engineering proceedings

MDPI

Proceeding Paper

Improving Remote Sensing Classification with Transfer Learning: Exploring the Impact of Heterogenous Transfer Learning †

Maria Rouba [1,*] and Mohammed El Amin Larabi [2]

[1] National Higher School of Geodesy Science and Spatial Techniques, Oran 31004, Algeria
[2] Algerian Space Agency, Alger 16032, Algeria; arabi.mohamed.lamine@gmail.com
* Correspondence: roubamaria645@gmail.com
† Presented at the 4th International Electronic Conference on Applied Sciences, 27 October–10 November 2023; Available online: https://asec2023.sciforum.net/.

Abstract: Deep learning (DL) has become increasingly popular in recent years, with researchers and businesses alike successfully applying it to a wide range of tasks. However, one challenge that DL faces in certain domains, such as remote sensing (RS), is the difficulty of creating large, well-annotated training datasets. This is due to the high cost of acquiring and labeling RS data. This challenge significantly limits the development of DL in RS. RS data can come from multiple sources, such as satellites, airplanes, and drones, and use different sensor technologies. Training DL models on data from one source may not produce the same accuracy on data from other sources, even if they cover the same region. Transfer learning (TL) can help to address this challenge by relaxing the requirement for large training datasets. Specifically, TL allows us to train a model on data from one source and then adapt it to data from another source, even with fewer training data. This makes TL a promising approach for solving both the problem of multisource adaptation and the problem of insufficient training data in the target domain. This paper evaluates the homogenous and heterogeneous TL approach that addresses model transfer across different domains. Transfer gain is measured through specific statistical metrics such as precision, kappa, recall, and F1-score, and a positive gain is empirically shown in the vast majority of cases. The proposed method is evaluated on the challenging task of Multispectral RS image (MSI) classification due to the complexity and variety of natural scenes. This work is examined in terms of its social, economic, and environmental consequences. Additionally, potential future directions for research and the achievement of established goals are explored.

Keywords: transfer learning; heterogenous learning; homogenous learning; multispectral remote sensing; Sentinel-2; Landsat-9

check for updates

Citation: Rouba, M.; Larabi, M.E.A. Improving Remote Sensing Classification with Transfer Learning: Exploring the Impact of Heterogenous Transfer Learning. *Eng. Proc.* **2023**, *56*, 316. https://doi.org/10.3390/ASEC2023-15505

Academic Editor: Simeone Chianese

Published: 31 October 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Most machine learning (ML) methods employed nowadays assume that training and test data are from the same feature space and distribution [1]. This implies that when the data distribution changes, models trained from scratch with newly collected data are required. However, there are situations where obtaining new data for training new models can be financially or logistically challenging, particularly in environments where data collection demands substantial computational power or financial resources. Hence, the ability to leverage existing knowledge becomes highly valuable, offering a means to circumvent the expensive endeavor of reconstructing a model from scratch and amassing sufficient new data to create a dependable system. This is where the concept of transfer learning (TL) becomes indispensable.

TL is a concept that allows information learned in one setting to be used in another, or at least part of it. This can improve the training process by avoiding the need to develop new models from scratch and instead adapting previously built models to a new setting. TL has proven to be a powerful technique with successful results in many different environments.

However, there are several challenges that need to be considered before deciding whether to use TL.

The classification of multispectral remote sensing (RS) images (MSI) is a critical and important topic that has been extensively studied. It is essential for a variety of applications, including land cover mapping, vegetation monitoring, urban analysis, resource management, and decision making. The complexity and variety of natural scenes pose significant challenges for MSI classification. This is due to the unique characteristics of RS data, such as spectral variability, spatial heterogeneity, and temporal dynamics. These characteristics make it difficult to distinguish between different land cover classes, and collecting data to train deep learning (DL) models can be expensive or even unfeasible. Therefore, using homogeneous and heterogeneous TL to reduce costs and labor time and enhance classification accuracies would be extremely useful. TL has shown promise in improving the performance of MSI classification. In recent years, a number of studies have used TL to achieve state-of-the-art results on a variety of MSI classification tasks. However, selecting the right source domain and task is essential for successful TL in RS. The performance of the target model is significantly influenced by the source model.

