

Apprentissage profond et traitement d'images pour la détection de fumée

Rabeb Kaabi

▶ **To cite this version:**

Rabeb Kaabi. Apprentissage profond et traitement d'images pour la détection de fumée. Traitement du signal et de l'image [eess.SP]. Université de Toulon; École nationale d'ingénieurs de Tunis (Tunisie), 2020. Français. NNT : 2020TOUL0017 . tel-03326897

HAL Id: tel-03326897

https://theses.hal.science/tel-03326897

Submitted on 26 Aug 2021

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.

Ecole Nationale Supérieure d'Ingénieurs de Tunis
En cotutelle avec
L'Université de Toulon

THÈSE

Présentée à

L'École Nationale Supérieure d'Ingénieurs de Tunis

En cotutelle avec

L'Université de Toulon

en vue de l'obtention du diplôme de

DOCTORAT

en

Génie Électrique

Spécialité : TRAITEMENT DU SIGNAL ET D'IMAGE

par

Rabeb Kaabi

Apprentissage profond et traitement d'images pour la détection de fumée

Responsables d'encadrement :

Mounir Sayadi	Professeur à l'ENSIT, Université de Tunis
Eric Moreau	Professeur et Directeur de Seatech, Université de Toulon
Moez Bouchouicha	Maitre de Conférences, Université de Toulon

Laboratoire de Recherche : Signal, Image et Maîtrise d'Energie (SIME)
ENSIT - LR13ES03- 5 Av. Taha Hussein 1008 Tunis, Tunisie
Equipe SIIM, Laboratoire d'Informatique et Système –Université de Toulon

Résumé

Cette thèse aborde le problème de la détection des fumées dans les vidéos avec des outils de traitement d'images et de l'apprentissage profond, dans le but de la détection précoce des incendies dans les forêts et les zones boisés. De nombreuses méthodes ont montré leur efficacité pour la détection des fumées, qui est une tâche difficile en raison de la grande variabilité de la couleur, de la texture et des formes de la fumée. La plupart de ces méthodes utilise les caractéristiques statistiques temporelles et spatiales classiques (couleur, analyse d'objet en mouvement, analyse de scintillement, énergie, analyse de désordre, etc.) avec des méthodes de classification intelligente basées sur l'apprentissage profond pour améliorer la détection de zones de fumée. On a proposé une méthode alternative neuronale basée sur l'apprentissage profond avec le DBN (Deep Belief Netowrk) en tenant compte d'autres caractéristiques de la fumée comme la couleur, le mouvement et l'énergie. En plus, un ensemble de techniques de prétraitement nous a aidé à construire une base de donnée importante pour tester la robustesse et l'efficacité de la méthode de détection des fumées basée proposée. Les résultats trouvés ont été comparés à ceux proposés dans d'autres travaux (Deep CNN, SVM…) et ont donné des performances acceptables, avec une localisation rapide et précise des zones de fumée.

Abstract

This thesis deals with the problem of forest fire detection using image processing and machine learning tools. A forest fire is a fire that spreads over a wooded area. It can be of natural origin (due to lightning or a volcanic eruption) or human. Around the world, the impact of forest fires on many aspects of our daily lives is becoming more and more apparent on the entire ecosystem.

Many methods have been shown to be effective in detecting forest fires. Most of these methods use classical temporal and spatial statistical characteristics (color, moving object analysis, flicker analysis, energy, disorder analysis, etc.) with intelligent classification methods based on deep learning to improve smoke detection. An alternative neural method based on deep learning with DBN (Deep Belief Netowrk) was proposed taking into account other smoke characteristics such as color, motion and energy. In addition, a set of pre-processing techniques helped us to build an important database to test the robustness and efficiency of the proposed based smoke detection method. The found results were compared to those proposed in other works (Deep CNN, SVM...) and gave acceptable performances, with a fast and accurate localization of smoke regions.

Remerciements

Ce travail a été fait au sein des deux Laboratoires SIME à l'ENSIT et LIS à Sea-Tech de Toulon, et dans le cadre du projet PHC Utique 41755XB, CMCU 19G1126, entre les deux Laboratoire. Je tiens à remercier en premier lieu Mr. Farhat Fnaiech, pour m'avoir accueilli dans son Labo et pour avoir assuré la direction de ma thèse de Doctorat dans sa première partie. Ses conseils et sa tolérance face à mes questions m'ont énormément aidé à surmonter de nombreux obstacles.

Je tiens à exprimer ma profonde gratitude à mes directeurs de thèse, Mr. Mounir Sayadi, Mr Moez Bouchouicha et Mr Eric Moreau, de m'avoir fait découvrir un domaine de recherche intéressant, pour leurs conseils, et leur soutient et pour toutes les discussions précieuses que nous avons eues pendant la période de ce travail.

Mes remerciements s'adressent également à Mr. Aymen Mouelhi pour son aide et son soutien et pour m'avoir fourni de nombreuses informations sur des points précis de mes travaux et publications.

Des remerciements particuliers sont adressés à mes collègues au département de génie électrique de l'ENSIT et à Seatech de Toulon où se construisent la plupart de mes bases théoriques. Je tiens également à remercier tous les membres du projet CMCU PHC, du Laboratoire SIME à l'ENSIT et du Laboratoire LIS à Toulon.

J'aimerais également remercier tous ceux qui m'ont prêté de l'aide pour terminer ce travail dans les conditions les plus favorables.

Liste des abréviations

ADAM :	Adaptive Moment Estimation Algorithm
CNN :	Convolutional Neural Network
DBN:	Deep Belief Network
DCNN:	Denoising Convolutional Neural network
DWT:	Discrete Wavelet Transform
Fast RCNN:	Fast Recurrent Convolutional Neural Network
IoU:	Intersection over Union
SSD:	Single Shot detectors
SVM:	Support Vector Machine
Yolo:	You only look once
SGD :	Stochastic Gradient Descent
RBM :	Restricted Boltzmann Machine (Machine de Boltzmann restreinte)

TABLE DE MATIERES

Introduction générale.....11

Chapitre I : Techniques de détection de fumée traitées dans la littérature.....14

I. Introduction.....15

II. Recueil des techniques utilisées dans la littérature pour la détection de fumée et des feux 16

1. Extraction de caractéristiques de la fumée dans les vidéos : 26

1.1. La caractéristique : Mouvement 26

1.2. La caractéristique : Energie.....29

2. Classification des régions de fumée et non fumée.....31

2.1. Méthodes SVM (Machine à vecteurs de support).....31

2.2. CS Adaboost.....32

2.3. Approche bayésienne 33

2.4. Algorithme KNN 34

2.5. Réseaux de neurones convolutifs CNN.....34

IV. Quelques méthodes de détection de feux et des flammes.....37

IV. Conclusion 40

Chapitre II : Techniques utilisées pour la classification neuronale intelligente et.....41

l'extraction des zones d'intérêt.....41

des zones d'intérêt.....42

I. Introduction 42

II. Rappel sur les systèmes d'apprentissage automatique 42

1. Systèmes à apprentissage automatiques et réseaux de neurones.....42

2. Types d'apprentissage automatique 43

2.1. Apprentissage supervisé.....44

2.2. Apprentissage non supervisé 45

2.3. Apprentissage semi supervisé.....46

2.4. Apprentissage par renforcement 46

III. Détails sur l'extraction et le nettoyage des données.....46

1. Extraction de données.....46

2. Données brutes et données transformées 47
3. Prétraitement et nettoyage des données 47
IV. Méthodes d'extraction des régions d'intérêts à appliquer 47
1. Extraction de région d'intérêt contenant la fumée par Mask-RCNN [Gir14] 48
2. Présentation du détecteur des objets en temps réel : YOLO [Red16] 50
V. Structure du réseau d'apprentissage profond utilisé le DBN (Deep Belief Network) 52
1. Présentation du réseau DBN : (Réseaux de conviction profonde).....52
3. Minimisation de la fonction coût et optimisation des poids [Hin02],[Hin06] 53
4. Comportement du réseau comme un perceptron multicouche après le prétraitement.....55
5. Etape de classification avec la régression logistique 56
VI. Conclusion 60
Chapitre III : Méthode proposée pour la détection de fumée avec le DBN et résultats expérimentaux.....61
I. Introduction et présentation des données 62
II. Détails de la méthode proposée pour de la détection de fumée 64
2. Etape du prétraitement et de localisation des fumées dans les images 66
3. Expériences réalisées pour le choix des paramètres du réseau DBN :.....68
III. Comparaison avec d'autres méthodes de détection de fumée 73
3. Comparaison entre notre approche et la méthode de Pundir et al [Pun19] basée sur le Deep CNN : 77
4. Robustesse de la méthode proposée en présence de bruit additif : 78
IV. Conclusion : 80
Conclusion générale et perspectives 82
Références 85
Liste des publications de Rabeb Kaabi.....91

LISTE DES FIGURES

FIGURE 1.1. PROCESSUS DE LA DETECTION DE LA FUMEE DANS LES METHODES CLASSIQUES 16
FIGURE 1.2. RESULTATS DE DETECTION DE FUMEE DEVELOPPEE PAR XIONG ET AL [3] 18
FIGURE 1.3. METHODE DE DETECTION DE FUMEE PROPOSEE PAR CALDERERA [CAL08] 19
FIGURE 1.4 .EXTRACTION DU MOUVEMENT DE LA FUMEE PROPOSEE PAR ZHAO [ZHA15] 20
FIGURE 1.5. COMPARAISON DE LA PRECISION ET RAPPEL ENTRE LA METHODE DE ZHAO [ZHA15] ET D'AUTRES METHODES 21
FIGURE 1.6. ALGORITHME PROPOSE POUR LA DETECTION DE FUMEE PAR WANG [YUA16] 22
FIGURE 1.7. EXTRACTION DE MOUVEMENT DE LA FUMEE PAR WANG [YUA16] 22
FIGURE 1.8. HISTOGRAMME DES IMAGES CONTENANT LA FUMEE AVANT ET APRES SEUILLAGE [YUA16] 23
FIGURE 1.9. PROCESSUS DE DETECTION DE FUMEE PROPOSEE PAR PUNDIR ET AL [PUN19] 25
FIGURE 1.10. SCHEMA RECAPITULATIF DE LA DETECTION DE FUMEE 26
FIGURE 1.11. EXEMPLE DE MELANGE GAUSSIEN (LES TROIS COMPOSANTES SONT EN COULEUR, ET EN NOIR LE MODELE GAUSSIEN) 27
FIGURE 1.12 : APPLICATION DE LA SOUSTRACTION DU FOND SUR QUELQUES TRAMES 28
FIGURE 1.13. PRINCIPE DU PASSAGE D'UN NIVEAU DE MULTI-RESOLUTION A UN AUTRE DANS LA TRANSFORMEE EN ONDELETTE DISCRETE DES IMAGES 29
FIGURE 1.14. EXEMPLE DE L'APPLICATION DE LA TRANSFORMEE EN ONDELETTES SUR UNE IMAGE, LES DETAILS VERTICAUX APPARAISSENT SUR LA COMPOSANTE HL, LES DETAILS HORIZONTAUX APPARAISSENT SUR LA COMPOSANTE LH 30
FIGURE 1.15. EXEMPLE DE RESULTATS DE DETECTION DE FUMEE PAR LA METHODE PROPOSEE PAR LEE ET AL [13] 32
FIGURE 1.16. APPLICATION DU GMM ET CS ADABOOST POUR LA DETECTION DE FUMEE [HUY14] 33
FIGURE 1.17. DES RESULTATS DE DETECTION DE FUMEE PAR LA METHODE DE TOREYIN ET AL [TOR06] 33
FIGURE 1.18. UN EXEMPLE DE RESULTATS DE DETECTION DE FUMEE PAR LA METHODE DE GUBBI ET AL [GUB09] 34
FIGURE 1.19. ARCHITECTURE DES RESEAUX DE NEURONES CONVOLUTIFS 35
FIGURE 1.20. QUELQUES RESULTATS DE DETECTION DE FUMEE PAR LA METHODE DE FRIZZI ET AL[FRI16] 36
FIGURE 1.21. ARCHITECTURE DU RESEAU UTILISE DANS [HUY18] 36
FIGURE 1.22. RESULTATS DE DETECTION DE FEUX REALISES PAR VERSTOCKT ET AL [VER12] 38
FIGURE 1.23. RESULTATS DE DETECTION DE FEUX OBTENUS PAR CELIK ET AL [CEL07] 39

