Lightweight and efficient octave convolutional neural network for fire recognition

Angel Ayala [1], Estanislau Lima [1], Bruno Fernandes [1], Byron L. D. Bezerra [1], and Francisco Cruz [2,3]
[1]Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil
[2]School of Information Technology, Deakin University, Geelong, Australia.
[3]Escuela de Ingeniería, Universidad Central de Chile, Santiago, Chile.
Emails: {aaam, ebl2, bjtf, byronleite}@ecomp.poli.br, francisco.cruz@deakin.edu.au

***Abstract*—Fire recognition from visual scenes is a demanding task due to the high variance of color and texture. In recent years, several fire-recognition approaches based on deep learning methods have been proposed to overcome this problem. However, building deep convolutional neural networks usually involves hundreds of layers and thousands of channels, thus requiring excessive computational cost, and a considerable amount of data. Therefore, applying deep networks in real-world scenarios remains an open challenge, especially when using devices with limitations in hardware and computing power, e.g., robots or mobile devices. To address this challenge, in this paper, we propose a lightweight and efficient octave convolutional neural network for fire recognition in visual scenes. Extensive experiments are conducted on FireSense, CairFire, FireNet, and FiSmo datasets. In overall, our architecture comprises fewer layers and fewer parameters in comparison with previously proposed architectures. Experimental results show that our model achieves higher accuracy recognition, in comparison to state-of-the-art methods, for all tested datasets.**

***Index Terms*—fire recognition, lightweight model, octave convolution, ResNet, cross-dataset**

I. Introduction

The presence of fire in some environments is capable of causing massive losses, hence, the early recognition of this kind of accidents is primordial. Early recognition of fire can be translated in a quick response to manage the accident and therefore, high accuracy of fire recognition it is also essential. In this regard, a system capable of triggering an alarm with high accuracy, it is crucial for the response team in charge of monitoring this kind of accidents.

Fire accidents can be present in many environments, e.g., open-air, private, or community use spaces, among others. Fire can be originated because of human intervention, piece of machinery malfunction, unstable state of some structures, or in many other cases as a consequence of other natural disasters. Uncontrolled fire, or blaze, can affect in economic, social, and environmental way principally. This damage could be restored or not, in case it could be restored, considerable effort and consequently, resources are required. A common type of fire accident is the forest fire, which can cause significant damage to the environment [1] and can increase its severity if it spreads.

In Latin America the forest fires are present in the Amazonia [2] and Chile [3] mainly, which have economical and environmental consequences such as mentioned at [4]. Chile, just in 2014, had more than 8000 fires which affected 130000ha. After the forest fire, the soil remains damaged [5] and it is difficult for the vegetation to grow again. When this type of accident occurs in the environment, all plants and animal life disappear from the affected zone as a perturbation of the environment itself. The problem with the fires is that they are unpredictable, in the way of when or where they will occur, especially for forest fires. Hence, an early alert system would help to manage these accidents or natural disasters.

Actual fire detection systems have a slow response time in fire recognition, primarily because most of them are built on sensors like thermal, smoke, or flame detectors. These sensors detection systems are not sufficiently reliable because of a high failure rate when the alarm is triggered. This kind of sensor-based detector also needs time for the internal chemical reaction of their material in order to trigger the signal. Thus, where the detection has triggered an alarm, the probability of damages it is already high.

In the last years, multiple methods have been proposed to automate fire recognition, most of them for video surveillance systems. The most recent methods are based on deep learning (DL) [6] methods, especially convolutional neural networks (CNNs). Furthermore, these deep models, such as AlexNet [7], VGG16 [8], Inception [9], ResNet [10] have been adjusted successfully to many tasks. Applying very deep networks to many real-world applications, such as robots, cars, smartphones, and mobile devices that have hardware limitations, reminds a challenging task.

In this paper, we propose a lightweight and efficient octave convolutional neural network for fire recognition, which can automatically recognize fire. Our network model is based on ResNet architecture with a few numbers of layers and the octave convolution, which reduces memory inference. In this regards, it is intended to be implemented as a monitoring system on mobile vehicles for fire detection, such as the one presented by Madhevan et al. [11]. Experiments show that our network with fewer layer and fewer parameters produces promising results in lower computational time, compared with recent state-of-the-art methods on images recognition.

