CNN-LSTM framework to automatically detect anomalies in farmland using aerial images from UAVs

Omprakash Dewangan*, and *Dr. Priya* Vij

Faculty of CS & IT, Kalinga University, Naya Raipur, Chhattisgarh, India

Abstract. Using aerial inspection techniques in farmlands can yield vital data instrumental in mitigating various impediments to optimizing farming practices. Farmland anomalies (standing water and clusters of weeds) can impede farming practices, leading to the improper utilization of farmland and the disruption of agricultural development. Utilizing Unmanned Aerial Vehicles (UAVs) for remote sensing is a highly effective method for obtaining extensive imagery of farmland. Visual data analytics in the context of automatic pattern recognition from collected data is valuable for advancing Deep Learning (DL) -assisted farming models. This approach shows significant potential in enhancing agricultural productivity by effectively capturing crop patterns and identifying anomalies in farmland. Furthermore, it offers prospective solutions to address the inherent barriers farmers encounter. This study introduces a novel framework, namely the hybrid Convolutional Neural Networks and Long Short-Term Memory (HCNN-LSTM), which aims to detect anomalies in farmland using images obtained from UAVs automatically. The system employs a Convolutional Neural Network (CNN) for deep feature extraction, while Long Short-Term Memory (LSTM) is utilized for the detection task, leveraging the extracted features. By integrating these two Deep Learning (DL) architectures, the system attains an extensive knowledge of farm conditions, facilitating the timely identification of irregularities such as the presence of water, clusters of weeds, nutrient deficit, and crop disease. The proposed methodology is trained and evaluated using the Agriculture-Vision challenge database. The results obtained from the experiment demonstrate that the proposed system has achieved a high level of accuracy, with a value of 99.7%, confirming the effectiveness of the proposed approach.

1 Introduction

UAVs have become increasingly prevalent in sensing and surveillance applications, experiencing remarkable growth over the past decade. This can be attributed to notable technological advancements, including rapid developments in hardware and software. The utilization of UAVs for smart monitoring has emerged as a crucial aspect in multiple sectors, such as Precision Agriculture (PA) [1], post-disaster evaluation, metropolitan monitoring,

* Corresponding author: ku.omprakashdewangan@kalingauniversity.ac.in

frontier inspection, ecological surveillance, and numerous civil and tactical situations [2]. These applications are facilitated by a resilient technological and interpersonal infrastructure that includes collecting information through Remote Sensing (RS) devices, analyzing signals, and evaluating data using sophisticated algorithms within cloud-based platforms. The rapid progress in RS-based information analysis for autonomous decision-making can be attributed to recent advances in multifunctional sensing approaches, wireless connectivity, IoT, AI techniques, and cloud computing.

UAVs present several advantageous features compared to conventional sensor deployment methods. These advantages include reduced costs, simplified deployment processes, enhanced proficiency in RS, and the capability to collect high-resolution photos and videos of landscapes in a non-intrusive manner. This ability is precious for obtaining feature representations of the surveyed area [3]. The monitoring of farmland on an extensive basis has gained significance in facilitating PA [4].

PA plays a pivotal role in the contemporary agricultural revolution. Integrating IoT, AI, and other advanced technologies has led to the refinement and increased precision of farming procedures. Incorporating contemporary practices and concepts in agriculture is a beneficial factor, as it enhances agricultural productivity and promotes enhanced sustainable development [5]. It facilitates the implementation of Site-Specific Management (SSM), which entails executing appropriate actions at the appropriate locations and instants. PA offers a framework for the integration of these concepts into agricultural practices. Implementing RS devices in agricultural fields enables the tracking of methodological parameters, facilitating the acquisition of current information. This continuous information stream offers a current view of farm and plant variables. Climate surveillance, crop surveillance, livestock monitoring, and greenhouse mechanization are among the prominent use cases derived from the IoT that significantly contribute to agricultural management and offer valuable insights to farmers.

