

Wildfire Mapping in Interior Alaska Using Deep Neural Networks on Imbalanced Datasets

Zachary L. Langford
Bredesen Center
University of Tennessee, Knoxville
Knoxville, TN, USA
zlangfor@vols.utk.edu

Jitendra Kumar
Environmental Sciences Division
Oak Ridge National Laboratory
Oak Ridge, TN, USA
jkumar@climatemodeling.org

Forrest M. Hoffman
Computational Sciences & Engineering Division
Oak Ridge National Laboratory
Oak Ridge, TN, USA
forrest@climatemodeling.org

Abstract—Wildfires are the dominant disturbance impacting many regions in Alaska and are expected to intensify due to climate change. Accurate tracking and quantification of wildfires are important for climate modeling and ecological studies in this region. Remote sensing platforms (e.g., MODIS, Landsat) are valuable tools for mapping wildfire events (burned or burning areas) in Alaska. Deep neural networks (DNN) have exhibited superior performance in many classification problems, such as high-dimensional remote sensing data. Detection of wildfires is an imbalanced classification problem where one class contains a much smaller or larger sample size, and performance of DNNs can decline. We take a known weight-selection strategy during DNN training and apply those weights to MODIS variables (e.g., NDVI, surface reflectance) for binary classification (i.e., wildfire or no-wildfire) across Alaska during the 2004 wildfire year, when Alaska experienced a record number of large wildfires. The method splits the input training data into subsets, one for training the DNN to update weights and the other for performance validation to select the weights based on the best validation-loss score. This approach was applied to two sampled datasets, such as where the no-wildfire class can significantly outweigh the wildfire class. The normal DNN training strategy was unable to map wildfires for the highly imbalanced dataset; however, the weight-selection strategy was able to map wildfires very accurately (0.96 recall score for 78,702 wildfire pixels (500×500 m)).

Index Terms—Deep Learning, MODIS, Wildfire, Imbalanced Classification

I. INTRODUCTION

Arctic tundra, boreal forests, and peatlands are already undergoing major changes, which have led to increased frequency and intensity of disturbances (e.g., wildfires) [1], [2]. Predictive ecosystem modeling capabilities for this region have large uncertainties due to the complexity of interacting ecosystem components [3] and creating accurate frameworks for quantification of disturbances are important to long-term modeling studies, such as those in DOE's NGEA Arctic [4] and NASA's ABoVE [3] projects. Wildfires are the dominant disturbance impacting the boreal forest [2]. Differences in fire frequency and size can occur between Alaska's ecoregions [5], [6]. Fire events in the boreal zone can repeat at a high frequency (105 year fire return interval for 1920–2009 [7]) and be very large in extent ($>1,000$ km²) [8]. While tundra fire events are typically rare and small in size (average of 30–55 km²) [7], evidence shows that climate change has led to an increase in fire occurrence and extent in tundra regions

[9]. One example is the 2004 wildfire season when extreme drought conditions caused a record wildfire year [10].

A popular method for wildfire mapping uses the Normalized Difference Vegetation Index (NDVI), a remotely sensed indicator of photosynthetic capacity [11]. NDVI is one of the most commonly used vegetation indices and is widely available from satellite sensors, such as the Moderate-Resolution Imaging Spectroradiometer (MODIS), a sensor on board NASA's Terra and Aqua satellites that has a spatial resolution of 250 m² and daily temporal resolution. The spatial and temporal resolution makes MODIS a valuable tool for mapping disturbances over large areas. Additionally, other MODIS products (i.e., land surface temperature (LST)) have been shown to be beneficial for wildfire mapping [12].

Deep learning methods have dramatically improved the state-of-the-art in many different classification applications [13]. These methods consist of multiple levels of representation, obtained by transforming the representation at one level into a representation at a higher, slightly more abstract level [13]. A deep neural network (DNN) can contain millions of weights, thousands of internal connected processors called nodes, and many hidden layers. However, most of the existing deep learning algorithms do not take the data imbalance problem into consideration [14]. Imbalanced data classification exists where one class (e.g., burned areas) contains a much smaller sample size than the others (e.g., unburned areas) in classification. It poses a great challenge for DNN architectures, due to the difficulty in recognizing the minority class [15]. However, there has been a significant amount of research performed on the class imbalance problems using dataset resampling [16], cost-sensitive weighting [17], and few-shot learning [18].

This paper describes application of DNN classification methods to the detection of wildfire events (i.e., recently burned land area or areas actively burning) for the 2004 wildfire year in Alaska. To identify the best approach for detecting wildfire events using imbalanced datasets, we follow the DNN training approach of Sze-To & Wong [15], where the original training data set is split into two subsets, one subset for training (updating weights) and the other for validation (selecting weights). The weights that render the best performance for the validation set are retained for subsequent classification.

II. STUDY AREA

The climate across Alaska is highly regionalized due to complex terrain, proximity to oceans and sea ice, and its large geographic extent [19]. Across Alaska, the length of the growing season period is an important factor for determining the spatial and temporal distribution of green vegetation (detected through satellite NDVI), which can vary significantly in Alaska from day of year (DOY) 92–300 for boreal ecosystems and DOY 130–273 for Tundra [20]. The study area (Figure 1) covered the interior regions of Alaska and was based on the climate division boundaries by Bieniek et al. [19]. Interior Alaska was chosen because of the large number of wildfires in the boreal region. The study area is bounded to the north by the Brooks Range and to the south by the Yukon–Tanana uplands and Alaska Range [19]. It is relatively far from ocean influences and has a continental climate with relatively low precipitation [19].

