Version of Record: https://www.sciencedirect.com/science/article/pii/S0957417422007497
Manuscript_c1fe57cae88376bbc70bfe3cf5130524

Official (Closed) - Non Sensitive

Automated accurate fire detection system using ensemble pretrained residual network

Sengul DOGAN[a], Prabal Datta BARUA[b,c,d], Huseyin KUTLU[e] , Mehmet BAYGIN[f] , Hamido FUJITA[g,h], Turker TUNCER[a], U Rajendra ACHARYA[i,j,k]

[a] Department of Digital Forensics Engineering, Technology Faculty, Firat University, Elazig, Turkey

{sdogan; turkertuncer}@firat.edu.tr

[b] School of Management & Enterprise, University of Southern Queensland, Australia

[c] Faculty of Engineering and Information Technology, University of Technology Sydney, Australia

[d] Cogninet Brain Team, Cogninet Australia, Sydney, NSW 2010, Australia

Prabal.Barua@usq.edu.au

[e] Department of Computer Technologies, Besni Vocational School, Adiyaman University, Adiyaman, Turkey

hkutlu@adiyaman.edu.tr

[f] Department of Computer Engineering, Engineering Faculty, Ardahan University, Ardahan, Turkey

mehmetbaygin@ardahan.edu.tr

[g] SOMET.org Associated Organization, Iwate, Japan

[h] Iwate Prefectural University, Iwate, Japan

HFujita-799@acm.org, hfujita@i-somet.org

[i] Ngee Ann Polytechnic, Department of Electronics and Computer Engineering, 599489, Singapore

[j] Department of Biomedical Engineering, School of Science and Technology, SUSS University, Singapore

[k] Department of Biomedical Informatics and Medical Engineering, Asia University, Taichung, Taiwan.

aru@np.edu.sg

© 2022 published by Elsevier. This manuscript is made available under the Elsevier user license
https://www.elsevier.com/open-access/userlicense/1.0/

Abstract

Nowadays, fires have been commonly seen worldwide and especially forest fires are big disasters for humanity. The prime objective of this work is to develop an accurate fire warning model by using images. In this work, two new deep feature engineering models are proposed to detect the fire accurately using images. To create deep features, residual networks (ResNet) are chosen since these networks are one of the highly accurate convolutional neural networks. In this work, four pretrained ResNets: ResNet18, ResNet50, ResNet101, and InceptionResNetV2 are used. These networks were trained using a cluster of ImageNet dataset and features were extracted using the last pooling and fully connected layers of these networks. Hence, eight feature vectors are chosen using these networks and the top 256 features of these networks are chosen using neighborhood component analysis (NCA). Support vector machine (SVM) classifier has been used for classification. Moreover, by using the eight feature vectors generated, two ensemble models have been presented. In the first ensemble model, generated all features are concatenated and the top 1000 features are chosen using a feature selector used (NCA), and these features are classified using SVM. In the second ensemble model, iterative hard majority voting (IHMV) has been applied to the generated results. The developed ensemble ResNet models attained 98.91% and 99.15% classification accuracies using an SVM classifier with a 10-fold cross-validation strategy. Our results obtained demonstrate the high classification accuracy of our presented ensemble pretrained ResNet-based deep feature extraction models. These developed models are ready to be tested with higher databases before actual real-world application.

Keywords: Fire detection; Ensemble ResNet; deep feature extraction; transfer learning; iterative hard majority voting; NCA.

1. Introduction

Fire is one of the most dangerous natural disasters that cause loss of life and property of people (Abedi Gheshlaghi, Feizizadeh, Blaschke, Lakes, & Tajbar, 2021; Nicolopoulos & Hansen, 2009). Due to this natural disaster, people suffer significant economic and ecological losses (Wang, et al., 2021). Fires are widely observed around the world (Sahal, Alsamhi, Breslin, & Ali, 2021). Forest fires are one of these types of fire. Many living things are damaged and the ecological balance is deteriorated due to forest fires, which significantly affect the natural ecosystem (Chaudhry, Sidhu, & Paliwal, 2021). Forest fires have been seen continuously throughout the world. One of the recent forest fires was seen in Australia (Pickrell, 2021). About 1 billion animals were affected by this fire, 28 people died and the fire lasted for 240 days (Norman, Newman, & Steffen, 2021).

Nowadays, early warning and automatic detection models are actively used in firefighting with the developing technology (Barmpoutis, Papaioannou, Dimitropoulos, & Grammalidis, 2020). Especially, traditional indoor fire detection technologies (such as smoke detectors) provide significant advantages in controlling large areas (Piera & Salva, 2019). In addition, situations where there is a fire or fire risk, can be detected by using imaging technologies. Nowadays, early detection models using machine learning methods are one of the important topics studied in the literature. In a study by Khatami et. al.(Khatami, Mirghasemi, Khosravi, Lim, & Nahavandi, 2017), K-medoids clustering and PSO-based approach were used for fire detection from the image. The K-medoids clustering method was used for fitness evaluation in their study. A feature matrix was created by manually selecting fire and non-fire images in the dataset. After that, Otsu (Otsu, 1979) thresholding method was applied and a binary image has been created. In the developed model, 86.24% true positive, 13.76% false-positive and 93.40 detection accuracy were achieved with the proposed approach. An unmanned aerial vehicle (UAV) based fire detection approach for forest fires was developed by Yuan et. al. (Yuan, Liu, & Zhang, 2017) using infrared images. Sharma et. al. (Sharma, Granmo, Goodwin, & Fidje, 2017) used deep CNN methods for fire detection in images. In their study, ResNet50 and VGG16 deep learning models were used. 651 images labeled as fire/non-fire were used in their experimental studies. Test results showed an accuracy of 91.18% using VCG16, and 92.15% accuracy using ResNet50. Gong et. al. (Gong, et al., 2019) proposed a real-time fire detection method from video. In the proposed method, motion-based four features are extracted and classified with support vector machines (SVM). Experimental results showed that the proposed method provided an accuracy of 95.29%. Similarly, Kim and Lee (Kim & Lee, 2019) reported a video-based fire detection method. In their method, areas with suspected fire were detected with faster region-based CNN. In addition, the long short-term memory method was used to detect whether there was a fire. In another study, Adaboost and feedforward neural network (Saeed, Paul, Karthigaikumar, & Nayyar, 2019) were preferred for fire detection. In the proposed method, 64170 images containing smoke and fire images were used. Majid et. al. (Majid, et al., 2022) proposed a CNN-based fire detection approach. In this study, four different datasets were combined and a total of 7977 siz fire/non-fire image dataset was obtained. In the developed model, four different deep learning models were tested and the highest accuracy of 95.4% was obtained with EfficentNetB0. A method for fire detection in surveillance videos was proposed by Muhammad et. al.(Muhammad, Ahmad, Mehmood, Rho, & Baik, 2018). In the proposed method, a custom-designed CNN inspired by the GoogleNet CNN model is used. Four different datasets were tested in their study. Hu et. al. (Hu, Tang, Jin, He, & Li, 2018) proposed deep convolutional long-recurrent networks (DCLRN) and optical flow methods for real-time fire detection. Videos containing fire images were used in the study, and an accuracy of 87.6% for RGB images and 90.3% for optical flow images was obtained. Mahmoud and Ren (Mahmoud & Ren, 2018) proposed a rule-based method. The proposed method mainly used image processing and temporal variation. The obtained results showed that the correct detection rate was 96.63%. A hand-crafted based flame detection method was developed by Liu et. al. (Liu, Yang, & Ji, 2016). Basically, YCbCr color space, saliency detection, and uniform local binary pattern (ULBP) were used in their method. The proposed method provided an accuracy of 84.50% using still images. Xu et. al. (Xu, Lin, Lu, Cao, & Liu, 2021) has developed a deep learning-based ensemble learning method. Yolov5, EfficientDet and EfficientNet models were used in their study. To test the proposed method, public fire datasets were combined to obtain 10,581 fire/non-fire images. They have obtained a false positive rate of 51.3%. Pereira et. al. (de Almeida Pereira, Minoro Fusioka, Tomoyuki Nassu, & Minetto, 2021) proposed an active fire detection model using Landsat-8 images. In this study, approximately 150,000 images were obtained from the satellite, then the obtained images were segmented. After that, a U-Net based CNN architecture was developed for active fire detection. They have reported a precision of 87.2% and recall rate of 92.4% using the proposed model. Similarly, Chanthiya and Kalaivani (Chanthiya & Kalaivani, 2021) classified Landsat images into two classes. In their method, first feature extraction was applied and then they were classified using SVM optimized by Krill herd. The developed model provided 99.21% accuracy and 98.41% precision.

