USING VIRTUAL SCENARIOS TO PRODUCE MACHINE LEARNABLE ENVIRONMENTS FOR WILDFIRE DETECTION AND SEGMENTATION

Telmo Adão [1,2,*], Tatiana M. Pinho [1], Luís Pádua [1,2], Nathalie Santos [2], António Sousa [1,2], Joaquim João Sousa [1,2], Emanuel Peres [1,2]

[1] INESC TEC Technology and Science, 4200-465 Porto, Portugal - (telmoadao,tatianap,luispadua,arms,jjsousa,eperes)@utad.pt
[2] University of Trás-os-Montes e Alto Douro, 5000-801 Vila Real, Portugal – nsguimaraes@utad.pt

Commission VI, WG VI/4

KEY WORDS: Fire Detection, Fire Monitoring, Deep Learning (DL), Feature Pyramid Network (FPN), Fire Virtual Simulation, Corsican Fire Dataset, Virtual-to-Real Learning Transference.

ABSTRACT:

Today's climatic proneness to extreme conditions together with human activity have been triggering a series of wildfire-related events that put at risk ecosystems, as well as animal and vegetal patrimony, while threatening dwellers nearby rural or urban areas. When intervention teams - firefighters, civil protection, police - acknowledge these events, usually they have already escalated to proportions hardly controllable mainly due wind gusts, fuel-like solo conditions, among other conditions that propitiate fire spreading.
Currently, there is a wide range of camera-capable sensing systems that can be complemented with useful location data - for example, unmanned aerial systems (UAS) integrated cameras and IMU/GPS sensors, stationary surveillance systems - and processing components capable of fostering wildfire events detection and monitoring, thus providing accurate and faithful data for decision support. Precisely in what concerns to detection and monitoring, Deep Learning (DL) has been successfully applied to perform tasks involving classification and/or segmentation of objects of interest in several fields, such as Agriculture, Forestry and other similar areas. Usually, for an effective DL application, more specifically, based on imagery, datasets must rely on heavy and burdensome logistics to gather a representative problem formulation. What if putting together a dataset could be supported in customizable virtual environments, representing faithful situations to train machines, as it already occurs for human training in what regards some particular tasks (rescue operations, surgeries, industry assembling, etc.)?
This work intends to propose not only a system to produce faithful virtual environments to complement and/or even supplant the need for dataset gathering logistics while eventually dealing with hypothetical proposals considering climate change events, but also to create tools for synthesizing wildfire environments for DL application. It will therefore enable to extend existing fire datasets with new data generated by human interaction and supervision, viable for training a computational entity. To that end, a study is presented to assess at which extent data virtually generated data can contribute to an effective DL system aiming to identify and segment fire, bearing in mind future developments of active monitoring systems to timely detect fire events and hopefully provide decision support systems to operational teams.

1. INTRODUCTION

Fire ignitions due to human activity and climatic changes are becoming increasingly challenging to manage and control. According to official sources from United States ("U.S. Wildfires," 2019) and Europe (Jesús San-Miguel-Ayanz et al., 2018), these events still have unpredictable impacts on several countries in what regards to both ecological and civil society perspectives.
The need for fire monitoring has been widely addressed among several works that encompass satellite-based approaches (Morisette et al., 2005), sensor networks (Bhattacharjee et al., 2012), mobile biological sensors (Sahin, 2007), unmanned aerial vehicles (UAV) and remote sensing (Yuan et al., 2015) and whose computation is handled by methods that might range from typical digital image processing (Celik et al., 2007) to neural networks (Muhammad et al., 2018).
Typical approaches for processing technological-based fire monitoring, namely the ones relying on deep learning (DL) require the use of existing datasets for models' training and validation. However, factors such as climate changes have the potential of bringing new disaster scenarios, desirable of being characterized or prevented even before they occur. Fire specialists/scientists might be valuable to point out directions for fire ignition and spreading in unseen scenarios and conditions, wherein virtual environments play a crucial role. More specifically, synthesizable hypothesis for fire occurrence prediction have the potential to constitute a relevant source for improving fire detection and preventing higher damages in effective threatening situations triggered by fire. Following this line, this paper argues that, like humans, machines are capable of learning out of virtual environments within a context that promotes a progressive objects' classification and segmentation accuracy enhancement, while simultaneously reducing the need for burdensome logistics to gather datasets worthy to be considered representative of a given problem. The motivation for such study lies on the latest graphic cards technology that has been allowing to reach computer-based synthetized environments of outstanding realism and convincingness, as well as on neural networks advances, namely regarding to backpropagation and, lately, region proposal capabilities.
To investigate machines' virtual-to-real transfer learning assumption, two main tasks were carried out: (a) a fire ignition and spreading simulator tool was developed to allow users to manipulate particles systems impersonating smoke and fire within a completely 3D virtual environment endowed with

[*] Corresponding author

https://doi.org/10.5194/isprs-archives-XLII-3-W8-9-2019

interactivity; (2) a Feature Pyramid Network (FPN) was applied to handle the computer-based learning of virtual images with fire collected from synthetized environments - manipulated with the previous mentioned simulator tool - and to make predictions on real images part of the Corsican fire dataset ("Corsican Fire Database," 2017; Toulouse et al., 2017). The little amount of virtual images homogeneously taken from the very same virtual environment - built with a low-detail photogrammetric model set up out of flight made over a chestnut area - used to train DL models allowed to conclude about the potential that machines have in earning knowledge from real-world by learning from synthetized environments.

