An Intelligent Crop Recommendation Framework Using a Deep Learning and Metaheuristic Optimization Approach

Parijata Majumdar[1], *Vishal Jain*[2]

[1] LUC Application Number: LUCPDF20251124004 Post Doctoral Research Fellow, Department of Computer Science and Engineering, Lincoln University College, Petaling Jaya, Malaysia ORCID ID: 0000-0002-2324-7685;

[2] Department of Computer Science and Engineering, School of Engineering & Technology, Vivekananda Institute of Professional Studies-Technical Campus, New Delhi, India

Email ID pdf.Parijata@lincoln.edu.my

Abstract: Crop recommendation systems examine variables like soil type, climate, and historical data to recommend appropriate crops for a given area with the aid of various machine learning and deep learning models. This study suggests a novel crop recommendation system that makes use of a Deep Convolutional Neural Network (DCNN). The Modified Salp Swarm Algorithm (MSSA), which depends on a competition mechanism and variable shifted windows, is used to optimize the DCNN's weight update. By adding a competitive mechanism that dynamically ranks salps depending on their fitness and permits stronger people to advise weaker ones, the MSSA improves the traditional Salp Swarm Algorithm and prevents early convergence. Concurrently, the variable shifting window technique improves exploration in the early stages and exploitation near convergence by adaptively modifying the search bounds of salps during iterations. Across several benchmark datasets, results from experiments verify that DCNN–MSSA performs noticeably better than traditional optimizers when measured in terms of performance metrics.

Keywords: Crop Recommendation System, Deep Convolutional Neural Network, Modified Salp Swarm Algorithm

Introduction

Efficient and sustainable agriculture practices are becoming increasingly important because of the population's exponential growth and the resulting increase in food consumption [1].

Latest advances in deep learning (DL) and machine learning (ML) have transformed agricultural operations by facilitating data-driven, intelligent decision-making. By analyzing large datasets that include soil nutrients, temperature, rainfall, and humidity, these methods have been successfully incorporated into precision agriculture to provide predictive insights that enhance resource use and productivity [2]. Because of their ability to capture complicated spatial and non-linear correlations, Deep Convolutional Neural Networks (DCNNs) have demonstrated remarkable promise in handling diverse agricultural data [3]. However, how effectively the parameters of DL models, such as weights and bias, are optimized during training has a great impact on how well they perform. Convergence rate, local minima, and sensitivity to hyperparameters are some of the common limitations of conventional optimization methods such as Stochastic Gradient Descent or Adam [4]. To optimize global searches on high-dimensional problems more effectively, researchers have turned to nature-inspired metaheuristic methods.

The Salp Swarm Algorithm (SSA) is a metaheuristic algorithm inspired by nature [5]. SSA is inspired by the swarm foraging behavior of salps in the ocean. SSA simulates how salps swarm and move in a chain pattern, with follower salps moving their positions relative to each other and the leader salp guiding the swarm to food sources. However, there are a number of intrinsic issues with SSA. One of these is that SSA has a high tendency to converge too quickly because of the high impact of the leader salp on swarm diversity. Secondly, rapid deterioration in the control coefficient related to movement causes SSA's ability to conduct an exploration process to degrade over time, thus failing to adequately cover the search space. Thirdly, SSA lacks sufficient ability in exploitation to adequately refine optimal zones and is characterized by delayed convergence to the global optimum [6].

To adequately enhance search efficiency and achieve a good trade-off between exploration and exploitation, Zhang et al. [7] proposed a Modified Salp Swarm Algorithm (MSSA) that incorporates a competition mechanism and a shifting window. With the competition mechanism in MSSA, individual salps can be ranked in terms of their fitness values. This adaptive leader-follower interaction prevents stagnation in local optima and maintains swarm diversity. In parallel to this, the variable shifting window method optimally changes the search space over time by gradually reducing the search space in later stages and increasing it in earlier stages to explore all potential areas for location identification.

