A Multi-Layer Fusion Approach For Real-Time Fire Severity Assessment Based on Multimedia Incidents

Gerasimos Antzoulatos*
Information Technologies Institute (ITI) - Centre for Research and Technology Hellas (CERTH)
gantzoulatos@iti.gr

Panagiotis Giannakeris
Information Technologies Institute (ITI) - Centre for Research and Technology Hellas (CERTH)
giannakeris@iti.gr

Ilias Koulalis
Information Technologies Institute (ITI) - Centre for Research and Technology Hellas (CERTH)
iliask@iti.gr

Anastasios Karakostas
Information Technologies Institute (ITI) - Centre for Research and Technology Hellas (CERTH)
akarakos@iti.gr

Stefanos Vrochidis
Information Technologies Institute (ITI) - Centre for Research and Technology Hellas (CERTH)
stefanos@iti.gr

Ioannis Kompatsiaris
Information Technologies Institute (ITI) - Centre for Research and Technology Hellas (CERTH)
ikom@iti.gr

ABSTRACT

Shock forest fires have short and long-terms devastating impact on the sustainable management and viability of natural, cultural and residential environments, the local and regional economies and societies. Thus, the utilisation of risk-based decision support systems which encapsulate the technological achievements in Geographical Information Systems (GIS) and fire growth simulation models have rapidly increased in the last decades. On the other hand, the rise of image and video capturing technology, the usage mobile and wearable devices, and the availability of large amounts of multimedia in social media or other online repositories has increased the interest in the image understanding domain. Recent computer vision techniques endeavour to solve several societal problems with security and safety domains to be one of the most serious amongst others. Out of the millions of images that exist online in social media or news articles a great deal of them might include the existence of a crisis or emergency event. In this work, we propose a Multi-Layer Fusion framework, for Real-Time Fire Severity Assessment, based on knowledge extracted from the analysis of Fire Multimedia Incidents. Our approach consists of two levels: (a) an Early Fusion level, in which state-of-the-art image understanding techniques are deployed so as to discover fire incidents and objects from images, and (b) the Decision Fusion level which combines multiple fire incident reports aiming to assess the severity of the ongoing fire event. We evaluate our image understanding techniques in a collection of public fire image databases, and generate simulated incidents and feed them to our Decision Fusion level so as to showcase our method's applicability.

Keywords

Crisis Management, Real-Time Fire Severity Assessment, Image Recognition, Object Detection, Semantic Segmentation.

*corresponding author

INTRODUCTION

According to the Natural Disasters report (Ritchie and Roser 2020), although the number of deaths from natural disasters represents 0.1% of global deaths, over the past decade approximately 60,000 people died from natural disasters each year, however, the devastating impacts of shock and intensive events have been increasing worldwide overcoming the 0.4% of deaths these years. Forest fires are considered one of the most dangerous natural disasters due to the uncertain and highly unpredictable factors that determine their behaviour and spread (Hamadeh et al. 2017, Pacheco et al. 2015). Difficulties in dealing with such phenomena have motivated the scientific community into the investigation of new approaches that will be enabling fire crisis managers to tackle the new challenges.

Over the last two decades, technological advances and progress in multidisciplinary fields lead to the development of more intelligent and accurate fire risk tools and risk-based decision support systems (DSS). The coupling of wildfire growth simulation models with DSS based on economic models and integrated them with GIS tools, entails the enhancement of both the early warning systems and response systems (Miller and A. A. Ager 2012, Pacheco et al. 2015, Sakellariou et al. 2017). Existing fire risk tools are constantly updated with new features, extended or hybridised with others and applied to new fire management problems. As research evolves and available computing power grows, it becomes more possible to combine factors and data, in order to address the increasingly complex challenges posed by destructive fires that threaten people and properties. These tools tend to focus on fire geospatial risk assessment however, the reference data that link to what is happening on the ground are often lacking. Traditional approaches generally lack the ability of combining data and evidences from various sources.

Aiming at filling this gap, we propose a multi-layer fusion approach that can process and combine, dynamically, heterogeneous data from various sources, such as weather data, social media, drone imagery and sensory data, but also real-time information obtained from citizens who become active agents in the process by sending reports that may include visual, audio or text data related to fire events. In this work, we begin by applying cutting-edge computer vision techniques to form an image understanding module which will act as the first fusion level and will be responsible to discover *fire incidents* and metadata from images. Next, we feed the generated *incident reports* to the next fusion level, which, as the main focus of our work, is responsible to assess the severity of the ongoing fire event.

The paper is organized as follows: The next section presents a literature review of previous works on (a) wildfire crisis management systems, and (b) image understanding techniques that are relevant to the ones deployed in this paper. The overall methodology of our multi-layer fusion approach is presented in the following section. Next, the details of our evaluation strategy are given and experimental results are reported. Finally, conclusions and future enhancements are drawn in the last section.

RELATED WORK

Wildfire Crisis Management Systems

In the framework of fire management, noteworthy efforts have been made towards the creation of comprehensive wildfire risk assessment and crisis management tools so as to provide high quality services to fire managers in the decision-making process at various levels and scales (Miller and A. A. Ager 2012, Pacheco et al. 2015, Sakellariou et al. 2017, Naderpour et al. 2019). In this direction, numerous decision support GIS platforms have been developed which support wildfire prevention as well as monitoring capabilities related to fire detection, weather analysis, fire risk analysis and fire behaviour modelling. These systems enable fire protection agencies to spatially forecast and identify areas with high fire risk potential, enhancing to plan the necessary preventive and control actions as well as the preparedness decisions (Lee et al. 2002, Taylor and Alexander 2006, Markatos et al. 2007, Fiorucci et al. 2008, Davies et al. 2009, Barber et al. 2010, Chuvieco et al. 2010, Calkin et al. 2011).

The first attempt to develop a fully integrated and distributed system for Greece was Virtual Fire (Kalabokidis, Athanasis, et al. 2013), a web-based GIS platform designed to assist on early fire warning, fire control and coordination of firefighting forces. AEGIS (Kalabokidis, A. Ager, et al. 2016) is a wildfire prevention and management information system that leverages recent technical advances in HPC and Cloud Computing to speed-up the spatial fire hazard calculations and fire behavior modeling. The efforts to develop universal fire management decision support tools resulted in the creation of the European Forest Fire Information System (EFFIS). It supports forest protection against wildfires for over 30 countries in the European and Mediterranean regions and aims to cover the full cycle of forest fire management, providing up-to-date, reliable, and harmonized information on forest fires in prevention and preparedness phase as well as in post-fire damage analysis phases San-Miguel-Ayanz et al. 2012. The Wildland Fire Decision Support System (WFDSS) developed to support fire managers to make comprehensive, risk-informed, strategic and tactical decisions for fire incidents in the US (Noonan-Wright et al. 2011).

