

中图分类号:

论文编号:

博士学位论文

基于多模态数据融合的生菜表型 估算方法研究

作者姓名 蔚栓

一级学科 电子科学与技术

二级学科 电路与系统

指导教师 赵春江/梁栋

The study of lettuce phenotype estimation method based on multi-modal data fusion

A Dissertation Submitted for the Degree of Doctor of Philosophy

Candidate: YU Shuan

Supervisor: ZHAO Chunjiang/LIANG Dong

中图分类号:

论文编号:

博 士 学 位 论 文

基于多模态数据融合的生菜表型估算方法研究

作者姓名	蔚栓	申请学位级别	工学博士
指导教师姓名	赵春江/梁栋	职 称	教授
学科专业	电路与系统	研究方向	多模态植物表型分析

学习时间自 2020 年 9 月 1 日起至 2024 年 6 月 30 日止

论文提交日期 2024 年 3 月 24 日 论文答辩日期 2024 年 5 月 19 日

摘 要

在植物表型研究中，精准估算植物表型特性具有极高的科学价值，这些特性综合体现了植物对遗传背景及环境因素的响应。传统测量方法，包括视觉观察与破坏性取样，以及直接的传感器数据采集，具有主观性强和效率低等限制。此外，单一模态数据在某些场景下可能有效，但往往无法全面揭示植物表型的复杂性。对此，多模态数据融合技术提供了一种从多尺度深入解析植物表型特性的方法。虽然在图像识别等方面机器学习展现了出色效果，但在植物表型分析领域的应用尚处于探索阶段。传统机器学习方法依赖于繁杂的特征工程，难以适应高通量植物表型分析（HTPP）的数据处理需求。在此背景下，深度学习作为一种更先进的机器学习范式，能够处理大规模复杂数据并自动提取有用模式和特征。本研究聚焦于生菜表型估算，结合高光谱成像（HSI）技术与 HTPP 平台获取多模态数据。通过深度学习技术自动提取多模态数据特征，成功估算了生菜的六个关键表型特性，涵盖品质表型（可溶性固形物含量 SSC 和 pH 值）、光合效能表型（SPAD 值）及营养元素表型（电导率 EC、钾离子 K^+ 、钠离子 Na^+ ），从而全面评估了生菜的品质、健康与营养价值，同时展示了多模态数据融合与深度学习技术的应用潜力。主要研究内容如下：

（1）提出了一种针对生菜光谱数据的自动特征提取方法，目的是解决传统植物品质表型特性估算的低效问题。通过 HSI 技术采集光谱数据，并基于 Inception 模块与全连接神经网络，设计了两种深度学习模型（SpecNet2D 和 SpecNet1D）以自动提取光谱特征，并以此估算生菜的 SSC 和 pH 值。在预测集中，SpecNet2D 和 SpecNet1D 模型分别实现了 0.9030（SSC）和 0.8490（pH 值）的决定系数（ R_p^2 ）。为深入评估深度学习模型的性能，建立了传统机器学习模型进行比较。尽管传统机器学习模型经过了细致的特征工程步骤，其估算精度（SSC: R_p^2 为 0.8587，pH 值: R_p^2 为 0.7477）仍低于深度学习模型。研究结果体现了在简化特征工程的过程中深度学习模型可以同时提高表型特性的估算精度与效率。

（2）为了进一步提升生菜品质表型特性的估算精度，将光谱与时间序列数据进行了融合。通过整合光谱数据和 HTPP 平台收集的时间序列数据，设计了一种光谱和时间序列的多输入特征融合模型（SpecTimeNet）。SpecTimeNet 结合了改进的 SpecNet2D 与循环神经网络，有效地融合了光谱特征和时间序列特征，实现了对生菜在水分胁迫条件下 SSC 和 pH 值的精确估算。在 SSC 的估算上，改进的 SpecNet2D 模型取得了

0.8743 的 R_p^2 ，而融合时间序列数据的 SpecTimeNet 模型将 R_p^2 提升至 0.8900。此外，SpecTimeNet 在检测生菜水分胁迫状态时，实现了 98.86% 的准确率。研究表明，结合多模态数据融合与深度学习的方法不仅显著提高了生菜表型特性估算的准确度，而且还揭示了其在促进抗逆性生菜品种选育方面的巨大潜力。

(3) 为了提升生菜光合效能表型特性的估算性能，将光谱与图像数据进行了融合。通过设计一种光谱和图像的多输入特征融合模型 (SpecImageNet)，SpecImageNet 结合了经过改进的 SpecNet2D 和 Big Transfer 模型，有效地融合了光谱与图像两种数据类型的特征，用于估算生菜的 SPAD 值，并成功取得了 0.8650 的 R_p^2 ，显著优于改进的 SpecNet2D 模型所达到的 0.8118 的 R_p^2 。为了深入挖掘光谱和图像数据的潜在特征，研究中还应用了类激活映射 (CAM) 技术进行针对性的光谱特征选择，并结合超分辨率生成对抗网络 (SRGAN) 以提高图像质量。CAM 和 SRGAN 的联合使用进一步提升了 SpecImageNet 模型的 R_p^2 (0.9033)。研究结果不仅体现了 CAM 和 SRGAN 在提升数据分析精度方面的卓越能力，同时也验证了将多模态数据融合于深度学习模型中，在精确、无损的植物表型特性估算领域具有重要的应用潜力。

(4) 为了提高生菜营养元素表型特性估算的准确度与效率，设计了一个基于光谱、时间序列和图像数据融合的深度学习模型 (SpecTimeImageNet)，全面地融合了三种数据模态的特征，并将其应用于估算生菜的 EC、 K^+ 和 Na^+ 。在 EC 估算中，与改进的 SpecNet2D 模型 ($R_p^2 = 0.7602$) 相比，SpecTimeNet 和 SpecImageNet 模型分别将 R_p^2 提升至 0.7806 和 0.8098，而 SpecTimeImageNet 模型进一步提升 R_p^2 至 0.8450。为了降低构建特征融合模型过程中的特征工程与模型调优的复杂度和时间成本，引入了自动机器学习 (AutoML) 技术 AutoKeras。尽管 AutoKeras 在估算精度上略低于 SpecTimeImageNet 模型 (EC、 K^+ 和 Na^+ 的 R_p^2 分别降低了 0.6%、0.7% 和 1.5%)，但其显著简化了深度学习特征融合模型的设计过程，为 HTPP 平台的数据处理提供了有力支持。研究结果展示了多模态数据融合与深度学习在提升生菜营养元素表型估算准确性中的有效性，同时突出了 AutoML 技术在简化模型设计过程中的关键作用。

关键词：多模态数据融合，高光谱，深度学习，生菜表型估算

Abstract

In plant phenotypic research, accurately estimating plant phenotypic traits is of great scientific value as these traits collectively reflect the plant response to its genetic background and environmental factors. Traditional measurement methods, including visual observation, destructive sampling, and direct sensor data collection, face limitations due to their subjectivity and low efficiency. Moreover, single-modal data might be effective in certain scenarios but often fails to fully reveal the complexity of plant phenotypes. Against this backdrop, multi-modal data fusion technology offers a method for an in-depth analysis of phenotypic traits from multiple dimensions. While machine learning has shown remarkable results in areas such as image recognition, its application in plant phenotypic analysis is still in the exploratory stage due to the complex feature engineering required by traditional machine learning methods, making them ill-suited for high-throughput plant phenotypic analysis (HTPP) data processing needs. In this context, deep learning, as an advanced machine learning paradigm, can handle large-scale complex data and automatically extract useful patterns and features. This study focuses on lettuce phenotype estimation, combining hyperspectral imaging (HSI) technology with the HTPP platform to acquire multi-modal data. By using deep learning technology to automatically extract features from multi-modal data, we have successfully estimated six key phenotypic traits of lettuce, encompassing quality phenotype (soluble solid content SSC and pH value), photosynthetic efficiency phenotype (SPAD value), and nutrient element phenotype (electrical conductivity EC, potassium K^+ , sodium Na^+), thereby comprehensively assessing the quality, health, and nutritional value of lettuce. This also demonstrates the application potential of multi-modal data fusion and deep learning technology. The main research contents are as follows:

(1) A method for automatic feature extraction from lettuce spectral data was proposed, aiming to solve the inefficiency issue of traditional plant quality phenotype characteristic estimation. Spectral data were collected using HSI technology, and two deep learning models (SpecNet2D and SpecNet1D) based on the Inception module and fully connected neural networks were designed to automatically extract spectral features for estimating SSC and pH values of lettuce. SpecNet2D and SpecNet1D achieved determination coefficients of

prediction set (R_p^2) of 0.9030 (SSC) and 0.8490 (pH value), respectively. To further assess the performance of the deep learning models, traditional machine learning models based on multivariate analysis methods were also established for comparison. Despite the detailed feature engineering steps undertaken by traditional machine learning models, their best estimation accuracy (SSC: $R_p^2 = 0.8587$, pH value: $R_p^2 = 0.7477$) still fell below that of the deep learning models. The results demonstrate the advantages of deep learning models in simplifying the feature engineering process while simultaneously enhancing the accuracy and efficiency of estimating phenotypic traits.

(2) To further enhance the estimation accuracy of lettuce quality phenotypic traits, spectral and time-series data were fused. By integrating spectral data and time-series data collected from the HTPP platform, a spectral and time-series multi-input feature fusion model (SpecTimeNet) was designed. SpecTimeNet, combining the improved SpecNet2D with a recurrent neural network, effectively fused spectral features and time-series features to achieve precise predictions of SSC and pH values under water stress conditions for lettuce. In estimating SSC, the improved SpecNet2D model achieved an R_p^2 of 0.8743, while the fusion of time-series data in SpecTimeNet elevated the R_p^2 to 0.8900. Additionally, SpecTimeNet realized an accuracy of up to 98.86% in detecting water stress conditions in lettuce. The research results demonstrate that the integration of multi-modal data fusion and deep learning methods not only significantly enhances the accuracy of estimating phenotypic traits of lettuce but also reveals its substantial potential in facilitating the breeding of stress-resistant lettuce varieties.

(3) To improve the predictive performance of lettuce photosynthetic efficiency phenotypic traits, spectral and image data were fused. A spectral and image multi-input feature fusion model (SpecImageNet) was designed, combining the improved SpecNet2D with the Big Transfer model. SpecImageNet effectively fused features from both spectral and image data to estimate the SPAD value of lettuce, achieving an R_p^2 of 0.8650, significantly better than the 0.8118 R_p^2 achieved by the improved SpecNet2D model using only spectral data. To further explore the potential features in spectral and image data, class activation mapping (CAM) technology was applied for targeted spectral feature selection, combined with super-resolution generative adversarial networks (SRGAN) to enhance image quality.

The joint application of CAM and SRGAN further elevated the R_p^2 of SpecImageNet to 0.9033. The results not only reflect the exceptional capabilities of CAM and SRGAN in enhancing data analysis accuracy but also verify the important application potential of deep learning models based on multi-modal data fusion for accurate, non-destructive estimation of plant phenotypic traits.

(4) To enhance the accuracy and efficiency of estimating lettuce nutritional phenotype characteristics, a deep learning model (SpecTimeImageNet) based on the fusion of spectral, time-series, and image data was designed. This model comprehensively integrated the features of three data modalities and was applied to estimate the EC, K^+ , and Na^+ of lettuce. In estimating EC, compared to the improved SpecNet2D model ($R_p^2 = 0.7602$), the SpecTimeNet and SpecImageNet models increased the R_p^2 to 0.7806 and 0.8098, respectively, while the SpecTimeImageNet model further improved the R_p^2 to 0.8450. To reduce the complexity and time cost associated with feature engineering and model tuning in constructing the feature fusion model, the automated machine learning (AutoML) technology AutoKeras was introduced. Although AutoKeras slightly underperformed the SpecTimeImageNet model in estimation accuracy (with R_p^2 reductions of 0.6% for EC, 0.7% for K^+ , and 1.5% for Na^+), it significantly streamlined the design process of the deep learning feature fusion model, providing strong support for data processing on the HTPP platform. The research results demonstrate the effectiveness of multi-modal data fusion and deep learning in enhancing the accuracy of lettuce nutrient element phenotype estimation and highlight the key role of AutoML technology in simplifying the model design process.

Keywords: Multi-modal data fusion, Hyperspectral, Deep learning, Lettuce phenotype estimation

目 录

第一章 绪论.....	1
1.1 研究背景与意义.....	1
1.2 国内外研究现状.....	2
1.2.1 高通量植物表型分析平台.....	3
1.2.2 表型分析方法.....	6
1.3 研究目的与研究内容.....	10
1.3.1 研究目的.....	10
1.3.2 研究内容.....	10
第二章 材料与方法	14
2.1 实验材料.....	14
2.2 高光谱图像获取与处理.....	15
2.2.1 高光谱图像获取.....	15
2.2.2 光谱反射率提取.....	17
2.3 数据分析方法.....	19
2.3.1 特征工程.....	20
2.3.2 模型建立.....	20
2.3.3 模型性能评估.....	21
2.4 本章小结.....	23
第三章 基于光谱的生菜品质表型指标的估算	24
3.1 引言.....	24
3.2 数据获取与处理.....	24
3.2.1 实验设计.....	25
3.2.2 深度学习方法.....	25
3.2.3 传统机器学习方法.....	27
3.2.4 SSC 和 pH 值的估算模型建立.....	28
3.3 结果分析与讨论.....	29
3.3.1 SSC 和 pH 值的参考值.....	29
3.3.2 光谱反射率分析.....	29
3.3.3 基于光谱的估算模型.....	31
3.4 本章小结.....	37
第四章 基于光谱和时间序列的生菜品质表型指标的估算	39
4.1 引言.....	39
4.2 数据获取与处理.....	39
4.2.1 实验设计.....	40

4.2.2 时间序列数据的获取.....	41
4.2.3 光谱和时间序列的融合方法.....	41
4.2.4 传统机器学习方法.....	44
4.2.5 SSC 和 pH 值的估算模型建立.....	44
4.3 结果分析与讨论.....	45
4.3.1 不同水分处理的 SSC 和 pH 值的参考值	45
4.3.2 不同水分处理的光谱反射率分析.....	46
4.3.3 基于光谱的估算模型.....	47
4.3.4 不同水分处理的时间序列数据分析.....	51
4.3.5 基于光谱和时间序列融合的估算模型	52
4.4 本章小结.....	57
第五章 基于光谱和图像的生菜光合效能表型指标的估算	59
5.1 引言.....	59
5.2 数据获取与处理.....	60
5.2.1 实验设计.....	60
5.2.2 RGB 图像获取与处理	61
5.2.3 高光谱图像的超分辨率重建.....	61
5.2.4 光谱和图像的融合方法.....	65
5.2.5 SPAD 值的估算模型建立.....	67
5.3 结果分析和讨论.....	67
5.3.1 SPAD 值的参考值.....	67
5.3.2 不同生菜品种的光谱反射率分析.....	68
5.3.3 基于光谱的估算模型.....	69
5.3.4 基于光谱和图像融合的估算模型.....	71
5.3.5 基于增强的光谱和图像融合的估算模型	74
5.4 本章小结.....	82
第六章 基于光谱、时间序列和图像的营养元素表型指标的估算 ...	84
6.1 引言.....	84
6.2 数据获取与处理.....	85
6.2.1 实验设计.....	85
6.2.2 多模态数据获取.....	85
6.2.3 光谱、时间序列以及图像的融合方法	86
6.2.4 AutoML 技术.....	88
6.2.5 EC、K ⁺ 和 Na ⁺ 的估算模型建立	90
6.3 结果分析和讨论.....	90
6.3.1 EC、K ⁺ 和 Na ⁺ 的参考值	90
6.3.2 基于特征级融合和定制化深度学习的估算模型	91
6.3.3 基于数据级融合和传统机器学习的估算模型	94

6.3.4 基于多模态数据融合和 AutoML 的估算模型.....	96
6.4 本章小结.....	99
第七章 总结与展望	101
7.1 总结.....	101
7.2 展望.....	102
参考文献.....	105
附录 A.....	120
附录 B.....	126
附录 C.....	132
附录 D.....	136

图目录

图 1.1 研究内容框架图.....	11
图 1.2 技术路线图.....	11
图 2.1 LQ-GHPheno 平台获取的七个品种的生菜图片.....	15
图 2.2 暗箱成像室高光谱系统 (A)、高光谱相机 (B)、灯具 (C) 和计算机采集装置 (D)	16
图 2.3 流水线成像室系统。(A) 系统外观、(B) 高光谱成像室 (红框内为高光谱相机)、(C) 多角度 RGB 成像室 (红框内为顶部 RGB 相机) 和 (D) 流水线控制系统	17
图 2.4 植物光谱提取的过程。(A) 校准的高光谱图像、(B) 二值掩模、(C) RGB 掩膜、(D) 256 个波长的 ROI 图片和 (E) 提取的植物平均反射光谱	19
图 3.1 估算生菜的 SSC 和 pH 值的数据处理流程图	24
图 3.2 田间实验现场布局 (A) 和田间生菜种植图 (B)	25
图 3.3 原始版本 Inception 模块 (A) 和 SpecNet2D (B) 的架构	26
图 3.4 FCNN 原理图 (A) 和 SpecNet1D 的架构 (B)	27
图 3.5 植物的原始平均光谱反射率以及经 MWS、FDR 和 SDR 预处理的平均光谱反射率	30
图 3.6 不同 SSC 和 pH 值的植物平均光谱反射率	31
图 3.7 基于 SpecNet2D 和 SpecNet1D 的 SSC 和 pH 值估算	32
图 3.8 基于 LWR 和 SDR 的 SSC 估算和基于 LWR 和 MWS 的 pH 值估算	36
图 4.1 估算生菜的 SSC, pH 值和水分胁迫的实验流程图	40
图 4.2 LQ-GHPheno 图片 (A) 和温室植物图片 (B)	41
图 4.3 深度学习模型 (SpecTimeNet) 的架构	42
图 4.4 SSC 和 pH 值的参考值范围	46
图 4.5 群体生菜、正常水分生菜和水分胁迫生菜的平均光谱反射率	47
图 4.6 基于 Inception 模块的品质表型估算	48
图 4.7 基于机器学习的品质表型估算	49
图 4.8 使用残差和注意力模块估算品质表型	51
图 4.9 不同水分处理的时间序列数据	52
图 4.10 使用 RNN 估算品质表型	54

图 4.11 正常水分和水分胁迫生菜的混淆矩阵。Optimal-Spec 模型的校准集 (A) 和预测集 (B) 和 Optimal-Spec-TD 模型的校准集 (C) 和预测集 (D)	55
图 4.12 pH 值的分布图	56
图 5.1 估算生菜的 SPAD 值的实验流程图	60
图 5.2 SRGAN 原理图.....	62
图 5.3 SRGAN 生成器和判别器的结构图.....	62
图 5.4 深度学习光谱图像融合模型 (SpecImageNet) 的架构	67
图 5.5 群体生菜及七个生菜品种的平均光谱反射率	69
图 5.6 基于 Inception、残差和注意力模块的 SPAD 值估算	71
图 5.7 基于初级图像特征、ResNet 和大型模型的 SPAD 值估算.....	73
图 5.8 使用 SRGAN 和插值法对原始高光谱图像上采样后的结果.....	74
图 5.9 基于原始 HSI、差值法和 SRGAN 的 SPAD 值估算结果	77
图 5.10 基于 CAM 方法的光谱特征选择结果	79
图 5.11 使用四种光谱特征选择方法合成的伪彩色图像及其相应的上采样后的图像	80
图 5.12 基于四种光谱特征选择方法的 SPAD 值估算	81
图 6.1 估算生菜的 EC、K ⁺ 和 Na ⁺ 的数据处理流程图	85
图 6.2 特征级 (A) 和数据级 (B) 融合原理图	87
图 6.3 TPOT 的数据处理流程	89
图 6.4 AutoKeras 系统流程图	89
图 6.5 EC、K ⁺ 和 Na ⁺ 的参考值范围	91
图 6.6 基于特征级融合的营养元素表型估算	92
图 6.7 基于数据级融合的营养元素表型估算	95
图 6.8 基于 TPOT 的营养元素表型估算	97
图 6.9 基于 AutoKeras 的营养元素表型估算.....	98

表目录

表 3.1 生菜的 SSC 和 pH 值的参考值测量	29
表 3.2 使用 SpecNet2D 和 SpecNet1D 估算 SSC 和 pH 值	31
表 3.3 基于机器学习方法的生菜 SSC 估算	33
表 3.4 基于机器学习方法的生菜 pH 值估算	34
表 3.5 结合 CARS 估算生菜的 SSC 和 pH 值	37
表 4.1 生菜中 SSC 和 pH 值的参考值	46
表 4.2 基于 Inception 模块估算生菜的 SSC 和 pH 值	48
表 4.3 基于机器学习估算生菜的 SSC 和 pH 值	49
表 4.4 Inception 模块与残差和注意力模块结合估算生菜的 SSC 和 pH 值	50
表 4.5 使用光谱和时间序列估算生菜的 SSC 和 pH 值	54
表 4.6 基于 Optimal-Spec 和 Optimal-Spec-TD 模型的生菜水分胁迫分类	55
表 5.1 各生菜品种的 SPAD 值的参考值测量	68
表 5.2 Inception、残差和注意力模块的结合估算生菜的 SPAD 值	70
表 5.3 初级特征提取、深度学习模型结合 RGB、HSI 相机估算生菜的 SPAD 值	72
表 5.4 基于 SRGAN 和插值法估算生菜的 SPAD 值	76
表 5.5 基于机器学习方法的光谱特征选择结果	79
表 5.6 基于四种光谱特征选择方法估算生菜的 SPAD 值	81
表 6.1 生菜中 EC、K ⁺ 和 Na ⁺ 的参考值	90
表 6.2 基于特征级融合估算生菜的 EC、K ⁺ 和 Na ⁺	92
表 6.3 基于数据级融合估算生菜的 EC、K ⁺ 和 Na ⁺	95
表 6.4 基于 AutoML 估算生菜的 EC、K ⁺ 和 Na ⁺	97

第一章 绪论

1.1 研究背景与意义

在全球化的背景下，人口的不断增长和资源的逐渐紧张，使得精准农业的发展显得尤为重要，它对于满足全球不断增长的食品需求发挥着关键作用^[1]。在这一过程中，植物表型分析成为确保作物改良和提高产量的重要技术。生菜（拉丁学名：*Lactuca sativa*），作为人们普遍喜爱的叶类蔬菜之一，因其高含水量和生食的清脆口感而广泛种植^[2]。生菜不仅营养价值丰富，富含维生素 B、C 和 E，以及钙、铁和纤维等矿物质和营养素。其低热量的特性更是使生菜成为维持健康饮食习惯的理想选择。作为菊科植物，生菜的叶片在形状和颜色上展现出多样性，颜色从深绿到浅绿不等，某些品种的叶尖甚至呈现出红色^[2]。

在种植过程中，生菜具有很强的环境适应能力，既可以在开放土地上种植，也可以利用水培等现代农业技术在室内实现全年无季节限制的生产^[3]。随着社会对健康饮食关注度的提高，生菜的消费量也在不断增长，促进了对其质量和抗逆性的科学研究。生菜的质量评估不仅根据外观层面的表型特性，如颜色，株高和叶面积等，更包括了可溶性固形物含量^[4-6]（SSC）、pH 值^[5]、SPAD 值^[7]、电导率^[8]（EC）、钾离子^[4]（K⁺）和钠离子^[9]（Na⁺）等内在生化表型指标。这些表型指标综合体现了生菜的品质、健康状况和营养价值。

SSC 是指植物体内溶解的糖分、酸、维生素、矿物质及其他固形物的总量，为果实和蔬菜的成熟度及品质提供了重要指标。高 SSC 值通常与优良的风味和营养价值相联系^[6]。pH 值反映了植物的酸碱度，并影响植物吸收营养的能力^[5]。SPAD 值是通过非破坏性地测量来评估叶绿素含量，进而反映植物的氮营养状况及其生长健康程度^[7]。EC 值作为反映植物体内离子浓度的指标，可以揭示植物体内各种离子含量的变化^[8]。K⁺对于植物的生长尤为重要，调节水分、酶活性及光合作用^[4]。Na⁺在植物中可作为微量元素吸收，尽管它在植物中通常不是必需的营养元素，但在某些情况下可以对植物生长产生积极影响，尤其是在一些对盐敏感或盐抗逆性较强的植物中^[9]。这些内在的品质表型特性之间相互也有关联^[4]，SSC 和 pH 值直接影响果实的风味品质。pH 值还影响着许多生命活动，包括代谢过程以及营养物质的吸收和运输，进而影响植物对 K⁺等离子的吸收。EC 值反映的盐分含量会影响植物对水和营养元素的吸收，特别是对于

Na^+ 的敏感性, 过高的 EC 值可能导致 Na^+ 积累, 进而影响植物对 K^+ 和其他营养元素的吸收和利用^[8]。

因此, 准确且快速地监测这些关键表型指标不仅有助于指导生菜的种植和管理, 提升其市场价值, 也能保障市场和消费者的需求。在此背景下, 植物表型精准分析尤为关键, 它是解读作物性状在基因型与环境相互作用下的变化的重要方法, 对促进作物遗传改良、提高资源利用效率、实现农业可持续发展均有着重要影响^[10-13]。然而, 传统的表型分析方法通常费时、费力且具有破坏性, 无法对大规模种植的生菜进行高精度且实时的监测, 这限制了精准农业技术的进一步发展^[14, 15]。

为应对这些挑战, 近年来, 配备多模态传感器的高通量植物表型分析 (HTPP) 平台为提高表型分析的效率与准确度开辟了新的可能性^[16]。多模态传感器的综合使用 (如高光谱成像、RGB 图像和激光雷达等), 使得研究人员能够从多个维度全面捕获植物的生理和形态特征, 为作物生长监测、病虫害识别等提供了更为丰富和准确的信息基础。此外, 深度学习算法的出现进一步加速了这一进程, 这些算法能够有效处理和分析大规模的多模态数据, 通过学习数据内在的复杂模式, 为植物表型的精准估算提供了强大的技术支持。尤其是在特征提取、模式识别等方面, 深度学习超越了传统机器学习方法的限制, 显著提高了表型分析的自动化和智能化水平^[17, 18]。

在实际应用中, 为了适应快速发展的 HTPP 平台和处理来自这些平台的海量数据, 亟需应用有效的自动化表型分析技术^[19, 20]。这不仅需要算法的支持, 更需要整合数据预处理、特征提取、模型训练等环节的自动化技术, 以实现从数据采集到表型解析的全流程自动化, 从而大幅度提高分析的效率和可靠性。因此, 探索基于多模态数据融合和深度学习在植物表型估算中的应用, 不仅可以提升表型数据处理的精度和效率, 还能为精准农业研究提供强大的技术支撑, 对促进作物改良和提高农业生产力具有深远的意义^[19, 20]。本研究旨在通过这些技术方法, 为植物表型分析提供更高效、更准确的解决方案, 为未来农业研究和实践开辟新的路径。

1.2 国内外研究现状

植物表型涵盖了一系列多样的特性, 包括生物物理特性 (如颜色、株高、叶面积和生物量)、生化特性 (如 SSC、pH 值、叶绿素含量、氮含量以及各种矿物元素) 以及生理特性 (如光合作用效率、呼吸速率和对环境胁迫的抗性能力)。这些特性共同构建了一个综合框架, 反映出植物细胞-器官-单株-群体尺度的结构和功能特点, 并且是

基因型与环境相互作用影响下的空间和时间维度上的动态表达^[10, 11]。通过深入的植物表型分析，能够揭示植物抗逆性的遗传基础，识别出表现最佳的基因组，这对作物育种和精准农业具有重要意义^[10, 13, 21]。这一过程不仅促进了研究人员对植物生长发育机制的理解，也对优化农业生产和应对全球变化挑战提供了重要的科学支撑。

传统上，植物表型分析主要依赖于人工观测和化学分析方法，这种方法简单直接，但效率低下且具有破坏性，还容易受到主观判断的影响，限制了大规模和精细的表型分析需求^[14, 15]。近十年来，HTPP 平台的兴起，辅以 RGB 相机、成像光谱、热红外成像和激光雷达等非破坏性传感器检测技术，极大地提升了植物表型数据的采集效率和精确度。这些传感器技术与 HTPP 平台的结合，实现了对作物图像、光谱和激光雷达等数据的快速连续采集，显著提升了表型数据获取的质量和效率。通过运用机器学习和深度学习分析方法分析这些数据，能够实时监测作物的品质、健康以及营养状况，为精准农业研究与作物改良提供了有力的技术支撑^[16]。

1.2.1 高通量植物表型分析平台

随着现代化和智能化农业装备的进步，研究人员越来越注重在大量植株中探索物种功能的多样性、评估不同品种的表现以及分析植株对环境变化的响应。这些研究有助于揭示表型性状间的相互关联、基因与环境因素对表型的影响等重要信息。应对这一需求，HTPP 平台得到了发展和应用^[22-24]。与传统人工表型分析相比，高通量技术能够测量更多的样本或数据点，不仅提升了单位时间内的采样量，还确保了数据处理和参数提取的同步性与效率，与硬件平台的数据扫描能力紧密匹配。鉴于高通量植物表型信息采集技术在农业科研中展现出的巨大潜力，全球各地的科研机构和企业正积极构建 HTPP 平台。这些平台按照搭载传感器的方式，可以分为暗箱、流水线、轨道、悬索以及无人机等不同类型^[25]。另外，植物表型分析平台根据使用环境的不同，可分为室内平台和室外平台。室内主要包括温室等环境，适用于对环境条件进行精细控制的研究。室外主要包括田间等自然环境，适用于在实际生长环境下研究植物的表型特性。这些平台使研究人员能够在更加多样化的条件下研究植物的表型特性，更好地理解植物对环境变化的响应^[25]。

暗箱植物表型分析平台设计用于小规模作物或植物器官以及小批量样本的精密分析。这种平台通过在受控环境下使用多种传感器，如 RGB 相机和高光谱相机等，提供了一种高度精确的方式来评估植物的特定性状，例如抗虫性。Goggin 等人就通过配备

RGB 相机和高光谱相机的暗箱表型平台来分析植物的抗虫性能^[26]。类似的, Young 等人利用搭配深度相机的暗箱表型平台来测量高粱的株高和茎宽^[27]。这种平台因其能够在受控环境中进行精确测量而受到科研人员的关注。

流水线植物表型分析平台通过在温室内部署多样化的成像装置和传感器, 以及将植物盆栽放置在与牵引装置相连的承载组件上, 或者直接安置在传送带上, 形成了一个高效率的表型数据收集流程^[22, 25]。随着植物盆栽沿着传送带进行旋转或往复运动, 它们将会精确地移动至配备有成像设备的暗箱或其他传感器的位置, 以进行详尽的表型信息采集。数据收集完成后, 植物继续沿传送带移动, 形成一个闭环的流程。这一设计不仅最大化了数据采集的效率和连续性, 而且通过精确控制植物样本的移动, 确保了数据的一致性和可重复性, 为深入研究植物生长发育和表型变异提供了有力支持。此平台已应用于多项研究, 如评估大米的籽粒投影等 15 个表型性状^[28], 以及对小麦的非生物胁迫检测^[29]。

轨道表型平台通过建立固定的轨道系统, 并使用电动机驱动的传感器系统来对植物进行综合监测, 提供了一种高效的移动平台。该平台设计允许在各种环境条件下进行全天候的测量活动, 且其稳定的机构设计确保了机械振动不会对成像质量产生不利影响, 使其成为获取高清晰度表型信息的理想选择。轨道式平台的灵活性使其既可以在温室内部署, 也适用于户外田间环境。此类平台已经在多项研究中显示出其价值, 例如用于追踪小麦在其整个生命周期中的发育情况^[30], 以及研究珍珠小米的叶面积生长曲线和蒸腾作用的响应机制^[31]。

悬索植物表型平台通过在特定的田间地块上固定, 依靠自然环境条件下采集的表型数据进行分析, 提供了高可靠性的数据来源。这种平台已经成功应用于评估玉米和大豆的叶绿素含量和水分含量^[23], 同时也用于提取冬小麦的冠层覆盖度和冠层高度等关键表型指标^[32]。

用于表型分析的无人机平台涵盖了多旋翼^[33]、固定翼^[34]、直升机^[35]、飞翼^[36]、降落伞^[37]、滑翔机^[38]以及飞艇^[39]等多种类型。这些平台搭载的传感器是实现低空遥感技术数据采集的关键。传感器分为两大类: 辅助传感器和表型传感器。辅助传感器确保无人机能够安全且高效地执行飞行任务, 其中惯性测量单元^[40]用于检测无人机机体的运动, 保持传感器在平衡位置, 而气压式高度表则优化了飞行高度的测量, 广泛应用于飞行控制系统中。表型传感器的主要功能是采集作物的表型信息。此类无人机平台已广泛应用于多个研究领域, 包括检测植物叶片的病害^[41]以及测量植物株高、生物量、

叶面积指数和植物密度等关键指标^[10, 11]。

专门为温室环境设计的 HTPP 平台因其能够在受控环境下执行日常测量作业而受到了特别的重视^[28]。尤其值得强调的是，一些先进的 HTPP 平台还整合了智能植物表型分析软件。这些软件能够自动处理传感器收集到的图像、光谱和点云数据，从而准确地提取出植物颜色、形态、植被指数等关键表型信息^[22, 28]。这项技术的发展不仅使得表型数据采集过程更为快速，而且极大地增加了研究人员对植物表型多样性的理解。

在配置植物表型分析平台的传感器时，通常需综合考虑传感器的光谱分辨率、空间分辨率、成本等关键因素以及应用需求^[10]。常见于这些平台的传感器主要包括 RGB 相机、多光谱相机、高光谱成像（HSI）相机、热红外相机、叶绿素荧光成像以及激光雷达等。这些传感器各有其独特的功能和优势，为植物表型分析提供多角度和多尺度的测量能力。

RGB 相机捕集红色（R）、绿色（G）、蓝色（B）三个颜色通道的图像。其图像数据易于处理，且结果直观易懂。相对低廉的价格、广泛的可用性和简单的操作使得其成为最通用的视觉传感器。RGB 相机已经成功用于检测大麦的株高、鲜生物量和干生物量^[42]。多光谱相机能捕捉超出 RGB 相机范围的特定波段，如近红外（NIR）波段。这使其能获取与作物健康状况密切相关的信息，比如用于计算植被指数、监测作物生长和病虫害。多光谱相机已用于检测葡萄叶片的病害^[43]。HSI 相机捕捉连续的、窄波段的光谱信息，通常覆盖从可见光到近红外甚至中红外的范围。它提供的详细物质特征光谱能揭示作物的生化和生理信息，如叶绿素含量、水分和营养状况等^[44]。HSI 相机已用于检测番茄叶片的病害^[45]。热红外相机捕捉物体发出的热辐射，用于测量温度分布。它能监测植物的水分胁迫状况，因为受胁迫的植物蒸腾作用受限，导致叶片温度升高。热红外相机已成功用于检测柑橘的水分胁迫情况^[46]和棉花的冠层温度^[47]。叶绿素荧光成像是一种通过测量叶绿素分子在脉冲光照射后释放的荧光强度来获取信息的技术。这种荧光测量通常在植物叶片受到特定光脉冲照射后进行，其中部分激发能量以荧光形式释放。叶绿素荧光的强度能提供关于光合作用中光系统 II（PSII）的活性和效率的重要信息。荧光参数可以反映植物对环境胁迫（如干旱、高温或盐胁迫）的响应，这些胁迫会影响光合作用并进一步改变荧光发射特性。通过测量叶绿素荧光，能评估非光化学淬灭和光化学淬灭的相对比例，帮助了解植物是如何管理吸收的光能，是用于光合作用还是通过其他途径耗散。叶绿素荧光成像技术已用于评估农药对拟南芥代谢的干扰^[48]，展现了其在揭示植物生理状态方面的重要应用价值。激光雷达通过

发射激光脉冲并捕捉其反射光来测量目标与传感器之间的距离，能生成高分辨率的三维点云数据，用于精确测量作物的几何结构和体积。激光雷达已用于草地的地上生物量估算^[49]和葡萄园中葡萄的分布情况^[50]。

1.2.2 表型分析方法

在上述的传感器技术中，特别是成像光谱学技术，包括多光谱相机和 HSI 相机，因为它们能同时提供植物光谱和图像数据，受到了特别的关注^[10, 11]。多光谱成像虽在捕获特定光谱波段信息方面极有价值，但其不连续的波段设定限制了全面光谱信息的采集。相比之下，高光谱成像技术突破了这一局限，提供连续且细密的光谱特征信息，涵盖从可见光至近红外乃至中红外的宽广光谱区域。这一能力为深入分析植物生理状态提供了有力的数据支持，因而在植物表型分析的应用中越发普及，赋予研究者获取精准、细节丰富的数据的能力^[10, 45]。获得这些高质量的传感器数据之后，如何有效地挖掘和应用这些丰富的植物表型数据便成为了研究的关键。利用高光谱成像技术提供的丰富光谱细节，机器学习在解读这类复杂数据集中展现出了巨大的潜能^[10, 11]。机器学习可以根据模型的结构和学习方式大致分为传统的机器学习和深度学习模型两大类^[10, 11]。这些数据分析技术不仅能从数据中提取深层次的、有价值的信息，还有助于揭示植物表型-基因型-环境条件之间的复杂交互作用。

Weng 等人采用 HSI 结合机器学习模型如偏最小二乘回归 (PLSR)、支持向量回归 (SVR) 和局部加权回归 (LWR)，成功估算草莓果实的新鲜度和维生素 C 含量^[51, 52]。类似地，机器学习技术结合 HSI 应用于评估土豆和黑麦的水分胁迫状况^[53, 54]。在对生菜表型的深入分析中，Sun 等人利用 PLSR、多元线性回归 (MLR)、人工神经网络 (ANN) 和自适应提升法精确估算了生菜的水分含量和氮素水平^[55, 56]。此外，Lara 和 Osco 等人通过结合主成分分析和 ANN，探测了生菜的盐分和水分胁迫^[57, 58]，而 Mo 等人采用单因素方差分析在线评估生菜品质^[59]。由于原始 HSI 中的光谱数据存在噪声以及相邻波长特征之间的高度相关性^[51, 52]，处理 HSI 的光谱数据时机器学习模型通常需执行诸如预处理和光谱特征选择等特征工程步骤。例如，Zhou 等人在使用 PLSR 模型估算生菜水分含量时，采用了小波变换 (WT) 对光谱数据进行预处理，提升了模型的估算性能^[60]。Saputro 等人运用竞争性自适应重加权采样 (CARS) 方法挑选出光谱数据中的关键特征，并发现光谱特征选择明显优化了 PLSR 模型估算香蕉 SSC 的精度^[61]。同样，多元散射校正 (MSC)、标准正态变量变换 (SNV)、移动窗口平滑 (MWS)、

Savitzky-Golay 平滑 (SG)、一阶导数 (FDR) 和二阶导数 (SDR) 等光谱预处理技术^[62], 以及连续投影算法 (SPA)、无关信息变量消除 (UVE) 和遗传算法 (GA) 等光谱特征选择技术^[63], 常用于高光谱数据处理。虽然这些方法在不同植物光谱数据处理中广泛应用, 但是特征工程的结果并不总是一致, 这意味着在传统机器学习模型中达成优异的估算性能可能需要大量时间去寻找最合适的特征工程策略, 限制了这些模型在 HTPP 领域的快速应用。

近期研究展现了深度学习算法特别是卷积神经网络 (CNN) 和全连接神经网络 (FCNN) 在处理光谱数据上的显著优势。这些技术通过自动从原始数据中提取特征, 显著提高了多种光谱分析任务的性能, 开拓了植物表型精准和高效分析的新途径^[17, 18, 64, 65]。例如, Xiao 等人成功地将高光谱数据与深度迁移学习结合, 准确评估了棉花叶片的叶绿素含量^[66], 而 Zhou 和 Sun 等人通过堆叠卷积自编码器和深度信念网络等深度学习模型, 对生菜中的镉和铅含量进行了深入研究^[67-70]。CNN 因其优秀的特征提取能力而受到特别重视。随着 Inception 模块的引入, CNN 在处理高光谱数据方面的能力得到了进一步强化, 该模块通过多尺度核学习优化了特征捕获的细致度^[71]。Rehman 等人通过改进 Inception 模块, 提高了对玉米叶片相对水分含量 (RWC) 估算的准确性^[72]。然而, 网络的扩展和深入带来了重复特征提取和梯度消失的挑战, 残差和注意力模块的引入有效缓解了这些问题, 提升了网络的学习效率和性能^[73], 如 Weng 等人在提升小麦赤霉病估算准确性方面的应用^[74]。随着对深度学习模型可解释性的关注^[75, 76], 类激活映射 (CAM) 技术随之提出以增强模型的可解释性。CAM 能够直观显示网络在做出决策时关注的区域, 帮助理解模型如何识别和处理数据中的关键特征^[75-77], 并已成功应用于遥感图像的目标识别^[78]。

成像光谱技术不仅提供了详细的光谱信息, 还带来了图像数据, 但是其相对较低的空间分辨率限制了数据的深入利用。在这方面, 超分辨率重建 (SR) 技术, 尤其是基于生成对抗网络 (GAN) 的超分辨率生成对抗网络^[79-83] (SRGAN) 以及传统的图像插值方法^[84-89], 提供了有效的解决途径。SRGAN 已经在提高遥感领域多光谱图像分辨率上展现出其强大的能力^[90-92]。正如光谱数据能够揭示植物生理状态信息一样, 图像数据提供了植物的外在表型信息, 包括形态特性、发育状况及病害表现等^[93-95]。例如, Serouart 等人运用 U-net 对 RGB 图像中的绿色植被与衰老植被进行分割^[93], 而 Zheng 等人利用 Mask R-CNN、ResNet 和 VGG16 来分析草莓冠层并估算其生物量^[94]。Okoyere 等人则提出了一种 RGB 图像自动分割方法, 适用于温室和田间植物的识别^[95]。同时,

随着大型计算机视觉模型的发展，需要挖掘这些模型处理大规模图像数据集和学习复杂视觉模式的潜能^[96-98]，尽管它们在农业领域的研究和应用还相对有限，主要由于农业图像数据的特殊性和大型模型所需的大量标注数据及计算资源的难以获取。所以需要深入研究大型模型在农业领域的适应性和效率，以及探寻更有效地利用这些技术进行精准农业分析和决策支持的方法。在分析时间序列数据方面，循环神经网络（RNN）通常比 CNN 和 FCNN 表现得更优异^[99]。时间序列数据同样在植物表型特性的理解上提供了独特的视角。它记录了植物生长及其对环境变化的动态反应，揭示了植物的生长速率、表型变化的模式以及环境适应策略等关键信息^[100-106]。例如，Chen 等人结合双向循环神经网络（BRNN）与气候和大气环流信息有效检测了棉花病虫害^[100]，BRNN 也成功用于监测木薯的糖化和发酵过程^[101]。此外，利用 CNN 和长短期记忆网络（LSTM）不仅估算了温室作物的产量^[102]，LSTM 还模拟了树木的个体生长^[103]。Khaki 等人通过 LSTM 估算了玉米和大豆的产量^[104]，而 RNN 成功用于区分干旱和半干旱草原的归一化植被指数^[105]（NDVI）。这些研究中向 RNN 输入的大多数时间序列数据都来自外部环境。同样值得关注的是，植物的生物物理表型（如颜色和形态）与生化表型（如内部品质）以及生理表型（如水分胁迫）之间存在着紧密的联系^[106]。将 RNN 应用于 HTPP 平台收集的庞大时间序列数据，特别是那些涉及生物物理表型的数据，用于估算植物的生化和生理表型，这一领域仍有待进一步探究。

上述研究围绕单一数据模态展开，包括光谱数据、图像数据和时间序列数据等，已经在各自领域取得了显著成果。然而，尽管单模态数据分析在表型特性估算方面已显示出其有效性，其固有的局限性仍可能阻碍了对植物表型特性更深层次的理解。因此，为了克服这些局限并实现对表型信息的深入挖掘，引入多模态数据融合技术成为了一条新的途径，它通过整合来自不同数据模态的互补信息，大幅增强了表型分析的综合性和准确性^[107-117]。多模态数据融合的关键在于结合不同数据源的优势。例如，光谱数据能够提供丰富的化学成分信息，图像数据能够捕捉到植物的形态和结构特征，而时间序列数据则能够反映植物的动态变化过程。通过将这些数据融合在一起，可以实现对植物表型特性的全面、精准估算。数据级融合^[107]与特征级融合^[78]作为两种主要的融合方法，各自展现了其在推动研究进展方面的独特优势。例如，Li 等人通过结合激光雷达数据和 RGB 数据，实现了对玉米株高的精准估算^[107]。Mishra 等人通过融合高光谱数据的光谱信息与图像纹理信息，成功实现了对茶叶的分级^[108]。Li 等人将多光谱数据和激光雷达数据结合，估算了小麦的产量^[109]。同时，Shu 和 Chang 等人利用无

人机获取的多光谱与 RGB 图像,评估了玉米的冠层覆盖度、株高和生物量等表型特性及番茄的产量^[110, 111]。Zhang 等人通过高光谱与叶绿素荧光数据的融合,识别了水稻的水分胁迫状态^[112],而 Li 等人使用无人机搭载的高光谱与 RGB 图像来估测马铃薯的地上生物量和产量^[113]。多模态数据融合不仅仅是数据的简单叠加,更重要的是其在信息互补和数据协同方面的潜力。融合技术可以有效利用不同模态数据的优势,弥补单一模态数据的不足,从而提升模型的鲁棒性和泛化能力。具体而言,数据级融合能够在早期阶段结合不同模态的数据,提供更为全面的输入信息;而特征级融合则通过在模型内部结合不同模态的特征信息,能够更好地捕捉到数据之间的复杂关联。针对生菜表型分析的数据融合应用主要围绕利用高光谱相机、多光谱相机、RGB 相机以及叶绿素荧光成像等,进行生菜的水分胁迫、重金属胁迫和新鲜度的估算^[114-117]。尽管如此,利用多模态数据融合技术来探索生菜的品质、光合效能以及营养元素等生化表型的研究还相对较少,显示出这一领域仍有广阔的研究空间。多模态数据融合技术的应用,将有助于更全面地理解植物的生理状态和生长环境,为精准农业和植物育种提供更加可靠的科学依据。另外,考虑到不同植物和表型特性具有独特的生理和生化背景,这就需要设计针对性的估算模型,并组合不同模态的特征以适应这些差异。虽然基本的多模态机器学习技术(如融合光谱和图像数据的方法)可以跨多种植物应用,但模型的特定配置、训练数据集的构成以及最终的估算目标可能需要根据特定的应用场景进行大幅调整。此外,植物表型估算研究时所依赖的数据集大小极为不等,范围从几十到数千个样本不等,这进一步复杂了模型的训练过程和泛化能力的评估。因此,准确地调整和优化这些模型以适应不同数据集的规模和特性是实现高效、可靠估算的关键。随着 HTPP 平台数据量的持续增长和表型分析需求变得日益复杂,建立性能卓越的植物表型估算模型变得更加困难。处理单一模态数据就需要耗费大量时间,而来自 HTPP 平台的多模态数据处理则更加复杂且耗时^[10-12]。在这种背景下,快速确定最优模型及其参数配置成为了新的挑战。自动机器学习(AutoML)技术可以作为解决此问题的一种有效方法^[118, 119],AutoML 技术在自动化模型选择和参数优化过程,为大数据时代的表型分析提供一种高效、智能化的解决策略。AutoML 技术分为两类:一类是针对传统机器学习模型优化的自动化技术^[120, 121],另一类是专注于深度学习网络优化和参数调整的自动化技术^[122, 123]。

总而言之,植物表型分析的技术和研究方法正处于快速发展期。从最初的传统人工观测,发展到运用 HTPP 平台,然后是机器学习和深度学习的广泛应用。特别是分

析方法已经从依赖单一模态数据，比如仅限于光谱或 RGB 图像的分析，扩展到了多模态数据的综合利用，表型分析技术正在向着自动化、效率化和精准化的方向发展。随着技术的持续进步和创新，可以预期植物表型分析在精准农业、作物改良和生态研究等多个领域将发挥重要作用。

1.3 研究目的与研究内容

1.3.1 研究目的

本研究致力于克服传统单一数据模态在植物表型特性评估中的限制，通过深度学习技术综合光谱、图像和时间序列等多模态数据，实现了对生菜表型特性的快速、精确和无损估算。这一方法不仅极大提高了生菜表型估算的准确率和效率，而且加深了对生菜生长与环境因素相互作用的理解，为生菜作物遗传改良、栽培管理和环境应对提供了科技支撑。此项研究的成果展示了多模态数据融合在生菜品质、光合效能和营养元素表型精确估算方面的应用价值。同时，该研究提供的视角有助于推动精准农业技术进步，支持农业生产朝着可持续方向发展，为实际生产实践提供了技术支持。

1.3.2 研究内容

本研究通过结合深度学习和多模态数据融合方法来估算生菜中的品质表型（SSC 和 pH 值）、光合效能表型（SPAD 值）以及营养元素表型（EC、 K^+ 和 Na^+ ）。首先设计了一个基于光谱数据分析 SSC 和 pH 值的基线深度学习网络架构，确定了深度学习模型在自动提取光谱特征以及提升植物表型估算精度的有效性。在此基础上，研究分别探索了两条数据融合路径：一是通过光谱数据与时间序列数据的融合，提升了单一光谱模态对生菜在水分胁迫条件下 SSC 和 pH 值的估算精度；二是结合光谱数据与图像数据，提升了单一模态光谱模型估算 SPAD 值的精度，并利用 CAM 和 SRGAN 选择光谱特征和提升图像质量，进一步增强了深度学习融合模型的估算能力。最后，通过整合光谱数据、时间序列数据和图像数据来估算 EC、 K^+ 和 Na^+ 含量，旨在提升表型特性估算的准确度，并探究不同数据模态的融合对估算性能的影响。研究内容框架图如图 1.1 所示，技术路线图见图 1.2。

