

---

# Rotation Equivariant Deforestation Segmentation and Driver Classification

---

**Joshua Mitton\***

School of Computing Science  
Glasgow University  
Glasgow, UK  
j.mitton.1@research.gla.ac.uk

**Roderick Murray-Smith**

School of Computing Science  
Glasgow University  
Glasgow, UK  
roderick.murray-smith@glasgow.ac.uk

## Abstract

Deforestation has become a significant contributing factor to climate change and, due to this, both classifying the drivers and predicting segmentation maps of deforestation has attracted significant interest. In this work, we develop a rotation equivariant convolutional neural network model to predict the drivers and generate segmentation maps of deforestation events from Landsat 8 satellite images. This outperforms previous methods in classifying the drivers and predicting the segmentation map of deforestation, offering a 9% improvement in classification accuracy and a 7% improvement in segmentation map accuracy. In addition, this method predicts stable segmentation maps under rotation of the input image, which ensures that predicted regions of deforestation are not dependent upon the rotational orientation of the satellite.

## 1 Introduction

Deforestation has been greatly accelerated by human activities with many drivers leading to a loss of forest area. Deforestation has a negative impact on natural ecosystems, biodiversity, and climate change and it is becoming a force of global importance (Foley et al., 2005). Deforestation for palm plantations is projected to contribute 18-22% of Indonesia’s CO<sub>2</sub>-equivalent emissions (Carlson et al., 2013). Furthermore, deforestation in the tropics contributes roughly 10% of annual global greenhouse gas emissions (Arneth et al., 2019). In addition, over one quarter of global forest loss is due to deforestation with the land being permanently changes to be used for the production of commodities, including beef, soy, palm oil, and wood fiber (Curtis et al., 2018). Climate tipping points are when a small change in forcing, triggers a strongly nonlinear response in the internal dynamics of part of the climate system (Lenton, 2011). Deforestation is one of the contributors that can cause climate tipping points (Lenton, 2011). Therefore, understanding the drivers for deforestation is of significant importance.

The availability and advances in high-resolution satellite imaging have enabled applications in mapping to develop at scale (Roy et al., 2014; Verpoorter et al., 2012, 2014; Janowicz et al., 2020; Karpatne et al., 2018). A range of prior works have used decision trees, random forest classifiers, and convolutional neural networks for the task of classifying and mapping deforestation drivers (Phiri et al., 2019; Descals et al., 2019; Poortinga et al., 2019; Hethcoat et al., 2019; Sylvain et al., 2019; Irvin et al., 2020). However none of these previous methods leverage advances in group equivariant convolutional networks (Cohen & Welling, 2016a,b; Weiler & Cesa, 2019) and as such the methods are not stable with respect to transformations that would naturally occur during the capture of such data.

---

\*<https://github.com/JoshuaMitton/Rotation-Equivariant-Deforestation-Segmentation>

In this work we train models to classify drivers of deforestation and generate a segmentation map of the deforestation area. For this we build a convolutional and group equivariant convolutional model to assess the impact on classification accuracy and both segmentation accuracy and stability of the segmentation maps produced. We show that not only does the group equivariant model, with translation and rotation equivariant convolutions, improve classification and segmentation accuracy, but it has the desired property of stability of the segmentation map under natural transformations of the data capture method, namely rotations of the satellite imaging.

## 2 Equivariance

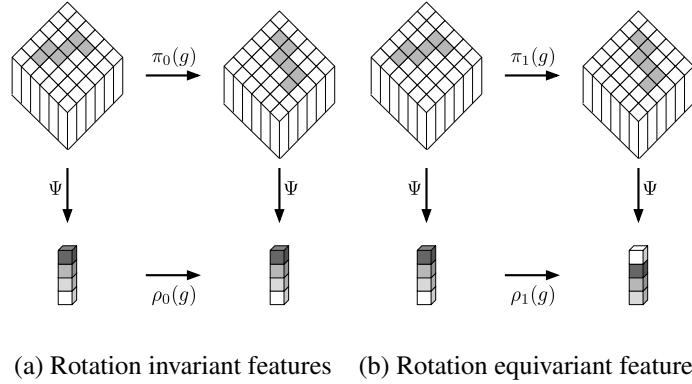


Figure 1: An image is rotated by  $g$  using  $\pi(g)$ , where  $\pi$  is some transformation law and  $g$  is any angle of rotation. The filters,  $\Psi$ , of the layer produce some output features, here a single fiber is shown. The representation  $\rho$  specifies how the feature vectors transform. (a) The representation  $\rho_0$  is the trivial representation, where  $\rho_0(g) = 1 \forall g \in G$ . This is used for scalar features that do not change under rotation, for example, The three RGB channels of an image are each a scalar feature and they do not mix under rotation. Therefore, typically, the input representation is the direct sum,  $\bigoplus_{i=1}^3 1 = \text{id}_{3 \times 3}$ , of three trivial representations. (b) The representation  $\rho_1$  is here used to represent the regular representation, where  $\rho_1(g) = g \forall g \in G$ . In this example the image is rotated by  $90^\circ$  which corresponds to a cyclic shift of the features in the output fiber.

