Exploring selective image matching methods for zero-shot and few-sample unsupervised domain adaptation of urban canopy prediction

Francis, John[1]; Law, Stephen[2]

[1] The Alan Turing Institute, London, UK
[2] Department of Geography, University College London

Introduction

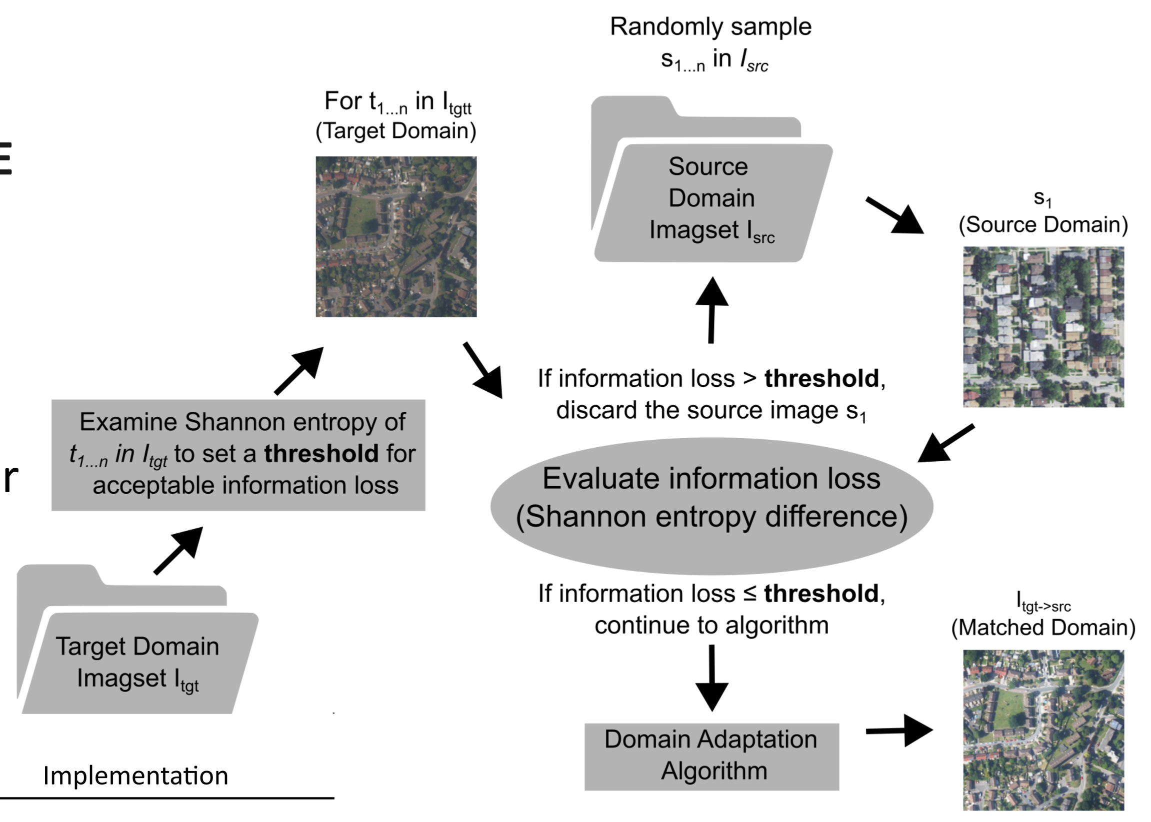
Machine learning algorithms utilizing remote sensing data have been able to estimate numerous features of the environment. While these algorithms are often able to produce reliable estimates for a particular geographical area, when images from a new environment, or images captured by different instruments at potentially different resolutions are introduced, the reliability of these algorithms falters [1].

UDA techniques, which can be categorised into data-based and model-based approaches, leverage models trained in one source domain to perform accurately on a different, unlabeled target domain. While model-based adaptations modify the underlying source model to improve its performance on the target domain, primarily through adversarial or self-training methods, data-based adaptation focuses on altering images either in the source or target domain so that the two are statistically similar.

Here, we examine a multi-task UNet model which simultaneously predicts urban canopy cover as well as canopy height from RGB aerial imagery. Instead of fine-tuning or training a domain-adaptive-classifier, our novelty lies in experimenting with a variety of simple data-based UDA approaches in a zero-shot setting, which doesn't require any training, or with a small amount of fine-tuning.

Data and Methods

Data for this study includes 812 1m resolution RGB images from each of Chicago, USA (**SOURCE** domain) and London, UK (**TARGET** domain). Images were aligned with LiDAR point clouds to generate estimates of ground truth canopy cover and height for evaluation. We utilize a randomized image matching process, and similar to [2] we use the Shannon entropy measure to compare pairs of images, to ensure a viable match leading to minimal information loss.