FIGURE 2. 1. DIFFERENTS TYPES D'APPRENTISSAGE AUTOMATIQUE 43
FIGURE 2. 2. SCHEMA ILLUSTRATIF DU PROCESSUS D'APPRENTISSAGE AUTOMATIQUE (SUPERVISE ET NON SUPERVISE) POUR L'OBTENTION D'UN MODELE PREDICTIF ET CLASSIFICATION 44
FIGURE 2. 3. SCHEMA DE PRINCIPE DE L'APPRENTISSAGE SUPERVISE [RAY15] 44
FIGURE 2. 4. SCHEMA DE PRINCIPE DE L'APPRENTISSAGE NON SUPERVISE [WIT14] 45
FIGURE 2. 5. SCHEMA DE PRINCIPE DE L'APPRENTISSAGE PAR RENFORCEMENT [ZEN20] 46
FIGURE 2. 6. ETAPES DE PREPARATION DE DONNEES POUR LES MODELES A APPRENTISSAGE AUTOMATIQUE 48
FIGURE 2. 7. ARCHITECTURE DE MASK-RCNN 50
FIGURE 2. 8. SCHEMA DE PRINCIPE DE L'ALGORITHME YOLO (YOU ONLY LOK ONCE) [RED16] 51
FIGURE 2. 9. RAPPEL DU SCHEMA DE PRINCIPE DU PERCEPTRON MULTICOUCHES [KHO98] 56
FIGURE 2. 10. ARCHITECTURE UTILISEE DU RESEAU DBN 57

FIGURE 2. 11. COMPARAISON DE LA VARIATION DE LA FONCTION COUT PENDANT L'APPRENTISSAGE EN UTILISANT LE SGD ET L'OPTIMISEUR ADAM 59
FIGURE 2. 12. COMPARAISON DE VARIATION DE LA FONCTION DE PRECISION PENDANT L'APPRENTISSAGE EN UTILISANT LE SGD ET L'OPTIMISEUR ADAM » 60

FIGURE 3. 1. EXEMPLES D'IMAGES DE LA BASE DE DONNEES CONTENANT DE LA FUMEE 63
FIGURE 3. 2. EXEMPLES D'IMAGES DE LA BASE DE DONNEES CONTENANT DE LA FUMEE, DES NUAGES, DES PERSONNES EN MOUVEMENT, DU BROUILLARD ETC. 63
FIGURE 3. 3. EXEMPLES D'IMAGES DE LA BASE DE DONNEES SANS FUMEE 63
FIGURE 3. 4. ORGANIGRAMME DE LA TECHNIQUE PROPOSEE 66
FIGURE 3. 5. SCHEMA DE RECAPITULATION DE LA METHODE PROPOSEE POUR LA DETECTION DE LA FUMEE 66
FIGURE 3. 6. EXEMPLES DE RESULTATS DE L'APPLICATION DU GMM POUR L'EXTRACTION DU MOUVEMENT DE LA FUMEE 67
FIGURE 3.7. EXEMPLES DE RESULTATS OBTENUS DE LOCALISATION DE FUMEE 68
FIGURE 3.8. EXEMPLE DE L'INFLUENCE DE LA VARIATION DU COEFFICIENT D'APPRENTISSAGE *LR* SUR LA FONCTION COUT DU PRETRAITEMENT 69
FIGURE 3. 9. INFLUENCE DE LA VARIATION DU NOMBRE DE COUCHE CACHEES *H* SUR LA CONVERGENCE DE LA FONCTION COUT D'APPRENTISSAGE 70
FIGURE 3. 10. VARIATION DE LA FONCTION COUT D'APPRENTISSAGE ET DE VALIDATION EN FONCTION DE L'EPOQUE 72
FIGURE 3. 11. VARIATION DU TAUX DE DETECTION EN FONCTION DES ITERATIONS DE L'APPRENTISSAGE EN PHASE D'APPRENTISSAGE ET DE VALIDATION 73
FIGURE 3. 12. EXEMPLE D'EVALUATION DE LOCALISATION PAR LA METRIQUE IOU 76

LISTE DES TABLEAUX

TABLEAU 1. 1. TABLEAU RECAPITULATIF DES METHODES ET TECHNIQUES UTILISEES POUR LA DETECTION DE LA FUMEE 37

TABLEAU 2. 1. QUELQUES DOMAINES D'APPLICATION DE L'APPRENTISSAGE SUPERVISE 45
TABLEAU 2. 2. QUELQUES DOMAINES D'APPLICATIONS DE L'APPRENTISSAGE NON SUPERVISE 45
TABLEAU 2. 3. RESUME DES ETAPES PROPOSEES POUR L'EXTRACTION DES REGIONS D'INTERET 49
TABLEAU 2. 4. CHOIX ET DESIGNATION DES HYPER-PARAMETRES 53

TABLEAU 3. 1. INFLUENCE DU NOMBRE DE COUCHES CACHEES SUR LE TEMPS D'APPRENTISSAGE *TP* , LE TEMPS DE FIN DE L'APPRENTISSAGE *TF* ET LE TAUX DE DETECTION *DR* 71
TABLEAU 3. 2. COMPARAISON DES METRIQUES DE CLASSIFICATION PAR LA METHODE PROPOSEE PAR WANG YUANBIN [YUA16] (M1) ET NOTRE METHODE (M2) 75
TABLEAU 3. 3. VARIATION DU TAUX DE DETECTION EN FONCTION DE DIFFERENTS NIVEAUX DU BRUIT GAUSSIEN AJOUTE 79
TABLEAU 3. 4. COMPARAISON ENTRE LES DIFFERENTS CLASSIFICATEURS UTILISES SANS ET AVEC AJOUT D'UN BRUIT DE SNR 5 DB 79

Introduction générale

Introduction générale

Les feux influencent de plusieurs façons la diversité biologique. Ils sont une importante source d'émissions de carbone et participent au réchauffement de la planète et aux changements dans la biodiversité. Les incendies changent le volume de la biomasse et altèrent le cycle hydrologique avec des retombées sur les systèmes marins comme les récifs coralliens. La fumée dégagée par les forêts en feux peut minimiser de façon remarquable l'activité photosynthétique et influence notamment la santé des êtres humains et des animaux. Par exemple, l'Amazonie présente le meilleur sanctuaire de la biodiversité au monde. Des milliers d'espèces de plantes, d'animaux d'insectes et des milliers d'espèces jamais recensées y cohabitent. Des incendies répétitifs menacent alors cette biodiversité. Tous ces dégâts ont introduit de grands défis dans la lutte contre les incendies. La conscience du risque ne cesse de s'internationaliser.

Une alarme précoce de l'apparition des feux dans les forêts serait primordiale pour réduire les pertes dues aux incendies. Cette alarme doit être basée sur une détection avec le maximum de précision et de robustesse. Mais par contre, les technologies traditionnelles de détection des incendies telles que les détecteurs de fumée, ne sont pas adaptées aux grands espaces. La détection classique des fumées est difficile en raison de limites de technologies cités précédemment, des fausses alarmes, des retards de détection et d'autres soucis qui se manifestent fréquemment. La lutte contre le risque des incendies est d'autant plus efficace qu'il est possible d'anticiper son départ et sa propagation. Par ailleurs, les chercheurs ont travaillé sur la détection précoce de feux de forêts en manipulant sur des caméras détectant les images des incendies dans le domaine de visible et dans le domaine d'Infra Rouge.

Par conséquence, la détection d'incendie par traitement des images et des vidéos de surveillance est un sujet de recherche motivant. Cette technologie assure plusieurs avantages comme la détection précoce des incendies avec une grande précision, une installation pratique facile et une capacité de détecter efficacement les incendies dans les grands espaces. Les images captées par des caméras embarquées sur des avions de surveillance ou des drones doivent être traités par des algorithmes efficaces et précis de détection et de localisation de la fumée. Mais cette détection doit être dans un temps minimal avec une bonne détection et

avec le minimum de données pour assurer la facilité de l'implémentation de ce système en temps réel.

La problématique de détection consiste à prédire à quel endroit se trouve la fumée dans une image, tandis que la classification vise à classifier les régions de fumée et non fumée. Ces deux problèmes sont intimement liés et seront tous les deux traitées. Le défi à relever est la « bonne » détection des régions d'intérêts avec un minimum de temps de calcul et un minimum de ressources.

La détection manuelle de la fumée ne peut pas répondre aux exigences d'un taux élevé de détection et a un long temps de traitement. Par ailleurs, les modifications complexes de la forme, de la texture et de la couleur de la fumée restent un défi important pour identifier la fumée dans une image donnée [Fri17]. D'autre part, la majorité des approches proposées dans la littérature tentent de détecter et classifier les régions de fumées dans les trames des vidéos, ce qui nécessite un temps de calcul assez long. En plus, parfois les équipements de contrôle sont très chers (comme les caméras thermiques). En plus, ces méthodes de détection reposent sur l'extraction manuelle des caractéristiques et l'implémentation des classificateurs qui ont parfois démontré des limites et ont causé de l'existence de fausses alarmes. De plus, même si les méthodes conventionnelles présentent un pas important dans le développement du traitement d'image avec l'apprentissage profond, ces techniques demandent beaucoup de ressources et se caractérisent par un temps de calcul élevé. C'est ce qui fait la nécessité de proposer une méthode qui combine entre l'extraction des caractéristiques (couleur, analyse d'objet en mouvement, analyse de scintillement, énergie, analyse de désordre, etc.) et l'apprentissage profond. Cette technique nous emmène à détecter efficacement et rapidement la fumée en limitant les caractéristiques.

Dans cette thèse, on vise donc à développer une nouvelle méthode de détection et de classification appliqués au contexte particulier des fumées surtout dont les images sont à faible résolution. Nous proposons donc une nouvelle méthode basée sur le DBN (Deep Belief Network) pour la détection de la fumée. Contrairement aux réseaux convolutifs profonds classiques (par exemple, Alex-Net, VGG-Net, Res-Net, Dense-Net, [Yin17],[Huy18] et le réseau neuronal à convolution de débruitage DNCNN [Yin17]), qui suivent le processus de classique de reconnaissance des formes avec une extraction et une classification des caractéristiques. En plus, les méthodes de détection de la fumée basées sur le réseau neuronal convolutif CNN [Fri16], [Ter19] sont nettement plus lentes en raison de

l'opération de maxpooling. En outre, la phase de formation peut prendre beaucoup de temps si l'ordinateur n'est pas équipé d'une puissante unité de traitement graphique GPU.

Notre idée est d'effectuer à la fois la reconnaissance de la fumée et la classification des régions avec et sans fumée en utilisant les caractéristiques statiques et dynamiques de la fumée. Ainsi, la contribution de ce travail est le développement d'une étape de prétraitement comprenant une nouvelle combinaison de caractéristiques : couleur, mouvement et énergie de la fumée, pour extraire les régions d'intérêt qui sont insérées dans le DBN [Hin06]. Tout d'abord, le modèle GMM (Modèle de mélange de gaussiennes) est utilisé pour capturer les trames contenant une grande quantité de mouvement. Ensuite, après avoir appliqué les règles RGB aux pixels dans les régions de fumée et analysé l'attitude énergétique de ses régions, les caractéristiques extraites sont ensuite utilisées pour alimenter un DBN de détection de la fumée. Cette méthode est capable de localiser de manière fiable les régions de fumée, avec un temps de calcul minimal et des performance améliorées par rapport à l'existant.

Organisation du rapport de la thèse

Dans le premier chapitre, on présente plusieurs méthodes et techniques issues de la littérature pour la détection de la fumée et du feu, en détaillant les éléments clefs de ces méthodes et en présentant quelques exemples de résultats. Ensuite, les techniques utilisées pour la classification neuronale intelligente et pour l'extraction des zones d'intérêt avec la combinaison des caractéristiques de la fumée et du DBN, sont présentés dans le chapitre 2. Les détails de la méthode proposée pour la détection de fumée avec le DBN et les résultats des expériences réalisées en plus des comparaisons avec les résultats des méthodes classiques sont présentés et discutés dans le Chapitre 3. Enfin, les conclusions et les perspectives sont présentées dans la dernière section du rapport.

Chapitre I : Techniques de détection de fumée traitées dans la littérature

Chapitre I : Techniques de détection de fumée traitées dans la littérature

I. Introduction

La fumée est le premier indicateur de la présence des incendies. Les chercheurs travaillant sur détection de feux de forêts ont focalisé leurs travaux sur la détection de fumée. Au départ, ils ont précisé les différentes caractéristiques de la fumée qui sont citées par la suite. Ces caractéristiques sont de type statiques ou dynamiques (couleur, mouvement, énergie, texture, taille.) [Fri17].

