Bridging the Domain Gap in Remote Sensing: A Comparative Study of Deep CORAL, DAN, and DANN on Multi-Sensor Aerial Image Datasets

Sumit Kumar[1, a)] and Kalpesh Jadav[2, b)]

[1]Research Scholar, Parul Institute of Engineering and Technology, Parul University, Vadodara, Gujarat, India
[2]Faculty of Engineering and Technology, Parul Institute of Engineering and Technology, Parul University, Vadodara, Gujarat, India

[a)]ys.sumitkumar@gmail.com
[b)]kalpesh.jadav@paruluniversity.ac.in

Abstract. Domain adaptation (DA) is essential in remote sensing image analysis, where models frequently encounter substantial performance drops due to domain shift—a result of distributional discrepancies across datasets sourced from heterogeneous sensors, platforms, and imaging conditions. This comparative study evaluates three leading unsupervised DA algorithms—Deep CORAL, Domain Adaptation Network (DAN), and Domain-Adversarial Neural Network (DANN)—on four major aerial image benchmarks, each presenting unique sensor characteristics. Six semantically aligned scene classes are selected to facilitate a rigorous assessment of how each technique manages sensor-induced and platform-driven variability. Through extensive experimentation and interpretability analyses, we quantify strengths and limitations of each approach, aiming to offer clear guidelines for DA algorithm selection in practical, sensor-diverse remote sensing tasks.

Keywords: Remote sensing, domain adaptation, Deep CORAL, DAN, DANN, aerial image classification, cross-domain learning, sensor variability.

INTRODUCTION

Remote sensing imagery has become indispensable to disciplines such as land cover mapping, urban development monitoring, disaster assessment, and global environmental change analysis. Advances in satellite and airborne imaging systems have produced vast amounts of high-resolution, multi-sensor data, greatly benefiting machine learning-driven automation of image classification tasks [1][2]. However, a central challenge persists: the phenomenon of "domain shift". This describes the degradation in model accuracy when algorithms trained in one domain (source) fail to generalize to new domains (target) caused by differences in imaging sensors, acquisition platforms, spatial and spectral characteristics, atmospheric conditions, or even societal and temporal changes [3].

For example, an image classification model trained on the AID dataset (Google Earth multisensory imagery) may not succeed when deployed on the UCMerced dataset (collected by USGS aircrafts), due to differences in resolution, compression artifacts, color palettes, and viewpoint angles. The domain gap is therefore not merely semantic, but fundamentally rooted in sensor and acquisition heterogeneity—posing a substantial obstacle to the operational scalability of remote sensing models [3][4].

Recent years have witnessed the emergence of domain adaptation (DA) techniques designed to counteract such distribution mismatches between source and target data. This paper investigates three mainstream unsupervised DA strategies—DeepCORAL[5], DAN[6], and DANN[7][8]—and systematically benchmarks them on four aerial scene classification datasets with six harmonized semantic classes. Our goal is to provide insight into how such methods perform across significant sensor and platform-related domain shifts, a frontier that remains underexplored in the remote sensing community [3][4].

RELATED WORK

The growing interest in domain adaptation for remote sensing has led to a taxonomy of approaches:

- Feature Alignment Methods: Early work, such as Correlation Alignment (CORAL), focused on matching the distribution of shallow or deep feature spaces by aligning second-order statistics (covariance) between source and target domains. Deep CORAL extends this concept to deep neural network features [4][5].
- Distribution Distance Minimization: Domain Adaptation Network (DAN) utilizes measures like Maximum Mean Discrepancy (MMD) at multiple neural layers, enabling alignment of more complex, nonlinear data distributions [3].
- Adversarial Methods: Domain-Adversarial Neural Network (DANN) advances a paradigm where a domain discriminator tries to distinguish source from target features, while the feature extractor is trained (via a gradient reversal layer) to produce domain-invariant representations, facilitating robust transfer learning [7][8].