1.1. Motivation and our model

Fires have generally caused disasters worldwide. Especially, forest fires are very crucial and they have affected the world and the ecosystem. Therefore, an effective early detection model should be introduced to contribute to firefighting. In this work, a computer-based model has been introduced to detect fire images correctly.

In the literature, various transfer learning and feature engineering methods have been introduced to contribute to computer vision research area (Barua, Chan, et al., 2021; Barua, Muhammad Gowdh, et al., 2021; Karadal, Kaya, Tuncer, Dogan, & Acharya, 2021; Poyraz, Dogan, Akbal, & Tuncer, 2022). The major motivation to this work is to suggest new deep feature engineering models and show the performances of the proposed deep feature engineering models on the fire detection problem. Also, a successful learning model should have effective feature creation, feature selection, and classification phases. In this work, four pre-trained deep ResNet (He, Zhang, Ren, & Sun, 2016) feature extractors are used. The neighborhood component analysis NCA (Goldberger, Hinton, Roweis, & Salakhutdinov, 2004) is one of the widely used features selectors. Therefore, NCA is used as a selector to attain high classification accuracy with fewer features. Furthermore, two ensemble models have been proposed in this work. In the first version, these features are concatenated and NCA chooses the top 1000 features. The classification results of this feature vector are presented. In this second model, IHMV is applied to the results calculated using the eight predicted vectors by calculating eight feature vectors generated using four ResNets.

1.2. Contributions

The key contributions of this work are given below:

- ResNet is one of the flagships of CNNs and there are ResNet-based CNNs in the literature. In this work, the commonly used four pretrained ResNets have been chosen to extract the deep features. The main objective of this selection is to use the effectiveness of these networks together. Moreover, the deep features are created using the last pooling and fully connected layers of these networks.
- Two effective ensemble learning models have been proposed using these pretrained ResNets features. Our introduced ensemble ResNet based image classification models are applied to fire images have yielded more than 98% accuracies.

2. Dataset

In this work, we have used the combined two open fire image datasets downloaded from Kaggle (Kumar, 2021; Saied, 2021). The used images are stored in JPEG with variable sizes. These datasets contain images with two categories (fire and non-fire). The number of images used in the combined two public fire image dataset is tabulated in Table 1.

Table 1. Details of the dataset used in the study.

Class	Number of images
No fire	785
Fire	865
Total	1650

3. Residual Network

ResNet architecture demonstrated its success in computer vision in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition in 2015 (He, et al., 2016). The residual layer provides a link between layers, bypassing one or more layers. With ResNet, a CNN architecture with fewer parameter updates, faster training, and superhuman success has been created. Network depth (number of layers) is an important parameter in CNNs. In theory, it is thought that with the increase in the number of layers, the network can extract stronger features from the data, and the capacity to represent the data will increase. However, due to two problems encountered in practice, increasing the number of layers does not increase the classification ability of the features. The first problem is that the values of neurons that do not have strong activation during backpropagation are destroyed due to the long layer chain and cannot contribute to the training of the network. This problem is the Vanishing/Exploding Gradients. The second problem is that the number of parameters increases with the number of layers and the optimization becomes difficult. The ResNet architecture largely solves these two problems with residual connections. The ResNet is an effective architecture and variable ResNet architecture-based networks have been presented in the literature.

4. Ensemble ResNet based learning models

Two new ensemble pretrained ResNet- based learning model have been presented in this work. Four pretrained ResNet-based learning networks: ResNets are ResNet18, ResNet50, ResNet101 and InceptionResNetV2 have been used as feature extractors in this work. A block diagram showing the layers of these CNN architectures is presented in Figure 1.

