Journal of Cognitive Human-Computer Interaction

♦ ISSN: 2771-1463 Vol: 09, No. 01, 2025 ♦ PP. 01–06

Article DOI: https://doi.org/10.54216/JCHCI.090101

FULL LENGTH ARTICLE

Intelligent Remote Sensing Scene Classification Model for On-Board Training of Resource-Constrained Devices

Ahmad Khaldi[1,*] **Josef Al Jumayel**[2]

[1] Mutah University, Faculty of Science, Jordan

[2] Faculty of Science, Beirut Arab University, Beirut, Lebanon

Emails: khaldiahmad1221@gmail.com · Josefjumayel113@gmail.com

Received: October 19, 2024 Revised: December 05, 2024 Accepted: January 10, 2025 ⋆ Corresponding author

ABSTRACT

Remote Sensing Scene Classification (RSSC) assigns remote sensing images to scene classes according to image content and is important for land mapping, agriculture, disaster analysis, and resource monitoring. Satellite imagery is frequently degraded by haze, fog, aerosol dispersion, noise, and limited on-board computational resources. This paper presents a Leveraging Tiny Convolutional Neural Networks with a Water Cycle Algorithm for Remote Sensing Scene Classification (LTCNN-WCRSSC) model. The method first applies median filtering for image preprocessing, then extracts compact and discriminative representations using ConvNeXt-Tiny enhanced with efficient channel attention and label smoothing. A spatiotemporal attention bidirectional long short-term memory (STA-BiLSTM) classifier is used for scene recognition, while the Water Cycle Algorithm (WCA) selects hyperparameters to improve classification performance. Experiments on the UC-Merced and EuroSAT datasets show that the proposed LTCNN-WCRSSC model obtains strong accuracy, precision, recall, and F1-score compared with several lightweight and deep learning baselines.

Keywords: ConvNeXt-Tiny ▪ Remote sensing scene classification ▪ Water Cycle Algorithm ▪ Resource-constrained devices ▪ Image preprocessing

1. INTRODUCTION

Rapid advances in communication, collaborative technologies, robotics, the Internet of Things (IoT), drones, artificial intelligence, edge/cloud computing, aerial access networks, and big data analytics have enabled many smart-city and remote sensing applications [1, 2, 3]. Aerial platforms, satellites, unmanned vehicles, and sensor systems improve the collection of real-time observations and support intelligent security, monitoring, and emergency response services [4].

Remote sensing imaging (RSI) has developed substantially, with higher spatial resolution and richer semantic information. RSI is valuable in agricultural production, national defense, disaster warning, environmental monitoring, and land-use interpretation. RSSC remains challenging because scene categories can be visually similar, intra-class variation can be large, and practical applications require accurate and efficient processing of large-scale images on resource-constrained devices [5, 6].

Deep learning (DL) has attracted attention in the remote sensing community because it can learn hierarchical feature representations and often outperforms handcrafted approaches [7, 8]. However, deploying heavy models directly on satellites, drones, or IoT devices is difficult because of memory, computation, and energy constraints. This motivates lightweight RSSC systems that combine compact convolutional backbones, robust temporal or attention-based classification, and meta-heuristic hyperparameter tuning.

This study develops the LTCNN-WCRSSC model for efficient remote sensing scene classification with on-board training capability. The model integrates median filtering, ConvNeXt-Tiny feature extraction, STA-BiLSTM classification, and WCA-based hyperparameter selection. The main workflow is shown in Fig. 1.

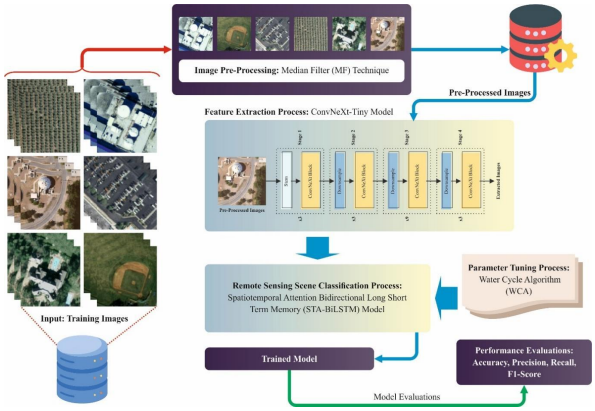

Figure 1. Overall workflow of the LTCNN-WCRSSC model.

