Seasonal Domain Shift in the Global South: Dataset and Deep Features Analysis

Georgios Voulgaris, Andrew Phillipides, Jonathan Dolley, Jeremy Reffin, Fiona Marshall, Novi Quadrianto.

Department of Informatics, University of Sussex, Brighton, UK

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

ABSTRACT

Domain shifts during seasonal variations are an important aspect affecting the robustness of aerial scene classification and so it is crucial that such variation is captured within aerial scene datasets. This is more evident in geographic locations in the Global South, where aerial coverage is scarcer and the rural and semi-urban landscape varies dramatically between wet and dry seasons. As current datasets do not offer the ability to experiment with domain shifts due to seasonal variations, this work proposes a labelled dataset for classifying land use from aerial images, comprising both wet and dry season data from Ghaziabad in India. Moreover, we conduct a thorough investigation into how image features, namely colour, shape, and texture, influence the accuracy of scene classification. We demonstrate that a combination of an architecture that extracts salient features, with the implementation of a larger receptive field improves classification performance when applied to both shallow or deep architectures by extracting invariant feature representations across domains.

Colour changes in wet (coloumns 1 and 3) vs dry (coloumns 2 and 4) seasons

Methods

- **Dataset**: Aerial images obtained from Google Maps with a spatial resolution of 0.6 meters per pixel. Image size: 256 x 256 pixels.

Mixed, Wet, Dry Season Scene Dataset.

Class	Wet	Dry	Mixed
Agriculture	3,278	1,772	5,621
Barren Land	632	301	924
Brick Kilns	968	292	1,008
Forest	936	713	1,516
Orchard Industry	711	726	1,144
Roads	-	-	565
Urban	1,241	566	1,420
Urban Grn Space	584	306	838
Water	404	662	141

- **Dilated Convolution**: The proposed architecture utilises dilated convolutional layer with varying dilated rates (d = 1, 3, 6, 9) to localise scene-related regions. By summing the averaged localisation maps with the localisation map of the smallest receptive field, the network avoids overemphasising scene-irrelevant areas [1].

$$H = H_0 + \frac{1}{n_d}\sum_{i=1}^{n_d} H_i$$

- **Gabor CNN Parameters:** Gabor function parameters are updated during backpropagation [2].

  Gabor function:

$$g(x, y, \omega, \theta, \psi, \sigma) = exp - \left(\frac{x'^2 + y'^2}{2\sigma^2}\right) \cos(\omega x' + \psi)$$

$$x' = x\cos\theta + y\sin\theta; \quad y' = -x\sin\theta + y\cos\theta$$

  Where, (x, y) stands for the pixel spatial domain position, θ the filter orientation, σ the standard deviation, and ω the frequency. Following Meshgini et al. [3] we use a bank of Gabor filters with frequencies ω_n and orientations θ_m:

$$\omega_n = \frac{\pi}{2}\sqrt{2}^{-(n-1)} \text{ where, } n = 1, 2, \ldots, 5$$

$$\theta_m = \frac{\pi}{8}(m - 1) \text{ where, } m = 1, 2, \ldots, 8$$

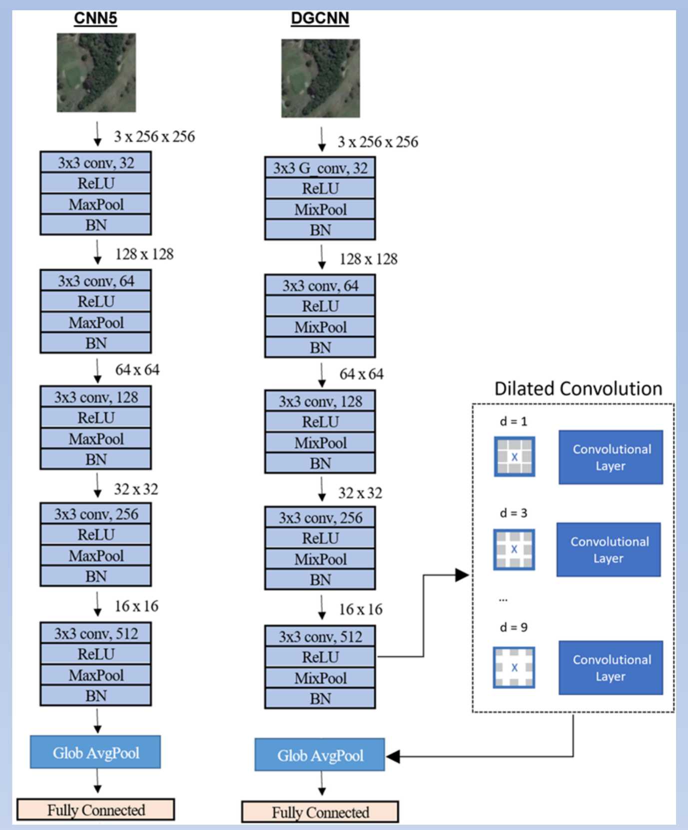
  Gabor layer weights initialization:

