



Article

Probabilistic Wildfire Segmentation Using Supervised Deep Generative Model from Satellite Imagery

Ata Akbari Asanjan ^{1,2,*}, Milad Memarzadeh ^{1,2} , Paul Aaron Lott ^{2,3} , Eleanor Rieffel ³ and Shon Grabbe ³

¹ Data Science Group (DSG), NASA Ames Research Center, Moffett Field, CA 94035, USA; milad.memarzadeh@nasa.gov

² USRA Research Institute for Advanced Computer Science (RIACS), Washington, DC 20024, USA; plott@usra.edu

³ Quantum Artificial Intelligence Laboratory (QuAIL), NASA Ames Research Center, Moffett Field, CA 94035, USA; eleanor.rieffel@nasa.gov (E.R.); shon.grabbe@nasa.gov (S.G.)

* Correspondence: ata.akbariasanjan@nasa.gov

Abstract: Wildfires are one of the major disasters among many and are responsible for more than 6 million acres burned in the United States alone every year. Accurate, insightful, and timely wildfire detection is needed to help authorities mitigate and prevent further destruction. Uncertainty quantification is always a crucial part of the detection of natural disasters, such as wildfires, and modeling products can be misinterpreted without proper uncertainty quantification. In this study, we propose a supervised deep generative machine-learning model that generates stochastic wildfire detection, allowing fast and comprehensive uncertainty quantification for individual and collective events. In the proposed approach, we also aim to address the patchy and discontinuous Moderate Resolution Imaging Spectroradiometer (MODIS) wildfire product by training the proposed model with MODIS raw and combined bands to detect fire. This approach allows us to generate diverse but plausible segmentations to represent the disagreements regarding the delineation of wildfire boundaries by subject matter experts. The proposed approach generates stochastic segmentation via two model streams in which one learns meaningful stochastic latent distributions, and the other learns the visual features. Two model branches join eventually to become a supervised stochastic image-to-image wildfire detection model. The model is compared to two baseline stochastic machine-learning models: (1) with permanent dropout in training and test phases and (2) with Stochastic ReLU activations. The visual and statistical metrics demonstrate better agreements between the ground truth and the proposed model segmentations. Furthermore, we used multiple scenarios to evaluate the model comprehension, and the proposed Probabilistic U-Net model demonstrates a better understanding of the underlying physical dynamics of wildfires compared to the baselines.

Keywords: wildfire detection; generative machine-learning; stochastic modeling; remote sensing; segmentation; uncertainty analysis



Citation: Akbari Asanjan, A.; Memarzadeh, M.; Lott, P.A.; Rieffel E.; Grabbe S. Probabilistic Wildfire Segmentation Using Supervised Deep Generative Model from Satellite Imagery. *Remote Sens.* **2023**, *15*, 2718. <https://doi.org/10.3390/rs15112718>

Academic Editor: Gwanggil Jeon

Received: 4 April 2023

Revised: 4 May 2023

Accepted: 8 May 2023

Published: 24 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Wildfires are one of the necessary dynamic components of terrestrial ecosystems, and they provide significant ecological benefits [1,2]. Natural wildfires offer significant ecological benefits through promoting forest rejuvenation, nutrient cycling, and habitat diversity, all of which contribute to the overall health and resilience of ecosystems [3].

However, it is important to acknowledge the growing trends in wildfire size, frequency, and intensity, which are largely influenced by human activities and interventions. Factors such as wildfire suppression efforts and urban/wildlife encroachment have contributed to increases in wildfire size, frequency, and intensity [4,5]. These anthropogenic influences have transformed wildfires into a global problem in recent decades [2,6]. Consequently, wildfires have emerged as one of the most destructive natural hazards, with severe consequences for both human and ecological systems.

The scale of devastation caused by wildfires is alarming. In the United States alone, wildfires have burned over 6 million acres of land [7], while globally, the figure exceeds 1 billion acres [8]. These staggering numbers underscore the urgent need for effective wildfire management strategies, enhanced understanding of fire behavior, and improved decision-making processes to mitigate the devastating impacts of wildfires on communities, ecosystems, and economies. By leveraging advanced technologies, such as remote sensing, data analytics, and predictive modeling, researchers and practitioners can gain valuable insights into wildfire dynamics and develop proactive measures to reduce risks and enhance resilience in fire-prone regions.

Preventive and mitigative decision-making and proper resource management are tied to the availability of insightful, accurate, and timely wildfire monitoring, along with a deep understanding of wildfire dynamics. For this reason, wildfire detection is an active research field focusing on understanding wildfire's complex processes and correlated factors (e.g., fuel load and structure, vegetation health and types) [9–11]. Many studies have been conducted to improve the accuracy and detection latency [12–15] for the benefit of acute and agile decision-making. Wildfire studies can be categorized into two main directions: (1) deterministic and (2) stochastic models. Deterministic models are a category of simulation that assumes the simulation process is fully resolved and the simulations can be conducted with negligible errors [16]. Many studies use such an assumption to address this problem using deterministic tools [17–19]. For instance, Toan et al., 2019 [18] developed a deterministic machine-learning model that uses geostationary satellites (GOES-16) as input to detect wildfires at the pixel level. The paper reports robustness against different wildfire types and adversarial conditions. In another study, Sayed et al., 2019 [19] created a wildfire dataset from satellite imagery and used a feed-forward neural network and Support Vector Machine (SVM) to detect wildfire events. Although deterministic models can solve complex highly non-linear scenarios, they are still insufficient in fully resolving the process behaviors [20]. Additionally, deterministic models are limited from the uncertainty quantification perspective [13,21,22].

Stochastic models, as opposed to deterministic models, acknowledge the presence of unresolved subprocesses and seek to incorporate them into the modeling framework. These stochastic approaches introduce randomness into the inference process, resulting in varying outcomes even under identical conditions. This variability gives rise to a distribution of possible outcomes, providing a comprehensive view of the wildfire segmentation. Such stochastic models prove to be well-suited for wildfire analysis due to their ability to capture the inherent uncertainty and complexity associated with these natural phenomena [23,24]. By considering multiple plausible generated outcomes, stochastic models offer informative insights into the characteristics of potential wildfire patterns and aid in assessing the uncertainty and variability associated with the segmentation results.

The modeling characteristics of stochastic models enable a more comprehensive understanding of the inherent uncertainties and complexities associated with wildfires [25,26]. By considering the variability and uncertainty in the data, stochastic models can capture the inherent stochasticity in wildfire processes and provide valuable insights into the spatial dynamics of wildfire events. This is important in wildfire analysis as it allows for the exploration of various scenarios and the assessment of the likelihood of different outcomes. Moreover, stochastic models facilitate probabilistic-based decision-making processes, enabling more informed and robust wildfire management strategies. The utilization of stochastic models in wildfire analysis has shown promising results in improving our understanding of fire behavior, predicting fire spread patterns [25,26]. Through the incorporation of stochastic modeling approaches, we can enhance our ability to effectively understand and mitigate the risks associated with wildfires, ultimately contributing to more resilient and sustainable fire management practices [25].

Additionally, wildfire ground-truth segmentations are arbitrary and/or noisy to some level, due to human labeling, instrument differences, and other artifacts, affecting the wildfire segmentation shapes. A good example of the wildfire discrepancies can be seen

in the comparison of MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) fire radiative power products [27]. Even subject matter experts (SMEs), assigned to wildfire delineation tasks, often disagree on the active fire's spatial extent. These plausible but discrepant takes from the same events prompt a different look at wildfire detection where wildfire segmentations are considered a distribution of events instead of a single unified segmentation.