Regarding organization, the remainder of this paper is divided into five other sections. A brief background addressing automatic fire detection is presented in the second section. Afterwards, an overview of the proposed virtual-to-real transfer learning approach is proposed in the third section, followed by Section 4 that encompasses implementation details regarding the virtual fire simulator as well as the FPN used as deep learning-based segmentation architecture to assess machine's virtual-to-real learnability. The last two sections (5 and 6) are reserved for results, discussion, conclusions and future work.

2. BACKGROUND IN AUTOMATIC FIRE MONITORING/DETECTION

Fire ignitions in general and wildfire in particular have been greatly addressed by the community of scientists/practitioners interested in studying this type of hazards/disasters and reducing/mitigating their destructive effects.

Whilst international efforts to sense and provide responses to fire events have been witnessed for a while Olenick and Carpenter, 2003 and Alkhatib (2014) suggest that the main systems for the purpose of detecting fire (in this case, within forest context) rely on satellites, optical sensors, digital cameras, and wireless sensor networks.

Satellites have the capability of remotely monitoring fires at large distances, in a macro-analytical way. For example, Morisette et al. (2005) resorts to data produced by *Terra*'s Moderate Resolution Imagine Spectroradiometer (MODIS) along with Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) to study and assess fire pixels analysis algorithms. Weaver et al. (2004) studied wildfire detection and short-range forecast in geostationary satellite imagery. Studies within the context of fire and satellites usage are still prevailing (Johnston et al., 2018).

Within optical sensors/digital cameras, infrared (IR) sensitive devices are suitable to detect heat regions as pointed out by Töreyin et al. (2007), who also characterized flame flickering through hidden Markov models. Least mean square errors active learning was a process proposed by Günay et al. (2010), who characterized fire by colour, motion, flicker handling and contour analysis. Also highlighting colour and motion, Zhang et al. (2014) suggested an improved probabilistic approach for fire detection in videos, based on candidate fire regions generation and fire area change. A block-based approach over low quality cameras' imagery using discrete cosine transforms (DCT) and wavelets, and in combination with k-Nearest neighbor (k-NN) and support vector machines (SVM) demonstrated usefulness in detecting smoke, which is used as an early indicator for fire hazard (Gubbi et al., 2009). DCT was also one of the key processes to early detect fire through internet protocol (IP) cameras imagery in a later work, more specifically by tracking smoke (Millan-Garcia et al., 2012). In the work of Muhammad et al. (2018), convolutional neural networks (CNN) fine-tuned for closed-circuit television cameras were applied towards fire detection. An approach to boost smoke detection on wildland forests through Region-CNN (R-CNN) and synthesized smoke images was recently proposed (Zhang et al., 2018).

Regarding wireless sensor networks, Bhattacharjee et al. (2012) reported a life-saving system to early detect fire in coal mines, resorting to temperature and gas sensors. The idea of using animals as mobile biological sensors was proposed by Sahin (2007), wherein some species were equipped with thermal and radiation sensors with Global Positioning System (GPS) features towards early detection forest fires. A similar concern with the forest protection was encompassed by Fernández-Berni et al., 2012, but with a vision-enabled wireless sensor network supported by a robust algorithm for smoke detection.

Others addressed fire detection with a radio-acoustic sound system and thermal maps (Sahin and Ince, 2009). Within soft computing techniques for fire detection, Mahdipour and Dadkhah (2014) proposed an extensive review that considers the categorization made by Alkhatib (2014) and intelligent techniques (artificial neural networks, cellular automata, etc.), as well as false alarms reduction. A relatively recent survey made by Yuan et al. (2015) highlights the benefits and opportunities of using UAV platforms for monitoring, detecting and fighting fire in forests, as well as key technologies for the automatic performance of such tasks. Also focusing UAVs, Cruz et al. (2016) proposed a high precision fire index with low computational penalty.

Regarding CNN architectures, a considerable amount of them have been proposed: VGG - named after Visual Geometry Group labortatory - (Simonyan and Zisserman, 2014), ResNet (He et al., 2015), Inception family (Szegedy et al., 2014) (Szegedy et al., 2015) (Szegedy et al., 2016), Google's MobileNet (Howard et al., 2017), Xception (Chollet, 2016) and DenseNet (Huang et al., 2016). Segmentation networks oriented to object detection include – but are not confined to - Fully Convolutional Network (FCN) (Long et al., 2014), ParseNet (Liu et al., 2015), U-Net (Ronneberger et al., 2015) and Feature Pyramid Network (FPN) (Zhao et al., 2016).

The following section presents a virtual-to-real transfer learning approach that aims to endow a machine with the capabilities to recognize fire in real imagery, based on previously learned virtual environments, specifically set up with particle systems to simulate smoke and flames.