The importance of selecting the right source dataset and task for TL (TL) in multispectral RS image (MSI) classification for both homogeneous and heterogeneous TL was investigated using Sentinel-2 (S2) and Landsat-9 (L9) MSI imagery as heterogeneous domains. Homogeneous TL was evaluated for S2 to S2 and L9 to L9, and heterogeneous TL was evaluated for S2 to L9 and L9 to S2. A variety of metrics were used for evaluation, including precision, kappa, recall, and F1-score. The results showed the importance of TL and suggested using homogeneous TL whenever possible.

This paper is organized as follows. Section 2 delves into the fundamental concepts of TF, including domain and task definitions, and differentiates between homogeneous and heterogeneous TL. Section 3: The methodology outlines the meticulous data preparation process, dataset details, and the architecture of our neural network models used in the experiments. Section 4: The experimental section is the heart of our paper, where we present the results and analyses of two distinct TF experiments: homogeneous TL and heterogeneous TL. This paper concludes with Section 5.

2. Deep Transfer Learning

In TL, knowledge is transferred from one domain to another to improve the training process, either in terms of model performance or training speed. TL can be used to address the scarcity and cost of collecting training data. When discussing TL, it is necessary to define some concepts. According to [1], a domain D is defined as a collection of data with a shared feature space X and marginal probability distribution P(X), which can be represented as D = {X, P(X)}. Similarly, a task T is defined as a set of data with a shared target space Y and objective predictive model M, which can be represented as T = {Y, M}. TL can be defined as the process of reducing the cost of learning a predictive model in a target domain DT by leveraging knowledge from a source domain DS and a learning task TS, where the source and target domains may have different feature spaces X and/or target spaces Y [2]. This leads to the definition of two fundamental types of TL: homogeneous and heterogeneous.

2.1. Homogenous Transfer Learning

Homogeneous TL is described as a situation where the feature spaces of data in both the source and target domains match precisely (XS = XT), the corresponding outcome spaces are identical (YS = YT), and the dimensions of these spaces are also equivalent (dS = dT) [3]. This alignment simplifies the task of transfer learning (TL), as it establishes a seamless match between the data attributes and the target objectives. When transitioning knowledge from one domain to another, the transferred model simply requires specialization in tackling the particular task within the target domain. This process can be accomplished through fine-tuning the model, wherein the model is trained using data from the target domain while maintaining the weights acquired during training in the source domain. This is more efficient than randomly initializing and optimizing the model from scratch, as the model already has some knowledge from the previous task. Homogeneous TL has proven to be a successful technique in bolstering model performance and expediting training. To illustrate, in reference [4], the authors show that knowledge transfer can be used to improve the performance of image classification on the ImageNet dataset [5].

2.2. Heterogenous Transfer Learning

Heterogeneous TL [6] is distinguished by the presence of dissimilar feature spaces and/or target spaces between the source and target domains, as depicted in Figure 1. In other words, XS ≠ XT and/or YS ≠ YT. This signifies that the domains may lack shared features, and the dimensions of these features may also differ. Therefore, heterogeneous TL is more challenging than homogeneous TL, as it necessitates bridging the gap between disparate features and their respective quantities. While knowledge transfer remains a possibility in heterogeneous TL, it becomes more challenging due to the need to translate valuable information originally represented in terms of the source domain into an appropriate format for the target domain. Heterogeneous TL is not always possible or advisable, as it can be more difficult to implement and may not lead to the same performance improvements as homogeneous TL.