With the advent of terrestrial and atmospheric remote sensing, mainly supported by satellite and aviation platforms, the means to monitor and detect wildfires have been more accessible [28]. Advances in observation sensors, and specific enhancement of spatial, temporal, and spectral resolution, allow more in-depth studies and reveal some of the unknown dynamics of fires such as holdover fires [28,29]. However, with the increase in the number of satellites/aviation missions, and the increase of retrieved information, efficient and effective land management through remote sensing has been challenging [15].

Machine learning proposes an opportunity to extract useful information from a large volume of remote sensing datasets. Unsupervised methods are popular for such processes due to generally limited labels for remote sensing in Earth science. Methods such as Auto-Encoders (AEs) are widely used for such tasks where the network is an encoder–decoder architecture and the model aims to learn a compressed representation of the data with minimum information loss [30]. The main issue of these deterministic models in image-to-image translation is their loss of resolution problem. The encoder part of the model subsamples the spatial information to compress the data and due to such an operation, the decoder is not able to recover the spatial information effectively [31,32]. To remedy this issue, [32] proposed U-NET which is encoder–decoder architecture with skip-connections in all spatial resolution levels from encoded activations to the corresponding decoding layers to preserve the spatial information. Despite the wide applications of AEs and U-NETs, they are not capable of learning distributions around events which limits their expressibility of data. Generative models such as variational inference methods enable characterizing stochastic behaviors in data [33], such as ones in wildfire processes.

Variational Auto-Encoders (VAEs) are among the most popular unsupervised variational inference techniques in machine learning. We propose a supervised version of VAEs, developed by [34], where the model consists of four submodels: (1) prior network in charge of learning the latent prior distribution of input data, (2) posterior network in charge of learning the latent posterior distribution of input and target data, (3) U-NET network in charge of feature extraction of inputs, and (4) Combination network that uses the U-NET features and samples from latent distribution to generate stochastic wildfire segmentations.

The main contributions of this work are: (1) developing a stochastic machine-learning model with accurate and fast probabilistic inference on target wildfire segmentation, (2) conducting uncertainty quantification by drawing a significant number of samples, and (3) performing what-if scenarios to understand the impact of inputs variability.

The rest of the paper is structured as follows: Section 2 presents the methodology and proposed model, Section 3 shows the obtained results, uncertainty quantification, and comparison with baseline along with discussions, and Section 4 focuses on the conclusion and summary of findings.

2. Methodology

2.1. Variational Autoencoder

To gain a comprehensive understanding of the proposed methodology, it is imperative to establish a foundational understanding of variational autoencoders (VAEs). VAEs represent a key component in elucidating the intricacies of the proposed approach. Variational autoencoders (VAEs) are powerful *unsupervised* generative models that combine the concepts of autoencoders and variational inference. They are designed to learn a low-dimensional latent space representation of complex high-dimensional input data. The latent space is a continuous multivariate distribution that captures the underlying structure and variations within the data. VAEs consist of two main components: (1) an encoder

and (2) a decoder. The encoder maps input data into the latent space, while the decoder reconstructs the data from the latent space back to the original input space.

In VAEs, instead of directly encoding input data into a single point in the latent representation, the data is encoded into probability distributions over the latent variables [35]. This probabilistic representation allows for more flexibility and uncertainty modeling. It enables VAEs to not only reconstruct the input data but also generate new samples by sampling from the learned probability distributions in the latent representation and decoding them using the decoder network.

The fundamental idea underlying VAEs is to approximate the input data distribution (i.e. marginal likelihood) noted by $P_\theta(x)$. VAEs achieve this goal by maximize the evidence lower bound (ELBO), which serves as the objective function during the training process (1). The ELBO consists of two main components: the reconstruction loss, which measures how well the VAE can reconstruct the input data, and the regularization term that encourages the latent space to adhere to a predefined prior distribution, often a multivariate Gaussian distribution. By maximizing the ELBO, VAEs achieve a delicate balance between accurately reconstructing the input data and regularizing the latent space to follow the prior distribution.

$$\log P_\theta(x) \geq \mathcal{L}(\theta, \phi; x) = \mathbb{E}_{Q_\phi(z|x)} [-\log Q_\phi(z|x) + \log P_\theta(x, z)] \quad (1)$$

Equation (1) can be re-written as equation below with the ELBO loss on left. Right hand side of the equation presents the regularization term, called Kullback-Leibler divergence, and reconstruction term.

$$\mathcal{L}(\theta, \phi; x) = -D_{KL}(Q_\phi(z|x)||P_\theta(z)) + \mathbb{E}_{Q_\phi(z|x)}[\log P_\theta(x|z)] \quad (2)$$

$D_{KL}(Q_\phi(z|x)||P_\theta(z))$ term represents the Kullback-Leibler (KL) divergence between the posterior distribution $Q_\phi(z|x)$ and the prior distribution $P_\theta(z)$. It measures the discrepancy or difference between these two distributions. $\mathbb{E}_{Q_\phi(z|x)}[\log P_\theta(x|z)]$ represents the expected log-likelihood of the reconstruction, where x is the input data and z is a latent variable sampled from the posterior distribution $Q_\phi(z|x)$. It measures how well the VAE can reconstruct the input data given a sampled latent variable.

This regularization process encourages the latent space of the VAE to capture meaningful and continuous representations of the data. It facilitates various tasks, including data generation and interpolation, by ensuring that similar input data points are mapped to nearby regions in the latent space. As a result, VAEs provide a powerful framework for learning complex data distributions and exploring the latent space in a probabilistic manner.

VAEs focus on unsupervised learning and aim to learn meaningful representations of the input data by modeling the underlying probability distributions. Probabilistic U-Net extends the variational capabilities of VAEs to supervised learning and create possibilities to perform tasks such as segmentation in variational context.

2.2. Proposed Approach

Image segmentation is the process of identifying and isolating objects or features of interest in input images. One of the commonly used techniques for the segmentation of instances is the U-NET model, initially developed for biomedical image segmentation but also applicable to other fields such as Earth sciences and space exploration. U-NET is a deep convolutional neural network that performs image-to-image translation by taking an image as input and generating a segmentation map as output. The model is trained using supervised learning, which involves providing accurate segmented images to train the network to map input images to their corresponding segmentations. Despite the impressive performance of U-NET in image segmentation tasks, its deterministic nature poses a limitation. The mapping from input images to output segmentation maps is fully deterministic and fails to consider sources of uncertainty and stochasticity, which can lead

to overfitting and poor generalization to new data. Moreover, the deterministic nature of U-NET limits its ability to perform “what-if” analysis and provide probabilistic segmentations.

Kohl et al. [34] proposed a novel Probabilistic U-Net model for image segmentation that combines the U-NET with a Conditional Variational Auto-Encoder (CVAE) [36,37]. The CVAE framework allows the model to generate plausible hypotheses and explore “what-if” scenarios. The architecture of the proposed model is depicted in Figure 1. Specifically, the U-NET generates segmentations that are conditioned on the samples drawn from the latent feature space of the VAE. This low-dimensional space captures the range of possible segmentation variations and can be used to evaluate “what-if” scenarios during the evaluation phase. By conditioning the segmentation generation on the latent space, the model can produce multiple segmentation maps for a single input image, corresponding to different regions of the latent feature space that are sampled. According to the authors, this capability enables the model to “learn hypotheses that have a low probability and to predict them with the corresponding frequency” (Kohl et al., 2018).

