



Fire images classification based on a handcraft approach

Houda Harkat^{a,b,*¹}, José M.P. Nascimento^{a,c,2}, Alexandre Bernardino^{d,3}, Hasmath Farhana Thariq Ahmed^{e,4}

^a Instituto de Telecomunicações, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

^b Center of Technology and Systems (CTS-Uninova), NOVA University of Lisbon, 2829-516 Monte de Caparica, Portugal

^c Instituto Superior de Engenharia de Lisboa, IPL, Lisbon, Portugal

^d ISR – Instituto de Sistemas e Robotica, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

^e School of Computer Science Engineering, Vellore Institute of Technology, Chennai, Tamilnadu 600 127, India



ARTICLE INFO

Keywords:

Wildfire
Higher-order features
Feature selection
Support Vector Machine (SVM)
Radial Basis Function (RBF)

ABSTRACT

In recent years, wildfires and forest fires have ravaged millions of hectares of the forest all over the world. Recent technological breakthroughs have increased interest in computer vision-based fire classification that classifies fire and non-fire pixels from image or video datasets. Fire pixels from an image or video can be classified using either a traditional machine learning approach or a deep learning approach. Presently, the deep learning approach is the mainstream in forest fire detection studies. Although deep learning algorithms can handle vast amounts of data, they ignore the variation in complexity among training samples and as a result, their training model performance is limited. Furthermore, deep learning approaches with little data and features perform poorly in real-world challenging fire scenarios. As a result, the current study adopts a machine learning technique to extract higher-order features from the processed images from the publicly available datasets: Corsican dataset and FLAME, and a private dataset: Firefront_Gestosa, for classifying fire and non-fire pixels. It should be emphasized that in machine learning applications, handling multidimensional data to train a model is challenging. Feature selection is used to overcome this problem by removing redundant or irrelevant data that has an impact on the model's performance. In this paper, information-theoretic feature selection approaches are used to choose the most important features for classification while minimizing the computational cost. The traditional machine classifier, Support Vector Machine (SVM) is adopted in the present work, that works on the discriminative features input selected from the feature selection technique. The SVM performs the classification of fire and non-fire pixels with a Radial Basis Function (RBF) kernel, and the model's performance is measured using assessment measures such as overall accuracy, sensitivity, specificity, precision, recall, F-measure, and G-mean. The model draws an overall accuracy of 96,21%, a sensitivity of 94,42%, a specificity of 97,99%, a precision of 97,91%, a recall of 94,42%, an f-measure and g-mean values of 96,13% and 96,19% respectively.

1. Introduction

In recent times, millions of hectares of forest are devastated globally by wildfire or forest fire (“The Amazon in Brazil is on fire - how bad is it?”, Bhujel, Maskey-Byanju, & Gautam, 2017; “CNN (2020) California wildfires have burned an area almost the size of Connecticut,”; Mannan et al., 2017; Nolan et al., 2020). Forests safeguard the earth’s natural

equilibrium. Unfortunately, forest fires are typically detected after they have spread over a broad region, making their control and extinguishment challenging, although not impossible, at times (Alkhateeb, 2014). With recent technological advancements, computer vision-based fire detection gains research interest in detecting surface fire and crown fires (Guan, Min, He, Fang, & Lu, 2022). Detection of the latter seems challenging as surface fire occurring in the forest surface can easily be

* Corresponding author at: Instituto de Telecomunicações, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal.

E-mail addresses: houda.harkat@usmba.ac.ma (H. Harkat), jose.nascimento@isel.pt (J.M.P. Nascimento), alex@isr.tecnico.ulisboa.pt (A. Bernardino), hasmath.farhana@vit.ac.in (H.F.T. Ahmed).

¹ ORCID: 0000-0002-7827-1527.

² ORCID: 0000-0002-5291-6147.

³ ORCID: 0000-0003-3991-1269.

⁴ ORCID: 0000-0002-9159-7804.

detected using smoke sensors. However, crown fires occur if a surface fire is not extinguished quickly enough and produces large flames. In such a situation no sensor can resist the extreme heat created by a crown fire, thus the sensing task becomes complex (Chowdary, Gupta, & Singh, 2018). Therefore, the early detection of crown fire necessitates monitoring the forest area via satellites or UAV through which videos or images are captured continuously. Furthermore, the vision-based approaches perform the detection task from the video or image input and help in alerting the firefighter's team to reach the destination. Besides, avoiding the widespread of fires, minimizing the economic and financial loss to the human and forest lives.

Fire detection from the image or video can be performed either via a classical machine learning approach (Bot & Borges, 2022) or a deep learning approach (Bouguettaya, Zarzour, Taberkit, & Kechida, 2022; Houda Harkat, Nascimento, Bernardino, & Thariq Ahmed, 2022; Majid et al., 2022). The classical machine learning approach perquisite appropriate feature extraction and selection techniques for better performance. On the other hand, the deep learning approach can be adopted for automatic feature extraction and selection through which it performs the classification (Farhana Thariq Ahmed et al., 2019). With massive volume of data, manual feature extraction becomes unfeasible due to its inefficiency of capturing discriminative feature information. Furthermore, handcrafted techniques perform inefficiently and are unreliable due to their low classification performance with increasing data size. Contrarily, deep learning methods can handle large data size, however, these approaches do not account for the change in complexity across training samples. Hence, limiting the effectiveness of their training models and increasing misclassification that degrade the model's performance (Guan et al., 2022).

Moreover, in real-world complex fire scenarios, deep learning approach with limited data and features, shows declined performance. Thus, the present research adopted machine learning technique to extract higher order features (Swami, Mendel, & Nikias, 1998) from the image dataset for classification of fire pixels from non-fire pixels.

The present work implemented higher order cumulant features as it is resilient to Gaussian noise that is present in the initial data. Higher order statistical features are widely used in medical diagnosis from images ("National Institute of Open Schooling, Ministry of HRD, Govt. of India," 2010; Vijithananda et al., 2022), however not explored much in remote sensing applications like fire monitoring and detection. Thus, higher order cumulants, order 3, are evaluated in the present work, that extract the cumulant coefficients from the images relying on an unbiased approach.

It is to be noted, that handling multidimensional data to train a model is challenging in machine learning applications. To solve this, feature selection is used to eliminate redundant or irrelevant data that has a marginal impact on the model's performance. In this work, information theoretic feature selection methods are adopted to selects the significant features for classification, meanwhile reduce the computational cost. The selected set of discriminative features are supplied as input to the Support Vector Machine (SVM) classifier. The images are classified as fire and non-fire images adopting an SVM classifier with Radial Basis Function (RBF) kernel and measure the performance of the model based on overall accuracy, sensitivity, specificity, precision, recall, F-measure and G-mean as assessment metrics.

The present work implements the detection task through the fire images acquired from publicly available datasets, specifically the Corsican dataset (Toulouse, Rossi, Campana, Celik, & Akhloufi, 2017) and FLAME (Shamsoshoara et al., 2021) dataset. Further, performed fire detection utilizing images acquired by the Firefront project team (M. Brown, Szeliski, & Winder, 2005; The FIREFRONT Project: Objectives and First Steps) during the Gestosa mission, Firefront_Gestosa (<https://firefront.pt/>).

The remainder of the paper is organized as follows: Section 2 reviews the related work, Section 3 explains the materials and methods of the proposed framework, detailing the Higher-Order Statistics (HOS)

feature extraction approach, the experimented feature selection algorithms, and the classification approach. Besides, the datasets adopted in the present study are explained with data preparation and processing techniques. Section 4 discusses the experimental setup, and the results of the flame detection are visualized and evaluated in detail. Section 5 summarizes and concludes the present work.

2. Related work

Existing Fire or smoke detection approaches through the literature are grouped into broadband categories: classification techniques and segmentation models (Bouguettaya et al., 2022). In a binary mode, the former technique classifies patches or miniaturized images as fire or non-fire. While the latter employs pixel-based clustering to generate a mask wherein the white pixels symbolize the fire distribution, and the remainder is considered as background. Segmentation based detection approach is time-consuming, as it necessitates highly configured GPU for training the deep learning model. Thus, the alternate solution is to reduce the dimensionality of the images prior feeding it as input to the deep learning model. Further, a challenging aspect is to locate the fire pixels from the aerial images from a few numbers of acquired images. Consequently, it will be challenging to apply dimensionality on such images as it might cause the training data unbalanced, affecting the classification performance. To address the challenge, a pre-classification procedure should be adopted for narrowing down the portions of fire or smoke pixels and localize them in a well-defined area prior the segmentation step. The original images will be scanned utilizing a sliding window-based method, with the selected windows classifying the images as fire or non-fire. The windows representing flame/smoke will then be fed into the segmentation model. Various classifiers have been suggested in the literature for the former task.

Fire detection is done adopting various techniques and features, for example, a fire detection study within a tunnel is performed with Histograms comparisons of the image dataset (Noda & Ueda, 1994). In another work, texture analysis is performed, fire and non-fire images are classified using SVM (Dimitropoulos, Barmpoutis, & Grammalidis, 2014). In a recent study forest fire detection is performed using wireless sensor network, that identify forest fires in their early stage through the implementation of a machine learning regression model for accurate detection of forest fire (Guan et al., 2022). Extracted features from the dataset have a high impact on the machine learning model performance. A rule-based classification approach is also adopted in some studies that compare the color intensity levels of the image for accurate detection of fire images (Celik & Demirel, 2009). A combination of texture analysis and color intensities are also adopted for the classification of fire images adopting Naïve Bayes and k-nearest neighbors (KNN) classifiers (Chino, Avalhais, Rodrigues, & Traina, 2015). However, Color-based fire detection techniques are prone to false alarms, due to the impact of fluctuating light intensities influenced by external environmental conditions. To minimize the false alarms, red (R) and saturation (S) components values are compared for the HSI (hue, saturation, intensity) and RGB (red, green and blue) threshold values that predefined to classify fire pixels from the colored pixels (Gong et al., 2019). The studies that explicitly specified a feature selection approach accomplished it by applying iterative feature elimination to identify the best features.