Plusieurs chercheurs ont joué un rôle important dans le développement des méthodes implémentées dans des systèmes de détection vidéo de fumée utiles. Nous nous sommes concentrés sur trois d'entre eux. La première méthode est destinée à la détection de fumée dans le domaine du visible [Tor06], [Xio07]. Dans ce domaine, la distance entre la caméra et la cible est inférieure à 100 m. D'autre part, la deuxième méthode [Ver12] traite des images infrarouges. La dernière [Cet13] combine des images visibles et infrarouges pour détecter la fumée.

Le cout élevé des caméras infrarouge nous pousse à focaliser nos recherches sur la détection de la fumée de feux de forêts dans le domaine du visible. Les méthodes de détection de fumée dans le domaine du visible seront exposées et détaillées par la suite. Cette détection est basée sur le critère du taux de détection et d'autres critères que nous allons citer et détailler par la suite. Plus le taux est élevé, plus la méthode est efficace et fiable.

La méthode de détection de la fumée est récapitulée dans le schéma suivant dans la Figure 1.1. L'extraction de caractéristiques est suivie par la classification du vecteur qui représente une combinaison de différentes caractéristiques. Ensuite, nous allons distinguer les régions de fumées avec l'utilisation des techniques classique de classification et segmentation.

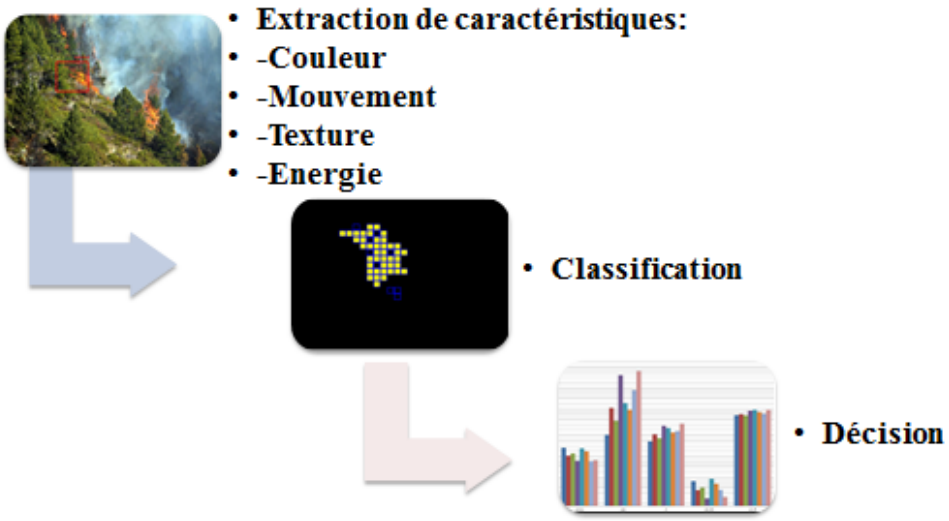


Figure 1.1. Processus de la détection de la fumée dans les méthodes classiques

II. Recueil des techniques utilisées dans la littérature pour la détection de fumée et des feux

Le nombre de documents traitant de la détection de fumée 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 des trames de vidéos.

Ainsi, Il sera très bénéfique si le système de détection d'incendie et de fumée par traitement d'images peut avertir de l'état du feu dès son apparition, et il peut être possible de minimiser les dommages même lorsque les personnes ne sont pas devant le moniteur de contrôle. La détection des incendies et de la fumée est différente de la détection des feux de forêt car certains éléments, comme les phares des voitures et des tunnels, sont différents de l'environnement forestier, ce qui nécessite le développement d'un algorithme spécifique.

L'idée est donc de combiner alors des caractéristiques temporelles et spatiales (couleur, analyse d'objet en mouvement, analyse de scintillement, énergie, analyse de désordre, etc.) avec des algorithmes de classification dans le même bloc pour détecter exactement les régions de fumée.

Récemment, les chercheurs ont minimisé le nombre de caractéristiques extraits des images et ont implémenté des classifieurs plus performants. En effet, il s'agit d'une tâche difficile

que de reconnaître la fumée sur les images en raison de la grande variabilité de la couleur, de la texture et des formes de la fumée. Il existe des méthodes de détection de la fumée qui ont été proposées, mais la plupart d'entre elles sont basées sur des caractéristiques classiques.

Toreyin et al [Tor06] propose une nouvelle méthode de détection de la fumée par vidéo en supposant que la caméra qui surveille la scène est fixe. La fumée est semi-transparente aux premiers stades d'un incendie. Par conséquent, les bords présents dans les images commencent à perdre de leur netteté et cela entraîne une diminution du contenu haute fréquence de l'image. L'arrière-plan de la scène est estimé et la diminution de l'énergie haute fréquence de la scène est surveillée en utilisant les transformées en ondelettes spatiales des images courantes et d'arrière-plan. Les bords de la scène produisent des extrema locaux dans le domaine des ondelettes. Une diminution du contenu énergétique de ces bords est un indicateur important de la présence de fumée dans le champ de vision de la caméra [Tor06]. En plus, la scène devient grisâtre lorsqu'il y a de la fumée et cela entraîne une diminution des valeurs de chrominance des pixels. Le comportement périodique des bords de la fumée est également analysé à l'aide d'un modèle de Markov caché (HMM) imitant le comportement temporel de la fumée. En outre, les limites des régions de fumée sont représentées dans le domaine des ondelettes et la nature haute fréquence des limites des régions de fumée est également utilisée comme indice pour modéliser le scintillement de la fumée. Tous ces indices sont combinés pour parvenir à une décision finale.

Par ailleurs, Xiong et al [Xio07] extraient en premier lieu les caractéristiques suivantes des images : le fond de l'image, la fréquence de scintillement, et les contours. Pour la vérification ultérieure, ils classent les zones de fumée en fonction des critères de séquence de périmètre et de zone de fumée. Des résultats de détection de fumée élaborés par Xiong et al [Xio07] sont affichée dans la Figure 1.2.



Figure 1.2. Résultats de Détection de fumée développée par Xiong et al [3]

Celik et al [Cel07] ont développé deux modèles: un pour la détection d'incendie et l'autre pour la détection de fumée. Un modèle basé sur les règles de la logique floue a été utilisé à la place des règles heuristiques existantes. Ce choix a rendu la classification plus robuste en distinguant efficacement la fumée 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 et al [Bor08] ont utilisé un vecteur multidimensionnel de caractéristiques comme entrée à un classificateur Bayes. Les caractéristiques utilisées sont : la rugosité, les limites des régions de feu, le moment statistique d'ordre 3 des régions de feu qui définit l'asymétrie, la variance et enfin la quantité de feu d'un cadre à l'autre (qui varie en raison de la flamme). D'autres travaux plus récentes ont été aussi publiés dans ce domaine [Gir13], [Gub09].

Comme dit précédemment, toutes ces méthodes sont basées sur des règles et exigent l'extraction de quelques caractéristiques discriminantes pour détecter le feu. Une autre alternative efficace serait l'utilisation des algorithmes d'apprentissage profond [Fri16], [fri17], [Huy18].

Calderera et al dans [Cal08] ont présenté un système de détection de fumée qui utilise un capteur commun de caméra CCD pour détecter la fumée dans les images et déclencher des alarmes. Tout d'abord, un modèle d'arrière-plan approprié a été proposé pour extraire les zones de fumée de manière fiable et éviter la sur-segmentation et les faux positifs, surtout dans les vidéos où de nombreux distracteur sont présent, comme des arbres en mouvement

à cause du vent ou des réflexes lumineux. Dans ce travail [Cal08], une approche bayésienne a été adoptée pour détecter les régions de fumée dans la scène traitée en analysant l'énergie de l'image au moyen des coefficients de transformation en ondelettes et de l'information sur la couleur. Un modèle statistique de l'énergie de l'image a été construit, en utilisant un mélange gaussien temporel, pour analyser la décroissance de l'énergie qui se produit généralement lorsque de la fumée couvre la scène. Ensuite, la détection a été renforcée en évaluant la différence des couleurs entre des fumées de référence et l'image testée. Le système proposé est capable de détecter les événements de fumée, de jour comme de nuit, avec un nombre réduit de fausses alarmes. Il est particulièrement adapté à la surveillance de grands scénarios extérieurs dans lesquels les capteurs courants seraient défaillants. Une vaste campagne expérimentale, à la fois sur des vidéos enregistrées et aussi en direct, a évalué l'efficacité de ce système de détection dans de nombreux scénarios, tels que les forêts.

La méthodologie proposée par Calderera [Cal08] est résumée dans la Figure 1.3.

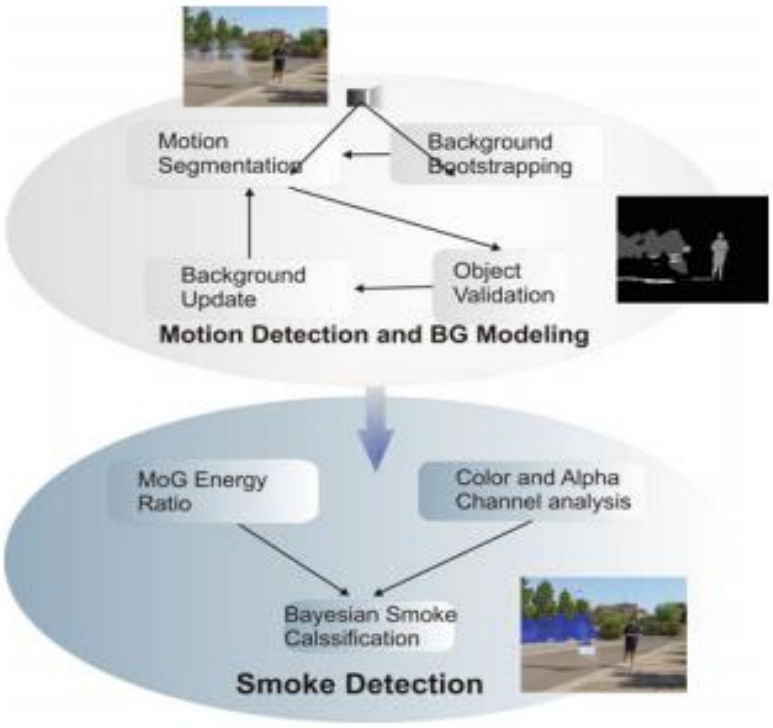


Figure 1.3. Méthode de détection de fumée proposée par Calderera [Cal08]

Malgré que la fumée est un indicateur important de présence d'incendies de forêt, cependant, de nombreux facteurs de fausses alarmes tels qu'un brouillard épais et des objets en

mouvement ressemblant à de la fumée dégradent considérablement la précision de la reconnaissance de la fumée.

Le travail de Zhao et al [Zha15] a présenté une nouvelle méthode de détection de la fumée dans les vidéos des feux de forêt à l'aide de l'algorithme CS Adaboost. Tout d'abord, les régions de mouvement ont été extraites de deux images adjacentes par un modèle d'arrière-plan approprié qui peut éviter les faux positifs de certains distracteurs statiques, tels que le ciel bleu et les feuilles grises.

Ensuite, un algorithme CS Adaboost est utilisé pour reconnaître les régions de fumée en utilisant le mouvement des centroïdes à l'aide du calcul de la vibration de la fumée, de l'énergie de l'image sur la base des coefficients de la transformation en ondelettes et des informations de couleur de la fumée entre une couleur de fumée de référence et la trame d'entrée.

Les résultats expérimentaux de ce travail ont montré que la méthode proposée peut non seulement détecter des images de fumée, mais aussi distinguer le brouillard dense de la fumée comme affiché dans la Figure 1.4. Une comparaison des valeurs de la précision et du rappel entre les différentes méthodes et la méthode de Zhao [Zha15] est affichée dans la Figure 1.5.