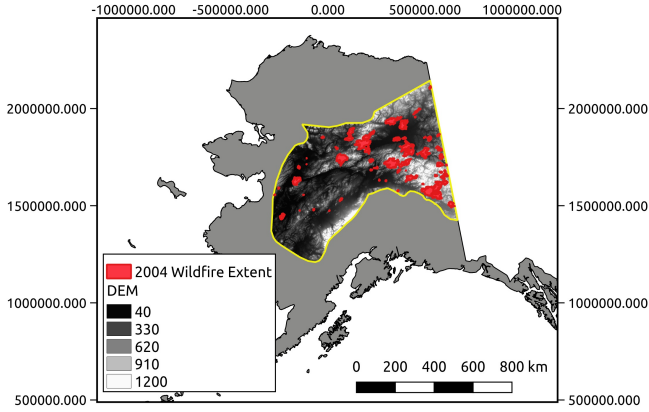


Fig. 1. Study area (yellow lines) and the 2004 wildfire boundaries (red polygons).

III. MATERIALS AND METHODS

A. Deep Neural Networks

A deep neural network (DNN) is an artificial neural network that consists of multiple hidden layers between the input and output layers. This study is focused on feedforward networks. Each hidden layer consists of many units that act in parallel, each representing a vector-to-scalar function [21]. If the dataset $D = \{(x^{(n)}, y^{(n)})\}_{n=1}^N$, where x is the n -dimensional vector and y is the class label associated with the instance x , then a feedforward neural network models the data as a nonlinear function of

$$p(y^{(n)} = 1 | x^{(n)}, \theta) = \sigma \left(\sum_i \theta_i x_i^{(n)} \right), \quad (1)$$

where θ represents the parameters of the network (e.g., weights) and σ represents the activation function that is used to determine the activation of the output node. A DNN learns

the value of the parameters θ that result in the best function approximation [21]. This can be represented as

$$y^{(n)} = \sum_j \theta_j^{(2)} \sigma \left(\sum_i \theta_{ji}^{(1)} x_i^{(n)} \right) + \epsilon^{(n)}, \quad (2)$$

where ϵ represents the learning rate. DNN learning also requires computing the gradients using the back-propagation algorithm, which calculates the direction and magnitude during training that is used to update the network weights. Training a DNN also requires making decisions such as choosing the optimizer, cost function, activation functions (which are used to compute the hidden layer values), and the form of the output units [21].

B. DNN Architecture

The standard strategy of DNN training consists of running the back-propagation algorithm on the training data to obtain a smaller loss value, and in each epoch an updated set of weights are obtained. Sze-To & Wong [15] proposed a strategy known as the Validation-Loss (VL) strategy that splits the input training data into two sets, one for training the DNN to update weights and the other for performance validation to select the weights [15]. The validation loss is calculated by the summation of the errors made for each example in the validation sets by cross entropy. The interpretation is how well the model is performing, with the main objective to reduce the validation loss score with respect to the model's parameters by changing the weights through backpropagation. The best weights updated by the VL strategy are selected during DNN training and saved as the final model for testing. Figure 2 shows the current approach of incorporating the VL strategy. The validation dataset was partitioned equally across classes, in order for the weights that are updated to reflect this balanced dataset.

The deep learning library Keras (<http://keras.io/>) with the TensorFlow [22] backend was used to implement DNN in this study. The DNN has five layers, with the 1st layer being the MODIS variables. The 2nd, 3rd, and 4th layers are hidden layers of rectified linear unit (ReLU) as the activation function [23]. The final layer is an output layer consisting of the *softmax* function, which is a categorical probability distribution that represents the probability that any of the classes are true. Figure 3 shows a visual representation of the DNN architecture.

C. Baseline Method

An ensemble model of gradient boosted trees is used to compare with the DNN methods. Specifically, the XGBoost algorithm is used, which has been highly successful in a number of machine learning and data mining challenges [24]. XGBoost is a scalable implementation of gradient boosting machines, which is an ensemble technique that builds upon many weak successive trees to produce a strong model [24]. XGBoost provides a new sparsity-aware algorithm for sparse data and weighted quantile sketch for approximate tree learning [24]. Also, an additional regularization term and model

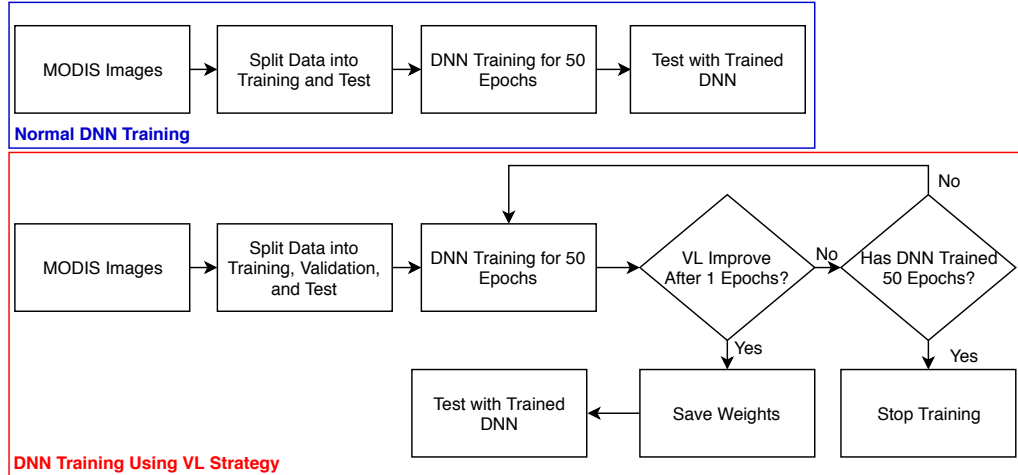


Fig. 2. Flowchart of study when incorporating a standard and Validation-Loss strategy (red box).

