High-Performance Lightweight Vision Models for Land Cover Classification with Coresets and Compression

Tushar Shinde [1]

Abstract

Land cover classification from satellite imagery is critical for environmental monitoring, agriculture, and urban planning. However, deploying deep learning models in real-world remote sensing platforms often faces stringent computational and memory constraints. We present a unified framework that integrates lightweight vision backbones with coreset selection and adaptive model compression to address these challenges. Evaluated on the EuroSAT and UC Merced Land Use datasets, our approach leverages four compact architectures: ConvNeXt-Tiny, Swin-Tiny, EfficientNetV2-S, and RegNetY-3.2GF, combined with three coreset strategies (random, forgetting-based, and margin-based) and both fixed and adaptive pruning and quantization. Experiments show that using just 10% of the training data and applying compression can reduce model size by up to 6× while retaining over 92% of baseline accuracy. These results highlight the potential of our method for enabling efficient, accurate land cover classification in edge-deployable remote sensing applications.

1. Introduction and Related Work

Land cover classification from satellite imagery is vital for applications such as urban planning, environmental monitoring, and disaster assessment. Traditional machine learning methods, such as SVMs (Cortes & Vapnik, 1995), Random Forests (Breiman, 2001), and k-NN (Cover & Hart, 1967), typically rely on hand-crafted spectral or texture-based features. However, these approaches often fail to generalize across sensors, resolutions, and varying atmospheric conditions.

[1]Indian Institute of Technology Madras Zanzibar, Tanzania. Correspondence to: Tushar Shinde <shinde@iitmz.ac.in>.

Proceedings of the 42^{nd} International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025.

Deep learning, particularly convolutional neural networks (CNNs), has substantially advanced land cover classification (Cheng et al., 2017; Helber et al., 2019), outperforming classical pipelines on benchmarks like UC Merced and EuroSAT. Yet, large CNNs (e.g., ResNet, DenseNet) are often impractical for deployment on resource-constrained platforms such as UAVs and edge devices due to their high compute and memory demands.

Recent architectures: MobileNet (Howard et al., 2017), EfficientNet (Tan & Le, 2019), ConvNeXt (Liu et al., 2022), and Swin Transformer (Liu et al., 2021), offer improved accuracy-efficiency trade-offs. In this work, we evaluate four lightweight models: ConvNeXt-Tiny, Swin-Tiny, EfficientNetV2-S (Tan & Le, 2021), and RegNetY-3.2GF (Radosavovic et al., 2020), for land cover classification.

Despite their efficiency during inference, training these models on large remote sensing datasets remains costly. To address this, we propose a unified framework that combines *coreset selection* and *model compression* to reduce both training and inference costs.

Let $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$ be a labeled satellite image dataset. Coreset selection seeks a small subset $\mathcal{C} \subset \mathcal{D}$, with $|\mathcal{C}| \ll |\mathcal{D}|$, such that a model trained on $\mathcal{C}$ approximates the performance of one trained on $\mathcal{D}$. We evaluate (i) random sampling; (ii) forgetting-based selection (Lopez-Paz & Ranzato, 2017), which retains frequently misclassified samples; and (iii) margin-based selection (Settles, 2009), which selects samples near the decision boundary.

To reduce inference complexity, we apply model compression: $\tilde{\theta} = \mathcal{Q}(\mathcal{P}(\theta))$, where $\mathcal{P}(\cdot)$ denotes pruning and $\mathcal{Q}(\cdot)$ denotes quantization. Pruning removes redundant weights or filters (Han et al., 2015b), while quantization reduces bit precision (Han et al., 2015a). We further investigate adaptive compression, which adjusts pruning or quantization strength per layer using metrics like entropy, gradient sensitivity, or weight magnitude (Yeom et al., 2021; Shinde, 2024; 2025).