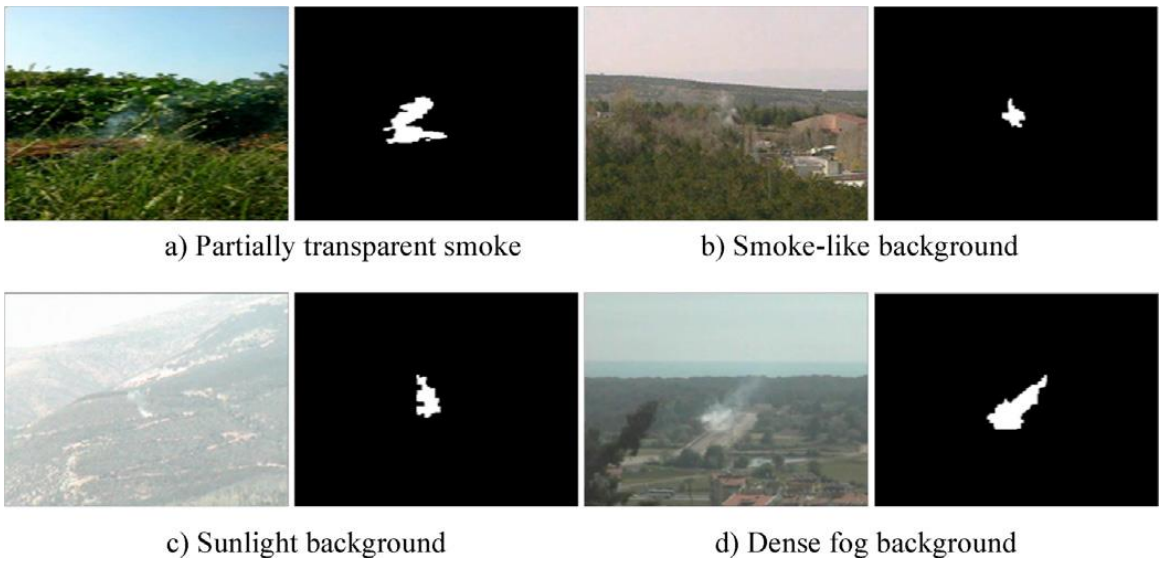


Figure 1.4 .Extraction du mouvement de la fumée proposée par Zhao [Zha15]

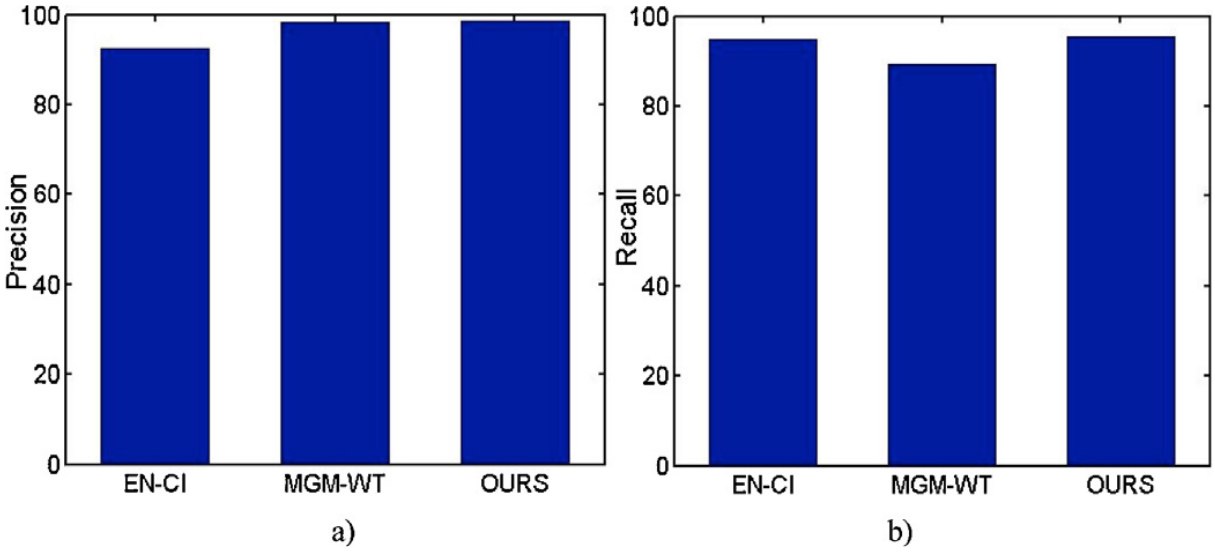


Figure 1.5. Comparaison de la précision et Rappel entre la méthode de Zhao [Zha15] et d'autres méthodes

Une autre approche de détection de la fumée a été proposée par [Yua16]. Elle est basée sur la technologie du traitement de l'image dans un système de surveillance intelligent pour l'alerte précoce des incendies. Premièrement, pour améliorer la qualité de l'image, la logique floue a été utilisée pour l'amélioration de l'image. Ensuite, la zone de fumée est extraite par le modèle mixte gaussien. En fonction des caractéristiques de la fumée, des caractéristiques statiques et dynamiques sont extraites. Troisièmement, ils fixent les hyper paramètres de la méthode de classification SVM. Enfin, les vecteurs de caractéristiques sont insérés dans le SVM pour la reconnaissance de la fumée. Des expériences sont effectuées par plusieurs vidéos dans des conditions variées et il est démontré que l'algorithme proposé est réalisable et efficace. Il présente également une forte robustesse et un temps réel élevé.

La présentation de l'algorithme de la méthode proposée par Wang [Yua16] est résumée dans la Figure 1.6. et quelques résultats de simulation sur l'extraction de mouvement et les histogrammes des images contenant la fumée sont présentées dans la Figure 1.7 et la Figure 1.8.

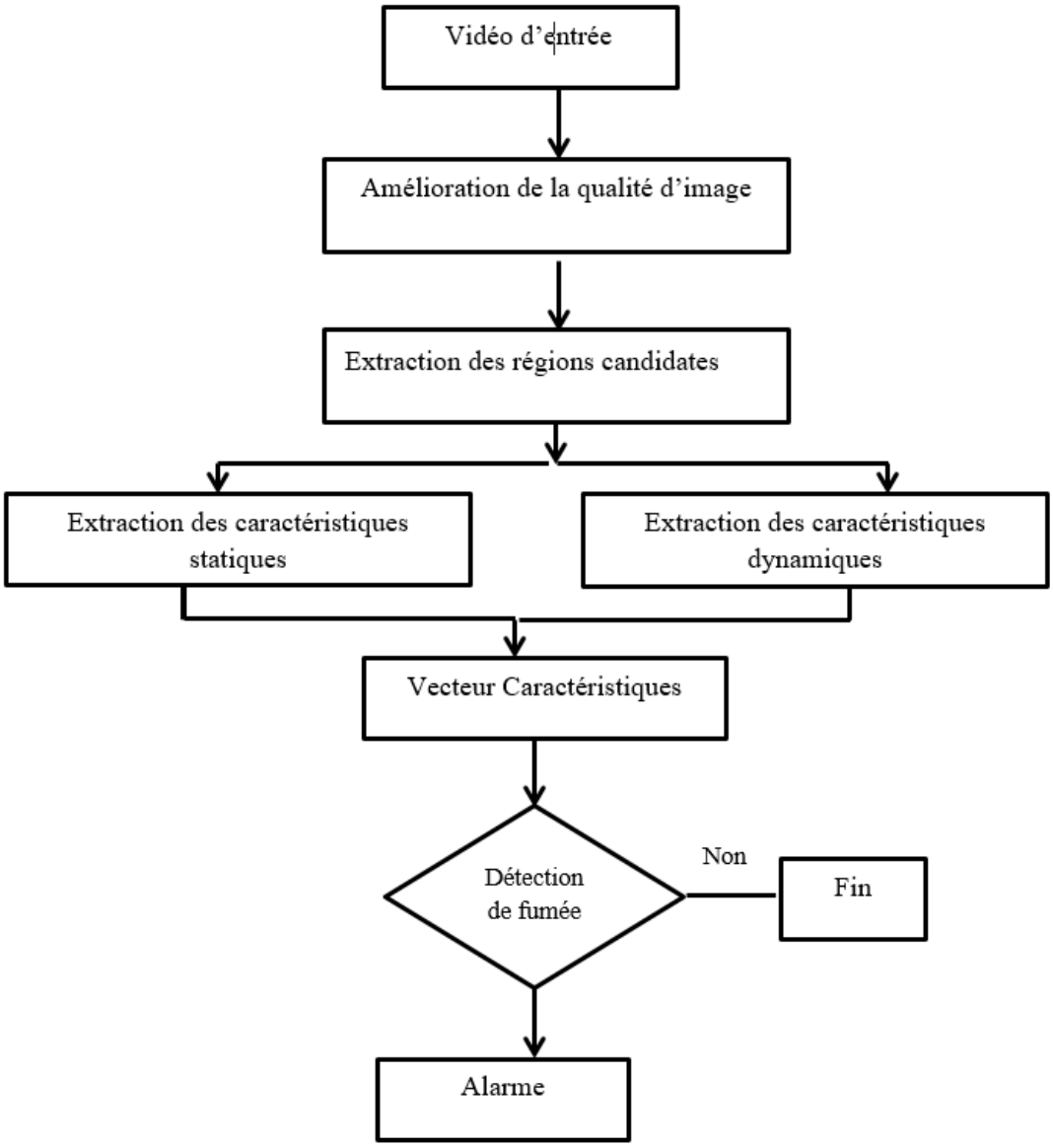


Figure 1.6. Algorithme proposé pour la détection de fumée par Wang [Yua16]

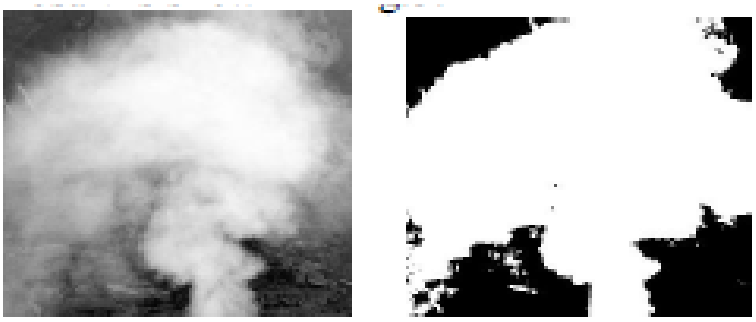


Figure 1.7. Extraction de mouvement de la fumée par Wang [Yua16]

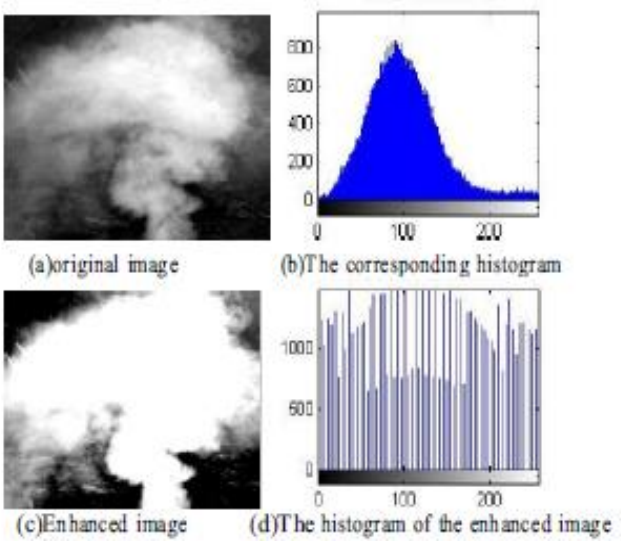


Figure 1.8. Histogramme des images contenant la fumée avant et après seuillage [Yua16]

Les scènes réelles peuvent inclure des objets similaires à la fumée avec un comportement dynamique dû à des caméras de faible résolution, au flou ou aux conditions météorologiques. Par conséquent, la vérification de la détection de la fumée est une étape nécessaire dans les systèmes de surveillance. La vérification confirme les véritables zones de fumée, lorsque les zones similaires à la fumée sont déjà présentes dans une séquence vidéo.

Les contributions réalisées par Favaroskaya et al [Fav16] dans la détection de la fumée sont multiples. Premièrement, de nombreux types de motifs binaires locaux (LBP) avec des généralisations en 2D et 3D ont été étudiés lors des expériences en tenant compte de l'évolution des caractéristiques de la fumée pendant la propagation du feu. Deuxièmement, une carte des différences de luminosité, une carte des bords et une carte Laplaciaire ont été étudiées dans la spécification des LBP spatio-temporelles (STLBP) [Fav16].

Les descripteurs sont basés sur des histogrammes, et une classification en trois classes : fumée dense, fumée transparente et non-fumée a été mise en œuvre en utilisant la divergence de Kullback-Leibler. Les résultats de bonne reconnaissance sont entre 84% et 96% de précision pour la fumée dense, en fonction des différents types de LPB utilisés.

Dans les travaux de Lee et al [Lee07], ils proposent une technique de traitement d'image pour la détection automatique en temps réel des incendies et des fumées dans l'environnement des tunnels. Afin d'éviter les dommages importants causés par le feu dans

le tunnel, ils ont utilisé un système permettant de minimiser et de découvrir l'incident le plus rapidement possible. Cependant, il est impossible de maintenir l'observation humaine de la télévision en circuit fermé dans le tunnel pendant 24 heures.

