



A Sparrow Search Algorithm – Driven Vision Transformer – Temporal Convolutional Network Hybrid Framework for Precision Smart Irrigation Prediction Using Environmental and Soil Condition Analysis

S. Prakash²

²Assistant professor, Department of Electrical and Electronics Engineering,
Aarupadai Veedu Institute of Technology Vinayaka Missions Research Foundation,
Chennai, India - 603104. prakash.yours@gmail.com

ABSTRACT: Smart irrigation prediction uses Deep Learning (DL) and weather data to forecast exact watering needs, helping farms save water, cut costs, soil condition and improve crop health. Therefore, in this paper, a hybrid Vision Transformer based Temporal Convolutional Network (ViT-TCN) with Sparrow Search Optimization (SSA) is proposed for smart irrigation prediction using environmental and soil condition analysis. Data visualization transforms complex sensor and prediction data into clear, interactive dashboards that help operators to monitor field health and optimize watering schedules in real time. The integration of a ViT-TCN framework effectively reduces the challenge of precision agriculture by processing both complex environmental imagery and continuous time-series data simultaneously. The SSA improves smart irrigation prediction by smart feature selector and hyperparameter tuner. The performance analysis is carried out using python software and assesses the model's performance using key metrics such as confusion matrix, and ROC-AUC. The proposed ViT-TCN-SSA uses three classes, they are low irrigation, medium irrigation and high irrigation. The proposed ViT-TCN-SSA achieves high accuracy of 98% and better convergence speed. The proposed method outperforms conventional techniques, contributing a robust and reliable tool for smart irrigation prediction using environmental and soil condition analysis.

Keywords: ViT-TCN-SSA, smart irrigation prediction, environmental and soil condition analysis, and accuracy.

1. Introduction

The agricultural sector uses over 70% of fresh water due to population increase and the requirement to produce food [1]. Irrigation systems lose a significant amount of energy and water. Due to evaporation, inadequate irrigation systems, and poor water management, 40% [2,3] of water is wasted [4]. A sustainable ecosystem can be achieved through the interdependent water, energy, and food (WEF) strategy [5]. In particular, an efficient irrigation system maximizes water use and lowers energy use. Long-term environmental,

economic, and social objectives can be met through WEF ecosystem nexuses [6]. Water usage is decreased and irrigation systems are able to use water in an efficient and sustainable manner [7]. A significant amount of water is also needed for plants as a result of climate change, decreased rainfall, and global shifts [8]. In today's agricultural civilization, research on need analysis and plant water supply management is crucial. An irrigation system is used to control plant water requirements based on changing environmental conditions. In order to boost production yields, various water-optimization approaches and

efficient water-management systems are employed to lower water consumption and create an efficient water-requirements prediction system for plants. Society needs efficient irrigation and scheduling systems since they boost output and use less water. The DL is essential for smart irrigation prediction because it automatically extracts complex, nuanced patterns from diverse data types without manual feature engineering. The conventional techniques such as Convolutional ShuffleNetV2 (Conv-ShuffleNetV2) [9], YOLO [10] and Convolutional Long Short-Term Memory with Vision Transformer (ConvLSTM-ViT) [11] have been used for smart irrigation prediction using environmental and soil condition analysis. These techniques have some disadvantages like low convergence speed, less reliability and High response time. To avoid these issues, a ViT-TCN-SSA is proposed for smart irrigation prediction using environmental and soil condition analysis. The contribution of the paper is as follows,

- Data visualization is used to transform complex numerical metrics and categorical model predictions into intuitive visual formats.
- ViT-TCN is proposed to achieve highly accurate, interpretable, and computationally efficient DL framework capable of differentiating multiple levels of irrigation level (such as low, medium and high) from agricultural data.
- SSA is proposed to optimize complex model hyperparameters and to maximize classification accuracy while minimizing computational complexity by identifying the most soil textures, temperature, or humidity based features.

