Utilizing Deep Learning for Mapping Dozer Lines from Aerial Imagery

Jonathan Calderon Chavez, Matthew Clark, Gurman Gill

Sonoma State University, Rohnert Park, United States

Abstract

This article explores the utility of using aerial imagery and deep learning to automate dozer line mapping. Due to the lack of dozer line data, we leveraged dirt and gravel road data to estimate dozer line characteristics and trained deep learning networks. We examined the performance of deep neural networks composed of UNET decoders with different ResNet encoders. Preliminary results demonstrated that a deep neural network composed of ResNet-18 and UNET provided an overall IOU score of 0.91 when used to map dirt and gravel roads in Sonoma County during the 2017 North Tubbs Fire.

1. Introduction

1.1. Wildfire impact

California faces destructive wildfires which have impacted the lives of its residents. In 2017 alone, wildfires destroyed over 10 thousand buildings and caused the deaths of around 50 residents [1]. The Tubbs fire became one of California's most destructive and deadliest wildfires [2]. The loss of jobs, homes, and lives due to wildfires drives the need for effective methods of combating wildfires.

1.2. Dozer line usage and its environment impact

CAL FIRE works to mitigate the impact of wildfires in affected areas throughout California. CAL-Fire employs a variety of tactics such as controlled fires, air-based attacks, and removal of fire-fueling objects such as vegetation. A widely employed tactic used during a wildfire is constructing dozer lines: fire barriers created by bulldozers through removing vegetation in strategic areas with the intent to stop wildfire spread. While an effective tactic, the creation of dozer lines causes erosion, harms soil fertility, and pollutes nearby water sources [3]. The impact on the environment can cascade down to harm the local wildlife and communities [3].

The reusing and mapping of dozer lines are highly desirable to minimize the environmental impact and reduce the time to contain a wildfire. The current methods of mapping dozer lines rely on inefficient methods such as GPS mapping and prior personnel knowledge. Mapping dozer lines using GPS mapping requires costly manual labor. The data collected from these methods become outdated as firefighters create new dozer lines while others become overgrown, making these maps unreliable. Some property owners in affected areas construct dozer lines on their property during wildfires, and collecting such data becomes difficult due to private property rights. Combating the issue of outdated and inaccurate data requires a rapid and less invasive method of mapping dozer lines.

1.3. Deep learning as a viable solution

Deep learning focuses on algorithms designed to solve complex problems traditionally thought to require human intelligence, e.g., object classification and detection. When the goal is to delineate an object's boundary in the image, it is called a semantic segmentation problem. These problems can be solved using Artificial Neural Networks, which are data structures similar to the human brain neural network. Deep learning (DL) algorithms that leverage artificial neural networks can learn to solve such complex problems without being explicitly programmed to solve them by learning and recognizing data patterns [4]. Deep learning achieves good results in several different areas (medicine, security, gaming, environmental) with the added benefit of being significantly faster than humans [5].

The task of mapping dozer lines from aerial imagery relates to other researchers' ongoing works in using deep learning for mapmaking. Related work includes using deep learning to map rural road networks to update government-maintained maps in Western Canada [14]. The researchers leverage high-resolution satellite imagery to train deep learning networks such as SegNet [15] and obtain promising results. However, the networks struggled with distinguishing between roads and other structures such as gas pipelines, rivers, and clouds [14]. In another study, researchers have explored the utility of using deep learning to map roads of different surfaces such as gravel, asphalt, and dirt throughout Spain [16]. Their findings indicated other challenges of mapping roads, such as lack of defined road edges, differences in road widths, and significant curvature changes. The researchers noted U-Net [11] as the best performing segmentation architecture in their study [16].

DL algorithms' ability to solve new problems has excellent potential to automate dozer line mapping. If mapping through deep learning is feasible, the method can promote reusing dozer lines and monitoring environment impact by rapidly creating reliable maps without requiring strenuous manual labor.

2. Dataset

2.1. Sources used to create dataset

An annotated dataset was crucial to train and evaluate a deep neural network. A dataset comprising image data of dozer lines with corresponding images depicting their location, called a segmentation mask, was required to solve a semantic segmentation problem (Fig. 2). The dataset for this project used Post-fire RGB-IR aerial imagery of the Sonoma County 2017 wildfire season [6]. The aerial imagery captures areas in Sonoma County affected by the Tubbs, Nunns, and Pocket fires (Fig. 1). The dataset used imagery from within the fire boundaries of each fire to avoid irrelevant data. The dataset utilized a false-color version of the post-fire imagery to help pronounce helpful features such as vegetation using the near-infrared data. Vegetation significantly reflects near-infrared light, which can make vegetation-less roads noticeable.