Yin et al [Yin17] ont proposé un DNCNN : Deep Normalization and Convolutional Neural Network, c.a.d. un réseau de neurones avec normalisation profonde et convolutive, contenant 14 couches pour mettre en œuvre l'extraction et la classification automatique des caractéristiques et améliorer les performances de la détection de fumée. Dans le DNCNN, les couches convolutives traditionnelles sont remplacées par des couches de normalisation et convolutives afin d'accélérer le processus d'apprentissage et d'accroître les performances de la détection de la fumée. Pour réduire l'éventuel sur-apprentissage causé par des échantillons d'apprentissage déséquilibrés et insuffisants, ils génèrent davantage d'échantillons d'apprentissage à partir des ensembles de données d'apprentissage originaux en utilisant diverses techniques d'enrichissement des données. Les résultats expérimentaux ont montré que cette méthode a permis d'obtenir des taux de fausses alarmes très faibles, inférieurs à 0,6 %, avec des taux de bonne détection de à 96,37 % [Yin17].

Cependant, il reste difficile d'extraire des caractéristiques stables pour la reconnaissance de la fumée, en grande partie à cause de ses variations de couleur, de forme et de texture. Les réseaux neuronaux convolutifs classiques peuvent apprendre automatiquement des représentations de caractéristiques à partir d'une seule image, mais ne parviennent pas à capturer des informations de mouvement entre les images.

Pour résoudre ce problème, Hu et al ont proposé dans leur article [Huy18] un réseau neuronal convolutif spatio-temporel pour la détection de la fumée dans les vidéos et ils proposent une architecture qui utilise une stratégie d'apprentissage multitâche pour reconnaître conjointement la fumée et estimer le flux optique, en capturant simultanément les caractéristiques d'apparence intra-image et les caractéristiques de mouvement entre les images. L'efficacité et l'efficience de la méthode ont été validées par des expériences menées sur l'ensemble de données, qui atteint un taux de détection de 97% et un taux de fausses alarmes de 3,5% avec un temps de traitement de 5 ms par image, ce qui est évidemment supérieur aux méthodes existantes. Le souci de cette technique est le problème de ressources.

D'autre part, Pundir et al. [Pun17], [Pun19] ont proposé une méthode qui vise à extraire les caractéristiques éminentes de la fumée en utilisant l'apprentissage profond et ils ont essayé de classer les régions de fumée et de non fumée en utilisant différents classificateurs.

L'architecture qu'ils ont proposée est basée sur les réseaux neuronaux convolutifs profonds double, qui a prouvé leur supériorité dans les applications de reconnaissance d'objets. Le premier modèle d'apprentissage est utilisé pour extraire des caractéristiques d'image des zones de fumée, qui sont extraits en utilisant un algorithme de segmentation en super-pixel. Ces caractéristiques sont : la couleur, la texture de la fumée et le désordre du périmètre. Le deuxième modèle d'apprentissage est utilisé pour extraire des caractéristiques telles que le mouvement de la fumée. La méthode du flux optique a été utilisée, afin de capturer le mouvement aléatoire de la fumée. Ces flux optiques extraits sont ensuite introduits dans le « Deep CNN ». Les caractéristiques extraites sont combinées par la suite pour classifier les régions en utilisant la machine à support de vecteurs SVM. Après les tests effectués sur les différents types de vidéos (avec fumée, sans fumée, avec brouillard et nuages), la méthode a donné une précision sur la fumée proche et la fumée lointaine est de 98,29% et 91,96% respectivement. La contribution de ce travail réside dans le fait que les auteurs ont traité 20000 images basées sur des données de fumée et de non-fumée et que méthode surpasse plusieurs autres méthodes existantes dans des conditions d'imagerie difficiles (brouillard, nuages, …). La procédure d'apprentissage est résumée dans la Figure 1.9.

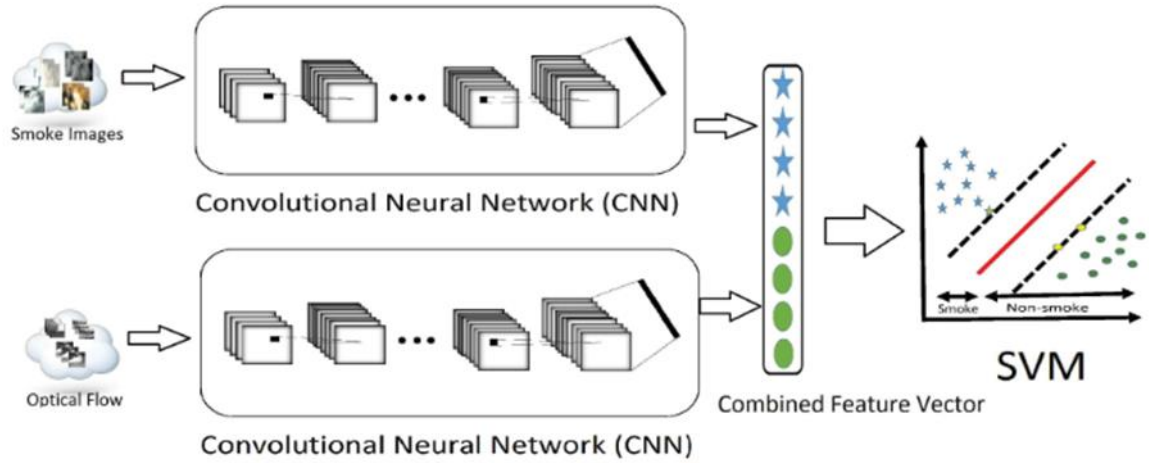


Figure 1.9. Processus de détection de fumée proposée par Pundir et al [Pun19]

III. Présentation des caractéristiques de la fumée et des systèmes de classification utilisés :

On présente dans cette section les différentes caractéristiques qu'on peut extraire des vidéos de fumée et aussi plusieurs systèmes de classification utilisés dans ce domaine. On commence par donner dans la Figure 1.10 un schéma du principe général des méthodes classiques de caractérisation, détection et classification de la fumée.

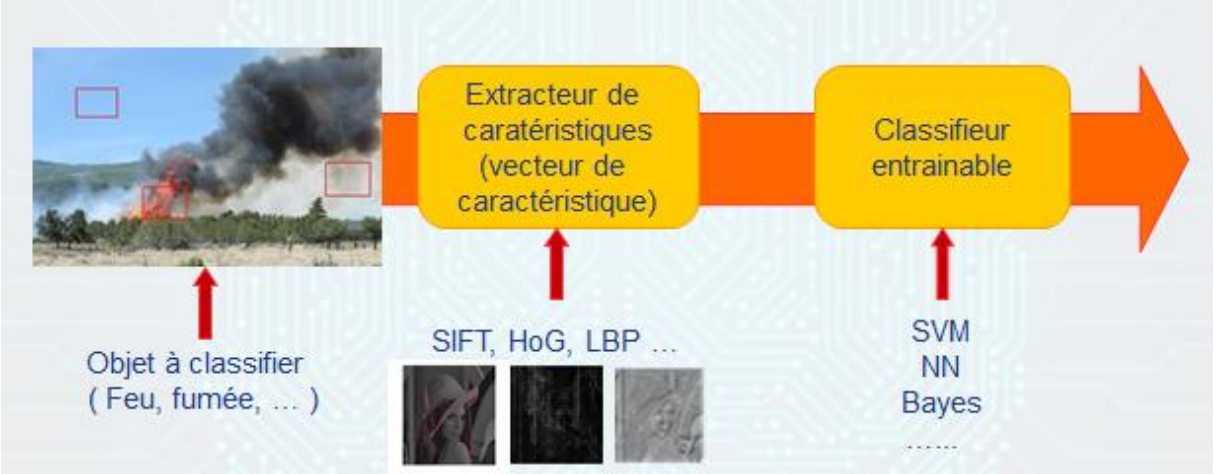


Figure 1.10. Schéma Récapitulatif de la détection de fumée

1. Extraction de caractéristiques de la fumée dans les vidéos :

1.1. La caractéristique : Mouvement

L'étiquetage des mouvements causés par la fumée est une étape cruciale pour extraire les régions d'intérêt et pour connaître la nature du mouvement (ordinaire ou chaotique). La détection de mouvement peut s'effectuer selon trois méthodes : soustraction du fond, différenciation temporelle et flux optique. Le but est de déterminer la position de la fumée en mouvement dans la séquence vidéo à chaque instant. Ce processus de localisation est basé sur la représentation de l'objet cible en utilisant un ensemble de caractéristiques discriminantes telles que la couleur, le mouvement, etc. [Vij14] [Fel10]. Ces méthodes sont basées sur le fait que s'il y a du mouvement dans une image, la valeur de l'intensité d'un pixel sera différente par rapport à celle correspondante dans la trame suivante ou la trame de référence. La technique la plus utilisée pour le traitement de la soustraction d'arrière-plan

est le modèle par un mélange de gaussiennes [Yan14], [Mck00], [Sta00]. Dans cette approche, la caméra est immobile.

Le modèle de mélange de gaussiennes (GMM) est une méthode statistique utilisée pour estimer une distribution aléatoire en utilisant une somme pondérée des densités de composantes gaussiennes (Voir l'exemple de la Figure 1.11).



Figure 1.11. Exemple de mélange gaussien (les trois composantes sont en couleur, et en noir le modèle gaussien)

Cette méthode est largement utilisée pour la détection de mouvement. Elle est utilisée pour modéliser la variation d'intensité des pixels dans les séquences vidéo. Ainsi, on peut identifier les pixels appartenant à des objets en mouvement qui n'appartiennent pas aux modèles de pixels d'arrière-plan déjà calculés. La modélisation d'une distribution aléatoire en utilisant un mélange de *K* distributions gaussiennes équivaut à trouver les poids w_i, les moyennes μ_i et les variance σ_i *pour* chaque gaussienne d'indice *i* pour chaque pixel x_t soit modéliser avec :

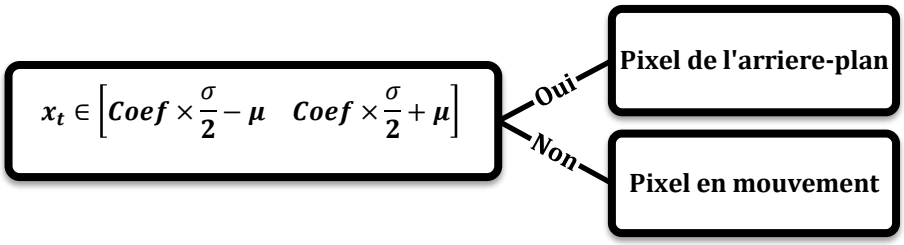
$$P(x_t) = \sum_{i=1}^{K} w_i \times \mathcal{N}(x_t|\mu_{i,t}, \sigma_{i,t}) \quad (1.1)$$

Où $\mathcal{N}$ est la fonction densité de probabilité gaussienne, donnée par:

$$N(x_t\ /\mu_i, \sigma_i^2) = \frac{e^{-1/2\sigma^2\ (x_t\ -\mu)^T(x_t\ -\mu)}}{(2\pi)^{1/2}\sigma} \quad (1.2)$$

Face aux changements de luminosité au fil du temps, les paramètres du modèle d'un pixel doivent être mis à jour. Le processus est divisé en divisé le processus en deux étapes ; la

classification des nouvelles valeurs de pixels et la mise à jour du modèle correspondant en se basant sur les résultats de l'étape précédente. Ainsi, La première étape détermine pour chaque pixel si son intensité suit ou non le modèle du pixel actuel, c'est-à-dire si le pixel appartient ou non aux pixels d'arrière-plan. Dans [Mck00], [Sta00], les auteurs proposent de vérifier chaque nouveau pixel avec les K distributions existantes du modèle. Si sa valeur d'intensité est appartienne à l'intervalle de longueur égale à un coefficient multiplié par l'écart type de la distribution correspondante pour chaque composante RVB, le pixel est considéré comme étant un pixel de l'arrière-plan et appartenant à la distribution comme il est indiqué ci-dessous :



(1.3)

Un exemple de soustraction de fond est représenté dans la Figure 1.12.

