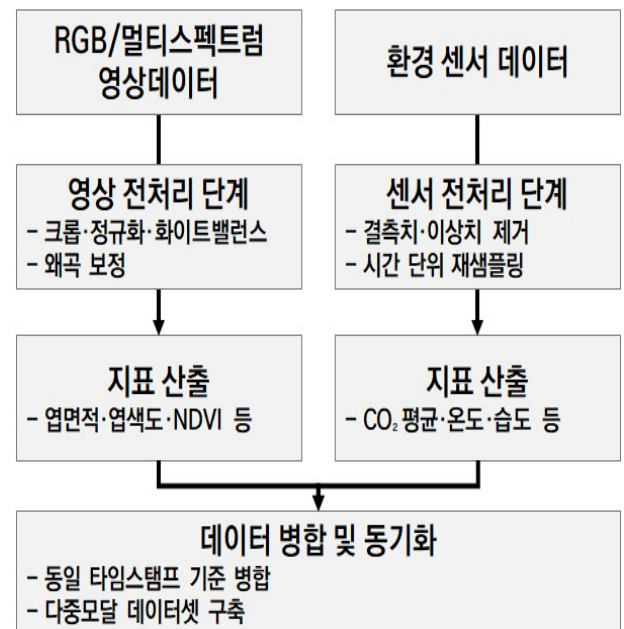


그림 3. 영상·센서 데이터 전처리 및 병합 과정

환경 센서값은 CO₂ 농도, 온도, 조도, 상대습도를 중심으로 수집하였고, 모든 데이터에 타임스탬프를 부여하여 영상과 시공간적으로 동기화하였다. 센서값의 결측치나 이상치는 평균·표준편차 기반의 간단한 필터링으로 제거하였으며, 영상 데이터와 동일한 시간 단위로 재샘플링(resampling)하여 분석에 활용하였다. RGB 영상으로부터 엽면적, 엽색, 형상 등 기본 지표를 추출하였으며, proxy VI (Normalized Difference Vegetation Index), 녹색 채널 비율 등 광합성 활성도 지표도 함께 산출하였다.

표 1. 주요 생육·환경 지표 정의 및 산출 방법

지표명	계산식 / 산출 방법	단위	사용 목적
엽면적	영상에서 잎 영역을 추출해 면적 계산	cm ²	작물의 생육 정도 및 성장률 평가
엽색도	영상에서 잎의 색상값을 활용해 계산	-	엽록소 함량·건강 상태 추정
형상지표	잎의 모양·둘레 등을 이용해 산출	-	엽형 변화 및 생육 단계 판별
proxy VI	멀티스펙트럼 영상의 적색·근적외 값 이용	-	광합성 활성도 및 건강 상태 평가
녹색채널 비율	영상의 녹색값 비율로 계산	-	엽색 변화 모니터링
CO ₂ 농도	센서 측정값 사용	ppm	환경 요인 및 광합성 반응 분석
온도		℃	생육 적정 환경 유지 확인
상대습도		%	증산 및 수분 상태 추정
조도		lux	광량 및 LED 운영 상태 확인

각 지표는 표준화된 계산식과 단위로 변환하여 모델 입력 피처로 사용하였다[13].

본 연구에서는 멀티스펙트럼 NDVI 대신 RGB 카메라 영상만을 활용하여 Proxy Vegetation Index(대체 녹색성 지표, Proxy VI)를 산출하였다.

사용된 주요 지표는 $NGRDI = (G-R)/(G+R)$, $VARI = (G-R)/(G+R-B)$, $ExG = 2G-R-B$ 이며, 일반적인 NDVI 구분하기 위해 'Proxy VI'로 표기하였다.

RGB 기반 지표의 유효성을 검증하기 위해, 일부 샘플에 대해 휴대형 proxy VI 센서 측정값과의 스피어만 상관관계수(ρ)를 산출하였다.

본 연구는 무작위 분할 대신 배치 단위 교차 검증을 수행하였다. 또한, 조명 강도와 색온도의 변화에 따른 편차를 완화하기 위해 y보정, hue 및 brightness 조정 등의 포토메트릭 증강 기법을 적용하였다. 영상 추출 과정에는 Instance Normalization을 도입하여 조명 조건 변화에 대한 강건성을 확보하였다.

