

DEEP LEARNING

Ensembled Deep Learning-Based Optimisation of Multi-Robot Communication Networks in Agricultural Environments

Jawahir Mohammad Hashir, M. A. Baba, M. Tela, J. Adamu A. Umar

Centre for Embedded Artificial Intelligence and Smart Energy Systems, Abubakar Tafawa Balewa University, Bauchi, Nigeria

Published October 2026

Abstract

In this paper, a novel ensemble deep learning framework is presented for communication management in agricultural multi-robot systems. The framework uses Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM), and Graph Neural Networks (GNN) to classify link quality, predict the future evolution of link quality, and inform routing decisions in wireless networks used in agricultural scenarios. The models were trained on the Vineyard-2021 LoRa RSSI dataset, comprising real wireless measurements gathered during deployment in vineyards in Northern Italy, preprocessed and annotated across seven sensor nodes and 1,877 hourly timesteps. To prevent target leakage, the prediction task was framed as a future-prediction task: models predict the link-quality class at time $t+1$ using only measurements available up to time t . The CNN reached a test accuracy of 83.69%, the LSTM 81.14%, and the GNN 85.82%. An unweighted ensemble of the three models reached a test accuracy of 89.32%. All classification errors occurred between adjacent classes, and no Weak-to-Strong or Strong-to-Weak errors were observed. This ensemble was then used to drive routing decisions in a multi-robot simulation built on the NetworkX platform, with fleet sizes ranging from 5 to 30 robots. Adaptive routing driven by the ensemble reached a Packet Delivery Ratio of 99.68%, compared with 95.33% for a distance-only baseline, while slightly increasing latency (18.03 ms vs. 12.80 ms) and energy consumption (1.81 vs. 1.36 units). The paper presents a documented baseline for deep-learning-based link-quality classification on agricultural LoRa data, a novel ensemble design combining spatial, temporal, and topological features, and an evaluation methodology for communication management in simulated multi-robot networks. Limitations include the static deployment of the sensors used for training, the geographic and seasonal specificity of the dataset, and the use of literature-derived packet-success probabilities in the simulation. Future work will gather data from moving robots in Nigerian agricultural environments and validate the framework using ROS/Gazebo.

Keywords: convolutional neural networks; ensemble learning; graph neural networks; long short-term memory; multi-robot communication

1 Introduction

Agriculture in sub-Saharan Africa, and Nigeria in particular, is under pressure from population growth, climate change, reducing soil quality, and a shrinking farm workforce, and the old ways of producing food may not be enough to feed everyone in the coming years. Agriculture contributes approximately 24% of Nigeria's GDP and remains the source of income for the majority of the rural population [1]. The Nigerian National Agricultural Technology and Innovation Policy (NATIP) 2022-2027 identifies precision agriculture, robotics, and digital technologies as tools that can help modernise farming [2], and among these, locally deployed autonomous multi-robot systems have great potential to support large-scale precision farming.

A multi-robot system is a group of robots that work together in the same area; unlike a single robot working alone, many robots can share tasks, cover a larger area, and work continuously. In agriculture, groups of robots can be used for crop monitoring, accurate spraying, soil sampling, and harvesting because they can carry out these tasks more effectively than other methods [3], [4]. The main problem affecting the use of multi-robot systems on real farms is not the capability of the robots to work, but their ability to communicate effectively: when communication fails, the robots cannot coordinate, they act on old or missing information, and coordinated tasks fail.

Farms are among the most difficult environments for wireless communication. Crops and plants weaken and block radio signals, and the amount of loss depends on the type of crop and how much it has grown [5]. Farm machinery produces electrical noise, terrain creates areas of poor connectivity, and rain further affects signal propagation - all of which change constantly over time. Common wireless protocols such as CSMA/CA and TDMA were not designed for such a rapidly changing environment: they mainly react to a problem after it has occurred, which is too slow for a channel that can change within seconds because of movement, machinery, and weather.

Deep learning offers an alternative: rather than reacting after a problem has occurred, a well-trained neural network can learn the relationship between what can be observed in the wireless environment and how well the communication channel is working, and use this to predict problems before they happen. Convolutional Neural Networks (CNNs) are effective at finding patterns in structured, recurring data such as interference signatures [6]. Long Short-Term Memory (LSTM) networks are suited to data that changes over time and can use past channel measurements to predict future changes [7]. Graph Neural Networks (GNNs) can represent how robots are connected in a network, allowing the system to reason about the network as a whole rather than about a single link in isolation [8].

