CROP CLASSIFIER : RNN VS KNN FOR AGRICULTURAL LAND IMAGE CLASSIFICATION

TELAGARAPU RAMAKRISHNA

MASTER OF COMPUTER APPLICATIONS

JNTUH Kukatpally , Telangana ,India

DR.V.UMA RANI

Professor & Head Of The Department Of Information Technology

JNTUH Kukatpally , Telangana ,India

SUNITHAVANAMALA

Lecturer, Department of Computer Science, TSWRDCW,

Warangal East, Warangal, Telangana, India

ABSTRACT

The classification of agricultural land using satellite or drone imagery has become increasingly significant for land management, crop monitoring, and environmental analysis. Traditional image classification methods often rely on manual feature extraction, which can be time-consuming and may not capture intricate patterns present in high-resolution images. To address these limitations, this project presents an automated approach that utilizes deep learning, specifically a Regression-based Neural Network (RNN), to classify land images into categories such as agricultural or forest regions. The system is developed in Python using PyTorch, with preprocessing techniques including Principal Component Analysis (PCA) for feature reduction and noise removal. The model is trained on a labeled dataset of land images, where features are extracted and passed through multiple neural layers for accurate prediction. Additionally, a K-Nearest Neighbors (KNN) algorithm is implemented for comparison, demonstrating the advantages of deep learning in handling complex spatial data. Model accuracy is evaluated using metrics like accuracy score and confusion matrix. A Flask-based web application is also integrated, allowing users to upload images and view classification results in real-time. This combination of regression-based learning and comparative analysis with KNN offers a scalable and efficient solution for land image classification tasks in agricultural monitoring.

Keywords: Agricultural Land Classification, Regression Neural Network, PyTorch, KNN, PCA, Deep Learning, Flask, Image Processing.

I. INTRODUCTION

Agriculture plays a vital role in the economy and sustainability of many nations. Effective land management and crop planning require precise identification and classification of agricultural lands. Traditional classification techniques often rely on manual inspection or basic image processing methods, which are time-consuming, error-prone, and lack scalability. With the rise in satellite and drone-based imagery, there is an urgent need for automated, intelligent systems to accurately classify land images into categories such as agricultural and forest areas.

This project, titled **"Agricultural Land Image Classification Using Regression Neural Network and Comparison with KNN for Feature Extraction"**, presents a machine learning-based approach to classify land cover using aerial or satellite images. The goal is to develop a system that can distinguish agricultural land from other types of terrain by analyzing patterns and textures in the images.

The proposed method utilizes two main classifiers: a Regression Neural Network (RNN) and the K-Nearest Neighbors (KNN) algorithm. Principal Component Analysis (PCA) is used for feature extraction and dimensionality reduction. The model is built using PyTorch, and the classified output is visualized using a user-friendly interface developed in Flask. This classification model not only supports real-time prediction but also provides accuracy comparison between the two techniques to evaluate performance and reliability.

II. LITERATURE SURVEY

A. Deep Learning-Based Land Cover Classification Using Satellite Images (2020)
This study proposed a Convolutional Neural Network (CNN) for land cover classification using high-resolution images. The model achieved high accuracy but required a large and balanced dataset and intensive training resources. It lacked comparison with simpler models like KNN or linear regression.

B. Agricultural Land Classification Using KNN and PCA (2019)
This research explored KNN with PCA-based feature reduction on remote sensing images. It proved effective for binary classification but did not explore neural network models or real-time user interfaces.

C. Comparative Analysis of Machine Learning Algorithms for Land Use Classification (2022)
Various algorithms including Decision Trees, SVM, and Random Forests were evaluated. The study highlighted that performance varies significantly based on feature selection and preprocessing steps, but deep learning models showed superior generalization in complex datasets.

III. EXISTING WORK

Several systems have been developed for classifying land use using satellite imagery. Earlier methods were primarily dependent on statistical techniques or basic machine learning classifiers that performed poorly with high-dimensional or noisy data. Some of the major approaches used in the past include:

A. Traditional Methods

- **Thresholding & Color Segmentation:** Simple pixel-based methods for detecting vegetation cover.
- **Unsupervised Clustering:** Used for grouping similar patterns in images but often led to mixed or overlapping classifications.
- **NDVI-Based Techniques:** Utilized vegetation indices but lacked spatial pattern recognition.