3. 다중모달 융합 모델 설계 및 학습

본 연구에서는 영상 데이터와 환경 센서 데이터를 동시에 활용하기 위해 CNN-LSTM 기반의 다중 모달 융합 모델을 설계하였다[14].

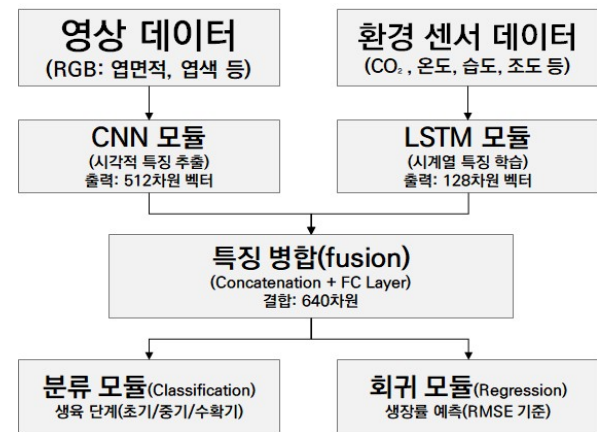


그림 4. 다중모달 융합 모델 구조도

영상 데이터는 합성곱신경망(Convolutional Neural Network, CNN)을 이용하여 엽면적, 엽색, 형상 등의 시각적 특징(feature map)을 추출하였고, 환경 데이터는 장단기기억망(LSTM)을 통해 시간에 따른 CO₂, 온도, 조도, 습도의 변화를 학습하였다. 두 모듈의 출력 feature vector는 병합층(concatenation layer)에서 결합된 후, 완전연결층(Fully Connected Layer)을 거쳐 생육 단계 분류(분류 모듈)와 생장률 예측(회귀 모듈)을 동시에 수행하도록 구성하였다.

CNN은 사전학습된 ResNet-18 모델을 기반으로 하였으며, 마지막 합성곱층(conv5)에서 추출된 512차원 특징 벡터를 사용하였다.

LSTM은 입력 시퀀스 길이(window size) 10, hidden size 128, dropout 0.2로 구성하였고, 최종 hidden state를 128차원 feature vector로 변환하여 CNN feature와 결합하였다.

결합된 640차원(512+128) feature vector는 256차 fully connected layer를 거친 후 Softmax 분류층(생육 단계 예측)과 Linear 회귀층(생장률 예측)으로 분기된다.

환경 시계열 데이터(온도, CO₂, 광량)는 LSTM으로 임베딩한 후, 영상 특징 벡터에 대해 Feature-wise Linear Modulation(FiLM) 계층을 적용하여 비선형 상호작용을 반영하였다.

추가적으로, Quadratic Layer를 삽입하여 ($T \times CO_2$, $T \times$ 조도, $CO_2 \times$ 조도 등) 교차항을 학습할 수 있도록 설계함으로써 환경 변수 간 상호영향을 내재적으로 포착하였다.

표 4. 단일모달과 다중모달 성능 비교

모델	입력 데이터	생육 단계 분류 정확도(%)	F1-score(%)	생장률 예측 MSE	생장률 예측 MAE
CNN 단일모달	영상 특징	85	83.5	0.035	0.145
LSTM 단일모달	환경 특징	78.2	77.4	0.042	0.167
다중모달 융합(CNN+LSTM)	영상+환경 특징	91.3	90.8	0.028	0.122

표 2. 모델 입력·출력 변수 및 평가 지표

변수 구분	항목	내용(예시)	단위 / 지표
입력	영상 특징	엽면적, 엽색도, 형상지표, proxy VI, 녹색채널 비율 등	cm ² , - -
입력	환경 특징	CO ₂ 농도, 온도, 상대습도, 조도	ppm, °C, %, lux
출력	생육 단계	초기, 중기, 수확기(3단계 분류)	Class
출력	생장률	생육 단계별 생장률 예측값	%/day 또는 g/day
평가	분류 성능	정확도(Accuracy), F1-score	%
평가	회귀 성능	평균제곱오차(MSE), 평균절대오차(MAE)	단위 없음(정규화)

본 절에서 기술한 절차에 따라 모델 구조와 학습 방법을 명확히 규정함으로써, 이후 결과(3장)에서 단일모달 대비 성능 향상을 비교 분석할 수 있도록 하였다.