This research brings these three network types together into a single ensemble communication-management framework for agricultural robots. The CNN, LSTM, and GNN components were trained and tested on the Vineyard-2021 LoRa RSSI dataset, a real dataset recorded from sensors in a functioning vineyard, and the three trained models were then combined into an ensemble and evaluated further in a larger multi-robot network simulation. The specific objectives of the study are to (i) develop CNN, LSTM, and GNN models for interference detection, link-quality prediction, and route selection; (ii) combine the three models into a single ensemble framework and compare its performance against conventional routing on packet delivery ratio, latency, and energy consumption; and (iii) evaluate the ensemble framework across different fleet sizes and communication conditions to determine how well it scales.

2 Related work

Research into multi-robot systems in agriculture has grown rapidly over the last two decades. Dorigo et al. [3] reviewed the development of swarm robotics from early coordination concepts to field experiments, and found that unreliable communication remains a key barrier to the wide-spread use of multi-robot systems on real farms. Early multi-robot field trials in European vineyards demonstrated that autonomous robot fleets could perform useful agricultural tasks but were affected by wireless communication problems while operating in and around crop vegetation [9]. Queralta et al. [4] similarly found, in a systematic review of collaborative multi-robot systems, that most studies either assume reliable communication or simplify the wireless environment in ways that limit the applicability of their findings to real agricultural deployments.

Signal-propagation studies confirm that radio signals in agricultural environments behave differently from those in open or indoor spaces. Gielis et al. [5] found that mature crops caused greater LoRa signal loss than young crops, mainly because of increased moisture content and leaf density, and that this loss varies with crop growth stage. Similar heterogeneity has been reported for other agricultural LoRa deployments [10], [11], reinforcing the case for models that can learn directly from field measurements rather than relying on fixed propagation formulas.

Deep learning has been applied to wireless link-quality problems in other domains: recurrent networks have been used for link-quality prediction in community wireless networks [12], and deep models have been used for link-quality estimation in industrial wireless towers [13]. Ensemble methods, in general, are well established as a way of combining diverse learners to reduce error relative to any single model [14], [15]. However, at the time of writing, no published study combines CNN, LSTM, and GNN architectures for communication management in agricultural multi-robot networks, and prior link-quality prediction work has generally used present-time RSSI values as predictors, which risks target leakage. The present study addresses this gap by (i) combining spatial, temporal, and topological deep-learning components into a single ensemble, and (ii) framing the prediction task as a genuine future-prediction problem.

3 Research methodology

The study follows an experimental, data-driven design carried out in four stages: (1) dataset preparation and CNN training; (2) LSTM training; (3) GNN training on the graph structure of the sensor network; and (4) combination of the three trained models into an ensemble framework evaluated against a baseline in a multi-robot network simulation. Python 3.10 was used throughout; TensorFlow 2.12 was used to build the CNN and LSTM models, PyTorch 2.0 with PyTorch Geometric was used for the GNN, and scikit-learn was used for preprocessing and baseline comparison. Figure 1 summarises the methodology from problem identification through to the fleet-level simulation and conclusions.

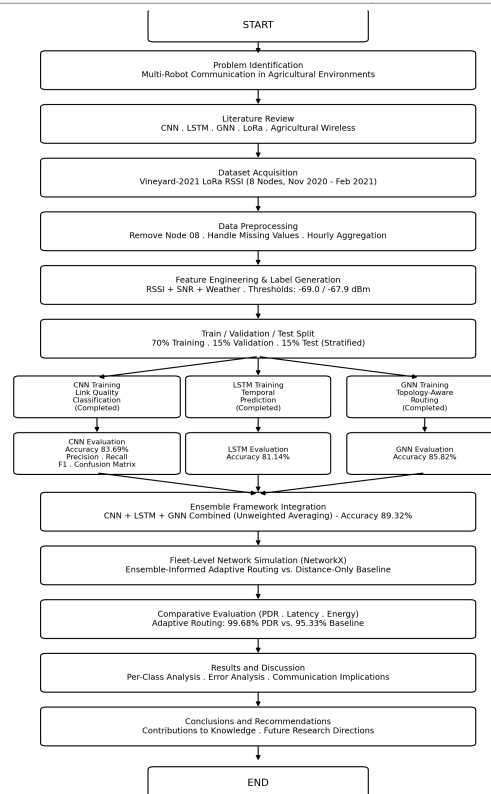


Figure 1. Research methodology flowchart.