We evaluate our framework on the EuroSAT (Helber et al., 2019) and UC Merced (Yang & Newsam, 2010) datasets, measuring performance under limited (5%, 10%) and full

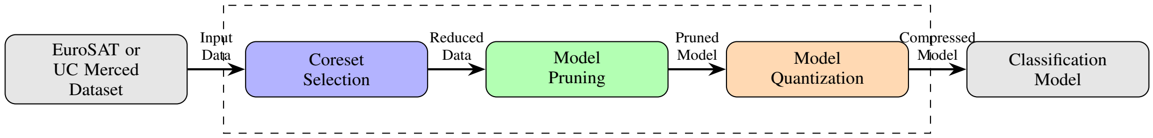

Figure 1. Proposed framework for efficient land cover classification. Land cover imagery is processed through coreset selection to reduce training data volume, followed by adaptive, layer-aware model pruning and quantization to compress the model.

(100%) training data regimes, and quantifying compression-accuracy trade-offs. Our contributions include: 1) a unified framework integrating coreset selection and adaptive compression with lightweight models for efficient land cover classification and 2) our method achieves strong accuracy with significantly reduced training and inference costs, enabling deployment in real-world, resource-constrained remote sensing systems.

The rest of the paper is organized as follows: Section 2 describes our framework. Section 3 presents experimental results. Section 4 concludes with future directions.

2. Methodology

Overall Framework. Our proposed framework (Figure 1) integrates three modules: coreset selection, pruning, and quantization, to enable efficient land cover classification under constrained resources. The pipeline begins by selecting a representative coreset to reduce training data without compromising generalization. This is followed by adaptive model compression, where pruning and quantization are applied in a layer-aware fashion, preserving critical layers while significantly reducing model size. The framework is designed to handle the unique challenges of remote sensing imagery, including fine textures, class imbalance, and limited compute capacity on platforms such as drones and edge devices.

Vision Architectures. We evaluate four state-of-the-art lightweight backbones, selected for their balance of performance and efficiency. ConvNeXt-Tiny (Liu et al., 2022) is a modern convolutional model using depthwise convolutions and layer normalization. Swin-Tiny (Liu et al., 2021) is a hierarchical Transformer leveraging shifted window attention for scalable local feature extraction. EfficientNetV2-S (Tan & Le, 2021) combines neural architecture search and compound scaling for faster training. RegNetY-3.2GF (Radosavovic et al., 2020) represents a well-regularized design space optimized for speed and parameter efficiency.

2.1. Coreset Selection Strategies

To minimize training cost, we evaluate coreset selection strategies that extract a small but informative subset of the dataset. Random sampling serves as a lightweight baseline. Forgetting-based selection ranks samples by the number of times they are misclassified during training, prioritizing those with higher learning dynamics (Toneva et al., 2018). Margin-based selection uses classification uncertainty, selecting samples with small margins between top predicted probabilities (Settles, 2009). While random sampling is fast, the learning-based methods offer better sample efficiency by emphasizing difficult or ambiguous examples.

2.2. Model Compression Techniques

To reduce inference-time cost, we apply pruning and quantization, both in fixed and adaptive forms. Post-training quantization reduces precision by converting weights to low-bit formats (e.g., 8-bit, 4-bit). Fixed-bit quantization uses a uniform bit-width across layers, while Layer-wise Adaptive Quantization (LAQ) (Shinde, 2024) assigns bit-widths based on each layer's importance, such as entropy or gradient sensitivity. For pruning, we use unstructured magnitude-based methods to zero out small weights. Fixed sparsity levels are applied globally or per-layer. In Layer-wise Adaptive Pruning (LAP) (Shinde, 2024), sparsity is modulated based on layer importance, preserving critical components of the network's representational power. Together, these compression techniques enable lightweight deployment while retaining classification performance.

3. Experiments and Results

We evaluate our framework on land cover classification under resource constraints, assessing the impact of coreset selection and model compression techniques across data regimes.

