



Détection de la fumée et du feu par réseau de neurones convolutifs

Sébatien Frizzi, Rabeb Kaabi, Moez Bouchouicha, Jean-Marc Ginoux, Farhat Fnaiech, Eric Moreau

► To cite this version:

Sébatien Frizzi, Rabeb Kaabi, Moez Bouchouicha, Jean-Marc Ginoux, Farhat Fnaiech, et al.. Détection de la fumée et du feu par réseau de neurones convolutifs. Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle, Jul 2017, Caen, France. Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle. <hal-01561518>

HAL Id: hal-01561518

<https://hal.archives-ouvertes.fr/hal-01561518>

Submitted on 12 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Détection de la fumée et du feu par réseau de neurones convolutifs

Sébastien Frizzi¹ Rabeb Kaabi^{2,3,4} Moez Bouchouicha^{2,3} Jean-Marc Ginoux^{2,3} Farhat Fnaiech⁴ Eric Moreau^{2,3}

¹Département Génie Biologie- IUT de Toulon, 83957 La Garde, France

²Aix Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, 13397 Marseille, France

³Université de Toulon, CNRS, LSIS, UMR 7296, 83957 La Garde, France

⁴Université de Tunis, ENSIT, LR13ES03, SIME, 1008, Montfleury, Tunisie

Résumé

La recherche dans le domaine de la détection du feu et de la fumée devient un sujet de plus en plus traité. Les algorithmes conventionnels utilisent exclusivement des modèles basés sur des vecteurs de caractéristiques. Ces vecteurs restent difficiles à définir et dépendent largement du type de feu observé. Ces méthodes donnent des résultats avec un faible taux de détection et un haut taux de fausse classification. Une approche innovante pour résoudre ce problème est d'utiliser un algorithme permettant de déterminer automatiquement les caractéristiques utiles pour classifier le feu et la fumée. Dans cet article, nous proposons un réseau de neurones convolutif pour identifier le feu et la fumée d'une vidéo en temps réel. Les réseaux de neurones convolutifs ont montré leur grande performance dans le domaine de la classification des objets. Testée sur des séquences vidéos réelles, l'approche proposée atteint une meilleure performance de classification que les méthodes conventionnelles. Ces résultats indiquent clairement que l'utilisation des réseaux de neurones convolutifs pour la détection du feu et de la fumée est très prometteuse.

Mots Clef

Feu, fumé, détection, réseau de neurones convolutif, apprentissage profond, cartes de caractéristiques, dropout, maxpooling.

Abstract

Research on video analysis for fire detection has become a hot topic in computer vision. However, the conventional algorithms use exclusively rule-based models and features vector to classify whether a frame is fire or not. These features are difficult to define and depend largely on the kind of fire observed. The outcome leads to low detection rate and high false-alarm rate. A different approach for this problem is to use a learning algorithm to extract the useful features instead of using an expert to build them. In this paper, we propose a convolutional neural network (CNN) for identifying fire in videos. Convolutional neural network are shown to

perform very well in the area of object classification. This network has the ability to perform feature extraction and classification within the same architecture. Tested on real video sequences, the proposed approach achieves better classification performance as some of relevant conventional video fire detection methods and indicates that using CNN to detect fire in videos is very promising. It's the English version of the abstract.

Keywords

Fire and smoke detection, deep learning, convolutional neural network, feature maps, max pooling, dropout

1 Introduction

La détection du feu et de la fumée est cruciale pour la protection des personnes. Des dispositifs ont été développés pour prévenir les dommages provoqués par un feu. La majorité de ces dispositifs sont basés sur des capteurs non optiques et se limitent à la détection dans les locaux d'habitations ou commerciaux. La détection de la fumée par ionisation requiert une grande proximité avec le feu. Par conséquent il est impossible de couvrir une large zone et de donner des informations sur la localisation de la naissance d'un incendie, la direction de propagation, l'importance du feu, etc.

La rapide évolution des caméras numériques a permis aux techniques de détections vidéos via l'évolution des ordinateurs de remplacer les méthodes standard de détection du feu et de la fumée. Les techniques de détection vidéo de la fumée et du feu sont particulièrement adaptées en espace ouvert. De plus, ces systèmes offrent la possibilité d'analyser le comportement et la localisation en 3 dimensions du feu. Actuellement les circuits de surveillance vidéos tant en intérieur qu'en extérieur se développent rapidement. Dans ces circonstances, il serait judicieux de développer un dispositif de détection du feu et de la fumée pouvant utiliser ces équipements sans frais.