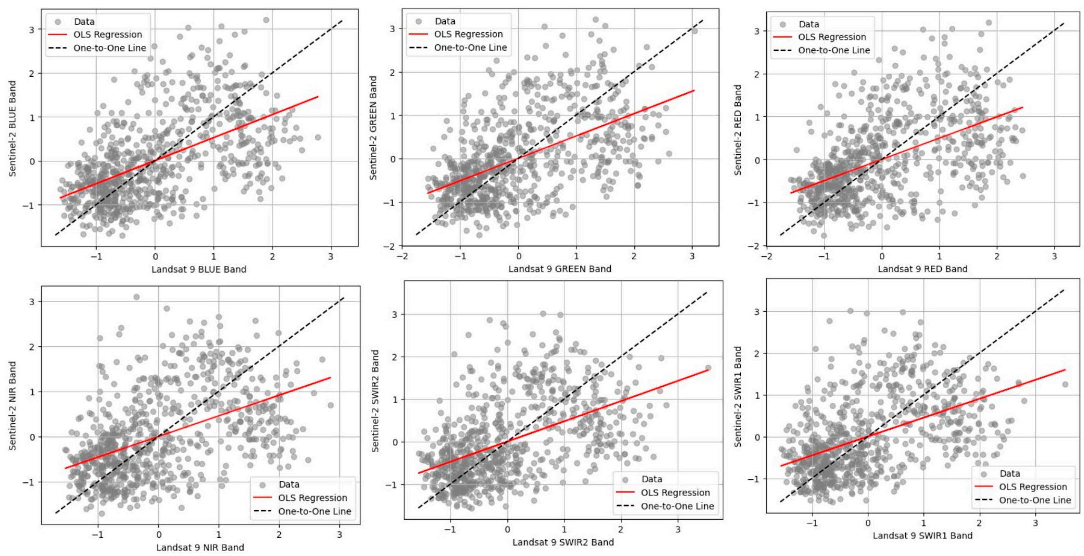

Figure 1. Comparing surface reflectance values from S2 and L9 data over the same geographic area. Common bands such as blue, green, red, NIR, SWIR-1, and SWIR-2 are typically selected in both domains. This uniform band selection ensures that the spectral information used for comparison is consistent across different data sources. In all plots, the x-axis represents the L_9 values, and the y-axis represents the S_2 values.

When the domains, features, or tasks exhibit substantial dissimilarity, a phenomenon known as negative transfer learning (TL) can occur [6]. Negative TL arises when the process of transferring knowledge from the source domain to the target domain has an adverse impact on the model's performance within the target domain. In such cases, it is often more prudent to opt for training a model from the ground up, utilizing advanced techniques like data augmentation, active/meta learning, and others, rather than attempting to transfer a model from the source domain. The different types of TL methods can be distinguished based on their feature space, the difference between the domains, and the tasks that the predictive models are intended to perform.

3. Methodology

3.1. Source and Target Data Preparation

The first step in TL is to pre-process the data and create the appropriate training and validating splits. S2 and L9 imageries are downloaded and saved. Then, only the bands (features) that are considered most relevant to the task are kept. These features, which are briefly described in Table 1, are the same for both domains. Once the final datasets are created, they are split into training and test subsets. A 70:30 data split is used, meaning that 70% of the data samples are used for training, and 30% are used for testing. Finally, the training and test sets are scaled. At this point, the data is fully processed, and the training and test datasets can be used to train and validate the predictive models. The entire data processing stage is performed on both the source and target domain data independently.

Table 1. Selected MSI bands.

	Sentinel-2		Landsat-9	
	Band Resolution	**CV**	**Band Resolution**	**CV**
	(m)	**(nm)**	**(m)**	**(nm)**
Blue	10 m	490 nm	30 m	452 nm
Green	10 m	560 nm	30 m	561 nm
Red	10 m	665 nm	30 m	665 nm
NIR	10 m	842 nm	30 m	865 nm
SWIR1	20 m	1610 nm	30 m	1609 nm
SWIR2	20 m	2190 nm	30 m	2200 nm

As discussed in [6–10], the key to achieving successful TL and avoiding negative effects is to discover and exploit shared underlying structures between DS (X; Y) and DT (X; Y).

3.2. Dataset Description

Our dataset comprises two subsets derived from satellite images acquired by S2 and L9, encompassing seven spectral bands (blue, green, red, near-infrared, shortwave infrared 1, shortwave infrared 2), along with calculated indices: NDVI (Normalized Difference Vegetation Index) and NDWI (Normalized Difference Water Index).