IV. 시스템 적용 및 실증 결과

1. 데이터셋 구축 결과

본 절에서는 2장에서 기술한 실험 시스템과 전처리 과정을 거쳐 구축된 데이터셋의 규모와 특성을 제시한다. 본 연구에서는 수직농장 실증지에서 수집한 영상 및 환경 센서 데이터를 사용하였다. 연구기간동안 각 재배층에서 RGB 영상과 CO₂ 농도, 온도, 상대습도, 조도 데이터를 동기화하여 수집하였으며, 전처리와 지표 산출을 거쳐 다중모달 학습용 데이터셋을 구축하였다[10].

수집된 영상 데이터는 전체 360장으로, 초기·중기·수확기 세 생육 단계로 분류하였다. 환경

센서 데이터는 동일 기간 동안 분 단위로 측정된 총 1,200건으로 구성되었다. 전처리 과정에서 결측치와 이상치가 제거된 후 영상과 센서 데이터가 동일한 타임스탬프 기준으로 매칭되어 최종적으로 360개의 다중모달 샘플이 확보되었다.

표 3. 데이터셋 규모 및 구성

구분	내용	건수
영상 데이터	RGB 영상 (초기·중기·수확기 포함)	360장
환경 센서 데이터	CO ₂ 농도, 온도, 상대습도, 조도	1,200건
생육 단계 분류	초기	120샘플
	중기	120샘플
	수확기	120샘플
최종 다중모달 데이터셋	영상과 센서 데이터 매칭 결과	360샘플

2. 단일모달과 다중모달 모델의 성능 비교

본 절에서는 단일모달(CNN 또는 LSTM 단독)과 다중모달 융합 모델의 성능을 비교한 결과를 제시한다.

영상 데이터만을 입력으로 사용하는 CNN 단일모달 모델과 환경 센서 데이터만을 입력으로 사용하는 LSTM 단일모달 모델을 각각 학습·평가하였으며, 동일한 데이터셋을 사용하여 다중모달 융합 모델(CNN+LSTM)을 학습·평가하였다[16]. 세 모델 모두 3.1절에서 구축한 데이터셋을 학습(70%)·검증(15%)·테스트(15%) 세트로 구분하여 사용하였다.

분류 성능 평가에는 정확도(Accuracy)와 F1-score를, 회귀 성능 평가에는 평균제곱오차(MSE)와 평균절대오차(MAE)를 사용하였다. 모델별 성능은 표 5에 요약하였으며, 그림 5에는 정확도 및 RMSE의 비교 그래프를 나타냈다.

