

NORTHWEST NAZARENE UNIVERSITY

Wildland Fire Tree Mortality Mapping from Hyperspatial Imagery Using Machine Learning

THESIS

Submitted to the Department of Mathematics and Computer Science  
in partial fulfillment for the degree of  
BACHELOR OF ARTS

Samuel Jones  
2021

THESIS  
Submitted to the department of Mathematics and Computer Science  
in Partial fulfillment of the requirements  
for the degree of  
BACHELOR OF ARTS

Samuel Jones  
2021

Wildland Fire Tree Mortality Mapping from Hyperspatial Imagery Using Machine Learning

Author: Samuel Jones  
Samuel David Jones

Approved: Dale A Hamilton  
Dr. Dale A. Hamilton, Ph.D., Associate Professor, Department of Mathematics and Computer Science, Faculty Advisor

Approved: Benjamin L. Fischer  
Dr. Benjamin Fischer, Ph.D., Associate Professor of Language and Literature, Second Reader

Approved: Barry Myers  
Dr. Barry L. Myers, Ph.D., Chair, Department of Mathematics and Computer Science

## ABSTRACT

Wildland Fire Tree Mortality Mapping from Hyperspatial Imagery Using Machine Learning

JONES, SAMUEL (Department of Mathematics & Computer Science),  
HAMILTON, DR. DALE (Department of Mathematics and Computer Science)

The use of small Unmanned Aircraft Systems (sUAS) for obtaining wildland imagery has enabled the production of more accurate data regarding the effects of fire on forested land. This increase in precision enables accurate detection of trees from hyperspatial imagery, and thus the calculation of canopy cover. When pre-fire data is compared with post-fire data from existing canopy cover products such as the LANDFIRE project, a measure of tree mortality, which is a measure of burn severity, can be calculated from the difference between the two. A mask region-based convolutional neural network (MR-CNN) was trained to classify trees as groups of pixels from a hyperspatial orthomosaic. The tree classification is summarized at 30m in a canopy cover raster. Canopy reduction allows the mapping of burn severity while also identifying where surface, passive, and active crown fire occurred within the observed region. The calculated canopy cover derived from hyperspatial imagery was found to be significantly more accurate when compared with LANDFIRE canopy cover in all observed regions. The hyperspatially derived canopy cover reflected visual observations from respective sUAS imagery and ground truthing.

## **Acknowledgements**

I would like to thank my family for supporting me through my academic career. I also want to thank Dr. Dale Hamilton for the opportunity to work with the FireMAP team on this project and to co-author the resulting paper. I also want to thank Kamden Brothers for being a great co-worker while working on this project remotely during the COVID-19 pandemic, Jacob Winters for taking the time to assist me in troubleshooting the Mask R-CNN, Braelyn Boerner for helping us create validation data, as well as the previous students that worked on previous FireMAP projects that laid the groundwork to make this project possible.

## Table of Contents

Abstract .....	iii
Acknowledgements .....	iv
1. Overview .....	1
2. Publication .....	2
2.1. Introduction .....	4
2.1.1 Background .....	5
2.2. Materials and Methods: Mapping Canopy Cover and Tree Mortality .....	8
2.2.1 Classification of Tree Crowns from Hyperspatial Imagery .....	8
2.2.2 Calculating Canopy Cover .....	11
2.2.2.1 Calculate Five Centimeter Class Densities for each 30 Meter Pixel .....	11
2.2.2.2 Canopy Cover Analysis .....	12
2.2.3 Canopy Cover: sUAS Derived vs LANDFIRE .....	13
2.2.3.1 Canopy Cover Comparison Methods .....	14
2.2.3.2 Canopy Cover Comparison Analysis .....	16
2.2.4 Mapping Tree Mortality via Canopy Reduction .....	18
2.2.4.1 Canopy Reduction Methods .....	19
2.2.4.2 Canopy Reduction Analysis .....	20
2.3. Results .....	21
2.3.1 Tree Crown Classification Validation .....	21
2.3.2 Canopy Cover Analysis .....	26
2.3.3 Canopy Cover Comparison: Hyperspatial Derived vs LANDFIRE .....	27
2.3.4 Tree Mortality Mapping Assessment .....	29
2.4. Discussion .....	31
2.5. Conclusions .....	33
2.5.1 Future Work .....	33
3. Finale .....	39
References .....	41

## Table of Figures

Figure 1: Mask R-CNN training data.....	10
Figure 2: (a) Unclassified 5cm sUAS image (b) Classified 5cm sUAS image .....	11
Figure 3: HDCC Methods Flowchart.....	16
Figure 4: South Placersville HDCC .....	27
Figure 5: Comparison between the LANDFIRE and HDCC data .....	28
Figure 6: (a) 5cm observation area (b) HDCC of observation area .....	30
Figure 7: (a) 5cm observation area (b) HDCC of observation area .....	30
Table 1: Confusion matrices of MR-CNN tree crown classifications.....	25
Table 2: MR-CNN tree crown classification statistical summary .....	26
Table 3: Mean difference and the standard deviation.....	28

## 1. Overview

The ecoinformatics research at NNU began in 2014 with the research Dr. Hamilton worked on in obtaining his PhD. Since then, the NNU FireMAP (Fire Mapping and Assessment Platform) project has built on this work to develop multiple solutions for assessing wildfire effects and/or severity using sUAS (small Unmanned Aircraft Systems) data. Among these is a machine learning classifier using a Mask Regional-Convolutional Neural Network (Mask R-CNN) model, developed by Jacob Winters, which can accurately analyze forest canopy cover from five-centimeter drone imagery.

The goal of this research project is to build off Jacob Winters work with the trained Mask R-CNN classifier to calculate current canopy cover with the use of sUAS data and evaluate it against calculated canopy cover recorded before a wildland fire event. This pre-fire data is available through the LANDFIRE program, which is derived from nationwide 30-meter spatial resolution Landsat imagery. By comparing pre-fire canopy cover data from LANDFIRE and post-fire canopy cover data from sUAS imagery, we can calculate the canopy reduction caused by wildfires.

A snapping tool (referred to as the “snapper”) was developed by Nick Hamilton to process the sUAS drone imagery, which has a resolution of 5cm, to the same extent of the calculated canopy cover derived from satellite data in the LANDFIRE program. Without snapping the sUAS data to the extent of the LANDFIRE data, a pixel-to-pixel comparison would not be accurate [22].

Using 30-meter Landsat pixels based on 5-centimeter classification from drone imagery, we can calculate canopy cover of a given area, then compare the pre-fire and post-fire data sets to see if a canopy reduction has occurred in the observation area. If partial reduction in forest crown is found, then a passive crown fire was present. If a large

reduction in forest crown is found, then an active crown fire was present. If the total reduction in forest crown is within the margin of error, then an active ground fire was present. These measures can give a glimpse of how a wildland area was affected by fire without the need of a ground crew to examine the entire sight.

This project was a collaborative effort between Dr. Dale Hamilton, Dr. Jason Colwell, Kamden Brothers, Jacob Winters, and myself during our tenured research period over the summer of 2020. My role in this project focused on preparing the sUAS data for classification data using the Mask R-CNN, classifying sUAS data, running density calculations, along with other data processing tasks.

sUAS data preparation involved splitting orthomosaics and matching their spatial referencing to the spatial referencing of the LANDFIRE data set using the snapper. Batch files were created to streamline the classification process of the sUAS data with the Mask R-CNN. Kamden and I worked together on creating the canopy reduction layers using ArcGIS tools outlined in section 2.2. I was also a co-author on a research paper that was written about our findings. It was accepted for publication in MDPI's Remote Sensing research journal this last January.

## **2. Publication**

The following paper was the result of the research completed by NNU's summer FireMAP research team from May-August 2020 and was published in MDPI's Remote Sensing research journal on January 15<sup>th</sup>, 2021. This study establishes the accuracy of calculating canopy cover with a Mask R-CNN from sUAS data as well as the limits and capabilities of using LANDFIRE canopy cover as a pre-fire canopy cover dataset. The paper should be cited as the following:

D. A. Hamilton, K. L. Brothers, S. D. Jones, J. Colwell, and J. Winters, “Wildland Fire Tree Mortality Mapping from Hyperspatial Imagery Using Machine Learning,” *Remote Sensing*, vol. 13, no. 2, p. 290, Jan. 2021.