3. VIRTUAL-TO-REAL TRANSFER LEARNING APPROACH

Inspired by the way that virtual reality systems have been training humans for the most varied tasks (Cha et al., 2012), an approach to transfer learning from 3D synthetized environments to train machines in identifying fire events in real scenarios is proposed. It is composed of 3 main functional blocks: a basic virtual fire simulation tool, a processing component working with a CNN-based FPN, and a sensing layer.

The virtual fire simulation tool allows to build a faithful virtual environment that can integrate convincing manually produced or photogrammetrically computed landscape models, upon which particle systems simulating smoke and dynamic flames can seed. The functionalities planned for this basic version include:

- Loading of virtual landscapes;
- Placement and manipulation of particle systems to produce fire-like effects;
- Basic virtual lightning parameterization;
- Virtual photographic registration (print-screens of regions of interest within the synthetized environment).

FPNs became a popular convolutional-based image segmentation approach, developed with the participation of Facebook Artificial Intelligence Research (FAIR).It is known for properly dealing with different scales, due to it's pyramidal bottom-up and top-

https://doi.org/10.5194/isprs-archives-XLII-3-W8-9-2019 | © Authors 2019. CC BY 4.0 License.

down strategies for determining feature maps. As such, a decision was taken to integrate FPNs in the processing component of this proposal, with applicability in two main steps: 1) training stage, to create the convolutional model out of virtual environments simulating burning fire situations with particle systems; and 2) prediction stage, to estimate fire masks upon imagery depicting real-world scenarios.

Collecting *in situ* imagery for monitoring and detecting fire deflagrations requires the integration of a sensing layer, eventually composed of smartphone-based, fixed and/or UAV cameras for image acquisition and processing plus additional sensors/devices, such as positional systems and communication infrastructure.

Regarding the working pipeline, fire simulator constitutes a tool for setting up convincing fire deflagrations and extracting images and classification region masks out of it. These outcomes are then sent to a processing component that applies a FPN architecture to create an estimation model for classifying and segmenting regions of interest in incoming images, depicting real scenarios. Such images are provided by an abstract sensing layer mainly composed by camera-capable devices that collect digital snapshots from the real-world to be delivered to the processing component, for consistent prediction and monitoring. Ideally, this data flow ends up feeding a decision support system, benefiting several actors (citizens, forest ranges, farmers, security authorities, etc.) with context-aware information, according to each one's activities. Thereby, resulting intervention guidelines have the potential to advise citizens on how to avoid burning areas, to guide firefighters by providing fire striking strategies, highlight paths to safely help farmers to secure livestock, and so on. Conceptually, data flows among a communication infrastructure preferably configured by a wireless-based transmission/reception architecture, due to the obvious risk of damage for tangible hardware, such as cables. Figure 1 sums up the proposed virtual-to-real transfer learning approach in a general architecture scheme.

4. MATERIALS AND METHODS

The implemented tools, as well as the landscape models and FPN details will be provided in this section.

4.1 Fire Simulator

The tool developed to build faithful environments for fire simulations upon landscape virtual models (Figure 2) was developed in Unity 3D (Unity Technologies, California, United States).

The user interface includes a navigable 3D environment composed of a digital landscape, a top bar providing instructions and applicational proprieties' status. There is also a small previewer at the bottom-right corner of the applicational window that gives a glimpse of the content that is being captured by the virtual camera that collects imagery within the virtual environment, in two modes: (1) attached to the user as a first-person shooter (FPS) camera; and (2) attached to a virtual drone, simulating a birds eye perspective.

Regarding interaction and functionalities, a user can select the virtual landscape to load from the hard drive, set some areas on fire by clicking in the mouse scroller, enable or disable smoke particle system, by pressing "F" key, and make the acquisition virtual camera switch between FPS and birds eye perspective mode by selecting "C" key. Each time "P" key is pressed with the referred virtual camera pointed to a region of interest within the virtual environment, a couple images in portable network graphics (PNG) format are printed out to files and saved in the hard drive: one of them represents the region of interest as-is, while the other constitutes the ground-truth mask segmenting the fire flames.