The essential characteristic of these DSS for fire risk management is the utilisation of fire weather indicators which rely on mathematical models for estimation of fire danger level (Hamadeh et al. 2017). These forecasting models usually analyse meteorological data and are more precise when they are based on weather forecast of the previous evening or previous day (Hamadeh et al. 2017), such as the simple fire index (Sharples et al. 2009), or they may combine weather data with fuel moisture codes that follow daily changes, such as the Canadian Forest Fire Weather Index (FWI) (Wagner 1987). These methods numerically estimate the fire danger indicator which can be then translated as a level of alarm that is activating when the probability of fire occurrence conditions is high.

Image Understanding

Fire **image classification** has been studied before in the works of Muhammad et al. 2017 and Sharma et al. 2017. Both works follow state-of-the-art Deep Convolutional Neural Network (CNN) approaches (Krizhevsky et al. 2012). Deep CNNs automatically craft features that encode the semantic content of the image in the first convolutional layers, while the final fully connected layers are responsible to classify the extracted features to the appropriate class. After the investigation on deeper models (Simonyan and Zisserman 2014), where smaller filters were examined for the feature extraction layers, the introduction of Inception (Szegedy et al. 2015) and ResNet (He et al. 2016) architectures proposed novel network wiring methods that led to significant improvements. The most recent methods explored the tasks of learning the model architecture automatically (Zoph et al. 2018, Xie et al. 2019), as well as recognition from limited data learning (Qiao, C. Liu, et al. 2018) and semi-supersived learning (Qiao, Shen, et al. 2018).

Semantic image segmentation state-of-the-art techniques follow the CNNs trend as well in the works of Long et al. 2015 and Ronneberger et al. 2015. These methods tackle the task of semantic segmentation by altering the objective of the classifier using pixel level labels in the image, leading to a segmentation mask instead of a recognition class for the whole image. More recent techniques like Deeplab's architecture in L.-C. Chen et al. 2018, exploit atrous separable convolutions coupled with a decoder module to refine the segmentation results around object boundaries. Techniques for fire segmentation are scarcely found, as there are not many public datasets with groundtruth fire masks. A worth-to-note technique which performs fire detection in social images with the use of color and texture attributes was presented in Chino et al. 2015.

Regarding **Object detection**, early CNN based works (Girshick 2015) proposed to deploy a multi-scale bounding box proposal generation technique first as a feeding mechanism of candidate boxes to deep classifiers. Later, single shot object detectors appeared with an end-to-end deep architecture like the ones proposed in the works of Ren et al. 2015, Redmon et al. 2016 and W. Liu et al. 2016. Those models achieved a better trade-off between accuracy and speed. In more recent studies, Neural Architecture Search has been proposed so as to discover a better feature pyramid network architecture (Ghiasi et al. 2019), and the problem of scale variation in objects has been explored (Li et al. 2019).

METHODOLOGICAL APPROACH

The overall objective of our framework is to assess the severity of an extreme fire event. We propose a Multi-Layer Fusion approach that can process data from heterogeneous sources. An overview of the proposed framework is illustrated in Figure 1.

Specifically, in the first layer, named the *Early Fusion level*, the acquired data are analysed so as to discover fire incidents and extract key concepts and metadata that are informative towards the severity of the situation. The innovative aspect of this approach lies in the fact that data can be obtained from various sources. Moreover, several analysis modules may exist in this level designed specifically to process each type of data, independently. Possible data sources include social media, real-time Wireless Sensor Networks (WSNs) and IoT systems, unmanned aerial vehicles (UAVs), or citizens and first responders directly, using a dedicated mobile application so as to upload to our system multimedia captured by their cell phones (images, videos, audio messages or text reports). Depending on the type of data, certain concepts can be extracted that can contribute to the assessment of the severity. For example, besides fire and smoke detection in images, pedestrian and vehicle detection may provide valuable information about the impact of fire to human beings. Moreover, important metadata can be extracted, such as the GPS coordinates of the captured files, the geolocation of tweets, or the altitude, angle and speed of flight in the case of data captured from UAVs. As soon as a piece of data is processed (e.g. an image or a social media post), the corresponding module generates an *incident* in the form of a report that contains all the extracted concepts in a structured format and forwards it to the next layer.

In the second layer, named the *Decision* or *Late Fusion level*, we formulate a rule-based approach that is responsible to gather fire incident reports from the previous layer and fuse them. The *Real-Time Monitoring and Severity*