A CNN is, in general, comprised of multiple convolutional layers, alongside other layers. These convolutional layers are translation equivariant. This means that if the input signal is translated, the resulting output feature map is translated accordingly. Translation equivariance is a useful inductive bias to build into a model for image analysis as it is known that there is a translational symmetry within the data, i.e. if an image of an object is translated one pixel to the left, the image is still an image of the same object. This translational symmetry can be expressed through the group  $(\mathbb{R}^2, +)$  consisting of all translations of the plane  $\mathbb{R}^2$ .

This leads us to consider if the data has additional symmetries, such that we can look at these symmetry groups and utilise them in a model. Steerable CNNs define feature spaces of steerable feature fields  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^c$ , where a  $c$ -dimensional vector  $f(x) \in \mathbb{R}^c$  is linked to each point  $x$  of the bases space (Cohen & Welling, 2016b). Steerable CNNs are equipped with a transformation law that specifies how the features transform under actions of the symmetry group. The transformation law is fully characterized by the group representation  $\rho$ . A group representation  $\rho : G \rightarrow \text{GL}(\mathbb{R}^c)$  specifies how the channels,  $c$ , of the feature vector,  $f(x)$ , mix under transformations. For a network layer to be equivariant it must satisfy the transformation law, see Figure 1. This places a constraint over the kernel, reducing the space of permissible kernels to those which satisfy the equivariance constraint. As the goal is to build linear layers that combine translational symmetry with a symmetry of another group for use in a model, the vector space of permissible kernels forms a subspace of that used in a conventional CNN. This increases the parameter efficiency of the layers, similar to how a CNN increases parameter efficiency over an MLP (Weiler & Cesa, 2019).

One particular group of interest for satellite imagery is the orthogonal group  $O(2) = \{O \in \mathbb{R}^{2 \times 2} | O^T O = \text{id}_{2 \times 2}\}$ . The orthogonal group consists of all continuous rotations and reflections leaving the origin invariant. In addition to the orthogonal group, the cyclic group,  $C_N$ , and the

dihedral group,  $D_N$ , consisting of discrete rotations by angles of multiples of  $\frac{2\pi}{N}$  and in the case of the dihedral group reflections also. These rotational symmetries are of interest for analysing satellite imagery as there is no global orientation of the images collected, i.e. if an image of a forest is captured it is still the same image of the same forest if it is rotated by an angle or reflected.

### 3 Methods

The dataset used is the same as that used by Irvin et al. (2020), where forest loss event coordinates and driver annotations were curated by (Austin et al., 2019). Random samples of primary natural forest loss events were obtained from maps published by Global Forest Change (GFC) at 30m resolution from 2001 to 2016. These images were annotated by an expert interpreter (Austin et al., 2019). The drivers are grouped into categories determined feasible to identify using 15m resolution Landsat 8 imagery, while ensuring sufficient representation of each category in the dataset (Irvin et al., 2020). The mapping between expert labelled deforestation driver category and driver group used as a classification target is provided in Table 3. The dataset consists of 2,756 images, segmentation maps, and class labels; we follow the training/validation/testing set splits as provided by Irvin et al. (2020).

We use a U-Net (Ronneberger et al., 2015) architecture for the task of segmentation and attach an MLP to the lowest dimensional feature space for classification. In one model we use translation equivariant convolutional layers, while in the other we use translation rotation equivariant convolutional layers. For the rotation equivariant version we choose the group  $C_8$  of discrete rotations by  $45^\circ$  as the symmetry group. The input to the model is therefore three trivial representations, while hidden layers are multiple regular representations of the group, chosen similarly to the size of feature spaces in the non-rotation equivariant model, and the output is a single trivial representation. An example of how a trivial representation and a regular representation transform the output feature space is given in Figure 1 (a) and (b) respectively. Building a model in this way will ensure that the output segmentation map is stable under  $45^\circ$  rotations of the input image.

