

Extraction of Forest Road Information from CubeSat Imagery using CNNs

Lukas Winiwarter, Nicholas C. Coops, Alex Bastyr, Daisy Zhao, Adam Ford

Faculty of Forestry, University of British Columbia, Vancouver Campus,
2424 Main Mall, Vancouver, BC, V6T 1Z4, Canada.

Faculty of Science, University of British Columbia, Okanagan Campus

March 31st, 2023

Submitted to Scientific Authority

Canadian Wildlife Service

A digital report including additional figures and code snippets to reproduce the results is available online:

<https://github.com/lwiniwar/roadCNN/blob/main/report/report.md>

Abstract

Forest roads provide access to remote wooded areas, serving as key transportation routes and contributing to human impact on the local environment. Large animals such as bears, moose, and caribou are affected by their presence. To address sparse and outdated location data, we employ CubeSat Imagery from the Planet constellation to predict road probabilities using a Convolutional Neural Network. The classification is refined through post-processing, which considers spatial misalignment and road width variability. On a withheld test subset, we achieve an overall accuracy of 99.1%, a precision of 76.1% and a recall of 91.2% (F1-Score: 83.0%) after considering these effects. We investigate the performance with respect to canopy coverage using NDVI, topography (slope and aspect) and land cover metrics. Predictions are best in flat areas, low to medium canopy coverage, and in the forest (coniferous and deciduous) land cover class. The results are vectorized into a drivable road network, allowing for vector-based access and coverage analyses. Our approach digitized 79,624 km of roads in an ~25,000 km² area in B.C., Canada. Compared to reference data provided by the B.C. government, we are able to detect an additional 2,500 km of roads connected to the network.

Introduction

The provision of up to date and spatially accurate information on the location of roads is increasingly fundamental for a range of conservation and industrial activities globally (Barrington Leigh and Millard-Ball 2017), and there is growing recognition that the ecological implications of roads can be far-reaching (Bennett et al., 2011). As a result, there is an emerging field of road ecology to improve our understanding of both the direct and indirect effects of roads on for example wildlife populations (Kearney et al., 2020). Roads are critical for intensifying exploitation of timber, minerals, fossil fuels and arable land with new road construction highly dynamic as infrastructure is developed to access sites and transport materials. In a forestry context, roads are required for effective forest management, providing access for harvesting and silvicultural operations, recreational, wildlife and wildfire management.

Roads however can have long lasting and pervasive impacts on ecosystems (Sherba et al., 2014). Road networks, through increases in compaction of road surfaces which in turn reduces infiltration, can modify the hydrological cycle (Sherba et al., 2014). Roads can have large impacts on environmental integrity and wildlife health through habitat fragmentation, and alterations of landscape structure, which can result in cascading ecological implications such as increased risk of mortality (Selva et al., 2015). Roads can also be a major sink for animal mortality particularly for large species such as bears (*Ursus arctos*), moose (*Alces alces*) and caribou (*Rangifer tarandus caribou*), who use roads networks for movement, but also food, as roadside verges, with more sunlight and water than within the forest canopy. Roads also facilitate motorized human access to remote areas, causing habitat disruption (Proctor et al., 2020).

The mapping and monitoring of forest roads can be difficult and databases that contain road information can be challenging to build and maintain due to road networks being highly dynamic with new roads developed, and older roads decommissioned, annually depending on use. British Columbia, Canada offers a complex case study in road design, building, and maintenance due to its extensive road network covering over 800,000 km (Forest Practices Board, 2002), of which 74% is related to the forest industry, as well as range of environmental conditions including highly variable terrain, climate and landform features as well as many being in remote areas (Waga et al., 2020).

In Canada, and in fact, globally, most provincial, state or territory governments maintain vector information on the forest roads under their jurisdiction (Roussel et al., 2022). Road locations and associated attributes are conventionally mapped using aerial photography. Roads are manually delineated using key features on the photography themselves, often augmented with field information, road

planning documents and local geographic information systems. Road positional information can be quite accurate assuming road features are clearly visible on the photography. In forested environments, the interpreter is limited in what can be observed underneath forest canopies resulting in significant position error.

An alternative approach is to extract road information from remote sensing data which can be acquired from either spaceborne or airborne sensors. A large number of road extraction approaches have been investigated, since the 1980's, for use with digital remote sensing imagery (Wang et al., 2016). The last 5 years has seen a rapid interest in the use of deep learning approaches, such as convolutional neural networks (CNNs) which are showing an improvement over conventional image classification for road extraction related computer vision tasks (Kearney et al., 2020). This has been attributed to the ability to train the CNN on many thousands of examples of road segments across a range of conditions, allowing the CNN to use both variations in the brightness of road surfaces and local morphology to learn typical road features without significant pre-processing (e.g., Alshehhi et al., 2017; Das et al., 2011; Gao et al., 2019; Xu et al., 2018), and to the CNN's ability to learn feature representation through annotated training data (Goodfellow et al., 2016).

Most deep learning studies to date have focused on the detection of paved urban roads and most often utilized very high spatial resolution imagery (i.e., ≤ 1 m pixels) which may limit their applicability to unpaved road networks (Xu et al., 2018). As a result, when building training datasets, it is often assumed that road segments are well formed and wide, with a permanent impervious surface with highly contrasting surrounding land cover, and regular road patterns (such as long straight roads in a grid formation). While successful in urban environments, road extraction techniques applicable to the complexities of forested regions are still lacking (Kearney et al., 2020).

Once accurate road information is available, additional information such as road age is also important to understand the temporal dimension of road use over the landscape. Information on the age is crucial for effective forest management, as these roads play a vital role in planning, conservation, and economic activities (Barrington-Leigh & Millard-Ball, 2017; Ferraz et al., 2016; Waga et al., 2020). While improvements in road design and construction can reduce erosion on newly built roads, assessing the impacts of existing roads can be challenging due to the lack of information on their age, condition, and spatial distribution, particularly in remote and densely forested regions.

Here we evaluated a CNN road extraction routine for forest roads using high-spatial resolution Cubesat imagery (3 m pixels) over large areas of montane and boreal forests in northern BC. The Planet constellation of CubeSats enables almost daily imaging over the globe at fine spatial resolution, making it a strong candidate for extraction and monitoring of road networks information. We chose a large forested region due to the dynamic network of unpaved roads serving local industry including forestry, mining and oil and gas activities, and concerns over the implications of current and expanding road networks for the recovery of caribou present in the region. The secondary goal of this analysis is to date road segments within the road network derived from the CNN utilizing a combination of the spectral indices, Normalized Difference Vegetation Index (NDVI) and the Normalized Burn Ratio (NBR), and a time series of annual Landsat best-available pixel (BAPs) composites available at 30m spatial resolution.

The developed approach, when expanded, could provide additional information about road position through space and time enabling managers and planners to better understand roads over large forest management areas and improve understanding of well used, vs decommissioned roads for recreation and species conservation purposes.

Materials and Methods

1. Study Area

To speed up processing and to enable clear visualisations, we processed and presented data on two spatial subsets. One is defined by a polygon outlining Caribou herds and spans a total of approximately 25,000 km². The second one is a regular grid defining the training, validation, and test data sets used for the deep learning approach. This smaller study area investigated in this study is depicted in Figure 1. It spans an area of 71 x 145 km² (total area of 10,295 km²) in Eastern BC. The area overlaps with the habitat of Mountain Caribou (Hart Ranges herd, pop. 405) as well as Northern Caribou (Burnt Pine, Quintette, and Narraway herds, pop. 1, 120 and 35, respectively) (BC Ministry of Environment, 2010; Environment Canada, 2010). Populations are estimates by the Province of British Columbia for October 2021 (BC Government, 2021).

While the smaller area of interest (black raster in Figure 2) was covered by a single dataset downloaded from Planet as a best-pixel-product, the larger area was covered by multiple satellite overpasses. We used 23 partially overlapping overpasses to cover the entire dataset, including data from 2017-07-05 until 2017-09-04.

