Analysis of Imagery Sensors and Methodologies Towards Early Fire Detection

Maria João Santos Lopes de Sousa
maria.joao.sousa@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

November 2017

Abstract

Due to climate change, fire risk conditions are more frequent which precipitates the greater occurrence of these phenomena. With the escalation in the severity of these events, there is an increased urgency of diminishing the response time from emergency teams, so the rapid detection of fire outbreaks becomes imperative. This work aims to study different types of sensors and explore computer vision techniques to detect these incidents at an early stage. Two main lines of research are examined, to understand the merits and limitations of thermal and visible range cameras. For the first approach, a color segmentation heuristic was devised to understand the sensor response and the RGB encoding of the camera. Second, an intelligent systems approach centered on the study of visible range images using a deep neural network. Since this work is inserted in Project Firecamp2, to evaluate the performance of both methodologies, the test results are analyzed for a set of experimental scenarios with a focus on fire risk in a camping context. While the thermal imaging approach extends the sensing ability enabling faster detection, the deep learning approach with color images has the potential to provide more robust algorithms with greater generalization. Overall, the comparative analysis between the two approaches taken demonstrated that for a fire detection application the development of algorithms should encompass both thermal and visible range images.
Keywords: fire detection, thermal imaging, RGB imaging, deep learning, color segmentation, aerial images.

1. Introduction

Fire is a concerning natural hazard on a global scale, that has a brutal impact on the environment by disturbing the climate and natural ecosystems. This phenomenon has drastic effects on communities on a social and economic level, since it can lead to the loss of lives and material damages. Every year the Mediterranean basin is invariably under high fire risk, with recurrent wildfires that amount to a large number of hectares of area burned [3].

Over the years, there has been continued research focused on fire incidence in this region, to address fire risk assessment and identification [2]. However, despite placing certain areas under high alert based on quantification of risk due to past events, many studies overlook the urgency of diminishing the delay in response time from when a fire starts until first responders attend to it. Following this line of reasoning, there has been considerable interest in applying computer vision techniques to identify fire using color imaging as well as from the thermal infrared range.

Due to climate change, conditions of high ignition propensity are more frequent, which results in an increase in fire incidents. These are escalating in severity, which prompts the necessity of mitigating its devastating effects. By nature, fire occurrence is a process with a high degree of uncertainty, both on a spatial and temporal level, making it difficult to pinpoint the threat. Therefore, the immediate identification and prompt action from emergency response teams are key for minimizing its consequences. Currently, there are several projects implemented with the objective of reducing fire propagation by early detection, yet their success is still limited.

In this context, camping parks present an elevated risk because, on top of material damages, fire related incidents can have fatal consequences. The present work, is developed under the scope of project Firecamp 2, which centers on studying fire risk in areas dedicated to camping and caravanning activities. The investigation scope of this project is not limited to camping parks, and is extensible to fire risk situations in the wildland-urban interface. In this context, this work centered on analyzing different imaging sensory data and exploring computer vision techniques for detection of fire in-

cidents in an initial stage. The principal objective centers on evaluating the advantages of both color and thermal imaging sensors, and assess how image variability affects the performance of classification algorithms for fire detection. To accomplish this goal a set of experimental tests was conducted for image acquisition purposes.

2. Data Acquisition System

2.1. Hardware Overview

With the purpose of acquiring thermal and visual range aerial images, an integrated acquisition system was developed to be taken onboard a multirotor. Its principal components are:

- On-board computer - Raspberry Pi 3 Model B
- Color Camera - Raspberry Pi Camera Module V2
- Thermal Camera - FLIR Vue Pro 336 9mm

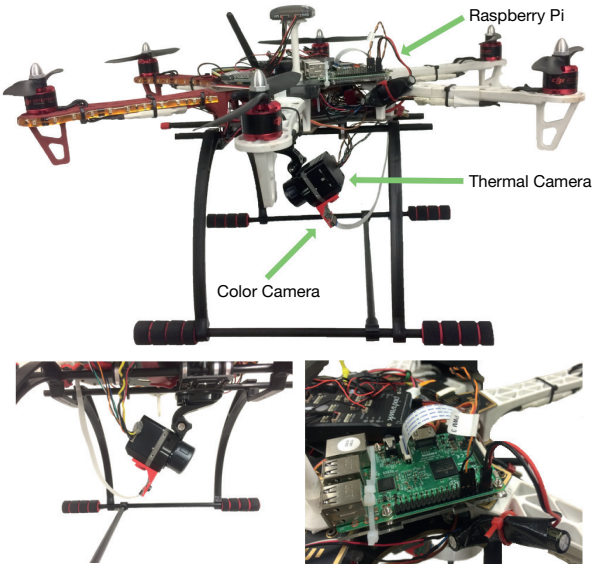


Figure 1: Image acquisition system mounted on a drone.

The Raspberry Pi 3 Model B was selected to be the on-board computer to handle image acquisition tasks. Its 8.5 x 5.6 cm^2 size facilitates the installation on the drone frame, and with 42 g it is easy to accommodate in the drone payload budget. For the videos from both cameras to be as synced as possible, both cameras are directly controlled by the Raspberry Pi. The thermal camera is hardware-triggered via PWM signal, and the color camera is software-triggered using the Picamera Python library. To avoid interference with drone sensors, wireless connectivity is turned off and Python scripts are run through secure shell (SSH) on the ground using an ethernet cable and stopped with the same protocol upon landing.