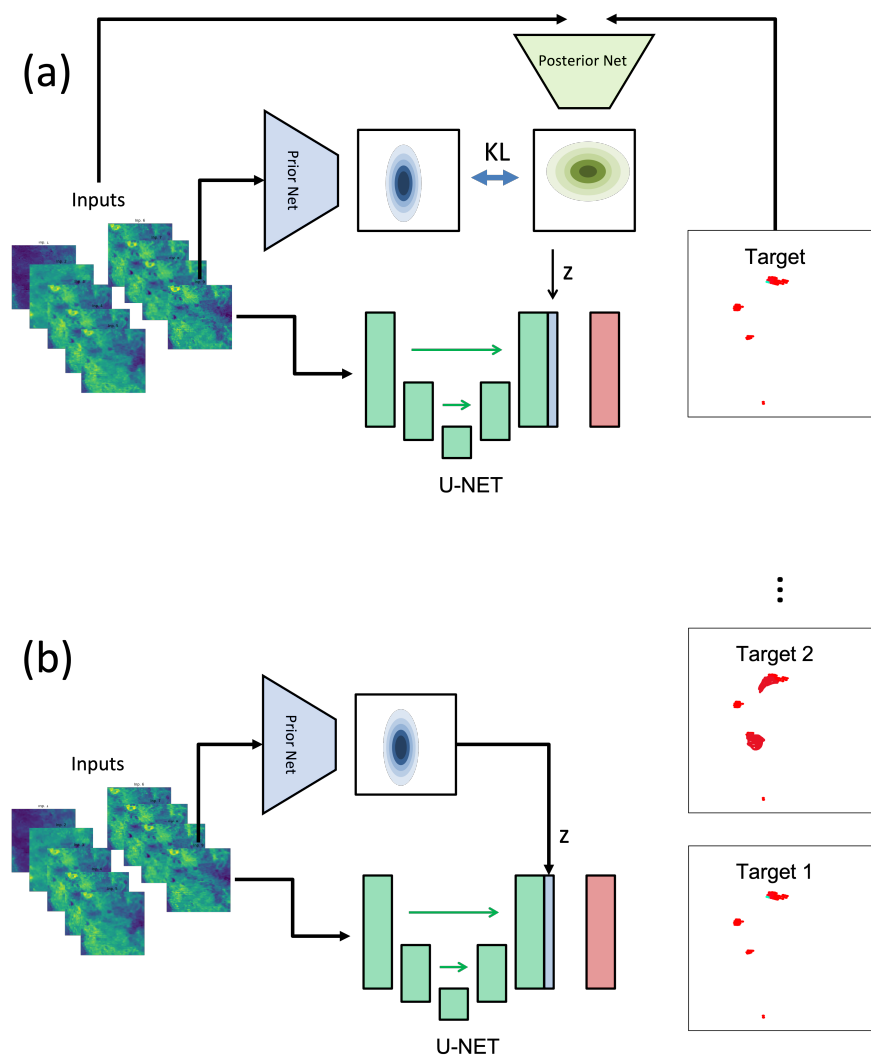


Figure 1. Graphical illustration of the proposed Probabilistic U-Net framework. The inputs are NDVI, NDVI difference with long-term NDVI, and MODIS MCD43A4 channels for Land/Cloud/Aerosols. (a) presents the training scheme where the prior network encodes inputs and the posterior network encodes the inputs and target data together into multivariate Gaussian distributions. The samples from the unified multivariate Gaussian distribution are concatenated with U-Net outputs to produce stochastic events of target data. (b) demonstrates the inference scheme where samples are drawn from the prior network.

The output of the U-NET (the green block) and the drawn sample from latent space (the blue block labeled as z) are concatenated and passed to the red block \mathcal{F} , which generates the corresponding segmentation $S_i = \mathcal{F}(f_{\text{U-NET}}(X, \theta), z_i; \psi)$, where S_i represents the segmentation corresponding to the latent space sample z_i , and θ and ψ are the model parameters of the U-NET model and the \mathcal{F} , respectively. The model is trained using two objectives, namely (1) generating accurate wildfire segmentation from the input data and (2) generalizing well to unseen or rare scenarios. The model is enforced to meet the first objective by minimizing the supervised cross-entropy loss between the generated segmentation, $S(X, z)$, and the ground truth, Y . The model generalizes its understanding by minimizing the Kullback–Leibler divergence between the prior, $P(z | X)$, and posterior, $Q(z | Y, X)$, distributions of the variables in the latent feature space. Thus, the total loss function is a combination of the two losses as follows:

$$\mathcal{L}(Y, X) = \mathbb{E}_{z \sim Q(\cdot|Y, X)} [-\log P(Y | S(X, z))] + \beta \text{KL}[Q(z | Y, X) || P(z | X)] \quad (3)$$

The parameter β serves as a hyper-parameter that governs the extent to which the KL-divergence term, also known as the regularization term, influences the model's output. The model (illustrated in detail in Figure 2) is trained end-to-end. Hyper-parameter optimization was performed using logarithmic scaling from 10^{-6} to 10^{-3} and the optimum value for β is 0.0001.

2.3. Baseline Methods

To compare the performance of our proposed approach and evaluate its capabilities in generating consistent and contextual information, we developed two baseline methods with the similar stochastic nature. These baseline models are similar to the proposed model in capturing distributions over multi-modal segmentation. The introduced models are designed to accommodate the similarity in network architecture and to investigate the nature of stochasticity. By developing these baselines, we were able to analyze the effect of each stochasticity approach on (1) learning the underlying distribution of the wildfires, and (2) the performance of the network architectures under similar conditions. The baselines will shed light on the efficacy of the different stochasticity similar varieties of U-Net.

2.3.1. U-Net with Dropout

Dropout in a U-Net architecture can perform as a special case of the delta rule in which we introduce noise in the transmission of information [38] by randomly masking weights of the network. Dropout is presented as an especial case of delta rule called stochastic delta rule [39] in which each weight in the model is assigned as a random variable from a Gaussian distribution with the mean $\mu_{w_{ij}}$ and standard deviation of $\sigma_{w_{ij}}$ [38]. Dropout, as a special case of stochastic delta rule, introduces a form of regularization that aids in escaping poor local minima. By randomly deactivating a subset of neurons during each training iteration, dropout prevents the network from relying too heavily on specific neurons or features. This selective deactivation encourages the remaining neurons to compensate and learn more robust representations, leading to a broader exploration of the weight space and increasing the odds of finding the optimum solution [38]. Additionally, keeping the dropout in the inference process will introduce stochasticity by generating results from a randomly selected sub-network and will result in an approximation of posterior distribution [40]. Dropout obtains these advantages by removing hidden neurons according to a Bernoulli distribution with a probability parameter p . The dropout probability of the baseline model is set to be $p = 0.3$ meaning that at each pass of the network, only 70% of the neurons will be activated via a random selection.

