Igniting Precision: Amplifying Wildfire Prediction in Diverse Regions via Teacher-Student Model Fusion

Michael Lindemann[1], Kathleen DeMichele[2], Masoumeh Heidari Kapourchali[2], Mohammad Heidari Kapourchali[1,4], Christine Waigl[3], Erin Trochim[4], Long Zhao[5]

[1] *Department of Electrical Engineering, University of Alaska Anchorage, Anchorage, AK*
[2] *Department of Computer Science and Engineering, University of Alaska Anchorage, Anchorage, AK*
[3] *International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AK*
[4] *Alaska Center for Energy and Power, University of Alaska Fairbanks, Fairbanks, AK*
[5]*Department of Electrical Engineering and Computer Science, South Dakota School of Mines and Technology*
[1,2,3,4]{mlindemann, kmdemichele, mheidari2, mhkapourchali, cwaigl, edtrochim}@alaska.edu
[5]long.zhao@sdsmt.edu

***Abstract*—Accurate wildfire prediction in diverse and geographically dispersed areas is crucial for effective wildfire management. However, the limited availability of labeled data in data-challenged regions, along with the unique characteristics of these areas, poses challenges for training robust prediction models. This study investigates the performance of a convolutional neural network (CNN) on datasets comprising Landsat images from Canada and Alaska. Through principal component analysis (PCA), the study uncovers distinct differences in data distribution between the two regions. It is observed that the reduced data size of the Alaskan dataset, along with its distinct data distribution, leads to a decrease in the CNN's accuracy to 75% compared to an impressive 98% achieved on the Canadian dataset. To address this limitation, we propose a teacher-student model approach, transferring knowledge from a CNN trained on the larger Canadian dataset. The results demonstrate a significant accuracy improvement to 88.96% on the Alaskan dataset. Our findings highlight the effectiveness of the teacher-student model in mitigating data scarcity challenges, enhancing wildfire prediction capabilities in regions with limited training data. This research contributes to improved wildfire monitoring and prevention strategies in challenging geographical locations.**

***Index Terms*—wildfire prediction, teacher-student, knowledge transfer, Landsat, remote sensing**

I. Introduction

Wildfires have increased in frequency and intensity in recent years, posing new and emerging risks to infrastructures and communities. According to the National Climate Assessment, the size of the area burned in Alaska's wildfires is projected to double by 2050 and triple by 2100 under continued emissions and further warming [1]. The ability of emergency management professionals and decision makers to determine if an area is at higher risk of a fire breaking out has never been greater [2]. In northern high-latitude regions like Alaska, where fires impact remote and sparsely populated areas, decisions regarding site access and fire suppression are marked by considerable complexity and cost [3]. Monitoring forest conditions in Alaska also presents a significant challenge due to its vast expanse of over 200,000 square miles of forest and limited road infrastructure. In such areas, satellite remote sensing is the sole avenue for acquiring data to support decision-making [3]. Remotely sensed data has emerged as a valuable tool for modeling risk zones of forest fires and provide a fast, noninvasive, and expansive coverage compared to traditional field-based observations [4], [5]. Landsat data has become one the most comprehensive and longest running Earth observation datasets in existence [6]. Therefore, it would be extremely beneficial to utilize this extensive dataset to predict areas of land in extreme northern latitudes that are at risk of burning due to wildfires.

The availability of remotely sensed data is experiencing a remarkable increase due to the reduction in launching costs and the proliferation of satellites and mini-satellites equipped with high-quality sensors [7]. Despite their immense potential, these resources have been minimally utilized in Alaska and other fire management settings in high latitude northern regions [8]. While satellite technology holds great promise, it is not without its limitations. Challenges such as atmospheric opacity, revisit time, and sensor characteristics still pose hurdles in its effective utilization [9]. During a forest fire event, the presence of smoke and clouds can disrupt the accurate observation of the burned area, thereby limiting the selection of suitable satellite images. Additionally, the use of coarse-resolution sensors in many satellites further compounds these limitations [10].

In recent years, transfer learning [11]–[13] has gained significant interest as an effective approach for improving prediction models in diverse applications. It involves leveraging

This work was supported by the U.S. National Science Foundation (NSF) under Grants RISE-2220624, RISE-2220627 and RISE-2022705.