There are several other works reported that minimize the false alarms by utilizing the Color values in hybrid to machine learning algorithms (Borges & Izquierdo, 2010; Foggia, Saggese, & Vento, 2015; Mueller, Karasev, Kolesov, & Tannenbaum, 2013; Sudhakar et al., 2020; Yuan, Ghamry, Liu, & Zhang, 2016).

The existing studies mostly emphasize front view datasets, which are subtly distinct from aerial datasets. With the classification approach, it is convenient to train the model for such images, as the area of fire is large and visually distinct. Nonetheless, in aerial images, the fire/smoke pixels are quite limited and sometimes difficult to distinguish without zooming in. There are only few works that examine the case of aerial

data sets so far. (Shamsoshoara et al., 2021) deploys an Xception architecture (Chollet, 2017) to distinguish fire and non-fire pictures. The model draws 76.23 % for test sets over the FLAME data (Shamsoshoara et al., 2021). Nonetheless, the model is computationally demanding and require an adequate GPU material to train the model.

(Dutta & Ghosh, 2021) had proposed a hybrid technique based on a simple separable convolutional neural network (CNN) model adopting regularization combined with an image processing phase with thresholding and segmentation. The tests over the FLAME dataset (Shamsoshoara et al., 2021) shows promising results. It records sensitivity of 98.10 % and a specificity of 87.09 %. However, this technique relies over an adequate choice of the threshold. Moreover, for the real implementation of the framework, more data is required, and the upper layers of the model need to be retrained.

Though the idea was to implement a low-cost fast prediction model with proper selection of feature extraction techniques to maximize the fire detection performance. Our model prerequisite manual intervention, and prior insights on the data. Further, demands a well-defined exploratory data analysis to tune the features in a way that the classifier can differentiate fire and non-fire images.

3. Methods and dataset

This section presents an overview of the techniques adopted in the present work for image-based classification of fire and the adopted dataset. Besides, an outline of the methodology adopted in the present work is explained in detail. Fig. 1 illustrate the present work methodology for classifying the fire pixels adopting handcrafted feature extraction and selection approach.

The present study adopted three different publicly available image datasets Corsican (Toulouse et al., 2017), FLAME (Shamsoshoara et al., 2021), and Firefront_Gestosa (<https://firefront.pt/>). Besides, for the

implemented machine learning model the three different datasets are combined to obtain a larger number of training samples. The dataset images are pre-processed before the classification task through mask generation and flame localization. The images of the adopted datasets are labelled manually, with the help of some experts in the field, and annotated accordingly with the data portioning technique. In fact, the FLAME and Corsican datasets come with ground truths pictures. However, the Firefront_Gestosa ground truths were generated manually with binary pixel labels based on human observation deploying “MATLAB ImageLabeler”. Afterwards, we had proceeded to a zooming over the flame positions (after localizing it based on the ground truths masks). So, at these positions we had extracted the fire patches. We had randomly patched the pictures again and annotated some non-fire patches. Images are then cropped and resized to 300x300 pixels with patch creation technique. The Firefront_Gestosa pictures were simply resized due to their limited area of fires, no further patching was applied on it. Thereafter, the resized pictures are partitioned into training, testing, and validation test. The resized pictures are then transformed using the Radon transform (Stanley, 1983) to reduce the dimensionality of the image from 2D to 1D.

The Radon transform is performed using a projection angle $\theta = 0^0$ to 360^0 and with a step size of 10^0 . This results in the formation of 37 angles for every image. The current study extracts HOS cumulant features from each projection and then concatenates the extracted features from all 37 angles into a feature vector. As it is impractical to directly feed a larger feature vector to the classifier, the current work adopted a feature selection step prior to classification. Consequently, the computational complexity of the deployed model is decreased. The present work experimentally evaluated the extracted features with the combination of information theoretic feature algorithms such as Mutual Information Feature Selection (MIFS) (Battiti, 1994), Min-Redundancy Max-Relevance (mRMR) (Nguyen, Chan, Romano, & Bailey, 2014; Peng,

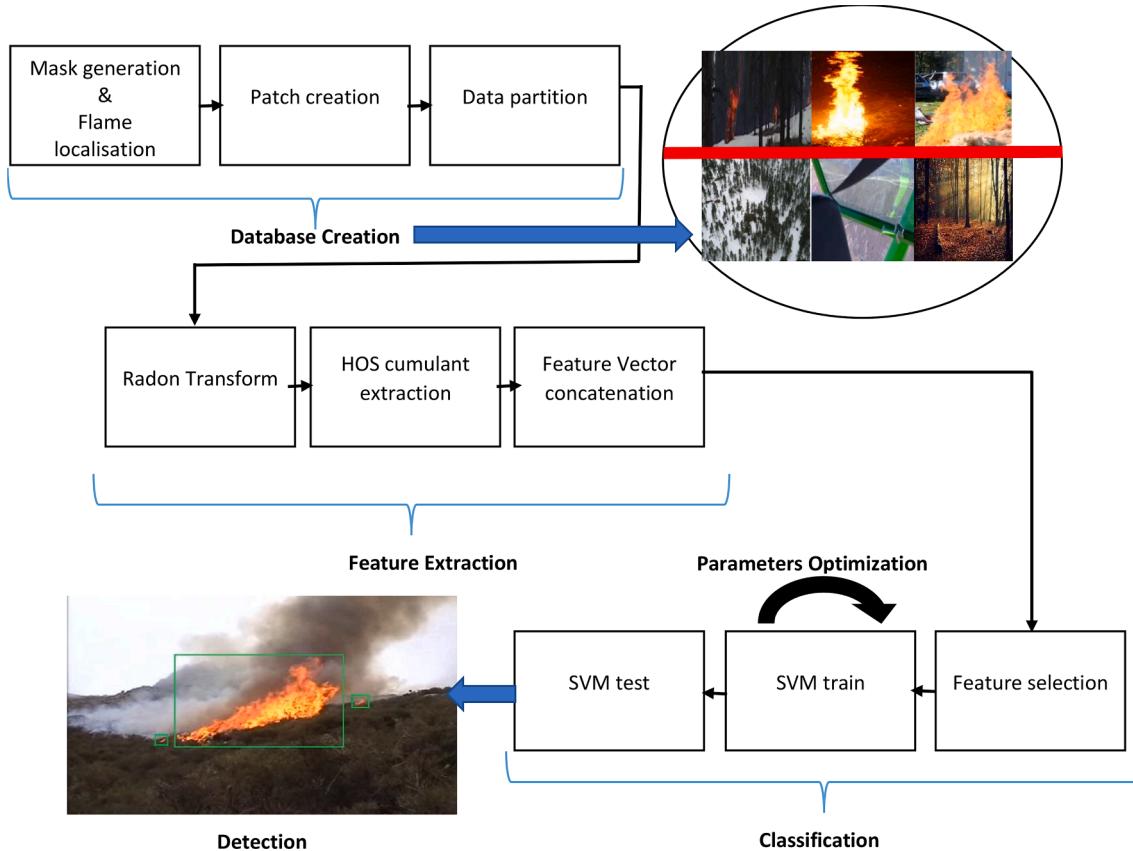


Fig. 1. The proposed framework for fire detection.

Long, & Ding, 2005), Conditional Infomax Feature Extraction (CIFE) (Lin & Tang, 2006), Joint Mutual Information (JMI) (Yang & Moody, 1999), Conditional Mutual Information Maximization (CMIM) (Fleuret, 2004), Double Input Symmetrical Relevance (DISR) (Meyer & Bontempi, 2006), Interaction Capping (ICAP) (Jakulin, 2005), and Conditional Redundancy (CONDRED) (G. Brown, Pocock, Zhao, & Luján, 2012). Lastly, the reduced set of optimal features are supplied as input to the SVM classifier that perform classification of the fire pixels. The SVM parameters are optimized in a loop procedure. Finally, the localization of fire pixels is done with the machine learning model, through sliding window approach. The classification performance of SVM is measured with overall accuracy, sensitivity, specificity, precision, recall, f-measure and g-mean as the assessment metric. More details on the dataset and data preparation technique are explained in detail with feature extraction, selection, and classification approach in the subsequent sections.

3.1. The dataset

The present study utilizes image data from the publicly available datasets, such as the Corsican dataset (Toulouse et al., 2017) and FLAME (Shamsoshoara et al., 2021), for classifying the fire pixels from non-fire pixels from the image. In addition, the classification task is incorporated over the images gathered by the FireFront project team during the Gestosa mission.

3.1.1. Description

Nearly 2000 wildfires were captured with multiple camera arrangements in the original Corsican dataset (Toulouse et al., 2017). The images were captured in the visible and near-infrared spectral bands at a resolution of 1024x768 pixels and saved in the portable network graphics (png) file format. In the present work 1175 images were selected from the total image data set to find heterogeneous flame color or with appropriate background settings (textures and light). A collection of multimodal pictures obtained with a 'JAI AD-080GE' camera are included in the dataset (Toulouse et al., 2017). Using the same aligned optics, this camera can capture images in the visible and near-infrared spectra at the same time. Every image in the Corsican dataset is labeled with its associated segmentation mask, which is created using a homography matrix transform-based picture registration technique.