The S2 data were collected in 2023 using a single raster image, along with a cloud mask to detect and mask cloudy areas. In contrast, the L9 data were gathered over three years, from 2021 to 2023, using a series of five raster images, and they also incorporate a cloud mask for the identification and exclusion of cloudy regions. Each subset contains three CSV files, each with 10,000 samples, resulting in a total of 30,000 samples. The first file focuses on distinguishing olive-bearing land from other land types, featuring two classes: olive and non olive. The second file targets the differentiation between palm vegetation and other land cover types, comprising palm and non-palm classes. Lastly, the third file aims to detect the presence of buildings within land sections, with classes labeled as Building and non building. This dataset offers a valuable resource for land classification tasks, including land cover mapping, urban planning, and agriculture monitoring, harnessing the richness of spectral bands and computed indices to facilitate accurate and efficient land feature classification in various applications.

3.3. Model Training

After processing the data samples and constructing train and test datasets for both the source and target domains, the next step involves the creation and training of a neural network specifically designed for the source domain. However, to gain a comprehensive understanding of how the proposed solution operates, it is necessary to explain the neural network architectures selected to accomplish this task.

3.4. Neural Network Architectures

1D convolutional neural networks (CNNs) are used here as shown in Figure 2. They consist of two consecutive 1D convolutional layers, followed by one dropout and one max-pooling layers; then, features are flattened to be injected in classical classifiers multiple layer perceptron (MLP).

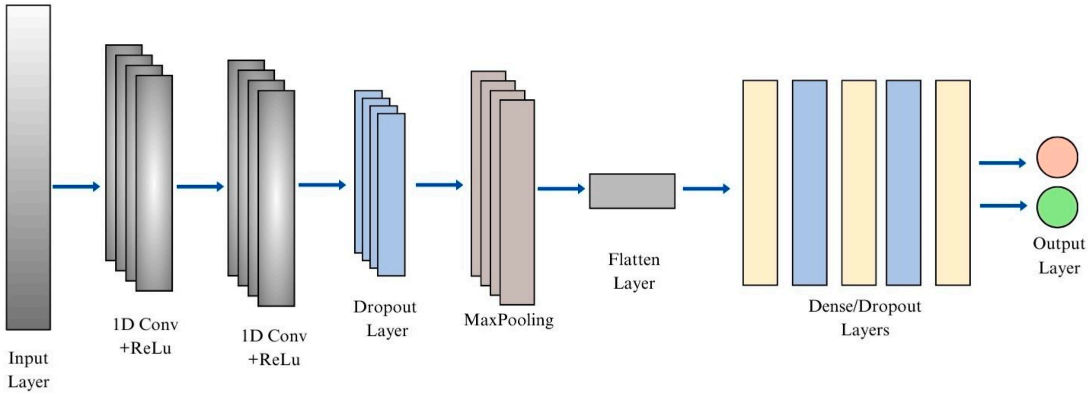

Figure 2. Convolutional neural-network model architecture. (Find the datasets and the code used in this study on the GitHub repository: https://github.com/mrouba/TL_RS (accessed on 1 November 2023)).

3.5. Source Model Training Process

CNN weights are randomly initialized via Xavier initialization [11]. All layers use rectified linear unit (Relu) as the activation function and Adam [12] as the optimizer with a learning rate of 0.001. The training was run for 100 epochs with batch size of 32.

3.6. Target Model Training (Fine-Tuning)

The TL idea is based on the conclusions discussed in [5], which state that neural networks have a tendency to acquire more general features in their initial layers and progressively more specific features in the later layers of their architecture. Consequently, even when the features received by the network differ, as long as there is a sufficient degree of similarity between the source domain (SD) and the target domain (TD), the neural network should have the capacity to adapt and leverage this shared information to successfully perform the assigned task. It is generally not advisable to remove and randomly initialize the weights of the input layer when aiming to maximize the transfer of information across domains.

The information contained in the first layer of a neural network can be pivotal for the effective functioning of subsequent layers. Altering or removing this information may lead to a significant decrease in the network's performance. In TL, the final classification layer of a pretrained CNN (SoftMax) is reconfigured to meet the new classification task, and the remaining layers are frozen during training and later used as feature extractors. However, the last two layers are unfrozen and trained to learn the new classification task. Once the TL is performed and the transferred model is capable of processing the input features of the target domain, the performance of the transferred model is evaluated. To do this, an exact copy of the transferred model is created, but with reinitialized weights. This allows the results obtained by the transferred model to be compared to the results that a model with the same architecture trained from scratch would produce.