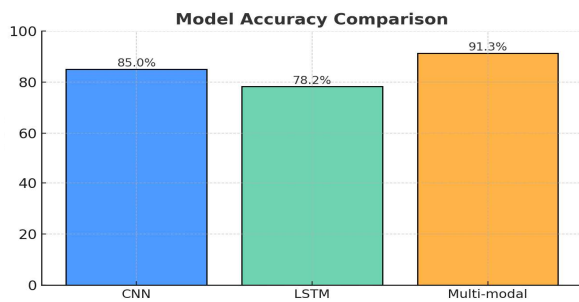


그림 5-1. 모델별 정확도 비교 그래프

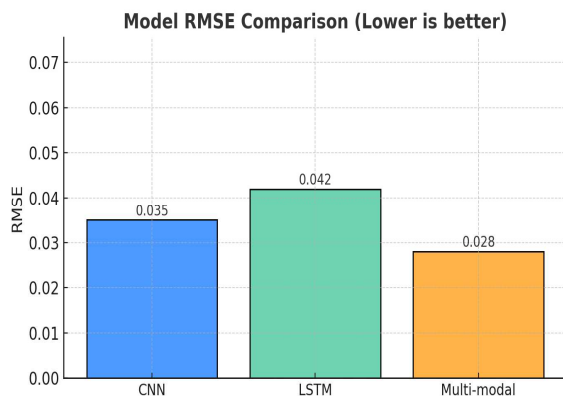


그림 5-2. 모델별 RMSE 비교 그래프

그림 5-1에 나타난 바와 같이, CNN 단일모달 모델의 생육 단계 분류 정확도는 85.0%였으며, LSTM 단일모달 모델은 78.2%로 상대적으로 낮게 나타났다. 반면, 다중모달 융합 모델은 91.3%의 정확도를 보여 두 단일모달 모델보다 우수한 성능을 나타냈다. 그림 5-2의 생장률 예측 RMSE 비교에서도 CNN 단일모달 모델은 0.035, LSTM 단일모달 모델은 0.042인 반면, 다중모달 융합 모델은 0.028로 가장 낮은 값을 보였다. 이러한 결과는 영상과 환경 데이터를 통합해 학습하는 다중모달 접근법이 단일모달보다 생육 단계 분류와 생장률 예측에서 모두 높은 성능을 제공함을 시사한다.

3. 생육 단계 분류 및 생장률 예측 결과

본 절에서는 다중모달 융합 모델을 적용하여 얻은 생육 단계 분류 결과와 생장률 예측 결과를 제시한다. 3.1절에서 구축한 데이터셋을 이용하여 학습된 모델을 테스트 세트에 적용한 결과, 생육 단계(초기·중기·수확기) 분류의 전체 정확도는 91.3%로 나타났다. 혼동행렬을 분석한 결과, 초기

단계는 94.0%, 중기 단계는 89.5%, 수확기 단계는 90.3%로 분류되었으며, 특정 단계 간 경계 구간에서 일부 오분류가 발생하였다.

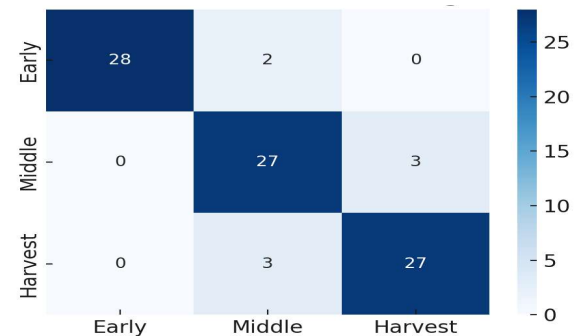


그림 6. 다중모달 모델의 생육 단계 분류 혼동행렬

생장률 예측의 경우, 테스트 세트에서 평균제곱오차(MSE)는 0.028, 평균절대오차(MAE)는 0.122로 나타나 단일모달 모델 대비 낮은 예측 오차를 보였다. 실제값과 예측값을 비교한 산점도에서도 두 값이 1:1 선 근처에 분포하여 예측의 일관성이 확인되었다.

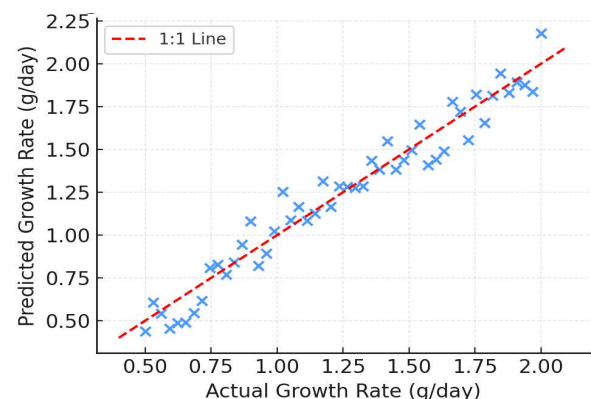


그림 7. 생장률 예측 실제값과 예측값 비교 그래프

그림 6에 나타난 바와 같이, 다중모달 융합 모델의 생육 단계 분류 결과는 초기 단계 94.0%, 중기 단계 89.5%, 수확기 단계 90.3%로 높은 정확도를 보였으며, 일부 경계 구간에서만 오분류가 발생하였다. 이는 영상 특징과 환경 변수를 동시에 학습함으로써 각 생육 단계의 패턴을 보다 정밀하게 구분할 수 있었기 때문으로 판단된다.