In remote sensing, DA has been explored for multispectral, hyperspectral, and panchromatic scenes, with prior works illustrating both the challenges and the promise of such approaches [3]. However, systematic and comparative studies focusing on aerial scene classification using data from multiple, disparate sensors are relatively scarce [4].

DATASETS AND SENSOR CHARACTERISTICS

This study uses four benchmark datasets, each with a distinct sensor and collection profile, illustrating the spectrum of domain variability encountered in remote sensing.

Dataset Summary

TABLE 1: Dataset description

Dataset	Sensor/Platform	Resolution	Classes Used / Total	Description
AID [10]	Google Earth (multi-sensor)	0.5–8 m	6 / 30	Diverse locations, varied sources
NWPU [11]	Google Earth (multi-source)	0.2–30 m	6 / 45	Greater spatial detail than AID
PatternNet [12]	Aircraft/Aerial	~0.06 m	6 / 38	Ultra-high spatial resolution, rich texture
UCMerced [13]	USGS aircraft	0.3 m	6 / 21	Widely used aerial scene benchmark

Common Class Selection

To ensure consistency, we extract six common classes namely Airplane, Beach, Dense Residential, Forest, Sparse Residential and Storage Tank from each dataset as shown in figure 1.

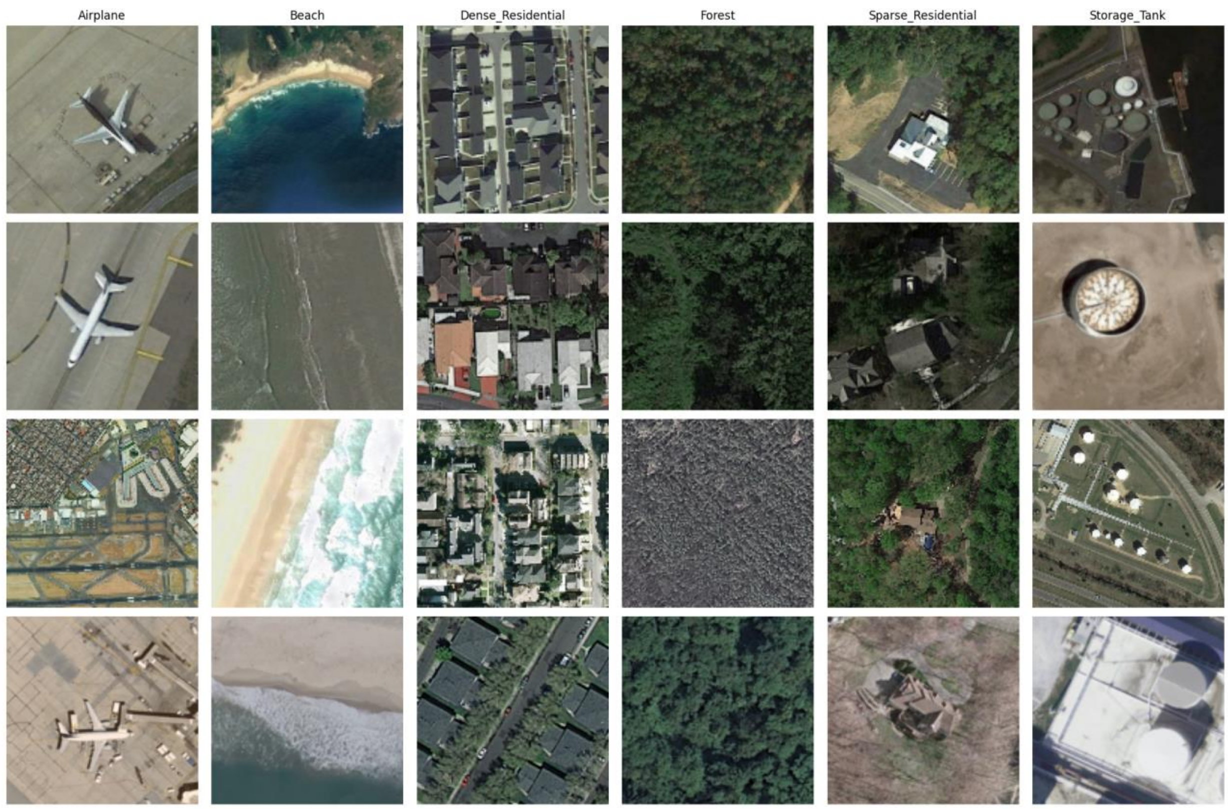


FIG 1. Sample image from each class across the different datasets. The first row shows images from the NWPU dataset, followed by samples from PatternNet, AID, and UCMerced

Sensor Variability

Each dataset reflects sensor differences:

- Spatial resolution: PatternNet presents ultra-high resolution, whereas AID and NWPU have broader and lower resolution ranges.
- Acquisition platform: UCMerced is based on aircraft imagery; AID and NWPU rely on aggregated satellite/aerial imagery from Google Earth.
- Compression and visual artifacts: Google Earth datasets introduce color tone variations and probable compression artifacts, enhancing domain discrepancies.

Such diverse acquisition conditions represent ideal testbeds for benchmarking DA methods, as they closely reflect real-world operational scenarios.

