Neural Architecture Search for Crop Yield Prediction Using Multimodal Data

Heera Lal Khatri*, Jade Lyons, Thomas Dowling, and Paul Greaney*

Department of Computing, Atlantic Technological University, Letterkenny, Donegal, Ireland

Abstract
Accurate crop yield prediction is essential for sustainable agriculture and food security. Traditional methods often fall short in addressing the complex factors influencing crop growth. This paper explores the use of neural architecture search (NAS) combined with multimodal data integration to improve prediction accuracy. By incorporating diverse data sources such as unmanned aerial vehicle (UAV) imagery, and weather data, the research develops a framework for crop yield prediction. NAS techniques systematically identify optimal neural network architectures to handle these varied datasets effectively. The approach is validated with real-world agricultural data, showing that NAS-optimised models significantly outperform traditional methods. This work enhances precision agriculture, enabling better resource allocation and sustainable farming practices.

Keywords
Crop Yield Prediction, Neural Architecture Search, Multimodal Data Integration, Machine Learning

1. Introduction

In today's agriculture sector, accurately predicting crop yields is essential for making well-informed decisions and promoting sustainable farming practices. Precise yield forecasts can significantly impact the success and flexibility of agricultural operations, aiding in efficient resource allocation and adaptation to market fluctuations [1]. Traditional prediction methods often struggle to consider the complex array of factors that influence crop growth and productivity [2]. This challenge has sparked a wave of enthusiasm for using cutting-edge technologies such as machine learning and data science to enhance predictive models.

One of the encouraging approaches is integrating neural architecture search (NAS) techniques with multimodal data. NAS methods have the potential to revolutionise how we perceive and forecast crop yields [3]. By utilising various data sources such as UAV imagery, weather data, soil conditions, and agronomic practices, multimodal approaches offer a more comprehensive understanding of the factors shaping crop outcomes. This paper explores the potential of NAS-based methodologies in enhancing crop yield prediction through multimodal data fusion, ultimately aiming to advance sustainable agricultural practices and provide stakeholders with valuable insights to guide their decisions.

Crop yield prediction is incredibly important in modern agriculture. It's not just about estimating how much food we'll have; it's about ensuring global food security, maintaining economic stability, and promoting sustainable development. As the world's population continues to grow and is expected to exceed 9 billion by 2050, the demand for food will skyrocket, putting huge pressure on our agricultural systems [4]. Accurately forecasting crop yields becomes crucial in this scenario, as it allows farmers and stakeholders to plan ahead, allocate resources effectively, and manage risks proactively. By knowing what to expect from future harvests, farmers can make smart decisions about water usage, fertilisers, and pesticides, ensuring they get the most out of their crops while minimising waste and environmental harm. Moreover, reliable yield predictions

AICS'24: 32nd Irish Conference on Artificial Intelligence and Cognitive Science, December 09—10, 2024, Dublin, Ireland
* Corresponding Author
l00178511@atu.ie (Heera Lal Khatri); paul.greaney@atu.ie (Paul Greaney)

enable policymakers to create well-informed policies, identify areas that are particularly vulnerable, and develop specific actions to make agriculture stronger and fight against food insecurity.

Machine learning (ML) has revolutionised agriculture, offering powerful tools to analyse vast datasets and address challenges such as predicting yields, detecting diseases, and optimising resource management [5]. Neural networks, known for capturing complex agricultural patterns, enable precision farming and sustainable practices across diverse conditions. Despite technological advancements, accurately predicting crop yields remains complex due to varied environmental factors and data challenges [6]. Integrating multimodal data enhances prediction accuracy and strength, yet requires advanced computational techniques such as NAS for optimal model design [7]. This research aims to improve crop yield prediction through NAS and multimodal data fusion, encouraging sustainable agriculture by enabling informed decisions customised to diverse farming needs.

2. Related work

Understanding the landscape of crop yield prediction requires exploring various techniques and emerging trends that shape agricultural research [8]. Deep learning, a subset of AI, trains neural networks on large datasets to learn and predict without explicit programming [9]. These networks, with deep layers of interconnected nodes, automatically extract complex data patterns, transforming fields like computer vision, natural language processing, and agriculture. In farming, they enhance tasks such as crop yield prediction, disease detection, and precision farming, offering efficient solutions to agricultural challenges.

Convolutional neural networks (CNNs) excel in spatial data analysis, crucial for tasks like analysing UAV imagery in agriculture [10]. They extract hierarchical features from images, aiding in crop health monitoring and yield prediction by finding complex details in vegetation and land use. Recurrent neural networks (RNNs), on the other hand, specialise in sequential data, ideal for analysing time-series data such as weather patterns and soil moisture [11]. RNNs capture dynamic environmental changes over time, enhancing yield prediction accuracy by modelling long-term trends.

Despite their effectiveness, the success of deep learning, CNNs, and RNNs in agriculture depends on data quality and diversity [12]. Representative training datasets are vital to developing reliable models. Challenges like data pre-processing, feature engineering, and model interpretability require attention to maximise model applicability.

Neural architecture search (NAS) has gained considerable attention in recent years as a means to automate the design of neural network architectures [13]. NAS algorithms employ search strategies, such as reinforcement learning, evolutionary algorithms, or gradient-based optimisation, to explore the vast space of possible architectures and identify configurations that optimise performance on a given task. By automating the design process, NAS enables the discovery of architectures that are tailored to specific applications and data characteristics, leading to improved performance and efficiency compared to manually designed architectures.