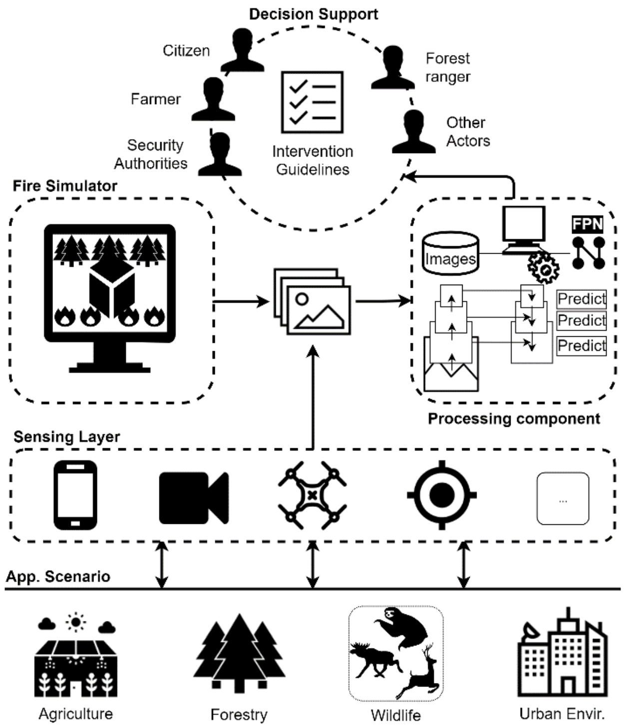

Figure 1. General architecture of the virtual-to-real transfer learning approach: a fire simulator enables to hypothesize virtual scenarios with particle systems spawining flames/smoke, while a FPN-based processing component learns to identify those particle-based elements; the imagery/data produced by a sensing layer can use that artifical knowledge to identify probable fire hazard situations in real-world, through the processing component that is ideally linked to a decision support system to provide intervention guidelines to different stakeholders.

4.2 Virtual landscape model integrated in the fire simulator

The fire simulation tool supports the integration of virtual landscapes selected from user's hard drive. In this study, a particular photogrammetric model, produced based on UAV images of a flight mission performed over an area located in north-eastern Portugal (41°22'43.8"N, 7°35'00.8"W) was used. The area is populated with clusters of pine (*Pinus pinaster*) and chestnut (*Castanea sativa Mill.*) trees, shrubland communities, some grassland and agricultural properties, and few man-made infrastructures. Imagery acquisition was done using a DJI Phantom 4 (DJI, Shenzhen, China), a cost-effective multi-rotor UAV equipped with a Global Navigation Satellite System (GNSS) receiver and a 12.4 MP RGB sensor mounted in a 3-axis electronic gimbal (Pádua et al., 2018). Pix4Dcapture (Pix4D SA, Lausanne, Switzerland) allowed to plan the flight that was carried out at 100 m height, covering 3.8 ha, with a frontal imagery overlap of 80% and a longitudinal overlap of 70%, resulting in a 4 cm ground sampling distance. Then, the high-resolution RGB imagery was submitted to a photogrammetric processing using Pix4Dmapper Pro (Pix4D SA, Lausanne, Switzerland), which resorts to Structure from Motion (SfM) algorithms to identify common (tie) points in the images, enabling the generation of three-dimensional dense point clouds, from which textured

https://doi.org/10.5194/isprs-archives-XLII-3-W8-9-2019

meshes are produced and made available to use as virtual models in external programs.

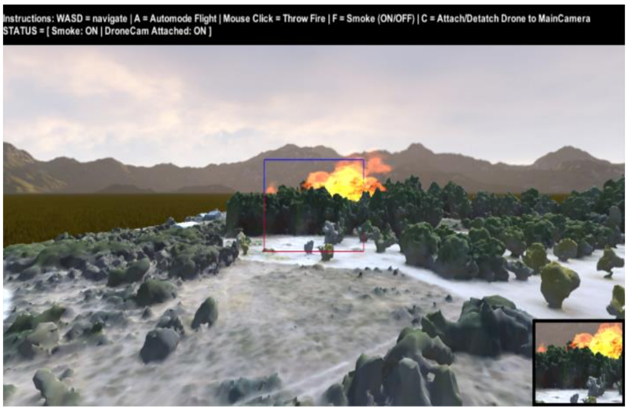

Figure 2. Fire simulator ambient displaying a virtual landscape and a sinthetized fire focus of small dimension.

4.3 Datasets

Two datasets were used in this study: one virtual for training purposes and another constituted by real-world occurrences for assessing the virtual-to-real transfer learning effectiveness.

4.3.1 Real-world dataset

The laboratory "Sciences Pour l'Environnement" UMR CNRS 6134 SPE of the University of Corsica is the main stage for the "Fire" project, which is dedicated to the modeling and experimentation of fires in natural environments populated with vegetation ("Corsican Fire Database," 2017; Toulouse et al., 2017).
Following the project goals, a wide database composed of fire imagery was established, containing thousands of wildfire pictures and image sequences acquired in the visible and near-infrared spectral range, considering various conditions of shooting, type of burning vegetation, climatic conditions, brightness and distance to fire. Within this dataset, two types of images can be found: a RGB picture and a corresponding binary mask annotating the active flame in the image (Figure 3).

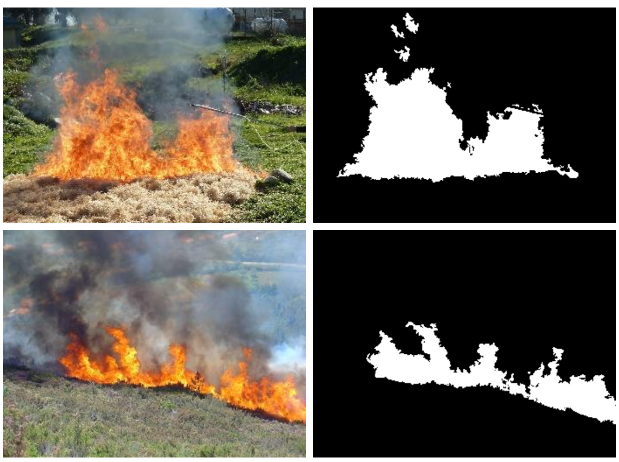

Figure 3. Examples of RGB and binary images retrieved from Corsican dataset ("Corsican Fire Database," 2017; Toulouse et al., 2017).