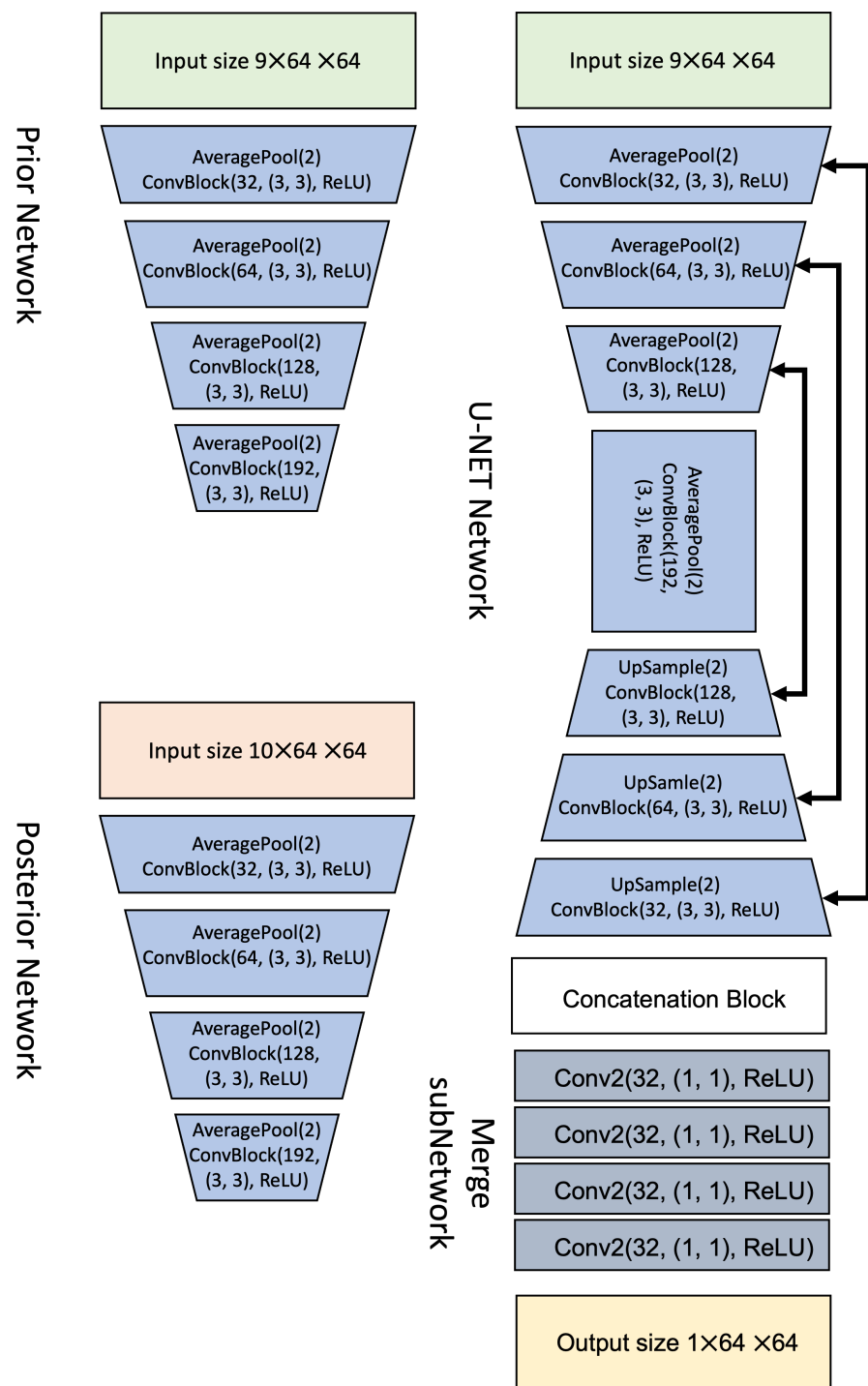


Figure 2. Network Architecture of Probabilistic U-Net consisting of three main networks: U-Net (plus a merge sub-network), prior network, and posterior network. The Convolution Blocks in the blue hidden layers consist of three sub-blocks of 2D convolutional layers with the same size features and kernel size and ReLU activation. The darker blue layers in U-Net represent the merger sub-network where the samples from latent distribution are concatenated with U-Net output and flow through the merger network to generate wildfire masks.

2.3.2. U-Net with Stochastic Activations

The concept of stochastic non-linear activations was first proposed by [41] to improve models by resolving the degenerative behavior of deterministic activation functions. Another study by Shridhar et al., 2020 [42] introduced a probabilistic activation definition

which makes the model behavior stochastic. The activation function, regardless of its type, will gain stochasticity by introducing Gaussian noise to its value [42]. In this architecture, instead of using a deterministic activation (e.g., ReLU), a Gaussian noise trick will apply perturbation to the forward and backward processes Figure 3. The parameters of the Gaussian perturbation can stay fixed or trained as a trainable parameter via backpropagation. Obtained from several experiments, the optimum sigma is found to be 5.

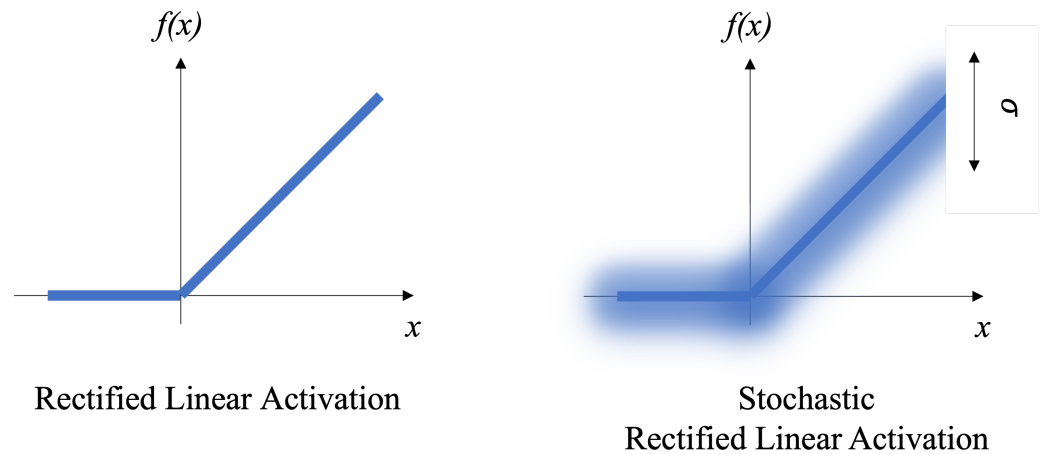


Figure 3. Rectified Linear Activation (ReLU) and stochastic ReLU. σ represents the standard deviation of the Gaussian noise.

2.4. Statistical Metrics

We have used multiple statistical metrics to evaluate the segmentation quality and assign lower and higher bounds for multiple draws of the same events. In particular, we used

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1\text{-score} = \frac{2 TP}{2 TP + FP + FN} \quad (6)$$

$$Jaccard\ Index = \frac{TP}{TP + FP + FN} \quad (7)$$

where TP , TN , FP , and FN are the true positive, true negative, false positive, and false negative, respectively. Precision, Recall, and F1-score are popular metrics that provide valuable Jaccard Index, also known as Intersect of Union (IoU), which is a good measure to calculate the overlap of predicted and target wildfire segmentation. Due to the stochasticity of each sample, we represented the statistics with lower and upper bounds of performance for each metric.

3. Experiments

In this section, we present the segmentation performance of the proposed method along with the two baselines, but first, we introduce the dataset used in this study and then explain the statistical metrics used in comparisons.

3.1. Dataset

In this study, we focus on the discrepancies in the fire products of MODIS constellation and VIIRS instruments onboard the joint NASA/NOAA Suomi National Polar-Orbiting

Partnership (Suomi NPP) and NOAA-20 satellites [43]. We aim to frame this problem to (1) offer an alternative fire product to resolve the MODIS' patchy and inconsistent segmentation, and (2) develop a distribution-over-event-based model to obtain epistemic uncertainty quantification and run what-if scenarios on input variables. For this purpose, we have collected MODIS MCD43A4, a daily product with 250 m spatial resolution, and collocated VIIRS fire product, with a daily 375 m spatial resolution, as target data. We used the Land/Cloud/Aerosol boundaries and properties channels with bandwidths of 620–670, 841–876, 459–479, 545–565, 1230–1250, 1628–1652, 2105–2155 nanometer (Table 1). We added the Normalized Difference Vegetation Index (NDVI) as a reliable proxy for estimating the fuel loads available for fires [44,45] using the following equation:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (8)$$

The NDVI ranges from -1 (not vegetation) to 1 (healthy vegetation) and is obtained from near-infrared (841–876 nm) and red (620–670 nm) bands. Multiple alternatives to NDVI have tried to address some of its issues using additional band [46,47]; however, due to the less noise sensitivity of NDVI and wide application of NDVI in the literature [48–51], NDVI has been considered to be the reference index for fuel-load analysis. Despite the usefulness of NDVI, it cannot be useful directly for fire detection due to the location dependency of NDVI values. For instance, NDVI values can be lower in arid zones compared to subtropical regions, but still, wildfires happen in subtropical regions due to abnormally low vegetation moisture. To tackle this issue, relative NDVI is calculated by subtracting the NDVI of each day from the mean NDVI of the same location for the whole period of study. This will give us a sense of abnormal vegetation conditions potent for wildfires.