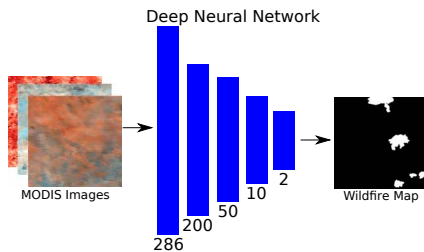


Fig. 3. DNN architecture for classifying MODIS images. The input represents the 286 MODIS bands (described in section III-D1).

formalization helps to smooth the final weights to avoid overfitting [24].

D. Datasets

Table I lists the raster datasets used in this study. The MODIS products were processed in the Google Earth Engine (GEE) platform (<https://earthengine.google.com/datasets/>). The MOD11A2 level-3 MODIS global Land Surface Temperature (LST) and Emissivity 8-day data are composed from the daily 1-kilometer LST product (MOD11A1) and stored on a 1-km Sinusoidal grid as the average values of clear-sky LSTs during an 8-day period. The MOD11A2 products are comprised of daytime and nighttime LSTs, quality assessment, observation times, view angles, bits of clear sky days and nights, and emissivities. The MODIS Surface-Reflectance Product (MOD09A1) provides an estimate of the surface spectral reflectance values at 500 m resolution in a gridded format. Each pixel for both datasets contains the best possible observation during an 8-day period, selected as the one with large observation coverage, a high viewing angle, an absence of clouds or cloud shadow, and aerosol loading [25]. The Monitoring Trends in Burn Severity (MTBS) product (<https://www.mtbs.gov>) includes all fires 1000 acres or greater

in the western United States, which is used as the target class for the MODIS variables. MTBS was chosen because of the accuracy in its representation of the extent of wildfires.

TABLE I
RASTER PREDICTOR VARIABLES.

Description	Variable	Resolution
MOD09A1	NDVI	500 m at 8 days
MOD09A1	EVI	500 m at 8 days
MOD09A1	SAVI	500 m at 8 days
MOD09A1	Bands 1–7	500 m at 8 days
MOD11A2	Daytime LST (Kelvin)	1 km at 8 days
MTBS	Wildfire Extent	500 m

1) *Data Processing*: Figure 4 shows the main processing steps for the MODIS variables and MTBS dataset. All MODIS products were processed from early-April through late-October in 2004. The MTBS vector layers were converted to raster and resampled to 500 m. Additionally all products were resampled to 500 m using the nearest neighbor resampling method in GEE. Contaminated pixels can often lead to false disturbance signatures when analyzing MODIS satellite time series datasets [26]. Thus, it is important to remove such data before performing the analysis. The MODIS products include a quality control (QC) file. Per-pixel QC information in MODIS products allows for removal of most contamination of the NDVI signal related to clouds, aerosol and snow [25]. For all MODIS products, only the good data quality information from the QC file was kept and the rest were removed. For the MOD11A2, all pixels that have average LST errors less than 1 K (i.e., QC = 0, 1, 5, 17, 21) were kept, while pixels with other QC values were removed.

Noise reduction and gap filling in the time series of MODIS data are active areas of research and needed for accurate time series analysis [27]. The Savitzky-Golay algorithm was applied on the filtered MODIS products to fill in the missing pixels

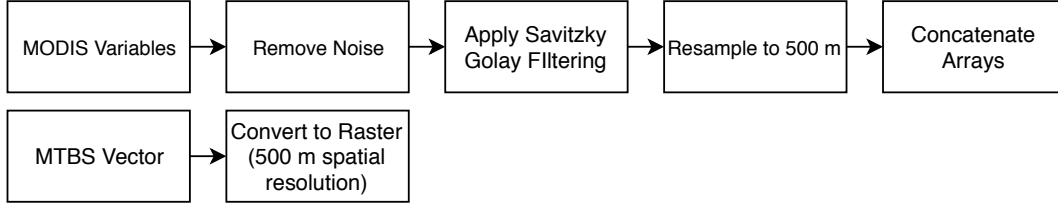


Fig. 4. Main processing steps for the MODIS variables and MTBS dataset.

and smooth the data. The smoothing is performed to capture the overall trend of spectral signatures after a wildfire. It has been shown that the Savitzky-Golay smoothing algorithm is a popular choice based on MODIS datasets [28]. Savitzky-Golay applies a moving-window quadratic polynomial function to the original time-series data and estimates new values for the center point of each moving-window [28]. Figure 5 shows the averaged NDVI values of a large wildfire occurring on July 7, 2004, where the dotted line represents the Savitzky-Golay smoothing. Finally, the MODIS images were concatenated into one image consisting of 288 variables.

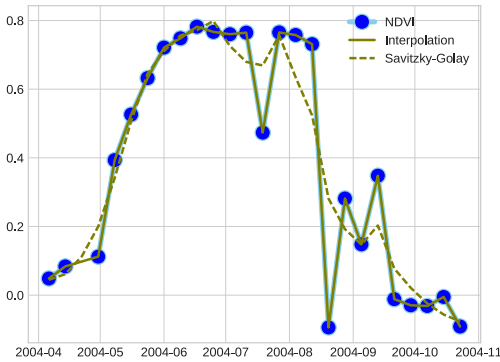


Fig. 5. Plot showing the Savitzky-Golay algorithm averaged over a large wildfire occurring on July 7, 2004. NDVI values drop significantly after the wildfire started.

2) *Data Splitting and Validation Metrics:* The MODIS and MTBS datasets were split into training and testing categories for building and validating the model. Overall, the MODIS and MTBS pixels contained 1,742,618 no-wildfire pixels and 105,072 wildfire pixels, with both classes covering an area of 461,922.5 km² in Interior Alaska. The original dataset was split into 2 categories varying the amount of data for each category, where dataset-0 contains a large amount of training data and dataset-1 contains a small amount of training data and a large amount of testing data. The validation set was used only when the VL strategy was applied and was selected from the training dataset. It was partitioned to have equal amounts for each category. Additionally, a single wildfire was removed from the original dataset and tested using models trained from

dataset-0 and dataset-1. The purpose was to create visual maps of the single wildfire covering an area of 70 km² to assess and illustrate model performance. Table II shows the amount of data used for the training, testing, validation, and a single wildfire.