1946-0759/23/$31.00 ©2023 IEEE
DOI 10.1109/ICMLA58977.2023.00079

knowledge acquired from one task or domain and applying it to a related but different task or domain. In the context of wildfire prediction, transfer learning enables the transfer of learned features, representations, or models from areas with abundant data to areas with limited data availability or unique characteristics. By utilizing transfer learning techniques, valuable knowledge can be adapted to enhance prediction accuracy and decision-making capabilities in data-challenged regions, such as high latitude northern areas like Alaska. In order to address the challenges posed by the unique characteristics of these areas, we developed a teacher-student model for knowledge transfer. The model leverages the expertise gained from a CNN trained on the extensive Canadian dataset, comprising Landsat images of regions with historical fire occurrences and regions without fire occurrences. The performance evaluation demonstrated a significant improvement in accuracy, with the model achieving 88.96% accuracy on the Alaskan dataset. These findings highlight the effectiveness of the teacher-student model in mitigating challenges associated with limited training data, contributing to improved wildfire monitoring and prevention strategies in challenging geographical locations.

While the smaller size of the Alaskan dataset demonstrates the effectiveness of transfer learning, it is important to note that the higher performance observed in the larger Canadian dataset cannot be solely attributed to its size. The PCA analysis revealed distinct data distribution patterns, suggesting that the effectiveness of the model is influenced by a combination of factors. These factors potentially include not only the larger dataset size but also the presence of better characterization features in the Canadian dataset. These features encompass higher data quality, finer-resolution imagery, different vegetation types, the influence of historic fire management practices, and a wider variety of land-use types. Therefore, our study emphasizes the significance of both dataset size and the availability of better characterization features in achieving improved wildfire prediction accuracy.

The rest of this paper is organized as follows. Section II presents a concise literature review, discussing relevant studies in the field. In Section III, we outline our methodology, including details on the dataset used and the implementation of the teacher-student model for knowledge transfer. Section IV presents the experimental results, highlighting the performance and accuracy achieved.

II. Literature review

In this section, a brief literature review is provided to explore the existing research and studies on wildfire prediction and the application of transfer learning techniques in diverse and geographically dispersed areas. The purpose of this section is to provide an overview of the relevant literature and methodologies employed in both wildfire prediction and transfer learning. Wildfires pose significant challenges and have far-reaching consequences in various regions, while transfer learning offers a promising approach to leverage knowledge from well-resourced regions for improved predictions in data-challenged areas.

In the field of wildfire prediction, numerous studies have explored the utilization of satellite images and remote sensing data to enhance prediction accuracy. A comprehensive survey of wildfire prediction and detection is provided in [14], [15]. Traditional approaches, including support vector machine [16], decision trees [17], random forest [18], and logistic regression [19] have been applied to assess risk and predict wildfire occurrences. Additionally, deep learning techniques [20] and CNNs [21] have been explored for risk assessment and wildfire prediction. However, these methods suffer from certain limitations. Traditional approaches often rely on manually engineered features derived from satellite images, which may not fully capture the complex and diverse characteristics of fire-prone regions. The scarcity of labeled data in specific geographic areas also hinders model generalization and accuracy. Recent studies have proposed CNN architectures inspired by well-established models like AlexNet, incorporating video-based datasets [22]. Alternatively, they explore transfer learning approaches using pretrained models such as VGG or ResNet, primarily through fine-tuning the network parameters rather than adapting them to different locations [23]. However, these studies primarily focus on active fire detection rather than risk prediction, limiting their applicability in proactive wildfire management.

Transfer learning [11] has emerged as a valuable technique in various domains, offering the potential to improve prediction models by leveraging knowledge from well-resourced areas.

The teacher-student model, a notable approach in transfer learning, involves distilling knowledge from a pre-trained teacher model into a student model, resulting in enhanced prediction capabilities across different domains [24], [25]. The teacher-student model has been successfully employed in various domains, such as computer vision and natural language processing, to transfer knowledge and improve model performance [12], [26]. However, its application in the domain of wildfire prediction, particularly for transferring knowledge between regions, remains largely unexplored. Additionally, the teacher-student model has been utilized for self-training purposes, where an unlabeled dataset is used to train a larger or equally-sized student model [27].

In contrast to existing approaches, our methodology goes beyond retraining a single model and extends the application of teacher-student models for knowledge transfer between different regions. We also use a model fusion technique involving two specialized student models, each focusing on distinct aspects, thereby enhancing the overall predictive capabilities in wildfire prediction. By leveraging the teacher-student model and incorporating transfer learning techniques, this research contributes to the advancement of wildfire monitoring and prevention strategies in high-latitude northern regions. It provides valuable insights into the application of transfer learning methods for enhancing prediction accuracy and decision-making capabilities in data-challenged areas. Moreover, this study is the first known instance of employing a teacher-student model to transfer knowledge between regions in the domain of wildfire prediction, making a notable contribution to the field.