In agriculture, NAS has demonstrated its potential in fine-tuning neural network architectures for a range of tasks, spanning from crop classification and disease detection to yield prediction [14]. By adapting architectures to the unique characteristics of agricultural data, NAS enables the creation of models that are better suited to capture relevant features and patterns. However, the application of NAS in crop yield prediction using multimodal data remains relatively unexplored. Few studies have investigated the potential of NAS to automatically design architectures that effectively leverage diverse data sources for accurate and strong yield predictions [15].

NAS techniques have found applications across a wide range of domains, including computer vision, natural language processing, healthcare, and finance [16]. In computer vision, NAS has been used to automatically design architectures for image classification, object detection, and image segmentation tasks. These methods have led to the discovery of novel network architectures that outperform handcrafted designs on benchmark datasets such as ImageNet. In natural language processing, NAS has been applied to tasks such as machine translation, text generation, and sentiment analysis, leading to the development of state-of-the-art language models with improved performance [17].

Despite the potential benefits, several challenges exist in the integration of NAS with multimodal data for crop yield prediction [18]. These include the complexity of agricultural systems, the heterogeneity of data sources, and the need for scalable and efficient search algorithms. Addressing these challenges requires interdisciplinary collaboration between researchers in machine learning, agriculture, and data science. Moreover, opportunities exist for developing novel NAS algorithms tailored to the specific requirements of agricultural applications, as well as for exploring innovative approaches to multimodal data fusion and feature representation [19].

NAS continues to evolve rapidly, with new techniques and approaches being proposed to overcome existing limitations and expand its applicability to diverse domains [20]. In the context of crop yield prediction, NAS offers the potential to unlock deeper insights by optimising model architectures that effectively combine the strengths of multiple data sources, such as UAV images, weather data, soil conditions, and historical yield data [21]. By automating this process, researchers can systematically evaluate a wide range of architectures and fine-tune them to maximise accuracy, efficiency, and scalability.

One of the key advantages of NAS is its ability to automate hyperparameter tuning, a traditionally labour-intensive task [22]. This includes not only the search for the most suitable architecture but also fine-tuning elements such as the learning rate, batch size, and layer configurations, which are critical for achieving optimal performance in yield prediction tasks. Automating this aspect significantly reduces the manual intervention required, allowing researchers and practitioners to focus more on interpreting results and improving model deployment strategies.

Additionally, the use of NAS can help address scalability concerns in agriculture [23]. Agricultural data is often vast and continuously expanding, especially as more farms adopt precision agriculture techniques and deploy Internet of Things (IoT) devices to gather real-time data [24]. NAS can assist in identifying architectures that not only provide high accuracy but also scale effectively with larger datasets, reducing computational costs while maintaining prediction accuracy [25]. This scalability is essential for real-world applications where processing efficiency can be as important as predictive performance.

Another challenge is the effective integration and interpretation of multimodal data [26]. While NAS excels at automating the design of neural networks, the unique challenges presented by multimodal data – such as the need for appropriate feature extraction, data fusion techniques, and handling missing or imbalanced data – still require significant attention. Future research should explore how NAS can be further tailored to address these challenges, particularly in agricultural contexts where data is often noisy or incomplete [27].

The integration of cloud-based platforms and NAS could further enhance scalability and accessibility [28]. Cloud-based platforms offer the computational power necessary to perform large-scale NAS operations, enabling smallholder farmers or agricultural organisations with limited resources to access sophisticated yield prediction models [29]. By using cloud-based NAS frameworks, model deployment can also become more seamless, allowing updates and improvements to be applied dynamically as new data becomes available [30].

NAS represents a frontier in agricultural technology, offering the potential to revolutionise the way crop yield predictions are made [31]. Its ability to automatically design architectures that are tailored to complex, multimodal datasets holds significant promise for improving accuracy, scalability, and efficiency in predictive modelling. Nonetheless, addressing the computational, data integration, and interpretability challenges associated with NAS will be crucial to unlocking its full potential in this domain. Through continued innovation and interdisciplinary collaboration, NAS could become a pivotal tool in driving sustainable agricultural practices and enhancing global food security.

3. Methodology

NAS automates the design of neural network architectures for crop yield prediction using multimodal data. The search space includes configurations like CNNs for image data, RNNs for time-series data, and hybrid models. Architectural parameters vary, including the number of layers, layer types, filter sizes, learning rate, batch size, and dropout rate. Performance is evaluated using root mean squared error (RMSE), mean absolute error (MAE), and R-squared (R^2), along with computational efficiency and model generalisability. The NAS process involves defining the search space, initialising the search strategy, generating candidate architectures, training and evaluating them, and iterating until a preferred model is identified as shown in Figure 3.1.