4. Experimental Section

This project has been developed using Python and TensorFlow as the DL framework. The developed solution was coded in Jupyter Notebook, and executed with CPU Intel Core i7-9700k with 16 GB of RAM and GPU NVIDIA 1080-Ti.

4.1. MSI Homogenous TL

Sentinel-2A and Landsat-9 data were leveraged as the target and source domains to enhance the model's capabilities for classifying olive trees. First, a model was trained on Sentinel-2A and Landsat-9 data with a palm classification task [13,14]. This foundational training enabled the model to grasp the intricacies of identifying palm trees in satellite imagery. Then, the model was fine-tuned on Sentinel-2A and Landsat-9 data with an olive classification task. Fine-tuning the model with data from a source domain with remarkable similarity to the target domain was found to be strategic. By employing a source domain that mirrored the characteristics of the target domain, the fine-tuning process was highly effective, enabling the model to adapt quickly to the nuances of the olive classification task and resulting in a substantial boost in accuracy.

According to Table 2, impressive accuracy was achieved when our model for olive tree classification was fine-tuned using a source model trained for building classification. This experiment broadened our perspective on the capabilities of machine learning models by showing that the traditional boundaries between source and target tasks can sometimes be more fluid than expected. This prompted us to rethink our preconceived notions about domain dissimilarity in machine learning. The ability of our model to fine-tune effectively suggests that there may be latent similarities or transferable knowledge across seemingly distinct domains.

Table 2. Performance metrics for homogeneous transfer learning from different tasks using Landsat-9/Sentinel-2A data.

	Landsat-9				Sentinel-2			
	Accuracy	**Precision**	**Recall**	**F1-Score**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
TL (Palm $\rightarrow$ Olive)	0.98	0.98	0.98	0.98	0.95	0.95	0.95	0.95
TL (Buildings $\rightarrow$ Olive)	0.96	0.96	0.96	0.96	0.85	0.88	0.85	0.84

The distribution of the three classes (olive, building, and palm) in Sentinel and Landsat data was visualized using t-SNE plots in Figures 3 and 4, which condensed the multidimensional information into a clear two-dimensional representation. This revealed how the classes were distributed and related. The olive class was notably positioned in close proximity to the building and palm classes, suggesting that it shares certain common features or characteristics with the other two classes.

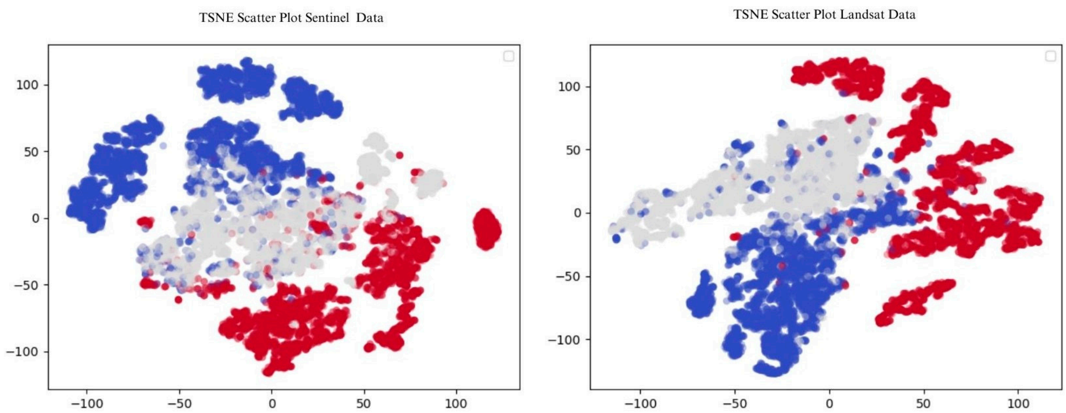

Figure 3. t-SNE visualization for the frequency vectors of the S2 dataset on right side and the L9 dataset on the left side, revealing olive in gray, palm in blue, and building in red.