### 4 Results

The model trained with rotation equivariance outperforms the non rotation equivariant model for classification of the drivers of deforestation, shown in Table 1. Given that the convolutional kernels are constrained to be rotation equivariant in the better performing model it is possible for the model to use the features more efficiently and hence model parameters are not used learning similar features at different orientations. As a result the model is able to better distinguish between the different deforestation drivers. In addition to classification accuracy, the rotation equivariant model achieves better test segmentation accuracy, demonstrated in Table 2. One cause of this benefit is that the model can share learned segmentation features across different orientations that occur across the different images in the dataset.

Table 1: Comparison between a model with translation equivariant convolutions and a model with both translation and rotation equivariant convolutions. Results are displayed as percentages for the classification accuracy of driver of deforestation.

Model	Train	Validation	Test	Rotated Test
UNET - CNN	90.3	60.6	57.9	56.3
UNET - C8 Equivariant	82.7	67.1	63.0	64.3

Table 2: Comparison between a model with translation equivariant convolutions and a model with both translation and rotation equivariant convolutions. Results are displayed as percentages for the segmentation accuracy of per pixel prediction averaged between the true deforestation and non-deforestation areas to account for the class imbalance towards non-deforestation areas.

Model	Train	Validation	Test	Rotated Test
UNET - CNN	72.9	68.7	67.8	67.9
UNET - C8 Equivariant	84.1	71.3	72.3	72.3

Furthermore, the segmentation map predictions for the non rotation equivariant model and rotation equivariant models are shown to compare the stability of segmentation under rotation in Figure 2. This highlights, in Figure 2, that the segmentation map prediction for the non-rotation equivariant model changes as the image is rotated, which would be highly undesirable if used in practice as the rotation orientation of the satellite should not effect the segmentation map prediction of deforestation. On the other hand, the rotation equivariant models segmentation map prediction is stable under rotation, which is a desirable property of the model.

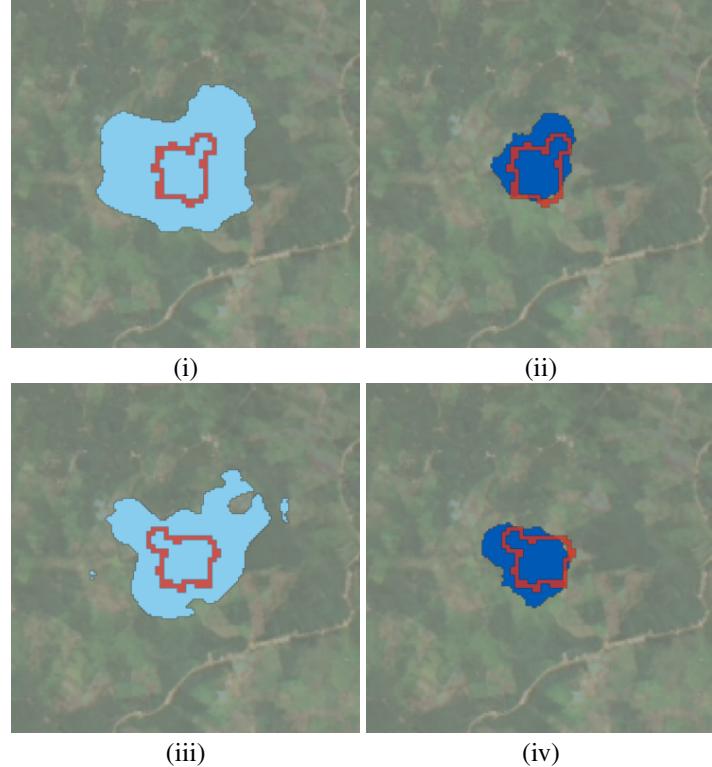


Figure 2: A comparison of predicted segmentation maps under rotation for both the non-rotation equivariant model and the rotation equivariant model. The original image is shown in (i) and (ii) with the edge of the true segmentation map in red. (i) shows the predicted segmentation map for the non-rotation equivariant model in light blue. (ii) shows the predicted segmentation map for the rotation equivariant model in dark blue. The 90° rotated image is shown in (iii) and (iv) with the edge of the true segmentation map in red. (iii) shows the predicted segmentation map for the non-rotation equivariant model in light blue. (iv) shows the predicted segmentation map for the rotation equivariant model in dark blue.