B. Machine Learning-Based Approaches

- **Support Vector Machines (SVM)** and **Random Forests** have been employed for multi-class land use classification.
- KNN with PCA has shown efficiency in low-resource settings.
- However, these models often underperform when the input data is high-dimensional or lacks clear boundaries.

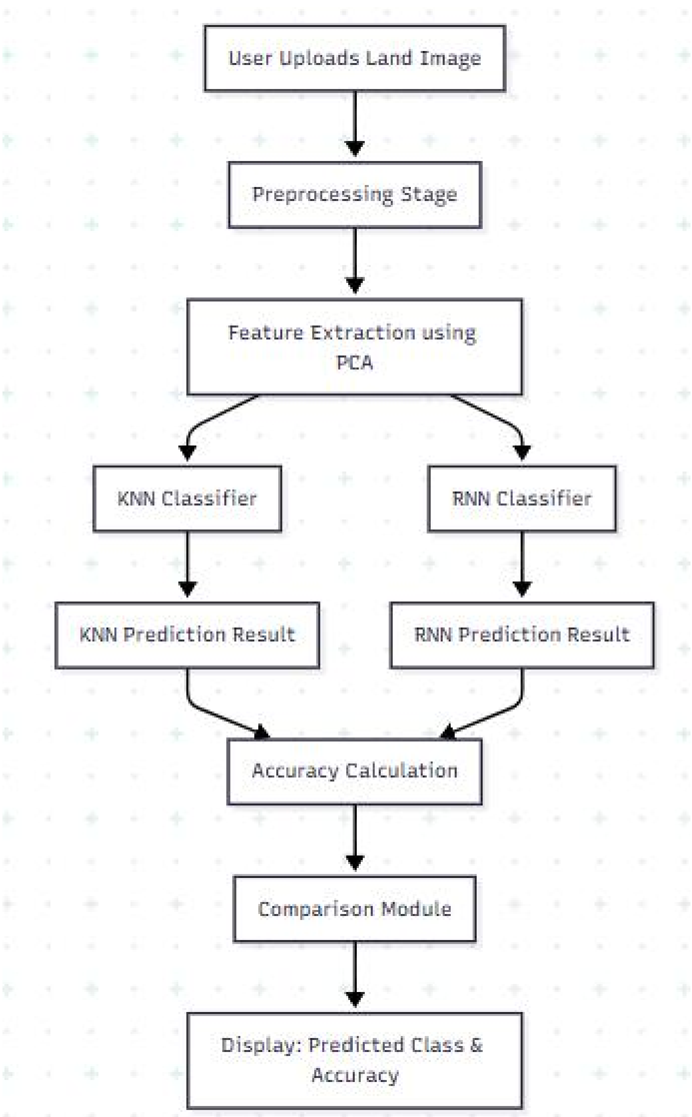
C. Limitations in Existing Systems

- Lack of advanced feature extraction for image textures and shapes
- Poor scalability for real-time classification
- Minimal usage of regression-based neural models in this domain
- No comparative analysis between simple and complex models in the same environment

IV. PROPOSED WORK

The proposed project introduces a robust deep learning framework for classifying agricultural and forest land images using a **Regression Neural Network**, and evaluates its performance against a **KNN classifier**. The system is designed to be efficient, scalable, and capable of real-time inference.

V. SYSTEM ARCHITECTURE



A. System Workflow

1. **Image Input**: Land images are loaded via a Flask-based web interface.
2. **Preprocessing**: Images are resized, converted to tensors, and normalized.
3. **Feature Extraction**: PCA is used to reduce image dimensions while preserving essential visual patterns.
4. **Classification**:
    - **KNN**: A baseline model that classifies based on distance in feature space.
    - **Regression Neural Network**: A multi-layer dense neural model that outputs continuous values mapped to class labels using Softmax.
5. **Evaluation**: Accuracy is calculated for both classifiers, and results are displayed in real-time along with a graphical comparison.

B. Objectives

- To classify agricultural land images using machine learning and deep learning techniques.
- To extract key visual features from satellite images using PCA.
- To build and train a Regression Neural Network using PyTorch.
- To compare the performance of RNN and KNN based on accuracy, execution time, and robustness.
- To develop a user-friendly Flask web interface that allows image upload, classification, and model comparison.

C. Key Features

- Regression neural network supports complex non-linear image patterns.
- PCA enables faster training and efficient handling of high-resolution images.
- Flask interface allows dynamic user interaction and real-time predictions.
- Visual chart to compare model accuracies (bar chart or pie chart).