2. RELATED WORK

Recent RSSC research has explored deep convolutional networks, knowledge distillation, lightweight architectures, few-shot learning, and on-board image interpretation. CNN-based scene classification can transfer learned image representations to high-resolution remote sensing data [7]. Data augmentation and multilayer feature integration have also improved the discriminative ability of RSSC systems [8, 5].

Resource-aware remote sensing is increasingly important for satellites, UAVs, and edge IoT devices. Prior studies considered onboard autonomy, aerial object detection, emergency monitoring, compact neural networks, and satellite image classification using pretrained vision models [1, 2, 4, 9]. Lightweight change detection and compact multimodal models further demonstrate the need for small yet accurate models in constrained deployments [10, 11].

ConvNeXt-Tiny provides a compact convolutional backbone with modern design choices suitable for efficient image representation. Attention mechanisms can enhance feature selection by emphasizing informative channels or locations, while label smoothing reduces over-confidence during classification. Recurrent and attention-based classifiers such as STA-BiLSTM can model contextual relations in the extracted feature sequence. Finally, meta-heuristic optimization methods such as WCA can tune hyperparameters without exhaustive search [12, 13, 14].

3. PROPOSED METHODOLOGY

The proposed LTCNN-WCRSSC framework contains four main stages: image preprocessing, ConvNeXt-Tiny feature extraction, STA-BiLSTM based classification, and WCA-based hyperparameter tuning.

3.1 Image Preprocessing

The input remote sensing image is first processed using a median filter (MF), which reduces impulsive noise while preserving local edges better than simple averaging. For a pixel location (x,y) and neighbourhood Ω, the filtered image can be written as

$$I_{\mathrm{MF}}(x,y) = \mathrm{median}\{I(i,j) : (i,j) \in \Omega_{x,y}\}. \tag{1}$$

This stage improves the robustness of subsequent feature extraction under degraded imaging conditions.

3.2 ConvNeXt-Tiny Feature Extraction

The preprocessed images are passed to a ConvNeXt-Tiny backbone to extract compact features. The method uses efficient channel attention (ECA) to recalibrate feature channels without heavy fully connected layers. The ECA operation is expressed as

$$F_{\mathrm{ECA}}(x) = \sigma\left(\mathrm{Conv1D}_k(\mathrm{GAP}(x))\right) \otimes x, \tag{2}$$

where $\mathrm{GAP}(\cdot)$ is global average pooling, σ is the sigmoid function, k is the one-dimensional convolution kernel size, and $\otimes$ denotes element-wise multiplication. The kernel size is adaptively determined by

$$k = \psi(C) = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right|_{\mathrm{odd}}, \tag{3}$$

where C denotes the number of channels and $|\cdot|_{\mathrm{odd}}$ maps the value to the nearest odd integer. Label smoothing regularization (LSR) is also used:

$$y_i = \begin{cases} 1-\varepsilon, & i = c, \\ \dfrac{\varepsilon}{k-1}, & i \neq c. \end{cases} \tag{4}$$

3.3 STA-BiLSTM Classification

For RSSC, the feature sequence is classified using an STA-BiLSTM model. The LSTM cell uses memory and gates to preserve long-range dependency and reduce gradient vanishing or explosion. The main equations are

$$f_t = \sigma\left(W_f^T s_{t-1} + U_f^T x_t + b_f\right), \tag{5}$$
$$i_t = \sigma\left(W_i^T s_{t-1} + U_i^T x_t + b_i\right), \tag{6}$$
$$C_t' = \tanh\left(W_c^T s_{t-1} + U_c^T x_t + b_c\right), \tag{7}$$
$$C_t = f_t \times C_{t-1} + i_t \times C_t', \tag{8}$$
$$O_t = \sigma\left(W_o^T s_{t-1} + U_o^T x_t + b_o\right), \tag{9}$$
$$s_t = O_t \times \tanh(C_t), \tag{10}$$
$$\min J(\theta) = \sum_{t=1}^{T} \mathrm{loss}\left(\hat{y}^{(t)}, y^{(t)}\right). \tag{11}$$

BiLSTM processes features in forward and backward directions, while spatial and temporal attention assign greater weights to informative features and hidden states. The STA-BiLSTM structure is illustrated in Fig. 2.