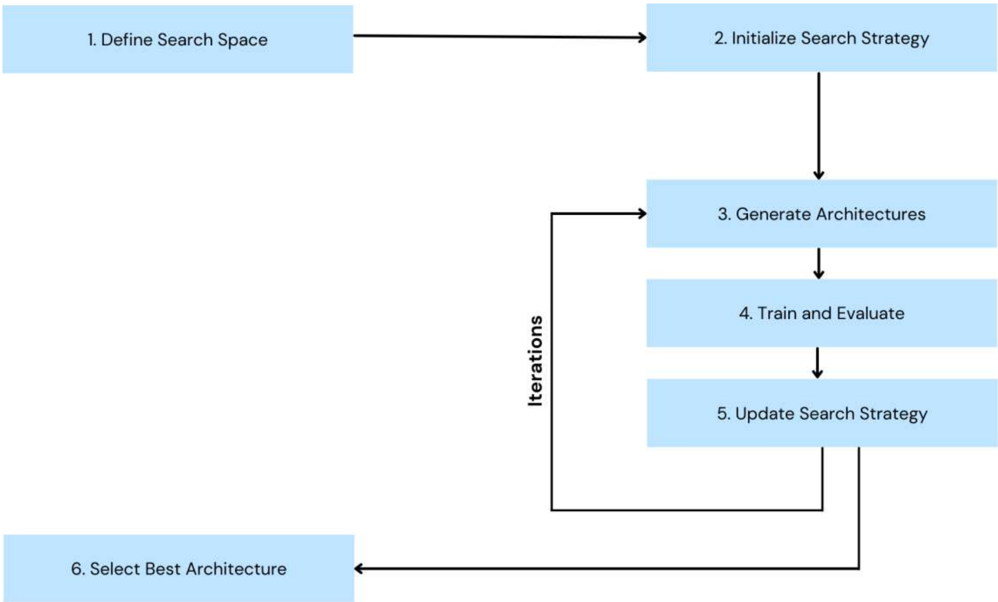

Figure 3.1 NAS Process Workflow

The chosen methodology addresses the complexity and variability of factors affecting crop yield, including weather conditions, soil characteristics, and plant health. Integrating multimodal data, such as weather, soil measurements, and UAV imagery, captures the full scope of variables influencing agricultural productivity, leading to more accurate and reliable prediction.

3.1. Data Acquisition Process

Data acquisition involves identifying relevant sources like weather stations, soil sensors, and UAV imagery. For this study, pre-existing data from a recent study [32] was acquired. This includes six years of image and weather data gathered in several regions of Japan, which we use in developing model architectures using NAS. Sample images captured by UAVs as shown in *Figure 3.2* and multispectral cameras provide critical information about crop health.

The weather data used in this project includes several key features that help analyse crop yields. Features recorded here are Date, Year, Yield, Normal and Observed. The Date and Year features record when observations were made, allowing for the study of seasonal trends and yearly variations. Location specifies the geographical area, which is vital as different regions have unique climatic and soil conditions affecting yield. The Yield feature contains the actual crop yield, the primary target for the model to predict. Normal refers to the expected average yield based on historical data, offering a baseline for comparison. Observed represents the real, recorded yield for that specific time and place, allowing direct comparison with both the predicted and expected outcomes. Together, these features provide a detailed view for predicting and analysing agricultural productivity.

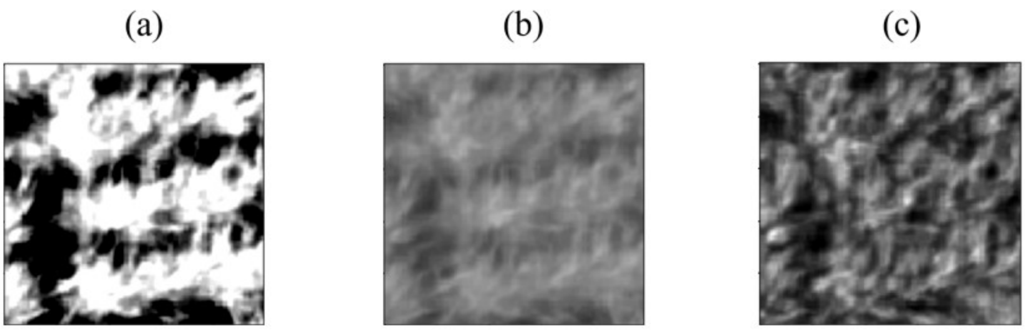

Figure 3.2 Images captured through UAV and multispectral cameras: (a) green, (b) NIR (near-infrared), and (c) red

3.2. Data Processing and Integration

Following the acquisition of multimodal datasets, including UAV imagery and weather data, a preprocessing step was applied to ensure consistency, reliability, and compatibility across the different data types. For image data, UAV-captured images were directly used, preserving the original resolution. For weather data, temporal features such as date and year were incorporated to capture seasonal patterns, while spatial features like Location were included to account for regional differences. Yield-related attributes, including normal yield (expected average yield based on historical data) and observed yield (actual recorded yield), were standardised and used as primary inputs for the predictive model.

Once preprocessed, the datasets were aligned and integrated into a unified format, ensuring temporal and spatial synchronisation between weather data and UAV imagery. This combined dataset was then partitioned into training, validation, and test subsets, following an 80:20 split. The training set, comprising 80% of the data, was used to train neural network models by iteratively optimising parameters. The validation set was reserved for hyperparameter fine-tuning and monitoring overfitting, while the independent test set served as a reliable benchmark for evaluating the model's real-world predictive accuracy.

Python scripts were employed throughout the data processing and integration pipeline. Libraries such as Pandas were used for data cleaning and manipulation, while TensorFlow and Keras supported the development and fine-tuning of the predictive models. This precise data preparation approach established a solid foundation for building and testing high-performing models for crop yield estimation.

3.3. Research Architecture

A high-level overview of the method is illustrated in *Figure 3.3*. This integrates both image and weather data into a unified model. By combining visual and environmental inputs, this approach enhances the accuracy of yield forecasts and provides more comprehensive insights.