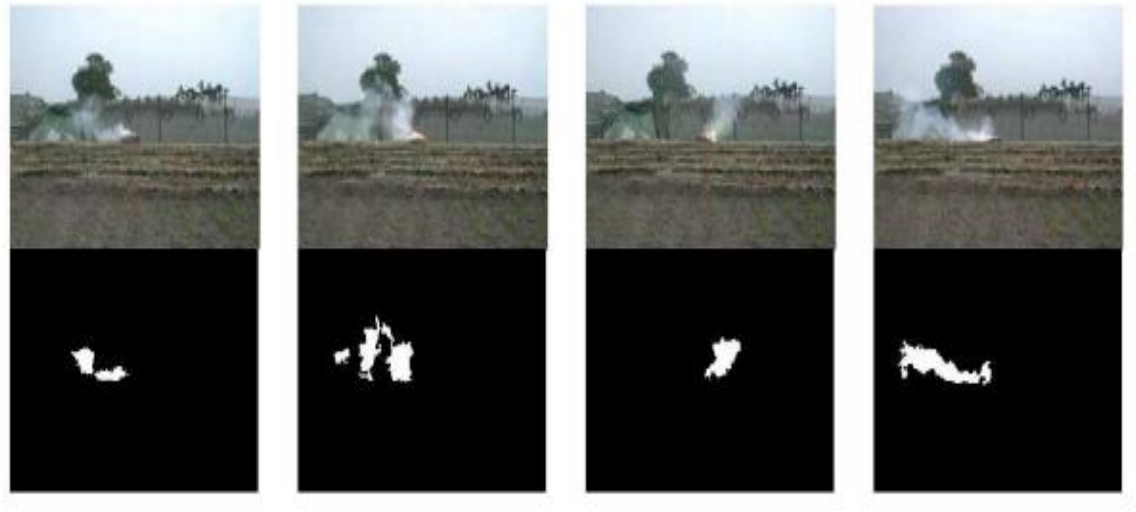


Figure 1.12 : Application de la soustraction du fond sur quelques trames

1.2. La caractéristique : Energie

Au fur et au mesure que la fumée adoucit les bords d'une image, le rapport de l'énergie de la trame actuelle divisée par l'énergie du fond diminue. Ceci est dû aux caractéristiques de la transparence de la fumée comme indiqué par Toreyin et al. [Tor06]. Pour calculer la diminution d'énergie du bloc sur les bords de l'image, plusieurs chercheurs [Cal08], [Gub09] ont appliqué la transformée en ondelettes discrète au trame actuel et à l'image de l'arrière-plan avant l'apparition de la fumée et après l'apparition de la fumée. Cette analyse conduit à l'analyse du comportement temporel de la fumée. Le principe de la décomposition en ondelettes discrètes (ou la transformée en ondelettes discrète) est résumé sur la figure 1.13 lors du passage d'une résolution à une autre, en montrant les 4 images résultantes obtenues :

HH, (High High, où apparaissent les détails diagonaux de l'image),

LH (Low High, où apparaissent les détails horizontaux de l'image),

HL (High Low, où apparaissent les détails verticaux de l'image),

LL (Low Low ou apparait une approximation de l'image).

Ceci est réalisé en convolant l'image d'intensité par des filtres appropriés. Un exemple de résultat de cette décomposition est donné sur la figure 1.14.

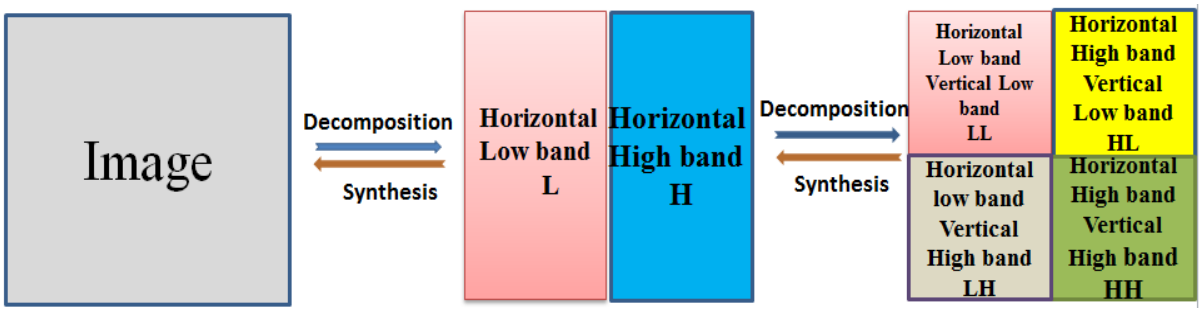


Figure 1.13. Principe du passage d'un niveau de multi-résolution à un autre dans la transformée en ondelette discrète des images

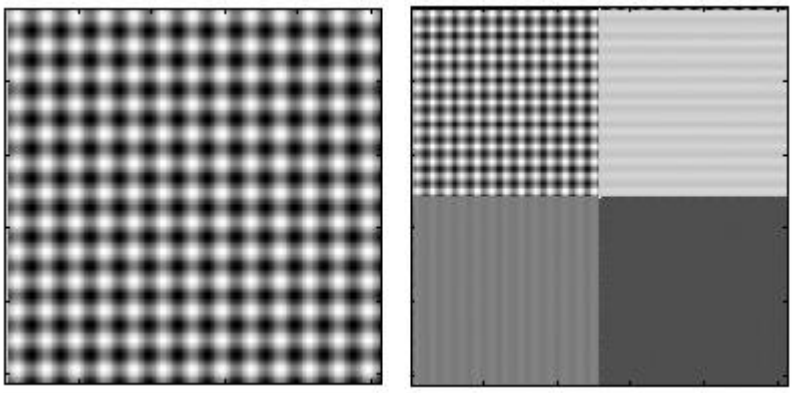


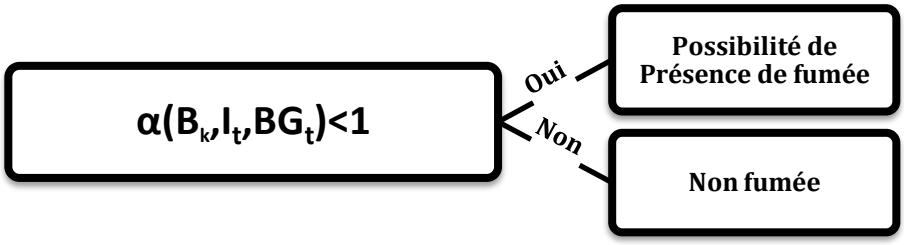
Figure 1.14. Exemple de l'application de la transformée en ondelettes sur une image, les détails verticaux apparaissent sur la composante HL, les détails horizontaux apparaissent sur la composante LH

L'idée proposée dans [Tor06] est que l'image I_t de la fumée à l'instant *t* est divisée en blocs de taille arbitraire noté B_k alors que l'image de l'arrière-plan est noté BG_t, ensuite pour chaque niveau de décomposition, l'énergie des blocs contenant la fumée est calculée avec :

$$E(B_k, I_t) = \sum_{m,n \in B_k} LH^2(m,n) + HL^2(m,n) + HH^2(m,n) \tag{1.4}$$

$$\alpha(B_k, I_t, BG_t) = \frac{E(B_k, I_t)}{E(B_k, BG_t)} \tag{1.5}$$

Où $E(B_k, BG_t)$ représente l'énergie de l'image de référence (ou l'arrière-plan) avant l'apparition de la fumée.



(1.6)

1.3. La caractéristique : Couleur

La détection des couleurs est l'un des critères les plus importants et les plus utilisés pour la détection de la fumée. La couleur de fumée peut être gris, blanc, noir et gris foncé. Cette caractéristique peut être traitée dans différents espaces (RVB, YUV, YCbCr, etc.).

Dans les zones de fumée, les pixels dans l'espace RVB sont caractérisés par des composantes de valeurs proches r, v, et b :

$$\begin{cases} r = \frac{R}{R+V+B}, g = \frac{G}{R+V+B}, b = \frac{B}{R+V+B} \\ r + v + b = 1 \end{cases} \quad (1.7)$$

$$\begin{aligned} r(I_t, b_k) &\simeq r(I_{t+\Delta t}, b_k) \\ v(I_t, b_k) &\simeq v(I_{t+\Delta t}, b_k) \\ b(I_t, b_k) &\simeq b(I_{t+\Delta t}, b_k) \end{aligned} \quad (1.8)$$

Par contre, dans l'espace YUV, les valeurs de chrominance U et V diminuent dans les régions de fumée. De plus, la valeur de luminance *Y* des régions affectées doit être élevée par rapport aux seuils définis expérimentalement :

$$\begin{cases} \mathrm{Y} > \mathrm{S_y} \\ |\mathrm{U} - 128| < \mathrm{S_u} \qquad |\mathrm{V} - 128| < \mathrm{S_v} \end{cases} \quad (1.9)$$

$Y, U, V \in [0, 255]$, S_y, S_u et S_v sont les seuils, expérimentalement fixés .

Le but de l'extraction de caractéristiques de fumée est de limiter les régions d'intérêts et de les insérer dans un vecteur de caractéristiques pour classifier les régions de fumée et non fumée par la suite.

2. Classification des régions de fumée et non fumée

2.1. Méthodes SVM (Machine à vecteurs de support)

Les SVM [Chi02], [Gub09] (Support Vector Machine, séparateurs à vaste marge) sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de discrimination et de régression. Ils sont une généralisation des classifieurs linéaires et ont un potentiel considérable en tant que classificateurs. Les SVM ont des similarités avec les réseaux de neurones et peuvent fournir des approximations universelles pour toute fonction multivariée avec le degré de précision souhaité. Au lieu de minimiser une fonction objective, les SVM tentent de minimiser l'erreur de prédiction. Par ailleurs, les SVM sont de plus en plus utilisés en raison de leur grande capacité de généralisation dans un large éventail d'applications et de leurs performances améliorées d'apprentissage automatique traditionnels, tels que les réseaux à fonctions de base radiales et les réseaux de neurones de rétro propagation.

Le SVM a été utilisé par Lee et al dans [13] pour la classification de la fumée. Ils ont opté pour trois caractéristiques qui sont la détection de mouvement, la couleur et l'analyse de l'énergie par blocs pour les combiner dans un vecteur appelé vecteur de caractéristiques. Ce vecteur sera un vecteur d'entrée pour le classificateur SVM. Un exemple de résultats de la classification par SVM d'images de fumée est illustré dans la Figure 1.15.



Figure 1.15. Exemple de résultats de détection de fumée par la méthode proposée par Lee et al [13]

2.2. CS Adaboost

CS AdaBoost un algorithme d'apprentissage développé par Viola et al [Vio01] et amélioré par Mikolajczyk et al [Mik04] et Laptev et al [Lap04].

L'algorithme de CS Adaboost peut être associé à autre algorithme d'apprentissage pour améliorer les performances de classification [Hyu14]. Le principe est issu de la combinaison de classifieurs. Ce classificateur est utilisé dans la détection de fumée après l'extraction du mouvement comme caractéristique de la fumée. Cette technique de détection a été développée par Jensen et al [Jen09]. Un exemple de résultats de détection de fumée est illustré dans la Figure 1.16.



Figure 1.16. Application du GMM et CS Adaboost pour la détection de fumée [Huy14]

2.3. Approche bayésienne

Les réseaux bayésiens sont des outils puissants pour l'extraction, la classification des informations incertaines et complexes [Jen09]. Un réseau Bayésien est composée d'une composante graphique qui permet de définir les variables du problème et leurs relations et d'une composante probabiliste qui permet de quantifier l'incertitude relative aux relations de dépendance entre les variables. Ils sont utilisés comme outil de classification.

Une approche intéressante qui utilise cette technique de classification est celle de Toreyin et al [Tor06] pour détecter les régions de fumée. Cette méthode extrait tout d'abord les caractéristiques suivantes : Mouvement, vibration, énergie et texture. Quelques exemples de résultats de détection de fumée sont présentés sur Figure 1.17.



Figure 1.17. Des résultats de détection de fumée par la méthode de Toreyin et al [Tor06]

2.4. Algorithme KNN

L'algorithme KNN [Gub09] figure parmi les plus simples algorithmes d'apprentissage artificiel. Dans un contexte de classification d'une nouvelle observation *x*, l'idée fondatrice simple est de déterminer les plus proches voisins de cette observation. La classe de *x* est déterminée en fonction de la classe majoritaire parmi les *K* plus proches voisins de l'observation *x*. La méthode KNN est donc une méthode non-paramétrique à base de voisinage. Ceci signifie que cet algorithme permet de faire une classification sans faire d'hypothèse sur la fonction qui relie la variable dépendante aux variables indépendantes.
Dans notre application, cet algorithme de classification a été implémenté par Gubbi et al [Gub09] pour la classification des régions de fumée et non fumée. Cette étape est précédée par une extraction de caractéristique de l'énergie. Un exemple de résultats de détection de fumée est illustré dans la figure Figure 1.18.