La recherche dans ce domaine a commencé dans les années 1990. On dénote quelques algorithmes vidéos de détection du feu et des flammes dans la littérature. La majorité de ces algorithmes se concentrent sur les

caractéristiques de couleurs et de formes du feu et de la fumée [1],[2],[3],[4],[5], [6] et [7]. Ensuite, l'objectif fut de construire un vecteur de caractéristiques multidimensionnel utilisé en entrée d'un algorithme conventionnel de classification : SVM, Réseau de neurones, Adaboost, etc. Par conséquent, des experts deviennent indispensables pour créer les règles de base permettant de déterminer les vecteurs de caractéristiques propres à la détection recherchée. Une approche différente et innovante réside dans le fait de ne pas utiliser d'expert pour la construction du vecteur de caractéristique en faisant extraire automatiquement celui-ci par un algorithme d'apprentissage. Les algorithmes d'apprentissage profond possèdent la capacité d'apprendre les caractéristiques utiles à une classification du feu et de la fumée. Les réseaux de neurones convolutifs (CNN) sont de variantes des réseaux profonds et permettent d'extraire les propriétés topologiques d'une image.

Notre approche est conceptuellement simple. Nous utilisons un réseau de neurones convolutifs comme un puissant détecteur du feu et de la fumée dans des images d'une vidéo. L'avantage de cette approche est que cette méthode est insensible au mouvement de la caméra. Le Réseau convolutif opère directement sur les images brutes en RGB dans le spectre du visible. Le réseau convolutif apprend automatique par un ensemble d'images de la base d'entraînement. L'article est organisé de la manière suivante : le paragraphe suivant relate les travaux connexes. Le paragraphe 3 introduit brièvement les réseaux de neurones convolutifs. Le paragraphe 4 quant à lui décrit en détail l'architecture de notre réseau convolutif. Les résultats expérimentaux et la performance se trouvent dans le paragraphe 5. Enfin, le dernier paragraphe discute des limitations de notre méthode et conclut cet article.

2 Travaux connexes

Le nombre de documents traitant de la détection de feu vidéo dans la littérature augmente de façon exponentielle. Plusieurs chercheurs ont joué un rôle important dans le développement d'algorithmes de détection de feu par traitement vidéo [6]. Verstockt [1] a proposé un détecteur d'incendie à capteurs multiples qui fusionne les caractéristiques de flammes visuelles et non-visuelles des objets en mouvement. Il a utilisé des images thermiques ordinaires vidéo et infrarouge à ondes longues (LWIR). Tout d'abord, il exploite une soustraction de fond dynamique pour extraire des objets en mouvement. De plus, les objets en mouvement LWIR sont filtrés par segmentation d'objets à chaud basée sur l'histogramme. Un ensemble de caractéristiques de flamme analyse ces objets en mouvement en mettant l'accent sur les caractéristiques géométriques, temporelles

et spatiales distinctives des désordres des régions de flamme. Ensuite, une probabilité de flamme LWIR est calculée en combinant la probabilité du désordre de la fenêtre de délimitation, du désordre d'orientation principal et de la rugosité de l'histogramme des objets en mouvement chaud dans LWIR. Toreyin [2] utilise un algorithme de détection vidéo à quatre étapes. D'abord, il a estimé les pixels et les régions en mouvement en utilisant une méthode de fond hybride: une opération de différenciation à trois images est effectuée pour déterminer des régions de mouvement légitime, suivies d'une soustraction de fond adaptative pour extraire toute la région en mouvement. Deuxièmement, il a utilisé un modèle de mélange gaussien dans l'espace de couleur RVB pour détecter les pixels de couleur de feu. La répartition des couleurs du feu est obtenue à partir d'images échantillonnées contenant des régions de feu. Dans la troisième étape, une transformation en ondelettes temporelles est effectuée pour analyser le scintillement de la flamme. Enfin, une analyse spatiale en ondelettes de régions en mouvement contenant des pixels de masque de feu pour évaluer des variations de couleur en valeurs de pixel est effectuée. Des variations spatiales significatives présupposent la région du feu. Celik [3] a développé deux modèles: un pour la détection d'incendie et l'autre pour la détection de fumée. Un modèle de logique floue basé sur des règles a été utilisé à la place des règles heuristiques existantes. Ce choix a rendu la classification plus robuste en distinguant efficacement le feu et le feu comme des objets colorés. Pour la détection des fumées, une analyse statistique a été effectuée sur la base de l'idée que la fumée présente une couleur grisâtre avec un éclairage différent. Borges [4] a utilisé un vecteur de caractéristiques multidimensionnel comme entrée à un classificateur Bayes. Les caractéristiques sont: la rugosité limite des régions de feu potentiel, le moment statistique du troisième ordre des régions de feu potentiel qui définit l'asymétrie, la variance et enfin la quantité de feu d'un cadre à l'autre (varie en raison de la flamme). Il existe des travaux plus récents inspirés par des recherches antérieures connexes [5], [6]. Comme dit précédemment, toutes ces méthodes sont basées sur la règle ou exigent la construction de caractéristiques discriminantes pour détecter le feu. Une autre solution serait l'utilisation de l'algorithme d'apprentissage en profondeur. Dans la section suivante, nous présentons une variante de l'apprentissage profond: Réseau de neurones convolutifs.