2. Recent Works

Smart irrigation prediction using Conv-ShuffleNetV2 networks have been presented in [9] to analyse irrigation system. Sensors placed at different observation sites can be used to measure

the soil moisture in the crop field, indicating the amount of water retained. Unfortunately, communication problems or sensor failure prevent the smart farming system from receiving the soil moisture data supplied by the irrigation management. The Internet of Things (IoT) technologies can support creative farming methods. [10] suggests YOLO-Berry, a lightweight and effective object detection model, for identifying the growing stage of blueberries in real time. First, created an extensive field-image collection of blueberries that included several stages of growth under various environmental circumstances. By adding a Bidirectional Feature Pyramid Network (BiFPN) for better multi-scale feature fusion, a Convolutional Block Attention Module (CBAM) to concentrate on crucial areas, and lightweight Ghost and C3Ghost structures to reduce computational costs, the suggested model improves the YOLO11 baseline. [11] presents a method for predicting soybean yield by combining vision transformer (ViT), three-dimensional convolutional neural network (3D-CNN), and convolutional long short-term memory (ConvLSTM). Our model uses multispectral remote sensing data to capture intricate patterns in agricultural datasets by utilizing the global context analysis of ViTs, the temporal sequencing capabilities of ConvLSTM, and the spatial hierarchy of 3D-CNNs. By combining these cutting-edge techniques, a thorough examination of the temporal and spatial dimensions of crop growth is made possible, leading to more precise and reliable forecasts. However, it produces high computational costs, making real-time deployment difficult and model optimization complex. To overcome these drawbacks, a ViT-TCN-SSA is proposed in this paper smart irrigation prediction using environmental and soil condition analysis.

3. Proposed Description

Smart irrigation prediction using environmental and soil condition analysis optimizes water use by

delivering precise amounts of water based on real-time data like soil moisture, temperature, and humidity. The proposed block diagram of ViT-TCN-SSA for smart irrigation prediction using environmental and soil condition analysis is represented in figure 1. The process starts with an input dataset that works thorough preprocessing. This block includes data preparation, a consistency check, missing value imputation to handle gaps in data, and outlier suppression to

filter out anomalous noise. In feature normalization & visualization, the data passes through a data visualization phase and enters a feature normalization block. Here, Min-Max scaling is performed with a feature correlation analysis to refine and select key features. The SSA acts as a tuner, interacting directly with the feature block and the classification model to optimize parameters or feature selection.