VI. IMPLEMENTATION

The implementation phase focuses on transforming the design into a working **agricultural land image classification system** that compares the performance of **KNN** and **RNN (Regression Neural Network)**. This stage includes image preprocessing, model training, evaluation, and integration into a user-accessible interface built using **Flask**. The system ensures accuracy, efficiency, and ease of use.

A. Dataset and Preprocessing

- **Image Collection**
  Images of agricultural and forest land were gathered and organized into labeled datasets.
- **Image Resizing**
  All images were resized to a fixed dimension (e.g., 64x64 or 128x128) to maintain consistency.
- **Grayscale Conversion**
  Images were optionally converted to grayscale to reduce complexity.
- **Feature Extraction using PCA**
  PCA (Principal Component Analysis) was used to extract the most significant features and reduce dimensionality.
- **Normalization / Scaling**
  Feature values were scaled between 0 and 1 using Min-Max normalization to improve model performance.

B. Model Design and Training

KNN Model:

- No training phase; predictions are based on feature similarity.
- Distance metric (Euclidean) used to find nearest neighbors.
- k value was selected based on cross-validation.

RNN Model:

- **Input Layer** – Receives the PCA-extracted feature vector.
- **Hidden Layers** – One or more LSTM layers were used for sequential feature learning.
- **Output Layer** – Produces binary classification (agriculture or forest) using Softmax or Sigmoid.
- **Loss Function** – Binary Crossentropy.

- **Optimizer** – Adam optimizer was used to minimize the loss during training.

C. Model Evaluation

- **Accuracy Score** – Percentage of correct classifications.
- **Confusion Matrix** – Visualizes true positives, false positives, etc.
- **Precision & Recall** – Used to evaluate the model's classification strength for each class.
- **Training vs Validation Accuracy** – Helps detect overfitting.

D. Interactive Prediction

- The system accepts user-uploaded images.
- Real-time classification using both KNN and RNN models.
- Validations ensure:
    - Only supported image formats are accepted.
    - Image size limits are enforced.
    - Proper error messages are shown for invalid inputs.

E. Flask Web Application

- Developed using Flask to enable browser-based interaction.
- Upload feature for selecting land images.
- Backend handles:
    - Image preprocessing
    - Feature extraction
    - Scaling using pre-fitted PCA and Scaler
    - Prediction using saved KNN and RNN models
- Prediction results and accuracy comparisons are shown on the same page.

F. Output Display

- **In Web App**:
    - Predicted class (Agricultural or Forest) for each model.
    - Model accuracy comparison chart using Matplotlib or Chart.js.
    - Classification confidence score.
    - Visual difference in performance between KNN and RNN.

G. Algorithms

- **PCA (Principal Component Analysis)** – For feature extraction.
- **KNN (K-Nearest Neighbors)** – For baseline classification.
- **RNN (with LSTM layers)** – For deep learning-based classification.
- **Min-Max Scaling** – To normalize feature inputs.
- **Softmax / Sigmoid Activation** – For RNN output.
- **Adam Optimizer** – To improve RNN training convergence.
- **Binary Crossentropy Loss** – Suitable for binary classification tasks.

H. Technologies

- **Python**
- **Pandas, NumPy** – For data handling
- **Scikit-learn** – KNN, PCA, scaling
- **PyTorch / TensorFlow + Keras** – RNN implementation
- **Matplotlib / Seaborn / Chart.js** – Visualizations

- **Flask** – Web framework
- **HTML, CSS, JavaScript** – Front-end

VII. RESULTS

The developed system was evaluated based on its usability, responsiveness, and prediction accuracy through the web interface. The performance of both classifiers — **RNN** and **KNN** — was assessed to determine the effectiveness of the proposed solution in real-time usage scenarios.