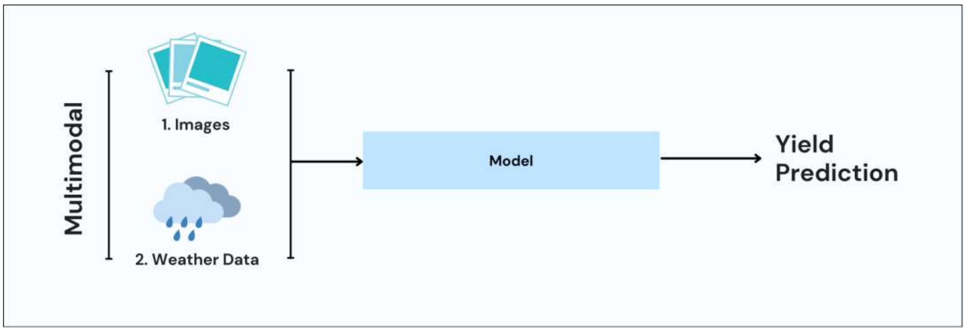

Figure 3.3 Research Design Architecture

Neural architecture search refines the prediction process by exploring different configurations of convolutional neural networks (CNNs) for image analysis and optimising the integration of multimodal data, such as combining image and weather information. By automating the search for the best-performing architectures, NAS tailors the CNN to effectively capture both spatial and environmental factors, improving yield prediction.

Data mapping aligns temporal and spatial dimensions of weather data with UAV imagery, ensuring accurate reflection of conditions and time frames. This alignment is crucial for the model to learn relationships between data types. By processing image and weather data pairs, the integrated model leverages complementary strengths of both data types, improving predictive performance.

4. Experiments and Results

4.1. Initial Model Training

The initial model training utilised a multimodal dataset comprising images and weather data. This dataset was split into training, validation, and test sets following an 80:20 ratio. Specifically, 80% of the data was dedicated to training, where the model's parameters were optimised, and its architecture refined. The remaining 20% was divided between validation and test sets. The validation set played a critical role in monitoring performance and fine-tuning hyperparameters to avoid overfitting, ensuring the model could generalise effectively to unseen data.

After training, the model was validated and tested on test datasets. The validation process allowed for iterative adjustments to hyperparameters by comparing the model's predictions with actual outcomes, ensuring strong performance without overfitting. Finally, the model was evaluated on test set to measure its ability to generalise beyond the data seen during training and validation. The test results provided key performance metrics, offering insights into the model's predictive accuracy and practical applicability in real-world crop yield prediction tasks.

4.2. Loss Function and Evaluation Metrics

Loss is a measure of how well the model's predictions match the actual data. In our case, it reflects how closely the predicted crop yields align with the true yields. The loss used in this study is mean-squared error (MSE), which calculates the average squared difference between the predicted and actual values. MSE was chosen for its sensitivity to larger errors, making it a suitable loss function for regression tasks like crop yield prediction.

The performance of the model was evaluated using three key metrics: root mean squared error (RMSE), mean absolute error (MAE), and R-squared (R^2).

RMSE measures the average magnitude of the errors in a set of predictions. Lower values indicate better predictive accuracy. It is calculated as follows:

$$RMSE = \sqrt{\frac{1}{2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where y_i is the true value and $\hat{y}_i$ is the value predicted by the model.

MAE calculates the average of the absolute differences between predicted and actual values. It provides a measure of the model's accuracy in terms of absolute error. It is calculated as

$$\mathrm{MAE} = \frac{1}{n}\Sigma_{i=1}^{n}|y_i - \hat{y}_i|$$

Finally, the R^2 metric indicates how well the model explains the variation in the dataset. A higher R^2 value indicates a better fit of the model to the data. It is calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

4.3. Architecture Development and Results

During the development of the architecture, the integrated multimodal data was processed to generate predictions. Initially, we began with a smaller number of epochs to quickly evaluate the model's baseline performance. Following this, the number of epochs was progressively increased from 10 to 140 to allow the model more time to potentially learn complex patterns within the data and improve its predictive accuracy. This gradual increase in epochs was crucial for fine-tuning the model, ensuring it had sufficient training time to enhance performance without overfitting or under fitting. Throughout this process, we closely monitored key performance metrics to assess improvements and optimise the model effectively.

The training loss graph in Figure 4.1 plots the loss value against the 10 epochs during the model training process.