Another publicly available aerial dataset with fire images used in this work is FLAME dataset (Shamsoshoara et al., 2021). The images in the FLAME dataset are collection of pictures captured with drones and are mostly made up of videos; some were converted into frames for segmentation propose. A Zenmuse X4S and a phantom 3 camera with a resolution of 640×512 pixels and a frame rate of 30 frames per second were used to create the required video. The collection consists of 2003 images that have already been annotated with ground truths for each image.

Firefront Gestosa is indeed another aerial dataset, however private and unlabeled, acquired during the Gestosa mission by FireFront project team's (<https://firefront.pt/>). This dataset is considered to make public for the researchers in the near future. This dataset consists mostly of five primary recorded videos that range in length from two to three minutes. For more details about each video, please see Table 1 below. The very last video is a compilation of several perspectives from the first five videos, rather than an actual recording of the mission. Furthermore, this dataset presents a substantial challenge because it contains pictures that have been entirely hidden by smoke, making it difficult to discern the location of the flame from the rest of the image. For the experiments, the present work chose, two hundred and thirty-eight fire frames from the videos that are already available. 'MATLAB Image Labeler' was used to manually label the fire pictures.

3.1.2. The final dataset

The present work, firstly convert the image in the datasets into a

Table 1

Firefront_Gestosa dataset: The first five videos consist of five main recorded videos. The last video is a compilation of multiple perspectives from the first five videos. The videos have 24 bits per pixels and a video format type RGB24.

Video	Pixel Resolution	Number of Frames per second (Frame Rate)	Durations in seconds
PIC_081.MP4	1920x1080	50	182.88
PIC_082.MP4			163.68
PIC_083.MP4			178.56
PIC_085.MP4			66.24
PIC_086.MP4			191.04
Gestosa2019. MP4	1280x270	29.97	218.18

greyscale with the patch size 300x300. For the Corsican and FLAME datasets, a simple patching approach was employed to zoom through the fire position and compress the data to analyze. Starting with the ground truths, the algorithm localizes the fire pixels and flame positions. However, such an approach was not appropriate for Firefront Gestosa since the fire pixels in the pictures are quite limited. An example of the patches of the dataset is given in Fig. 2.

Table 2 summarizes dataset adopted in the present work. The Corsican dataset consists of only 1775 fire pictures, whereas the Flame dataset consists of 4785, in which there are 2003 fire pictures and 2782 non-fire pictures. The Firefront_Gestosa dataset, comprises of 1352 pictures, with 238 fire and 1113 non-fire pictures. Besides, sun pictures consist of 65 non-fire pictures. Sun pictures are introduced to simulate high intensity pictures (mainly sunrise, sunset, and some nature pictures) that could make a confusion for the classifier, and hence reduce the False positives (FP).

The data of these datasets are fused together to form a new dataset with 4016 fire pictures and 4020 non-fire pictures for performing the classification task. Thus, the present work contributed a fused dataset and perform the classification task for a total of 8036 pictures.

Fire pictures that are captured from exceptionally high altitudes are also included in aerial datasets, providing a very distant picture of the flame. As a result, the task of classification seems more challenging. Besides, the images were patched with resolution of 300x300 for covering the entire flame pixels that eliminated inconsistent data processing, preserving essential information.

3.2. Classification approach

Fire pictures are classified based on machine learning approach. For accurate classification the images in the adopted dataset are passed through several steps of processing for making the classification task computationally optimal. The images are pre-processed with Radon transform and then higher order cumulant feature extraction and selection is performed for making the input suitable for SVM classifier. The mathematical formulation is introduced the subsequent sections, while the implementation details on the pre-processing phase, feature extraction, selection and classification are outline in section 4.

3.2.1. Pre-processing

Radon transform is performed as a pre-processing step on the images of the adopted dataset. An image's Radon transform (Stanley, 1983) is the sum of each pixel's Radon transformations. This method separates the pixels in the image into four subpixels and then projects each of the subpixel separately. According on the distance between the projected location and the bin centers, each subpixel's contribution is proportionally split between the two nearest bins. The bin on the axes computes the full value of the subpixel, or one-fourth the value of the pixel if the subpixel projection strikes the center point of a bin.

3.2.2. Cumulants feature extraction

Higher-order cumulants allows to extract statistical features from

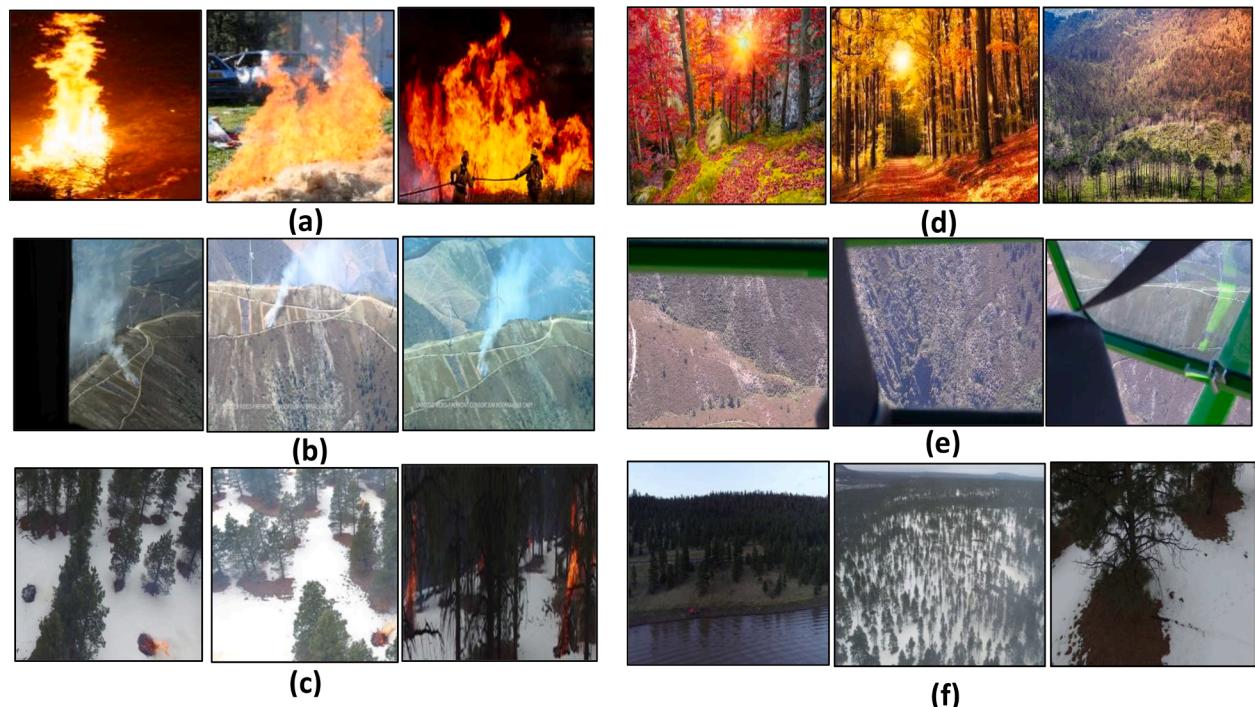


Fig. 2. Examples of fire patches included in our simulation dataset. Patches depicting flames from (a) Corsican data, (b) Firefront_Gestosa, and (c) FLAME. Some sunset pictures were injected into the final set of data, some samples are depicted in (d). Patches representing negative samples from (e) Firefront_Gestosa, and (f) FLAME.

Table 2

Statistics about the final dataset used in this work, in terms of the number of samples.

Dataset	Corsican Dataset	FLAME Dataset	Firefront_Gestosa Dataset	Sun Pictures	Total
Fire pictures	1775	2003	238	0	4016
Non-Fire Pictures	0	2842	1113	65	4020
Total	1775	4845	1351	65	8036

non-gaussian time series distributions (Mendel, 1991; Nikias & Mendel, 1993).

Let $\{s(t)\}$ be an m^{th} order stationary time series of a random signal; the m^{th} order cumulants are expressed as the joint m^{th} order cumulants of $s(t) + s(t+\tau_1), \dots, s(t+\tau_{m-1})$. The mathematical formulation, formulated relying on the Leonov-Shiryaev theory (Leonov & Shiryaev, 1959), is given by the following equation:

$$C_{m,s}(\tau_1, \tau_2, \dots, \tau_{m-1}) = \text{Cum}[s(t) + s(t+\tau_1) + \dots + s(t+\tau_{m-1})] \quad (1)$$

where $\tau_1, \tau_2, \dots, \tau_{m-1}$ are defined as time lags or shifts and $\text{Cum}[\cdot]$ refers to the cumulants operator.

The n^{th} time lag is a multiple of the sampling period T_s : $\tau_n = n * T_s$. Following the Leonov-Shiryaev theory, which mathematically

proves that the cumulants could be expressed in function of their corresponding moments (Leonov & Shiryaev, 1959), the 2nd, 3rd, and 4th order cumulants for $s(t)$ could be expressed as:

$$C_{2,s}(\tau_1) = E[s(t)s(t+\tau_1)] \quad (2)$$

$$C_{3,s}(\tau_1, \tau_2) = E[s(t)s(t+\tau_1)s(t+\tau_2)] \quad (3)$$

$$\begin{aligned} C_{4,s}(\tau_1, \tau_2, \tau_3) = & E[s(t)s(t+\tau_1)s(t+\tau_2)s(t+\tau_3)] - E[s(t)s(t+\tau_1)]E[s(t+\tau_2)] \\ & + E[s(t)s(t+\tau_3)] - E[s(t)s(t+\tau_2)]E[s(t+\tau_1)s(t+\tau_3)] \\ & - E[s(t)s(t+\tau_3)s(t+\tau_1)s(t+\tau_2)] \end{aligned} \quad (4)$$

Admitting the assumption that the time series $s(t)$ is a real, random, and zero-mean process.

