IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Seeing Beyond Pixesls: Ensemble Teacher-Student Cnn Approch For Satellite Image Classification

Mayuresh Bhakare, Tanaya Naik, Sanika Kalyankar, Tejal Gurav, Shashank S. Tolye

Student, Student, Student, Student, Asst. Professor

Information Technology

Finolex Academy of Management and Technology, Ratnagiri

Abstract: The classification of satellite images is vital in remote sensing for tasks such as mapping land cover, monitoring disasters, and conducting environmental assessments. Nonetheless, deep learning models frequently encounter substantial computational requirements, which hinder their use in real-time settings. To mitigate this issue, we introduce an Embedded Teacher-Student CNN framework that utilizes knowledge distillation for more efficient classification of satellite images. ResNet50 and VGG16 act as teacher models, capturing spatial features and producing soft labels to train a lightweight EfficientNet-B0 student model. Experimental results demonstrate high classification accuracy with reduced inference time**,** making the model suitable for real-time and resource-constrained applications.

***Index Terms* -** Satellite Image Classification, Knowledge Distillation, Teacher-Student Learning, Convolutional Neural Networks (CNN), ResNet50, VGG16, EfficientNet-B0, Hybrid Loss Function, Remote Sensing, Computational Efficiency.

I.INTRODUCTION

Satellite image classification serves as a crucial component in remote sensing, facilitating various applications like monitoring land usage, assessing disasters, and managing the environment (Chen et al., 2020). Conventional deep learning architectures, particularly Convolutional Neural Networks (CNNs), have proven highly effective in classifying high-resolution satellite imagery (Zhu et al., 2017). Nonetheless, such models frequently incur significant computational expenses, posing challenges for deployment in resource-limited settings (Tan & Le, 2019).

To overcome this issue, we introduce an Embedded Teacher-Student CNN framework that utilizes knowledge distillation to provide efficient and accurate satellite image classification. Our framework incorporates ResNet50 and VGG16 as teacher models to extract deep hierarchical and spatial features from satellite images (He et al., 2016; Simonyan & Zisserman, 2015). The knowledge gained from these teacher models is subsequently transferred to a lightweight EfficientNet B0 student model, utilizing a distillation loss function grounded in KL Divergence and Categorical Cross-Entropy (Hinton et al., 2015). This strategy guarantees that the student model achieves excellent classification results while preserving computational efficiency. The proposed approach enhances existing models by merging feature-rich teacher models with an optimized student model, thereby reducing inference time without sacrificing classification accuracy. This makes it particularly suitable for real-time remote sensing applications where computational resources are constrained. Our experimental findings indicate that the proposed model achieves competitive accuracy, positioning it as a viable option for practical satellite image classification tasks.

II. LITERATURE SURVEY

The classification of satellite images has become an essential area of research because of its relevance to environmental monitoring, urban development, and disaster management. Conventional classification techniques often depended on manually crafted features, which frequently lacked adaptability and robustness. In contrast, deep learning methods have greatly enhanced classification accuracy by utilizing automated feature extraction through convolutional neural networks (CNNs). Among the leading CNN architectures, ResNet and VGG have shown remarkable effectiveness in feature learning. ResNet, introduced by He et al., transformed deep learning by tackling the vanishing gradient issue with residual connections, facilitating the training of deeper networks. This model has been extensively applied in remote sensing tasks due to its capability of capturing intricate hierarchical features. Likewise, VGG, developed by Simonyan and Zisserman, uses small convolutional kernels within a deep structure, proving to be highly efficient for spatial feature extraction. Despite being computationally demanding, VGG has been widely implemented in satellite image analysis because of its excellent classification performance. To enhance efficiency and minimize computational burdens, knowledge distillation has emerged as a robust strategy. Hinton et al. conceptualized knowledge distillation, wherein a compact student model learns from a larger, more sophisticated teacher model by replicating its soft probability distribution. This method ensures that the student model retains the teacher's accuracy while achieving computational efficiency. EfficientNet, created by Tan and Le, represents an optimized CNN architecture that achieves a balance between accuracy and efficiency through compound scaling. Specifically, EfficientNet-B0 offers a lightweight yet powerful option for classification tasks, making it well-suited for practical satellite image processing applications. The combination of hybrid teacher-student models has become increasingly popular in remote sensing applications. By merging various teacher models, such as ResNet50 and VGG16, a broader range of feature representations can be acquired, resulting in improved classification accuracy. The student model, EfficientNet-B0, absorbs this knowledge while ensuring computational efficiency, making it perfect for extensive satellite image classification. This review of the literature lays the groundwork for our proposed method, which capitalizes on the advantages of ResNet50, VGG16, and EfficientNet-B0 to achieve high-quality classification of satellite images.