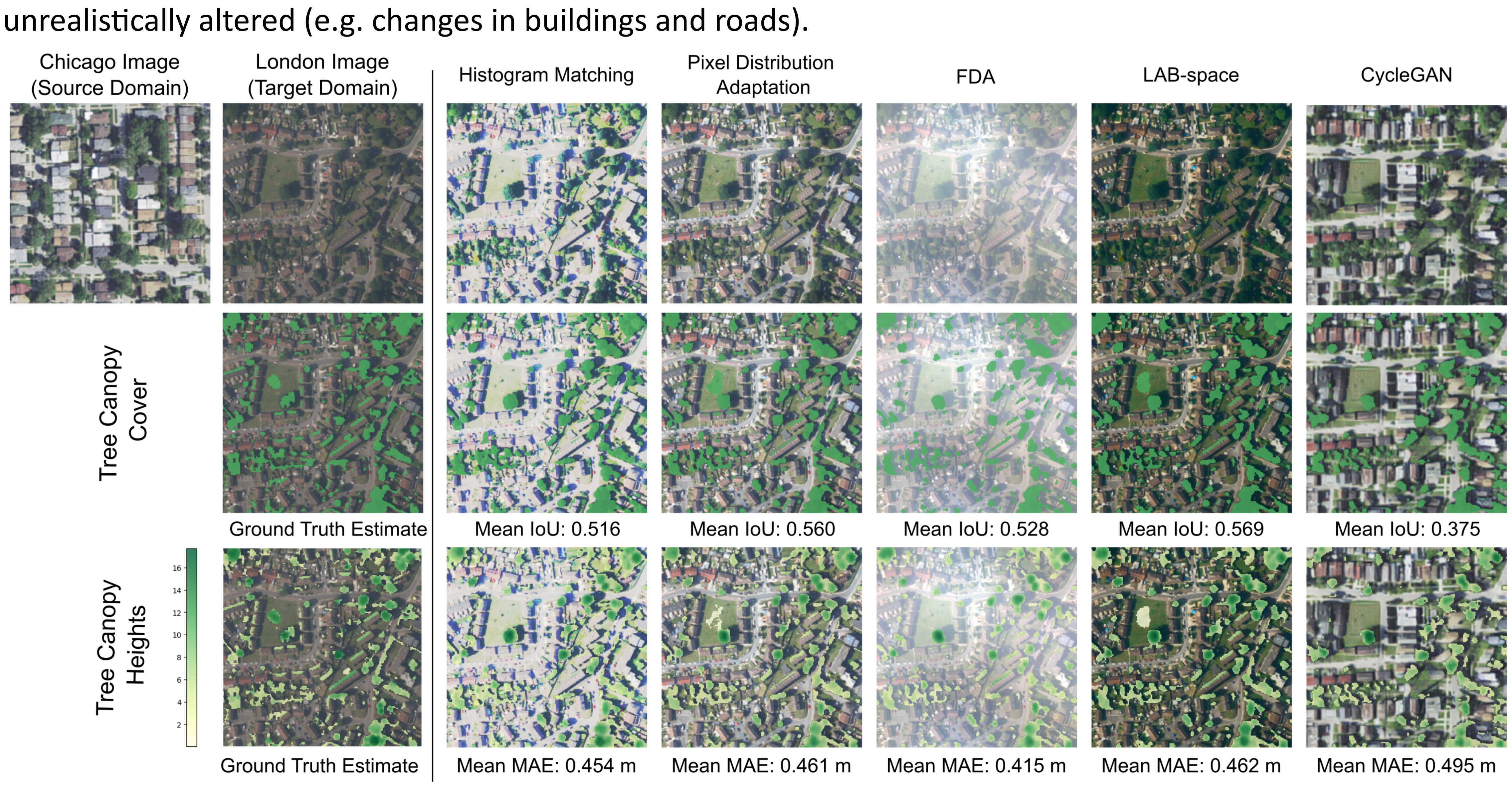
Unsupervised Domain Adaptation methods

Method	Description	Implementation
Histogram Matching in RGB space (HM)	Adjusts an image so that its cumulative histogram at the pixel level matches that of another image. The adjustment is applied separately for each RGB channel of the image.	scikit-image
Histogram Matching in LAB space (HM-LAB)	Similar to hisotogram matching, expect images are first transformed into CIE Lab colorspace before matching.	scikit-image
Pixel Distribution Adaptation (PDA)	Aligns pixel value distributions of source and target images by fitting a simple PCA transformation to both images, and then applying the inverse transformation of the source image to the target image.	Albumentations python package
Fourier Domain Adaptation (FDA)	FDA manipulates the frequency components of images to reduce the domain gap between source and target datasets. FDA achieves domain alignment by swapping low-frequency components of the Fourier transform between the source and target images.	Albumentations python package

We tested four simple data-based image matching approaches (left). Additionally one image-to-image translation model, CycleGAN [3], as well as the raw target images, were tested as baselines. We evaluate each of these methods in a zero-shot setting, testing a pre-trained UNet [4] in the source domain on the transformed images, as well as on fine-tune separate versions of the algorithm using the transformed datasets.