3 Réseau de neurones convolutifs

Les réseaux convolutifs ont été introduits pour la première fois par Fukushima [8], il a dérivé une architecture de réseau nerveux hiérarchique inspirée par

le travail de recherche de Hubel [9]. Lecun [10] les a généralisés pour classer les chiffres avec succès et pour reconnaître les numéros de contrôle manuscrit par LeNet-5 qui est montré à la Fig. 1. Ciresan [11] a utilisé les réseaux convolutifs et a réalisé les meilleures performances dans la littérature pour la reconnaissance d'objets multiples pour des bases de données d'images multiples: MNIST, NORB, HWDB1.0, CIFAR10 et l'ensemble de données ImageNet.

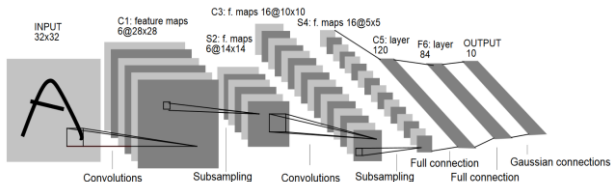


Figure 1. LeNet-5, un réseau neuronal convolutif pour la reconnaissance des chiffres

Un réseau neuronal convolutif se compose de plusieurs couches. Figure. 2, montre ces différentes couches.

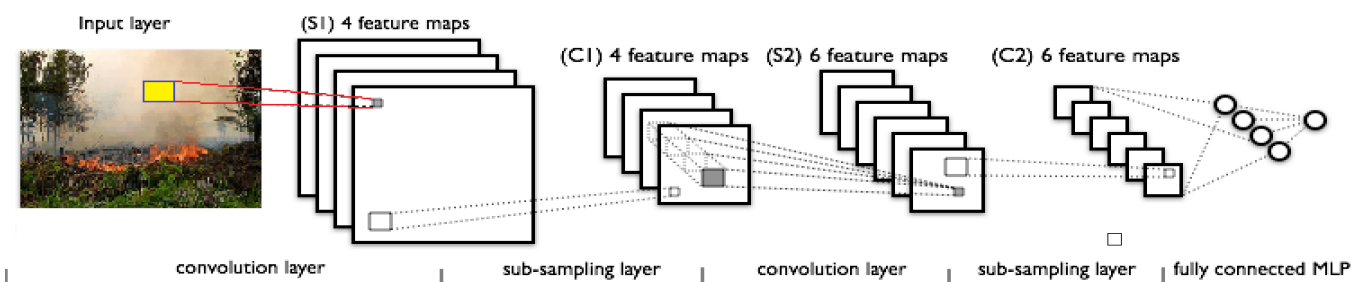


Fig. 2. CNN layers

3.1 Couches convolutives

Les couches convolutives constituent le noyau du réseau convolutif. Ces couches se composent d'une grille rectangulaire de neurones qui ont un petit champ réceptif étendu à travers toute la profondeur du volume d'entrée. Ainsi, la couche convolutionnelle est juste une convolution d'image de la couche précédente, où les poids spécifient le filtre de convolution.

3.2 Couches de pooling

Après chaque couche convolutive, il peut y avoir une couche de pooling. La couche de pooling sous-échantillonne leur entrée. Il y a plusieurs façons de faire cette mise en commun, comme prendre la moyenne ou le maximum, ou une combinaison linéaire apprise des neurones dans le bloc. Par exemple, la Fig. 3. montre max pooling sur une fenêtre 2×2 .

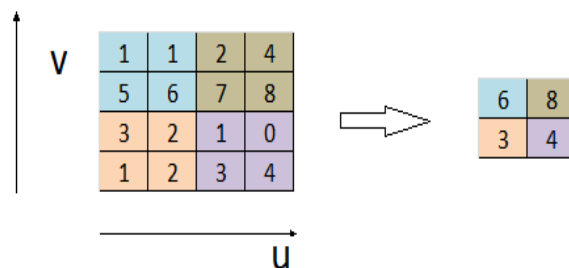


Fig. 3. Max pooling

3.3 Couches totalement connectées

Enfin, après plusieurs couches de convolution et pooling, le raisonnement de haut niveau dans le réseau neuronal se fait via des couches totalement connectées.