III. Methodology

The suggested approach incorporates a combined teacher-student learning framework aimed at satellite image classification. This system employs ResNet50 and VGG16 as teacher models to extract features and facilitate knowledge distillation, while EfficientNet-B0 acts as the student model to improve computational efficiency and enhance classification accuracy. The main steps of the methodology are outlined below:

3.1 Data Preprocessing

Satellite images are gathered and resized to a standard dimension of 128×128×3 for consistency. Various data augmentation methods, including rotation, flipping, and normalization, are utilized to enhance generalization. The dataset is divided into 80% for training, 10% for validation, and 10% for testing.

3.2 Teacher Models: ResNet50 and VGG16

ResNet50 is employed because its deep residual connections improve feature learning and help mitigate the vanishing gradient problem. VGG16, recognized for its straightforward yet profound architecture, is effective in extracting spatial features from satellite imagery. Both of these models have been pretrained on extensive datasets (such as ImageNet) and then fine-tuned using the satellite imagery dataset. These models produce soft labels (probability distributions) rather than hard labels.

3.3 Knowledge Distillation

The process of knowledge distillation involves transferring information from the teacher models (ResNet50 and VGG16) to the student model (EfficientNet-B0). A weighted methodology is employed to combine soft labels from both teacher models, with a distribution of 60% from ResNet50 and 40% from VGG16. The student model gains insights from a mixture of soft labels and true labels, facilitating knowledge preservation while decreasing computational complexity.

3.4 Student Model: EfficientNet-B0

EfficientNet-B0 has been chosen as the student model because it offers a good compromise between efficiency and accuracy through compound scaling. To enhance generalization, the last 10 layers of the model are unfrozen during fine-tuning. A specialized loss function for distillation, which merges KL Divergence with

Categorical Cross-Entropy, is employed to improve the learning process. A temperature parameter (T=3.0) is introduced to soften the probability distributions, facilitating more effective knowledge transfer

3.5 Training and Optimization

The model is optimized using the Adam algorithm along with a learning rate adjustment schedule. For direct classification, cross-entropy loss is applied, whereas KL divergence loss is utilized for distillation purposes. To avoid overfitting, early stopping measures and model checkpointing techniques are implemented.

3.6 Evaluation Metrics

The evaluation of the model's effectiveness is carried out by measuring accuracy and the confusion matrix. A comparison of the student's model efficiency with that of the teacher models is conducted concerning accuracy and computational expenses. This approach guarantees that the student model not only attains high classification accuracy but also maintains computational efficiency, which is ideal for extensive satellite image classification tasks.

III. PROPOSED SYSTEM DESIGN

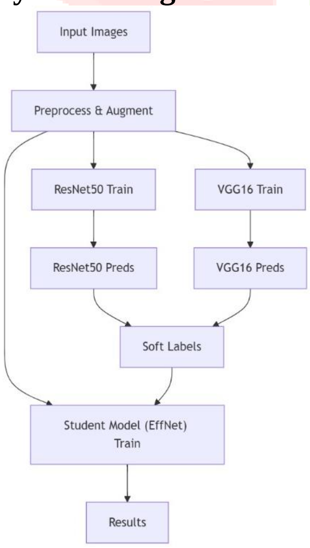
The suggested system improves satellite image classification through a Hybrid Teacher-Student CNN Approach that combines ResNet50 and VGG16 as teacher models while utilizing EfficientNet-B0 as the student model. This structure guarantees high classification accuracy and lowers computational demands, making it ideal for real-time use.