3.4 WCA-Based Hyperparameter Tuning

The WCA is used to optimize the hyperparameters of the STA-BiLSTM classifier. Candidate solutions are represented as raindrops. The best solution is considered the sea, selected high-quality solutions are rivers, and the remaining solutions are streams. Streams flow toward rivers and the sea, while evaporation and raining operations help avoid local optima. This optimization process improves the final classification

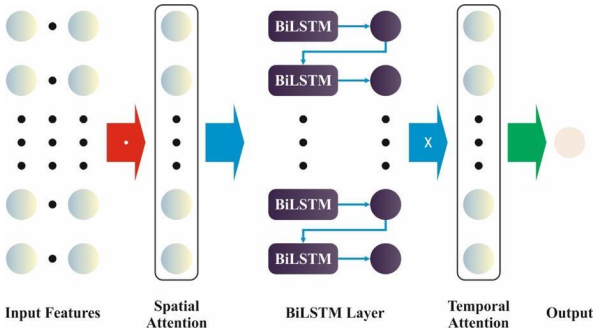

Figure 2. Structure of the STA-BiLSTM model.

results without exhaustive manual tuning.

4. EXPERIMENTAL SETUP

The experimental evaluation uses the UC-Merced land-use dataset and the EuroSAT dataset. The UC-Merced dataset contains 21 land-use classes, while EuroSAT contains 10 land-cover classes. Sample images from the datasets are shown in Fig. 3. Data availability is reported through the public dataset links in the original study.

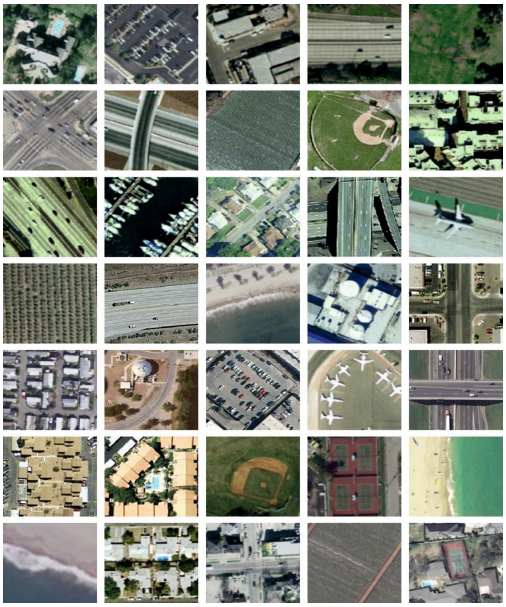

Figure 3. Sample images from the benchmark remote sensing datasets.

Performance is measured using accuracy, precision, recall, and F1-score:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \tag{12}$$

$$\text{Precision} = \frac{TP}{TP+FP}, \tag{13}$$

$$\text{Recall} = \frac{TP}{TP+FN}, \tag{14}$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{15}$$

5. RESULTS AND DISCUSSION

The classification behaviour of the proposed model is evaluated using confusion matrices, precision–recall curves, ROC curves, training/validation accuracy, training/validation loss, and comparative analysis with existing models.

Fig. 8 shows the classification results for the EuroSAT dataset,

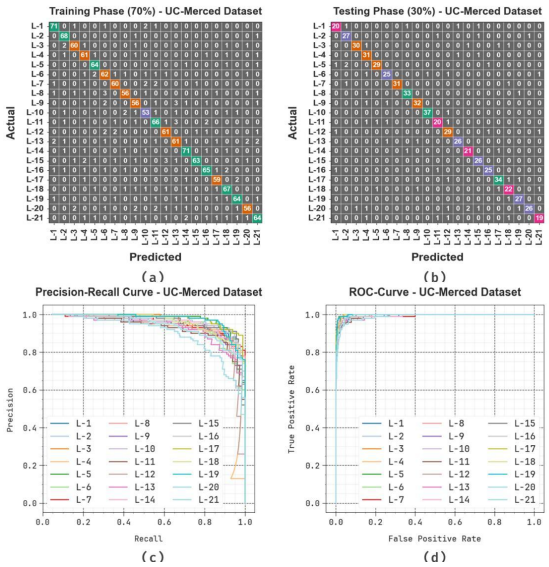

Figure 4. UC-Merced dataset: confusion matrices, precision–recall curve, and ROC curve.