3.1 Dataset: Vineyard-2021 LoRa RSSI

The Vineyard-2021 dataset is a real-world dataset collected from a LoRa sensor network in a vineyard in Northern Italy between November 2020 and February 2021, using eight Tinovi PM-IO-5-SM sensor nodes communicating with a Mikrotik wAP LR9 kit LoRa gateway at distances of 40 to 340 metres. The nodes reported RSSI, SNR, and weather variables (temperature, humidity, rainfall). After hourly aggregation, the dataset provided about 2,026 timesteps across the seven nodes retained after preprocessing. Although the dataset was collected from fixed sensors rather than moving robots, it captures real vegetation, terrain, weather, and interference effects across nodes with different distances and line-of-sight conditions to the gateway - conditions broadly comparable to the link diversity found in a multi-robot network.

3.2 Data preprocessing and label generation

Node 08 was removed because of extensive missing data, leaving 2,026 hourly records from seven nodes; a further 149 records with incomplete RSSI across all seven nodes were removed, leaving 1,877 fully labelled records. These were split chronologically (to avoid leakage from the strong autocorrelation of hourly time series) into 70% training (1,314), 15% validation (282), and 15% test (281) sets. Labels were generated by averaging RSSI across all seven nodes and applying thresholds from the Vineyard-2021 documentation: Strong (Class 2) for average RSSI greater than -67.9 dBm, Medium (Class 1) for -69.0 to -67.9 dBm, and Weak (Class 0) for RSSI at or below -69.0 dBm. Because the prediction target is the class at $t+1$ using only data up to time t , the RSSI value used to form the label is never available as an input for the same instance. The resulting classes are close to

balanced: 619 Weak, 620 Medium, and 638 Strong samples, so a majority-class classifier would achieve only about 34% accuracy.

3.3 CNN architecture

The CNN treats the 18-value feature vector (RSSI and SNR values plus engineered features) as a one-dimensional signal. It uses three convolutional layers of 32, 64, and 128 filters (3×3 kernels, ReLU activation) each followed by max pooling, then two fully connected layers of 256 and 128 neurons with 30% dropout after the first, and a three-unit softmax output layer. Training used categorical cross-entropy loss and the Adam optimiser [16] with a learning rate of 0.001, batch size 32, and early stopping with a patience of 15 epochs; the model has 245,891 trainable parameters. Training stopped at epoch 61 with the best checkpoint restored from epoch 46.

3.4 LSTM architecture

The LSTM predicts link-quality at $t+1$ from the previous 24 hourly readings ($t - 23$ to t), using two stacked LSTM layers of 128 hidden units each, a feed-forward layer of 64 units, and a three-unit output layer, with 30% dropout between layers [7]. It was trained with the same loss and optimiser as the CNN for up to 150 epochs with early stopping, and has 198,531 trainable parameters.

3.5 GNN architecture

The GNN represents the seven sensor nodes as a graph $G = (V, E)$ with $|V| = 7$. Each node's 8-dimensional feature vector combines its RSSI, SNR, distance to gateway, shared weather variables, and a sine/cosine hour encoding. Nodes within 100 metres of each other are connected, with an edge weight

$$w_{ij} = \exp\left(-\frac{d_{ij}}{50}\right) \quad (1)$$

so that closer nodes are weighted more strongly, following the standard GCN self-loop and normalisation convention [8], [17]. The architecture uses two GCN layers (GCNConv) followed by global mean pooling and a linear classifier, with 30% dropout between the GCN layers, trained with cross-entropy loss; it has 12,547 trainable parameters.

3.6 Ensemble framework design

The CNN, LSTM, and GNN play complementary roles: the CNN classifies instantaneous link quality from RSSI/SNR, the LSTM predicts link quality one step ahead using 24 hours of history, and the GNN incorporates network topology to support routing-level decisions. The softmax outputs of the three trained models are combined by simple unweighted averaging to produce a single probability distribution over the three link-quality classes. Unweighted averaging was chosen for its simplicity and interpretability, and because it requires no additional training while still allowing each model's contribution to be examined separately [14].