**Abstract:** The use of imagery from small unmanned aircraft systems (sUAS) has enabled the production of more accurate data about the effects of wildland fire, enabling land managers to make more informed decisions. The ability to detect trees in hyperspatial imagery enables the calculation of canopy cover. Comparison of hyperspatial post-fire canopy cover and pre-fire canopy cover from sources such as the LANDFIRE project enables the calculation of tree mortality, which is a major indicator of burn severity. A Mask Region-based Convolutional Neural Network was trained to classify trees as groups of pixels from a hyperspatial orthomosaic acquired with a small unmanned aircraft system. The tree classification is summarized at 30 meters, resulting in a canopy cover raster. A post-fire canopy cover is then compared to LANDFIRE canopy cover preceding the fire, calculating how much the canopy was reduced due to the fire. Canopy reduction allows mapping of burn severity, while also identifying where surface, passive crown, and active crown fire occurred within the burn perimeter. Canopy cover mapped through this effort was lower than the LANDFIRE Canopy Cover product, which literature indicated is typically over reported. Assessment of canopy reduction mapping on a wildland fire reflects observations made both from ground truthing efforts as well as observation made of the associated hyperspatial sUAS orthomosaic.

## **2.1. Introduction**

This study investigates the use of machine learning to identify burn severity as a measure of canopy reduction. Tree crowns are classified from hyperspatial (sub-decimeter resolution) imagery using a Mask Region-based Convolutional Neural Network (MR-CNN). The tree crown classification is then summarized at 30-meter resolution, resulting in a canopy cover map. A post-fire canopy cover mapping product was compared to a pre-fire canopy cover mapping layer from the LANDFIRE project. The difference between the pre-fire and post-fire canopy cover enables a calculation of canopy reduction resulting as an effect of the fire, which gives an indication of tree mortality which is a primary indicator of post-fire effects resulting from a wildland fire.

Forest management practices, including fire suppression efforts, have led to forests with higher density than were found under pre-European settlement conditions. As a result, wildlands are experiencing a higher frequency of catastrophic fires, with millions of acres burned across the western US every year, often burning with much higher severity than experienced under historic conditions. These higher intensity fires result in increased incidence of soil erosion, hydrologic degradation, and deteriorating ecosystem resilience. The ability to detect trees in hyperspatial imagery enables the calculation of canopy reduction, which is a measure of wildland fire severity. The availability of canopy reduction data enables better decision making for ecosystem management and fire effect mitigation.

### **2.1.1 Background**

The term “wildland fire severity” can refer to many different effects observed through a fire cycle, from how intensely an active fire is burning to the response of the ecosystem to the fire over the subsequent years. As opposed to the intensity of an active fire which is often referred to as fire severity, this study investigates direct or immediate effects of a fire as observed in the days and weeks after the fire is contained, which we refer to as burn severity [1], [2]. Previous studies in this effort examined the effect of the fire on biomass consumption which indicates how much surface vegetative material has been consumed by the fire [1], [3]–[5]. By contrast, this study investigates the effect that fire has had on canopy vegetation in forested ecosystems. In situations where the amount of current post-fire canopy has not been significantly altered from what can be observed from the pre-fire conditions, it can be assumed that a fire was only consuming surface vegetation (or possibly the lower canopy) but did not consume the upper portion of the tree crowns in the canopy [6]. This low intensity surface fire results in no mortality of the trees in the upper canopy, which would result in no noticeable change of canopy cover between pre-fire and post-fire conditions. Another situation results from where all the upper tree crowns in the forest canopy have been either been consumed by a fire or the crown has been scorched by a more intense fire on the surface which extends into the tree crowns [6]. In these active crown fires, the canopy has been consumed or scorched by the fire, resulting in the death of all the trees provided that the trees are not resprouting deciduous species. This high severity active crown fire will result in a complete reduction of canopy cover as compared from pre-fire conditions. In between surface fires and crown fires, lie a situation where fire burning through surface vegetative material having enough intensity to burn or scorch the crowns of individual trees or clusters of

trees, resulting in a medium severity passive crown fire where some of the tree crowns are completely consumed or scorched [7]. While the post-fire canopy cover will be reduced from pre-fire canopy conditions as a result of these passive crown fires, the post-fire canopy cover will not be reduced to zero because some of the trees that made up the pre-fire canopy will survive, having been neither burned nor scorched.

Most land managers that use canopy cover mapping layers currently rely on the canopy cover layers that are available through the LANDFIRE project, a collaborative effort between the United States Forest Service and the US Department of the Interior. LANDFIRE produces geospatial products that describe wildland fire fuel conditions across the United States. From the LANDFIRE portal, land managers have access seamless nation-wide geospatial layers with a spatial resolution of 30 meters (30m) [8]. These canopy cover layers were generated from imagery acquired by the Enhanced Thematic Mapper Plus sensor on board the Landsat 7 satellite using a decision tree algorithm that was trained using plot data stored in the LANDFIRE Reference Database which was collected from across the US [9].

The utilization of small unmanned aircraft systems (sUAS) as a remote sensing tool provides land managers an affordable way to acquire imagery with much higher spatial resolution (sub-decimeter) than was previously available. This hyperspatial imagery produces homogeneous pixels which are comprised of a single class, resulting in higher mapping accuracy. By contrast, heterogeneous pixels which are large enough to spatially contain more than one class are common with lower resolution imagery [1], [4].

Burned vegetation, unburned canopy vegetation (conifer and deciduous trees), and unburned surface vegetation (brush and grass) have been shown to have distinct spectral

responses in both the visible and near infra-red spectra, though the spectral difference between canopy vegetation and surface vegetation is not as large as that between burned and general unburned vegetation [10]. This spectral difference is apparent when observing the two unburned vegetation classes in hyperspatial imagery. Persons who had participated in the acquisition of the sUAS imagery found that they could observe the visual difference between surface and canopy vegetation and that it matched what they had observed in the field while conducting image acquisition flights with the sUAS.

Hamilton [1] found that tree crowns obstruct surface features in imagery acquired with a sUAS, preventing classification of surface features. Obstructed features include surface biomass consumption, a primary indicator of surface fire severity. When a tree crown is surrounded by burned surface fuel in the image, it is reasonable to assume that the ground under the tree has also burned but is obscured by the unburned crown [1], [11]. Unburned tree crowns within a burned area cause an under-reporting of fire extent due to the misclassification of unburned tree crowns within the fire perimeter as being unburned areas. The first step to identifying and mapping obscured burned surface fuels is to identify trees within the image. When a tree surrounded by burned surface fuels is identified, it can be inferred that the surface fuels under the tree are in fact burned and should be included within the fire extent. Wildland fire can also affect an ecosystem through tree mortality. Mortality may result from either the ignition and consumption of a tree crown or from scorching of the tree crown caused by heat from surface fire beneath the tree. The extent of tree mortality can be estimated by calculating the change in canopy cover before and after the fire. Post-fire canopy cover can be calculated as the percentage

of an image that consists of tree crowns. Pre-fire canopy cover can be retrieved from the Canopy Cover rasters that were published by the LANDFIRE Fire Fuels project [8].

## **2.2. Materials and Methods: Mapping Canopy Cover and Tree Mortality**

Canopy cover describes the percentage area occupied by the upper tree crowns in forested ecosystems. In this effort, canopy cover is calculated as the percent of upper layer canopy cover within a stand, which is represented as a 30-meter pixel. Once the tree crown classification has been completed, canopy cover is determined by summarizing the percentage of hyperspatial (5cm) pixels that were classified as canopy vegetation within each 30-meter Landsat pixel. Finally, the post-fire canopy cover is compared against pre-fire canopy cover from the LANDFIRE project to determine how much upper layer canopy cover reduction occurred due to the fire, giving an indication of tree mortality, which is a primary indicator of post-fire effects in forested ecosystems.

### **2.2.1 Classification of Tree Crowns from Hyperspatial Imagery**

The initial step in mapping canopy cover and tree mortality enables the mapping of individual tree crowns within a hyperspatial orthomosaic acquired with the use of an sUAS. By classifying the tree crowns within an orthomosaic, we will be able to distinguish between landcover lifeforms that are indicative of canopy vegetation as opposed to surface vegetation.