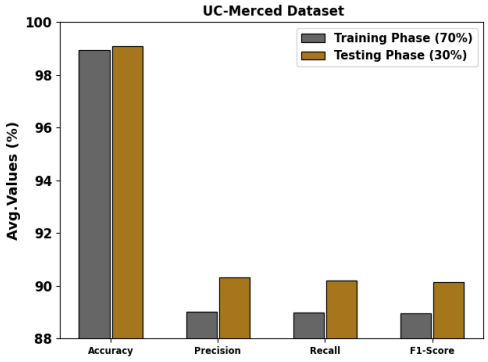

Figure 5. Average performance of the LTCNN-WCRSSC technique on the UC-Merced database.

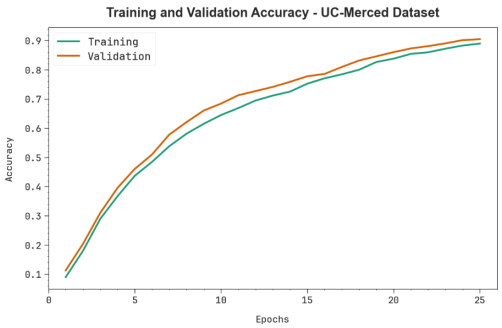

Figure 6. Accuracy curve of the LTCNN-WCRSSC model on the UC-Merced database.

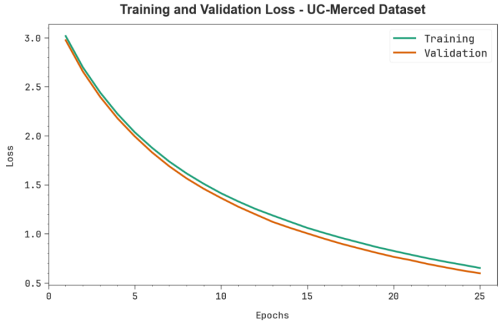

Figure 7. Loss curve of the LTCNN-WCRSSC model on the UC-Merced database.

Table 1. RSSC detection outcomes of LTCNN-WCRSSC on the UC-Merced database.

70% TRAPH

Class	Acc.	Prec.	Rec.	F1
L-1	99.25	92.21	93.42	92.81
L-2	99.32	89.47	97.14	93.15
L-3	99.25	92.31	90.91	91.60
L-4	98.78	83.56	91.04	87.14
L-5	99.32	90.14	95.52	92.75
L-6	98.64	84.93	87.32	86.11
L-7	99.05	89.55	89.55	89.55
L-8	98.98	90.32	86.15	88.19
L-9	98.64	84.85	84.85	84.85
L-10	98.78	85.48	85.48	85.48
L-11	98.64	86.84	86.84	86.84
L-12	98.98	88.41	89.71	89.05
L-13	98.57	85.92	84.72	85.31
L-14	98.98	88.75	92.21	90.45
L-15	99.12	96.92	85.14	90.65
L-16	98.98	90.28	89.04	89.66
L-17	99.32	90.77	93.65	92.19
L-18	98.91	88.16	90.54	89.33
L-19	99.12	92.75	88.89	90.78
L-20	98.84	90.32	83.58	86.82
L-21	98.50	87.67	83.12	85.33
Average	98.95	89.03	88.99	88.96