3.7 Fleet-level simulation setup

To evaluate the ensemble beyond static classification, a discrete-event network simulator was built in Python using NetworkX. For fleet sizes of 5, 10, 15, 20, 25, and 30 robots, robots were placed randomly within an area scaled to the vineyard site, with five random topologies generated per fleet

size. Packet transmission outcomes were simulated independently of the classifier, using a log-distance path-loss model with a log-normal shadowing term ($\sigma = 4$ dB), to avoid circular validation:

$$P_{rx} = P_{tx} - (20\log_{10}d + 20\log_{10}f - 147.55) - X_{\sigma} \quad (2)$$

At each simulation step, the actual simulated RSSI readings were fed to the trained ensemble to obtain a link-quality classification, which was then used to compare an ensemble-informed adaptive routing policy against a conventional distance-only baseline. For each topology and fleet size, 300 packets were routed between randomly chosen source-destination pairs, and results were averaged over five repeats with 95% confidence intervals.

4 Results and discussion

4.1 Dataset overview

The cleaned Vineyard-2021 dataset comprises 1,877 labelled hourly observations across seven LoRa sensor nodes. Figure 2 shows the per-node RSSI time series, the average-RSSI link-quality classification, the per-node SNR time series, and the final class distribution. RSSI levels vary widely across nodes, from about -55 dBm for the closest node (40 m from the gateway) to -80 to -85 dBm for the weakest node (160 m, likely with partial obstruction), confirming that the network regularly contains links spanning all three quality classes simultaneously. SNR remained mostly within 8.0–9.5 dB, with occasional sharp drops that were retained in the dataset as genuine interference rather than removed as outliers.

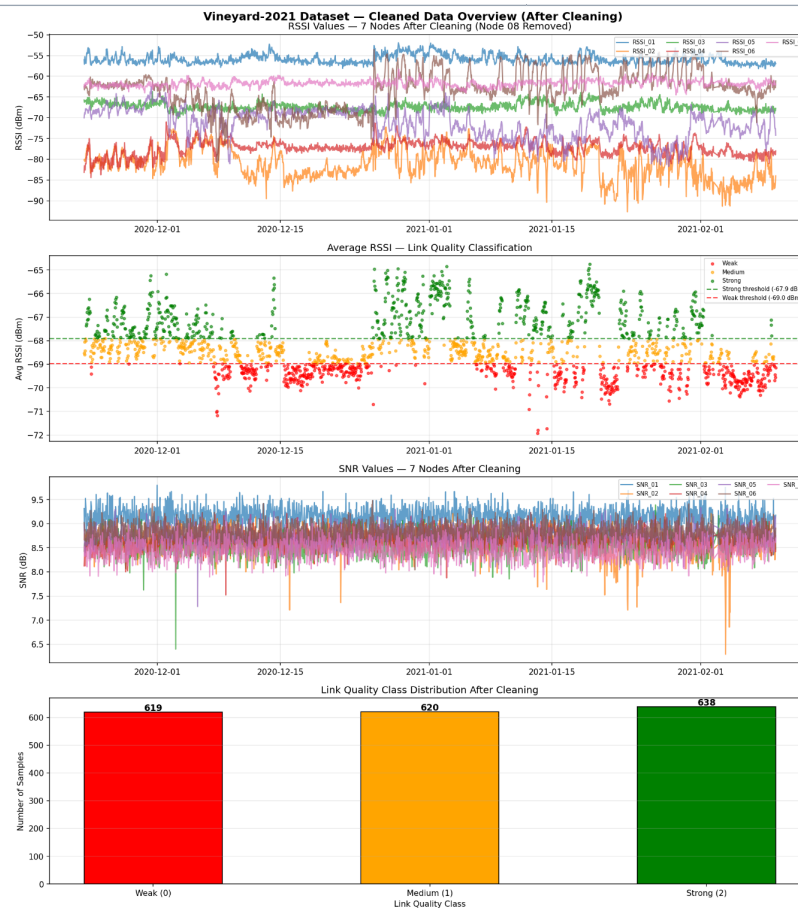


Figure 2. Vineyard-2021 dataset overview after cleaning: per-node RSSI time series, average-RSSI link-quality classification, per-node SNR time series and the final class distribution.

4.2 CNN training and classification performance

The CNN was trained for up to 150 epochs with early stopping; training concluded at epoch 61 with the best checkpoint restored from epoch 46 (Figure 3). The model reached a test accuracy of 83.69% on the balanced three-class problem - more than 50 percentage points above the approximately 33% expected by chance. Precision was highest for the Strong class (0.95), reflecting the well-separated signals from nodes close to the gateway, while the Weak class achieved 0.88 precision, 0.84 recall, and an F1 score of 0.86, with no Weak samples misclassified as Strong. The Medium class was the most difficult, with precision of 0.72 and F1 of 0.77, because its RSSI range spans only 1.1 dBm (-69.0 to -67.9 dBm), making it easy for borderline signals to drift across the boundary between one hourly reading and the next. Critically, all CNN errors occurred between adjacent classes: no Weak sample was classified as Strong and no Strong sample was classified as Weak, which is the most operationally important type of error to avoid in a routing system.