DL techniques have been effectively employed in a multitude of computer vision applications. As a result, numerous studies have also proposed visual-based approaches for the agricultural field [6]. Integrating deep learning techniques with RS and imaging technology can significantly enhance field maintenance practices by facilitating the evaluation and forecasting of various farmland and crop variables. Farmers can obtain data on various parameters, such as humidity and water content, using sensors mounted on UAVs. In addition, DL algorithms can evaluate the optimal conditions for sowing crops and offer valuable insights regarding appropriate timing for cultivation. These advancements can potentially improve agricultural productivity and promote adopting environmentally friendly farming practices [7]. The utilization of pattern recognition on farmland images holds significant potential for enhancing the forecasting of agricultural yields, determining appropriate crop patterns, assessing biological parameters associated with crops, and identifying anomalies in farmland.

Incorporating advanced technologies has become a fundamental aspect of transforming conventional farming practices in the ever-changing field of modern agriculture [8]. An innovation that garnered significant attention is the application of UAVs to capture high-resolution aerial imagery of agricultural land. CNNs and LSTMs have demonstrated significant efficacy in visual analysis and sequence modeling. This study explores the integration of these technologies, introducing an innovative CNN-LSTM framework specifically tailored for the automated identification of irregularities in agricultural land through the analysis of aerial imagery captured by UAVs.

Detecting anomalies in farmland is a crucial component of PA, with the objective of early identification of abnormalities such as crop illnesses, pest outbreaks, or irrigation problems. A high demand for labor characterizes the conventional manual techniques employed for monitoring extensive agricultural fields, and they are susceptible to human errors. The CNN-

LSTM framework proposed in this study aims to tackle these obstacles by harnessing the spatial comprehension abilities of CNNs and the temporal interactions recorded by LSTMs. The integration of these complementary elements enables the identification and examination of complex patterns and deviations in the temporal progression of agricultural landscapes by analyzing aerial images obtained from UAVs over time.

2 Related works

This literature review uses CNN-LSTM to study farmland anomaly detection using high-resolution UAV aerial images. The introduction emphasizes the convergence of these technologies and their potential for proactive farmland anomaly identification. The survey examines the key studies and methods contributing to CNN-LSTM frameworks for automatic anomaly detection.

Using data, Hu et al. (2023) assessed cattle farm air quality and environmental conditions. Sensors strategically placed on the farm collect environmental data, focusing on air quality factors. The research covers ecological factors in bovine farms [9]. The analysis includes many air quality and environmental measurements. Real-time monitoring and targeted interventions improve farm management. However, the initial financial investment for sensor deployment and data management challenges should be considered.

Deep learning was used by Kavithamani and Uma Maheswari (2023) to detect whitefly infestations in coconut tree leaves [10]. The method involves training a DL model with labeled images of various whitefly conditions. The output provides a mechanized coconut tree-whitefly detection system. The results show that the model detects whitefly infestations well. One benefit of this method is its quick and accurate target identification. It requires a large dataset with proper labeling and computational resources, which may be a drawback.

Ali et al. (2023) proposed a new pest detection system for large farms. This system performs well using AI and IoT. The authors used sound analytics to find pests. The implementation uses IoT devices with sound sensors to detect and monitor pest activity [11]. The system provides current agricultural pest data. This study accurately identifies pest sounds and sends notifications quickly. The benefits include early pest detection and reduced pesticide use. The need for seamless integration with farm management systems and the possibility of false positives are drawbacks.

Di Lorenzo et al. (2022) used predictive methods to identify photovoltaic system performance and deterioration [18]. The methodology involves systematic observation and analysis of photovoltaic system performance data. The output predicts system issues. The outcome values include early detection of poor performance and decline. Prognostic algorithms improve system maintenance and lifespan. However, the complexity of these algorithms may cause implementation issues.

Chaudhary et al. (2024) examined a hybrid framework for precise farming that integrates ML and the IoT [13]. Implementation requires the synthesis of hybrid framework literature in precision agriculture. The output provides a comprehensive analysis of current methodologies' pros and cons. The findings provide insights into precision agriculture's ML and IoT technologies use. Hybrid frameworks are comprehensive but can be difficult to interoperate due to the many technologies involved.