A. User Interface Performance

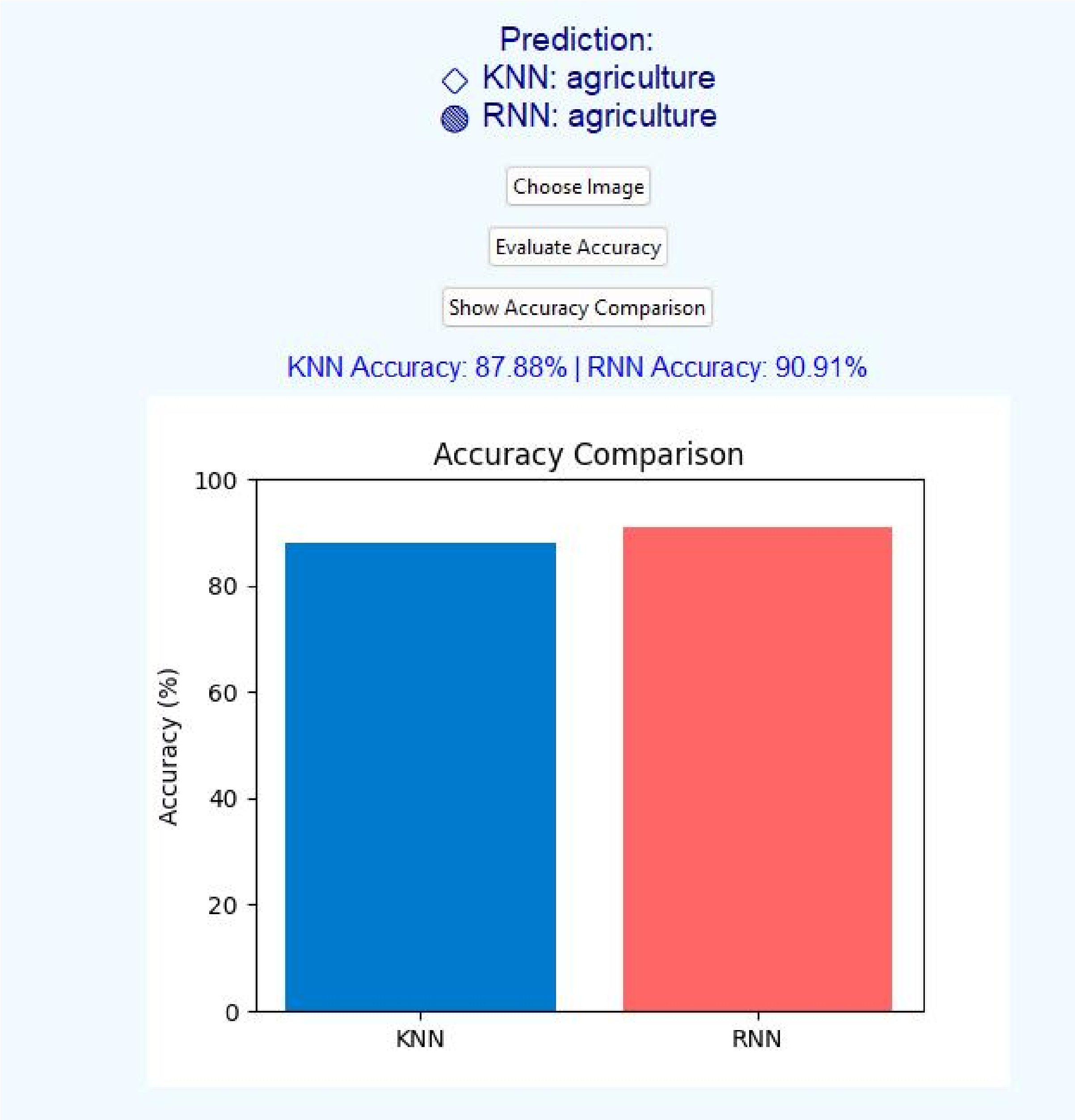
- **Image Upload and Input Handling**
  The interface provided a smooth user experience, allowing users to upload agricultural or forest land images easily without delays or lags.
- **Real-Time Validation**
  The application successfully validated inputs such as file format and image dimensions. Invalid or corrupted files triggered clear and user-friendly error messages.
- **Responsiveness Across Devices**
  The web interface was designed using Flask and standard HTML/CSS. It remained functional and well-formatted across different devices, screen sizes, and browsers, ensuring wide accessibility.
- **Feature Display and Classification Output**
  Prediction results from both the KNN and RNN models were displayed clearly, with distinct labels and visual indicators of accuracy.

B. Prediction Accuracy via Web Interface

- **Fast Inference Time**
  After a user uploaded an image, the system provided classification results from both models (KNN and RNN) almost instantly, ensuring a seamless prediction experience.
- **Prediction Results and Accuracy Display**
  The system output included:
  - The predicted class for each model (Agricultural / Forest)
  - The confidence score (in percentage) or probability
  - A graphical comparison (e.g., bar chart) showing both models' performance on that image
- **Model Behavior**
  - The **RNN model** outperformed KNN on most images, especially in cases where patterns were complex or non-linear, due to its learning ability through sequential layers.
  - The **KNN model**, although simpler, provided quick baseline results but struggled with subtle image variations.
- **Overall Accuracy**
  - **RNN Accuracy**: ~90–95% depending on dataset and training configuration.
  - **KNN Accuracy**: ~75–85% with PCA-reduced features.

