QUANTUM-COMPATIBLE VARIATIONAL SEGMENTATION FOR IMAGE-TO-IMAGE WILDFIRE DETECTION USING SATELLITE DATA

Ata Akbari Asanjan[2,3], *Milad Memarzadeh*[2,3], *P. Aaron Lott*[1,3], *Thomas Templin*[2], *Eleanor Rieffel*[1]

(1) Quantum Artificial Intelligence Laboratory (QuAIL), NASA Ames Research Center
(2) Data Science Group (DSG), NASA Ames Research Center
(3) USRA Research Institute for Advanced Computer Science (RIACS)

ABSTRACT

Wildfire occurrences have been increasing for the past decade, leaving devastating traces across the world. In the recent efforts, remote sensing and airborne missions have been utilized to better understand and manage wildfires. This has resulted in an exponential increase in volume of remote sensing data, which has pushed the need for intelligent automation of data extraction for wildfire studies. Machine learning offers accurate automation in detecting such natural anomalies and enable decision-makers to take actions in a timely manner. Recent advances in machine learning algorithms, namely probabilistic generative methods, allow researchers and decision-makers to step beyond detection and study "what-if" scenarios for wildfire occurrences. Additionally, they offer better imitations to the stochastic behavior of nature, and wildfire events. However, optimizing the performance of these probabilistic generative models is a computationally expensive process, specially using digital computers. On the other hand, quantum computers have recently shown a promise to reduce computationally costly training of such models and provide performance improvements. There is a body of research investigating the potential for improved machine learning methods in which key operations are performed on a quantum computer. In this study, we propose a probabilistic image-to-image segmentation approach combining a very well-known segmentation method, U-NET, with a Conditional Variational Auto-Encoder (CVAE) to not only detect wildfires but also describe the stochasticity of the phenomenon and be capable of running "what-if" scenarios. Our proposed model is compatible with training on quantum computers, which results in a quantum-assisted image-to-image segmentation approach and can be used to benchmark the potential benefit of quantum computing over the classical one.

Index Terms— Wildfire Detection, Remote Sensing, Image-to-Image Segmentation, Quantum Machine Learning

1. INTRODUCTION

Wildfires are an essential part of terrestrial ecosystem and provide significant ecological benefits [1, 2]. However, the fire intensities, sizes, frequencies have increased across the world, with an exceptional impact on the western United States in the past decades [3, 2]. Proper resource management and timely decision-making require accurate monitoring and understanding of the wildfire nature. With the advent of terrestrial and atmospheric remote sensing, mainly supported by satellite and aviation vessels, the means to monitor and detect wildfires have been more accessible [4]. Advances in observation sensors, specifically 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 [5, 4]. However, with the increase in the number of satellites/aviation missions, and the enhancement of spatial, temporal, and spectral resolution, efficient and effective land management through remote sensing has been challenging [6]. Automated intelligent approaches such as machine learning propose an opportunity to extract useful information from a large volume of remote sensing datasets. From the spectrum of machine learning approaches, variational methods outshine others for wildfire applications due to their capabilities in learning stochastic behaviors [7], such as ones in wildfire processes. Recent advances in quantum computations allow classical machine learning models to be trained on quantum computers, with the potential for significant advantages in computing costly operations and simulating stochastic behavior [8]. However, optimizing the performance of such models is a computationally expensive process, specially using digital computers. Quantum computers on the other hand, make use of quantum effects such as quantum superposition, entanglement, and interference to compute in ways that have no classical analog. Quantum algorithms making use of these effects have been proven to outperform classical algorithms in a variety of settings, including for certain sampling problems [9], though much work remains to be done to understand how best to take advantage of these effects. In this study, the authors propose a quantum-compatible variational machine learning approach that takes images of lightning-based wildfire predictors and segments wildfire events. Replacing components of this approach with quantum compatible computations makes the quantum-

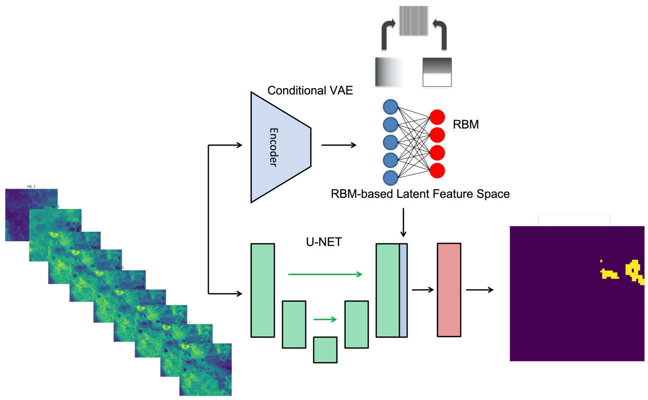

Fig. 1. Graphical illustration of the proposed probabilistic U-NET framework. The inputs are NDVI, NDVI difference with long-term NDVI and MODIS MCD43A4 channels for Land/Cloud/Aerosols. The model uses an RBM to represent the latent space as a Boltzmann distribution.

assisted machine learning possible and can benchmark its superiority to classical computers.