Khan et al. (2022) used machine and deep learning to analyze hyperspectral imaging technology for agricultural applications [14]. The methodology involves reviewing hyperspectral imaging technology applications and advancements literature. The output is a comprehensive assessment of hyperspectral imaging in agriculture. Results show machine and deep learning's effectiveness in hyperspectral data analysis. This technology's strengths are its comprehensive understanding of its potential, while its weaknesses include data preprocessing and model interpretability.

Kou et al. (2022) implemented the Software-Defined Networking (SDN)-based network intrusion detection model for drone communication. SDN is used to create a model to identify drone communication network intrusions [15]. An effective intrusion detection system for UAV networks is produced. Outcome measures include intrusion detection accuracy. Drone communication networks improve security. However, this approach may have drawbacks, such as adapting the model to new threats.

Fu et al. (2023) proposed a machine learning-based intrusion detection system for agricultural information security and UAVs. ML algorithms are integrated into UAV-supported agricultural information systems during implementation [16]. The output improves agricultural data and system security. The outcomes measure intrusion detection precision and security architecture resilience. Implementing agricultural information systems improves cybersecurity. This can improve agricultural data security from unauthorized access or malicious activity. However, drawbacks should be considered. For instance, updating to combat new cybersecurity threats may be difficult or inconvenient.

In conclusion, this literature survey used CNN-LSTM to detect farmland anomalies, demonstrating the strength of UAV-captured aerial imagery and advanced deep learning. The surveyed studies demonstrate how this approach has revolutionized agricultural monitoring. Converging CNNs and LSTMs will shift anomaly detection from reactive to proactive, giving farmers and agricultural practitioners a powerful tool to optimize resource management and farm productivity [19].

3 Proposed HCNN-LSTM to detect anomalies in farmland using images obtained from UAVs

The HCNN-LSTM framework is an advanced approach for the automated identification of anomalies in agricultural land by analyzing images captured from UAVs. The proposed methodology integrates the spatial feature extraction capabilities of CNNs with the temporal sequence modeling proficiency of LSTMs. The CNN module demonstrates exceptional proficiency in detecting complex patterns within high-resolution aerial images, enabling the identification of spatial intricacies that are essential for detecting abnormalities in crops and soil conditions [12]. Concurrently, the LSTM component processes these features temporally, thereby discerning dynamic patterns and deviations. Utilizing a hybrid framework facilitates a comprehensive approach to anomaly detection, thereby contributing to the advancement of precision agriculture. This approach enables timely and accurate insights into the overall health of farmland, thereby enhancing its effectiveness. The integration of CNN and LSTM technologies represents a notable advancement in proactive, data-driven farm management. This development enables the rapid detection of anomalies and facilitates precise interventions that optimize agricultural outcomes.