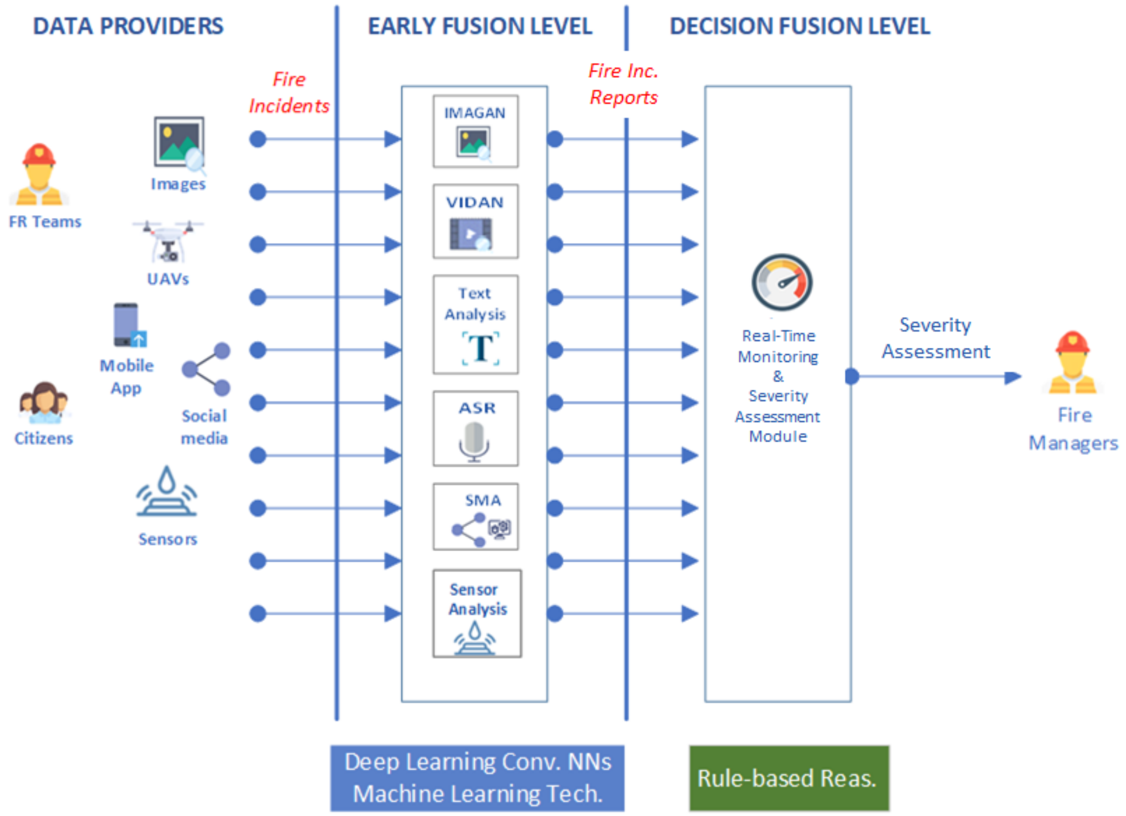


Figure 1. The proposed Multi-Layer fusion approach for multimedia analysis and severity assessment

Assessment module in this level encapsulates functionalities that are responsible for: (a) clustering the reports, (b) assessing the severity of each cluster and (c) assessing the severity of the whole region. To this aim, an online spatial clustering algorithm has been developed, to cluster fire incidents in a region of interest. Then, a rule-based reasoning approach estimates the severity level of each cluster based on concepts detected in each incident and their status. Finally, the fire severity level of the whole region of interest is estimated based on the severity levels of the set of formed clusters.

Early Fusion level

As explained previously, we perceive the Early Fusion level as a collection of autonomous modules whose purpose is to discover fire events as they happen, through the analysis of raw data. In this work, we focus on analysing images that may come from either social media or directly from citizens and firts responders using a mobile application. Therefore, an Image Analysis module constitutes our Early Fusion level. The input at this level besides the images is geolocation in the case of social media and GPS coordinates from the mobile application. The output, as discussed, is an incident report per image. In this section we delve deeper into the details regarding techniques and methods deployed in the Image Analysis module.

The Image Analysis module has two distinct objectives: (a) to detect images where fire or smoke is present and (b) to detect living targets, such as people or animals that may be in danger. The first objective is carried out by an image classification component that can separate fire and smoke images from other non-related scenery, in addition to a semantic segmentation component that can extract fire textures. The second objective is carried out by an object detection component that can locate the 2D pixel coordinates of objects that exist in the images and then recognize their semantic content.

Fire and Smoke Recognition and Segmentation

In order to recognize whether fire or smoke appears in an image the Emergency Classification (**EmC**) component has been deployed. It is based on a state-of-the-art image recognition technique that utilizes Deep Convolutional Neural Networks (CNNs), inspired by the recent success that deep learning has shown in image understanding (Simonyan and Zisserman 2014) and scene recognition (Zhou et al. 2014). The pre-trained weights of the VGG-16 architecture on the Places365 dataset was the starting point of the model's final training. The scene recognition oriented task that the pre-trained model was optimized to solve, is expected to be suitable prior knowledge in the form of distinctions between various visual clues that relate to generic scenery images. Starting our training from there we chose to fine-tune the model in order to recognize one of the following classes: "smoke", "fire" or "other". The first two classes correspond to smoke or fire images and the "other" class corresponds to images that neither smoke or fire appear. To do so, the final Fully Connected (FC) layer was removed and replaced with a new FC layer of width 3, while freezing the weights up to the previous layer. Finally, a softmax layer was deployed on top so as to enable multi-class recognition.

We have selected and compiled several datasets that contain fire and smoke images in order to fine-tune our model. More specifically, the BowFire dataset (Chino et al. 2015), the Corsican Fires database[1], the Fire-Detection-Image-Dataset[2], the Fire-Flame-Dataset[3], the Forest Fire dataset[4] and part of the FiSmo dataset Cazzolato et al. 2017. These datasets contain images that are taken mainly from fire forests, but also fire outbreaks in urban environments. Although our use cases are oriented towards forest fires, we include images that show fires in another context for two reasons:

- To train our model to learn visual clues of the flame texture irrespective of the environment that it may exist.
- To provide a sufficient number of instances for training. The public datasets that specialize on forest fires are not many.

Note that the instances that contain smoke and flames are annotated with the "fire" label and only the instances that solely contain smoke without any flames are annotated with the "smoke" label. We distinguish in this way between images that the fire is visible and the ones that only smoke is visible because they might have been taken from a location more distant to the fire center.

Simultaneously with EmC, an Emergency Localization (EmL) component is deployed, which is responsible to semantically segment the regions where fire pixels exist in case EmC's result indicates an emergency situation. The DeepLab architecture of "atrous convolution" was adopted (L. Chen et al. 2016), which uses convolution with up-sampled filters, so as to solve the fire localization problem in images. The EmL component is trained using the BowFire fire segmentation dataset (Chino et al. 2015), which provides masks of fire regions. We are interested in extracting the fire regions in order to correlate them later with the object bounding boxes and find the overlap that they may have in the 2D image plane. We aim to discover burning objects in this manner, notably vehicles that may contain humans. Note that we are not interested in extracting smoke regions in this work.

Object Detection Since we are especially interested to infer the level of impact that a fire event may have upon human targets and animals, we aim to extract pedestrians, animals, as well as vehicles that may contain humans from the images. The Object Detector (ObD) component is responsible to provide a set of bounding boxes that correspond to the image plane coordinates of the persons, animals and vehicles that exist in the images. Animal classes contain instances of cats and dogs and vehicle classes contain the following categories: cars, trucks, buses, bicycles and motorcycles. The architecture of our object detection component adopts the Faster R-CNN Ren et al. 2015, trained on OIDv4 dataset Kuznetsova et al. 2018. During inference we only keep the relevant classes and discard all other bounding boxes. More specifically, based on Huang et al. 2017, the ResNet101 feature extractor has been deployed for the extraction of deep features and then a Region of Interest (RoI) pooling scheme is used to classify candidate boxes.