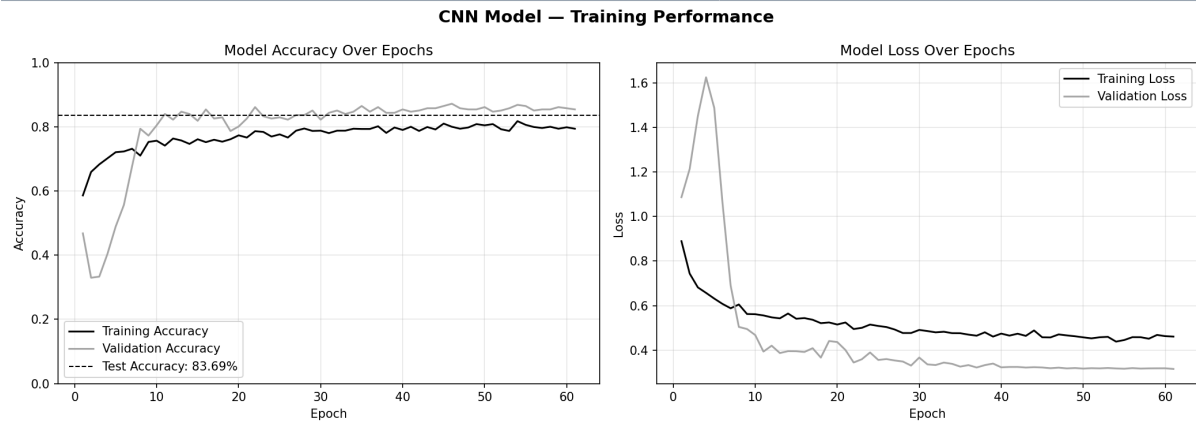


Figure 3. CNN training and validation accuracy and loss curves.

4.3 LSTM and GNN performance

The LSTM, trained to predict link quality one hour ahead from the preceding 24 hourly readings, reached a test accuracy of 81.14% (Figure 4). The GNN, which incorporates the spatial topology of the sensor network, achieved the best standalone accuracy of 85.82% (Figure 5), suggesting that network-level structural information is at least as informative as instantaneous or temporal features alone for this task.

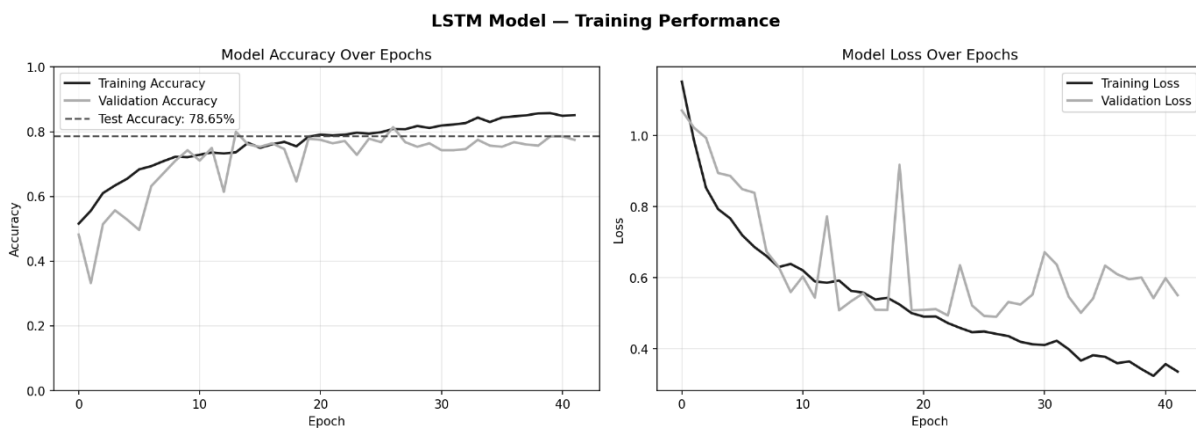


Figure 4. LSTM training and validation accuracy and loss curves.



Figure 5. GNN training and validation accuracy and loss curves.