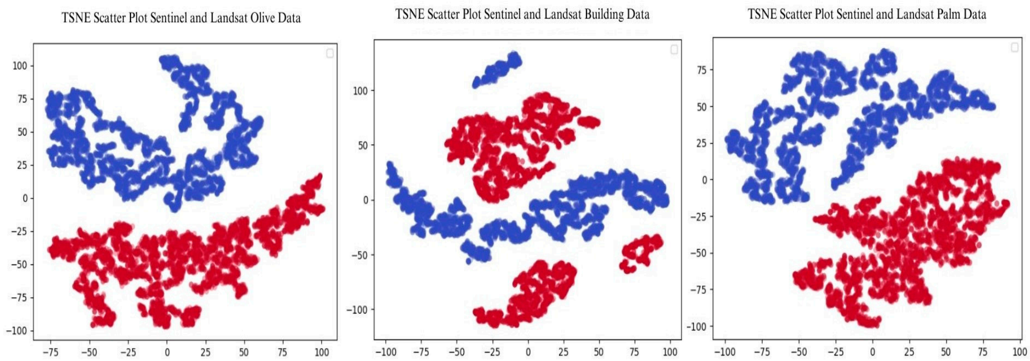

Figure 4. t-SNE visualization for the frequency vectors of L9 and S2 data; S2 in blue, and L9 in red.

4.2. MSI Heterogenous TL

A model for olive tree classification was fine-tuned on S2 data using a building/palm task classification model trained on S2 images as the source model. The goal was to create a robust model capable of accurately identifying olive trees in L9 data. However, the dissimilarity between the two domains (L9 for target model and S2 for source model) posed a significant challenge. Even though bands that were shared between the two domains were carefully selected, they exhibited a low correlation, making it difficult for the model to generalize effectively.

We undertook the task of resampling L9 data to match the spatial resolution of S2, which is 10 m. It is worth noting that L9 originally possessed a spatial resolution of 30 m. The resampling procedure involved the utilization of cubic resampling techniques, and the outcomes were truly noteworthy. The resampled dataset exhibited a remarkable enhancement in classification accuracy when compared to the original data. This discovery underscores the critical significance of carefully selecting appropriate spatial resolution images for RS applications, given their profound influence on the performance and adaptability of machine learning models.

According to Table 3, the utilization of L9 imageries as the source domain for fine-tuning on S2 data as the target domain yields a notably strong performance. This remarkable outcome can be primarily attributed to the quality of the L9 data. The high quality of the source data significantly contributes to the effectiveness of the TL process, providing a solid foundation for the model's adaptation to the target domain.

Table 3. Performance metrics for HTL from different tasks using L9/S2 data.

	Sentinel-2 to Landsat-9				Landsat-9 to Sentinel-2			
	Accuracy	**Precision**	**Recall**	**F1-Score**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
TL (Palm → Olive)	0.50	0.25	0.50	0.33	0.67	0.80	0.67	0.63
TL (Buildings → Olive)	0.50	0.25	0.50	0.33	0.86	0.89	0.86	0.86

Table 4 reveals the considerable improvements resulting from the resampling of L9 data to match the spatial resolution of S2 data in the realm of HTL. Accuracy, precision, recall, and F1-score metrics emphasize the substantial impact of spatial resolution alignment on model performance.

Table 4. Performance metrics for HTL from different tasks using L9 (resampled)/S2 data.

	Sentinel-2 to Landsat-9				Landsat-9 to Sentinel-2			
	Accuracy	**Precision**	**Recall**	**F1-Score**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
TL (Palm → Olive)	0.51	0.69	0.51	0.35	0.97	0.98	0.97	0.97
TL (Buildings → Olive)	0.50	0.25	0.50	0.33	0.86	0.89	0.89	0.85

The quality of the L9 data is a critical factor influencing the success of fine-tuning in the context of HTL. This observation holds true for both Tables 3 and 4. When L9 data exhibit higher quality compared to S2, they significantly contribute to the effectiveness of the TL process. This underscores the importance of having high-quality source data, as it provides a solid foundation for the model's adaptation to the target domain. Additionally, taking spatial resolution alignment into consideration, as demonstrated in Table 4, plays a vital role in enhancing model performance. Therefore, ensuring high-quality source data and thoughtful consideration of resolution are paramount for achieving superior performance in such transfer learning scenarios.