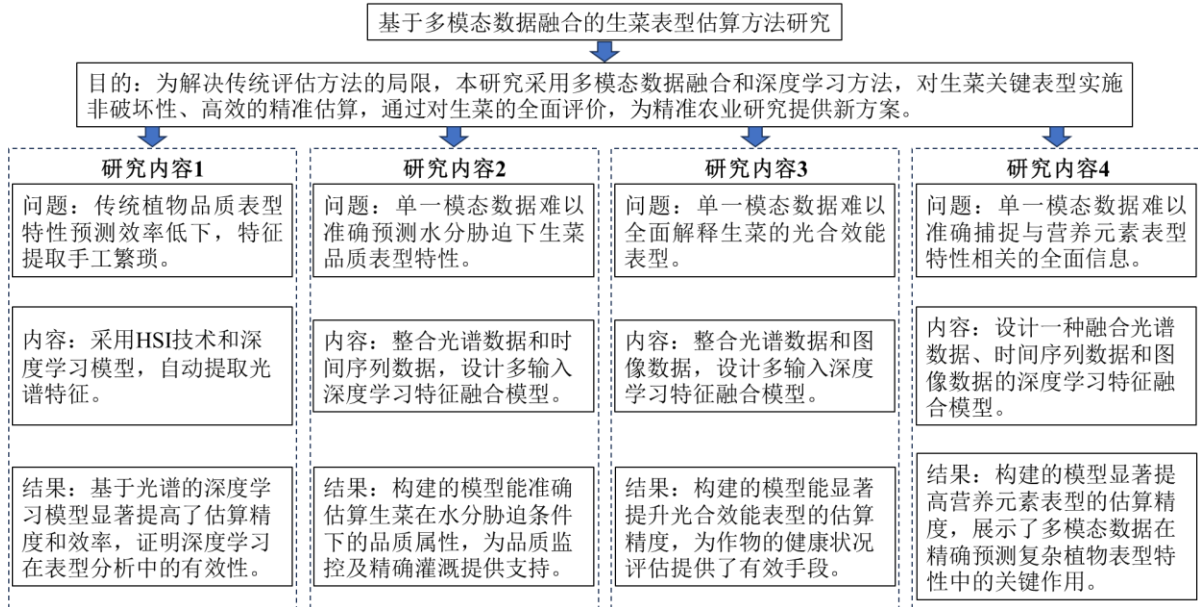


图 1.1 研究内容框架图

Figure 1.1 The framework of research content

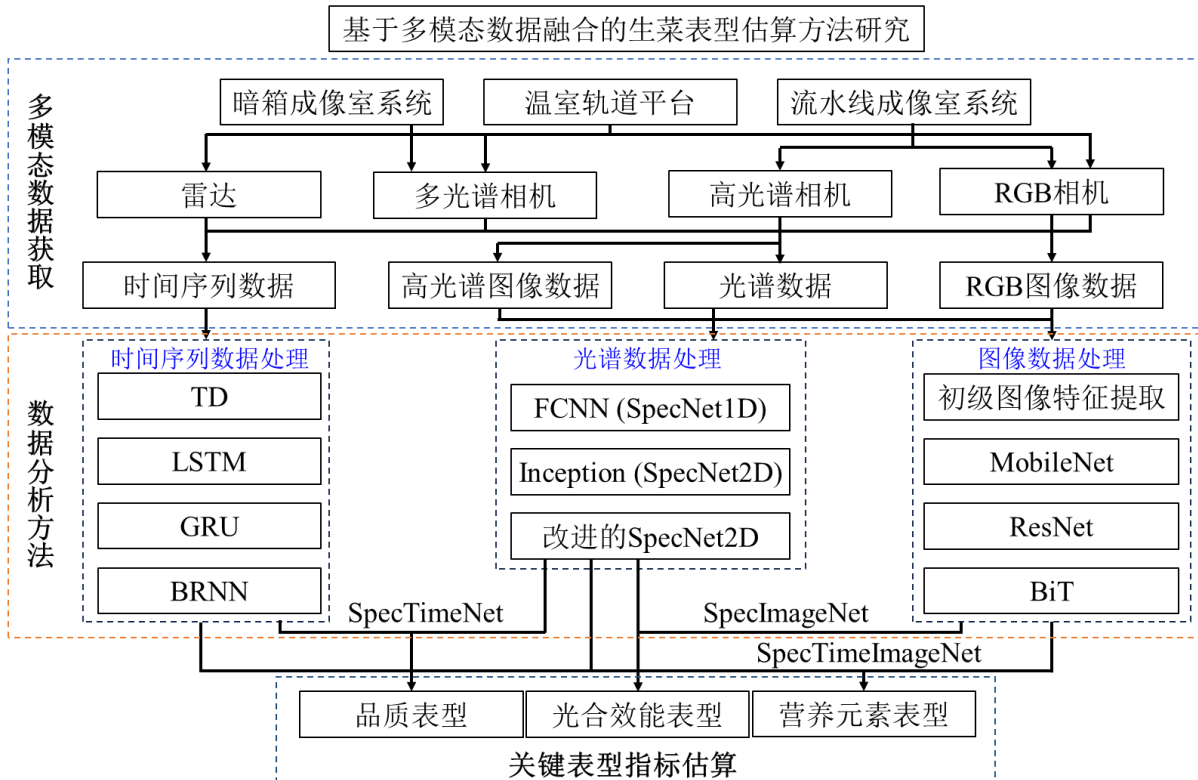


图 1.2 技术路线图

Figure 1.2 The framework of research content

具体研究内容包括：

(1) 首先对生菜进行了高光谱成像，以获取其光谱数据。在此基础上，设计了两

种深度学习模型：一种是结合了二维 CNN（2DCNN）和 Inception 模块的 SpecNet2D 模型，其特点在于能够在多个尺度上捕捉光谱数据的特征；另一种是基于 FCNN 的 SpecNet1D 模型，其特点在于模型结构精简。这两个模型均为端到端设计，能够直接从生菜的原始光谱数据中自动提取关键特征，并用于估算生菜的 SSC 和 pH 值。为了评估提出的深度学习模型的效能，构建了基于两种线性多元回归方法（PLSR 和 MLR）以及三种非线性多元回归方法（LWR、ANN 和 SVR）的机器学习模型，对生菜的 SSC 和 pH 值进行了分析。鉴于机器学习模型在处理高光谱数据中的噪声以及识别与 SSC 和 pH 值相关的光谱特征时可能遇到的挑战，在机器学习模型的建模过程中引入了特征工程技术。这包括应用五种光谱预处理技术（MWS、SG、FDR、SDR 和 WT）以降低原始光谱数据中的噪声，并采用 CARS 方法来筛选关键光谱特征。通过与传统机器学习模型的比较，观察所提出的深度学习模型在特征提取方面的性能表现。

（2）在上述工作的基础上，研究进一步扩展了深度学习模型架构，将残差和注意力模块集成到 SpecNet2D 模型中，以评估不同模块组合对 SpecNet2D 模型在光谱数据特征提取方面的性能影响，筛选出最优的模块组合来改进 SpecNet2D 模型，以提高生菜在水分胁迫条件下 SSC 和 pH 值的估算精度。为了进一步提升模型对生菜品质表型的估算能力，引入来自 HTPP 平台的时间序列数据，以增强模型对生菜生长动态的解读能力。这是基于生菜外部生长特征与内部品质表型之间存在紧密联系的理解。通过融合时间序列数据到深度学习架构中，并采用时间分布式（Time-Distributed, TD）、LSTM、门控循环单元（GRU）和 BRNN 四种 RNN 技术，探究了光谱与时间序列数据融合对模型性能的影响及各种 RNN 技术对模型性能的具体贡献，以确定最佳的时间序列特征提取方法。此外，使用性能最优的深度学习模型对生菜水分胁迫状态进行监测，并通过像素级光谱技术和深度学习模型可视化生菜 pH 值的空间分布，进而探讨了使用生菜 pH 值的空间分布图来判定水分胁迫早期发生情况的实际应用潜力。

（3）在上述的研究基础上，本阶段确定了最佳模块组合改进的 SpecNet2D 模型，用于处理光谱数据并估算生菜的 SPAD 值。鉴于 SPAD 值与植物颜色等外观特征紧密关联，HSI 相机和 RGB 相机提供的图像数据可以成为光谱数据的关键补充。因此，分别对 HSI 和 RGB 提供的图像数据进行处理，采用初级特征提取技术、小型模型（MobileNet 和 ResNet）及大型模型（BiT），以确定最有效的图像数据类型和处理策略。面对高光谱图像数据在空间分辨率上的局限，采用 SRGAN 和传统的插值方法来提高图像的分辨率，深入探讨了这些技术对于 SpecImage 模型估算性能的影响。此外，探

究了 SpecImage 模型的可解释性, 采用 CAM 方法、CARS 和 UVE 从光谱数据中筛选出三个关键光谱特征。基于这三个特征对应的高光谱图像数据, 合成了伪彩色图像。采用最优图像处理策略对这些伪彩色图像进行特征提取, 以考察这些关键特征对于提高 SpecImage 模型性能的影响。

(4) 在前三项工作的研究基础上, 确定了光谱数据、时间序列数据和图像数据的最优处理策略, 构建了 SpecTimeImage 模型, 该模型综合这三种数据模态的特征以估算生菜的 EC、 K^+ 和 Na^+ 。这些营养元素的表型特性不仅可以从光谱数据中的吸收或反射特性中得到体现, 也可以从图像数据所揭示的外部视觉信息以及时间序列数据所捕捉的植物生长过程中的动态变化中获得反映。因此探究了不同数据模态组合对营养元素表型估算性能的具体影响, 以评估各模态数据对模型的贡献度, 并确定最佳的数据模态组合。考虑到深度学习模型通常需要较长的训练和调试周期, 为了快速验证不同模态数据融合对营养元素表型估算的效用, 并拓展多模态数据融合技术在更广泛的植物表型分析领域的适用性, 引入了计算资源需求较低的传统机器学习模型作为辅助分析方法。这些辅助的机器学习模型 (FusionML) 采取数据级融合方法, 可以快速探索多模态数据融合潜力。为提升植物表型估算效率, 评估了 AutoML 技术在植物表型分析中的应用潜力。通过使用 TPOT 生成机器学习模型和 AutoKeras 构建深度学习模型, 实现了模型选择和参数优化的自动化。将 AutoML 技术、SpecTimeImage 模型和 FusionML 模型进行比较, 探寻一条高效且可靠的技术路径, 用于精确估算作物关键表型指标。

本文通过深入探索深度学习与多模态数据融合技术, 为生菜表型特性的精确估算提供了新方法。从 HSI 获取的单一模态的光谱数据, 到利用 HTPP 平台收集的时间序列数据以及 RGB 和 HSI 相机提供的图像数据, 每一步都是为了丰富数据信息, 进而为实现生菜的品质、光合效能和营养元素表型特性的快速、精确、无损估算提供有力支持。

第二章 材料与方法

2.1 实验材料

本研究所使用的实验材料主要包括田间种植的生菜和温室种植的生菜两种类型。田间生菜在北京市农林科学院露天环境进行种植，具体位置位于北纬 39.9438°，东经 116.2876°。为确保实验的准确性和高效性，现场部署了一套配备高光谱相机的暗箱植物表型分析平台（暗箱成像室系统）。植物生长在专门设计的垄上，每个垄间设有排水沟渠，以保持土壤湿度的适宜和均匀。温室生菜则在北京市农林科学院设施温室内种植。温室内环境条件控制在温度 18 至 25°C、相对湿度 75%至 95%，配备了轨道 HTPP 平台（LQ-GHPheno, NERCITA, 中国）、与田间相同配置的暗箱成像室系统，以及集成了高光谱相机和 RGB 相机的流水线 HTPP 平台（Line-GHPheno, NERCITA, 中国），也称作流水线成像室系统。这些平台能够高效地获取生菜的生长数据，包括形态结构、生长速率和生物量等重要表型指标。此外，各品种生菜都按照统一的标准进行管理，以确保实验的可重复性和数据的可比性。

在两种不同的环境条件下，共种植了七个不同的生菜品种，涵盖了生菜的主要品种和类型。具体包括奶油生菜、结球生菜、散叶生菜、橡叶生菜、罗马生菜、茎用生菜以及野生近缘种生菜，如图 2.1 所示。奶油生菜叶片松散、质地柔软，结球生菜则以其紧密的头部和脆嫩的叶片而著称。散叶生菜具有多汁的叶片，而橡叶生菜则以其独特的波纹边叶片和丰富的色彩而受到欢迎。罗马生菜拥有较为整齐的长椭圆形叶片排列，而茎用生菜则因其肉质的茎部和用于食用的叶柄而被重视。此外，野生近缘种则为基因多样性和表型特性的研究提供了宝贵的参考^[2]。这些品种的选择旨在涵盖遗传多样性的广谱，为跨不同生菜类型的表型特性提供全面评估。



图 2.1 LQ-GHPheno 平台获取的七个品种的生菜图片

Figure 2.1 Lettuce pictures of seven cultivars acquired by the LQ-GHPheno platform

2.2 高光谱图像获取与处理

2.2.1 高光谱图像获取

生菜植株的可见光至近红外（Vis-NIR）高光谱图像是在暗箱成像室和流水线成像室两个平台上以反射模式获取的。

暗箱成像室长 1.5 m，宽 1.5 m，高 2.5 m，如图 2.2 所示。暗箱中的高光谱成像系统由以下几个模块组成：（1）高光谱仪（GaiaField-V10E，双利合谱，中国），（2）CCD 相机，（3）四个钨卤素灯，以及（4）装有图像采集软件（Specview，双利合谱，中国）的计算机（Ins 15-7510-R1645S，Dell，美国）。高光谱仪是内置的推扫式线扫描设备，具有 400-1000 nm 的光谱范围，光谱分辨率设置为 2.8 nm，共 256 个波长。CCD 相机的空间分辨率为每次扫描线 696 像素，并配备了 23 mm 镜头。其中，高光谱仪和 CCD 相机集成于一体，被称为高光谱相机。灯具提供 350-2500 nm 波长的光，功率为 50 W。生菜冠层与镜头之间的距离设定为 1 m，灯具与相机之间的角度设定为 45°，以提供足够的光线到成像区域进行图像采集。高光谱图像（维度：775×696×256）获取的仪器设置参数：曝光时间为 30 ms，帧率为 14 FPS，增益为 1，空间合并为 2，光谱合并为 4。

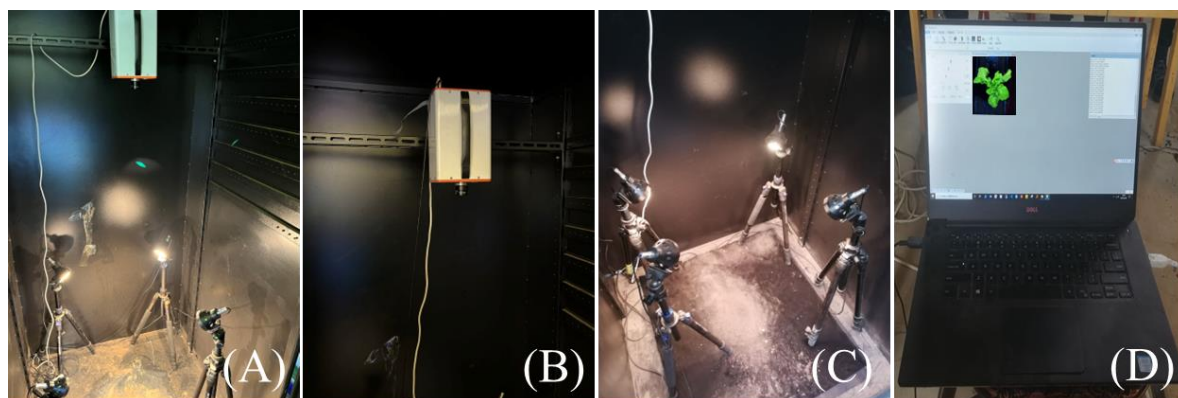


图 2.2 暗箱成像室高光谱系统 (A)、高光谱相机 (B)、灯具 (C) 和计算机采集装置 (D)

Figure 2.2 Hyperspectral system in a darkroom imaging chamber (A), hyperspectral camera (B), lamps (C), and computer acquisition device (D)

流水线成像室系统 (LQ-LinePheno, NERCITA, 中国) 作为一种 HTPP 平台, 如图 2.3 所示, 具备在 2 分钟内同步获取单株植物多角度 RGB 图像与高光谱图像数据的能力。此系统利用多角度 RGB 图像数据, 配合专业软件 (Context Capture 18, Bentley Systems, 法国) 进行三维重建, 有效生成植物的点云数据。为保证数据处理的一致性, 专门选用位于植物冠层正上方的顶部相机所拍摄的 RGB 图像, 以确保其拍摄角度与高光谱图像相匹配。流水线成像室中的高光谱成像系统模块与暗箱成像室中的配置相同, 使用的高光谱相机参数设置也完全一致, 唯一不同之处在于生菜冠层与相机镜头之间的设定距离为 1.8 m。

LQ-LinePheno 系统支持多种作物 (包括辣椒、茄子、番茄、生菜、草莓和玉米等) 的自动化高通量数据采集与解析, 同时也支持蔬菜单株尺度的三维建模与可视化, 为品种展示、株型评价、品种鉴定、结构功能分析、表型组学研究以及数字育种等多方面的应用提供了高精度与自动化的数据采集工具。该系统重建的蔬菜植株点云模型精度高达每株不少于 10 万点, 点云密度小于 0.05 cm, 实现了亚厘米级的高精度测量。平台能够自动解析包括株高、叶面积、分层投影面积、凸包体积、植株正侧向面积、容积率、茎秆基部直径及植株紧凑度等多维形态形状指标, 以及颜色性状指标, 如各颜色通道的均值及方差。此外, 该平台还可选配搭载多光谱传感器, 自动化地获取蔬菜光谱表型数据, 其高效的数据采集能力显著增强了大规模样本表型数据采集的支持能力。

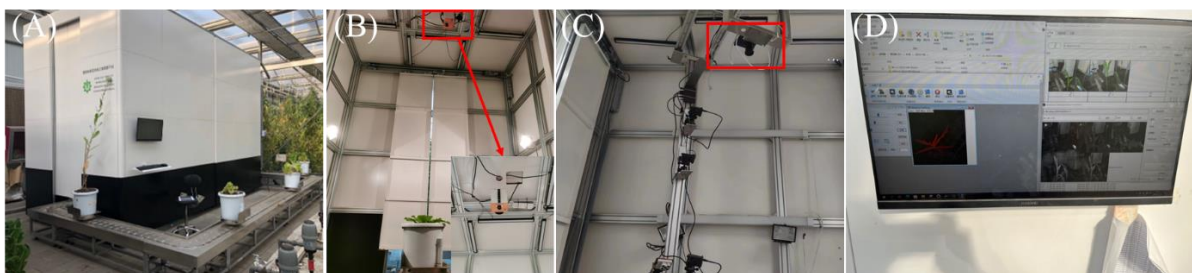


图 2.3 流水线成像室系统。(A) 系统外观、(B) 高光谱成像室（红框内为高光谱相机）、(C) 多角度 RGB 成像室（红框内为顶部 RGB 相机）和 (D) 流水线控制系统

Figure 2.3 Conveyor belt imaging chamber system. (A) external view of the system, (B) hyperspectral imaging chamber (the hyperspectral camera is within the red frame), (C) multi-angle RGB imaging chamber (the top RGB camera is within the red frame), and (D) conveyor control system

2.2.2 光谱反射率提取

由于高光谱系统的成像模式设定为推扫式，且非平面的相机镜头与光谱仪是分离的，所以获取的原始高光谱图像存在畸变^[124]。因此，使用 Specview 软件提供的镜头校准功能对原始高光谱图像进行了镜头校准。高光谱图像中的每一个像素的原始记录值为数字量（Digital Number，DN 值）。DN 值本身是传感器对接收到的光信号量化后的结果，它受到传感器自身特性、成像条件、目标物体反射和辐射特性的影响。虽然 DN 值可以反映相对的光照强度和物质的光谱特性，但由于受到诸多不确定因素的影响，如仪器噪声如相机暗电流、光照条件如不均匀的光分布和观测角度等，DN 值无法直接用于表征目标物体的精确光谱特性。为了从 DN 值中提取出有用的科学信息，需要进行校准处理，将其转换为更具有物理意义的反射率值。目的是为了剔除影响量化值的非目标因素，从而得到真实反映目标物体光谱特性的信息^[124]。反射率校准的具体操作：扫描一块 99% 反射率的聚氯乙烯（PVC）板作为白色参考值图像来校准图像中的光照变化，然后，通过关闭灯光并覆盖相机镜头收集暗电流参考值图像，以消除高光谱相机的暗电流。采用公式（2.1）进行反射率校准：

$$R = \frac{R_{\text{raw}} - R_d}{R_w - R_d} \quad (2.1)$$

在此公式中， R 、 R_{raw} 、 R_w 和 R_d 分别代表校准后的图像、原始高光谱图像、从 PVC 板获得的白色参考值图像以及关闭光源和覆盖镜头情况下获得的暗电流参考值图像。

植株的高度差异导致其不同部位的光谱反射率和亮度存在显著变化，这会给感兴趣区域（ROI）的精确提取带来挑战。传统的阈值分割方法通过设定一个特定的数值来区分目标和背景，但该方法在处理高度差异导致的光谱和亮度不均匀时往往效果不佳。具体来说，阈值分割法在以下几个方面存在缺点：对于变化复杂的背景，如不同高度和不同角度的植株叶片，单一的阈值很难适应所有情况，导致分割结果不准确；对光照条件敏感，光照不均匀时容易产生误分割；很大程度上依赖于经验，缺乏自适应性，当图像特征变化时，手动调整阈值非常耗时；无法有效区分重叠或相邻的目标，容易造成目标间的错分或漏分。相对于阈值分割方法，光谱角制图（SAM）具有独特优势，特别适合于高光谱图像中 ROI 的精确提取。SAM 基于光谱角度的概念，通过比较像素光谱与参考光谱之间的角度，从而区分不同材质，对不均匀光照和复杂背景更加鲁棒。它允许对 ROI 的光谱特征进行更加精细和动态的匹配，使得 ROI 提取不受单一阈值限制。SAM 可以自动适应不同植物部位的光谱差异，减少人为选择阈值的需求，提高了分割效率和准确性。由于 SAM 考虑了更加全面的光谱信息，因此它能够更好地处理植物冠层中的立体结构和重叠叶片问题^[125-127]。所以，SAM 将用于从反射率校准后的高光谱图像中精确提取整个生菜冠层的 ROI。通过 ENVI 5.3 软件（ITT Visual Information Solutions，美国）从生菜的顶部点、顶部区域、中部点、中部区域、底部点和底部区域提取的六种光谱作为生菜冠层的波谱库，如图 2.4 所示。通过波谱库来获得生菜冠层的掩膜，将此掩膜与反射率校准后的图像相乘，获取了 ROI 图像。通过对 ROI 图像中所有像素的光谱求平均，获得平均光谱，用于进一步的建模分析。

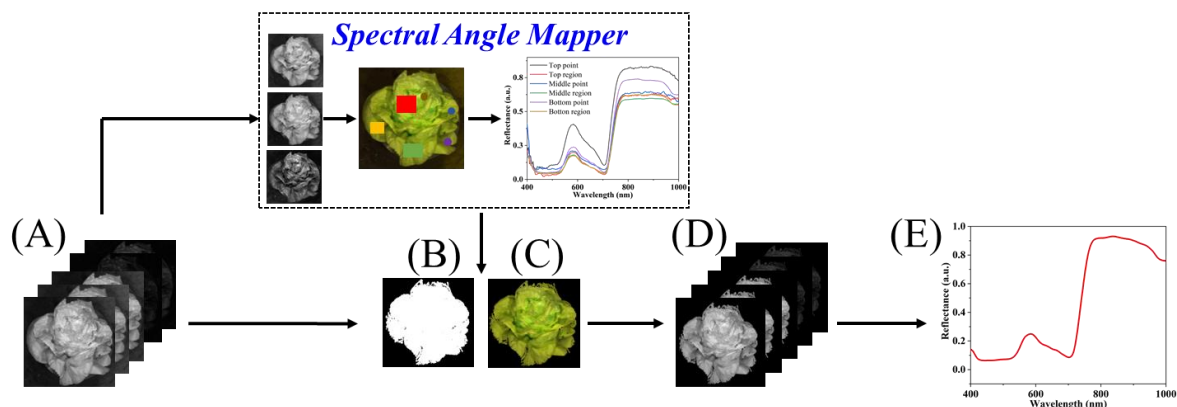


图 2.4 植物光谱提取的过程。(A) 校准的高光谱图像、(B) 二值掩模、(C) RGB 掩模、(D) 256 个波长的 ROI 图片和 (E) 提取的植物平均反射光谱

Figure 2.4 The process of plant spectra extraction. (A) calibrated hyperspectral images, (B) binary plant mask, (C) RGB plant mask, (D) 256 wavelengths corresponding to the ROI, and (E) extracted plant mean reflectance spectrum

2.3 数据分析方法

在本研究中，多模态数据是从三种表型平台收集而来，涵盖了光谱、时间序列以及图像三种数据模态。具体来说，暗箱成像室提供了 HSI 数据，LQ-GHPheno 平台提供了时间序列数据，而流水线成像室则同时提供了高光谱数据和 RGB 图像数据，其中高光谱数据包含了丰富的光谱及图像信息。在植物表型分析中，选择光谱数据、时间序列数据，以及图像数据作为研究的三种主要数据模态，是为了全面捕获和理解植物的生长状态及其对环境变化的响应。这三种数据模态各自揭示了植物表型的不同方面：光谱数据提供了关于植物生理状态的详细信息，包括叶绿素含量、水分状态和营养状况等，这些特征反映了植物的健康状况和生长潜力^[4]。光谱数据通过分析植物对光的吸收和反射特性，能够在分子水平上揭示植物表型的细微变化^[10, 11]。时间序列数据记录了植物生长和发育过程中的动态变化，以及植物对环境条件变化的适应性。这种数据模态有助于观察植物生长速率、表型变化规律，以及环境应激响应等长期趋势和周期性变化，从而为理解植物的生长动态提供了时间维度的视角^[100-105]。图像数据，包括 RGB 和高光谱图像，能够直观展现植物的外观特征，如形态结构、颜色变化和病害症状等^[93, 108]。这些外在表型特征反映了植物的生长状况和健康水平，同时图像数据的空间分辨率高，能够提供植物表型的精细观察^[90-92]。

在收集多模态数据后，研究工作集中于设计高效的数据处理和分析方法，以便精

确估算植物的关键表型指标。处理方法涉及机器学习和深度学习两个主要方向。机器学习方法的步骤包括数据预处理、光谱特征的选择、模型选择和参数优化；深度学习方法则侧重于设计端到端的模型架构和调整模型参数。数据处理的核心目标是在庞杂的原始数据中筛选出对生菜表型特性估算起到重要作用的信息，进而提升估算精度和效率。通过这些精细化处理步骤，不仅能够优化数据分析流程，提升分析速度和结果的质量，同时也为精准农业的进步贡献力量。

2.3.1 特征工程

特征工程，尤其是在处理高维度的光谱数据时，对于机器学习方法来说，是建立有效估算模型的关键步骤。原始光谱数据通常含有由背景干扰和光散射引起的噪声成分，这些噪声会对模型的估算性能产生负面影响。为了有效削弱这些影响，特征工程阶段采用了一系列光谱预处理技术。预处理技术的选择是为了从不同角度解决数据噪声问题，提高模型的估算准确性。例如，MWS 和 SG 技术主要用于平滑数据，而 FDR 和 SDR 技术则用于增强光谱曲线的特征峰，WT 能够在多尺度下分析光谱信号，揭示更细微的特征。这些预处理技术的应用，能够有效改善光谱数据质量，为后续建模提供更准确的输入特征。在光谱特征选择方面，由于光谱数据的高维度特性，直接建模可能会导致维度灾难或模型过拟合。因此，采用了 CARS 和 UVE 等特征选择方法来识别和保留最具代表性和估算能力的光谱波长，这些方法通过统计分析和机器学习算法确定对目标变量估算最为关键的特征，简化了模型并提升了估算效率和准确性。

对于深度学习模型，尽管其具有从大量数据中自动提取复杂特征的能力，但结合特征选择方法仍能进一步提升模型性能。特别是，引入 CAM 不仅加强了模型对光谱数据中重要特征的识别能力，也增加了模型的可解释性，使研究人员能够更清晰地理解模型的决策过程。这种方法可以通过精确控制和选择输入特征，实现对植物表型特性的高效精确估算。

2.3.2 模型建立

在建模阶段，针对不同类型的数据，综合运用了机器学习和深度学习技术。机器学习模型，如 PLSR、MLR、LWR、ANN 和 SVR，用于分析经过预处理和特征选择的光谱数据、时间序列数据和图像数据。而对于深度学习模型，使用 FCNN 和 Inception 模块处理光谱数据^[65, 71]，RNN 处理时间序列数据^[99]，这些模型在自动学习原始数据中的复杂特征表达方面，显示出极佳的潜力。在图像数据处理方面，不仅采用了颜色直

方图、灰度共生矩阵和 Hu 矩来手动提取颜色、纹理和形状特征等图像初级特征^[128, 129], 还部署了预训练的深度学习模型, 如 ResNet^[94]和 Big Transfer^[98] (BiT) 等, 以实现高级抽象特征的自动提取。

在数据融合的实验, 探索了数据级融合与特征级融合两种策略, 目的是提升模型对植物表型特性的整体理解和估算能力。数据级融合, 主要应用于机器学习模型中, 涉及将光谱数据、时间序列数据和图像数据在输入层面进行整合, 融合后的数据直接用于训练机器学习模型, 从而估算表型特性^[107]。而特征级融合, 特别适用于深度学习模型, 它允许各模态数据先独立进入专为其设计的深度网络架构中进行特征提取和优化, 再将这些独立提取的特征在模型的后阶段进行综合融合, 形成一个全面的特征集以进行最终估算^[78]。特征级融合的策略充分发挥了深度学习模型自动提取复杂特征的能力, 通过模型不同层次上的特征提取和融合, 实现了更为细致的数据处理。这种策略不仅加深了模型对每种数据模态的理解, 还提升了模型在处理复杂和高维度数据时的估算准确率和解释力。因此, 特征级融合适用于处理多模态数据的深度学习模型, 尤其适用于需要深入理解不同数据模态之间复杂交互的场景, 从而在精准农业领域的表型分析中提供有力技术支持。

综合而言, 深度学习模型能够直接利用大量的原始数据自动学习特征, 而机器学习模型在某些情况下则更依赖于人为设计的特征工程。通过充分发挥这两种模型的各自优势, 不仅能够精确估算生菜的表型特性, 还能深入探究不同模态数据对估算性能的具体影响。通过这样的多元化方法, 模型不仅在表型估算性能上有所提升, 同时也能在特征的解释性和生物学意义上提供更多见解。

2.3.3 模型性能评估

为了建立生菜表型估算模型, 样本数据集以 2:1 的比例分为校准集和预测集。在校准集中, 进一步随机选取了 10%至 20%的数据, 作为验证集以用于模型参数的优化。具体选取的数据比例将根据模型在调优过程中的表现来确定。三个数据集中, 校准集主要用于模型的训练; 而验证集则用于调整和优化模型参数, 以提升模型性能; 预测集的作用则是评价最终建立模型的估算性能。

所有建立的回归模型的性能均使用校准集的决定系数和均方根误差 (R_c^2 , RMSEC)、验证集的决定系数和均方根误差 (R_v^2 , RMSEV)、预测集的决定系数和均方根误差 (R_p^2 , RMSEP) 以及预测集的相对百分比差异 (RPD) 来评估。 R^2 、RMSE

和 RPD 的计算公式如下：

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_r - y_p)^2}{\sum_{i=1}^n (y_r - \bar{y})^2} \quad (2.2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_r - y_p)^2}{n}} \quad (2.3)$$

$$RPD = \frac{SD}{RMSE} \quad (2.4)$$

y_p 、 y_r 和 $\bar{y}$ 分别是表型（如 SSC、pH 值和 SPAD 值等）的估算值、参考值和参考值的平均值， n 是样本数量，SD 是参考值的标准差。RMSEC、RMSEV 和 RMSEP 的值越小，表示模型的估算准确度越高，因为这些指标代表了估算值与实际值之间的误差。 R_c^2 、 R_v^2 、 R_p^2 和 RPD 的值越大，表示模型的拟合度和估算能力越强^[113]。其中，RPD 值在模型性能评估中扮演着关键角色，它直观地反映了模型估算准确度的高低。RPD 值低于 1.5 意味着模型具有较弱的估算能力，其估算结果的可靠性较低；当 RPD 值处于 1.5 至 2 之间，表明模型具备一定的估算能力，尽管存在一定程度的不确定性；RPD 值在 2 至 2.5 之间，模型展现出较好的估算性能，能够提供相对准确的估算结果；而 RPD 值大于或等于 2.5 时，则说明模型具有极强的估算能力，估算结果的精确度和可靠性极高^[113]。因此，RPD 作为衡量模型估算能力的重要指标，对于评价和选择高性能的估算模型具有重要的参考价值，尤其在需要高精度估算的应用场景中显得尤为重要。

分类模型的评价指标为混淆矩阵（Confusion Matrix）和准确率（Accuracy）。混淆矩阵提供了一种直观的方式来展示模型估算结果与实际标签之间的对应关系，尤其适用于分析二分类问题的模型性能。混淆矩阵由以下四个基本元素构成：真正例（True Positives, TP）：模型正确估算为正类的样本数；假正例（False Positives, FP）：模型错误估算为正类的样本数；真负例（True Negatives, TN）：模型正确估算为负类的样本数；假负例（False Negatives, FN）：模型错误估算为负类的样本数。准确率则通过公式（2.5）计算得出，它反映了模型正确估算的样本占总样本数的比例，是评估模型整体性能的一个直观指标。

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2.5)$$

准确率作为一个快速衡量分类模型优劣的指标，易于理解。混淆矩阵不仅能帮助研究人员直观理解模型在不同类别上的估算能力，还为计算其他关键性能指标（如敏感度和特异性）提供了基础^[74]。

2.4 本章小结

本章详细介绍了实验材料与数据分析方法，说明了在不同条件下生长的多个生菜品种的样本准备过程，以及通过植物表型平台获得的高光谱图像数据的收集方法，并阐述了高光谱数据进行镜头校准和反射率校准的必要性和具体操作过程。在光谱反射率提取方面，解释了传统阈值分割方法的局限性，以及采用 SAM 进行 ROI 提取的优势。通过专业软件 ENVI 的使用，实现了高精度的波谱库生成和 ROI 图像获取，为后续的建模分析提供了基础。在数据处理方法中，讨论了特征工程、数据融合方法以及建模方法。在数据预处理环节，强调了 MWS、SG、FDR、SDR 和 WT 等光谱预处理技术的重要性。在特征选择中，介绍了不同的特征选择技术，包括基于机器学习的 CARS 和 UVE，以及基于深度学习的 CAM，用于提取对模型估算影响最显著的光谱特征。在建模方面，展示了如何使用机器学习和深度学习模型来处理光谱数据、时间序列数据和图像数据。最后，对模型性能评估方法进行了描述，使用 R^2 、RMSE 和 RPD 等统计指标来评估回归模型的性能，以及 Confusion Matrix 和 Accuracy 在分类模型评估中的应用。

第三章 基于光谱的生菜品质表型指标的估算

3.1 引言

HSI 技术与深度学习模型的结合为非破坏性地监测作物生长状态、病害诊断及品质评估提供了新的可能性，尤其在广泛消费的叶菜作物如生菜的品质表型特性估算方面展现了其应用潜力^[55, 56]。本章将专注于探讨 HSI 技术与深度学习模型在生菜表型分析中的具体运用，重点关注光谱数据的处理与分析。目的是探索深度学习在光谱特征提取方面的能力，并为接下来的多模态数据分析提供理论与实验基础。

研究将围绕深度学习构建以及其性能评价展开。首先，设计两种端到端深度学习模型，用于快速且非破坏性地估算生菜的表型特性。其次，验证所提出的深度学习模型能够直接利用原始反射光谱数据作为输入，准确估算生菜的 SSC 和 pH 值。此外，通过比较不同的线性与非线性多元分析方法构建的机器学习模型与深度学习模型之间的性能差异，选择适合应用于 HTPP 平台的模型。

3.2 数据获取与处理

图 3.1 展示了本章的数据处理流程图。在获取并处理高光谱图像数据后，这些数据将输入到相应的模型中，用于估算 SSC 和 pH 值。具体的数据处理方法将在本节的后文部分进行详细阐述。

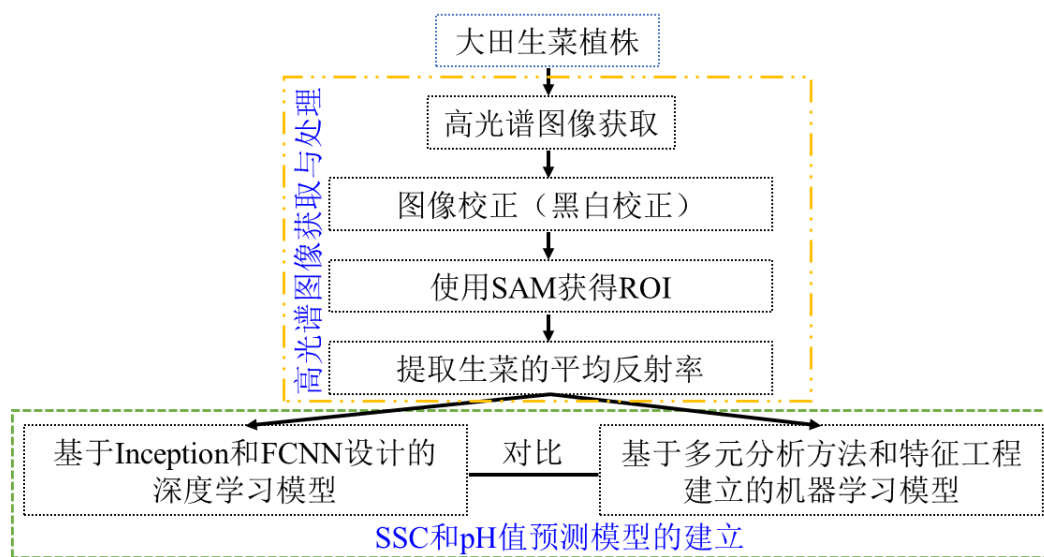


图 3.1 估算生菜的 SSC 和 pH 值的数据处理流程图

Figure 3.1 Flow diagram of data processing used to estimate SSC and pH value in lettuce

3.2.1 实验设计

生菜于 2021 年 4 月 12 日在北京市农林科学院露天种植。其中，三个生菜品种（奶油生菜、散叶生菜和罗马生菜）用于后续实验。实验现场布局 and 生菜材料如图 3.2 所示。植物的生长环境为：（1）土壤未施加肥料；（2）滴灌在专业人员监督下进行；（3）所有植物的水分处理相同。2021 年 5 月 15 日、20 日和 25 日，每个品种挑选了 45 株生菜放入花盆中，然后送至田边实验室进行高光谱成像。成像后，按以下步骤获取生菜汁液：（1）去除生菜根部以获取叶片；（2）使用手摇榨汁机（LKM-ZZ01，力可美科技有限公司，中国）提取汁液；（3）将汁液储存在离心管中（容量：50 ml）。由于实验过程中 18 个生菜样本的标签丢失，最终分析了 387 份生菜汁液样本（丢失样本见附录 A 表 A1）。分别使用数字折光仪（PAL-1，ATAGO，日本）和 pH 计（206pH1，Testo，德国）测量生菜汁液的 SSC 和 pH 值（测量仪器实物图见附录 A 图 A1）。每个样本的 SSC 和 pH 值测量三次，最后取平均值作为该样本 SSC 和 pH 值的参考值。

本实验采用暗箱成像室系统拍摄生菜的高光谱图像。高光谱图像具体的获取方法见 2.2.1 节，光谱反射率提取过程见 2.2.2 节。



图 3.2 田间实验现场布局 (A) 和田间生菜种植图 (B)

Figure 3.2 Field experiment site layout (A) and field lettuce planting map (B)

3.2.2 深度学习方法

在 HTPP 中，深度学习模型展示了其卓越的数据挖掘能力^[72]。这些模型已成功应用于众多领域，例如，它可用于图像分割^[93-95]以及多种表型特性的量化或定性分析^[65-70]。然而，将深度学习模型与光谱数据相结合用于量化生菜冠层品质表型的研究还相对少见。在本章中，针对生菜的 SSC 和 pH 值估算，设计了基于 Inception 和 FCNN 两种网络的深度学习模型，以评估不同网络模型在植物表型特性估算准确性方面的表现。

CNN 以其仿生物视觉感知机制而闻名，它能够识别和处理像素和音频等网状结构数据^[132]。CNN 的一大优势在于其隐藏层的卷积核参数共享和层间连接的稀疏特性，这不仅减少了计算量，还使得模型对额外的特征工程需求降至最低。在早期尝试中，注意到仅通过线性叠加卷积层构成的模型性能不尽人意，这可能是由于未能充分挖掘原始光谱数据中蕴含的信息。为此，本章在原始版本的 Inception 模块的基础上引入了经过优化的 Inception 模块（如图 3.3A-B 所示），通过这种结构增强了模型的特征提取和学习能力。SpecNet2D 模型中的 CNN 和 FCNN 层分别采用了 ELU 和线性激活函数，选择 Adam 作为优化器，并以 RMSE 和平均绝对误差（MAE）分别作为损失函数和模型在训练时的性能评估指标。Inception 模块的设计主要基于以下几点：多尺度卷积核的使用能捕捉不同尺度的特征信息；通过合并这些多尺度特征，模型能够综合不同层次的信息；同时，合理平衡网络的深度和宽度，避免模型过于复杂导致的饱和。如图 3.3B 所展示的，通过结合三种规模的特征，经合并操作后进行特征整合，得到了丰富的特征表示。由于合并后的特征维度较高，引入了全连接层模块来降维，这一策略显著提高了 SpecNet2D 模型的稳健性。

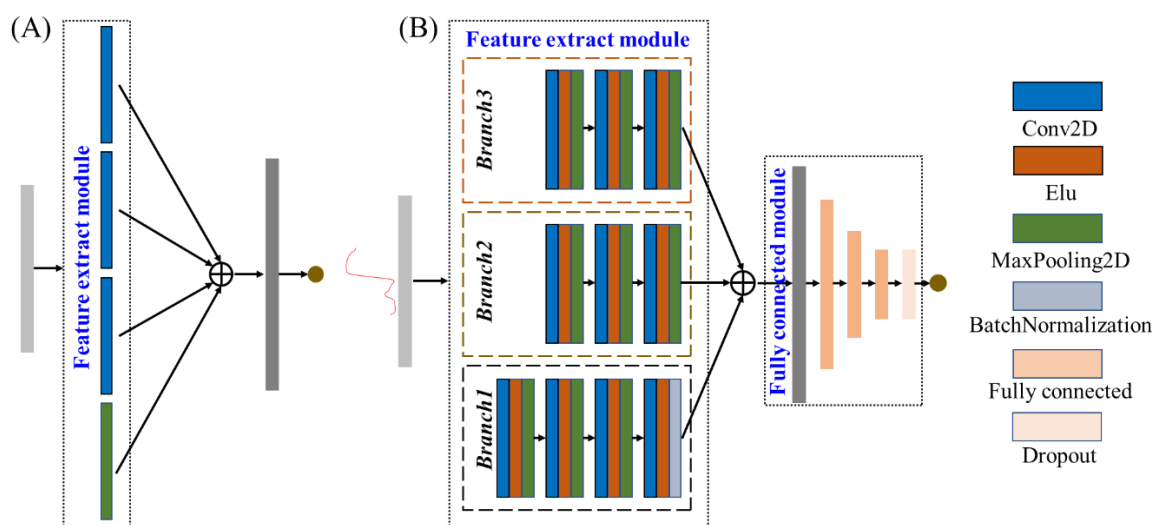


图 3.3 原始版本 Inception 模块 (A) 和 SpecNet2D (B) 的架构

Figure 3.3 Architecture of the naive version Inception module (A) and SpecNet2D (B)

FCNN 是一种进阶版的单层感知器^[65, 131]，它能在数据的不同类别间找到划分的超平面。与经典感知器的单层结构不同，FCNN 在每一层都进行层内节点与层间节点的全面连接（如图 3.4A 所示）。虽然增加多层结构有利于提取更细致的数据特征，但同时也带来了训练时间延长和过拟合的风险。为解决这一问题，引入了随机失活策略（Dropout），避免过拟合现象。如图 3.4B 所揭示的 SpecNet1D 模型结构，在数据进入

模型后的底部三层添加了 Dropout，并在隐藏层使用了线性激活函数。SpecNet1D 模型的优化器、损失函数和评价指标的设置与 SpecNet2D 保持一致。

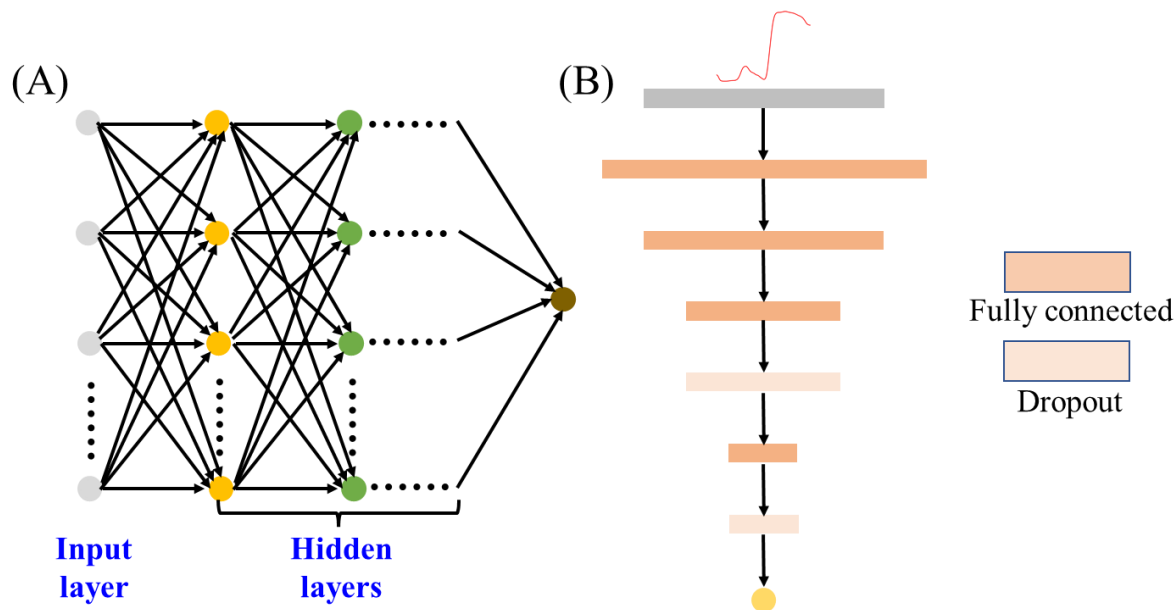


图 3.4 FCNN 原理图 (A) 和 SpecNet1D 的架构 (B)

Figure 3.4 FCNN schematic chart (A) and architecture of the SpecNet1D (B)

3.2.3 传统机器学习方法

(1) 建模方法

为了与设计的深度学习模型进行对比，本章使用了两种线性多元分析方法（PLSR、和 MLR），以及三种非线性多元分析方法（LWR、ANN 和 SVR）来建立机器学习估算模型。

PLSR 是一种处理多响应变量 Y 与多解释变量 X 间关系的建模技术。在模型建立过程中，PLSR 侧重于从响应矩阵 Y 和解释矩阵 X 中提取关键的主成分，并寻求这些成分之间的最大相关性，从而实现对复杂数据结构的简化和有效估算。在此场合， X 和 Y 分别代表光谱数据和所需估算的参数（比如 SSC 或 pH 值）。MLR 方法致力于揭示多个估算变量与响应变量之间的线性组合关系。在现实问题分析中，由于研究人员常常需要处理多维特征空间，MLR 能较好地适应这种情况，通过为每个特征赋予权重，实现对目标变量的综合估算。

LWR 则是一种非参数回归方法，它依赖于本地信息进行回归分析。具体来说，LWR 将数据集划分成多个小区间，对每个区间内的样本进行多项式回归拟合，并赋予不同的权重，通过反复迭代这一过程，在各个小区间内获取加权回归线，最终通过连

接这些局部回归线的中心点，建立起一个覆盖整个数据范围的完整回归模型。ANN 模仿了人脑神经元的信息处理机制，建立了一系列简化的模型，其中的每个神经元节点通过不同的连接方式组网。每个节点代表一个输出函数，亦即激活函数，而节点之间的连接权重则决定了信息传递的强度。SVR 是支持向量机的一个关键应用分支，主要应用于回归问题的处理。SVR 通过发现一个最优超平面，以最小化数据点相对于该平面的偏差总和，并依据此超平面对数据集进行整体性的拟合。这种方法在处理非线性和高维度问题时，表现出了强大的泛化能力。

（2）光谱预处理与特征选择方法

在从高光谱图像中提取反射光谱时，光照不均匀性和背景干扰会引入噪声，影响光谱数据的质量。此外，高维光谱中存在与 SSC 和 pH 值无关的光谱特征，以及相邻波长的光谱特征之间高度相关的问题，这些因素都可能降低通过多元分析方法建立的机器学习模型的准确性。然而，通过光谱预处理方法，例如 MWS、SG、FDR、SDR 和 WT，以及光谱特征选择方法如 CARS，可以削弱甚至消除这些问题，从而提高模型的性能。

MWS 技术采用特定宽度的平滑窗口逐点移动于光谱曲线之上，通过对窗口内的光谱值进行平均处理，有效滤除了光谱噪声，增强了有效数据的信号强度。这种技术通过平滑处理，在减少随机噪声的同时，得以保留光谱曲线的基本形态。SG 通过在一定时间窗口内，对光谱数据进行局部多项式拟合并进行最小二乘运算，实现了在保持光谱特性（如峰位和宽度）不变的前提下，有效去除噪声。这是一种将数学精细度与实验数据平滑处理相结合的技术。导数光谱，如 FDR 和 SDR 通过对光谱信号进行求导，有助于剥离背景信号的干扰，强化了有用信号，提高了光谱数据的信噪比，有利于捕捉信号中更为细微的变化。CARS 通过回归系数的绝对值评价各个波长的光谱特征对估算模型的贡献大小，有效地筛选出与目标估算变量关联最密切波长的光谱特征，以实现光谱信息的最大化提取和冗余信息的最小化。

3.2.4 SSC 和 pH 值的估算模型建立

采用 Kennard–Stone (KS) 算法^[133]将 387 个生菜的反射光谱数据分为校准集和预测集，比例为 2:1。在校准集中，随机选择了 15% 的光谱数据作为验证集，校准集、验证集和预测集的具体样本数量分别为 219、39 和 129。在通过多元分析方法建立的机器学习模型中，数据集在预处理或光谱特征选择后进行划分。模型性能评价指标及其描

述见 2.3.3 节。

SpecNet2D 和 SpecNet1D 使用 Python 3.7.10 和 TensorFlow 2.4.1 环境中实现的。多元分析方法、光谱预处理和光谱特征选择过程是在 MATLAB2020 (Mathworks, 美国) 软件中进行的。所有计算实验是在 DELL OptiPlex 7080 (Dell, 美国) 计算机上进行的, 该计算机配置有 2.90 GHz 的英特尔酷睿 i7-10700 处理器 (CPU), 32 GB 随机存取内存 (RAM) 和英伟达 Quadro P2200 图形处理单元 (GPU), 操作系统为 Windows 10 家庭版 20H2。

3.3 结果分析与讨论

3.3.1 SSC 和 pH 值的参考值

生菜的 SSC 和 pH 值数据列于表 3.1 中。SSC 的测量范围介于 0.8750%至 5.8250%, 平均值为 3.1040%, 标准差为 0.7528。标准差较大, 说明样本的 SSC 值分布相对较散。pH 值的测量范围介于 6.3125 至 6.8175, 平均值为 6.5945, 标准差为 0.0860。相比 SSC, pH 值的变化范围较小, 标准差较低, 这表明生菜样本的 pH 值相对比较集中, 波动小。

表 3.1 生菜的 SSC 和 pH 值的参考值测量

Table 3.1 Reference measurement of SSC and pH value in lettuces

Phenotypes	Range	Mean	SD
SSC (%)	0.8750-5.8250	3.1040	0.7528
pH value	6.3125-6.8175	6.5945	0.0860

3.3.2 光谱反射率分析

图 3.5 展示了从高光谱图像中获得的生菜在 400 到 1000 nm 波长范围内的平均光谱反射率。植物的光谱反射率特征能够提供其生物物理、生理和化学特性的信息^[72]。例如, 约 580 nm 处的波段与叶黄素有关^[57, 66]。710-760 nm (红边) 波段和 700 nm 附近的波段与叶绿素有关^[57, 66]。在近红外区域的 980 nm 附近的波段与水的 O-H 键有关^[134]。所有经过 MWS、FDR 和 SDR 预处理的生菜植物的高光谱图像的光谱反射率在图 3.5 中展示。所有植物在四个波段区域 (400-435 nm、515-650 nm、690-780 nm 和 960-1000 nm) 显示出明显的差异。这里, MWS 的反射值与原始反射值之间没有显著差异, 这可能是因为原始光谱本身就相对平滑, 因此平滑处理并不会导致肉眼可见的差异。与原始反射值相比, FDR 和 SDR 的反射值变化明显增强, 尤其是 SDR 变化。之前的研究也发现了类似的结果^[4], 导数强调了数据的变化率, 使得轻微的变化 (波峰和波谷) 更加明显。在光谱分析中, 它可以突出光谱曲线的变化细节, 即光谱特征的局部变化,

从而间接地提高光谱的分辨率，使得邻近的吸收峰更容易区分。在复杂光谱中，不同物质的吸收峰可能会重叠，导数处理可以帮助分离这些重叠峰，以便于识别和量化。

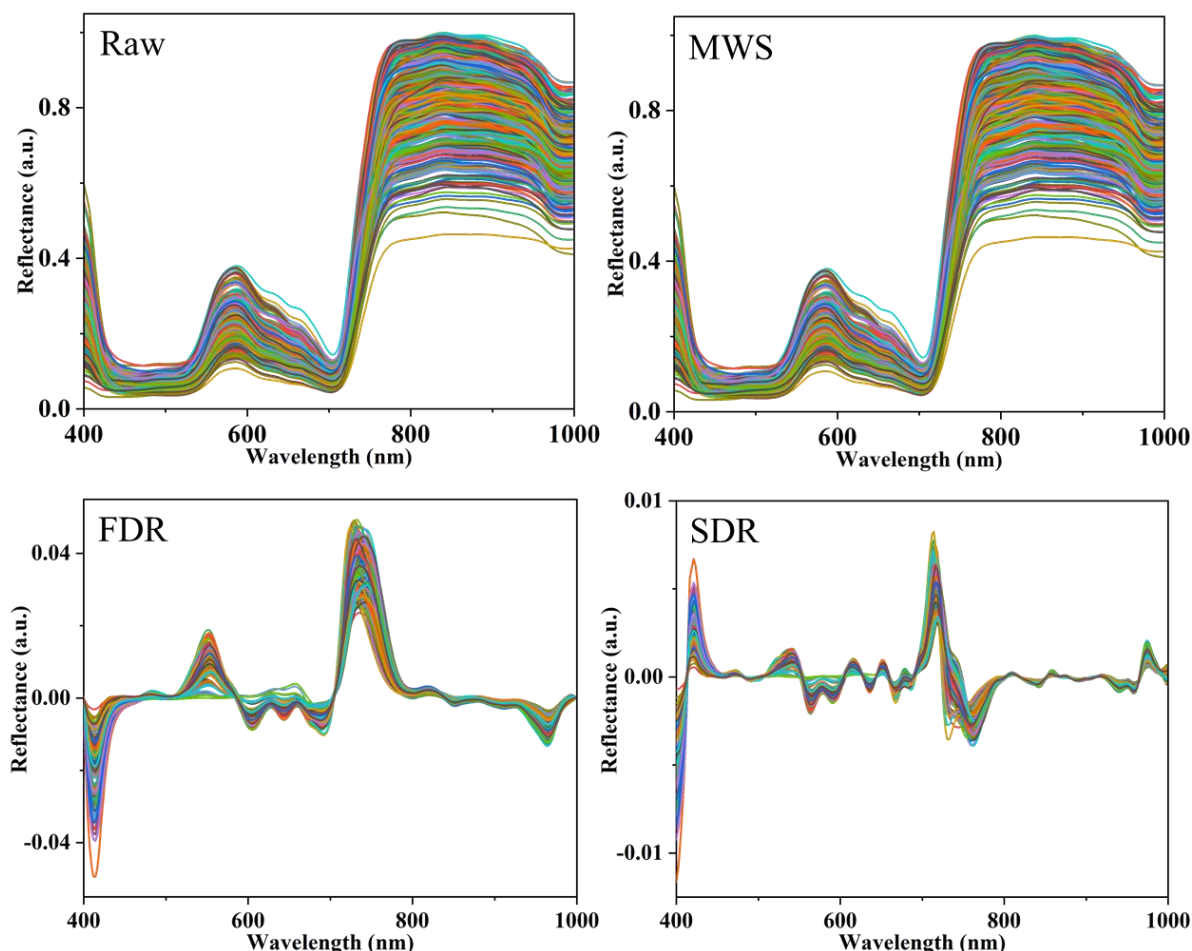


图 3.5 植物的原始平均光谱反射率以及经 MWS、FDR 和 SDR 预处理的平均光谱反射率

Figure 3.5 Raw mean spectral reflectance of plants and mean spectral reflectance preprocessed by MWS, FDR, and SDR