2.2. Thermal and RGB Cameras

All objects above absolute zero (0° Kelvin) emit thermal infrared energy. Since fire is a source of extreme temperatures, thermal imaging sensors are a clear choice for a fire detection application. For this reason, the system designed employs a thermal camera, i.e., FLIR Vue Pro. In order to acquire images from the visible range the Raspberry Pi Camera Module V2 was selected. This camera module allows for perfect integration with the Raspberry Pi, and due to its reduced weight it is a good choice for this application, since payload is a concern. The main specifications of these two cameras are summarized in table 1.

Table 1: Summary of the specifications of the thermal visible range camera.

	RPi Camera Module V2	FLIR Vue Pro
Size (mm)	25 x 23 x 9	63 x 44.4 x 44.4
Weight (g)	3.4	92.1 - 113.4
Sensor Resolution (px)	3280 x 2464	336 x 256
Focal Length (mm)	3.04	9.0
Horizontal FOV (°)	62.2	35
Vertical FOV (°)	48.8	27
Spectral Band (μm)	0.35 - 0.74	7.5 - 13.5

2.3. Experimental Tests

Since this work is developed under the scope of project Firecamp 2, the trials have an emphasis on fire risk situations related to the camping activity. While there was the intention of performing tests in an outdoor setting, due to the early arrival of fire season in 2017, for safety reasons, these tests to be conducted in June, were only carried out in indoor laboratory conditions, at the Laboratory of Forestry Fire Studies in Lousã.

Using a controlled environment in these experiments allowed for the acquisition of video sequences from static and moving platforms, as well as from different types of thermal cameras. In order to have a ground truth for this preliminary study, in addition to the two cameras presented in the previous sections, data were also recorded from an elevated platform using a radiometric thermal camera, FLIR SC660.

For the situations depicted in figure 4, an array of tests was performed under different conditions. Table 2 presents the summary of the tests performed.

The use of both radiometric and nonradiometric thermal cameras in the laboratory tests enables the development of an analysis of the sensor response, and understand how image statistics change for the set of experimental scenarios.

3. Thermal Imaging

Typically, thermal cameras only produce monochromatic images, unlike color cameras, which make use of 3 color channels, e.g., RGB. That is precisely the case of FLIR Vue Pro, which

(a) Setup 1 (b) Setup 2

Figure 2: Experimental setups used in the indoor trials.

Table 2: Summary of the experimental tests.

Experimental Tests				Data Acquired		
ID	Description	Conditions	Platform	Color	Thermal Radiometric	T. Nonradiometric
Test 1	Straw Burning	Static	Elevated Platform	✓	✓	✓
Test 2	Tent Burning	Static	Elevated Platform	✓	✓	✓
Test 3	Straw Burning	Moving	Drone	✓	✓	✓
Test 4	Boom Festival	Moving	Balloon			✓

in the raw format produces a 14-bit monochromatic image in grayscale with 16384 levels of intensity. By employing an array of proprietary camera features, the camera firmware processes the video from the monochromatic raw format to a video encoded in pseudocolor, according to the filter selected. First, the amounts of radiance registered by the sensor are converted to a grayscale image with intensity levels. Then, the firmware attributes to these levels a scale of colors, according to a specified palette.

In this work, the GrayRed filter was selected, which helps to draw attention to the hottest objects in a scene, by applying high-contrast color values with a divergent color scheme. While *a priori* it is unclear if employing this color palette specifically will yield any benefit in classification performance, FLIR preset filters will be used as it facilitates the development of future work, providing a basis for comparison of results. For the same reason, this color palette was applied to the raw data acquired with FLIR SC660, using FLIR ResearchIR software.

For both cases, the intensity level is given by a digital number assigned by the sensor analog to digital converter, the A/D counts. The difference is that for radiometric data this raw signal are converted to a temperature value according to the camera calibration. Nonetheless, these measurements should be corrected *a posteriori*, since the setup for experiments is not always the same. Consequently, the encoded radiometric data will not be accurate if the measurements are not properly adjusted.

3.1. Sensor Measurements

Using the FLIR ResearchIR software, the metadata encoded in the images can be visualized interactively, and we have access to the scale of temperatures for the radiometric case. These temperatures vary in each frame as the scale adjusts to the maximum temperature in the scene.

One of the main difficulties of interpreting thermal imaging data is the constant color adjustment of the image. In order to get a better sense of what is the output of the sensor prior to RGB encoding, it is convenient to first analyze how the raw digital signals and the respective temperature measurements evolve over time. Figure 3 presents this comparison for a fire ignition captured from an elevated platform using a radiometric camera.

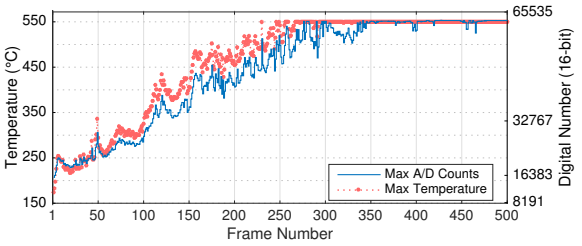


Figure 3: Temperature versus A/D Counts.