그림 7에서는 생장률 예측의 실제값과 예측값을 비교한 결과를 나타낸다. 산점도가 1:1 기준선 부근에 밀집되어 있어 모델이 실제 생장률을 안정적으로 추정하고 있음을 확인할 수 있으며,

RMSE 및 MAE도 단일모달 모델보다 낮게 나타나 예측 성능이 향상되었음을 보여준다.

4. 특징 중요도 및 상관분석

본 절에서는 다중모달 융합 모델의 입력 변수들(영상 지표와 환경 변수)이 모델 예측에 미치는 상대적 중요도를 분석하고, 주요 지표 간의 상관관계를 제시한다. 모델 학습 완료 후, SHAP(Shapley Additive exPlanations) 값을 이용해 각 입력 변수의 기여도를 산출하였다. 분석 결과, 영상 지표 중에서는 엽면적과 proxy VI가, 환경 변수 중에서는 CO₂ 농도와 조도가 생육 단계 분류 및 생장률 예측에 가장 큰 영향을 미친 것으로 나타났다.

또한 피어슨 상관계수를 이용하여 주요 영상·환경 지표 간 상관관계를 분석하였다. proxy VI와 엽면적은 생장률과 양(+)의 높은 상관관계를 보였으며, CO₂ 농도와 조도 역시 일정 수준 이상의 양(+)의 상관을 나타냈다. 표 5는 각 변수별 상관계수를 요약하고, 그림 8은 입력 변수별 상대적 중요도를 시각화한 결과를 나타낸다.

표 5. 생장률 상관계수 주요지표

변수	상관계수 (r)
Leaf Area	0.82
proxy VI	0.78
CO ₂	0.65
Light Intensity	0.58
Temperature	0.4
Humidity	0.35

표 5는 주요 영상 및 환경 지표와 생장률 간의 상관계수를 나타낸다. 엽면적과 proxy VI는 각각 0.82, 0.78로 높은 양(+)의 상관관계를 보여 생장률을 설명하는 핵심 변수임을 확인할 수 있었다. CO₂ 농도와 광량 역시 0.65, 0.58로 유의한 양(+)의 상관을 보였으며, 온도(0.40)와 상대습도(0.35)는 비교적 낮은 상관을 나타냈다. 이러한 결과는 영상 지표가 환경 변수보다 생장률과의 직접적인 연관성이 더 높음을 시사한다.

그림 8은 다중모달 융합 모델 학습 후 산출된 입력 변수별 상대적 중요도를 나타낸 것이다. 엽면적과

proxy VI의 중요도가 각각 0.25와 0.22로 가장 높았으며, 그 뒤를 CO₂ 농도(0.18), 광량(0.15), 온도(0.12), 상대습도(0.08)가 차지하였다. 이는 모델이 영상과 환경 데이터를 함께 사용하되, 특히 영상 지표(엽면적·proxy VI)를 주요 단서로 활용하여 생육 단계 분류 및 생장률 예측을 수행하고 있음을 보여준다.

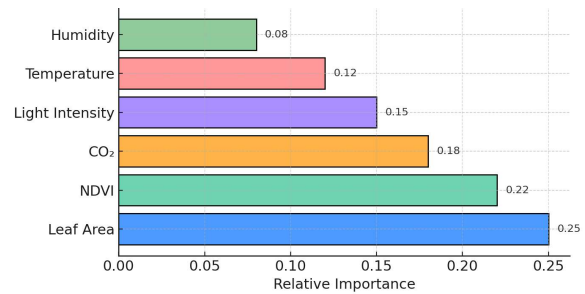


그림 8. 다중모달 모델 입력 변수별 상대적 중요도

본 연구에서 제안한 다중모달 융합 모델은 단순히 생육 상태를 예측하는 데 그치지 않고, 예측 결과를 수직농장 내 환경제어 시스템과 연계할 수 있다는 점에서 실무적 의미를 가진다. 예를 들어, 모델이 특정 생육 단계(증가→수확기 전환기)를 탐지할 경우, 제어 시스템은 LED 광량을 단계적으로 감소시키거나 CO₂ 주입량을 조정하여 에너지 효율을 높일 수 있다. 또한 예측된 생장률이 기준치를 하회할 경우, 냉난방 및 조도 제어 파라미터를 자동 보정하는 피드백 루프를 구성함으로써 운영자의 개입을 최소화할 수 있다.