Early Risk Assessment and Incident Report Generation Each image is fed to each component independently and the processing is done simultaneously to save time. Then, all the targets (i.e. pedestrians, animals and vehicles) that have been extracted from a single image are fed to an early Risk Assessment component. This component is responsible to assign an early risk estimation upon each detected target based on the existence of fire or smoke in the image, as well as the potential overlap its bounding box has with a fire region. The early risk indicator is calculated per target as a percentage value ranging from 0 to 1 as follows:

- If the image does not contain fire or smoke the risk indicator is set to 0%.
- If the image has fire or smoke the risk indicator is set to 50%.

[1] `http://cfdb.univ-corse.fr/index.php?menu=1`
[2] `https://github.com/cair/Fire-Detection-Image-Dataset`
[3] `https://github.com/DeepQuestAI/Fire-Smoke-Dataset`
[4] `https://dataturks.com/projects/qdyzl1013/Forest%20Fire`

- If the target's bounding box and an extracted fire region overlap, the percentage of fire pixels inside the bounding box is calculated. The overlap percentage contributes to the other 50%, raising the risk indicator up to 100% in the extreme case of full overlap.

In other words, a risk value of 0.5 indicates that a target has been found inside an image that shows fire or smoke. Consequently, a value higher than 0.5 indicates that a target is possibly close to the fire or burning. In the case of targets appearing in images with no fire or smoke, although the early risk indicator is set to 0, their contribution to the overall severity assessment may still be significant after clustering. At the final step, the extracted information, i.e. the EmC result, the extracted targets and their estimated risk indicators, along with the image GPS location, are gathered together to form an incident report per image which is immediately forwarded to the Decision Fusion level.

Decision Fusion level

During the emergency phase, the *Real-Time Monitoring and Severity Assessment* module enables fire/crisis managers to monitor the ongoing situation and assess its severity dynamically. Hereafter, we mention, and use interchangeably, as *Severity* of an incident (cluster) the magnitude of the incoming incident (cluster) and as *Severity Level* the corresponding mapping in five distinct categories, namely "Low", "Moderate", "High", "Very High" and "Extreme". This approach is similar with the one that used by European Forest Fire Information System (EFFIS) (San-Miguel-Ayanz et al. 2012). Generally, the severity levels imply the significance of an incident towards the overall severity assessment of the ongoing fire crisis event.

The proposed fire severity assessment algorithm relies on the results of the Early Fusion level (i.e. here, the Image Analysis module). Every time where a new incident report from the previous level appears, the *Real-Time Monitoring and Severity Assessment* module is activated and a sequence of steps are applied including the clustering of the incident and the estimation the severity levels: a) of the incident; b) of its cluster; and c) of the region that the cluster belongs to, following a bottom-up approach. Specifically, the steps of the fire severity assessment algorithm are the following:

Step 1. Data Acquisition: includes the processes to retrieve the outcomes of the Image Analysis. For each one of the incidents (i.e. images), the module obtains, as input data from the previous level, the incident's location, its type (i.e. fire, smoke or other), the class of each one of the targets that it contains, as well as the risk indicator and detection confidence for each one of those targets.

Step 2. On-line Clustering: a naive distance-based clustering algorithm is applied in order to cluster the imported incident in an on-line manner: each incident comes in, is processed and clustered into an existing cluster, or a new one is created depending on a distance criterion satisfaction. The clustering approach is described thoroughly in the following subsection *On-line Clustering algorithm*.

Step 3. Incident Severity level assessment: a rule-based reasoning based on a linear weighted fusion approach is performed, so as to estimate the severity of the incident and its corresponding severity level.

Step 4. Cluster Severity level assessment: a linear weighted fusion approach is performed in order to assess the severity of a new or updated cluster, based on the estimated severity of its members (incidents) and their corresponding severity levels.

Step 5. Overall Severity level assessment: a majority voting fusion algorithm is utilised so as to estimate the severity level of the whole region of interest, based on the severity levels of its clusters.

On-line Clustering algorithm

Clustering is commonly used as a method for data exploration, description, and summarization. Clustering the capturing incidents could help to reveal and understand the trends and mechanisms of natural devastating events. Furthermore, the clustering process can enhance the crisis managers' capacities to tackle a fire crisis event as they can estimate immediately the number of people and other assets in danger as well as the severity level of an ongoing fire event. Thus, they can manage and lead their available resources to respond appropriately and de-escalate the crisis situation. Towards this direction, a naive on-line clustering paradigm for spatial increasing volumes of data, which have low dimensionality, is applied in this work.

Initially, the geographical location of the first inserted incident is considered as the centre of the first cluster. As new incidents arrive, their distance from the existing centres are estimated. Each incident is assigned to the closest cluster according to its distance from the centre of clusters. We assume that incidents that have captured the same fire crisis event are located reasonably close to the crisis event. When a new incident is imported to a cluster, then the centre of the cluster is updated in order to include the new member. Contrariwise, an incident that is located far from the existing centres, probably captures another fire event. Thus, a new cluster should be created, with the incident's location as the centre. In this manner, the number of clusters incrementally increases as new incidents from various locations arrive.

It should be mentioned, that the distance threshold is a tunable parameter, to be determined by the users (i.e. fire managers, first responders and responsible authorities) so as to better reflect their intuition, knowledge and experience in dealing with such crisis events. Determining the distance threshold depends heavily on the operational capabilities of the crisis response forces to cover the fire incidents in a timely and efficient manner, as well as the geomorphological characteristics of the region where the fire incident occurs.