The most immediate observation from the graph is that the change in maximum temperature is highly correlated to the increase in the maximum value of A/D counts for each frame. Furthermore, from the graph in figure 3, it is interesting to note that around frame 350, the sensor output of maximum raw signals is close to the limit of the sensor resolution, $2^{16} = 65536$ levels, and this corresponds to a maximum temperature of 550°C.

In order to clarify what is the actual working temperature range, we will resort to the metadata encoded in the image files for both cameras. In the case of the radiometric camera, these are SEQ files, that include multiple sequential frames with the raw

digital values. Conversely, for the nonradiometric camera, both TIFF and JPEG files are used. The first can be of single or multiple frames, and the latter include both the RGB encoded image as well as the raw TIFF file. The metadata from these files can be extracted using Exiftool, an application capable of handling a variety of metadata formats.

By using Exiftool, it was possible to discover the temperature range of measurements for both cameras, as listed in table 3. These camera parameters are characteristic of each particular sensor.

Table 3: Camera Temperature Limits

Camera	FLIR SC660	FLIR Vue Pro 336 9mm
Temperature Range Max (°C)	500	135
Temperature Range Min (°C)	0	-25
Temperature Max Saturated (°C)	550	150
Temperature Min Saturated (°C)	-60	-60

Observing the temperature intervals of both cameras it is simple to verify that they have quite distinct sensing capabilities. This new information listed above is not provided in the official datasheets, but is extremely important. In what regards FLIR Vue Pro, the fact that the sensor saturates at 150°C can potentially be a limitation for a fire identification application. Considering that fire reaches much higher temperatures, a detection algorithm solely based on temperature would most likely be prone to false alarms.

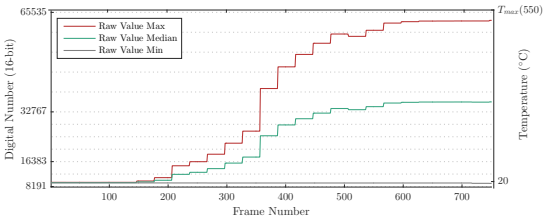
3.2. Adaptive Color Scale

As mentioned previously, understanding how the camera adapts the color scale is one of the principal difficulties in interpreting thermal imaging data, which is essential to devise a robust computer vision algorithm. For FLIR Vue Pro the adaptive color assignment depends on an array of camera features such as Digital Detail Enhancement (DDE), Active Contrast Enhancement (ACE) and Smart Scene Optimization (SSO). These algorithms tune the color assignment and can be adjusted manually using FLIR UAS mobile app. From the analysis of the metadata of the image and video files, two exif tags were identified that have paramount influence in adapting the color scale, the Raw Value Median and Raw Value Range. With the median and raw values from the metadata tags, the maximum and minimum values of the color scale are computed as follows:

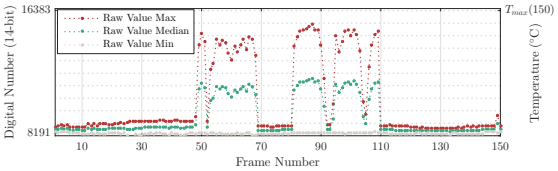
$$\text{Raw Max} = \text{Raw Median} + \frac{\text{Raw Range}}{2} \quad (1)$$

$$\text{Raw Min} = \text{Raw Median} - \frac{\text{Raw Range}}{2} \quad (2)$$

To investigate further the response of the sensor, and the influence of these parameters, two examples are given for the radiometric and nonradiometric camera in real operation conditions. The first case (fig. 4a) demonstrates the sensor response in a fire situation with extreme temperatures, for the radiometric camera. The second case (fig. 4b) illustrates the sensor response retrieved from nonradiometric data acquired with a balloon at high altitude at Boom Festival.



(a) Raw measured with FLIR SC600 for a controlled fire situation.



(b) Raw measured with FLIR Vue Pro for community kitchen area.

Figure 4: Sensor response in raw values for two distinct situations.

Figure 4a reveals the radiometric camera has a processing level that does not update the scale in every frame but rather in steps of 30 frames. Taking into account this device works at 15 Hz, the color scale is updated once every 2 seconds. Regarding figure 4b, considering this camera was programmed to take photographs with a 5 second time-lapse, it is unknown the time the firmware takes to adjust the color scale. Nevertheless, these results give insight into how the color scale expands when objects with high temperatures enter the field of view of the camera.

This initial analysis illustrates to some extent the principle behind how these devices work in and gave a better understanding of how the color scale adjusts. To further explore this aspect, the analysis will continue with a color segmentation approach.

4. Color Segmentation Heuristic

From the literature review, a common computer vision approach to fire detection is through color segmentation [1]. Before this work, under the scope of Project Firecamp 2, a preliminary analysis was performed using a segmentation heuristic to attempt to detect fire presence in a real-world setting. Now, with access to more data as well as radiometric information, this approach will be extended.