The precision-recall metric was used to evaluate the wildfire classifications, which is a useful measure of success of prediction when the classes are very imbalanced. A high recall but low precision returns many results, but most of the predicted labels are incorrect when compared to the training labels [29]. A system with high precision but low recall is just the opposite, returning very few results, but most of its predicted labels are correct when compared to the training labels [29]. An ideal system will have a high precision and high recall [29].

Precision (P) is defined as the number of true positives (T_p) over the number of true positives plus the number of false negatives (F_p), which can be defined as

$$P = \frac{T_p}{T_p + F_p}. \quad (3)$$

Recall (R) is defined as the number of true positives (T_p) over the number of true positives plus the number of false negatives (F_n), which can be defined as

$$R = \frac{T_p}{T_p + F_n}. \quad (4)$$

Additionally, a normalized confusion matrix between 0 and 1 was produced to evaluate the quality of the classifier. This was performed by taking the percentage of elements of the correct class i that was classified into each class and taking row fixing the i and divide each element by the sum of the elements in the row.

IV. RESULTS

A. Conventional DNN Training

Table III shows the results for dataset-0 and dataset-1 when the conventional DNN training method is used (Figure 2). Dataset-0 performs the best for mapping wildfires with precision and recall scores of 0.90 and 0.90, respectively, for the wildfire class. Dataset-1 performed the worst and was unable to classify the wildfire class using the test dataset. This is mainly due to the small amount of training data and the large amount of test data (Table II) for dataset-1. Figure 6 shows

TABLE II
NUMBER OF PIXELS (500×500) USED FOR TRAINING, TESTING, AND VALIDATION OF THE DNN ALGORITHM FOR EACH DATASET. THE VALIDATION COLUMN WAS ONLY APPLIED WHEN USING THE VL STRATEGY.

Dataset	No-Fire	Fire	Percentage
Dataset-0 Train	1,154,333	70,493	75%
Dataset-0 Test	427,115	26,356	25%
Dataset-0 Validation	7,947	7,947	10%
Dataset-1 Train	384,375	23,477	25%
Dataset-1 Test	1,282,862	78,702	75%
Dataset-1 Validation	2,617	2,617	10%
Single Wildfire	9,724	276	<1%

TABLE III
CONVENTIONAL DNN TRAINING METHOD PRECISION, RECALL, AND NUMBER OF TEST SAMPLES

Dataset	Class	Precision	Recall	Samples
0	Fire	0.90	0.90	26,356
	No-Fire	0.99	0.99	427,115
1	Fire	0.00	0.00	78,702
	No-Fire	1.00	1.00	1,282,862

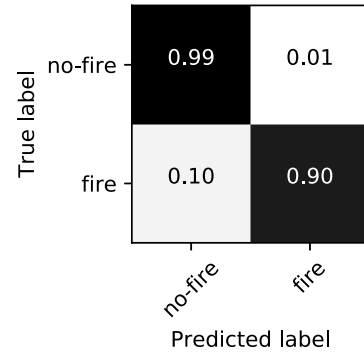
the normalized confusion matrices for dataset-0 and dataset-1. Dataset-0 performs well with a normalized score of 0.90 for the wildfire class, with a small number of false positives. Dataset-1 was unable to classify the wildfire class, with a large number of false positives.

1) *Single Wildfire Mapping*: Figure 7 shows the model results using the standard DNN training applied to a single wildfire (100×100 pixel grid) that was held out of the training, testing, and validation dataset. Figure 8 shows the normalized confusion matrix for dataset-0 and dataset-1 using the standard DNN training. Dataset-0 performs the best with a normalized score of 0.74 for the wildfire class. Dataset-1 performs the worst and was unable to map the wildfire.

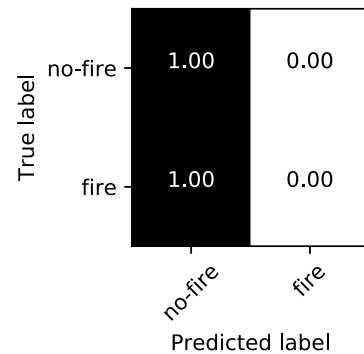
B. VL DNN Training

TableIV shows the results for dataset-0 and dataset-1 using the VL DNN training method (Figure 2). Dataset-0 had a precision and recall of 0.68 and 0.95, respectively, for the wildfire class. Dataset-1 had similar scores with precision and recall of 0.61 and 0.96, respectively, for the wildfire class. Dataset-0 and dataset-1 both improved scores from using the standard DNN training (Table III). This is a significant improvement for dataset-1 (i.e., most imbalanced dataset) increasing the recall score from 0.00 to 0.96. Figure 9 shows the normalized confusion matrices for dataset-0 and dataset-1. Dataset-1 performs significantly better with a normalized score of 0.96 for the wildfire class, with a small amount of false positives. Dataset-0 also performs better with a normalized score of 0.95 for the wildfire class.

Figure 10 shows the training accuracy scores when using the VL strategy. The red line indicates validation-loss scores based on the amount of validation samples (Table II) being



(a) Dataset-0



(b) Dataset-1

Fig. 6. Confusion matrices results using the standard DNN Training for dataset-0 and dataset-1.