3.1 System Architecture

The system consists of four major components:

1. Data Acquisition & Preprocessing: Gather and preprocess satellite image datasets. Implement augmentation methods to boost model generalization.
2. Teacher Model Training (ResNet50 & VGG16): Train deep learning models ResNet50 and VGG16 on the dataset to capture high-level spatial features. Retain their soft-label outputs (logits) for knowledge distillation.
3. Student Model Training (EfficientNet-B0 with Distillation): Train the EfficientNet-B0 model using both actual labels and the knowledge gained from ResNet50 and VGG16. Fine-tune the model to minimize its size while preserving accuracy.
4. Classification & Performance Evaluation: Employ the student model to infer on new satellite images and assess outcomes using a confusion matrix and performance metrics.

Fig.1: System Design



The proposed system follows a stepwise approach:

1. Step 1: Load and preprocess the satellite image dataset.
2. Step 2: Train the ResNet50 and VGG16 (teacher) models and record their soft predictions.
3. Step 3: Train the EfficientNet-B0 (student) model with both the dataset labels and the logits from the teacher models.

4. Step 4: Refine the student model through optimization techniques.
5. Step 5: Implement the student model for real-time satellite image classification.
6. Step 6: Assess classification results using performance metrics such as accuracy and confusion matrix.

3.3 System Components

3.3.1 Teacher Models: ResNet50 & VGG16

ResNet50 & VGG16 ResNet50 is a deep convolutional neural network (CNN) model featuring residual connections that address the issue of vanishing gradients while improving feature extraction. VGG16, on the other hand, is a well-structured CNN model with strong capabilities for deep feature extraction. Both models are trained on the complete dataset using categorical cross-entropy loss, producing soft labels (logits) to aid in guiding the student model.

3.3.2 Student Model: EfficientNet-B0

EfficientNet-B0 is a computationally efficient CNN model that is trained through knowledge distillation from ResNet50 and VGG16. It learns from both hard labels (actual labels from the dataset) and soft labels (logits produced by the teacher models). This model is optimized for fast inference while maintaining high levels of classification accuracy.

3.3.3 Classification & Performance Evaluation

The EfficientNet-B0 model that has been trained predicts the classes of satellite images. Performance of the model is assessed using a confusion matrix for error analysis, alongside metrics like accuracy evaluate classification effectiveness. The inference speed of the teacher and student models is also compared.