Domain Gap Quantification via Fréchet Inception Distance (FID)

While domain shift in remote sensing is often attributed to differences in sensor platforms, resolutions, and acquisition conditions, we quantitatively assess these shifts using the Fréchet Inception Distance (FID) [14] [15]. FID measures the distributional distance between feature embeddings of two datasets, capturing both first and second-order statistics. Lower FID indicates higher similarity between datasets.

We computed FID scores for all 12 source-target combinations using InceptionV3 activations on the six aligned classes across datasets (Airplane, Beach, Dense Residential, Forest, Sparse Residential, Storage Tank). Results are summarized in Table 2.

TABLE 2: FID Score Matrix Between Datasets

Source → Target	NWPU	PatternNet	AID	UCMerced
NWPU	—	165.68	136.06	252.68
PatternNet	165.68	—	242.45	212.71
AID	136.06	242.45	—	296.53
UCMerced	252.68	212.71	296.53	—

Based on FID scores in Table 2 we observe that AID and NWPU show the smallest domain gap (FID=136.06), making them a relatively easy adaptation pair. Conversely, AID and UCMerced (FID=296.53) exhibit the largest distributional discrepancy, highlighting a hard domain transfer case. PatternNet displays moderate similarity with UCMerced but diverges significantly from AID, likely due to its ultra-high spatial resolution and low background clutter.
To better interpret the relative difficulty of adaptation tasks, we categorize the domain shifts using FID thresholds. Table 3 provides a ranked summary of selected source-target pairs, their FID values, and qualitative difficulty levels based on empirically observed accuracy drops and distributional divergence.[16][17]

TABLE 3: FID Score Matrix Between Datasets

Source → Target	FID Score	Difficulty Level
AID → UCMerced	296.53	**Hard**
UCMerced → NWPU	252.68	Hard
PatternNet → AID	242.45	Hard
PatternNet → UCMerced	212.71	Medium
NWPU → PatternNet	165.68	Medium
NWPU → AID	136.06	**Easy**