The Mask Region-based Convolutional Neural Network (MR-CNN) is an algorithm for instance segmentation. It detects objects in an image and determines which pixels comprise each object. When classifying an image, MR-CNN begins by applying a Convolutional Neural Network (CNN) to extract features from the image. The features

extracted from the image are used as input for a Region Proposal Network, which slides over the feature map and detects Regions of Interest (RoIs) which are likely to contain objects. Each RoI is evaluated further to determine what type of object, if any, is inside (in this case, tree and non-tree); the object's bounding box; and what pixels inside the bounding box comprise the object [12].

Training data for the MR-CNN model consists of images in which the boundaries of all trees are marked with individual polygon. Unmarked areas are assumed to be non-tree. The model is trained using mini-batch gradient descent to minimize a loss function that measures the difference between the model's output in response to a training image and the training data.

A MR-CNN model with a ResNet-101 backbone that had been pre-trained on the Microsoft COCO dataset [13] was retrained to detect trees. Training images were cropped from orthomosaics of study areas flown with sUAS over the Boise National Forest. These forested sites primarily contained *Pinus Ponderosa* (Ponderosa Pine), *Populus Trichocarpa* (Black Cottonwood), *Pseudotsuga Menziesii* (Douglas Fir) and *Pinus Contorta* (Lodgepole Pine). A polygon was traced around each tree occurring in each training image using the VGG Image Annotator [14] as shown in Figure 1. A total of 1148 trees were labeled over 14 training images. Questions of whether visually ambiguous objects were trees or surface vegetation were answered by referring to a 3D model created from the same imagery as the orthomosaics. To discourage overfitting, training images were rotated and mirrored randomly during the training process.



*Figure 1 Mask Region-based Convolutional Neural Network training data*

As an instance segmentation algorithm, MR-CNN can answer questions that a pixel-based approach cannot, such as, “How many trees are in the image?” and “What is the distribution of tree sizes?” To facilitate calculation of canopy cover, the MR-CNN’s output was used to perform semantic segmentation. All pixels that were determined to be part of a tree were marked as tree crown and all other pixels were marked as surface.



*Figure 2 (a)*



*Figure 2 (b)*

(a) is pre-classified image. (b) is an image that has been classified by the MR-CNN

### **2.2.2 Calculating Canopy Cover**

Canopy cover describes the vertical projection of the tree canopy onto a horizontal plane in forested ecosystems. In this effort, canopy cover is calculated as the percent canopy cover within a stand. Once the tree crown classification has been completed as described earlier, canopy cover is determined by summarizing the percentage of 5cm pixels that were classified as canopy vegetation within each 30 meter (30m) Landsat pixel.

#### **2.2.2.1 Calculate Five Centimeter Class Densities for each 30 Meter Pixel**

The density of the tree class is calculated for each 30-meter Landsat pixel from the 5cm tree classification pixels obtained from the MR-CNN. A tool was written which calculates the density of 5cm pixels classified as tree crown within each 30m pixel, outputting a 30m resolution grayscale TIFF image. This Density Tool counts the total

number of hyperspatial tree class pixels occurring in the 30m pixel. The total number of tree pixels that occurred is then divided by the total number of 5cm pixels within the 30m pixel and multiplied by 100. The class densities for each 30m pixel are recorded in the class density output raster. This class density raster contains the percentage of each 30-meter pixel that is comprised of 5cm pixels that are classified as being part of a tree crown, which effectively records Canopy Cover for each 30 meter pixel.

#### **2.2.2.2 Canopy Cover Analysis**

Previous efforts including Hamilton [1], [4] and Matese [15] have shown that mapping with higher resolution imagery results in increased accuracy than using lower resolution imagery. This section summarizes the analysis performed to assess the effectiveness of the methods specified previously for calculating canopy cover.

**Setting and Notation.** Suppose that there is a 30m pixel depicting a region of ground. The same  $30\text{m} \times 30\text{m}$  region of ground is viewed as a  $600 \times 600$  region of 5cm pixels in the hyperspatial orthomosaic. Then the canopy cover is calculated from the hyperspatial orthomosaic, and this calculation is subject to error. The probability of correctly identifying a “tree” pixel (sensitivity) is observed from training data to be at least 0.70, and the probability of correctly identifying a “non-tree” pixel (specificity) is observed from training data to be at least 0.95. Both values are usually higher (see Table 2 in Section 3.2 for examples)., but we will use the values of 0.70 and 0.95 as conservative estimates.

**Question:** What level of accuracy can one expect for the value of canopy cover calculated from the  $600 \times 600$  pixel region of the hyperspatial orthomosaic corresponding

to a Landsat pixel? Precisely, what is the standard error of this value? Is this significantly smaller than what one might hope to obtain from LANDFIRE data?

**Answer:** Using The value of canopy cover can be estimated with bias close to 0 and standard error less than 0.001080. In most cases, the standard error is less than 0.000824. These values of standard error are significantly smaller than what one might hope to obtain from LANDFIRE data. They are obtained using mathematics beyond just routine statistical inference. The details are omitted here so as not to disturb the flow of the article. However, they are included in the Appendix for the interested reader.

### **2.2.3 Canopy Cover: sUAS Derived vs LANDFIRE**

The LANDFIRE project is a collaborative effort between the United States Forest Service and the US Department of the Interior, producing geospatial products that describe wildland fire fuel conditions supporting wildland fire modeling across the United States. LANDFIRE's Existing Vegetation Cover and Forest Canopy Cover products record the vertically projected percentage cover of the upper vegetation canopy layer [8]. For this effort in evaluating tree mortality, the Existing Vegetation Cover raster was updated to only contain canopy cover in forested ecosystems. Additionally, the Existing Vegetation Cover was updated (via the ArcGIS Spatial Analyst Con tool) with the Forest Canopy Cover for forested pixels where the predominant lifeform appeared to be either shrub or herbaceous. This conversion was accomplished by reclassing the vegetation cover for all non-forested lifeforms to zero, then substituting the Forest Canopy Cover for the Existing Vegetation Cover on pixels that had a non-forested life form, but also had a Forest Canopy Cover value. This resulted in a geospatial layer with

spatial resolution of 30 meters representing vertically projected percentage forest canopy cover.

The comparison of the hyperspatial derived canopy cover (HDCC) as accomplished using the ArcGIS combine raster tool. This resulted in an attribute table that had each combination of canopy cover that occurred within pixels in the two rasters as well as how many pixels the combination occurred in. From the combine table, the difference between the HDCC and LANDFIRE forest Canopy Cover (LFCC) can be calculated. This difference can be used to analyze the statistical difference between the sUAS derived and LANDFIRE products as well as the determination of canopy reduction as a measure of tree mortality in areas where the sUAS canopy cover was calculated from imagery acquired after a wildland fire.

### **2.2.3.1 Canopy Cover Comparison Methods**

When classifying sUAS imagery through the Mask Regional-Convolutional Neural Network (MR-CNN), a new image is created as an output image in which pixels are based on the original image is part of a tree crown (tree) or does not comprise part of a tree crown (no tree). During this process, however, spatial referencing and null data (where a pixel represents no data) values are lost and must be added back.

Spatial referencing is added back using an ArcPy script which takes the world file spatial referencing from the original images that were aligned with the 30m imagery, inputted into the MR-CNN and adds the spatial referencing to the MR-CNN output.

Band 4 of the classified output from the MR-CNN contains the tree/no tree classification. This band is extracted using the Copy Raster Tool on Band 4 of the image.

This newly extracted classification raster is given as input to the density tool, sums the number of 5cm tree pixels within the 30m pixel and divides it by the total number of 5cm pixels within a 30m pixel. Multiplying this number by one hundred gives the percent of tree pixels or the canopy cover in the 30m pixel area as expressed in Equation 6. This process is repeated until each pixel in the image has been analyzed giving a 30m resolution canopy cover (CC) output raster.

$$\% \text{ of } CC = \left( \frac{\text{tree pixels}}{\text{total pixels}} \right) * 100 \quad (6)$$

Before adding the null data back into the orthomosaic, the 5cm resolution imagery is resampled to 30m resolution. Resampling before adding the null data speeds up the process of adding the null data back in. Null data is recorded with a pixel value of 0 (not a tree) by the MR-CNN. To restore its value to NULL, ArcPy's conditional toolset is used to compare pixels in the source image (the MR-CNN output) and with the original data to check if the value was NULL. The density tools output is changed to 255 for every pixel where the source image's value is NULL. 255 is used to represent NULL because it is outside the range of possible values (0% to 100%).