图 3.6 显示了具有不同 SSC 和 pH 值的生菜的光谱反射率。具有不同 SSC 和 pH 值的生菜反射率呈现显著差异和变化。这种现象表明，使用反射光谱来检测生菜中的 SSC 和 pH 值具有可行性。在 600-800 nm 区间的反射率变化随 SSC 和 pH 值的增加表现出不规则性。在 400-500 nm 和 800-1000 nm 区间，光谱反射率首先随 SSC 增加而增加，然后降低，最后再次升高。与此相反，在 800-1000 nm 区间，光谱反射率随着 pH 值的升高而呈现出相反的趋势，即先下降，然后上升，最后再次下降。

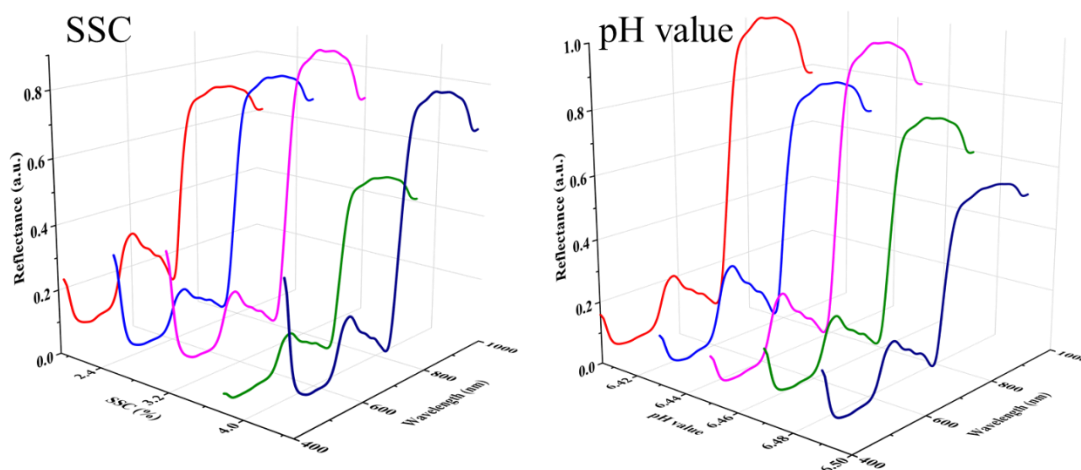


图 3.6 不同 SSC 和 pH 值的植物平均光谱反射率

Figure 3.6 Mean spectral reflectance of plants of different SSC and pH value

3.3.3 基于光谱的估算模型

(1) 基于深度学习的估算模型

本章采用了 SpecNet2D 和 SpecNet1D 模型来估算生菜的 SSC 和 pH 值。在优化批处理大小和学习率等超参数时,以校准集、验证集和测试集的 RMSE 作为评估标准。实验结果表明,批处理大小设置为 4 且学习率为 0.0001 时,模型的 RMSE 相对较低,因此本章选用了这组超参数配置来训练模型。SpecNet2D 和 SpecNet1D 的详细超参数展示在附录 A 表 A2-A3 中,生菜 SSC 和 pH 值的估算结果见表 3.2。

表 3.2 使用 SpecNet2D 和 SpecNet1D 估算 SSC 和 pH 值

Table 3.2 Estimation of SSC and pH value using SpecNet2D and SpecNet1D

Phenotypes	Models	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
SSC (%)	SpecNet2D	0.9642	0.1500	0.8974	0.1671	0.9030	0.1969	3.2237
	SpecNet1D	0.9031	0.2470	0.8527	0.2002	0.8385	0.2541	2.4980
pH value	SpecNet2D	0.8842	0.0319	0.8241	0.0296	0.7807	0.0313	2.1445
	SpecNet1D	0.8670	0.0342	0.8674	0.0257	0.8490	0.0260	2.5839

在估算生菜的 SSC 方面, SpecNet2D 模型超过了 SpecNet1D 模型,表现出显著的优势 (SpecNet2D, R_c^2 : 0.9642, RMSEC: 0.1500; R_v^2 : 0.8974, RMSEV: 0.1671; R_p^2 : 0.9030, RMSEP: 0.1969; RPD: 3.2237; SpecNet1D, R_c^2 : 0.9031, RMSEC: 0.2470; R_v^2 : 0.8527, RMSEV: 0.2002; R_p^2 : 0.8385, RMSEP: 0.2541; RPD: 2.4980), SSC 的 SpecNet2D 模型的 RPD 值大于 3, 这表明 SpecNet2D 提取的多尺度特征能够捕捉到 SSC 相关的更多信息,相较于 SpecNet1D 提取的单一尺度特征,其性能更加出色。尽管在 R_c^2 与 R_p^2 的差异上, SpecNet2D 与 SpecNet1D 在估算 SSC 的情况下展现出了相似的鲁棒性,但在 pH 值分析上, SpecNet1D 却展现出较高的精确度和鲁棒性

(SpecNet1D, R_c^2 : 0.8670, RMSEC: 0.0342, R_v^2 : 0.8674, RMSEV: 0.0257, R_p^2 : 0.8490, RMSEP: 0.0260, RPD: 2.5839; SpecNet2D, R_c^2 : 0.8842, RMSEC: 0.0319, R_v^2 : 0.8241, RMSEV: 0.0296, R_p^2 : 0.7807, RMSEP: 0.0313, RPD: 2.1445)。SpecNet2D 在提取 pH 相关的多尺度特征时可能引入了冗余信息,降低了模型的鲁棒性。基于 SpecNet2D 和 SpecNet1D 对 SSC 和 pH 值的估算结果在校准集和预测集上的散点图显示于图 3.7。图中可以看出,基于 SpecNet2D 的 SSC 估算值通常更贴近其实际参考值。在 pH 值分析中,SpecNet2D 在校准集的误差较小,而 SpecNet1D 在预测集的误差较小,而预测集的误差对于实际应用的影响更为关键。

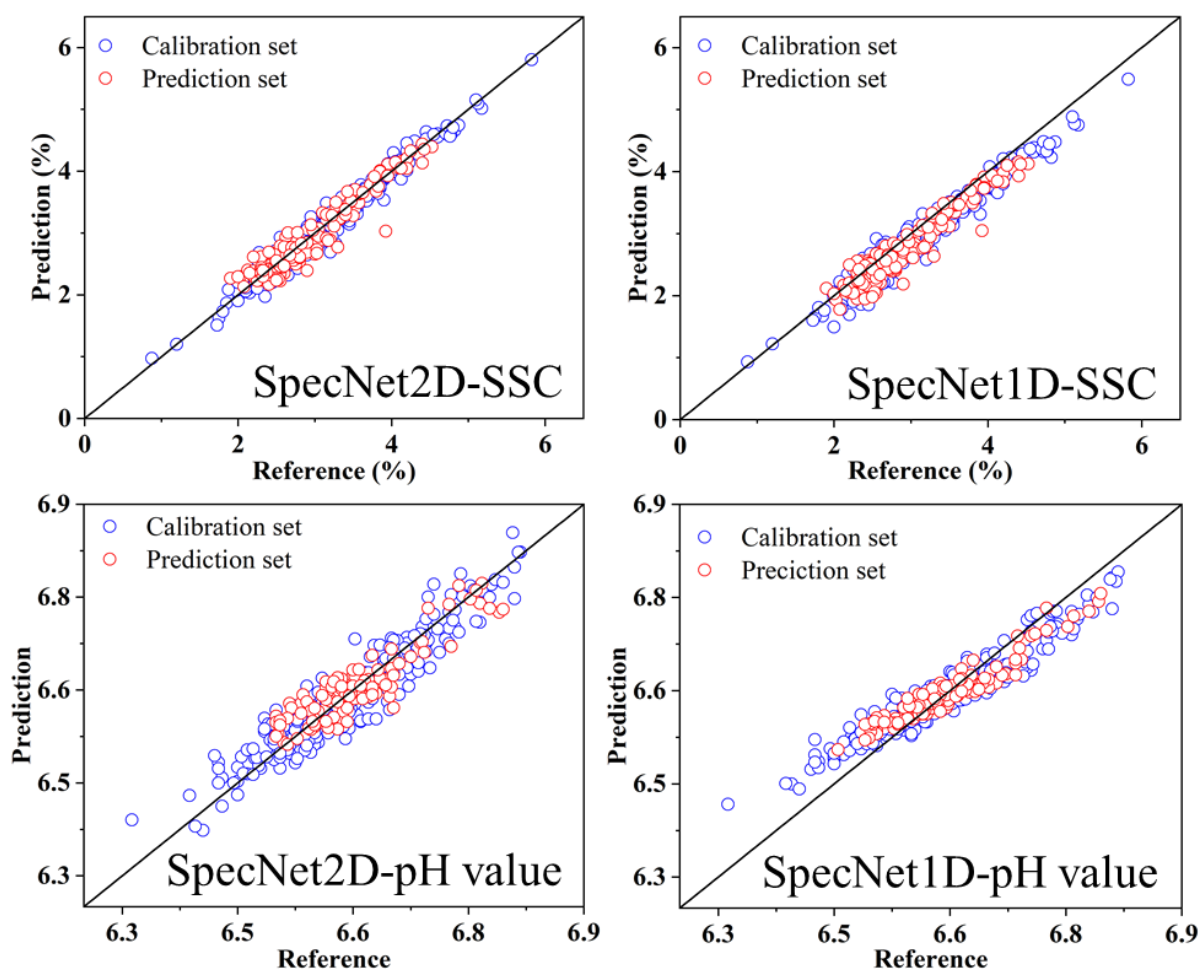


图 3.7 基于 SpecNet2D 和 SpecNet1D 的 SSC 和 pH 值估算

Figure 3.7 Estimation of SSC and pH value based on SpecNet2D and SpecNet1D

两种深度学习模型框架在估算不同表型特性方面的能力存在显著差异。据此,未来的研究应当着重寻找适应不同表型特性估算需求的网络架构,并定义一个能够同时准确估算多个表型特性的通用深度网络架构。与此同时,一些研究已经基于 HSI 技术建立了深度学习模型,用于估算植物的表型特性,并取得了良好的结果。例如,

Rehman 等人使用端到端的深度模型估算了玉米植株的 $RWC^{[72]}$, 并获得了 0.8720 的 R_p^2 。此外, Zhou 等人使用堆叠的卷积自编码器检测生菜中的重金属^[68], 取得了不错的估算效果 (R_p^2 : 0.9418, RMSEP: 0.04123, RPD: 3.214)。Sun 等人使用深度信念网络估算了生菜叶片中的镉含量^[135], 并获得了优异表现 ($R_p^2 = 0.9234$, RMSEP = 0.5423, RPD = 3.5894)。虽然先前关于生菜的研究已经取得了一定成果, 这些研究主要聚焦于叶片尺度, 但在实际种植和市场消费中, 生菜的冠层表型特性同样关键, 相关研究却较为缺乏。本章弥补了生菜冠层表型特性研究的空白, 并进一步证实, 融合高光谱成像技术的深度学习模型在准确估算植物表型特性方面具有重大潜力。

(2) 基于传统机器学习的估算模型

为更直观地评估 SpecNet2D 和 SpecNet1D 模型的性能, 运用了多元变量分析方法 (PLSR、MLR、LWR、ANN 和 SVR) 来建立机器学习模型, 对生菜的 SSC 和 pH 值进行了分析。相较于深度学习模型, 这些传统的机器学习模型在区分光谱中的噪声和关键信息方面显得能力有限。为此, 在建立机器学习估算模型前, 采用了包括 MWS、SG、FDR、SDR 和 WT 在内的五种光谱预处理技术, 目的是减少这些负面效应。生菜 SSC 和 pH 值的估算结果见表 3.3 和表 3.4, 相应的参数配置见附录 A 表 A4-A5。

表 3.3 基于机器学习方法的生菜 SSC 估算

Table 3.3 Estimation of the SSC in lettuces based on machine learning methods

Phenotype	Models	Pretreatment	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
SSC (%)	PLSR	Raw	0.9471	0.1846	0.9296	0.1677	0.8170	0.2658	2.3464
		MWS	0.9274	0.2163	0.9274	0.1703	0.8346	0.2527	2.4684
		SG	0.9284	0.2149	0.9212	0.1775	0.8404	0.2483	2.5126
		FDR	0.9246	0.2203	0.8542	0.2621	0.8400	0.2556	2.5098
		SDR	0.9465	0.1826	0.9332	0.1895	0.8581	0.2510	2.6648
		WT	0.9370	0.2014	0.9169	0.1822	0.8365	0.2512	2.4831
	MLR	Raw	0.9683	0.1430	0.9480	0.1442	0.7063	0.3367	1.8524
		MWS	0.9350	0.2046	0.9314	0.1657	0.8062	0.2735	2.2803
		SG	0.9399	0.1968	0.9293	0.1681	0.8056	0.2740	2.2769
		FDR	0.9537	0.1727	0.9333	0.1773	0.8055	0.2818	2.2763
		SDR	0.9476	0.1801	0.9320	0.1987	0.8148	0.2910	2.3327
		WT	0.9542	0.1718	0.9396	0.1554	0.8072	0.2728	2.2865
	LWR	Raw	0.9358	0.2034	0.9125	0.1870	0.8273	0.2582	2.4159
		MWS	0.9288	0.2142	0.9219	0.1767	0.8348	0.2526	2.4698
		SG	0.9278	0.2157	0.9106	0.1890	0.8389	0.2494	2.5012
		FDR	0.9178	0.2300	0.8403	0.2743	0.8387	0.2566	2.4999
		SDR	0.9183	0.2209	0.9249	0.2040	0.8587	0.2657	2.6705
		WT	0.9210	0.2255	0.8835	0.2158	0.8183	0.2648	2.3552
	ANN	Raw	0.9113	0.2391	0.8549	0.2409	0.7889	0.2855	2.1849
		MWS	0.8913	0.2646	0.8376	0.2548	0.7763	0.2939	2.1227
		SG	0.9014	0.2520	0.8633	0.2338	0.7864	0.2872	2.1722
		FDR	0.9584	0.1615	0.9473	0.1614	0.8170	0.2794	2.3468
		SDR	0.9627	0.1516	0.9532	0.1583	0.8170	0.2908	2.3470

表 3.3 (续)

SVR	WT	0.9266	0.2175	0.8912	0.2086	0.7968	0.2801	2.2270
	Raw	0.7186	0.4259	0.7142	0.3397	0.6615	0.3658	1.7053
	MWS	0.6838	0.4552	0.6839	0.3557	0.5721	0.4150	1.5030
	SG	0.7073	0.4367	0.7172	0.3501	0.5868	0.4075	1.5308
	FDR	0.8230	0.3391	0.7343	0.3031	0.8009	0.2921	2.2045
	SDR	0.8768	0.2799	0.8893	0.2635	0.7928	0.3004	2.1765
	WT	0.7121	0.4337	0.6990	0.3491	0.5835	0.4148	1.5036

基于机器学习所建立的 SSC 分析模型中, 经 FDR 和 SDR 预处理的反射光谱相比其他光谱数据显示出更优的结果, 这可能是因为 FDR 和 SDR 增强了不同 SSC 水平之间光谱的区分度。所有预处理方法都提高了 MLR 模型的精度, 其中 SDR 预处理所得模型具备最高的准确性, 其 R_c^2 为 0.9476, RMSEC 为 0.1801, R_v^2 为 0.9320, RMSEV 为 0.1987, R_p^2 为 0.8148, RMSEP 为 0.2910, RPD 为 2.3327。尽管在 FDR 预处理光谱数据的基础上, SVR 模型的 R_p^2 值超过 0.8, 但整体表现依然不及其他模型。PLSR 和 LWR 模型展现出较高的稳健性, 特别是在 SDR 预处理光谱下的 LWR 模型表现最为出色 ($R_c^2 = 0.9183$, RMSEC = 0.2209, $R_v^2 = 0.9249$, RMSEV = 0.2040, $R_p^2 = 0.8587$, RMSEP = 0.2657, RPD = 2.6705)。不过, 即便是机器学习模型中的最优 SSC 估算结果, 也未能达到 SpecNet2D 模型的估算水平。

表 3.4 基于机器学习方法的生菜 pH 值估算

Table 3.4 Estimation of the pH value in lettuces based on machine learning methods

Phenotype	Models	Pretreatment	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
pH value	PLSR	Raw	0.8617	0.0344	0.7778	0.0342	0.7092	0.0373	1.8623
		MWS	0.8667	0.0338	0.8105	0.0316	0.7092	0.0373	1.8624
		SG	0.8753	0.0327	0.8170	0.0310	0.7205	0.0366	1.8996
		FDR	0.8881	0.0309	0.8714	0.0216	0.6884	0.0395	1.7990
		SDR	0.8828	0.0314	0.7105	0.0322	0.6635	0.0419	1.7312
		WT	0.8840	0.0315	0.8482	0.0283	0.7259	0.0363	1.9180
	MLR	Raw	0.9434	0.0220	0.9025	0.0227	0.6182	0.0472	1.5162
		MWS	0.8811	0.0321	0.8342	0.0341	0.6692	0.0393	1.7460
		SG	0.8780	0.0323	0.8305	0.0299	0.6976	0.0381	1.8262
		FDR	0.8879	0.0308	0.8080	0.0278	0.6665	0.0415	1.7389
		SDR	0.8831	0.0306	0.8663	0.0317	0.6768	0.0442	1.7665
		WT	0.9437	0.0220	0.9108	0.0217	0.6317	0.0447	1.6013
	LWR	Raw	0.8696	0.0334	0.8378	0.0326	0.7274	0.0362	1.9233
		MWS	0.8734	0.0329	0.8511	0.0313	0.7380	0.0354	1.9619
		SG	0.8646	0.0341	0.8194	0.0344	0.7309	0.0359	1.9357
		FDR	0.8782	0.0324	0.8682	0.0239	0.7120	0.0375	1.8711
		SDR	0.8651	0.0334	0.7783	0.0372	0.7355	0.0381	1.9525
		WT	0.8740	0.0329	0.8667	0.0296	0.7368	0.0355	1.9574
	ANN	Raw	0.9115	0.0275	0.8375	0.0332	0.6695	0.0398	1.7469
		MWS	0.8723	0.0331	0.7958	0.0366	0.6569	0.0406	1.7145
		SG	0.9090	0.0279	0.8476	0.0316	0.6724	0.0396	1.7546
		FDR	0.8866	0.0313	0.7910	0.0317	0.5900	0.0443	1.5683
		SDR	0.8802	0.0318	0.8801	0.0274	0.6654	0.0419	1.7362
		WT	0.9253	0.0253	0.9087	0.0245	0.6656	0.0400	1.7367

表 3.4 (续)

SVR	Raw	0.7629	0.0455	0.5847	0.0479	0.6618	0.0421	1.6508
	MWS	0.7496	0.0464	0.5473	0.0504	0.6515	0.0424	1.6380
	SG	0.7556	0.0455	0.6859	0.0420	0.6257	0.0453	1.5703
	FDR	0.8505	0.0361	0.8486	0.0246	0.7084	0.0386	1.7832
	SDR	0.8814	0.0321	0.7906	0.0359	0.6624	0.0410	1.6856
	WT	0.7761	0.0440	0.6885	0.0416	0.6267	0.0443	1.5687

在 pH 值的估算上, ANN 模型的性能最低。比起原始光谱, 仅有 SG 预处理对模型的准确性有所改进, 获得的 R_c^2 为 0.9090, RMSEC 为 0.0279, R_v^2 为 0.8476, RMSEV 为 0.0316, R_p^2 为 0.6724, RMSEP 为 0.0396, RPD 为 1.7546 的结果。虽然所有预处理技术都增加了 MLR 模型的鲁棒性, 但最佳结果的 R_p^2 依然低于 0.7。SVR 模型在大部分情况下的估算准确度都比 MLR 模型低, 但是它的最佳结果却比 MLR 模型好, 经 FDR 预处理的 SVR 模型的 R_p^2 达到了 0.7084。PLSR 和 LWR 的结果与它们在 SSC 估算中的效果类似, 在所有模型中表现出较高的性能。具体来说, MWS、SG 和 WT 都提高了 PLSR 和 LWR 模型的准确性, 其中以 MWS 预处理光谱的 LWR 模型表现最优 (R_c^2 : 0.8734, RMSEC: 0.0329, R_v^2 : 0.8511, RMSEV: 0.0313, R_p^2 : 0.7380, RMSEP: 0.0354, RPD: 1.9619), 该最佳模型的 RPD 值小于 2, 还是低于 SpecNet1D 和 SpecNet2D 模型的估算水平。

图 3.8 呈现了利用机器学习模型获得的 SSC 和 pH 值的最佳估算结果, 与图 3.7 的最优结果相比, 贴近拟合线的数据点数量有了显著减少。机器学习模型的结果表明了不同的预处理手段对同一模型的促进作用并非总是一致, 有时甚至会起到负面作用, 且同一预处理方法对不同模型的影响也呈现出多样性。鉴于此, 选择恰当的多元分析技术和预处理对于机器学习模型的性能至关重要。与机器学习模型相比, 深度学习模型有能力自主地从原始数据中提取特征, 并建立起更高精度的模型。类似的结论也在早前的研究中得到说明^[64, 72]。总体来说, 相对于深度学习模型, 机器学习模型建立之前所需的特征工程降低了 HTPP 在数据处理上的流畅度, 且其建立的模型性能显著不如研究中提出的深度学习模型。因此, 在高通量表型平台数据分析的实践中, 应优先考虑深度学习模型的应用。

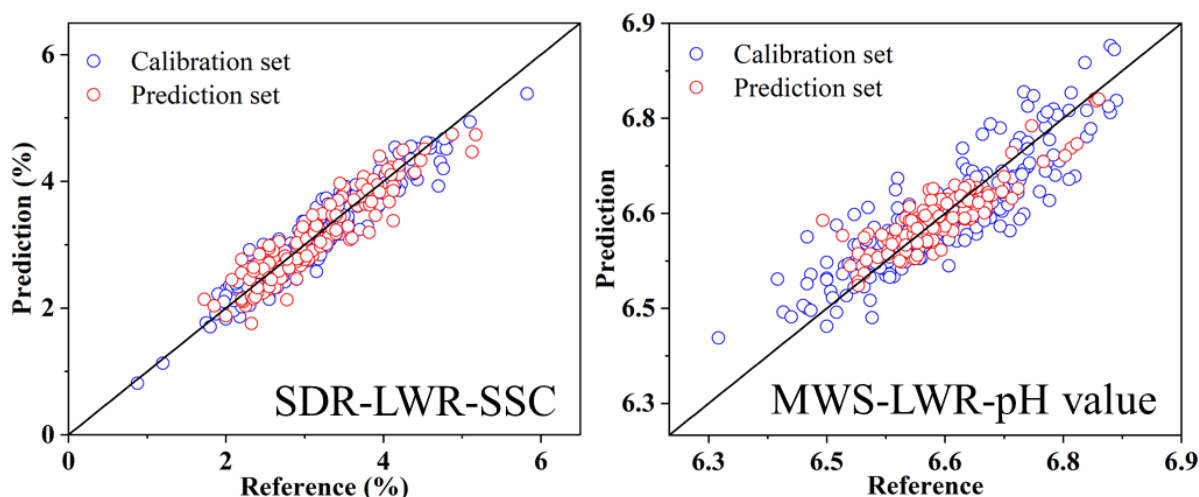


图 3.8 基于 LWR 和 SDR 的 SSC 估算和基于 LWR 和 MWS 的 pH 值估算

Figure 3.8 Estimation of SSC based on LWR and SDR, and pH value based on LWR and MWS

(3) 基于光谱特征选择的估算模型

从高光谱图像中提取的反射光谱具有波段窄和维度高的特点，这种特点使得邻近波长的光谱特征之间的相关性比较高，并且光谱特征中夹杂了与生菜的 SSC 和 pH 值无关的信息^[136]。这一现象可能会削弱机器学习模型的精确度，并提高模型计算上的复杂度。为了解决这个问题，运用了 CARS 技术，从经过预处理的光谱中筛选出关键波长的光谱特征，以期获得更佳的实验结果。根据表 3.3 和表 3.4 的结果，挑选了三种优选的预处理方法以进行光谱特征的选择，分别是对 SSC 估算有更好效果的 SG、FDR 和 SDR 以及 pH 值的 MWS、SG 和 WT。接着，基于这些筛选出来的关键波长的光谱特征，采用了表现卓越的多元分析方法（PLSR 和 LWR）来建立机器学习估算模型，并对比全波段光谱的估算结果，观察模型性能的变化。

估算结果如表 3.5 所示，具体参数设置和具体选定光谱特征见附录 A 表 A6-A7。研究中发现，CARS 为不同的预处理光谱和模型挑选了不同数量和特定的波长的光谱特征。在估算生菜 SSC 方面，只有应用了 LWR 和 SG 的模型显示出性能提升 (R_c^2 : 0.9210, RMSEC: 0.2225, R_v^2 : 0.8865, RMSEV: 0.2233, R_p^2 : 0.8450, RMSEP: 0.2475, RPD: 2.5400)。此外，在运用 CARS 之后，变量的数量 (NVs) 减少了大约 90%。与全波段光谱模型相比，基于筛选光谱特征的 SSC 估算模型性能大部分略有下降。而基于 CARS 选择的光谱特征的 PLSR 模型在 pH 值分析上展现了更好的性能。WT 预处理的光谱不仅用于建模的光谱特征的数量最少，而且取得了更好的结果 (R_c^2 : 0.7739, RMSEC: 0.0449, R_v^2 : 0.7589, RMSEV: 0.0331, R_p^2 : 0.7404, RMSEP: 0.0322, RPD: 1.9626)。与此同时，基于 LWR 和 WT 的模型拥有最低的估算误差 ($R_c^2 = 0.7628$,

RMSEC = 0.0459, $R_v^2 = 0.6461$, RMSEV = 0.0382, $R_p^2 = 0.7477$, RMSEP = 0.0330, RPD = 2.0532)。值得注意的是, 在所有基于多元分析方法的 pH 估算模型中, 仅有 LWR 和 WT 模型的 RPD 值超过了 2, 但是该模型的性能仍低于 SpecNet1D 的性能。总体来看, 相比于全波段光谱模型, 基于 CARS 挑选光谱特征建立的模型并不总是能够带来性能的提高。因此, 与基于多元分析方法的机器学习模型相比较, 深度学习模型大幅减少了手动特征提取所需的时间和错误率, 更适用应用于 HTTP 平台的数据分析。

表 3.5 结合 CARS 估算生菜的 SSC 和 pH 值

Table 3.5 Estimation of the pH value in lettuces combining with CARS

Phenotypes	Models	Pretreatment	NVs	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
SSC (%)	PLSR	SG	33	0.9196	0.2274	0.9093	0.1940	0.8378	0.2409	2.4831
		FDR	27	0.8794	0.2791	0.8474	0.2401	0.8397	0.2458	2.4977
		SDR	30	0.8888	0.2667	0.8687	0.2467	0.8447	0.2529	2.5377
	LWR	SG	25	0.9210	0.2225	0.8865	0.2233	0.8450	0.2475	2.5400
		FDR	24	0.9203	0.2291	0.8934	0.2448	0.8370	0.2355	2.4772
		SDR	30	0.8755	0.2822	0.8602	0.2546	0.8504	0.2482	2.5858
pH value	PLSR	MWS	42	0.7366	0.0488	0.7679	0.0337	0.7104	0.0340	1.8584
		SG	43	0.7320	0.0490	0.6380	0.0351	0.7249	0.0337	1.9065
		WT	36	0.7739	0.0449	0.7589	0.0331	0.7404	0.0332	1.9626
	LWR	MWS	50	0.7954	0.0429	0.6856	0.0380	0.7145	0.0339	1.8714
		SG	48	0.7756	0.0451	0.7157	0.0331	0.7277	0.0326	1.9165
		WT	41	0.7628	0.0459	0.6461	0.0382	0.7477	0.0330	2.0532

3.4 本章小结

在本章中, 基于 Inception 模块和 FCNN 设计了两种深度学习模型, 它们通过端到端的学习流程准确地估算了生菜的 SSC 与 pH 值。在传统的机器学习模型中, 为了精确估算生菜的表型特性, 在将光谱数据输入估算模型之前, 需要执行一系列复杂的特征工程步骤, 包括数据的预处理和精细的光谱特征筛选等, 才能获得一个比原始数据输入相对较优的结果。然而, 深度学习模型能够直接处理从高光谱图像中提取的原始反射光谱数据, 并直接估算生菜的 SSC 与 pH 值, 省去了繁杂的特征工程。

将 SpecNet1D 和 SpecNet2D 模型与基于五种多元分析方法建立的机器学习模型 (PLSR、MLR、LWR、ANN 和 SVR) 进行了比较。并使用了五种预处理方法 (MWS、SG、FDR、SDR 和 WT) 对光谱进行去噪, 同时使用 CARS 来移除光谱中冗余的光谱特征, 以提高基于这五种方法建立的模型的性能。结果证明提出的 SpecNet2D 和 SpecNet1D 模型在估算性能上仍然优于这些基于多元分析方法的机器学习模型。SpecNet2D 模型在估算生菜的 SSC 方面取得了最佳结果 (R_p^2 : 0.9030, RMSEP: 0.1969, RPD: 3.2237)。而 SpecNet1D 模型则在 pH 值的估算上表现最为突出 (R_p^2 : 0.8490,

RMSEP: 0.0260, RPD: 2.5839)。这一结果不仅凸显了深度学习模型在直接从原始数据中自动提取关键特征方面的高效性,同时也显示了其在植物表型特性分析中减少特征选择误差和时间消耗的潜力。不过,为了更精准地分析和量化植物的多样表型特性,发掘与之相匹配的或通用的网络结构依然是未来研究的重点。另外,了解植物表型特性随时间的动态变化对于确定最佳收获时机具有重要帮助,考虑到不同生菜品种间形态特征的多样性,将时间序列和图像数据等其他数据模态纳入模型输入显得尤其有价值。多模态数据的引入可以提供更全面的特征视角,不仅增加了模型处理的信息维度,而且有助于从不同侧面评估和提升模型的性能。通过这种方式,模型将能够更有效地捕捉和理解植物生长的细微差别,尤其对于在不同环境条件下生长表现出显著差异的植物,能够实现更为精确的表型分析。

第四章 基于光谱和时间序列的生菜品质表型指标的估算

4.1 引言

在第三章的研究中，通过探索和应用深度学习技术处理光谱数据，本研究成功估算了大田生菜的品质表型特性，包括 SSC 和 pH 值。这一工作不仅展示了深度学习在数据特征自动提取方面的应用潜力，也为进一步探索多模态数据融合技术在植物表型分析中的价值奠定了基础。本章在第三章的基础上进一步拓展，将研究范围从大田环境中的生菜扩展到不同水分处理条件下温室内生长的生菜，并且增加了研究涉及的生菜品种，为探索生菜品种间的表型差异性和适应性提供了更为丰富的数据基础。考虑到不同水分条件可能对生菜的品质表型特性产生影响，本章引入了时间序列数据与光谱数据的融合方法，以期获得更加精确的生菜品质表型估算结果。

为了实现研究连贯性和便于后续工作的展开，本章选择以 Inception 网络架构为核心来进一步处理光谱数据。选择 Inception 模块的原因在于其能够在多个尺度上捕获图像特征，从而为复杂表型特性提供更丰富的信息。此外，Inception 架构的灵活性和高效性使其在模型参数调优中有着更大的操作空间。通过固定使用 Inception 网络架构，还可以确保研究的一致性，从而更好地适用于植物表型估算任务的需求。

本章的研究目标是通过整合时间序列数据和光谱数据，捕捉和分析水分胁迫对生菜品质表型特性的影响，进一步优化和提升模型的估算能力。引入时间序列数据的原因在于其能够提供生菜在特定时间段内生长变化的连续观测，这为理解生菜对水分胁迫的响应提供了重要视角^[100-105]。这种多模态数据融合方法有望在细节上提供更丰富的信息，帮助构建更为准确的生菜表型特性估算模型^[106]。

因此，本章将详细介绍在不同水分处理条件下温室生菜品质表型特性估算的研究方法，包括数据收集、模型建立与评估过程。通过与单一光谱数据模态的品质表型估算效果进行比较，验证多模态数据融合技术的有效性。研究结果有利于进一步推动精准农业和植物表型分析领域的技术进步，并为接下来的研究提供方向和思路。

4.2 数据获取与处理

图 4.1 展示了本章的实验设计与数据处理的完整流程图。实验数据的采集包括获取 SSC 和 pH 值的参考值以及不同表型分析平台采集的数据。数据处理环节涉及对高光谱

图像和时间序列的处理，最终将处理后的数据输入到模型中，用以估算 SSC 和 pH 值。具体的实验步骤和数据处理方法将在本节的后文部分进行详细阐述。

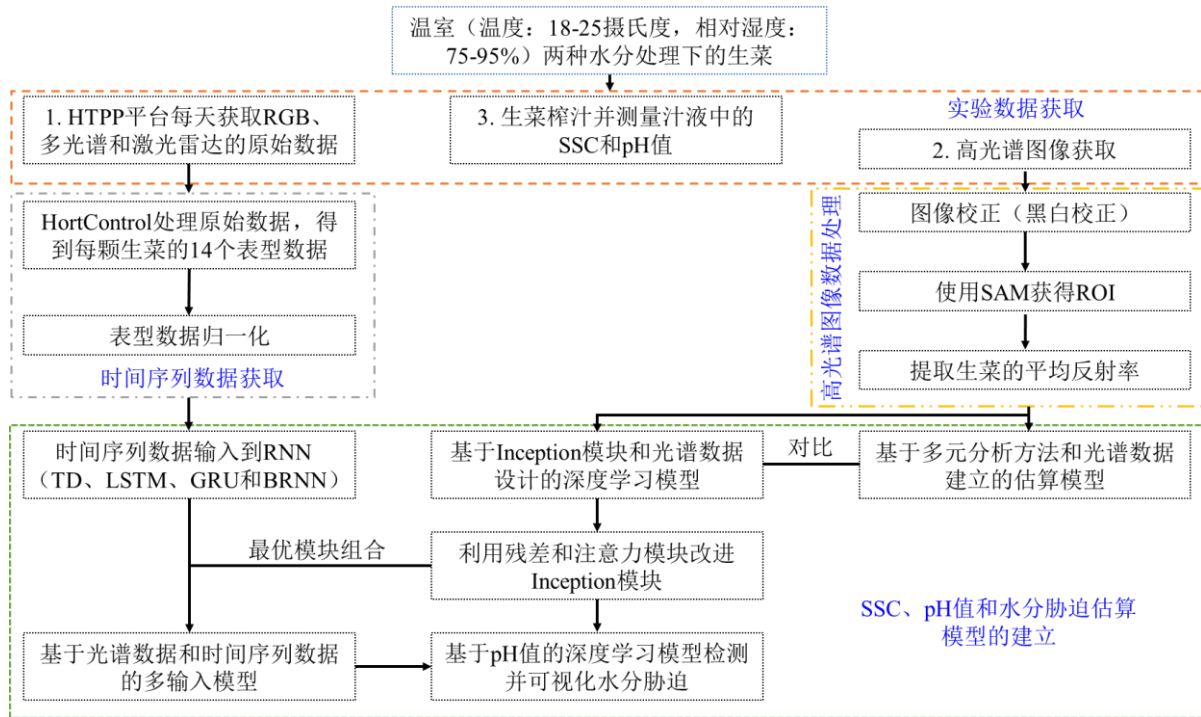


图 4.1 估算生菜的 SSC，pH 值和水分胁迫的实验流程图

Figure 4.1 Flowchart of the experiment used to estimate SSC, pH value, and water stress in lettuces

4.2.1 实验设计

在北京市农林科学院设施温室中选取了四个不同品种的生菜（奶油、结球、散叶以及罗马生菜）进行实验。生菜种植时间为 2021 年 9 月 18 日。实验现场部署的 LQ-GHPheno 如图 4.2A 所示。生菜在花盆中种植，如图 4.2B 所示，并分别施以正常浇水和水分胁迫两种处理。在专业人员的指导下，正常水分处理生菜自种植起每 7 至 10 天浇水一次，而水分胁迫处理生菜在种植后则不再进行浇水。两种处理均未施加肥料。选取了 160 株奶油生菜、200 株结球生菜、100 株散叶生菜和 100 株罗马生菜进行实验，每种水分处理生菜数量均分。实验时间为 2021 年 11 月 16 日至 19 日，分别对奶油、结球、散叶和罗马生菜的品种每日采样 40、50、25 和 25 株。实验数据的获取包括以下几个步骤：首先，收集生菜的时间序列数据；其次，采集生菜的高光谱图像；最后，对所有生菜叶片进行榨汁处理。测定生菜汁液中的 SSC 和 pH 值的仪器同 3.2.1 节。每个样本的 SSC 和 pH 值测量三次，最后取平均值作为该样本 SSC 和 pH 值的参考值。

本章使用暗箱成像室系统来获取高光谱图像。高光谱图像具体的获取方法见 2.2.1 节，光谱反射率提取过程见 2.2.2 节。



图 4.2 LQ-GHPheno 图片 (A) 和温室植物图片 (B)

Figure 4.2 Picture of LQ-GHPheno (A) and plants in the greenhouse (B)

4.2.2 时间序列数据的获取

位于温室内的高通量表型平台，集成了 RGB 相机、多光谱相机和激光雷达传感器，这一系统每 24 h 自动采集一次数据。通过配套软件 HortControl (PhenoSpex, 荷兰) 的协助，可以从每株生菜中提取了 14 种表型参数，包括植物数字生物量、绿度、株高、最大株高、色调、叶片角度、叶面积、叶面积指数、投影叶面积、叶片倾斜度、激光穿透深度、归一化植被指数 (NDVI)、归一化色素叶绿素指数 (NPCl) 和植物衰老反射指数 (PSRI)。其中，数字生物量反映了植物叶面积与株高的关联，绿度与植物在绿色波段的反射率密切相关，而最大株高值则补充了株高测定在精确度上的不足。株高的测量重在稳定性，最大株高值的测量则注重精确性。这些表型参数分为颜色、形态和植被指数三大类别，详细描述见附录 B 表 B1。基于生菜的生长状况，选取了采摘前五天内的时间序列数据进行分析，为每株生菜累积了 70 个时间序列数据记录。

4.2.3 光谱和时间序列的融合方法

(1) Inception 模块

植物的光谱反射数据蕴含丰富信息，这些信息反映了植物的生长状态及其生理和生化特性。与传统的多元分析技术相比，深度学习模型在处理属性差异性较大的样本时表现出更高的稳健性^[72]。本章在第三章的研究基础上，建立了一个基于 Inception 模块的深度学习模型 (如图 4.3 所示)，这一模型能够直接采用生菜的平均光谱反射率作为输入，并据此估算生菜的品质表型，包括 SSC 和 pH 值。

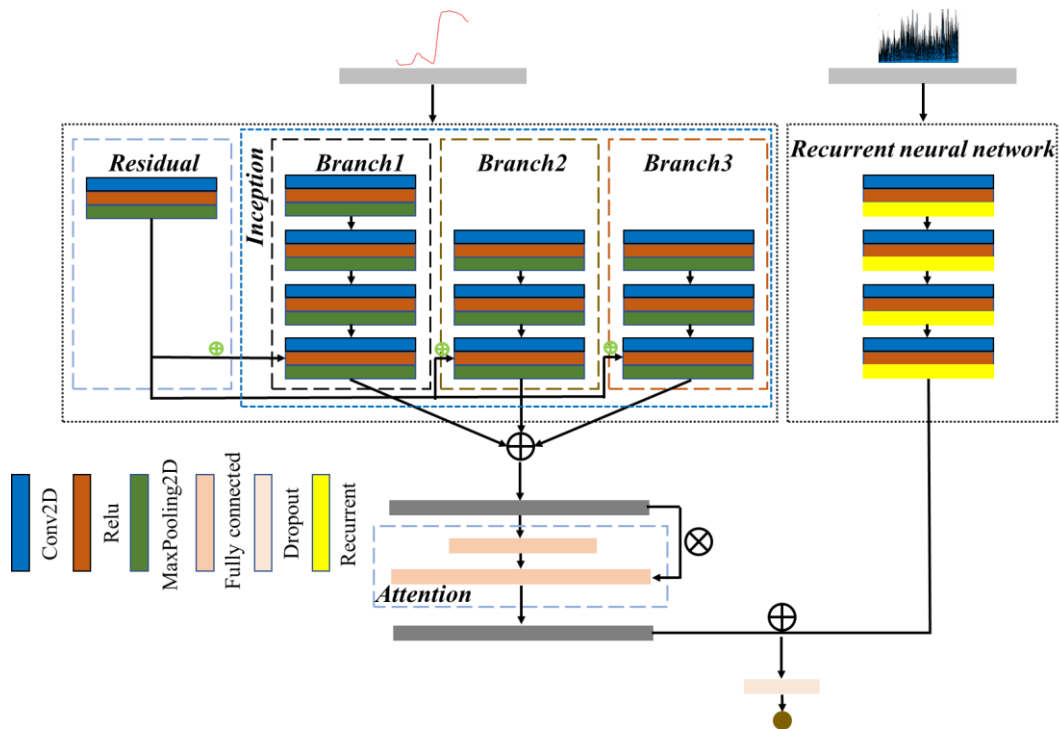


图 4.3 深度学习模型（SpecTimeNet）的架构

Figure 4.3 Architecture of deep learning model (SpecTimeNet)

CNN 本质上是一种多层感知器，它可以自动从原始数据中学到与样本标签间的映射关系。得益于其稀疏连接和权重共享的特点，CNN 能有效提取局部特征，并且与 FCNN 相比，大幅减少了参数的数量，从而降低了过拟合的风险^[138]。基于其优越的特征提取能力，CNN 已经在植物品质分析和胁迫估算方面得到了应用^[65, 66, 68-70]。然而，有关水分胁迫对生菜品质表型影响的研究仍相对较少。Inception 模块是一种轻量级且有效的 CNN 架构，其通过增加网络的宽度和深度来设计，从而在深度与宽度之间取得平衡，避免了模型饱和现象的发生^[71]。该模块在特征映射上并行执行多个卷积操作，以捕获不同尺寸的信息。然后，这些信息将合并成深度特征，实现了多尺度特征的整合。基于 CNN 的 Inception 模块在深度学习模型框架中的数据流程如附录 B 图 B1 所揭示，展示了输入数据通过卷积层序列处理的过程。本章中，Inception 模块的第一分支包含一个不同卷积核尺寸（ $1 \times n$ ， n 取值可为 7、5 或 1）的单卷积层。第二分支由卷积核尺寸为 1×3 的两个卷积层组成。第三分支包括一个池化层，紧接着是两个 1×5 的卷积层。最后，第四分支包含三个卷积核尺寸为 1×3 或 1×5 的卷积层。每个分支的配置都是根据特定的品质表型和实验结果设计的，以获得最优的模型表现。深度学习模型参数设置，如卷积核数量、卷积核尺寸、池化操作类型和激活函数等，在后续的结果讨论部分和附录 B 资料中有详细说明。Inception 模块参数的调整通常采用“超参数

调优”或“超参数优化”的技术，这一过程涉及系统地尝试不同的参数组合，并通过验证集来评价它们的表现。

（2）残差和注意力模块

在 Inception 模块中堆叠神经网络层会加剧深度学习模型的训练难度和所需时间，并可能导致梯度消失问题和冗余特征的产生。针对这些问题，采用了残差模块和注意力模块（如图 4.3 所示）进行改善^[73, 74]。残差模块通过巧妙地设计跳跃连接，使得原始输入信息能够直接传输到更深层次的网络结构中，从而有效避免了在神经网络深度增加时关键信息的丢失。而注意力模块，能够根据不同任务的输入自适应地为每个特征图分配权重，聚焦于显著特征。在残差模块中，需要关注的参数包括卷积层的卷积核数量与尺寸、是否执行下采样以及是否采用批量标准化。至于注意力模块，特征输入头的数量、每个头部的线性变换权重、查询与键之间的映射函数，以及激活函数，都是关键参数。

（3）RNN

CNN 和 FCNN 的共同显著特性是它们均不具备记忆能力。这意味着在处理输入时，它们对每个输入都是独立操作的，并不在各个输入之间保持状态。因此，这类网络在处理数据点组成的序列时，需要将整个序列一次性地输入网络，换言之，就是将序列转化为单一的数据点。在处理时间序列数据等序列数据时，CNN 和 FCNN 的性能通常较差，这在很大程度上是因为它们无法捕捉输入数据的上下文关联，易受梯度消失问题的困扰。

相较之下，RNN 能够有效应对这一难题。RNN 通过顺序访问序列中的各个数据点，并维持一个含有所观测信息的状态来处理序列数据^[100-105, 137]。当开始处理新的独立序列时，RNN 会重置其状态，从而便于将序列作为单独的数据点来处理。不同于 FCNN 和 CNN，RNN 不是一次性处理所有数据点，而是通过遍历序列内的数据点来逐步执行。在本章中，运用了 TD、LSTM、GRU 和 BRNN 这四种 RNN 架构，以处理来自 HTTP 平台的时间序列数据。

TD 架构能够对每个输入序列的时间步使用卷积层处理，在实现所有时间步的完整连接的同时，还能在不同卷积层的特征图之间共享权重^[139]。LSTM 具有跨多个时间步传递和保存信息的能力，有效预防了在处理时间序列数据的过程中早期信号的衰减问题。GRU 与 LSTM 遵循相同的操作原则，但它对 LSTM 进行了简化处理，减少了相应的门控操作，导致其表示能力可能略逊于 LSTM，但其运行起来的计算成本更低^[99]。

在机器学习中，计算成本与表示能力之间的这种权衡极为常见。BRNN 的核心结构是由两个单向循环神经网络组成的。在每个时间步上，输入将从正向和反向同时喂入这两个 RNN，并且各自生成在该时间步的状态和输出。BRNN 的输出将是这两个单向网络输出的直接组合。除了方向性，这两个网络在结构上是一致的，并且具有结构对称性。任何 RNN，包括 LSTM 和 GRU，都能够建立循环结构。

TD 的配置参数涵盖了每个时间步所应用层的具体设置，包括卷积层中卷积核的数量和尺寸等特有的参数设定。LSTM、GRU 和 BRNN 的参数设置在一定程度上是相似的，它们都包括层内的单元数量或隐藏状态的数量、选用的激活函数类型，以及权重初始化的方式。具体来说，LSTM 还包括记忆单元数量的参数以及每个门控和输出所使用的激活函数，GRU 的参数设置则涉及重置门和更新门的配置，及其偏置的初始方法。同时，BRNN 的参数设定还包括了对于循环结构类型的选择。TD、LSTM、GRU 和 BRNN 在 RNN 体系内的运作机制和具体应用详见附录 B 图 B2。

在本章中，建立了一种多输入模型，其中光谱数据输送至 CNN，而时间序列数据则由 RNN 处理。这两种网络各自提取的光谱特征和时间序列特征经过融合，并一同输入到一个全连接层中，最终实现对品质表型的估算（见图 4.3）。本研究中，所有的基于光谱数据和时间序列数据的深度学习特征融合模型统称为 SpecTimeNet。

4.2.4 传统机器学习方法

由于 Inception 模型是本章深度学习模型中的基线模型，本章采用了线性（PLSR）与非线性（SVR）两种多元分析方法来建立机器学习估算模型，以此评估 Inception 模块在处理光谱数据上的有效性。在运用 SVR 时，需要挑选合适的核函数、正则化参数以及核参数。常用的 SVR 核函数包含线性核（linear）、多项式核以及径向基函数核（RBF）。本章对每个品质表型的 SVR 模型的不同核函数进行了探究，并通过性能指标来评价它们的效果，以确定最合适的核函数。最终，SSC 选用了 linear，而 pH 值则采用了 RBF。

4.2.5 SSC 和 pH 值的估算模型建立

在最终的分析中，受样本标签遗失及仪器量程限制的影响，分别有 524 个样本用于 SSC 和 pH 值的建模（四个生菜品种具体的样本数量见附录 B 表 B2）。两种不同的水分处理条件下，采集了等量的样本。这些样本将随机分配至校准集和预测集^[67]，比例为 2:1。接着，从校准集中随机挑选出来 10% 的样本，构成验证集。无论是正常水分

处理还是水分胁迫处理，校准集和预测集都是按照 2:1 的比例随机选取的。为保证评价不同建模方法性能的公正性，随机数发生器的状态在所有建模中保持一致。模型性能评价指标及其描述见 2.3.3 节。

在深度学习模型和基于多元分析方法的机器学习模型参数调优过程中，通过在校准集上对不同参数组合进行训练，并在验证集上考察其表现，以挑选出最佳参数配置。特别是，倾向于选择那些具有较高决定系数和较低均方根误差的参数组合。对于基于 Inception 模块、残差模块、注意力模块和 RNN 的深度学习模型，在 Python 3.8.13 和 TensorFlow 2.6.0 环境下进行建模操作，而多元分析方法则使用 MATLAB2020（Mathworks 公司，美国）来完成。硬件配置与第三章相同。

4.3 结果分析与讨论

4.3.1 不同水分处理的 SSC 和 pH 值的参考值

表 4.1 和图 4.4 详细展示了 SSC 和 pH 值的统计数据及其分布范围。正如表 4.1 和图 4.4 所展现的，两种水分处理条件下各个品质表型的参考值明显不同，这提示需要考虑分层抽样对建模性能可能带来的影响^[140]。在生菜遭受水分胁迫的情况下，观察到 pH 值出现了显著的下降趋势，同时 SSC 有所增加。这可能是生菜经历缺水胁迫时，复杂的生理响应导致的。首先，为了适应水分不足，植物会增加呼吸作用来产生更多能量，但这也导致更多二氧化碳的产生，使得细胞内酸性物质积累，致使细胞质 pH 值下降。同时，植物累积了更多的有机酸，如苹果酸和柠檬酸，不仅帮助维持渗透平衡，也进一步导致 pH 值的下降。缺水还导致气孔关闭，光合作用变得低效，减少了对二氧化碳的利用，进一步影响了 pH 的平衡。此外，随着水分胁迫的加剧，植物可能会通过合成更多如蔗糖和脯氨酸等渗透调节物质来提高细胞液的渗透浓度，储存更多的水分，从而引起 SSC 上升。这一系列的生理反应，不仅揭示了植物如何调整自身机制来应对缺水胁迫，也为在实践中依据 SSC 和 pH 值的变化来优化灌溉策略提供了宝贵的指导，有助于提升作物品质和农业生产效率。

表 4.1 生菜中 SSC 和 pH 值的参考值

Table 4.1 Reference measure of SSC and pH value in lettuces

Phenotypes	Normal			Water Stress		
	Range	Mean	SD	Range	Mean	SD
SSC (%)	0.95-3.8	1.8987	0.5204	1.4-9.05	3.0267	1.1562
pH value	4.89-6.82	6.1141	0.6037	4.75-6.76	5.1609	0.2249