4.4 Ensemble framework evaluation

Combining the CNN, LSTM, and GNN outputs by unweighted averaging produced a test accuracy of 89.32%, exceeding each individual model, with a macro-averaged F1 score and balanced accuracy of 0.89. Of the 281 test timesteps, the ensemble correctly classified 85 Weak, 79 Medium, and 87 Strong links, with precision/recall of 0.94/0.91 for Weak, 0.82/0.86 for Medium, and 0.92/0.91 for Strong (Figure 6). Because the CNN, LSTM, and GNN attend to different aspects of the communication environment - instantaneous readings, temporal trends, and network topology, respectively - their errors are not strongly correlated, so combining them reduces the overall error relative to any single model.

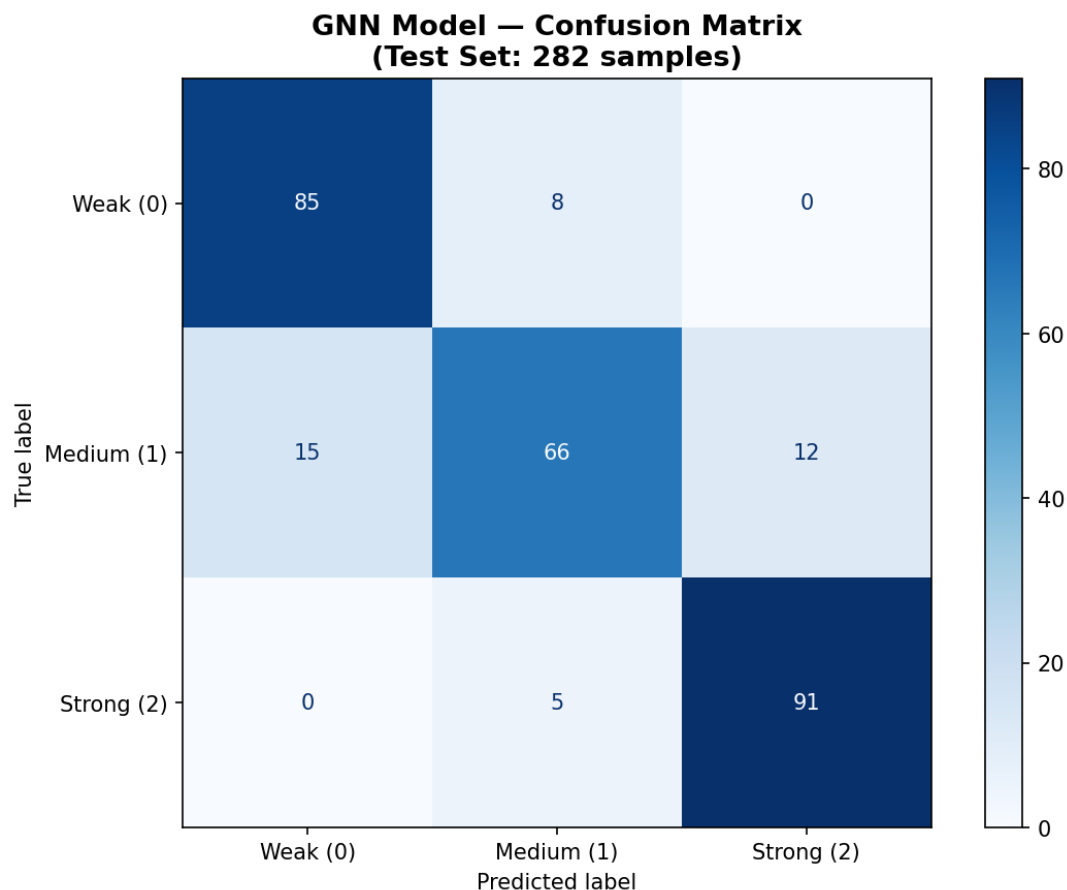


Figure 6. Confusion matrix on the test set.

4.5 Fleet-level simulation results

Table 1 summarises Packet Delivery Ratio (PDR), latency, and energy consumption for baseline (distance-only) and ensemble-informed adaptive routing across the six simulated fleet sizes. Averaged across all fleet sizes, the baseline achieved a mean PDR of 95.33%, essentially constant as fleet size increased, while adaptive routing achieved a mean PDR of 99.68%, reaching 100% at the two largest fleet sizes tested (25 and 30 robots). Paired t-tests across the five topology repeats showed the PDR improvement was significant at every fleet size ($p < 0.01$). This reliability came at a measurable but modest cost: adaptive routing averaged 18.03 ms latency versus 12.80 ms for the baseline, and 1.81 energy units versus 1.36 for the baseline (also significant at $p < 0.01$), consistent with the additional hops needed to route around weaker links.