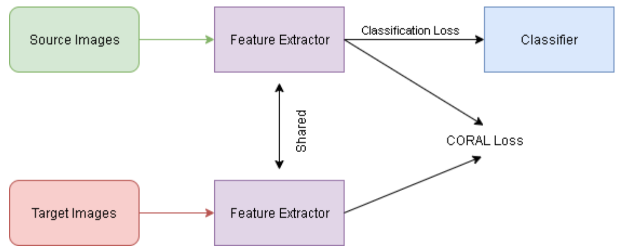
METHODOLOGY

A. Problem Setup

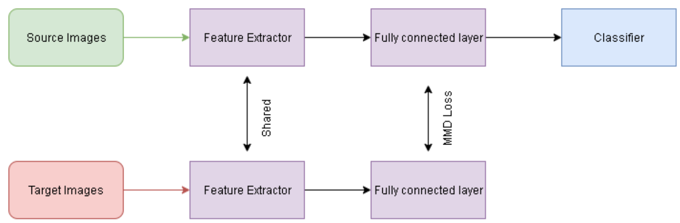
Let the labeled source domain be: $D_s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$, and the unlabeled target domain be: $D_t = \{(x_j^t)\}_{j=1}^{N_t}$. The task is to optimize a classifier using D_s that performs effectively on D_t regardless of the domain shift.

B. Techniques Compared

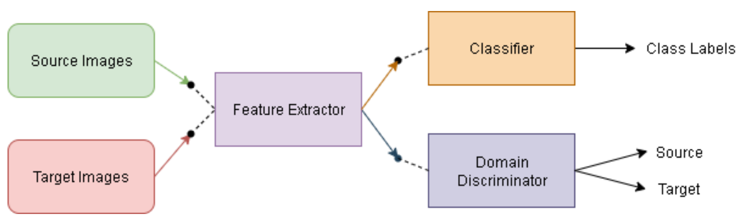
1. **Deep CORAL:** As shown in Figure 2(a) it aligns second-order statistics (covariance) of deep features between source and target domains; it is computationally lightweight and easy to integrate [13].
2. **DAN**: As shown in Figure 2(b) it employs multi-kernel MMD across several network layers to achieve more flexible, nonparametric distribution alignment between source and target representations [12].
3. **DANN:** As shown in Figure 2(c) it incorporates a domain discriminator connected via a Gradient Reversal Layer, enforcing feature extractor outputs that are both discriminative for the classification task and indistinguishable between domains, through adversarial training [12].



(a) Deep CORAL



(b) DAN



(c) DANN

Fig ***2****. Block diagram of Domain Adaptation Techniques namely (a) Deep CORAL, (b) DAN and (C) DANN.*

C. Architecture

All models employ Resnet50 as the backbone feature extractor, balancing computational efficiency with proven effectiveness in remote sensing applications. The classification head uses softmax activation, and for adversarial methods (e.g., DANN), a domain classification branch is introduced. All images are resized and standardized to 128 × 128 pixels.

EXPERIMENTAL SETUP

A. Training Protocol

We used the Adam optimizer with a learning rate of 1 x 10^-4 to improve the model. This is a conventional number that balances speed of convergence and stability. The training used a batch size of 32 over 10 epochs, which allowed for enough iterations for the model to converge and kept the compute efficient. The loss structure was designed as a composite loss function, encompassing cross-entropy loss for classification tasks and integrating supplementary domain adaption components based on the specific approach utilized. These encompassed Maximum Mean Discrepancy (MMD) for the Deep Adaptation Network (DAN), CORAL loss for DeepCORAL, and domain adversarial loss for Domain-Adversarial Neural Networks (DANN). For DANN, a progressive domain loss weight was utilized to stabilize the adversarial training process by incrementally enhancing the impact of the domain classifier throughout training.[18][19]

B. Evaluation Metrics

We primarily employed classification accuracy and confusion matrices to evaluate the model's performance on domain adaptation tasks. The primary indicator of the model's efficacy in transferring knowledge across domains was its performance in classifying data within the unlabeled target domain. Confusion matrices offer a comprehensive analysis of the transferability of each class, beyond only assessing overall correctness. This enabled the assessment of the model's ability to generalize specific class-level attributes from the source domain to the target domain. This facilitated our understanding of the model's advantages and disadvantages in adapting to new domains.[20]