The various methods of processing the data outlined in the previous sections creates a HDCC raster that is aligned with the LFCC raster. This allows the rasters to be compared pixel by pixel, using the ArcGIS Combine Raster tool. This creates a new raster with a raster attribute table column for each of the combined rasters. Next, the LFCC column was subtracted from the HDCC column into a newly added Difference column. With sUAS data of unburned areas, this shows how close HDCC and LFCC data are. When comparing pre-fire LFCC against post-fire HDCC, this difference can be used

in the calculation of canopy reduction, an indicator of tree mortality which is a primary effect of wildland fire.

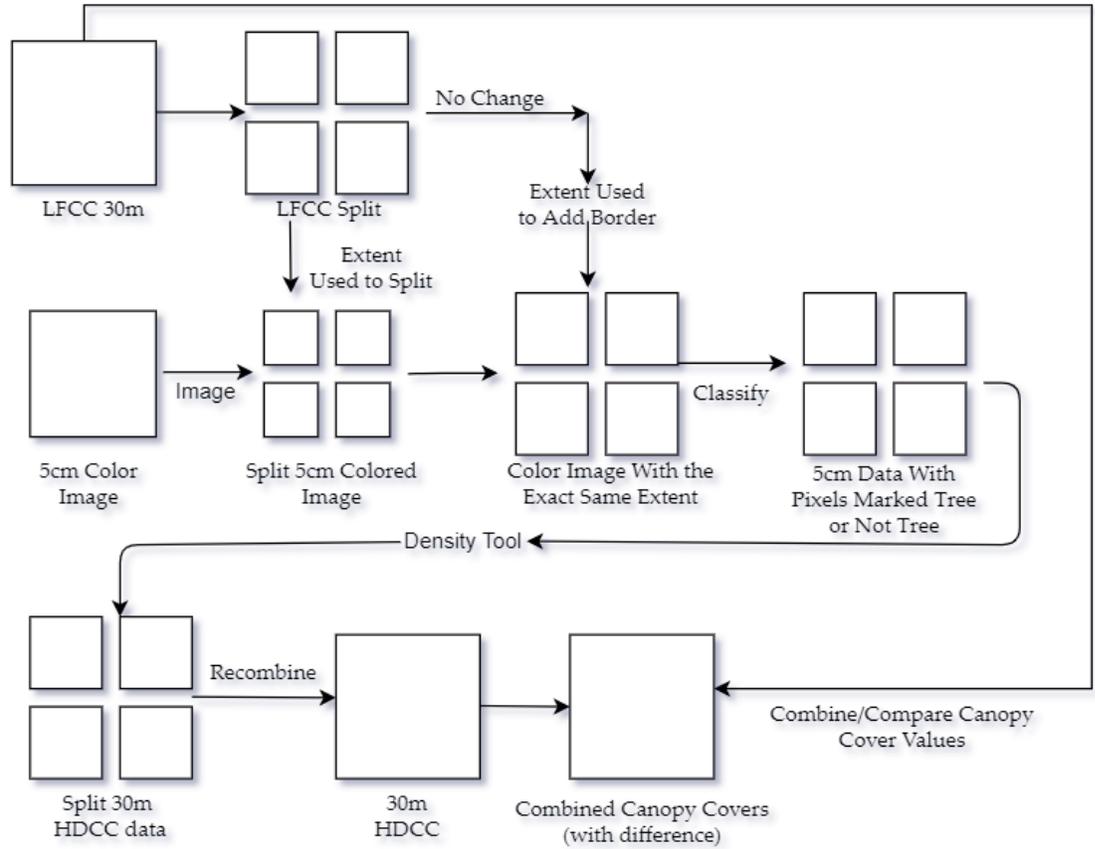


Figure 3: An overview of the processes used to create the HDCC

### 2.2.3.2 Canopy Cover Comparison Analysis

**working assumption** In section 2.2.2, it was observed that (assuming 5cm pixel identification better than 0.80 correct for pixels representing ground with canopy cover, better than 0.95 for pixels representing ground without canopy cover, and no more than 0.80 canopy cover) using adjusted hyperspatial derived data, canopy cover can be calculated with a standard error of less than 0.000824. This is a much smaller standard error than we could hope to obtain from 30m LANDFIRE pixels. Accordingly, it is

reasonable, when comparing LANDFIRE data to hyperspatial derive data, to make the working assumption that the adjusted HDCC value is correct.

**objective** The method of maximum-likelihood estimation is used to construct a model of the error of the value of canopy cover derived from LANDFIRE data.

**setting and notation** There are  $n$  LANDFIRE 30m pixels, and for each, the estimated value  $B \in [0, 1]$  of canopy cover. For each there is a corresponding region of 5cm pixels in the hyperspatial orthomosaic, and the adjusted value of canopy cover  $C \in [0, 1]$  calculated from the  $600^2$  pixels in corresponding region of the hyperspatial orthomosaic. As stated above, the latter value ( $C$ ) is presumed correct, and the objective is to specify a model of the error in the former value ( $B$ ). Write

$$X = B - C \quad (7)$$

for the error. The data consists of  $n$  pairs  $(b_1, c_1), (b_2, c_2), \dots, (b_n, c_n)$ , from which are obtained  $n$  error values

$$x_1 = b_1 - c_1, \dots, x_n = b_n - c_n. \quad (8)$$

**model for error** A normal model for  $X$  is used, whose density is

$$fX(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (9)$$

with parameters  $\mu$  (mean) and  $\sigma$  (standard deviation). (Note that  $\mu$  and  $\sigma$  are the respective mean and standard deviation of the error  $B - C$  only, and not of the canopy cover itself or any estimate thereof.) The values of the parameters are unknown, and the next task is to choose the best values for them (so that the model best accounts for the present data).

**maximum-likelihood estimators for parameters** The maximum-likelihood estimations for  $\mu$  and  $\sigma$  are

$$\hat{\mu} = \bar{X} = \frac{\sum X}{n} \quad (10)$$

and

$$\hat{\sigma} = \hat{S} = \sqrt{\frac{\sum(X-\bar{X})^2}{n}} \quad (11)$$

respectively, the latter being the “population standard deviation”  $\hat{S}$  with denominator  $n$  instead of the “sample standard deviation”  $\bar{S}$  with denominator  $n - 1$ .

**conclusion** The maximum-likelihood estimation for a normal error  $X$  of the LANDFIRE value for canopy cover has mean  $\bar{X}$  and standard deviation  $\hat{S}$ , calculated from the  $n$  values  $x_1 = b_1 - c_1, \dots, x_n = b_n - c_n$  of the difference  $X = B - C$  between LANDFIRE Canopy Cover and (adjusted) Hyperspatial Derived Canopy Cover.

#### 2.2.4 Mapping Tree Mortality via Canopy Reduction

While it is much more accurate to map canopy cover from hyperspatial sUAS imagery, the mapping extent is limited to how large of a study area can be flown with a sUAS. Current regulatory and technical constraints limit flight operations with a single sUAS in forested environments to under 400 hectares in a single day [16]. For mapping post-fire canopy cover, this constraint can be overcome with more flights on all but the largest fires. Due to the unpredictable nature of fires resulting from unplanned ignitions, it is not feasible to obtain hyperspatial sUAS derived canopy cover data reflecting pre-fire conditions. Fortunately, the LANDFIRE products are generated for the entire United States and are updated every two years.

For mapping canopy reduction, which is a proxy for tree mortality, a comparison was made between the LFCC (pre-fire canopy cover) against the HDCC (post-fire) layer, as described in the previous section.