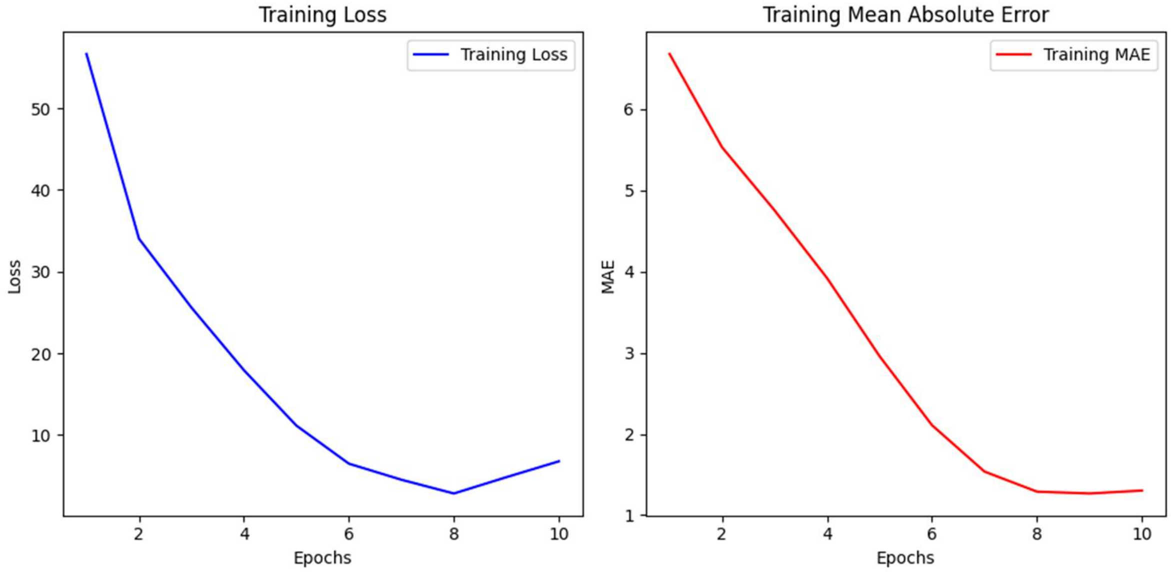

Figure 4.1 Training Loss and Training MAE using 10 Epochs

The training MAE graph in Figure 4.1 above shows the average absolute difference between the predicted and actual values over the epochs. Similar to the training loss, the MAE starts high initially due to the model's lack of familiarity with the data patterns. As the model learns, the MAE decreases, signifying that the predictions are getting closer to the actual values. Eventually, the MAE stabilises, showing that the model's average prediction error is consistent and has reached its minimum possible value given the data and model complexity.

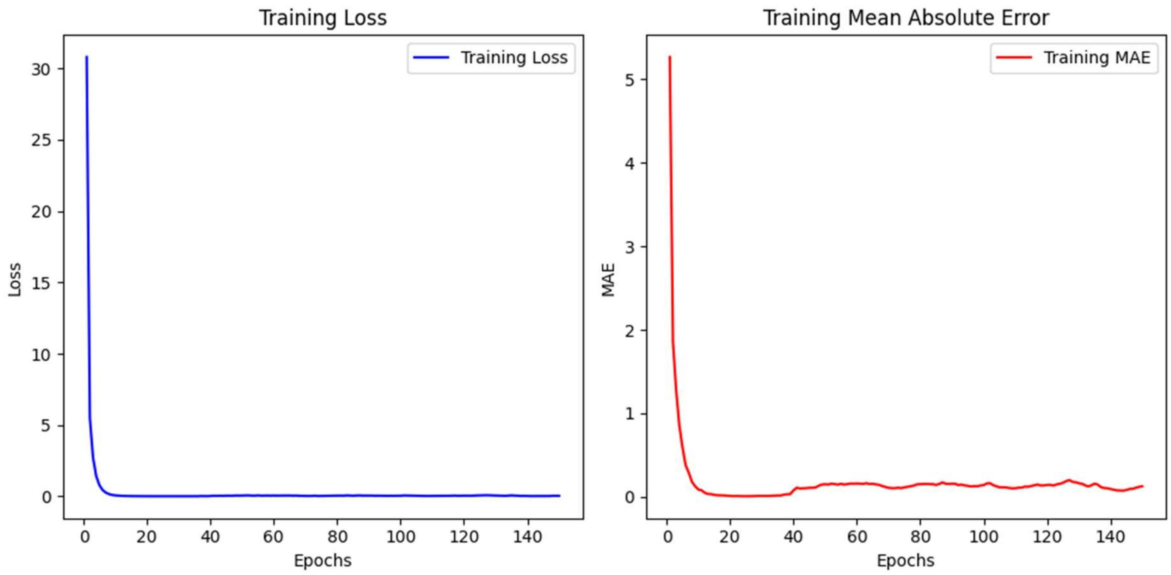

Figure 4.2 Training Loss and Training MAE using 140 Epochs

Both graphs in Figure 4.1 above show a decreasing trend, indicating successful learning as the model improves its predictions by adjusting parameters. Continuous decreases in loss and MAE without levelling off could signal overfitting, where the model memorises training data instead of generalising. However, similar decreasing trends and stabilisation in validation loss and MAE suggest overfitting is less likely.

The training loss graph on the left side of Figure 4.2 shows the loss value against the number of epochs, extending over 140 epochs. This extension of the number of epochs reveals that training the model beyond 140 epochs does not yield any further reduction in loss or MAE. The loss then stabilises near zero, indicating effective learning without significant improvement from further training. The consistent low loss suggests the model avoids overfitting.

The graph in Figure 4.3 illustrates the training and validation loss on the left and the training and validation mean absolute error (MAE) on the right, providing insights into the model's performance over 120 generations of NAS. In this process, instead of training a single CNN architecture, the neural architecture search (NAS) generates and evaluates multiple CNNs across generations. Each generation involves training a set of candidate architectures for a specified number of epochs, after which the best-performing architectures are selected for the next generation. This evolutionary approach refines the architectures, leading to improved models over time.

In the left plot, the loss values for both training (blue line) and validation (orange line) are shown. This decline in loss suggests that the NAS process is effectively generating better architectures over successive generations. Importantly, the close alignment between the training and validation losses indicates that the model generalises well to new data, with no significant overfitting. This generational improvement reflects the ability of NAS to explore and optimise different architecture configurations for enhanced prediction accuracy.

Both the training MAE (red line) and validation MAE (green line) show a steep decline in the early stages, similar to the loss curves, before levelling off as training progresses. The consistent decrease in MAE suggests that the successive models are becoming increasingly accurate in their predictions. Notably, the validation MAE remains consistently lower than the training MAE, highlighting the model's strong performance on unseen data.