III. Models and Methods

In this section, we outline the models and methods employed in our study for wildfire prediction using transfer learning and the teacher-student model. Our approach aims to leverage the knowledge acquired from a well-resourced region to enhance wildfire prediction capabilities in data-challenged areas. The proposed model consists of several steps. Firstly, a teacher model f_t is trained using a large labeled dataset D_t. This model is then utilized to generate pseudo-labels by applying it to a smaller and more challenging unlabeled dataset D_s. These pseudo-labels indicate whether an area is at risk of a wildfire or not. Next, two student models, f_{s_1} and f_{s_2}, are trained using the combined datasets, D_s and D_t, where D_s now includes labels based on the teacher model's predictions. Combining datasets provides a richer and more diverse training set for the student models. The student model f_{s_1} focuses on predicting wildfires, while f_{s_2} is trained for non-wildfire areas. By incorporating the knowledge obtained from the teacher model, the teacher-student model aims to enhance the accuracy of wildfire predictions on the challenging and unlabeled dataset D_s. The use of two student models allows for specialized training and independent analysis for each class, leading to improved performance evaluation and model insights. An overview of our model is given in Fig. 1.

A. Training Teacher Model with Labeled Dataset

The teacher model, f_t, employed in our work is a convolutional neural network (CNN) that serves as a pivotal component in the knowledge transfer process for wildfire prediction. This model consists of multiple layers designed to capture spatial features inherent in the input images. During the training process, f_t effectively acquires knowledge and expertise in distinguishing areas at a higher risk of wildfires from those with limited fire incidents. This acquired knowledge serves as a foundation for the subsequent knowledge transfer to the student models, empowering them to enhance their wildfire prediction capabilities. By leveraging the insights gained by the teacher model, the student models become equipped with improved abilities to identify and assess the likelihood of wildfires in different geographic regions. The architecture of the CNN model is shown in Fig. 2. It includes a Conv2D layer that performs two-dimensional convolution to extract important features from the input data. The MaxPooling2D layer applies downsampling, reducing the spatial dimensions of the data. The Flatten layer converts the multidimensional data into a one-dimensional vector for further processing. Dense layers are fully connected layers where neurons in one layer are connected to neurons in the previous layer. The model architecture has been empirically determined and optimized for our task, with ReLU activation functions introducing non-linearity. In the output layer, a sigmoid activation function is used for binary classification. The model is trained using the 'adam' optimizer and the binary cross-entropy loss function.

B. Pseudo-Label Generation using the Teacher Model

In the pseudo-label generation step, we leverage the trained teacher model, f_t, to assign pseudo-labels to an unlabeled dataset, D_s, indicating the likelihood of wildfire presence in each area. These pseudo-labels are obtained by applying f_t to D_s and thresholding the predictions. The pseudo-labeled samples are then combined with the labeled dataset, D_t, to create a more diverse training set for the student models, f_{s_1} and f_{s_2}. By incorporating the teacher model's knowledge, the student models enhance their wildfire prediction capabilities. This approach improves the accuracy of wildfire predictions on the challenging and unlabeled dataset, D_s, and enables specialized training and independent analysis for each class.

C. Training the Student Models

The student models, denoted as f_{s_1} and f_{s_2}, are trained using the combined datasets, which include both labeled data D_t and pseudo-labeled data D_s. The labeled dataset D_t contains a large number of labeled samples, while the smaller and more challenging unlabeled dataset D_s is used to generate pseudo-labels using the teacher model. During the training process, the student models benefit from both the labeled and pseudo-labeled data. The labeled data provides ground truth information for training, while the pseudo-labeled data generated by the teacher model provides additional training samples for the student models. This combination of labeled and pseudo-labeled data enriches the training set and improves the generalization capability of the student models.

Both student models, f_{s_1} and f_{s_2}, employ a similar CNN architecture to the teacher model, f_t. This design choice is based on evaluating different architectures and selecting the one that yields the best performance for the student models. By utilizing a similar architecture, the student models can effectively benefit from the learned representations and spatial features captured by the teacher model. The training of the student models involves optimizing their respective CNN architectures using the combined dataset. This training process allows the student models to learn and refine their predictive abilities based on the combined knowledge from the teacher model and the labeled and pseudo-labeled data.

Our approach aims to significantly enhance the accuracy and reliability of wildfire predictions across diverse geographic regions by harnessing the specialized capabilities of the student models. By employing two distinct student models within our framework, we leverage their unique characteristics and expertise to address the challenges and variations present in different areas. This enables us to effectively capture the complex patterns and behaviors associated with wildfires, leading to more precise and reliable predictions. Through the combined efforts of the student models, our approach offers an advanced solution for wildfire prediction that exhibits improved performance and adaptability across various geographic regions.