This research has been inspired by the optimized results achieved in MSSA for location identification and has proposed a smart crop recommendation system based on DCNN and MSSA (DCNN–MSSA). The proposed framework optimally increases the prediction accuracy and convergence rate of the DCNN model by employing MSSA to optimize the weight parameters of the DCNN model. The agro-environmental characteristics are analyzed based on three publicly available Kaggle datasets, namely soil nutrients (N, P, K), pH value, temperature, humidity, rainfall, and precipitation. To guarantee high-quality inputs, extensive preprocessing is carried out, comprising data cleansing, normalization, and duplicate elimination.

In terms of performance metrics across several benchmark datasets, experimental study shows that the DCNN–MSSA model consistently performs better than traditional ML, DL, and various other metaheuristic-based optimization models. The proposed model offers precise, data-driven, and sustainable crop recommendations by fusing the computational power of DL with the adaptive optimization strength of the modified salp swarm algorithm. This enables farmers to take well-informed cultivation choices that are specific according to local soil and surroundings.

The proposed methodology for the development of the crop recommendation system using the DCNN-MSSA technique is provided in detail in Section 2. The experimental design, performance evaluation criteria, and comparison of the suggested model with current approaches using actual agricultural datasets are covered in Section 3. The study's main conclusions are summarized in Section 4.

2. Proposed Methodology

The proposed Deep Convolutional Neural Network with Modified Salp Swarm Algorithm (DCNN–MSSA) model for crop recommendation is presented in this section. The overall methodology consists of dataset integration, data preprocessing, model construction, and Modified Salp Swarm Algorithm (MSSA) optimization to improve prediction accuracy.

2.1 Dataset Description

To enhance the generalization ability of the proposed model, three publicly available crop recommendation datasets were used.

Dataset 1: Crop Recommendation Dataset (2021) by Siddharth S.S., available at https://www.kaggle.com/datasets/siddharthss/crop-recommendation-dataset [8].

Dataset 2: Crop Recommendation Dataset (2020) by Atharva Ingle, available at https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset [9].

Dataset 3: Crop Recommendation Dataset (2024) by Varshita Nalluri, available at https://www.kaggle.com/datasets/varshitanalluri/crop-recommendation-dataset [10].

Each dataset contains agricultural attributes describing different crop classes, including:

- Temperature
- Humidity
- Soil pH
- Rainfall
- Nitrogen (N)
- Phosphorus (P)
- Potassium (K)

After merging the datasets, data preprocessing was performed which included:

- Removal of missing values
- Elimination of duplicate records
- Data normalization

The final dataset was divided into training and testing sets, where 80% of the data was used for training and 20% for testing.

2.2 Deep Convolutional Neural Network (DCNN)

The proposed model uses a Deep Convolutional Neural Network (DCNN) architecture for crop classification. The architecture consists of:

- Three Convolutional Layers
- Two Max-Pooling Layers
- One Batch Normalization Layer
- One Fully Connected Layer
- One Softmax Output Layer

The DCNN extracts spatial and feature-level relationships among environmental and soil parameters to recommend suitable crops.

The model input includes agricultural parameters such as Temperature, Humidity, Soil pH, Soil moisture, Nitrogen, Phosphorus, Potassium and Rainfall.

The first convolutional layer ($Conv_1$) filters the input data and extracts primary feature maps representing relationships among crop growth variables.

The convolution operation is defined as:

$$F = \sigma(W * X + b)$$

where

- X= input data
- W= weight matrix
- b= bias term
- $\sigma(\cdot)$= activation function (ReLU)

The extracted feature maps are then processed through max-pooling layers, which reduce dimensionality while preserving important information.

Further convolution layers ($Conv_2$ and $Conv_3$) refine the extracted features and help identify complex patterns associated with crop growth conditions.

The feature maps generated from convolution and pooling layers are flattened and passed to a Fully Connected (FC) layer, which performs high-level feature integration and classification.