Once the difference between pre-fire LFCC and post-fire HDCC, canopy cover is calculated, canopy cover reduction (CCR) can be calculated by calculating the difference between LFCC and HDCC, minus the estimated error standard deviation (multiplied by the Standard Normal density) and estimated error mean as shown in Eq 12.

#### **2.2.4.1 Canopy Reduction Methods**

Due to LANDFIRE's canopy cover being less accurate than the HDCC, it cannot simply be stated that reduction is the discrepancy between sUAS Derived data and LANDFIRE data. Instead, we must use the data from unburned areas to see how much this discrepancy varies. Then any numbers outside of the normal variance is a change in canopy cover to a 95% degree of certainty. Once it is determined that there has been a change there are two possibilities for what that change could be, active crown fire or passive crown fire. If the HDCC is 0% it was an active crown fire in this area otherwise it was a passive crown fire.

The main difficulty with calculating the type of fire based on this method is that there is some uncertainty due to LANDFIRE being 30-meter data. For this reason, it is difficult to tell what happened to pixels where there was not enough of a difference to be certain if there was a change. The discrepancy between the data could just be due to LANDFIRE's overreporting or it could be due to a passive crown fire.

#### 2.2.4.2 Canopy Reduction Analysis

**working assumption** As before, a 30m-by-30m square of land is being considered, which corresponds to a single LANDFIRE pixel as well as to a 600 x 600 pixel region in a hyperspatial orthomosaic. Again,  $B \in [0, 1]$  is the canopy cover according to LANDFIRE, and  $C \in [0, 1]$  is the adjusted value of canopy cover obtained from the hyperspatial orthomosaic, which is assumed to be correct. In the previous section, maximum-likelihood estimation was used to construct a model of the error  $X = B - C$ , which was a normal density with estimated parameter values  $\mu = \hat{\mu} = \bar{X}$  and  $\sigma = \hat{\sigma} = \hat{S}$ .

**objective** A criterion is developed by which it can be concluded, with the desired level of confidence  $1 - \alpha$ , that canopy cover, from past LANDFIRE image to present hyperspatial orthomosaic, has decreased by at least some particular value  $r$ .

**the difficulty** The difficulty here is that there is no available pre-fire value of  $C$  (canopy cover calculated from a hyperspatial orthomosaic), but only a past value of  $B$  (canopy cover calculated from LANDFIRE) to compare with the present value of  $C$ . The notation  $X = B - C$ , is still used, while now it is recognized that  $B$  is a past value and  $C$  is a present value. This means that  $X$  (difference in LFCC and (adjusted) HDCC values) may include change in canopy cover (where a positive value might result from a crown fire), or inaccuracy of LANDFIRE data, or both.

**the null hypothesis** The null hypothesis  $H_0$  that we wish to test is that there was a reduction in canopy cover of no more than  $r$ . This  $H_0$  will be rejected if, even allowing for a canopy reduction of  $r$ , the value of  $B - C$  is too high to be explained by chance; that is, if

$$B - C > r + \mu + z_{\alpha}\sigma, \quad (12)$$

or, equivalently,

$$X > r + \mu + z_{\alpha}\sigma.$$

Replacing the inequality with an equality and solving for  $r$ , the following criterion is obtained.

**criterion** It may be inferred that canopy cover has decreased by at least

$$B - C - \mu - z_{\alpha}\sigma, \quad (13)$$

where  $B$  is past canopy cover according to LANDFIRE,  $C$  is the adjusted value of canopy cover from a hyperspatial orthomosaic,  $\mu$  and  $\sigma$  are the parameter values obtained as described previously, and  $z_{\alpha}$  is the critical value of the Standard Normal density at significance level  $\alpha$ . Intuitively, the amount that we can conclude by which canopy cover has been reduced is the difference “past LANDFIRE minus present hyperspatial”  $B - C$ , minus the average discrepancy  $\mu$  between the LANDFIRE  $B$  and hyperspatial-derived  $C$ , minus an additional margin of error  $z_{\alpha}\sigma$  which is larger, the higher the level of confidence we wish to have.

## 2.3. Results

### 2.3.1 Tree Crown Classification Validation

To assess the performance of the MR-CNN in mapping tree crowns, validation data was prepared using ArcGIS. Sample areas of canopy and surface vegetation were recorded as polygons in shapefiles digitized from existing orthomosaics. Validation data was made for all orthomosaics used in the research.

The MR-CNN was run on the orthomosaics used for validation. Agreement between a classifier's output and the validation data was calculated using ArcMap's Tabulate Area tool as shown in

. A classifier's accuracy on an orthomosaic was computed as the area where the classifier agreed with the class marked in the validation data divided by the total area marked in the validation data.

Orthomosaics were chosen with a variety of different conditions to find any anomalies where the classifier underperforms. Trees in the chosen orthomosaics were comprised primarily of *Pinus Ponderosa* (Ponderosa Pine), *Populus Trichocarpa* (Black Cottonwood), *Pseudotsuga Menziesii* (Douglas Fir) and *Pinus Contorta* (Lodgepole Pine).

The distinguishing characteristics of the ten chosen orthomosaics were as follows:

- NewStart: Surface vegetation comprised of brush and herbaceous vegetation in the riparian zone along Grimes Creek. Low canopy cover.
- GrimesCreek: Surface vegetation comprised of brush and herbaceous vegetation both inside and outside the riparian zone along Grimes Creek. Moderate to low canopy cover.
- South Placerville: Surface vegetation comprised of brush and herbaceous vegetation both inside and outside the riparian zone. Low canopy cover.
- East Placerville: Mix of herbaceous vegetation and canopy cover. Moderate to low canopy cover.
- West Placerville: Wide open area covered in trees with surface vegetation along the ground around trees. Moderate to high canopy cover
- Northwest Placerville: Flat area with trees and tree-like brush. Moderate canopy cover.

- Edna Creek: Lots of unforested areas or areas with dead trees. Patches of forest throughout the orthomosaic. Low canopy cover.
- Belshazzar - Very dense surface vegetation. Not in a riparian zone. Certain areas experienced larger-than-normal orthomosaic stitching artifacts, which we hypothesized could cause trouble for object detection techniques. Predominantly high canopy cover with some meadows.
- North Experimental Forest: Herbaceous vegetation around water sources. Sparse forest. Moderate to low canopy cover.
- South Experimental Forest: Not much other vegetation than trees. Ground is white in many areas. Moderate canopy cover.

Table 1: Confusion matrices of MR-CNN tree crown classifications. Classified surface and canopy are shown in the rows while human labeled surface and canopy are in columns.

	South Placerville		NW Placerville		Edna Creek		South Exp. Forest	
	Surface	Canopy	Surface	Canopy	Surface	Canopy	Surface	Canopy
Surface	47.5%	19.9%	46.6%	6.2%	49.0%	0.6%	49.1%	10.6%
Canopy	0.6%	32.0%	2.3%	44.9%	0.0%	50.3%	0.4%	40.0%
	North Exp. Forest		East Placerville		West Placerville			
	Surface	Canopy	Surface	Canopy	Surface	Canopy		
Surface	50.0%	10.5%	49.29%	4.50%	49.98%	7.29%		
Canopy	0.3%	39.2%	0.00%	46.21%	0.00%	42.73%		
	Belshazzar		Newstart		Grimes Creek			
	Surface	Canopy	Surface	Canopy	Surface	Canopy		
Surface	47.1%	3.6%	50.1%	16.9%	51.5%	2.9%		
Canopy	3.0%	46.3%	0.2%	32.8%	0.2%	45.4%		

The MR-CNN model’s overall accuracy was 90%. The specificity (correctly identified tree pixels/total number of tree pixels) was 84% and the sensitivity (correctly identified non-tree pixels/total number of non-tree pixels) was 99%. The MR-CNN was found to correctly identify a high number of tree pixels while only missing a low number of tree pixels as shown in **Error! Reference source not found.**

Table 2: MR-CNN tree crown classification statistical summary

	<b>Accuracy</b>	<b>Specificity</b>	<b>Sensitivity</b>
<b>South Placerville</b>	79.5%	98.2%	70.5%
<b>NW Placerville</b>	91.5%	95.1%	88.3%
<b>Edna Creek</b>	99.3%	100.0%	98.8%
<b>South Exp. Forest</b>	89.1%	99.0%	82.2%
<b>North Exp. Forest</b>	89.2%	99.2%	82.6%
<b>East Placerville</b>	95.5%	100.0%	91.6%
<b>West Placerville</b>	92.7%	100.0%	87.3%
<b>Belshazzar</b>	93.4%	93.9%	92.9%
<b>Newstart</b>	82.9%	99.4%	74.8%
<b>Grimes Creek</b>	96.9%	99.6%	94.7%

### 2.3.2 Canopy Cover Analysis

Running an orthomosaic through the Snapper Tool, MR-CNN, and Density Tool gave a 30m orthomosaic with canopy cover that also is coregistered with LANDFIRE data. Then the HDCC numbers were adjusted based on the bias of the MR-CNN. The values used to calculate the bias were the accuracy of correctly identifying tree pixels (sensitivity) and the odds of a false positive (specificity). These numbers ended up being 84% and 99% respectively. Figure 4 shows the HDCC layer for South Placerville.