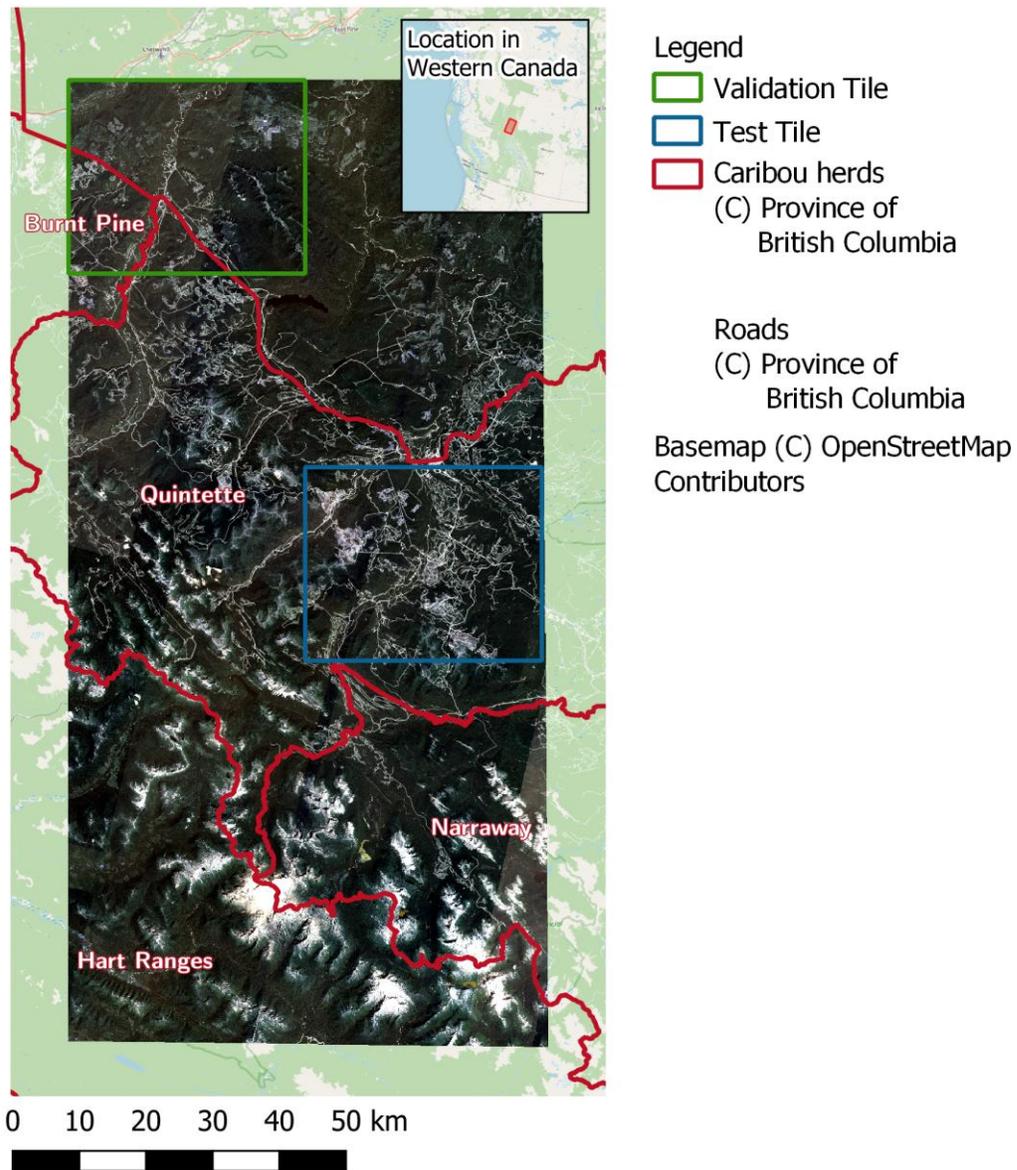


Figure 1: Overview map of the study area (raster Aoi in Fig. 2). The Caribou herds are marked in red, and the validation and test tiles in green and blue, respectively. The RGB composite of the satellite imagery shows the full extent of the investigated area, with parts (in the south) with little to no roads (white lines) and parts with many roads (center and northwest).

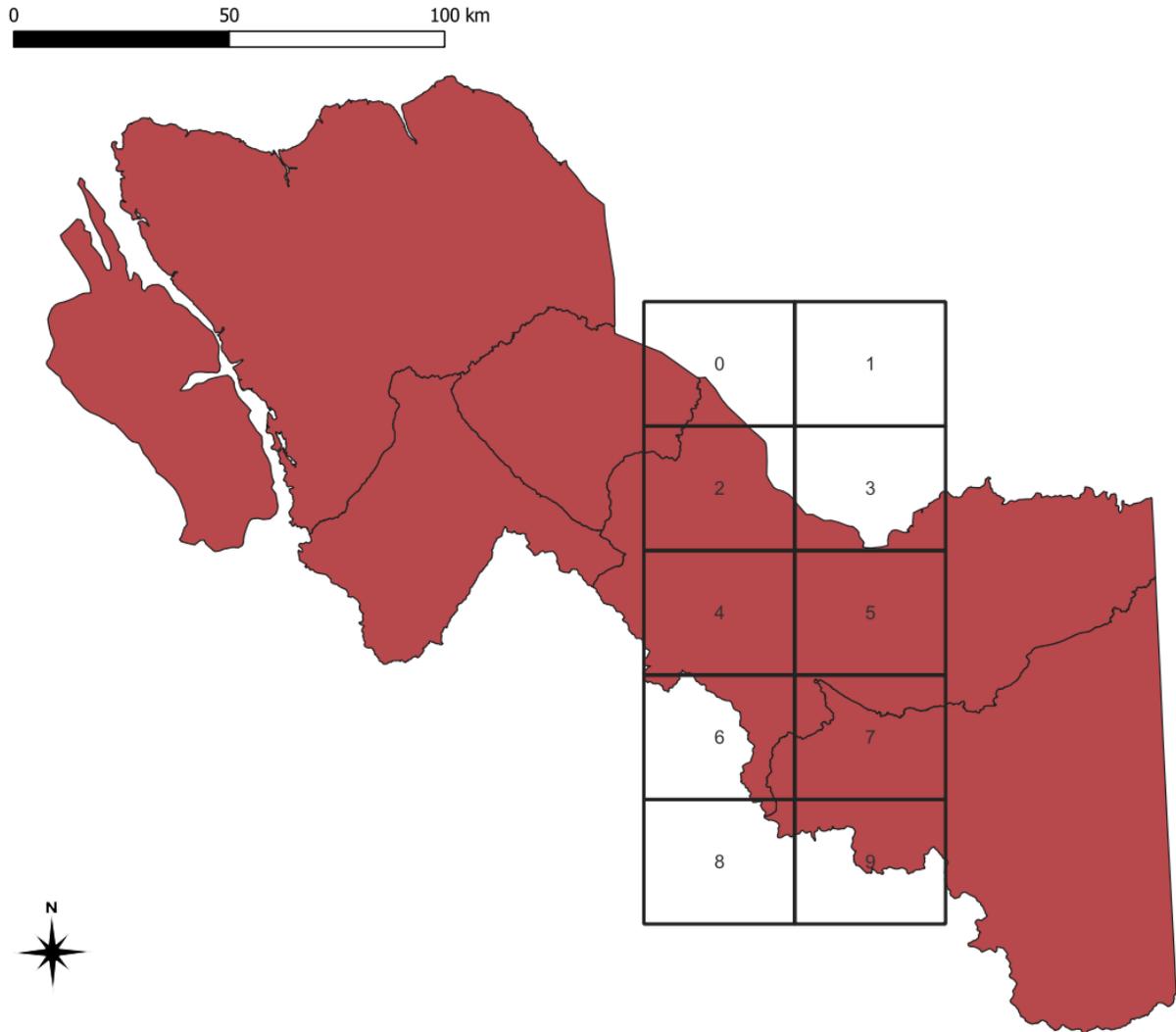


Figure 2: Larger (red polygons) and smaller (black grid) areas of interest. For training the method, polygons labelled 1-6, 8, and 9 were used. Polygon 0 acted as validation set, and polygon 7 was used for the finer-scale analyses.

2. CubeSat Data ALS acquisition and processing

We utilized data acquired from the Dove Classic satellites from the PlanetScope CubeSat constellation, which consists of over 130 CubeSats in sun-synchronous orbit at altitudes of approximately 475 km (Frazier & Hemingway, 2021). The constellation of satellites is such that it provides near daily image acquisition globally, at a ground-sampling distance (GSD) of approximately 3 m. Imagery from the constellation archives is available beginning 2016 and is acquired in 4 spectral bands: Blue (455 – 515 nm), Green (500 – 590 nm), Red (590 – 670 nm) and Near-Infrared (780 – 860 nm).

To ensure that the image data and the reference road data (cf. Section 4) represent a similar state, we use imagery from the summer of 2017. We selected tiles from Planet’s Basemap product¹, which are mosaiced from multiple acquisitions within a calendar month. Where tiles for the month of August 2017 showed smoke or clouds, we manually replaced them with a cloud- and smoke-free mosaic from either July or September 2017.

3. Landsat historical information for stand age

We utilized annual Landsat image composites generated through the Composite2Change (C2C) approach (Hermosilla et al., 2016) to conduct a temporal analysis for determine the age of road segments within the derived CNN road network. This approach utilizes the free and open-access Landsat archive (White & Wulder, 2014) to produce yearly, gap-free, surface-reflectance image composites with a spatial resolution of 30 meters (Hermosilla et al., 2015b). The C2C approach also allows for the detection and characterization of forest disturbances (Hermosilla et al., 2015a). For image compositing, a target date of August 1st ± 30 days is used, as this date coincides with the growing season for most of Canada’s forested ecosystems. All atmospherically correct Landsat images (Masek et al., 2006; Schmidt et al., 2013) acquired within this target time frame from 1984 to 2021 are used, along with pixel scoring functions, to create the best available pixel surface reflectance image composites. The pixel scoring functions rank and select the optimum annual pixel observation to be used in an image composite based on sensor type, date of acquisition, distance to cloud or cloud shadow, and atmospheric opacity (White et al., 2014). The initial annual surface reflectance image composites generated suffer from data gaps due to varying data availability at different latitudes and persistent clouds in certain regions. To mitigate this issue, anomalous values are removed, and the data gaps are infilled using proxy values assigned through a spectral trend analysis (Hermosilla et al., 2015a). This results in the generation of seamless Landsat surface-reflectance image composites from 1984 to 2021.