4.1. Color Segmentation

Considering that thermal images in pseudocolor are used to highlight differences in temperature, a segmentation heuristic should take into account the color bands used for such purpose. Recalling the images in figure 3.4, these colors are red, green and gray, as assigned by the GrayRed filter. The access to the color assignment of the filter is essential to devise the segmentation heuristic.

Figure 5: Division in color segmentation classes

The complete scale of 120 discrete color values, which depicts increasing temperature from gray to red, was divided into three segments. First, the gray segment includes 18 color levels, which represent 15% of the full scale. The mid-range is composed of 48 colors of green tones, which correspond to 40% of the scale. The remainder is the red part, that comprises the last 45% of the color palette, and is defined by 54 color levels.

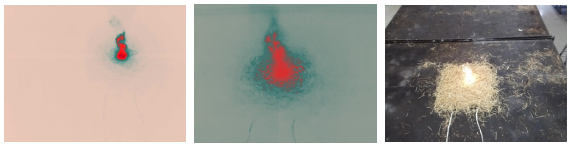
Table 4: Segmentation thresholds

Gamma	Gray			Green			Red		
Channels	R	G	B	R	G	B	R	G	B
Upper limit	253	199	185	143	169	157	255	73	71
Lower limit	149	171	160	98	90	86	103	89	85

To assess the advantages and limitations of thermal imaging the segmentation heuristic will be evaluated for the situations of the experiments tests. These cover three distinct scenarios of fire risk situations in camping parks. These situations encompass burning of forest fuels like straw, but also include highly combustible materials characteristic to the camping context.

4.2. Straw Example

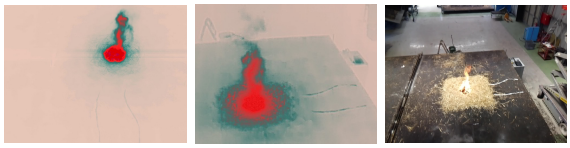
For the controlled ignition of straw, two types of tests were conducted. For the first case, videos were recorded with both radiometric and nonradiometric cameras from a static elevated platform. Fig. 6 depicts the data captured with these devices.



(a) Radiometric (b) Nonradiom. (c) Color

Figure 6: Images acquired with (a) FLIR SC660, (b) FLIR Vue Pro and (c) Pi Camera V2.

A second test again considers the burning of straw, with the radiometric camera positioned in an elevated platform. The nonradiometric image acquisition was, in this case, performed from a moving platform, using the UAV as presented earlier in figure 2a. Fig. 7 illustrates the data from the different points of view.



(a) Radiometric (b) Nonradiom. (c) Color

Figure 7: Perspectives with 3 different cameras.

Both cases presented so far comprise the same type of test, and data were recorded with the three devices, with the only difference being the point of view for the second case. These instances account for situations where the fire is visible to the naked-eye, thus could, in theory, be detected using color images also. Notwithstanding, one of the main advantages of using thermal imaging is the ability to see beyond the visible spectrum, thus being capable of detecting sources of thermal radiation to which there is no direct view. The following example aims to harness this advantage in another set of conditions.

4.3. Tent Example

In a camping environment, fires can originate from inside tents which means it is important to study these cases due to their elevated propagation potential. An important aspect of this type of instances is that the fire sources are obstructed by a somewhat opaque material. To evaluate this situation, a test was performed under static conditions on an elevated platform, again using the radiometric, nonradiometric and color cameras, as illustrated previously in fig. 2b. Fig. 8 illustrates the differences of data recorded with these devices.

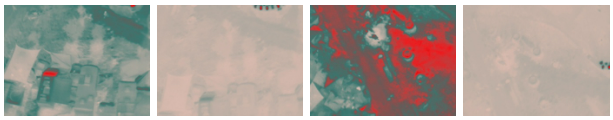
(a) Radiometric (b) Nonradiom. (c) Color

Figure 8: Image data from the burning of camping tent.

Comparing with the previous tests, now, at an early stage, the area in red does not depict the actual fire, but the heat built-up inside the tent. Whereas this is visible in fig. 8a and 8b, for the color image (fig. 8c) the fire will only be visible when it starts consuming the exterior of the tent.

4.4. Boom Festival Example

In addition to the previous tests, the segmentation heuristic was also tested on a dataset acquired at Boom Festival. Since these sequences are of thermal nonradiometric images at high altitude and of difficult interpretation, only a few selected samples are shown in fig. 9.



(a) Frame 48 (b) Frame 49 (c) Frame 80 (d) Frame 81

Figure 9: Heat detection with a nonradiometric thermal camera.

By inspecting the corresponding sequence of frames, presented in figure 9, what stands out is that the color scale adjusts when the same area enters the FOV of the camera. With knowledge of the venue premisses, it is possible to identify this place as the community kitchen area with rocket stoves. This is an area where fire detection is to be expected, in addition to other heat sources such as objects at elevated temperatures.