TABLE IV
VL DNN TRAINING METHOD PRECISION, RECALL, AND NUMBER OF TEST SAMPLES

Dataset	Class	Precision	Recall	Samples
0	Fire	0.68	0.95	26,356
	No-Fire	1.00	0.97	427,115
1	Fire	0.61	0.96	78,702
	No-Fire	1.00	0.96	1,282,862

used to select the weights. Dataset-1 shows higher variation in the validation loss scores during training compared to dataset-0. This is due to the amount of data being used during training and validation for updating the weights (Table II). Both datasets show the validation-loss reaching the lowest scores after 20 epochs.

1) *Single Wildfire Mapping*: Figure 11 shows the model results using the VL DNN training applied to a single wildfire (100×100 pixel grid) that was held out of the training, testing, and validation dataset. Figure 12 shows the normalized confusion matrix for dataset-0 and dataset-1 using the VL DNN training. Dataset-0 performs the best with a normalized score of 0.71 for the wildfire class. Dataset-1 improved on the standard DNN training, with a normalized score of 0.55 for the wildfire class, significantly improving the score from 0.00 (Figure 8). Both methods for dataset-0 also identify wildfire

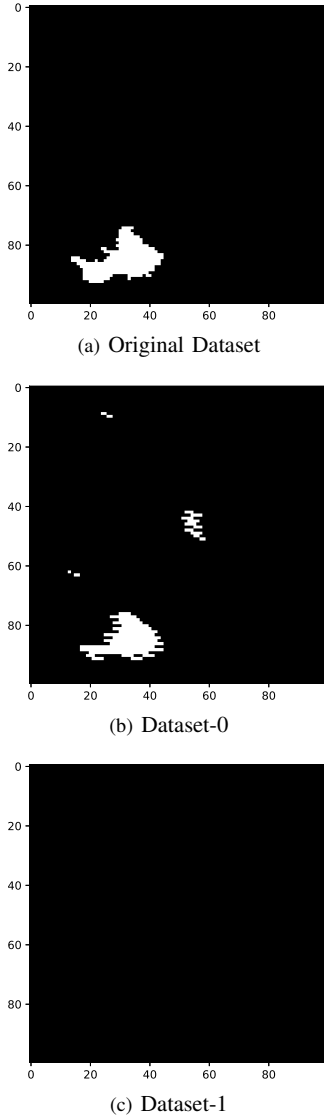


Fig. 7. Mapping results for a single wildfire (a) of dataset-0 (b) and dataset-1 (c) using the standard DNN training models.

pixels (Figure 7 (b) and Figure 11 (b)) that are outside the main perimeters delineated by MTBS (Figure 11 (a)), indicating possible burnt area that were not captured by MTBS.

C. Comparisons with XGBoost Algorithm

Table V shows the results for dataset-0 and dataset-1 for using the XGBoost method for the entire study region and the single wildfire. Both datasets perform equally, with precision and recall scores of 0.95 and 0.85, respectively, for the wildfire class over the study region. For the single wildfire, dataset-0 performed slightly better with precision and recall scores of 0.88 and 0.76, respectively. Overall, XGBoost performed well given the imbalanced dataset when compared to the standard DNN training III. However, the VL DNN training method showed better results IV when compared to XGBoost.

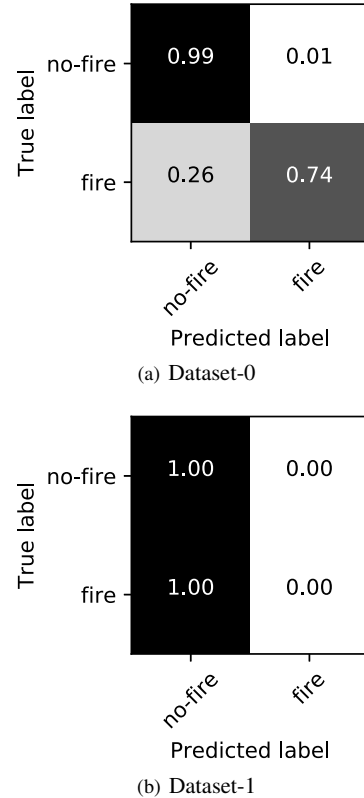


Fig. 8. Confusion matrices results for the single wildfire using the standard DNN Training for dataset-0 and dataset-1.

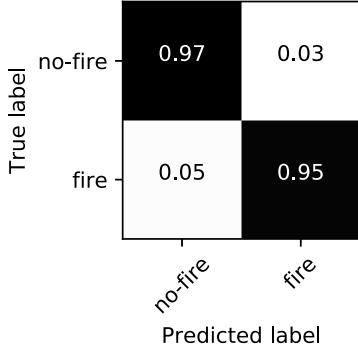
TABLE V
XGBOOST METHOD SCORES FOR PRECISION, RECALL, AND NUMBER OF TEST SAMPLES

Dataset	Class	Precision	Recall	Samples	Region
0	Fire	0.94	0.85	26,356	Study Region
	No-Fire	0.99	1.00	427,115	Study Region
	Fire	0.88	0.77	276	Single Wildfire
	No-Fire	0.99	1.00	9,724	Single Wildfire
1	Fire	0.94	0.85	78,702	Study Region
	No-Fire	0.99	1.00	1,282,862	Study Region
	Fire	0.88	0.76	276	Single Wildfire
	No-Fire	0.99	1.00	9,724	Single Wildfire

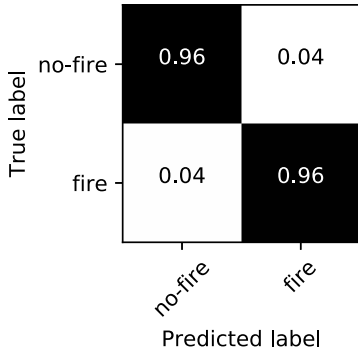
V. DISCUSSION

The VL DNN training strategy showed significant improvement over the standard DNN training for dataset-1 when testing against the wildfires. The scores only reflect comparisons with the perimeters delineated by MTBS, while this method could also identify other small wildfire burn areas (Figure 11), which were perhaps not mapped since MTBS only maps wildfire above 1000 acres [30]. While the results in this paper were applied only for Interior Alaska, the VL strategy could be applied for global-scale wildfire mapping, since the method does not rely on parameters specific to this region.