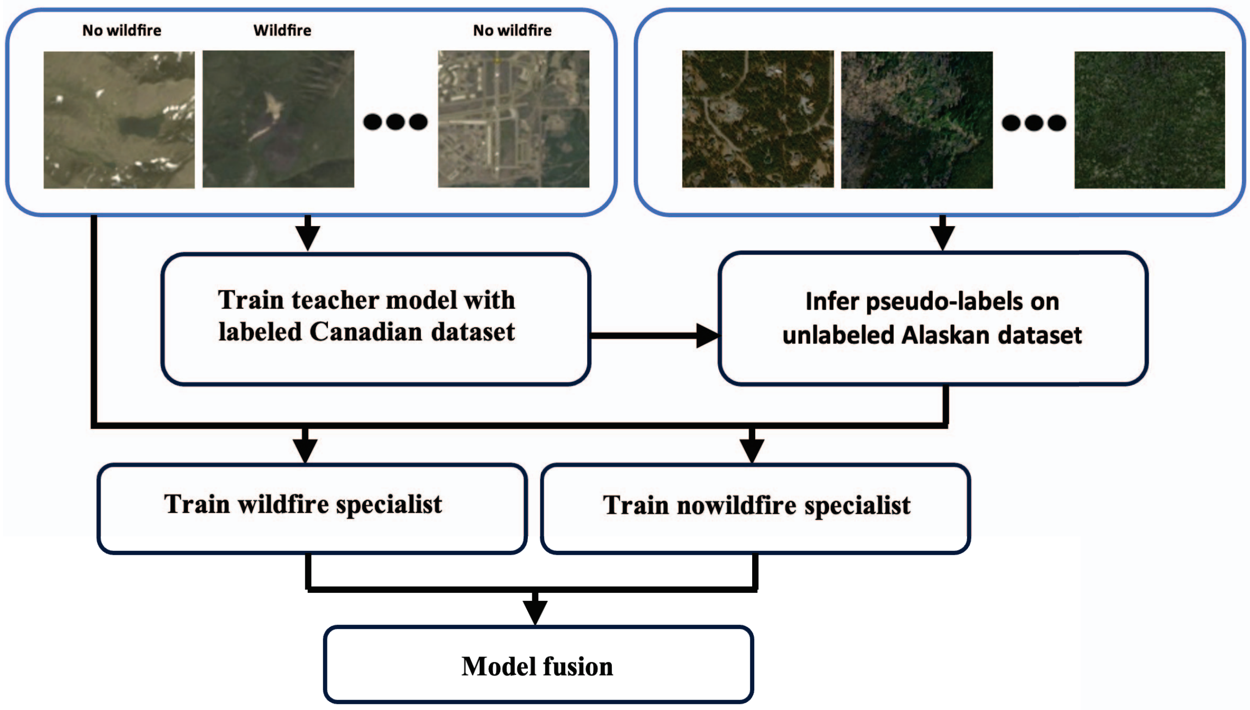

Fig. 1. Block diagram of the proposed model for transferring knowledge from the Canadian dataset to the Alaskan dataset.

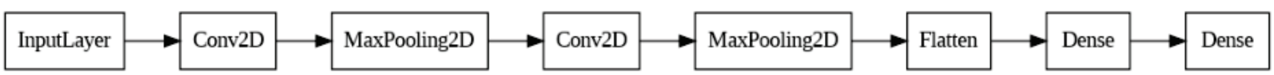

Fig. 2. Architecture of the Convolutional Neural Network (CNN) model for the teacher model

D. *Fusion of Student Models*

Our late fusion technique combines the predictions of two student models, represented by f_{s_1} and f_{s_2}, specializing in wildfires and non-wildfire areas, respectively, to enhance the accuracy of our wildfire prediction system. Initially, each model independently makes a hard decision, denoted as y_{s_1} and y_{s_2}, based on their respective expertise. In the case where $y_{s_1} \neq y_{s_2}$, indicating a discrepancy or uncertainty, a fusion process is initiated. The model that made the hard decision, f_{s_1}, incorporates the soft decision made by f_{s_2}, refining its prediction by combining the posterior probabilities. This fusion-based approach harnesses the strengths and insights of both models, resulting in a comprehensive and more reliable prediction for wildfire detection.

Algorithm 1 summarizes the process of Section III.

IV. Experimental results

A. *Dataset*

The dataset used in this study consists of satellite images from two different regions: Canada and Alaska. The Canadian dataset [28] comprises 42,848 satellite images that have been divided into training, testing, and validation sets. These images were generated using MapBox API and uploaded to Kaggle [28]. Each image is a 350 pixel by 350 pixel RGB image centered on either the location of a wildfire or a location where there has never been a wildfire. The original wildfire points data for Canada was obtained from the Canadian government [29]. The dataset includes satellite images captured over a period spanning from April 30, 1972, to October 31, 2021.