Where $E[\cdot]$ denotes the statistical expectation value operator.

In fact, every cumulant could be rewritten as a correlation between the original signal and its time-shifted versions. Thus, the unbiased estimates of the 2nd, 3rd, and 4th cumulants are given below:

$$\widehat{C}_{2,s}(k) = \frac{1}{N} \sum_{n=0}^{N-1} s(n)s(n+k) \quad (5)$$

$$\widehat{C}_{3,s}(k, l) = \frac{1}{N} \sum_{n=0}^{N-1} s(n)s(n+k)s(n+l) \quad (6)$$

$$\begin{aligned} \widehat{C}_{4,s}(k, l, m) = & \sum_{n=0}^{N-1} \frac{1}{N} s(n)s(n+k)^* s(n+l)^* s(n+m)^* - \sum_{n=0}^{N-1} \frac{1}{N^2} [s(n)s(n+k)^*][s(n+l)^* s(n+m)^*] \\ & - \sum_{n=0}^{N-1} \frac{1}{N^2} [s(n)s(n+m)^*][s(n+k)^* s(n+l)^*] \end{aligned} \quad (7)$$

The signal $s(n)_{n=0, \dots, N-1}$ is a finite N -sample vector. In this work, the Higher-Order Spectral Analysis Matlab Toolbox (Swami, Mendel, & Nikias, 1993; Swami et al., 1998) was deployed for feature extraction.

3.2.3. Feature selection

3.2.3.1. Information theory based ranking algorithms. The key reason for incorporating a feature selection step is that using the full set of extracted features has an adverse influence on the computing performance. In this case, feature selection algorithms aid by selecting a subset of appropriate feature inputs (G. Wang, Zou, Zhou, Wu, & Ni, 2016; Y. Wang, Jiang, Cao, & Wang, 2015; Zhang, Wei, Hu, & Kanhere, 2016).

Information theory feature ranking methods like MIFS (Battiti, 1994), mRMR (Nguyen et al., 2014; Peng et al., 2005), CIFE (Lin & Tang, 2006), JMI (Yang & Moody, 1999), CMIM (Fleuret, 2004), DISR (Meyer & Bontempi, 2006), ICAP (Jakulin, 2005), and CONDRED (G. Brown et al., 2012) permits reducing the features set to a specific K elite indices. The feature selection algorithms are incorporated using MATLAB feature selection toolbox (FEAST) (G. Brown et al., 2012). A brief insight about the feature ranking techniques, based on mutual information theory, is given in Table 3.

3.2.4. SVM classifier

SVM (Chang & Lin, 2011) is a widely used supervised learning technique that uses kernel strategies to perform classification tasks. Kernels aid in the transformation of data points and the formulation of an optimum decision boundary between class labels. SVM separates the linear data for binary classification using its linear and polynomial kernels. SVM with a Gaussian Radial Basis Function (RBF) kernel, on the other hand, enables multi-class classification and non-linear mapping of samples in n -dimensional space. SVM-RBF can also deal with any non-linear relationship between the more than one class labels and the feature variables.

In some input domains, RBF encodes the feature variables of two

Table 3
Description of feature selection algorithms.

Algorithms	Description
MIFS (Battiti, 1994)	MIFS is a greedy approach, in which only extremely informative features are chosen and form an optimal subset. To eliminate redundancy and uncertainty in the feature vector, it identifies the non-linear association between the feature and the output class (H Harkat, Ruano, Ruano, & Bennani, 2019; Hoque, Bhattacharyya, & Kalita, 2014; Mariello & Battiti, 2018).
mRMR (Nguyen et al., 2014; Peng et al., 2005)	It is an incremental feature selection algorithm that picks the features with the minimal redundancy and the most relevance and form an optimal subset of features.
CIFE (Lin & Tang, 2006)	Through reducing redundancies, it creates an optimal feature subset that maximises the class-relevant information between the features.
JMI (Yang & Moody, 1999)	It identifies conditional mutual information and eliminates irrelevant features to define joint mutual information among the features.
CMIM (Fleuret, 2004)	Only selects features if they include additional information and make the output class's classification easier.
DISR (Meyer & Bontempi, 2006)	Instead of looking at specific features and the information contained in it, DISR measures symmetrical relevance and merges all feature variables that express more data about the class label or output.
ICAP (Jakulin, 2005)	Features are chosen based on interactions and a grasping the feature set's regularities.
CONDRED (G. Brown et al., 2012)	Determines whether there is conditional redundancy between the features.

classes v_1, v_2 as:

$$K(v_1, v_2) = \exp\left(-\frac{\|v_1 - v_2\|^2}{2\zeta^2}\right) \quad (8)$$

where $\|v_1 - v_2\|^2$ denotes the model fit parameter and the squared Euclidean distance between the feature variables ζ . The present study uses SVM-RBF as an effective kernel choice because of its multi-class classification capability. Using Gamma and cost parameters, the SVM classifier analyses the performance of the feature selection process. Non-linear classification is performed with the Gamma and cost parameters, which measure the classification loss.

3.2.5. Assessment metrics

The metrics used to access the performance of the models are:

- Overall accuracy, expresses the number of samples correctly classified through the all the dataset, given by the following formula:

$$\text{Overallaccuracy} = \frac{TP}{TP + TN + FP + FN} \quad (9)$$

It is an expressive metric for balanced data, but it become misleading for imbalanced datasets. Where TP denotes the true positives and FP the false positives.

- Sensitivity:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (10)$$

- Specificity:

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (11)$$

- Precision: is a measure of correctness of a positive prediction, given by the following equation:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

- Recall: Called also the true positive rate, is the proportion of true positives correctly predicted. The metric is defined as:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

The precision/recall is directly affected by the parameters/hyper-parameters of the model. Nonetheless, a having a high precision means a low recall value and vice versa.

- F-measure: It is a kind of dilemma between recall and precision. The metric is defined as:

$$F - \text{measure} = 2 \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

- G-mean: The geometric mean is the root of the product of the sensitivity and specificity.

$$G - \text{mean} = \sqrt{\text{Precision} * \text{Recall}} \quad (15)$$

This measure considers imbalanced classes problem.

4. Implementation and evaluation

The present work performed classification of fire pixels from the image dataset adopting SVM classifier using the optimal feature subset

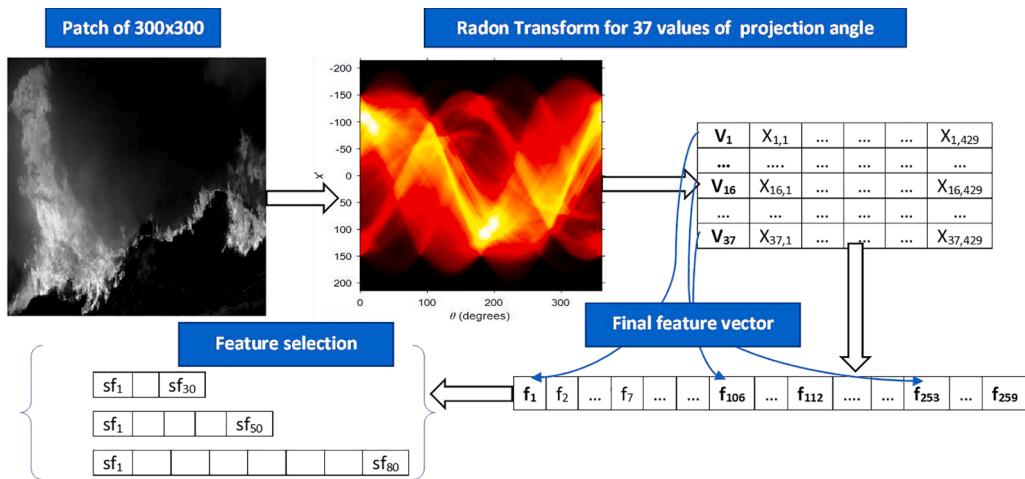


Fig. 3. System Overview: $V_i = [X_{i,1}, \dots, X_{i,429}]$ $i = \{1, \dots, 37\}$ represents a single radon projection corresponding to a defined angle in the set $[0^\circ:10^\circ:360^\circ]$, $f_1, \dots, 259$ is the final feature vector and sf is the set of selected features.

as illustrated in Fig. 3. The 2D images, of 300x300 pixels, are transformed to a single dimension times series using radon transform V_i $i = \{1, \dots, 37\}$, for different projection angles (namely 37 angles in the range of 0° to 360° with a step of 10° , resulting in 37 angles). 3rd HOS features are

extracted for every angle and the features are all concatenated in a single feature vector $f_1, \dots, 259$ with a corresponding 259 features. Then the feature set is reduced using a ranking approach, for 30, 50 or 80 features (sf). An SVM is used for classification.