Google Earth Engine (GEE) was used to process the data



(a) Dataset-0

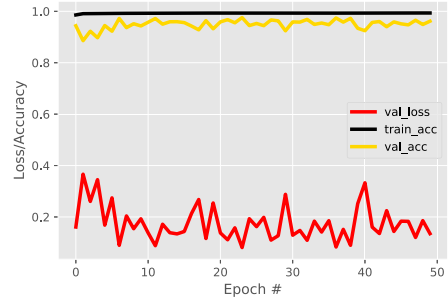


(b) Dataset-1

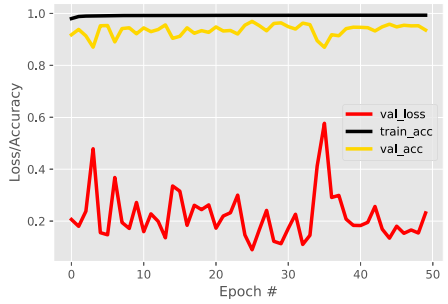
Fig. 9. Confusion matrices results using the VL DNN Training for dataset-0 and dataset-1.

and is a valuable resource for data processing. GEE has a vast database of remote sensing products and this method could be used at higher spatial (i.e., Landsat) and temporal resolutions (i.e., daily). Other datasets, such as land cover and climate variables, within GEE could provide opportunities for new machine learning approaches. For example, it has been shown that prolonged periods of warm and dry conditions enables large wildfires, while wetting precipitation within a week of ignition enables smaller wildfires [31]. Additionally, it is important to validate these methods to existing MODIS-based active fire maps. It has been shown that MODIS surface reflectance products (MOD09A1) has been used to successfully map wildfires using the MODIS active fire product (MOD14A2) as training data [32]. Future research will investigate using MOD14A2 and other fire-based products instead of the MTBS dataset. New machine learning approaches that incorporate these dynamics will be beneficial for additional strategies to map wildfire extent.

XGBoost showed high accuracies (Table V) in mapping wildfires, this could be due to an additional regularization term, and model formalization helps to smooth the final weights to avoid overfitting [24]. More advanced deep learning methods should be investigated for wildfire mapping in Alaska. It has been shown that convolutional neural networks (CNNs) are powerful algorithms for generating feature vectors in this



(a) Dataset-0



(b) Dataset-1

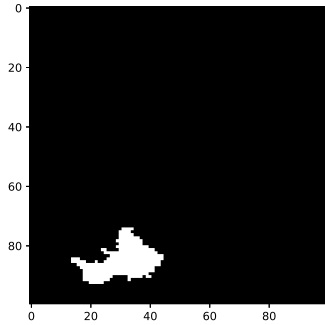
Fig. 10. Accuracy and validation loss scores using the validation dataset (Table II) during VL DNN Training for dataset-0 and dataset-1 using 50 epochs.

region [33]. The VL training strategy could also be applied to CNNs along with other deep learning architectures that help with overfitting (e.g., dropout, regularization techniques). Architectures that leverage sequential data, such as recurrent neural networks (RNNs), have been shown to provide accurate classification performance of MODIS datasets [34], [35]. Additionally, sequential deep learning architectures have shown that image preprocessing (noise removal) is handled by the networks which could alleviate the need for performing such tasks [36].

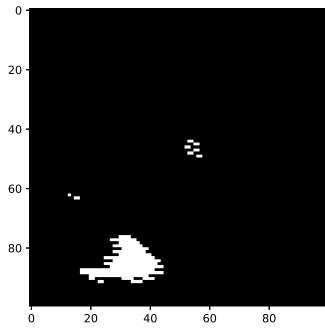
Model learning from imbalanced datasets is a significant challenge for remote sensing datasets. This method splits the input training data into sets, one for training the DNN to update weights and the other validation set (split equally among classes) for performance validation to select the weights [15]. While this method was shown to significantly improve performance for the highly imbalanced dataset, other techniques for model learning that is independent from the global model optimization should also be investigated.

VI. CONCLUSION

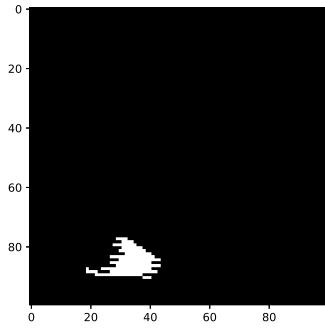
We present a method for training deep neural networks (DNN) using a validation-loss (VL) weight selection strategy on imbalanced datasets. The method selects the weights based on the validation-loss scores during training when using equal amounts per class to reflect a balanced dataset (Figure 2). The



(a) Original Dataset



(b) Dataset-0



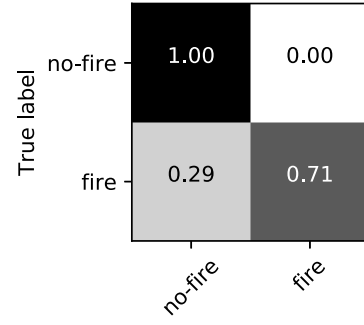
(c) Dataset-1

Fig. 11. Mapping results for a single wildfire (a) of dataset-0 (b) and dataset-1 (c) using the VL DNN training models.