*Figure 4: Values are in percent canopy cover*

### **2.3.3 Canopy Cover Comparison: Hyperspatial Derived vs LANDFIRE**

Ten orthomosaics where there was no fire were compared using the methods stated above. LANDFIRE was over reporting by an average of 2.5% over the HDCC layer. This agrees with what Scott [17] stated, that the “canopy cover values [for LANDFIRE] are too high.” The standard deviation for the difference between each pixel was 14. This number is a little high and will hopefully be driven down by future improvements. Figure 5 shows a piece of the results for Edna Creek. White represents a higher canopy cover for the HDCC layer while white represents a higher canopy cover reported by LANDFIRE. The grey composing most of the figure is for when the difference is equal or close to zero.

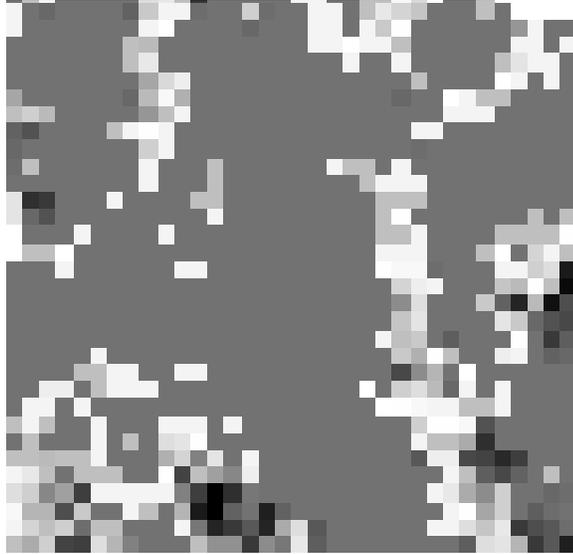


Figure 5: A raster representing the comparison between the LANDFIRE and HDCC data. White indicates LANDFIRE reported more canopy cover. Black indicates the inverse

Table 3: Mean difference and the standard deviation of that difference for each area

	Mean	Standard Deviation
<b>Belshazzar</b>	17%	14
<b>New Start</b>	4%	13
<b>Grimes Creek</b>	6%	15
<b>NW Placerville</b>	1%	11
<b>0530 Placerville</b>	10%	9
<b>Edna Creek</b>	4%	1
<b>South Experimental Forest</b>	3%	12
<b>North Experimental Forest</b>	7%	12
<b>East Placerville</b>	7%	12
<b>West Placerville</b>	2%	13

**Error! Reference source not found.** shows that the LFCC over reported for e very orthomosaic.

### **2.3.4 Tree Mortality Mapping Assessment**

The standard deviation and mean collected from the ten non-burnt orthomosaics show that there needs to be at least a change of 27% to know there has been a change with 95% confidence. If the difference is high enough to say there has been a change next the HDCC layer can be looked at to see whether it was a passive crown fire or an active crown fire. Pixels that did not contain adequate canopy reduction (more than 26%) from comparing pre-fire LFCC to post-fire HDCC are considered to be inconclusive. This would include unburned areas, surface fire, passive crown fire, and active crown fire. If an area had less than 27% canopy cover then we cannot say whether any canopy reduction occurred. This method was used on orthomosaics acquired over portions of the Mesa fire (a 16,000 hectare mixed severity fire) to calculate burn intensity. This fire worked well as there were areas of active crown fire, passive crown fire, and surface fire. Figure 6 shows the Southwest area of the Mesa fire which had a lot of Active Crown fire. While Figure 7 the east section which had more passive crown fire and possible surface fire.



**(a)**

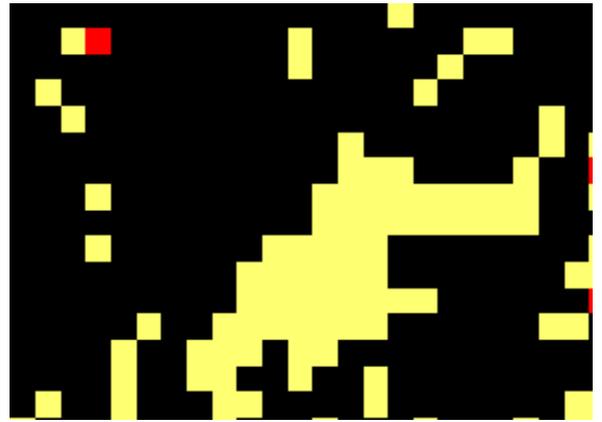


**(b)**

*Figure 6: (a) Hyperspatial data after a burn. (b) Types of fire: Red=Active, Yellow=Passive, Black=inconclusive crown fire activity*



**(a)**



**(b)**

*Figure 7 (a) Hyperspatial data after a burn. (b) Types of fire: Red=Active, Yellow=Passive, Black=inconclusive crown fire activity*

## 2.4. Discussion

Post-fire mapping products are commonly used by land managers as they are developing post-fire management plans which specify strategies for mitigating the effects of the fire. In order for this data to be most useful, it needs to be readily available to managers. The advantage of using imagery acquired by sUAS is that it can be available as quickly as managers can get out to the site and conduct image acquisition flights. The timing of the flights can be tailored to optimize meteorological conditions that are conducive to acquiring image acquisition by enabling managers to acquire imagery as soon as the scene is not obscured by smoke and there is low enough cloud cover to ensure adequate illumination of the scene.

The MR-CNN was found to be very effective in identifying tree crowns. The classifier was found to have virtually no false positives with a specificity of 99%. The false negatives were still minimal, with a sensitivity of 84%.

Scott [17] showed that LANDFIRE Canopy Cover typically is over reported. As expected, when comparisons were made of study areas that had not experienced wildland fire, the HDCC was on average 2.5 percentage points lower than the LANDFIRE Canopy Cover. While the mean difference between the HDCC and LFCC was only a few percentage points, the higher standard deviation which typically ranged a bit more than 10 percentage points indicates that the canopy cover between the HDCC and LFCC products differed more significantly than indicated solely by the means.

The canopy reduction mapping over a burned area was very promising. When ran on the Mesa fire most pixels correctly identified surface fire as well as both active and

passive crown fire. These methods work best on areas that had higher pre-fire canopy cover and canopy reduction. There needs to be at least a 27% difference for there to be a known change. Various theories to bring this number down are brought up in the future work section (5.1).

In addition to higher accuracy mapping of canopy cover as well as tree mortality as measured by canopy reduction, the use of sUAS allows for much more current data acquisition than is possible with the LANDFIRE products. Using imagery acquired with a sUAS, it would be possible to have the HDCC product available within a day of flying a study area. By contrast, the current LANDFIRE 2.0.0 Remap effort is using Landsat imagery that was acquired in 2016. The initial product releases which covered the majority of the Contiguous US did not occur until 2019, meaning even the initial product releases were derived from imagery that was already three years out of date [8]. While canopy cover can be mapped with much higher accuracy using hyperspatial imagery acquired with a sUAS, this research team found that current technical and regulatory constraints on drone usage realistically will only allow for the acquisition of up to 600 hectares per day [16]. To effectively map large class F (> 400 ha) and G (> 5,000 ha) fires [18], larger and more efficiently acquirable imagery satellite system still needs to be utilized to obtain a large enough analysis extent to map the whole fire. As the spatial resolution of satellite imagery increases with some of the new high resolution sensors that are being put into orbit, it may be possible to leverage this methodology, applying it to new generations of imagery as they approach hyperspatial resolution.