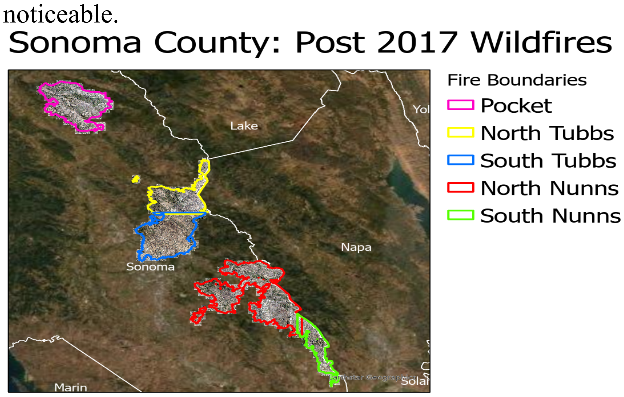

Figure 1. Extent of the aerial imagery.

The low accuracy of GPS data of the dozer lines locations created in 2017 creates a mismatch between the GPS data points and the imagery. Training a DL algorithm based on such data was not ideal because it includes many mislabeled data. Consequently, we utilized an artificial impervious surface map of the county in 2013 [7]. The map locates most dirt and gravel roads during 2017 that are wide enough for vehicle traffic, closely resembling dozer lines' properties. We are estimating dozer lines through the use of dirt and gravel roads. However, the map may miss roads and dozer lines made between 2013 and 2017. The map was modified to remove road segments covered by canopy using a land cover map as a reference[8]. Roads covered by canopy can mislead a deep learning algorithm associating canopy features as dirt and gravel road features.

We used *Esri's ArcPy*, a library of GIS tools [9], to create the annotated dataset. Using the *export training data for the deep learning* method from *ArcPy* made 256 x 256 image chips and segmentation mask chips from the map and aerial imagery within the fire perimeter.

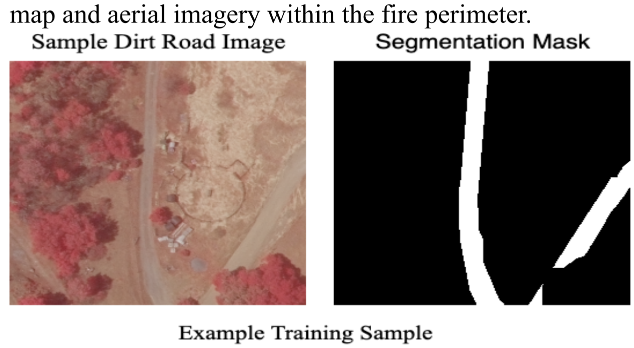

Figure 2. An example segmentation mask where black regions are labeled as “background” and white regions are labeled as “dirt/gravel road.”

2.2. Data splitting for model evaluation

After creating the dataset, the dataset was partitioned into three different sets: training, validation, and test. The DL network uses the training set during the training process to find features and distinguish between the two main classes: “dirt/gravel roads” and “background.” The validation set was used during the training to monitor the performance of the deep networks and adjust their parameters. The test set was the final assessment of the network's performance and used to compare and contrast different models. The data samples in the test set are entirely held out from the training process to avoid bias.

The training data comprise the North Nunns fire, South Nunns fire, South Tubbs fire, and Pocket fire containing dirt and gravel roads for a total of 8,821 samples. Thirty percent of this total were randomly selected and held out to be used as the validation set. The test set comprises all data from North Tubbs fire regardless of the presence of dirt and gravel roads to emulate a real-world use of the DL network. The motive for separating the test region from the training region was to measure the performance of the deep network when exposed to new geographical and regional landscapes.

3. Methods

3.1. Deep neural networks overview

Deep neural networks (DNN) have trainable parameters that allow the network to find features that help distinguish dirt and gravel roads from the background class. These parameters are fine-tuned to our dataset during the training process, which is a cycle of predicting the validation set and parameter adjustments. When a DNN is done predicting the testing, a loss function calculates its performance. An optimizer is then used to take the loss function value and adjust the DNN's parameters. A low loss function value indicates convergence and hyper-parameter quality. We will be using the Adam optimizer and testing different learning rates to find the optimal configuration for our dataset.