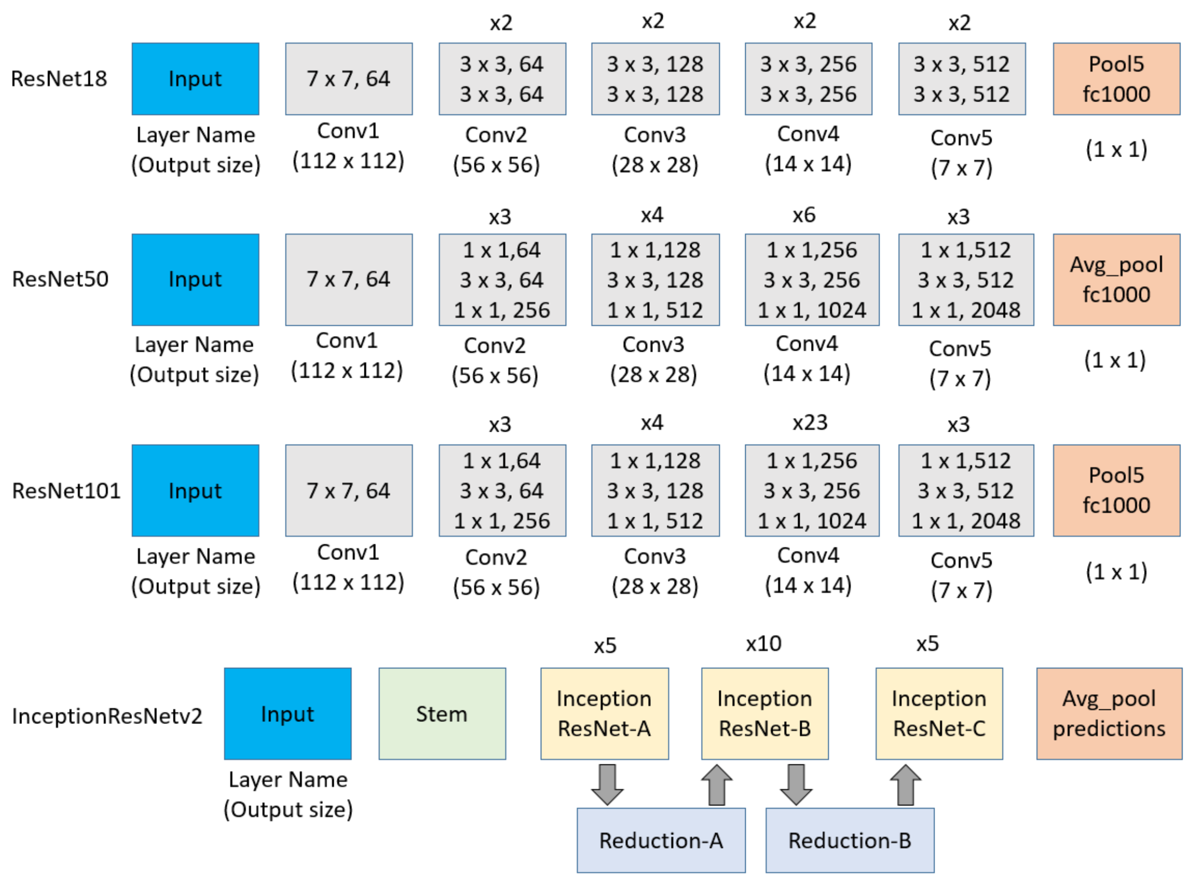


Fig. 1. Block diagram of ResNet architectures.

By using the last pooling and fully connected layers of each pretrained network, eight feature vectors are generated. For the first ensemble model, these features are merged and the most significant 1000 features are chosen by applying NCA. In the classification phase, SVM (Vapnik, 1998, 2013) is employed for automated classification. For the second version, NCA is applied to each feature vector and the most significant 256 features (like local binary pattern) have been chosen. All feature vectors are classified using an SVM classifier and eight results are obtained. In the last phase of defined Case 2, IHMV is applied to calculate the general results. The two ensemble ResNet versions presented in this work are explained below.

4.1. The proposed Ensemble ResNetV1

In version 1, the methodology is created based on feature concatenation. To use the effectiveness of four ResNets used, NCA is employed. The visual explanation of the proposed Ensemble ResNetV1 is denoted in Figure 2.

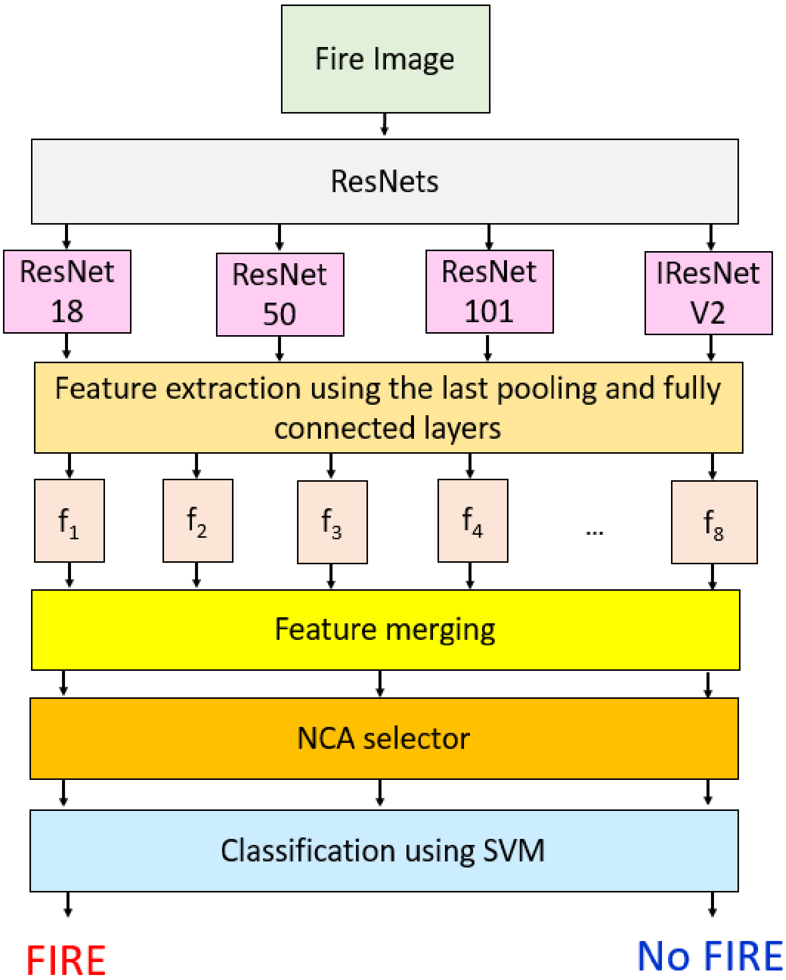


Fig. 2. Block diagram of the ensemble ResNetV1.

Figure 2 depicts the visual illustration of the Ensemble ResNetV1. The steps of this model are given below:

- ***Step 0:*** Load the fire images.
- ***Step 1:*** Load pretrained ResNet18, ResNet50, ResNet101 and InceptionResNetV2 (it is denoted in Figure 2 as IResNetV2).
- ***Step 2:*** Use the last pooling layer and the last fully connected layers of these networks to generate eight feature vectors with variable sizes (these feature vectors are depicted in Figure 2 as f1, f2, …, f8). Table 2 lists the extracted feature vectors, length of each network, and the used layers.