II. Related Work

To endow a machine with the ability to recognize the presence of fire, it is a highly demanding task, mainly due

978-1-7281-5666-8/19/$31.00 ©2019 IEEE

to the recognition of the texture, color, and the fire's phenome representation by itself [12]. Initially, the task of fire recognition in computer vision was addressed using techniques based on color space such as presented at [13] which presents a flame detection algorithm. This algorithm employs features as color probability with contour irregularity, among others. Another technique based in color is presented at [14] which was implemented in a CCTV system where the color detection method uses the information of the RGB space to detect the foreground at video sequences of fire-like objects. Additionally, another technique uses spectral color such as presented at [15] where it is proposed a multi-sensor surveillance system using optical and infrared cameras for remote monitoring. Also, a spatio-temporal technique is presented at [12] where frame by frame is analyzed for flickering and also, with color probability, for fire detection. Moreover, at [16] characteristics of textures technique are applied, where a dynamic analysis of the textures takes advantage of the knowledge of fire presence in prior frames using linear dynamical systems.

More recently methods are based in DL [6] neural networks. These methods are widely used to learn the features of images where, at first layers, the simplest features are recognized and, at top layers, the more complex ones. In [17] is proposed a model based on SqueezeNet. The authors present a custom layers block for signal processing, as well as model compressing to achieve a light-size model. The VGG16, ResNet, and Simple CNN models are tested in [18]. This work also presents some variation of the VGG16 and ResNet models for test purpose in the dataset. The Simple CNN was used to check the profundity level required to acquire features of fire images. In [19] is presented an algorithm capable to recognize types of flame, based on single-shot multibox detector where the task is separated in fire recognition and fire location. The work presented in [20] proposes a model based in VGG16 trained with a limited quantity of images using a generative adversarial network and data augmentation techniques for increasing the number of the images in the dataset. One of the most recent works presented in [21] proposes a lightweight neural network to be applied in the internet of things, ensuring a high frame rate for a fire recognition in limited hardware devices such as Raspberry Pi, as a fire and smoke detection system.

III. Proposed Approach

As mentioned in the previous section, deep learning methods for neural networks present many advantages in object recognition. The number of elements to be recognized or the complexity of the image, in which DL is used, defines the number of layers. During training, the feature complexity which is learned from the images varies at different layer levels of the network. As fire images present characteristics which are hard to be hand-craft extracted, in this work, a DL approach is used for this purpose.

The proposed model is based on the ResNet architecture, which is composed of levels of blocks with convolutional layers, followed by a batch normalization layer and an activation layer. Furthermore, each block connects its input independently with the output of the block, which is known as residual. Therefore, the next block obtains as input the processed signal by the current block as well as the raw signal. The first block of each level, unlike the consecutive blocks, processes this residue by convolution and batch normalization layers. To connect each block uses an activation layer that processes the previous block for the next one.

To obtain a lightweight model capable of reaching high performance on portable hardware, it needs to have as few parameters as possible. In order to reduce the parameters model number, a low deep level architecture is proposed in this work. This low deep level is suited for fire recognition as mentioned at [19], where, due to its characteristics, just a few layers are used. For the proposed model to work with as minimal hardware as possible, another aspect addressed it is the application of octave convolution [22], which modifies the usual way how convolutions are done. The octave convolution separates the input signal into two channels, one for high frequencies to acquire more detailed features, and another for low frequencies to more general features. This technique allows the model to work with less memory and a fewer number of FLOPs in comparison to a vanilla convolution layer.

Combining the ResNet-like architecture with octave convolutions leads to a lightweight model with a low computational cost. The input for the network is a 96x96 pixels image at RGB channel. This architecture is composed of 2 ResNet levels with 4 and 2 block, respectively. On top of the network, it has a global average pooling followed by a fully-connected layer with 2 units and softmax activation. The α value of the ratio for the octave convolution is 0.25 and with 64 initial filters. This configuration turns out a model with 956226 trainable parameters (size on disk ~12MB). A simplified version of the architecture can be seen in Fig. 1, additional details of the implementation can be seen on the project repository[1].