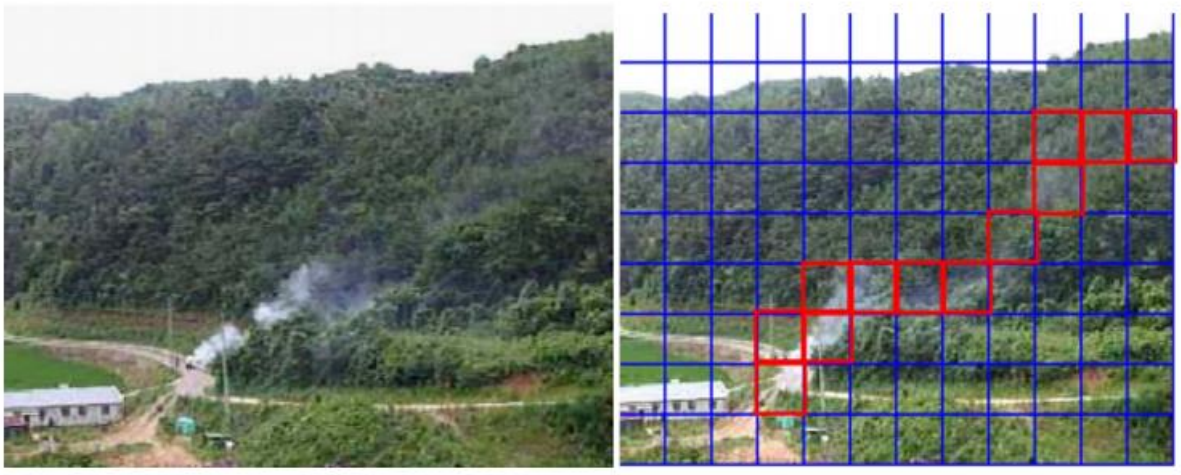


Figure 1.18. Un exemple de résultats de détection de fumée par la méthode de Gubbi et al [Gub09]

2.5. Réseaux de neurones convolutifs CNN

Les réseaux neuronaux convolutifs [Ter09], [Yin17] sont des réseaux neuronaux artificiels profonds utilisés principalement pour classer les zonez dans les images, les regrouper par similarité et effectuer la reconnaissance d'objet au sein de scènes. Ce sont des algorithmes permettant d'identifier des visages, des individus, des panneaux de signalisation, des tumeurs, des ornithorynques et de nombreux autres aspects des données visuelles.

L'efficacité des réseaux de neurones convolutifs (ou à convolution, ConvNets ou CNN) dans la reconnaissance d'images est l'une des raisons principales pour lesquelles plusieurs chercheurs se sont rendu compte de l'efficacité de l'apprentissage profond. Ils alimentent des avancées majeures dans la vision par ordinateur qui ont des applications évidentes pour les voitures autonomes, la robotique, les drones, la sécurité, les diagnostics médicaux et les traitements pour les malvoyants.

Les réseaux de neurones convolutifs ingèrent et traitent les images en tant que tenseurs, et les tenseurs sont des matrices de nombres de dimensions supplémentaires. En effet, mathématiquement, une convolution est l'intégrale mesurant la surface de chevauchement de deux fonctions lorsqu'elles passent l'une sur l'autre. Par ailleurs, elles constituent le fondement des opérations d'algèbre linéaire utilisées pour traiter les images.

Un réseau de convolution manipule des zones carrées de pixels et les fait passer à travers un filtre. Ce filtre est également une matrice carrée plus petite que l'image elle-même et de la même taille que le patch. Le filtre a pour tâche de rechercher des motifs dans les pixels.

La couche suivante dans un réseau de convolution a trois noms: pooling maximum, sous-échantillonnage et sous-échantillonnage. Les cartes d'activation sont introduites dans une couche de sous-échantillonnage et, comme pour les convolutions, cette méthode est appliquée patch par patch. Dans ce cas, la mise en pool maximale prend simplement la valeur la plus élevée d'un patch d'une image, et ensuite la replace dans une nouvelle matrice à côté des valeurs maximales des autres patchs et supprime le reste des informations contenues dans les cartes d'activation.

La Figure 1.19 montre la séquence de transformations impliquées dans un réseau de convolution typique.

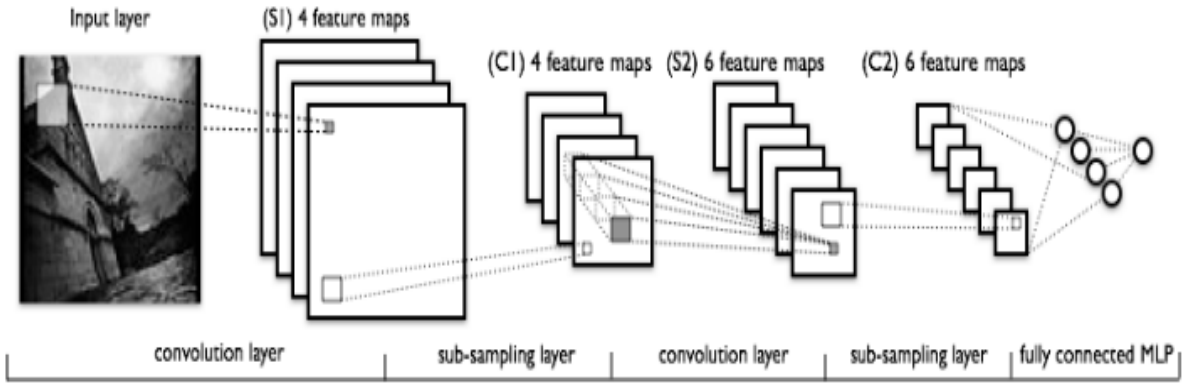


Figure 1.19. Architecture des réseaux de neurones convolutifs

Cette méthode a été récemment appliquée sur la détection de la fumée : Frizzi et al [Fri16], et Hu et al [Huy18]. Ces travaux ont abouti à des résultats intéressants en générant des valeurs élevées de taux de bonne classification.

Cette technique a permis aux chercheurs de détecter et localiser la fumée dans les frames comme affiché dans la Figure II.20.

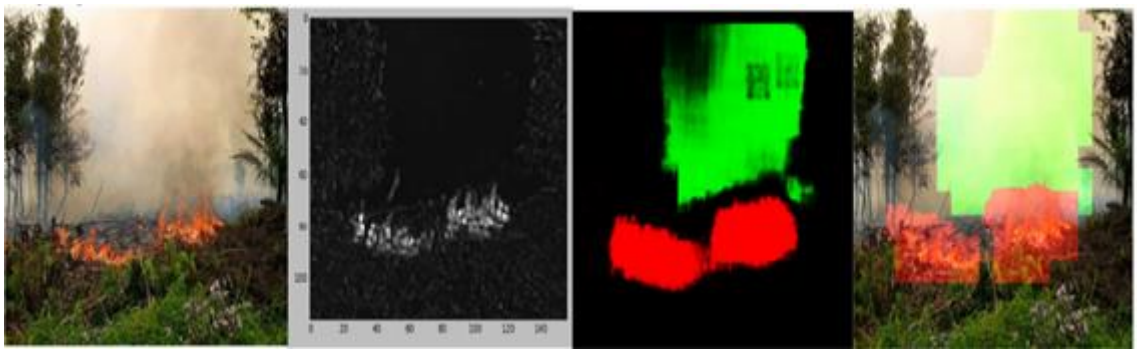


Figure 1.20. Quelques résultats de détection de fumée par la méthode de Frizzi et al[Fri16]

L'utilisation des réseaux convolutifs dans la détection de la fumée a été améliorée par Hu et al [Huy18] par la combinaison et la normalisation des couches de ces réseaux. Pour améliorer les performances de détection de fumée, ils proposent une nouvelle profondeur de réseau de neurones de normalisation et de convolution (DNCNN) en utilisant un total de 14 couches pour mettre en œuvre la fonctionnalité automatique d'extraction et de classification.

Dans le DNCNN, les couches convolutives traditionnelles sont remplacées par la normalisation et des couches convolutives pour accélérer le processus de formation et améliorer les performances de la détection de fumée. Le taux de détection de fumée atteint par cette méthode est égal à 96.37. L'architecture du réseau est présentée par la Figure 1.21.

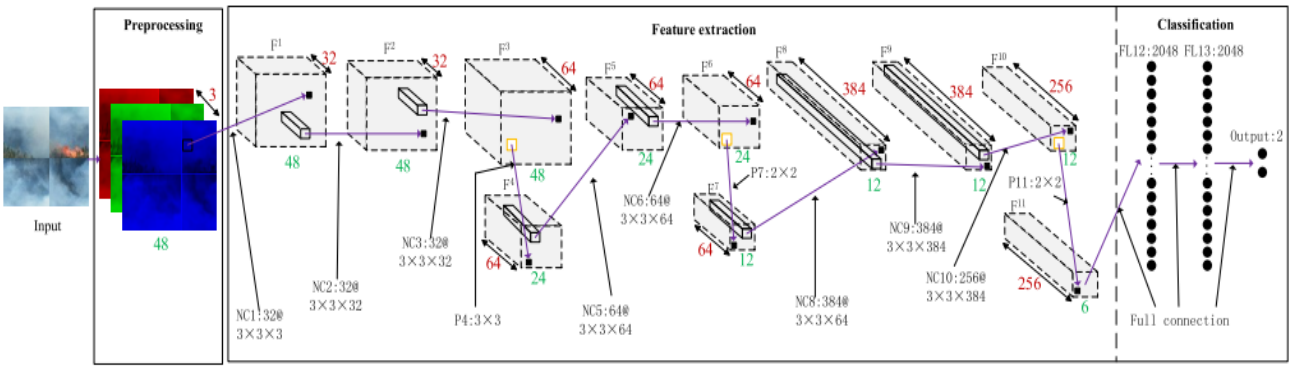


Figure 1.21. Architecture du réseau utilisé dans [Huy18]

Les méthodes citées précédemment destinées pour la détection des régions de fumée et non fumée suivent le même processus pour la détection : Extraction et Classification. Toutes ces méthodes sont récapitulées dans le Tableau 1.1.

Tableau 1. 1. Tableau Récapitulatif des méthodes et techniques utilisées pour la détection de la fumée

Paper	Couleur	Mouvement	Energie	Désordre	Vibration	Classification (SVM, KNN, HMM…)
Toreyin [Tor06] (2006)	YUV	×	×	×		×
Xiong [Xio07] (2007)		×	×	×		
Calderara[Cal08],2008	RGB	×	×		×	×
Zhao [Zha15] (2015)	RGB	×	×		×	×
Wang Yuanbin [Yua16] (2016)		×				×
Yin [Yin17] (2017)	RGB	×				DNCNN
Hu [Hu18] (2018)	RGB	×				CNN
Pundir [Pun19] (2019)	RGB	×				Deep CNN
Méthode Proposée	RGB	×	×			DBN

IV. Quelques méthodes de détection de feux et des flammes

Les recherches traitant la détection de feu dans la littérature augmente de façon exponentielle. Plusieurs chercheurs ont développé des algorithmes de détection de feu par traitement vidéo. Verstockt et al [Ver12] ont 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 basée sur l'histogramme. Un ensemble de caractéristiques est calculés à partir des images de flamme 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 dans LWIR.

La Figure 1.22 présente quelques résultats de l'application de la méthode de Verstockt et al [Ver12] sur des vidéos tests.