The final Softmax layer generates probability values for each crop class:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}}$$

where

- p_i= probability of the i^{th}crop class
- C= total number of crop classes

The model is trained by minimizing the cross-entropy loss function:

$$L = -\sum_{i=1}^{C} y_i \log(p_i)$$

where

- y_i= true label of the i^{th}class

The DCNN parameters are optimized using the Modified Salp Swarm Algorithm (MSSA) to minimize the loss function:

$$\Theta^* = \arg\min_{\Theta} L(\Theta)$$

Where Θrepresents the set of network parameters (weights and biases).

2.3 Modified Salp Swarm Algorithm (MSSA) for DCNN Training

The Modified Salp Swarm Algorithm (MSSA) is an improved version of the standard Salp Swarm Algorithm designed to prevent premature convergence, maintain a balance between exploration and exploitation and accelerate optimization convergence.

Inspired by Zhang et al. [7], the proposed MSSA incorporates several improvements.

Key Enhancements

Competition Mechanism

Leader salps compete based on fitness values to maintain population diversity. Strong followers can replace poorly performing leaders, which accelerates convergence toward optimal solutions.

Variable Shifted Windows

The number of dimensions updated during each iteration is dynamically adjusted to balance exploration and exploitation.

Adaptive Generalized Opposition-Based Learning (AGOBL)

Opposition-based candidate solutions are periodically generated to increase the probability of escaping local optima.

Fitness Function

The fitness function used for optimization is defined as:

$$\eta = \frac{1}{r}\sum_{\delta=1}^{r}(N_r^* - O)^2$$

where

- r= number of training samples
- N_r^*= target output
- O= DCNN predicted output

Algorithm 1: MSSA for DCNN Weight Optimization

```
Input: Training dataset
Output: Optimized DCNN weights
    1. Initialize population of salps $X_i$ where $i = 1,2,...,N$ in a $d$-dimensional search space.
    2. Evaluate fitness $f(X_i)$ for each salp using the DCNN loss function.
    3. Identify leader and follower salps based on fitness ranking.
While termination criterion is not satisfied
    4. Update leader position:
       $X_1^{t+1} = X_1^t + c_1(F_{best} - X_1^t) + c_2 \times rand \times (X_{rand} - X_1^t)$
    5. Apply competition mechanism.
       If $f(X_j) < f(X_1)$, exchange leader and follower roles.
    6. Update follower positions using variable shifted window:
       $X_i^{t+1} = \frac{X_i^t + X_{i-1}^{t+1}}{2} + \lambda \times rand \times Shift(w)$
where Shift(w) controls the number of dimensions updated per iteration.
    7. Apply AGOBL-based opposition learning for selected salps.
    8. Recalculate fitness and update the best solution $X_{best}$.
End While
    9. Return optimized DCNN weights $X_{best}$.
```

3. Experimental Results and Analysis

The performance of the proposed DCNN–MSSA model was evaluated using Python 3.6.5 on a system with the following configuration:

- Processor: Intel i5-8600K
- GPU: NVIDIA GeForce 1050Ti (4 GB)
- RAM: 16 GB
- Storage: 250 GB SSD and 1 TB HDD

Three publicly available crop recommendation datasets were used in the experiments:Crop Recommendation Dataset (2021), Crop Recommendation Dataset (2020), and Crop Recommendation Dataset (2024).

The proposed DCNN–MSSA model was compared with several existing crop recommendation frameworks, including XAI–CROP [11], IDCSO–WLSTM [12], RFOERNN [13], MMML [14] and DCNN–SSA. The comparative analysis demonstrates that the DCNN–MSSA model performs effectively across different datasets and evaluation metrics. The performance results of DCNN–MSSA on the Crop Recommendation Datasets (2021), (2020), and (2024) are presented in Tables 1, 2, and 3, respectively. Here, MCC refers to the Matthews Correlation Coefficient.