5. Conclusions

This study explores the concept of TL, emphasizing its vital role in domain and task considerations. It categorizes knowledge transfer into two forms: homogeneous TL and heterogeneous TL, illustrating how source–target domain similarities impact transfer effectiveness. Using S2 and L9 data, we explore these TL scenarios, highlighting the importance of task and source data choice for efficient transfer. Our work underscores the need for thoughtful domain and task selection to optimize TL outcomes. Additionally, our findings reveal the complexity of domain adaptation, showing that the reverse approach does not guarantee success. This is influenced by domain nuances, dataset characteristics, and model adaptability. Our study emphasizes the importance of strategic domain selection for effective TL. This is influenced by domain nuances, dataset characteristics, and model adaptability. Our study emphasizes the importance of strategic domain selection for effective TL.

6. Declaration of Competing Interest

This study explores the concept of Transfer Learning (TL), emphasizing its vital role in domain and task considerations. It categorizes knowledge transfer into two forms: homogeneous and heterogeneous TL, illustrating how source-target domains similarities impact TL effectiveness. Using S2 and L9 data, we explore these TL scenarios, highlighting the importance of task and source data choice for efficient TL. Our work underscores the need for thoughtful domain and task selection to optimize TL outcomes. Additionally, our findings reveal the complexity of domain adaptation, showing that the reverse approach doesn't guarantee success. This is influenced by domain nuances, dataset characteristics, and model adaptability. Our study emphasizes the importance of strategic domain selection for effective TL.

Author Contributions: Conceptualization, M.R., M.E.A.L.; software, M.R., M.E.A.L.; validation, M.R., M.E.A.L.; formal analysis, M.R., M.E.A.L.; investigation, M.R., M.E.A.L.; data curation, M.R., M.E.A.L.; writing—original draft preparation, M.R., M.E.A.L.; writing—review and editing, M.R., M.E.A.L.; visualization, M.R., M.E.A.L.; supervision, M.E.A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [CrossRef]
2. Day, O.; Khoshgoftaar, T.M. A survey on heterogeneous transfer learning. *J. Big Data* **2017**, *4*, 29. [CrossRef]
3. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *arXiv* **2014**, arXiv:1411.1792.
4. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis. (IJCV)* **2015**, *115*, 211–252. [CrossRef]

5. Wang, Z.; Dai, Z.; Póczos, B.; Carbonell, J. Characterizing and avoiding negative transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 11293–11302.
6. Zhao, H.; Liu, F.; Zhang, H.; Liang, Z. Convolutional Neural Network Based Heterogeneous Transfer Learning for Remote-Sensing Scene Classification. *Int. J. Remote Sens.* **2019**, *40*, 8506–8527. [CrossRef]
7. Ganin, Y.; Lempitsky, V. Unsupervised domain adaptation by backpropagation. *Int. Conf. Mach. Learn.* **2015**, *37*, 1180–1189.
8. Pires de Lima, R.; Marfurt, K. Convolutional neural network for remote-sensing scene classification: Transfer learning analysis. *Remote Sens.* **2020**, *12*, 86. [CrossRef]
9. Long, M.; Cao, Y.; Wang, J.; Jordan, M.I. Learning transferable features with deep adaptation networks. In Proceedings of the 32nd International Conference on International Conference on Machine Learning, Lille, France, 6–11 July 2015.
10. Tzeng, E.; Hoffman, J.; Saenko, K.; Darrell, T. Adversarial discriminative domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2017, Honolulu, HI, USA, 21–26 July 2017; Volume 1, pp. 7167–7176.
11. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.
12. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
13. Alhammadi, M.S.; Glenn, E.P. Detecting Date Palm Trees Health and Vegetation Greenness Change on the Eastern Coast of the United Arab Emirates Using SAVI. *Int. J. Remote Sens.* **2008**, *29*, 1745–1765. [CrossRef]
14. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 778–782. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.