30% TESPH

Class	Acc.	Prec.	Rec.	F1
L-1	98.89	86.96	83.33	85.11
L-2	98.89	87.10	90.00	88.52
L-3	99.21	96.77	88.24	92.31
L-4	98.57	81.58	93.94	87.32
L-5	99.05	93.55	87.88	90.62
L-6	98.89	89.29	86.21	87.72
L-7	99.68	100.00	93.94	96.88
L-8	99.37	94.29	94.29	94.29
L-9	99.37	94.12	94.12	94.12
L-10	99.68	97.37	97.37	97.37
L-11	98.89	86.96	83.33	85.11
L-12	98.57	82.86	90.62	86.57
L-13	99.37	92.86	92.86	92.86
L-14	99.21	87.50	91.30	89.36
L-15	99.37	86.67	100.00	92.86
L-16	99.37	92.59	92.59	92.59
L-17	99.05	91.89	91.89	91.89
L-18	98.57	81.48	84.62	83.02
L-19	99.37	90.00	96.43	93.10
L-20	98.73	96.30	78.79	86.67
L-21	98.89	86.36	82.61	84.44
Average	99.09	90.31	90.21	90.13

including confusion matrices and curve-based evaluation.

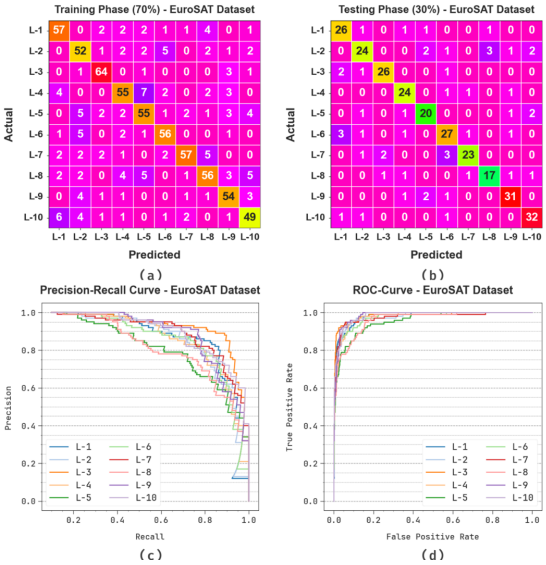

Figure 8. EuroSAT dataset: confusion matrices, precision–recall curve, and ROC curve.

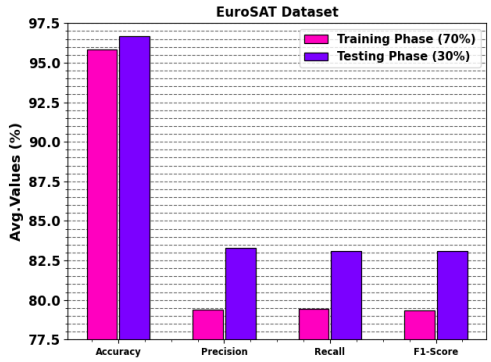

Figure 9. Average performance of the LTCNN-WCRSSC model on the EuroSAT database.

The comparative results show that the proposed method outperforms EfficientNet, MobileNet, ResNet-50, ShuffleNet, TGRRS-DK, DLB-LULCC, and RSSC-DBOEDL on the reported metrics for both benchmark datasets.

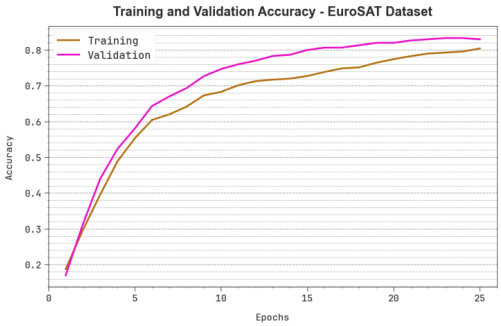

Figure 10. Accuracy curve of the LTCNN-WCRSSC model on the EuroSAT database.

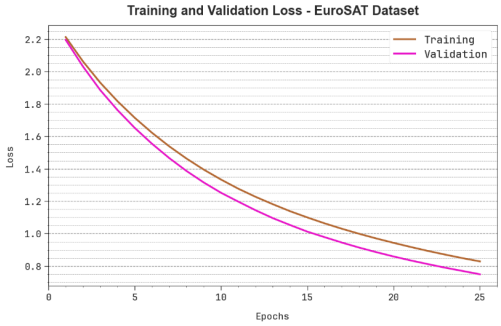

Figure 11. Loss curve of the LTCNN-WCRSSC model on the EuroSAT database.