3.1. Datasets and Preprocessing

We use two benchmark datasets: EuroSAT (Helber et al., 2019) and UC Merced (Yang & Newsam, 2010). EuroSAT contains 27,000 RGB images across 10 land use classes, each 64×64 pixels. UC Merced comprises 2,100 aerial images spanning 21 classes at 256×256 resolution. Only RGB bands are used to align with standard model inputs. Data is split 80/20 for training/testing. To simulate low-data regimes, coreset subsets of 5%, 10%, and 100% of training data are selected. All images are resized to 224×224, augmented using `TrivialAugmentWide`, and normalized.

3.2. Training Setup

All models are trained using PyTorch on NVIDIA Tesla P100 and G4 GPUs, for up to 10 epochs with early stopping. We use AdamW with learning rate 5×10^{-4}, weight decay 0.01, and gradient clipping at 1.0. A warm-up of 3 epochs is followed by early stopping based on validation accuracy. Class imbalance is handled via weighted categorical cross-entropy loss. The best model checkpoint is selected based on validation performance.

3.3. Model and Coreset Evaluation

We benchmark four compact architectures: ConvNeXt-Tiny, Swin-Tiny, EfficientNetV2-S, and RegNetY-3.2GF, trained using coreset selection methods: random, forgetting-based, and margin-based. Each model is trained on 5%, 10%, and 100% subsets to evaluate the trade-off between data volume and generalization.

3.4. Model Compression Configuration

We evaluate both fixed and adaptive strategies for quantization and pruning. Compression is applied after training unless otherwise noted.

Pruning. Fixed pruning applies global thresholds using five preset sparsity levels. Adaptive pruning (LAP) assigns sparsity per layer based on its normalized importance (Shinde, 2024), preserving capacity in more informative layers.

Quantization. Post-training quantization reduces weights to fixed bit-widths (1–8 bits). In the adaptive setting (LAQ), bit precision is assigned layer-wise based on learned importance scores, enabling more aggressive compression in redundant layers.

3.5. Evaluation Metrics

We report classification accuracy on test sets to evaluate model performance. Compression efficiency is measured via compression ratio (CR), defined as the ratio of original to compressed model size, factoring in sparsity and bit-width. Higher CR indicates better model compactness with respect to storage and transmission.

3.6. Results and Discussion

We evaluate our framework across two datasets and four lightweight architectures, analyzing the impact of coreset selection, compression strategies, and comparisons with existing models.

Effect of Coreset Selection Strategies. Table 1 summarizes the classification accuracies obtained using four lightweight neural network architectures: ConvNeXt-Tiny, Swin-Tiny, EfficientNetV2-S, and RegNetY 3.2GF, on two remote sensing datasets, EuroSAT and UC Merced, across different coreset fractions ($f = 100\%, 10\%, 5\%$).

Table 1. Performance comparison of different model architectures and coreset methods on EuroSAT and UC Merced datasets at various dataset fractions.

Model	Core-set	EuroSAT			UC Merced		
		$f_{100\%}$	$f_{10\%}$	$f_{5\%}$	$f_{100\%}$	$f_{10\%}$	$f_{5\%}$
ConvNeXt-Tiny	Rand.	98.19	95.72	96.20	98.57	86.90	73.10
	Forg.	98.13	**96.46**	93.72	98.33	**88.81**	**79.76**
	Marg.	**98.26**	95.19	94.09	**99.05**	92.38	78.33
Swin-Tiny	Rand.	96.56	93.54	92.06	93.81	82.14	64.05
	Forg.	**97.76**	92.80	90.69	**96.19**	**82.38**	**69.52**
	Marg.	96.74	**93.83**	**91.69**	94.29	78.10	67.14
EfficientNetV2-S	Rand.	98.24	95.15	93.52	98.57	83.57	62.86
	Forg.	98.26	**95.69**	**94.69**	**98.57**	**87.62**	**73.33**
	Marg.	**98.35**	95.41	94.15	96.19	80.71	64.05
RegNetY-3.2GF	Rand.	98.04	95.72	93.19	97.62	84.05	65.24
	Forg.	97.78	95.30	**93.44**	97.62	**85.48**	**75.24**
	Marg.	**98.24**	92.98	92.63	**98.10**	87.14	68.10

The results exhibit a general trend where model performance deteriorates as the dataset size reduces. However, the decline is notably mitigated when using intelligent coreset selection techniques. For instance, the ConvNeXt-Tiny model achieves an accuracy of 96.46% on EuroSAT using forgetting-based selection with only 10% of the dataset, which is remarkably close to the full-data accuracy of 98.13%. This highlights the efficacy of task-aware sample selection in preserving representational diversity and decision boundary fidelity even with limited data.