In this work, for the calculation of distance between two locations, the Haversine formula is utilised Sinnott 1984. It employs the law of haversine, which assumes that the Earth is spherical with radius R ($\approx 6371km$). The distance between two locations A, in spherical coordinates (longitude and latitude) lon_A and lat_A, and B in spherical coordinates lon_B and lat_B is calculated as follows:

$$dlon = lon_B - lon_A \tag{1}$$

$$dlat = lat_B - lat_A \tag{2}$$

$$a = sin^2(\frac{dlat}{2}) + cos(lat_A) * cos(lat_B) * sin^2(\frac{dlon}{2}) \tag{3}$$

$$c = 2 * arcsin(sqrt(a)) \tag{4}$$

$$Dist_{Haversine}(A, B) = R * c \tag{5}$$

The intermediate result c is the great circle distance in radians. The $Dist_{Haversine}(A, B)$ will be in the same units as R. The centre of the k-th cluster is updated by the following formula:

$$C_k = \left(\frac{1}{n} \sum_{i=1}^{n} x_{i,lon}, \frac{1}{n} \sum_{i=1}^{n} x_{i,lat} \right) \tag{6}$$

where, n is the cardinality of the k-th cluster and as $x_{i,lon}, x_{i,lat}$ denote the longitude and latitude of the i-th member of the particular cluster.

Incident Severity assessment

In Step 3, the severity and the severity level of each incident is estimated. The decision fusion algorithm classifies the targets in images into the following classes depending on their expected capacity to human lives.

- $class_1$ = ["human", "cat", "dog", "bicycle", "motorcycle"]
- $class_2$ = ["bus", "car", "truck"]

We assume that the members of the 2nd class have the capacity to carry on more than one human being. Thus, in case a target of this class, e.g. a "bus", is detected, then, the algorithm presumes that this has the potential of being equivalent to the detection of multiple human targets, and should hold higher weight when estimating the incident severity compared to members of the 1st class. It is worth mentioning here, that this classification is indicative and not compulsory to the execution of the whole process. The end-users could alter the members of classes, even increase the number of classes in order to include various kinds of targets.

In order to calculate an incident's severity based on the detected targets' properties, pre-defined fixed weights have been assigned to all possible cases. The magnitude of the weight depends on the type of incident (i.e. "fire", "smoke" or "other"), and the risk indicator and detection confidence given for each target in the incident report. The latter are quantified into 3 classes ("low", "medium" and "high"). The weight assigning process effectively encodes into the model to the reliability of the target detection and the targets' overlap or co-existence with fire or smoke instances in the image. The weights are given for a fire incident in Table 1a, a smoke incident in Table 1b and other incidents in Table 1c. It should be mentioned here, that in the case where no targets are detected into an incident, then for the "fire" type incident, the algorithm assigns a Severity value equal to 0.75, while for the "smoke" one, it assigns the value 0.65.

Henceforth, we are going to use the following notations:

Table 1. Weights assigned to targets for every incident type.

(a) Table of weights for a fire incident

Risk		Confidence		
		Low	Medium	High
Low	$class_1$	0.3	0.55	0.65
	$class_2$	0.4	0.65	0.75
Medium	$class_1$	0.55	0.65	0.75
	$class_2$	0.65	0.75	0.85
High	$class_1$	0.65	0.75	0.9
	$class_2$	0.75	0.85	1.0

(b) Table of weights for a smoke detected incident

Risk		Confidence		
		Low	Medium	High
Low	$class_1$	0.25	0.5	0.6
	$class_2$	0.35	0.6	0.7
Medium	$class_1$	0.5	0.6	0.7
	$class_2$	0.6	0.7	0.8
High	$class_1$	0.6	0.7	0.85
	$class_2$	0.7	0.8	0.95

(c) Table of weights for the incident type "other"

Risk		Confidence		
		Low	Medium	High
Low	$class_1$	0.1	0.2	0.3
	$class_2$	0.25	0.35	0.45

- I: a list of current incidents that they have already arrived and analysed by the modules
- $|I|$: the number of current incidents in the list I
- i: the i-th individual of the current analysed incidents, where $i = 1, \ldots, |I|$
- T_i: a list of detected targets in the i-th incident
- $|T_i|$: the number of detected targets in the i-th incident
- t_i: the $t-th$ individual of the targets in the i-th incident, where $t_i = 0, \ldots, |T_i|$. $|T_i| = 0$ implies the lack of detected targets in the i-th incident (image)

Let assume that the Image Analysis module has analysed the i-th incident and identified its type of *fire* or *smoke* or *other*. Moreover, a list of targets ($|T_i| > 0$) has been detected in it. Then, the following steps are performed:

(i) For each one of the t_i-th target in the list of targets T_i:

$\alpha)$ Determine the class, $class_c$, $c = 1, 2$, that the target belongs to.

$\beta)$ Determine the weight, w_{class_c, t_i} of the specific target in terms of the risk indicator and detection confidence in the incident. The weight is defined by employing the above tables, Table 1a if the EmC module has detected fire in the incident, Table 1b if the EmC module has detected smoke in the incident, or Table 1c if the EmC module neither detects fire nor smoke.

(ii) Estimate the severity of the i-th incident, $Severity_i$, by employing the following linear weighted fusion approach and aggregating over the weights of the targets in the particular incident:

$$Severity_i = \frac{1}{|T_i|} * \sum_{i=1}^{|T_i|} w_{class_c, t_i} \tag{7}$$

(iii) The severity level of an incident is estimated following the rule below:

if $Severity_i \leq 0.25$ **then**
$SeverityLevel_i$ = **Low**
else if $0.25 < Severity_i \leq 0.5$ **then**
$SeverityLevel_i$ = **Moderate**
else if $0.5 < Severity_i \leq 0.7$ **then**
$SeverityLevel_i$ = **High**
else if $0.7 < Severity_i \leq 0.8$ **then**
$SeverityLevel_i$ = **Very High**
else if $0.8 < Severity_i \leq 1.0$ **then**
$SeverityLevel_i$ = **Extreme**
end if

Cluster Severity assessment

In Step 4, a linear weighted strategy is applied for the cluster's severity assessment. Each time that a new incident inserted in an existing cluster, or a new cluster is created, this process is triggered in order to estimate the severity level of the cluster. Hence, the Severity of the k-th cluster is estimated by the following formula:

$$Severity_k = \frac{1}{N_{inc}} \sum_{i=1}^{N_{inc}} w_i * Severity_{k,i} \tag{8}$$

where N_{inc} is the number of incidents in the k-th cluster and $Severity_{k,i}$ denotes the severity of the i-th incident in the particular cluster, based on the equation (Eq. 7). Finally, w_i denotes the weight of the i-th incident. Here, we can consider that each one of the incidents has equal contribution to the estimation of the severity level of the cluster. Thus, the w_i is set to 1. The Severity Level of the cluster is estimated following the above rule, which is employed to map the incidents into 5 categories from "Low" to "Extreme".