IV. Experimental Setup

Due to the complexity of the fire recognition task, mainly because some similarities with fire-like objects, a suitable dataset is required for training the proposed model. In this regard, a fair trade-off between images with fire presence and images without is needed to carry out a good training process. Moreover, it has to be considered images with objects with fire-like color labeled as no fire. From the literature review, four datasets have been selected to compare the performance of the proposed model. Three of the datasets have been used in previous works, and another one has been recently compiled. Custom names has been given to the three previously used datasets, FireSense[2] [16], CairFire[3] [18], and the most recently released FireNet[4] [21]. The fourth one, never used before,

[1] Github repository https://github.com/angel-ayala/fire_recognition

[2] Online available at https://zenodo.org/record/836749 [Accessed: July, 2019]

[3] Online available at https://github.com/cair/Fire-Detection-Image-Dataset [Accessed: July, 2019]

[4] Online available at https://github.com/arpit-jadon/FireNet-LightWeight-Network-for-Fire-Detection, [Accessed: July, 2019]

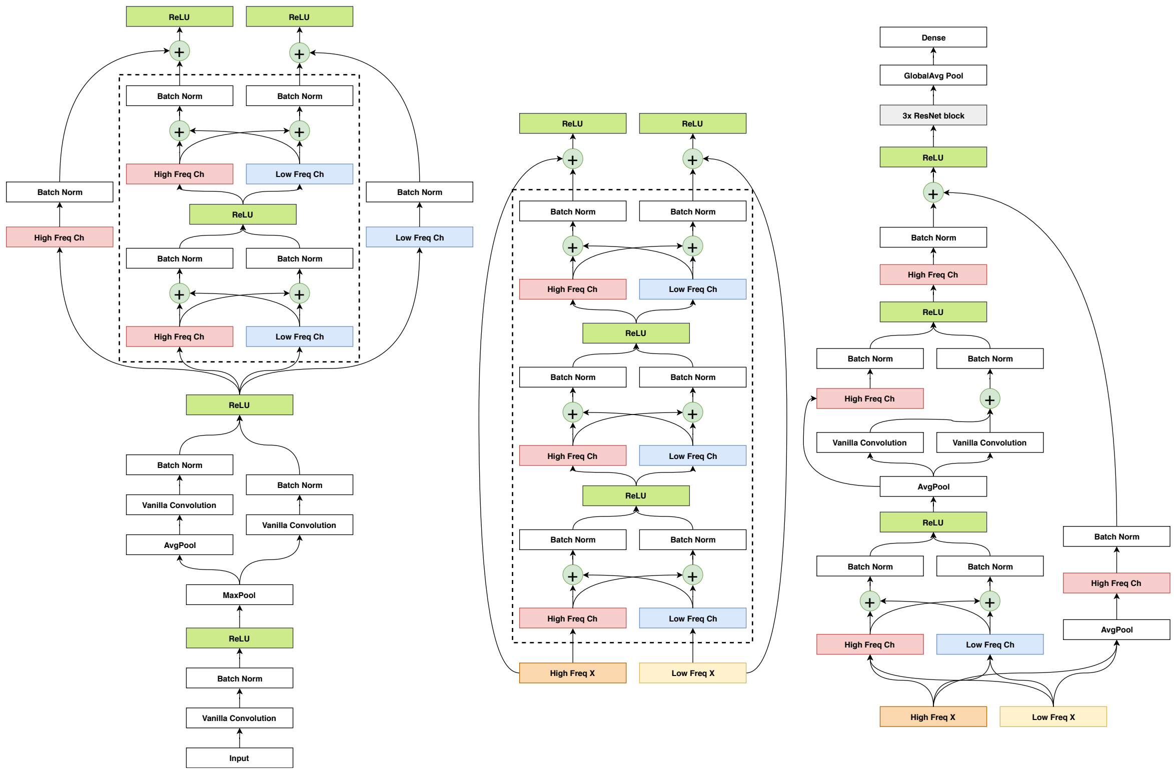

Fig. 1. Simplified architecture of the proposed model blocks of ResNet with octave convolution. To the left is the initial block, where the residual passes through a convolution and batch normalization. In the middle, it is the consecutive block with no processed residual, which is repeated 3 times with 2 levels. Finally, to the right is encountered the top layers which merge the octave convolution into one. After this, the signal is processed by a ReLU activation consecutive with a ResNet block of 3 levels with vanilla convolution, batch normalization, and ReLU activation.