4. Existing roads database and reference field data

The BC provincial road database provides contemporary information about the road network in BC. In its most recent update in 2018, 719,000 km of road were defined, with approximately 92% of them unpaved and used by the natural resource sector. The dataset is based on the Digital Road Atlas of British Columbia, which combines map data from historical survey data and satellite imagery. However, both the spatial accuracy (i.e., position of the road itself) and the accuracy of the presence or omittance of roads

¹ <https://www.planet.com/products/basemap/>, last access 2023-03-03

(completeness and correctness) is unknown. The government of BC has acknowledged this lack of information on roads suggesting that up to 150,000 km of resource roads are missing and the current data is inaccurate (GeoBC, 2017). In the study area presented in Figure 1, 6,600 km of road are present in the database. We exclude roads of the type “Trail” from our analyses, as they typically constitute non-drivable paths of subordinate relevance. Furthermore, they are often not detectable in the satellite imagery.

5. CNN classification

Recent advancements in deep learning and CNN’s development in computer science has led wide variety of applications utilising remote sensing datasets. Developed from traditional neural networks, deep neural nets can outperform traditional machine learning methods used in environmental remote sensing (Yuan et al., 2020). A review of current advances in CNN road extraction confirms the majority of studies to date, extracted urban rather than rural road networks. Studies were predominately from China and the United States with very few studies conducted in Europe, Canada, Asia, Africa or South America. For new algorithm development many studies used the Massachusetts Road Dataset (Mnih, 2013), or the DeepGlobe Road Extraction Challenge 2018 data set (Demir, 2018) for both model training and validation. With respect to accuracy of the existing approaches, most papers define their own evaluation metrics and datasets, making state of the art accuracies difficult to derive.

A recent example is Microsoft’s RoadDetections dataset, using a Residual U-Net neural network (Zhang et al., 2018) on Bing Imagery. As this network has been trained mainly on paved roads in urban environments, the performance in remote, wooded landscapes is inferior.

We therefore employ a domain transfer technique, where a neural network with pre-trained weights is adopted. The use of this pre-trained network reduces the amount of training required as the solution space is pre-selected. We use the SegNet approach (Badrinarayanan et al., 2015), which consists in an encoder- and a decoder network. For the encoder network, we apply weights from the pre-trained VGG-16 network (Simonyan and Zisserman, 2015), which used the ImageNet dataset (Deng et al., 2009). The decoder network is initialized with random weights.

The SegNet encoder consists in four convolutional blocks, each followed by a max pooling layer. Similarly, the decoder is a sequence of upsampling layers, followed by convolutional blocks. The output of the last layer is fed through a softmax layer to create pseudo-probabilities for the two classes (non-road and road). The dimensions and filter sizes for each of the layers are shown in the accompanying digital report.

Table 1: Tuned hyperparameters used in the final network

HYPERPARAMETER	VALUE	EXTRA INFORMATION
LEARNING RATE	0.0001	Decreased after each epoch ($\gamma = 0.1$)
OPTIMIZER	Adam	Betas: $\beta_1 = 0.9, \beta_2 = 0.999$, Weight decay: $wc = 0.0005$
LOSS FUNCTION	Weighted cross-entropy	
WEIGHT BALANCE	0.95 (road), 0.05 (non-road)	
DATA AUGMENTATION	8-fold	Rotations by 90, 180 and 270 degrees, vertical and horizontal flip, and combinations (following (Sonogashira et al., 2020))
NUMBER OF EPOCHS	3	30,559 patches each
BATCH SIZE	16	Resulting in 163,408 batches
INPUT PATCH SIZE	244 × 244 pixels	
PATCH OVERLAP	122 pixels	Each pixel is classified 4 times
MERGING FUNCTION	Gaussian weighted average	

To ensure that weight updates are focused on the decoder network, we set the learning rate of the encoder to $\frac{1}{2}$ of the value for the decoder. The hyperparameters for the network are tuned manually and final values are presented in Table 1: Tuned hyperparameters used in the final network. We use weighted cross-entropy as a loss function to ensure that the network focusses on learning road pixels over non-road pixels, anticipating that false positives (road pixels where there are no roads) are less problematic than false negatives (missing road pixels).

As the network architecture is limited to processing images of input size 244 × 244 pixels, patches of that size are created from the satellite imagery. To minimize edge effects, we create these patches with an overlap of 50%, i.e., 122 pixels. For training and loss calculation, we use these patches, but when inferring the road network on the validation and test areas, we combine overlapping patches by a weighted average. Weights are calculated by a 2D Gaussian centered in the image center with a standard deviation of $\frac{1}{4}$ the image width. The reasoning here is that the CNN is able to make better predictions in the image center, as the neighbourhood context is clearer. Thus, each pixel’s road probability is the result of four

passes through the network. To obtain a final binary map, simple thresholding with a cut-off value of 0.5 is used.

In the training process, we limit input patches to ones that contain at least 10 pixels classified as roads. This data-pre-selection speeds up training significantly, as patches with no roads present provide little information on how to detect roads. For validation and test, we run the network on the full data, as no a-priori information on road presence is assumed.

6. Post-Processing and Evaluation

As the training, test and validation road network data is not perfectly aligned to the satellite imagery we anticipate that initial performance scores are quite low, as a large number of false positives and false negatives occur at the edges of the detected roads. More specifically, with a 3m pixel size, the Planet imagery is likely to detect changes in radiance within 1 – 3 pixels wide the equivalent of a 9m side road, which is unlikely to occur. When compared to a rasterised vector road coverage, the road itself will be 3m wide, resulting in a miss-classification along road segments, an example can be seen in Figure 3.

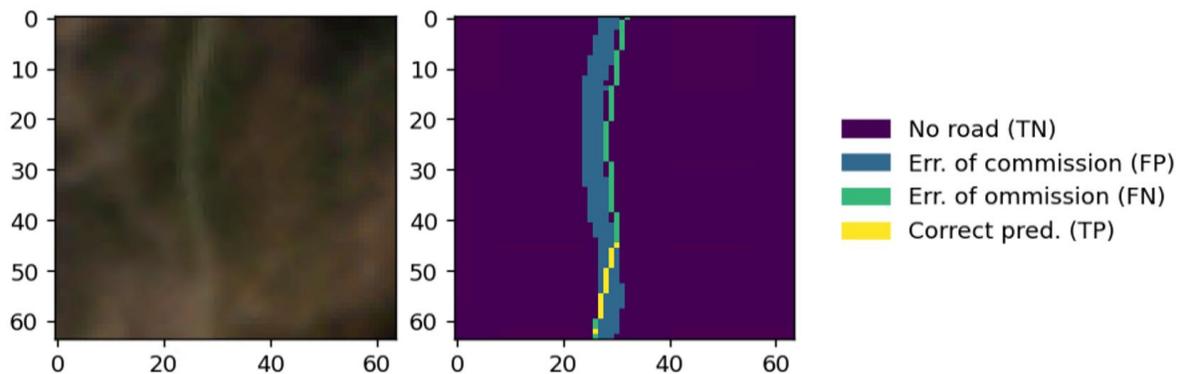


Figure 3: Left: RGB imagery with a road visible. Right: Classification result (True Positive and False Positives) of the road. It can clearly be seen that the detected road follows the reference road at this position, yet only a small part of the road is considered correctly predicted ("True Positive").

As we are principally concerned with the detection of a road, and less with its exact position within a 9m buffer, we employ morphological operations to disregard these misclassifications. To that end, we first buffer the reference roads as well as the predicted roads by n pixels using morphological dilation. We then select all the pixels in these buffered maps that are true positives, i.e., correctly predicted roads. From the map of true positives, we again apply a buffer, this time of $2n$ pixels. The resulting map is used as a mask: In this area, false negatives and false positives are disregarded as being the result of alignment

and road-width mismatches. Subsequently, the number of true positives (TP, correctly identified road pixels), true negatives (TN, correctly identified non-road pixels), false positives (FP, non-road pixels incorrectly identified as road) and false negatives (FN, road pixels incorrectly identified as non-road) is evaluated, and metrics of accuracy, precision, recall, and F1-Score are calculated according to Equations 1 – 4.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (2)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (3)$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{(Precision+Recall)} \quad (4)$$

The buffer size n is determined in accordance with the expected alignment error, maximum road width and pixel size. We investigate different values for n and observe a convergence limit for the evaluation scores.