Table 2. Generated feature vectors using the four pretrained networks.

Network	Layer	Name	Length
ResNet18	fc1000	f_1	1000
	pool5	f_2	512
ResNet50	fc1000	f_3	1000
	avg_pool	f_4	2048
ResNet101	fc1000	f_5	1000
	pool5	f_6	2048
InceptionResNetV2	predictions	f_7	1000
	avg_pool	f_8	1536

- ***Step 3:*** Concatenate the feature vectors generated in Table 2.

$$X = concat(f_1, f_2, \ldots, f8) \tag{1}$$

Herein, X is the merged features with a length of 10,144 and $concat$ is the concatenation function.

- ***Step 4:*** Choose the most discriminative/significant 1000 features of the 10, 44 features generated by deploying NCA selector.

$$id = fsNCA(X, y) \tag{2}$$

$$X^s(k, i) = X\big(k, id(i)\big), k \in \{1,2, \ldots, Dim\}, i \in \{1,2, \ldots, 1000\} \tag{3}$$

Herein, id is the qualified indexes of the features by generating NCA feature selector ($fsNCA(.,.)$), y is actual/real labels, X^s is selected feature vector with a size of $Dim \times 1000$ and Dim defines a number of images.

- ***Step 5:*** Classify the selected features (X^s) using polynomial kernelled SVM. The used SVM is called cubic SVM since uses 3rd-degree polynomial kernel. The other setting of this classifier is; kernel scale is automatic, box constraint is one and standardize is true.

The given five steps above are defined in the presented Ensemble ResNetV1.

4.2. The proposed Ensemble ResNetV2

In the second version, we have calculated individual results of each feature vector and the general classification accuracy is obtained using the IHMV technique. The schematic depiction of the Ensemble ResNetV2 is shown in Figure 3.

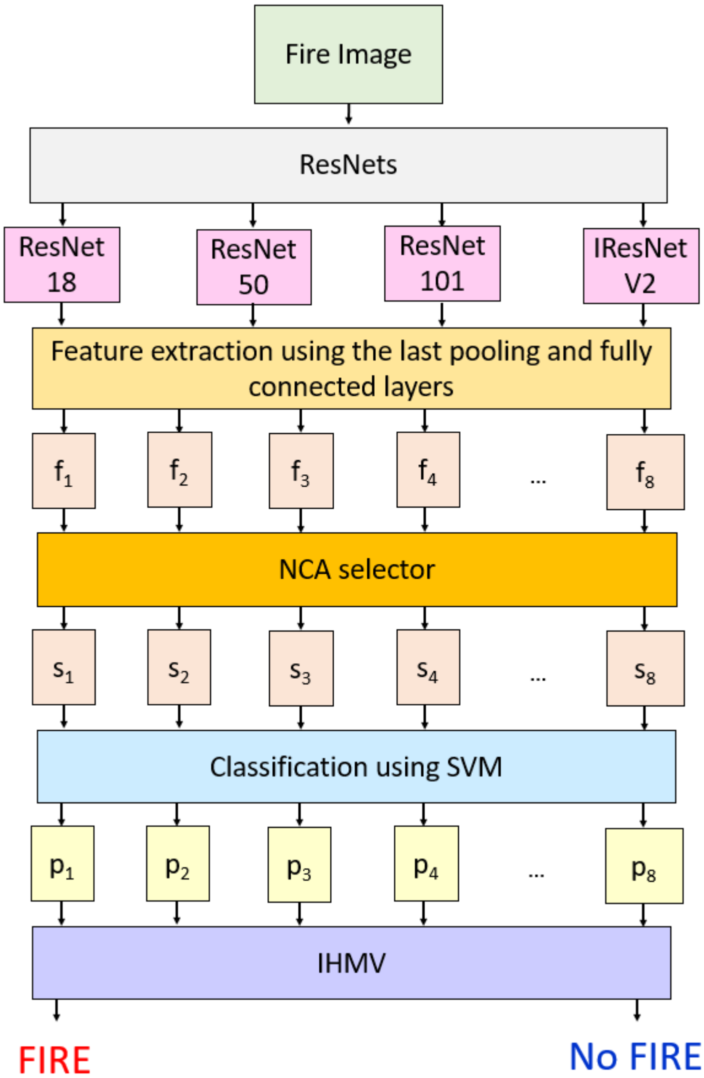


Fig. 3. Block diagram of the ensemble ResNetV2.

Herein, eight feature vectors are generated like our proposed Ensemble ResNetV1. The most significant 256 features of each feature vector are chosen using NCA (the selected vector are denoted in this figure using s). Individual results of each vector are calculated deploying SVM with 10-fold CV and the prediction vectors calculated are shown using p1, p2, …, p8. By using p vectors, IHMV is conducted and final classification results are obtained.

Steps of the presented Ensemble ResNetV2 are;

- ***Step 0-2:*** Use the same steps of the Ensemble ResNetV1.
- ***Step 3:*** Choose the most informative 256 features from the generated feature vectors.

$$id^j = fsNCA(f_1, y), j \in \{1,2,\dots,8\} \quad (4)$$

$$s_j(k,i) = X\left(k, id^j(i)\right), k \in \{1,2,\dots,Dim\}, i \in \{1,2,\dots,256\} \quad (5)$$

Herein, id^j is the qualified indexes of the j^{th} feature vector and s_j is the selected feature vector from j^{th} feature vector and the length of each s vector is 256.

- ***Step 4:*** Calculate prediction vectors using SVM classifier. The attributes of the used SVM are given in Step 5 of the presented Ensemble ResNetV1.
- ***Step 5:*** Apply IHMV to eight prediction vector calculated and obtain general result. Pseudocode of the IHMV used is given in Algorithm 1.

Algorithm 1. Detailed flow of the IHMV technique.