The Alaskan dataset used in this study comprises satellite images collected by the research team from Landsat, a series of Earth observation satellites operated by NASA and the U.S. Geological Survey (USGS). The dataset covers a substantial period from 2015 to 2021 and includes wildfire events. Filtering criteria were applied to select wildfire locations based on factors such as minimum acreage burned and exclusion of false alarms, including points that are expected to experience wildfire occurrences. ArcGIS Pro [30], a geographic information system (GIS) software developed by Esri, was utilized for data processing, including the filtering of wildfire locations and performing a nearest neighbor search. The resulting dataset includes Landsat 8 satellite images obtained for each selected wildfire location using the Google Earth Engine API. Images captured during the subsequent summer season

Algorithm 1 Pseudocode for the proposed model.

Require:
Labeled images $D_t = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$ and unlabeled images $D_s = \{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_m\}$

$\theta^* \leftarrow$ Adam $(\frac{1}{n}\sum_{i=1}^{n} loss(y_i, f_t(x_i, \theta)))$

Generate pseudolabels for D_s using f_t :

$$\tilde{y}_i \leftarrow f_t(\tilde{x}_i, \theta^*) \ \forall i = 1, ..., m$$

Combine labeled dataset D_t with pseudo-labeled dataset D_s to create combined datasets D_cs:

$D_{c_1} \leftarrow D_s \cup D_{t_1} = \{(x_i, y_i) \mid (x_i, y_i) \in D_t, y_i = \text{“}Wildfire\text{”}\}$

$D_{c_2} \leftarrow D_s \cup D_{t_2} = \{(x_i, y_i) \mid (x_i, y_i) \in D_t, y_i = \text{“}Nowildfire\text{”}\}$

$\theta^*_{s_j} \leftarrow$ Adam $(\frac{1}{n}\sum_{(x_i,y_i)\in D_{c_1}} loss(y_i, f_t(x_i, \theta_{s_j})),\ j = 1, 2)$

for each $(x_i, y_i) \in D_c$ **do** ▷ Model fusion

  if $f_{s_1}(x_i, \theta^*_{s_1}) \neq f_{s_2}(x_i, \theta^*_{s_2})$ **then**

    $p_1 \leftarrow f_{s_1}(x_i, \theta^*_{s_1})$ ▷ p stands for posterior probability

    $p_2 \leftarrow f_{s_2}(x_i, \theta^*_{s_2})$

    $p_{\hat{y}_i} \leftarrow \alpha \cdot p_1 + (1 - \alpha) \cdot p_2$

  else

    $\hat{y}_i \leftarrow f_{s_1}(x_i, \theta^*_{s_1})$

  end if

end for

with minimal snow and cloud cover were chosen. Additionally, "no-wildfire" images were generated using a nearest neighbor search approach implemented in ArcGIS Pro. This approach involved identifying locations in close proximity to recorded wildfire locations, ensuring the selection of suitable areas without wildfire occurrences.

B. Data Distribution Analysis

To gain insights into the data distribution of the satellite images from Alaska and Canada, we performed PCA on the dataset. PCA allows us to reduce the dimensionality of the data while retaining the most important variations present in the dataset. The PCA analysis was conducted separately for the Alaska and Canadian datasets. Fig. 3 presents a scatter plot of the dataset projected onto the two principal components. The x-axis represents the first principal component, and the y-axis represents the second principal component. Each point on the plot corresponds to an image in the dataset, and the color of the point indicates its label (fire or no wildfire).

From the scatter plot, we can observe distinct data distribution differences between the Alaska and Canadian datasets. In the Canadian dataset, the fire and no wildfire classes exhibit relatively well-separated clusters, indicating a clear distinction between the two classes. However, in the Alaska dataset, the separation between the fire and no wildfire classes is less pronounced, with more overlap between the clusters. This suggests that the classification task for Alaska is more challenging due to the similarity in data distribution between the fire and no wildfire classes. The smaller size of the Alaskan dataset demonstrates the effectiveness of transfer learning, but the higher performance in the larger Canadian dataset cannot be solely attributed to its size. The PCA analysis reveals distinct patterns in the data distribution which may be influenced by superior characterization features in the Canadian dataset. These features potentially include enhanced data quality, finer-

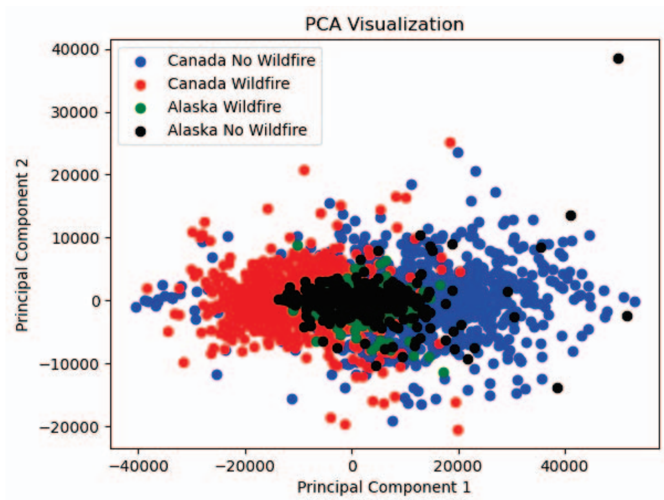

Fig. 3. Scatter plot of the Alaska and Canadian datasets projected onto the first two principal components.

resolution imagery, diverse vegetation types, the impact of historical fire management practices, and a broader range of land-use types.