To further understand the impact of different topographic factors on the classification quality, we investigate the four metrics above stratified by different slope angle ranges, slope aspects, land cover classes and Normalized Difference Vegetation Index (NDVI) values. The slope angle and aspect is calculated from a 30 m DEM (Abrams et al., 2020; NASA/METI/AIST/Japan Spacesystems and U.S./Japan ASTER Science Team, 2019), which is subsequently resampled to fit the 3 m cell size of the RGB data using nearest-neighbour interpolation. Similarly, the 30 m landcover map (Hermosilla et al., 2022, 2018, 2016) is resampled using nearest-neighbour interpolation. The NDVI is calculated from the input RGB(I) data, and therefore does not need to be resampled.

7. Vectorization

For certain applications, including navigation purposes, a vectorized and topologically intact road network is required. To automatically derive such a network from satellite imagery, we use the *VecNet* approach outlined by (Roussel et al., 2023). This approach operates on the probability raster outputted by the CNN.

The vectorization employs a “driving” approach, where roads are followed along the most likely path, with intersections and turnoffs recoded to be driven at a later point. To start this algorithm, input segments (“seeds”) are required. We automatically derive these seeds from the existing reference data for test- and validation tiles, by overlaying the road network with a 100 m buffer around the boundary polyline of the respective patch. We assume that road networks will be connected, i.e., there will be no ‘islands’ of roads within the analysis area that are not connected to roads leading into the area.

Due to single patches not being able to predict roads correctly, gaps in the raster probability map appear. As *VecNet* can only leap over small gaps, an iterative manual approach was used, where the output from *VecNet* was taken, gaps were closed manually, and then *VecNet* was re-run to obtain a topologically correct road network.

8. Road Age

The annual Landsat image composites were clipped to the derived CNN road network. We derived spectral profiles for each pixel through time by calculating their associated NDVI and NBR values. NDVI has been widely utilized to track changes over time series of satellite imagery as it is a reliable measure of vegetation productivity and health (Helman et al., 2015; Hmimina et al., 2013; Tucker, 1979). Time series of NBR pixels has been demonstrated to be a sensitive and consistent method for the retrieval of disturbance events over forest environments (Kennedy et al., 2010). The construction of a road through a forested environment is associated with a sudden shift in the spectral properties, resulting in NDVI and NBR values to decrease. To detect this change, a pixel-wise change detection algorithm was developed and applied over the time series (Figure 4). When there is no drop in spectral values, it is assumed that the road is either too small for a noticeable drop in indices to occur, or that the road was built in or prior to 1984 (the start of the time series).

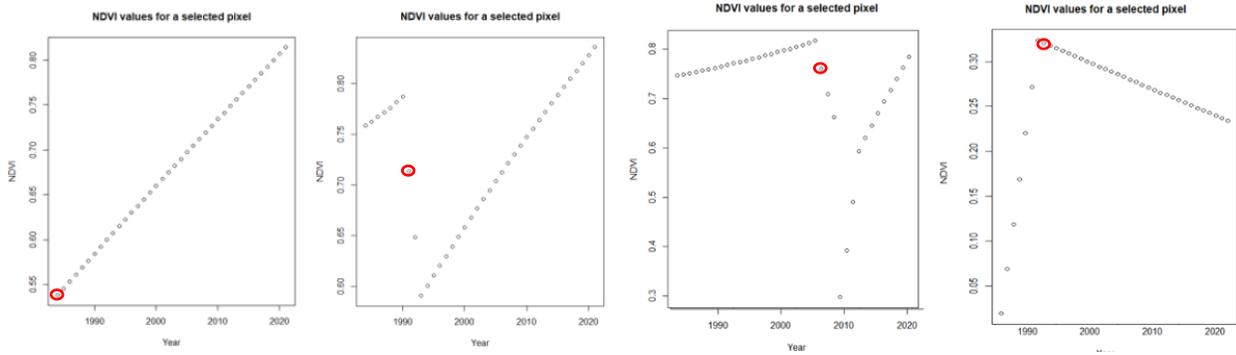


Figure 4: Spectral profiles of selected pixels through time. The point highlighted by the red circle indicates the year the change detection algorithm would return.

Results

Selected areas of the RGB test data are shown in Figures 5-7 along with the ground-truth and the predicted road probability. While most roads, both paved and unpaved, can be identified reasonably well in the imagery (Fig. 5), the CNN estimate gets poorer when roads disappear beneath the canopy. Roads located in active mining areas (Fig. 6) and/or recent clear-cuts (Fig. 7) are typically not detected. Especially in these areas, the quality of the reference data is also poorer, as roads often change over short periods, and the road database is not updated frequently.

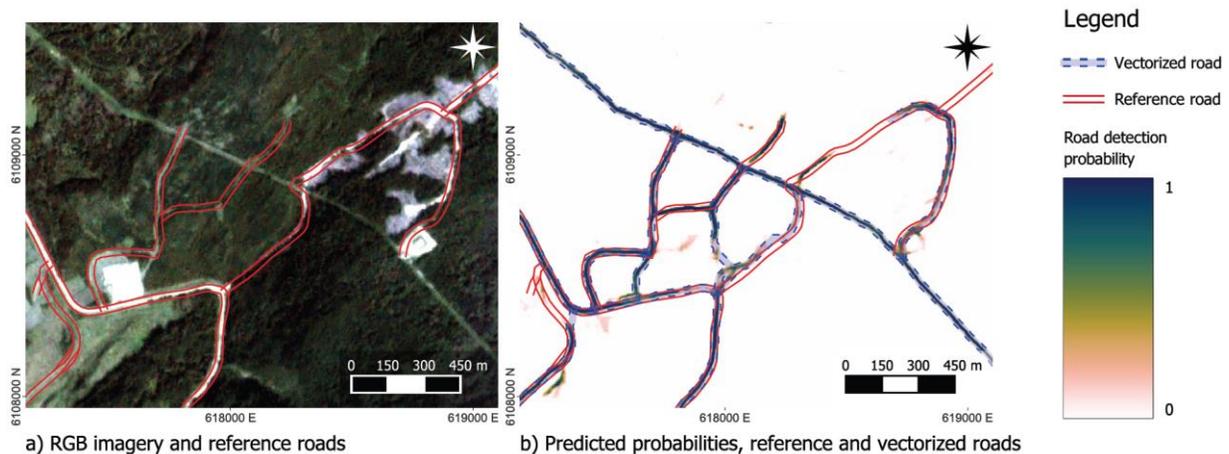


Figure 5: Detailed vectorization result showing a seismic line incorrectly classified as a road, with good classifications otherwise. (a) RGB imagery and reference road network, (b) predicted probabilities, reference road network and predicted road network after vectorization.

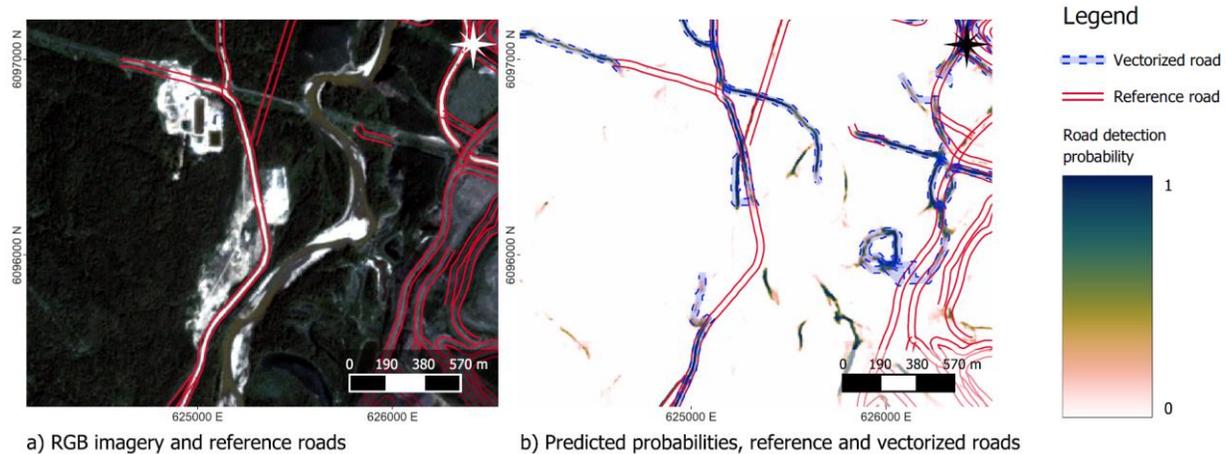


Figure 6: Detailed vectorization result showing a mining area (eastern part) and barren ground around infrastructure (center, along the road), where roads are generally missed. (a) RGB imagery and reference road network, (b) predicted probabilities, reference road network and predicted road network after vectorization.