3.2. Using transfer learning

Fine-tuning a DNN's trainable parameters from scratch is not a straightforward task. A network may need thousands or even millions of diverse labeled images to find suitable parameter settings. With the small size of the dataset, it is not feasible to train deep neural networks from scratch. However, leveraging pre-trained deep learning neural networks offers a great workaround through a process called transfer learning [10]. Transfer learning is the concept of using parameter settings/ weights of pre-trained networks as a starting point in the training process of similar problems such as dozer line mapping. This approach requires less data for training to converge.

3.3. Using UNET decoder and different ResNet encoders

The process of creating an effective deep neural network required searching and testing a wide array of deep network architectures and hyper- parameters. For this project, we focused on using a UNET decoder with different residual neural network (ResNet) encoders [11].

An encoder in a UNET model is tasked with feature learning, while the decoder uses the learned feature from the encoder to distinguish our two classes. The different ResNet encoders tested were initially trained to classify objects from the ImageNet dataset: a dataset containing over 1 million images of 1000 different objects [12]. The encoders are frozen, meaning that the encoder parameter values were not changed during training.

The difference in the different encoders was their length. The longer the encoder, the more trainable parameters that can create more refined features. However, using larger encoders comes at the cost of increasing time and computational resources for training. This project experimented with ResNet-18, ResNet-34, ResNet-50, and ResNet-101.

4. Results

4.1. Metrics and Search Area

We trained various networks with different ResNet encoders, learning rate values, and epochs to search for the optimal parameters. We used learning rate values of 0.0001, 0.001, and 0.01, and epochs values of 5 to 50 in intervals of 5 in order to limit the number of different possible combinations. With these values, we trained a total of 120 different DL networks. After training, we evaluated the network's performance on the test set using the metric called the intersection over the union (IOU) score. The IOU score indicates the amount of overlap between the ground truth label and the network's predictions, with IOU Scores of 0.5 or greater considered good predictions [13].

4.2. UNET performance on the entire test set

The results demonstrated that the highest performing UNET model was obtained using a ResNet-18 encoder, a learning rate of 0.1, and trained for 15 epochs (Figs. 3-5). The optimal UNET network received an IOU score of .91 on the test set. It is important to note that the test set was imbalanced since more images contained only the background class than images with dirt/gravel roads. To better understand the different UNET networks' performance, we further separated the test set into two groups, "images with dirt/gravel road" and "image no dirt/ gravel roads." We computed the IOU score on both groups separately.