Input: Prediction vectors (p)

Output: General result.

```
Load prediction vector.
Calculate accuracy of each prediction vector using real labels (y).
Sort these vectors according to the accuracy calculated and obtain qualified indexes (ind).
for n=2 to 8 do // Iteratively apply majority voting from the 2 best vectors to all vectors.
    for i=1 to Dim do // Use each value of each p vector.
        for j=1 to n do
            arr(j) = p_ind(j)(i); // Creating array (arr) using the qualified prediction vectors
        end for j
        p^voted_(n-1)(i) = mode(arr); // Apply mode operator to calculate voted vectors.
    end for i
end for n
Calculate acof curacies each voted vector using p^voted and y
Choose the best p^voted as a result
```

5. Performance evaluation

Our proposed two ensemble models have used four pretrained ResNets: ResNet18, ResNet50, ResNet101, and InceptionResNetV2 which were trained millions of images (approximately 1.2 million images) with 1000 clusters. In the proposed work, we have used the optimal weights to extract deep features. Two lightweight ensemble models have been presented using the transfer technique. The suggested two ensemble ResNet models have used NCA as a feature selection model. In the NCA selector, stochastic gradient descends (SGD) has been considered as an optimizer. Moreover, a shallow classifier SVM is used for classification. The proposed ensemble ResNets were implemented on a simple configured computer with an Intel @i9 processor,48 GB memory, and 1 TB hard disk.

The performance parameters namely specificity, sensitivity, geometric mean, precision, F1-score, and accuracy metrics have been used and mathematical equations of these metrics are given below (Chicco & Jurman, 2020; Warrens, 2008).

$$specificity = \frac{TN}{TN + FP} \tag{6}$$

$$sensitivity = \frac{TP}{TP + FN} \tag{7}$$

$$geometric\ mean = \sqrt{specificity \times sensitivity} \tag{8}$$

$$precision = \frac{TP}{TP + FP} \tag{9}$$

$$F1 - score = 2\frac{precision \times sensitivity}{precision + sensitivity} \tag{10}$$

$$accuracy = \frac{TN + TP}{TN + FP + TP + FN} \tag{11}$$

It can be seen from Eqs. (6) - (11) that, the used metrics are calculated using TN, FP, TP and FN values and these values are obtained in Figure 4.

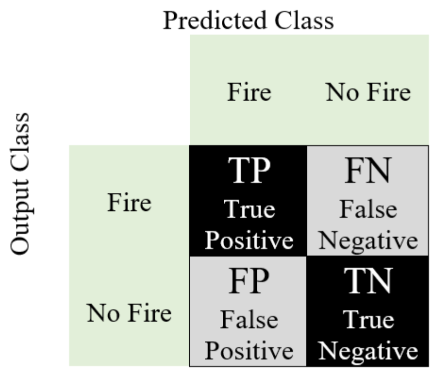


Fig. 4. Description of confusion matrix.

In the first step confusion matrices of the proposed both ensemble ResNets are calculated and the TN, FP, TP and FN values are tabulated in Table 3.

Table 3. Confusion matrices were obtained for both Ensemble ResNetV1 and Ensemble ResNetV2 models.

Model	TP	FN	TN	FP

Ensemble ResNetV1	857	8	779	6
Ensemble ResNetV2	855	10	777	8

Summary of performance obtained using two presented ensemble ResNet models is tabulated in Table 4.

Table 4. Summary of performance obtained using two presented ensemble ResNet models.

Performance metric	Ensemble ResNetV1	Ensemble ResNetV2
Specificity	99.24	98.98
Sensitivity	99.08	98.84
Geometric mean	99.16	98.91
Precision	99.30	99.07
F1-score	99.19	98.96
Accuracy	99.15	98.91

Table 4 denotes that the proposed Ensemble ResNetV1 attained 0.24% higher classification accuracy than the proposed Ensemble ResNetV2 for this dataset.

6. Discussion

Two novel ensemble ResNets have been proposed in this work. In the presented Ensemble ResNetV1 model, generated features from pretrained ResNets are merged and the NCA selector is employed to choose the most effective features. In the second model (Ensemble ResNetV2), IHMV is employed to use the advantages of the considered ResNets. The accuracy and other performance metrics obtained for each deep learning model and layer are given in Table 5.

Table 5. Summary of results (%) obtained using four pretrained networks.

Deep model	Feature vector	Specificity (%)	Sensitivity (%)	Geometric mean (%)	Precision (%)	F1-score (%)	Accuracy (%)
ResNet18	s_1	97.83	98.15	97.99	97.96	97.90	98.00
	s_2	97.96	98.27	98.11	98.15	98.21	98.12
ResNet50	s_3	98.34	98.61	98.48	98.50	98.56	98.48
	s_4	98.47	98.61	98.54	98.61	98.61	98.55
ResNet101	s_5	98.34	98.38	98.36	98.50	98.44	98.36
	s_6	97.96	98.38	98.17	98.15	98.27	98.18
IResNetV2	s_7	98.09	97.80	97.95	98.26	98.03	97.94
	s_8	98.22	98.03	98.13	98.38	98.20	98.12

*s_{odd} = last pooling layer, s_{even} = ffully connected layer

While creating Table 5, the "last pooling" and "fully connected" layers were used as feature extractors. In addition, as in the developed model, features were selected by the NCA method and classified by SVM. In this way, performance metric values were calculated for each deep model and the layers in the models. It can be noted from Table 5 that, results of the chosen feature vectors (s) with a length of 256 and the best accurate features are extracted from the fc1000 layer of the ResNet50 and it attained 98.55% classification accuracy. The individual results are ranged from 97.94% to 98.55%.

After the feature extraction phase, ReliefF, mRMR, NCA and Chi2 methods were tested for feature selection, respectively. Accuracy values were calculated for each of these methods. As a result of these calculated values, it was decided to use the NCA method in the study. The accuracy values obtained according to the feature selection methods are shown in Figure 5.

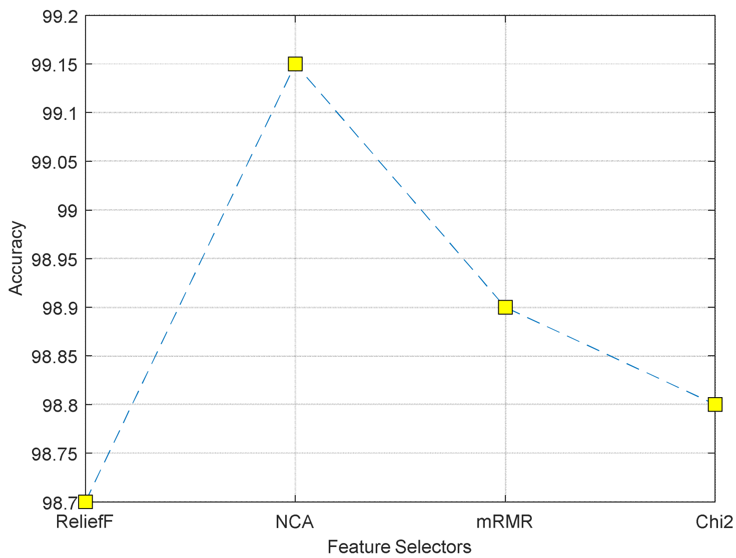


Fig. 5. Accuracy values according to feature selectors.