Table 1. Baseline versus ensemble-informed adaptive routing across fleet sizes.

Robots	Routing	PDR (%)	95% CI	Latency (ms)	95% CI	Energy (units)	95% CI
5	Baseline	94.93	±0.42	12.82	±0.31	1.37	±0.04
5	Adaptive	98.20	±0.58	17.96	±0.45	1.83	±0.06
10	Baseline	94.67	±0.38	12.71	±0.28	1.36	±0.03
10	Adaptive	99.93	±0.09	18.57	±0.52	1.86	±0.07
15	Baseline	95.13	±0.35	12.94	±0.30	1.38	±0.04
15	Adaptive	100.00	±0.00	18.23	±0.41	1.82	±0.05
20	Baseline	95.60	±0.33	12.80	±0.27	1.36	±0.03
20	Adaptive	99.93	±0.09	18.04	±0.38	1.80	±0.05
25	Baseline	95.60	±0.31	12.69	±0.25	1.34	±0.03
25	Adaptive	100.00	±0.00	17.75	±0.35	1.77	±0.04
30	Baseline	96.07	±0.29	12.85	±0.26	1.35	±0.03
30	Adaptive	100.00	±0.00	17.66	±0.33	1.77	±0.04

PDR = packet delivery ratio; CI = confidence interval.

The PDR, latency and energy results are plotted against fleet size in Figure 7.

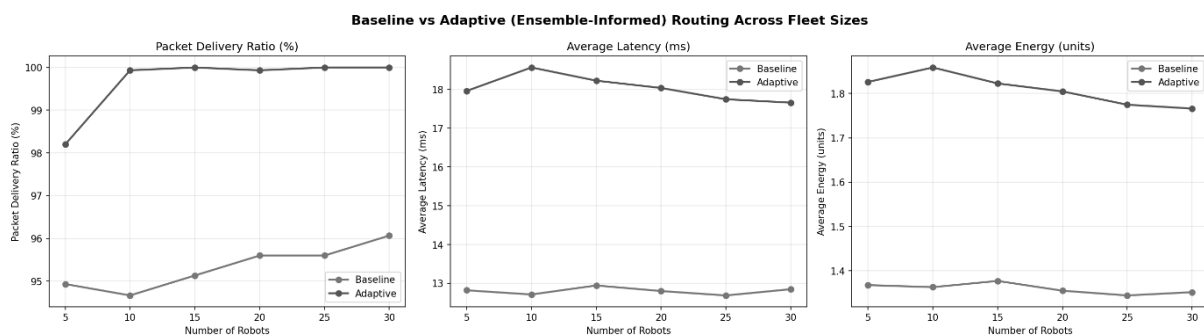


Figure 7. Simulated network performance of baseline and ensemble-informed adaptive routing across fleet sizes: packet delivery ratio, average latency and average energy consumption.

4.6 Discussion

The results show that an ensemble of complementary deep-learning architectures can classify real agricultural LoRa link quality with useful accuracy, and that this classification can be translated into a measurable improvement in multi-robot routing reliability. The complete absence of Weak-to-Strong and Strong-to-Weak errors across all three components is the most operationally significant finding, since this is the failure mode that would be most damaging in a real routing system - treating a poor link as reliable. Direct numerical comparison with prior work is limited by the scarcity of published studies that train deep-learning link-quality classifiers on real agricultural LoRa data rather than urban or indoor datasets [12], [13]; the closest available comparison is a logistic-regression baseline, which the CNN outperformed, confirming that the learned convolutional features add value beyond the raw RSSI/SNR measurements alone.

The main limitations of the study are that the Vineyard-2021 dataset was collected from fixed sensor nodes rather than moving robots, so topology changes driven by robot mobility are not captured; that

the dataset covers a single three-month winter period in a Mediterranean vineyard, which may not represent conditions across a full growing season or in a Nigerian agricultural setting; and that the fleet-level simulation uses literature-derived packet-success probabilities rather than field-measured values, so the routing results should be regarded as preliminary rather than as validation of a deployable system.