## 5 Conclusion

We develop a U-Net style model for classification and segmentation of deforestation that makes use of translation rotation equivariant convolutions. To the best of our knowledge this is the first study to make use of rotation equivariance in deforestation segmentation. The improved weight sharing through consideration of known symmetries in the data improves the classification accuracy of the model by 9%. Furthermore, the rotation equivariant model predicts segmentation maps that are stable under rotation. In a practical application of this model this would ensure that deforestation segmentation would not be dependent upon the rotational orientation of the satellite, which does not hold true for other models. Finally, the rotation equivariant model is 7% more accurate than the non-rotation equivariant model for the segmentation maps it produces when compared to ground truth segmentation. The improvement gain in both classification and segmentation of deforestation drivers will allow for conservation and management policies to be implemented more routinely based on model predictions from satellite data

## 6 Acknowledgements

Joshua Mitton is supported by a University of Glasgow Lord Kelvin Adam Smith Studentship. Roderick Murray-Smith acknowledges funding from the QuantIC project funded by the EPSRC Quantum Technology Programme (grant EP/MO1326X/1) and the iCAIRD project, funded by Innovate UK (project number 104690). Roderick Murray-Smith acknowledges funding support from EPSRC grant EP/R018634/1, Closed-loop Data Science.

## References

Almut Arneth, Fatima Denton, Fahmuddin Agus, Aziz Elbehri, Karl Heinz Erb, B. Osman Elasha, Mohammad Rahimi, Mark Rounsevell, Adrian Spence, Riccardo Valentini, et al. Framing and Context. In *Climate change and land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*, pp. 1–98. Intergovernmental Panel on Climate Change (IPCC), 2019.

Kemen G. Austin, Amanda Schwantes, Yaofeng Gu, and Prasad S. Kasibhatla. What Causes Deforestation in Indonesia? *Environmental Research Letters*, 14(2):024007, 2019.

Kimberly M. Carlson, Lisa M. Curran, Gregory P. Asner, Alice McDonald Pittman, Simon N. Trigg, and J. Marion Adeney. Carbon Emissions from Forest Conversion by Kalimantan Oil Palm Plantations. *Nature Climate Change*, 3(3):283–287, 2013.

Taco Cohen and Max Welling. Group Equivariant Convolutional Networks. In *International conference on machine learning*, pp. 2990–2999, 2016a.

Taco S. Cohen and Max Welling. Steerable CNNs. *arXiv preprint arXiv:1612.08498*, 2016b.

Philip G. Curtis, Christy M. Slay, Nancy L. Harris, Alexandra Tyukavina, and Matthew C. Hansen. Classifying Drivers of Global Forest Loss. *Science*, 361(6407):1108–1111, 2018.

Adrià Descals, Zoltan Szantoi, Erik Meijaard, Harsono Sutikno, Guruh Rindanata, and Serge Wich. Oil Palm (*Elaeis Guineensis*) Mapping with Details: Smallholder Versus Industrial Plantations and Their Extent in Riau, Sumatra. *Remote Sensing*, 11(21):2590, 2019.

Jonathan A. Foley, Ruth DeFries, Gregory P. Asner, Carol Barford, Gordon Bonan, Stephen R. Carpenter, F. Stuart Chapin, Michael T. Coe, Gretchen C. Daily, Holly K. Gibbs, et al. Global Consequences of Land Use. *science*, 309(5734):570–574, 2005.

Matthew G. Hethcoat, David P. Edwards, Joao MB Carreiras, Robert G. Bryant, Filipe M. Franca, and Shaun Quegan. A Machine Learning Approach to Map Tropical Selective Logging. *Remote sensing of environment*, 221:569–582, 2019.

Jeremy Irvin, Hao Sheng, Neel Ramachandran, Sonja Johnson-Yu, Sharon Zhou, Kyle Story, Rose Rustowicz, Cooper Elsworth, Kemen Austin, and Andrew Y. Ng. Forestnet: Classifying Drivers of Deforestation in Indonesia Using Deep Learning on Satellite Imagery. *arXiv preprint arXiv:2011.05479*, 2020.

Krzysztof Janowicz, Song Gao, Grant McKenzie, Yingjie Hu, and Budhendra Bhaduri. GeoAI: Spatially Explicit Artificial Intelligence Techniques for Geographic Knowledge Discovery and Beyond, 2020.

Anuj Karpatne, Imme Ebert-Uphoff, Sai Ravela, Hassan Ali Babaie, and Vipin Kumar. Machine Learning for the Geosciences: Challenges and Opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8):1544–1554, 2018.