Table 1. List of data used as inputs in the model.

Input	Bandwidth
Land/Cloud/Aerosols Boundary	620–670
	841–876
Land/Cloud/Aerosols Properties	459–79
	545–565
	1230–1250
	1628–1652
	2105–2155
NDVI	N/A
NDVI Derivation	N/A

The target fire dataset is obtained from thermal anomalies/active fire products with two fire-associated properties; brightness temperatures (in Kelvin), and fire radiative power (in Megawatts) among others. The dataset is provided in individual point locations with a spatial resolution of 375 m which is converted into gridded maps using the nearest neighbor method.

The training, validation, and testing sets consist of patches of data described above over wildfire events detected across the Continental United States. We shifted patches randomly to generate augmented patches and prevent the artifact of always having wildfire pixels in the center of the patch. The data were collected and patched for 2018 and filtered to only keep events with more than 20 pixels of wildfire inside. We used a 60–20–20 percentage for training, validation, and testing to obtain the hyper-parameter values. Then we retrained the models using training and validation sets and then evaluated the unbiased estimate of the performance using the testing set.

To generate multiple inference segmentations from the same input data, we feed the inputs to the U-NET model to obtain relevant spatial features. Simultaneously the

inputs are fed into the prior network and obtain latent space samples (z in Figure 1b). Combining the U-NET features with each sample z will provide a unique variation of corresponding segmentation. Multiple samples drawn from prior networks will provide multiple segmentations for that specific event.

3.2. Results

Throughout this section, we focus on evaluating the proposed Probabilistic U-Net model and compare the performances to the two baseline models: Dropout U-Net and Stochastic ReLU U-Net. We will first present a visual comparison benchmark for wildfire detection and quantify the visual uncertainty, and then present a more comprehensive performance using the metrics discussed in Section 3.3. Figure 4 consists of two independent wildfire incidents that describe two different wildfire dynamics. Each incident ((a) and (b)) demonstrates the visual consistency of the Probabilistic U-Net and the two baselines by drawing five random samples for a specific event. In Figure 4a, first five columns from left for the first, second, and third rows present the samples from Dropout U-Net, Stochastic ReLU U-Net, and Probabilistic U-Net models, respectively. Overall, all samples from all models are consistent with the target segmentation (last row, far left column). It is noticed that Dropout U-Net has less spatial coherency compared to the other two. Stochastic ReLU U-Net detects consistent wildfire in the circular area but misses the bottom left region of fire. The Probabilistic U-Net on the other hand shows a diverse range of detections capturing both patches of circular and bottom left fires. Comparing the detected wildfires by all models with NDVI indicates that all models understand the dynamics of vegetation and wildfire, where fire spreads in the surrounding of low vegetation (burned area). The NDVI deviation from the historical mean is very similar to the current NDVI in terms of burned area shape and size meaning the NDVI has not significantly changed, but the region is still experiencing wildfire activities. The far right column illustrates the spatial stochasticity of each model from 1000 independent samples. The Dropout U-Net model demonstrates low confidence in the bottom left region and left semi-circle of the circular region. Stochastic ReLU U-Net is confident in its detections and does not anticipate any fire in the bottom left region. The Probabilistic U-Net model produces a reasonable uncertainty map, covering most of the observed region with high confidence; however, the model is uncertain about the wildfire shape, specifically in the bottom left region.

Figure 4b demonstrates second independent incident where, similar to Figure 4a, the first, second and third rows belong to Dropout U-Net, Stochastic U-Net and Probabilistic U-Net, respectively. Performing similar to Figure 4a, all the segmentations are consistently close to target mask (bottom row, left column). Dropout U-Net presents higher variability compared to the other two models and result in higher uncertainty, especially in the wildfire border areas. Stochastic U-Net detects a more consistent segmentation pattern with less variability. It is noteworthy that Stochastic U-Net segmentations are undercomplete and do not fully cover the target segmentation area. Probabilistic U-Net demonstrates coherent patterns as target data, with uncertainty in the boundary of burning regions. The incident is slightly different in dynamics compared to Figure 4 due to NDVI behavior. In this incident, the NDVI deviation from historical mean is different, meaning the area is losing vegetation health quality due to the wildfire.

Figure 5 indicates the statistical performance of the three models over 1000 runs. We present the model performances for the testing set (including 1500 non-overlapping wildfire events) in box plots to incorporate the uncertainty level for each model. The precision statistics show similar sentiment to the visual samples, where Dropout U-Net under-detects the fire pixels (causing lower True Positives) with high variability, Stochastic ReLU U-Net detects a significant area of wildfire with high confidence and Probabilistic U-Net that has moderate detection capability with a similar range of Dropout U-Net variability. However, the results shift in the recall, showing a higher range for Probabilistic U-Net compared to the baselines. This stems from the lower FN values of this model, compared to the other two baselines. As a result of this, the F1-score which is a harmonic mean of precision and

recall shows higher performance for the proposed Probabilistic U-Net compared to the baselines. Lastly, the Jaccard index or IOU indicates higher agreement between the target segmentations and the Probabilistic U-Net segmentation variants. Stochastic ReLU U-Net is the second-best model with low variability, and the lowest IOU belongs to the Dropout U-Net model.