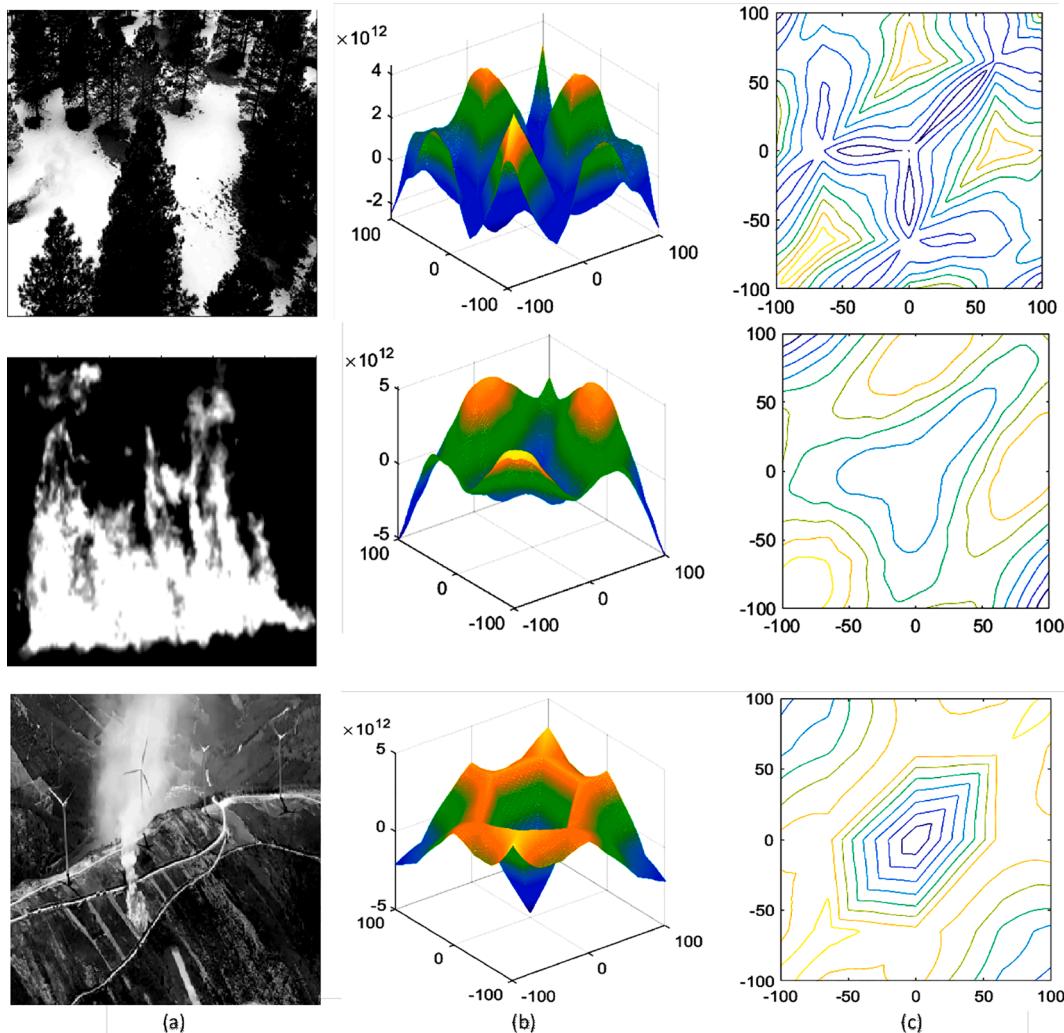


Fig. 4. Examples 3rd Order Cumulants estimates of fire patches from the three datasets: (a) represents the samples, the first is from FLAME, the second from Corsican and the last one from Firefront_Gestosa. (b) The magnitude and (c) contour plot of HOS cumulants spectra for Radon transform for angle $\theta = 360^\circ$.

4.1. Pre-processing

The subpixel value is shared evenly amongst the bins if the subpixel projection meets the border across two bins. In the present work, Radon transform is performed on the 2-D grayscale images of 300x300 as in Fig. 3, for angles in range of 0 to 360°, that project the image intensity along 37 values of radially oriented line at some specific angle. One can note that the Radon transform is performed for 37 angles (for every angle, it computes a 1D time series vector of size 429 elements), for every projection a feature vector of 7 features is extracted. Later, the computed feature vectors (features extracted from time series V1-V37) are concatenated to a single feature vector consisting of 259 elements. Lastly, feature selection is performed to extract a predefined optimal feature subset of 30/50/80 features. The subsequent sections explain in detail on feature extraction and selection with the classification approach, that measure the performance of the implemented model.

4.2. Feature extraction

In this work, the 3rd order cumulants was adopted since it allows an accurate characterization of texture variance in the pictures besides to their robustness against Gaussian noise (Etemad & Chellappa, 1998; Mendel, 1991; Swami et al., 1993, 1998). This was practically approved by conducting several experiments while tuning the parameter configurations.

The overall feature extraction procedure is summarized in Fig. 3. For every instance, a radon transform is performed for 37 projection angles ($0 = 0^\circ$ to 360° with a step of 10°). For every projection, a feature vector of seven elements is computed.

The final feature vector is the result of concatenation of all the feature vectors of the 37 projections of radon transform.

The final feature vector, for every instance, is normalized by z-score formula:

$$f'_{1-259} = \frac{f - \text{mean}(f)}{\text{std}(f)} \quad (16)$$

where f represents the final feature vector, with a size of 259 elements, $\text{mean}(f)$ and $\text{std}(f)$ are the mean value and standard deviation of the vector respectively, and f' the normalized vector.

An example of the two-dimensional feature distribution is shown in Fig. 4 for fire samples from the three datasets. We had adopted a time lag number equal to 100 for the example to show the geometrical structure of the distribution, while for our case we had used a time lag number equal to 3 to reduce the dimensionality of the features.

The Time-series are segmented into records of 128 samples each, avoiding overlap. An unbiased estimate of the 3rd order cumulants is computed for every segment and then the results are averaged.

4.3. Feature selection

The feature selection step is introduced on the extracted set of large number of features to reduce the computational complexity. The reduced feature set size K was chosen as 30, 50 or 80, in the present work; based on a trial-and test method. In fact, the tests shows that if $K \leq 30$ the classification performance drops significantly. Otherwise, if $50 \leq K \leq 80$, the SVM performance are higher, and the model presents more stable results. Nonetheless, with a higher value of K ($K < 80$), the performance (mainly the model accuracy) does not increase significantly however, the model complexity increases significantly. The model's performance is monitored based on the feature selection algorithm and the optimal feature subset size. The experimental trials reported better performance with ICAP and MIFS than the other selection techniques, hence explained below.

4.3.1. Interaction Capping (ICAP)

ICAP aims to gather redundancies among the feature pairs redundancy and penalizes features depending on the amount of redundant information as follows,

$$F_x = MI(F_x; C_O) + \sum_{F_y \in E} \min(0, I(F_x; F_y; C_O)) \quad (17)$$

If the interaction between the features and the output label or class seems to have $I(F_x; F_y; C_O) < 0$, that indicated information redundancies then feature F_x , will be penalized as in Equation (17) Nevertheless, the complementary association or relationship between the features and output class is neglected if $I(F_x; F_y; C_O) > 0$.

4.3.2. MIFS

Mutual information is used to measure the relationship between features in MIFS (Battiti, 1994), which is a greedy feature selection technique. With a regularization parameter, this selection paradigm determines the non-linear relationship between the selected features and the associated output class. This value eliminates redundant features and minimizes the uncertainty associated with the optimal and reduced feature subset. Finally, for multi-class classification tasks, the optimal subset of features containing pre-defined feature information will be fed into the SVM classifier.

MIFS performs better than other linear transformation-dependent approaches, however next to ICAP, as it considers the non-linear relationship between feature variables and the class labels. As a result, MIFS delivers strong generalization for training while also reducing computing time.

The MIFS algorithm adopts the following steps to ensure that each selected feature variable is informative:

1. Empty feature set is initialized with extracted set of features, \hat{E} .
2. Within the features present in the initial feature set, mutual information is computed between the feature variable, \hat{e} and class labels C .
3. The feature is selected as an optimal feature if the features are highly mutually related to the class labels.
4. A greedy feature selection strategy is employed for the pre-defined set of features – 30/50/80 as follows:

For every feature that are included in the initial feature set, mutual information is computed.

The features that compute high values of mutual information are finally chosen:

$$I(C : F) - \beta \sum_{\hat{e} \in E} I(F, \hat{e}) \quad (18)$$

Thus, the final feature subset with predefined feature value consists of features ranked on the order of mutual information present in it.

MIFS assures that it contains only relevant features by progressively adding them to the feature subset. Experiments were carried out by changing the values of β from 0 to 1 to reduce redundancy among the features.

4.4. Performance evaluation

The present study adopts a classical machine learning classifier, SVM with Gaussian RBF Kernel to measure the classification performance of the proposed framework. SVM classifier analyses the performance of feature selection strategies with the regularization and cost parameters. At a time, one value from {0.003, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100, 300} is set to cost, and each iteration takes all values in sequence, yielding $11 \times 11 = 121$ combinations, which are validated using 5×5 cross-validation. For the adopted datasets, Corsican dataset (Toulouse et al., 2017), FLAME (Shamsoshaara et al., 2021) and Firefront_Gestosa (<https://firefront.pt/>) dataset the classification accuracy of SVM incorporating all feature selection techniques such as MIFS, mRMR, CIFE, JMI, CMIM, DISR, ICAP, and CONDRED ($k = 30 / 50 / 80$ cumulant

Table 4

Performance of SVM classifier using different information theory based ranking approaches, and three different K values.