  - standard deviation to $\sigma = \frac{\pi}{\omega}$.
  - ψ is set by uniform distribution $(0, \pi)$.

- **Mix Pooling**: On the lowest layer Max pooling was applied, and as increment to highest layers, it was decreased by 0.2:

$$f_{min}(x) = a_l \cdot f_{max}(x) + (1 - a_l) \cdot f_{avg}(x) \quad [4]$$

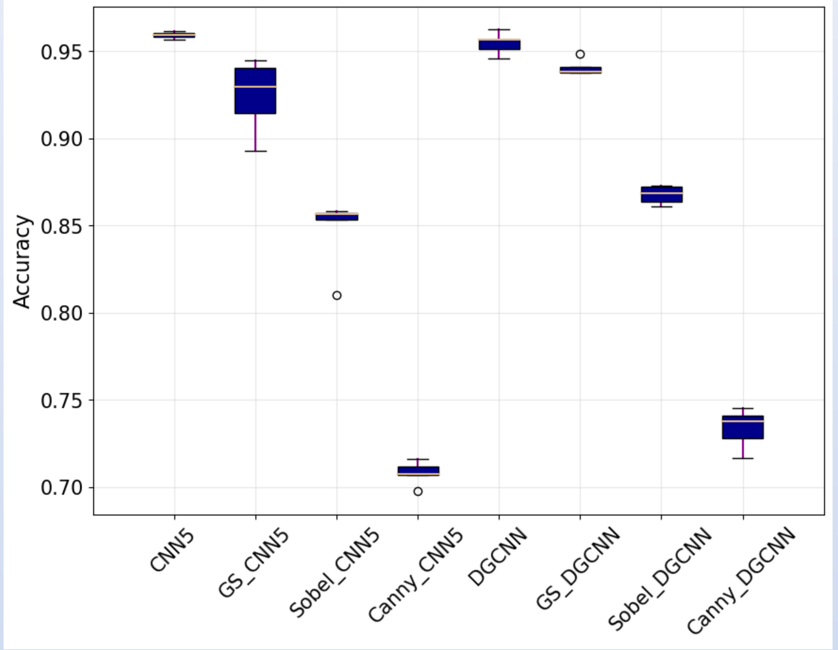
- **Models:**



PRELIMINARY RESULTS

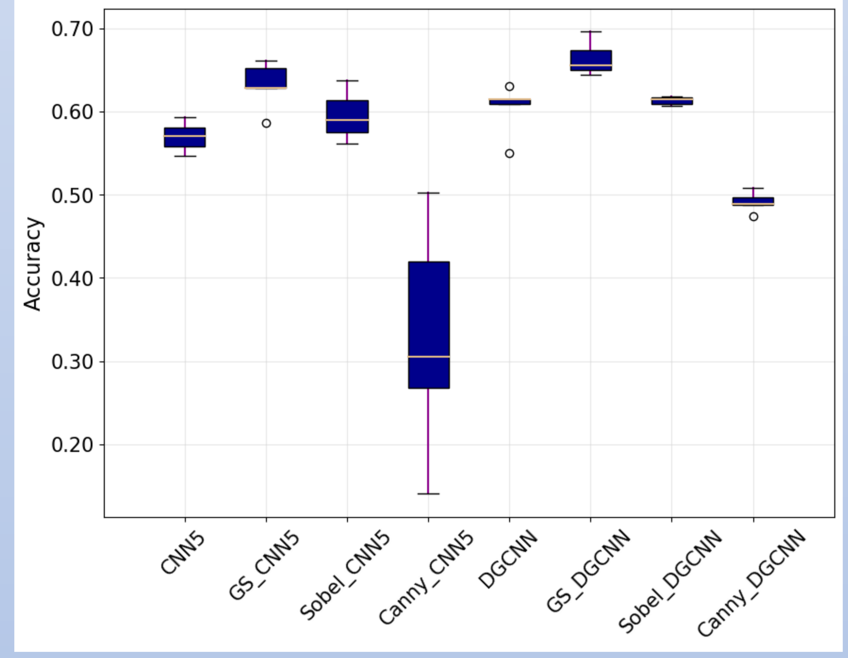
Mixed Season Dataset

Feature bias performance on the mixed season scene dataset. Dilated Gabor-CNN is less biased towards colour features and extracts more salient features.