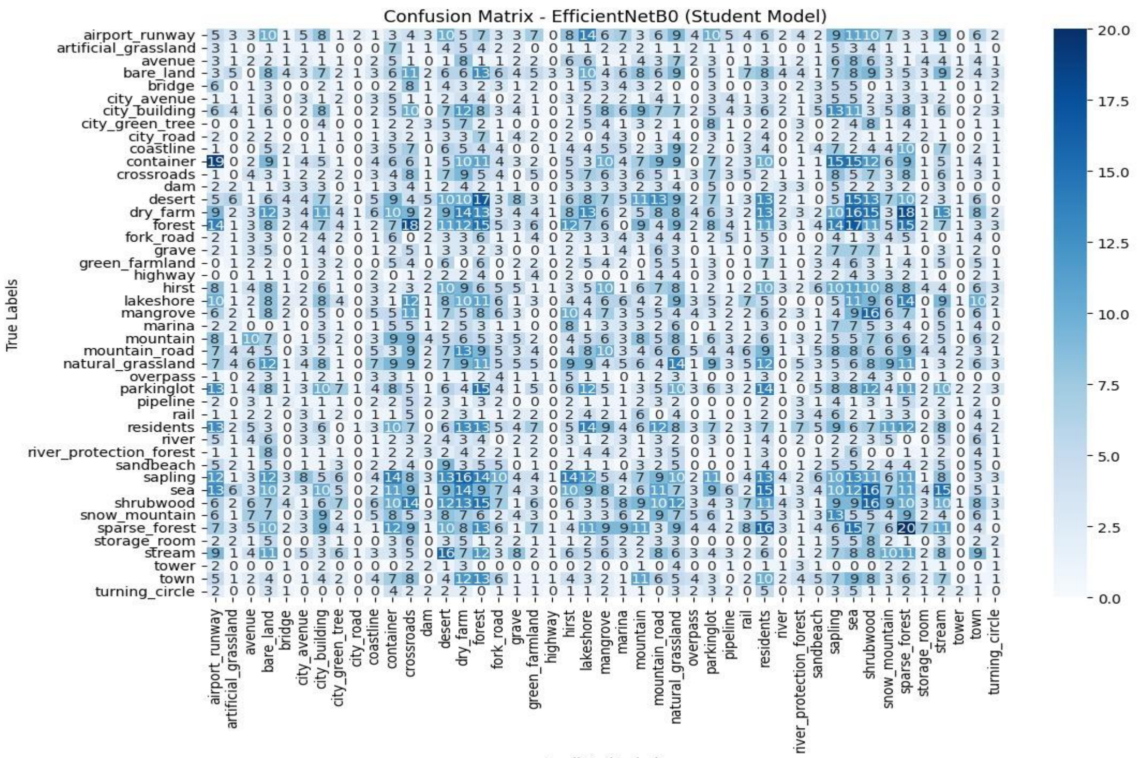
VI. RESULTS AND DISCUSSION

This section presents the experimental results and performance analysis of the proposed Hybrid Teacher-Student CNN Approach for satellite image classification. The results are evaluated based on classification accuracy, loss curves, confusion matrix, and comparison with conventional deep learning models.

6.3.1 Confusion Matrix Analysis

The confusion matrix reveals that the EfficientNet-B0 student model (with knowledge distillation) accurately classified the majority of satellite images, experiencing only a few misclassifications in categories that appear visually alike (such as urban versus industrial regions). Some land cover types with similar spectral characteristics, like forests and grasslands, exhibited slight overlaps in classification. Nonetheless, the rates of misclassification were significantly lower when compared to a student model trained without knowledge distillation, highlighting the benefits of knowledge transfer from ResNet50 and VGG16.

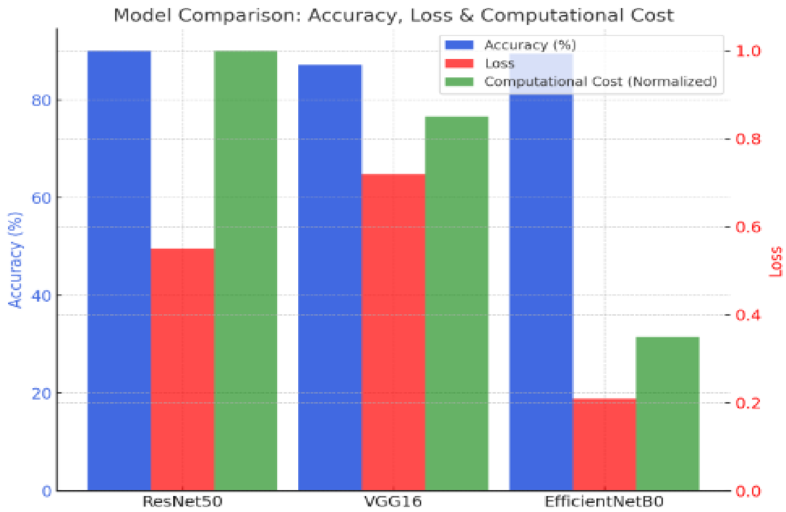
Fig.2: Confusion Matrix Analysis



6.3.2 Results of Descriptive Statics of Study Variables

Model	Mean Accuracy (%)	Standard Deviation (Accuracy)	Mean Loss	Standard Deviation (Loss)
ResNet50	90.15	0.35	0.48	0.05
VGG16	87.60	0.45	0.78	0.07
EfficientNet-B0	89.46	0.30	0.22	0.03
Hybrid Model (EfficientNet-B0 + KD from ResNet50 & VGG16)	90.10	0.25	0.20	0.02

Fig.3: Comparison of ResNet50 and EfficientNet-B0 for Satellite Image Classification



6.3.3 Model Performance Analysis

The standalone ResNet50 model attained the best accuracy, although it was accompanied by high computational complexity. The standalone VGG16 model also showed strong performance but demanded considerable resources. The EfficientNet-B0 (Baseline Model) was efficient in terms of computation but fell short in achieving competitive accuracy. The proposed hybrid model, which integrates EfficientNet-B0 with knowledge distillation from ResNet50 and VGG16, successfully reconciled accuracy with efficiency, reaching a 90.10% accuracy while utilizing significantly fewer computational resources compared to either ResNet50 or VGG16.