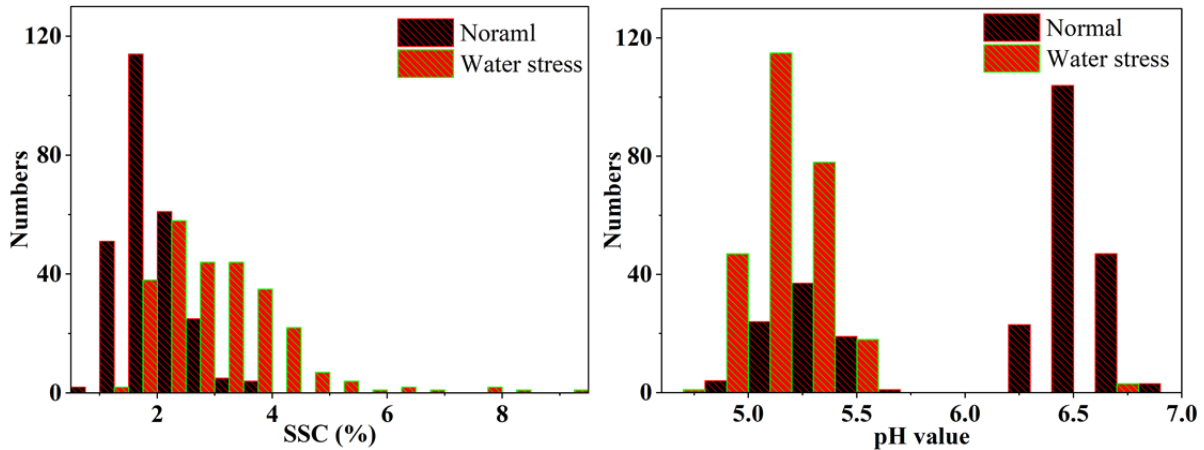


图 4.4 SSC 和 pH 值的参考值范围

Figure 4.4 The reference values range of SSC and pH value

4.3.2 不同水分处理的光谱反射率分析

图 4.5 中展现了群体生菜、正常水分生菜和水分胁迫生菜在可见光至近红外 (Vis-NIR) 光谱范围 (400-1000 nm) 的平均光谱反射率。植物光谱的特征细节丰富, 包含有关其生长状况的重要信息。可以观察到温室生菜和大田生菜光谱反射率的变化趋势是相似的。在可见光区域, 450 nm 与 700 nm 附近的波段与叶绿素的吸收特性相关联。受到叶绿素吸收效应较小的影响, 在这两个叶绿素吸收波段之间, 约在 580 nm 处形成了一个反射峰^[57, 66]。而在 760 nm 波段附近, 光谱反射率的急剧增加形成了红边现象, 这是植物光谱中最突出的特点^[44]。在近红外区域, 植物光谱的信号主要受其叶片内部结构的影响。而在 980 nm 附近的小吸收峰则与水的 O-H 键相联系^[134]。尽管群体生菜的光谱反射率趋势相似, 但它们的反射率值在不同样本之间却有所变化, 这意味着利用光谱数据来估算生菜的品质表型是可行的。

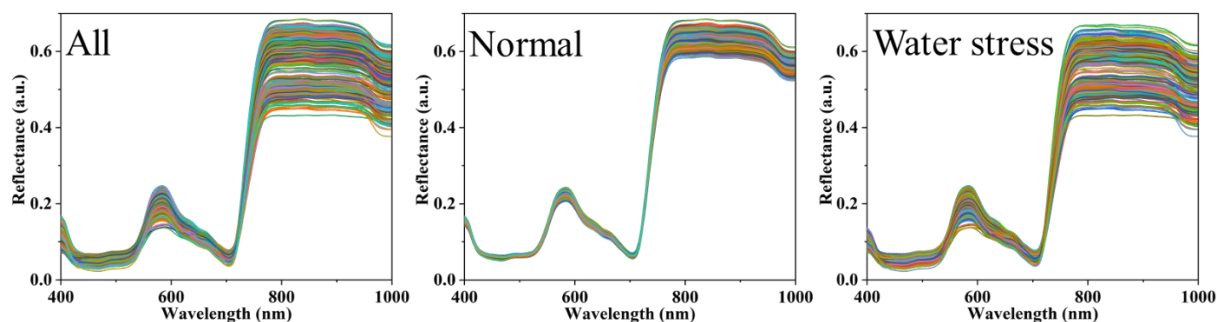


图 4.5 群体生菜、正常水分生菜和水分胁迫生菜的平均光谱反射率

Figure 4.5 Mean spectral reflectance of all lettuces, normal, and water stress

在经历了不同的水分处理之后，生菜反射率的变化展现出明显的差异，这一现象指示了光谱反射率在探测生菜水分胁迫状态方面的应用潜力。虽然两种水处理条件下的生菜光谱曲线大体呈现相同的趋势，但在水分胁迫处理下，生菜的光谱曲线相对于正常水分处理呈现出更为宽广的分布。特别是在近红外光谱区段，正常水分处理的生菜反射率值主要集中在 0.55 至 0.7 之间，而在水分胁迫条件下，反射率值则扩展至 0.4 至 0.7 的范围。在可见光区域的叶绿素吸收频带（580 nm 附近），也观察到了相似的分布趋势。这一现象揭示了在正常供水条件下，不同生菜品种的生长表现出一致性，而在水分胁迫条件下，则表现出显著的生长差异，这表明了光谱分析技术在鉴别抗水分胁迫生菜品种方面的潜在应用价值。

4.3.3 基于光谱的估算模型

（1）基于 Inception 模块的估算结果

基于先前研究结果，在进行神经网络结构优化时，特别关注了卷积层与全连接层的层数、卷积核的尺寸以及卷积核的数量。为了精准提高估算结果的准确度，网络结构会根据不同的品质表型进行适当调整。为了确保对所有模型进行有效的评估，在初步的研究中已经确定了验证集的比例、验证方法、训练的周期数以及批次大小，这些设置的详细信息见附录 B 表 B3。

通过使用基于 Inception 模块建立的深度学习模型，对生菜的 SSC 和 pH 值进行了估算。另外，还特别考虑了不同水分处理条件下生菜品质表型的变异性，分析了分层抽样方法对估算模型表现的影响。表 4.2 显示了生菜 SSC 和 pH 值估算的结果，而表 B3 则详细列出了各个品质表型估算模型的参数设置。

表 4.2 基于 Inception 模块估算生菜的 SSC 和 pH 值

Table 4.2 Estimation of the SSC and pH value in lettuces based on Inception module

Model	Phenotypes	Stratified sampling	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
Inception	SSC (%)	Yes	0.8866	0.3436	0.8968	0.2977	0.8307	0.4629	2.4301
		No	0.8643	0.4047	0.7982	0.4797	0.8279	0.4030	2.4105
	pH value	Yes	0.9290	0.1760	0.8785	0.2128	0.9116	0.1944	3.3642
		No	0.9253	0.1759	0.8665	0.2305	0.9079	0.2072	3.2954

估算结果显示，分层抽样方法显著提升了基于 Inception 模块建立的深度学习模型的估算准确性，因此在随后的模型分析中，都将采用分层抽样策略。在 pH 值的分析中，最佳结果为 $R_c^2 = 0.9290$ ， $RMSEC = 0.1760$ ， $R_v^2 = 0.8785$ ， $RMSEV = 0.2128$ ， $R_p^2 = 0.9116$ ， $RMSEP = 0.1944$ ， $RPD = 3.3642$ ，这明显表明基于 Inception 模块设计的深度学习模型在估算生菜的 pH 值上表现卓越。虽然 SSC 的估算精度相对较低，但其 R_p^2 均超过 0.8 且 RPD 超过 2，估算结果符合预期的误差范围。图 4.6 展示了基于 Inception 模块的估算散点图，直观地展现了校准集和预测集中 SSC 和 pH 值的最优估算结果，从估算点的分布来看，pH 值的数据点比 SSC 的数据点更加分散，这说明评价不同的表型特性时除了要观察 R^2 ，还需要根据每个表型特性的具体度量单位来分析其估算误差。

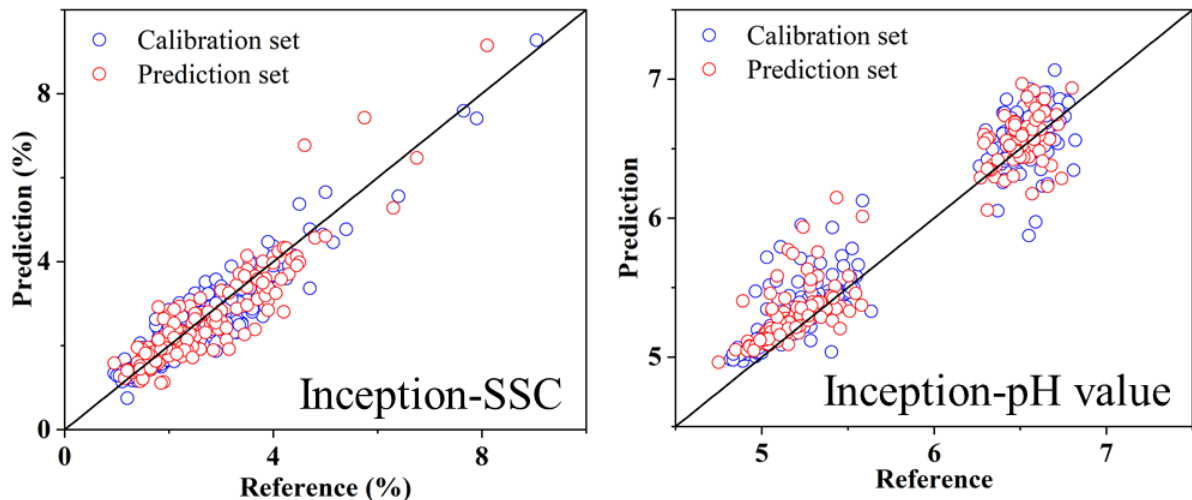


图 4.6 基于 Inception 模块的品质表型估算

Figure 4.6 Estimation of quality phenotypes based on Inception module

(2) 基于传统机器学习的估算结果

为深入探究基于 Inception 模块的深度学习模型表现，使用线性（PLSR）和非线性（SVR）的多元分析方法建立了机器学习模型来估算品质表型。早期研究表明，PLSR 和 SVR 能有效应用植物的光谱反射数据，对 SSC 和 pH 值等品质表型做出精确估算^[4]。

表 4.3 展现了 PLSR 和 SVR 估算 SSC 和 pH 值的结果，附录 B 表 B4 详细呈现了这

两种模型的参数设置。在估算两种品质表型时，SVR 模型的表现都比 PLSR 好，并且 PLSR 模型对 SSC 的估算有过拟合现象。其中，对于 SSC，获得的 R_c^2 为 0.8184，RMSEC 为 0.4672， R_v^2 为 0.8573，RMSEV 为 0.3501， R_p^2 为 0.8023，RMSEP 为 0.5599，RPD 为 2.0148。对于 pH 值，获得的 R_c^2 为 0.9180，RMSEC 为 0.1920， R_v^2 为 0.7115，RMSEV 为 0.3279， R_p^2 为 0.8753，RMSEP 为 0.2374，RPD 为 2.7625。相对于 PLSR 和 SVR 模型，基于 Inception 模块的估算结果在 SSC 和 pH 值分析方面更为理想。Rehman 等人对玉米叶片的相对含水量进行研究也得出了类似结论，即深度学习模型相对于 PLSR 和 SVR 拥有更佳的性能^[72]。在第三章的研究中，观察到 PLSR 在分析生菜品质时通常比 SVR 表现得更好，但本章的结果却与此不同。此外，Eshkabilov 等人的研究中，发现在估算生菜品质的准确性上，采用线性多元分析方法建立的模型胜过了非线性多元分析方法^[4]。这些结果揭示了机器学习模型在估算不同品质表型时泛化能力的不足，而深度学习方法则因其卓越的泛化能力而值得优先考虑。图 4.7 展示了基于机器学习对 SSC 和 pH 值的最优估算结果的估算散点图，相较于图 4.6，数据点更少聚集在拟合线附近。在后续的品质表型估算中选择使用基于 Inception 模块建立的深度学习模型。

表 4.3 基于机器学习估算生菜的 SSC 和 pH 值

Table 4.3 Estimation of the SSC and pH value in lettuces based on machine learning								
Phenotypes	Models	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
SSC (%)	PLSR	0.8931	0.3337	0.8969	0.2975	0.7913	0.5139	2.1954
	SVR	0.8184	0.4672	0.8573	0.3501	0.8023	0.5599	2.0148
pH value	PLSR	0.8037	0.2926	0.7800	0.2863	0.7920	0.2983	2.1989
	SVR	0.9180	0.1920	0.7115	0.3279	0.8753	0.2374	2.7625

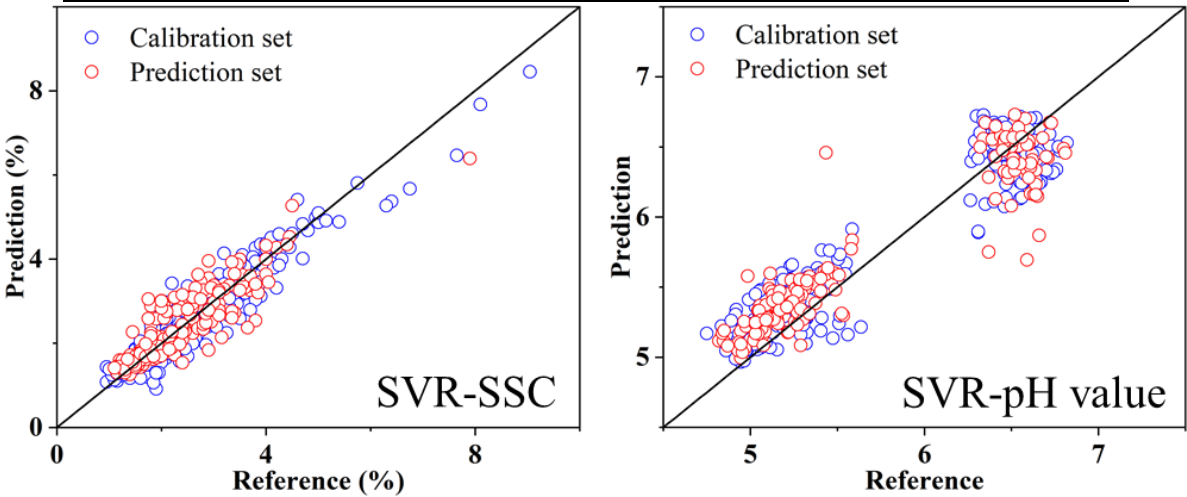


图 4.7 基于机器学习的品质表型估算

Figure 4.7 Estimation of quality phenotypes based on machine learning

(3) 基于 Inception、残差和注意力模块的估算结果

Inception 模块中网络宽度的扩展和卷积层的堆叠可能会导致特征冗余和梯度消失问题的出现。为了提升基于 Inception 模块建立的深度学习模型的表现，引入了残差模块和注意力模块进行优化。本章探究了 Inception 模块、残差模块（residual）和注意力模块（attention）的三种不同组合策略，即 Inception-residual、Inception-attention 和 Inception-residual-attention。

表 4.4 展示了利用残差模块和注意力模块进行 SSC 和 pH 值的估算结果。残差模块和注意力模块的具体参数详见附录 B 表 B5。在品质表型的估算任务中，由 Inception 模块、残差模块和注意力模块相互组合建立的模型性能均显著优于单一 Inception 模块。

表 4.4 Inception 模块与残差和注意力模块结合估算生菜的 SSC 和 pH 值

Table 4.4 Estimation of the SSC and pH value in lettuces combining Inception module with residual and attention module

Phenotypes	Models	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
SSC (%)	Inception-residual	0.8825	0.3497	0.9141	0.2715	0.8413	0.4481	2.5102
	Inception-attention	0.8911	0.3367	0.8923	0.3041	0.8694	0.4065	2.7674
	Inception-residual-attention	0.9027	0.3183	0.9064	0.2835	0.8743	0.3989	2.8202
pH value	Inception-residual	0.9565	0.1377	0.9383	0.1516	0.9472	0.1502	4.3535
	Inception-attention	0.9596	0.1327	0.9539	0.1310	0.9399	0.1603	4.0798
	Inception-residual-attention	0.9317	0.1726	0.9255	0.1666	0.9240	0.1804	3.6263

Inception-residual-attention 模型在 SSC 上取得了最佳的估算结果（ $R_c^2 = 0.9027$ ， $RMSEC = 0.3183$ ， $R_v^2 = 0.9064$ ， $RMSEV = 0.2835$ ， $R_p^2 = 0.8743$ ， $RMSEP = 0.3989$ ， $RPD = 2.8202$ ）。对于 pH 值的估算，Inception-residual 模型表现了最佳性能，其 R_c^2 为 0.9565， $RMSEC$ 为 0.1377， R_v^2 为 0.9383， $RMSEV$ 为 0.1516， R_p^2 为 0.9472， $RMSEP$ 为 0.1502， RPD 为 4.3535。这些结果表明残差模块和注意力模块的性能提升并不总是相互累积的。例如，在 pH 值的估算中，Inception 模块与残差模块和注意力模块两两组合的效果都超过了三种模块的组合，这可能是 Inception 模块和残差模块的组合并没有产生冗余的特征。这一发现与 Weng 等人使用深度学习模型及高光谱图像特征来识别赤霉病小麦中残差和注意力模块具有互补效应的结论不同^[74]。造成这种结果的原因可能是光谱数据和图像数据所揭示的关于植物状况的信息程度存在差异。因此，为了深入探究残差和注意力模块对深度学习模型的影响，值得尝试一种融合光谱数据和图像数据的方法，用于建立植物表型估算模型。同时，为了达到不同品质表型的最高估算准确度，Inception、残差和注意力模块的有效组合至关重要。在实际部署中，考虑到数据处理的时效性，将残差模块和注意力模块同时整合至 Inception 模块中是可行的，因为

这种组合的模型在估算精度上表现出色。在 HTPP 领域, 为了提升数据分析的速度, 适当牺牲一些精度是可以接受的^[141]。图 4.8 描绘了应用残差模块和注意力模块在校准集和预测集上得到的 SSC 和 pH 值的最佳估算效果。与图 4.6 和图 4.7 相较, 图 4.8 明显表明品质表型的估算误差有了显著降低, 尤其是在 SSC 数据点上的改进尤为突出。接下来的分析将采用这些基于光谱数据的最优深度学习模型架构。

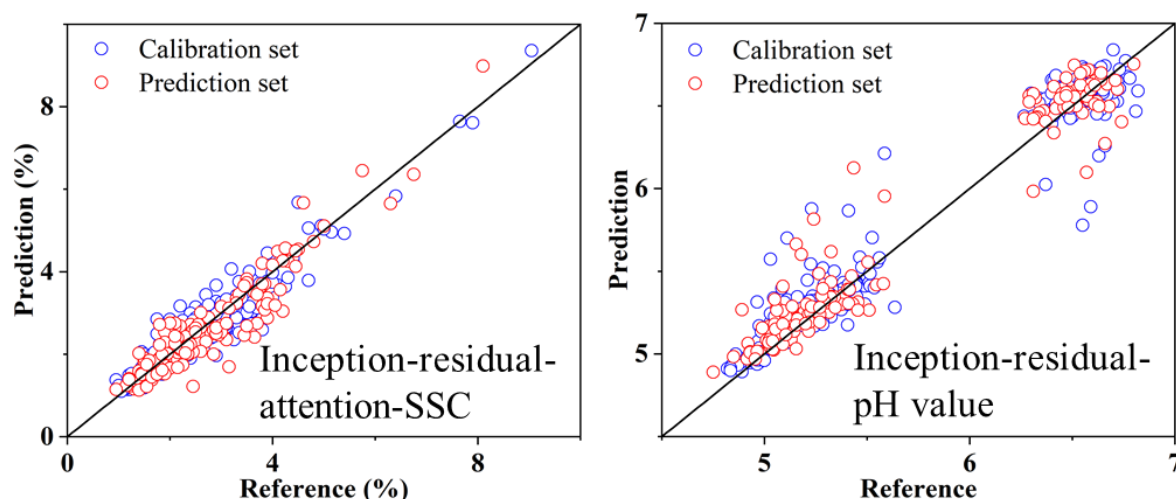


图 4.8 使用残差和注意力模块估算品质表型

Figure 4.8 Estimation of quality phenotypes using residual and attention module

4.3.4 不同水分处理的时间序列数据分析

植物的颜色和形态等生物物理表型与其内部品质表型之间有着密切的联系^[106]。在植物生长的过程中, 随着其颜色和形态特性的变化, 内部品质也会随之发生变化。为了深入探究颜色和形态表型对品质表型建模的影响, 研究中将这些表型的时间序列数据融入深度学习模型。在实验启动前的一周内, 在专业指导下对生菜的两种植品质表型进行了精确测定, 观察这些品质表型的日变化, 当品质表型在一段时间内展现出明显的一致性时, 开始实验。最终, 选取了生菜采收前五日的时间序列数据, 用于后续的建模分析。

在收集到的时间序列数据中, 各个表型的数值分布于各自的独特范围之内, 比如植物绿度的数值介于 0 到 1 之间, 而植株株高则在 100 到 700 mm 之间。直接使用原始的时间序列数据, 会显著影响基于 RNN 的深度学习模型的估算性能^[142]。所以, 在将这些时间序列数据输入 RNN 之前, 需要对这些数据进行归一化处理, 这是为了剔除数据中的异常值, 并确保不同维度上的表型数值呈现出相对的稳定性。

随着生菜的成长, 其叶子的颜色和形态包括叶面积与株高, 均呈现出明显的变化。

在 HTPP 平台获得的表型中, 生菜的绿度能够反映其颜色信息, 而数字生物量则与叶面积和株高紧密相关。图 4.9 展示了经过两种不同水分处理的 20 个生菜样本, 在五天内时间内的数字生物量、绿度和株高的变化趋势。在不同的水分处理条件下, 时间序列的变化趋势表现出了不一致性, 相同水分处理下不同样本的时间序列也显示出差异。在正常水分处理下, 生菜的数字生物量和株高基本呈现出逐渐增加的趋势, 这通常是因为生菜得到了适宜的水分供给, 有利于其正常生长和发育, 而其绿度则在第五天前都呈下降趋势, 可能是因为随着生菜叶片的成熟, 叶绿素分解速度可能超过了新的叶绿素的合成速度, 在第五天绿度上升, 可能是由于外界光照强度的变化导致。在水分胁迫处理下, 生菜绿度在第二天达到高峰, 可能是因为植物为了适应水分胁迫环境, 会在短期内增加叶绿体数量或叶绿素含量, 以提高光合作用效率, 作为一种应对初始胁迫的机制, 而大部分样本的数字生物量与株高则维持在一个较为稳定的区间内, 这可能是随着水分胁迫的持续, 植物可能会采取保守策略, 减少生长速率, 导致其叶面积和株高可以维持在较稳定的状态。

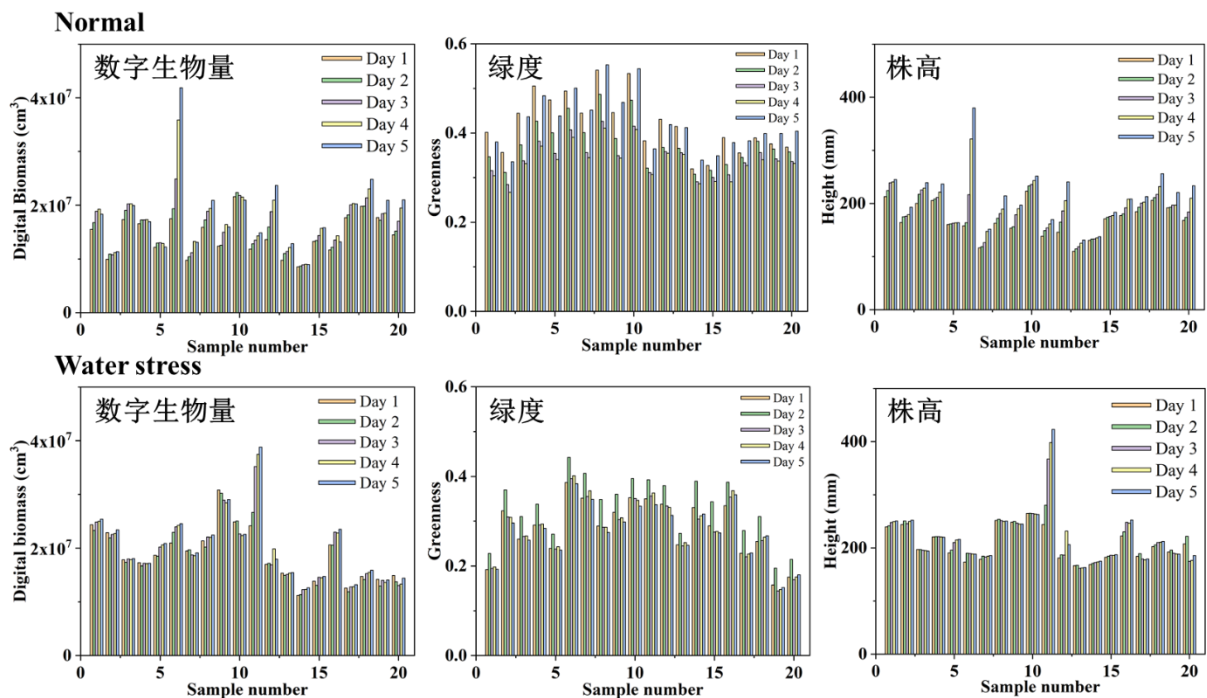


图 4.9 不同水分处理的时间序列数据

Figure 4.9 Time-series data of different water treatment

4.3.5 基于光谱和时间序列融合的估算模型

(1) 品质表型的估算

相较于传统的前馈网络结构如 CNN 与 FCNN, RNN 在处理时间序列数据方面的表

现通常更为卓越^[99]。在本章中，将时间序列数据输入到各类 RNN 模型，包含 TD、LSTM、GRU 以及 BRNN。这些模型与基于光谱数据所训练出的最优深度学习模型（Optimal-Spec）进行了结合，目的是为了更精准地估算生菜的品质表型。

表 4.5 展示了利用时间序列数据估算 SSC 和 pH 值的结果，而附录 B 表 B6 则提供了 RNN 的详细参数设置。在 SSC 的分析中，四种 RNN 模型与 Optimal-Spec 模型的结合显著提升了估算准确性。具体而言，Optimal-Spec-TD 模型在 SSC 的估算上取得了最优效果（ $R_c^2 = 0.9042$ ，RMSEC = 0.3158， $R_v^2 = 0.8723$ ，RMSEV = 0.3312， $R_p^2 = 0.8900$ ，RMSEP = 0.3731，RPD = 3.0153）。在利用时间序列数据及 TD 的情况下，SSC 估算模型的 RPD 值超过了 3。在 pH 值的估算中，除 TD 外，其他三个 RNN 模型均未能提升 Optimal-Spec 模型的表现，其中 Optimal-Spec-TD 模型在 pH 值的估算精度上表现卓越，其 R_c^2 为 0.9637，RMSEC 为 0.1258， R_v^2 为 0.9551，RMSEV 为 0.1293， R_p^2 为 0.9583，RMSEP 为 0.1336，RPD 为 4.8971。这可能是因为 pH 值的变化规律较为直接，与时间序列数据的关联不像 SSC 那样需要复杂的时间依赖建模，即不需要 LSTM、GRU 和 BRNN 在时间步骤之间传递状态，所以这三种 RNN 模型没有给 Optimal-Spec 模型带来提升。总体来说，TD 在品质表型的估算中比其他三种 RNN 表现得更好。另外，在估算 SSC 和 pH 值时，Inception 模块与残差和注意力模块的结合，较之时间序列数据和 RNN，更能有效提升模型的整体性能。这一现象可能说明光谱数据对品质表型估算的影响更为显著，未来的研究可以考虑通过改变时间序列数据的时间分辨率来进一步观察和分析模型性能的变化。图 4.10 显示了 RNN 模型在校准集和预测集上估算 SSC 和 pH 值的最优结果。与图 4.8 相比，图 4.10 中 SSC 和 pH 值的数据点变化都较小。这一趋势与实验中各估算模型性能提升的实际观测结果相符。

表 4.5 使用光谱和时间序列估算生菜的 SSC 和 pH 值

Table 4.5 Estimation of the SSC and pH value in lettuces using spectra and time-series

Phenotypes	Models	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
SSC (%)	Optimal-Spec-TD	0.9042	0.3158	0.8723	0.3312	0.8900	0.3731	3.0153
	Optimal-Spec-LSTM	0.8867	0.3435	0.8803	0.3206	0.8764	0.3955	2.8440
	Optimal-Spec-GRU	0.8788	0.3552	0.8562	0.3514	0.8772	0.3941	2.8542
	Optimal-Spec-BRNN	0.8843	0.3471	0.8639	0.3418	0.8828	0.3851	2.9210
pH value	Optimal-Spec-TD	0.9637	0.1258	0.9551	0.1293	0.9583	0.1336	4.8971
	Optimal-Spec-LSTM	0.9646	0.1243	0.9366	0.1537	0.9420	0.1575	4.1534
	Optimal-Spec-GRU	0.9626	0.1277	0.9516	0.1343	0.9404	0.1597	4.0948
	Optimal-Spec-BRNN	0.9756	0.1031	0.9675	0.1101	0.9464	0.1514	4.3199

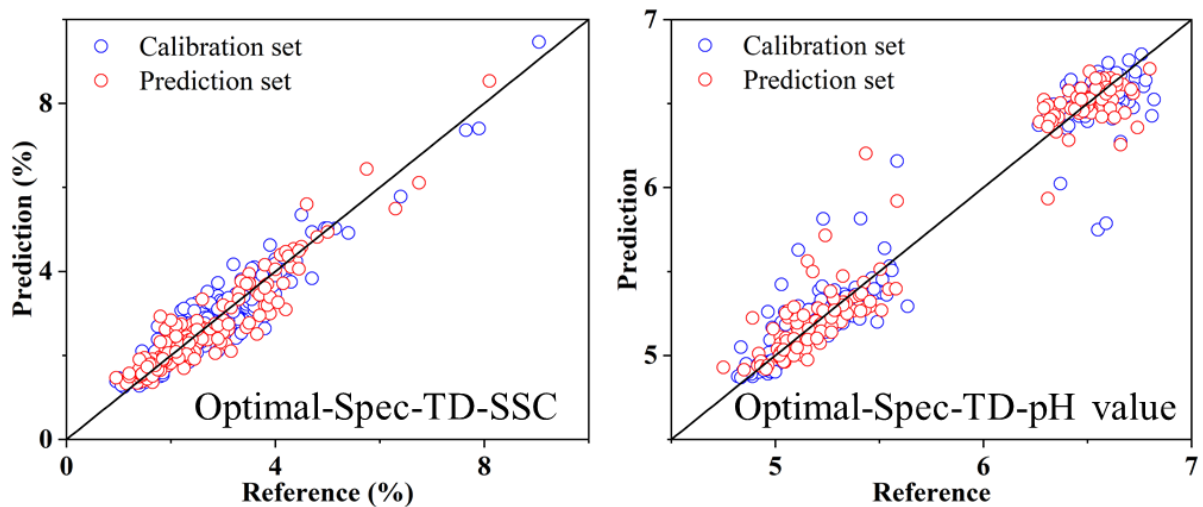


图 4.10 使用 RNN 估算品质表型

Figure 4.10 Estimation of quality phenotypes using RNN

(2) 水分胁迫的估算和可视化

在不同的水分处理下，品质表型、光谱反射率以及时间序列都表现出明显的差异。这说明基于生菜的颜色、形态、光谱特征以及品质表型来探测水分胁迫具有可行性。在所研究的品质表型中，pH 值在不同的水分处理条件下表现出最为显著的差异，并且其估算结果也最为精确。所以，尝试了两种 pH 估算模型（Optimal-Spec 模型和 Optimal-Spec-TD 模型）来估算水分胁迫的状态，并通过观察 pH 值的变化来判断水分胁迫迹象的发生。

为了评估水分胁迫检测模型的性能，采用了混淆矩阵和准确率指标^[74]。在模型训练过程中，校准集、验证集和预测集的划分方式与估算模型的划分相同。水分胁迫检测结果见表 4.6 和图 4.11。对于正常水分处理，两种模型的检测准确率都达到了 100% ($\text{Accuracy}_c = 100\%$, $\text{Accuracy}_v = 100\%$, $\text{Accuracy}_p = 100\%$)。在水分胁迫处理的情况下，Optimal-Spec-TD 模型展现了更加出色的性能，准确率为 $\text{Accuracy}_c = 98.73\%$,

$Accuracy_v = 100\%$, $Accuracy_p = 98.86\%$ 。这与 Optimal-Spec 模型和 Optimal-Spec-TD 模型在估算 pH 值时的表现是一致的。图 4.11 揭示了 Optimal-Spec-TD 模型将正常水分处理误认为水分胁迫处理的具体情况，在预测集中有一个样本，而在校准集中有两个样本发生了此类误判。从光谱的角度分析，发生误判的原因可能是正常水分处理的光谱分布范围有一部分是包含在水分胁迫处理的范围之内。

表 4.6 基于 Optimal-Spec 和 Optimal-Spec-TD 模型的生菜水分胁迫分类

Table 4.6 Classification of water stress of lettuces based on Optimal-Spec and Optimal-Spec-TD

		model		
Models	Class	Accuracy _c	Accuracy _v	Accuracy _p
Optimal-Spec	Normal	100%	100%	100%
	Water stress	96.82%	100%	96.59%
Optimal-Spec-TD	Normal	100%	100%	100%
	Water stress	98.73%	100%	98.86%

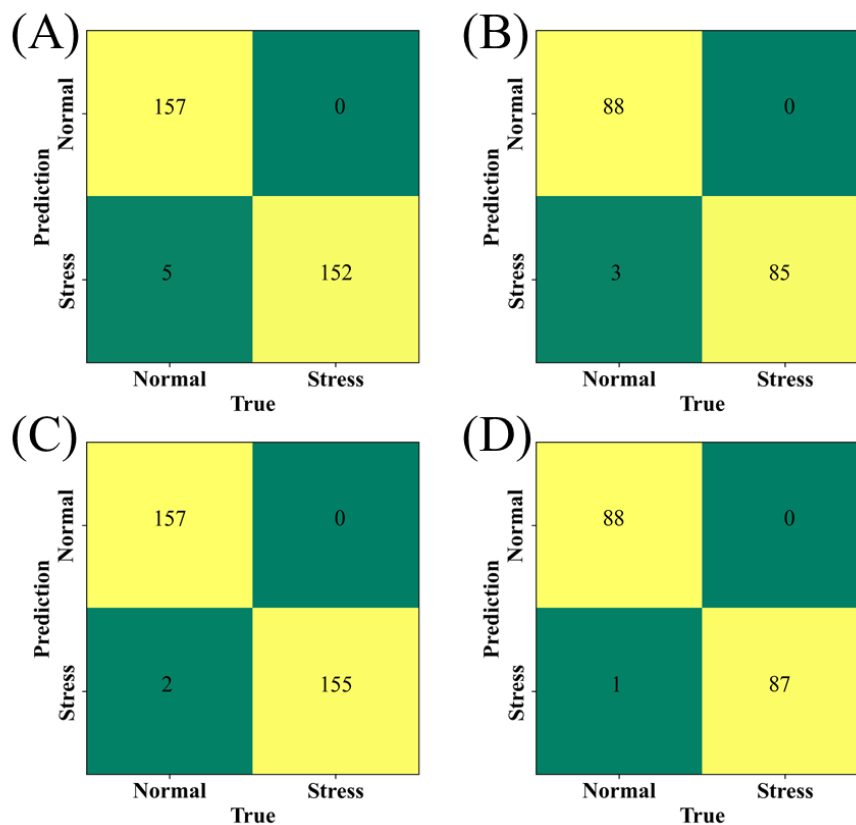


图 4.11 正常水分和水分胁迫生菜的混淆矩阵。Optimal-Spec 模型的校准集 (A) 和预测集 (B) 和 Optimal-Spec-TD 模型的校准集 (C) 和预测集 (D)

Figure 4.11 Confusion matrix of lettuces of normal and water stress. Calibration set (A) and prediction set (B) of Optimal-Spec model, and calibration set (C) and prediction set (D) of Optimal-Spec-TD model

生菜植株不同叶片对水分胁迫的响应各异。在实验中生菜受到水分胁迫后，观察到冠层周边的叶片最先干枯。说明生菜冠层不同部位叶片对水分胁迫的反应各异，进而表明水分胁迫程度能够更进一步地进行分级，这有助于早期识别水分胁迫，以尽可能降低对作物品质的不良影响。在两种水分处理下，生菜的 pH 值的变化最为显著，因此本章研究了估算生菜冠层不同部位 pH 值的可能性，并以此作为间接评估水分胁迫程度的方法。鉴于时间序列数据是基于单株生菜个体尺度获取的，而高光谱图像的平均光谱则是基于像素光谱获得的，这意味着时间序列数据不能反映生菜冠层在不同部位的品质表型。所以，应用了基于光谱数据建立的 Optimal-Spec 模型来估算生菜高光谱图像中每个像素点的 pH 值^[143-151]。图 4.12 显示了在两种水分处理条件下生菜的 pH 值分布情况。从图中可见，经历水分胁迫的生菜呈现较冷的色调（pH 值低），而正常水分处理的生菜则呈现出较暖的色调（pH 值高）。另外，生菜冠层外围的 pH 值较低，这可以与水分胁迫条件下生菜的外观变化相关联，即外层叶片率先出现枯萎现象。因此，通过对 pH 值的监测，或许能够在早期及时发现生菜所遭受的水分胁迫。一旦观察到 pH 值的降低，这可能预示着水分胁迫的发生。Yan 等人使用膳食纤维来评估芹菜的新鲜度，并通过膳食纤维分布图来展示新鲜度水平^[143]，这进一步证明了基于高光谱数据建立的分布图，可以作为一种非破坏性品质检测的有效辅助手段。

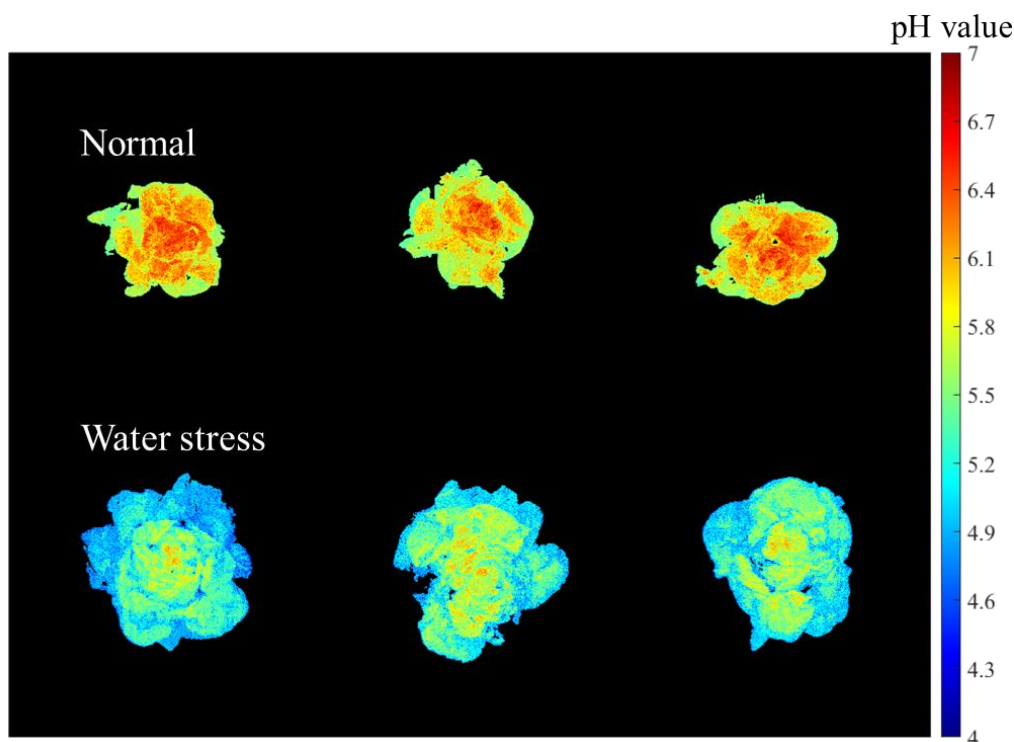


图 4.12 pH 值的分布图

Figure 4.12 The distributed map of pH values

4.4 本章小结

本章针对生菜在水分胁迫条件下品质表型特性的估算进行了研究。与传统的测量手段与单一数据模态相比，本章提出的基于光谱和时间序列数据的深度学习方法展现了无损、快速且准确估算生菜品质表型特性和水分胁迫状态的潜力。虽然高光谱技术已被先前研究用于监测生菜的品质和胁迫状态，但直接应用于商业栽培仍然面临挑战，特别是在样本代表性和模型泛化能力方面。植物的颜色和形态等外观特征与其内部品质密切相关，并在生长过程中展现出一系列变化，这些变化反映了植物品质的动态演进。尽管利用时间序列数据估算水分胁迫条件下的生菜品质表型特性研究仍较少，但 HTPP 平台集成的多种传感器开启了估算植物品质表型和辨别抗逆性品种的新途径。本章通过运用 RNN 深入挖掘时间序列数据中的信息，作为光谱数据估算植物品质表型特性的有效补充。

本章首先深入探讨了通过残差和注意力模块的组合对 SpecNet2D 模型（基于 Inception）在光谱数据特征提取方面的影响，强调了针对不同表型特性选择最优模块组合的重要性。研究发现，Inception-residual-attention 获得了最优的 SSC 估算结果（ $R_c^2 = 0.9027$, $RMSEC = 0.3183$, $R_v^2 = 0.9064$, $RMSEV = 0.2835$, $R_p^2 = 0.8743$, $RMSEP = 0.3989$, $RPD = 2.8202$ ），而 Inception-residual 在 pH 值的估算中表现最好（ $R_c^2 = 0.9565$, $RMSEC = 0.1377$, $R_v^2 = 0.9383$, $RMSEV = 0.1516$, $R_p^2 = 0.9472$, $RMSEP = 0.1502$, $RPD = 4.3535$ ）。

然后，通过在光谱数据基础上引入时间序列数据，并运用四种不同的 RNN 技术处理时间序列数据，构建了 SpecTimeNet 模型。该模型可以通过细致的时间序列特征提取，进一步提高生菜 SSC 和 pH 值估算的准确度，并探索各种 RNN 技术的有效性。实验结果显示，SpecTimeNet 模型显著提升了估算精度。在 SSC 估算中，Optimal-Spec-TD（Inception-residual-attention-TD）模型将 R_p^2 提升至 0.8900，而在 pH 值的估算方面，Optimal-Spec-TD（Inception-residual-TD）模型以 0.9583 的 R_p^2 达到了最佳性能。此外，为了进一步探索水分胁迫对生菜品质的影响，采用了 Optimal-Spec 和 Optimal-Spec-TD 模型针对 pH 值进行估算，分别在预测集上达到了 96.59%和 98.86%的准确率。基于 Optimal-Spec 模型和像素级光谱技术生成的 pH 值空间分布图，展示了其在早期检测生菜水分胁迫方面的潜力。这些发现不仅验证了深度学习技术与多模态数据融合在生菜种植业中的应用价值，而且在精确估算植物表型特性、评估水分胁迫状况及筛选抗逆

性生菜品种方面展现了显著的潜力。

未来研究可进一步优化实验设计，如增加水分处理梯度和调整时间序列数据的时间分辨率，以更细致地观察模型效果。还可以将模型应用于更广泛的植物抗逆性品种识别，为精准农业和植物表型分析提供更全面的技术支持。

第五章 基于光谱和图像的生菜光合效能表型指标的估算

5.1 引言

在精准农业和植物表型分析领域，快速且准确地估算作物的光合效能表型，如通过叶绿素含量指示的 SPAD 值，对于提升农作物管理效率和产量具有关键意义^[7]。尽管传统测定方法在特定条件下有效，但它们在大规模和实时数据采集方面存在明显局限，难以满足现代农业生产的需求。

继第四章对光谱与时间序列数据融合在提高品质表型估算精度的成功探索后，本章转向 SPAD 值的研究，聚焦于光谱与图像数据融合的潜力。选择这两种数据模态是因为叶绿素含量既在光谱数据中直接体现，也能通过图像数据间接展示其变化。因此，本章旨在通过深入研究光谱与图像的融合效果，探索更多数据模态对植物表型特性估算的贡献，目的是提升估算模型的准确度和应用范围。这不仅建立在前章工作的基础之上，更拓宽了对多模态数据融合技术在精准农业中应用的理解和实践，为接下来的章节探索更复杂的多模态数据融合方式提供了思路。

本章中，采用了基于不同尺度的图像特征提取技术处理图像数据，同时，利用经过优化的 SpecNet2D 模型来处理光谱数据，进而将两者的特征结合起来，以期达到对 SPAD 值更准确的估算。此过程中，通过比较不同图像处理策略对模型性能的影响，以筛选出最高效的图像特征提取方法。同时，引入 SR 技术以提升高光谱图像的空间分辨率，进一步丰富了图像数据中的植物表型信息，再通过 CAM 技术可视化的方法对光谱数据中的关键区域进行了细致的探索与选取，以优化特征提取过程。最终，评估这些方法对于光谱和图像融合模型性能的具体影响。

此外，本研究通过扩展到七个生菜品种，进一步探讨了光谱与图像融合技术在不同遗传背景下的普适性和准确性。同时，深度学习模型受益于更丰富的样本量，不仅有助于增强模型的泛化能力，也为筛选具备高光合效能的生菜品种提供了一个有效的方法，从而推动精准农业技术在品种优化和选择方面的应用。

本章聚焦于光谱与图像数据的融合方法，旨在深入理解作物光合效能的细微变化，并推动更为精确高效的植物表型估算技术的发展。此种方法的应用，不仅扩展了植物表型估算的技术手段，而且对于优化作物管理具有重要意义。

5.2 数据获取与处理

图 5.1 展示了本章中实验设计与数据处理的完整流程图。实验数据的采集包括获取 SPAD 值的参考值以及传感器采集的数据。数据处理环节涉及对高光谱图像和 RGB 图像的处理，最终将处理后的数据输入到模型中，用以估算 SPAD 值。具体的实验步骤和数据处理方法将在本节的后文部分进行详细阐述。

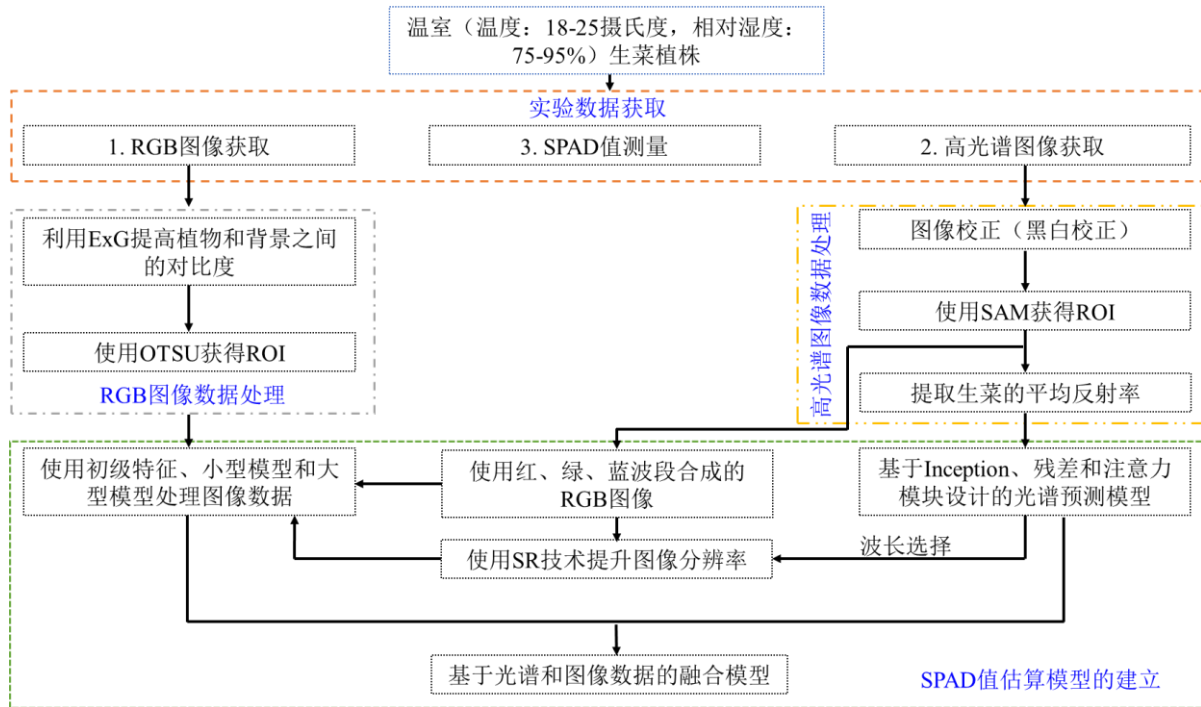


图 5.1 估算生菜的 SPAD 值的实验流程图

Figure 5.1 Flowchart of the experiment used to estimate SPAD value in lettuces

5.2.1 实验设计

2022 年 12 月 1 日在北京市农林科学院的温室中种植了七个生菜品种的幼苗，分别是奶油生菜（Butter）、结球生菜（Crisphead）、散叶生菜（Leaf）、橡叶生菜（Oak leaf）、罗马生菜（Roman）、茎用生菜（Stem）以及野生近缘种生菜（Wild relatives）。实验现场部署了流水线成像室系统。生菜在花盆中种植。在专业监督下，生菜在种植后每 7 到 10 天浇一次水，种植过程中都不施肥。实验选择了 365 棵奶油生菜、394 棵结球生菜、370 棵散叶生菜、49 颗橡叶生菜、507 棵罗马生菜、83 颗茎用生菜和 118 颗野生近缘种生菜，共 1886 颗生菜样本。实验分别在 2023 年 1 月 15 日、16 日、17 日、18 日和 19 日进行，每天采样的生菜由相关专家根据生菜的实际长势获取，每天获取生菜样本约为 380 颗。获取实验数据的步骤如下：首先，获取生菜的 RGB 图像数据，然

后获取生菜的高光谱图像，最后使用叶绿素计（SPAD-502 plus, Konica Minolta, 日本）测量生菜的 SPAD 值，测量 SPAD 值时选择生菜冠层最外层，中间层，最里层各 3 片叶子，共 9 片叶子的 SPAD 值的平均值。叶绿素计实物图见附录 C 图 C1。

本章使用流水线成像系统来获取高光谱数据。高光谱图像具体的获取方法见 2.2.1 节，光谱反射率提取过程见 2.2.2 节。

5.2.2 RGB 图像获取与处理

在处理 RGB 图像以提取 ROI 时，使用了一种结合超绿特征（Excess Green, ExG）与最大类间方差法（OTSU 算法）的策略^[95]。ExG 作为一种高效的植物绿色区域突显技术，通过增加绿色通道的权重并降低红色与蓝色通道的权重，有效地在 RGB 图像中高亮植物部分。此方法不仅加强了图像中植物绿色部分的视觉效果，而且提升了其与背景的对比度，为后续处理奠定了基础。另一方面，OTSU 算法通过统计学方法自动寻找图像的最优二值化阈值，可以将图像像素划分为前景（目标）与背景两大类，实现最大化两类间方差，从而获得出色的分割效果。该算法无需手动设置阈值，能够适应多样的图像环境，并在自然背景下精确提取植物区域。

将 ExG 与 OTSU 算法相结合，能够充分利用两者的优势：ExG 对植物绿色部分的增强作用为 OTSU 算法提供了更为明显的前景与背景对比，而 OTSU 算法的自动阈值选择功能则能高效地从背景中分离出增强后的植物区域。这种方法显著提升了 ROI 提取的精确度与自动化程度，对于进一步的表型分析具有重要的意义。

5.2.3 高光谱图像的超分辨率重建

SRGAN 为图像分辨率提升领域带来了一种有效的解决方案，特别是在处理高光谱图像时，其能力变得尤为重要。高光谱图像虽然在光谱分辨率上具有优势，提供了丰富的光谱信息，但其空间分辨率较低，这限制了其在需要丰富空间细节的应用场景中的有效性。通过应用 SRGAN，以有效提升高光谱图像的空间分辨率。SRGAN 原理图见图 5.2。