Table 1 Results of DCNN–MSSA on Crop Recommendation Dataset (2021)

Performance Metric	Value (%)
Accuracy	96.12
Precision	96.28
Recall	96.24
Specificity	95.81
PR-Score	96.05
ROC-Score	96.34
F1-Score	96.21
MCC	95.19

Table 2 Results of DCNN–MSSA on Crop Recommendation Dataset (2020)

Performance Metric	Value (%)
Accuracy	98.84
Precision	98.91
Recall	98.88
Specificity	98.42
PR-Score	98.72
ROC-Score	98.95
F1-Score	98.79
MCC	97.82

Table 3 Results of DCNN–MSSA on Crop Recommendation Dataset (2024)

Performance Metric	Value (%)

Accuracy	99.73
Precision	99.71
Recall	99.75
Specificity	99.34
PR-Score	99.46
ROC-Score	99.86
F1-Score	99.53
MCC	98.72

Table 4 summarizes the comparative performance of DCNN–MSSA and existing algorithms across the Crop Recommendation datasets (2020, 2021, and 2024). The results clearly indicate that the DCNN–MSSA model achieves higher prediction accuracy and lower execution time compared with other methods. Here, Acc. denotes Accuracy.

Table 4 Comparison of DCNN–MSSA and Existing Algorithms Across Crop Recommendation Datasets

Model	**Acc (2020)**	**F1 (2020)**	**R^2 (2020)**	**Time (s)**	**Acc (2021)**	**F1 (2021)**	**R^2 (2021)**	**Time (s)**	**Acc (2024)**	**F1 (2024)**	**R^2 (2024)**	**Time (s)**
XAI–CROP	94.87	94.56	93.92	240.28	95.12	95.86	94.08	240.05	94.65	94.23	93.81	239.72
IDCSO–WLSTM	91.94	91.68	91.33	241.48	92.68	91.79	92.39	241.04	91.57	91.40	91.10	240.98
RFO ERNN	98.22	98.18	98.79	240.13	98.45	98.46	99.88	240.04	98.38	98.25	98.67	239.84
MMML	97.46	97.32	97.89	242.23	97.91	97.92	98.39	242.05	97.60	97.52	98.04	241.91
DCNN–SSA	95.43	95.25	95.81	242.34	91.52	91.46	91.54	242.07	91.10	91.02	90.83	241.76
DCNN–MSSA	**98.84**	**98.79**	**98.90**	**238.76**	**96.12**	**96.21**	**96.24**	**238.97**	**99.73**	**99.53**	**99.86**	**238.54**

The DCNN–MSSA framework consistently outperforms the benchmark models across all three datasets (2020, 2021, and 2024). The model achieves the best performance on the 2024 dataset, obtaining 99.73% accuracy and 99.86% R^2 score, which are the highest values among all evaluated methods.

Compared with existing hybrid and deep learning models, the proposed DCNN–MSSA demonstrates improved convergence speed, higher prediction accuracy, and reduced computational time. These improvements confirm the effectiveness of the dynamic exploration–exploitation mechanism of MSSA, which enhances optimization during DCNN training.

Therefore, the proposed method provides a scalable, robust, and efficient solution for intelligent crop recommendation systems.

3.1 Statistical Analysis

To verify the statistical significance of the improvements, statistical tests were conducted: Paired t-test At a 95% confidence level, the paired t-test produced low p-values ($p < 0.05$) across all comparisons. This indicates that the improvements achieved by the DCNN–MSSA model in terms of Accuracy, F1-score, and R^2-score are statistically significant. Furthermore, the model exhibited low variance in accuracy ($\sigma^2 = 0.0042$), demonstrating the robustness and consistency of the proposed method. The results of the paired t-test comparing DCNN–MSSA with benchmark models are summarized in Table 5.