Dans les réseaux de neurones convolutifs, chaque couche agit comme un filtre de détection pour la présence de caractéristiques spécifiques ou de motifs présents dans les données d'origine. Les premières couches d'un convolutif détectent des caractéristiques qui peuvent être reconnues

et interprétées relativement facilement. Les couches ultérieures détectent de plus en plus des caractéristiques plus abstraites. La dernière couche du réseau convolutif est capable de faire une classification ultra-spécifique en combinant toutes les caractéristiques spécifiques détectées par les couches précédentes dans les données d'entrée. Dans la section suivante, l'architecture proposée du réseau convolutif pour le feu vidéo et la détection de fumée est présentée.

4 Réseau de neurones convolutifs pour la détection du feu et de la fumée dans une vidéo

4.1 La structure

Notre architecture de classification est classique [17], elle combine convolution et Max pooling. Cependant, pour obtenir une classification rapide permettant une classification et localisation en temps réel, nous avons choisi un réseau léger. Figure. 4 montre les neuf couches de notre réseau convolutif.

Une image couleur RVB dans le spectre visible passe successivement par deux opérations convolutives avec noyau de taille 3x3. La même structure est appliquée après la troisième couche. Un Max pooling 3x3 avec la pas 2 suit les couches convolutives deux et cinq. Les couches un à quatre possèdent 16 cartes de caractéristiques. Les couches cinq et six n'ont qu'une seule carte de caractéristiques. La couche sept et huit sont entièrement connectées. La sortie de la dernière couche totalement connectée alimente un Softmax 3 voies produisant une distribution sur 3 classes.

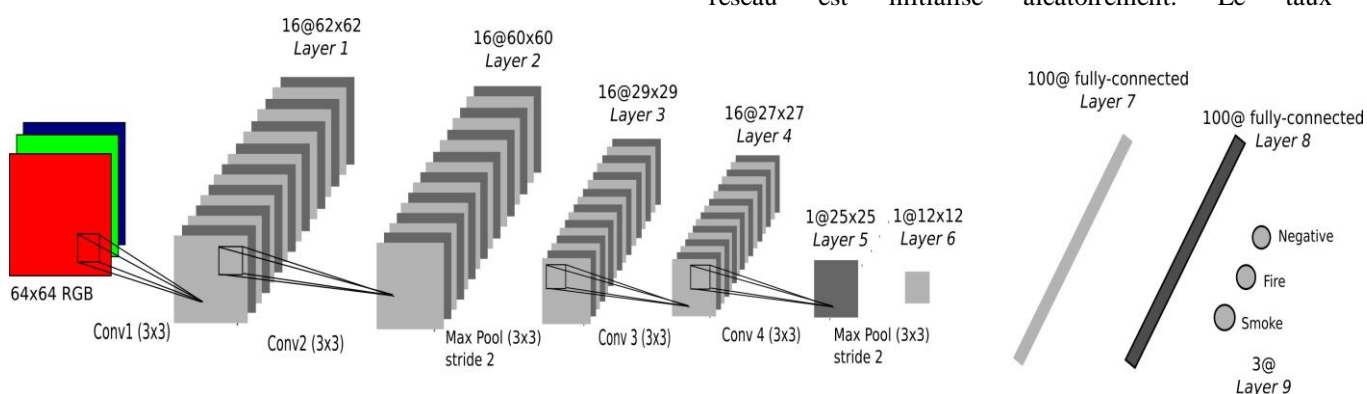


Fig. 4. Architecture de notre réseau convolutionnel

Nous avons choisi [12] et [13] pour les couches convolutives et une entièrement connectée avec une fonction d'activation Leaky ReLu avec coefficient $a = 1/3$ (voir figure 5).

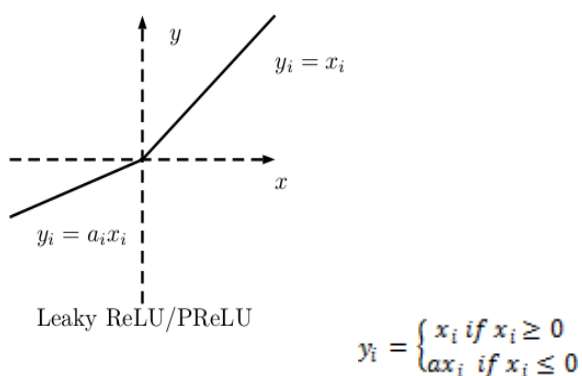


Fig. 5. Leaky ReLu

4.2 Entraînement du réseau

L'objectif de notre classification est de décider si une image contient du feu et / ou de la fumée. Pour résoudre ce problème, l'apprentissage du classifieur est réalisé à partir d'une collection d'images dans le spectre du visible en 3 canaux (RVB) étiquetées. En outre, nous voulons