The observed variations in data distribution and the complexities in wildfire classification tasks highlight the importance of exploring alternative approaches to address the challenges of predicting wildfire in diverse environments.

C. Performance Analysis and Comparative Evaluation of Wildfire Prediction: Study on Canadian and Alaskan Datasets:

This section presents the analysis and comparison of different methods for predicting wildfires on the Canadian and Alaskan datasets. In our experimental setup, we allocated 30% of the dataset for testing purposes, ensuring a reliable evaluation of the models' performance. Additionally, we assigned equal importance (with alpha=0.5) to both wildfire specialists and non-wildfire specialists during the training process. The accuracy (%) results of various approaches, including Logistic Regression (LR), Multilayer Perceptron (MLP), VGG19, ResNet-50, MobileNet, EfficientNet-B4, EfficientNet-B7, CNN1, and the Teacher model, are summarized in Table I. In the Canadian dataset, all models achieved relatively high accuracies. Notably, the CNN1 model, consisting of three convolutional layers with max pooling, followed by a flatten layer, two dense layers, and a sigmoid output layer, achieved the highest accuracy of 98.40% on the Canadian dataset. This demonstrates the effectiveness of convolutional neural networks in accurately predicting wildfires using satellite images in the Canadian region.

However, when these models were applied to the Alaskan dataset, the performance significantly dropped. LR and MLP achieved accuracies of 30.40% and 30.10%, respectively, indicating poor performance. The deep learning models also struggled to perform well on the Alaskan dataset, with accuracies ranging from 51.23% to 60.89%. The CNN1 model achieved an accuracy of 75.60%, which is higher compared to other models but still relatively low. These findings underscore the need for improved generalization capabilities of wildfire prediction models, as even strong models demonstrate limited performance due to variations in data distribution and the unique characteristics of the Alaskan region. To address the limitations of the models when applied to the Alaskan dataset, we proposed the use of a teacher-student model for knowledge transfer. By leveraging the knowledge acquired from the teacher model trained on the larger Canadian dataset, we aim to enhance the prediction capabilities on the Alaskan dataset. The next section will discuss the evaluation of the teacher-student model, demonstrating its effectiveness in mitigating the challenges associated with limited training data and improving wildfire prediction accuracy in challenging geographical locations.

TABLE I
COMPARISON OF ACCURACY (%) FOR DIFFERENT METHODS ON CANADIAN AND ALASKAN DATASETS. LR CORRESPONDS TO LOGISTIC REGRESSION, MLP REPRESENTS MULTILAYER PERCEPTRON.

Method	Canada data	Alaska data
LR	88.90	30.40
MLP	89.40	30.10
VGG19 [31]	94.64	51.23
ResNet-50 [32]	95.84	50.00
MobileNet [33]	96.14	59.67
EfficientNet-B4 [34]	79.96	60.89
EfficientNet-B7 [34]	85.43	58.12
CNN1	**98.40**	**75.60**
Teacher model	**94.23**	**72.23**

D. Enhancing Wildfire Prediction in Data-Challenged Regions: Results of Teacher-Student Model:

The teacher-student model was developed to tackle the challenges stemming from limited training data and to enhance wildfire prediction accuracy in data-challenged regions. Leveraging the knowledge obtained from the teacher model trained on the larger Canadian dataset, our objective was to improve prediction capabilities on the Alaskan dataset. The teacher model served as a valuable source of knowledge, while the student model was specifically designed to learn from the teacher's predictions. In order to assess the performance of the teacher-student model on the Alaskan dataset, we first present the confusion matrix analysis.

The development of the teacher-student model aimed to address the challenges arising from limited training data and improve the accuracy of wildfire prediction in regions with data limitations. By leveraging the knowledge gained from the teacher model trained on the larger Canadian dataset, our objective was to enhance prediction capabilities on the Alaskan dataset. The teacher model played a crucial role as a valuable source of knowledge, while the student model was designed to learn from the teacher's predictions. To evaluate the performance of the teacher-student model on the Alaskan dataset, the confusion matrix analysis is presented in Fig. 4.