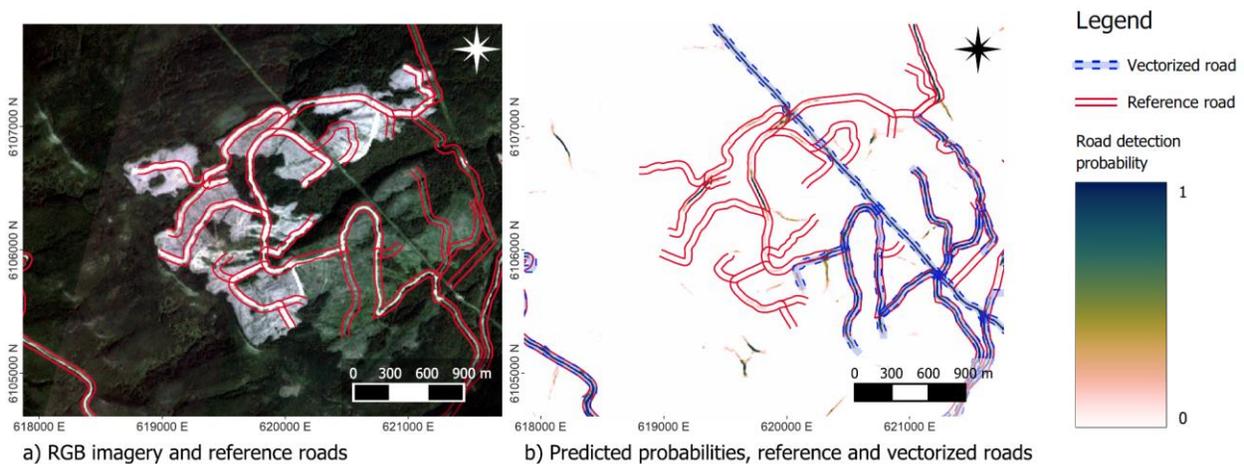


Figure 7: Detailed vectorization result showing a freshly clear-cut area (white), where roads are generally missed, and an older clear-cut area (light green), where the detection works well. (a) RGB imagery and reference road network, (b) predicted probabilities, reference road network and predicted road network after vectorization.

Spatial offsets between the reference vector data and the imagery can be seen e.g., in Figure 3. These offsets cause a large proportion of the false positives and false negatives as shown in Table 2. Consequently, the measures of precision, recall, and F1 are influenced by these offsets. Using morphological dilation operations, we acknowledge this spatial offset and can obtain improved estimates of the performance scores. By further increasing the allowed offset, we also incorporate parts of the mining- and freshly clear-cut areas (cf. Figs. 6 and 7), where false positives and false negatives are ignored. Figure 8 shows the progression of the performance metrics for increasing buffer sizes, which are

calculated from the number of true positives, true negatives, false positives, and false negatives as shown in Table 2. While recall values (how many of the roads are detected) are very low for small dilation values, they increase quickly and are already > 50% after 2-pixel buffer (corresponding to 6 m offset). The precision (how many of the detected pixels are roads) has a slower increase, indicating that the number of false positives is less a result of spatial misalignment. Still, precision converges to a value of about 95%, while recall reaches 99% at 15 pixels (45 m) buffer around the roads.

Table 2: Prediction quality for different buffer sizes, measured in pixels for the test area (polygon ID 7). The large number of false positives for small buffer sizes mostly corresponds to spatial alignment problems of the datasets.

BUFFER SIZE	TRUE POSITIVES	TRUE NEGATIVES	FALSE POSITIVES	FALSE NEGATIVES
0 PX (CNN OUTPUT)	172666	111843439	2012961	360545
1 PX (3 M)	584586	112262164	1232338	310523
5 PX (15 M)	2489847	110880575	779587	239602
15 PX (45 M)	7859615	105864302	532243	133451

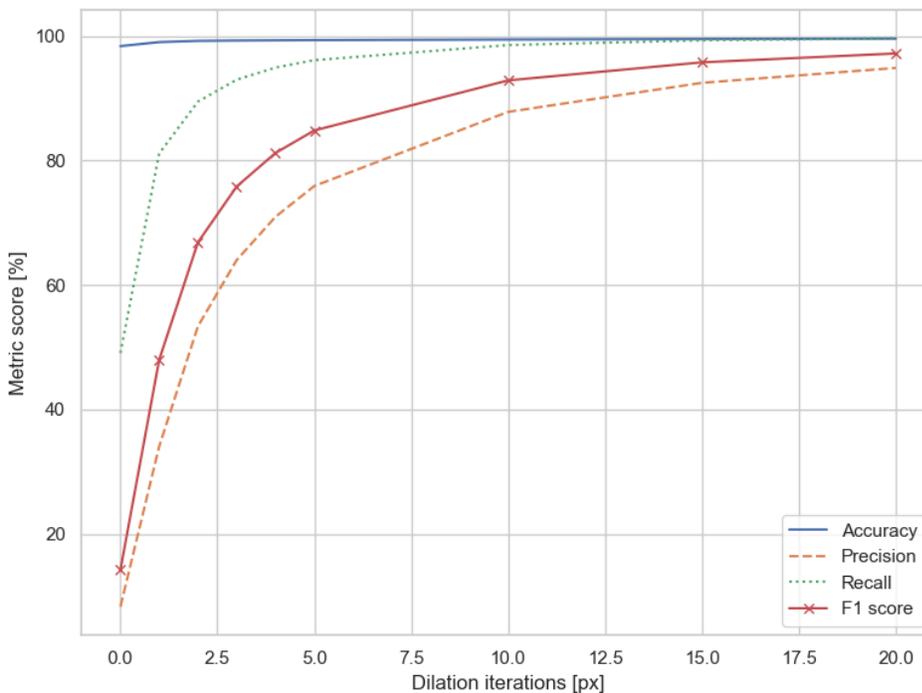


Figure 8: Accuracy, Precision, Recall and F1-Score for the binary classification, shown for different buffer sizes (dilation iterations). One pixel corresponds to 3 m of allowed offset.

In Table 3, we show how the CNN road extraction performs on different terrain features, i.e., slope, aspect, elevation, as well as crown coverage (indicated by the NDVI) and land cover classes.

Table 3: Performance values stratified by different land cover types, vegetation coverage, slope, and aspect. Values in parentheses are after allowing for a 5 px (15 m) dilation. N/A values correspond to no true positives in the respective strata.

	ACCURACY [%]	PRECISION [%]	RECALL [%]	F1 [%]
NDVI < 0	N/A (98.6)	N/A (60.0)	N/A (100.0)	N/A (75.0)
0 <= NDVI < 0.5	95.0 (97.8)	13.6 (68.1)	50.0 (88.2)	21.4 (76.8)
0.5 <= NDVI < 0.75	91.7 (97.6)	11.5 (81.2)	61.7 (96.4)	19.4 (88.1)
0.75 <= NDVI < 1	98.9 (99.5)	6.1 (74.6)	40.9 (96.6)	10.6 (84.2)
SLOPE < 5 DEG	98.1 (99.3)	9.2 (79.9)	51.3 (96.4)	15.6 (87.4)
5 <= SLOPE < 10 DEG	97.9 (99.2)	8.9 (79.2)	50.4 (96.4)	15.2 (86.9)
10 <= SLOPE < 20 DEG	98.5 (99.3)	7.8 (73.5)	46.9 (95.8)	13.4 (83.2)
20 <= SLOPE < 30 DEG	99.1 (99.4)	5.0 (54.8)	38.9 (93.9)	8.8 (69.2)
30 <= SLOPE	99.3 (99.4)	1.7 (25.6)	24.9 (90.2)	3.2 (39.9)
ASPECT: NORTH	98.2 (99.3)	8.5 (75.9)	47.6 (95.7)	14.4 (84.7)
ASPECT: EAST	98.3 (99.3)	8.7 (77.2)	51.6 (96.5)	15.0 (85.8)
ASPECT: SOUTH	98.4 (99.4)	8.5 (77.0)	51.4 (96.5)	14.6 (85.6)
ASPECT: WEST	98.5 (99.4)	7.9 (74.4)	48.3 (96.1)	13.6 (83.9)
LC: 20 (WATER)	99.5 (99.6)	1.0 (23.1)	40.6 (100.0)	1.9 (37.5)
LC: 31 (SNOW/ICE)	98.2 (98.9)	5.1 (55.4)	54.5 (97.5)	9.3 (70.7)
LC: 32 (ROCK/RUBBLE)	98.0 (98.1)	0.0 (4.0)	1.0 (64.9)	0.1 (7.5)
LC: 33 (EXPOSED/BARREN LAND)	96.0 (98.0)	6.6 (67.0)	28.2 (88.0)	10.8 (76.1)
LC: 50 (SHRUBS)	94.4 (98.5)	10.2 (86.3)	55.4 (97.4)	17.3 (91.5)
LC: 80 (WETLAND)	97.1 (99.0)	9.8 (80.3)	66.9 (98.5)	17.1 (88.5)
LC: 81 (WETLAND-TREED)	98.2 (98.8)	4.3 (53.9)	60.1 (99.2)	8.0 (69.8)
LC: 100 (HERBS)	92.0 (98.6)	12.0 (91.8)	56.7 (97.4)	19.8 (94.5)
LC: 210 (CONIFEROUS)	99.2 (99.6)	7.1 (68.2)	51.4 (96.7)	12.5 (80.0)
LC: 220 (BROADLEAF)	98.1 (99.1)	6.9 (70.3)	45.7 (95.9)	12.0 (81.1)
LC: 230 (MIXED WOOD)	N/A (99.9)	N/A (60.0)	N/A (75.0)	N/A (66.7)

When stratified by NDVI (indicative of canopy cover), higher accuracy values correspond to high NDVI values. However, precision in particular decreases with high NDVI, which corresponds to incorrectly labelled roads under dense vegetation. This holds mostly true when a 15 m buffer is considered. The recall value is highest for medium NDVI values between 0.5 and 0.75.