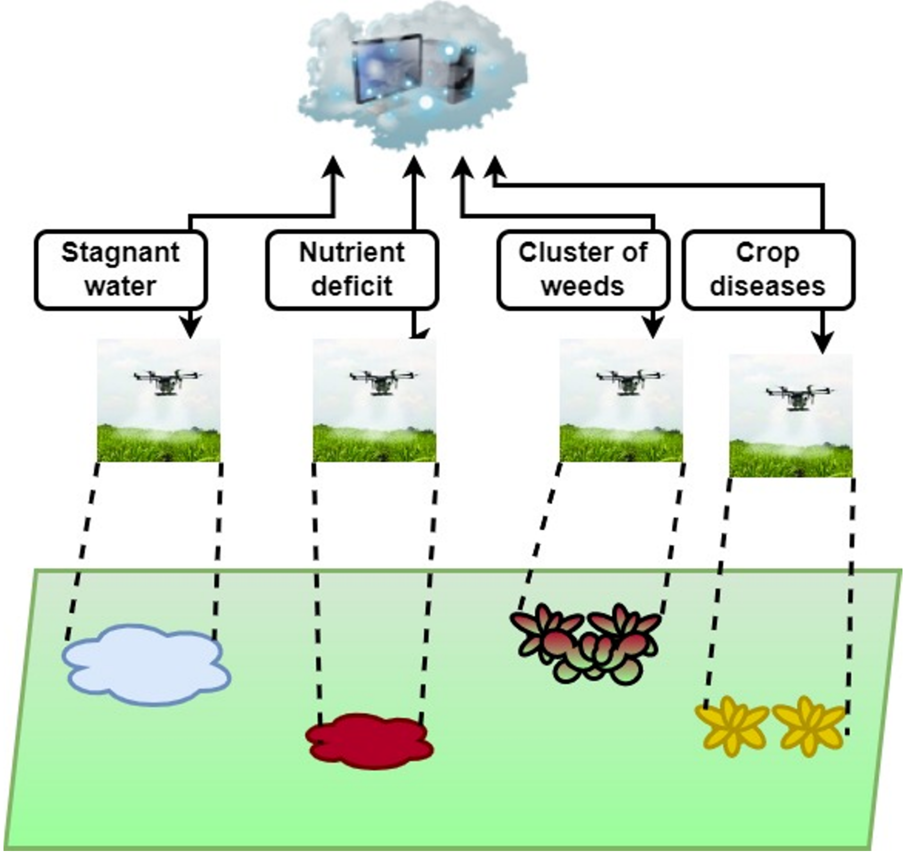

Fig. 1. IoT framework with cloud server to detect anomalies in farmland using images obtained from UAVs

Fig. 1 illustrates an IoT framework that incorporates a cloud server to detect farmland anomalies. This detection is achieved through the utilization of images acquired from UAVs. This study introduces an IoT framework that incorporates a cloud server. A deep aerial semantic segmentation framework is also proposed for agricultural pattern recognition. This framework utilizes UAV-acquired aerial images. The overall objective of this research is to explore the potential of IoT-assisted PA. In this study, we present a resilient multi-scale hierarchical attention network designed for semantic segmentation on images of farmland.

3.1 CNN

CNNs, which belong to the category of deep learning models, have demonstrated significant efficacy in diverse computer vision applications such as image recognition, object detection, and semantic segmentation. The design of a CNN draws inspiration from the neural processing mechanisms observed in the human visual system, making it highly suitable for tasks that involve structured data representations, such as images. The architecture of a CNN comprises essential components such as Convolutional Layers (CL), Pooling Layers (PL), and Fully Connected (FC) layers. Every layer within the system performs a vital function in extracting and abstracting features. CL is a fundamental component of CNNs; the primary operation is the convolution operation. The process entails applying a filter, commonly referred to as a kernel, to the input image to extract localized features. In mathematical terms, the convolution operation can be represented as:

$$S(i,j) = (I * K)(i,j) = \sum m \sum n I(m,n) K(i-m, j-n) \tag{1}$$

Here, $S(i,j)$ is the output of the convolution operation, I is the input image, and K is the filter.
Activation Function:
In conventional practice, employing a non-linear activation function such as the Rectified Linear Unit (ReLU) after the convolution operation is common. The rectified linear unit (ReLU) function is mathematically defined as follows:

$$f(x) = max(0,x) \tag{2}$$

PL: PL is responsible for downsampling the input's spatial dimensions, resulting in reduced computational requirements and improved translational invariance. One frequently used pooling operation in computer vision is known as max pooling. This operation is defined as:

$$O(i,j) = \max \begin{bmatrix} I(2i,2j) & I(2i+1,2j) \\ I(2i,2j+1) & I(2i+1,2j+1) \end{bmatrix} \tag{3}$$

FC layers: Following a series of CL and PL, it is customary for the network to conclude with one or more FC layers. The interconnections between neurons in each layer establish a comprehensive connectivity pattern wherein every neuron is linked to every neuron in both the preceding and succeeding layers.

3.2 LSTM Classifier

The input to the LSTM in the proposed approach consists of the feature maps obtained from the Softmax layer, which are utilized to improve the classification performance. To enhance and regulate the data performance of the proposed system, the LSTM model incorporates a gate architecture consisting of an Input Gate (IG), Forget Gate (FG), and Output Gate (OG). The numerical equation corresponding to each gate is elucidated as follows:
The FG (Fg^t) aids the LSTM in deciding the data that must be provided and dismissed through the cell state based on the prior hidden structure.