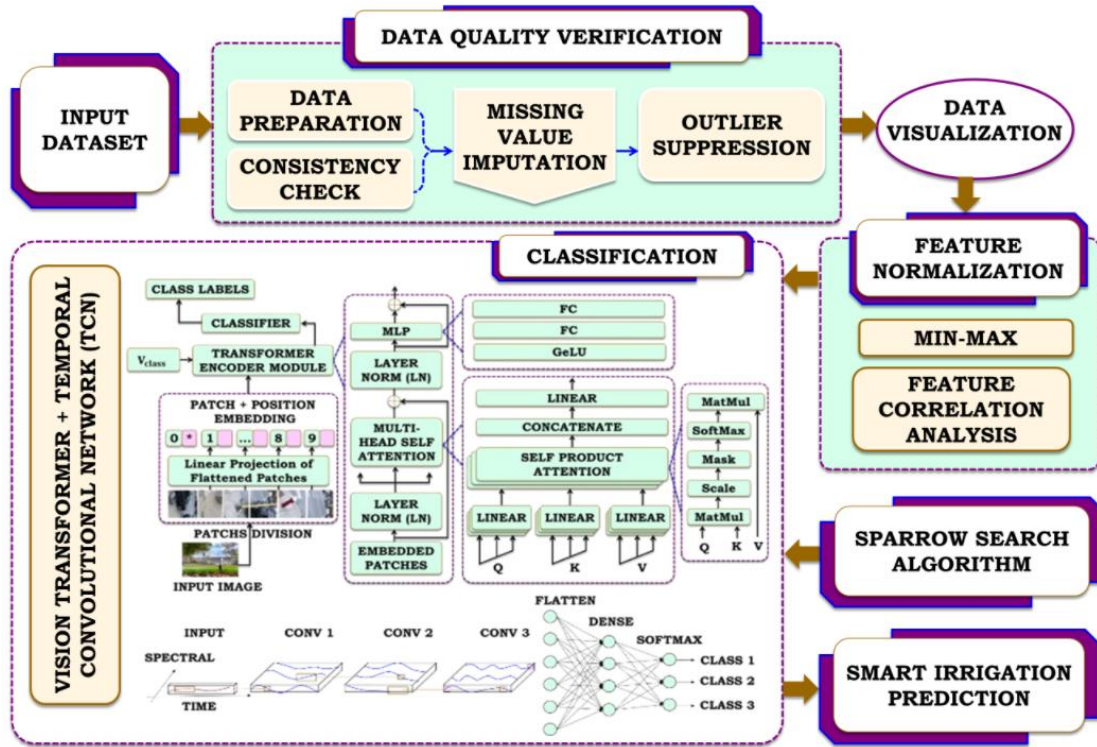


Figure 1: Proposed block diagram of ViT-TCN-SSA for smart irrigation prediction

The ViT-TCN classification processes the refined features through a hybrid neural network split into two main streams. The ViT takes an input image, divides it into patches, applies a linear projection for patch and position embedding, and passes it through standard Transformer components to decode spatial relationships and TCN processes spectral and time-series multi-dimensional data through successive 1D/2D convolutional layers, a flattening layer, a dense network, and a Softmax classifier. The final combined output delivers the smart irrigation prediction, categorizing the target conditions into discrete classification outcomes.

3.1. Modelling of Data Visualization

The process of turning raw numbers into clear data like charts and graphs is named as data visualization.

Mathematical expression for mean (average) to check data center:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

Mathematical expression for min-max feature scaling:

$$x_{normalized} = \frac{x_i - \min(X_{train})}{\max(X_{train}) - \min(X_{train})} \quad (2)$$

Crucially, the min and max values are calculated only from the training data (X_{train}). The same parameters are then used to transform the test data. A ViT-TCN is used after initial visualization or feature extraction to dynamically capture global context and complex long-range dependencies. It refines visual or spatial data by adjusting to varying shapes, noise, or multi-scale structures.

3.2. Mathematical modelling of ViT-TCN

A Hybrid ViT-TCN system for smart irrigation blends spatial visual analysis (such as drone/satellite canopy images or leaf health data) with multivariate long-term time-series tracking (such as IoT environmental logs. Figure 2 shows the architecture of ViT-TCN for smart irrigation prediction.