4.5. Results

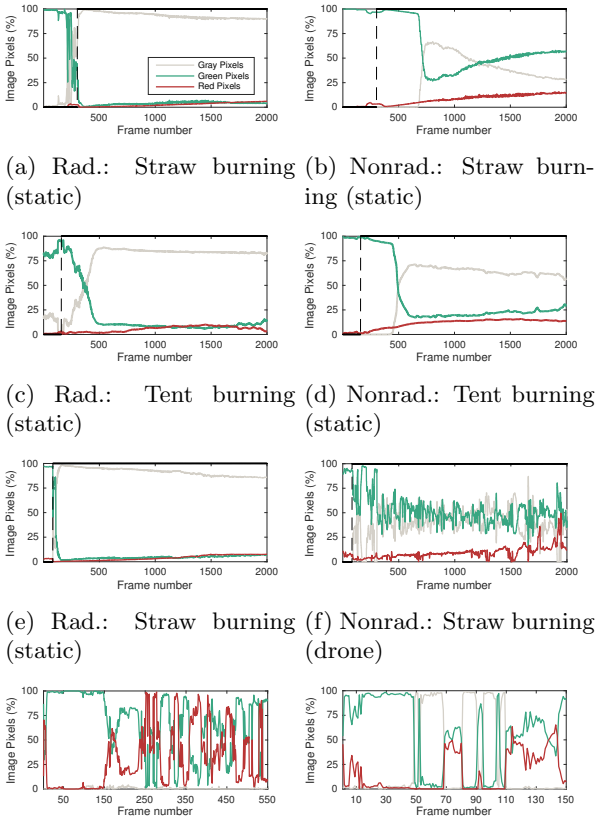
The results from this section pertain the study of the thermal cameras in terms of RGB encoding for the segmented classes. These results of all the tests are presented side-by-side in fig. 10, since it facilitates their interpretation. Yet, by comparing data of images from the two cameras, the effect of the difference in spatial resolution should be taken into account.

Starting with figure 10a and 10b, these correspond to the first straw example (Test 1) captured with FLIR SC660 and FLIR Vue Pro, respectively. Observing the variation in both graphs, it is evident the ignition provokes an abrupt change in the gray and green levels. However, considering the ignition occurs around frame 300, it is noticeable that the nonradiometric camera responds with some delay.

Moreover, examining the graphs from figures 10c and 10d (tent example, Test 3) the change in green and gray levels happens, once again, in an almost symmetric fashion. But, for this instance, the adaptation of the color scale to the fire incident takes effect more slowly. The increase in the percentage of red pixels occurs gradually but still at relatively low levels.

Looking at both of these tests, in which images were captured from a static platform, the variation of the color levels develops in an incremental manner, so it would be difficult to establish a threshold based on the change in consecutive frames. Additionally, the percentage of pixels in gray seems the most promising feature to identify the occurrence of fire. Still, the image does not adjust so abruptly in all cases, so if a low percentage of gray pixels is defined as a threshold, the classification algorithm would be prone to false alarms.

While the results of examples covered so far were obtained from laboratory tests in a controlled environment, under real operating conditions the movement of an aerial platform would cause far greater inconsistency in the color levels. This can be observed by comparing figures 10e and 10f, which depict the second straw example (Test 2) captured from a static platform and a drone, respectively.



(a) Rad.: Straw burning (static) (b) Nonrad.: Straw burning (static)

(c) Rad.: Tent burning (static) (d) Nonrad.: Tent burning (static)

(e) Rad.: Straw burning (static) (f) Nonrad.: Straw burning (drone)

(g) Boom Festival (baseline) (h) Boom Festival: community kitchen area

Figure 10: Color segmentation results.

Finally, examining the graphs from figures 10g and 10h, relative to the tests at the Boom Festival with the tethered balloon, the movement of the camera, as well as the set 5 second timelapse, have a considerable effect on the sensor response. Now, the variation is no longer incremental, and consecutive frames display abrupt changes in the percentage of either of the segmented colors.

Overall, the principal insight gained from this analysis is that the gray and green percentages adjust in a complementary way, as the color scale adjusts to the hottest objects in the scene. Furthermore, since for fire identification the ignition starts

with low spatial resolutions, the red percentage is usually the one with least influence for a fire detection heuristic.

These results demonstrate that the variability inherent to dealing with thermal images, which depict relative temperature differences, is a considerable challenge to the development of a robust classification algorithm. Although a set of heuristic rules could be fine-tuned to work for the situations presented, that type of algorithm would fail, if applied to real-world scenarios. As the tests from Boom Festival illustrated, real conditions account for much higher variability in image content. Granted that the data from the limited amount of tests is insufficient to build a reliable algorithm at this stage, the data available can still be explored using other techniques. Taking into account the complexity of this classification task, and its feature-rich characteristics, more powerful techniques should be employed in order to achieve better generalization.

To address this issue, the next section proposes an intelligent systems approach using deep neural networks. This type of architecture has recently achieved state-of-the-art results in computer vision tasks. These models are prepared to handle high-dimensional data, and spatial features of high complexity. Hence, the approach will depart from thermal imaging, and deal primarily with color images without prior feature extraction.

5. Deep Neural Networks

The second approach to address the fire detection problem explored the color imaging data, without prior feature extraction. This was achieved by resorting to a state-of-the-art deep neural network, i.e. the Inception V3 [5], and applying a transfer learning scheme. Transfer Learning is a machine learning method that consists of leveraging a previously optimized network, a pretrained model, and use it to learn an entirely new task, using a different dataset. Inception V3 has 5 types of layers for feature extraction purposes, i.e., fully-connected, convolutional, two pooling layers (max and average), the dropout layer, and a softmax classifier.