## **2.5. Conclusions**

Detection and mapping of trees from hyperspatial drone imagery enables more accurate mapping of canopy cover, a measure of forest density which is used to determine ecosystem departure from historic fire regimes, a primary measure of ecosystem resilience [19]. A comparison of a pre-fire LANDFIRE Canopy Cover product against post-fire HDCC resulted in a very effective measure of canopy fire reduction, which is a measure of tree mortality resulting as an effect of the wildland fire.

### **2.5.1 Future Work**

While this effort was highly effective in establishing a timely, effective and cost-efficient means of mapping canopy cover and tree mortality, there are many areas of future research which could result in an improvement in the effectiveness of this methodology.

An improvement in wildland fire extent mapping can be achieved by identifying unburned trees that are surrounded by surface vegetation. Current classification methods developed through this effort [1], [3] which map burn extent are not capable of detecting that a surface fire burned underneath an unburned tree crown due to the view of the surface vegetation being obscured from the sensor on the sUAS by the tree crown. While it may not be possible to determine whether the biomass consumption under the tree crown is high or low due to the presence of white ash, it would be possible to identify that the surface vegetation was burned and update the burn extent and aerial mapping accordingly based on assumptions made based on whether surface vegetation visible next to the crown was burned.

While the specificity of the MR-CNN was quite high at 99%, the accuracy of the classifier could be improved by refining the training data for the MR-CNN, based on observations of false negatives during validation. This refinement would likely improve tree crown classification, reducing the number of false negatives, thereby increasing the classifier sensitivity, which was 84% in the validation dataset used for this effort.

An improvement in tree crown classification accuracy might be able to be achieved by using pixels' heights as additional input to the classifiers. Per-pixel height values were available from a digital surface model produced during the orthomosaic creation process but were not used as an input to the MR-CNN in this effort.

Due to the need to split each orthomosaic, the MR-CNN classifier can miss trees that are on the border of these splits since they look like half trees. This problem could be eliminated by modifying the snapping and density tool to run larger files than what they are currently limited to due to OpenCV's file size limits. This would allow entire orthomosaics to be processed at once.

A more major problem that arises with splitting up the image is when 30m pixels overlap. This happens because the LANDFIRE and hyperspatial image cannot line up perfectly. Because of this the 5cm data has an extra pixel on the bottom and right when it is clipped. The density tool creates a whole 30m pixel from these few extra 5cm pixels on the edges. When the 30m rasters are recombined the data at the overlap is corrupted. This problem can be fixed either creating a more direct method to handle the extra pixel or by getting rid of the split. Getting rid of the split would require OpenCV to be able to handle larger orthomosaics. This problem is most likely a big factor in the high standard

deviation. If fixed the difference needed to be confident there has been a change should drop.

**Supplementary Materials:** The following are available online from the authors at <https://firemap.nnu.edu/tree-mortality>: source code and data.

**Author Contributions:** Conceptualization, D. Hamilton; Methodology, D. Hamilton, K. Brothers, S. Jones, J. Winters; software, D. Hamilton, K. Brothers, S. Jones, J. Winters; validation, K. Brothers.; formal analysis, J. Colwell.; investigation, D. Hamilton, K. Brothers, S. Jones; data curation, D. Hamilton.; writing, D. Hamilton, J. Colwell, K. Brothers, S. Jones, J. Winters; visualization, K. Brothers.; supervision, D. Hamilton.; project administration, D. Hamilton.; funding acquisition, D. Hamilton. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the USDA Forest Service Boise National Forest, Forest Service Agreement No, 18-CS-11040200-0025.

**Acknowledgments:** We would like to acknowledge the students in the Northwest Nazarene University Department of Math and Computer Science who have helped with different aspects of this effort, including current and former students Nicholas Hamilton, Jonathan Hamilton, Braelyn Boerner, Gabriel Murphy, Enoch Levandovsky, Gabriel Johnson, Austin White, Aleesha Chavez and Andrew Welk. Additionally, we would like to acknowledge Northwest Nazarene University who funded the research efforts of many of the students mentioned.

### **Appendix: Accuracy of Calculated Value for Canopy Cover**

In this appendix, the accuracy of the value for canopy cover calculated from the hyperspatial orthomosaic is shown to be better than that calculated from LANDFIRE. This conclusion was stated in Subsection 2.2.2, but depends on mathematics beyond just routine statistical inference. The details are given here to justify the conclusion.

Setting and notation. Suppose that there is a 30m pixel depicting a region of ground with canopy cover value  $c \in [0, 1]$ . (The actual value  $c$  would be unknown in practice.) The same  $30\text{m} \times 30\text{m}$  region of ground is viewed as an  $n \times n$  region of smaller pixels in the hyperspatial orthomosaic. (In practice,  $n$  would be equal to 600, and the  $600 \times 600$  orthomosaic would consist of 5cm pixels, but the symbol  $n$  will be used for now for greater clarity.) The number of true “tree” pixels is  $n^2 c$ , and the number of “non-tree” pixels is  $n^2(1 - c)$ . Each of the  $n^2$  pixels is assigned by the MR-CNN a binary value  $A = 1$  (tree) or  $A = 0$  (non-tree). Then the canopy cover is calculated from the hyperspatial orthomosaic as

$$\bar{A} = \frac{\sum A}{n^2}, \quad (1)$$

where the sum is taken over all  $n^2$  pixels.

The assignments are subject to error. Suppose the probability of correctly assigning  $A = 1$  to a “tree” pixel (sensitivity) is known to be  $p \in [0, 1]$ , and the probability of correctly assigning  $A = 0$  to a “non-tree” pixel (specificity) is known to be  $q \in [0, 1]$ .

Question. What level of accuracy can one expect for the value of canopy cover calculated from the  $n \times n$  pixel region of the hyperspatial orthomosaic corresponding to a

Landsat pixel? Precisely, what is the standard error of this value? Is this significantly smaller than what one might hope to obtain from LANDFIRE data?

Mean and standard deviation of calculated value. The expected value (mean) of the calculated value  $\bar{A}$  is

$$\begin{aligned} E[\bar{A}] &= cp + (1 - c)(1 - q) \\ &= 1 - q + c(p + q - 1). \end{aligned} \quad (2)$$

The standard deviation is

$$\begin{aligned} &\sqrt{\frac{n^2 cp(1-p) + n^2(1-c)q(1-q)}{(n^2)^2}} \\ &= \frac{\sqrt{cp(1-p) + (1-c)q(1-q)}}{n} \end{aligned} \quad (3)$$

Adjustment of value for canopy cover. Recall that  $c$  is the proportion of small pixels representing ground that has canopy cover. It is the number wanted but not known.

One solves for  $c$  in the expression for  $E[\bar{A}]$  above to obtain

$$c = \frac{E[\bar{A}] + q - 1}{p + q - 1} \quad (4)$$

And thus

$$C = \frac{\bar{A} + q - 1}{p + q - 1}$$

is an unbiased estimator of  $c$ . This will be called the “adjusted canopy cover”.

And the standard error of  $C$  is just the standard deviation of  $\bar{A}$  divided by  $p + q - 1$ ,

$$\text{which is } \frac{\sqrt{cp(1-p) + (1-c)q(1-q)}}{n(p+q-1)}.$$

Next, this expression will be examined more closely to obtain an upper bound.