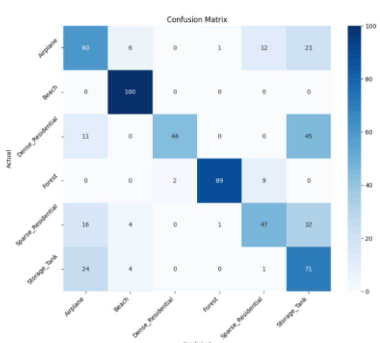
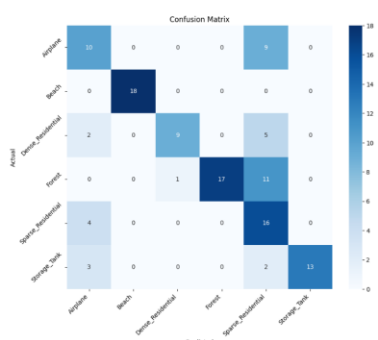
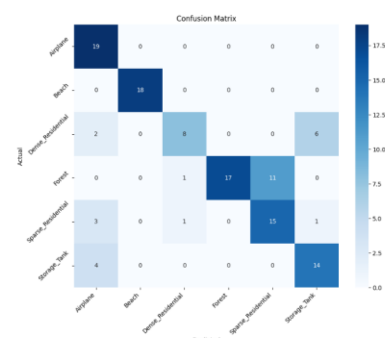
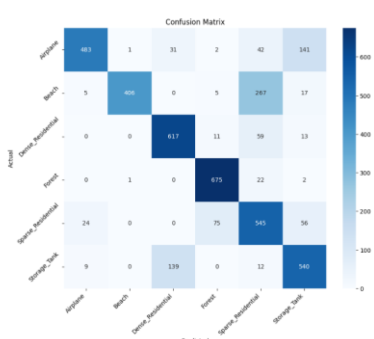
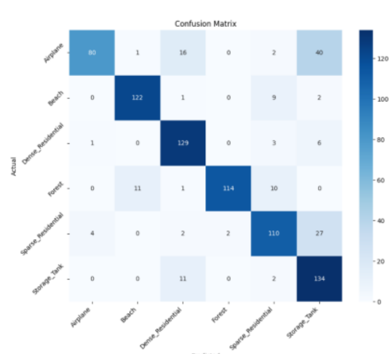
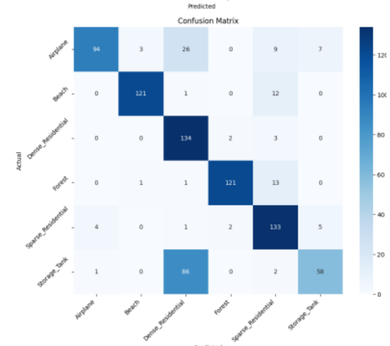
RESULTS AND ANALYSIS

A. Accuracy Comparison

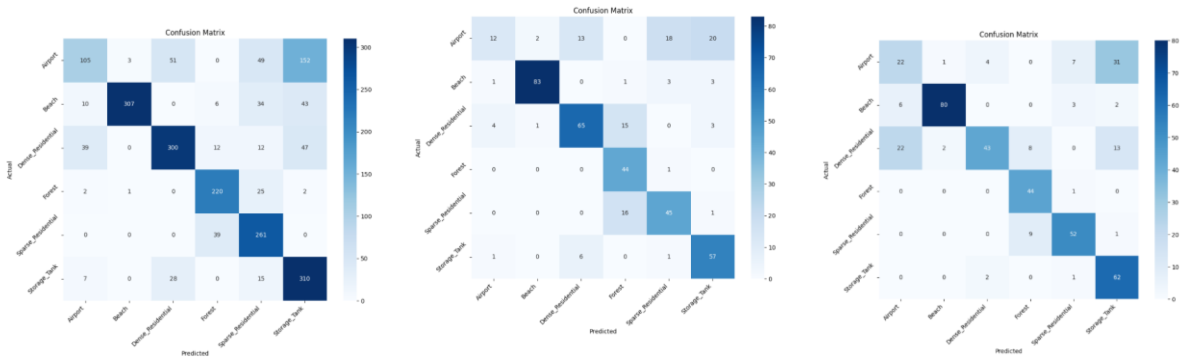
TABLE 4. Accuracy comparison for Source to Target transfer