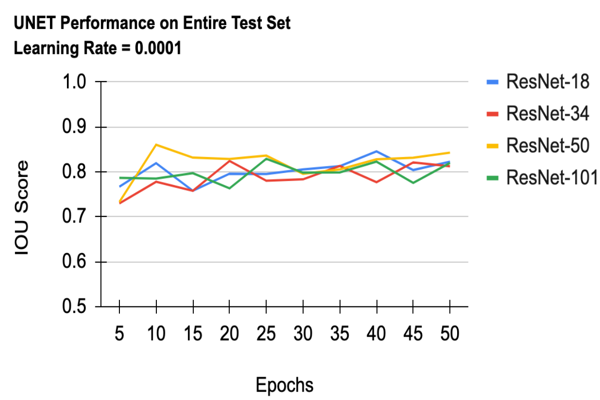

Figure 3. Performance of UNET networks trained with a learning rate of 0.0001

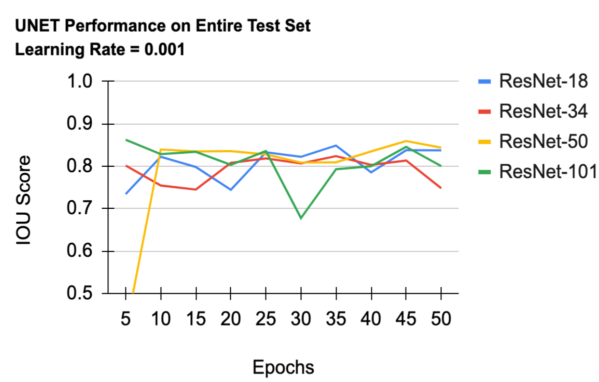

Figure 4. Performance of UNET networks trained with a learning rate of 0.001

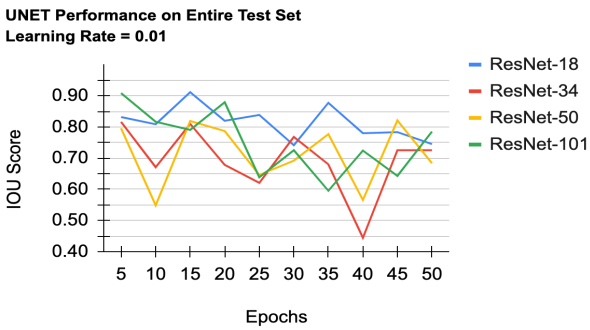

Figure 5. Performance of UNET networks trained with a learning rate of 0.01

4.3. Performance on images with dirt or gravel roads

When testing the different UNET networks on only images that contained dirt/gravel roads, the resulting IOU scores were significantly different from those obtained from testing on the entire test set (Fig. 6). The IOU score across all the UNET networks ranged between .2 and .43, significantly lower than the scores obtained when testing on the entire test set. However, it is important to note that in various examples, the prediction quality is good, but the width of the estimated track does not match consistently, leading to a low IOU score. The model that obtained the highest IOU score on the entire test set obtained one of the lowest IOU scores, a score of 0.25, when evaluated on only images containing dirt/gravel roads.

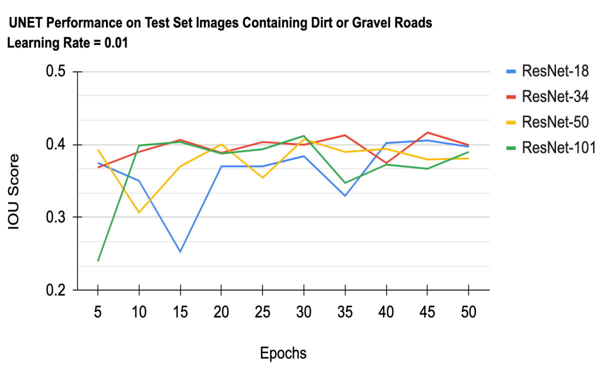

Figure 6. Performance of UNET networks trained with a learning rate of 0.01 on images containing dirt or gravel roads.

4.4. Performance of images with no dirt or gravel roads

The ability to avoid predicting dirt and gravel roads is an important objective, considering the vast amount of artificially made structures resembling dirt and gravel roads, e.g., canals, hiking trails, paved roads, and farming fields. When evaluating the different UNET Networks using only images with no dirt/images, the IOU scores obtained are similar to those tested on the entire test set (Fig. 7).

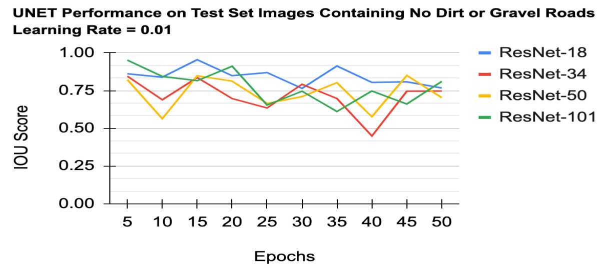

Figure 7. Performance of UNET networks trained with a learning rate of 0.01 on images not containing dirt or gravel roads.

4.5. Examples Predictions

The UNET Network that obtained the highest IOU score on the entire test demonstrated some promising qualities in mapping dirt and gravel roads. One such example is that the network avoided labeling linear artificial structures on farmland as dirt or gravel roads (Fig. 8). Another excellent quality is its ability to correctly predict curved roads in some test images (Fig. 9). However, one of the downfalls of this network is when it comes to mapping dirt and gravel roads in areas where it intersects or branches into two or more roads (Fig. 10).