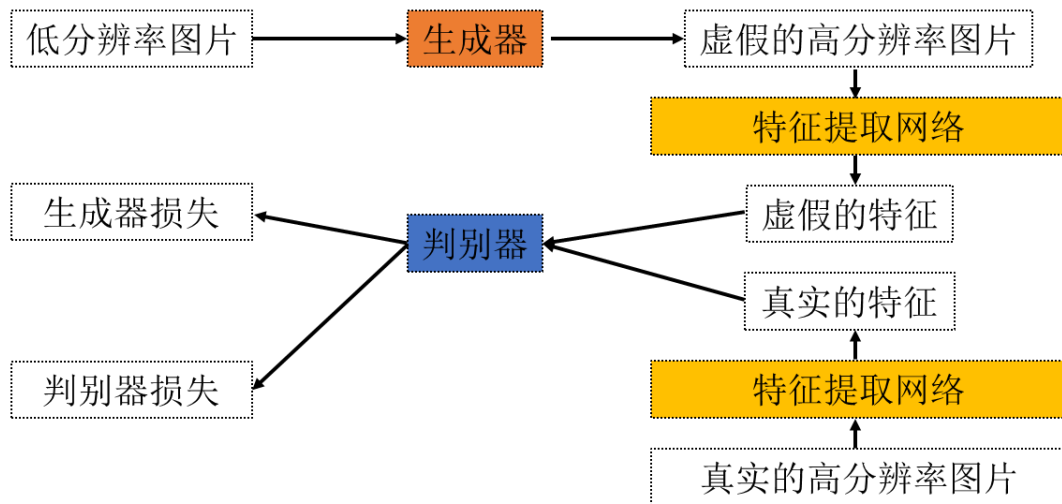


图 5.2 SRGAN 原理图

Figure 5.2 Schematic diagram of SRGAN

SRGAN 通过生成器和判别器的对抗过程，能够从低分辨率图像生成视觉上具有高质量的高分辨率图像。在此框架中，生成器负责产生尽可能接近真实高分辨率图像的输出，而判别器则尝试区分生成的图像与实际的高分辨率图像。SRGAN 的生成器和判别器的结构图见图 5.3。这一过程的核心在于通过不断的迭代训练，使生成器能够最终产生高质量的超分辨率图像。生成器和判别器的参数设置见附录 C 表 C1。

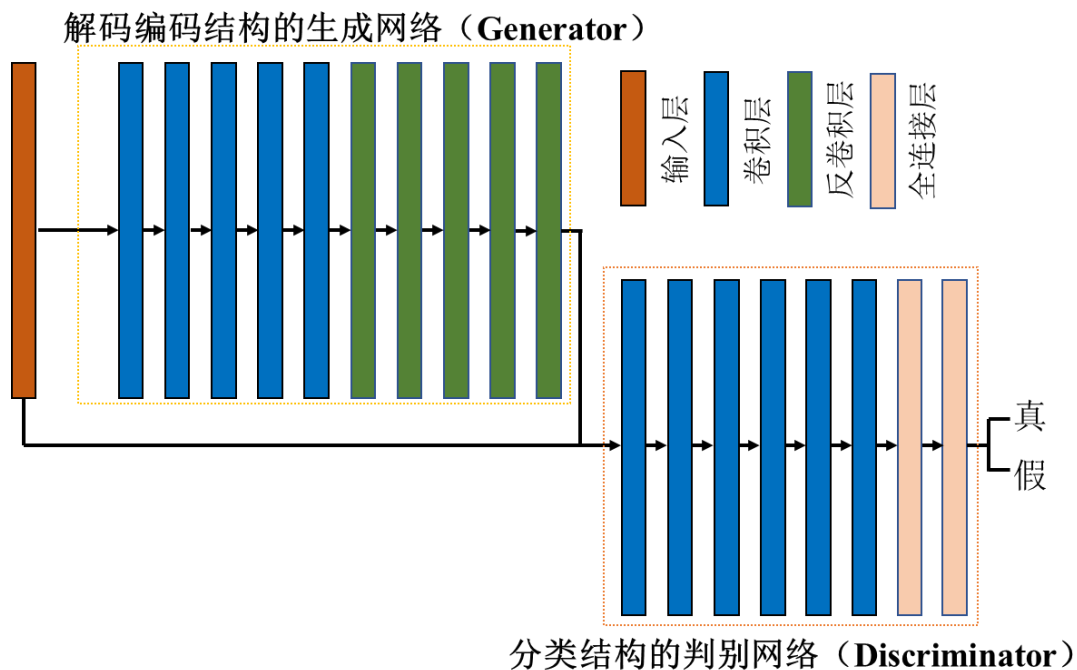


图 5.3 SRGAN 生成器和判别器的结构图

Figure 5.3 Structure diagram of generator and discriminator of SRGAN

为了进一步提升 SRGAN 在高光谱图像超分辨率任务中的性能，比较了四种不同

的特征提取网络：参数量较大的 VGG16^[94] 和 VGG19^[153]，以及轻量级网络 MobileNet^[154] 和 EfficientNet^[155]。VGG16 和 VGG19 作为图像处理领域的经典模型，以其深层结构和强大的特征提取能力著称，但同时也伴随着较高的计算成本。相比之下，MobileNet 和 EfficientNet 通过优化设计，在保持高效图像处理能力的同时减少计算资源的消耗，使其特别适合于资源受限的环境。在 SRGAN 中，特征提取网络扮演了重要的角色，它负责从图像中提取出有助于构建高分辨率图像的关键特征。这些特征对于 SRGAN 在执行图像超分辨率任务中的表现及最终生成图像的质量具有直接影响。通常，网络底层提取的特征倾向于包含更多局部细节的信息，而网络顶层则倾向于提取具有全局意义和高级语义的特征^[71]。鉴于 SRGAN 的目标是将低分辨率图像转换为具有丰富细节和高清晰度的高分辨率图像，因此模型需要同时考虑图像的局部细节和全局上下文信息。所以，选择特征提取网络中间层的输出作为 SRGAN 的特征提取基础，以实现图像全局语义信息和局部细节信息的有效平衡。这种方法使得 SRGAN 模型不仅重建了图像的视觉内容，同时保留了关键的细节信息，从而生成了视觉上更为清晰的高分辨率图像。通过此设计，能够确保 SRGAN 在提升图像分辨率的同时，兼顾了图像的整体语义连贯性和细节的精确再现，进一步提升了模型在图像超分辨率领域的性能和应用价值。特征提取网络的参数设置见附录 C 表 C1。

为了与基于深度学习的 SRGAN 进行对比，本章还探索了传统的图像插值（image interpolation）方法在提升图像空间分辨率方面的能力，细致分析了四种广泛应用的插值算法：三次插值（Cubic）、线性插值（Linear）、最近邻插值（Nearest）和 Lanczos 插值^[84-89]。这些方法在处理空间分辨率低的图像时，展现出了各自的特点和适用性。Cubic 方法通过建立三次多项式模型并考虑目标像素周围的 16 个像素点，能够在放大图像时提供平滑且连续的视觉效果，适合需要平衡图像细节与平滑度的应用。Linear 方法虽然计算过程简单，但它依靠最近的两个像素点进行直线估计，可能导致插值结果较为锐化，适用于对计算效率有较高要求的场景。Nearest 方法作为基础的插值策略，直接采用最邻近的像素值，尽管其处理速度快，但可能引入较明显的锯齿效应，适合对处理速度要求远大于图像质量的场合。Lanczos 方法通过广泛考虑像素邻域并应用特定窗口函数进行处理，能够有效维持图像的锐度，减少放大过程中的伪影，特别适用于对图像质量要求较高的应用。

本章的目标是通过对比图像插值方法与 SRGAN 的性能，评估它们在提升高光谱图像空间分辨率上的有效性。高光谱成像以其在光谱信息方面的丰富性，在多个领域

得到了广泛的应用，但其较低的空间分辨率限制了其进一步的应用潜力。通过对比分析，不仅揭示了各插值技术在特定应用场景下的优势与局限，而且深入了解了基于深度学习的 SRGAN 在显著提升图像质量方面的强大能力，为未来的高光谱图像处理技术的发展提供了宝贵的见解和指导。

在研究的初期阶段，用于后续分析的高光谱图像所选定波长的光谱特征是位于红、绿、蓝三个颜色波段范围内，目的是通过合成的 RGB 图像进行分辨率提升，以便与传统 RGB 相机拍摄的图像进行对比。Specview 软件可以自动从生成的高光谱图像中挑选波长图像合成 RGB 图像，选定红、绿、蓝波长分别为 667 nm、570 nm、450 nm。这种方法的选择基于将高光谱图像处理技术与常规的 RGB 图像处理技术进行直接比较的需求，从而评估超分辨率技术在提升经典 RGB 图像质量方面的能力。

随后，在进一步的研究中，通过对高光谱数据的深入分析，特别是基于光谱特征选择（CARS、UVE 和 CAM）的研究，选择了三个关键波长特征对应的高光谱图像进行进一步的实验。这种选择旨在合成一幅伪彩色图像，该图像能够突出特定的物理或生物特征，如植被的健康状况或水分含量，这些特征在传统 RGB 图像中可能不那么明显。CARS 主要应用于光谱数据分析，如 NIR 光谱分析，用于快速识别和量化样品中的化合物。UVE 是一种用于消除无信息变量的技术，常用于光谱分析和化学计量学中。该方法通过建立模型并评估各个变量的重要性来识别和排除那些对模型几乎没有贡献的变量。UVE 能有效降低模型的复杂度，提高估算性能和解释性。它特别适用于变量数量远大于样本数量的情况。CAM 用于深度学习中，特别是在 CNN 中进行图像识别和分类任务时，来解释模型决策的依据。通过 CAM，可以生成一张热力图来直观显示模型在做出决策时，输入数据所起到关键作用的区域^[75-78]。这对于提高模型透明度、理解模型如何工作以及进一步改进模型都非常有帮助。

使用基于 RGB 图像训练的 SRGAN 来提升伪彩色图像分辨率的有效性，代表了一种探索性尝试。虽然 SRGAN 最初是针对 RGB 图像设计和训练的，但其核心是学习图像从低分辨率到高分辨率的映射关系。这种学习过程并不严格限制于特定的颜色波段。因为 SRGAN 学习的是图像纹理、边缘和细节等特征的提升方法，这些特征在伪彩色图像中同样存在。因此，尽管伪彩色图像在波长特征的组成上与训练时的 RGB 图像不同，但 SRGAN 仍有潜力通过其学到的超分辨率映射关系来提升伪彩色图像的分辨率。此外，这种方法也提供了一种探索深度学习模型跨波长特征泛化能力的机会，为将来的研究开辟了新的道路。

总而言之，通过挑选关键的波长特征对应的高光谱图像进行图像合成和分辨率提升，不仅展示了超分辨率技术在传统 RGB 图像与高光谱伪彩色图像上的应用潜力，还探索了深度学习模型在处理多波长特征图像方面的广泛适用性，为高光谱图像的分析和应用提供了新的视角和方法。

5.2.4 光谱和图像的融合方法

基于第四章的研究结果，寻找 Inception、残差和注意力模块的最佳组合方式建立深度学习模型来处理光谱数据。

在图像处理和分析的领域内，特征提取的首要目标之一便是确保所提取特征的稳定性，不论是在噪声、尺度、旋转，还是光照变化的条件下。这些不变性的保障，使得特征能够在多变的环境与条件中保持一致性，从而显著提升了图像分析任务的准确性及其鲁棒性^[156]。具体来说，噪声不变性确保了在图像受到随机噪声干扰时，仍能准确地提取特征信息。尺度不变性意味着特征提取过程能够适应图像大小的变化，保证了特征的一致性，从而支持了图像之间的比较和识别。旋转不变性确保图像即便经过旋转处理，仍可准确识别其特征，这对于形状分析尤为关键。光照不变性则克服了光照条件变化对特征提取的影响，确保了在不同光照条件下提取到的特征依然稳定。

在本章中，选择了九种初级图像特征（primary feature）来捕捉图像信息。颜色特征，包括颜色直方图和颜色矩，虽然对光照变化较敏感，但通过综合应用，能够详尽描述颜色的分布与变化，为颜色分析提供了丰富维度的信息。纹理特征，如灰度共生矩阵、Laws 纹理能量测量法、Gabor 滤波器以及局部二值模式（LBP），尽管需注意噪声影响，但其能够从多角度揭示图像的纹理结构。特别是，LBP 为纹理特征提供了光照不变性的分析，增强了特征的鲁棒性。形态特征，包括 Hu 矩、周长及面积，通过描绘图像的形状信息，提供了尺度和旋转的不变性，其中 Hu 矩为形状识别提供了稳定的描述子。所挑选的初级图像特征共同组成了一个全方位捕捉图像颜色、纹理和形态信息的框架。在选择这些特征时，充分考虑它们在面对噪声、尺度、旋转以及光照变化时的表现，确保了特征的稳定性和可靠性。这种图像特征选择策略体现了图像分析中特征不变性的重要性，有助于后续图像数据分析任务的开展。

为了深入评估初级图像特征提取方法的效果，采用了深度学习模型来实施端到端的自动特征提取。在此过程中，选取了不同规模的模型进行探索，包括在小规模的数据集上预训练的模型（小型模型）如 ResNet 和 MobileNet，以及大规模数据集上预训

练的模型（大型模型），例如 BiT 模型（BiT-S 和 BiT-M）。ResNet，也称残差网络，通过其残差学习框架有效解决了深度网络训练的挑战，显著提高了处理深层信息的能力，并确保训练过程的稳定性。MobileNet 专门针对移动和嵌入式设备设计，采用深度可分离卷积技术大幅降低了计算量和模型尺寸，保证了在资源有限环境中的高效运行。BiT 是一个基于 ResNet 的深度学习框架，专为迁移学习而设计。BiT 模型通过在广泛且多样化的大型数据集（如 ILSVRC-2012 和 ImageNet-21K）上进行预训练，来学习丰富的特征表示。这种大规模预训练策略显著提升了模型在各种下游任务上的性能，即使这些任务的可用数据相对较少^[98]。BiT 模型的关键优势在于其简化的迁移学习过程，它允许研究者通过简单的微调来适应特定的任务需求，极大地增强了模型在处理不同视觉任务时的灵活性和通用性。

通过深度学习技术，这些模型能自动识别并提取图像中的复杂特征，克服了传统手动特征提取可能引入的主观偏差与限制。小型模型由于其较少的参数和计算需求，非常适合资源受限环境，而大型模型则凭借其深层网络结构和卓越的特征提取能力，更适用于对精度要求较高的应用场景。本章通过比较不同规模深度学习模型在自动特征提取方面的性能和适用性，旨在图像分析任务中提供更加丰富和精确的特征表示。此外，此项比较研究也展示了深度学习技术在图像处理和分析领域，尤其是在高通量表型分析中的应用潜力和价值。

图像数据分为两种处理流程：一是依赖于传统方法的初级图像特征提取，二是利用深度学习技术进行的端到端自动特征提取。为了满足下游 SPAD 值估算任务的需求，这两类提取的特征均需经过两个卷积层的进一步微调。同时，光谱数据将输入 Inception 模块的组合中，与图像特征共同构成了一个多输入模型，如图 5.4 所示。由于光谱特征和图像特征之间存在关联性，特别是来自高光谱相机的光谱和图像数据，在特征融合阶段加入了一个注意力模块（见图 5.4）。该策略的引入可以使模型更专注于对估算任务具有决定性影响的关键特征。本章将此技术环节称为融合注意力模块（Fusion Attention）。深度学习光谱图像融合模型的相关参数设置见附录 C 表 C3-C4。本研究中，所有的基于光谱数据和图像数据的深度学习特征融合模型统称为 SpecImageNet。

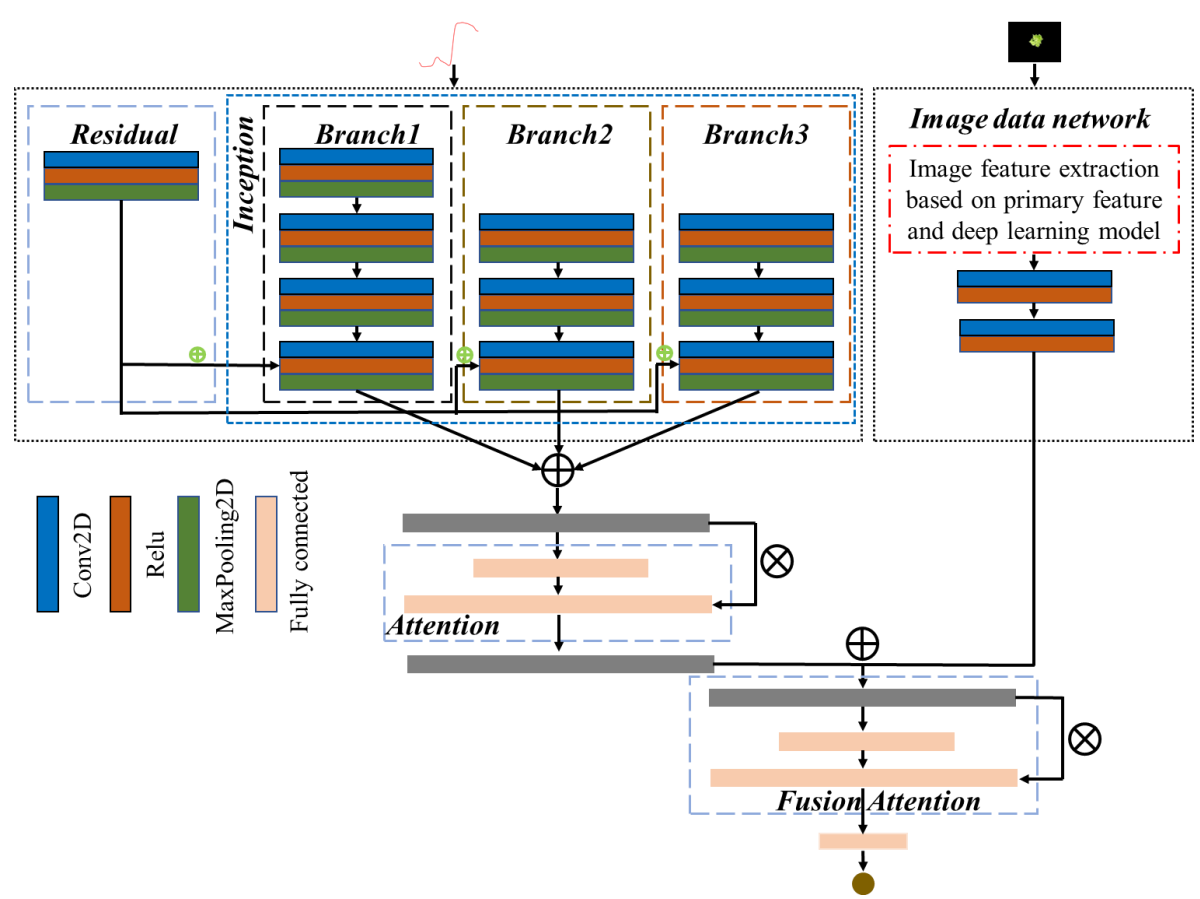


图 5.4 深度学习光谱图像融合模型（SpecImageNet）的架构

Figure 5.4 Architecture of deep learning spectra image fusion model (SpecImageNet)

5.2.5 SPAD 值的估算模型建立

在最终分析中，使用 1886 个实验样本来建立生菜 SPAD 值的估算模型。这些样本经过随机分配^[67]，形成了校准集和预测集，分配比例为 2:1。再随机选取校准集的 20% 样本，用于组成验证集。模型性能评价指标及其描述见 2.3.3 节。为了确保不同建模方法性能评估的公平性，所有建模过程中随机数生成器的状态保持不变。所有模型都是在 Python 3.8.13 和 TensorFlow 2.6.0 环境下进行。所有计算实验是在 HP Z8 G4 Workstation（HP, 美国）计算机上进行的，该计算机配置有 2.40 GHz 的英特尔至强 Gold 6148 CPU，256 GB RAM 和两块英伟达 Quadro P6000 GPU，操作系统为 Windows 10 企业版。

5.3 结果分析和讨论

5.3.1 SPAD 值的参考值

表 5.1 详细呈现了包括群体生菜以及分为七个类别的生菜 SPAD 值的统计数据 and 分

布范围。通过对表 5.1 的分析可见,不同生菜品种的 SPAD 值分布存在较为明显差异。具体来说,散叶生菜和橡叶生菜在 SPAD 值的变异性方面表现得尤为突出,这表明这两种生菜品种在叶绿素含量或叶片健康状态上的差异可能较大,反映出其对环境条件或生长阶段的敏感性。相比之下,罗马生菜的 SPAD 值分布范围相对较窄,其数据的变异性也是所有品种中最小的。这种现象可能说明罗马生菜在不同环境条件下维持较为稳定的叶绿素含量或叶片健康状况的能力较强,表现出较好的适应性。值得注意的是,尽管罗马生菜的样本数量最多,但其稳定的数据分布和较小的变异性强调了该品种在光合效率和营养吸收方面的一致性。这些差异可能与各品种的遗传特性、生长习性以及对环境条件的适应能力紧密相关。此外,各品种生菜的平均 SPAD 值也存在差异,其中罗马生菜和野生近缘种的平均值较高,可能指示这些品种具有较强的光合作用能力或更优良的叶片营养状态。总体而言,这项分析不仅揭示了不同生菜品种在叶绿素含量和健康状态方面的差异,也为进一步研究植物生理特性、改善品种选择和优化种植管理提供了有价值的信息。

表 5.1 各生菜品种的 SPAD 值的参考值测量

Cultivars	Range	Mean	SD
All	8.3-44.4	26.3821	7.8159
Butter	8.5-40.9	19.7321	5.5310
Crisphead	12.4-41.9	26.8850	5.4617
Leaf	8.3-37.3	21.5589	6.7056
Oak leaf	14.8-40.5	25.9490	6.8831
Roman	19.4-44.1	33.2006	5.2462
Stem	14.3-36.7	24.5699	5.3175
Wild relatives	10.6-44.4	32.5551	5.7015

5.3.2 不同生菜品种的光谱反射率分析

图 5.5 详细展示了群体生菜样本及其七个品种在 400 至 1000 nm 波长区间内的平均光谱反射率。可以发现,尽管所有品种的光谱反射率曲线在大体上随波长的增加显示出类似的趋势,但是不同品种之间,乃至同一品种内不同样本间在具体的反射率值上存在显著的差异。这种变异性揭示了各品种及其个体在光谱特征上的独特性。进一步分析发现,不同生菜品种在光谱反射率的分布上存在显著差异,特别是在近红外区域表现得尤为突出。例如,奶油生菜和结球生菜的反射率分布相对较为紧凑,主要集中在 0.7 至 0.9 的区间内。这一现象表明,这两种生菜在近红外波段的反射和吸收特性上显示出较高的一致性和特异性,这可能与其叶绿素含量、水分状况和组织构造紧密相关。相比之下,其他五种生菜,尤其是罗马生菜和野生近缘种生菜,其反射率值的分

布范围更广，介于 0.4 至 0.9 之间，反映了这些生菜品种在光谱特性上的广泛多样性，这种多样性可能源自于遗传变异和生理状态的差异。

总之，通过对生菜光谱反射率的细致分析，不仅能够识别出品种间的差异，而且还能够为理解植物的光合作用效率、水分利用和营养成分的生物学基础提供关键的信息。这些光谱数据在农业生产实践中也具有重要的价值，比如用于品种选择和种植条件的优化等方面。

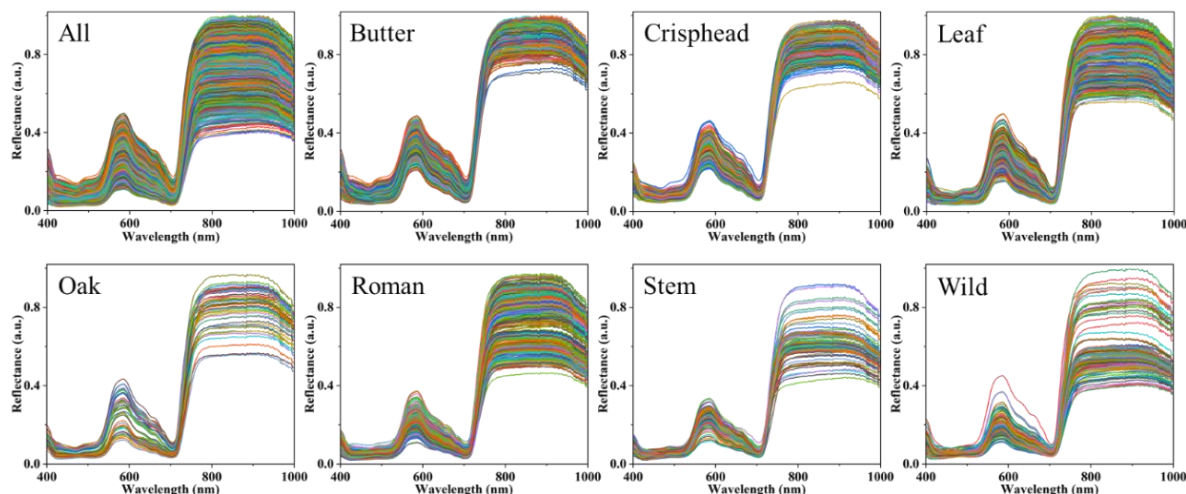


图 5.5 群体生菜及七个生菜品种的平均光谱反射率

Figure 5.5 Mean spectral reflectance of all lettuces and seven lettuce cultivars

5.3.3 基于光谱的估算模型

基于第四章的实验结果，可以发现残差和注意力模块可以提升 Inception 模块的建模精度，但是对于不同的表型估算模型来说，最优模块组合存在不确定性。所以需要不同的模块组合进行消融实验，以确定 SPAD 值估算模型的最佳模块组合。结果如表 5.2 所示，当 Inception 模型中仅加入残差模块，观察到模型的性能有所提升， R_c^2 从 0.7635 增加到 0.7785， R_v^2 从 0.7386 增加到 0.7593， R_p^2 从 0.7594 增加到 0.7743。这表明残差模块有助于改进模型对生菜 SPAD 值的估算能力，可能是因为它促进了梯度的流动，减少了训练过程中的信息丢失。只使用注意力机制时，性能得到了更显著的提升， R_c^2 增加到 0.8192， R_v^2 增加到 0.7929， R_p^2 增加到 0.8118。这表明了注意力模块在提升估算准确度方面的有效性，这可能是因为注意力模块允许模型集中学习数据中最有用的特征。当同时使用残差模块和注意力模块时，结果显示 R_c^2 、 R_v^2 和 R_p^2 分别为 0.7992、0.7781 和 0.7940，表明残差和注意力模块组合并没有达到单独使用注意力模块时的性能水平。说明两种技术的结合并不总是产生叠加效应，这一发现与第四章的研

究结果是一致的。在所有模型配置中，单独使用注意力机制时的 RPD 值最高（2.3050），这意味着模型的估算能力在这种配置下最为可靠。而组合使用残差模块和注意力模块的 RPD 值略低（2.2034），但仍然高于没有使用这些技术的 Inception 模型。图 5.6 展示了基于 Inception、残差和注意力模块的估算散点图。对于局部比较小的 SPAD 值数据点，注意力模块对于模型的优化效果比残差模块更好。研究表明，即使是基于已经非常强大的 Inception 模型，通过引入神经网络技术如残差模块和注意力模块，依然有可能进一步提升模型性能。残差模块和注意力模块的引入在提高模型的估算准确性方面发挥了积极作用，特别是注意力机制在模型性能提升方面显得尤为重要。这些结果也再次表明，深度学习架构的改进并不总是线性叠加的，各种技术的相互作用可能复杂，需要通过实验来确定最有效的配置。所以，在接下来的分析中，针对光谱的深度学习模型将采用 Inception 模块和注意力模块的组合。

表 5.2 Inception、残差和注意力模块的结合估算生菜的 SPAD 值

Table 5.2 Estimation of SPAD value in lettuces using the combination of Inception, residual, and attention module

Model	Improve		R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
	Residual	Attention							
Inception	×	×	0.7635	3.7588	0.7386	3.7997	0.7594	3.9182	2.0386
	√	×	0.7785	3.6370	0.7593	3.6460	0.7743	3.7945	2.1050
	×	√	0.8192	3.2858	0.7929	3.3821	0.8118	3.4653	2.3050
	√	√	0.7992	3.4631	0.7781	3.5004	0.7940	3.6251	2.2034

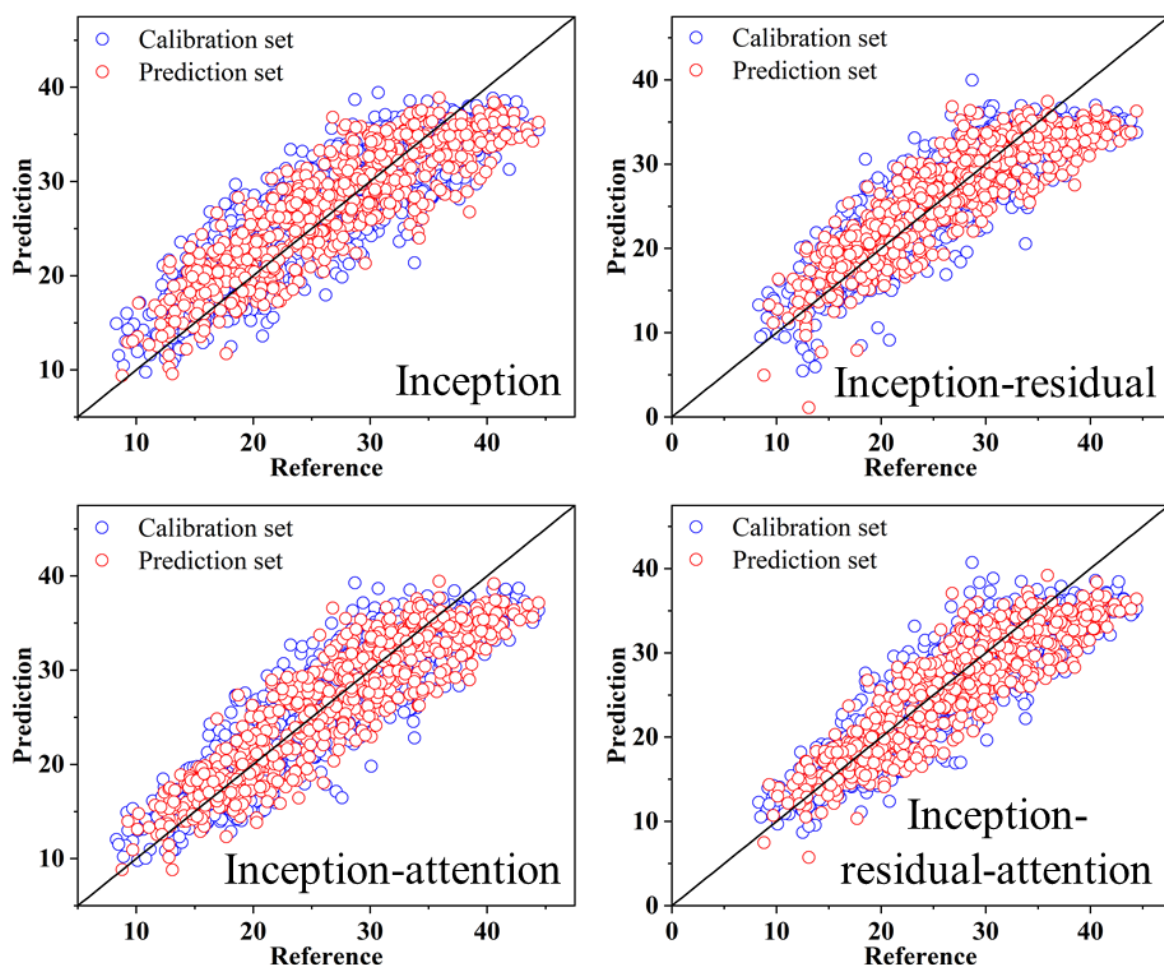


图 5.6 基于 Inception、残差和注意力模块的 SPAD 值估算

Figure 5.6 Estimation of SPAD value based on Inception, residual, and attention module

5.3.4 基于光谱和图像融合的估算模型

为进一步提升 SPAD 值的估算精度，在光谱数据的基础上引入了图像数据。在处理图像数据时，对比了初级图像特征提取、深度学习小型模型（ResNet 和 MobileNet），以及大型模型（BiT-M 和 BiT-S）在 RGB 和 HSI 数据上估算生菜的 SPAD 值表现。深度学习小型模型使用 MobileNet 和 ResNet，主要有两方面设计，一是 MobileNet 作为超分辨率重建模型的特征提取网络表现优异，以及 ResNet 是 BiT 模型的基础模型；二是因为 MobileNet 作为一种轻量级深度学习模型，可以与 ResNet 模型之间做个对比，因为 ResNet 模型的参数量是 MobileNet 模型的 7 倍以上。图谱融合最终目的在于确定图像网络分支中最优的图像处理方法，并将获得最优性能的方法作为后续研究的焦点。结果如表 5.3 所示。

表 5.3 初级特征提取、深度学习模型结合 RGB、HSI 相机估算生菜的 SPAD 值

Table 5.3 Estimation of SPAD value in lettuces using primary feature extraction and deep learning

model combined RGB and HSI camera								
Sensors	Models	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
RGB	Primary feature	0.9652	1.4420	0.8162	3.1861	0.8196	3.3925	2.3545
	ResNet	0.8914	2.5464	0.8578	2.8019	0.8334	3.2599	2.4502
	MobileNet	0.8758	2.7241	0.8355	3.0140	0.8237	3.3542	2.3813
	M	R50*1	0.8909	2.5530	0.8554	2.8255	0.8470	2.5562
		R50*3	0.8892	2.5726	0.8303	3.0616	0.8650	2.9343
		R101*1	0.9497	1.7340	0.8313	3.0520	0.8404	3.1905
		R101*3	0.9062	2.3670	0.8581	2.7997	0.8519	3.0742
		R152*4	0.8777	2.7029	0.8330	3.0373	0.8510	3.0836
	S	R50*1	0.8705	2.7816	0.8421	2.9532	0.8436	3.1584
		R50*3	0.9358	1.9576	0.8340	3.0274	0.8480	3.1136
		R101*1	0.8500	2.9929	0.8333	3.0344	0.8376	3.2189
		R101*3	0.8645	2.8443	0.8377	2.9941	0.8446	3.1492
		R152*4	0.8526	2.9674	0.8221	3.1348	0.8435	3.1594
	Primary feature	0.9590	1.5656	0.7871	3.4289	0.8204	3.3848	2.3598
	ResNet	0.8625	2.8658	0.8469	2.9078	0.8390	3.2051	2.4921
	MobileNet	0.8654	2.8356	0.8317	3.0486	0.8291	3.3023	2.4188
HSI	M	R50*1	0.9145	2.2600	0.8594	2.7870	0.8599	2.9902
		R50*3	0.8744	2.7386	0.8524	2.8553	0.8620	2.9672
		R101*1	0.8933	2.5248	0.8524	2.8547	0.8533	3.0590
		R101*3	0.8815	2.6604	0.8686	2.6941	0.8588	3.0016
		R152*4	0.8971	2.4797	0.8457	2.9188	0.8523	3.0698
	S	R50*1	0.8959	2.4931	0.8555	2.8247	0.8517	3.0756
		R50*3	0.8739	2.7449	0.8467	2.9093	0.8555	3.0362
		R101*1	0.9474	1.7733	0.8310	3.0554	0.8442	3.1524
		R101*3	0.8708	2.7776	0.8513	2.8659	0.8482	3.1118
		R152*4	0.8812	2.6636	0.8585	2.7951	0.8579	3.0109

对于 RGB 数据，初级图像特征提取方法表现出较高的校准集相关系数（ $R_c^2 = 0.9652$ ），表明模型在训练数据上学习得相当充分。然而，其在验证集（ $R_v^2 = 0.8162$ ）和预测集（ $R_p^2 = 0.8196$ ）上的性能显著下降，反映出该方法容易导致过拟合。相比之下，对于 HSI 数据，初级图像特征提取的泛化性能有所提升（ $R_p^2 = 0.8204$ ）。但依然不及深度学习模型，特别是在 R_p^2 指标上。对于两种深度学习小型模型来说，不管是在 RGB 数据上，还是 HSI 数据上，ResNet 的性能（RGB: $R_c^2 = 0.8914$, $R_v^2 = 0.8578$, $R_p^2 = 0.8334$; HSI: $R_c^2 = 0.8625$, $R_v^2 = 0.8469$, $R_p^2 = 0.8390$ ）都优于 MobileNet（RGB: $R_c^2 = 0.8758$, $R_v^2 = 0.8355$, $R_p^2 = 0.8237$; HSI: $R_c^2 = 0.8654$, $R_v^2 = 0.8317$, $R_p^2 = 0.8291$ ）。这表明在估算生菜 SPAD 值时，ResNet 模型的图像特征提取能力优于轻量级的 MobileNet 模型。另外，根据 R_c^2 与 R_p^2 的差值分析，RGB 数据的效果不仅略差于 HSI 数据，还有更大的过拟合。

对于 BiT-M 模型，除了 R50*3 大型模型在 RGB 数据上的效果（ $R_p^2 = 0.8650$ ）优

于 HSI 数据 ($R_p^2 = 0.8620$), 剩余 4 种大型模型都是 HSI 数据的效果更好。对于 BiT-S 模型而言, 5 种大型模型都是 HSI 数据的效果更好。考虑到 HSI 数据在模型中的优异表现, 尤其是大型模型, 引入超分辨率重建技术来提升高光谱图像的空间分辨率, 可能会更进一步提升模型性能。

在 BiT-M 和 BiT-S 模型处理相同的图像数据的情况下, 除了 S-R152*4-HSI 的效果 ($R_p^2 = 0.8579$) 略优于 M-R152*4-HSI ($R_p^2 = 0.8523$), 所有的 BiT-M 模型都优于 BiT-S 模型。这可能是因为 BiT-M 相对于 BiT-S 拥有更大的模型容量, 因此可以捕捉到更多的图像特征和语义信息。更大的模型通常具有更强的表示能力, 能够处理更复杂的任务和更丰富的数据。另外, BiT-M 是在比 BiT-S 更大、更多样化的数据集上进行了预训练。预训练阶段的数据集越大且多样性越高, 学到的特征表示就越具有泛化能力, 从而使得模型在不同任务上的适应性更强。对于同类大型模型, 大型模型的网络宽度适当增加有利于提升模型性能, 但不能一直增加, *3 相对于*1 和*4, 效果更好, 就大型模型网络结构而言, ResNet50 的效果要优于 ResNet101。其中, M-R50*3 在 RGB 数据 ($R_c^2 = 0.8892$, RMSEC = 2.5726, $R_v^2 = 0.8303$, RMSEV = 3.0616, $R_p^2 = 0.8650$, RMSEP = 2.9343, RPD = 2.7221) 和 HSI 数据 ($R_c^2 = 0.8744$, RMSEC = 2.7386, $R_v^2 = 0.8524$, RMSEV = 2.8553, $R_p^2 = 0.8620$, RMSEP = 2.9672, RPD = 2.6919) 上都获得了最优的结果。因此, 在后续工作中, 将专注于应用 M-R50*3 大型模型, 并探索超分辨率重建技术在 HSI 数据上的应用潜力。图 5.7 展示了使用 primary feature、ResNet 以及大型模型进行 SPAD 值估算的最优结果对应的散点图。相较于图 5.6 中展示的 Inception-attention 模型结果, 这三种方法的估算数据点均表现出更高的集中度, 更紧密地贴近似合线。特别是大型模型的估算效果最为突出, 绝大多数数据点都显著靠拢于拟合线, 展现了高度的估算精度和一致性, 尽管仍有少数数据点的精度未显著提升。

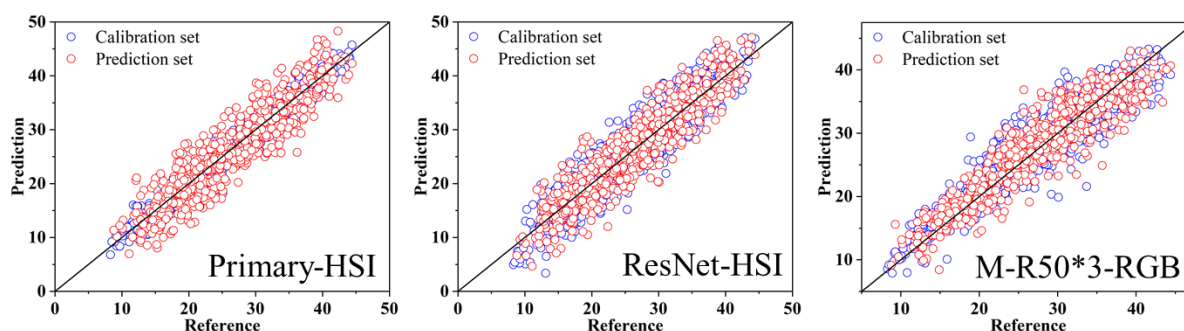


图 5.7 基于初级图像特征、ResNet 和大型模型的 SPAD 值估算

Figure 5.7 Estimation of SPAD value based on primary feature, ResNet, and large model

总而言之，图谱融合相对于单一的光谱特征，提升了模型的性能。大型模型在估算效果上均优于小型模型和初级图像特征方法。因为大型深度学习模型是通过多个层次的抽象表示学习图像的特征。底层的卷积层可以捕捉低级别的特征，如边缘和纹理，而高层的卷积层则能够捕捉更抽象和语义化的特征。这使得模型能够理解图像中的复杂结构和关系。并且大型模型是在大规模且多样化的数据集上进行训练的。这使得它们能够学到更通用、更具有代表性的特征，而不仅仅是针对特定任务的特定特征。在具有足够多样性的数据集上进行训练有助于模型更好地泛化到其他任务。

5.3.5 基于增强的光谱和图像融合的估算模型

(1) 高光谱图像空间分辨率的提升

为了深度挖掘高光谱图像的潜力，采用了八种超分辨率重建方法对高光谱图像的空间分辨率进行了提升，其中包括传统的图像插值方法（Cubic、Linear、Near 和 Lanczos）和基于深度学习的 SRGAN（特征提取网络：EfficientNet、MobileNet、VGG16 和 VGG19）。为了确定最优的上采样倍率，基于 SRGAN 在训练时所用的高分辨率 RGB 图像尺寸（ 2560×1920 ）以及估算时采用的原始分辨率高光谱图像尺寸（ 696×775 ），设定了 2 倍和 4 倍的上采样选项。使用 SRGAN 和插值法，对原始高光谱图像进行上采样后的结果如图 5.8 所示（以奶油生菜为例）。

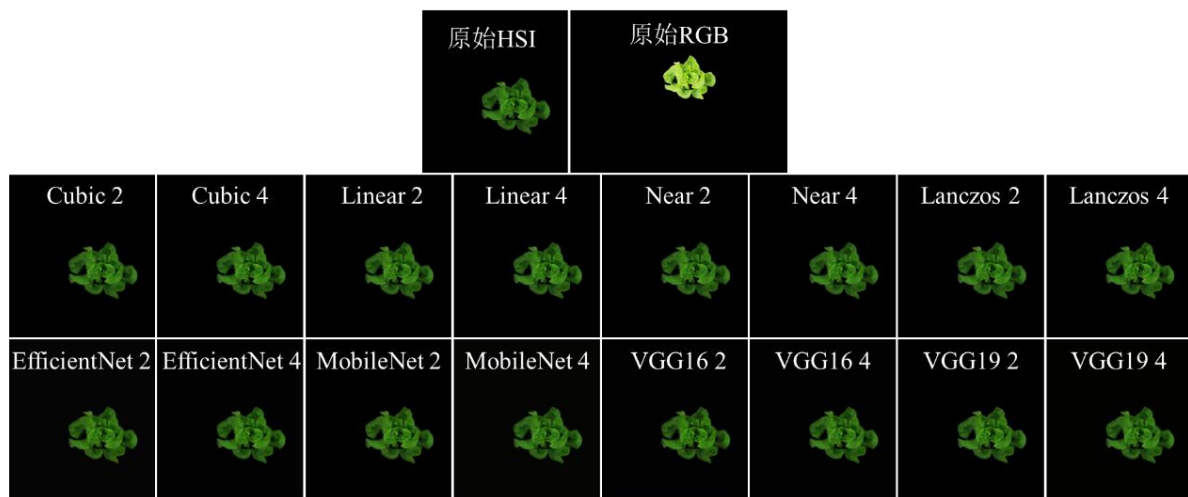


图 5.8 使用 SRGAN 和插值法对原始高光谱图像上采样后的结果

Figure 5.8 The result after upsampling the raw hyperspectral image using SRGAN and interpolation

通过观察可以明显区分原始的高光谱图像合成的 RGB 图像与直接由 RGB 相机拍摄的 RGB 图像之间存在的显著差异。这种差别主要源于两种成像技术的根本区别：高光谱成像系统能够捕获的光谱范围远超传统 RGB 相机，通常涵盖从可见光到近红外甚

至更广泛的光谱区域。高光谱图像包含数十至数百个窄波段通道，每个通道捕获特定波长的光谱特征信息，提供了丰富的光谱数据。相比之下，RGB 相机仅捕获红、绿、蓝三个宽波段通道的信息，这三个通道大体上对应人眼的颜色感知范围，其提供的光谱信息量远不及高光谱成像系统。此外，两种成像技术在成像原理及图像处理方面也存在区别，高光谱图像的合成涉及选择特定接近人眼感知波段的多个窄波段进行组合，此过程及其后续的图像处理显著影响了图像的最终外观。而 RGB 相机图像的生成直接基于传感器对红绿蓝光的响应及相机内部的图像处理算法，这些算法可以优化图像的视觉效果，如对比度、饱和度和锐度等。

尽管从模型建立的角度看，不同的图像分辨率提升方法展现出了一定的性能差异，但这些差异在视觉上难以识别。即便经过仔细观察，由不同超分辨率技术处理后的图像间的区别对于肉眼几乎是不可辨识的，特别是在生菜图像这种视觉上颜色相对单一（主要为不同阴影的绿色）的情境下。这种现象限制了通过视觉感知来区分和评价超分辨率技术带来的改进，强调了在植物表型分析等特定应用场景中，单色主导的图像可能制约了视觉评估超分辨率技术改进的能力。此外，尽管生菜图像中的细微颜色变化对量化模型性能评估至关重要，但这些微妙的变化很可能超出肉眼的识别范围，尤其是在背景颜色较为单一且对比度较低的条件。因此，这一现象凸显了评估超分辨率技术有效性时，依赖于量化的模型性能指标的方法的重要性。在当前技术水平下，由于不同处理方法之间的视觉差异微乎其微，传统的主观感知评价的局限性变得尤为明显。因此，更加精细和全面的量化评估方法，如图像质量评价指标峰值信噪比（Peak Signal-to-Noise Ratio, PSNR）和结构相似性指数（Structural Similarity Index, SSIM），成为了判断模型性能优化的关键^[79]。SRGAN 的训练结果见附录 C 图 C2。插值法重建后的图像质量评价见附录 C 表 C2。在植物表型分析中应用超分辨率技术时，需依赖于这些量化的表型估算模型性能，而非仅仅依赖于视觉感知，以确保评估结果的准确性和公正性。

（2）基于 SR 的光谱和图像特征融合的估算模型

在图像上采样后，两种倍率在生菜 SPAD 值估算性能方面的表现将用作评估标准。相关结果详见表 5.4。与表 5.3 中提供的基线模型（M-R50*3-HSI）在生菜 SPAD 值估算性能相比，可以发现，对于 R_c^2 和 R_v^2 ，几乎所有使用超分辨率技术的方法都取得了比基线模型更高的值，显示出了超分辨率技术在改善模型训练和验证集上的性能方面的潜力，但对于 R_p^2 ，仅有 SRGAN 方法实现的模型超过了基线模型。插值法中使用

Lanczos 方法将原始高光谱图像放大了 4 倍获得了最好的结果 ($R_c^2 = 0.9429$, $RMSEC = 1.8461$, $R_v^2 = 0.8521$, $RMSEV = 2.8582$, $R_p^2 = 0.8559$, $RMSEP = 3.0324$, $RPD = 2.6340$), 但这个结果仍然比基线模型的结果差。所有的差值方法都使得模型的性能有所下降。

表 5.4 基于 SRGAN 和插值法估算生菜的 SPAD 值

Table 5.4 Estimation of SPAD value in lettuces based on SRGAN and interpolation

Models	SR	Methods	Scale	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
M-R50*3	Image interpolation	Cubic	2	0.9153	2.2490	0.8349	3.0197	0.8470	3.1244	2.5565
			4	0.9251	2.1157	0.8556	2.8238	0.8523	3.0695	2.6022
		Linear	2	0.9014	2.4263	0.8450	2.9254	0.8513	3.0800	2.5934
			4	0.9050	2.3826	0.8473	2.9039	0.8539	3.0533	2.6160
		Near	2	0.9296	2.0505	0.8308	3.0572	0.8428	3.1672	2.5219
			4	0.8942	2.5143	0.8563	2.8171	0.8523	3.0700	2.6018
		Lanczos	2	0.9395	1.9003	0.8532	2.8473	0.8515	3.0783	2.5948
			4	0.9429	1.8461	0.8521	2.8582	0.8559	3.0324	2.6340
		EfficientNet	2	0.9051	2.3803	0.8600	2.7802	0.8752	2.8215	2.8309
			4	0.9404	1.8864	0.8719	2.6600	0.8697	2.8831	2.7705
	SRGAN	MobileNet	2	0.9359	1.9561	0.8773	2.6027	0.8900	2.6489	3.0154
			4	0.9012	2.4294	0.8501	2.8770	0.8793	2.7749	2.8785
		VGG16	2	0.8994	2.4508	0.8509	2.8700	0.8799	2.7676	2.8861
			4	0.9119	2.2941	0.8535	2.8442	0.8759	2.8141	2.8384
		VGG19	2	0.9051	2.3810	0.8528	2.8516	0.8795	2.7723	2.8812
			4	0.9351	1.9682	0.8652	2.7282	0.8679	2.9034	2.7511

与插值法相比, 使用 SRGAN 方法的模型都提升了模型的性能。其中分辨率提升 2 倍的 MobileNet 的效果最好 ($R_c^2 = 0.9359$, $RMSEC = 1.9561$, $R_v^2 = 0.8773$, $RMSEV = 2.6027$, $R_p^2 = 0.8900$, $RMSEP = 2.6489$, $RPD = 3.0154$)。在 SRGAN 中, 特征提取网络通常是用来提取图像的特征表示, 这些特征表示会用于生成超分辨率图像, 所以特征提取网络对 SRGAN 生成图像质量起着关键的作用。在 SRGAN 的实际应用中, 特征提取网络的选择将基于所需的性能和可用的计算资源。复杂且深度较大的网络可能捕捉更细致的特征, 但也会增加计算的负荷。轻量级网络虽然计算效率更高, 但可能需要更精细的调整来达到同样的性能水平。这里选择 2 种轻量级网络 (EfficientNet、MobileNet) 和 2 种重量级网络 (VGG16、VGG19) 来对比。可以发现, MobileNet 的效果更好, 另外 VGG16 的效果 (scale 2: $R_p^2 = 0.8799$; scale 4: $R_p^2 = 0.8759$) 也比 VGG19 (scale 2: $R_p^2 = 0.8795$; scale 4: $R_p^2 = 0.8679$) 优异。所以, 在实际应用中, SRGAN 的特征提取网络并不是越复杂越好, 选择合适的特征提取网络对模型最终的估算性能有着很大影响。图 5.9 展示了基于原始高光谱图像, 以及性能最优的插值法和 SRGAN 技术的 SPAD 值估算结果散点图。从图中可以观察到, 使用插值法处理后的数据点相较于原始 HSI 的分布更为离散, 表明估算误差有所增加。相反, SRGAN 技术带

来的估算结果则显著改善，数据点更紧密地聚集在拟合线周围，反映了其在提升估算精度方面的显著效果。

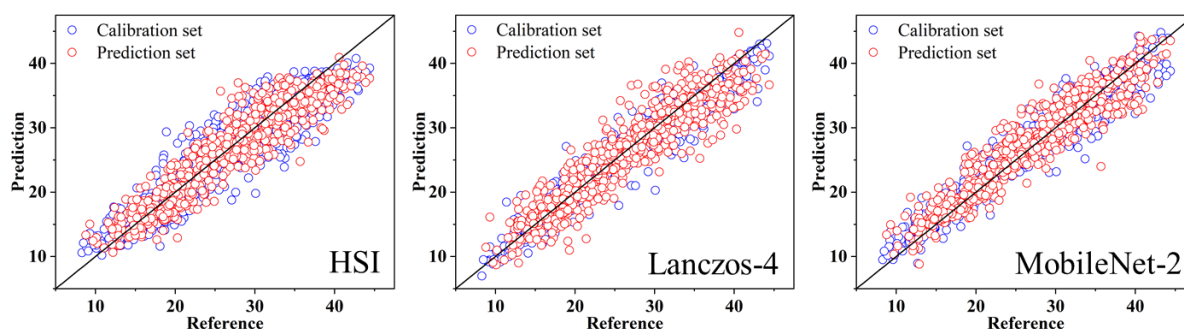


图 5.9 基于原始 HSI、差值法和 SRGAN 的 SPAD 值估算结果

Figure 5.9 Estimation of SPAD value based on raw HSI, interpolation, and SRGAN

可以发现，上采样倍数的增加并不总会提升估算模型的性能。在使用 SRGAN 的情况下，图像放大 4 倍后的效果相比 2 倍都变差了，例如使用 SRGAN 和 MobileNet 组合时，4 倍放大比例下的 R_p^2 值（0.8793）比 2 倍放大比例下的 R_p^2 值（0.8900）低。而插值法的 4 倍的效果都在 2 倍的基础上得到了提升。值得注意的是，增加放大倍数意味着模型需要在放大后的图像上填充更多的细节。理论上，当放大倍数较大时，模型必须生成更多未在原始低分辨率图像中存在的信息。差值法在 4 倍放大时的性能优于 2 倍，这可能是因为差值法依赖统计方法来估算新像素的值，这种方法在放大倍数较大时不会引入显著的错误，尤其是在图像的局部结构比较规则时效果较好。SRGAN 在 2 倍放大时的效果通常优于 4 倍，这可能是因为当放大倍数较小（例如 2 倍）时，生成器需要推断的新内容相对较少，SRGAN 能够更好地恢复高质量细节，因此生成的图像更加准确。而在放大倍数增加到 4 倍时，生成器需要填充更多的细节，这可能导致生成的图像出现更多的伪影或不真实的纹理，影响模型的估算性能。另外，由于训练使用的高分辨率图像的尺寸限制，如果模型生成超出其在训练时见过的尺寸范围的图像。这可能导致生成器无法有效地学习到足够的高分辨率内容，从而影响其性能。