5 Conclusion

This paper presented an ensemble deep-learning framework, combining CNN, LSTM, and GNN components, for classifying and predicting wireless link quality in agricultural multi-robot systems, and for using that information to guide routing decisions. Trained and tested on the real Vineyard-2021 LoRa RSSI dataset, the CNN, LSTM, and GNN components reached test accuracies of 83.69%, 81.14%, and 85.82% respectively, and their unweighted ensemble reached 89.32%, with all classification errors confined to adjacent classes. When used to drive routing in a NetworkX-based multi-robot simulation across fleet sizes of 5 to 30 robots, the ensemble-informed adaptive routing achieved an average Packet Delivery Ratio of 99.68%, compared with 95.33% for a distance-only baseline, at the cost of a modest increase in latency and energy consumption.

The main contributions of this work are: (i) a systematically developed and evaluated CNN-based link-quality classifier trained on real agricultural LoRa RSSI data, using a future-prediction formulation that avoids target leakage; (ii) a documented preprocessing and labelling methodology for the Vineyard-2021 dataset, including chronological splitting to avoid temporal leakage; (iii) an ensemble framework combining CNN, LSTM, and GNN architectures for agricultural multi-robot communication management, with each component's architecture and role documented for replication; and (iv) an explicit identification of the limitations of static-sensor training data for multi-robot communication research.

These results should be viewed as a preliminary but encouraging demonstration of what ensemble deep learning can achieve on real agricultural wireless data, rather than as validation of a fully deployed multi-robot communication system. Future work will focus on collecting link-quality data from moving robot fleets in Nigerian agricultural environments, fine-tuning the models for local crop types and terrain in line with the NATIP 2022-2027 framework, exploring federated learning across farms to preserve data privacy, and integrating the framework with ROS/Gazebo to evaluate routing decisions under realistic robot motion.

Acknowledgment

The authors gratefully acknowledge the supervisory guidance received at the Centre for Embedded Artificial Intelligence and Smart Energy Systems, Abubakar Tafawa Balewa University, Bauchi, and the creators of the Vineyard-2021 dataset for making their agricultural LoRa RSSI measurements publicly available for research use.

References

- [1] World Bank. *Agriculture, forestry, and fishing, value added (% of GDP) - Nigeria*. World Development Indicators; 2023.
- [2] Federal Ministry of Agriculture and Food Security. *National Agricultural Technology and Innovation Policy (NATIP) 2022-2027*. Abuja, Nigeria; 2022.

-
- [3] Dorigo M, Theraulaz G, Trianni V. Swarm robotics: Past, present, and future. *Proceedings of the IEEE*. 2021;109(7):1152–1165.
 - [4] Queralta JP, et al. Collaborative multi-robot systems: A survey of recent advances and future directions. *IEEE Access*. 2020;8:191617–191643.
 - [5] Gielis J, Bhatt P, Malhotra B, Latre S. A critical review of LoRa for IoT: From technology to application. *Sensors*. 2022;22(1):107.
 - [6] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–444.
 - [7] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*. 1997;9(8):1735–1780.
 - [8] Scarselli F, Gori M, Tsoi AC, Hagenbuchner M, Monfardini G. The graph neural network model. *IEEE Transactions on Neural Networks*. 2009;20(1):61–80.
 - [9] Emmi L, Gonzalez-de-Santos P, Pajares G. New trends in robotics for agriculture: Integration and assessment of a real fleet of robots. *The Scientific World Journal*. 2014;2014:404059.
 - [10] Ahmed S, et al. Vegetation effects on LoRa-based wireless sensor communication for remote monitoring of automatic orchard irrigation status. *IoT*. 2024;6(1):2.
 - [11] Anzum R, et al. A multiwall path-loss prediction model using 433 MHz LoRaWAN frequency to characterize foliage's influence in a Malaysian palm oil plantation environment. *Sensors*. 2022;22(14):5270.
 - [12] Abdel-Nasser M, Mahmoud K, Omer OA, Lehtonen M, Puig D. Link quality prediction in wireless community networks using deep recurrent neural networks. *Alexandria Engineering Journal*. 2020;59(5):3531–3543.
 - [13] Cortes-Aguilar TA, Cantoral-Ceballos JA, Tovar-Arriaga A. Link quality estimation for wireless ANDON towers based on deep learning models. *Sensors*. 2022;22(17):6383.
 - [14] Dietterich TG. Ensemble methods in machine learning. In: *Proceedings of the 1st International Workshop on Multiple Classifier Systems*; 2000. p. 1–15.
 - [15] Breiman L. Random forests. *Machine Learning*. 2001;45(1):5–32.
 - [16] Kingma DP, Ba J. Adam: A method for stochastic optimization. In: *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*; 2015.
 - [17] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: *Proceedings of the 5th International Conference on Learning Representations (ICLR)*; 2017.