4.3.2 Synthesized dataset

Within the previously addressed fire simulator, imagery datasets can be produced by using an FPS camera, under user control. Each time a shoot is requested from that FPS camera, two PNG files are outputted regarding a virtual area of interest under focus: a clean print-screen of the scene and a mask annotating synthetized flames (Figure 4). The process for producing a ground-truth mask consists in coloring in black the materials of all objects other than particle systems-based fire, change their shaders to disable diffuse reflections triggered by ambient lightning, clear sky-boxes to black, hide synthetized smoke and then produce the image and binarize it according to a straightforward black/non-black pixels criterion. In the end, objects features are completely restored to the state they were before ground-truth mask production process, in an operation that is imperceptible to users' interaction.

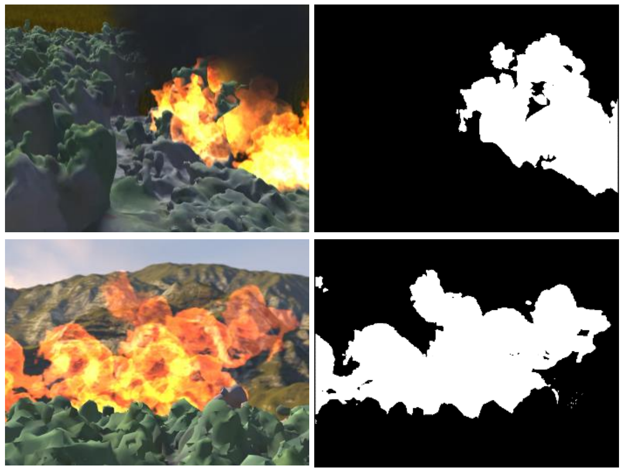

Figure 4. Examples of RGB and binary images retrieved from the fire simulator tool.

4.4 FPN approach

For a faster integration of FPN capabilities, an implementation made available by Matterport, Inc (Sunnyvale, California, USA) was used and properly adapted to the fire problem posed in this paper (*Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow*, 2019). This is an implementation of mask region-CNN (R-CNN) for Python 3, built upon Keras and TensorFlow. Models resulting from this approach generate bounding boxes and segmentation masks for each object instance in the images. A ResNet architecture provides support in the backbone. It allows to visualize every step related to the anchor boxes refinement towards final detection boxes, generates masks and promotes control over activation layers. TensorBoard plotting in one of the supported features, as well as the possibility of dealing with both Common Objects in Context (COCO) and customized datasets preparation.

The next section presents the results of the preliminary tests to assess the potential of FPN to infer real fires out of virtual knowledge.

5. PRELINARY RESULTS AND DISCUSSION

The FPN network was trained with the virtual fire images took from the virtual environment. Regarding configurations settings, here are some of the used parameters:

- Image dimension: 1024;
- Image minimum dimension: 800;
- Image resize mode: square;
- Learning momentum: 0.9;
- Learning rate: 0.001;
- Weight decay: 0.0001;
- Loss: bounding box/mask loss;

- Mask pool size: 14;
- Class number: 2 (Fire and background);
- RPN train anchors per image: 128;
- Steps per epoch: 10;
- Top-down pyramid size: 256.
- Classifier layer size: 1024

Moreover, among several region proposal network (RPN) settings specifying anchors' scales, ratios, strides, also a value for detection confidence can be defined, thresholding the objects that will be outputted by the FPN in use.
Learning stage presented a convergence tendency, visible in loss values for both training and validation, as evidenced by Table 1. Besides global loss, 5 other parameters were retrieved: bounding box, class, mask, RPN bounding box and RPN class loss values. As usual, training showed a faster convergence inasmuch images' features are "freely" acquired and mapped, according to the behavior imposed by ResNet architecture, in this case. On the other hand, validation stage had an expected lower convergence rate in time, due to the underlying matching process that tests learned features against imagery reserved for proofing/checking (validation set), which is, in turn, implicated in the calibration of the weights and biases of the neural network in use.

Parameter (Loss)	Training		Validation	
	Epoch 1	Epoch 50	Epoch 1	Epoch 50
Loss Value	3.4	0.2	2.2	1.9
Bounding Box	0.9	0.06	0.7	0.3
Class	0.3	0.02	0.08	0.06
Mask	1.3	0.1	0.6	0.1
RPN Bounding Box	0.6	0.01	0.7	1.3
RPN Class	0.08	0.01	0.3	0.2

Table 1. Results of the FPN training stage for both training and validation, considering loss metrics.