2. PROPOSED FRAMEWORK

Image segmentation refers to the task of identifying and segmenting objects/phenomena of interest in the input image. One of the popular methods for instance segmentation is U-NET [10], originally developed for biomedical image segmentation problems and adopted to many domains including space exploration and Earth sciences. U-NET is an image-to-image translation deep neural network with a convolutional architecture that takes an image as input and outputs the segmentation map of the image. It is trained in a supervised fashion, meaning that accurate segmented images need to be provided for training the model to perform such mapping. Although U-NET has shown significant performance in image segmentation, one of the drawbacks of the framework is its deterministic nature. The mapping of the input image to the output segmentation map is fully deterministic and does not incorporate sources of uncertainty and stochasticity into account. This can cause the model to overfit to the training data and generalize poorly to an unseen regions, and also cannot be used to perform "what-if" scenarios and provide probabilistic segmentation.

Kohl et al. [11] extends the U-NET framework and proposes a probabilistic U-NET model for image segmentation. They achieve this by combining the U-NET with a Conditional Variational Auto-Encoder (CVAE) [12, 13] that allows the model to produce plausible hypotheses and "what-if" scenarios. Figure 1 shows the overall architecture of the proposed model. As it can be seen, the segmentation generation of the U-NET model is conditioned on the sample obtained from the latent feature space of the VAE. This low-dimensional latent feature space encodes the possible segmentation variants and can be used to evaluate "what-if" scenarios in the evaluation mode (once the model is trained). This addition allows the model to generate multiple segmentation maps for a single input image, conditioned on the region of the latent feature space that is sampled. As suggested by the authors, this capability allows the model *"to learn hypotheses that have a low probability and to predict them with the corresponding frequency"*.

Once the output of the U-NET (the green block) and the sampled latent variable (blue block) are concatenated, a function, $\mathcal{F}$ (red block), generates a segmentation for each sample. i.e., $S_i = \mathcal{F}\left(f_{\text{U-NET}}\left(X, \theta\right), z_i; \psi\right)$, where S_i is the segmentation corresponding to the sampled latent variable z_i, and θ and ψ are the parameters of the U-NET model and the generation function, respectively. The model is trained based on two objectives: (1) generating accurate segmentation for detecting wildfire in the input image, and (2) generalize well to unseen or rare scenarios. The first objective is obtained by minimizing a supervised cross-entropy loss between the generated segmentation, $S(X, z)$, and the ground truth, Y, and the second objective is obtained by minimizing the KL-divergence between the prior, $P(z \mid X)$, and posterior, $Q(z \mid Y, X)$, distributions of the variables in the latent feature space. The overall loss function for training the machine learning model is defined as follow,

$$\begin{aligned}\mathcal{L}\left(Y, X\right) = &\mathbb{E}_{z\sim Q(.|Y,X)}\left[-\log P\left(Y \mid S\left(X, z\right)\right)\right] + \\ &\beta \mathrm{KL}\left[Q\left(x \mid Y, X\right) || P\left(z \mid X\right)\right]\end{aligned} \tag{1}$$

Where, β is a hyper-parameter controlling the effect of KL-divergence term (i.e., the regularization term). The model is trained end-to-end. In order to utilize the proposed method in a quantum-compatible environment, we propose a discrete latent feature space parameterized by an energy-based model such as a Quantum-compatible Restricted Boltzmann Machine (RBM) (Figure 1). The RBM model with binary units in visible and hidden variables represents an Ising model with the following energy function,

$$E_z = -\sum_a b_a z_a - \sum_{a,b} w_{ab} z_a z_b \tag{2}$$

Where, b_a and w_{ab} are the RBM parameters and z_a and z_b are the binary units. The goal for training the RBM is to find the Hamiltonian parameters (i.e. the weights of RBM) optimally for resembling the probability distribution of the real latent feature space. We aim to achieve this objective by minimizing the average negative log-likelihood

$$\mathcal{L}_{RBM} = -\sum_v P_v(z) \log \frac{\sum_h e^{-E_z}}{\sum_{z'} e^{E_{z'}}} \tag{3}$$

Where, v and h represent states of visible and hidden variables. The sampling scheme used in training these hidden variables is a Monte Carlo simulation which can be either quantum-assisted or a quantum-inspired algorithm such as parallel tempering [14, 15]. The quantum-oriented improvements, introduced by [16, 17] demonstrate state-of-the-art performance and efficient model parameterizations that can leverage near-term quantum architectures in training latent discrete variables within generative models.