In the classification phase of the developed model, various kernel types of the SVM method were tested and their accuracy values were calculated. The results obtained according to SVM kernel types are given in Figure 6.

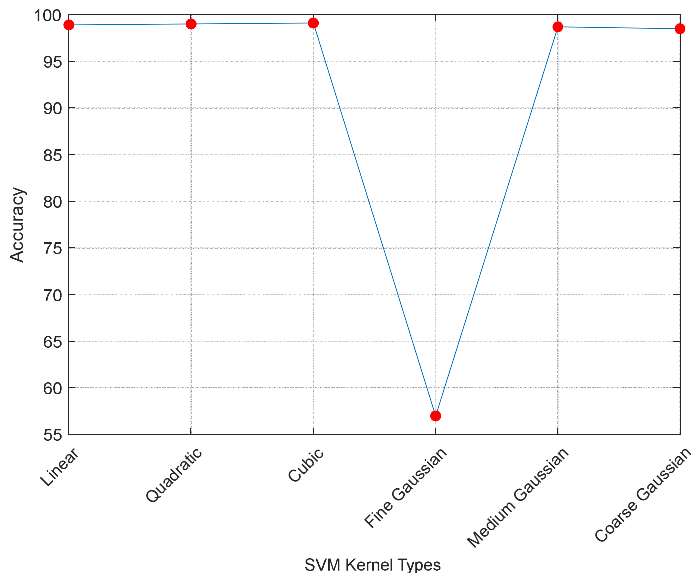


Fig. 6. Accuracy values according to SVM kernel types.

As given in Figure 6, the best accuracy value was obtained with Cubic SVM (Polynomial 3rd kernel). In addition to these results, ensemble classifiers were also tested in the study. The core of the proposed model is the ensemble architecture. Therefore, the accuracy value has been calculated for ensemble classifiers. The accuracy values obtained according to the Ensemble classifiers are presented in Figure 7.

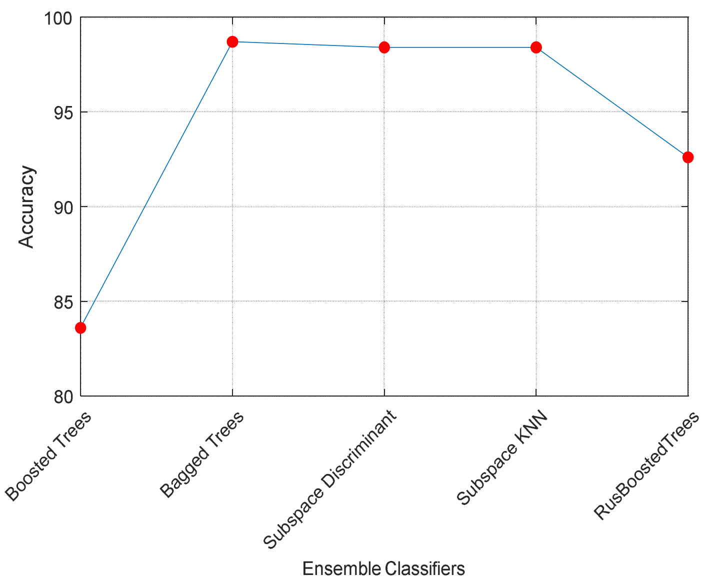


Fig. 7. Accuracy values according to ensemble classifiers.

When the results in Figures 6 and 7 were examined, Cubic SVM (Polynomial 3rd Kernel) was chosen as the best classifier. To increase classification accuracy, IHMV is applied. The plot of best accuracy obtained versus the generated eight prediction vectors is shown in Figure 8.

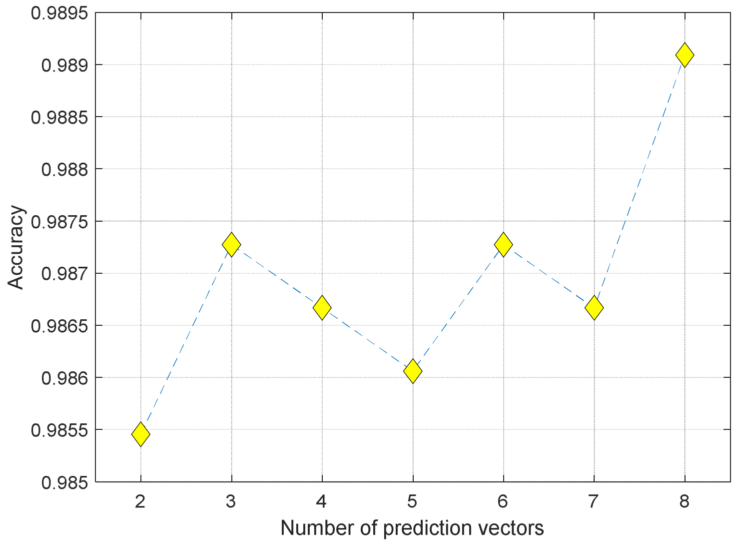
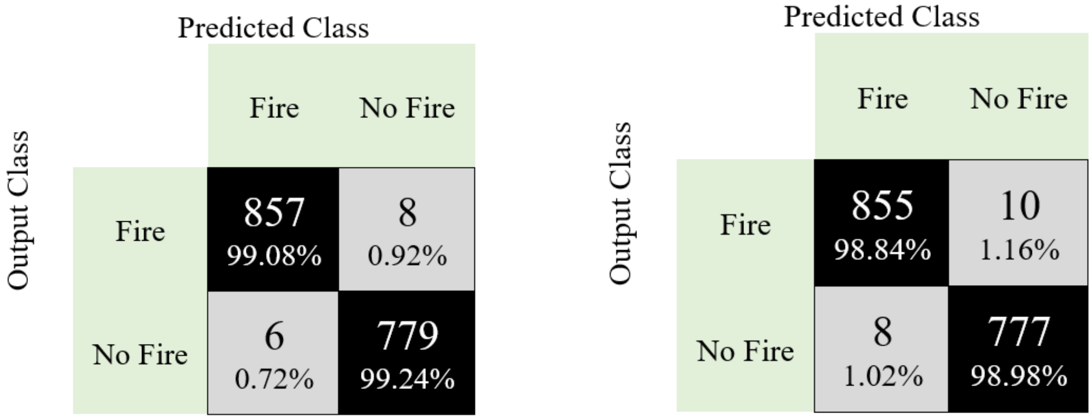


Fig. 8. Plot of best accuracy obtained versus the generated eight prediction vectors.