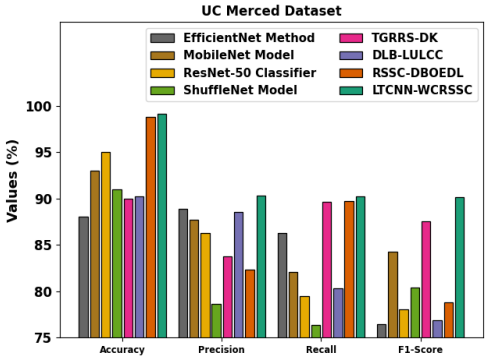

Figure 12. Comparative analysis of the LTCNN-WCRSSC model on the UC-Merced database.

Table 2. RSSC detection outcomes of LTCNN-WCRSSC on the EuroSAT dataset.

70% TRAPH

Class	Acc.	Prec.	Rec.	F1
L-1	96.00	79.17	81.43	80.28
L-2	94.57	69.33	77.61	73.24
L-3	97.86	87.67	91.43	89.51
L-4	95.29	78.57	75.34	76.92
L-5	94.57	75.34	73.33	74.32
L-6	96.86	82.35	84.85	83.58
L-7	97.00	89.06	80.28	84.44
L-8	94.71	78.87	71.79	75.17
L-9	96.29	78.26	83.08	80.60
L-10	95.43	75.38	75.38	75.38
Average	95.86	79.40	79.45	79.35

30% TESPH

Class	Acc.	Prec.	Rec.	F1
L-1	96.67	81.25	86.67	83.87
L-2	95.00	80.00	72.73	76.19
L-3	97.67	89.66	86.67	88.14
L-4	97.33	82.76	88.89	85.71
L-5	96.67	80.00	80.00	80.00
L-6	95.00	77.14	79.41	78.26
L-7	97.00	88.46	79.31	83.64
L-8	97.00	80.95	77.27	79.07
L-9	97.00	86.11	88.57	87.32
L-10	97.33	86.49	91.43	88.89
Average	96.67	83.28	83.09	83.11

Table 3. Comparative analysis of LTCNN-WCRSSC on the UC-Merced database.

Method	Accuracy	Precision	Recall	F1-score
EfficientNet	88.00	88.87	86.25	76.42
MobileNet	93.00	87.67	82.07	84.28
ResNet-50	95.00	86.31	79.44	78.07
ShuffleNet	91.00	78.60	76.36	80.37
TGRRS-DK	89.99	83.74	89.62	87.51
DLB-LULCC	90.25	88.54	80.29	76.86
RSSC-DBOEDL	98.75	82.31	89.68	78.81
LTCNN-WCRSSC	99.09	90.31	90.21	90.13

Table 4. Comparative analysis of LTCNN-WCRSSC on the EuroSAT dataset.

Method	Accuracy	Precision	Recall	F1-score
EfficientNet	85.23	75.52	73.73	82.06
MobileNet	87.52	82.19	77.91	81.48
ResNet-50	90.34	70.33	71.13	75.05
ShuffleNet	88.68	72.63	70.66	77.38
TGRRS-DK	85.12	73.74	82.97	74.68
DLB-LULCC	90.55	74.56	78.55	71.77
RSSC-DBOEDL	95.07	73.18	79.20	82.05
LTCNN-WCRSSC	96.67	83.28	83.09	83.11

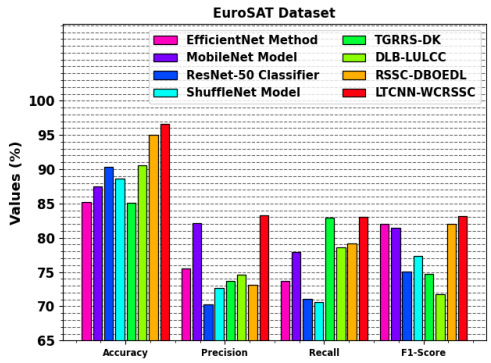

Figure 13. Comparative analysis of the LTCNN-WCRSSC approach on the EuroSAT database.