3. EXPERIMENTS

In this section, we present U-Net, Probabilistic U-Net with Gaussian latent space (as proposed in [11]), Probabilistic U-Net with Bernoulli latent space as baselines. We also present the proposed Probabilistic U-Net with RBM latent space model, and compare it with the baseline models.

3.1. Training and Implementation

We trained all the models using the same architecture for U-Net sub-model with three convolutional blocks, each consisting of three convolutional layers with ReLU activation followed by a Average Pooling layer, in the encoder part. The bottleneck layer consists of three convolutional layers with ReLU activation, and decoder contains three convolutional blocks, each consisting of three convolutional layers with ReLU activation followed by a nearest-neighborhood upsampling layer. The feature size for the encoder, bottleneck, decoder blocks are 32, 64, 128, 192, 128, 64, 32, respectively. In the probabilistic models, the prior and posterior networks have identical architecture except the final layer of distribution. The prior and posterior networks have the same architecture as the encoder plus bottleneck in the U-Net. The latent space of the probabilistic models consist of 32 latent units. The U-Net model is trained by satisfying three losses at the same time: a weighted binary cross-entropy for pixel-wise detection with 750-to-1 ratio which significantly penalizes missing the fire pixels, a perceptual loss for encouraging neighborhood similarity, and a F-2 score loss to further penalize the false positives and false negatives in the binary predictions. Similar combination is deployed in probabilistic models where the perceptual loss and F-2 score loss is used together with the Evidence Lower Bound (ELBO). The Probabilistic models are trained using a β value of 10, and the Probabilistic U-Net with Bernoulli and with RBM are trained using a relaxed Bernoulli distribution with temperature of 0.5.

3.2. Results

Table 1 lists the performances of the baselines and the proposed model in terms of Precision, Recall, F-1 score and Jaccard score (which shows the intersect over union of fire pixels between groundtruth and predictions) on the test set. The results demonstrates higher performances from RBM U-Net compared to the other Bernoulli U-Net.

	U-Net	Prob. U-Net (Gaussian)	Prob. U-Net (Bernoulli)	Prob. U-Net (RBM)
Precision	0.536	0.431	0.235	0.654
Recall	0.987	0.955	0.752	0.473
F-1 score	0.695	0.594	0.358	0.549
Jaccard score	0.532	0.422	0.318	0.378

Table 1. Statistical performances for the baseline models and Prob. U-Net with RBM latent space.

Figure 2 shows a sample wildfire predicted by the models in this study. The Probabilistic models are sampled 4 times to better illustrate their stochastic nature. The visual comparison shows close to groundtruth predictions with small but observable variations for the probabilistic models.

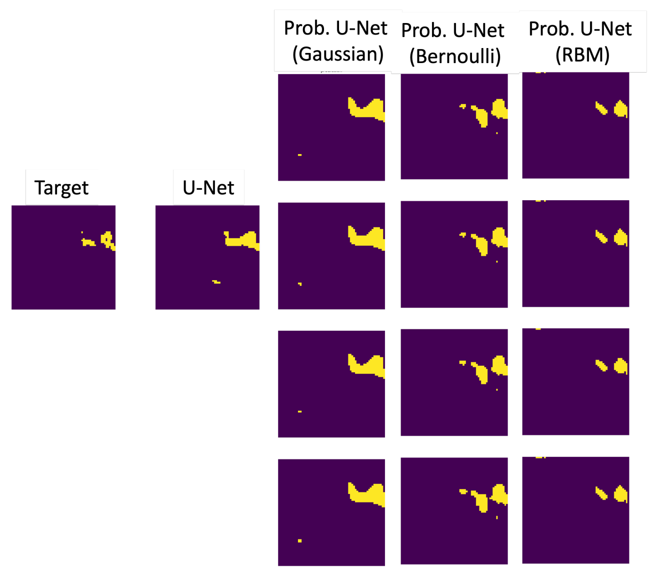

Fig. 2. Performance of the presented model for a sample wildfire detection.