	DeepCORAL	DAN	DANN
AID → UCMerced	68.50%	69.17%	**75.83**
UCMerced → NWPU	77.76%	82.02%	78.69
PatternNet → AID	72.26%	73.56%	72.83
PatternNet → UCMerced	91.83%	95.83%	92.5
NWPU → PatternNet	94.90%	95.73%	95.52
NWPU → AID	88.26%	90.14%	**90.15**

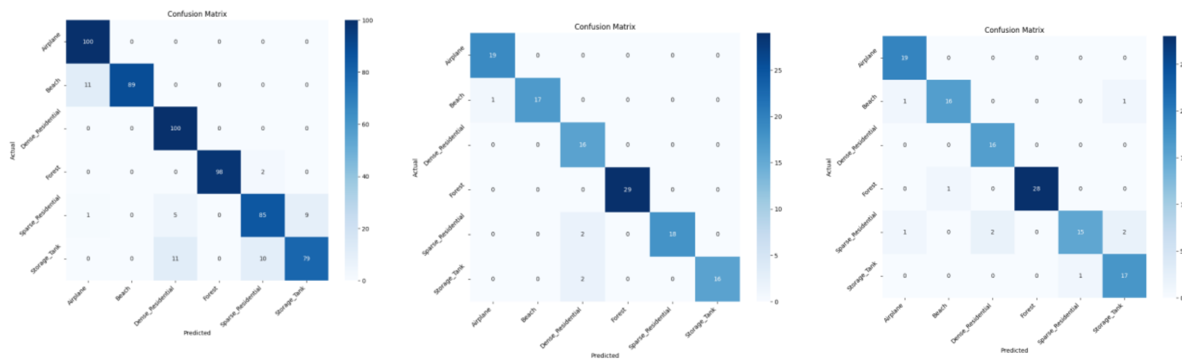
TABLE 5. Confusion matrix Source to Target transfer (A-AID, U-UCMerced, N-NWPU, P-PatternNet)