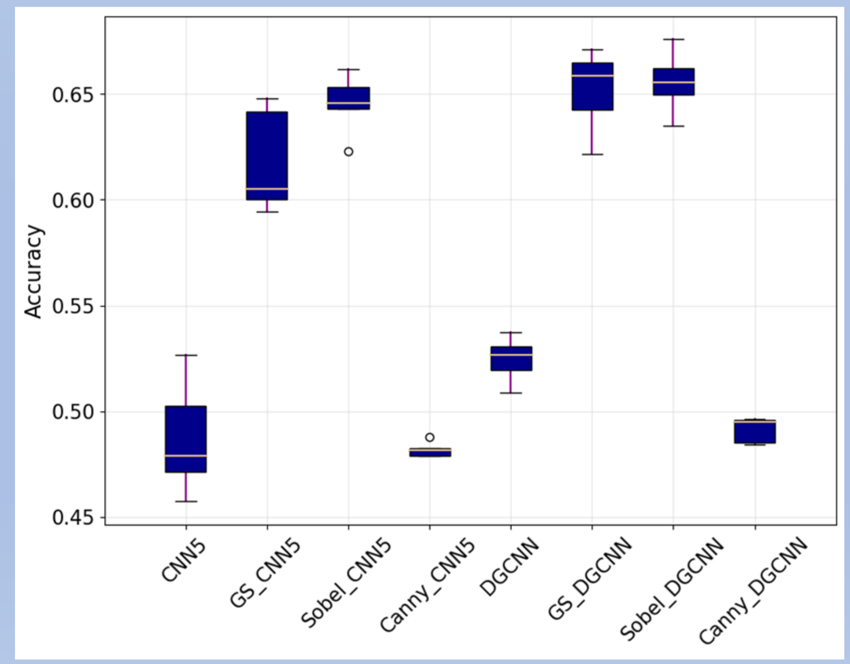


Seasonal Domain Shift

Train: Dry Season Data, Test: Wet Season Data Feature Bias Performance. The proposed architecture outperforms the original by 6% on colour and by 10% on grey-scale images.

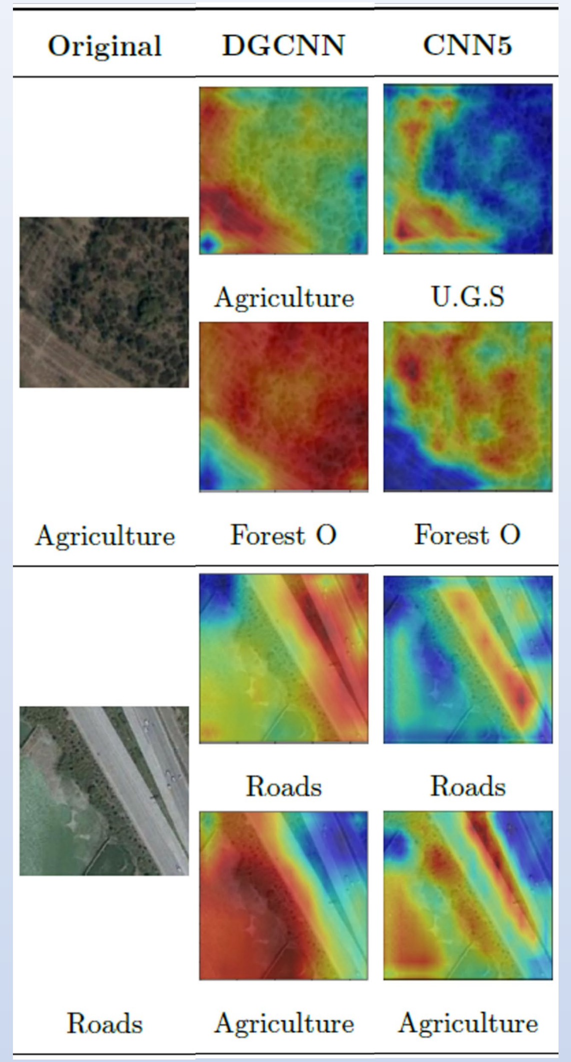


Train: Wet Season Data, Test: Dry Season Data Feature Bias Performance. The proposed architecture outperforms the original by 6% on colour and by 10% on gray scale images.