Across different slope values, there is little variation in the accuracy, especially when considering the 15 m buffer. However, both with and without buffer, higher slope values correspond to a decrease in precision and recall, in turn resulting in lower values of F1 score. This is one of the strongest relationships observed among the stratification. In contrast, road aspect has a minimal influence on the road detection capabilities, as values are almost constant across the different cardinal directions indicating the detection is not influenced by, for example, shaded, north-facing hills.

The highest detection accuracy is achieved in the coniferous land cover class, which is the most prominent non-road classes. For land cover, recall scores are especially interesting, as they explain how likely a road is to be detected in a specific land cover. Overall, the lowest performance is achieved on Rock/Rubble and Exposed/Barren Land, which correspond to mining sites and clear cuts.

In the vectorization step, 870.5 km of roads are vectorized from 75 initial, automatically derived seeds for the smaller area of interest (cf. Figure 2). The resulting road density map derived from the vectorized road data is presented in Figure 9. For validation purposes, the road density map of the reference dataset is displayed next to it. Furthermore, we calculate the difference in road densities (Figure 10). In contrast to

the absolute values provided by the classification metrics, these maps show the overall spatial prediction quality.

In Figure 10, differences in road density between the reference data and the result after the vectorization

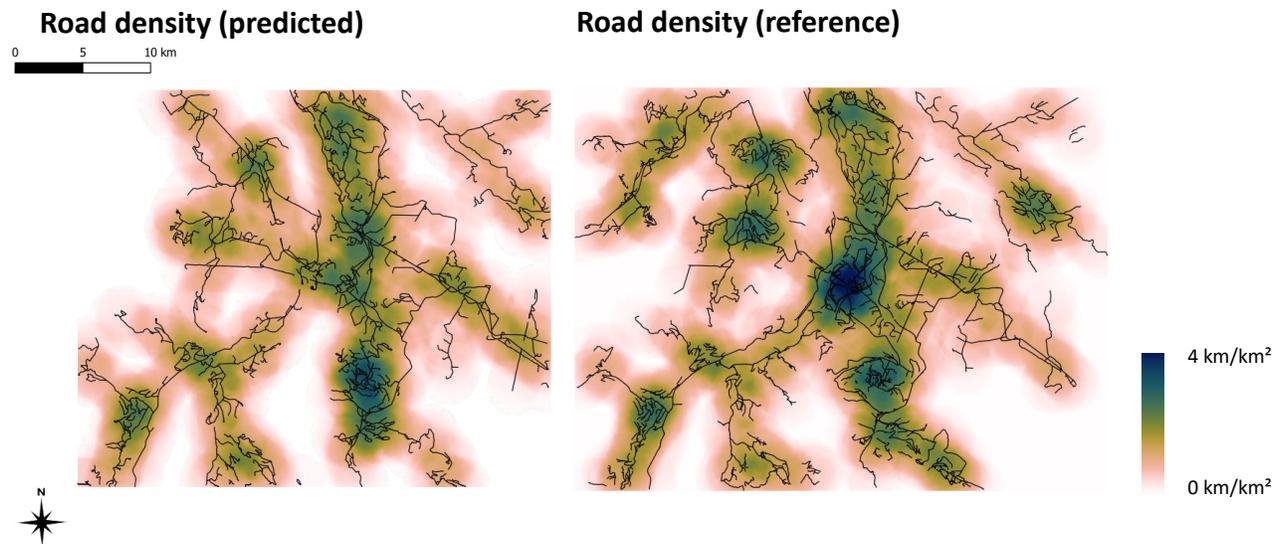


Figure 9: Predicted (left) and reference (right) road densities. The road densities were calculated on a 250 m grid using a search radius (range) of 2 km.

step are shown. In general, difference values are low (within +/- 1 km/km²) over the investigated area. The two main differences are missing roads in the North-West, where a single road link is missing and the vectorization is interrupted, and in the center, where a mining operation has many roads in the reference data, which are not present in the predicted dataset due to their radiometric footprint.

On the larger area of interest, segments were derived from the reference road layer for multiple overlapping satellite data strips. After merging these data, a total length of 79,624 km of roads were digitized automatically from the imagery. To investigate the difference to the reference road dataset, a buffer of 45 m around reference roads was established, and predicted road lengths were summed up when they intersected the larger area of interest but not the buffer (i.e., excluding roads close to the ones present in the reference data). This resulted in 2,554 km of roads that were additionally detected.

When applying the change detection algorithm over the NDVI time series, it was found that 51% of roads were built post 1985, with 30% of roads being built prior to 1984 and 19% in 1985. The output and distribution of road age when applying the change detection method over the NDVI time series is illustrated in Figure 11. When applying the change detection algorithm over the NBR time series, it was found that 49.5% of roads were built post 1985, with 26.5% of roads being built prior to 1984 and 25% in

1985. The output and distribution of road age when applying the change detection method over the NDVI time series is illustrated in Figure 12.

Road density differences (Predicted – Reference)

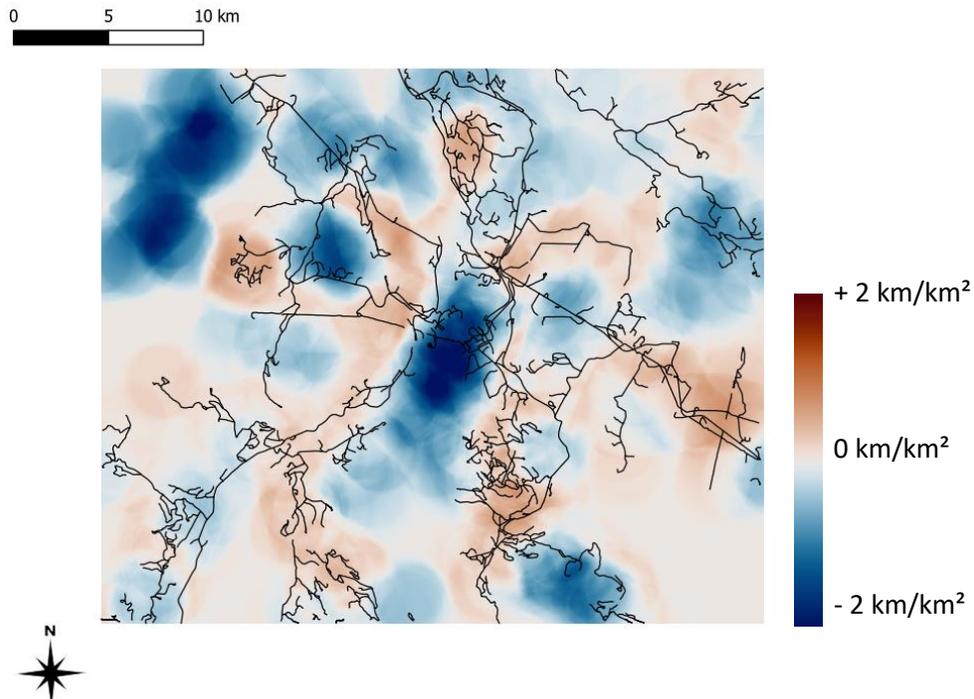


Figure 10: Differences in derived road densities. Negative values (blue) correspond to roads present in the reference dataset, but not in the predicted data. Positive values correspond to more detected roads than what was present in the reference dataset. The displayed road network is the predicted vectorized data.