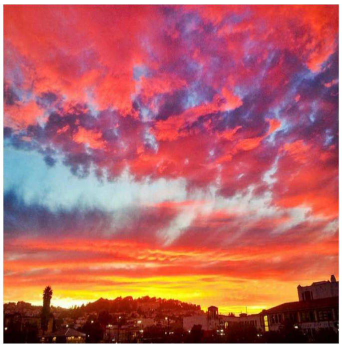
Moreover, the confusion matrices obtained using our two ensemble ResNets are denoted in Figure 9.



(a) Ensemble ResNetV1 (b) Ensemble ResNetV2

Fig. 9. Confusion matrices were obtained using two proposed methods.

Figure 9 is the visualization version of Table 3 and the results are denoted comprehensively. Furthermore, some misclassified images are denoted in Figure 10 and 11 for Ensemble ResNetV1 and Ensemble ResnetV2, respectively.


Real= No fire, Predicted: Fire


Real= Fire, Predicted: No Fire

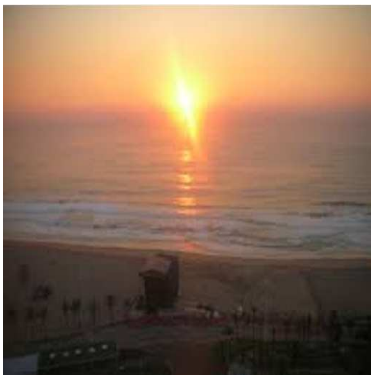

Real = No Fire, Predicted: Fire


Real= No Fire, Predicted: Fire


Real= Fire, Predicted: No Fire

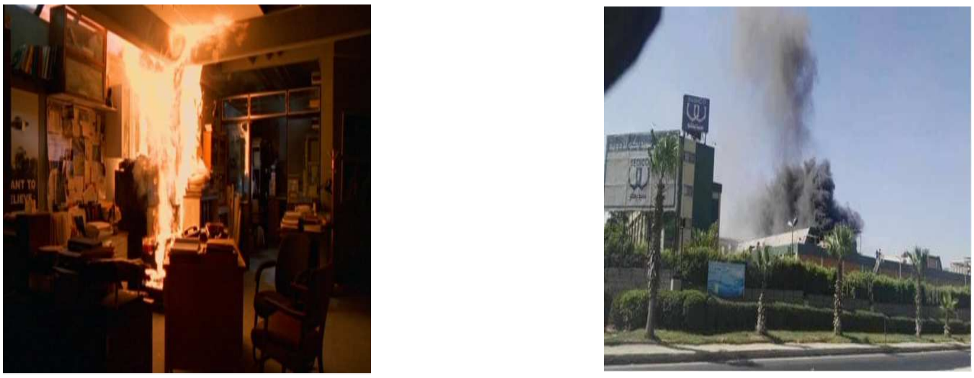
Fig. 10. Few misclassified images using Ensemble ResNetV1.

1


Real= No fire, Predicted: Fire


Real= No fire, Predicted: Fire


Real = No Fire, Predicted: Fire



Real= Fire, Predicted: No Fire   Real= Fire, Predicted: No Fire

Fig. 11. Few misclassified images using Ensemble ResNetV2.

Figure 10 and 11 depicts the false predicted sample images per the used ensemble network and these images are hard images for classification. Figure 9 clearly illustrates the high classification ability of the introduced ensemble ResNets.

In our proposed method, deep feature extraction was implemented using the transfer learning approach. Additionally, in our proposed method, the false positive rate is 0.0076 and 0.0092 for ensemble ResNetV1 and ensemble ResNetV2, respectively (see Figure 9-a and b). Considering these results, our results justify that the proposed ensemble ResNetV1 approach is accurate, reliable, and yield the highest automatic fire detection results.

The main advantages of our proposed methods are given below:

- Two novel ensemble deep learning mode have been proposed. In this work, novel ensemble ResNet models are presented.
- The presented two networks have been tested on the combined fire image dataset and these models have yielded 99.15% and 98.91% classification accuracies, for Ensemble ResNet V1 and Ensemble ResNet V2, respectively. In Ensemble ResNetV1, 1000 features have been used for classification while 256 features have been used in Ensemble ResNetV2. Moreover, we have used a shallow classifier (SVM).
- 10-fold CV technique is employed to develop a robust model.
- Programming of both these models is very easy. Therefore, users can use these models to solve their intended computer vision problems.

The limitation of this is given below:

- More fire images/ categories can be used.

- The presented model can detect fire using color and smokes (see Figure 8). More cognitive classification strategies can be proposed in the future.

7. Conclusions

In this work, we have proposed two novel ensemble models Ensemble ResNetV1 and Ensemble ResNetV2 to automatically detect the fire. In the first version (Ensemble ResNetV1), eight feature vectors are extracted using four pretrained ResNets and these features are merged. In order toe choose the most significant features, NCA selector has been used. In the second version (Ensemble ResNetV2) the generated eight feature vectors have been used individually and IHMV is conducted to use the advantages of the used ResNets. The performance of the developed model is validated using the combined two open-source fire image datasets downloaded from the Kaggle database. The introduced ensemble models attained 99.15% and 98.91% accuracies for Ensemble ResNetV1 and Ensemble ResNetV2 models, respectively. The main drawback of the proposed model is the use of a limited number of images to develop these two models. In the future, we plan to explore the possibility of using more images obtained from other databases.

In the future, a big fire image dataset can be collected or the available fire images can be increased using generative adversarial networks (GANs). Moreover, attention transformers-based fire detection models can be proposed to overcome limitations of this work.

Conflict of interest: The authors of this manuscript declare no conflicts of interest.

Funding: The authors state that this work has not received any funding.

References

Abedi Gheshlaghi, H., Feizizadeh, B., Blaschke, T., Lakes, T., & Tajbar, S. (2021). Forest fire susceptibility modeling using hybrid approaches. *Transactions in GIS, 25*, 311-333.

Barmpoutis, P., Papaioannou, P., Dimitropoulos, K., & Grammalidis, N. (2020). A review on early forest fire detection systems using optical remote sensing. *Sensors, 20*, 6442.