localiser la position du feu et de la fumée dans une image ou dans une vidéo. L'ensemble d'entraînement est composé de 27919 images RVB étiquetées de taille 64x64 pixels. 8915 pour la fumée, 7257 pour le feu et 11752 négatives (pas de feu ou de fumée). Nous créons 3 sous-ensembles: training 60% des images, validation 20% et test 20%. L'entraînement a été réalisé avec un ordinateur composé d'un microprocesseur Intel Xeon (fréquence CPU 3,1Ghz, RAM 16Go) et une carte graphique GTX 980 Ti (2816 cœurs, 6 Go de mémoire). Nous avons utilisé une descente de gradient stochastique (SGD) avec des mini-lots de taille 100. Le poids dans le réseau est initialisé aléatoirement. Le taux

d'apprentissage initial est de 0,01 et le momentum de 0,9. Le taux d'apprentissage diminue d'un facteur 0,95 chaque 5 itérations (epochs). En revanche, le momentum augmente pour atteindre 0,9999. Pour obtenir la meilleure précision possible pour ces paramètres, plusieurs essais ont été effectués. Nous avons implémenté CNN avec Theano [14], [15] et Lasagne [16]. Nous utilisons un dropout de 0,5 dans les deux couches entièrement connectées pour éviter le surapprentissage. L'apprentissage des paramètres du réseau est réalisé sur 100 cycles ou itérations.

5 Résultats

La précision de classification sur l'ensemble test est de 97,9%. L'ensemble d'essais est composé de 1427 images d'incendie, 1758 images de fumée et 2399 images négatives, donc 5584 images. Les images de fumée représentent plus une texture que l'objet « fumée » en lui-même. Les tableaux 1 à 3 donnent la matrice de confusion pour chaque classe. Sur la matrice de confusion du feu, les faux négatifs et les faux positifs ne contiennent pas d'images de fumée. De la même manière, la matrice de confusion des fumées ne contient pas d'image de feu pour les faux négatifs et les faux positifs. Nous pouvons conclure que les paramètres de notre modèle de réseau convolutif permettent une bonne distinction de classification entre le feu et la fumée. En

outre, les surfaces sous la courbe ROC (Receiver Operating Characteristic) de la Fig. 6, pour les trois classes sont proches de l'unité, indiquant une bonne classification sur l'ensemble d'essai. La courbe ROC feu a une aire plus grande que les autres, ce qui indique une meilleure classification des performances pour le feu.

TABLE I. CONFUSION MATRIX POUR LE FEU

Fire	True class		
	True	False	
Hypothesis class	True	1400	3 ^a
	False	27 ^a	4154

^anot smoke images

TABLE II. CONFUSION MATRIX POUR LA FUMÉE

Smoke	True class		
	True	False	
Hypothesis class	True	1698	26 ^a
	False	60 ^a	3800

^anot fire images

TABLE III. CONFUSION MATRIX POUR LES NEGATIFS (NI FIRE/SMOKE)

No Fire/Smoke	True class		
	True	False	
Hypothesis class	True	2370	87 ^a
	False	29 ^b	3098

^aImage Fire 27 – image Smoke 60

^bImage Fire 3 – image Smoke 26

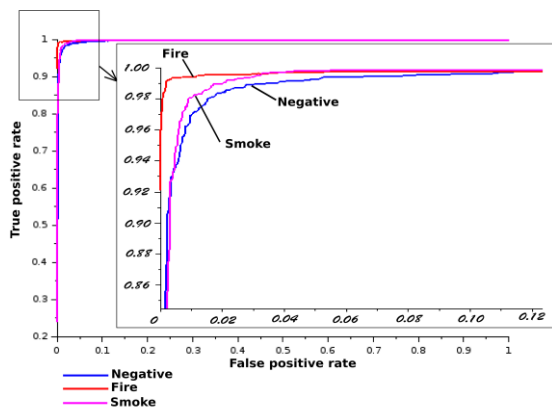


Fig. 6. Courbe ROC pour les 3 classes: fumée, feu et négative

Notre objectif est de détecter en temps réel le démarrage du feu ou de caractériser un incendie sur une vidéo. Le temps de traitement pour la détection est un facteur clé avec la précision. Par conséquent, nous décidons d'utiliser la "structure légère" décrite dans la Fig. 4. Les méthodes actuelles utilisent des fenêtres coulissantes pour détecter et classer l'objet sur des images originales

ou redimensionnées. Chaque fenêtre passe par le réseau de neurones convolutif et les couches totalement connectées pour finir par le classifieur. Pour analyser l'image entière d'un cadre vidéo, la position de la fenêtre doit changer et aller de nouveau à travers le réseau neuronal convolutif. Notre approche est tout à fait différente, au lieu de glisser une fenêtre de 64x64 pixels dans l'image d'origine RGB pour localiser le feu et la fumée; Nous décidons de travailler sur la dernière carte de caractéristiques. Nous divisons le réseau en deux parties. La première partie est composée de 6 couches: 1 à 6 (couches convolutives et Max pooling), la deuxième partie est composée uniquement des entièrement connectées Fig. La figure 7 montre cette architecture CNN.