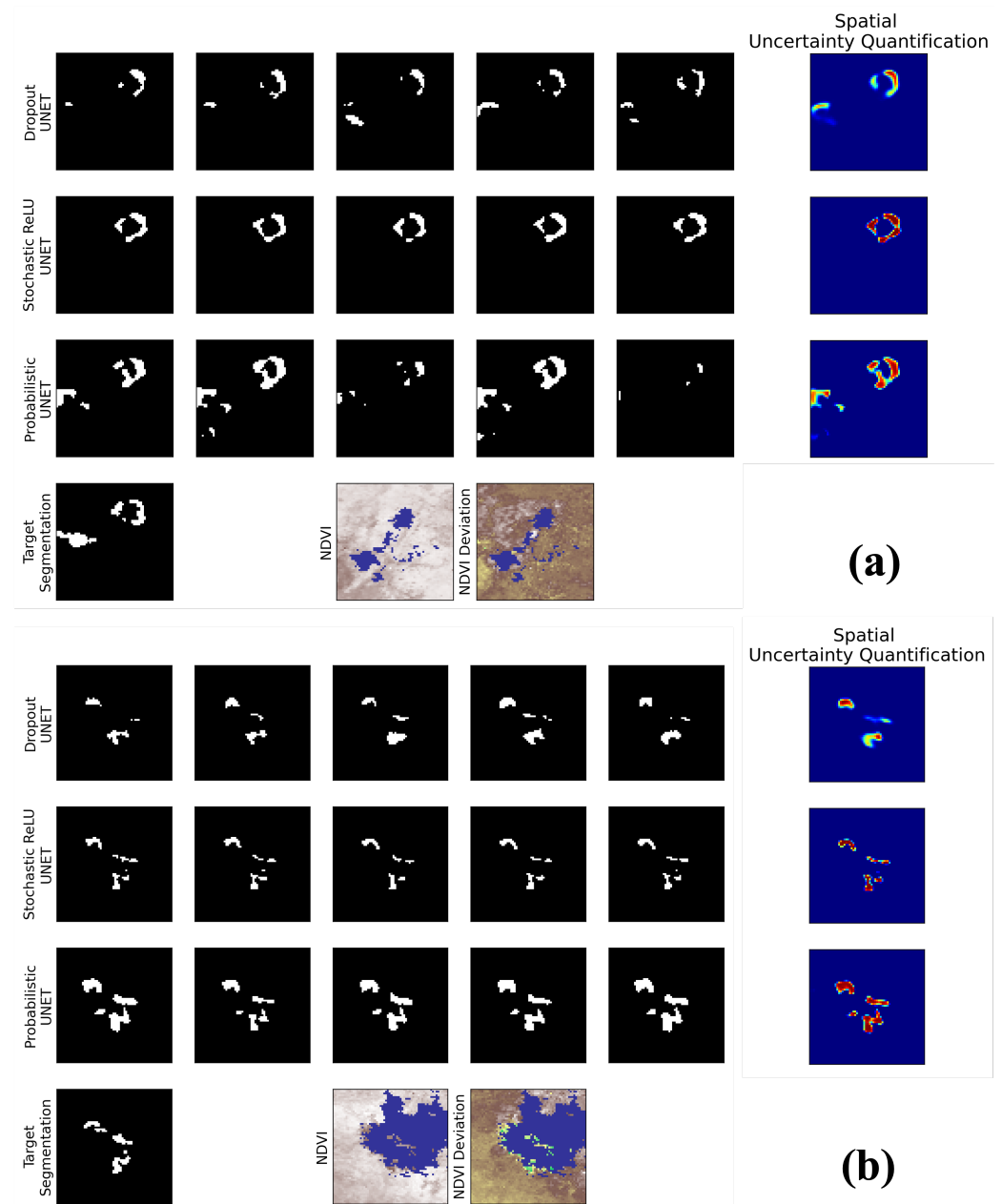


Figure 4. The figure demonstrates two independent wildfire incidents (subplot (a,b)) consisting of 5 drawn samples (first 5 columns) from the proposed Prob. U-Net (first row) and the baseline models (second and third rows) along with spatial uncertainty quantification for the same event using 1000 runs (last column). The last row, shows the target segmentation, corresponding NDVI, and NDVI deviation from historical, from left to right, respectively.

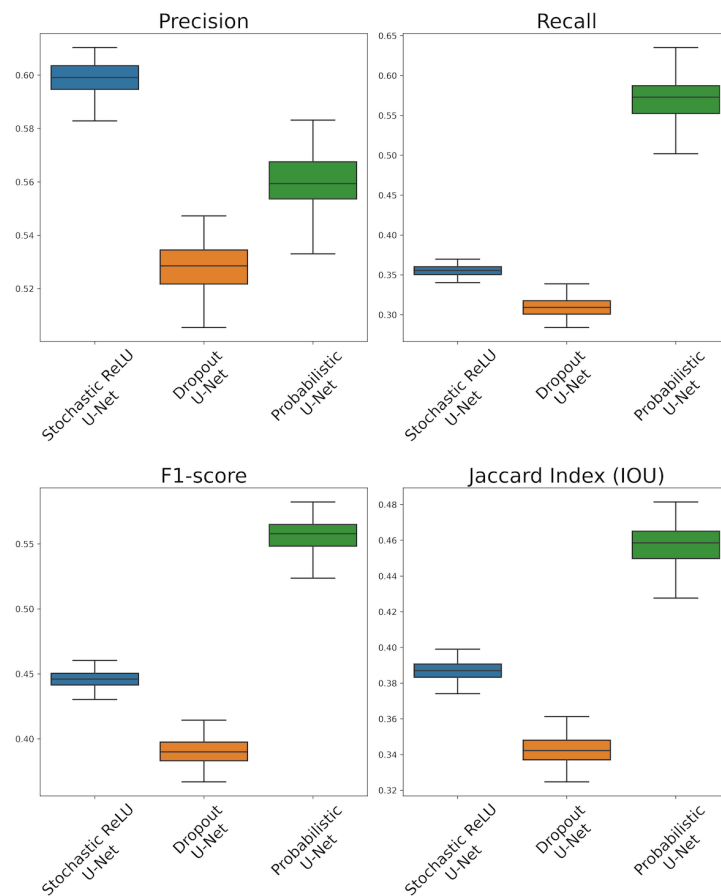


Figure 5. Statistical comparison between Probabilistic U-Net, Dropout U-Net, and Stochastic ReLU U-Net over 1000 runs.

3.3. Discussion

Furthermore, we investigated the semantic variation of the models by changing the dynamic of NDVI. In this experiment, we aimed to better understand the grasp of each model in understanding the NDVI dynamics. The experiments continued to investigate the physical comprehension of each model by changing NDVI and observing the changes in wildfire segmentations. We follow the following reasonings: (1) an increase in NDVI will not trigger wildfire (at least not as severe as before), (2) a spotty decrease in NDVI allows the wildfire to spread toward lower NDVI (unhealthy vegetation) area, (3) a significant decrease in NDVI in a region will not provide enough fuel for the fire to spread. We tested these hypotheses on the sample data we had in Figure 6. In the first three rows, we have the model detections from original NDVI values which are similar to the samples shown in Figure 4. The second three rows demonstrate the model responses to an increase of NDVI within and surrounding low NDVI area (burned region). Based on the results, we see that Dropout and Probabilistic U-Nets will not detect a burning segment and Stochastic ReLU U-Net will detect smaller segmentations. The third three rows investigate the idea of sparsely lowering the vegetation in regions close to burning scars. The results show that Dropout U-Net and Stochastic ReLU U-Nets will not capture the ignitions toward new places, especially in the bottom left region. However, Probabilistic U-Net is understanding spread reasoning and detecting segments in the bottom left area. Lastly, we show significant NDVI reduction for a large area in the last three rows. The NDVI decrease mainly impacts the bottom left region and spotty locations in the circular segment region. All models are correctly ruling out the possibility of wildfire in the bottom left region. Dropout U-Net has difficulty understanding the circular shape affected by spotty NDVI decreases. Stochastic ReLU U-Net is persistently detecting the circular segment, but Probabilistic

U-Net has slightly adjusted the circular segment according to the spotty changes. It is noteworthy that the model hyperparameters (σ , β , and dropout rate for Stochastic ReLU, Probabilistic, and Dropout U-Nets) are selected based on best precision and recall. It seems that Stochastic ReLU U-Net performs better under lower variability and deteriorates under higher σ values.