Evaluation metrics were based on known evaluation indicators, more specifically, true positive/negative (*TP*,*TN*) and false positive/negative (*FP*,*FN*) pixels counting, as well as Dice coefficient and Jaccard Index. The formulas for latter pair of metrics are presented in equations (1 and 2), under a set theory perspective, wherein *DC* and *JI* represent Dice Coefficient and Jaccard Index, both operating with two sets specified by *X* and *Y*.

$$DC = \frac{2\,|X \cap Y|}{|X| + |Y|} \qquad (1)$$

$$JI = \frac{|A \cap B|}{|A \cup B|} \qquad (2)$$

While Figure 5 provides a visual insight of the predictions made over real open-air fire images using knowledge acquired from a virtual dataset, Table 2 presents the quantitative analysis using the previously mentioned set of metrics. As it can be observed, the prediction of fire could be achieved successfully in all of the images submitted to FPN. Regarding similarity to ground-truth, one can easily conclude that accuracy is not high. The following aspects might be contributing for this outcome: only 16 images depicting virtual fire were used, 10 for training and 6 for validation; data augmentation was not considered in this stage; only a single virtual environment was used, with prejudice to the heterogeneity of examples required for an effective feature learning; photogrammetric processing was not set to output a top quality landscape virtual model, considering a balance between available computational resources and smoothness regarding simulator tool usage, at runtime; finally, some parts of the training/validation masks were annotating virtual smoke as fire - due to virtual simulator's occlusion aspects requiring revision -, with visible impact in FPN performance.

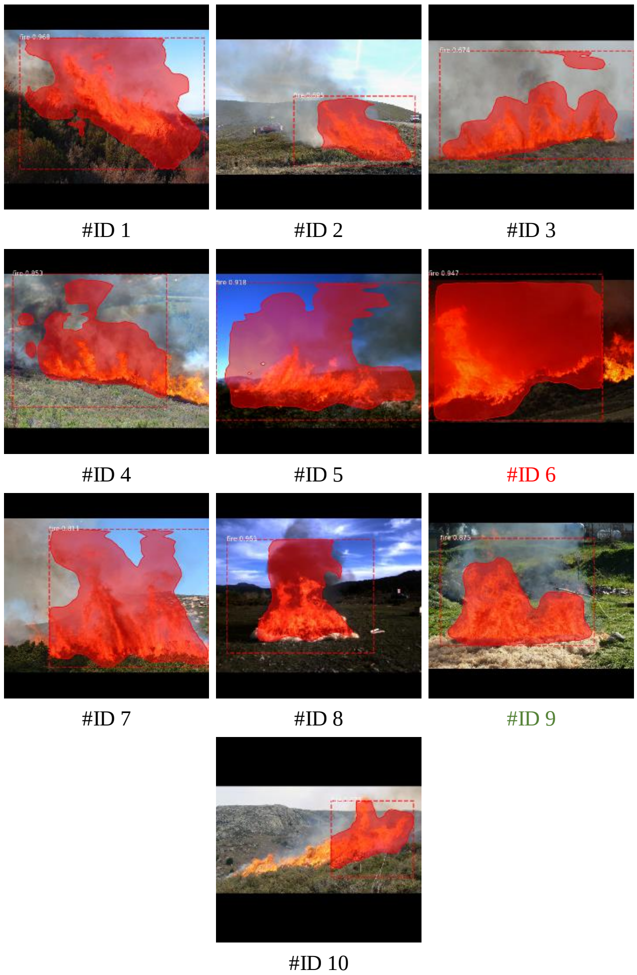

Figure 5. Prediction masks over the 10 images used to test the virtual-to-real transfer learning approach.

#ID	Estimation Data					
	TP(px)	**FP(px)**	**TN(px)**	**FN(px)**	**DC(%)**	**JI(%)**
1	104112	253039	691425	0	45%	29%
2	62117	32524	953935	1289	79%	66%
3	116828	91075	840673	2184	72%	56%
4	62769	176532	809275	26058	42%	26%
5	99495	300687	648394	384	40%	25%
6	100828	416767	530981	29388	32%	19%
7	227805	172423	648348	13716	72%	57%
8	85198	77522	885856	1656	69%	52%
9	153000	47467	848109	6942	87%	76%
10	57730	61510	929336	38668	65%	48%

Table 2. Results of the 10 images used to test the virtual-to-real transfer learning approach. *TP*, *FP*, *TN* and *FN* correspond to *True positive*, *False Postive*, *True Negative* and *False Negative*, respectivelly; all of them are measured in numbers of pixels. *DC* stands for *Dice Coefficient*, while *JI* represents *Jaccard Index*, both presented as percentual quantities.

In spite of the discussed issues, correlations rated over 70% can be found in 4 images (according to Dice Coefficient), pointing out the potential of the proposed virtual-to-real transfer leaning approach. Indeed, images identified with the IDs 5 and 6 had the worst correlations in this trial, probably, due to the conditions in

https://doi.org/10.5194/isprs-archives-XLII-3-W8-9-2019

which they were acquired. While the former can be characterized by the presence of a great extent of smoke causing entropy in FPN detectors - due to the aforementioned issue related to training with smoke annotated as fire -, the latter is influenced not only by the same aspects, but also by poor luminosity conditions and background occlusion with potential impact on global context perception. Thereby, a training process encompassing influence conditions likely to integrate a fire deflagration situation is of major importance for an effective detection of the element of interest. As in most of supervised deep learning problems, dataset is a key component for success.