FiSmo [23], is a compiled dataset of fire and smoke images called this way by the authors themselves, which is also publicly available to use at the authors' cloud folder. All datasets contain numerous images with fire and without it. More details of each image dataset will be discussed below.

During the learning process, all datasets were used for training, validation, and testing. The validation was settled with images from the same dataset partitioned and others with the entire dataset.

A. Datasets

Choosing an image dataset is fundamental for training a neural network. The quality and variety of these images will define the generalization that would be capable of achieving the neural network. In this regard, four datasets were selected for training, validation and testing the proposed model. More than 10000 images were used for this purpose, which is described below.

The FireSense dataset [16], is a video compilation which contains 27 videos for fire detection and, 22 videos for smoke detection. From the fire detection videos are 11 videos with fire presence and 16 videos without it. Moreover, the smoke detection videos are 13 videos with smoke presence and 9 videos without it. From this compilation, just the fire detection videos were used. Therefore, frames extraction has been performed for the training of the model. Just one frame per second was obtained from these videos, getting a total of 329 frames with fire presence and 577 frames without it. Some samples are shown in Fig. 2(a).

For the CairFire dataset [18], the authors generated the dataset by selecting images from internet. They present images with different fire scenarios, indoor and outdoor, and different type of illumination as fire-like color. The dataset is highly unbalanced and contains 110 images with fire presence and 541 images without it. Examples of images are shown in Fig. 2(b).

The FireNet dataset [21], it is a recent compilation of challenging images with and without fire presence. The authors complement the datasets used in previous works with internet images to make them more diverse. They produce the dataset summarizing a total of 1124 images with fire presence and 1301 images without it. Examples images are shown in Fig. 2(c). This dataset additionally includes 871 images for testing purpose, where 593 images have fire presence and 278 do not have.

Finally, the FiSmo dataset [23] has also been used. In this dataset, the authors also create a compilation of images from other datasets obtaining more than 9000 images. The source datasets used for FiSmo are in the context of the RESCUER Project[5]. One of the subsets is called Flickr-FireSmoke by

[5]Project FP7-ICT-2013-EU-Brazil - "RESCUER - Reliable and Smart-Crowdsourcing Solution for Emergency and Crisis Management"

the authors, which have in total more than 5000 images. Another subset is the Flickr-Fire, which present balanced quantities of images between fire and no fire from Flickr-FireSmoke, adding 281 other images from other sources with the presence of fire. Additionally, from the BoWFire dataset [24], which is included in the compilation, just the testing subset is used as part of the dataset. This selection is made because the training subset is meant to training a pixel-value fire recognition algorithm. Since the SmokeBlock contains only images with smoke presence, this subset is omitted in this work. A total of 2004 images with fire presence and 4059 images without it are used from FiSmo dataset.

Some example images are shown in Fig. 2(d). Summarizing all the datasets, a total of 10916 images are used in this work. From these, 4160 images have the presence of fire and 6756 do not have. In Table I is shown a detailed trade-off summary of images with fire and no fire from the used datasets.

TABLE I
SUMMARY OF DATASETS USED FOR THIS WORK.

Dataset	Fire	No Fire	Total
FireSense	329	577	906
CairFire	110	541	651
FireNet	1124	1301	2425
FireNet (testing)	593	278	871
FiSmo	2004	4059	6063
Total	4160	6756	10916

B. *Implementation details*

For the training process, all the aforementioned datasets were used. For this purpose, all the algorithms were implemented with the Python programming language, using the framework Keras with TensorFlow as a back-end for the neural network approach. The images were normalized to make their values $\in [0, 1]$ before being processed by the model. The model was trained during 100 epochs with Adam optimizer with Nesterov momentum with a learning rate of $\alpha = 0.0001$.

Each one of the datasets was used for training, validation, and testing of the proposed model to achieve the best possible generalization. For example, using cross-dataset validation, the entire FireSense dataset was used for training and, Cairfire, FireNet, and FiSmo were used separately for validation. Therefore, each dataset obtains four different training results, one by splitting the dataset itself for training and validation, and the other three using cross-dataset validation.