En utilisant la première partie du réseau, nous évaluons la dernière carte de caractéristiques (couche 6) de l'image entière. Nous savons que sur la base de la structure du réseau de la CNN, une fenêtre coulissante de taille 64x64 pixels dans l'image RVB correspond à une taille de fenêtre de 12x12 pixels dans la dernière carte de caractéristiques. Fig 8.

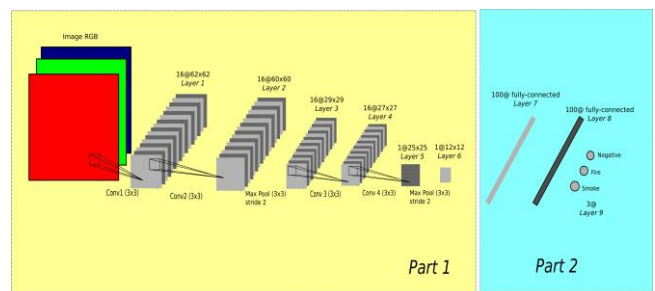


Figure 7. Deux parties de notre réseau convolutif. Partie 1: 6 couches (convolutives, maxpooling). Partie 2: Deux couches totalement connectées et la couche de sortie.

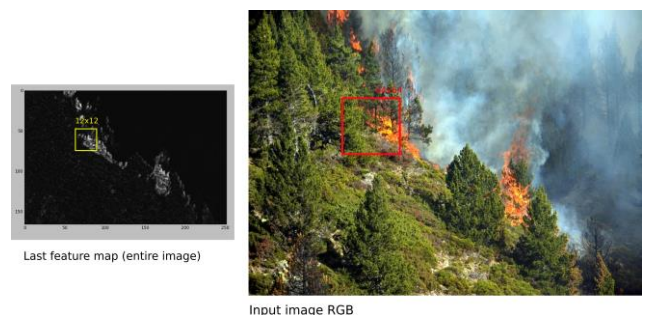


Figure 8. Fenêtre glissante: Image originale vs Carte des caractéristiques

Pour détecter le feu et la fumée dans l'image d'une vidéo, nous appliquons une fenêtre glissante de taille 12x12

pixels sur la dernière carte de caractéristiques (voir la figure 8). Pour accélérer la prédiction pour chaque fenêtre 12x12 de l'image entière, nous réalisons un tenseur 12x12x1xN (N: nombre de fenêtres permettant la détection du feu ou de la fumée) à partir de la dernière carte de caractéristiques et nous utilisons le GPU de la carte graphique. Avec cette méthode, la précision semble inchangée et la vitesse de détection et de prédiction augmentent en fonction de la taille d'origine de l'image et du nombre de fenêtres à prédire comme représenté sur la Fig. 9. et Fig. 10.

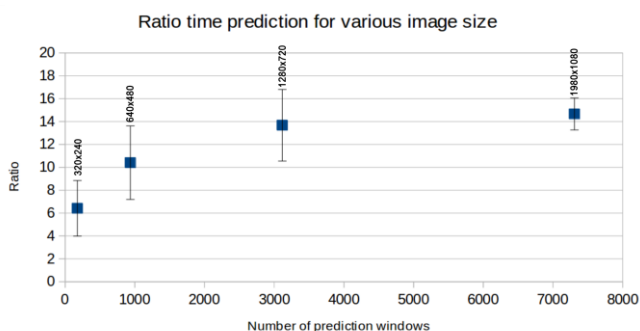


Figure. 9. Ratio du temps de prédiction pour différentes tailles d'image.

Ratio = (temps de prédiction pour l'image originale entière) / (temps de prédiction sur la dernière carte de fonction + temps pris pour réaliser la dernière carte de caractéristique). Fenêtres glissantes par pas de 16px

sur l'image originale et 4px sur la dernière carte de caractéristiques. Expérience réalisée sur 200 images d'une vidéo.