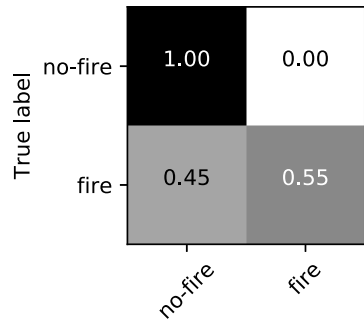
VL approach showed a significant improvement when using the heavily imbalanced dataset (i.e., dataset-1), increasing the recall score from 0.00 to 0.96 when performing on 78,702 wildfire pixels (500×500 m). It provides a remote sensing based approach for large scale mapping of wildfires and other classification problems in ecology that suffers from the issue of imbalanced data.

ACKNOWLEDGMENT

The Next-Generation Ecosystem Experiments (NGEE Arctic) project is supported by the Office of Biological and Environmental Research in the DOE Office of Science. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and



(a) Dataset-0



(b) Dataset-1

Fig. 12. Confusion matrices results for the single wildfire using the VL DNN Training for dataset-0 and dataset-1.

the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

REFERENCES

- [1] A. J. Soja, N. M. Tchepakova, N. H. French, M. D. Flannigan, H. H. Shugart, B. J. Stocks, A. I. Sukhinin, E. Parfenova, F. S. C. III, and P. W. S. Jr., "Climate-induced boreal forest change: Predictions versus current observations," *Global and Planetary Change*, vol. 56, no. 34, pp. 274 – 296, 2007, northern Eurasia Regional Climate and Environmental Change. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0921818106001883>
- [2] R. H. Fraser, I. Olthof, S. V. Kokelj, T. C. Lantz, D. Lacelle, A. Brooker, S. Wolfe, and S. Schwarz, "Detecting Landscape Changes in High Latitude Environments Using Landsat Trend Analysis: 1. Visualization," *Remote Sensing*, vol. 6, no. 11, p. 11533, 2014. [Online]. Available: <http://www.mdpi.com/2072-4292/6/11/11533>
- [3] J. B. Fisher, D. J. Hayes, C. R. Schwalm, D. N. Huntzinger, E. Stofferahn, K. Schaefer, Y. Luo, S. D. Wullschleger, S. Goetz, C. E. Miller, P. Griffith, S. Chadburn, A. Chatterjee, P. Ciais, T. A. Douglas, H. Genet, A. Ito, C. S. R. Neigh, B. Poulter, B. M. Rogers, O. Sonnentag, H. Tian, W. Wang, Y. Xue, Z.-L. Yang, N. Zeng, and Z. Zhang, "Missing pieces to modeling the Arctic-Boreal puzzle,"