When the same dataset was used, it was split into two different sets, one for training and another for validation. For the FireSense and FiSmo datasets, the 80% of the images were used for training purpose and 20% for validation. For CairFire and FireNet datasets, the same divisions made by the authors at [18] and [21] were used respectively. In this regard, for the CairFire dataset, 549 images were used for training and, 102 images for validation. From the training subset, 59 images had fire presence, and 490 did not have. The validation subset was composed of 51 images with fire presence and 51 images without it. To the FireNet dataset, a 70% was used for training, while the other 30% of the images was used for validation.

As previously mentioned, for validation different datasets are also used for testing. For example, if CairFire is used for training, and FireSense for validation, FireNet, and FiSmo are used for testing. The only exception is the FireNet dataset, which already contains different images for training and testing. For statistical analysis, the algorithm has been executed ten times for each dataset.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The obtained results for the pre-processed datasets during the training were achieved with the following configurations. For each dataset, the training was repeated ten times. Each training had a duration of 100 epochs for each of those run. The validation during training was performed with images of the same dataset and also with cross-dataset validation. The average results are shown in Table II. Fig. 3 shows the best models obtained from all the runs for each validation.

TABLE II
AVERAGE RESULTS FROM 10 TIMES OF TRAINING EXECUTION OBTAINED WITH CROSS-DATASET VALIDATION. FOR THE RESULTS OBTAINED IN FIRESENSE AND FISMO WITH THEMSELVES, A DATASET SPLIT OF $80\%/20\%$ WAS USED FOR TRAINING AND VALIDATION, RESPECTIVELY. FOR THE CAIRFIRE, 549 IMAGES WERE USED FOR TRAINING AND, 110 IMAGES FOR VALIDATION. FOR FIRENET A $70\%/30\%$ SPLIT WAS USED.

Dataset	FireSense	CairFire	FireNet	FiSmo
FireSense	**100%** **±0.0%**	82.87% ±0.89%	67.84% ±2.87%	69.71% ±3.21%
CairFire	83.43% ±2.08%	**90.78%** **±1.61%**	90.65% ±2.08%	79.59% ±1.21%
FireNet	85.14% ±2.29%	100% ±0.0%	**95.47%** **±0.34%**	81.90% ±0.69%
FiSmo	88.29% ±2.58%	94.09% ±0.50%	84.49% ±0.81%	**87.44%** **±0.42%**

In order to check generalization, cross-dataset testing is carried out with the best models obtained during the validation. For the testing step, the accuracy, precision, recall, and f1-score metrics are used. Table III show the testing metrics for the model using the FireNet dataset for training. Different datasets are used for validation (abbreviated as "Val."), and testing. CF and FS are the abbreviations for CairFire and FireSense, respectively.

TABLE III
NO-FIRE LABEL CLASSIFICATION TESTING METRICS FOR TRAINING WITH FIRENET DATASET USING CROSS-DATASET VALIDATION.

Training: FireNet			**No Fire**		
Val.	**Testing**	**Accuracy**	**Precision**	**Recall**	**f1-score**
FireNet	FireSense	75.06%	90.72%	67.76%	77.58%
	CairFire	99.69%	100%	99.63%	99.81%
	FireNet_test	96.33%	92.41%	96.40%	94.37%
	FiSmo	77.88%	93.09%	72.33%	81.41%
CF	FireSense	83.44%	93.48%	79.55%	85.95%
	FireNet_test	92.19%	84.31%	92.80%	88.36%
	FiSmo	78.91%	89.47%	77.63%	83.13%
FS	CairFire	100%	100%	100%	100%
	FireNet_test	95.87%	96.54%	90.29%	93.31%
	FiSmo	80.34%	90.08%	79.38%	84.39%
FiSmo	FireSense	85.87%	85.81%	93.24%	89.37%
	FireNet_test	86.11%	70.60%	96.76%	81.64%
	CairFire	100%	100%	100%	100%

(a) Images samples of FireSense dataset

(b) Images samples of CairFire dataset

(c) Images samples of FireNet dataset

(d) Images samples of FiSmo dataset

Fig. 2. Images samples of each used dataset. From (a) to (d), first rows of each subfigure, show images with the label of fire, and the second rows show images with the label no fire.