5.1. Fully-Connected Layer

The structure of the fully-connected (FC) layer is based on the architecture of classic neural network models like the multi-layer perceptron. In fully-connected layers, the neurons in adjacent layers have pairwise connections, but despite their name, neurons within a single layer do not share connections between them. Due to their amount of connections FC layers are very computationally expensive. Therefore, these are used sparingly in convolutional neural networks (CNNs), and usually at the penultimate layer before the classifier.

5.2. Convolutional Layer

Convolutional layers are the main building block of CNNs and are responsible for extracting representations from the data as it flows through the network. As illustrated in fig. 11, this type of layer computes the dot product between patches of the input volume and the weights of the filters. The dimension of the input and output volumes reflects the usual application of this type of layer that reduces data in spatial size and increases its depth.

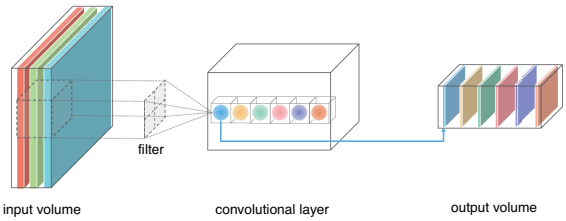


Figure 11: Example of a convolutional layer.

Due to their parameter sharing scheme (fig. 12), where neurons in each depth slice share the same weights, the network translation invariance. The output volume of this layer, in fig. 11, stacks the activation maps of the different slices of neurons, which extract distinct features.

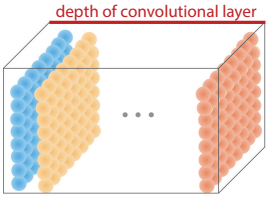


Figure 12: Parameter sharing scheme.

5.3. Pooling Layer

Pooling layers are introduced in the network to perform dimensionality reduction on the activation maps, by extracting only the most relevant features. Typically two types of pooling can be employed, max pooling and average pooling.

Pooling puts an emphasis on the relative location of features rather than their exact location, thus reducing the spatial size of the representations which consequently decreases the computational cost.

5.4. Dropout Layer

Deep neural networks, due to their complexity, are prone to overfitting when the amount of data available for training is limited. To address this issue, a regularization technique called dropout is used. By dropping a percentage of the hidden units at each training iteration, by zeroing its activations and discarding the influence of the respective connections to the output, the network learns redundant ways of mapping the inputs to the outputs, yielding improvements in generalization [4].

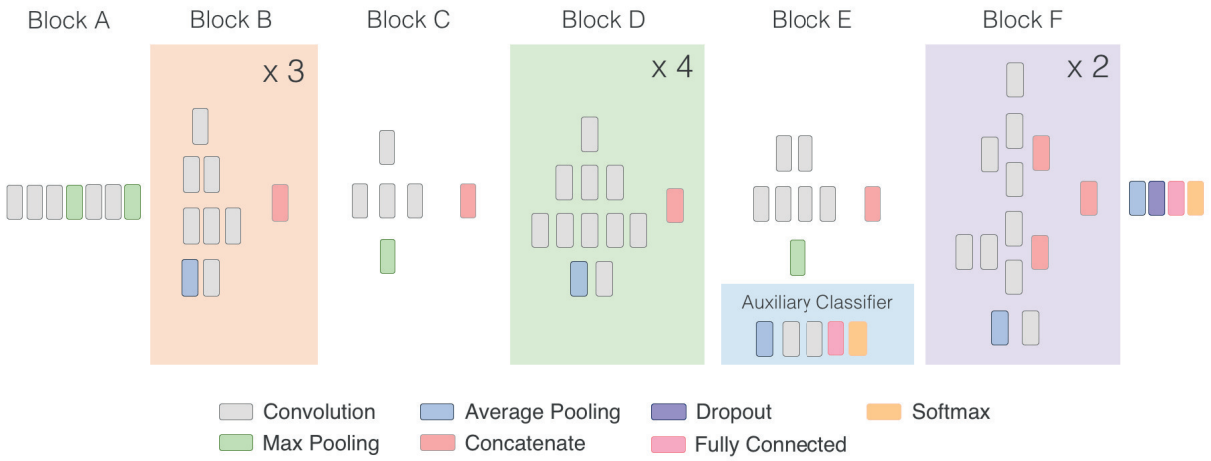


Figure 13: Inception V3 network architecture[1].

5.5. Softmax Classifier

The softmax is the score function responsible for generating the predictions of the model. This function is given by:

$$\hat{y}_k = \sigma(z)_k = \frac{\exp(z_k)}{\sum_i^K \exp(z_i)} \quad , \quad k = 1, ..., K \qquad (3)$$

with k representing the index of the class label. For every input, z_k, of the output layer, the softmax will generate K class prediction scores, $\hat{y}$.

5.6. Inception V3

The Inception v3 model (fig. 13) is a deep neural network that is built modularly, by means of a stacking of parallel convolution structures known as Inception modules. These employ several different convolution factorization schema depending on their relative depth location in the network and provide different hierarchical levels of feature extraction. The diagram with the compressed view of the network divides it in two parts, where the first is responsible for extracting spatial features (A-F) and the second is a classifier. To apply transfer learning the classifier part of the network will be retrained.