	DeepCORAL	DAN	DANN
A → U			
U → N			

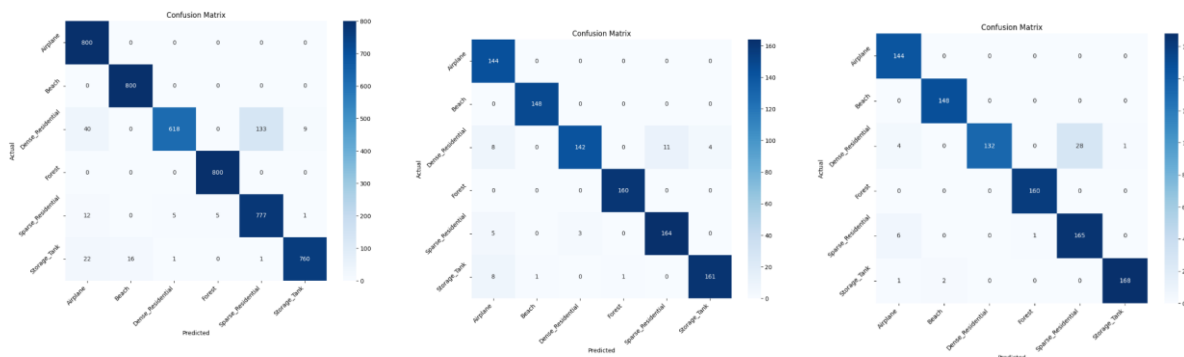
P → A



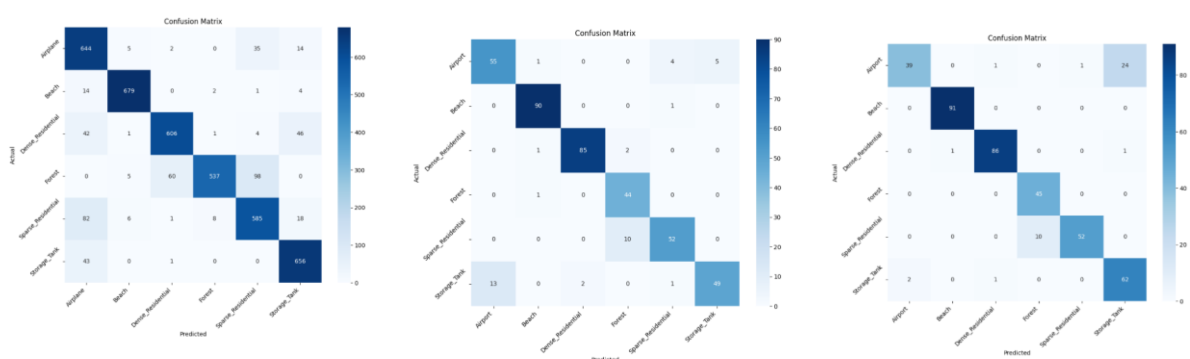
P → U



N → P



N → A



From Table 3 and 4 we can observe that as the FID score (as per table 3) decreases domain gap between datasets reduces and hence we are finding higher accuracy in domain adaptation (as per table 4). Table 5 shows confusion matrix for all domain adaptation cases which supports the accuracy scores of table 4.

B. Qualitative Insights

Our results highlight the importance of domain adaptation in remote sensing tasks that encounter significant cross-sensor or cross-platform transitions. Among the methods that were looked at:

- DANN gets the best results when the domain shift is very strong since it has strong adversarial feature alignment.
- DAN is more stable and consistent when there are only a few domain gaps.
- Deep CORAL is still appealing since it is simple and doesn't need a lot of computing power, which makes it good for applications that don't have a lot of resources.

So, the level of projected domain discrepancy and resource availability should help you choose the right DA technique.

CONCLUSION

This thorough study shows how important domain adaptation (DA) is for remote sensing applications, where there are often big domain gaps caused by sensors and platforms. Domain-Adversarial Neural Networks (DANN) exhibited the most reliable performance enhancements in the presence of significant domain shift among the assessed methodologies. On the other hand, the Deep Adaptation Network (DAN) showed more steady benefits in situations where the domains were only slightly different. DeepCORAL also became a good choice because it worked well and was easy to use with little extra processing power. These results indicate that the choice of a DA approach ought to be guided by the extent of the domain gap and the computational resources at hand. Looking ahead, there are a number of paths that need more research. These include using transformer-based architectures to better capture spectral-spatial dependencies, using contrastive learning and self-supervised methods to reduce the need for labeled source data, and extending to multi-source domain adaptation to make the system more robust in different sensor environments. Also, adding sensor metadata, like geographic coordinates, to the model adaption process can make it more useful in the real world and better fit with how remote sensing is used in the field.