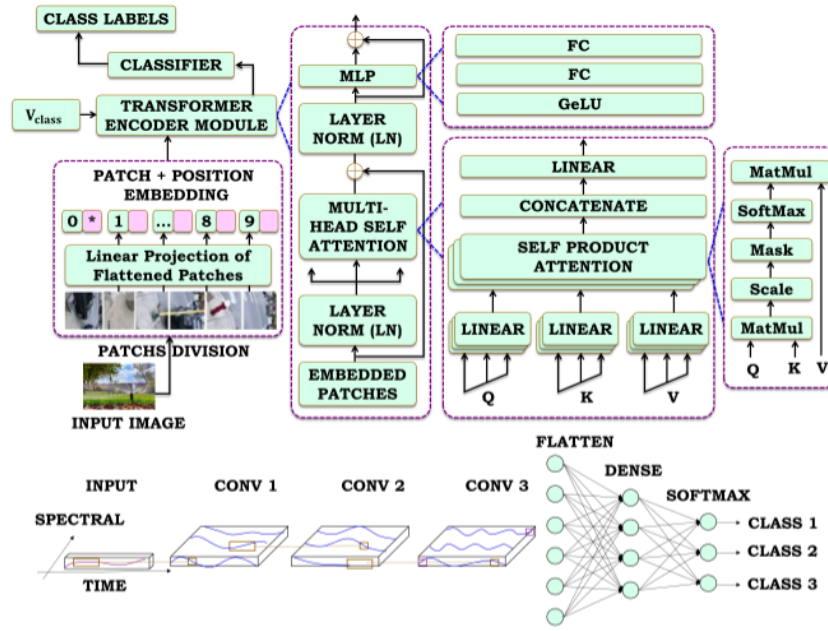


Figure 2: Architecture of ViT-TCN

3.2.1 Input Data Definition and Shapes

Initialize temporal environmental feature matrix (X) such as multivariate records from continuous IoT ground sensors and features such as soil moisture (%), ambient temperature (°C), relative humidity (%), and solar radiation (W/m²).

3.2.2 Vision Transformer

The ViT stream models spatial context, structural soil variations, and canopy stress distributions.

Patch Extraction: The image I is divided into a grid of non-overlapping patches of size.

$$\text{Number of patches} = \frac{H \times W}{p^2} \quad (3)$$

Linear Projection and Class Token: Flattened patches are mapped to an embedding dimension $D = 768$ using a learnable projection matrix. A

learnable [CLS] token is prepended to capture global spatial features.

Positional Encoding: 1D learnable positional embeddings are added to retain coordinates:

$$E_{spatial} = [X_{cls}; patch_1E; patch_2E, \dots, patch_NE] + E_{pos} \quad (4)$$

Multi-Head Self-Attention (MHSA): Fed through Multi-Head Attention blocks. The system calculates Scaled Dot-Product Attention:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

Output Vector: The extracted [CLS] token yields a pure, high-level spatial representation vector:

$$Z_{spatial} \in R^{32 \times 768} \quad (6)$$

3.2.3 The Temporal Convolutional Network

The TCN handles continuous multivariate sensor logs, capturing local and multi-scale multi-day trends without sequence leakage.

Dilated Causal Convolutions: Standard convolutions suffer from narrow receptive fields. The TCN uses 1D causal convolutions with expanding dilation rates ($d = 1, 2, 4, 8$) to exponentially look back into historical cycles. The operation for a filter f on sequence elements is:

$$y(t) = \sum_{i=1}^{k-1} f(i) \cdot X_{t-d \cdot i} \quad (7)$$

Residual Blocks: To combat vanishing gradients across deep temporal lookbacks, the outputs are passed through absolute residual mappings:

$$o = \text{Activation}(x + \text{TCN}(x)) \quad (8)$$

Output Vector: After progressive layers, the sequence is pooled across the timeline to compress temporal patterns:

$$Z_{\text{temporal}} \in R^{32 \times 256} \quad (9)$$

3.2.4 Multi-Modal Feature Fusion & Regression Head

The distinct properties from the earth's surface (spatial) and atmospheric conditions (temporal) are bound together.

Concatenation: The vectors are combined into a joint feature space vector.

$$Z_{\text{fused}} = [Z_{\text{spatial}} || Z_{\text{temporal}}] \quad (10)$$

Dense Layers: The vector passes through fully connected layers with Dropout ($p=0.3$) and ReLU activations:

$$y = \text{dense}\left(\text{ReLU}\left(\text{dense}(Z_{\text{fused}})\right)\right) \quad (11)$$

The hybrid ViT-TCN allows the system to analyze the entire field simultaneously, precisely isolating crop water stress, soil patch variations, and total canopy context while resisting occlusion or noise. Optimization is required after classification to adjust model weights and minimize error and for tuning hyperparameters.

3.3. Modelling of SSA

The SSA is a swarm intelligence technique inspired by the foraging and anti-predation behaviors of sparrows, divided into producers (discoverers), scroungers (followers), and aware/warning sparrows (scouts).

Here is the step-by-step process of using COA for hyperparameter tuning, complete with the mathematical equations that drive each phase. Flow chart of SSA for smart irrigation prediction is represented in figure 3.