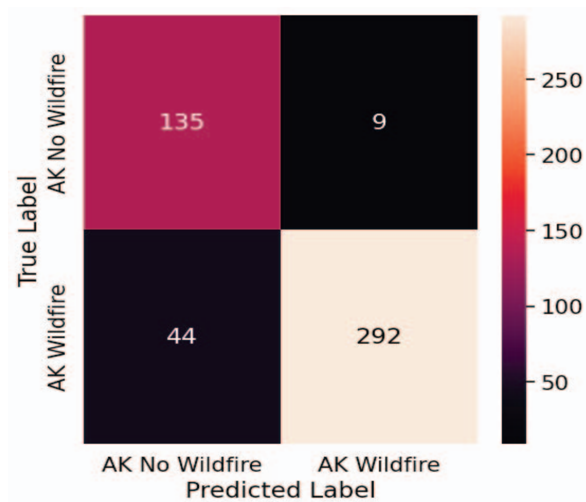

Fig. 4. Confusion matrix of the Teacher-Student model on the Alaskan dataset

Table II presents the results of different model configurations for wildfire prediction on the Alaskan dataset, along with various performance metrics. These metrics provide insights into the effectiveness of each configuration and its implications for wildfire prediction.

The first and third configurations involve a teacher-student model with one student model learning from the teacher. These configurations demonstrate the impact of incorporating knowledge transferred from the teacher model, resulting in improved prediction accuracy. The second and fourth configurations utilize two student models with different specializations and fusion. These configurations further enhance the accuracy and F1-score compared to the first and third configurations. The fourth configuration, the proposed model, achieves the highest accuracy of 88.96% and the highest F1-score of 91.68%. This configuration, with two student models and

TABLE II
RESULTS OF DIFFERENT MODEL CONFIGURATIONS

Model Configuration	Accuracy (%)	F1-Score (%)	Precision (%)	Recall (%)
CNN1 (Teacher) + 1 CNN1 (Student)	81.59	87.81	81.12	94.32
CNN1 (Teacher) + 2 CNN1 (Students) + Fusion	81.87	85.13	100	74.19
Teacher Model + 1 Student	83.42	84.40	83.42	86.60
Teacher Model + 2 Students + Fusion (the proposed model)	**88.96**	**91.68**	97.00	86.90

fusion, demonstrates the strongest performance in terms of accuracy and precision.

These results highlight the effectiveness of the teacher-student model in enhancing wildfire prediction capabilities in data-challenged regions like Alaska. By leveraging the knowledge transferred from the teacher model to the student models, substantial improvements in accuracy, precision, and overall performance are achieved. The proposed model contributes to the advancement of wildfire monitoring and prevention strategies in challenging geographical locations.

V. CONCLUSION

The results of our study highlight the effectiveness of our proposed model in improving wildfire prediction capabilities in challenging geographical locations. By leveraging the teacher-student model and transfer learning techniques, we have successfully enhanced prediction accuracy by incorporating knowledge from a teacher model trained on a larger Canadian dataset. Our research addresses the challenges of accurate wildfire prediction in diverse environments, with a specific focus on the unique characteristics of the Alaskan region. This research significantly contributes to the advancement of wildfire monitoring and prevention strategies by providing a valuable framework for knowledge transfer and model adaptation in areas with varying data availability and distinct geographical features. Furthermore, our Alaskan dataset and the corresponding results serve as a valuable benchmark for evaluating wildfire prediction models in diverse environments.

REFERENCES

[1] T. Sanford, R. Wang, and A. Kenwa, "The age of alaskan wildfires," *Climate Central, Princeton*, 2015.

[2] J. Balson, M. Chinchilla, C. Lu, J. Washburn, and N. Lohia, "Identification and characterization of forest fire risk zones leveraging machine learning methods," *SMU Data Science Review*, vol. 5, no. 2, p. 3, 2021.

[3] C. F. Waigl, M. Stuefer, A. Prakash, and C. Ichoku, "Detecting high and low-intensity fires in alaska using viirs i-band data: An improved operational approach for high latitudes," *Remote Sensing of Environment*, vol. 199, pp. 389–400, 2017.

[4] B. Leblon, L. Bourgeau-Chavez, and J. San-Miguel-Ayanz, "Use of remote sensing in wildfire management," *Sustainable development-authoritative and leading edge content for environmental management*, pp. 55–82, 2012.

[5] M. A. Enoh, U. C. Okeke, and N. Y. Narinua, "Identification and modelling of forest fire severity and risk zones in the cross–niger transition forest with remotely sensed satellite data," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 24, no. 3, pp. 879–887, 2021.

[6] M. C. Hansen and T. R. Loveland, "A review of large area monitoring of land cover change using landsat data," *Remote sensing of Environment*, vol. 122, pp. 66–74, 2012.

[7] M. Gazzea, L. M. Kristensen, F. Pirotti, E. E. Ozguven, and R. Arghandeh, "Tree species classification using high-resolution satellite imagery and weakly supervised learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2022.