Figure 8. Network's prediction on linear artificial structure from the test set

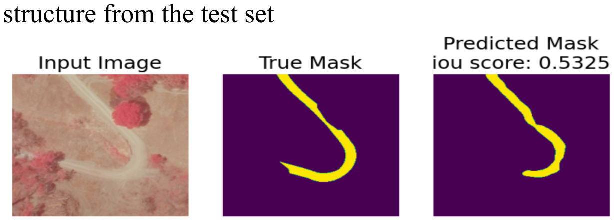

Figure 9. Network's prediction on a road with large curvature from the test set

Figure 10. Network's prediction on a road that self-intersects from the test set

5. Overview of results

The results demonstrate the feasibility of creating an automated dozer line mapping system using aerial imagery and deep learning. Despite the lack of usable dozer line data for DL, estimating dozer lines with dirt and gravel road data gave us a view of some traits that make unpaved roads challenging to delineate their boundaries from images.

6. Discussion and Future Work

6.1. Outdated dirt and gravel road data

We used dirt and gravel road data from 2013 with 2017 aerial imagery to train and evaluate the DNNs. The outdated road data will cause inaccuracies in the overall IOU score. In cases where a DNN will correctly predict a dirt or gravel road created after 2013, the DNN will be penalized as that road will be mislabeled as background in the ground truth label. In the training process, the mislabeling of roads as the background will cause the DNN to incorrectly associate dirt and gravel road features as the background, lowering the performance.

6.2. Estimating dozer lines through dirt and gravel roads

Not using dozer lines to train the DNNs will cause them not to learn the unique properties of dozer lines. Not learning to find these properties can make it difficult to distinguish between dozer lines and other types of dirt and gravel roads. We plan to digitize and accurately label the GPS data of dozer lines made in 2017 to match the post-fire aerial imagery. Successfully correcting the dozer line data will eliminate the issues associated with using outdated dirt and gravel road data.

6.3. Leave-one-Fire-out validation

The current evaluations of the model are not exhaustive to indicate the actual performance of the different neural networks explored in this research paper. The selection of using North Tubbs data as a test set may be unrealistic of real-world examples, causing the obtained IOU to either underestimate or overestimate the real IOU when used in the real world.

Future evaluations of model performance will use Leave-one-Fire-out validation to combat such biases. This technique will allow the deep network to be trained and tested on different partitions of the dataset such that each partition is an independent fire region, giving a greater insight into the real-world performance case.

References

[1] Fire C. Wildfire activity statistics. California Department of Forestry and Fire Protection.

[2] Fire C. Top 20 deadliest California wildfires.

[3] Ingalsbee T, Beasley M, Plummer D, Cowen M. Carr Fire CATlines: The Environmental Impacts of Bulldozers in Wildfire Suppression.

[4] Y. LeCun, Y. Bengio, and G. Hinton. 2015. Deep learning. Nature 521(7553):436–444.

[5] Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. InArtificial Intelligence in healthcare 2020 Jan 1 (pp. 25-60). Academic Press.

[6] Sonoma County Water Agency, Sonoma County Agricultural Preservation and Open Space District, Sonoma County Vegetation Mapping and LiDAR Program. 2013.

[7] Sonoma Veg Map, Sonoma County Agricultural Preservation and Open Space District, Sonoma County Information Services, Sonoma County Transportation and Public Works, University of Vermont Spatial Informatics Group, and Tukman Geospatial. Sonoma County Impervious 2013

[8] Sonoma County Water Agency, Sonoma County Agricultural Preservation and Open Space District, Sonoma County Vegetation Mapping and LiDAR Program. Sonoma County Lifeform 2013

[9] Environmental Systems Research Institute, Inc., 2010 ArcGIS ArcPy Redlands, CA

[10] J.Yosinski, J. Clune, Y. Bengio, and H. Lipson. 2014. How transferable are features in deep neural networks? Pages 3320–3328 in Advances in neural information processing systems.

[11]Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham.

[12] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei. 2009. ImageNet: A Large-Scale Hierarchical Image Database. Pages 248-255 in IEEE Computer Vision and Pattern Recognition

[13] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," arXiv preprint arXiv:1411.4038, 2014. [Online].

[14]Kearney SP, Coops NC, Sethi S, Stenhouse GB. 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. 2020 May 1;87:102031.

[15] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence. 2017 Jan 2;39(12):2481-95.

[16] Cira CI, Alcarria R, Manso-Callejo MÁ, Serradilla F. A deep learning-based solution for large-scale extraction of the secondary road network from high-resolution aerial orthoimagery. Applied Sciences. 2020 Jan;10(20):7272.

Address for correspondence:
Jonathan Calderon Chavez.
1801 East Cotati Ave, Rohnert Park, CA 94928
Jonathan.c98@yahoo.com