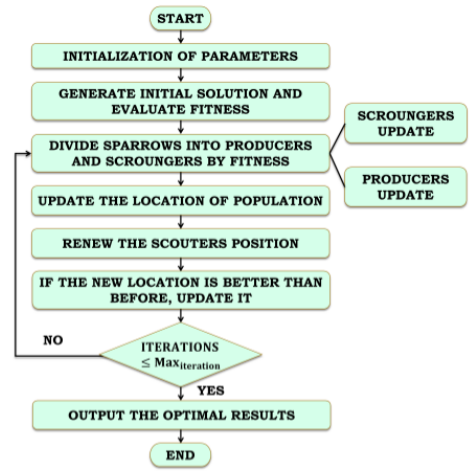


Figure 3: Flowchart of SSA

3.3.1 Initialize the Population and Parameters

Initialize a population of n sparrows randomly in a d -dimensional search space.

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,d} \\ x_{2,1} & x_{2,2} & x_{2,d} \\ x_{n,1} & x_{n,2} & x_{n,d} \end{bmatrix} \quad (12)$$

Evaluate the fitness value $f(X_i)$ for each sparrow i .

3.3.2 Evaluate Fitness

Evaluate all sparrows, rank them based on fitness, and identify the current best producer X_g , global best) and worst position (X_w).

3.3.3 Update Producer Positions

Producers have high fitness values and search a wide range to guide the population. Their positions are updated using:

$$X_{i,j}^{t+1} = X_{i,j}^t \cdot \exp\left(-\frac{i}{\alpha \cdot T}\right) \quad (13)$$

Where, t is the current iteration, i is the maximum number of iterations, and α is a constant parameter to control food abundance.

3.3.4 Update Scrounger Positions

Scroungers follow producers with better food sources to improve their own energy. If a scrounger has low fitness:

$$p(t) = C_3 \times \left(1 - \frac{t}{T_{Max}}\right) \quad (14)$$

3.3.5 Update Warning/Aware Sparrow Positions (Scouts)

A percentage (10% to 20%) of the population is chosen randomly to detect danger and sound an alarm.

$$X_{i,j}^{t+1} = X_{best}^t + \beta \cdot |X_{i,j}^t - X_{best}^t| \quad (15)$$

3.3.6 Evaluate and Iterate

Calculate the new fitness values for all updated positions, compare them with previous iterations, and update the global best X_{best} if a superior position is found. Increment $t = t + 1$ and repeat from Step 3.3.2 until $t = T$.

The ViT-TCN-SSA models are accurately classifying smart irrigation prediction by analysing temperature, humidity and intensity. Reduces water consumption by 30% to 70% by eliminating runoff, deep percolation, and unnecessary watering based on real-time soil moisture thresholds.

4. Results and Discussion

The ViT-TCN-SSA is proposed for smart irrigation prediction, which is evaluated using python software. The name of dataset used for simulation is smart agricultural dataset and it includes three classes, such as low irrigation, medium irrigation and high irrigation.

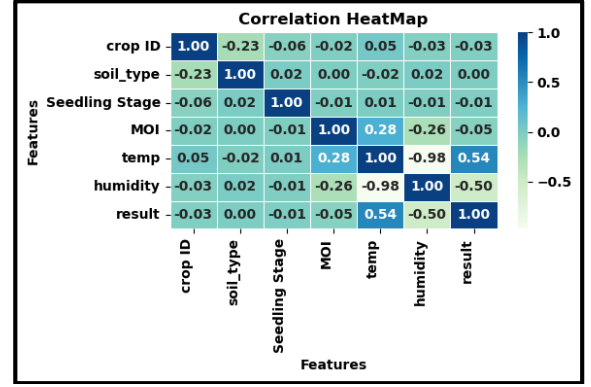


Figure 4: Correlation Heatmap

Figure 4 displays a correlation heatmap, which shows its strongest positive correlation with temperature of 0.54 and its strongest negative correlation with humidity (-0.50). Conversely, variables like crop ID (-0.03), soil_type (0.00), and Seedling Stage (-0.01) exhibit near-zero linear correlation with the result. A critical multi-collinearity pattern is evident between the predictors, where temp and humidity share an almost perfect negative correlation of -0.98.