The formulation of the optimization problem for readjusting the parameters of last layers consists in finding the weights and biases, that define the mapping between the inputs and the correct outputs. To that end, the objective function will minimize the error between the prediction of the model and the ground truth labels.

The loss function quantifies the distance between the predicted outputs, $\hat{y}$, and the ground truth labels, y. For this model the loss is computed using the cross-entropy function, which is calculated as the sum of the ground truth labels, represented by $\mathbf{y}$, times the natural logarithm of the prediction scores, $\hat{\mathbf{y}}$:

$$L(\mathbf{y}, \hat{\mathbf{y}}) \; = \; -\sum_k y_k \ln(\hat{y}_k) \qquad (4)$$

[1]The diagram presented in the figure is a compressed view of the full architecture. This diagram makes a correction from the representation available in [6].

6. Implementation

Since those experiments account for three video sources only, additional color videos were added to the database to enable training the deep neural network. Then, in addition to the thermal image dataset, an augmented color dataset was created, and some samples can be observed in fig. 14.

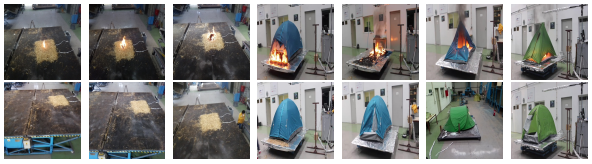


Figure 14: Samples from Firecamp2 Dataset.

6.1. Firecamp2 Dataset

To identify fire the problem is formulated as a binary classification where one class is *fire* and the second is *not fire*. In table 5 the division of the dataset is listed for these two classes.

Table 5: Dataset division

Camera	Fire	Not Fire	Total
Database	30682	27584	58266
Training Data	16000	16000	32000
Testing Data	5143	6407	11550
Straw burning	1699	301	2000
Tent burning	1624	5926	7550
Straw burning (drone)	1820	180	2000

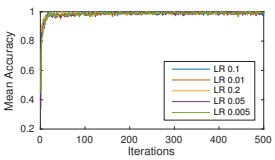
6.2. Data Preprocessing

Since in this work the focus is on evaluating performance of models in offline classification, the

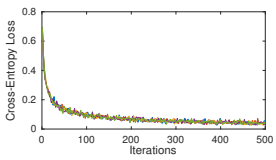
frames were extracted from video sources previously recorded. In VLC, using the scene filter the images were reduced to 299x299 with the original temporal resolution of the videos which varies depending of the video source. The Inception V3 model normalizes the input images to a [-2; 0] interval. This is done by dividing the input by 255, the maximum of a RGB channel , which normalizes the inputs to a 0 to 1 range. To apply a transfer learning approach, the images of training dataset are run through the network to extract the feature representation of the penultimate layer. This is a pooling layer with a tensor output of 2047 dimension, which is usually referred to as the bottleneck layer.

6.3. Model Retraining

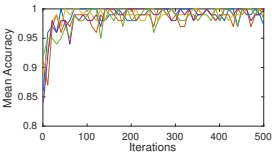
Since hyperparameter selection is an iterative process, a series of five distinct values of learning rate were tested: 0.005, 0.01, 0.05, 0.1, and 0.2. For these, training batch size was changed between 100 and 250, but it did not yield benefits in classification performance.



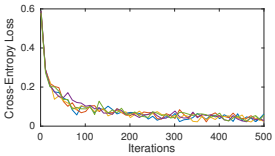
(a) Training accuracy



(b) Train. cross-entropy



(c) Validation accuracy



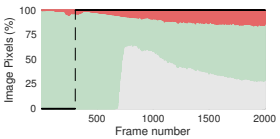
(d) Valid. cross-entropy

Figure 15: Comparison of models with different learning rate and batch size of 100.

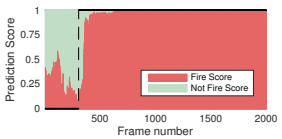
As can be observed in the training and validation curves in figure 15 for accuracy and cross-entropy, the variation of the performance of the models with the change in learning rate is not significant. Furthermore, the convergence of the cross-entropy loss to small values in the training and validation sets demonstrates the training procedure is successful.

6.4. Comparison of Results

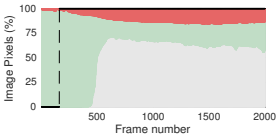
To test the performance of the model the situations analyzed are the same considered for the evaluation of the color segmentation heuristic, presented in section, with the exception of the Boom Festival example for which there were no color images. Whereas that approach used thermal imaging, now the color images captured in those experimental trials are used to assess if an approach using this type of data would be effective. Note that now, the levels of gray, green, and red are stacked to for ease of comparison.



(a) Nonradiometric: Straw burning (static)

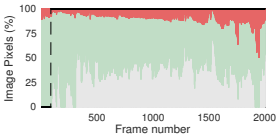


(b) Inception: Straw burning (static)

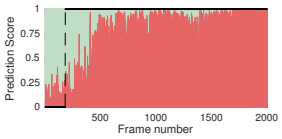


(c) Nonradiometric: Tent burning (static)

(d) Inception: Tent burning (static)



(e) Nonradiometric: Straw burning (drone)



(f) Inception: Straw burning (drone)

Figure 16: Comparison of the segmentation heuristic for thermal images (left) versus the inception classification for color images (right).