6. CONCLUSIONS AND FUTURE WORK

In a rapid climate changing conjuncture, establishing tools for dealing with eventual catastrophe events becomes increasingly relevant. With particular focus in open-sky fire deflagrations in rural/forested area, this paper aims to propose a tool that aims to simulate fire events within virtual environments - endowed, for example, by photogrammetric models - and use the knowledge of hypothetical synthetized situations to endurance machine-based detection in real-world. An FPN deep learning approach was applied to learn the synthetized scenarios in a first instance, and then perform predictions in real images of burning fire.

More specifically, the study in this paper involved 16 images of a virtual scenario enriched with a north-eastern Portugal forested area photogrammetric model and fire simulated through particle systems for learning purposes. After 50 training epochs, the generated predictive model was used to make estimations upon real images of burning fire in rural/forested areas. In spite of the not so high precisions in segmenting fire, the element of interest was detected in the 10 tested images with satisfactory results, considering the small number of training images and the lack of heterogeneity covering different possibilities. In one image of the testing dataset, more complying with the conditions of the used virtual environment, a similarity rate of 87% (estimated segmentation vs. ground truth mask) was reached, somehow pointing out the potential of the virtual-to-real transfer learning approach and justifying the need for a deeper study.

Future work must encompass a wider study with a broader dataset, techniques for data augmentation and stratification, as well as more heterogeneous hypothetical conditions using, for example, different photogrammetric models regarding distinct contexts. Also, a revision to the simulator should be carried out to improve the fire masking mechanisms, avoiding the probable entropy that might be rising from smoke particles partially annotated as as fire.

ACKNOWLEDGEMENTS

The authors would like to acknowledge project “CHIC – Cooperative Holistic View on Internet and Content” (N° 24498), financed the European Regional Development Fund (ERDF) through COMPETE2020 - the Operational Programme for Competitiveness.

REFERENCES

Alkhatib, A.A.A., 2014. A Review on Forest Fire Detection Techniques. Int. J. Distrib. Sens. Netw. 10, 597368. https://doi.org/10.1155/2014/597368

Bhattacharjee, S., Roy, P., Ghosh, S., Misra, S., Obaidat, M.S., 2012. Wireless sensor network-based fire detection, alarming, monitoring and prevention system for Bord-and-Pillar coal mines. J. Syst. Softw., Novel approaches in the design and implementation of systems/software architecture 85, 571–581. https://doi.org/10.1016/j.jss.2011.09.015

Celik, T., Demirel, H., Ozkaramanli, H., Uyguroglu, M., 2007. Fire detection using statistical color model in video sequences. J. Vis. Commun. Image Represent. 18, 176–185. https://doi.org/10.1016/j.jvcir.2006.12.003

Cha, M., Han, S., Lee, J., Choi, B., 2012. A virtual reality based fire training simulator integrated with fire dynamics data. Fire Saf. J. 50, 12–24. https://doi.org/10.1016/j.firesaf.2012.01.004

Chollet, F., 2016. Xception: Deep Learning with Depthwise Separable Convolutions. ArXiv161002357 Cs.

Corsican Fire Database [WWW Document], 2017. URL http://cfdb.univ-corse.fr/index.php?menu=1 (accessed 7.14.19).

Cruz, H., Eckert, M., Meneses, J., Martínez, J.-F., 2016. Efficient Forest Fire Detection Index for Application in Unmanned Aerial Systems (UASs). Sensors 16, 893. https://doi.org/10.3390/s16060893

Fernández-Berni, J., Carmona-Galán, R., Martínez-Carmona, J.F., Rodríguez-Vázquez, Á., 2012. Early forest fire detection by vision-enabled wireless sensor networks. Int. J. Wildland Fire 21, 938–949. https://doi.org/10.1071/WF11168

Gubbi, J., Marusic, S., Palaniswami, M., 2009. Smoke detection in video using wavelets and support vector machines. Fire Saf. J. 44, 1110–1115. https://doi.org/10.1016/j.firesaf.2009.08.003

Günay, O., Taşdemir, K., Uğur Töreyin, B., Çetin, A.E., 2010. Fire Detection in Video Using LMS Based Active Learning. Fire Technol. 46, 551–577. https://doi.org/10.1007/s10694-009-0106-8

He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition. ArXiv151203385 Cs.

Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. ArXiv170404861 Cs.

Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q., 2016. Densely Connected Convolutional Networks.

Jesús San-Miguel-Ayanz, Tracy Durrant, Giorgio Libertà, Alfredo Branco, Daniele de Rigo, Davide Ferrari, Pieralberto Maianti, Tomàs Artés Vivancos, Hugo Costa, Fabio Lana, 2018. Forest Fires in Europe, Middle East and North Africa 2017 (No. EUR 29318 EN). European Union.