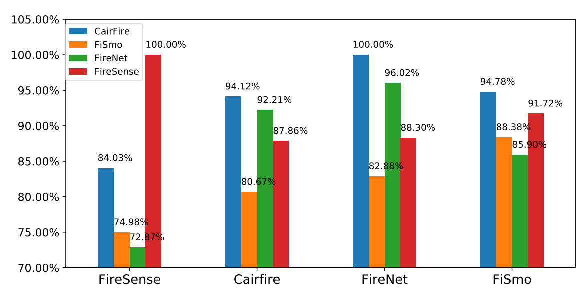

Fig. 3. Best accuracy obtained during the training of cross-dataset validation from 10 runs, during 100 epochs each.

A. Performance Comparisons

After training, the overall obtained results show excellent performance for each dataset. These results represent an interesting improvement, considering the previously reported outcomes in the original works where each dataset was presented. Following, the main differences are presented.

For FireSense dataset, an accuracy of 95.27% has been previously reported on [16]. In their method, the authors make sections at the image with a size of $8x8$ for evaluation. Our proposed model, using FireSense dataset, achieves a 100% of accuracy for the 20% of validation during training.

Using the CairFire dataset, the authors present a modified version of the ResNet50, which obtains an accuracy of 92.15% at most [18]. For this dataset, our proposed model achieves a maximum accuracy of 94.12%.

For the next dataset, used in FireNet [21], the authors presented a lightweight model for an internet-of-things application. This model shows a testing accuracy of 93.91%, comprising more than 600,000 parameters. In comparison, our model achieves a testing accuracy of 96.33% using 950,000 parameters approximately. Although our model comprises more parameters, the proposed model uses less memory and it is computationally efficient due to the octave convolutions.

Finally, for FiSmo, the authors have only presented the dataset. To the best of our knowledge, no work has been presented yet using this dataset, and therefore, no comparison has been performed. Regardless, our proposed model obtains an average accuracy of 87.44% and a maximum of 88.38% for these images.

B. Discussion

Given the obtained results, our model is capable of achieving a high precision on datasets with a reduced number of images. However, datasets with a greater variety of fire-like objects in the images are more difficult to generalize. Regardless of these problems, the proposed model is still capable of achieving high rates of accuracy in testing. Furthermore, the reduced number of parameters does not affect the overall performance negatively. From the addressed datasets, FiSmo presents a more challenging number of images in comparison to the others. For this dataset, during the validation process, the proposed model obtains 87.44% of accuracy.

VI. Conclusions

In this work, we present a fire-recognition model to achieve better performance in comparison to previously presented works. Moreover, we introduce a cross-dataset validation as a baseline to compare fire-recognition algorithm performances with a great number of images and variety. In this regard, the FiSmo dataset is an excellent approach to test the generalization of our algorithm.

The obtained training outcomes achieved by our model show excellent performance. High testing accuracy is obtained with great precision for no-fire recognition. Additionally, our model includes a reduced number of parameters as well as low computational costs. Thus, our approach is suitable for being implemented in mobile devices.

As future works, we propose to implement the model on a limited-hardware device for better performance testings. Furthermore, we plan to extend our work to fire detection using a bounding box approach as well.

Acknowledgment

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, the Brazilian agencies FACEPE, CNPq, and Universidad Central de Chile under the research project CIP2018009.

References

[1] J. Moreno, M. Arianoutsou, A. González-Cabán, F. Mouillot, W. Oechel, D. Spano, K. Thonicke, V. Vallejo, and R. ed.) Vélez, "Forest fires under climate, social and economic changes in europe, the mediterranean and other fire-affected areas of the world," *FUME : lessons learned and outlook*, 1 2014.

[2] J. Barlow and C. A. Peres, "Avifaunal responses to single and recurrent wildfires in amazonian forests," *Ecological Applications*, vol. 14, no. 5, pp. 1358–1373, 2004. [Online]. Available: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/03-5077

[3] X. Úbeda and P. Sarricolea, "Wildfires in chile: A review," *Global and Planetary Change*, vol. 146, pp. 152 – 161, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0921818116302090

[4] N. V. Urzúa Valenzuela and M. F. Cáceres Bueno, "Incendios forestales: principales consecuencias económicas y ambientales en chile," *Revista Interamericana de Ambiente y Turismo - RIAT*, vol. 7, no. 1, pp. 18 – 24, 2011. [Online]. Available: http://riat.utalca.cl/index.php/test/article/view/108