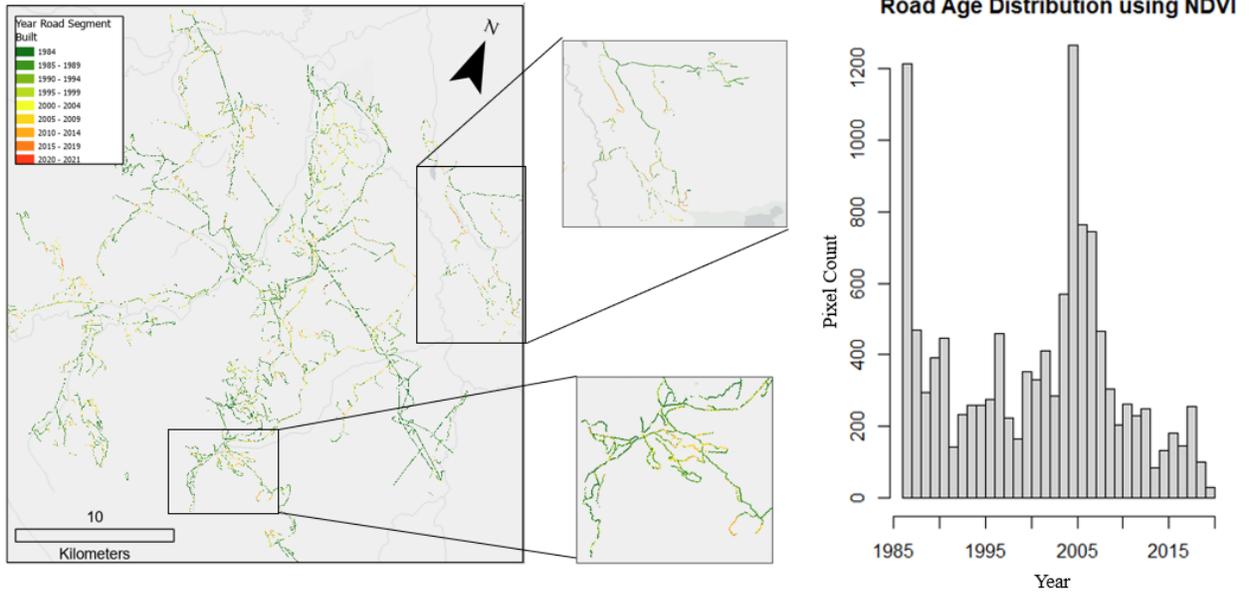


Figure 11: Output of change detection algorithm applied over NDVI time series, with supporting insets. Road age is grouped into 5-year time steps. 51% of roads were built post 1985, the histogram displays the distribution of roads built in this period

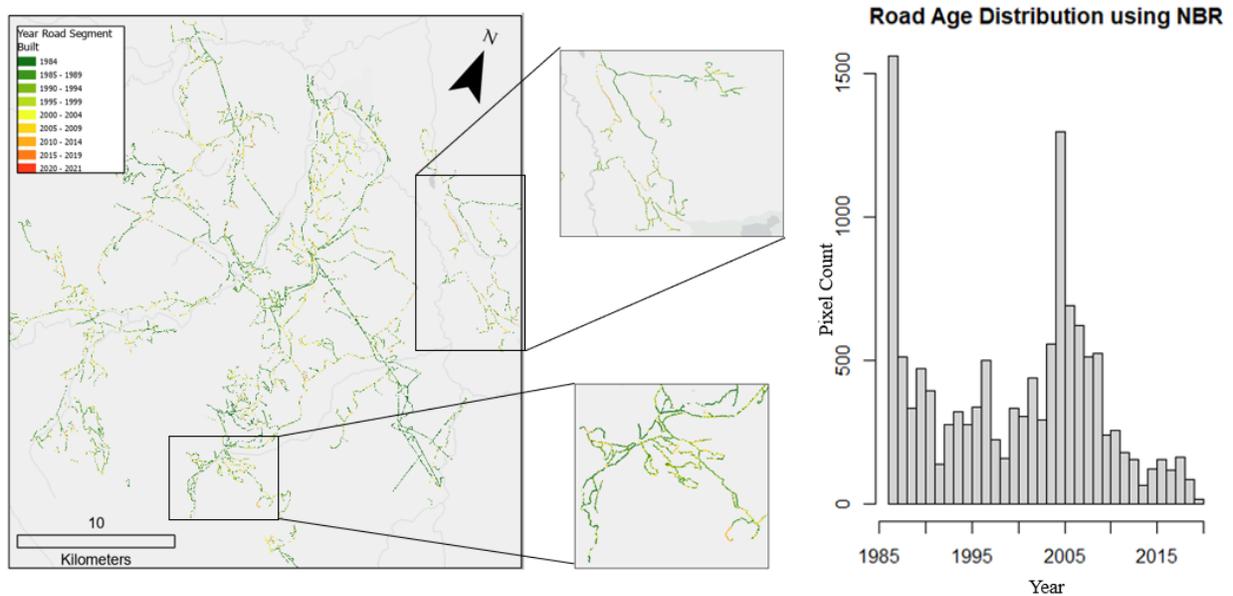


Figure 12: Output of change detection algorithm applied over NBR time series, with supporting insets. Road age is grouped into 5-year time steps. 49.5% of roads were built post 1985, the histogram displays the distribution of roads built in this period

Discussion

The CNN-based road extraction approach developed in this study and trained on PlanetScope imagery, achieved reasonably high performance on both the paved and unpaved rural roads typical of the boreal regions of Canada. *Recall*, *Precision* and *F1* quality metrics were comparable to studies using higher resolution imagery, such as Rapideye, to extract paved roads across much smaller areas using deep learning (e.g., Gao et al., 2019; Xu et al., 2018) and other approaches (e.g., Miao et al., 2015; Zhou et al., 2019). A key benefit of using data from CubeSats, such as PlanetScope, is a high spatial and temporal resolution. While PlanetScope imagery is typically available at daily or near-daily timesteps, clouds and cloud shadows increase the revisit period which can, in cloudy areas, exceed a number of weeks (Keay et al., 2023). In most cases this delayed temporal resolution is unlikely to cause issues, as the updating of road information is unlikely to occur more often than a few times per year. PlanetScope data products use anomalous brightness values to detect clouds. However, recently constructed roads with little remaining vegetation will also have anomalous brightness values and may be misclassified as cloud, thereby reducing the number of valid pixels from which to detect a change. With improvements in the UDM2 product and addition of new spectral bands to PlanetScope satellites, this misclassification of clouds may be reduced (Keay et al., 2023). In addition, while users can download PlanetScope scenes with a daily temporal revisit, Planet now offers image composites at coarser temporal resolutions (e.g., biweekly or monthly), which typically have lower cloud cover than that of the average scene over a given area.

A key benefit of the approach used in this study is the use of a pre-existing training set with which to build the initial CNN model. The use of deep learning algorithms such as CNN has expanded greatly within the past 5 years and advances in computing power and the ability to apply deep learning algorithms in a spatial context allows users previously unable to utilise the power of CNN, to now be able to do so. Roads are difficult objects to classify due to their spectral similarity to other high albedo objects such as bare soil and rock, and their morphological similarity to other linear features such as rivers and pipelines making a CNN approach logical to apply.

Another key outcome of the methods developed in this paper is the development of a fully connected road network, rather than separate, discrete, road segments. Even after post-processing, a pixel-based road extraction approach can be difficult to convert into a fully connected vectorized road network for navigation (Xu et al., 2018). The approach, developed by Rousell et al. (2023) and applied here, allows the

detected road segments to be connected in a vector dataset to ensure a topologically correct road network, which is required for most applications such as road planning, and conservation and needed in order to ingest this road information into existing vector databases.

Lastly our analysis indicates how long-term archival satellite data can be used to determine the age of forest roads in northern British Columbia using time series of spectral indices and annual Landsat best-available pixel image composites. The road age results produced from the NDVI and NBR time series are comparable, with minor variations year to year. Both methods, however, result in road segments along a road to fluctuate in age. This fluctuation in age is in part attributed to road segments within Landsat pixels being too small or affected by spectral noises from background land cover, resulting in no changes being detected. To account for this, future work will incorporate a minimum change detection threshold.

Up-to-date information on the rural road network is critical for the management of threatened and endangered species and is becoming a critical piece of information to aid in conservation activities. In British Columbia, with increasing levels of disturbance due to insect infestation, harvest, fire as well as ongoing anthropogenic disturbances associated with exploration and mining, the density of roads across the region remains of paramount concern. Quantifying the density of the road network within individual caribou herd ranges is one metric that can provide some context to how much activity may be occurring within each of these caribou herd ranges. This idea is similar to that undertaken with the management of the grizzly bear in Alberta where the Albertian provincial government recommended road densities below certain thresholds to encourage recovery of the species. While these thresholds are important, it is therefore critical that spatial coverages showing the extent of the road coverage are accurate and complete which is often not the case in remote areas. The technique presented here to detect roads from high spatial resolution Planet imagery, and then provide an indication as to their age from historic Landsat imagery, provides approaches where an accurate road network can be built. In addition, the ability to vectorized these predictions to produce a linked road network, whilst in its infancy, is the critical final step in ensuring these layers can then effectively be used for management.

Looking to the future, for the management potential of these types of technologies, one possible strategy would be to integrate GNSS data acquired from in vehicle loggers, which map the use of the landscape by vehicles, combined with geospatial coverages from satellite data such as demonstrated here could result in a database of both accurate road location information but also use. Information on the speed and diurnal use of roads could also be extracted from these databases to provide a comprehensive analysis of the road infrastructure in an area. Whilst these types of databases are starting to be developed within the

forest industry using GNSS receivers on harvesting operations, there broader use in a conservation setting has not yet been fully examined and may provide additional insights into road networks particularly for endangered species.