[8] A. York and R. Jandt, "Opportunities to apply remote sensing in boreal," in *Arctic Wildfire Management & Science: A Workshop Report www. frames. gov/catalog/57849 (University of Alaska, Fairbanks, 2019)*, 2019.

[9] A. Y. Cho, S.-e. Park, D.-j. Kim, J. Kim, C. Li, and J. Song, "Burned area mapping using unitemporal planetscope imagery with a deep learning based approach," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 242–253, 2022.

[10] K. Ahmad, K. Pogorelov, M. Riegler, O. Ostroukhova, P. Halvorsen, N. Conci, and R. Dahyot, "Automatic detection of passable roads after floods in remote sensed and social media data," *Signal Processing: Image Communication*, vol. 74, pp. 110–118, 2019.

[11] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[12] J. Yim, D. Joo, J. Bae, and J. Kim, "A gift from knowledge distillation: Fast optimization, network minimization and transfer learning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4133–4141.

[13] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, 2015.

[14] F. Abid, "A survey of machine learning algorithms based forest fires prediction and detection systems," *Fire technology*, vol. 57, no. 2, pp. 559–590, 2021.

[15] P. Jain, S. C. Coogan, S. G. Subramanian, M. Crowley, S. Taylor, and M. D. Flannigan, "A review of machine learning applications in wildfire science and management," *Environmental Reviews*, vol. 28, no. 4, pp. 478–505, 2020.

[16] O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, and J. Aryal, "Forest fire susceptibility and risk mapping using social/infrastructural vulnerability and environmental variables," *Fire*, vol. 2, no. 3, p. 50, 2019.

[17] R. Sharma, S. Rani, and I. Memon, "A smart approach for fire prediction under uncertain conditions using machine learning," *Multimedia Tools and Applications*, vol. 79, pp. 28 155–28 168, 2020.

[18] P. Janiec and S. Gadal, "A comparison of two machine learning classification methods for remote sensing predictive modeling of the forest fire in the north-eastern siberia," *Remote Sensing*, vol. 12, no. 24, p. 4157, 2020.

[19] Y. Michael, D. Helman, O. Glickman, D. Gabay, S. Brenner, and I. M. Lensky, "Forecasting fire risk with machine learning and dynamic information derived from satellite vegetation index time-series," *Science of The Total Environment*, vol. 764, p. 142844, 2021.

[20] B. Lattimer, J. Hodges, and A. Lattimer, "Using machine learning in physics-based simulation of fire," *Fire Safety Journal*, vol. 114, p. 102991, 2020.

[21] G. Zhang, M. Wang, and K. Liu, "Forest fire susceptibility modeling using a convolutional neural network for yunnan province of china," *International Journal of Disaster Risk Science*, vol. 10, pp. 386–403, 2019.

[22] K. Muhammad, J. Ahmad, and S. W. Baik, "Early fire detection using convolutional neural networks during surveillance for effective disaster management," *Neurocomputing*, vol. 288, pp. 30–42, 2018.

[23] J. Sharma, O.-C. Granmo, M. Goodwin, and J. T. Fidje, "Deep convolutional neural networks for fire detection in images," in *Engineering Applications of Neural Networks: 18th International Conference, EANN 2017, Athens, Greece, August 25–27, 2017, Proceedings*. Springer, 2017, pp. 183–193.

[24] L. Wang and K.-J. Yoon, "Knowledge distillation and student-teacher learning for visual intelligence: A review and new outlooks," *IEEE*

transactions on pattern analysis and machine intelligence, vol. 44, no. 6, pp. 3048–3068, 2021.
[25] C. Hu, X. Li, D. Liu, X. Chen, J. Wang, and X. Liu, "Teacher-student architecture for knowledge learning: A survey," *arXiv preprint arXiv:2210.17332*, 2022.
[26] X. Jiao, Y. Yin, L. Shang, X. Jiang, X. Chen, L. Li, F. Wang, and Q. Liu, "Tinybert: Distilling bert for natural language understanding," *arXiv preprint arXiv:1909.10351*, 2019.
[27] Q. Xie, M.-T. Luong, E. Hovy, and Q. V. Le, "Self-training with noisy student improves imagenet classification," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 687–10 698.
[28] A. ABA, "Wildfire prediction dataset," https://www.kaggle.com/datasets/abdelghaniaaba/wildfire-prediction-dataset, accessed [Date Accessed].
[29] Government of Canada, "Wildfire points data for canada," https://open.canada.ca/data/en/dataset/9d8f219c-4df0-4481-926f-8a2a532ca003, accessed [Date Accessed].
[30] "ArcGIS Pro," https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview, accessed July 14, 2023.
[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
[33] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
[34] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.