$$Fg^t = \sigma(M_{Fg}.[H^{t-1}, Y^t] + b_{Fg}) \tag{4}$$

Where M_{Fg} and b_{Fg} are the weight and bias vectors, respectively. H^{t-1} is the previous gate, Y^t is the input data, and $\sigma(.)$ is the sigmoid function. The candidate's cell state $\tilde{S}^t$ determined by IG (Ig^t) is given as:

$$\tilde{S}^t = \tanh(M_S.[H^{t-1}, Y^t] + b_S) \tag{5}$$

$$Ig^t = \sigma(M_{Ig}.[H^{t-1}, Y^t] + b_{Ig}) \tag{6}$$

$\tanh(.)$ denotes the hyperbolic tangent function. The updated cell state S^t has been given as:

$$S^t = Fg^t * S^{t-1} + Ig^t * \tilde{S}^t \tag{7}$$

S^{t-1} is the past cell state, and the new candidate cell state is given by $\tilde{S}^t$. The output of the LSTM cell has been regulated by the OG, which is given as follows:

$$Og^t = \sigma(W_{Og}.[H^{t-1}, Y^t] + b_{Og}) \tag{8}$$

$$\text{where } H^t = Og^t.\tanh(S^t) \tag{9}$$

The LSTM model effectively tackles the issue of persistent dependencies in data caused by prolonged addiction and overcomes the problem of gradient vanishing or exploding by utilizing memory storage cells and gate mechanisms. Consequently, the LSTM model mitigates the limitations of Recurrent Neural Networks (RNNs).

3.3 Proposed HCNN-LSTM

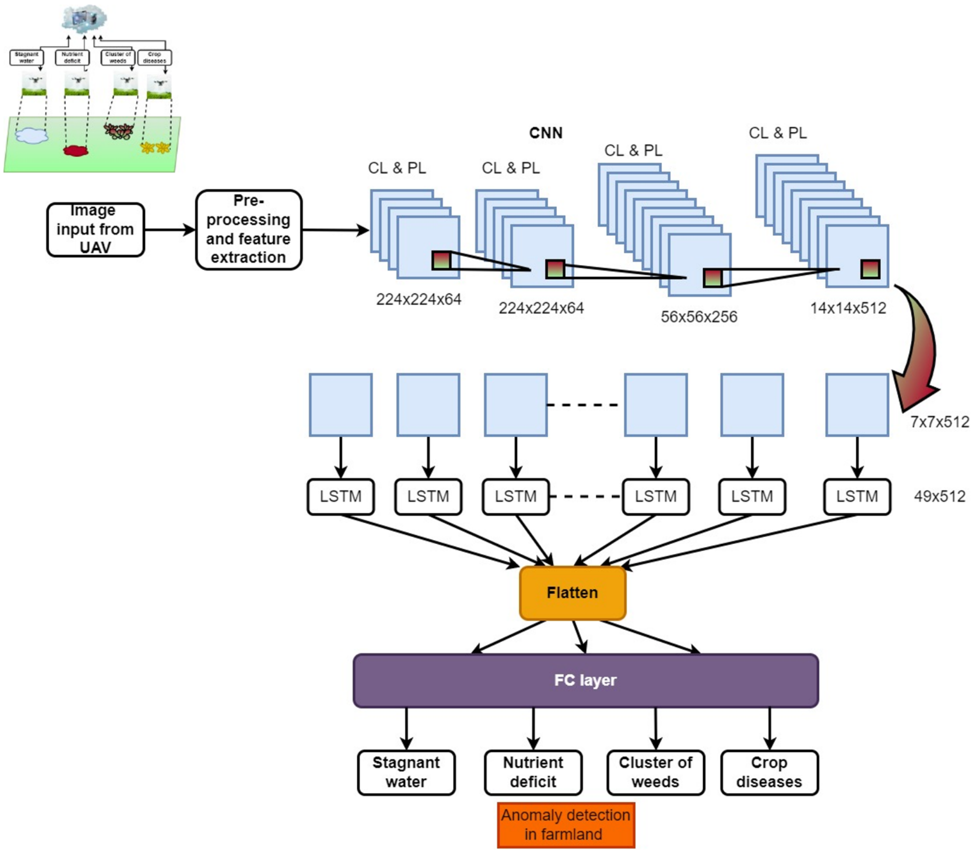

Fig. 2. Proposed HCNN-LSTM framework for automatic detection of anomalies in farmland using images obtained from UAVs