REFERENCES

1. Tuia, Devis & Persello, Claudio & Bruzzone, Lorenzo. (2016). Domain Adaptation for the Classification of Remote Sensing Data: An Overview of Recent Advances. IEEE Geoscience and Remote Sensing Magazine. 4. 41-57. 10.1109/MGRS.2016.2548504.
2. Ma, Lei & Liu, Yu & Zhang, Xueliang & Ye, Yuanxin & Yin, Gaofei & Johnson, Brian. (2019). Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing. 152. 166-177. 10.1016/j.isprsjprs.2019.04.015.
3. Tuia, D., Persello, C., & Bruzzone, L. (2021). Recent Advances in Domain Adaptation for the Classification of Remote Sensing Data. arXiv: Computer Vision and Pattern Recognition. https://doi.org/10.1109/MGRS.2016.2548504
4. Valerio Marsocci, Nicolas Gonthier, Anatol Garioud, Simone Scardapane, Clément Mallet.(2023) GeoMultiTaskNet: remote sensing unsupervised domain adaptation using geographical coordinates. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Jun 2023, Vancouver, Canada.
5. Sun, Baochen & Saenko, Kate. (2016). Deep CORAL: Correlation Alignment for Deep Domain Adaptation. 10.1007/978-3-319-49409-8_35.
6. M. Long et al., “Learning Transferable Features with Deep Adaptation Networks,” ICML, 2015.
7. Y. Ganin et al., “Domain-Adversarial Training of Neural Networks,” JMLR, 2016.
8. Z. Zhong et al., “Deep Adversarial Learning for Cross-Domain Scene Classification,” Remote Sensing, 2020.
9. Zeng, J., Gu, Y., Qin, C. et al. Unsupervised domain adaptation for remote sensing semantic segmentation with the 2D discrete wavelet transform. Sci Rep 14, 23552 (2024). https://doi.org/10.1038/s41598-024-74781-y
10. G. Xia et al., “AID: A Benchmark Dataset for Performance Evaluation of Aerial Scene Classification,” IEEE TGRS, 2017.
11. G. Cheng et al., “Remote Sensing Image Scene Classification: Benchmark and State of the Art,” Proc. IEEE, 2017.

12. W. Zhou et al., “PatternNet: A Benchmark Dataset for Remote Sensing Image Retrieval,” ISPRS, 2018.
13. Y. Yang and S. Newsam, “Bag-of-Visual-Words and Spatial Extensions for Land-Use Classification,” ACM GIS, 2010.
14. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). GANs trained by a two time-scale update rule converge to a local Nash equilibrium. Advances in neural information processing systems.
15. Alegavi, S., & Sedamkar, R. (2025), “Optimizing Remote Sensing Image Retrieval Through a Hybrid Methodology”, Journal of Imaging, MDPI, 11(6), 179. https://doi.org/10.3390/jimaging11060179
16. Akshay Utane, Sharad Mohod, Ashay Rokade, Yogesh Thakare, Hemant Kasturiwale, “An Ensemble Learning with Deep Feature Extraction Approach for Recognition of Traffic Signs in Advanced Driving Assistance Systems” International Journal of Intelligent Systems and Applications in Engineering, 12(3), 1222–1229. https://ijisae.org/index.php/IJISAE/article/view/5402
17. S. Alegavi, R.R. Sedamkar, “Implementation of Deep Convolutional Neural Network for Classification of Multiscaled & Multiangled Remote Sensing Scene”, International Journal on Soft Computing approaches for image analysis in practical scenario: Challenges, Solutions and Applications, Intelligent Decision Technologies, 14 (2020) 21–34, DOI: 10.3233/IDT-190076 IOS Press February 2020. (Scopus, Web of Science)
18. Swati Bhisikar, Shreya Sawant, Tanvi Sawant , Sayali Narale , Hemant Kasturiwale, HAND GESTURE CONTROLLED POWERPOINT PRESENTATION USING OPENCV” European Chemical Bulletin , Volume -12 , Special Issue-3 : Page: 5137 – 5145, June 2023; DOI:10.31838/ecb/2023.12.s3.572
19. S. Alegavi, P. Janrao, C. Mahajan, R. Thakkar, V. Pandya, H. Kasturiwale, “Weather forecasting using IoT and neural network for sustainable agriculture”, American Institute of Physics, AIP Conf. Proc. 2842, 040003 (2023), Volume 2842, Issue 1, 12 October 2023, https://doi.org/10.1063/5.0176348
20. Preeti R. Lawhale, Sujata N Kale, Hemant Kasturiwale, Yogesh N. Thakare, “FPGA Implementation of Compact Architecture for Lightweight Hash Algorithm for Resource Constrained Devices” Communications on Applied Nonlinear Analysis, ISSN: 1074-133XVol 32 No. 2(2025)DOI: https://doi.org/10.52783/cana.v32.1861.