SRGAN 在提升图像质量及估算生菜 SPAD 值方面表现出了明显的优势，尤其是在 2 倍上采样时。然而，在 4 倍或更高倍数的上采样中，图像插值法的表现出了意外的韧性，这揭示了该方法在图像放大过程中的潜在适应性。因此，在未来的工作中，除了进一步研究和优化 SRGAN 在处理更高倍数放大时的细节生成问题，如使用更高分辨率的训练图像来增强模型学习高质量细节的能力；对生成器进行改进以更好地处理高倍数放大下细节生成的问题，或者采用增强学习等方法来优化 4 倍或在此之上放大倍

数下的性能。还应考虑探讨和深入挖掘插值方法在极端放大情况下的可行性和性能极限。具体而言，这可能涉及测试不同的插值算法，理解其在极端放大情况下保持图像质量的内在机制，并评估其对提高原始分辨率下图像质量的贡献。

（3）基于光谱特征选择的光谱和图像特征融合的估算模型

与 RGB 图像只含有红绿蓝三个波长特征的图像相比，高光谱图像包含了可见-近红外范围内（400-1000 nm）256 个波长特征的图像。从光谱学的角度来分析，这 256 个波长的光谱特征对 SPAD 值的估算准确度的贡献是不一样的，所以这些光谱特征对应的图像中可能蕴含有更能准确反应生菜 SPAD 值的图像信息。接下来将采用基于机器学习的光谱特征选择方法（CARS 和 UVE）以及基于深度学习的 CAM 方法来挑选关键波长的光谱特征。由于上述基于光谱建立的模型采用了 Inception-attention 模块，为了增强该模型的可解释性，在注意力模块之前（Before-attention）和之后（After-attention）分别使用 CAM 方法来挑选光谱特征，以此观察在注意力模块前后，模型对光谱的关注区域发生了什么变化。

根据每种方法选择的关键三个光谱特征对应的高光谱图像，组合成伪彩色图像，使用特征提取网络为 MobileNet 的 SRGAN 对伪彩色图像进行 2 倍上采样，然后结合 M-R50*3 大型模型用于估算生菜的 SPAD 值。表 5.5 和图 5.10 展现了基于机器学习（按光谱特征的权重顺序排列）和 CAM 方法进行光谱特征选择的结果。观察显示，机器学习方法倾向于选取的光谱特征主要集中在可见光区域，而 CAM 方法不仅覆盖了可见光区域，还扩展到了红边和近红外区域。这一发现可以从光谱学角度得到解释，尽管叶绿素的主要吸收波段位于可见光区域，红边和近红外区域的光谱信息也对反映叶绿素含量提供了额外的视角。红边区域的变化与植物的氮含量、叶片结构和水分状态紧密关联，使其在估算 SPAD 值时展现出显著的响应性，且叶绿素含量的变化会影响红边的位置和形状^[4]。植物叶片通常对近红外光有高反射率，这与叶片结构属性有关，在光合作用较弱的叶片中，近红外反射率甚至更高，因此，近红外区域的信息在 SPAD 值的估算中同样显示出关键性的响应^[124]。从模型性能分析来看，深度学习模型在特征提取上展现出其强大的能力，能够自动发掘并利用目标叶绿素吸收波段之外的其他光谱信息。这表明，在训练过程中，模型可能识别出红边和近红外区域的光谱特征与 SPAD 值之间存在强烈的相关性。特别是在应用了注意力模块的深度学习方法中，模型对光谱信息的关注点变得更加集中，尤其是在可见光区域（紧邻叶绿素吸收波段）和近红外区域，进一步强化了这些区域在 SPAD 值估算中的重要性。这一结果强调了深

度学习技术在植物生理特性研究中的应用潜力，尤其是在光谱数据处理领域。

表 5.5 基于机器学习方法的光谱特征选择结果

Table 5.5 Results of spectral feature selection based on machine learning method	
Wavelength selection	Wavelengths (nm)
CARS	484, 486, 564, 582, 646, 562, 489, 735, 529, 534, 738, 651, 580, 674, 531, 721, 585, 648
UVE	482, 484, 487, 480, 489, 646, 644, 679, 681, 677, 648, 674, 684, 491, 651, 641, 524, 477, 580, 527, 522, 653, 583

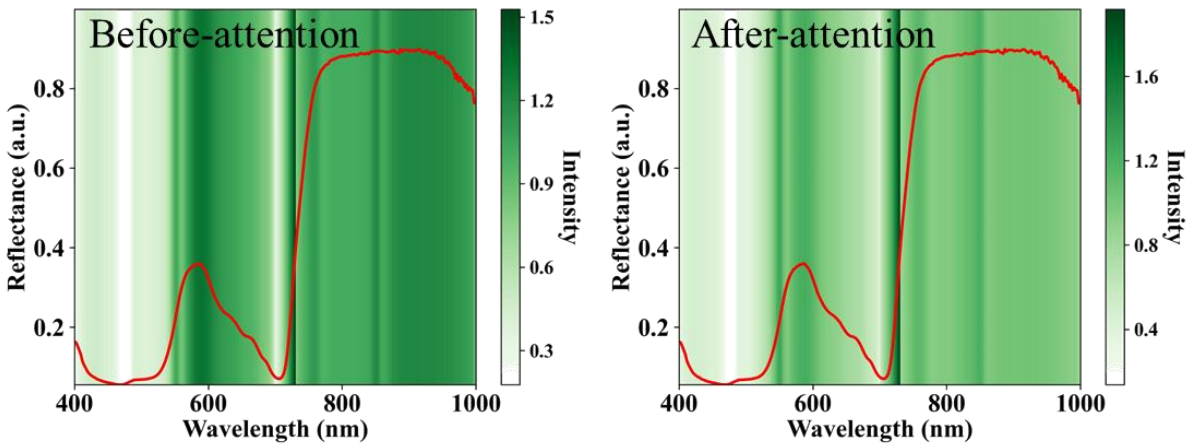


图 5.10 基于 CAM 方法的光谱特征选择结果

Figure 5.10 Results of spectral feature selection based on CAM method

考虑到光谱数据的相邻窄波段间具有高度的共线性，相邻波段之间只选择模型最关注的波长位置。利用 CAM 方法分析光谱特征时，选择了近红外区域、红边区域和可见光区域最为显著的光谱特征位置。具体而言，注意力模块应用之前，选定的显著光谱特征位置分别为 913 nm、728 nm 和 590 nm；而注意力模块应用之后，则为 852 nm、728 nm 和 548 nm。相应地，CARS 和 UVE 方法选择的光谱特征分别是 646 nm、564 nm、484 nm 以及 679 nm、646 nm、482 nm。图 5.11 展示了基于这些光谱特征生成的伪彩色原始图像及其 2 倍上采样后的图像。从分辨率提升前后的图像来看，视觉上依旧难以识别明显的差异，但不同伪色图之间的颜色和亮度变化却是显而易见的。CARS 方法生成的图像相比于 UVE 方法的更为明亮，注意力模块应用前后的图像在色彩上也展现出了明显的差异。

伪彩色图像通过将选定的光谱特征对应的高光谱图像映射到 RGB 颜色空间中，使得不同特征的光谱信息以颜色的形式直观展现。比如，红边区域（728 nm）和近红外区域（如 913 nm 或 852 nm）的反射率变化，可能在伪彩色图像中以特定的颜色变化呈现，这些颜色变化可以与植物的叶绿素含量、氮含量以及水分状态等生理特征相关联。

图像的亮度反映了目标区域光谱反射率的强度。高反射率区域在图像中可能表现为更亮的颜色，这在分析植物叶片的结构和健康状况时尤为重要。例如，健康的植物叶片通常在近红外区域有较高的反射率，因此，这些区域在伪彩色图像中可能呈现为更亮的颜色。通过细致分析基于特定光谱特征生成的伪彩色图像，可以深化对植物生理特征及其在不同光谱区域的反应的理解。

原始图片



分辨率提升



图 5.11 使用四种光谱特征选择方法合成的伪彩色图像及其相应的上采样后的图像

Figure 5.11 Pseudo color images synthesized using four different spectral feature selection methods and corresponding upsampled images

四种光谱特征选择方法得到的估算结果如表 5.6 所示，除了 UVE 以外，另外 3 种方法选择的光谱特征组成的伪彩色图像的结果，都比红绿蓝波长特征组成的 RGB 图像的好。其中 CAM 方法的效果比 CARS 更好，特别是在注意力模块后使用 CAM（CAM After-attention）方法获得了最优的性能， R_c^2 达到了 0.9553， R_v^2 达到了 0.8942， R_p^2 达到了 0.9033，RPD 达到了 3.2158。结果表明 CAM 方法能够将模型内部决策过程的行为可视化，增强了模型的可解释性。四种光谱特征选择方法的估算散点图如图 5.12 所示。从散点图中可以观察到，尽管 CARS 方法相较于 UVE 方法展现出略优的估算性能，但两者之间的差异并不非常显著。相比之下，基于 CAM 方法得到的结果，无论是在引入注意力模块之前还是之后，都显示出拟合线附近数据点密度的显著增加，表明了更

高的估算准确度。

表 5.6 基于四种光谱特征选择方法估算生菜的 SPAD 值

Table 5.6 Estimation of SPAD value in lettuces based on four spectral feature selection methods

Wavelength selection	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
CARS	0.9291	2.0576	0.8733	2.6448	0.8923	2.6215	3.0470
UVE	0.9380	1.9238	0.8859	2.5101	0.8872	2.6828	2.9773
CAM Before-attention	0.9513	1.7055	0.8655	2.7250	0.8984	2.5456	3.1377
CAM After-attention	0.9553	1.6333	0.8942	2.4167	0.9033	2.4838	3.2158

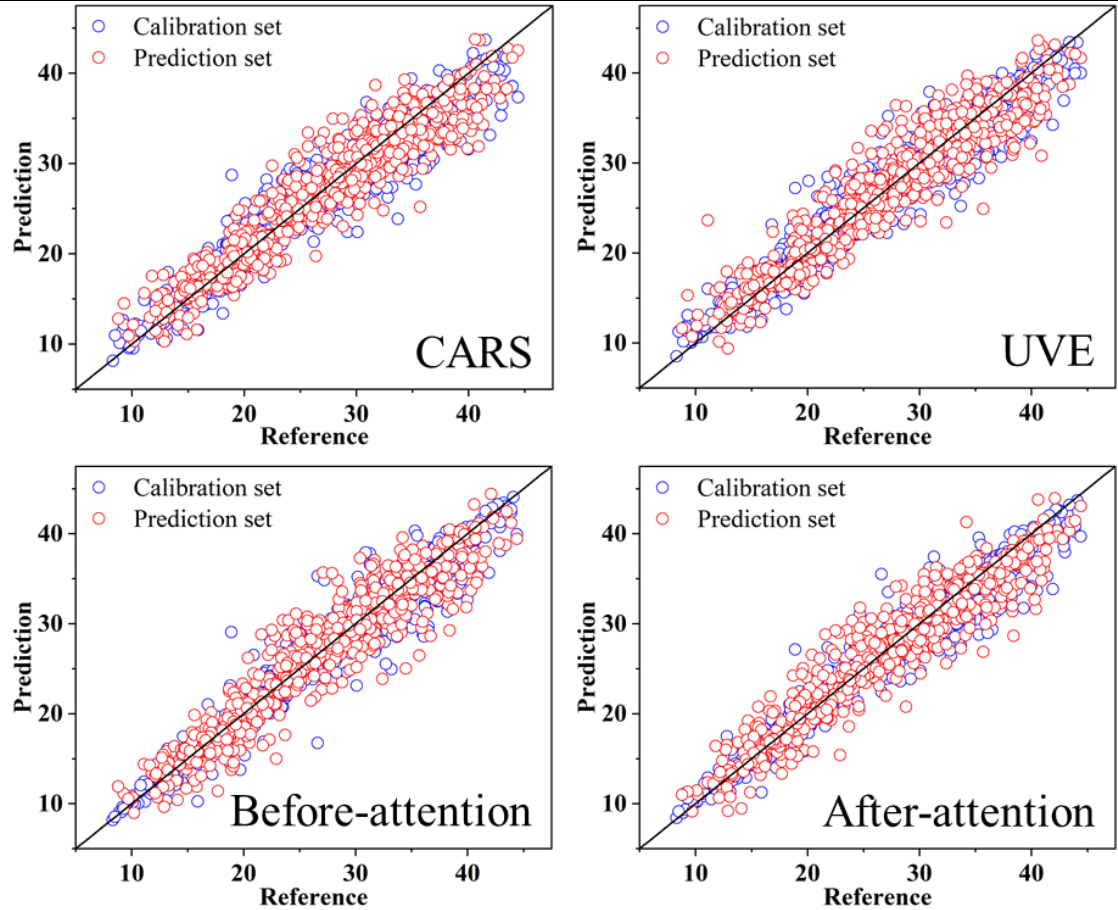


图 5.12 基于四种光谱特征选择方法的 SPAD 值估算

Figure 5.12 Estimation of SPAD value based on four spectral feature selection methods

上述分析显示，通过选择的光谱特征以及结合超分辨率技术和深度学习模型，可以显著提升生菜 SPAD 值估算的精度。这说明 HSI 数据以其光谱信息的丰富，确实能够提供比 RGB 数据更加详尽的特征，这些特征与深度学习模型的结合，可以显著提升生菜 SPAD 值分析的精度。特别是使用了 CAM After-attention 光谱特征选择方法后，估算性能得到了显著提升，验证了 HSI 数据在植物表型分析中的应用价值以及深度学习方法在处理复杂数据时的优势。在接下来的章节中，将专注于使用高光谱图像数据。

值得注意的是，传统方法挑选的最优光谱特征的子集，除了所选的三个关键光谱

特征,还包括了一些波长特征,它们与高光谱图像合成的原始 RGB 图像中选用的波长特征十分接近。对比 CARS 和 UVE 所选的三个光谱特征,可以看出基于机器学习方法选择的光谱特征还有进一步提高 SPAD 值估算精度的潜力。研究中基于三个重要光谱特征组成的伪彩色图像不一定是最优子集内的最佳特征组合,这说明需要探索最优子集内不同光谱特征组合的效果,但是这个过程会非常耗时。相比之下,深度学习方法利用 CAM 分析的关注区域可以快速进行光谱特征选择,这不仅提高了效率,而且增强了模型的可解释性,这在高通量植物表型分析领域具有重要意义,尤其是在分析植物昼夜节律以及光合作用日变化等细分的时间尺度变化时。未来的研究可以继续探索不同光谱特征选择方法与深度学习结合的策略,以进一步优化模型性能。

5.4 本章小结

本章通过应用深度学习图像光谱融合技术,旨在提高生菜 SPAD 值估算的准确性,进而为精准农业领域提供更为高效和准确的表型分析技术。本章使用流水线成像室系统,在温室中收集了七种生菜的高光谱图像数据和 RGB 图像数据,并利用这些数据进行了深入的建模与分析。首先探索了基于光谱数据的不同深度学习模块组合。研究结果显示,融合注意力模块的 Inception 模块在提升效果方面表现尤为突出,其 R_c^2 为 0.8192, RMSEC 为 3.2858, R_v^2 为 0.7929, RMSEV 为 3.3821, R_p^2 为 0.8118, RMSEP 为 3.4653, RPD 为 2.3050。此后,引入了 RGB 相机和 HSI 相机获取的图像数据,并将其通过初级图像特征提取、小型模型与大型模型进行处理。比较分析表明,大型模型在特征提取方面表现出显著的优势。其中,结合 RGB 相机数据的 M-R50*3 模型表现最佳 ($R_c^2 = 0.8892$, RMSEC = 2.5726, $R_v^2 = 0.8303$, RMSEV = 3.0616, $R_p^2 = 0.8650$, RMSEP = 2.9343, RPD = 2.7221)。相应地,使用 HSI 相机数据的 M-R50*3 模型也表现了良好的估算能力 ($R_c^2 = 0.8744$, RMSEC = 2.7386, $R_v^2 = 0.8524$, RMSEV = 2.8553, $R_p^2 = 0.8620$, RMSEP = 2.9672, RPD = 2.6919)。

为了进一步探究图像分辨率对融合模型的影响,采用了图像插值方法和 SRGAN 技术来提升高光谱图像的空间分辨率。其中,采用 MobileNet 的 SRGAN 将高光谱图像的空间分辨率提升 2 倍对图谱融合模型的性能提升最为明显,其 $R_c^2 = 0.9359$, RMSEC = 1.9561, $R_v^2 = 0.8773$, RMSEV = 2.6027, $R_p^2 = 0.8900$, RMSEP = 2.6489, RPD = 3.0154。结果证明 SRGAN 可以为光谱和图像融合模型的性能提升提供有效的技术路径。此外,通过比较不同特征提取网络的性能,为选择适合特定应用需求的网络提供了重

要参考,进一步推动了图像处理技术的发展。此外,为了提升模型的可解释性,应用了 CAM 技术进行光谱特征选择。在注意力模块后使用 CAM 技术进行光谱特征选择的方法提高了图谱融合模型对 SPAD 值估算的精度,其 R_p^2 和 RPD 分别达到了 0.9033 和 3.2158。

综上所述,通过图谱融合模型的应用,相较于单一光谱模型,显著提高了生菜 SPAD 值估算的精度。CAM 和 SRGAN 技术的结合不仅有效提升了融合模型的性能,同时也有助于推动其他植物表型估算和 HTPP 技术的进步。未来研究可进一步增加特定生菜品种的数量,以更细致地观察模型效果,为精准农业提供更多有力支持。

第六章 基于光谱、时间序列和图像的营养元素表型指标的估算

6.1 引言

在先前章节的研究基础上，本章深入探讨了通过光谱数据、时间序列数据和图像数据的融合，对生菜营养元素表型（EC、 K^+ 和 Na^+ ）估算准确性的影响。这一研究是为了延续和扩展第四章与第五章的工作，通过探索更加复杂的多模态数据融合方法，以深入理解不同数据模态对植物营养元素表型估算的贡献。

本章的实验材料与第四章相同，并且营养元素表型和品质表型都是一同测定的。在第四章中，通过融合光谱数据和时间序列数据，发现对生菜品质表型（SSC 和 pH 值）的估算准确性得到了提升。虽然这种融合方法也提升了营养元素表型的估算精度，但是和品质表型相比，营养元素表型的估算精度都比较低。这启发了对包括图像数据在内的更广泛数据模态融合效果的进一步探索。继而，在第五章中，通过光谱与图像数据的融合分析，证明了图像数据在提升单一光谱数据对光合效能表型估算精度上的显著作用，为本章研究方向的选择提供了依据。

营养元素表型估算精度相比品质表型估算精度较低的原因，可能涉及到营养元素本身在植物体内分布的复杂性和动态变化。营养元素表型通常受到多种生理和环境因素的影响，如土壤条件、水分状态和植物吸收能力等，这些因素的复杂互动使得营养元素表型在空间和时间上的变异性更大^[9]，从而增加了估算的难度。相较而言，品质表型如 SSC 和 pH 值，虽然也受到多种因素的影响，但其在植物叶片或果实中的表现更为直接和稳定，因此通过光谱数据进行估算时，相关光谱特征与表型特性之间的关联性可能更加明确和一致。加入时间序列数据，理论上是为了捕捉植物生长过程中营养元素表型的动态变化，以提升模型对时间序列变化敏感表型特性的估算能力。然而，时间序列数据的引入也带来了额外的变异性 and 复杂性，特别是当时间序列数据的采集频率、精度或代表性无法完全匹配营养元素表型的真实动态变化时，可能不会显著提升估算效果。因此，加入图像数据和进一步地优化多模态数据融合方法，有助于从更多维度捕捉植物表型特性的相关信息，从而提升多模态数据融合模型对营养元素表型复杂变异性的解读能力和估算精度，尤其是在探索植物营养状况与外观特征之间的细微关联时。

据此，本章的数据融合方法设计主要分为三个部分：光谱和时间序列融合、光谱

和图像融合，以及光谱、时间序列和图像的融合，以深入比较和分析不同融合方法对营养元素表型估算准确性的影响。此外，探讨了 AutoML 技术在简化模型构建和优化融合模型过程中的应用，以实现更高效、更精确的营养元素表型估算。

通过这一系统性的多模态数据融合研究，可以进一步推动精准农业技术的发展，尤其是在营养元素表型估算方面。这不仅展示了多模态数据融合和 AutoML 技术在农业科研中的应用价值，也为未来的农业生产管理提供了有效支撑。

6.2 数据获取与处理

图 6.1 展示了本章的数据处理流程图。多模态数据涵盖了光谱数据、时间序列数据以及图像数据，其中光谱与图像数据均源自高光谱相机。接下来，这些多模态数据将通过定制化的深度学习模型、传统机器学习模型以及 AutoML 技术进行处理，目的是估算 EC、K⁺和 Na⁺。具体的数据处理方法将在本节的后文部分进行详细阐述。

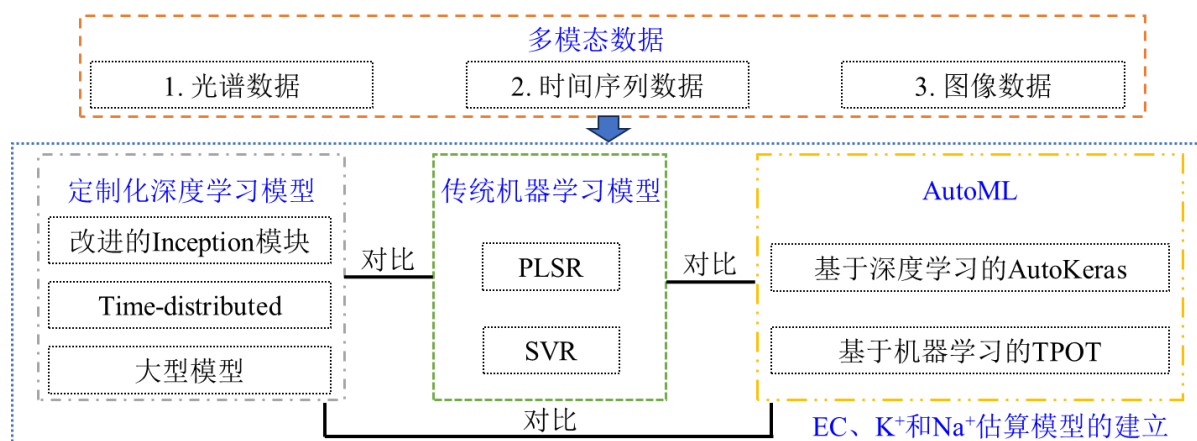


图 6.1 估算生菜的 EC、K⁺和 Na⁺的数据处理流程图

Figure 6.1 Flow diagram of data processing used to estimate EC, K⁺, and Na⁺ in lettuce

6.2.1 实验设计

本章的实验设计与第四章相同，具体描述见 4.2.1 节。分别使用 EC 测定仪（EC-33B，HORIBA，日本）、K⁺测定仪（K-11，HORIBA，日本）和 Na⁺测定仪（NA-11，HORIBA，日本）来测定生菜汁液中的 EC、K⁺和 Na⁺的含量（仪器图见附录 D 图 D1）。本章使用暗箱成像室系统来获取高光谱图像，高光谱图像具体的获取方法见 2.2.1 节，光谱反射率提取过程见 2.2.2 节。

6.2.2 多模态数据获取

本章所用的时间序列数据获取见 4.2.2 节描述。本章使用图像数据来源于高光谱，

并使用 SRGAN 提升高光谱图像分辨率，具体方法见 5.2.3 节描述。

6.2.3 光谱、时间序列以及图像的融合方法

(1) 特征级融合方法

在第三章、第四章和第五章中，深入探讨了如何基于不同数据模态对生菜的 SSC、pH 值和 SPAD 值进行精准估算的问题。在研究的初期阶段，面临的主要挑战在于选择合适的数据融合方法。在尝试直接采用数据级融合进行深度学习建模时，发现这种方法难以确定一个具有广泛适用性和稳定估算能力的网络框架。这可能有以下几个原因：在所使用的多模态数据中，光谱数据通常是一维的，包含不同光谱特征下的反射强度或辐射强度值，图像数据则是二维的，包含空间信息，而时间序列数据则是一维的时间序列，这些数据在维度上的差异意味着它们在空间或时间分布上有不同的特性，需要不同的网络架构来提取特征；不同数据模态通常具有不同的尺度和分辨率，例如时间序列的时间分辨率比光谱和图像的要高；光谱和图像数据是连续值，而时间序列数据是离散的；不同模态的数据可能含有不同类型和量级的噪声，例如，光谱数据中存在着光的散射影响，图像数据中含有相机的暗电流噪声，时间序列数据来源于三种传感器（RGB 相机、多光谱相机和激光雷达），这些传感器中都含有一些噪声。所以，不再寻求在数据层面直接融合不同模态的数据，而是为每种数据模态寻找最适合的网络架构，以确保能从光谱数据、时间序列数据和图像数据中提取出最有价值的信息。这一过程涉及了使用 Inception 网络处理光谱数据，利用 TD 网络处理时间序列数据，以及采用 MobileNet 作为特征提取网络的 SRGAN 技术对原始高光谱图像数据进行上采样，最终通过 BiT-M-R50*3 大型模型精细图像特征提取。然后，对这些深度学习网络提取的三种模态特征进行融合，建立特征融合模型（见附录 D 图 D2），用于后续的表型分析任务。所以本章基于前文探索的网络架构，使用光谱数据、时间序列数据和图像数据来对生菜的 EC、K⁺和 Na⁺进行估算。特征级融合方式的原理图见图 6.2A。本研究中，所有的基于光谱数据、时间序列数据和图像数据的深度学习特征融合模型统称为 SpecTimeImageNet。

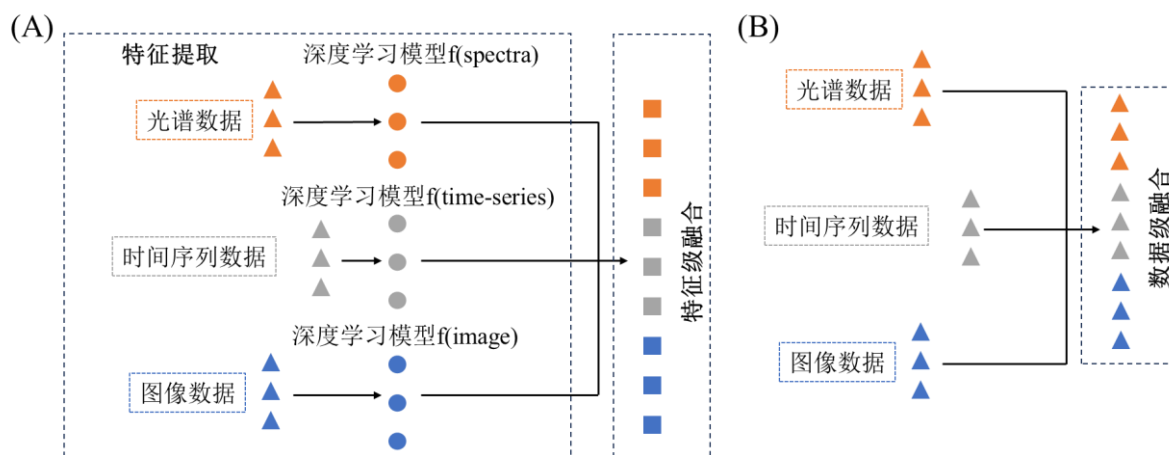


图 6.2 特征级 (A) 和数据级 (B) 融合原理图

Figure 6.2 Schematic of feature-level (A) and data-level (B) fusion

(2) 数据级融合方法

在实验的后期阶段，尽管深度学习模型在多模态数据融合方面展现了显著优势，但其建模过程通常需要较长的训练时间和更多的实验资源。为了确保研究结果的普适性和实用性，特别是针对那些可能缺乏深度学习经验或无法访问高性能计算资源的研究人员，本章在深度学习建模完成后，引入了传统机器学习模型作为一种辅助方法来探索多模态数据融合的潜力。这种策略旨在通过快速建立的机器学习模型，验证多模态数据融合相较于单一数据模态的优势，同时降低经济成本和技术门槛，使得更广泛的研究群体能够从这一研究结果中受益。

相比深度学习模型，传统机器学习模型通常对计算资源的需求更低，可以在不具备高性能 GPU 和广泛并行处理能力的标准计算机上有效地进行训练和部署。此外，对于数据量不大的情况，机器学习模型能够实现更快的训练和迭代周期，这对于需要快速验证研究假设的项目尤为重要。因此，采取了使用基于机器学习的模型来确定最有效的多模态数据融合方法，以此作为深度学习建模工作的补充和验证。值得注意的是，深度学习与传统机器学习在数据处理和特征提取方面存在本质区别，主要体现在处理数据和提取特征的方法上。深度学习能够通过端到端的学习实现数据级和特征级融合，而机器学习通常不具备这种端到端特征提取的能力，因此本实验中基于机器学习的模型采用了数据级融合的方法。通过这种方法，不仅可以深入地探讨多模态数据融合在精准农业领域的应用潜力，还确保了数据融合技术在更广泛的研究和应用场景中的可接受性和实用性。数据级融合方式的原理图见图 6.2B。本研究中，所有的基于光谱数据、时间序列数据和图像数据的机器学习融合模型统称为 FusionML。

6.2.4 AutoML 技术

在深度学习建模的过程中,为每种数据模态挑选最合适的网络架构是一个既耗时又充满挑战的任务。这项任务对于那些专长于农业科学领域、而非计算机科学的研究人员而言尤其艰巨。此外,当涉及到应用传统机器学习算法时,这些研究人员还面临着一个关键挑战:必须找出一种方法,在数据级融合的基础上有效地整合来自不同模态的数据,建立统一的分析框架。这些多模态数据中蕴含的丰富的线性与非线性特征对模型性能的影响至关重要,正确地解析这些信息是优化模型性能的关键。确保多模态数据的独特特性得以维持,同时准确地挑选和调整能够有效处理这些特征的机器学习模型,这一过程要求高度的技术知识和复杂的操作。这通常涉及到复杂且技术性强的手动模型选择和参数调优工作。随着 HTPP 平台在获得大量实时数据方面的应用日益增多,这些挑战提高了在 HTPP 领域应用深度学习和机器学习的门槛。对于时效性强的实验数据,将大量时间花费在选择网络架构或优化参数上不仅会造成数据资源的浪费,也会阻碍实验研究的进展。

为了应对深度学习和机器学习在植物分析应用中的复杂性和高门槛,特别是在面对大量、多模态的数据时,采用了 AutoML 技术,如 TPOT^[118]和 AutoKeras^[123],以简化模型的选择和调参过程。这些技术利用算法自动化探索最佳的数据处理流程和网络架构,显著减少了从数据预处理到模型优化所需的时间和专业知识要求。TPOT 是一个利用遗传算法来探索各种数据预处理和模型管道的 AutoML 技术,它能够在不同的机器学习算法中找到一个相对较优的模型来处理特定的数据集,尽管它可能不会找到一个同时适用于所有线性和非线性数据方面的单一模型。TPOT 数据处理流程如图 6.3 所示。AutoKeras 是一个针对深度学习的 AutoML 系统,它可以通过自动化搜索过程,优化网络架构以及调整参数,从而简化了手动网络设计的繁重工作。在这一过程中,它能够有效地探索多种可能的网络配置,识别出在特定数据集上表现最佳的深度学习模型。如图 6.4 所示,在 AutoKeras 系统中,用户通过调用 API 启动流程,随后搜索器便会在 CPU 上生成神经网络的架构。在 Keras 框架支持下,图模型(Graph)根据这些神经架构数据构筑出实际可运行的神经网络,并在 RAM 中初始化网络参数。接着,为了训练这些神经网络,系统会将它们复制到 GPU 上执行训练过程。训练完成后,调校过的神经网络便将储存至存储设备中,以备未来使用。

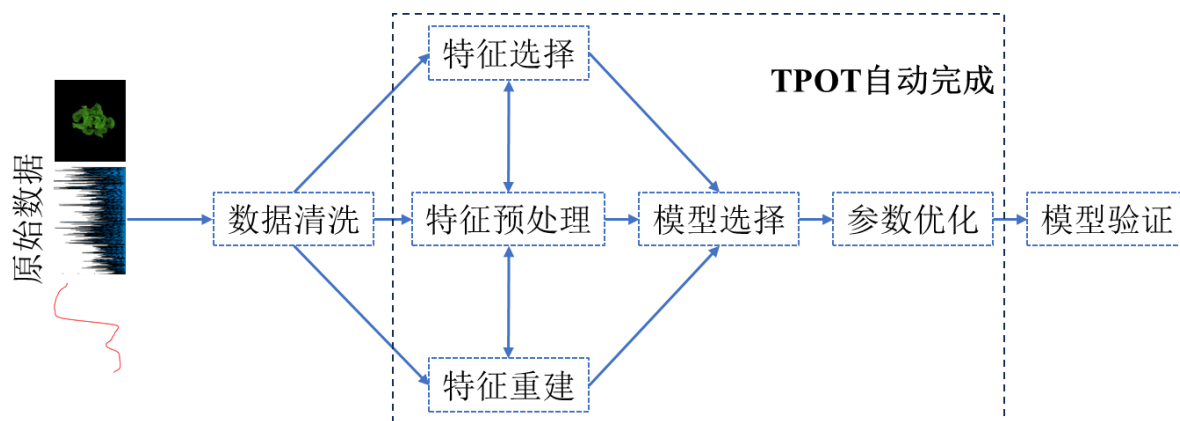


图 6.3 TPOT 的数据处理流程

Figure 6.3 Data process flow of TPOT

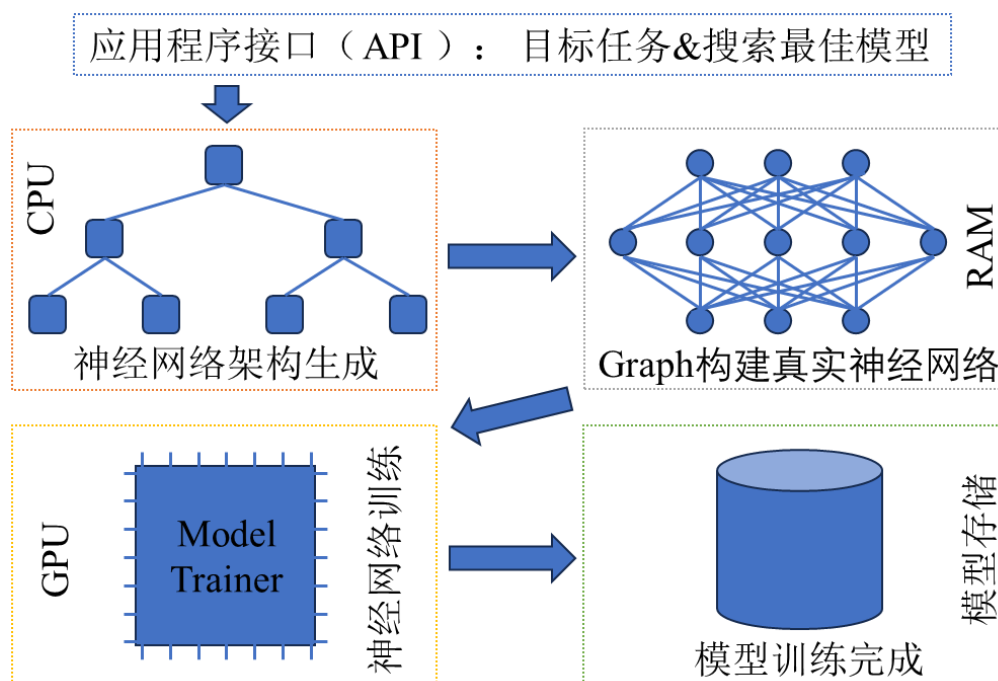


图 6.4 AutoKeras 系统流程图

Figure 6.4 Flow chart of AutoKeras system

AutoML 的应用可以提升研究效率，通过自动化流程优化模型选择和参数调整，显著减少建模过程中的人工干预需求。这一自动化策略能够快速有效地处理来自 HTTP 平台的多模态数据，加快表型信息的解析过程，并确保数据分析及表型估算的可靠性，为植物表型解析提供了一个可靠的技术支撑。通过 AutoML，即使是缺乏深厚计算机编程或数据科学背景的研究人员也能轻松地完成特征选择、模型选择、参数调优等复杂的数据科学任务，无需详细了解每个模型的技术细节。这种方法极大地降低了机器学习技术的应用难度，使研究人员可以更专注于模型的解读和应用，而不需深陷于模型选择和调优的技术复杂性中。

6.2.5 EC、K⁺和 Na⁺的估算模型建立

在最终的分析中，用于 EC、K⁺和 Na⁺建模的样本数量和第四章中 SSC 和 pH 值一样，都是 524 个。模型性能评价指标及其描述见 2.3.3 节。EC、K⁺和 Na⁺建模的样本划分方式和第四章中 SSC 和 pH 值一样，且随机数发生器的状态在所有建模中保持一致。深度学习建模是在 Python 3.8.13 和 TensorFlow 2.6.0 环境下进行，机器学习建模在 MATLAB2020（Mathworks, 美国）中进行。硬件配置与第五章相同。

6.3 结果分析和讨论

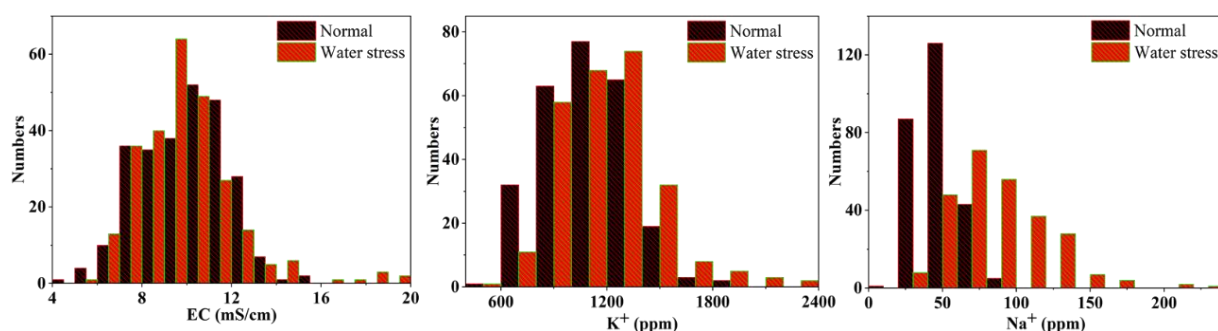
6.3.1 EC、K⁺和 Na⁺的参考值

表 6.1 和图 6.5 详细展示了 EC、K⁺和 Na⁺的统计数据及其分布范围，发现生菜在水分不足条件下，Na⁺的浓度有显著上升，而 K⁺的浓度和 EC 亦呈现上升趋势，但幅度不及 Na⁺那般明显。植物在水分匮乏的环境中会受到根部吸水能力的限制，导致土壤中的盐分相对浓度上升。这会造成植物细胞吸收 Na⁺，引起细胞内外 Na⁺浓度失衡。这一现象可能反映了植物在缺水条件下的抗逆性调节机制：通过调控 Na⁺的积累来维持细胞的渗透压，以稳定细胞内的水分含量。此外，K⁺通常在植物细胞内积聚，以保持细胞功能。在缺水的情形下，细胞内水分的降低可能导致 K⁺浓度的轻微上升，这是植物为减少 K⁺流失而作出的抗逆性响应。K⁺作为一种调节性离子，其功能包括减轻细胞内 Na⁺积累引发的离子毒性，通过置换 Na⁺以维持细胞内的离子平衡。缺水还可能导致植物细胞内的 EC 略有增加，因水分的减少可能会使离子浓度升高，从而增强了植物组织内电流的传导能力，进而轻微提高了 EC。

表 6.1 生菜中 EC、K⁺和 Na⁺的参考值

Table 6.1 Reference measure of EC, K⁺, and Na⁺ in lettuces

Phenotypes	Normal			Water Stress		
	Range	Mean	SD	Range	Mean	SD
EC (mS/cm)	4.915-15.45	9.9609	1.9418	5.7-19.31	9.8832	2.2590
K ⁺ (ppm)	520-1900	1085.2862	228.6807	550-2250	1172.7101	271.8956
Na ⁺ (ppm)	14.5-99	46.9275	13.9411	21-225	85.4351	31.7070

图 6.5 EC、K⁺和 Na⁺的参考值范围Figure 6.5 The reference values range of EC, K⁺, and Na⁺

6.3.2 基于特征级融合和定制化深度学习的估算模型

利用改进的 Inception 模块、TD、SRGAN 和 BiT-M-R50*3 模型分别处理光谱数据、时间序列数据和图像数据，估算生菜的 EC、K⁺和 Na⁺的结果如表 6.2 所示，Spec 代表光谱数据，Time 代表时间序列数据，Image 代表图像数据，SR 代表经过 SRGAN 提升分辨率后的图像数据。模型的相关参数设置见附录 D 表 D1-D3。可以看出，在估算生菜的 EC、K⁺和 Na⁺含量时，不同数据模态的融合对模型性能都产生了提升效果。这一点在将光谱数据、时间序列数据，以及经 SRGAN 提升分辨率后的图像数据融合时表现得最为突出。这也进一步证明了 SRGAN 方法的适用性，不仅能够有效提升流水线成像室拍摄的高光谱图像的空间分辨率，同样适用于提高暗箱成像室拍摄的高光谱图像的空间分辨率。融合这三种数据模态后，这三种表型的估算模型都达到了最优的效果（EC：R_c² = 0.8617，RMSEC = 0.7827，R_v² = 0.8654，RMSEV = 0.7346，R_p² = 0.8450，RMSEP = 0.8262，RPD = 2.5398；K⁺：R_c² = 0.8872，RMSEC = 85.4770，R_v² = 0.8620，RMSEV = 84.4233，R_p² = 0.8558，RMSEP = 96.4612，RPD = 2.6330；Na⁺：R_c² = 0.9607，RMSEC = 9.5691，R_v² = 0.8855，RMSEV = 9.2092，R_p² = 0.9030，RMSEP = 9.5582，RPD = 3.2104）。其中，Na⁺的最优模型是三种表型估算模型中 RPD 值唯一超过 3 的估算模型。图 6.6 展示了 EC、K⁺和 Na⁺的最佳估算结果的散点图，可以发现，即使是在 EC、K⁺和 Na⁺那些数值较大且位置偏离数据主体集群的离群点，它们与最佳拟合线的距离依然十分接近。这一现象印证了深度学习技术在处理量化有限、范围广的数据时的高效性。这表明，深度学习模型能够通过其学习算法精准捕捉到数据的潜在模式和关系，即便在数据稀疏或者分布不均的情况下依然能够做出准确的估算。这进一步验证了即使在样本数量有限的情况下，深度学习模型也具有优越的泛化能力，能够有效估算极端数值。

表 6.2 基于特征级融合估算生菜的 EC、K⁺和 Na⁺

Table 6.2 Estimating of EC, K⁺, and Na⁺ in lettuces based on feature-level fusion

Phenotypes	Data modalities	R _c ²	RMSEC	R _v ²	RMSEV	R _p ²	RMSEP	RPD
EC (mS/cm)	Spec	0.7847	0.9764	0.7774	0.9448	0.7602	1.0274	2.0423
	Spec+Time	0.8030	0.9340	0.7924	0.9124	0.7806	0.9829	2.1348
	Spec+Image	0.8223	0.8871	0.8168	0.8571	0.8098	0.9151	2.2931
	Spec+SR	0.8490	0.8179	0.8493	0.7774	0.8313	0.8619	2.4346
	Spec+Time+SR	0.8617	0.7827	0.8654	0.7346	0.8450	0.8262	2.5398
	Spec+Time+SR+Image	0.9332	0.7285	0.9332	0.6885	0.9332	0.7285	2.9332
K ⁺ (ppm)	Spec	0.8172	108.7870	0.7820	106.1278	0.7455	128.1266	1.9823
	Spec+Time	0.8637	93.9398	0.8467	88.9884	0.8235	106.6902	2.3805
	Spec+Image	0.8683	92.3407	0.8416	90.4519	0.8309	104.4309	2.4320
	Spec+SR	0.8789	88.5360	0.8600	85.0284	0.8400	101.5820	2.5003
	Spec+Time+SR	0.8872	85.4770	0.8620	84.4233	0.8558	96.4612	2.6330
Na ⁺ (ppm)	Spec	0.8419	12.4539	0.8267	11.3311	0.8286	12.7048	2.4153
	Spec+Time	0.8656	11.4806	0.8534	10.4219	0.8460	12.0430	2.5480
	Spec+Image	0.8649	11.5144	0.8467	10.6566	0.8566	11.6181	2.6412
	Spec+SR	0.9127	9.2557	0.8847	9.2435	0.8791	10.6703	2.8758
	Spec+Time+SR	0.9067	9.5691	0.8855	9.2092	0.9030	9.5582	3.2104

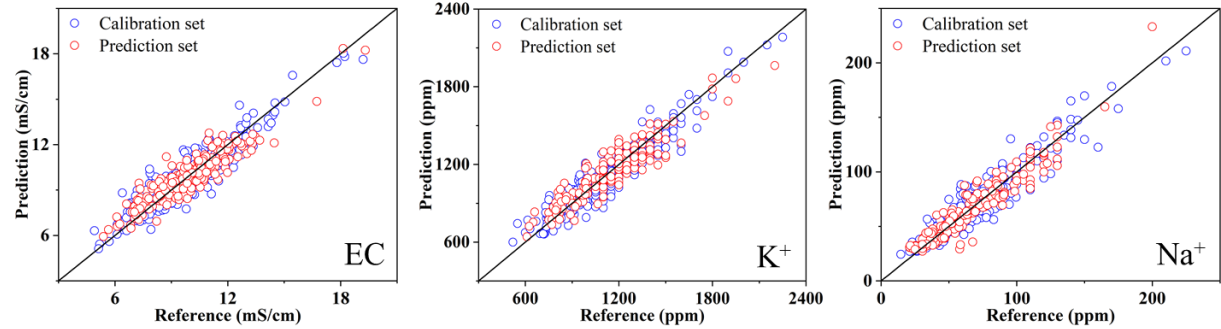


图 6.6 基于特征级融合的营养元素表型估算

Figure 6.6 Estimation of nutrient element phenotypes based on feature-level fusion

另外，随着数据模态数量的增加，观察到模型性能普遍提升，这表明多模态数据融合在提升建模精度方面起到了积极的作用^[106-117]。然而，不同数据模态对模型性能改善的贡献度各不相同^[106, 108, 113, 115, 117]。例如，在 EC 值的估算中，光谱数据的基线模型已经表现出了相对较低的 R_p^2 (0.7602)，这意味着当面对未经过训练的数据时，模型的估算能力较为有限。而在结合了时间序列数据之后，EC 的 R_p^2 提升了 2.68%。当加入图像数据后，EC 的估算性能进一步改善， R_p^2 提升了 6.52%。图像数据在经过 SRGAN 的处理后使得 EC 的 R_p^2 显著提升了 9.35%。最终三者融合带来了最大的提升 (11.15%)。与 EC 相比，K⁺的估算结果 (R_p^2) 在数据模态融合中的提升幅度更大，这可能表明 K⁺含量与数据模态中的特定特征有更为紧密的关联。对于 Na⁺而言，尽管其基线模型精度已经相当高，但融合多模态数据后的提升幅度并不如 EC，这可能说明 Na⁺的估算准确性受到了基线模型效果的上限制约。具体来说，在光谱数据的基础上，K⁺的时间序列数据，图像数据，SR 后的图像数据，三者融合的 R_p^2 分别提升了 10.46%、

11.46%、12.68%和 14.80%， Na^+ 分别提升了 2.1%、3.38%、6.09%和 8.98%。

在考虑单个图像数据模态的融合性能时，值得注意的是，SRGAN 能够显著提升三种表型的估算精度。这进一步证实了提升高光谱图像分辨率对于最大化利用 HSI 数据中包含的丰富光谱和空间信息的关键作用。因此，在随后的分析中，将专注于使用经 SRGAN 增强过的高光谱图像数据，而不再考虑其原始分辨率图像。此外，还观察到 SRGAN 对不同表型的估算影响呈现出一定的差异性。EC 和 Na^+ 的估算性能通过 SRGAN 提升后均有约 2.6% 的提升，而 K^+ 的提升则相对较小，仅为 1.1%。这说明，虽然超分辨率技术能够增强图像质量，从而提升模型性能，但其对不同生理表型的影响并非一致，可能与这些表型本身的生物学特性和图像数据中所包含的生物信息有关。将光谱数据、时间序列数据以及经 SRGAN 提高了分辨率的图像数据这三种不同数据模态融合起来时，可以发现融合后的数据模态对模型性能的贡献度并不是简单的线性叠加。这一观察提示了一个更为复杂的相互作用模式。可能的解释是，各个数据模态中可能含有重叠的植物表型特征，这些特征的复杂关系在融合后，现有的深度学习模型可能未能彻底揭示和充分利用这些关系。这表明在未来的建模过程中，不仅要探索新的数据融合技术，还需要探索新的模型架构和优化算法，以便能够更充分地挖掘和利用这些复杂数据中的所有有用信息。通过这些深入的分析，有助于更加准确地理解不同模型、不同数据模态、以及它们之间相互作用的复杂性，为未来研究的方向提供重要的指导。可以预期，通过在更广泛的数据集上进行验证，并应用更多样化的模型与融合方法，能够进一步提升模型的估算表现。

在分析不同数据模态对估算性能的相对贡献时，观察到在光谱数据中融入图像数据对于模型性能的提升作用普遍高于在光谱数据中融入时间序列数据。这一结果突显了图像数据在提供植物外部形态和纹理信息方面的独特价值，这些信息对于捕捉与目标表型密切相关的特征至关重要^[108]。时间序列数据的结合也提高了模型的估算性能，强调了其在捕捉生理层面关键信息方面的重要性。该实验结果证明了在植物表型分析中应用多模态数据融合方法的前景。尤其是在复杂的农业应用背景下，这种策略能显著提高估算精度和效率。未来可以探索更多的模型种类和优化方法，以进一步挖掘和利用融合后的数据模态中的有用信息，从而提升模型性能并实现在不同农业条件下的广泛应用。同时，考虑到实验设计的局限性，接下来需要在更广泛的数据集和更复杂的环境条件下验证所提出的模型，以确保其强大的泛化能力。

6.3.3 基于数据级融合和传统机器学习的估算模型

基于数据级融合的传统机器学习模型的结果如表 6.3 所示, 模型参数设置见附录 D 表 D4。对于每个给定的数据模态, 随着其他数据模态的融合, PLSR 和 SVR 两种模型的模型都得到了提升。和基于特征级融合的深度学习模型相同, 不管是 PLSR 模型还是 SVR 模型, 基于数据级融合的机器学习模型也是在光谱、时间序列和 SRGAN 重建后的图像三者融合的情况下, 每种表型都获得了最佳的估算结果 (EC: $R_c^2 = 0.8156$, RMSEC = 0.9036, $R_v^2 = 0.7500$, RMSEV = 1.0013, $R_p^2 = 0.7122$, RMSEP = 1.1257, RPD = 1.8640; K^+ : $R_c^2 = 0.7513$, RMSEC = 126.8863, $R_v^2 = 0.7335$, RMSEV = 117.3376, $R_p^2 = 0.6529$, RMSEP = 149.6278, RPD = 1.6974; Na^+ : $R_c^2 = 0.8626$, RMSEC = 11.6106, $R_v^2 = 0.8358$, RMSEV = 11.0300, $R_p^2 = 0.8467$, RMSEP = 12.0149, RPD = 2.5539)。然而, 基于机器学习的最优模型, 除了 Na^+ 外, EC 和 K^+ 的效果甚至没有仅使用光谱数据的深度学习模型好, 两种表型的最优模型的 RPD 值都低于 2。另外, 这两个模型在估算不同表型的表现有所差异。对于 EC, 每种数据模态的融合, 都是 PLSR 模型表现较好。而对于 Na^+ 含量的估算, 都是 SVR 模型较好。至于 K^+ , 除了光谱和时间序列的数据模态融合, PLSR 模型表现较好, 剩余三种融合方式都是 SVR 模型性能更为优异。基于数据级融合的机器学习模型估算 EC、 K^+ 和 Na^+ 的最优结果的散点图如图 6.7 所示, 与图 6.6 的结果相对照, 可以明显观察到 EC 和 K^+ 的数据点与拟合线之间的距离更大, 尤其是对于 K^+ , 估算误差有了明显的扩大。这表明, 虽然数据级融合技术在融合多种数据模态方面具有潜力, 但它在估算某些特定指标时可能还需要进一步的调整和优化以提高估算精度。特别是对于 K^+ , 可能需要更深入的特征工程来减少估算误差。