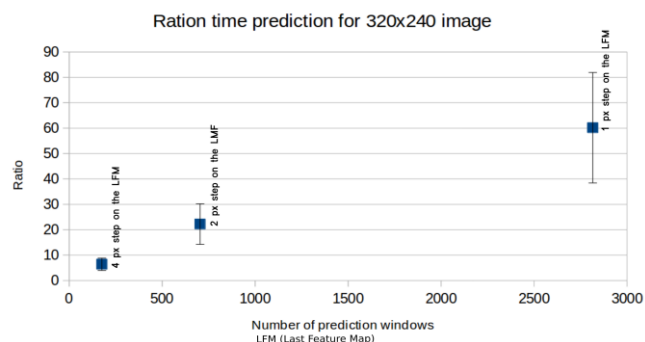


Figure. 10. Ratio du temps de prédiction pour une taille d'image 320x240 avec différent pas de glissement. Expérience réalisée sur 200 images d'une vidéo.

La classification et la localisation affichée à la Fig. 11 et 12, a été effectuée en faisant glisser une fenêtre de 12x12 sur la carte de caractéristique par pas de 2 pixels. Figure. 9.1 (c) et la Fig. 9.2. C) décrivent le masque de détection et de localisation. La couleur rouge représente l'incendie détecté et la couleur verte représente la fumée détectée. L'intensité des couleurs rouge et verte varie en fonction de la probabilité de détection de feu et / ou de fumée. La localisation du feu et de la fumée dans la carte des

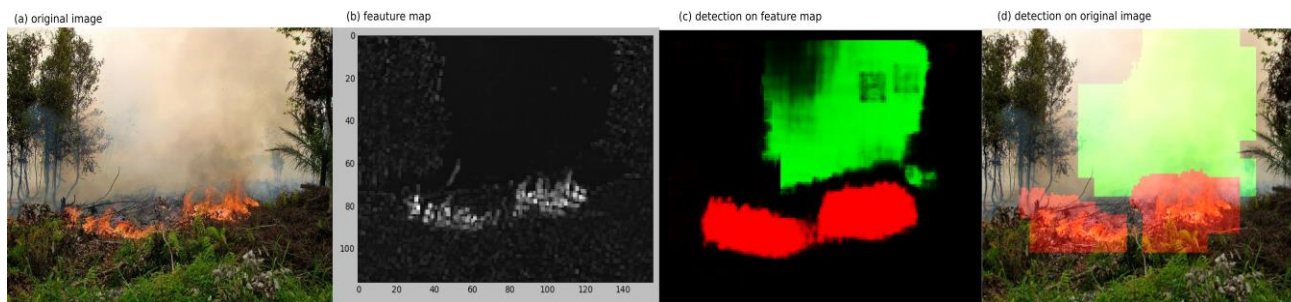


Fig. 11. Feu/fumée détection sur une image de forêt

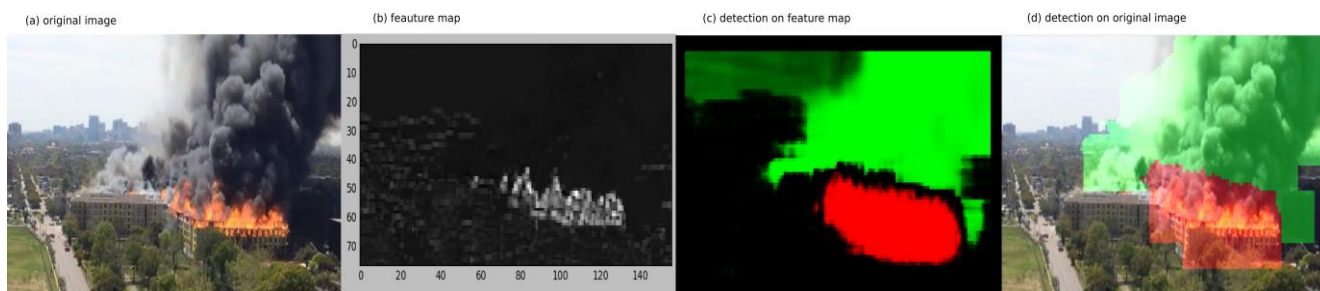


Fig. 12. Feu/fumée détection sur une image d'immeuble

caractéristiques permet de projeter ces positions sur l'image RVB originale (Rectangles rouge et vert).

6 Conclusion

Dans cet article, une méthode de détection de fumée et de feu dans le spectre du visible a été présentée. L'algorithme proposé utilise une approche d'apprentissage profond basée sur le réseau neuronal convolutif. La matrice de confusion et les courbes ROC indiquent une très bonne précision globale pour l'étape de détection. Nous avons montré que notre méthode de travail sur la dernière carte de caractéristique permet de diminuer le temps de prédiction d'un facteur allant de 6 à 60.