Barua, P. D., Chan, W. Y., Dogan, S., Baygin, M., Tuncer, T., Ciaccio, E. J., Islam, N., Cheong, K. H., Shahid, Z. S., & Acharya, U. R. (2021). Multilevel Deep Feature Generation

Framework for Automated Detection of Retinal Abnormalities Using OCT Images. *Entropy, 23*, 1651.

Barua, P. D., Muhammad Gowdh, N. F., Rahmat, K., Ramli, N., Ng, W. L., Chan, W. Y., Kuluozturk, M., Dogan, S., Baygin, M., & Yaman, O. (2021). Automatic COVID-19 detection using exemplar hybrid deep features with X-ray images. *International journal of environmental research and public health, 18*, 8052.

Chanthiya, P., & Kalaivani, V. (2021). Forest fire detection on LANDSAT images using support vector machine. *Concurrency and Computation: Practice and Experience*, e6280.

Chaudhry, S., Sidhu, G. P. S., & Paliwal, R. (2021). Restoring Ecosystem Services of Degraded Forests in a Changing Climate. *Handbook of Ecological and Ecosystem Engineering*, 353-375.

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics, 21*, 6.

de Almeida Pereira, G. H., Minoro Fusioka, A., Tomoyuki Nassu, B., & Minetto, R. (2021). Active Fire Detection in Landsat-8 Imagery: a Large-Scale Dataset and a Deep-Learning Study. *arXiv e-prints*, arXiv: 2101.03409.

Goldberger, J., Hinton, G. E., Roweis, S., & Salakhutdinov, R. R. (2004). Neighbourhood components analysis. *Advances in neural information processing systems, 17*, 513-520.

Gong, F., Li, C., Gong, W., Li, X., Yuan, X., Ma, Y., & Song, T. (2019). A real-time fire detection method from video with multifeature fusion. *Computational intelligence and neuroscience, 2019*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Hu, C., Tang, P., Jin, W., He, Z., & Li, W. (2018). Real-time fire detection based on deep convolutional long-recurrent networks and optical flow method. In *2018 37th Chinese Control Conference (CCC)* (pp. 9061-9066): IEEE.

Karadal, C. H., Kaya, M. C., Tuncer, T., Dogan, S., & Acharya, U. R. (2021). Automated classification of remote sensing images using multileveled MobileNetV2 and DWT techniques. *Expert Systems with Applications, 185*, 115659.

Khatami, A., Mirghasemi, S., Khosravi, A., Lim, C. P., & Nahavandi, S. (2017). A new PSO-based approach to fire flame detection using K-Medoids clustering. *Expert Systems with Applications, 68*, 69-80.

Kim, B., & Lee, J. (2019). A video-based fire detection using deep learning models. *Applied Sciences, 9*, 2862.

Kumar, A. (2021). Fire Detection Dataset, https://www.kaggle.com/atulyakumar98/test-dataset, 06.11.2021. In.

Liu, Z.-G., Yang, Y., & Ji, X.-H. (2016). Flame detection algorithm based on a saliency detection technique and the uniform local binary pattern in the YCbCr color space. *Signal, Image and Video Processing, 10*, 277-284.

Mahmoud, M. A., & Ren, H. (2018). Forest fire detection using a rule-based image processing algorithm and temporal variation. *Mathematical Problems in Engineering, 2018*.

Majid, S., Alenezi, F., Masood, S., Ahmad, M., Gündüz, E. S., & Polat, K. (2022). Attention based CNN model for fire detection and localization in real-world images. *Expert Systems with Applications, 189*, 116114.

Muhammad, K., Ahmad, J., Mehmood, I., Rho, S., & Baik, S. W. (2018). Convolutional neural networks based fire detection in surveillance videos. *IEEE Access, 6*, 18174-18183.

Nicolopoulos, N., & Hansen, E. (2009). How well prepared are Australian communities for natural disasters and fire emergencies? *Australian Journal of Emergency Management, The, 24*, 60-66.

Norman, B., Newman, P., & Steffen, W. (2021). Apocalypse now: Australian bushfires and the future of urban settlements. *npj Urban Sustainability, 1*, 1-9.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics, 9*, 62-66.

Pickrell, J. (2021). *Flames of Extinction: The Race to Save Australia's Threatened Wildlife*: Island Press.

Piera, P. J. Y., & Salva, J. K. G. (2019). A wireless sensor network for fire detection and alarm system. In *2019 7th International Conference on Information and Communication Technology (ICoICT)* (pp. 1-5): IEEE.

Poyraz, A. K., Dogan, S., Akbal, E., & Tuncer, T. (2022). Automated brain disease classification using exemplar deep features. *Biomedical Signal Processing and Control, 73*, 103448.

Saeed, F., Paul, A., Karthigaikumar, P., & Nayyar, A. (2019). Convolutional neural network based early fire detection. *Multimedia Tools and Applications*, 1-17.

Sahal, R., Alsamhi, S. H., Breslin, J. G., & Ali, M. I. (2021). Industry 4.0 towards Forestry 4.0: Fire detection use case. *Sensors, 21*, 694.

Saied, A. (2021). FIRE Dataset, https://www.kaggle.com/phylake1337/fire-dataset?select=fire_dataset, 06.11.2021. In.

Sharma, J., Granmo, O.-C., Goodwin, M., & Fidje, J. T. (2017). Deep convolutional neural networks for fire detection in images. In *International conference on engineering applications of neural networks* (pp. 183-193): Springer.

Vapnik, V. (1998). The support vector method of function estimation. In *Nonlinear Modeling* (pp. 55-85): Springer.

Vapnik, V. (2013). *The nature of statistical learning theory*: Springer science & business media.

Wang, D., Guan, D., Zhu, S., Mac Kinnon, M., Geng, G., Zhang, Q., Zheng, H., Lei, T., Shao, S., & Gong, P. (2021). Economic footprint of California wildfires in 2018. *Nature Sustainability, 4*, 252-260.

Warrens, M. J. (2008). On the equivalence of Cohen's kappa and the Hubert-Arabie adjusted Rand index. *Journal of classification, 25*, 177-183.

Xu, R., Lin, H., Lu, K., Cao, L., & Liu, Y. (2021). A Forest Fire Detection System Based on Ensemble Learning. *Forests, 12*, 217.

Yuan, C., Liu, Z., & Zhang, Y. (2017). Fire detection using infrared images for UAV-based forest fire surveillance. In *2017 International Conference on Unmanned Aircraft Systems (ICUAS)* (pp. 567-572): IEEE.