References

- Abrams, M., Crippen, R., Fujisada, H., 2020. ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). *Remote Sens.* 12, 1156. <https://doi.org/10.3390/rs12071156>
- Alberta Environment and Parks, 2016. Alberta Grizzly Bear (*Ursus arctos*) Recovery Plan,. Alberta Environment and Parks, Alberta Species at Risk Recovery Plan No. 38. Edmonton, AB.
- Alshehhi, R., Marpu, P.R., Woon, W.L., Mura, M.D., 2017. Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks. *ISPRS J. Photogramm. Remote Sens.* 130, 139–149. <https://doi.org/10.1016/j.isprsjprs.2017.05.002>
- Badrinarayanan, V., Handa, A., Cipolla, R., 2015. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling. *ArXiv150507293 Cs*.
- Barrington-Leigh, C., Millard-Ball, A., 2017. The world's user-generated road map is more than 80% complete. *PloS one* 12 (8), e0180698. <https://doi.org/10.1371/journal.pone.0180698>
- BC Government, 2021. Population Estimates for Caribou Herds of British Columbia [WWW Document]. URL https://www2.gov.bc.ca/assets/gov/environment/plants-animals-and-ecosystems/wildlife-wildlife-habitat/caribou/bc_caribou_herds_population_estimates.pdf (accessed 1.31.23).
- BC Ministry of Environment, 2010. Caribou in BC range by ecotype [WWW Document]. URL https://www.env.gov.bc.ca/wld/speciesconservation/caribou_by_ecotype.html#_ (accessed 1.31.23).
- Bennett, V.J., Betts, M.G., Smith, W.P., 2011. Toward understanding the ecological impact of transportation corridors, USDA Forest Service - General Technical Report PNW-GTR. <https://doi.org/10.2737/PNW-GTR-846>
- Das, S., Mirnalinee, T.T., Varghese, K., 2011. Use of salient features for the design of a multistage framework to extract roads from high-resolution multispectral satellite images. *IEEE Trans. Geosci. Remote Sens.* 49, 3906–3931. <https://doi.org/10.1109/TGRS.2011.2136381>
- Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., Hughes, F., Tuia, D., & Raskar, R. (2018). DeepGlobe 2018: A challenge to parse the earth through satellite images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 172-181).
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. ImageNet: A large-scale hierarchical image database, in: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. Presented at the 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>
- Environment Canada, 2010. Working Together to Recover Boreal Caribou [WWW Document]. URL https://publications.gc.ca/collections/collection_2010/ec/En14-16-2010-eng.pdf (accessed 1.31.23).
- Ferraz, A., Mallet, C., & Chehata, N. (2016). Large-scale road detection in forested mountainous areas using airborne topographic lidar data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 112, 23-36.
- Forest Practices Code (2002). *Forest Road Engineering Guidebook*. Ministry of Forests, British Columbia.
- Frazier, A. E., & Hemingway, B. L. (2021). A technical review of planet smallsat data: Practical considerations for processing and using planetscope imagery. *Remote Sensing*, 13(19), 3930.

- Gao, L., Song, W., Dai, J., Chen, Y., 2019. Road Extraction from High-Resolution Remote Sensing Imagery Using Refined Deep Residual Convolutional Neural Network. *Remote Sens.* 11, 1–16.
<https://doi.org/10.3390/rs11050552>
- GeoBC (2017). Resource Roads Spatial Data Update. Government of BC. <https://www2.gov.bc.ca/gov/content/data/geographic-data-services/topographic-data/roads/resource-roads>
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press. www.deeplearningbook.org
- Helman, D., Lensky, I. M., Tessler, N., & Osem, Y. (2015). A phenology-based method for monitoring woody and herbaceous vegetation in Mediterranean forests from NDVI time series. *Remote Sensing*, 7(9), 12314–12335.
- Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015a). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. *Remote Sensing of Environment*, 158, 220–234.
<https://doi.org/https://doi.org/10.1016/j.rse.2014.11.005>
- Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015b). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment*, 170, 121–132.
<https://doi.org/https://doi.org/10.1016/j.rse.2015.09.004>
- Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., & Campbell, L. B. (2016). Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring. *International Journal of Digital Earth*, 9(11), 1035–1054.
<https://doi.org/10.1080/17538947.2016.1187673>
- Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., 2022. Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes. *Remote Sens. Environ.* 268, 112780. <https://doi.org/10.1016/j.rse.2021.112780>
- Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2018. Disturbance-Informed Annual Land Cover Classification Maps of Canada's Forested Ecosystems for a 29-Year Landsat Time Series. *Can. J. Remote Sens.* 44, 67–87. <https://doi.org/10.1080/07038992.2018.1437719>
- Hmimina, G., Dufrêne, E., Pontailleur, J.-Y., Delpierre, N., Aubinet, M., Caquet, B., De Grandcourt, A., Burban, B., Flechard, C., & Granier, A. (2013). Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. *Remote Sensing of Environment*, 132, 145–158.
- Kearney, S. P., Coops, N. C., Sethi, S., & Stenhouse, G. B. (2020). Maintaining accurate, current, rural road network data: An extraction and updating routine using RapidEye, participatory GIS and deep learning. *International Journal of Applied Earth Observation and Geoinformation*, 87, 102031.
- Keay, L., Mulverhill, C., Coops, N. C., & McCartney, G. (2023). Automated Forest Harvest Detection With a Normalized PlanetScope Imagery Time Series. *Canadian Journal of Remote Sensing*, 49(1), 2154598.
- Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sensing of Environment*, 114(12), 2897–2910.
- Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., Gao, F., Kutler, J., & Lim, T.-K. (2006). A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geoscience and Remote Sensing Letters*, 3(1), 68–72.
<https://doi.org/10.1109/LGRS.2005.857030>

- Miao, Z., Shi, W., Gamba, P., Li, Z., 2015. An Object-Based Method for Road Network Extraction in VHR Satellite Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8, 4853–4862. <https://doi.org/10.1109/JSTARS.2015.2443552>
- Mnih, V. (2013). Machine learning for aerial image labeling. PhD thesis, University of Toronto (Canada).
- NASA/METI/AIST/Japan Spacesystems and U.S./Japan ASTER Science Team, 2019. ASTER Global Digital Elevation Model V003. <https://doi.org/10.5067/ASTER/ASTGTM.003>
- Proctor, M.F., McLellan, B.N., Stenhouse, G.B., Mowat, G., Lamb, C.T., Boyce, M.S., 2020. Effects of roads and motorized human access on grizzly bear populations in British Columbia and Alberta, Canada. *Ursus* 2019, 16–39. <https://doi.org/10.2192/URSUS-D-18-00016.2>
- Roussel, J.-R., Bourdon, J.-F., Morley, I. D., Coops, N. C., & Achim, A. (2022). Correction, update, and enhancement of vectorial forestry road maps using ALS data, a pathfinder, and seven metrics. *International Journal of Applied Earth Observation and Geoinformation*, 114, 103020.
- Roussel, J.-R., Bourdon, J.-F., Morley, I. D., Coops, N. C., & Achim, A. (2023). Vectorial and topologically valid segmentation of forestry road networks from ALS data using a self-driving agent. Submitted to *International Journal of Applied Earth Observation and Geoinformation*.
- Schmidt, G. L., Jenkerson, C., Masek, J. G., Vermote, E., & Gao, F. (2013). Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description. US Department of the Interior, US Geological Survey.
- Selva, N., Switalski, A., Kreft, S., Ibisch, P.L., 2015. Why keep areas road-free? The importance of roadless areas, in: Van Der Ree, R., Smith, D.J., Grilo, C. (Eds.), *Handbook of Road Ecology*. John Wiley & Sons, Ltd., pp. 16–26.
- Sherba, J., Blesius, L., & Davis, J. (2014). Object-based classification of abandoned logging roads under heavy canopy using LiDAR. *Remote Sensing*, 6(5), 4043-4060.
- Simonyan, K., Zisserman, A., 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. <https://doi.org/10.48550/arXiv.1409.1556>
- Sonogashira, M., Shonai, M., Iiyama, M., 2020. High-resolution bathymetry by deep-learning-based image superresolution. *PLOS ONE* 15, e0235487. <https://doi.org/10.1371/journal.pone.0235487>
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127–150. [https://doi.org/https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/https://doi.org/10.1016/0034-4257(79)90013-0)
- Waga, K., Tompalski, P., Coops, N. C., White, J. C., Wulder, M. A., Malinen, J., & Tokola, T. (2020). Forest road status assessment using airborne laser scanning. *Forest Science*, 66(4), 501-508.
- Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T., & Eklund, P. (2016). A review of road extraction from remote sensing images. *Journal of Traffic and Transportation Engineering (English Edition)*, 3(3), 271-282.
- White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., Coops, N. C., Hall, R. J., Hostert, P., Dyk, A., et al. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, 40(3), 192-212.
- Xu, Y., Xie, Z., Feng, Y., Chen, Z., 2018. Road extraction from high-resolution remote sensing imagery using deep learning. *Remote Sens.* 10. <https://doi.org/10.3390/rs10091461>
- Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., et al. (2020). Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment*, 241, 111716. Elsevier.
- Zhang, Z., Liu, Q., Wang, Y., 2018. Road Extraction by Deep Residual U-Net. *IEEE Geosci. Remote Sens. Lett.* 15, 749–753. <https://doi.org/10.1109/LGRS.2018.2802944>
- Zhou, T., Sun, C., Fu, H., 2019. Road Information Extraction from High-Resolution Remote Sensing Images Based on Road Reconstruction. *Remote Sens.* 11. <https://doi.org/10.3390/rs11010079>