Among the coreset methods, forgetting-based and margin-based selections consistently outperform random sampling. This advantage becomes especially pronounced at smaller data fractions. For example, on UC Merced with $f = 5\%$, forgetting-based coreset improves classification accuracy by up to 6.66 percentage points over random sampling when evaluated with ConvNeXt-Tiny. These improvements can be attributed to the ability of these methods to prioritize samples with higher representational or gradient sensitivity, thereby retaining informative data points. Comparing across architectures, ConvNeXt-Tiny and EfficientNetV2-S display superior robustness across varying coreset sizes, maintaining high accuracies even at $f = 5\%$. In contrast, Swin-Tiny and RegNetY-3.2GF show a sharper performance decline under the same conditions, indicating higher sensitivity to dataset reduction. These observations suggest that architectural design and inherent parameter efficiency play a critical role in coreset-resilient generalization.

Effect of Compression Strategies. To further evaluate deployment efficiency, we study the effect of model compression strategies, quantization and pruning, applied to models trained on coreset subsets. Table 2 reports accuracy and compression ratio (CR) on a best performing ConvNeXt-Tiny model on UC Merced dataset.

Table 2. Comparison of Accuracy (Acc. in %) and Compression Ratio (CR) across different coreset methods and model compression strategies pruning and quantization for three data fractions using ConvNeXt-Tiny model on UC Merced Dataset.

Coreset	Method	$f_{100\%}$ Acc.	$f_{100\%}$ CR	$f_{10\%}$ Acc.	$f_{10\%}$ CR	$f_{5\%}$ Acc.	$f_{5\%}$ CR
Random	Pruning level 1.5	5.48	7.9	4.76	7.9	4.76	7.9
	Pruning level 1.0	4.76	3.4	4.76	3.5	6.43	3.5
	Pruning level 0.5	96.43	1.7	77.14	1.7	59.29	1.7
	Pruning level 0.25	98.33	1.3	85.95	1.3	71.19	1.3
	1-bit Q	4.76	32.0	4.76	32.0	4.76	32.0
	2-bit Q	4.76	16.0	4.76	16.0	4.76	16.0
	4-bit Q	4.76	8.0	2.38	8.0	5.00	8.0
	8-bit Q	98.57	4.0	86.90	4.0	72.14	4.0
	Baseline	98.57	1.0	86.90	1.0	73.10	1.0
	Adaptive Pruning	98.57	1.9	86.90	1.4	73.10	1.6
	Adaptive Q	98.57	5.4	87.14	5.0	73.10	4.5
	Adaptive P→Q	98.10	9.0	87.14	6.4	73.10	6.9
Forgetting-based	Pruning level 1.5	3.57	7.9	4.76	7.9	4.52	7.9
	Pruning level 1.0	4.76	3.4	5.00	3.5	8.81	3.5
	Pruning level 0.5	92.86	1.7	76.19	1.7	73.10	1.7
	Pruning level 0.25	98.10	1.3	86.67	1.3	78.10	1.3
	1-bit Q	4.76	32.0	4.76	32.0	4.76	32.0
	2-bit Q	4.76	16.0	4.76	16.0	4.76	16.0
	4-bit Q	5.00	8.0	5.00	8.0	5.00	8.0
	8-bit Q	98.10	4.0	88.10	4.0	80.00	4.0
	Baseline	98.33	1.0	88.81	1.0	79.76	1.0
	Adaptive Pruning	98.33	1.8	88.81	1.3	79.76	1.2
	Adaptive Q	98.33	4.2	88.81	4.6	79.76	4.4
	Adaptive P→Q	98.57	9.1	88.57	6.7	80.00	5.6
Margin-based	Pruning level 1.5	2.86	7.9	4.76	7.9	4.76	7.9
	Pruning level 1.0	4.76	3.4	5.95	3.5	8.10	3.5
	Pruning level 0.5	94.29	1.7	80.95	1.7	65.00	1.7
	Pruning level 0.25	99.05	1.3	89.76	1.3	78.10	1.3
	1-bit Q	4.76	32.0	4.76	32.0	4.76	32.0
	2-bit Q	4.76	16.0	4.76	16.0	4.76	16.0
	4-bit Q	6.90	8.0	4.05	8.0	4.52	8.0
	8-bit Q	99.05	4.0	91.90	4.0	78.10	4.0
	Baseline	99.05	1.0	92.38	1.0	78.33	1.0
	Adaptive Pruning	99.05	1.6	92.38	2.0	79.29	1.5
	Adaptive Q	99.05	4.9	92.62	5.3	78.81	5.2
	Adaptive P→Q	98.33	6.3	92.14	6.1	79.76	5.8