4. CONCLUSION

In this study, we propose a variational image-to-image translation model with RBM latent space which opens the door for utilization of quantum computers in latent space sampling. The classical version of the model without quantum sampling demonstrates satisfying performances compared to comparable baselines and shows potential performance gains in stochastic image-to-image translation via quantum sampling. It is noteworthy that the presented results can still benefit from a hyper-parameter tuning for the β, temperature, RBM refining iterations which is part of our future work goals.

5. ACKNOWLEDGEMENT

This research is funded and made possible by NASA ROSES AIST-QRS-21. AAA, MM, PAL are thankful for support from NASA Academic Mission Services, Contract No. NNA16BD14C.

6. REFERENCES

[1] James K Agee, Mark Finney, and Roland De Gouvenain, "Forest fire history of desolation peak, washington," *Canadian Journal of Forest Research*, vol. 20, no. 3, pp. 350–356, 1990.

[2] Mohammad Reza Alizadeh, John T Abatzoglou, Charles H Luce, Jan F Adamowski, Arvin Farid, and Mojtaba Sadegh, "Warming enabled upslope advance in western us forest fires," *Proceedings of the National Academy of Sciences*, vol. 118, no. 22, 2021.

[3] Philip E Dennison, Simon C Brewer, James D Arnold, and Max A Moritz, "Large wildfire trends in the western united states, 1984–2011," *Geophysical Research Letters*, vol. 41, no. 8, pp. 2928–2933, 2014.

[4] Robert S Allison, Joshua M Johnston, Gregory Craig, and Sion Jennings, "Airborne optical and thermal remote sensing for wildfire detection and monitoring," *Sensors*, vol. 16, no. 8, pp. 1310, 2016.

[5] Christopher J Schultz, Nicholas J Nauslar, J Brent Wachter, Christopher R Hain, and Jordan R Bell, "Spatial, temporal and electrical characteristics of lightning in reported lightning-initiated wildfire events," *Fire*, vol. 2, no. 2, pp. 18, 2019.

[6] Dmitry Rashkovetsky, Florian Mauracher, Martin Langer, and Michael Schmitt, "Wildfire detection from multisensor satellite imagery using deep semantic segmentation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7001–7016, 2021.

[7] Rui Shu, James Brofos, Frank Zhang, Hung Hai Bui, Mohammad Ghavamzadeh, and Mykel Kochenderfer, "Stochastic video prediction with conditional density estimation," in *ECCV Workshop on Action and Anticipation for Visual Learning*, 2016, vol. 2, p. 2.

[8] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione, "An introduction to quantum machine learning," *Contemporary Physics*, vol. 56, no. 2, pp. 172–185, 2015.

[9] Austin P Lund, Michael J Bremner, and Timothy C Ralph, "Quantum sampling problems, boson sampling and quantum supremacy," *npj Quantum Information*, vol. 3, no. 1, pp. 1–8, 2017.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *in Proc. Int. Conf. Medical Image Comput. Comput.-Assisted Intervention*, pp. 234–241, 2015.

[11] S.A.A. Kohl, B. Romera-Parades, C. Meyer, J. De Fauw, J.R. Ledsam, K.H. Maier-Hein, S.M.A. Eslami, D.J. Rezende, and O. Ronneberger, "A probabilistic u-net for segmentation of ambiguous images," *in NeurIPS*, pp. 6965–6975, 2018.

[12] D.P. Kingma, D.J. Rezende, S. Mohamed, and M. Welling, "Semi-supervised learning with deep generative models," *Advances in Neural Information Processing Systems (NeurIPS)*, 2014.

[13] K. Sohn, H. Lee, and X. Yan, "Learning structured output representation using deep conditional generative models," *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 3483–3491, 2015.

[14] E. Marinari and G. Parisi, "Simulated tempering: A new monte carlo scheme," *Europhysics Letters*, vol. 19, no. 6, 1992.

[15] David J. Earlab and Michael W. Deema, "Parallel tempering: Theory, applications, and new perspectives," *Physical Chemistry Chemical Physics*, vol. 7, no. 23, 2005.

[16] Amir Khoshaman, Walter Vinci, Brandon Denis, Evgeny Andriyash, Hossein Sadeghi, and Mohammad H Amin, "Quantum variational autoencoder," *Quantum Science and Technology*, vol. 4, no. 1, pp. 014001, 2018.

[17] Hossein Sadeghi Amir Khoshaman3 Evgeny Andriyash Walter Winci, Lorenzo Buffoni and Mohammad H Amin, "A path towards quantum advantage in training deep generative models with quantum annealers," *Machine Learning Science and Technology*, vol. 1, no. 4.