Table 5 Paired t-test Results Comparing DCNN–MSSA and Benchmark Models

Model Compared	Metric	t-Statistic	p-Value
DCNN–MSSA vs XAI–CROP	Accuracy	5.412	0.0028
DCNN–MSSA vs IDCSO–WLSTM	F1-Score	4.957	0.0043
DCNN–MSSA vs RFOERNN	R^2-Score	3.889	0.0076
DCNN–MSSA vs MMML	Accuracy	4.284	0.0058
DCNN–MSSA vs DCNN–SSA	**Accuracy**	**3.745**	**0.0091**

The obtained p-values (< 0.01) confirm that the superior performance of the DCNN–MSSA model is statistically significant. Therefore, the observed improvements are not due to random variation but are the result of the proposed optimization strategy.

4. Conclusion and Future Work

The novelty of the current research is the development of a sophisticated hybrid model named DCNN–MSSA, which is employed in crop prediction. The improved MSSA avoids stagnation, maintains an efficient balance of exploration and exploitation, and accelerates convergence to optimize the weight update of the DCNN model. The efficacy of the model in crop prediction has been proved by the 94\% training accuracy of the model, the reduction in the execution time of the model, and the consistent performance of the model on the benchmark datasets. To better comprehend the intricate non-linear relationships among the environmental and soil factors that affect crop prediction, the application of sophisticated deep learning models such as the encoder-decoder model, Vision Transformers, and Graph Neural Networks can be explored in the future. The model's trustworthiness is further supported by statistical validation, and its practical value for intelligent, data-driven agricultural decision-making is highlighted by its potential for real-world application.

References

1. Hemathilake D, Gunathilake D (2022) Agricultural productivity and food supply to meet increased demands. Future Foods pp 539–553

2. Bhoj J, Bharte G, Bhalerao C, et al (2023) Crop recommendation system using machine learning algorithms. International Research Journal of Modernization in Engineering Technology and Science 05(03)
3. Ullah F, Ullah I, Khan K, et al (2025) Advances in deep neural network-based hyperspectral image classification and feature learning with limited samples: a survey. Applied Intelligence 55(6):370
4. Barve T, Samant A, Kulaye M (2025) A comparative study of optimization algorithms in deep learning: Sgd, adam, and beyond. International Journal of Computer Technology and Electronics Communication 8(4):11018–11021
5. Mirjalili S, Gandomi A, Mirjalili S, et al (2017) Salp swarm algorithm: A bio-inspired optimizer for engineering design problems. Advances in Engineering Software 114:163–191
6. Abualigah L, Hawamdeh W, Zitar R, et al (2024) Salp swarm algorithm: survey, analysis, and new applications. In: Metaheuristic Optimization Algorithms. Morgan Kaufmann, p 241–258.
7. Zhang H, Qin X, Gao X, et al (2024) Modified salp swarm algorithm based on competition mechanism and variable shifted windows for feature selection. Soft Computing. https://doi.org/10.1007/s00500-024-09876-9
8. Siddharth S (2021) Crop recommendation dataset. Kaggle https://www.kaggle.com/datasets/siddharthss/crop-recommendation-dataset (Accessed: 2 November 2025)
9. Ingle A (2020) Crop recommendation dataset. Kaggle https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset (Accessed: 2 November 2025)
10. Nalluri V (2024) Crop recommendation dataset. Kaggle https://www.kaggle.com/datasets/varshitanalluri/crop-recommendation-dataset (Accessed: 2 November 2025)
11. Shams M, Gamel S, Talaat F (2024) Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making. Neural Computing and Applications 36(11):5695–5714
12. Kiruthika S, Karthika D (2023) Iot-based professional crop recommendation system using a weight-based long-term memory approach. Measurement: Sensors 27:100722
13. Gopi P, Karthikeyan M (2023b) Red fox optimization with ensemble recurrent neural network for crop recommendation and yield prediction model. Multimedia Tools and Applications pp 1–21.
14. Gopi P, Karthikeyan M (2023a) Multimodal machine learning based crop recommendation and yield prediction model. Intelligent Automation & Soft Computing 36(1).