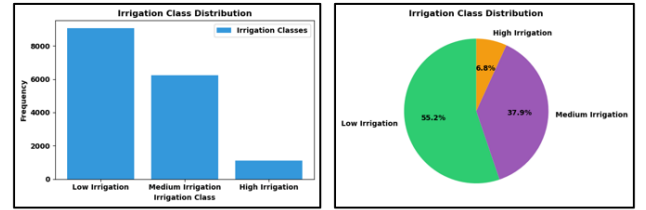


Figure 5: Irrigation class distribution

Figure 5 shows the irrigation class distribution. The low irrigation is the most frequent class with over 9,000 occurrences, followed by Medium Irrigation at approximately 6,200, and High Irrigation as the least frequent at around 1,100. The percentage of low irrigation accounts for the majority share at 55.2%, medium irrigation shows the 37.9%, and high irrigation represents the smallest portion at just 6.8%.

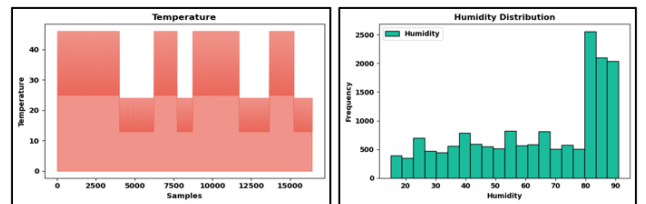


Figure 6: Temperature and humidity distribution

Figure 6 shows the temperature and humidity distribution. The temperature shows that the temperature across 16,500 samples, showing distinct rectangular phases shifting between values of 13, 25, and 45. The humidity illustrates a massive frequency spike occurs at the higher end between 80 and 90 humidity, soaring to a maximum peak of over 2,500 occurrences.

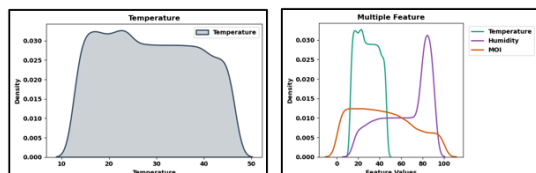


Figure 7: Temperature and multi feature curve

Figure 7 shows the temperature and humidity curve, the individual temperature shows from 10 to 48, reaching its peak density of about 0.033 between values 15 and 25. In the multi-feature curve, the temperature curve shows this same distribution profile. The humidity curve ranges from 10 to 100 and shows a highly concentrated, sharp peak of 0.031 at a feature value around 85.

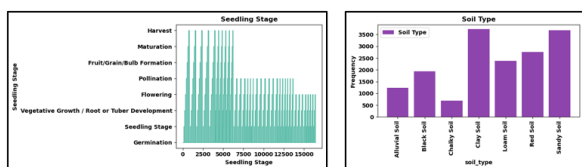


Figure 8: Seeding stage and soil type

Figure 8 shows the seeding stage and soil type. The harvest phase is prominently reached in the first 6,500 samples before the cycle levels off into pollination and flowering stages. The soil type highlights the frequency of seven distinct soil types, where Clay Soil and Sandy Soil share the highest distribution at approximately 3,700 counts each.

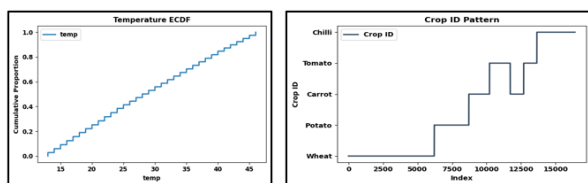


Figure 9: Temperature ECDF and crop ID pattern

The temperature ECDF and crop ID pattern are shown in figure 9. The temperature ECDF shows the cumulative proportion rises uniformly from 0.0 to 1.0 across a temperature range of approximately 13 to 46. The crop ID pattern illustrates a five crop types are distributed across a sequential index spanning from 0 to over 15,000. Wheat dominates the first part of the dataset from index 0 to 6,000, followed by steps representing Potato, Carrot, and Tomato. Finally, Chilli occupies the highest indices of the dataset, starting from around index 13,500 up to the maximum index value.