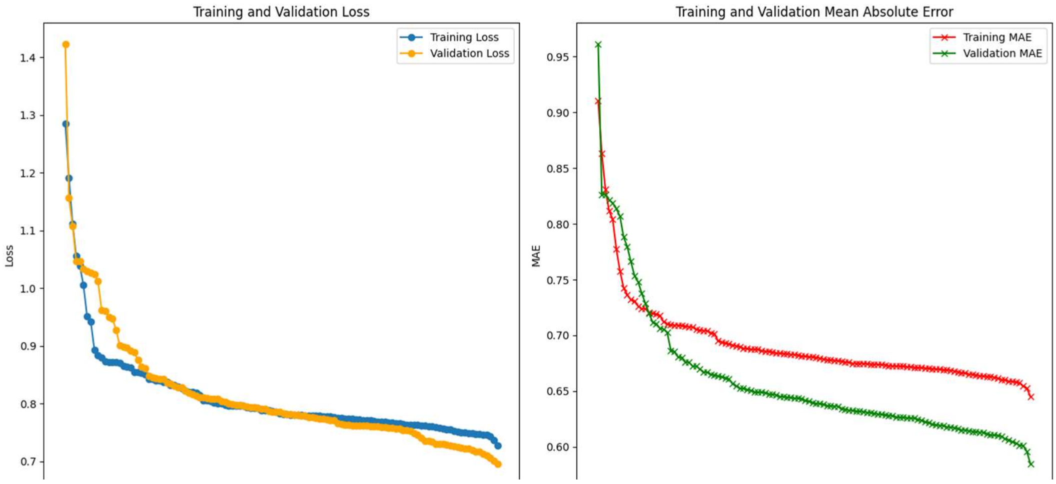

Figure 4.3 Training and Validation Loss and MAE with 120 numbers of generations

The NAS-optimised model significantly outperforms the best CNN architecture in [32] in predicting crop yields. It achieves a lower RMSE (0.70 vs. 0.821 t/ha), better accuracy with a higher R^2 (0.90 vs. 0.83–0.86), and faster convergence, requiring fewer epochs. The best traditional model in [32] had a validation RMSE of 0.831, while the NAS model had a lower validation RMSE of 0.75, showing improved generalisation. NAS automatically optimised the architecture for multimodal data (images and weather), resulting in superior performance and efficiency.

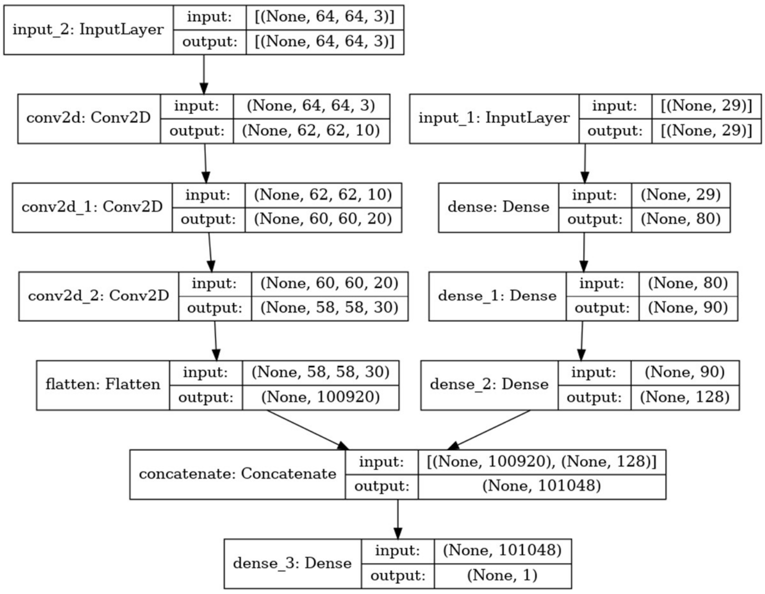

Figure 4.4 Best Architecture Design

The best architecture, as shown in Figure 4.4, combined convolutional layers for image data processing and dense layers for tabular input. The convolutional layers efficiently extracted spatial features, while the dense layers captured interactions between tabular data features. Outputs from both pipelines were merged using a concatenate layer to enable joint learning from multimodal data, improving model performance. Key optimisation techniques, such as dropout layers and pooling, played a role in earlier experiments by mitigating overfitting and reducing computational overhead. The model performed best with a learning rate of 0.001, batch sizes of 32 to 64, and convolutional filters ranging from 32 to 128, balancing feature extraction and avoiding overfitting. The network's moderate depth provided sufficient learning capacity without excess complexity. The iterative NAS process was crucial for refining the architecture, leading to an optimal final model. These design choices resulted in a well-balanced architecture that delivered excellent performance on the task.

Overall, these results indicate that neural architecture search (NAS) generated successful architectures for multimodal crop yield prediction. The loss and mean absolute error (MAE) metrics demonstrate significant improvement early on and stabilise at low values, confirming that the model has effectively learned and can generalise well. This suggests that the model is capable of capturing complex patterns in the multimodal data, contributing to accurate predictions of crop yield.

These results are based on data from specific regions in Japan, which may limit the model's generalisability. Differences in crop varieties, soil types, weather patterns, and farming practices in other regions could impact its performance. For example, a model trained on temperate climate data may not perform well in tropical regions due to varying crop growth cycles and environmental factors. To improve the model's adaptability, future work should test it in diverse geographic regions and crop types. Expanding the dataset with additional data sources, such as soil sensor data or remote sensing from different sensors, could further enhance its predictive capabilities and confirm its versatility in precision agriculture.