Timothy M. Lenton. Early Warning of Climate Tipping Points. *Nature climate change*, 1(4):201–209, 2011.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An Imperative Style,

High-Performance Deep Learning Library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems 32*, pp. 8024–8035. Curran Associates, Inc., 2019. URL <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.

Darius Phiri, Justin Morgenroth, and Cong Xu. Long-Term Land Cover Change in Zambia: An Assessment of Driving Factors. *Science of The Total Environment*, 697:134206, 2019.

Ate Poortinga, Karis Tenneson, Aurélie Shapiro, Quyen Nquyen, Khun San Aung, Farrukh Chishtie, and David Saah. Mapping Plantations in Myanmar by Fusing Landsat-8, Sentinel-2 and Sentinel-1 Data along with Systematic Error Quantification. *Remote Sensing*, 11(7):831, 2019.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional Networks for Biomedical Image Segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

David P. Roy, Michael A. Wulder, Thomas R. Loveland, Curtis E. Woodcock, Richard G. Allen, Martha C. Anderson, Dennis Helder, James R. Irons, David M. Johnson, Robert Kennedy, et al. Landsat-8: Science and Product Vision for Terrestrial Global Change Research. *Remote sensing of Environment*, 145:154–172, 2014.

Jean-Daniel Sylvain, Guillaume Drolet, and Nicolas Brown. Mapping Dead Forest Cover Using a Deep Convolutional Neural Network and Digital Aerial Photography. *ISPRS Journal of Photogrammetry and Remote Sensing*, 156:14–26, 2019.

Charles Verpoorter, Tiit Kutser, and Lars Tranvik. Automated Mapping of Water Bodies Using Landsat Multispectral Data. *Limnology and Oceanography: Methods*, 10(12):1037–1050, 2012.

Charles Verpoorter, Tiit Kutser, David A. Seekell, and Lars J. Tranvik. A Global Inventory of Lakes Based on High-Resolution Satellite Imagery. *Geophysical Research Letters*, 41(18):6396–6402, 2014.

Maurice Weiler and Gabriele Cesa. General E(2)-Equivariant Steerable CNNs. In *Advances in Neural Information Processing Systems*, pp. 14334–14345, 2019.

## A Appendix

### A.1 Dataset

Table 3 gives the mapping between the original labels provided for the dataset by Austin et al. (2019) and those labels used as classification targets for our models.

### A.2 Equivariance - Limitations and Alternative Approaches

Equivariance places a constraint over the kernels used by the model such that the model respects symmetries in the data. An alternative approach to this is to use data augmentation, which is generally easier to implement. On the other hand, data augmentation effectively increases the size of the dataset and therefore makes training slower. Building equivariant models guarantees the models behaviour under certain symmetries, whereas data augmentation does not. Furthermore, equivariance can reduce the number of parameters required in the model and increase training efficiency. Therefore, in this work, given that we have a known symmetry group equivariant models are a sensible choice.

### A.3 Model Architecture

For both the non-rotation equivariant and rotation equivariant models we use the same model architecture with the key difference that the convolutional layers are either rotation equivariant or non-rotation equivariant depending on the choice of model. The model architecture is a U-Net style model, which makes use of a convolutional block comprised of two convolutional layers, two batch normalisation layers, and two drop out layers. The model then consists of five convolutional blocks

Table 3: The mapping between deforestation driver groups as defined in (Irvin et al., 2020) and the expert labelled deforestation driver categories defined in (Austin et al., 2019). The deforestation driver groups are used as classification targets when training models.

Expert Labelled Deforestation Driver Category	Classification Target Driver Group
Oil palm plantation	
Timber plantaion	Plantation
Other large-scale plantations	
Grassland/shrubland	Grassland/shrubland
Small-scale agriculture	
Small-scale mixed plantation	Smallholder agriculture
Small-scale oil palm plantation	
Mining	
Fish pond	
Logging road	Other
Secondary forest	
Other	

with downsampling in-between each and five convolutional blocks with upsampling in-between each. Further, a skip connection is placed between each convolutional block connecting upsampled layers with the corresponding same shape downsampled layer. In addition, there is a flatten layer and three multi-layer perceptron layers providing the driver classification output from the lowest dimensional space. We build the model using PyTorch (Paszke et al., 2019) and for the rotation-equivariant layers we make use of E2CNN (Weiler & Cesa, 2019).

The non-rotation equivariant model has 3.7 million trainable parameters and the rotation equivariant model has 3.0 million trainable parameters. Each model was run on a Titan Xp GPU taking less than 30 minutes and requires approximately 3GiB of memory to train.