Feature selection technique	Number of features (K)	Overall accuracy	sensitivity	specificity	precision	recall	f-measure	g-mean
mRMR	30	95.78	94.00	97.56	97.47	94.00	95.70	95.76
	50	95.94	94.30	97.59	97.50	94.30	95.87	95.93
	80	96.06	94.42	97.69	97.61	94.42	95.99	96.04
CIFE	30	91.12	87.33	94.90	94.48	87.33	90.76	91.03
	50	93.92	91.71	96.12	95.94	91.71	93.78	93.89
	80	94.20	91.91	96.49	96.32	91.91	94.06	94.17
JMI	30	87.03	84.76	89.30	88.79	84.76	86.73	87.00
	50	89.92	88.52	91.32	91.06	88.52	89.77	89.91
	80	93.17	90.19	96.14	95.90	90.19	92.96	93.12
CMIM	30	95.31	92.95	97.66	97.54	92.95	95.19	95.28
	50	95.88	94.30	97.46	97.38	94.30	95.81	95.87
	80	96.08	94.32	97.84	97.76	94.32	96.01	96.06
DISR	30	83.40	79.58	87.21	86.15	79.58	82.73	83.31
	50	89.50	88.07	90.92	90.65	88.07	89.34	89.49
	80	92.89	89.59	96.19	95.92	89.59	92.65	92.83
ICAP	30	95.33	93.10	97.56	97.45	93.10	95.23	95.31
	50	96.17	94.37	97.96	97.88	94.37	96.10	96.15
	80	96.21	94.42	97.99	97.91	94.42	96.13	96.19
CONDRED	30	79.09	79.18	79.01	79.03	79.18	79.10	79.09
	50	84.63	86.95	82.31	83.08	86.95	84.97	84.60
	80	90.12	87.65	92.59	92.20	87.65	89.87	90.08
MIFS ($\beta=1$)	30	95.91	93.78	98.04	97.95	93.78	95.82	95.88
	50	96.03	94.07	97.99	97.90	94.07	95.95	96.01
	80	96.16	94.25	98.06	97.98	94.25	96.08	96.14

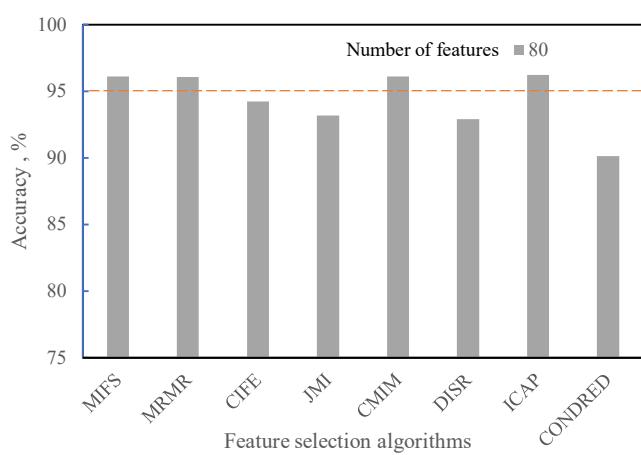


Fig. 5. Comparison of SVM performance with different feature selection algorithms.

feature) is measured and analyzed comparatively. Besides, for experimental evaluation the data is partitioned with 80 % for training and 20 % for testing. A multiclass SVM classification, the support vector machine (LIBSVM) software (Chang & Lin, 2011) was used.

4.4.1. Evaluation based on assessment metrics

The present study measures the performance using the SVM classifier with a range of information theoretic feature selection paradigm. Table 4 reports the measured assessment metric values of SVM with information theoretic feature selection algorithms for K set to 30, 50 and 80 optimal feature subsets. It is observed that with increasing number of feature inputs the performance of SVM increases significantly irrespective of the adopted feature selection criterion.

Fig. 5 shows the performance of SVM with the optimal feature subset derived from the information theoretic feature selection algorithms such as MIFS, MRMR, CIFE, JMI, CMIM, DISR, ICAP and CONDRED. The results are shown for the maximum feature subset value of 80 as maximum performance is attained for this value. SVM measured accuracy values more than 95 % with MIFS, MRMR, CMIM and ICAP. Nevertheless, ICAP and MIFS stand out from the other adopted feature selection paradigm.

Table 5 reports the maximum accuracy values attained by SVM with 80 feature inputs of ICAP and MIFS ($\beta = 1$). ICAP reported maximum accuracy values with 30, 50 and 80 optimal features. Secondly, MIFS with ($\beta = 1$) reported better accuracy with varying optimal feature subset. SVM with 80 feature inputs of ICAP reported an overall accuracy value of 96.90 %, 98.30 % and 99.13 % for Corsican + Sun, Firefont_Gestosa and FLAME datasets respectively. With MIFS ($\beta = 1$) feature input, SVM reported 96.90 %, 98.89 % and 99.01 % for Corsican

Table 5

SVM performance with ICAP and MIFS features over the three sets of data separately. We note that Corsican + Sun refers to Corsican data combined with the 65 samples of sun pictures for $k = 80$.

Feature selection technique	Dataset	Overall accuracy	sensitivity	specificity	precision	recall	f-measure	g-mean
ICAP	Corsican + Sun	96.90	100	-	96.89	100	98.42	98.43
	Firefont_Gestosa	98.30	92.02	99.64	98.21	92.02	95.01	95.75
	FLAME	99.13	98.30	99.72	99.60	98.30	98.95	99.01
MIFS ($\beta = 1$)	Corsican + Sun	96.90	100	-	96.89	100	98.42	98.43
	Firefont_Gestosa	98.89	95.38	99.64	98.27	95.38	96.81	97.49
	FLAME	99.01	98.05	99.68	99.54	98.05	98.79	98.87

+ Sun, Firefont_Gestosa and FLAME datasets respectively. Since it is not reasonable to measure the performance with only accuracy as a key metric, the present study computed the values for the assessment metric such as sensitivity, specificity, precision, recall, f-measure and g-mean. Those metrics take in matter the imbalanced problem and reflect more the real performance of the approach. The principal aim is to boost the recall and specificity without really affecting the precision and sensitivity values. In fact, recall could not determine the number of negatives incorrectly labeled as positives, as the same as precision that could not claim the number of positive incorrectly classified. In real situations, attaining a higher recall value doesn't directly lead to higher precision. Models with a huge margin between recall and precision are not well trained/optimized. F-measure gives a kind of dilemma to adjust recall and precision with a weighting policy. G-mean measure the degree of inductive bias while considering classes weighting to have balanced performance. Following the results presented in Table 4. The models trained using 80 features and the ranking technique ICAP draws a similar sensitivity and recall value of 94.42 % as well as the one trained with mRMR, but the model with ICAP draws a higher specificity, precision, f-measure, and g-mean values of 97.99 %, 97.91 %, 96.13 % and 96.19 % respectively. The same logic, for the model trained with 80 and the ranking technique CMIM, that draws higher accuracy value of 96.08 %, in comparison with mRMR. However, this model draws lower specificity and recall values and higher value of specificity, precision, f-measure, and g-mean of 97.84 %, 97.76 %, 96.01 %, and 96.06 %.

Several experiments were run using MIFS algorithms, with $K = 80$, and different β values. The curves of evolution of the different assessment metrics in function of the values of the parameter β , are given in Fig. 6. The overall accuracy reaches the maximum of 96.16 % value for $\beta = 0.3, 0.9$ and 1 (Fig. 6.a.). The sensitivity and the recall (Fig. 6.b. and Fig. 6.e.) have the same evolution curve shape and attend a peak (a value of 94.25 %) for $\beta = 0.9$ and 1. The specificity and precision (Fig. 6.c. and Fig. 6.d.) as well have the same behavior, with a maximum at for $\beta = 0.3$ (values of 98.11 % and 98.03 % respectively). The f-measure and g-mean (Fig. 6.f. and Fig. 6.g.) curves copy the behavior of the accuracy curve with an extremum for a value of $\beta = 1$. We had opted for a $\beta = 1$ since the higher values of f-measure and g-mean, 96.08 % and 96.13 % respectively, are recorded at this point.

The performance of this model was directly compared to the model trained with $K = 80$ and ICAP for the overall dataset and for every set of data separately (please refer to Table 4 and Table 5). For Corsican + Sun the two models perform the same. Firefront_Gestosa is a one-class dataset with a limited sample, it is a bit difficult to judge the performance of our model depending only on this dataset. However, it seems from the results over FLAME set that the model trained using ICAP performs better. We had judged that ICAP ranking technique performs better following the recorded values over the overall dataset (the three sets, results recorded in Table 4).

4.4.2. Results comparison

Several other classifier types were tested, using a subset of 80 features selected using ICAP algorithm. The experimental evaluation of classifiers was done adopting, the 80 selected features of ICAP algorithm. as these values found to perform better (refer Table 4) with SVM and for a reasonable comparison. The following experiments were

performed using WEKA software (Frank et al., 2009) with fivefold cross validation strategy. For the different classifiers we had opted for the default setting proposed by the software.

The performance of a Naïve Bayesian (NB) network (Friedman, Geiger, & Goldszmidt, 1997), a Random Forest (RF) (Breiman, 2001) and a Decision Table (DT) (Kohavi, 1995) were assessed with the same metrics, as demonstrated in Table 6. The SVM outperforms the NB and DT approach. However, RF approach outperforms SVM in terms of overall accuracy, sensitivity, recall-measure, and g-mean. Nonetheless, SVM draws better results in terms of specificity and precision.

In fact, RF is more designated for multiclass problems, hence the current problem could be extended to three or four classes: flame, non-flame, smoke, and non-smoke. However, more experiments need to be run to further assess the performance of the RF classifier.

Even though SVM performs better, in the case of three or four classes scene, the initial problem required to be reduced to a set of binary classification problems.

4.4.3. State of art comparison

Accurate comparison of the results is difficult because the authors employ the same training parameters and validation modalities, even if they utilize the same data, but evidently not the same partition and assessment criteria. For Corsican data, we had regrouped in the Table 7 the most outstanding studies and their corresponding announced results.