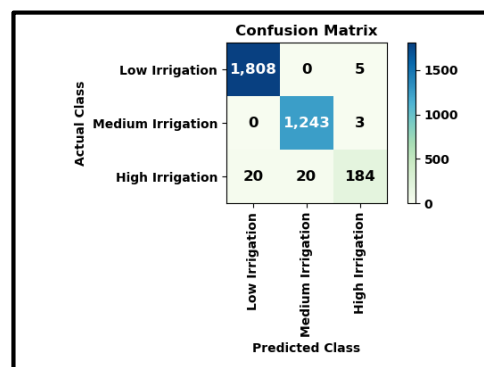


Figure 10: Confusion matrix

Figure 10 shows the confusion matrix. The model demonstrates high performance by correctly identifying 1,808 instances of low irrigation and 1,243 instances of medium irrigation. Misclassifications remain extremely low for these two classes, with only 5 low and 3 medium instances incorrectly predicted as high irrigation. However, the High Irrigation class shows slightly lower accuracy, with 184 instances correctly predicted while 20 cases each were misclassified as Low and Medium.

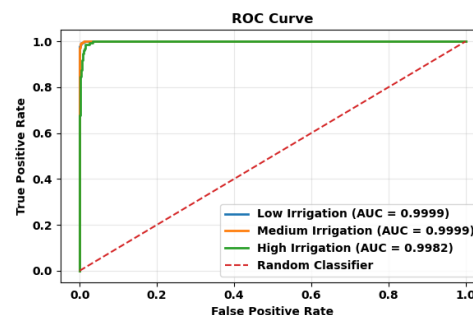


Figure 11: ROC Curve

Figure 11 shows the ROC Curve of proposed ViT-TCN-SSA. The AUC values of class low irrigation is 0.9999, medium irrigation is 0.9999, and high irrigation is 0.9982. The AUC values shows that the effectiveness of proposed ViT-TCN-SSA in smart irrigation prediction using environmental and soil condition analysis.

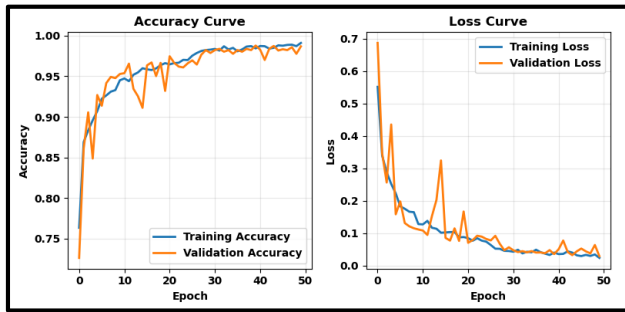


Figure 12: Accuracy and loss of proposed ViT-TCN-SSA

The accuracy and loss of proposed ViT-TCN-SSA are represented in figure 12. The accuracy of proposed ViT-TCN-SSA is 98% and the loss of proposed ViT-TCN-SSA is 10%, which demonstrates the effectiveness of proposed system.

Table 3: Comparison of response time and accuracy

Method	Response time	Accuracy
Conv-ShuffleNetV2 [8]	0.7s	97%
YOLO [9]	0.6s	97.34%
ConvLSTM-ViT [10]	0.5s	96%
Proposed ViT-TCN-SSA	0.1s	98%

Table 3 shows the comparison of response time and accuracy. The response time of proposed ViT-TCN-SSA is 0.1s. Lower response time in a ViT-TCN-SSA classification model enables real-time edge deployment, reduced compute costs, and instant decision-making. The accuracy of proposed ViT-TCN-SSA is 98%. High accuracy in a ViT-TCN-SSA for smart irrigation prediction ensures precise predictions of irrigation level.