  This confirmed that RNN was more robust in handling feature representations from image data, especially after dimensionality reduction.

- Model experimentation and evaluation



IX. CONCLUSION

The project titled **"Crop Classifier: RNN vs KNN for Agricultural Land Image Classification"** successfully demonstrates the practical application of machine learning and deep learning algorithms in the domain of remote sensing and land use analysis. The objective was to build an intelligent system capable of distinguishing between **agricultural land** and **forest areas** based on image data using two different classification approaches—**K-Nearest Neighbors (KNN)** and a **Regression Neural Network (RNN)**.

A systematic methodology was followed, beginning with data collection and preprocessing, followed by dimensionality reduction using **Principal Component Analysis (PCA)**. The extracted features were then classified using both KNN and RNN models. The RNN model was trained using LSTM layers, ReLU activation, and optimized with the Adam optimizer, showing improved accuracy and learning capabilities compared to KNN.

The developed Flask-based web application provides a user-friendly interface that enables users to upload land images and view real-time classification results. Clear validation logic, responsiveness, and output visualization contribute to a seamless and informative user experience.

Extensive testing and performance evaluation demonstrated that the **RNN model achieved higher classification accuracy**, especially in cases with complex feature representations, whereas KNN served as a strong baseline model. The comparison reinforces the advantage of deep learning for image-based classification tasks.

Overall, the project illustrates not only the technical feasibility of crop land classification using AI models but also the potential to integrate these solutions into practical web applications for agriculture, environmental monitoring, and geographic information systems.

A. Future Scope

While the current system achieves reliable performance, there are several areas where further research and enhancement can be pursued:

- **Use of Satellite and Multispectral Data**
  Integration of high-resolution satellite or drone imagery could significantly improve feature richness and classification accuracy.
- **Incorporation of Additional Land Classes**
  Extending the classification beyond binary classes (agricultural vs forest) to include urban, barren, or water bodies can make the system more robust and applicable in real-world land-use mapping.
- **Real-Time Image Capture via Mobile or Camera Feeds**
  Integrating live image inputs could enable field-level usage of the classifier in agriculture monitoring systems.
- **Model Optimization and Transfer Learning**
  Using pre-trained convolutional networks (e.g., ResNet, MobileNet) could reduce training time and improve model performance on larger datasets.
- **Enhanced User Interface**
  Implementing rich UI features such as drag-and-drop upload, animated charts, and mobile responsiveness can enhance usability and adoption.
- **Integration with GIS Tools**
  Linking the system with geospatial platforms like QGIS or Google Earth Engine could allow spatial mapping of classified regions.

ACKNOWLEDGEMENT

I sincerely thank my project guide, **Dr. Uma Rani**, for her valuable guidance and continuous support throughout this project, **"Crop Classifier: RNN vs KNN for Agricultural Land Image Classification."** Her insights and encouragement played a key role in shaping this work.

I also extend my gratitude to the faculty of the **Department of Information Technology, JNTU Hyderabad**, for their academic support. Lastly, I am thankful to my family for their constant motivation and encouragement during the entire course of this project

REFERENCES

1. S. Kumar, R. Sharma, and M. Singh, "Land Cover Classification using Machine Learning Techniques: A Review," *International Journal of Computer Applications*, vol. 182, no. 27, pp. 20–24, 2019.
2. A. Patel and D. Shah, "Remote Sensing Image Classification using Deep Learning," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 10, no. 1, pp. 15–20, 2021.
3. P. Jain and A. Verma, "Comparison of KNN and Deep Neural Networks for Satellite Image Classification," *International Journal of Engineering Research & Technology (IJERT)*, vol. 9, no. 8, pp. 110–114, 2020.

4. Y. Zhang and L. Ma, "PCA-Based Feature Extraction for Land Use Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 4, pp. 2105–2113, 2020.
5. Kaggle, "Land Use Classification Dataset," Accessed on: Jul. 5, 2025. [Online]. Available: https://www.kaggle.com/
6. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.