The FLAME dataset, original work (Shamsoshoara et al., 2021) with Xception model, draws a training, validation and test accuracy of 76.23 %, 94.31 %, and 76.23 % respectively (Shamsoshoara et al., 2021). Using a separable CNN technique (Dutta & Ghosh, 2021), the authors had recorded an accuracy of 92 % and a specificity of 87.09 % over the test set of Flame data. A single-shot detector (SSD) based approach was deployed by (Qi & Chen, 2022), the model draws a mean average precision of 94.3 %. Our approach draws 99.01 % overall accuracy, a specificity of 99.68 % and a precision of 99.54 %.

Inception-v3 model (Sousa, Moutinho, & Almeida, 2020) draws an overall accuracy, a sensitivity, a precision values of 98.6 %, 98.6 % and 100 % respectively over Corsican set of data. SqueezeNet (Perrolas, Niknejad, Ribeiro, & Bernardino, 2022), draws a 98.56 % accuracy for training and 95.98 % for validation. Our model classifies properly all the samples over Corsican data but draw an overall accuracy of 96.90 % results over Corsican + Sun set.

Firefront_Gestosa is private set, no recorded work yet to compare with it. But the results are very encouraging since the area of fire is very limited within the pictures.

5. Conclusions

Detection and fighting wildfires had become very preoccupying since the numbers of destroyed forests increase every year. The implementation of more intelligent systems for helping the firefighters had becomes a priority. Several prototypes had been presented in the last years. Firefront, is an initiated projects that aims to develops an intelligent behavior for localizing fires based on drones that scans the area with high risks to localize the fire source before alerting the firefighters. Hence reducing the cost of dislocation and improving their service. The presented work aims to test a handcraft classifier to determine and

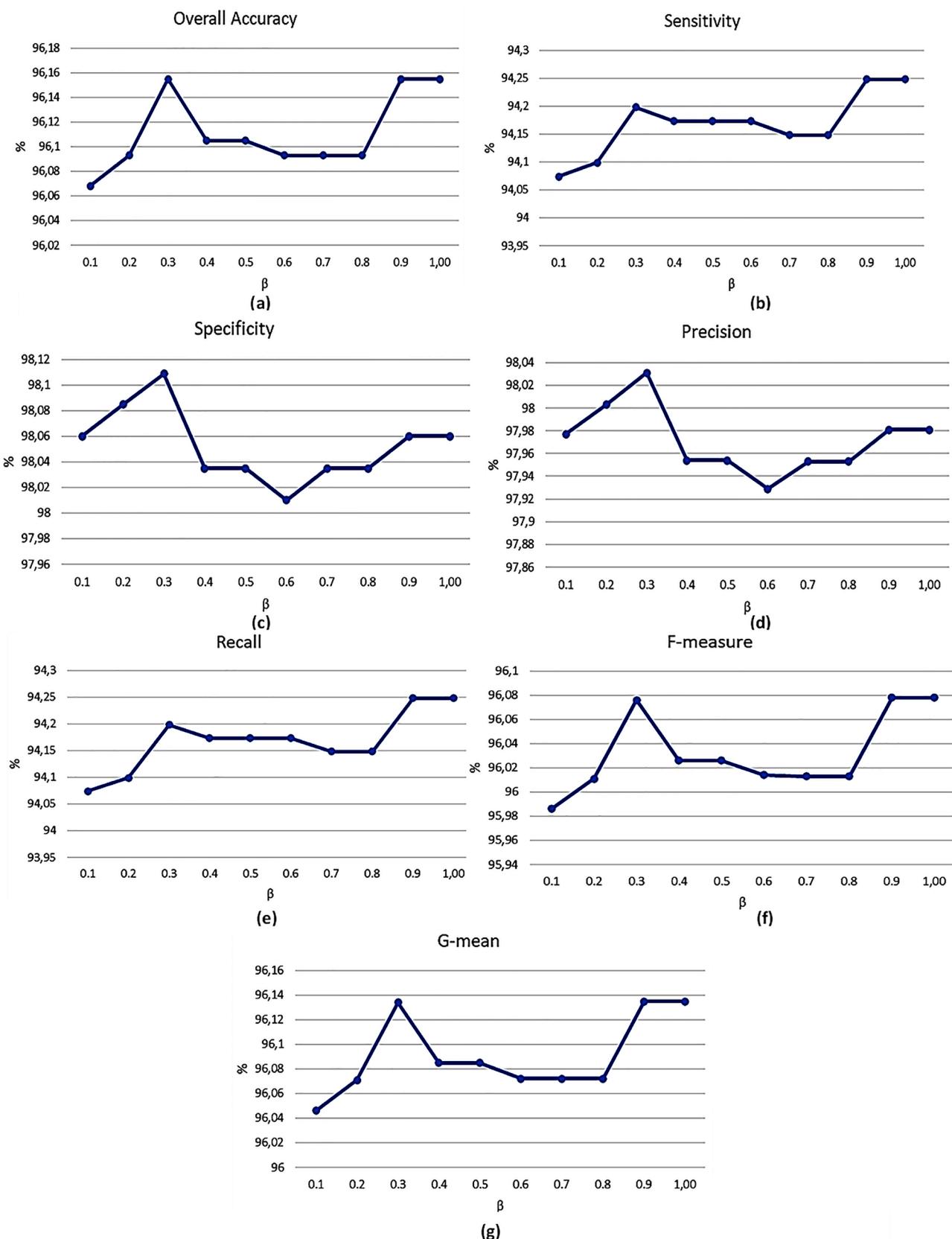


Fig. 6. Performance evolution in function of the parameter β , for the case of 80 features selected using MIFS algorithm: (a) Overall accuracy (b) Sensitivity (c) specificity (d) precision (e) recall (f) f-measure and (g) g-measure.

Table 6

Performance of the feature engineering technique using common classifiers.

Approach	Overall accuracy	sensitivity	specificity	precision	recall	f-measure	g-mean
SVM	96.21	94.42	97.99	97.91	94.42	96.13	96.19
NB	79.3	84.74	73.85	79.6	79.3	79.2	79.45
RF	96.60	95.86	97.26	96.60	96.60	96.60	96.60
DT	80.2	72.06	88.33	81	80.02	80.1	80.6

Table 7

Comparison of the state of art classification algorithms experimented over FLAME and Corsican datasets.

Approach	Dataset	Overall accuracy	Sensitivity	Specificity	Precision	Recall	F-measure	G-mean
Xception (Shamsoshoara et al., 2021)	FLAME	76.23 % for Test 94.31 % for Validation 96.79 % for Training	—	—	—	—	—	—
Separable CNN (Dutta & Ghosh, 2021)		Test 92 %	—	Test 87.09 %	—	—	—	—
SSD based model (Qi & Chen, 2022)		—	—	94.3 %	—	—	—	—
Our approach		99.01 %	98.05 %	99.68 %	99.54 %	98.05 %	98.79 %	98.87 %
Inception-v3 model (Sousa et al., 2020)	Corsican	98.6 %	98.6 %	—	100 %	—	—	—
SqueezeNet (Perrolas et al., 2022)		98.56 % for Training 95.98 % for Validation	—	—	—	—	—	—
Our approach: Corsican + Sun		96.90 %	100 %	96.89 %	100 %	98.42 %	—	—
Our approach	Firefront_Gestosa	98.89 %	95.38 %	99.64 %	98.27 %	95.38 %	96.81 %	97.49 %

classify pictures with probability of having flames already on. The tests were based on aerial pictures already known with their limited fire pixels. The used SVM classifier was based on HOS cumulant features ranked with ICAP to reduce the complexity of the system. With the same set of features, some tests were performed using other classifier types. Some preliminary results show a bit higher result using Random Forest classifier. Deep tests will be realized using Random Forest and SVM to further improve the current results.

CRediT authorship contribution statement

Houda Harkat: Conceptualization, Methodology, Software, Investigation, Writing – original draft, Writing – review & editing. **José M.P. Nascimento:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition. **Alexandre Bernardino:** Conceptualization, Writing – review & editing, Funding acquisition. **Hasmath Farhana Thariq Ahmed:** Methodology, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Houda Harkat reports financial support was provided by FCT. Houda Harkat reports a relationship with Foundation for Science and Technology that includes: funding grants.

Data availability

Data will be made available on request.

Acknowledgment

Houda Harkat acknowledges the financial support of the Portuguese Foundation for Science and Technology under the Project FIREFRONT with reference PCIF/SSI/0096/2017 and the Programmatic Financing of the CTS R&D Unit (UIDP/00066/2020), Jose Nascimento acknowledges the support of the Portuguese Foundation for Science and Technology under Project UIDB/50008/2020. The authors would like to thank UAVision and the Portuguese Air Force that run the flights for data acquisition.