- Environmental Research Letters*, vol. 13, no. 2, p. 020202, 2018. [Online]. Available: <http://stacks.iop.org/1748-9326/13/i=2/a=020202>
- [4] S. D. Wullschlegel, L. D. Hinzman, and C. J. Wilson, "Planning the Next Generation of Arctic Ecosystem Experiments," *Eos, Transactions American Geophysical Union*, vol. 92, no. 17, pp. 145–145, 2011. [Online]. Available: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011EO170006>
 - [5] F. M. Hoffman, J. Kumar, R. T. Mills, and W. W. Hargrove, "Representativeness-based sampling network design for the State of Alaska," *Landscape Ecology*, vol. 28, no. 8, pp. 1567–1586, Oct. 2013.
 - [6] L. K. Jenkins, L. L. Bourgeau-Chavez, N. H. F. French, T. V. Loboda, and B. J. Thelen, "Development of Methods for Detection and Monitoring of Fire Disturbance in the Alaskan Tundra Using a Two-Decade Long Record of Synthetic Aperture Radar Satellite Images," *Remote Sensing*, vol. 6, no. 7, p. 6347, 2014. [Online]. Available: <http://www.mdpi.com/2072-4292/6/7/6347>
 - [7] E. S. Kasischke, D. L. Verbyla, T. S. Rupp, A. D. McGuire, K. A. Murphy, R. Jandt, J. L. Barnes, E. E. Hoy, P. A. Duffy, M. Calef, and M. R. Turetsky, "Alaska's changing fire regime – implications for the vulnerability of its boreal forests," *Canadian Journal of Forest Research*, vol. 40, no. 7, pp. 1313–1324, 2010.
 - [8] E. S. Kasischke and M. R. Turetsky, "Recent changes in the fire regime across the North American boreal region – Spatial and temporal patterns of burning across Canada and Alaska," *Geophysical Research Letters*, vol. 33, no. 9, pp. n/a–n/a, 2006, 109703. [Online]. Available: <http://dx.doi.org/10.1029/2006GL025677>
 - [9] F. S. Hu, P. E. Higuera, J. E. Walsh, W. L. Chapman, P. A. Duffy, L. B. Brubaker, and M. L. Chipman, "Tundra burning in Alaska: Linkages to climatic change and sea ice retreat," *Journal of Geophysical Research: Biogeosciences*, vol. 115, no. G4, pp. n/a–n/a, 2010, g04002. [Online]. Available: <http://dx.doi.org/10.1029/2009JG001270>
 - [10] G. Pfister, P. G. Hess, L. K. Emmons, J.-F. Lamarque, C. Wiedinmyer, D. P. Edwards, G. Ptron, J. C. Gille, and G. W. Sachse, "Quantifying CO emissions from the 2004 Alaskan wildfires using MOPITT CO data," *Geophysical Research Letters*, vol. 32, no. 11, pp. n/a–n/a, 2005, 111809. [Online]. Available: <http://dx.doi.org/10.1029/2005GL022995>
 - [11] C. J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sensing of Environment*, vol. 8, no. 2, pp. 127 – 150, 1979. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0034425779900130>
 - [12] Z. Tan, S. Liu, B. Wylie, C. Jenkerson, J. Oeding, J. Rover, and C. Young, "MODIS-informed greenness responses to daytime land surface temperature fluctuations and wildfire disturbances in the Alaskan Yukon River Basin," *International Journal of Remote Sensing*, vol. 34, no. 6, pp. 2187–2199, 2013.
 - [13] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 5 2015.
 - [14] S. Wang, W. Liu, J. Wu, L. Cao, Q. Meng, and P. J. Kennedy, "Training deep neural networks on imbalanced data sets," in *2016 International Joint Conference on Neural Networks (IJCNN)*, July 2016, pp. 4368–4374.
 - [15] A. Sze-To and A. Wong, "A weight-selection strategy on training deep neural networks for imbalanced classification," in *International Conference Image Analysis and Recognition (ICIAR)*, June 2017.
 - [16] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
 - [17] K. M. Ting, "A comparative study of cost-sensitive boosting algorithms," in *Proceedings of the Seventeenth International Conference on Machine Learning*, ser. ICML '00. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000, pp. 983–990. [Online]. Available: <http://dl.acm.org/citation.cfm?id=645529.657944>
 - [18] S. Ravi and H. Larochelle, "Optimization as a model for few-shot learning," in *In International Conference on Learning Representations (ICLR)*, 2017.
 - [19] P. A. Bieniek, U. S. Bhatt, R. L. Thoman, H. Angeloff, J. Partain, J. Papineau, F. Fritsch, E. Holloway, J. E. Walsh, C. Daly, M. Shulski, G. Hufford, D. F. Hill, S. Calos, and R. Gens, "Climate Divisions for Alaska Based on Objective Methods," *Journal of Applied Meteorology and Climatology*, vol. 51, no. 7, pp. 1276–1289, 2012.
 - [20] M. Ueyama, H. Iwata, Y. Harazono, E. S. Euskirchen, W. C. Oechel, and D. Zona, "Growing season and spatial variations of carbon fluxes of Arctic and boreal ecosystems in Alaska (USA)," *Ecological Applications*, vol. 23, no. 8, pp. 1798–1816, 2013. [Online]. Available: <http://dx.doi.org/10.1890/11-0875.1>
 - [21] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
 - [22] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. A. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zhang, "TensorFlow: A system for large-scale machine learning," *CoRR*, vol. abs/1605.08695, 2016. [Online]. Available: <http://arxiv.org/abs/1605.08695>
 - [23] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in *Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML-10)*, 2010, pp. 807–814.
 - [24] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794. [Online]. Available: <http://doi.acm.org/10.1145/2939672.2939785>
 - [25] X. Lu, Z. Lin, X. Shen, R. Mech, and J. Z. Wang, "Deep Multi-patch Aggregation Network for Image Style, Aesthetics, and Quality Estimation," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec 2015, pp. 990–998.
 - [26] D. Verbyla, "Comparison of eMODIS and MOD/MYD13A2 NDVI products during 2012–2014 spring green-up periods in Alaska and northwest Canada," *International Journal of Applied Earth Observation and Geoinformation*, vol. 36, pp. 83 – 86, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0303243414002591>
 - [27] A. Moreno, F. J. García-Haro, B. Martínez, and M. A. Gilabert, "Noise reduction and gap filling of fapar time series using an adapted local regression filter," *Remote Sensing*, vol. 6, no. 9, pp. 8238–8260, 2014. [Online]. Available: <http://www.mdpi.com/2072-4292/6/9/8238>
 - [28] J. Ren, J. B. Campbell, and Y. Shao, "Estimation of SOS and EOS for Midwestern US Corn and Soybean Crops," *Remote Sensing*, vol. 9, no. 7, 2017. [Online]. Available: <http://www.mdpi.com/2072-4292/9/7/722>
 - [29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
 - [30] J. C. Eidsenshink, B. Schwind, K. Brewer, Z.-L. Zhu, B. Quayle, and S. M. Howard, "A project for monitoring trends in burn severity," *Fire Ecology*, vol. 3, no. 1, pp. 3–21, 2007. [Online]. Available: <http://pubs.er.usgs.gov/publication/70156000>
 - [31] J. T. Abatzoglou and C. A. Kolden, "Relative importance of weather and climate on wildfire growth in interior Alaska," *International Journal of Wildland Fire*, vol. 20, no. 4, pp. 479–486, 2011. [Online]. Available: <https://doi.org/10.1071/WF10046>
 - [32] V. Mithal, G. Nayak, A. Khandelwal, V. Kumar, R. Nemani, and N. C. Oza, "Mapping burned areas in tropical forests using a novel machine learning framework," *Remote Sensing*, vol. 10, no. 1, 2018. [Online]. Available: <http://www.mdpi.com/2072-4292/10/1/69>
 - [33] Z. L. Langford, J. Kumar, and F. M. Hoffman, "Convolutional Neural Network Approach for Mapping Arctic Vegetation using Multi-Sensor Remote Sensing Fusion," in *Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW 2017)*, Institute of Electrical and Electronics Engineers (IEEE). Conference Publishing Services (CPS), Nov. 2017.
 - [34] Y.-L. Kong, Q. Huang, C. Wang, J. Chen, J. Chen, and D. He, "Long short-term memory neural networks for online disturbance detection in satellite image time series," *Remote Sensing*, vol. 10, no. 3, 2018. [Online]. Available: <http://www.mdpi.com/2072-4292/10/3/452>
 - [35] A. X. Wang, C. Tran, N. Desai, D. Lobell, and S. Ermon, "Deep transfer learning for crop yield prediction with remote sensing data," in *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, ser. COMPASS '18. New York, NY, USA: ACM, 2018, pp. 50:1–50:5. [Online]. Available: <http://doi.acm.org/10.1145/3209811.3212707>
 - [36] M. Rußwurm and M. Körner, "Multi-temporal land cover classification with sequential recurrent encoders," *ISPRS International Journal of Geo-Information*, vol. 7, no. 4, 2018. [Online]. Available: <http://www.mdpi.com/2220-9964/7/4/129>