5. Conclusions and Future Work

This study explored the application of neural architecture search (NAS) combined with multimodal data for crop yield prediction. The dataset was partitioned into training, validation, and test sets, and a supervised neural network model was optimised and evaluated. To demonstrate the effectiveness of NAS, a direct comparison was made between the original, manually designed convolutional neural network (CNN) model and the NAS-generated CNN model.

The NAS-optimised model outperformed the traditional CNN in key performance metrics, particularly in reducing training loss and mean absolute error (MAE). The NAS approach explored various architectures, identifying those that were more suitable for the complexity of the multimodal data, which included both images and weather data. In contrast, the manually designed CNN, while effective, did not achieve the same level of optimisation in learning from the integrated data. This difference was evident in the lower MAE achieved by the NAS-optimised model, indicating more accurate predictions of crop yield.

The NAS-generated model demonstrated faster convergence, with a steeper reduction in training loss, indicating improved training efficiency and accuracy. This was confirmed by its strong performance on the validation and test sets, showing good generalisation and minimal overfitting due to careful hyperparameter tuning. NAS enabled the automated discovery of network architectures better suited for the multimodal dataset. While the original CNN performed well, the NAS-optimised model yielded significantly better results, highlighting the benefits of NAS for complex datasets.

Future research can refine NAS techniques for optimising CNN architectures and better capturing weather trends by incorporating advanced temporal models like recurrent neural networks (RNNs) or temporal convolutional networks (TCNs). Additional data types, such as soil composition and historical yields, needed to be integrated to enhance accuracy. Another focus should be deploying these models in real-time agricultural systems for dynamic decision-making and crop yield predictions, with agronomist collaboration for field validation.

Acknowledgments

We extend sincere thanks to the authors of [32] for sharing the dataset used in this study.

References

[1] P. K. Tripathi, C. K. Singh, R. Singh, and A. K. Deshmukh, "A farmer-centric agricultural decision support system for market dynamics in a volatile agricultural supply chain," *Benchmarking: An International Journal*, vol. 30, no. 10, pp. 3925–3952, Jan. 2022, doi: 10.1108/BIJ-12-2021-0780.

[2] E. Elbasi *et al.*, "Crop Prediction Model Using Machine Learning Algorithms," *Applied Sciences*, vol. 13, no. 16, Art. no. 16, Jan. 2023, doi: 10.3390/app13169288.

[3] N. Darra, B. Espejo-Garcia, A. Kasimati, O. Kriezi, E. Psomiadis, and S. Fountas, "Can Satellites Predict Yield? Ensemble Machine Learning and Statistical Analysis of Sentinel-2 Imagery for Processing Tomato Yield Prediction," *Sensors*, vol. 23, no. 5, Art. no. 5, Jan. 2023, doi: 10.3390/s23052586.

[4] K. E. Giller *et al.*, "The future of farming: Who will produce our food?," *Food Sec.*, vol. 13, no. 5, pp. 1073–1099, Oct. 2021, doi: 10.1007/s12571-021-01184-6.

[5] T. Ayoub Shaikh, T. Rasool, and F. Rasheed Lone, "Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming," *Computers and Electronics in Agriculture*, vol. 198, p. 107119, Jul. 2022, doi: 10.1016/j.compag.2022.107119.

[6] J. P. Bharadiya, N. T. Tzenios, and M. Reddy, "Forecasting of Crop Yield using Remote Sensing Data, Agrarian Factors and Machine Learning Approaches," *Journal of Engineering Research and Reports*, vol. 24, no. 12, Art. no. 12, Apr. 2023, doi: 10.9734/jerr/2023/v24i12858.

[7] T. Zhang, "Practical Solutions to Neural Architecture Search on Applied Machine Learning," 2024.

[8] S. Poggi *et al.*, "Chapter Seven - How can models foster the transition towards future agricultural landscapes?," in *Advances in Ecological Research*, vol. 64, D. A. Bohan and A. J. Vanbergen, Eds., in The Future of Agricultural Landscapes, Part II, vol. 64. , Academic Press, 2021, pp. 305–368. doi: 10.1016/bs.aecr.2020.11.004.

[9] S. Saha *et al.*, "Hierarchical Deep Learning Neural Network (HiDeNN): An artificial intelligence (AI) framework for computational science and engineering," *Computer Methods in Applied Mechanics and Engineering*, vol. 373, p. 113452, Jan. 2021, doi: 10.1016/j.cma.2020.113452.

[10] L. Wei, Y. Luo, L. Xu, Q. Zhang, Q. Cai, and M. Shen, "Deep Convolutional Neural Network for Rice Density Prescription Map at Ripening Stage Using Unmanned Aerial Vehicle-Based Remotely Sensed Images," *Remote Sensing*, vol. 14, no. 1, Art. no. 1, Jan. 2022, doi: 10.3390/rs14010046.

[11] Y. Chen, Q. Cheng, Y. Cheng, H. Yang, and H. Yu, "Applications of Recurrent Neural Networks in Environmental Factor Forecasting: A Review," *Neural Computation*, vol. 30, no. 11, pp. 2855–2881, Nov. 2018, doi: 10.1162/neco_a_01134.