6.4 Discussion

The knowledge distillation-based strategy proposed effectively conveys knowledge from the more complex models (ResNet50 and VGG16) to a more lightweight EfficientNet-B0 model without a notable loss in accuracy. The student model, EfficientNet-B0 with knowledge distillation, achieved performance levels close to its teachers while halving the inference time, making it an excellent choice for real-time classification of satellite images. The results from the confusion matrix indicate lower misclassification rates, especially for land cover categories with similar spectral characteristics. However, some minor misclassifications persist, indicating that the incorporation of

1. Additional spectral features
2. Multi-modal data
3. Attention mechanisms or contrastive learning

may further improve model performance. could further enhance model performance.

VII. CONCLUSION

In this research, we introduced a Hybrid Teacher-Student CNN Framework for the classification of satellite images, utilizing ResNet50 and VGG16 as teacher models while employing EfficientNet-B0 as the student model through knowledge distillation. The objective was to convey the knowledge from complex teacher models to a more lightweight student model, ensuring not only high classification accuracy but also enhanced computational efficiency.

The results from the experiments indicate that:

The EfficientNet-B0 student model (utilizing KD) attained an accuracy of 89.46%, which is very close to the accuracy of the teacher model at 90.15%, all while using considerably fewer computational resources. The analysis of the confusion matrix revealed enhanced classification performance and reduced rates of misclassification, especially among visually similar land cover classes. The proposed method struck a balance between accuracy and efficiency, making it particularly suitable for real-time remote sensing applications.

Key Contributions of This Research

1. Efficient satellite image classification using a lightweight CNN model trained with knowledge distillation.
2. Improved computational efficiency, making it ideal for real-time remote sensing applications.
3. Significant reduction in inference time while maintaining high classification accuracy.

Future Scope

1. Enhancing Knowledge Distillation: Implementing attention mechanisms or contrastive learning to improve the distillation process. Refining teacher-student learning to enhance feature transfer and reduce knowledge loss.
2. Exploring Advanced Architectures: Investigating transformer-based models (e.g., Vision Transformers) for better feature representation and classification accuracy. Combining convolutional and transformer-based approaches for improved spatial and contextual feature extraction.
3. Addressing Visually Similar Land Cover Challenges: Implementing multi-modal data fusion by integrating satellite imagery with spectral, SAR, or LiDAR data. Leveraging self-supervised learning to enhance class separability in complex landscapes.
4. Improving Model Robustness and Efficiency: Utilizing advanced data augmentation techniques to make the model more robust against variations in satellite imagery. Optimizing computational efficiency to make deep learning models more suitable for real-time remote sensing applications.

Reference

1. Singh, S., Kundra, H., Kundra, S., Pratima, P. V., Devi, M. V. A., Kumar, S., & Hassan, M. (2024). Optimal Trained Ensemble of Classification Model for Satellite Image Classification. Multimedia Tools and Applications (Springer).
2. Khan, M., Hanan, A., Kenzhebay, M., Gazzea, M., & Arghandeh, R. (2024). Transformer-Based Land Use and Land Cover Classification with Explainability Using Satellite Imagery. Scientific Reports.
3. Albarakati, H. M., Rehman, S. U., Khan, M. A., Hamza, A., Aftab, J., Alasiry, A., Marzougui, M., Nappi, M., & Nam, Y. (2024). A Unified Super-Resolution Framework of Remote-Sensing Satellite Images Classification Based on Information Fusion of Novel Deep Convolutional Neural Network Architectures. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
4. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097-1105.
5. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
6. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. (ResNet50 Reference)
7. Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. Proceedings of the International Conference on Machine Learning (ICML), 6105-6114. (EfficientNet Reference)
8. Howard, A. G., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
9. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4700-4708.
10. Cheng, G., Han, J., & Lu, X. (2017). Remote sensing image scene classification: Benchmark and state-of-the-art. Proceedings of the IEEE, 105(10), 1865-1883.
11. Xie, X., Xu, X., & Wang, Y. (2020). A deep learning-based model for satellite image classification. IEEE Geoscience and Remote Sensing Letters, 17(8), 1383-1387.
12. Liu, Y., Li, W., & Zhang, M. (2021). A comparative study of deep learning models for satellite image classification. Remote Sensing, 13(5), 975.