Overall Severity level assessment

Finally, the Overall Severity level (Step 5) can be assessed by the utilisation of a majority voting fusion strategy. In the majority voting based fusion, the overall severity level of the region of interest is the one where the majority of the clusters reach a similar severity level.

EXPERIMENTAL RESULTS

The purpose of this section is dual. Firstly, we present the experiments that have been carried out so as to evaluate the reliability and performance of the Early Fusion Level (i.e. the Image Analysis module). Next, we exhibit results from a simulated test that we have conducted using the synthetic incidents generated by the Image Analysis module. The goal is to evaluate the performance of the Decision Fusion level methodology in terms of its ability to make appropriate assessment of the severity of an on-going fire crisis event.

Image Analysis Evaluation

The compiled dataset for the EmC training contains more than 11000 images. We made a random split, where 70% of the images were selected for training and 30% for testing. The mean accuracy of the classifier in the test data across all three classes is 0.98%. This is shown in Figure 2 where the normalized confusion matrix for the emergency classification task is presented. As we can see, there are very good accuracy rates in general. Most of the errors are false negative cases in the case of the 'smoke' class. The performance drop on smoke images may indicate that generally smoke detection is a harder task than fire detection. This holds true especially in cases where smoke is confused with clouds or fog. Interestingly, there is very little confusion between the two emergency classes: just 0.02% overall confusion between fire and smoke.

As far as concept detection, our object detection and semantic segmentation models have been adopted from the works of Huang et al. 2017, L. Chen et al. 2016 and Giannakeris et al. 2018, therefore detailed quantitative performance evaluation can be found there. However, we chose to showcase their applicability by including some qualitative examples in Fig. 3.

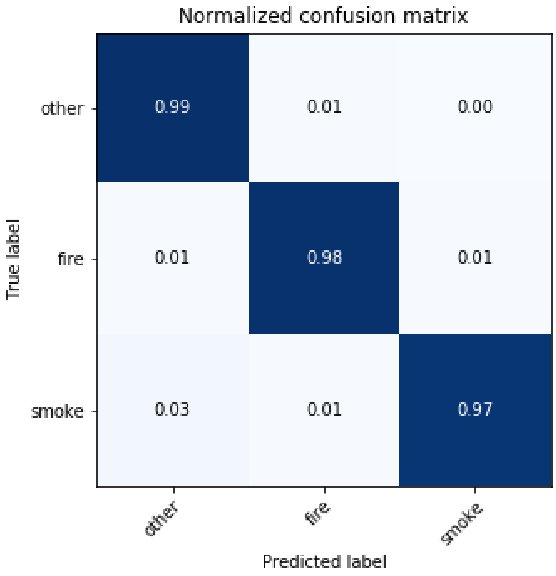
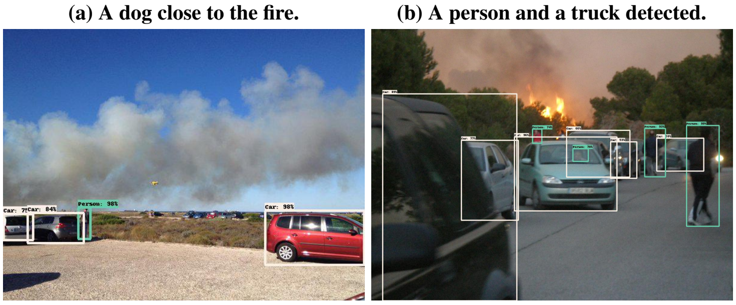


Figure 2. Confusion Matrix of EmC on test data

(a) A dog close to the fire. **(b) A person and a truck detected.**



(c) Vehicles detected in the presence of smoke. **(d) Vehicles detected in the presence of fire.**

Figure 3. Examples of detected concepts from the Image Analysis component.

Real-Time Monitoring and Severity Assessment Evaluation

The goal of this subsection is to evaluate the proposed algorithm for the severity assessment of the on-going fire, by exploiting the real-time information obtained from the analysis of images taken from the field. Hence, to achieve this goal, we assume that 60 incidents with multimedia content will be sent to the system for analysis.

The *Real-Time Monitoring and Severity Assessment* module obtains the incident reports with the analysis results from the previous layer, and fuses them with knowledge extracted from already existing incidents. In this experiment, the incidents generated by the image analysis module, are progressively assessed in terms of their severity and are clustered according to the methodology that governs this study. The geographical position of the images consumed by the Image Analysis module, follows the distribution pattern of incidents occurring on a close to real life scenario. We have adopted predefined thresholds such as radius, distance, and density of incidents, based on the knowledge gained by the fire use case testing in the beAWARE[5] project. In this evaluation experiment, initially, we assume

[5]https://beaware-project.eu/

Table 2. Experimental results of the Real-Time Monitoring and Severity Assessment module

Cluster	N_{inc}	Cluster Severity Level	Incident Severity Level				
			Low	Moderate	High	Very High	Extreme
1	16	High	0	1	12	3	0
2	20	Very High	0	0	1	19	0
3	17	Very High	0	0	1	16	0
4	3	Very High	0	0	1	1	1
5	2	Very High	0	0	0	2	0
6	2	High	0	0	2	0	0

three main fire sources to be assigned as the weighted centre of gravity for all the incidents. In total, 60 images were fed to the Early Fusion level one by one, resulting in 6 clusters of incidents. The majority of the incidents are clustered in 3 big groups, while other 3 smaller clusters were generated, which contained the isolated incidents. In Table 2, the results of the experiment are summarised.