[12] K. Alibabaei *et al.*, “A Review of the Challenges of Using Deep Learning Algorithms to Support Decision-Making in Agricultural Activities,” *Remote Sensing*, vol. 14, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/rs14030638.
[13] Z. Zhang, Z. Wang, and I. Joe, “CAM-NAS: An Efficient and Interpretable Neural Architecture Search Model Based on Class Activation Mapping,” *Applied Sciences*, vol. 13, no. 17, Art. no. 17, Jan. 2023, doi: 10.3390/app13179686.
[14] Y. Gulzar, Z. Ünal, H. Aktaş, and M. S. Mir, “Harnessing the Power of Transfer Learning in Sunflower Disease Detection: A Comparative Study,” *Agriculture*, vol. 13, no. 8, Art. no. 8, Aug. 2023, doi: 10.3390/agriculture13081479.
[15] Y. Akhauri and M. S. Abdelfattah, “Encodings for Prediction-based Neural Architecture Search,” Mar. 04, 2024, *arXiv*: arXiv:2403.02484. doi: 10.48550/arXiv.2403.02484.
[16] J. Waring, C. Lindvall, and R. Umeton, “Automated machine learning: Review of the state-of-the-art and opportunities for healthcare,” *Artificial Intelligence in Medicine*, vol. 104, p. 101822, Apr. 2020, doi: 10.1016/j.artmed.2020.101822.
[17] J. Sawicki, M. Ganzha, and M. Paprzycki, “The State of the Art of Natural Language Processing—A Systematic Automated Review of NLP Literature Using NLP Techniques,” *Data Intelligence*, vol. 5, no. 3, pp. 707–749, Aug. 2023, doi: 10.1162/dint_a_00213.
[18] X. Zhang *et al.*, “Multi-modal spatio-temporal meteorological forecasting with deep neural network,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 188, pp. 380–393, Jun. 2022, doi: 10.1016/j.isprsjprs.2022.03.007.
[19] Z. Ramzan, H. M. S. Asif, I. Yousuf, and M. Shahbaz, “A Multimodal Data Fusion and Deep Neural Networks Based Technique for Tea Yield Estimation in Pakistan Using Satellite Imagery,” *IEEE Access*, vol. 11, pp. 42578–42594, 2023, doi: 10.1109/ACCESS.2023.3271410.
[20] Y. Liu, Y. Sun, B. Xue, M. Zhang, G. G. Yen, and K. C. Tan, “A Survey on Evolutionary Neural Architecture Search,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 2, pp. 550–570, Feb. 2023, doi: 10.1109/TNNLS.2021.3100554.
[21] L. Zhang and L. Zhang, “Artificial Intelligence for Remote Sensing Data Analysis: A review of challenges and opportunities,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 10, no. 2, pp. 270–294, Jun. 2022, doi: 10.1109/MGRS.2022.3145854.
[22] M. S. Ismail, R. Samad, D. Pebrianti, M. Mustafa, and N. R. Hasma Abdullah, “Comparative Analysis of Deep Learning Models for Sheep Detection in Aerial Imagery,” in *2024 9th International Conference on Mechatronics Engineering (ICOM)*, Aug. 2024, pp. 234–239. doi: 10.1109/ICOM61675.2024.10652292.
[23] J. Monteiro, J. Barata, M. Veloso, L. Veloso, and J. Nunes, “A scalable digital twin for vertical farming,” *J Ambient Intell Human Comput*, vol. 14, no. 10, pp. 13981–13996, Oct. 2023, doi: 10.1007/s12652-022-04106-2.
[24] V. P. Kour and S. Arora, “Recent Developments of the Internet of Things in Agriculture: A Survey,” *IEEE Access*, vol. 8, pp. 129924–129957, 2020, doi: 10.1109/ACCESS.2020.3009298.
[25] X. Dong, L. Liu, K. Musial, and B. Gabrys, “NATS-Bench: Benchmarking NAS Algorithms for Architecture Topology and Size,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3634–3646, Jul. 2022, doi: 10.1109/TPAMI.2021.3054824.
[26] A. Singh and H. Nair, “A Neural Architecture Search for Automated Multimodal Learning,” *Expert Systems with Applications*, vol. 207, p. 118051, Nov. 2022, doi: 10.1016/j.eswa.2022.118051.
[27] Z. Jan *et al.*, “Artificial intelligence for industry 4.0: Systematic review of applications, challenges, and opportunities,” *Expert Systems with Applications*, vol. 216, p. 119456, Apr. 2023, doi: 10.1016/j.eswa.2022.119456.
[28] D. S. Adji, G. Eduardus, Michael, Minawati, and W. Budiharto, “Performance Analysis Between Cloud Storage and NAS to Improve Company’s Performance: A Literature Review,” in *2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI)*, Oct. 2021, pp. 263–268. doi: 10.1109/ICCSAI53272.2021.9609792.
[29] J. Liu, J. Xiang, Y. Jin, R. Liu, J. Yan, and L. Wang, “Boost Precision Agriculture with Unmanned Aerial Vehicle Remote Sensing and Edge Intelligence: A Survey,” *Remote Sensing*, vol. 13, no. 21, Art. no. 21, Jan. 2021, doi: 10.3390/rs13214387.
[30] K. Y. Chan *et al.*, “Deep neural networks in the cloud: Review, applications, challenges and research directions,” *Neurocomputing*, vol. 545, p. 126327, Aug. 2023, doi: 10.1016/j.neucom.2023.126327.

[31] R. Jakhmola-Mani *et al.*, “Exploring Machine Learning Algorithms for Gene Function Prediction in Crops,” in *Bioinformatics for Plant Research and Crop Breeding*, John Wiley & Sons, Ltd, 2024, pp. 159–183. doi: 10.1002/9781394209965.ch6.
[32] M. Mia *et al.*, “Multimodal Deep Learning for Rice Yield Prediction Using UAV-Based Multispectral Imagery and Weather Data,” *Remote Sensing*, vol. 15, p. 2511, May 2023, doi: 10.3390/rs15102511.