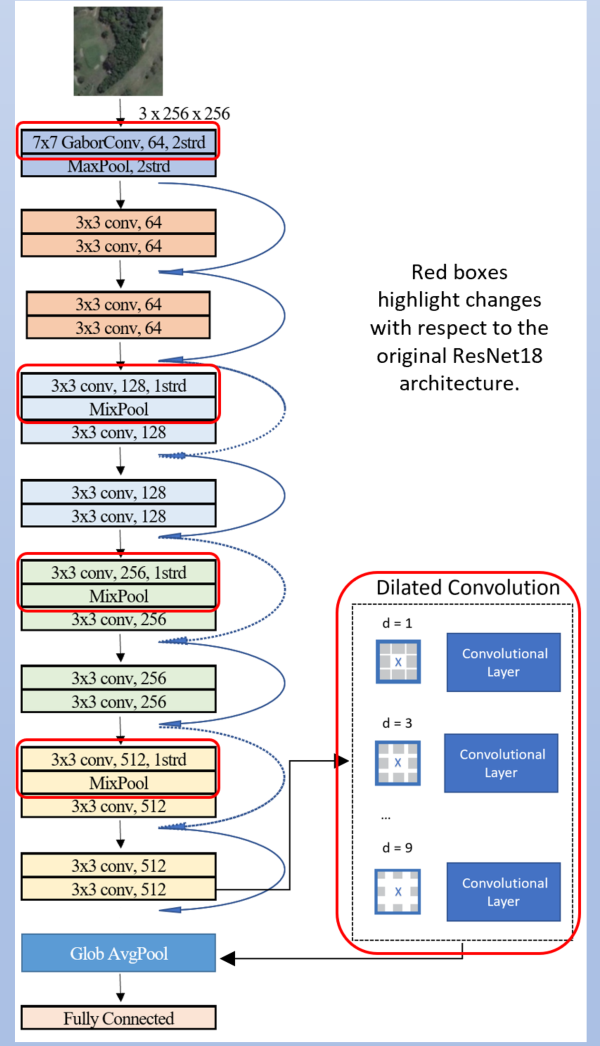
Class Activation Mapping

Two most confident predictions. For each original image, the top row displays the most confident prediction and the second row the second most confident prediction. The proposed model DGCNN, predicts correctly both scenes present in the image by focusing on more relevant parts, compared to the original CNN5 architecture.

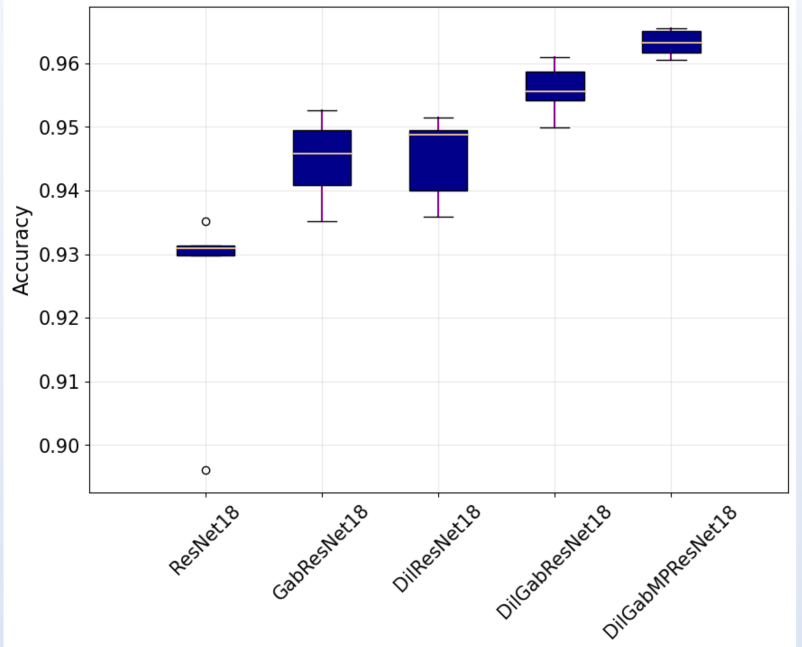


Ablation Study

Proposed model: implemented a Gabor convolutional layer, residual blocks 2, 3, and 4: replaced stride 2 convolutional layers with 1 stride while added a mixture of maximum & average pooling layers and added a dilated convolutional layer.