Fig. 2 depicts the proposed HCNN-LSTM framework for automatic detection of anomalies in farmland using images obtained from UAVs. The network consists of a total of 19 layers, which include 11 CL, five PL, one FC layer, one LSTM layer, and one output layer that utilizes the softmax function. Each convolutional block is augmented with two or three 2D-CNNs and one PL. This is subsequently followed by a dropout layer, which is characterized by a dropout rate of 20%. A CL with 3 × 3 kernels is employed for feature extraction, subsequently activated through the ReLU function. Utilizing a max-PL, employing kernels of size 2 × 2, diminishes the dimensions of an input image. In the final architectural design stage, the function map is passed on to the LSTM layer to extract time-related data. Following the convolutional block, the resulting output shape is determined to be (7, 7, 512). The input size of the LSTM layer has been transformed to (49, 512) using the reshape technique. By integrating two DL architectures (CNN and LSTM), the system attains an extensive knowledge of farm conditions, facilitating the timely identification of irregularities such as the presence of water, clusters of weeds, nutrient deficit, and crop disease.

4 Results and discussion

The proposed methodology is trained and evaluated using the Agriculture-Vision challenge dataset [17]. The dataset comprises a total of 20,997 images of agricultural land that were captured throughout the United States in the year 2020. Each image has four color channels, each with a resolution of 256×256 pixels. These channels include the Red, Green, and Blue (RGB) channels and the Near Infrared (NIR) channels. In addition, the images are accompanied by a perimeter map and a binary masking. The perimeter map delineates the spatial extent of the agricultural land depicted in the image, whereas the binary mask represents the pixels considered valid. The evaluation process does not encompass regions beyond the perimeter map or the mask. The markings encompass four distinct categories: identifying water stagnation, cluster of weeds, nutrient deficiency, and crop disease.

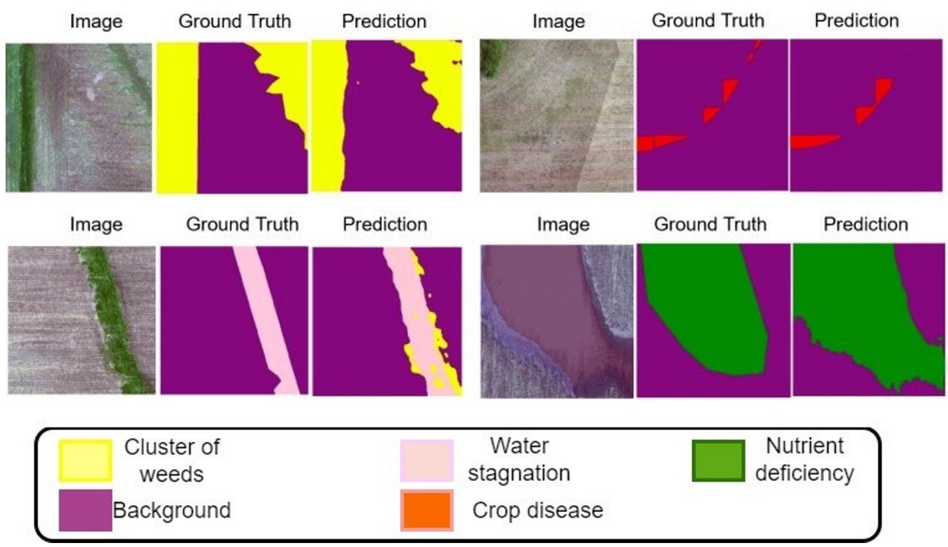

Fig. 3. Results of segmentation of images using the proposed method in the Agriculture-vision challenge dataset for automatic detection of farmland anomalies

Fig. 3 depicts the results of the segmentation of images using the proposed method in the Agriculture-vision challenge dataset for automatically detecting farmland anomalies. The segmentation predictions for classes such as weed clusters and cloud shadow exhibit a high level of precision, particularly in cases where the segmentation area is significantly larger. The model we have developed can accurately predict the presence of standing water on the farm field, even when the water is shallow. This prediction is based on the analysis of Fig. 3, which highlights the oversight of ignoring the presence of standing water during target segmentation for the corresponding image. Furthermore, there has been extensive analysis of cases involving complex shapes following the predictions made by the image analysis. A robust segmentation has been generated despite the similarity in image features between waterways and weed clusters. However, some erroneous predictions are along the corners of the background class. The achievement of high accuracy in results within the Agriculture-vision challenge dataset is noteworthy despite class imbalance.