[5] V. Quintanilla Pérez, "Perturbaciones a la vegetación nativa por grandes fuegos de 50 años atrás, en bosques nordpatagónicos. caso de estudio en chile meridional," *Anales de Geografía de la Universidad Complutense*, vol. 28, no. 1, pp. 85 – 104, 5 2008. [Online]. Available: https://revistas.ucm.es/index.php/AGUC/article/view/AGUC0808110085A

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 5 2015. [Online]. Available: http://www.nature.com/articles/nature14539

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017. [Online]. Available: http://doi.acm.org/10.1145/3065386

[8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.

[9] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[11] B. Madhevan, R. Sakkaravarthi, G. Mandeep Singh, R. Diya, and D. K. Jha, "Modelling, simulation and mechatronics design of a wireless automatic fire fighting surveillance robot," *Defence Science Journal*, vol. 67, no. 5, pp. 572–580, 9 2017.

[12] P. Barmpoutis, K. Dimitropoulos, and N. Grammalidis, "Real time video fire detection using spatio-temporal consistency energy," in *2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 8 2013, pp. 365–370.

[13] K. Dimitropoulos, F. Tsalakanidou, and N. Grammalidis, "Flame detection for video-based early fire warning systems and 3d visualization of fire propagation," 6 2012, pp. 18–20.

[14] Y.-H. Kim, A. Kim, and H.-Y. Jeong, "Rgb color model based the fire detection algorithm in video sequences on wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 10, no. 4, p. 923609, 2014.

[15] G. Nikos, A. Cetin, K. Dimitropoulos, F. Tsalakanidou, K. Kose, O. Gunay, B. Gouverneur, D. Torri, E. Kuruoglu, S. Tozzi, A. Benazza-Benyahia, F. Chaabane, B. Kosucu, and C. Ersoy, "A multi-sensor network for the protection of cultural heritage," 8 2011.

[16] K. Dimitropoulos, P. Barmpoutis, and N. Grammalidis, "Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 2, pp. 339–351, 2 2015.

[17] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, "Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 6 2018.

[18] J. Sharma, O.-C. Granmo, M. Goodwin, and J. T. Fidje, "Deep convolutional neural networks for fire detection in images," in *Engineering Applications of Neural Networks*, G. Boracchi, L. Iliadis, C. Jayne, and A. Likas, Eds. Cham: Springer International Publishing, 2017, pp. 183–193.

[19] D. Shen, X. Chen, M. Nguyen, and W. Q. Yan, "Flame detection using deep learning," in *Proceedings - 2018 4th International Conference on Control, Automation and Robotics, ICCAR 2018*. Institute of Electrical and Electronics Engineers Inc., 6 2018, pp. 416–420.

[20] A. Namozov and Y. I. Cho, "An efficient deep learning algorithm for fire and smoke detection with limited data," *Advances in Electrical and Computer Engineering*, vol. 18, pp. 121–128, 11 2018.

[21] A. Jadon, M. Omama, A. Varshney, M. S. Ansari, and R. Sharma, "Firenet: A specialized lightweight fire & smoke detection model for real-time iot applications," *CoRR*, vol. abs/1905.11922, 2019. [Online]. Available: http://arxiv.org/abs/1905.11922

[22] Y. Chen, H. Fan, B. Xu, Z. Yan, Y. Kalantidis, M. Rohrbach, S. Yan, and J. Feng, "Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution," 4 2019. [Online]. Available: http://arxiv.org/abs/1904.05049

[23] M. T. Cazzolato, L. P. S. Avalhais, D. Y. T. Chino, J. S. Ramos, J. A. d. Souza, J. F. Rodrigues Junior, and A. J. M. Traina, "Fismo: a compilation of datasets from emergency situations for fire and smoke analysis," in *Brazilian Symposium on Databases - SBBD*. SBC, 2017.

[24] D. Y. T. Chino, L. P. S. Avalhais, J. F. Rodrigues Jr., and A. J. M. Traina, "Bowfire: Detection of fire in still images by integrating pixel color and texture analysis," *CoRR*, vol. abs/1506.03495, 2015. [Online]. Available: http://arxiv.org/abs/1506.03495