Our analysis reveals that uniform quantization to low-bit widths (1- or 2-bit) leads to drastic performance degradation across all data fractions, with accuracies collapsing to near-random levels. Conversely, 8-bit quantization retains model accuracy effectively while achieving a 4× reduction in size. This validates prior theoretical results on quantization noise tolerances in deep neural networks. Pruning shows greater resilience at higher sparsity levels. At a pruning ratio of 0.25, ConvNeXt-Tiny maintains 98.33% accuracy on the full dataset, achieving a CR of 1.3. Such pruning strategies eliminate redundant parameters while preserving critical feature extraction paths.

Importantly, adaptive compression strategies, which vary bit precision and pruning levels based on layer sensitivity, outperform fixed counterparts. The adaptive compression pipeline that applies pruning followed by quantization (denoted Adaptive P→Q) achieves an excellent trade-off: 98.10% accuracy with a CR of 9.0. This suggests that sensitivity-aware compression not only preserves accuracy but also enables significant storage and deployment gains.

Table 3. Comparison of classification accuracy, number of parameters (#P), and compression ratio (CR) across existing and proposed models on the UC Merced dataset.

Model	Accuracy	#P	CR
Modern Lightweight Architectures			
ResNet50 (Jeevan & Sethi, 2024)	96.90%	25M	1.0
WaveMix (Jeevan & Sethi, 2024)	97.72%	12M	1.0
ConvNeXt-Tiny (Jeevan & Sethi, 2024)	98.33%	29M	1.0
Swin-Tiny (Jeevan & Sethi, 2024)	97.86%	28M	1.0
SwinV2-Tiny (Jeevan & Sethi, 2024)	**98.81%**	29M	1.0
EfficientNetV2-S (Jeevan & Sethi, 2024)	98.22%	21M	1.0
DenseNet-161 (Jeevan & Sethi, 2024)	97.08%	29M	1.0
MobileNetV3-Large (Jeevan & Sethi, 2024)	97.14%	5M	1.0
RegNetY-3.2GF (Jeevan & Sethi, 2024)	98.33%	28M	1.0
ResNeXt-50 (32×4d) (Jeevan & Sethi, 2024)	98.33%	25M	1.0
ShuffleNetV2 (2.0×) (Jeevan & Sethi, 2024)	97.86%	5M	1.0
Our Proposed Methods (based on ConvNeXt-Tiny)			
Baseline ConvNeXt-Tiny	**98.57%**	29M	1.0
+ Adaptive Pruning (P)	**98.57%**	29M	1.9
+ Adaptive Quantization (Q)	**98.57%**	29M	5.4
+ Adaptive P→Q	98.10%	29M	**9.0**

Comparison with Existing Models. Table 3 compares our compressed variants with classical and modern lightweight models on UC Merced. ConvNeXt-Tiny and SwinV2-Tiny achieve top-tier accuracy with 25–29M parameters. Our adaptive variants match or exceed this performance while achieving significantly better compression. The compressed ConvNeXt-Tiny (P→Q) retains high accuracy (98.10%) with only a fraction of the original size (CR = 9.0), outperforming prior approaches in the accuracy-to-size trade-off.