이러한 구조는 기존의 설정값 기반 자동화에서 데이터 기반 의사결정형 자동화로의 전환을 가능하게 하며, 향후 수직농장의 AI 기반 자율제어 기술 상용화에 기여할 수 있을 것으로 판단되었다.

V. 결론

본 연구에서는 수직농장 실증지에서 수집된 영상 및 환경 센서 데이터를 기반으로 다중모달 융합 모델을 설계·학습하여 작물의 생육 단계 분류와 생장률 예측을 수행하였다. 3장에서 제시한 실험 시스템과 전처리 과정을 통해 영상·환경 데이터가 시공간적으로 정렬된 다중모달 학습용 데이터셋을 구축하였으며, 4장에서 제시한 결과에 따르면 다중모달 모델은 단일모달(CNN 또는 LSTM)

대비 생육 단계 분류 정확도와 생장률 예측 오차에서 모두 우수한 성능을 보였다.

이러한 결과는 영상 지표(엽면적·proxy VI 등)와 환경 변수(CO₂, 광량 등)를 함께 활용하는 접근법이 개별 데이터만을 사용하는 것보다 작물의 상태를 보다 정밀하게 반영할 수 있음을 시사하였다. 특히 엽면적과 proxy VI가 모델의 주요 단서로 작용한 점과 CO₂ 농도·광량의 기여도가 확인된 점은, 향후 스마트 수직농장 운영에서 센서 설치와 데이터 수집 전략을 최적화하는 데 참고가 될 수 있을 것으로 보였다. 다만 본 연구는 한 실증지와 제한된 기간의 데이터를 사용하였다는 점에서 일반화에 한계가 있으며, 향후 연구에서는 다양한 작물과 환경 조건, 더 긴 기간의 데이터를 포함해 모델의 범용성과 안정성을 검증할 필요가 있다. 또한 환경 변수의 비선형 상호작용과 장기 예측 가능성을 높이기 위한 추가적인 모델링 기법의 활용이 필요할 것으로 사료된다.

다만 본 연구는 단일 실증지 및 단일 작물(상추)을 대상으로 진행되어 모델의 일반화에 제약이 있다.

향후 연구에서는 CO₂, 광량, 온도 등의 설정을 다양화하고, 여러 작물과 장기간 데이터를 포함하여 모델의 범용성과 안정성을 재검증할 예정이다.

또한 RGB 기반 녹색성 지표는 멀티스펙트럼 NDVI의 완전한 대체가 아니므로, 본문에서는 혼동을 방지하기 위해 'NDVI' 대신 'Proxy VI'로 표기하였다.

본 연구는 수직농장에서 확보한 다중모달 데이터를 활용해 생육 단계와 생장률을 동시에 예측할 수 있는 모델링 프레임워크를 제시함으로써, 향후 스마트 수직농장의 데이터 기반 의사결정과 제어 기술 개발에 기여할 수 있는 가능성을 확보할 수 있을 것으로 기대된다.

REFERENCES

[1] Z. Lin, R. Fu, G. Ren, R. Zhong, Y. Ying, and T. Lin, "Automatic Monitoring of Lettuce Fresh Weight by Multi-modal Fusion Based Deep Learning," *Frontiers in Plant Science*, vol. 13, article 980581, 2022, pp. 1-13.

[2] 조한진, "디지털 농업을 위한 딥러닝 기반의 환경 인자 추천 기술 연구," *스마트미디어저널*, 제12권 제5호, 65-72쪽, 2023년

[3] P. Morella, M. P. Lambán, J. Royo, and J. C.

Sánchez, "Vertical Farming Monitoring: How Does It Work and How Much Does It Cost?," *Sensors*, vol. 23, no. 7, article 3502, pp. 22-30, 2023.

[4] Kim, J. S. G.; et al. "Development of a Machine Vision-based Fresh Weight Estimation System to Monitor Crops from Transplant Stage to Harvest." *Frontiers in Plant Science*, vol. 15, article 1332489, pp. 1 - 10, 2024.

[5] Lee, J.; Han, S.; Park, H.; Kim, D. "Transformer-Based Multimodal Crop Growth Estimation in Indoor Vertical Farms." *Sensors*, vol. 24, no. 3, 1172, pp. 1 - 15, 2024.