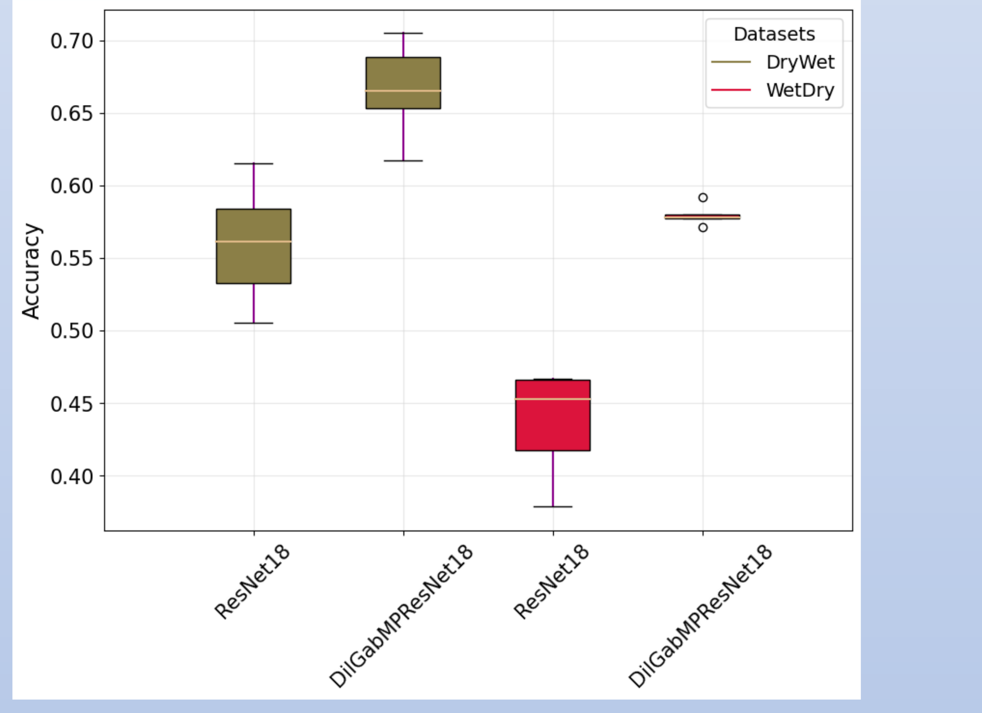


Models under review (from left to right): Original ResNet18, Gabor convolutional layer ResNet18, Dilated convolutional layer ResNet18, a combination of Gabor and dilated convolutional layer ResNet18, and a combination of Gabor, a mixture of max/average pooling layers and a dilated convolutional layer ResNet18 architectures trained on colour Mixed Season Data. The proposed architecture boosts performance by 3.7%.



Seasonal Domain Shift

Original Resnet18 and Gabor convolutional with a mixture of max/average pooling layers and a dilated convolutional layer trained with Dry and tested on Wet (gold); trained with Wet and tested on Dry, (red) colour Seasonal data. The proposed architecture boosts performance by 13% in seasonal domain-shifted data.



Conclusion

This work presents a novel dataset of aerial images that captures seasonal variations between wet and dry seasons, which impacts land use classification. We performed an aerial scene analysis on the proposed dataset and showed that classifiers tend to rely more on texture than colour or shape features. We therefore propose a novel architecture that combines salient feature extraction, with a wider receptive field to extract invariant feature representations and improve performance in the presence of domain shifts due to seasonal variations.

REFENCES & ACKNOWLEDGEMENTS

[1] Yunchao Wei, Huaxin Xiao, Honghui Shi, Zequn Jie, Jiashi Feng, and Thomas S Huang. Revisiting dilated convolution: A simple approach for weakly-and semi-supervised semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7268–7277, 2018.

[2] Andrey Alekseev and Anatoly Bobe, "Gabornet: Gabor filters with learnable parameters in deep convolutional neural network," in International Conference on Engineering and Telecommunication, 2019, pp. 1–4.

[3] Saeed Meshgini, Ali Aghagolzadeh, and Hadi Seyedarabi, "Face recognition using gabor filter bank, kernel principle component analysis and support vector machine," International Journal of Computer Theory and Engineering, vol. 4, no. 5, pp. 767, 2012.

[4] Chen-Yu Lee, Patrick W Gallagher, and Zhuowen Tu, "Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree," in Artificial intelligence and statistics. PMLR, 2016, pp. 464–472.