Dans les travaux futurs, nous souhaitons améliorer la méthode en utilisant un réseau de neurones convolutionnels 3D pour prendre en compte la dynamique temporelle dans notre modèle. En effet, le réseau convolutif est actuellement limité à gérer les entrées 2D qui nous amènent à traiter l'entrée vidéo seulement image par image. En revanche, le réseau convolutif 3D extrait des caractéristiques spatiales et temporelles en effectuant des convolutions 3D. Ainsi, l'information de mouvement du feu et de la fumée pourrait être encodée, ce qui permettrait de diminuer considérablement le temps de détection. De plus, pour optimiser la détection et la localisation de la fumée et du feu sur une vidéo, nous devons améliorer notre base d'image. La fumée est plus difficile à détecter et à localiser en raison de la nature de sa forme et de sa texture. Nous pourrions améliorer cette détection en augmentant la taille du réseau ou en ajoutant des images de l'objet « fumée » dans la base d'apprentissage. Notre modèle ne détecte que le feu rouge, pour détecter d'autres couleurs de feu, nous devons augmenter notre base d'image avec d'autres couleurs de feu telles que bleue, etc ... En outre, nous prévoyons de comparer notre algorithme aux méthodes conventionnelles sur une plus grande variété des images de feu vidéo: différents matériaux, sources et ventilations. Enfin, une approche intéressante serait de coupler notre réseau convolutionnel dans le spectre du visible avec une image infrarouge dans la bande LWIR afin de focaliser notre recherche dans une zone déterminée de l'image. Ceci permettrait également de différencier la fumée et les nuages.

References

- [1] S. Verstockt, A. Vanoosthuyse, S. Van Hoecke, P. Lambert, and R. Van de Walle, Multi-sensor fire detection by fusing visual and non-visual flame features, In Proceedings of International Conference on Image and Signal Processing, June 2010, pp. 333–341.
- [2] B. U. Toreyin, Y. Dedeoglu, U. Gudukbay, A. E. Cetin, Computer vision based method for real-time fire and flame detection, Pattern recognition letters, 2006, 27, 1, pp. 49-58.
- [3] T. Çelik, H. Özkaramanlı and H. Demirel, Fire and smoke detection without sensors: Image processing based approach, *Signal Processing Conference, 2007 15th European*, Poznan, 2007, pp. 1794-1798.
- [4] K. Borges, P. Vinicius, J. Mayer and E. Izquierdo, Efficient visual fire detection applied for video retrieval, *Signal Processing Conference, 2008 16th European*, IEEE, 2008.
- [5] K. Poobalan and S. Liew, Fire detection algorithm using image processing techniques, Proceedings of the 3rd International Conference on Artificial Intelligence and Computer Science (AICS2015), October 2015, pp. 160-168
- [6] P. Gomes, P. Santana and J. Barata, A vision-based approach to fire detection, *International Journal of Advanced Robotic Systems*, 09-2014.
- [7] E. Çetin et al, Video fire detection – Review, *Digital Signal Processing*, Volume 23, Issue 6, December 2013, pp. 1827-1843
- [8] K. Fukushima, Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 1980, 36(4), pp.193–202.
- [9] D. H. Hubel and T. N. Wiesel, Ferrier lecture: Functional architecture of macaque monkey visual cortex, *Proceedings of the Royal Society of London, Series B, Biological Sciences*, 1977, 198(1130):pp. 1–59.
- [10] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, Gradient-based learning applied to document recognition, in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov 1998.
- [11] D. Ciresan, U. Meier; J. Masci; L.M. Gambardella and J. Schmidhuber, *Flexible, High Performance Convolutional Neural Networks for Image Classification, Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume Volume Two 2: pp. 1237–1242, November 2013.*
- [12] X. Bing, N. Wang, T. Chen and M. Li, Empirical Evaluation of Rectified Activations in Convolutional Network, *CoRR abs/1505.00853 (2015): n. pag.*
- [13] A. L. Maas, A. Y. Hannun and A. Y. Ng, Rectify nonlinearities improve neural network acoustic model, *ICML 2013 Workshop on Deep Learning for Audio, Speech, and Language Processing*, June 2013, Atlanta
- [14] F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. Goodfellow, A. Bergeron, N. Bouchard, D. Warde-Farley and Y. Bengio. "Theano: new features and speed improvements". *NIPS 2012 deep learning workshop*.
- [15] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley and Y. Bengio. "Theano: A CPU and GPU Math Expression Compiler". *Proceedings of the Python for Scientific Computing Conference (SciPy) 2010. June 30 - July 3, Austin*
- [16] LASAGNE, Lightweight library to build and train neural networks in Theano, <https://github.com/Lasagne/Lasagne>, 13 August 2015.
- [17] S. Frizzi, R. Kaabi, M. Bouchouicha, J. M. Ginoux, E. Moreau and F. Fnaiech, "Convolutional neural network for video fire and smoke detection," *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, Florence, 2016, pp. 877-882.