Derivation of upper bound. The quantity  $x(1-x)$  is decreasing for on the interval  $[0.5, 1]$ , so that since  $p > 0.70$  and  $q > 0.95$ ,

$$p(1-p) < (0.70)(1-0.70) \text{ and } q(1-q) < (0.95)(1-0.95).$$

It also follows that

$$(0.70)(1-0.70) > (0.95)(1-0.95),$$

whereupon

$$c(0.70)(1-0.70) + (1-c)(0.95)(1-0.95)$$

(which is a weighted mean of  $(0.70)(1-0.70)$  and  $(0.95)(1-0.95)$ ) is an increasing function of  $c$ . Combining this fact with the first two inequalities, one obtains

$$\begin{aligned} & cp(1-p) + (1-c)q(1-q) \\ & < c(0.70)(1-0.70) + (1-c)(0.95)(1-0.95) \\ & < (0.80)(0.70)(1-0.70) + (1-0.80)(0.95)(1-0.95) \end{aligned}$$

for

$$p > 0.70, q > 0.95, \text{ and } c < 0.80.$$

The actual canopy cover ( $c$ ) in the training data used was less than 0.80, the sensitivity ( $p$ ) was better than 0.70, and the specificity ( $q$ ) better than 0.95. Using also  $p + q - 1 > 0.70 + 0.95 - 1$ , one obtains

$$\frac{\sqrt{cp(1-p)+(1-c)q(1-q)}}{n(p+q-1)} < \frac{\sqrt{(0.80)(0.70)(1-0.70)+(1-0.80)0.95(1-0.95)}}{600(0.70+0.95-1)} \approx 0.001080,$$

which compares favorably with results obtained with LANDFIRE 30m pixels. If one assumes that  $p > 0.80$  (which is true in most cases) rather than merely  $p > 0.70$ , the upper bound on the standard error of  $C$  becomes

$$\frac{\sqrt{(0.80)(0.80)(1 - 0.80) + (1 - 0.80)0.95(1 - 0.95)}}{600(0.80 + 0.95 - 1)} \approx 0.000824,$$

which would be impossibly low to achieve with LANDFIRE 30m pixels.

Conclusion. The adjusted value of canopy cover calculated from the hyperspatial orthomosaic has lower standard error than we could hope to obtain from LANDFIRE 30m pixels.

### **3. Finale**

This project found successful methods to compare pre-fire LANDFIRE and post-fire sUAS data to calculate a measure for tree mortality. However, there is still room for improvement. As described in the publication, LANDFIRE calculated canopy cover's margin for error is not as precise as the calculated canopy cover derived from sUAS data, which leads to a less accurate analysis of tree mortality. One solution to this problem is to use a pre-fire data set with a higher resolution such as Planet Labs SkySat, which provides 50cm imagery of almost any location on Earth. If SkySat imagery was available, the Mask R-CNN could be used to derive a pre-fire canopy cover data set with a much higher degree of accuracy than that of LANDFIRE's.

Another problem arises from splitting the image for classification. Rejoining the orthomosaic, for comparison with the LANDFIRE data, can create a one-pixel border on each edge of the split. The images are split because the computer that is used for classification has a limited amount of VRAM (video random access memory), which runs

out quickly when the orthomosaics are not trimmed down. There are a couple of solutions to this problem. The first is to upgrade the lab computer that runs the Mask R-CNN to a video card(s) with more VRAM. This is, however, a costly solution. A second solution would be to utilize a cloud computing service such as Amazon Web Services, Google Cloud, or Microsoft's Azure which all offer machine learning services.

Working on this project was a great learning experience. while studying geography at College of Western Idaho (CWI), I learned the skills needed for performing analysis on spatial data. This project allowed me to apply those spatial analysis and geographic information systems skills in a dynamic environment while also allowing me to apply what I learned at NNU in the machine learning and data structures courses.

Writing and publishing a peer reviewed paper is an opportunity that not many undergrads get, which I am grateful to have had the ability to be a part of the writing and publishing process.

## References

- [1] D. Hamilton, "Improving Mapping Accuracy of Wildland Fire Effects from Hyperspatial Imagery Using Machine Learning," PhD Dissertation, University of Idaho, Moscow, ID, 2018.
- [2] J. E. Keeley, "Fire intensity, fire severity and burn severity: a brief review and suggested usage," *Int.J.Wildland Fire*, vol. 18, no. 1, pp. 116–126, 2009.
- [3] D. Hamilton, R. Pacheco, B. Myers, and B. Peltzer, "kNN vs. SVM: a Comparison of Algorithms," in *Proceedings of the Fire Continuum - Preparing for the future of wildland fire. 2018 May 21-24*, Missoula, MT, 2020, p. 358, [Online]. Available: <https://www.fs.usda.gov/treearch/pubs/60581>.
- [4] D. Hamilton, N. Hamilton, and B. Myers, "Evaluation of Image Spatial Resolution for Machine Learning Mapping of Wildland Fire Effects," in *Proceedings of SAI Intelligent Systems Conference*, 2019, pp. 400–415.
- [5] D. Hamilton, B. Myers, and J. Branham, "Evaluation of Texture as an Input of Spatial Context for Machine Learning Mapping of Wildland Fire Effects," *Signal and Image Processing: An International Journal*, vol. 8, no. 5, 2017.
- [6] R. C. Rothermel, *Predicting Behavior and Size of Crown Fires in the Northern Rocky Mountains*. U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, 1991.
- [7] M. G. Cruz, "Development and testing of models for predicting crown fire rate of spread in conifer forest stands.," *Canadian Journal of Forest Research*, vol. 35, no. 7, pp. 1626–1639, 2005.
- [8] LANDFIRE, "Landfire Program," 2020. <https://landfire.gov/> (accessed Nov. 25, 2020).
- [9] M. C. Reeves, K. C. Ryan, M. G. Rollins, and T. G. Thompson, "Spatial fuel data products of the LANDFIRE Project," *Int. J. Wildland Fire*, vol. 18, no. 3, p. 250, 2009, doi: 10.1071/WF08086.
- [10] D. Hamilton, M. Bowerman, J. Collwel, G. Donahoe, and B. Myers, "A Spectroscopic Analysis for Mapping Wildland Fire Effects from Remotely Sensed Imagery," *Journal of Unmanned Vehicle Systems*, no. Journal Article, 2017, doi: 10.1139/juvs-2016-0019.
- [11] J. H. Scott and E. D. Reinhardt, "Assessing crown fire potential by linking models of surface and crown fire behavior," *USDA Forest Service Research Paper*, no. Journal Article, p. 1, 2001.
- [12] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," Oct. 2017.
- [13] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," in *European conference on computer vision*, 2014, pp. 740–755.
- [14] A. Dutta and A. Zisserman, "The VIA Annotation Software for Images, Audio and Video," New York, NY, USA, 2019, doi: 10.1145/3343031.3350535.
- [15] A. Matese *et al.*, "Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture," *Remote Sensing*, vol. 7, no. 3, pp. 2971–2990, 2015.
- [16] J. Goodwin and D. Hamilton, "Archaeological Imagery Acquisition and Mapping Analytics Development," Boise National Forest, Boise, ID, 2019.
- [17] J. Scott, "Review and assessment of LANDFIRE canopy fuel mapping procedures," LANDFIRE, LANDFIRE Bulletin, Apr. 2008. Accessed: Jul. 30, 2020. [Online]. Available: [https://landfire.cr.usgs.gov/documents/LANDFIRE\\_Canopyfuels\\_and\\_Seamlines\\_ReviewScott.pdf](https://landfire.cr.usgs.gov/documents/LANDFIRE_Canopyfuels_and_Seamlines_ReviewScott.pdf).
- [18] National Wildfire Coordinating Group (NWCG), "Size Class of Fire," 2020. /term/glossary/size-class-of-fire (accessed May 20, 2020).
- [19] S. Barrett *et al.*, *Interagency fire regime condition class guidebook. Version 3.0*. USDA Forest Service, US Department of Interior, and The Nature Conservancy, 2010.
- [20] Barrett, S.; Havlina, D.; Jones, J.; Hann, W.; Frame, C.; Hamilton, D.; Schon, K.; Demeo, T.; Hutter, L.; Menakis, J.; et al. Interagency Fire Regime Condition Class Guidebook; Version

3.0; USDA Forest Service, US Department of Interior, and The Nature Conservancy: Rolla, MO, USA, 2010

- [21] D. A. Hamilton, K. L. Brothers, S. D. Jones, J. Colwell, and J. Winters, "Wildland Fire Tree Mortality Mapping from Hyperspatial Imagery Using Machine Learning," *Remote Sensing*, vol. 13, no. 2, p. 290, Jan. 2021
- [22] N. A. Hamilton, "Classifying Wildland Fire Severity on Landsat Imagery Using Machine Learning Trained by Hyperspatial Imagery," 2019.