5. Conclusion

In this paper, a hybrid ViT-TCN-SSA has been proposed for smart irrigation prediction using environmental and soil condition analysis. The ViT-TCN-SSA are effective at classifying irrigation level, utilizing optimized feature sets allows for efficient, real-time detection without requiring massive, computationally expensive models. The developed method is compared with existing methods Conv-ShuffleNetV2, YOLO and ConvLSTM-ViT in terms of precision, recall and F1-Score. The performance analysis is carried out using python software. The DL models, ViT-TCN-SSA offer high accuracy of 98%, and less computational time of 0.1s in smart irrigation prediction. In future, this work will be implemented using advanced AI technologies and hybrid optimization techniques.

References

1. Ayesha Khaliq; Atif Khan; Salman Jan; Muhammad Umair; Asad Gulshair; Ahmad Ali, Usman Ali Shah, Year: 2025, "AI-Driven Smart Agriculture: An Integrated Approach for Soil Analysis, Irrigation, and Crop-Fertilizer Recommendations", IEEE Access, Vol: 13.
2. Priya Saha; Ditipriya Sinha, Year: 2025, "LightAgro: Lightweight Blockchain IoT Based Fabrication for Smart Irrigation", Journal of Mobile Multimedia, Vol: 2025, No: 21.
3. Ravi Kumar Munaganuri; Yamarthi Narasimha Rao, Year: 2024, "PAMICRM: Improving Precision Agriculture Through Multimodal Image Analysis for Crop Water Requirement Estimation Using Multidomain Remote Sensing Data Samples", IEEE Access, Vol. 12.
4. Lavinia Moscovini; Luciano Ortenzi; Simona Violino; Catello Pane; Simone Figorilli; Corrado Costa; Federico Pallottino, Year: 2025, "Vegetation Indices Prediction Using Machine-

- Learning and Deep-Learning Approaches From RGB Calibrated Images”, IEEE Access, Vol: 13, pp. 203216-203229.
5. V Thamilarasi; R. Harihara Krishnan; V. Vijayalakshmi, Year: 2025, “Artificial Intelligence Driven Kyphosis”, In Proceedings of Fourth International Conference on Computing and Communication Networks: ICCCN 2024, Vol: 4, pp. 177.
 6. V Thamilarasi, Year: 2026, “Neuro-Symbolic AI with Edge Computing and Reinforcement Learning Optimizing Autonomous Engineering Design Workflows”.
 7. Kulandaivel Maruthamuthu Paramasivam; Prabakaran MP; Jeya Bright Pankiraj, Year: 2025, “CycleGAN algorithm based enhancement of night vision halation images from multiple sources”, AIP Advances, Vol: 15, pp. 075031-1 – 075031.
 8. D Madhivadhani; S. Hema; Jeya Bright Pankiraj; M. Jagadeesh Kumar; V. Kavitha; C. Sankari, Year: 2025, “Enhanced Skin Cancer Detection Using a Hybrid Deep Learning and Particle Swarm Optimization Framework”, In Proceedings of 2025 Global Conference in Emerging Technology (GINOTECH), pp. 1-6.
 9. Xuehui Xu; R. S. M. Lakshmi Patibandla; Amit Arora; Muna Al-Razgan; Emad Mahrous Awwad; Vincent Omollo Nyangaresi, Year: 2024, “An Adaptive Hybrid (1D-2D) Convolution-Based ShuffleNetV2 Mechanism for Irrigation Levels Prediction in Agricultural Fields With Smart IoTs”, IEEE Access, Vol: 12, pp. 71901-71918.
 10. Yi Luo; Junxi Bian; Yun Yu, Year: 2026, “An Improved Lightweight YOLO Model for Real-Time Blueberry Growth Stage Detection to Facilitate Smart Irrigation”, IEEE Access, Vol: 14.
 11. Seyed Mahdi Mirhoseini Nejad; Dariush Abbasi-Moghadam; Alireza Sharifi, Year: 2024, “ConvLSTM–ViT: A Deep Neural Network for Crop Yield Prediction Using Earth Observations and Remotely Sensed Data”, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 17, pp. 17489-17502.