[6] Park, Y.; Choi, K.; Seo, H.; Lee, M. "Edge-AI Fusion Framework for Environmental and Image-Based Control in Smart Vertical Farms." *Agronomy*, vol. 14, no. 1, 85, pp. 1 - 12, 2023.

[7] Wang, Z.; Liu, X.; Zhang, Y.; Sun, J. "Temporal Transformer for Multi-Sensor Time-Series Forecasting of Crop Growth in Controlled Environments." *Frontiers in Plant Science*, 14, 1332489, 2024.

[8] Bwambale, E., Abagale F. K., Anornu G. K., "Model-based Smart Irrigation Control Strategy and Its Effect on Water Use Efficiency in Tomato Production," *Cogent Engineering* 10, 2259217, 2023.

[9] Upendra Roy B.P, Abdul Sattar. K. Ahmed A. Elngar, "A Smart Irrigation System Using the IoT and Advanced Machine Learning Model", *Journal of Smart Internet of Things(JSIoT)*, vol. 2024, no. 2, pp. 13-25, 2024.

[10] Juliana Ngozi Ndunagu, Kingsley Eghonghon Ukhurebor, Moses Akaaza, Robert Birundu Onyancha, "Development of a Wireless Sensor Network and IoT-based Smart Irrigation System," vol. 2022. 7678570, pp. 1 - 13, 2022.

[11] Allam, T. M.; Ali, S. E.; Sallam, E. A. "Machine Learning Algorithms to Classify Water Levels for Smart Irrigation Systems." *Journal of Intelligent Systems*, vol. 6, no. 3, pp. 45 - 58, 2022.

[12] Fatma Hamouda, Angela Puig-Sirera, Lorenzo Bonzi, Damiano Remorini, Rossano Massai, Giovanni Rallo, "Design and validation of a soil moisture-based wireless sensors network for the smart irrigation of a pear orchard," vol. 305, 109138, 2024.

[13] Ojo, M.; et al. "Estimating Hydroponic Lettuce Phenotypic Parameters for Precision Agriculture," *Agricultural and Forest Meteorology*, vol. 345, article 109268, pp. 1 - 14, 2024.

[14] Xu, J. X.; Chen, J.; Li, B.; Ma, J. "Improving

Lettuce Fresh Weight Estimation Accuracy through RGB-D Fusion.” *Journal of Agricultural Engineering*, vol. 54, no. 2, pp. 98 - 110, 2023.

- [15] 진영근, “수직 농업 자율 컨텍스트 결정을 위한 온톨로지 모델링에 관한 연구,” *스마트미디어저널*, 제13권, 제6호, 72-79쪽, 2024년
- [16] 김강섭, 홍영기, 김현중, 김국환, 김정철, 이명훈, “영상처리를 이용한 토마토 생육량 측정 시스템,” *한국통신학회논문지*, 제45권, 제8호, 1460-1471쪽, 2020년
- [17] 김찬민, 김락우, 석희웅, 이선형, 허윤규, 민경원, “글로벌 스마트농업 기술·정책 동향 분석을 통한 국내 농업 발전 방향 제시,” *스마트미디어저널*, 제14권, 제3호, 68-75쪽, 2025년

저 자 소 개



김현준(정회원)

2020년 국립순천대학교 정보통신공학과 졸업
2020년~ 2023년 한국스마트팜산업협회 대리
2024년 ~ 국립순천대학교 스마트농업전공 박사과정 재학

<주관심분야 : 스마트온실 환경제어, 빅데이터 활용, 스마트 농업, IoT, 빅데이터>



이명훈(정회원)

2006년 국립순천대학교 정보통신공학과 석사 졸업
2011년 국립순천대학교 정보통신공학과 박사 졸업
2010년 ~ 2013년 한국전자통신연구원 선임연구원
2017년 ~ 2021년 국립농업과학원 연구사
2021년 ~ 현재 : 국립순천대학교 융합바이오시스템기계공학과 교수

<주관심분야 : 빅데이터 활용, 스마트 농업, 인공지능, 전파 응용 기술, 표준화 >