References

- Alkhatib, A. A. (2014). A review on forest fire detection techniques. *International Journal of Distributed Sensor Networks*, 10(3), Article 597368.
- The Amazon in Brazil is on fire – how bad is it? . Retrieved from <https://www.bbc.com/news/world-latin-america-49433767>.
- Battiti, R. (1994). Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks*, 5(4), 537–550.
- Bhujel, K. B., Maskey-Byanju, R., & Gautam, A. P. (2017). Wildfire Dynamics in Nepal from 2000–2016. *Nepal Journal of Environmental Science*, 5, 1–8.
- Borges, P. V. K., & Izquierdo, E. (2010). A probabilistic approach for vision-based fire detection in videos. *IEEE Transactions on Circuits and Systems for Video Technology*, 20 (5), 721–731.
- Bot, K., & Borges, J. G. (2022). A systematic review of applications of machine learning techniques for wildfire management decision support. *Inventions*, 7(1), 15.
- Bouguettaya, A., Zarour, H., Taberkit, A. M., & Kechida, A. (2022). A review on early wildfire detection from unmanned aerial vehicles using deep learning-based computer vision algorithms. *Signal Processing*, 190, Article 108309. <https://doi.org/10.1016/j.sigpro.2021.108309>
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Brown, G., Pocock, A., Zhao, M.-J., & Luján, M. (2012). Conditional likelihood maximisation: A unifying framework for information theoretic feature selection. *Journal of Machine Learning Research*, 13(Jan), 27–66.
- Brown, M., Szeliski, R., & Winder, S. (2005). Multi-image matching using multi-scale oriented patches. *Paper presented at the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*.
- Celik, T., & Demirel, H. (2009). Fire detection in video sequences using a generic color model. *Fire Safety Journal*, 44(2), 147–158.
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–27. <https://doi.org/10.1145/1961189.1961199>
- Chino, D. Y., Avalhais, L. P., Rodrigues, J. F., & Traina, A. J. (2015). Bowfire: Detection of fire in still images by integrating pixel color and texture analysis. *Paper presented at the 2015 28th SIBGRAPI conference on graphics, patterns and images*.
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Chowdary, V., Gupta, M. K., & Singh, R. (2018). A Review on forest fire detection techniques: A decadal perspective. *Networks*, 4, 12.
- CNN (2020) California wildfires have burned an area almost the size of Connecticut. Retrieved from <https://edition.cnn.com/2020/09/14/us/california-wildfires-monday/index.html>.
- Dimitropoulos, K., Barmpoutis, P., & Grammalidis, N. (2014). Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(2), 339–351.
- Dutta, S., & Ghosh, S. (2021). Forest fire detection using combined architecture of separable convolution and image processing. *Paper presented at the 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*.
- Etemad, K., & Chellappa, R. (1998). Separability-based multiscale basis selection and feature extraction for signal and image classification. *IEEE Transactions on Image Processing*, 7(10), 1453–1465.
- Farhana Thariq Ahmed, H., Ahmad, H., Phang, S. K., Vaithilingam, C. A., Harkat, H., & Narasingamurthi, K. (2019). Higher Order Feature Extraction and Selection for Robust Human Gesture Recognition using CSI of COTS Wi-Fi Devices. *Sensors*, 19 (13), 1–23.

- The FIREFRONT Project: Objectives and First Steps. (February 14th , 2020). Workshop. Coimbra. Retrieved from <http://users.isr.ist.utl.pt/~alex/firefront/FirefrontWorkshop2020.pdf>.
- Fleuret, F. (2004). Fast binary feature selection with conditional mutual information. *Journal of Machine learning research*, 5(Nov), 1531–1555.
- Foggia, P., Saggesse, A., & Vento, M. (2015). Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(9), 1545–1556.
- Frank, E., Hall, M., Holmes, G., Kirkby, R., Pfahringer, B., Witten, I. H., & Trigg, L. (2009). Weka-a machine learning workbench for data mining. In *Data mining and knowledge discovery handbook* (pp. 1269-1277): Springer.
- Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine learning*, 29(2), 131–163.
- 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.
- Guan, Z., Min, F., He, W., Fang, W., & Lu, T. (2022). Forest fire detection via feature entropy guided neural network. *Entropy*, 24(1), 128.
- Harkat, H., Nascimento, J. M. P., Bernardino, A., & Thariq Ahmed, H. F. (2022). Assessing the impact of the loss function and encoder architecture for fire aerial images segmentation using Deeplabv3+. *Remote Sensing*, 14(9), 2023.
- Harkat, H., Ruano, A., Ruano, M., & Bennani, S. (2019). GPR target detection using a neural network classifier designed by a multi-objective genetic algorithm. *Applied Soft Computing*, 79, 310–325.
- Hoque, N., Bhattacharyya, D. K., & Kalita, J. K. (2014). MIFS-ND: A mutual information-based feature selection method. *Expert Systems with Applications*, 41(14), 6371–6385.
- Jakulin, A. (2005). *Machine learning based on attribute interactions*. Univerza v Ljubljani.
- Kohavi, R. (1995). The power of decision tables. *Paper presented at the European conference on machine learning*.
- Leonov, V. P., & Shiryaev, A. N. (1959). On a method of calculation of semi-invariants. *Theory of Probability & its applications*, 4(3), 319–329.
- Lin, D., & Tang, X. (2006). Conditional infomax learning: An integrated framework for feature extraction and fusion. *Paper presented at the European Conference on Computer Vision*.
- 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, Article 116114.
- Mannan, A., Feng, Z., Ahmad, A., Beckline, M., Saeed, S., Liu, J., ... Ullah, T. (2017). CO2 emission trends and risk zone mapping of forest fires in subtropical and moist temperate forests of Pakistan. *Applied Ecology and Environmental Research*, 17(2), 2983–3002.
- Mariello, A., & Battiti, R. (2018). Feature Selection Based on the Neighborhood Entropy. *IEEE transactions on neural networks and learning systems*(99), 1-10.
- Mendel, J. M. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications. *Proceedings of the IEEE*, 79(3), 278–305.
- Meyer, P. E., & Bontempi, G. (2006). On the use of variable complementarity for feature selection in cancer classification. *Paper presented at the Workshops on applications of evolutionary computation*.
- Mueller, M., Karasev, P., Kolesov, I., & Tannenbaum, A. (2013). Optical flow estimation for flame detection in videos. *IEEE Transactions on image processing*, 22(7), 2786–2797.
- National Institute of Open Schooling, Ministry of HRD, Govt. of India. (2010). Retrieved from <https://www.nios.ac.in/>.
- Nguyen, X. V., Chan, J., Romano, S., & Bailey, J. (2014). Effective global approaches for mutual information based feature selection. *Paper presented at the Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*.
- Nikias, C. L., & Mendel, J. M. (1993). Signal processing with higher-order spectra. *IEEE Signal processing magazine*, 10(3), 10–37.
- Noda, S., & Ueda, K. (1994). Fire detection in tunnels using an image processing method. *Paper presented at the Proceedings of VNIS'94-1994 Vehicle Navigation and Information Systems Conference*.
- Nolan, R. H., Boer, M. M., Collins, L., Resco de Dios, V., Clarke, H. G., Jenkins, M., ... Bradstock, R. A. (2020). Causes and consequences of eastern Australia's 2019-20 season of mega-fires. *Global change biology*.
- Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 27(8), 1226–1238.
- Perrolas, G., Niknejad, M., Ribeiro, R., & Bernardino, A. (2022). Scalable fire and smoke segmentation from aerial images using convolutional neural networks and quad-tree search. *Sensors*, 22(5), 1701.
- Qi, M., & Chen, B. (2022). *Forest Fire Detection Algorithm Based on Aerial Image*. Paper presented at the Artificial Intelligence in China, Singapore.
- Shamsoshoara, A., Afghah, F., Razi, A., Zheng, L., Fulé, P. Z., & Blasch, E. (2021). Aerial Imagery Pile burn detection using Deep Learning: The FLAME dataset. *Computer Networks*, 193, Article 108001.
- Sousa, M. J., Moutinho, A., & Almeida, M. (2020). Wildfire detection using transfer learning on augmented datasets. *Expert Systems with Applications*, 142, Article 112975. <https://doi.org/10.1016/j.eswa.2019.112975>
- Stanley, R. (1983). Deans, "The Radon Transform and Some of Its Applications.
- Sudhakar, S., Vijayakumar, V., Kumar, C. S., Priya, V., Ravi, L., & Subramaniyaswamy, V. (2020). Unmanned Aerial Vehicle (UAV) based Forest Fire Detection and monitoring for reducing false alarms in forest-fires. *Computer Communications*, 149, 1–16.
- Swami, A., Mendel, J. M., & Nikias, C. L. (1993). *Hi Spec Toolbox: for Use with MATLAB: User's Guide*: Math Works.
- Swami, A., Mendel, J. M., & Nikias, C. L. (1998). Higher order spectral analysis toolbox, for use with MATLAB, The MathWorks.
- Toulouse, T., Rossi, L., Campana, A., Celik, T., & Akhloufi, M. A. (2017). Computer vision for wildfire research: An evolving image dataset for processing and analysis. *Fire Safety Journal*, 92, 188–194. <https://doi.org/10.1016/j.firesaf.2017.06.012>
- Vijithananda, S. M., Jayatilake, M. L., Hewavithana, B., Gonçalves, T., Rato, L. M., Weerakoon, B. S., ... Dissanayake, K. D. (2022). Feature Extraction from MRI ADC Images for Brain Tumor Classification Using Machine Learning Techniques.
- Wang, G., Zou, Y., Zhou, Z., Wu, K., & Ni, L. M. (2016). We can hear you with wi-fi! *IEEE Transactions on Mobile Computing*, 15(11), 2907–2920. <https://doi.org/10.1109/TMC.2016.2517630>
- Wang, Y., Jiang, X., Cao, R., & Wang, X. (2015). Robust indoor human activity recognition using wireless signals. *Sensors*, 15(7), 17195–17208. <https://doi.org/10.3390/s150717195>
- Yang, H., & Moody, J. (1999). Feature selection based on joint mutual information. *Paper presented at the Proceedings of international ICSC symposium on advances in intelligent data analysis*.
- Yuan, C., Ghamry, K. A., Liu, Z., & Zhang, Y. (2016). Unmanned aerial vehicle based forest fire monitoring and detection using image processing technique. *Paper presented at the 2016 IEEE Chinese Guidance, Navigation and Control Conference (CGNCC)*.
- Zhang, J., Wei, B., Hu, W., & Kanhere, S. S. (2016). Wifi-id: Human identification using wifi signal. *Paper presented at the 2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*.