Results

Each of the data adaptation methods perceptually altered the pixel intensity of the target images. PDA in particular made the target image slightly brighter and more similar to the source domain. Conversely, although CycleGAN shifted the target image closest to the source domain, the semantics of the image have also been unrealistically altered (e.g. changes in buildings and roads).



In a zero-shot setting, PDA produced the best results for the canopy cover task. Conversely, for the canopy height task, the most performative method was FDA, which performed even better than the results of the algorithm on the source domain images. HM, PDA and CycleGAN all performed worse on the canopy height task than the original target images without any image transformations.

In a small sample fine-tuning setting, PDA was the best-performing method for both the canopy cover and canopy height task. All of the simple image matching methods outperformed the algorithm fine-tuned on the un-transformed target images for the canopy height task, while CycleGAN performed worse than the un-transformed images on both tasks.

Zero-shot UDA				UDA W/Fine-tuning (small samples)			
Data	Method	mIoU	MAE (m)	Data	Method	mIoU	MAE (m)
Source	NA	0.665	0.6538	Source	NA	0.6650	0.6538
Target	None	0.3709	0.8321	Target	None	0.6885	0.6340
Target	HM	0.4784	1.0302	Target	HM	0.6811	0.5944
Target	FDA	0.3659	**0.5864**	Target	FDA	0.6917	0.5839
Target	PDA	**0.5131**	0.9312	Target	PDA	**0.7014**	**0.5547**
Target	HM-LAB	0.3695	0.6062	Target	HM-LAB	0.6862	0.5627
Target	CycleGAN	0.401	1.2085	Target	CycleGAN	0.5714	0.8519

Conclusions

Overall, we find that selective-aligned simple image matching approaches can be an effective way to adapt a multi-task deep learning algorithm to a new geographic setting, especially when a small amount of fine-tuning is possible. FDA and PDA obtained the best results on our two tasks of canopy cover and canopy height prediction, although simple HM techniques still provided moderate boosts over not using any domain adaptation method.

Simple data-based domain adaptation approaches such as these provide easy, low-resource methods to enhance the utility of pre-trained models which utilize remote sensing imagery, when these algorithms are exposed to new, unseen domains. In many scenarios, fine-tuning or completely retraining an algorithm whenever it is exposed to a new domain can be either difficult to implement or prohibitively expensive. Researchers which are attempting to utilize pre-trained remote sensing algorithms in a new geography with different data sources should explore simple image matching techniques, before undertaking energy-extensive and costly retraining.

Future work would benefit from exploring domain shifts across a wider variety of geographic locales, more advanced model-agnostic methods to improve the results in a zero-shot setting, incorporating location information into the adaptation, or adaptation with large scale satellite image foundation models.

References

1. Devis Tuia, Claudio Persello, and Lorenzo Bruzzone. Domain adaptation for the classification of remote sensing data: An overview of recent advances. IEEE Geoscience and Remote Sensing Magazine, 4(2):41–57, June 2016.
2. Can Yaras, Kaleb Kassaw, Bohao Huang, Kyle Bradbury, and Jordan M Malof. Randomized histogram matching: A simple augmentation for unsupervised domain adaptation in overhead imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023.
3. Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision, pp. 2223–2232, 2017
4. John Francis, Mathias Disney, and Stephen Law. Monitoring canopy quality and improving equitable outcomes of urban tree planting using lidar and machine learning. Urban Forestry Urban Greening, 89:128115, 2023.
5. Yanchao Yang and Stefano Soatto. "Fda: Fourier domain adaptation for semantic segmentation." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4085-4095. 2020.

JFrancis@turing.ac.uk

Acknowledgements
This work was supported by Towards Turing 2.0 under the EPSRC Grant EP/W037211/1, the Ecosystem Leadership Award under the EPSRC Grant EP/X03870X/1 & The Alan Turing Institute.

Data Sources
https://digimap.edina.ac.uk/aerial
https://environment.data.gov.uk/survey
https://naip-usdaonline.hub.arcgis.com/