表 6.3 基于数据级融合估算生菜的 EC、K⁺和 Na⁺

Table 6.3 Estimating of EC, K⁺, and Na⁺ in lettuces based on data-level fusion

Phenotypes	Model	Data modalities	R _c ²	RMSEC	R _v ²	RMSEV	R _p ²	RMSEP	RPD
EC (mS/cm)	PLSR	Spec	0.7378	1.0776	0.7243	1.0514	0.6444	1.2513	1.6769
		Spec+Time	0.7806	0.9857	0.7232	1.0535	0.6588	1.2258	1.7118
		Spec+SR	0.7609	1.0291	0.7008	1.0953	0.6639	1.2165	1.7249
		Spec+Time+SR	0.8156	0.9036	0.7500	1.0013	0.7122	1.1257	1.8640
	SVR	Spec	0.7434	1.0660	0.7336	1.0335	0.6410	1.2573	1.6689
		Spec+Time	0.7631	1.0244	0.7362	1.0285	0.6458	1.2488	1.6802
		Spec+SR	0.7566	1.0383	0.6972	1.1018	0.6636	1.2170	1.7242
		Spec+Time+SR	0.8070	0.9246	0.7568	0.9875	0.7078	1.1342	1.8500
K ⁺ (ppm)	PLSR	Spec	0.7180	135.1196	0.7040	123.6666	0.6129	158.0259	1.6072
		Spec+Time	0.7284	132.6208	0.6849	127.5902	0.6309	154.2971	1.6461
		Spec+SR	0.7064	137.8799	0.6750	129.5671	0.6360	153.2294	1.6575
		Spec+Time+SR	0.7559	125.7204	0.7241	119.3881	0.6487	150.5317	1.6872
	SVR	Spec	0.7223	134.1019	0.6900	126.5450	0.6155	157.4831	1.6128
		Spec+Time	0.7299	132.2426	0.6864	127.2826	0.6263	155.2692	1.6357
		Spec+SR	0.7138	136.1254	0.6748	129.6202	0.6412	152.1313	1.6695
		Spec+Time+SR	0.7513	126.8863	0.7335	117.3376	0.6529	149.6278	1.6974
Na ⁺ (ppm)	PLSR	Spec	0.8617	11.6461	0.8198	11.5552	0.7584	15.0819	2.0346
		Spec+Time	0.7849	14.5263	0.7460	13.7171	0.7633	14.9284	2.0555
		Spec+SR	0.8591	11.7583	0.8625	10.0937	0.7809	14.3630	2.1364
		Spec+Time+SR	0.8489	12.1769	0.8224	11.4697	0.8133	13.2594	2.3142
	SVR	Spec	0.8515	12.0713	0.8298	11.2293	0.7820	14.3287	2.1415
		Spec+Time	0.8652	11.4998	0.8329	11.1283	0.7923	13.9838	2.1943
		Spec+SR	0.8317	12.8469	0.8086	11.9094	0.8086	13.4262	2.2855
		Spec+Time+SR	0.8626	11.6106	0.8358	11.0300	0.8467	12.0149	2.5539

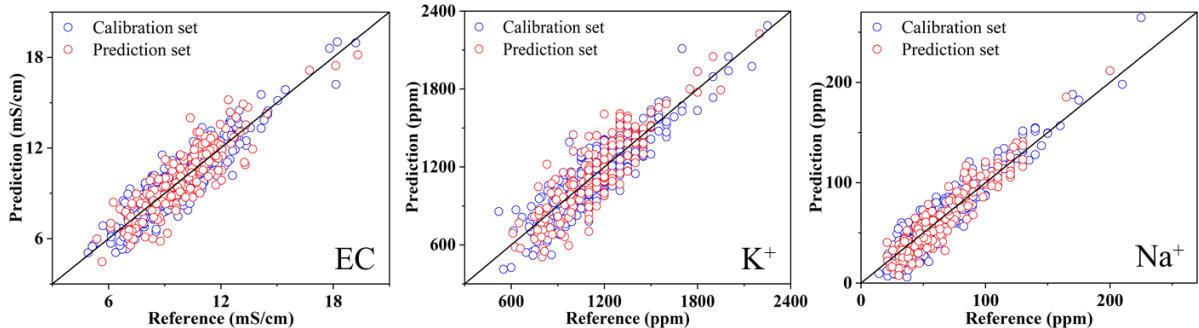


图 6.7 基于数据级融合的营养元素表型估算

Figure 6.7 Estimation of nutrient element phenotypes based on data-level fusion

另外，了解不同数据模态的相关性，并分析其线性或非线性特性，对于后续机器学习模型的选择具有重要作用。PLSR 模型擅长处理线性数据，尤其适用于处理变量众多或存在多重共线性的问题。而 SVR 模型则在处理复杂的非线性数据方面展现出优势，尤其是在传统线性模型难以捕捉数据的内在结构的情况下。在实际应用中，数据的线性或非线性特性可能并非完全独立，特别是在考虑多模态数据融合时，可能包含了线性和非线性关系的混合。这种情况下，集成学习方法可能是一个解决方案，它可以通过结合 PLSR 和 SVR 的优势，来充分发挥不同模型处理数据的潜力。因此，通过对数

据特性的深入分析和理解,可以更精准地选择适合特定数据集的机器学习模型,以优化模型性能,实现更可靠的估算结果。

对于每种数据模态对模型性能的贡献度,机器学习模型和深度学习模型,或者说数据级融合和特征级融合,结果都是一致的。图像数据的贡献度比时间序列数据的贡献度大,且不同模态的数据的贡献度不是彼此叠加的。以 R_p^2 为计算指标,对于 EC,时间序列对 PLSR 模型和 SVR 模型的性能分别提升了 2.24%、0.75%,SRGAN 重建后的图像则分别提升了 3.03%、3.53%;对于 K^+ ,时间序列对 PLSR 模型和 SVR 模型的性能分别提升了 2.94%、1.75%,SRGAN 重建后的图像则分别提升了 3.77%、4.18%;对于 Na^+ ,时间序列对 PLSR 模型和 SVR 模型的性能分别提升了 0.65%、1.32%,SRGAN 重建后的图像则分别提升了 2.97%、3.40%。可以发现,基于数据级融合的机器学习,时间序列数据和 SRGAN 重建后的图像数据对于模型精度的提升效果都低于基于特征级融合的深度学习。总体来说,数据级的融合提升了估算模型的性能,特别是在融合了 SRGAN 图像的情况下。这些发现强化了在植物表型分析中运用多模态数据融合潜力。通过这些比较,研究人员可以更好地理解不同数据模态在不同模型中如何影响估算的性能,并据此优化他们的估算模型。

通过对比特征级融合的深度学习模型和数据级融合的机器学习模型,可以得出结论,特征级融合的深度学习模型在所有情况下都优于数据级融合的机器学习模型。特征级融合通常能提供更高的估算准确性,这可能是因为特征级别上进行融合可以更有效地结合不同模态的信息,并发掘出更加复杂和深层的特征交互,这在建立估算模型时是非常宝贵的。这种比较分析有助于选择合适的数据融合方法。考虑到特征级融合的高性能,特别是当可用的数据类型和数量较多时,研究人员可能倾向于采用这种方法来进一步优化他们的估算模型。不过,这样的决策还需要同时考虑其他因素,如模型的复杂性、计算资源的可用性以及特定任务的需求。

6.3.4 基于多模态数据融合和 AutoML 的估算模型

为了有效管理和分析 HTPP 平台产生的海量数据,本章研究了 AutoML 在 HTPP 领域的实际应用潜力。表 6.4 中呈现了利用 AutoML 技术估算 EC、 K^+ 和 Na^+ 的结果,AutoML 的相关参数见附录 D 表 D5。

表 6.4 基于 AutoML 估算生菜的 EC、K⁺和 Na⁺Table 6.4 Estimating of EC, K⁺, and Na⁺ in lettuces based on AutoML

Phenotypes	AutoML	R_c^2	RMSEC	R_v^2	RMSEV	R_p^2	RMSEP	RPD
EC (mS/cm)	TPOT	0.8175	0.8991	0.7544	0.9924	0.7606	1.0267	2.0438
	AutoKeras	0.8547	0.8022	0.8542	0.7645	0.8399	0.8395	2.4994
K ⁺ (ppm)	TPOT	0.7526	126.5617	0.7265	118.8560	0.7045	138.0683	1.8395
	AutoKeras	0.8846	86.4447	0.8699	81.9899	0.8496	98.4882	2.5788
Na ⁺ (ppm)	TPOT	0.8666	11.4411	0.8546	10.3807	0.8538	11.7345	2.6150
	AutoKeras	0.9036	9.7256	0.8851	9.2280	0.8896	10.1938	3.0102

对于机器学习模型的估算能力, TPOT 得到的模型在三种不同表型上都展现出了性能的提升, 并且这种提升是基于表 6.3 中记录的最优结果而实现的 (EC: $R_c^2 = 0.8175$, RMSEC = 0.8991, $R_v^2 = 0.7544$, RMSEV = 0.9924, $R_p^2 = 0.7606$, RMSEP = 1.0267, RPD = 2.0438; K⁺: $R_c^2 = 0.7526$, RMSEC = 126.5671, $R_v^2 = 0.7265$, RMSEV = 118.8560, $R_p^2 = 0.7045$, RMSEP = 138.0683, RPD = 1.8395; Na⁺: $R_c^2 = 0.8666$, RMSEC = 11.4411, $R_v^2 = 0.8546$, RMSEV = 10.3807, $R_p^2 = 0.8538$, RMSEP = 11.7345, RPD = 2.6150)。在估算 EC 的模型中, TPOT 提供了唯一实现 RPD 值超过 2 的模型。基于 TPOT 的三种表型估算结果的散点图如图 6.8 所示。与图 6.7 展示的结果相比, Na⁺数据点的估算分布未见显著变化。在 EC 的估算中, 数据点更紧密地聚集于拟合线周围, 这显示出估算误差有所降低。对于 K⁺的估算, 虽然总体上数据点更趋向于拟合线, 表明了主体数据的估算准确性增强, 但离群点的估算误差却略有扩大, 这可能指出 TPOT 生成的 K⁺估算模型对于极端值的处理需进一步优化。TPOT 为三种表型选择的模型及相应模型参数见附录 D 表 D6。

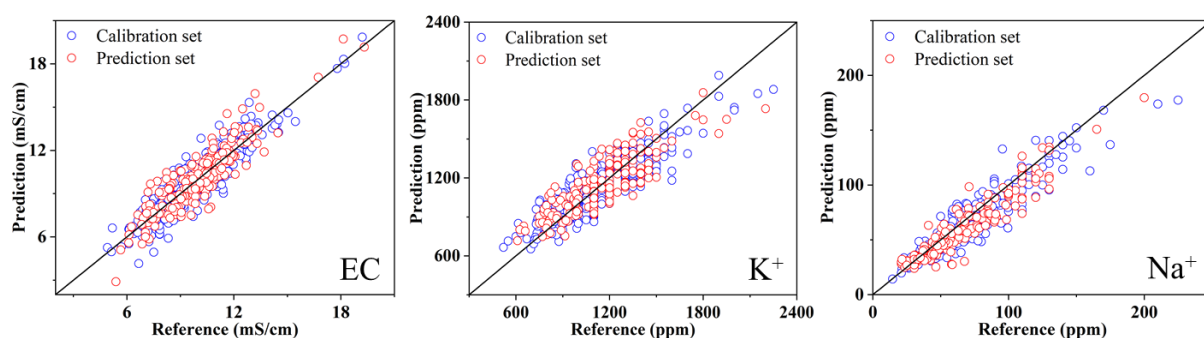


图 6.8 基于 TPOT 的营养元素表型估算

Figure 6.8 Estimation of nutrient element phenotypes based on TPOT

AutoKeras 通过其自动化搜索功能, 在估算三种表型方面展现出了卓越的性能, 其精度优于 TPOT 模型的结果。虽然这些由 AutoKeras 搜索得到的最佳模型并未能达到之前利用特定网络结构为不同数据模态量身定制的模型的高水平, 其 R_p^2 分别降低了 0.6%

(EC)、0.7% (K^+)、1.5% (Na^+)，但它们的表现仍紧随那些融合了光谱、时间序列数据和 SRGAN 重建图像的最优结果之后。AutoKeras 获得的 EC、 K^+ 和 Na^+ 的估算结果表现如下：EC 的 R_c^2 为 0.8547，RMSEC 为 0.8022， R_v^2 为 0.8542，RMSEV 为 0.7645， R_p^2 为 0.8399，RMSEP 为 0.8395，RPD 为 2.4994； K^+ 的 R_c^2 为 0.8846，RMSEC 为 86.4447， R_v^2 为 0.8699，RMSEV 为 81.9899， R_p^2 为 0.8496，RMSEP 为 98.4882，RPD 为 2.5788； Na^+ 的 R_c^2 为 0.9036，RMSEC 为 9.7256， R_v^2 为 0.8851，RMSEV 为 9.2280， R_p^2 为 0.8896，RMSEP 为 10.1938，RPD 为 3.0102。图 6.9 展示了基于 AutoKeras 的三种表型估算结果的散点图。相较于图 6.8 中的估算结果，所有表型的数据点都更紧凑地围绕着拟合线排列，无论是密集的数据群集还是偏离中心的离散点，这证明了深度学习在泛化方面的出色表现。与图 6.6 中深度学习模型的结果相比较，三种表型的数据点群集分布基本保持一致，但在离群点的估算准确性上，SpecTimeImageNet 模型要优于 AutoKeras 的生成的模型，这一差异凸显了 SpecTimeImageNet 模型在处理边缘案例时的潜在优势。值得强调的是，寻找合适的网络结构以适应这些多模态数据通常需要大量的时间。然而，AutoKeras 显示出了巨大的潜力，其与 SpecTimeImageNet 模型性能的接近性表明使用这种技术可以在不牺牲太多准确性的情况下能够在较短的时间内自动发现有效的解决方案，这一发现对于模型优化的效率提升有着重要意义^[158]。简而言之，AutoKeras 的结果缩小了与定制化模型在估算精度上的差距，同时其在时间效率方面的优势不容忽视，它能够快速提供与定制化模型近似的高精度估算结果，这一点对于 HTPP 领域的数据分析具有重要意义^[141]。AutoKeras 为三种表型选择的网络架构见附录 D 表 D7。

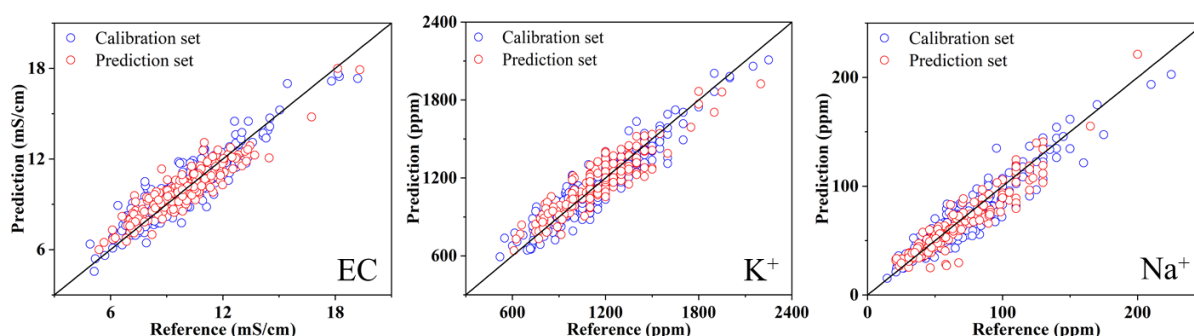


图 6.9 基于 AutoKeras 的营养元素表型估算

Figure 6.9 Estimation of nutrient element phenotypes based on AutoKeras

在农业科学领域，AutoKeras 和 TPOT 这两种 AutoML 技术在 HTPP 分析上展现了卓越的应用前景。AutoKeras 所提供的模型与定制深度学习模型在性能上十分接近，而

TPOT 能够自动识别出最优的机器学习模型并对参数进行优化, 其估算精度超越了传统的手动模型寻优方法。这两种 AutoML 技术, AutoKeras 的网络结构和参数调整策略, 似乎在探索多样化模型配置方面更为高效。尽管 TPOT 的性能未能达到 AutoKeras 的水平, 但它仍然证明了自身作为对传统机器学习方法的有效补充的价值。然而, AutoKeras 的通用设计并未充分考虑到农业数据的特殊性, 例如空间异质性、时间序列变化的复杂性以及与环境因素的紧密联系。这些特定的数据特征会影响模型的性能, 因此, 对于农业数据的分析, AutoKeras 构建的模型或许需要进一步的优化才能真正发挥其最大潜力。这意味着先使用 AutoKeras 快速搭建起模型框架, 然后通过细致的参数微调, 来增强模型的估算准确度。尽管 AutoML 简化了机器学习的过程, 但它并不能完全替代研究人员对于自身数据和研究问题的深入了解。研究人员需要利用 AutoML 技术合理构建模型、解释模型结果, 并将其应用于科学研究中^[19, 20]。在此过程中, AutoML 的易用性和高效性为那些非专业用户开放了机器学习技术的大门, 而对于专业用户来说, 则是实现快速模型构建和探索的有力平台。

综上, AutoML 作为一个自动化的模型搜索和参数寻优技术, 尽管在面对特殊数据集时可能需要进一步的调整, 但它提供的便利和效率优势, 在快速发展且数据量大幅增加的 HTPP 领域具有重要意义。同时, 它也鼓励用户在使用过程中加深对模型工作原理的理解, 并且根据特定任务需求进行必要的手动干预, 以达到最优的研究结果。

6.4 本章小结

本章深入探讨了利用光谱数据、时间序列数据和图像数据进行多模态数据融合以及 AutoML 技术应用于生菜营养元素 (EC 、 K^+ 和 Na^+) 估算的有效性。基于前三章关于生菜 SSC、pH 值和 SPAD 值精准估算的探索, 本章进一步深化了对多模态数据融合方法的理解, 开展了不同数据模态的融合方式及其对表型估算模型贡献的系统性研究。

在探索深度学习模型的过程中, 直接的数据级融合因数据模态间的维度、尺度和分辨率差异以及不同类型和量级的噪声而面临挑战, 导致难以捕捉到有效估算特征。因此, 第四章、第五章和本章的深度学习模型构建都采用了特征级融合方法, 针对每种数据模态选择最适合的网络架构提取有价值的特征, 有效克服数据模态间差异, 提升估算准确性。特别是, 综合应用光谱数据、时间序列数据和经 SRGAN 技术处理的图像数据的 SpecTimeImageNet 模型在营养元素估算上显著提升性能, 其 R_p^2 分别为 0.8450 (EC)、0.8558 (K^+) 和 0.9030 (Na^+), 其中 Na^+ 的 RPD 值超过了 3, 证明了深

度学习模型在处理多模态复杂数据时的强大能力。此外，对不同模态数据组合的探索揭示，光谱与图像数据的融合效果优于光谱与时间序列数据，尤其是 SRGAN 提升的图像数据对估算精度的贡献显著。

在深度学习和机器学习建模过程中，选择合适的模型以及后续的参数调优往往既耗时又复杂，这不利于其在 HTTP 数据处理方面的应用。为此，本章引入了 AutoML 技术，展示了它们在简化模型选择和参数调整过程中的应用价值。其中，TPOT 在提升 FusionML 模型性能方面表现突出，而 AutoKeras 虽精度略降但简化了模型构建过程，降低了时间成本。这一进展显示了 AutoML 技术在促进高效、精确的表型估算技术发展中的作用。

综上所述，本章不仅进一步证实了多模态数据融合在植物表型估算中的应用潜力，也展示了 AutoML 技术在简化模型构建和优化过程中的价值，为精准农业领域提供了有效的技术支持。未来的工作可以继续探索不同的融合方法和其他 AutoML 技术的应用，进一步提升农业研究的精度和效率，为精准农业的发展贡献更多力量。

第七章 总结与展望

7.1 总结

本文围绕精准农业和植物表型分析的需求，利用高光谱成像技术和 HTPP 平台，深入探索了多模态数据融合技术结合深度学习模型在快速、精确及非破坏性表型特性估算方面的应用。结果证明，本文不仅显著提升了估算生菜六种关键表型特性的准确度，如 SSC、pH 值、SPAD 值、EC、 K^+ 和 Na^+ ，而且还简化了传统的数据处理过程，加快了表型分析的效率。主要总结如下：

本文基于 Inception 的 SpecNet2D 模型和基于 FCNN 设计的 SpecNet1D 模型在估算生菜 SSC 和 pH 值中取得了 0.9030 和 0.8490 的 R_p^2 ，超越了经过复杂特征工程的五种机器学习模型，显示出深度学习模型在光谱数据特征提取方面的显著优势，为后续多模态数据的特征提取打下了良好的基础。

为进一步提升深度学习模型在表型估算方面的能力，本文将时间序列数据输入 RNN 架构，并与改进的 SpecNet2D 模型和光谱数据相结合，建立了 SpecTimeNet 模型，实现了对生菜品质表型的精确估算，并分析了不同 RNN 架构对 SpecTimeNet 模型的影响。实验结果揭示，虽然四种 RNN 架构均优化了模型对 SSC 估算的性能，但是 TD 以 0.8900 的 R_p^2 表现最佳，而在 pH 值的估算中，仅 TD 能显著提升模型性能，达到 0.9583 的 R_p^2 。这一发现突显了挑选合适 RNN 架构对提升融合模型性能的重要性，同时证实了 TD 在处理时间序列数据特征提取方面的高效性。这表明深度学习模型通过有效融合光谱数据和时间序列数据，相比仅使用光谱数据，能显著提升品质表型的估算性能。

在探索提升植物光合效能表型特性估算精度的过程中，本文深入挖掘了图像数据中蕴含的丰富信息。结果证明，图像数据的加入显著增强了模型的估算能力，SpecImageNet 模型对 SPAD 值的 R_p^2 达到了 0.8620。为了进一步提升 SpecImageNet 模型的估算性能，使用 SRGAN 提升高光谱图像的空间分辨率以增强图像内容的细节丰富度，为图像特征的提取提供更精细的视觉基础，并采用 CAM 方法可视化改进的 SpecNet2D 模型所关注的光谱区域，以识别关键的光谱特征。经过 SRGAN 和 CAM 联合优化的 SpecImageNet 模型在 SPAD 值的估算中取得了 0.9033 的 R_p^2 ，展示了 SRGAN 与 CAM 在促进光谱与图像数据深度融合方面的显著潜力。这不仅提高了模型的可解释

性，也为精确估算植物光合效能表型特性提供了有效的技术途径。

本研究中，将光谱数据、时间序列数据和图像数据经过各自优化的处理架构进行分析，获得了营养元素表型的最优估算结果，其 R_p^2 分别为 0.8450 (EC)、0.8558 (K^+) 和 0.9030 (Na^+)。在评估不同数据模态对模型性能的影响时，结果表明综合应用三种数据模态的融合模型超越了任何两种数据模态融合的效果，其中图像数据在提升模型估算性能方面的贡献显著优于时间序列数据。面对多模态数据背景下，深度学习与传统机器学习模型在 HTPP 应用中面临的模型选择和参数调优耗时问题，提出了优化策略。采用 AutoML 工具，包括 TPOT 和 AutoKeras，有效降低了模型构建和参数调优的复杂度，尤其是在处理 HTPP 平台的数据时，展示了其良好的应用价值。TPOT 全面提升了传统机器学习模型的性能，AutoKeras 的效果虽然比 SpecTimeImageNet 模型稍差，但其简化了模型构建过程和参数调优时间。

通过对多模态数据融合技术的深入探索和实践，本研究显著提升了植物表型特性估算的准确度，并验证了深度学习技术在处理复杂农业问题和多模态数据上的适应性和有效性。通过对不同数据模态处理方法的优化，不仅促进了快速、精确以及非破坏性植物表型估算领域的重要进展，也为 HTPP 领域中深度学习和 AutoML 技术的应用提供了有力的策略和参考，推动了精准农业技术的发展。

7.2 展望

未来的植物表型估算方法研究有着广阔的拓展空间和深化潜力，主要集中在以下几个方面：

未来研究的重点之一将是对数据融合方法的进一步探索。当前研究已经证明，采用多模态数据融合能显著提升模型性能。未来的研究将探讨包括决策级融合在内的其他数据融合方法，从而更深入地挖掘不同数据模态之间的相互联系，为精确估算植物表型特性提供更加详尽和细致的数据支持。此外，不同数据模态对模型性能的贡献度各不相同，因此在融合时可以考虑为不同数据模态分配不同的权重，从而最大化模型的估算性能和可靠性。

在多模态数据的获取和优化方面，本研究揭示了图像数据对模型性能的提升效果优于时间序列数据。因此，未来的研究可以探索通过提高时间序列数据的时间分辨率来观察其对模型的贡献，以便捕获更细致的时间序列变化，进一步丰富模型对植物生长动态的解读。通过优化时间序列数据的采集方法，有望在保持数据融合优势的同时，

平衡不同数据模态对模型性能的贡献，进一步提升植物表型估算的准确性和效率。同时，本研究主要采用了近地面的暗箱、轨道和流水线植物表型平台，这些近地平台提供了高分辨率的图像，适合捕捉植物的微观特征并在受控环境中进行精准数据采集。然而，这些方法在数据采集的范围和速度上有限，可能不适用于大规模田间研究。因此，未来工作中，将考虑结合更大尺度的数据采集手段，如无人机和卫星等，进行多平台的多模态数据融合。无人机和卫星等远地平台能够覆盖广阔的田间区域，提高数据采集效率，并提供了新的视角和尺度以捕捉作物的宏观特征和生长趋势。这种基于更大尺度的多平台的多模态数据融合不仅可以扩展数据采集的范围和尺度，而且可以为模型提供更加多维和全面的数据支持，从而进一步提升模型的估算能力和适用于更广泛的应用场景。通过充分利用近地平台的高分辨率优势与远地平台的广覆盖范围，可以实现对作物表型特性更全面和精确的分析，进而为精准农业提供有效的技术支撑。

特征工程对于挖掘植物表型与多模态数据间复杂关系的深入理解具有重要作用。尽管本研究已经探索并验证了针对光谱、图像和时间序列数据的最优处理方法，但在深度学习模型众多的架构中，可能还存在未被发掘的潜力。因此，未来研究的方向之一是借助 AutoML 技术优化特征工程，为模型架构的选择和优化提供辅助。此外，面对多样化的生态环境和气候条件，构建具有更强泛化能力的模型显得尤为重要。这不仅要求模型能够适应不同地理位置和环境因素的变化，也要求模型能够准确地估算和分析不同条件下的植物表型。在这一过程中 AutoML 技术可以帮助研究人员在模型架构和参数选择上做出更精确的决策，从而加速模型的构建和验证过程。

提高模型解释性也是未来研究的重要方向。尽管深度学习模型在估算性能上已取得显著成果，其“黑箱”特性却成为了科学家和农业工作者的一个重点关注领域。本研究已经在光合效能表型的估算中，探索了通过 CAM 技术来增强模型的可解释性，通过可视化深度学习模型关注的区域来揭示关键的光谱特征对模型的影响。未来的工作将继续沿这一方向发展，研究更多透明的模型解释方法，帮助用户更深入地理解模型的决策过程。这不仅能增加对模型估算结果的信任度和接受度，也为深度学习在精准农业中的应用提供 stronger 科学基础和实践指导。

最后，实时监控与自动化决策支持的实现，以及抗逆性品种的筛选和多尺度表型分析的深入，将是推动智慧农业发展的关键。这些研究方向不仅有助于实现对作物生长状态的实时评估和精细管理，同时为耐旱、抗病等逆境条件下表型特性的研究和新品种的培育提供了科学基础。通过这些方向的深入探索和技术创新，深度学习与多模

态数据融合方法有望为精准农业、植物育种以及环境应对策略提供更加强大的科学支撑，推动可持续农业生产向高效、精准和智能的方向发展。

参考文献

- [1] Zhang P, Guo Z, Ullah S, et al. Nanotechnology and artificial intelligence to enable sustainable and precision agriculture[J]. *Nature Plants*, 2021, 7(7): 864-876.
- [2] Noumedem J A K, Djeussi D E, Hritcu L, et al. *Lactuca sativa*[M]. *Medicinal spices and vegetables from Africa*. Academic Press, 2017: 437-449.
- [3] Gui X, Zhang Z, Liu S, et al. Fate and phytotoxicity of CeO₂ nanoparticles on lettuce cultured in the potting soil environment[J]. *PloS one*, 2015, 10(8): e0134261.
- [4] Eshkabilov S, Lee A, Sun X, et al. Hyperspectral imaging techniques for rapid detection of nutrient content of hydroponically grown lettuce cultivars[J]. *Computers and Electronics in Agriculture*, 2021, 181: 105968.
- [5] Bianchi G, Falcinelli B, Tosti G, et al. Taste quality traits and volatile profiles of sprouts and wheatgrass from hulled and non-hulled *Triticum* species[J]. *Journal of food biochemistry*, 2019, 43(7): e12869.
- [6] Pérez-Marín D, Torres I, Entrenas J A, et al. Pre-harvest screening on-vine of spinach quality and safety using NIRS technology[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, 207: 242-250.
- [7] Khan T A, Yusuf M, Ahmad A, et al. Proteomic and physiological assessment of stress sensitive and tolerant variety of tomato treated with brassinosteroids and hydrogen peroxide under low-temperature stress[J]. *Food chemistry*, 2019, 289: 500-511.
- [8] Cha S J, Park H J, Kwon S J, et al. Early detection of plant stress using the internal electrical conductivity of *Capsicum annuum* in response to temperature and salinity stress[J]. *Plant Growth Regulation*, 2021, 95: 371-380.
- [9] Keisham M, Mukherjee S, Bhatla S C. Mechanisms of sodium transport in plants—progresses and challenges[J]. *International journal of molecular sciences*, 2018, 19(3): 647.
- [10] Xie C, Yang C. A review on plant high-throughput phenotyping traits using UAV-based sensors[J]. *Computers and Electronics in Agriculture*, 2020, 178: 105731.
- [11] Feng L, Chen S, Zhang C, et al. A comprehensive review on recent applications of unmanned aerial vehicle remote sensing with various sensors for high-throughput plant

- phenotyping[J]. Computers and electronics in agriculture, 2021, 182: 106033.
- [12] Tao H, Xu S, Tian Y, et al. Proximal and remote sensing in plant phenomics: 20 years of progress, challenges, and perspectives[J]. Plant Communications, 2022, 3(6): 39.
- [13] 赵春江. 植物表型组学大数据及其研究进展[J]. 农业大数据学报, 2019, 1(2): 5-14.
- [14] Lebedev V G, Lebedeva T N, Chernodubov A I, et al. Genomic selection for forest tree improvement: Methods, achievements and perspectives[J]. Forests, 2020, 11(11): 1190.
- [15] Bian L, Zhang H, Ge Y, et al. Closing the gap between phenotyping and genotyping: review of advanced, image-based phenotyping technologies in forestry[J]. Annals of Forest Science, 2022, 79(1): 22.
- [16] Sun D, Robbins K, Morales N, et al. Advances in optical phenotyping of cereal crops[J]. Trends in plant science, 2022, 27(2): 191-208.
- [17] Sehwal V, Bhagoji A N, Song L, et al. Analyzing the robustness of open-world machine learning[C]. Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security. 2019: 105-116.
- [18] Arya S, Sandhu K S, Singh J, et al. Deep learning: As the new frontier in high-throughput plant phenotyping[J]. Euphytica, 2022, 218(4): 47.
- [19] Koh J C O, Spangenberg G, Kant S. Automated machine learning for high-throughput image-based plant phenotyping[J]. Remote Sensing, 2021, 13(5): 858.
- [20] Shoaib M, Sharma N, Kotthoff L, et al. AutoML: advanced tool for mining multivariate plant traits[J]. Trends in Plant Science, 2023.
- [21] Miao C, Pages A, Xu Z, et al. Semantic segmentation of sorghum using hyperspectral data identifies genetic associations[J]. Plant Phenomics, 2020, 2(1): 11.
- [22] Varshney R K, Thudi M, Pandey M K, et al. Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: integrating genomics, phenotyping, systems modelling and agronomy[J]. Journal of Experimental Botany, 2018, 69(13): 3293-3312.
- [23] Bai G, Ge Y, Soby D, et al. NU-Spidercam: A large-scale, cable-driven, integrated sensing and robotic system for advanced phenotyping, remote sensing, and agronomic research[J]. Computers and Electronics in Agriculture, 2019, 160: 71-81.
- [24] Mulugeta Aneley G, Haas M, Köhl K. LIDAR-based phenotyping for drought response

- p and drought tolerance in potato[J].
- Potato Research*
- , 2023, 66(4): 1225-1256.
- [25] 张慧春, 周宏平, 郑加强, 等. 植物表型平台与图像分析技术研究进展与展望[J]. *农业机械学报*, 2020, 51(3): 1-17.
- [26] Goggin F L, Lorence A, Topp C N. Applying high-throughput phenotyping to plant–insect interactions: picturing more resistant crops[J]. *Current Opinion in Insect Science*, 2015, 9: 69-76.
- [27] Young S N, Kayacan E, Peschel J M. Design and field evaluation of a ground robot for high-throughput phenotyping of energy sorghum[J]. *Precision Agriculture*, 2019, 20(4): 697-722.
- [28] Yang W, Guo Z, Huang C, et al. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice[J]. *Nature communications*, 2014, 5(1): 5087.
- [29] Bai G, Ge Y, Blecha S, et al. Phenotyping transgenic wheat in a greenhouse through multispectral and thermal imaging[C]. 2015 ASABE Annual International Meeting. American Society of Agricultural and Biological Engineers, 2015: 1.
- [30] Virlet N, Sabermanesh K, Sadeghi-Tehran P, et al. Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring[J]. *Functional plant biology*, 2016, 44(1): 143-153.
- [31] McMahon P, Vadez V, Kholova J, et al. Leasyscan-a Novel Semi-Field Platform to Phenotype Traits Controlling Plant Water Budget[J]. *ASA, CSSA and SSSA International Annual (2017)*, 2017.
- [32] Kirchgessner N, Liebisch F, Yu K, et al. The ETH field phenotyping platform FIP: a cable-suspended multi-sensor system[J]. *Functional Plant Biology*, 2016, 44(1): 154-168.
- [33] Mesas-Carrascosa F J, Rumbao I C, Berrocal J A B, et al. Positional quality assessment of orthophotos obtained from sensors onboard multi-rotor UAV platforms[J]. *Sensors*, 2014, 14(12): 22394-22407.
- [34] Quigley M, Goodrich M A, Griffiths S, et al. Target acquisition, localization, and surveillance using a fixed-wing mini-UAV and gimbaled camera[C]. *Proceedings of the 2005 IEEE international conference on robotics and automation*. IEEE, 2005: 2600-2605.
- [35] Wendel J, Meister O, Schlaile C, et al. An integrated GPS/MEMS-IMU navigation

- system for an autonomous helicopter[J]. Aerospace science and technology, 2006, 10(6): 527-533.
- [36] Lemko O L, Molodchyk O D. Aerodynamic and flight characteristics of prospective UAV scheme “flying wing”[C]. 2015 IEEE International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD). IEEE, 2015: 33-36.
- [37] Smith G C S, Pell J P. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials[J]. British Medical Journal, 2003, 327(7429): 1459-1461.
- [38] Leonard N E, Paley D A, Davis R E, et al. Coordinated control of an underwater glider fleet in an adaptive ocean sampling field experiment in Monterey Bay[J]. Journal of Field Robotics, 2010, 27(6): 718-740.
- [39] Yamasaki T, Goto N. Identification of blimp dynamics via flight tests[J]. Transactions of the Japan society for aeronautical and space sciences, 2003, 46(153): 195-205.
- [40] Jing R, Gong Z, Zhao W, et al. Above-bottom biomass retrieval of aquatic plants with regression models and SfM data acquired by a UAV platform—A case study in Wild Duck Lake Wetland, Beijing, China[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 134: 122-134.
- [41] Sugiura R, Itoh A, Nishiwaki K, et al. Development of high-throughput field phenotyping system using imagery from unmanned aerial vehicle[C]. 2015 ASABE Annual International Meeting. American Society of Agricultural and Biological Engineers, 2015: 1.
- [42] Bendig J, Bolten A, Bennertz S, et al. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging[J]. Remote sensing, 2014, 6(11): 10395-10412.
- [43] Kerkech M, Hafiane A, Canals R. Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach[J]. Computers and Electronics in Agriculture, 2020, 174: 105446.
- [44] Maes W H, Steppe K. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture[J]. Trends in plant science, 2019, 24(2): 152-164.
- [45] Xie C, Yang C, He Y. Hyperspectral imaging for classification of healthy and gray mold

- diseased tomato leaves with different infection severities[J]. Computers and electronics in agriculture, 2017, 135: 154-162.
- [46] Zarco-Tejada P J, González-Dugo V, Berni J A J. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera[J]. Remote sensing of environment, 2012, 117: 322-337.
- [47] Bian J, Zhang Z, Chen J, et al. Simplified evaluation of cotton water stress using high resolution unmanned aerial vehicle thermal imagery[J]. Remote Sensing, 2019, 11(3): 267.
- [48] Barbagallo R P, Oxborough K, Pallett K E, et al. Rapid, noninvasive screening for perturbations of metabolism and plant growth using chlorophyll fluorescence imaging[J]. Plant physiology, 2003, 132(2): 485-493.
- [49] Wang D, Xin X, Shao Q, et al. Modeling aboveground biomass in Hulunber grassland ecosystem by using unmanned aerial vehicle discrete lidar[J]. Sensors, 2017, 17(1): 180.
- [50] Andújar D, Moreno H, Bengochea-Guevara J M, et al. Aerial imagery or on-ground detection? An economic analysis for vineyard crops[J]. Computers and electronics in agriculture, 2019, 157: 351-358.
- [51] Weng S, Yu S, Dong R, et al. Nondestructive detection of storage time of strawberries using visible/near-infrared hyperspectral imaging[J]. International Journal of Food Properties, 2020, 23(1): 269-281.
- [52] Weng S, Yu S, Guo B, et al. Non-destructive detection of strawberry quality using multi-features of hyperspectral imaging and multivariate methods[J]. Sensors, 2020, 20(11): 3074.
- [53] Duarte-Carvajalino J M, Silva-Arero E A, Góez-Vinasco G A, et al. Estimation of water stress in potato plants using hyperspectral imagery and machine learning algorithms[J]. Horticulturae, 2021, 7(7): 176.
- [54] Alvarez J, Martinez E, Diezma B. Application of hyperspectral imaging in the assessment of drought and salt stress in magneto-primed triticale seeds[J]. Plants, 2021, 10(5): 835.
- [55] 孙俊, 武小红, 张晓东, 等. 基于高光谱图像的生菜叶片水分估算研究[J]. 光谱学与光

- 谱分析, 2013 (2): 522-526.
- [56] 孙俊, 金夏明, 毛罕平, 等. 基于 Adaboost 及高光谱的生菜叶片氮素水平鉴别研究[J]. 光谱学与光谱分析, 2013, 33(12): 3372-3376.
- [57] Lara M Á, Diezma B, Lleó L, et al. Hyperspectral imaging to evaluate the effect of irrigation water salinity in lettuce[J]. Applied Sciences, 2016, 6(12): 412.
- [58] Osco L P, Ramos A P M, Moriya É A S, et al. Modeling hyperspectral response of water-stress induced lettuce plants using artificial neural networks[J]. Remote Sensing, 2019, 11(23): 2797.
- [59] Mo C, Kim G, Kim M S, et al. On-line fresh-cut lettuce quality measurement system using hyperspectral imaging[J]. Biosystems engineering, 2017, 156: 38-50.
- [60] Zhou X, Sun J, Mao H, et al. Visualization research of moisture content in leaf lettuce leaves based on WT-PLSR and hyperspectral imaging technology[J]. Journal of Food Process Engineering, 2018, 41(2): e12647.
- [61] Saputro A H, Handayani W. Wavelength selection in hyperspectral imaging for estimation banana fruit quality[C]. 2017 International Conference on Electrical Engineering and Informatics (ICELTICs). IEEE, 2017: 226-230.
- [62] Yu S, Fan J, Lu X, et al. Hyperspectral technique combined with deep learning algorithm for estimation of phenotyping traits in lettuce[J]. Frontiers in plant science, 2022, 13: 927832.
- [63] Liu D, Sun D W, Zeng X A. Recent advances in wavelength selection techniques for hyperspectral image processing in the food industry[J]. Food and Bioprocess Technology, 2014, 7: 307-323.
- [64] Singh A K, Ganapathysubramanian B, Sarkar S, et al. Deep learning for plant stress phenotyping: trends and future perspectives[J]. Trends in plant science, 2018, 23(10): 883-898.
- [65] Furbank R T, Silva-Perez V, Evans J R, et al. Wheat physiology predictor: estimating physiological traits in wheat from hyperspectral reflectance measurements using deep learning[J]. Plant Methods, 2021, 17: 1-15.
- [66] Xiao Q, Tang W, Zhang C, et al. Spectral preprocessing combined with deep transfer learning to evaluate chlorophyll content in cotton leaves[J]. Plant Phenomics, 2022.

- [67] Zhou X, Sun J, Tian Y, et al. A deep learning based regression method on hyperspectral data for rapid estimation of cadmium residue in lettuce leaves[J]. *Chemometrics and Intelligent Laboratory Systems*, 2020, 200: 103996.
- [68] Zhou X, Sun J, Tian Y, et al. Hyperspectral technique combined with deep learning algorithm for detection of compound heavy metals in lettuce[J]. *Food chemistry*, 2020, 321: 126503.
- [69] Zhou X, Sun J, Tian Y, et al. Hyperspectral technique combined with deep learning algorithm for detection of compound heavy metals in lettuce[J]. *Food chemistry*, 2020, 321: 126503.
- [70] Sun J, Cao Y, Zhou X, et al. Detection for lead pollution level of lettuce leaves based on deep belief network combined with hyperspectral image technology[J]. *Journal of Food Safety*, 2021, 41(1): e12866.
- [71] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015: 1-9.
- [72] Rehman T U, Ma D, Wang L, et al. Predictive spectral analysis using an end-to-end deep model from hyperspectral images for high-throughput plant phenotyping[J]. *Computers and Electronics in Agriculture*, 2020, 177: 105713.
- [73] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018: 7132-7141.
- [74] Weng S, Han K, Chu Z, et al. Reflectance images of effective wavelengths from hyperspectral imaging for identification of Fusarium head blight-infected wheat kernels combined with a residual attention convolution neural network[J]. *Computers and Electronics in Agriculture*, 2021, 190: 106483.
- [75] Jung H, Oh Y. Towards better explanations of class activation mapping[C]. *Proceedings of the IEEE/CVF international conference on computer vision*. 2021: 1336-1344.
- [76] Poppi S, Cornia M, Baraldi L, et al. Revisiting the evaluation of class activation mapping for explainability: A novel metric and experimental analysis[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 2299-2304.
- [77] Yang W, Huang H, Zhang Z, et al. Towards rich feature discovery with class activation maps augmentation for person re-identification[C]. *Proceedings of the IEEE/CVF*

- conference on computer vision and pattern recognition. 2019: 1389-1398.
- [78] Fu K, Dai W, Zhang Y, et al. Multicam: Multiple class activation mapping for aircraft recognition in remote sensing images[J]. Remote sensing, 2019, 11(5): 544.
- [79] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.
- [80] Nagano Y, Kikuta Y. SRGAN for super-resolving low-resolution food images[C]. Proceedings of the Joint Workshop on Multimedia for Cooking and Eating Activities and Multimedia Assisted Dietary Management. 2018: 33-37.
- [81] Wang X, Yu K, Wu S, et al. Esrgan: Enhanced super-resolution generative adversarial networks[C]. Proceedings of the European conference on computer vision (ECCV) workshops. 2018: 0-0.
- [82] Pathak H N, Li X, Minaee S, et al. Efficient super resolution for large-scale images using attentional GAN[C]. 2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018: 1777-1786.
- [83] Liu B, Chen J. A super resolution algorithm based on attention mechanism and srgan network[J]. IEEE Access, 2021, 9: 139138-139145.
- [84] Thévenaz P, Blu T, Unser M. Image interpolation and resampling[J]. Handbook of medical imaging, processing and analysis, 2000, 1(1): 393-420.
- [85] Siu W C, Hung K W. Review of image interpolation and super-resolution[C]. Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference. IEEE, 2012: 1-10.
- [86] Patel V, Mistree K. A review on different image interpolation techniques for image enhancement[J]. International Journal of Emerging Technology and Advanced Engineering, 2013, 3(12): 129-133.
- [87] Han D. Comparison of commonly used image interpolation methods[C]. Conference of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013). Atlantis Press, 2013: 1556-1559.
- [88] Fadnavis S. Image interpolation techniques in digital image processing: an overview[J]. International Journal of Engineering Research and Applications, 2014, 4(10): 70-73.

- [89] Parsania P S, Virparia P V. A review: Image interpolation techniques for image scaling[J]. International Journal of Innovative Research in Computer and Communication Engineering, 2014, 2(12): 7409-7414.
- [90] Xiong Y, Guo S, Chen J, et al. Improved SRGAN for remote sensing image super-resolution across locations and sensors[J]. Remote Sensing, 2020, 12(8): 1263.
- [91] Dou X, Li C, Shi Q, et al. Super-resolution for hyperspectral remote sensing images based on the 3D attention-SRGAN network[J]. Remote Sensing, 2020, 12(7): 1204.
- [92] Zhang J, Wang X, Liu J, et al. Multispectral drone imagery and SRGAN for rapid phenotypic mapping of individual Chinese cabbage plants[J]. Plant Phenomics, 2022, 2022: 0007.
- [93] Serouart M, Madec S, David E, et al. SegVeg: Segmenting RGB images into green and senescent vegetation by combining deep and shallow methods[J]. Plant Phenomics, 2022.
- [94] Zheng C, Abd-Elrahman A, Whitaker V M, et al. Deep learning for strawberry canopy delineation and biomass estimation from high-resolution images[J]. Plant phenomics, 2022.
- [95] Okyere F G, Cudjoe D, Sadeghi-Tehran P, et al. Machine learning methods for automatic segmentation of images of field-and glasshouse-based plants for high-throughput phenotyping[J]. Plants, 2023, 12(10): 2035.
- [96] Zhu J, Hu C, Khezri E, et al. Edge intelligence-assisted animation design with large models: a survey[J]. Journal of Cloud Computing, 2024, 13(1): 48.
- [97] Lei Y, Li J, Li Z, et al. Prompt learning in computer vision: a survey[J]. Frontiers of Information Technology & Electronic Engineering, 2024, 25(1): 42-63.
- [98] Kolesnikov A, Beyer L, Zhai X, et al. Big transfer (bit): General visual representation learning[C]. Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16. Springer International Publishing, 2020: 491-507.
- [99] Zargar S. Introduction to sequence learning models: RNN, LSTM, GRU[J]. Department of Mechanical and Aerospace Engineering, North Carolina State University, 2021.
- [100] Chen P, Xiao Q, Zhang J, et al. Occurrence estimation of cotton pests and diseases by bidirectional long short-term memory networks with climate and atmosphere

- circulation[J]. Computers and Electronics in Agriculture, 2020, 176: 105612.
- [101] Wang B, Deng J, Jiang H, et al. Electronic nose signals-based deep learning models to realize high-precision monitoring of simultaneous saccharification and fermentation of cassava[J]. Microchemical Journal, 2022, 182: 107929.
- [102] Gong L, Yu M, Jiang S, et al. Deep learning based estimation on greenhouse crop yield combined TCN and RNN[J]. Sensors, 2021, 21(13): 4537.
- [103] Magalhaes J G S, Polinko A P, Amoroso M M, et al. The estimating tree growth app: an algorithmic approach to modelling individual tree growth[J]. Ecological Modelling, 2022, 467: 109932.
- [104] Khaki S, Wang L, Archontoulis S V. A CNN-RNN framework for crop yield estimation[J]. Frontiers in Plant Science, 2020, 10: 492736.
- [105] Wu T, Feng F, Lin Q, et al. Advanced method to capture the time-lag effects between annual NDVI and precipitation variation using RNN in the arid and semi-arid grasslands[J]. Water, 2019, 11(9): 1789.
- [106] Ge Y, Bai G, Stoerger V, et al. Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging[J]. Computers and Electronics in Agriculture, 2016, 127: 625-632.
- [107] Li Y, Wen W, Fan J, et al. Multi-source data fusion improves time-series phenotype accuracy in maize under a field high-throughput Phenotyping platform[J]. Plant Phenomics, 2023, 5: 0043.
- [108] Mishra P, Nordon A, Asaari M S M, et al. Fusing spectral and textural information in near-infrared hyperspectral imaging to improve green tea classification modelling[J]. Journal of Food Engineering, 2019, 249: 40-47.
- [109] Li Q, Jin S, Zang J, et al. Deciphering the contributions of spectral and structural data to wheat yield estimation from proximal sensing[J]. The Crop Journal, 2022, 10(5): 1334-1345.
- [110] Shu M, Fei S, Zhang B, et al. Application of UAV multisensor data and ensemble approach for high-throughput estimation of maize phenotyping traits[J]. Plant Phenomics, 2022.
- [111] Chang A, Jung J, Yeom J, et al. Unmanned aircraft system-(UAS-) based high-

- throughput phenotyping (HTP) for tomato yield estimation[J]. *Journal of Sensors*, 2021, 2021: 1-14.
- [112] Zhang C, Zhou L, Xiao Q, et al. End-to-end fusion of hyperspectral and chlorophyll fluorescence imaging to identify rice stresses[J]. *Plant Phenomics*, 2022.
- [113] Li B, Xu X, Zhang L, et al. Above-ground biomass estimation and yield estimation in potato by using UAV-based RGB and hyperspectral imaging[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 162: 161-172.
- [114] Gao H, Mao H, Zhang X. Inspection of lettuce water stress based on multi-sensor information fusion technology[C]. *Computer and Computing Technologies in Agriculture IV: 4th IFIP TC 12 Conference, CCTA 2010, Nanchang, China, October 22-25, 2010, Selected Papers, Part II 4*. Springer Berlin Heidelberg, 2011: 53-60.
- [115] Simko I, Jimenez-Berni J A, Furbank R T. Detection of decay in fresh-cut lettuce using hyperspectral imaging and chlorophyll fluorescence imaging[J]. *Postharvest Biology and Technology*, 2015, 106: 44-52.
- [116] Zhou X, Sun J, Tian Y, et al. Spectral classification of lettuce cadmium stress based on information fusion and VISSA-GOA-SVM algorithm[J]. *Journal of food process engineering*, 2019, 42(5): e13085.
- [117] Kumar P, Eriksen R L, Simko I, et al. Molecular mapping of water-stress responsive genomic loci in lettuce (*Lactuca* spp.) using kinetics chlorophyll fluorescence, hyperspectral imaging and machine learning[J]. *Frontiers in genetics*, 2021, 12: 634554.
- [118] Truong A, Walters A, Goodsitt J, et al. Towards automated machine learning: Evaluation and comparison of AutoML approaches and tools[C]. *2019 IEEE 31st international conference on tools with artificial intelligence (ICTAI)*. IEEE, 2019: 1471-1479.
- [119] Karmaker S K, Hassan M M, Smith M J, et al. AutoML to date and beyond: Challenges and opportunities[J]. *ACM Computing Surveys (CSUR)*, 2021, 54(8): 1-36.
- [120] LeDell E, Poirier S. H2o automl: Scalable automatic machine learning[C]. *Proceedings of the AutoML Workshop at ICML*. San Diego, CA, USA: ICML, 2020.
- [121] Wever M, Tornede A, Mohr F, et al. AutoML for multi-label classification: Overview and empirical evaluation[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2021, 43(9): 3037-3054.