Johnston, J.M., Johnston, L.M., Wooster, M.J., Brookes, A., McFayden, C., Cantin, A.S., 2018. Satellite Detection Limitations of Sub-Canopy Smouldering Wildfires in the North American Boreal Forest. Fire 1, 28. https://doi.org/10.3390/fire1020028

Liu, W., Rabinovich, A., Berg, A.C., 2015. ParseNet: Looking Wider to See Better. ArXiv150604579 Cs.

Long, J., Shelhamer, E., Darrell, T., 2014. Fully Convolutional Networks for Semantic Segmentation. ArXiv14114038 Cs.

Mahdipour, E., Dadkhah, C., 2014. Automatic fire detection based on soft computing techniques: review from 2000 to 2010. Artif. Intell. Rev. 42, 895–934. https://doi.org/10.1007/s10462-012-9345-z

Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow: matterport/Mask_RCNN, 2019. . Matterport, Inc.

Millan-Garcia, L., Sanchez-Perez, G., Nakano, M., Toscano-Medina, K., Perez-Meana, H., Rojas-Cardenas, L., 2012. An Early Fire Detection Algorithm Using IP

Cameras. Sensors 12, 5670–5686. https://doi.org/10.3390/s120505670

Morisette, J.T., Giglio, L., Csiszar, I., Setzer, A., Schroeder, W., Morton, D., Justice, C.O., 2005. Validation of MODIS Active Fire Detection Products Derived from Two Algorithms. Earth Interact. 9, 1–25. https://doi.org/10.1175/EI141.1

Muhammad, K., Ahmad, J., Baik, S.W., 2018. Early fire detection using convolutional neural networks during surveillance for effective disaster management. Neurocomputing, Learning System in Real-time Machine Vision 288, 30–42. https://doi.org/10.1016/j.neucom.2017.04.083

Olenick, S.M., Carpenter, D.J., 2003. An Updated International Survey of Computer Models for Fire and Smoke. J. Fire Prot. Eng. 13, 87–110. https://doi.org/10.1177/1042391503013002001

Pádua, L., Marques, P., Adáo, T., Hruška, J., Peres, E., Morais, R., Sousa, A., Sousa, J.J., 2018. UAS-based Imagery and Photogrammetric Processing for Tree Height and Crown Diameter Extraction, in: Proceedings of the International Conference on Geoinformatics and Data Analysis, ICGDA '18. ACM, New York, NY, USA, pp. 87–91. https://doi.org/10.1145/3220228.3220241

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv150504597 Cs.

Sahin, Y.G., 2007. Animals as Mobile Biological Sensors for Forest Fire Detection. Sensors 7, 3084–3099. https://doi.org/10.3390/s7123084

Sahin, Y.G., Ince, T., 2009. Early Forest Fire Detection Using Radio-Acoustic Sounding System. Sensors 9, 1485–1498. https://doi.org/10.3390/s90301485

Simonyan, K., Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv14091556 Cs.

Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2016. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. ArXiv160207261 Cs.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2014. Going Deeper with Convolutions. ArXiv14094842 Cs.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2015. Rethinking the Inception Architecture for Computer Vision. ArXiv151200567 Cs.

Töreyin, B.U., Cinbis, R.G., Dedeoglu, Y., Cetin, A.E., 2007. Fire detection in infrared video using wavelet analysis. Opt. Eng. 46, 067204. https://doi.org/10.1117/1.2748752

Toulouse, T., Rossi, L., Campana, A., Celik, T., Akhloufi, M.A., 2017. Computer vision for wildfire research: An evolving image dataset for processing and analysis. Fire Saf. J. 92, 188–194. https://doi.org/10.1016/j.firesaf.2017.06.012

U.S. Wildfires [WWW Document], 2019. . US Wildfire Stat. Provid. Natl. Interag. Fire Cent. NIFC Anom. Relat. 2000–2019 Aver. URL https://www.ncdc.noaa.gov/societal-impacts/wildfires/ (accessed 7.14.19).

Weaver, J.F., Lindsey, D., Bikos, D., Schmidt, C.C., Prins, E., 2004. Fire Detection Using GOES Rapid Scan Imagery. Weather Forecast. 19, 496–510. https://doi.org/10.1175/1520-0434(2004)019<0496:FDUGRS>2.0.CO;2

Yuan, C., Zhang, Y., Liu, Z., 2015. A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. Can. J. For. Res. 45, 783–792. https://doi.org/10.1139/cjfr-2014-0347

Zhang, Q., Lin, G., Zhang, Y., Xu, G., Wang, J., 2018. Wildland Forest Fire Smoke Detection Based on Faster R-CNN using Synthetic Smoke Images. Procedia Eng., 2017 8th International Conference on Fire Science and Fire Protection Engineering (ICFSFPE 2017) 211, 441–446. https://doi.org/10.1016/j.proeng.2017.12.034

Zhang, Z., Shen, T., Zou, J., 2014. An Improved Probabilistic Approach for Fire Detection in Videos. Fire Technol. 50, 745–752. https://doi.org/10.1007/s10694-012-0253-1

Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2016. Pyramid Scene Parsing Network. ArXiv161201105 Cs.