Overall, setting a threshold value of 0.5 (50%), the response of the Inception model is more robust and consistent in its predictions over time, whereas the color segmentation heuristic, since is dependent on the sensor response, would misclassify a greater percentage of the frames. To conclude this overview of the results it matters to quantify for both methodologies the time elapsed from the ignition to the first instance of fire detection. The results for each test are summarized in table 6.

Table 6: Time intervals for the first fire alarm for the experimental tests.

Data Type	Test 1	Test 2	Test 3
Thermal Nonradiometric	28 s	22.6 s	2.8 s
Color	3.4 s	6m48s	5 s

Despite the strengths and limitations of each approach, from the point of ignition to the first fire alarm, the detection is achieved in short timeframes for both approaches. Therefore, a framework for early fire detection would benefit from integrating these two types of sensory data into a single algorithm, in order to develop a more general solution.

7. Conclusions

The present work was developed under the scope of project Firecamp 2, which centers on studying fire risk in areas dedicated to camping and cara-

vanning activities. The investigation scope of this project is not limited to camping parks, and is extensible to fire risk situations in the wildland-urban interface. In this context, this work centered on analyzing different imaging sensory data and exploring computer vision techniques for detection of fire incidents in an initial stage.

To address a problem of this complexity, for which the available data were very limited, the first step required the development of experimental tests for image acquisition purposes. In order to perform realistic tests, a data acquisition system was designed to be integrated on a mobile aerial platform e.g. a drone, a balloon. While there was the intention of performing tests in an outdoor setting, due to the early arrival of fire season in 2017, for safety measures, these tests to be conducted in June, were only carried out in indoor laboratory conditions, at the Laboratory of Forestry Fire Studies in Lousã. The designed system included two different cameras to capture aerial images in the thermal and visible range. Since the image acquisition was performed synchronously, the tests resulted in a new dataset for fire detection composed by thermal nonradiometric and color images.

Having now different types of data to explore, various avenues of research were taken to evaluate its merits and limitations. Since the images from the thermal cameras are encoded in pseudocolor, a color segmentation heuristic was devised to partition the color palette into three different classes: gray, green, and red. The analysis of the variation of these classes provided insight into the response of the sensor in terms of RGB encoding. While this study enriched the understanding of the inner-workings of the thermal cameras, the analysis of the data from various experimental tests showcased the limitations of this type of approach. Due to excessive variability of the color levels, establishing an algorithm based on these three variables would have strong limitations in generalization for new scenarios. However, if a solution is required for fire detection for very specific contexts, a color segmentation heuristic could be fine-tuned to target a limited array of cases, by establishing a rule-based algorithm.

The second approach to address the fire detection problem explored the color imaging data, without prior feature extraction. This was achieved by resorting to a state-of-the-art deep neural network, i.e. the Inception V3 model, and applying a transfer learning scheme. Since this type of approach requires large amounts of data, the previous color image dataset was augmented with RGB images from additional tests. Although only offline classification was performed, the preprocessing procedure to extract frames from the video sequences was established having an online classification pipeline in mind. In what concerns the retraining of Inception V3, only the classifier part of the model was retrained using TensorFlow. Regarding the classification results, this approach demonstrated high accuracy for the cases where the fire is directly visible. From the results of this approach the test performed with images acquired from a drone, has special importance.

Although the two approaches proposed to this problem are very distinct in nature, the results provided a clear understanding of the limitations of each and the benefits using both can bring. Despite the fact that the tests presented in this work contemplate a limited set of situations, they demonstrated the value of designing a system that can capture different types of data. While the thermal imaging approach extends the sensing ability enabling faster detection, the deep learning approach with color images has the potential to provide more general and robust algorithms. Overall, the comparative analysis between the two approaches taken demonstrated that for a fire detection application the development of algorithms should encompass thermal and color image data.

References

[1] A. E. Çetin, K. Dimitropoulos, B. Gouverneur, N. Grammalidis, O. Günay, Y. H. Habiboğlu, B. U. Töreyin, and S. Verstockt. Video fire detection – review. *Digital Signal Processing*, 23(6):1827–1843, Dec. 2013.

[2] J. Keeley, W. Bond, R. Bradstock, J. G. Pausas, and P. Rundel. *Fire in Mediterranean Ecosystems. Ecology, evolution and management.* Cambridge University Press, 2012.

[3] J. San-Miguel-Ayanz, T. Durrant, R. Boca, G. Libertà, F. Boccacci, M. Di Leo, J. López Pérez, and E. Schulte. Forest Fires in Europe, Middle East and North Africa 2015. Technical report, Joint Research Centre - European Commission, 2016.

[4] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1):1929–1958, 2014.

[5] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the Inception Architecture for Computer Vision. *CoRR*, abs/1512.00567, 2015.

[6] TensorFlow Development Team. Inception in TensorFlow, Github Repository. Accessed: 2017-09-25.