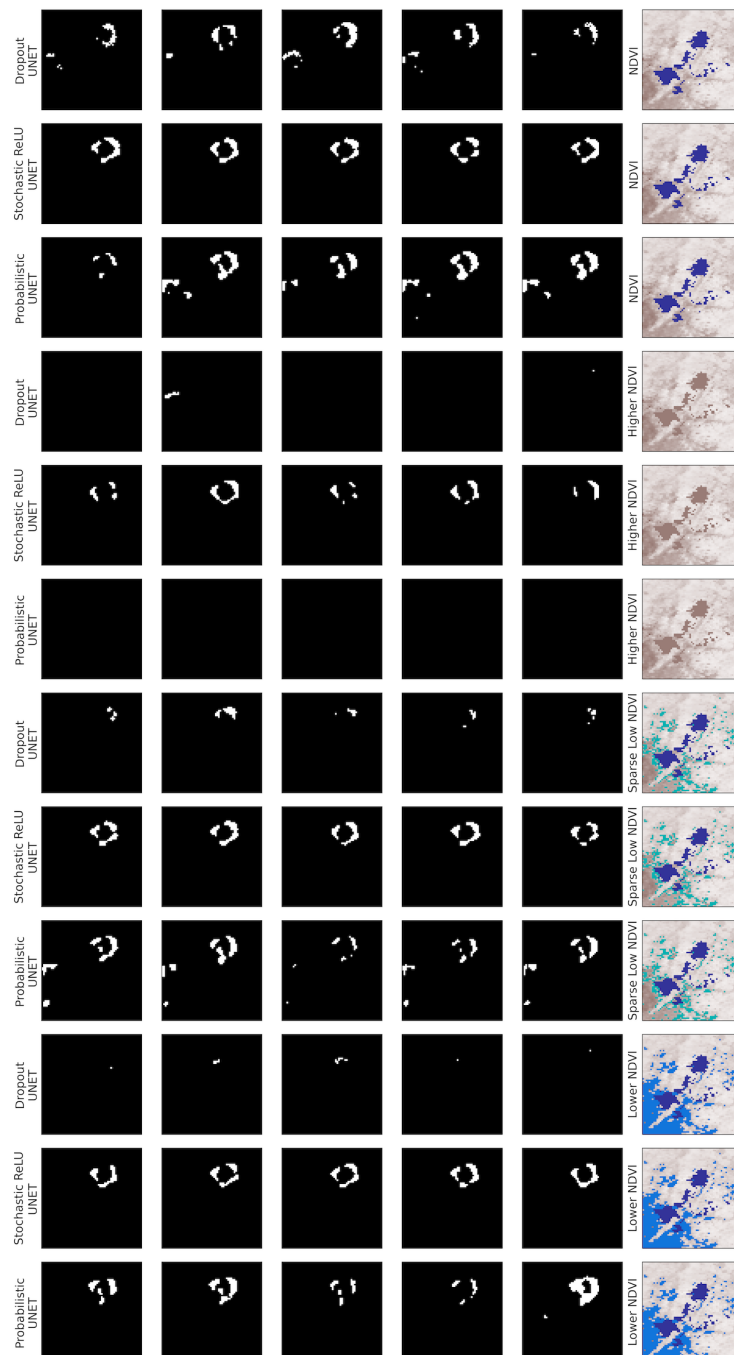


Figure 6. Empirical investigation of model comprehension from NDVI dynamics. The first three rows are the Stochastic ReLU, Dropout, and Probabilistic U-Net without a change in NDVI. The second three rows are the same order of models with greener NDVI within and in the surrounding of the bottom leaving a burned scar. The third three rows reduce NDVI sparsely, especially in the bottom left region. The last three rows present a significant decrease in NDVI in the vicinity of the bottom left region and spotty locations close to the circular scar.

4. Conclusions

In this study, we proposed a stochastic machine-learning approach that learns a latent distribution of wildfire events in a supervised manner and addresses the uncertainty quantification and inter-dataset discrepancies. We investigated the proposed method by segmenting active wildfires using the seven bands from MODIS and two derivatives (NDVI and historical deviation of NDVI) as inputs. The proposed model was compared with two stochastic baseline machine-learning models called Dropout U-NET, a U-NET with dropouts in training and test phases, and Stochastic ReLU U-NET, a U-NET with Stochastic ReLU activations. It was discovered through the conducted experiments that Probabilistic U-Net is more accurate and flexible compared to the other two models. The Stochastic ReLU U-Net seems to perform more accurately with lower variability, and Dropout U-Net is less accurate but demonstrates a wider range of variability. Additionally, we performed a scenario-based experiment to analyze the impact of physical changes on the response of the models. The probabilistic model showed a more comprehensive understanding of the physical relationship between NDVI and wildfire. However, the other two baseline models demonstrated partial alignment with the scenarios.

Author Contributions: Conceptualization, A.A.A., M.M. and P.A.L.; methodology, M.M. and A.A.A.; modeling and implementation, A.A.A.; formal analysis, A.A.A.; investigation, A.A.A.; data curation, A.A.A.; writing—original draft preparation, A.A.A. and M.M.; writing—review and editing, A.A.A., M.M., P.A.L., E.R. and S.G.; visualization, A.A.A.; supervision, P.A.L., E.R. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded and made possible by NASA ROSES AIST-QRS-21. A.A.A., M.M. and P.A.L. are thankful for the support from NASA Academic Mission Services, Contract No. NNA16BD14C.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Agee, J.K.; Finney, M.; Gouvenain, R.D. Forest fire history of desolation peak, Washington. *Can. J. For. Res.* **1990**, *20*, 350–356. [[CrossRef](#)]
2. Alizadeh, M.R.; Abatzoglou, J.T.; Luce, C.H.; Adamowski, J.F.; Farid, A.; Sadegh, M. Warming enabled upslope advance in western US forest fires. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2009717118. [[CrossRef](#)]
3. Barros, A. M.; Ager, A. A.; Day, M. A.; Krawchuk, M. A.; Spies, T. A. Wildfires managed for restoration enhance ecological resilience *Ecosphere* **2018**, *9*, e02161. [[CrossRef](#)]
4. Calkin, D.E.; Thompson, M.P.; Finney, M.A. Negative consequences of positive feedbacks in US wildfire management. *For. Ecosyst.* **2015**, *2*, 1–10. [[CrossRef](#)]
5. Chas-Amil, M.L.; Prestemon, J.P.; McClean, C.J.; Touza, J. Human-ignited wildfire patterns and responses to policy shifts. *Appl. Geogr.* **2015**, *56*, 164–176. [[CrossRef](#)]
6. Dennison, P.E.; Brewer, S.C.; Arnold, J.D.; Moritz, M.A. Large wildfire trends in the western United States, 1984–2011. *Geophys. Res. Lett.* **2014**, *41*, 2928–2933. [[CrossRef](#)]
7. Hoover, K.; Hanson, L.A. *Wildfire Statistics*; Technical Report; Congressional Research Service: Washington, DC, USA, 2021.
8. Giglio, L.; Boschetti, L.; Roy, D.P.; Humber, M.L.; Justice, C.O. The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sens. Environ.* **2018**, *217*, 72–85. [[CrossRef](#)]
9. Jain, P.; Coogan, S.C.; Subramanian, S.G.; Crowley, M.; Taylor, S.; Flannigan, M.D. A review of machine learning applications in wildfire science and management. *Environ. Rev.* **2020**, *28*, 478–505. [[CrossRef](#)]
10. Carmo, M.; Moreira, F.; Casimiro, P.; Vaz, P. Land use and topography influences on wildfire occurrence in northern Portugal. *Landsc. Urban Plan.* **2011**, *100*, 169–176. [[CrossRef](#)]
11. Narayanaraj, G.; Wimberly, M.C. Influences of forest roads on the spatial patterns of human-and lightning-caused wildfire ignitions. *Appl. Geogr.* **2012**, *32*, 878–888. [[CrossRef](#)]
12. Ghali, R.; Akhloufi, M.A.; Jmal, M.; Souidene Mseddi, W.; Attia, R. Wildfire segmentation using deep vision transformers. *Remote Sens.* **2021**, *13*, 3527. [[CrossRef](#)]
13. Green, M.E. *Some Results on a Set of Data Driven Stochastic Wildfire Models*; The University of Vermont and State Agricultural College: Burlington, VT, USA, 2020.
14. Khryashchev, V.; Larionov, R. Wildfire segmentation on satellite images using deep learning. In Proceedings of the 2020 Moscow Workshop on Electronic and Networking Technologies (MWENT), Moscow, Russia, 11–13 March 2020; pp. 1–5.
15. Rashkovetsky, D.; Mauracher, F.; Langer, M.; Schmitt, M. Wildfire detection from multisensor satellite imagery using deep semantic segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7001–7016. [[CrossRef](#)]