		Background	Crop diseases	Nutrient deficiency	Water stagnation	Weed Cluster
Actual	**Background**	**0.7**	0.13	0.03	0.03	0.11
	Crop diseases	0.28	**0.71**	0	0	0
	Nutrient deficiency	0.45	0.24	**0.19**	0.05	0.07
	Water stagnation	0.09	0.11	0.04	**0.75**	0.01
	Weed Cluster	0.25	0.09	0.01	0.01	**0.62**

Predicted

Fig. 4. Confusion matrix for the prediction of farmland anomalies using the proposed HCNN-LSTM framework

Fig. 4 gives the confusion matrix for predicting farmland anomalies using the proposed HCNN-LSTM framework. The diagonal elements of the matrix correspond to the true positive rates for each class, which signify the proportion of correctly predicted instances. An example of the model's effectiveness in accurately identifying instances of water stagnation can be observed in the high value 0.75 in the corresponding row and column. On the other hand, the presence of off-diagonal elements in the matrix demonstrates instances of misclassifications. For example, the value of 0.28, located in the intersection of the "Crop diseases" and the "Background," signifies a discernible degree of confusion between crop diseases and background. The matrix is a comprehensive instrument for assessing the model's precision, recall, and F1 score for each anomaly class. This aids in guiding subsequent enhancements and advancements in the HCNN-LSTM framework for detecting anomalies in farmland.

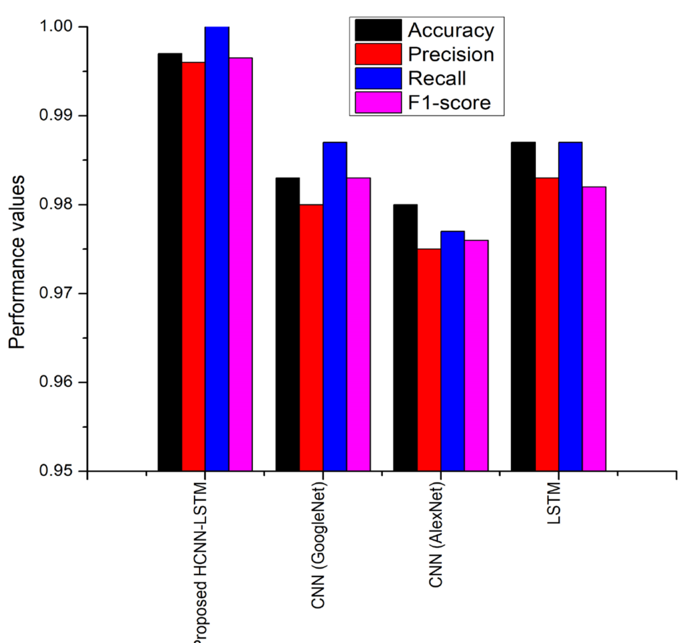

Fig. 5. Performance of the various DL models regarding accuracy, precision, recall, and F1-score for the prediction of farmland anomalies

Fig. 5 depicts the performance of the various DL models regarding accuracy, precision, recall, and F1-score for predicting farmland anomalies. The HCNN-LSTM model demonstrates a high level of accuracy, reaching an impressive 99.7%. This performance surpasses other models, highlighting its robustness in effectively capturing and classifying anomalies. The precision value of 99.6% signifies a low occurrence of false positives, highlighting the model's proficiency in accurately detecting anomalies. Moreover, achieving a perfect % recall rate of 100% indicates that the HCNN-LSTM model successfully detects all occurrences of anomalies without any false negatives. The F1-score, which stands at 99.65%, serves as a comprehensive measure of the model's performance by considering both precision and recall. When considering various models such as CNN (GoogleNet), CNN (AlexNet), and LSTM, it is evident that they demonstrate impressive performance. However, the proposed HCNN-LSTM model stands out as a particularly accurate and dependable solution for predicting anomalies in farmland.