The first cluster contains 16 incidents and "High" severity level was estimated. The cluster resulted holding 12 members classified as "High" severity, 3 as "Very High" and 1 as "Moderate". The outcome is rational, as the majority of the incidents were far from the location that fire had exploded and the majority of the captured images contains only smoke. Only 3 images out of 16 with fire were classified as "Very High", as shown in Figure 4a. The map shows blue points that represent each image (incident) sent to the Early Fusion level, and the numbers indicate the order of arrival. The second cluster contains 20 incidents, where, the majority of them were classified as "Very High" (19 out of 20) and one of them as "High" (Figure 4b). Image Analysis detected 3 cars and 1 human in the image shown at the bottom of the map, however the lack of fire degrades this incident's severity level to "High". Hence, the severity level of the whole cluster was assessed as "Very High", due to the frequent appearance of fire inside other images. The third cluster contains 16 incidents with "Very High" severity level and one incident which is classified as "High". Image Analysis manages to detect several instances of fire, as well as multiple targets. Thus, the overall cluster's severity was estimated as "Very High" (Figure 4c). The forth cluster is classified as "Very High", holding three members of "High", "Extreme" and "Very High" severity level (Figure 4d). The "Extreme" case, shows a car passing in a highway very close to a fire. The fifth and sixth clusters contain two members which were estimated as "Very High" and "High" severity incidents, accordingly (Figure 4e and Figure 4f). In the sixth cluster, the images contain only smoke and no targets are detected. Thus, the severity level of these incidents and of the whole cluster were estimated as "High". Finally, the algorithm assess as 'Very High' the overall severity of the Region of Interest, hence 4 out of 6 clusters classified as 'Very High' and the rest of them as 'High'.

(a) Cluster 1, estimated as "High" severity

(b) Cluster 2, estimated as "Very High" severity

(c) Cluster 3, estimated as "Very High" severity

(d) Cluster 4, estimated as "Very High" severity

(e) Cluster 5, estimated as "Very High" severity

(f) Cluster 6, estimated as "High" severity

Figure 4. Visualization of the created six clusters including an enumeration of the incoming fire incidents.

CONCLUSIONS

Overall, we have proposed a framework to tackle the current challenges in severity assessment, concerning the reveal of new or hidden knowledge from data that can be captured by multiple heterogeneous sources on the field. The simulated experiment that we conducted shows that the operation of the system was reasonable. The majority of the first incidents that arrived and were characterised as "High", created clusters which were also classified as of having "High" severity. As the severity of the incoming incidents increased to "Very High", the algorithm kept raising the severity levels of the created clusters, reflecting the new conditions and information obtained real-time from the field and providing a solid base environment for assisting the crisis manager in the decision-making process.

Despite these promising results, further work should be done to enhance the proposed approach in various directions. First of all, in the Decision Fusion layer the approach could be enriched taking under consideration information from heterogeneous sources. The analysis of textual and oral messages which are received from the citizens in the field through social media or mobile devices along with the video analysis from drones could be potential resources to extend our approach, by combining those types of information and empowering the dynamic severity assessment of the fire crisis. Moreover, the naive clustering algorithm which relies only on spatial criteria to group the incidents, could be further strengthen by employing more qualitative criteria to create the clusters. In this direction, state-of-the-art spatio-temporal clustering algorithms should be integrated in the proposed framework to improve the clustering process. In the above methodology, the employed weights for the estimation of the severity levels were preset and constant reflecting the end-users point of view and beliefs regarding the significance of the detected targets and the crisis events. However, this process should be changed by the utilisation more intelligent and dynamically adjustable approaches from the machine learning field. Finally, our framework can be customised to support multi-hazard scenaria or diverse type of data in the future.

ACKNOWLEDGMENTS

This work was supported by the EC-funded research and innovation programs H2020 beAWARE: "Enhancing decision support and management services in extreme weather climate events" under the grant agreement No.700475 and aqua3S: "Enhancing standardisation strategies to integrate innovative technologies for Safety and Security in existing water networks" under the grant agreement No.832876. The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

Barber, J., Bose, C., Bourlioux, A., Braun, W., Brunelle, E., Garcia, T., Hillen, T., and Ong, B. (Jan. 2010). "Burning issues with PROMETHEUS, the Canada's wildfire growth simulator". In: *Canadian Applied Math Quarterly* 16, pp. 337–378.

Calkin, D. E., Thompson, M. P., Finney, M. A., and Hyde, K. D. (2011). "A real-time risk assessment tool supporting wildland fire decisionmaking". In: *Journal of Forestry* 109 (5), pp. 274–280.

Cazzolato, M. T., Avalhais, L. P., Chino, D. Y., Ramos, J. S., Souza, J. A. d., Rodrigues Junior, J. F., and Traina, A. J. M. (2017). "FiSmo: A Compilation of Datasets from Emergency Situations for Fire and Smoke Analysis". In: *Proceedings of the satellite events*.

Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. (2016). "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs". In: *CoRR* abs/1606.00915. arXiv: 1606.00915.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. (2018). "Encoder-decoder with atrous separable convolution for semantic image segmentation". In: *Proceedings of the European conference on computer vision (ECCV)*, pp. 801–818.

Chino, D. Y., Avalhais, L. P., Rodrigues, J. F., and Traina, A. J. (2015). "Bowfire: detection of fire in still images by integrating pixel color and texture analysis". In: *Graphics, Patterns and Images (SIBGRAPI), 2015 28th SIBGRAPI Conference on*. IEEE, pp. 95–102.

Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M. P., Vilar, L., Martínez, J., Martín, S., Ibarra, P., et al. (2010). "Development of a framework for fire risk assessment using remote sensing and geographic information system technologies". In: *Ecological Modelling* 221.1, pp. 46–58.

Davies, D., Ilavajhala, S., Wong, M., and Justice, C. (2009). "Fire Information for Resource Management System: Archiving and Distributing MODIS Active Fire Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 47.1, pp. 72–79.

Fiorucci, P., Gaetani, F., and Minciardi, R. (2008). "Development and application of a system for dynamic wildfire risk assessment in Italy". In: *Environmental Modelling & Software* 23.6, pp. 690–702.

Ghiasi, G., Lin, T.-Y., and Le, Q. V. (June 2019). "NAS-FPN: Learning Scalable Feature Pyramid Architecture for Object Detection". In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Giannakeris, P., Avgerinakis, K., Karakostas, A., Vrochidis, S., and Kompatsiaris, I. (2018). "People and vehicles in danger-A fire and flood detection system in social media". In: *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*. IEEE, pp. 1–5.

Girshick, R. (2015). "Fast r-cnn". In: *arXiv preprint arXiv:1504.08083*.

Hamadeh, N., Karouni, A., Daya, B., and Chauvet, P. (2017). "Using correlative data analysis to develop weather index that estimates the risk of forest fires in Lebanon & Mediterranean: Assessment versus prevalent meteorological indices". In: *Case Studies in Fire Safety* 7, pp. 8–22.