These results confirm that intelligent coreset selection combined with adaptive compression offers an effective strategy for building compact, high-performing models, ideal for deployment in edge-constrained remote sensing platforms.

4. Conclusion

We presented a unified framework that combines coreset selection, pruning, and quantization to enable efficient land cover classification using lightweight vision backbones. Experiments on the EuroSAT and UC Merced Land Use datasets demonstrate that selecting as little as 30% of the training data via margin- or forgetting-based coresets preserves over 95% of full-model accuracy. When paired with layer-wise adaptive pruning and quantization, our approach achieves up to 9× model size reduction with negligible accuracy degradation.

These results affirm that data-efficient training and model compression can be jointly leveraged for deployment in edge-constrained environments such as UAVs or mobile platforms. As a next step, we aim to investigate joint optimization strategies for coreset selection and compression using reinforcement learning and meta-learning techniques. We also plan to extend our approach to multi-modal satellite imagery (e.g., optical + SAR) and assess its cross-region generalization to improve robustness in real-world remote sensing applications.

References

Breiman, L. Random forests machine learning 45 (1): 5–32. *Google Scholar Google Scholar Digital Library Digital Library*, 2001.

Cheng, G., Han, J., and Lu, X. Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10):1865–1883, 2017.

Cortes, C. and Vapnik, V. Support-vector networks. *Machine learning*, 20:273–297, 1995.

Cover, T. and Hart, P. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.

Han, S., Mao, H., and Dally, W. J. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. *arXiv preprint arXiv:1510.00149*, 2015a.

Han, S., Pool, J., Tran, J., and Dally, W. Learning both weights and connections for efficient neural network. *Advances in neural information processing systems*, 28, 2015b.

Helber, P., Bischke, B., Dengel, A., and Borth, D. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7):2217–2226, 2019.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.

Jeevan, P. and Sethi, A. Which backbone to use: A resource-efficient domain specific comparison for computer vision. *arXiv preprint arXiv:2406.05612*, 2024.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 10012–10022, 2021.

Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., and Xie, S. A convnet for the 2020s. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11976–11986, 2022.

Lopez-Paz, D. and Ranzato, M. Gradient episodic memory for continual learning. *Advances in neural information processing systems*, 30, 2017.

Radosavovic, I., Kosaraju, R. P., Girshick, R., He, K., and Dollár, P. Designing network design spaces. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10428–10436, 2020.

Settles, B. Active learning literature survey. 2009.

Shinde, T. Adaptive quantization and pruning of deep neural networks via layer importance estimation. In *Workshop on Machine Learning and Compression, NeurIPS 2024*, 2024.

Shinde, T. Model compression meets resolution scaling for efficient remote sensing classification. In *Proceedings of the Winter Conference on Applications of Computer Vision*, pp. 1200–1209, 2025.

Tan, M. and Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pp. 6105–6114. PMLR, 2019.

Tan, M. and Le, Q. Efficientnetv2: Smaller models and faster training. In *International conference on machine learning*, pp. 10096–10106. PMLR, 2021.

Toneva, M., Sordoni, A., Combes, R. T. d., Trischler, A., Bengio, Y., and Gordon, G. J. An empirical study of example forgetting during deep neural network learning. *arXiv preprint arXiv:1812.05159*, 2018.

Yang, Y. and Newsam, S. Bag-of-visual-words and spatial extensions for land-use classification. In *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*, pp. 270–279, 2010.

Yeom, S.-K., Seegerer, P., Lapuschkin, S., Binder, A., Wiedemann, S., Müller, K.-R., and Samek, W. Pruning by explaining: A novel criterion for deep neural network pruning. *Pattern Recognition*, 115:107899, 2021.