5 Conclusion

This paper presents a new framework, called the hybrid Convolutional Neural Networks and Long Short-Term Memory (HCNN-LSTM), to automatically detect anomalies in farmland using images acquired from UAVs. The employed system utilizes a CNN to perform deep feature extraction. Additionally, the system leverages LSTM for the detection task using the extracted features. By combining these two DL architectures, the system acquires a comprehensive understanding of agricultural conditions, enabling the prompt detection of anomalies such as the existence of water, clusters of weeds, nutrient deficiency, and crop diseases. The methodology under consideration is trained and assessed using the Agriculture-Vision challenge database. The experimental findings indicate that the proposed system has attained a significant degree of accuracy, with a value of 99.7%. This outcome validates the proposed approach's efficacy for detecting anomalies in farmland using UAV-acquired images.

References

1. M. A. Latif, IEEE Geoscience and Remote Sensing Magazine, **6(4),** 10-22 (2018).
2. R. Mishra, H. P. Gupta, & T. Dutta, IEEE Sensors Journal, **21(14),** 15527-15534 (2020).
3. M. J. Sobouti, Z. Rahimi, A. H. Mohajerzadeh, S. A. H. Seno, R. Ghanbari, J. M. Marquez-Barja, & H. Ahmadi, IEEE Sensors Journal, **20(13),** 7460-7471 (2020).
4. H. Tian, T. Wang, Y. Liu, X. Qiao, & Y. Li, Information Processing in Agriculture, **7(1),** 1-19 (2020).
5. D. Serebrennikov, F. Thorne, Z. Kallas, & S. N. McCarthy, Sustainability, **12(22),** 9719 (2020).
6. V. G. Dhanya, A. Subeesh, N. L. Kushwaha, D. K. Vishwakarma, T. N. Kumar, G. Ritika, & A. N. Singh, Deep learning based computer vision approaches for smart agricultural applications, Artificial Intelligence in Agriculture (2022).
7. T. A. Shaikh, T. Rasool, & F. R. Lone, Computers and Electronics in Agriculture, **198,** 107119 (2022).
8. P. Rai, & S. Maharjan, Quarterly Journal of Emerging Technologies and Innovations, **8(3),** 18-32 (2023).
9. J. Hu, R. Jagtap, R. Ravichandran, C. P. Sathya Moorthy, N. Sobol, J. Wu, & J. Gao, Atmosphere, **14(5),** 771 (2023).

10. V. Kavithamani, & S. UmaMaheswari, *Investigation of Deep learning for whitefly identification in coconut tree leaves*, Intelligent Systems with Applications, 200290, (2023).
11. M. A. Ali, R. K. Dhanaraj, & A. Nayyar, Microprocessors and Microsystems, **103,** 104946 (2023).
12. V. K. Stephen, V. Mathivanan, A. R. Manalang, & et al. Journal of Internet Services and Information Security, **13(2),** 128-145 (2023).
13. R. R. Chaudhary, S. Jain, & S. Gupta, Journal of Integrated Science and Technology, **12(2),** 730-730 (2024).
14. A. Khan, A. D. Vibhute, S. Mali, & C. H. Patil, Ecological Informatics, **69,** 101678 (2022).
15. L. Kou, S. Ding, T. Wu, W. Dong, & Y. Yin, Drones, **6(11),** 342 (2022).
16. R. Fu, X. Ren, Y. Li, Y. Wu, H. Sun, & M. A. Al-Absi, IEEE Internet of Things Journal (2023).
17. M. T. Chiu, X. Xu, Y. Wei, Z. Huang, A. G. Schwing, R. Brunner, ... & H. Shi, *Agriculture-vision: A large aerial image database for agricultural pattern analysis,* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2828-2838, (2020).
18. G. Di Lorenzo, E. Stracqualursi, L. Micheli, S. Celozzi, & R. Araneo, Energies, **15(17),** 6413 (2022).
19. I. Solikin, & D. Darmawan, Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, **14(2),** 82-93 (2023)