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

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S., et al. (2017). "Speed/accuracy trade-offs for modern convolutional object detectors". In: *IEEE CVPR*.

Kalabokidis, K., Ager, A., Finney, M., Athanasis, N., Palaiologou, P., and Vasilakos, C. (2016). "AEGIS: a wildfire prevention and management information system". In: *Natural Hazards and Earth System Sciences* 16, pp. 643–661.

Kalabokidis, K., Athanasis, N., Gagliardi, F., Karayiannis, F., Palaiologou, P., Savas, P., and Christos, V. (2013). "Virtual Fire: A web-based GIS platform for forest fire control". In: *Ecological Informatics* 16, pp. 62–69.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems*, pp. 1097–1105.

Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J., Kamali, S., Popov, S., Malloci, M., Duerig, T., et al. (2018). "The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale". In: *arXiv preprint arXiv:1811.00982*.

Lee, B., Alexander, M., Hawkes, B., Lynham, T., Stocks, B., and Englefield, P. (2002). "Information systems in support of wildland fire management decision making in Canada". In: *Computers and Electronics in Agriculture* 37.1, pp. 185–198.

Li, Y., Chen, Y., Wang, N., and Zhang, Z. (Oct. 2019). "Scale-Aware Trident Networks for Object Detection". In: *The IEEE International Conference on Computer Vision (ICCV)*.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., and Berg, A. C. (2016). "Ssd: Single shot multibox detector". In: *European conference on computer vision*. Springer, pp. 21–37.

Long, J., Shelhamer, E., and Darrell, T. (2015). "Fully convolutional networks for semantic segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431–3440.

Markatos, N., Kiranoudis, C., Vescoukis, V., and Balatsos, P. (2007). "Towards an integrated system for planning and decision support in forest fire management". In: *Proceedings of the 4th International Wildland Fire Conference*. Seville, Spain.

Miller, C. and Ager, A. A. (2012). "A review of recent advances in risk analysis for wildfire management". In: *International Journal of Wildland Fire* 22, pp. 1–14.

Muhammad, K., Ahmad, J., and Baik, S. W. (2017). "Early Fire Detection using Convolutional Neural Networks during Surveillance for Effective Disaster Management". In: *Neurocomputing*.

Naderpour, M., Rizeei, H. M., Khakzad, N., and Pradhan, B. (2019). "Forest fire induced Natech risk assessment: A survey of geospatial technologies". In: *Reliability Engineering & System Safety* 191.

Noonan-Wright, E. K., Opperman, T. S., Finney, M. A., Zimmerman, G. T., Seli, R. C., Elenz, L. M., Calkin, D. E., and Fiedler, J. R. (2011). "Developing the US Wildland Fire Decision Support System". In: *Journal of Combustion* 2011.

Pacheco, A. P., Claro, J., Fernandes, P. M., Neufville, R. de, Oliveira, T. M., Borges, J. G., and Rodrigues, J. C. (2015). "Cohesive fire management within an uncertain environment: A review of risk handling and decision support systems". In: *Forest Ecology and Management* 347, pp. 1–17.

Qiao, S., Liu, C., Shen, W., and Yuille, A. L. (June 2018). "Few-Shot Image Recognition by Predicting Parameters From Activations". In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Qiao, S., Shen, W., Zhang, Z., Wang, B., and Yuille, A. (Sept. 2018). "Deep Co-Training for Semi-Supervised Image Recognition". In: *The European Conference on Computer Vision (ECCV)*.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: Unified, real-time object detection". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788.

Ren, S., He, K., Girshick, R., and Sun, J. (2015). "Faster r-cnn: Towards real-time object detection with region proposal networks". In: *Advances in neural information processing systems*, pp. 91–99.

Ritchie, H. and Roser, M. (2020). "Natural Disasters". In: *Our World in Data*.

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical image computing and computer-assisted intervention*. Springer, pp. 234–241.

Sakellariou, S., Tampekis, S., Samara, F., Sfougaris, A., and Christopoulou, O. (Nov. 2017). "Review of state-of-the-art decision support systems (DSSs) for prevention and suppression of forest fires". In: *Journal of Forestry Research* 28.6, pp. 1107–1117.

San-Miguel-Ayanz, J., Schulte, E., Schmuck, G., Camia, A., Strobl, P., Liberta, G., Giovando, C., Boca, R., Sedano, F., Kempeneers, P., et al. (2012). "Comprehensive Monitoring of Wildfires in Europe: The European Forest Fire Information System (EFFIS)". In: *Approaches to Managing Disaster*. Ed. by J. Tiefenbacher. Rijeka: IntechOpen. Chap. 5.

Sharma, J., Granmo, O.-C., Goodwin, M., and Fidje, J. T. (2017). "Deep Convolutional Neural Networks for Fire Detection in Images". In: *International Conference on Engineering Applications of Neural Networks*. Springer, pp. 183–193.

Sharples, J., McRae, R., Weber, R., and Gill, A. (2009). “A simple index for assessing fire danger rating”. In: *Environmental Modelling & Software* 24.6, pp. 764–774.

Simonyan, K. and Zisserman, A. (2014). “Very Deep Convolutional Networks for Large-Scale Image Recognition”. In: *CoRR* abs/1409.1556. arXiv: 1409.1556.

Sinnott, R. W. (1984). “Virtues of the Haversine”. In: *Sky and Telescope* 68.2, p. 159.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., et al. (2015). “Going deeper with convolutions”. In: Cvpr.

Taylor, S. and Alexander, M. (Mar. 2006). “Science, technology, and human factors in fire danger rating: the Canadian experience”. In: *International Journal of Wildland Fire* 15, pp. 121–135.

Wagner, V. (1987). *Development and structure of the Canadian Forest Fire Weather Index System*. Tech. rep. Forestry Technical Report 35. Headquarters Ottawa: Canadian Forest Service, Northern Forestry Centre.

Xie, S., Kirillov, A., Girshick, R., and He, K. (2019). “Exploring randomly wired neural networks for image recognition”. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1284–1293.

Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., and Oliva, A. (2014). “Learning deep features for scene recognition using places database”. In: *Advances in neural information processing systems*, pp. 487–495.

Zoph, B., Vasudevan, V., Shlens, J., and Le, Q. V. (2018). “Learning transferable architectures for scalable image recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8697–8710.