- [122] Liang J, Meyerson E, Hodjat B, et al. Evolutionary neural automl for deep learning[C]. Proceedings of the genetic and evolutionary computation conference. 2019: 401-409.
- [123] He X, Zhao K, Chu X. AutoML: A survey of the state-of-the-art[J]. Knowledge-based systems, 2021, 212: 106622.
- [124] Gao D, Li M, Zhang J, et al. Improvement of chlorophyll content estimation on maize leaf by vein removal in hyperspectral image[J]. Computers and electronics in agriculture, 2021, 184: 106077.
- [125] Kumar A, Lee W S, Ehsani R J, et al. Citrus greening disease detection using aerial hyperspectral and multispectral imaging techniques[J]. Journal of Applied Remote Sensing, 2012, 6(1): 063542-063542.
- [126] 岑奕, 张良中, 张立福, 等. 地物光谱不确定性及光谱角制图算法适用性研究[J]. 光谱学与光谱分析, 2015, 35(10): 2841-2845.
- [127] Kim Y S. Detection of ecosystem distribution plants using drone Hyperspectral Spectrum and spectral angle mapper[J]. Journal of Environmental Science International, 2021, 30(2): 173-184.
- [128] Maheshwari M, Motwani M, Silakari S. Content Based Color Image Clustering[J]. International Journal of Computer Applications, 2012, 57(15): 38-43.
- [129] Abdulla B A, Ali Y H, Ibrahim N J. Extract the Similar Images Using the Grey Level Co-Occurrence Matrix and the Hu Invariants Moments[J]. Engineering and Technology Journal, 2020, 38(5A): 719-727.
- [130] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The journal of machine learning research, 2014, 15(1): 1929-1958.
- [131] Fan C, Yuan X, Zhang Y J A. Deep-learning-based signal detection for banded linear systems[C]. 2018 IEEE Global Communications Conference (GLOBECOM). IEEE, 2018: 1-6.
- [132] Fu L, Majeed Y, Zhang X, et al. Faster R-CNN-based apple detection in dense-foliage fruiting-wall trees using RGB and depth features for robotic harvesting[J]. Biosystems Engineering, 2020, 197: 245-256.
- [133] Saptoro A, Tadé M O, Vuthaluru H. A modified Kennard-Stone algorithm for optimal

- division of data for developing artificial neural network models[J]. Chemical Product and Process Modeling, 2012, 7(1).
- [134] Zhang L, Rao Z, Ji H. NIR hyperspectral imaging technology combined with multivariate methods to study the residues of different concentrations of omethoate on wheat grain surface[J]. Sensors, 2019, 19(14): 3147.
- [135] Sun J, Wu M, Hang Y, et al. Estimating cadmium content in lettuce leaves based on deep brief network and hyperspectral imaging technology[J]. Journal of food process engineering, 2019, 42(8): e13293.
- [136] Tian X, Li J, Wang Q, et al. A bi-layer model for nondestructive estimation of soluble solids content in apple based on reflectance spectra and peel pigments[J]. Food chemistry, 2018, 239: 1055-1063.
- [137] Jung D H, Kim H S, Jhin C, et al. Time-serial analysis of deep neural network models for estimation of climatic conditions inside a greenhouse[J]. Computers and Electronics in Agriculture, 2020, 173: 105402.
- [138] Moghimi A, Yang C, Anderson J A. Aerial hyperspectral imagery and deep neural networks for high-throughput yield phenotyping in wheat[J]. Computers and Electronics in Agriculture, 2020, 172: 105299.
- [139] Montaha S, Azam S, Rafid A K M R H, et al. TimeDistributed-CNN-LSTM: A hybrid approach combining CNN and LSTM to classify brain tumor on 3D MRI scans performing ablation study[J]. IEEE Access, 2022, 10: 60039-60059.
- [140] Sechidis K, Tsoumakas G, Vlahavas I. On the stratification of multi-label data[C]. Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2011, Athens, Greece, September 5-9, 2011, Proceedings, Part III 22. Springer Berlin Heidelberg, 2011: 145-158.
- [141] Lane H M, Murray S C. High throughput can produce better decisions than high accuracy when phenotyping plant populations[J]. Crop science, 2021, 61(5): 3301-3313.
- [142] Chollet F. Deep learning with Python[M]. Simon and Schuster, 2021.
- [143] Yan L, Xiong C, Qu H, et al. Non-destructive determination and visualisation of insoluble and soluble dietary fibre contents in fresh-cut celeries during storage periods using hyperspectral imaging technique[J]. Food Chemistry, 2017, 228: 249-256.

- [144] Zhao Y, Zhang C, Zhu S, et al. Shape induced reflectance correction for non-destructive determination and visualization of soluble solids content in winter jujubes using hyperspectral imaging in two different spectral ranges[J]. *Postharvest Biology and Technology*, 2020, 161: 111080.
- [145] Chu X, Wang W, Ni X, et al. Classifying maize kernels naturally infected by fungi using near-infrared hyperspectral imaging[J]. *Infrared Physics & Technology*, 2020, 105: 103242.
- [146] Jiang H, Jiang X, Ru Y, et al. Application of hyperspectral imaging for detecting and visualizing leaf lard adulteration in minced pork[J]. *Infrared Physics & Technology*, 2020, 110: 103467.
- [147] Jiang H, Jiang X, Ru Y, et al. Application of hyperspectral imaging for detecting and visualizing leaf lard adulteration in minced pork[J]. *Infrared Physics & Technology*, 2020, 110: 103467.
- [148] Zhang W, Cao A, Shi P, et al. Rapid evaluation of freshness of largemouth bass under different thawing methods using hyperspectral imaging[J]. *Food Control*, 2021, 125: 108023.
- [149] Yao K, Sun J, Chen C, et al. Visualization research of egg freshness based on hyperspectral imaging and binary competitive adaptive reweighted sampling[J]. *Infrared Physics & Technology*, 2022, 127: 104414.
- [150] Xuan G, Gao C, Shao Y. Spectral and image analysis of hyperspectral data for internal and external quality assessment of peach fruit[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2022, 272: 121016.
- [151] Chu X, Miao P, Zhang K, et al. Green Banana maturity classification and quality evaluation using hyperspectral imaging[J]. *Agriculture*, 2022, 12(4): 530.
- [152] Uto K, Seki H, Saito G, et al. Characterization of rice paddies by a UAV-mounted miniature hyperspectral sensor system[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, 6(2): 851-860.
- [153] Ravindaran R, Kasthuri N, Preethi S, et al. Performance Analysis of a VGG based Deep Learning Model for Classification of Weeds and Crops[C]. 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT).

IEEE, 2023: 1-5.

- [154] Howard A G, Zhu M, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv: 1704.04861, 2017.
- [155] Tan M, Le Q. EfficientNet: Rethinking model scaling for convolutional neural networks[C]. International conference on machine learning. PMLR, 2019: 6105-6114.
- [156] Khotanzad A, Hong Y H. Invariant image recognition by Zernike moments[J]. IEEE Transactions on pattern analysis and machine intelligence, 1990, 12(5): 489-497.
- [157] Ghassemian H. A review of remote sensing image fusion methods[J]. Information Fusion, 2016, 32: 75-89.
- [158] Guyon I, Chaabane I, Escalante H J, et al. A brief review of the ChaLearn AutoML challenge: any-time any-dataset learning without human intervention[C]. Workshop on Automatic Machine Learning. PMLR, 2016: 21-30.

附录 A

SSC 和 pH 值测量仪器实物图如图 A1 所示。每个生菜样本的汁液都单独放置在试管中，在测量 SSC 和 pH 值之前，用均质器对生菜汁液进行均质处理，确保生菜汁液中的成分能够均匀地混合，再用移液器取出生菜汁，并将其置于相应的仪器中进行测量。每个样本的 SSC 和 pH 值都会测量三次，最后取平均值作为该样本 SSC 和 pH 值的参考值。



图 A1 数字折光仪和 pH 值测量仪的实物图

Figure A1 Actual picture of digital refractometer and pH value meter

各品种在实验过程中因标签丢失而丢弃的具体样本数如表 A1 所示。最终用于建模的奶油、散叶和罗马品种样本数量分别为 125、127 和 135 个。

表 A1 实验过程中丢失的样品

Table A1 Samples lost during the experiment

Cultivars	May 15, 2021	May 20, 2021	May 25, 2021
Butter	-	4	6
Leaf	6	2	-
Roman	-	-	-

SpecNet2D 的超参数配置见表 A2。

表 A2 SpecNet2D 的配置

Table A2 Configuration of SpecNet2D

Branch1	Branch2	Branch3	Dense module
Convolution layer 64@1×3	Convolution layer 64@1×3	Convolution layer 64@1×1	Flatten 20480
MaxPooling layer 2	MaxPooling layer 2	MaxPooling layer 2	Dense layer 512
Convolution layer 128@1×3	Convolution layer 64@1×3	Convolution layer 64@1×1	Dense layer 256
MaxPooling layer 2	MaxPooling layer 2	MaxPooling layer 2	Dense layer 128
Convolution layer 256@1×3	Convolution layer 64@1×3	Convolution layer 64@1×1	Dropout 0.5
MaxPooling layer 2	MaxPooling layer 2	MaxPooling layer 2	
Convolution layer 512@1×3			

优化器为 Adam，loss 为 RMSE，batch size = 4，epochs = 2000，learning rate = 0.0001。其中，批量大小（batch size）是每次训练发送到神经网络的训练数据的数量。训练周期（epoch）是所有校准集的数据，全部发送到神经网络中的一次完整训练。学习率（learning rate）控制着模型在学习过程中权重更新的幅度

SpecNet1D 的超参数配置见表 A3。

表 A3 SpecNet1D 的配置

Table A3 Configuration of SpecNet1D

Model	Hyperparameters
SpecNet1D	Dense layer 512
	Dense layer 256
	Dense layer 128
	Dropout 0.1
	Dense layer 64
	Dropout 0.1

优化器为 Adam，loss 为 RMSE，batch size = 4，epochs = 1000，learning rate = 0.0001。

表 3.3-3.5 中使用的多元分析方法、预处理方法和 CARS 的参数设置如表 A4-6 所示。

表 A4 表 3.3 中各个模型参数设置

Table A4 Parameter setting of various models in Table 3.3

Models	Settings of parameters
PLSR	nLVs = 10
	nLVs = 10
	nLVs = 12
	nLVs = 11
	nLVs = 11
	nLVs = 11
LWR	a = 86, n_neigh = 258, distance = 1, weighting = 1
	a = 75, n_neigh = 232, distance = 1, weighting = 1
	a = 73, n_neigh = 229, distance = 1, weighting = 1
	a = 74, n_neigh = 258, distance = 1, weighting = 1
	a = 75, n_neigh = 258, distance = 1, weighting = 1
	a = 79, n_neigh = 258, distance = 1, weighting = 1
MLR	intercept = 1
	intercept = 1
	intercept = 1
	intercept = 1
	intercept = 1
	intercept = 1
ANN	topo = ['HH', 'L-'], epochs = 40
	topo = ['HH', 'L-'], epochs = 40
	topo = ['HH', 'L-'], epochs = 40
	topo = ['HH', 'H-'], epochs = 40
	topo = ['HH', 'H-'], epochs = 25
	topo = ['HH', 'H-'], epochs = 45
SVR	linear kernel function, C = 0.35, e = 0.2, tol = 0.001
	linear kernel function, C = 0.5, e = 0.2, tol = 0.0001
	linear kernel function, C = 0.5, e = 0.5, tol = 0.0001
	linear kernel function, C = 0.5, e = 0.5, tol = 0.001
	linear kernel function, C = 0.4, e = 0.5, tol = 0.001
	linear kernel function, C = 0.4, e = 0.5, tol = 0.001

预处理方法参数设置：MWS：窗口宽度 = 5；SG：计算点的数量 = 15，多项式阶数 = 2，导数阶数 = 0；FDR：计算点的数量 = 15，多项式阶数 = 2，导数阶数 = 1；SDR：计算点的数量 = 15，多项式阶数 = 2，导数阶数 = 2；WT：小波类型为 harr，变换的层数 = 1。

建模参数设置：PLSR：nLVs — 潜在变量的数目；LWR：a — 带宽参数；n_neigh — 最近邻的数量；distance — 1 = 欧氏距离，2 = 马氏距离；weighting — 1 = 一次方，2 = 三次方；MLR：intercept — 0 = 不计算截距，1 = 计算截距；ANN：topo — 激活函数，H = 双曲正切函数，L = 线性函数；SVR：C — 正则化参数；e — 惩罚参数的阈值；tol — 容忍度。

表 A5 表 3.4 中各个模型参数设置

Table A5 Parameter setting of various models in Table 3.4

Models	Settings of parameters
PLSR	nLVs = 11
	nLVs = 13
	nLVs = 13
	nLVs = 13
	nLVs = 12
	nLVs = 12
LWR	a = 59, n_neigh = 208, distance = 1, weighting = 1
	a = 63, n_neigh = 196, distance = 1, weighting = 1
	a = 63, n_neigh = 191, distance = 1, weighting = 1
	a = 61, n_neigh = 160, distance = 1, weighting = 1
	a = 68, n_neigh = 206, distance = 1, weighting = 1
	a = 80, n_neigh = 231, distance = 1, weighting = 1
MLR	intercept = 1
	intercept = 1
	intercept = 1
	intercept = 1
	intercept = 1
	intercept = 1
ANN	topo = ['HH', 'H-'], epochs = 40
	topo = ['HH', 'H-'], epochs = 35
	topo = ['HH', 'H-'], epochs = 40
	topo = ['HH', 'L-'], epochs = 20
	topo = ['HH', 'L-'], epochs = 15
	topo = ['HH', 'L-'], epochs = 45
SVR	linear kernel function, C = 0.35, e = 0.2, tol = 0.001
	linear kernel function, C = 0.5, e = 0.2, tol = 0.001
	linear kernel function, C = 0.5, e = 0.2, tol = 0.001
	linear kernel function, C = 0.5, e = 0.2, tol = 0.01
	linear kernel function, C = 0.4, e = 0.1, tol = 0.001
	linear kernel function, C = 0.4, e = 0.2, tol = 0.001

预处理方法设置与表 A4 相同。

表 A6 表 3.5 中各个模型参数设置

Table A6 Parameter setting of various models in Table 3.5

Phenotypes	Models	Settings of parameters
SSC (%)	PLSR	nLVs = 3
		nLVs = 3
		nLVs = 3
	LWR	a = 15, n_neigh = 231, distance = 1, weighting = 1
		a = 15, n_neigh = 231, distance = 1, weighting = 1
		a = 20, n_neigh = 231, distance = 1, weighting = 1
pH value	PLSR	nLVs = 4
		nLVs = 4
		nLVs = 4
	LWR	a = 8, n_neigh = 206, distance = 1, weighting = 1
		a = 8, n_neigh = 206, distance = 1, weighting = 1
		a = 5, n_neigh = 206, distance = 1, weighting = 1

CARS 参数设置: A = 9, fold = 5, method = 'center', num = 50。其中, A 为潜在变量的数量; fold 为交叉验证折数; method 为预处理方法; num 为蒙特卡洛采样次数。

CARS 选取的具体光谱特征如表 A7 所示，光谱特征顺序按权重降序排列。

表 A7 表 3.5 中使用 CARS 选择的具体光谱特征

Table A7 Concrete spectral feature selected using CARS in Table 5

Phenotypes	Models	Pretreatment	Wavelengths (nm)
SSC (%)	PLSR	SG	419;415;1000;439;942;435;448;972;456;674;954
			;628;995;992;828;944;472;637;656;630;937;487;
		FDR	619;639;825;646;677;463;667;559;665;621;568
			660;653;663;649;602;605;639;635;670;596;786;
			949;612;959;884;714;498;698;712;962;741;400;
	LWR	FDR	402;404;406;408;415
			972;977;942;801;982;487;765;472;852;948;874;
		SDR	899;974;847;889;503;937;985;760;494;796;505;
			448;483;755;904;912;769;786;463
			419;415;1000;439;942;435;448;972;456;674;954
pH value	PLSR	SG	;628;995;992;828;944;472;637;656;630;937;487;
			619;639;825
		FDR	660;653;663;649;602;605;639;635;670;596;786;
			949;612;959;
			884;714;498;698;712;962;741;400;402;404
	MWS	SDR	972;977;942;801;982;487;765;472;852;948;874;
			899;974;847;889;503;937;985;760;494;796;505;
		MWS	448;483;755;904;912;769;786;463
			600;602;612;621;632;609;463;591;487;490;593;
			614;501;532;619;639;476;811;498;808;510;598;
pH value	PLSR	SG	521;651;912;644;474;860;512;584;503;748;485;
			523;589;658;818;472;830;635;729;653
		WT	589;612;609;621;487;632;872;602;887;485;490;
	600;815;448;818;806;474;498;808;501;570;828;		
	WT	635;796;472;523;867;452;642;430;535;461;987;	
		582;912;745;492;644;463;789;512;496;510	
pH value	PLSR	SG	600;681;677;548;525;972;644;705;691;487;490;
			501;503;729;609;612;589;719;543;498;703;474;
		WT	476;724;580;818;820;755;630;568;656;658;698;
			738;483;485

表 A7（续）

LWR	MWS	600;602;612;621;632;609;463;591;487;490;593; 614;501;532;619;639;476;811;498;808;510;598; 521;651;912;644;474;860;512;584;503;748;485; 523;589;658;818;472;830;635;729;653;736;496; 441;705;432;492;580;957
		589;612;609;621;487;632;872;602;887;485;490; 600;815;448;818;806;474;498;808;501;570;828; 635;796;472;523;867;452;642;430;535;461;987; 582;912;745;492;644;463;789;512;496;510;459; 619;521;470;605
		600;681;677;548;525;972;644;705;691;487;490; 501;503;729;609;612;589;719;543;498;703;474; 476;724;580;818;820;755;630;568;656;658;698; 738;483;485;463;922;924;862;864
	SG	
	WT	

附录 B

图 B1 (A) 展示了卷积操作的基本原理。在这个过程中，将一个卷积核（在图中表示为“Filter W_n ”）用作滑动窗口，对每个像素矩阵进行扫描。滑动窗口的步长设定了窗口滑动的间隔。每到一处，窗口内的像素值会与卷积核的对应值进行逐个相乘，接着求和得到新的像素值。这样的操作在整个像素矩阵上重复进行，最终形成了卷积特征图。卷积核的作用类似于一个过滤器，它可以从输入的对象中提取出特定的特征。通过采用多种不同的卷积核，深度学习模型能够提取丰富的特征信息，进而有效学习输入对象的独特属性。CNN 在灵活性和有效性方面表现卓越，成为处理多种数据类型（包括光谱数据、图像数据等）的有效方法。

如图 B1 (B) 所示，Inception 模块可以融合多种尺寸的卷积核，目的是扩展网络的感知范围并显著提升神经网络的健壮性。该模块在同一层中堆砌了尺寸各异的核，使得网络能在一个平台上并行产生多样化的核尺寸。这一设计省去了手动挑选卷积操作的复杂过程，使得网络能够自行探索并识别最高效的卷积策略。这种策略极大地增强了网络的灵活性和调整能力，从而优化了它们捕捉和描绘输入数据中细腻特征的能力。

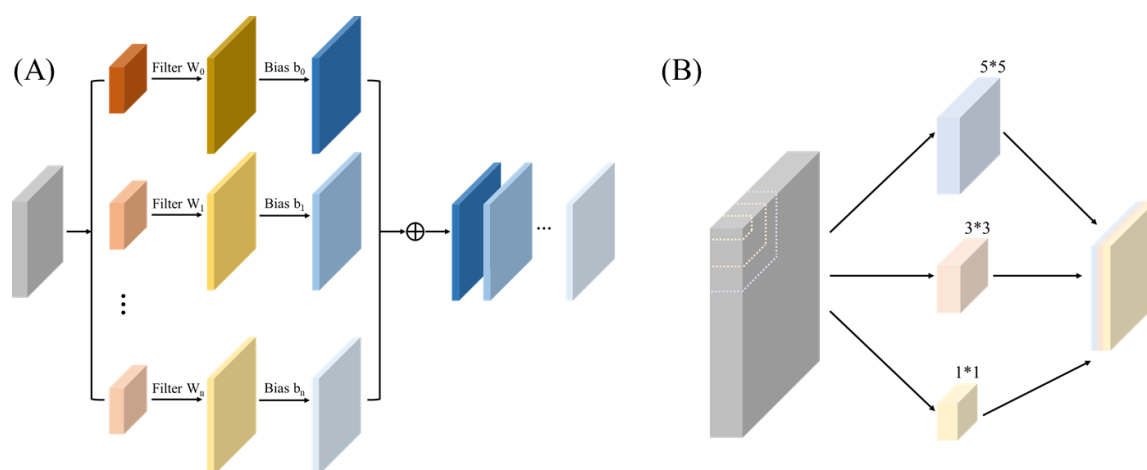


图 B1 基于 CNN 的 Inception 模块的数据流图。(A) CNN 数据流图和 (B) Inception 模块数据流图

Figure B1 The data flow of the CNN-based Inception module. (A) CNN data flow and (B) Inception module data flow

图 B2 详尽地阐述了在 RNN 架构中 TD、LSTM、GRU 和 BRNN 的基础理论及它们的具体运用。与利用卷积核进行特征提取的 CNN 相比，RNN 采用循环核（单元）

来实现特征提取任务。正如图 B2 (A) 所揭示的, 这些循环单元具备记忆功能, 能通过在不同时间步骤间共享参数, 来从时间序列数据中提取信息。每个循环单元包含若干记忆细胞, 图中以细小的圆柱体表示, 它们负责存储每个时间步的状态信息, 即 h_t 。权重矩阵 W_{xh} 、 W_{hh} 和 W_{hy} 构成了 RNN 的核心参数, 此外, x_t 表示当前时刻的输入特征, h_{t-1} 则代表上一时刻记忆中保留的状态。输出特征 y_t 是通过这些元素相互作用的一连串计算过程产生的。

正如图 B2 (B) 展示的那样, TD 在多个输入共用一个处理层并且为每个输入产生单独输出的过程中, 扮演了重要的角色。这种策略在模型处理时间序列数据时显得格外高效。而对于网络中的卷积层来说, 由于它们共享相同的权重集, 因此在功能上它们是等同的。这种权重共享机制确保了不同输入间能够进行一致且稳定的特征提取。更简洁地说, 可以将卷积层视作一系列克隆单元, 它们共同对输入数据的各个区域执行相同的操作集。

图 B2 (C) 展示的 LSTM 核心理念主要涉及单元状态和精巧的“门”结构设计。单元状态是信息在网络中流转的关键通道, 是网络的记忆核心。它在时间序列数据的处理过程中, 保留关键信息, 确保能够将早期的数据储存且有效地传播到后续的时间节点。这种设计巧妙地克服了短期记忆的限制。而“门”结构则由输入门、遗忘门和输出门构成, 它们共同协调信息的流入流出。这些门各自对输入的数据、保留的记忆和最终的输出进行精细调控。通过深度训练, 这些“门”结构能学习到何时该留存或抛弃信息, 从而增强网络在处理时间序列数据方面的精准度和效率。

与 LSTM 相较, GRU 如图 B2 (D) 所示, 呈现出更加精炼的结构特性。它由更新门和重置门两部分构成。重置门承担着与 LSTM 中输入门类似的职责, 而更新门则集 LSTM 中的遗忘门和输出门功能于一身。这种职能的合并使得 GRU 的体系结构更加精简, 同时也让其实现过程更为便捷高效。

BRNN (见图 B2 (E)) 的主要构架是由两个镜像对称的单向 RNN 所融合而成。这两个网络分享相同的架构设计, 它们唯一的不同在于各自的数据处理方向。每一个独立的网络单元, 记作 A_i , 都可以采用 LSTM 或者 GRU 等不同的循环结构实现。BRNN 的对称布局使其能够通过正向及反向分析输入序列, 全面捕获历史和未来的上下文信息。通过融合这两个网络的输出, BRNN 能有效地识别并整合输入数据中的前后依赖性, 显著提升其建模能力。这种特性使得 BRNN 特别适用于那些需要理解和捕捉输入序列中先前和随后元素间依赖关系的各类任务。

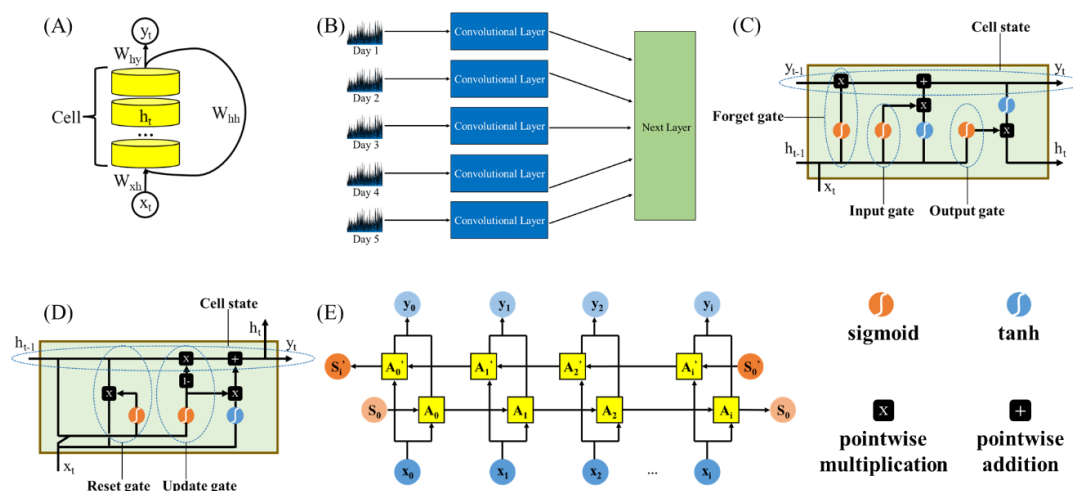


图 B2 RNN 框架的数据流图 (A) 以及 TD (B)、LSTM (C)、GRU (D) 和 BRNN (E) 在 RNN 内的具体用法

Figure B2 The data flow of RNN framework (A) and the specific usages of TD (B), LSTM (C), GRU (D), and BRNN (E) within the RNN

在温室 HTPP 平台中, 可在 1 h 内完成对 600 m² (相当于 2500 个花盆) 区域的植物冠层数据的详细采集工作, 数据处理效率为每分钟约 45 株植物。该平台能够保证每天连续不断地, 甚至在一天之内多次收集数据。该平台搭载了三种传感器, 包括 RGB 相机、多光谱相机和激光雷达来获取高质量的数据。这些数据随后输入到 HortControl 系统自动分析, 为每株植物精确量化了 14 种不同的表型参数, 包括数字生物量、绿度、株高、最大株高、色调、叶片角度、叶面积、叶面积指数、投影叶面积、叶片倾斜度、激光穿透深度、归一化植被指数 (NDVI)、归一化色素叶绿素指数 (NPCl) 和植物衰老反射指数 (PSRI)。其中, 前 11 种表型除了绿度和色调属于颜色表型, 其他都属于形态表型, 而后三种则归属于植被指数。这些表型参数的详细说明, 请参阅附表 B1。

表 B1 14 种表型的详细描述

Table B1 The detailed description of the 14 phenotypes

表型	描述
数字生物量 (cm^3)	数字生物量的计算基于一个假设，即植物拥有一个可以通过株高和宽度确定体积的规则形态，并将此体积与三维叶面积的乘积作为结果
绿度	绿度，亦即绿叶指数，指的是绿色光谱通道的反射率与其他两个可见光谱通道（红色与蓝色）的反射率所呈现的相对关系
株高 (mm)	株高的确定是通过分析沿 Z 轴排列的基本三角形的分布来实现的。首先步骤是计算出 Z 轴方向的直方图
最大株高 (mm)	最大株高参数可以精确测定植物最高点的绝对高度，其单位为毫米，保证了极高的准确度。这一参数并不是为了替代现有的株高测量参数，而是对株高参数的一个补充。相比之下，当前的株高测量更侧重于稳定性而不是精度，它通过减小风力、外部干扰或植物日夜运动等小幅动作的影响，来最小化误差
色调	将 RGB 颜色空间转换成 HSV 颜色空间后，提取色调通道的数值
叶片角度 ($^\circ$)	通过根据植物网格中每个面的法向量计算所有面角度的加权平均值，进而得出叶片的角度，该角度是各个面角度的平均值
叶面积 (cm^2)	源于三维点云数据，根据既定的准则对属于同一区域的点进行三角化处理。接着，通过累加这些基本三角形的面积，得出植物整体的叶面积总和
叶面积指数 (cm^2/cm^2)	叶面积指数通过对总叶面积与对应区域面积的比例进行计算，用以表征在给定面积内叶片层数的多寡
投影叶面积 (cm^2)	投影叶面积是指所有基本三角形在 X-Y 平面上的投影区域所覆盖的总面积
叶片倾斜度 (cm^2/cm^2)	叶片倾斜度反映了植物叶片平均立起的程度，通过将总叶面积除以所有基本三角形在 X-Y 平面上投影的总和来计算得出
激光穿透深度 (mm)	激光穿透深度是根据激光穿过植物冠层的深度来测定的
NDVI	计算公式 $(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$ ，与光合有效辐射拦截有关，NDVI 的值代表了地表植被的覆盖程度和健康状态，高的 NDVI 值通常意味着茂密的绿色植被，这样的植被能截获更多的光合有效辐射
NPCI	计算公式 $(\text{Red} - \text{Blue}) / (\text{Red} + \text{Blue})$ ，与叶绿素含量相关，在红光区域，叶绿素会吸收大量光能用于光合作用，因此该波段的反射率低。红边区域是叶绿素吸收开始下降的地方，反射率相对较高。NPCI 通过红光和红边区的反射率计算得出，它能够指示植物叶片中叶绿素的相对含量。NPCI 值随着叶绿素含量的增加而降低，因为叶绿素含量增加会导致红光区域的吸收加强。
PSRI	计算公式 $(\text{Red} - \text{Green}) / \text{NIR}$ ，与类胡萝卜素和叶绿素的比例相关，当植物衰老时，叶绿素含量会逐渐降低，叶子从绿色变为黄色或其他颜色，叶绿素含量减少的同时，类胡萝卜素含量可能会相对增加，以保护植物免受光氧化损害

由于样本标签的遗失，最终对 524 份样本进行了 SSC 和 pH 值的测定。各个生菜品种实际参与建模的样本数量详见表 B2。

表 B2 每个品种用于建模的数量

Table B2 The amount of each cultivar used for modeling

Phenotypes	Cultivars	Sample numbers
SSC (%) and pH value	Butter	152
	Head	190
	Leaf	90
	Roman	92

Inception 模块的参数设置如表 B3 所示。

表 B3 Inception 模块的配置

Table B3 Configuration of Inception module

Phenotypes	Branch1	Branch2	Branch3	Branch4	Dense	Training parameter
SSC (%)	Convolution layer 64@1×5	Convolution layer 64@1×3	MaxPooling layer 2	Convolution layer 128@1×5	GlobalAveragePooling 384	Epochs = 1500
		Convolution layer 128@1×3	Convolution layer 64@1×5 Convolution layer 64@1×5	Convolution layer 128@1×5 Convolution layer 128@1×5	Dense layer 64	Batch size = 6
pH value	Convolution layer 64@1×5	Convolution layer 64@1×3	MaxPooling layer 2	Convolution layer 128@1×5	GlobalAveragePooling 384	Epochs = 500
		Convolution layer 128@1×3	Convolution layer 64@1×5 Convolution layer 64@1×5	Convolution layer 128@1×5 Convolution layer 128@1×5	Dense layer 256	Batch size = 6

PLSR 和 SVR 的参数设置见表 B4。

表 B4 PLSR 和 SVR 配置

Table B4 Configuration of PLSR and SVR

Models	Phenotypes	Settings of parameters
PLSR	SSC (%)	nLVs = 13
	pH value	nLVs = 12
SVR	SSC (%)	linear kernel function, C = 0.5
	pH value	RBF kernel function, C = 0.42

残差和注意力模块参数设置如表 B5 所示。

表 B5 Optimal-Spec 模型的残差和注意力模块配置

Table B5 Configuration of residual and attention module of the Optimal-Spec models

Phenotypes	Branch3	Branch4	Attention	Training parameters
SSC (%)		Residual: Convolution layer 128@1×1	Dense (tanh) 192	Epochs = 1800
			Dense (tanh) 384	Batch size = 6
pH value	Residual: Convolution layer 64@1×1			Epochs = 800 Batch size = 6

RNN 的参数设置见表 B6。

表 B6 Optimal-Spec-RNN 模型配置

Table B6 Configuration of Optimal-Spec-RNN models

Phenotypes	RNNs	Settings of parameters	Training parameters
SSC (%)	TD	Convolution layer 16@1×2	Epochs = 2000
		Convolution layer 16@1×1	Batch size = 6
		Convolution layer 8@1×1	
pH value	TD	Convolution layer 64@1×2	Epochs = 950
		Convolution layer 64@1×1	Batch size = 6
		Convolution layer 128@1×1	

附录 C

叶绿素计的实物图如图 C1 所示。



图 C1 叶绿素计的实物图

Figure C1 Actual picture of chlorophyll meter

SRGAN 模型的训练细节如表 C1 所述。该模型的生成器由三大部分构成。首先，低分辨率图像通过一个卷积层，随后接入 ReLU 激活函数进行非线性变换；其次，图像数据传递至包含 16 个残差块的中间层，每个残差块内部都由两次卷积操作、归一化处理以及 ReLU 函数构成，并伴随着相应的跳跃连接（residual connection）；最终，数据流入上采样模块，在一至两次的上采样操作后，图像的宽度和高度分别放大至原尺寸的 2 倍或 4 倍，从而实现了图像分辨率的显著提升。至于 SRGAN 的判别器结构，则由 7 个卷积层组成，这些卷积层之间穿插着归一化处理和 LeakyReLU 激活函数，形成了一个深层网络来识别图像的真伪。在特征提取网络方面，采用了这些网络层中间层的权重进行训练，具体的层名称见表 C1。PSNR 和 SSIM 用于评价图像质量。PSNR 以分贝（dB）为单位，评价的是信号最大幅度与噪声幅度的比值，反映了图像重建的质量。具体来说，PSNR 值超过 40 分贝通常意味着图像质量极佳，即与原始图像高度相似；30 到 40 分贝范围内的 PSNR 值表示图像质量良好，尽管存在一定失真，但仍在可接受范围内；20 到 30 分贝的 PSNR 值则表明图像质量较差；而低于 20 分贝的 PSNR 值则通常表示图像质量不可接受。与 PSNR 不同，SSIM 是一个无单位的比率值，其取值范围从 -1 到 1。SSIM 测量的是两幅图像间的结构相似度，其中 1 代表两幅图像结构完全相同，而 -1 则意味着两者结构完全不同。通过这两个指标，可以精确地评估和比较图像处理、压缩或重建过程中的图像质量，为图像优化和分析提供了重要的量化依

据。

表 C1 SRGAN 的配置

Table C1 Configuration of SRGAN

Generator	Discriminator	Feature extraction network	Training parameters
		MobileNet ('conv_pw_6')	Epochs = 200
Number of residual networks is 16	Number of convolution, batchnormalization, and LeakyReLU is 7	EfficientNet ('block4c_project_conv')	Batch size = 4
		VGG16, VGG19 ('block3_conv3')	

同特征提取网络的 SRGAN 的 2 倍和 4 倍上采样训练结果的 PSNR 和 SSIM 如图 C2 所示。根据 PSNR 的训练曲线来看, 4 倍上采样的结果明显低于 2 倍上采样。另外, 4 倍上采样的所有结果都很相近。2 倍上采样中, VGG19 的图像重建质量低于另外三种, EfficientNet 在 120 到 170 个的训练周期中, 效果比 VGG16 和 MobileNet 差。从 PSNR 和 SSIM 的结果来看, 可以发现 4 种特征提取网络的 2 倍上采样结果在 175 个训练周期后, 效果有显著下降。在模型训练过程中, 会每隔 10 个训练周期保存一次 SRGAN 模型的权重, 所以最终选择训练了 170 个训练周期的模型用于最终的图像超分辨率提升。

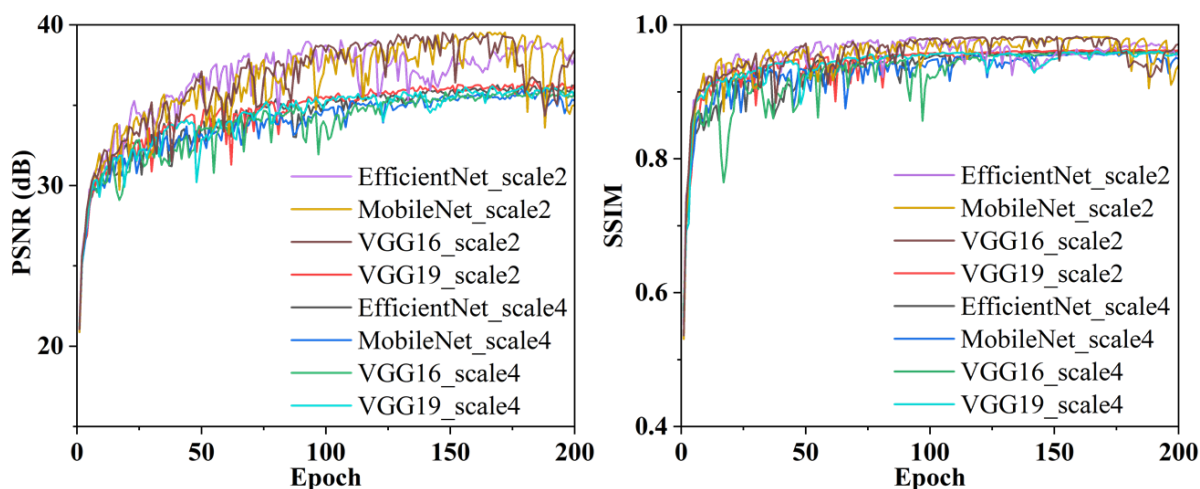


图 C2 不同特征提取网络的 SRGAN 的 2 倍和 4 倍上采样训练结果

Figure C2 Training results of 2x and 4x upsampling of SRGAN for different feature extraction networks

尽管图像插值方法并不涉及训练过程, 但是它们的效果仍然可以通过定量的指标进行评估, 而不只是依赖于主观的视觉观察。像 PSNR 和 SSIM (这样的性能指标, 不仅可以用于评估 SRGAN 这样的学习型方法, 也同样适用于评价图像插值方法在图像超分辨率处理中的表现。以 MobileNet 作为特征提取器的 SRGAN 在重建任务中表现最

为出色，因此，以此为基准，与四种插值方法重建后的图像进行了比较，计算了后者的 PSNR 和 SSIM 值。如表 C2 所示，4 倍上采样在所有指标上均优于 2 倍上采样，其中 Lanczos 插值方法获得的指标是最佳的，这与 SPAD 值的估算结果一致。除了 Lanczos 和线性插值方法在 4 倍上采样中获得的 PSNR 值超过 30 之外，其他方法的结果均低于 30，这表明插值法在图像重建上的效果相对较弱。

表 C2 插值法重建的图像质量评价

Table C2 Image quality assessment of the reconstruction by interpolation methods				
SR	Methods	Scale	PSNR (dB)	SSIM
Image interpolation	Cubic	2	26.6944	0.7598
		4	27.6172	0.7756
	Linear	2	29.0777	0.7893
		4	30.4736	0.7921
	Near	2	26.4844	0.7418
		4	28.6999	0.7810
	Lanczos	2	29.7861	0.7840
		4	31.6979	0.8173

基于光谱数据的深度学习模型的参数如表 C3 所示。

表 C3 光谱模型的配置

Table C3 Configuration of spectral model						
Branch1	Branch2	Branch3	Branch4	Attention	Dense	Training parameter
Convolution layer 64@1×1	Convolution layer 64@1×3	MaxPooling layer 5	Convolution layer 128@1×3	Dense (relu) 192	Dense layer 128	Epochs = 1000
	Convolution layer 128@1×3	Convolution layer 64@1×5	Convolution layer 128@1×3	Dense (relu) 384		Batch size = 6
		Convolution layer 64@1×5	Convolution layer 128@1×3			

针对基于图像数据的深度学习模型及其融合注意力模块的参数设置进行了详细说明，具体细节可参考表 C4。在模型建立过程中，特别关注了大型深度学习模型——BiT 模型的应用，该模型通过 `hub.KerasLayer` 接口加载。此外，还采用了 ResNet 和 MobileNet 这两种深度学习框架，使用了其在 ImageNet 数据集上预训练的权重。这种方法不仅利用了现有的强大深度学习架构，还通过引入预训练模型来提高了模型训练的效率 and 估算的准确性。

表 C4 图像模型的配置

Table C4 Configuration of image model

Methods	Fine-tuning	Feature Attention	Training parameters
BiT, ResNet, MobileNet, and Primary feature	Reshape	Concatenate	Epochs = 1100
	Convolution layer 512@1×5	Dense (relu)	Batch size = 6
	Convolution layer 256@1×5	Dense (softmax)	

附录 D

EC、K⁺和 Na⁺测量仪的实物图如图 D1 所示。

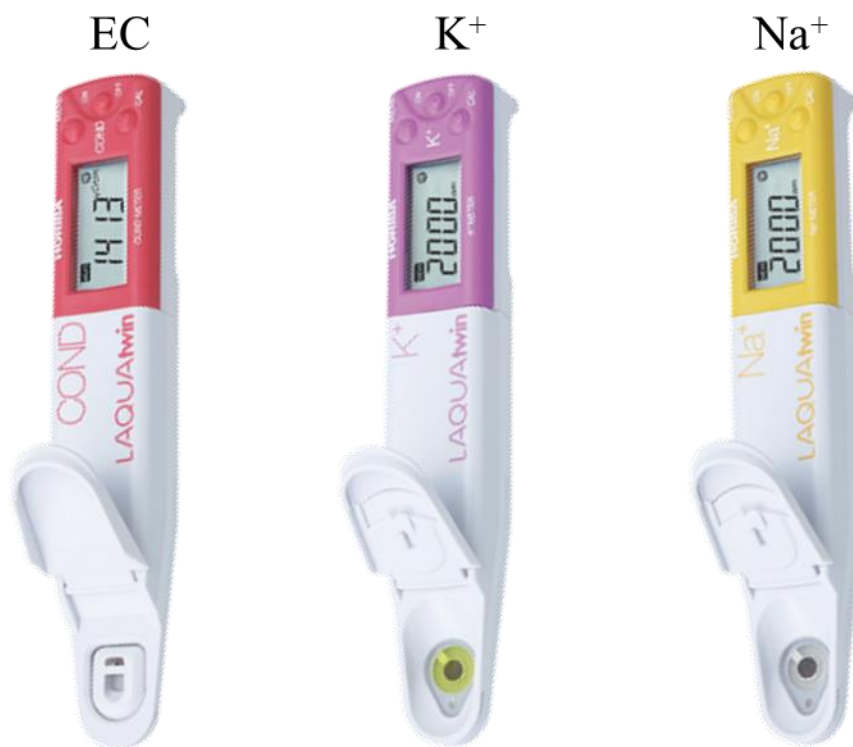


图 D1 EC 测量仪、K⁺测量仪和 Na⁺测量仪的实物图

Figure D1 Actual picture of EC meter, potassium meter, and sodium meter

SpecTimeImageNet 模型架构如图 D2 所示，模型架构综合了 SpecTimeNet 模型和 SpecImageNet 的优势。

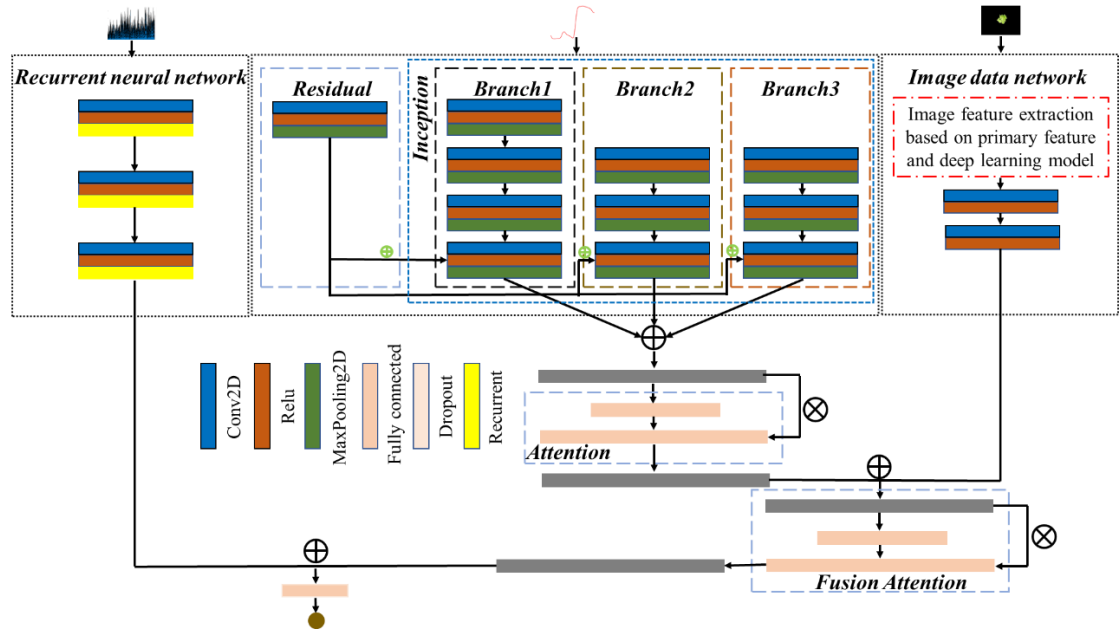


图 D2 定制化深度学习特征融合模型（SpecTimeImageNet）架构

Figure D2 Custom deep learning feature fusion model (SpecTimeImageNet) architecture

基于定制化深度学习模型的参数如表 D1-D3 所示，其中 EC，K⁺和 Na⁺的训练周期分别为 1500，1200 和 1200，batch size 都为 6。

表 D1 光谱模型配置

Table D1 Configuration of spectral model

Phenotypes	Branch1	Branch2	Branch3	Branch4	Residual & Attention	Dense
EC (mS/cm)	Convolution layer 64@1×7	Convolution layer 64@1×3	MaxPooling layer 3	Convolution layer 128@1×5	Branch4 & Residual Convolution layer 128@1×5	Dense layer 64
		Convolution layer 128@1×3	Convolution layer 64@1×3	Convolution layer 128@1×5		
	Convolution layer 64@1×5	Convolution layer 64@1×7	MaxPooling layer 3	Convolution layer 128@1×5	Branch4 & Residual Convolution layer 128@1×5	Dense layer 64
		Convolution layer 128@1×7	Convolution layer 64@1×3	Convolution layer 128@1×5		
K ⁺ (ppm)	Convolution layer 64@1×5	Convolution layer 64@1×7	MaxPooling layer 3	Convolution layer 128@1×5	Branch4 & Residual Convolution layer 128@1×5	Dense layer 64
		Convolution layer 128@1×7	Convolution layer 64@1×3	Convolution layer 128@1×5		
	Convolution layer 64@1×5	Convolution layer 64@1×7	MaxPooling layer 3	Convolution layer 128@1×5	Branch4 & Residual Convolution layer 128@1×5	Dense layer 64
		Convolution layer 128@1×7	Convolution layer 64@1×3	Convolution layer 128@1×5		

表 D1 (续)

	Convolution layer 64@1×1	Convolution layer 64@1×7	MaxPooling layer 3	Convolution layer 128@1×5	Branch4 & Residual Convolution layer 128@1×5	Dense layer 64
Na ⁺ (ppm)		Convolution layer 128@1×7	Convolution layer 64@1×3	Convolution layer 128@1×5	Dense (Mish) 192	
			Convolution layer 64@1×3	Convolution layer 128@1×5	Dense (Relu) 384	

表 D2 时间序列模型配置

Table D2 Configuration of time-series model

RNN	Phenotypes	Settings of parameters
	EC (mS/cm)	Convolution layer 64@1×2 Convolution layer 64@1×1
TD	K ⁺ (ppm)	Convolution layer 64@1×2 Convolution layer 128@1×1
	Na ⁺ (ppm)	Convolution layer 32@1×2 Convolution layer 64@1×1 Convolution layer 64@1×1

表 D3 图像模型模型配置

Table D3 Configuration of image model

Methods	Fine-tuning	Feature Attention
	Reshape	Concatenate
M-R50*3	Convolution layer 512@1×5 Convolution layer 256@1×5	Dense (relu) Dense (softmax)

PLSR 和 SVR 的参数设置见表 D4。

表 D4 PLSR 和 SVR 配置

Table D4 Configuration of PLSR and SVR

Models	Phenotypes	Settings of parameters
PLSR	EC (mS/cm)	nLVs = 12
	K ⁺ (ppm)	nLVs = 12
	Na ⁺ (ppm)	nLVs = 12
SVR	EC (mS/cm)	linear kernel function, C = 0.3
	K ⁺ (ppm)	linear kernel function, C = 0.4
	Na ⁺ (ppm)	RBF kernel function, C = 0.25

AutoML 的参数设置见表 D5。

表 D5 AutoML 配置

Table D5 Configuration of AutoML

Methods	Phenotypes	Settings of parameters
AutoKeras	EC (mS/cm)	max_trails = 30, loss = 'mse', metrics = 'mse', epochs = 500, batch size = 6
	K ⁺ (ppm)	
	Na ⁺ (ppm)	
TPOT	EC (mS/cm)	generation = 50, polulation size = 20, verbosity = 2
	K ⁺ (ppm)	
	Na ⁺ (ppm)	

AutoKears 参数设置: max_trails 为尝试的最大模型结构数量。TPOT: generation 为运行管道优化过程的迭代次数; population 为每代保留在基因编程人群中的个体数量; verbosity 为 TPOT 在运行时传达多少信息。

TPOT 为三种表型选择的模型见表 D6。

表 D6 TPOT 选择的模型

Table D6 The model selected by TPOT

Method	Phenotypes	Models
TPOT	EC (mS/cm)	RidgeCV
	K ⁺ (ppm)	LinearSVR
	Na ⁺ (ppm)	RidgeCV+ExtraTreesRegressor

EC 的模型是一种带有内置交叉验证的岭回归 (Ridge Regression) 模型 (RidgeCV), 它属于线性回归的范畴, 但在损失函数中添加了 L2 范数正则化项。岭回归是一种专门用于处理特征间存在高度共线性 (即输入特征之间相关度高) 的情况, 通过引入正则化项来减少模型复杂度并防止过拟合。在 RidgeCV 模型中, “CV” 代表交叉验证 (Cross-Validation), 这是一种评估统计模型泛化能力的方法, 可以系统地应用于不同的子数据集以选择最佳的正则化参数。交叉验证通过将数据集分割成多个小组, 并将其中一部分用作验证模型性能的验证集, 其余部分用作校准集, 通过这种方式来评估模型的稳定性和估算能力。RidgeCV 会自动执行这一过程来找到最佳的正则化参数 alpha。TPOT 选择的 RidgeCV 模型使用的是 scikit-learn 库中该模型默认参数设置。K⁺的模型是线性支持向量回归模型 (LinearSVR), 它是 SVM 用于回归问题的一个变种。与用于分类的 SVM 模型类似, 回归版本的 SVM 模型试图找到一个线性函数来估算实数值输出, 同时最小化估算误差和模型复杂度, 以达到良好的泛化能力。模型参数设置为 C = 1.0 (正则化参数)、dual = True (指示优化问题为求解对偶问题)、

$\epsilon = 0.1$ （惩罚参数的阈值）、 $\text{loss} = \text{'epsilon_insensitive'}$ （损失函数类型）和 $\text{tol} = 0.1$ （模型求解的容忍度）。 Na^+ 的模型由两个步骤组成，先使用 `RidgeCV` 对原始数据进行估算，并将估算结果作为额外的特征，然后在经过这种扩展后的特征集上训练一个极端随机树回归模型（`ExtraTreesRegressor`）。极端随机树是一种集成学习方法，它通过在树的每个分割上随机选择特征和分割点来拟合多个决策树，并对它们的估算结果进行平均以提高模型的估算精度和控制过拟合。这种结合使用不同模型的策略可以捕捉数据中的不同信息，并有可能提高最终模型的估算性能。第一个步骤中的 `RidgeCV` 依然使用的默认参数设置，第二个步骤中 `ExtraTreesRegressor` 参数设置为 $\text{bootstrap} = \text{True}$ （启用自助采样）、 $\text{max_features} = 0.65$ （查找最佳分割时要考虑的特征的比例）、 $\text{min_samples_leaf} = 4$ （叶节点必须有的最小样本数量）、 $\text{min_samples_split} = 5$ （拆分内部节点所需的最小样本数）和 $\text{n_estimators} = 100$ （模型集成中的评估器数量）。

在实验中，AutoKeras 为 EC、 K^+ 和 Na^+ 这三种植物表型特性的估算任务，分别探索了适合不同数据模态的网络结构。尽管表型不同，但 AutoKeras 为它们寻找到的网络框架类型保持了一致性。三种数据模态对应的网络架构见表 D7。

表 D7 AutoKeras 选择的网络框架

Table D7 The network framework selected by AutoKeras

Method	Data modalities	Networks
AutoKeras	Spectra	Convolution layer
		Convolution layer
		Pooling layer
	Time-series	BRNN BRNN
	Image	ResNet50

在处理光谱数据时，AutoKeras 构建了一个包含两个连续的卷积层后跟一个池化层的模块，这样的模块将重复两次。在这个模块中，针对 EC 估算的网络使用了 16 个卷积核，而 K^+ 和 Na^+ 估算的网络配置中，卷积核的数量为 32 个。在三个表型估算模型中，池化层均采用了大小为 2 的核进行下采样。对于时间序列数据，AutoKeras 构建了一个由两个 BRNN 堆叠而成的结构，其中每个 BRNN 均采用 LSTM 单元，每个 LSTM 单元配置了 28 个神经元。至于图像数据，AutoKeras 则直接采用了预训练的 ResNet50 神经网络。ResNet50 网络广泛应用于图像识别任务，其出色的特征提取能力已获得证明。