16. Poole, D.; Raftery, A.E. Inference for deterministic simulation models: The Bayesian melding approach. *J. Am. Stat. Assoc.* **2000**, *95*, 1244–1255. [[CrossRef](#)]
17. Cencerrado, A.; Cortés, A.; Margalef, T. Response time assessment in forest fire spread simulation: An integrated methodology for efficient exploitation of available prediction time. *Environ. Model. Softw.* **2014**, *54*, 153–164. [[CrossRef](#)]
18. Toan, N.T.; Cong, P.T.; Hung, N.Q.V.; Jo, J. A deep learning approach for early wildfire detection from hyperspectral satellite images. In Proceedings of the 2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA), Daejeon, Republic of Korea, 1–3 November 2019; pp. 38–45.
19. Sayad, Y.O.; Mousannif, H.; Al Moatassime, H. Predictive modeling of wildfires: A new dataset and machine learning approach. *Fire Saf. J.* **2019**, *104*, 130–146. [[CrossRef](#)]
20. Couce, E.; Knorr, W. Statistical parameter estimation for a cellular automata wildfire model based on satellite observations. *WIT Trans. Ecol. Environ.* **2010**, *137*, 47–55.
21. Quill, R.; Sharples, J.J.; Wagenbrenner, N.S.; Sidhu, L.A.; Forthofer, J.M. Modeling wind direction distributions using a diagnostic model in the context of probabilistic fire spread prediction. *Front. Mech. Eng.* **2019**, *5*, 5. [[CrossRef](#)]
22. Palmer, T.; Shutts, G.; Hagedorn, R.; Doblas-Reyes, F.; Jung, T.; Leutbecher, M. Representing model uncertainty in weather and climate prediction. *Annu. Rev. Earth Planet. Sci.* **2005**, *33*, 163–193. [[CrossRef](#)]
23. Artés, T.; Cencerrado, A.; Cortés, A.; Margalef, T. Time aware genetic algorithm for forest fire propagation prediction: exploiting multi-core platforms. *Concurr. Comput. Pract. Exp.* **2017**, *29*, e3837. [[CrossRef](#)]
24. Denham, M.; Laneri, K. Using efficient parallelization in graphic processing units to parameterize stochastic fire propagation models. *J. Comput. Sci.* **2018**, *25*, 76–88. [[CrossRef](#)]
25. Ramirez, J.; Monedero, S.; Silva, C.A.; Cardil, A. Stochastic decision trigger modelling to assess the probability of wildland fire impact. *Sci. Total Environ.* **2019**, *694*, 133505. [[CrossRef](#)]
26. Valero, M. M.; Jofre, L.; Torres, R. Multifidelity prediction in wildfire spread simulation: Modeling, uncertainty quantification and sensitivity analysis. *Environ. Model. Softw.* **2021**, *141*, 105050. [[CrossRef](#)]
27. Li, F.; Zhang, X.; Kondragunta, S.; Csiszar, I. Comparison of fire radiative power estimates from VIIRS and MODIS observations. *J. Geophys. Res. Atmos.* **2018**, *123*, 4545–4563. [[CrossRef](#)]
28. Allison, R.S.; Johnston, J.M.; Craig, G.; Jennings, S. Airborne optical and thermal remote sensing for wildfire detection and monitoring. *Sensors* **2016**, *16*, 1310. [[CrossRef](#)]
29. Schultz, C.J.; Nauslar, N.J.; Wachter, J.B.; Hain, C.R.; Bell, J.R. Spatial, temporal and electrical characteristics of lightning in reported lightning-initiated wildfire events. *Fire* **2019**, *2*, 18. [[CrossRef](#)]
30. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [[CrossRef](#)]
31. Akbari Asanjan, A.; Das, K.; Li, A.; Chirayath, V.; Torres-Perez, J.; Sorooshian, S. Learning instrument invariant characteristics for generating high-resolution global coral reef maps. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual Event, 6–10 July 2020; pp. 2617–2624.
32. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Springer International Publishing: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
33. Shu, R.; Brofos, J.; Zhang, F.; Bui, H.H.; Ghavamzadeh, M.; Kochenderfer, M. Stochastic video prediction with conditional density estimation. In Proceedings of the ECCV Workshop on Action and Anticipation for Visual Learning, Amsterdam, The Netherlands, 8–10 October 2016; Volume 2, p. 2.
34. Kohl, S.; Romera-Parades, B.; Meyer, C.; De Fauw, J.; Ledsam, J.; Maier-Hein, K.; Eslami, S.; Rezende, D.; Ronneberger, O. A Probabilistic U-Net for Segmentation of Ambiguous Images. *NeurIPS* **2018**, *31*, 6965–6975.
35. Kingma, D.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**, arXiv:1312.6114.
36. Kingma, D.; Rezende, D.; Mohamed, S.; Welling, M. Semi-supervised learning with deep generative models. *Adv. Neural Inf. Process. Syst. (NeurIPS)* **2014**, *27*.
37. Sohn, K.; Lee, H.; Yan, X. Learning structured output representation using deep conditional generative models. *Adv. Neural Inf. Process. Syst. (NeurIPS)* **2015**, *28*, 3483–3491.
38. Frazier-Logue, N.; Hanson, S.J. Dropout is a special case of the stochastic delta rule: Faster and more accurate deep learning. *arXiv* **2018**, arXiv:1808.03578.
39. Hanson, S.J. A stochastic version of the delta rule. *Phys. D Nonlinear Phenom.* **1990**, *42*, 265–272. [[CrossRef](#)]
40. Kendall, A.; Badrinarayanan, V.; Cipolla, R. Bayesian segnet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding. *arXiv* **2015**, arXiv:1511.02680.
41. Gulcehre, C.; Moczulski, M.; Denil, M.; Bengio, Y. Noisy activation functions. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 20–22 June 2016; pp. 3059–3068.
42. Shridhar, K.; Lee, J.; Hayashi, H.; Mehta, P.; Iwana, B.K.; Kang, S.; Uchida, S.; Ahmed, S.; Dengel, A. Probact: A probabilistic activation function for deep neural networks. *arXiv* **2019**, arXiv:1905.10761.
43. Schroeder, W.; Oliva, P.; Giglio, L.; Csiszar, I.A. The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sens. Environ.* **2014**, *143*, 85–96. [[CrossRef](#)]

44. Liu, Y.; Stanturf, J.; Goodrick, S. Wildfire potential evaluation during a drought event with a regional climate model and NDVI. *Ecol. Inform.* **2010**, *5*, 418–428. [[CrossRef](#)]
45. Dasgupta, S.; Qu, J.J.; Hao, X.; Bhoi, S. Evaluating remotely sensed live fuel moisture estimations for fire behavior predictions in Georgia, USA. *Remote Sens. Environ.* **2007**, *108*, 138–150. [[CrossRef](#)]
46. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
47. Matsushita, B.; Yang, W.; Chen, J.; Onda, Y.; Qiu, G. Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: A case study in high-density cypress forest. *Sensors* **2007**, *7*, 2636–2651. [[CrossRef](#)]
48. Pereira-Pires, J.E.; Aubard, V.; Ribeiro, R.A.; Fonseca, J.M.; Silva, J.M.; Mora, A. Fuel Break Vegetation Monitoring with Sentinel-2 NDVI Robust to Phenology and Environmental Conditions. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 6264–6267.
49. Mazeh, F.; El Sahili, J.; Zaraket, H. Low-Cost NDVI Platform for Land Operation: Passive and Active. *IEEE Sens. Lett.* **2021**, *5*, 1–4. [[CrossRef](#)]
50. Quan, X.; Xie, Q.; He, B.; Luo, K.; Liu, X. Corrigendum to: Integrating remotely sensed fuel variables into wildfire danger assessment for China. *Int. J. Wildland Fire* **2021**, *30*, 822–822. [[CrossRef](#)]
51. Holsinger, L.; Parks, S.A.; Miller, C. Weather, fuels, and topography impede wildland fire spread in western US landscapes. *For. Ecol. Manag.* **2016**, *380*, 59–69. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.