6. CONCLUSION

This work designed and developed the LTCNN-WCRSSC algorithm for efficient remote sensing scene classification in resource-constrained devices with on-board training capabilities. The method combines median-filter preprocessing, ConvNeXt-Tiny feature extraction, STA-BiLSTM classification, and WCA-based hyperparameter tuning. Experimental evaluation on benchmark image datasets indicates that the proposed model achieves superior performance compared with existing approaches. Future work can extend the model toward stronger real-time deployment, lower memory consumption, and additional remote sensing datasets.

DATA AVAILABILITY STATEMENT

The datasets supporting the findings of this study are openly available from the UC-Merced land-use dataset and the EuroSAT dataset.

REFERENCES

[1] A. Kothandhapani and V. Vatsal, "Methods to leverage onboard autonomy in remote sensing," in *2nd National Conference on Small Satellite Technology and Applications*, 2020.

[2] P. Aposporis, "Object detection methods for improving UAV autonomy and remote sensing applications," in *IEEE/ACM ASONAM*, pp. 845–853, 2020.

[3] J. J. Yebes, D. Montero, and I. Arriola, "Learning to automatically catch potholes in worldwide road scene images," *IEEE Intelligent Transportation Systems Magazine*, vol. 13, no. 3, pp. 192–205, 2020.

[4] C. Kyrkou and T. Theocharides, "EmergencyNet: Efficient aerial image classification for drone-based emergency monitoring using atrous convolutional feature fusion," *IEEE JSTARS*, vol. 13, pp. 1687–1699, 2020.

[5] E. Li, J. Xia, P. Du, C. Lin, and A. Samat, "Integrating multilayer features of convolutional neural networks for remote sensing scene classification," *IEEE TGRS*, vol. 55, no. 10, pp. 5653–5665, 2017.

[6] B. Sudharsan, J. G. Breslin, and M. I. Ali, "Edge2train: A framework to train machine learning models on resource-constrained IoT edge devices," in *Proceedings of the 10th International Conference on the Internet of Things*, pp. 1–8, 2020.

[7] F. Hu, G. S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for scene classification of high-resolution remote sensing imagery," *Remote Sensing*, vol. 7, no. 11, pp. 14680–14707, 2015.

[8] X. Yu, X. Wu, C. Luo, and P. Ren, "Deep learning in remote sensing scene classification: A data augmentation enhanced convolutional neural network framework," *GIScience & Remote Sensing*, vol. 54, no. 5, pp. 741–758, 2017.

[9] T. D. Le et al., "On-board satellite image classification for earth observation: A comparative study of pre-trained vision transformer models," *arXiv preprint arXiv:2409.03901*, 2024.

[10] H. A. Rashid et al., "TinyVQA: Compact multimodal deep neural network for visual question answering on resource-constrained devices," *arXiv preprint arXiv:2404.03574*, 2024.

[11] G. Wang et al., "Knowledge distillation-based lightweight change detection in high-resolution remote sensing imagery for on-board processing," *IEEE JSTARS*, 2024.

[12] R. A. Perdana and A. M. Arimurthy, "Remote sensing scene classification using ConvNeXt-Tiny model with attention mechanism and label smoothing," *Jurnal RESTI*, vol. 8, no. 3, pp. 389–400, 2024.

[13] S. Mu, B. Liu, J. Gu, C. Lien, and N. Nadia, "Research on stock index prediction based on the spatiotemporal attention BiLSTM model," *Mathematics*, vol. 12, no. 18, p. 2812, 2024.

[14] S. C. Das, F. Akhtar, A. F. Alrasheedi, and A. A. Shaikh, "Application of water cycle algorithm with demand follows green level and nonlinear power pattern of the product for an inventory system," *Scientific Reports*, vol. 14, no. 1, p. 20995, 2024.

[15] UC-Merced land-use dataset, http://weegee.vision.ucmerced.edu/datasets/landuse.html.

[16] EuroSAT dataset, https://www.kaggle.com/datasets/apollo2506/eurosat-dataset.

[17] M. Alamgeer et al., "Improving remote sensing scene classification using dung beetle optimization with enhanced deep learning approach," *Heliyon*, vol. 10, no. 18, 2024.

[18] A. Ghanbarzadeh and H. Soleimani, "Self-supervised in-domain representation learning for remote sensing image scene classification," *Heliyon*, 2023.