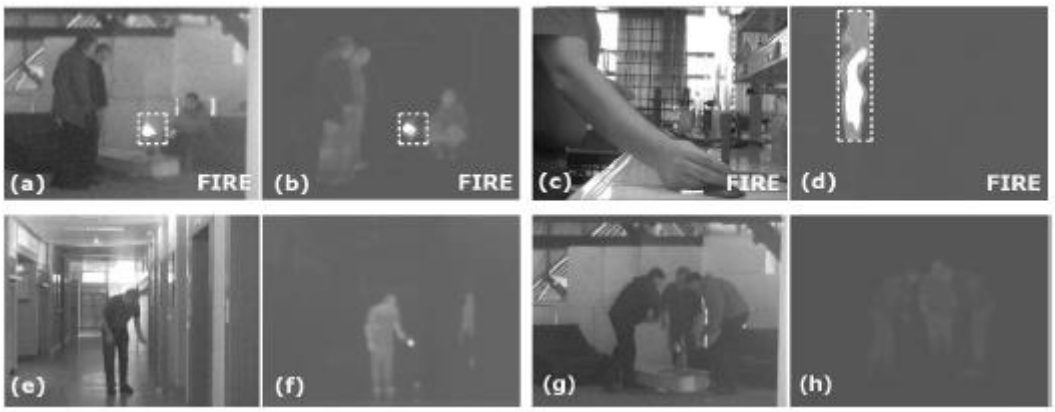


Figure 1.22. Résultats de détection de feux réalisés par Verstockt et al [Ver12]

D'autre part, Toreyin et al [Tor06] ont implémenté un algorithme de détection vidéo basé sur 4 principes. Ils ont estimé premièrement les pixels et les régions en mouvement en utilisant une méthode de fond hybride. Au fait, il s'agit d'une opération de différenciation à trois images effectuée pour extraire les régions de mouvement suivies d'une soustraction de fond adaptative pour extraire toute la région en mouvement. En plus, ils ont utilisé un modèle de mélange de gaussiennes dans l'espace de couleur RVB pour détecter les pixels de couleur de feu. Une transformation en ondelettes temporelles a été appliquée ensuite pour analyser le scintillement de la flamme. 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.

Celik et al [Cel07] a implémenté deux modèles : Le premier a été destiné pour la détection de flamme et le deuxième pour la détection de fumée. Un modèle de logique floue basé sur des règles a été exploité à la place des règles heuristiques existantes. Ce qui a rendu la classification plus consistante en séparant efficacement le feu et le feu comme des objets colorés. Pour la détection des fumées, une analyse statistique a été faite à partir de l'idée que

la fumée présente une couleur grisâtre avec un éclairage différent. Quelques résultats de l'application de la méthode proposée par Celik et al sur une vidéo est illustrée dans la Figure 1.23.



Figure 1.23. Résultats de détection de feux obtenus par Celik et al [Cel07]

On note enfin que Borges et al [Bor08] ont utilisé un vecteur de caractéristiques multidimensionnel comme entrée à un classificateur Bayes. Les caractéristiques de la flamme sont comme suivies :

-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 des régions de feu potentiel
- La quantité de feu d'un cadre à l'autre (varie en raison de la flamme).

Toutes ces méthodes sont fondées sur des règles mathématiques ou exigent la construction de caractéristiques discriminantes pour détecter le feu. Une autre alternative est l'utilisation des algorithmes d'apprentissage profond, qui sera traitée en détail dans le chapitre suivant.

IV. Conclusion

Dans ce chapitre, nous avons exposé des différentes méthodes de détection de la fumée, des flammes et de feu. On a aussi présenté les méthodes classiques de détection de fumée qui reposent essentiellement sur trois axes qui sont : l'extraction, la classification et la décision. Les techniques d'extraction varient entre les méthodes classiques aux méthodes à convolution. Parfois ils reposent essentiellement sur le choix des attributs et la deuxième est parfois sur des techniques d'apprentissage profond tels que le CNN, DBN. Puis, la classification des régions de fumées et non fumées par des classifieurs tels que le SVM, le KNN, …

Dans la section suivante, nous présentons en détail une autre alternative neuronale modifiée basée sur les systèmes à apprentissage automatique et spécialement l'apprentissage profond : le DBN, Deep Belief Netowrk, qui sera implémentée pour la détection de la fumée.

Chapitre II : Techniques utilisées pour la classification neuronale intelligente et l'extraction des zones d'intérêt

Chapitre II : Techniques utilisées pour la classification neuronale intelligente et l'extraction des zones d'intérêt

I. Introduction

Comme indiqué dans le chapitre précédent, il y a trois étapes principales dans le processus de détection de la fumée dans les image : le prétraitement d'image, l'extraction de caractéristiques et la détection de fumée. Parmi ces étapes, l'extraction de caractéristiques est la partie centrale des algorithmes. Les méthodes traditionnelles dépendent de la sélection manuelle des caractéristiques de la fumée suivie d'une classification avec un processus de classification à apprentissage. Dans ce chapitre, nous présentons un aperçu des techniques de classification basée sur l'apprentissage automatique et les réseaux de neurones classiques suivies des techniques d'apprentissage profond. Nous présentons ensuite les techniques d'extraction de régions d'intérêts que nous allons effectuer par la suite. La méthodologie de détection de fumée en utilisant le DBN comme outil de classification sera aussi détaillée.

II. Rappel sur les systèmes d'apprentissage automatique

1. Systèmes à apprentissage automatiques et réseaux de neurones

Les systèmes à apprentissage automatique (ou apprentissage machine, machine learning) offrent aux ordinateurs la possibilité d'apprendre à partir de données reçues et découvertes. Ils peuvent être appliqué dans des situations où il est très difficile, voire impossible, de définir des règles heuristiques et manuelles, telles que la détection de visage, la reconnaissance vocale, la prédiction de mouvement. Une grande famille des systèmes à apprentissage automatique est la famille des réseaux de neurones, qui utilisent un modèle informatique inspiré de la structure du cerveau. L'unité fondamentale de base d'un réseau de neurones est le neurone. Chaque neurone possède un ensemble d'entrées, chacune ayant un poids spécifique. Le neurone calcule une fonction sur ces entrées pondérées. Un neurone

reçoit une combinaison linéaire des entrées pondérées et applique une fonction d'activation (sigmoïde, tanh, etc.). Chaque couche est entièrement connectée à la couche suivante que ça soit en réseaux multicouches ou à architectures profondes.

La sortie de chaque neurone artificiel dans une couche est une entrée pour les neurones artificiels de la couche suivante. Ensuite, les techniques de rétro propagation est utilisée pour adapter les poids de tous les neurones du réseau.

2. Types d'apprentissage automatique

L'apprentissage automatique des réseaux de neurones peut être classé en plusieurs classes :

-Apprentissage supervisé

-Apprentissage non supervisé

-Apprentissage semi supervisé

-Apprentissage par renforcement

Le mécanisme de l'apprentissage automatique peut être résumé dans la Figure 2.1.

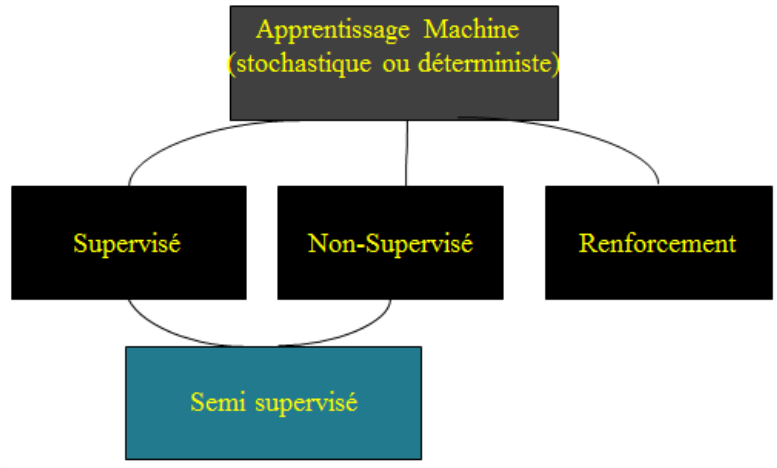


Figure 2. 1. Différents types d'apprentissage automatique

A partir d'un test d'apprentissge, on applique une méthode d'apprentissage supervisé, non supervisé ou une méthode par renforcement sur la base de donnée désirée. Au final, on obtiendra grace à ces méthodes un modèle prédictif et un groupe d'objets classifiés. Un schéma illustratif du processus d'apprentissage automatique est présenté dans la Figure 2.2.

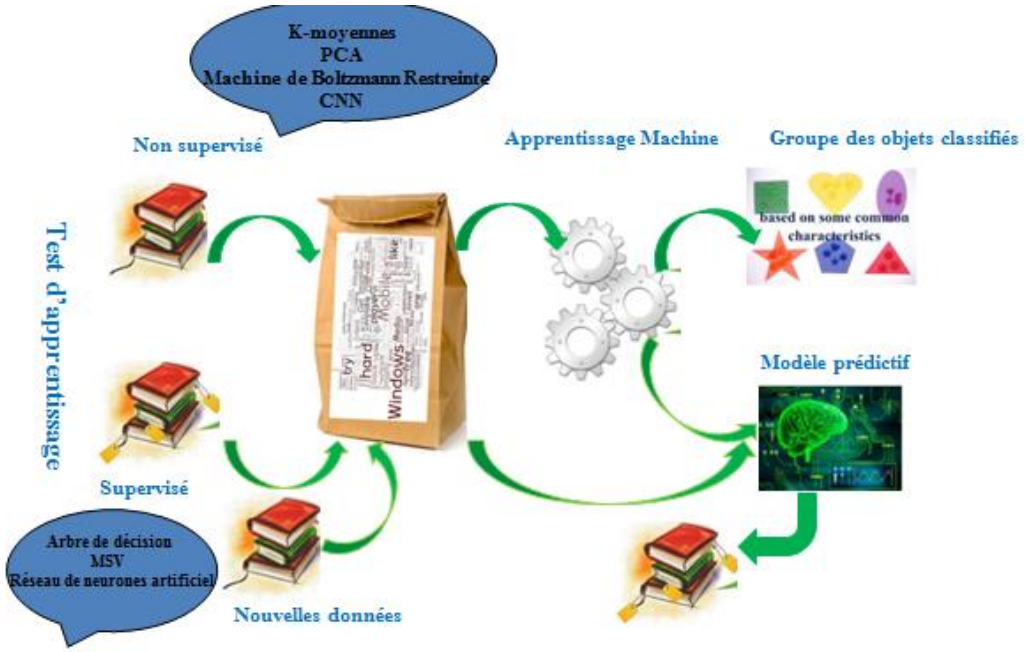


Figure 2. 2. Schéma illustratif du processus d'apprentissage automatique (supervisé et non supervisé) pour l'obtention d'un modèle prédictif et classification

2.1. Apprentissage supervisé

A partir d'une donnée d'entrée et d'une donnée désirée où les objets sont labélisés, la classification ou le modèle classification nous permet d'avoir une décision à la sortie en actualisant les poids à chaque couche. Le principe de ce type d'apprentissage est affiché dans la Figure 2.3.

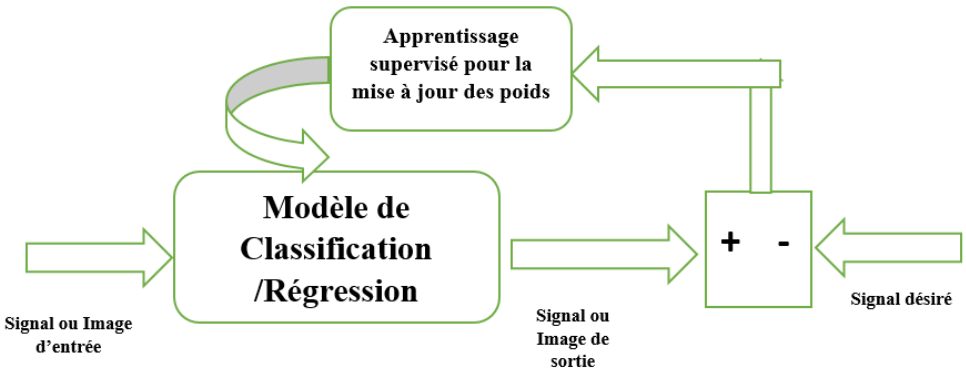


Figure 2. 3. Schéma de principe de l'apprentissage supervisé [Ray15]

Quelques domaines d'application de l'apprentissage supervisé sont présentés dans le Tableau 2.1.

Tableau 2. 1. Quelques domaines d'application de l'apprentissage supervisé

Algorithmes	Applications
Arbres de décision	Analyse financière
Classification Naïve bayésienne	Reconnaissance Faciale
Méthode des moindres carrées ordinaire	Prédiction métrologique
Machine à support de vecteurs SVM	Classification Image
Réseaux de Neurones Artificiels	Identification Faciale

2.2. Apprentissage non supervisé

En apprentissage non supervisé, les classes correctes des données à tester sont inconnues. Il peut à la fois extraire et sélectionner les caractéristiques. L'apprentissage non supervisé peut être sous la forme d'un regroupement (clustering) ou une réduction de dimensions, Les algorithmes d'apprentissage non supervisé les plus utilisés sont : K plus proches voisins, analyse en composantes principales ACP, Auto encodeurs, FCM [Say07], [Tli12].

Un schéma de principe de l'apprentissage non supervisé est présenté dans la Figure 2.4 et quelques domaines d'application de ce type d'apprentissage sont résumés dans le Tableau 2.2.

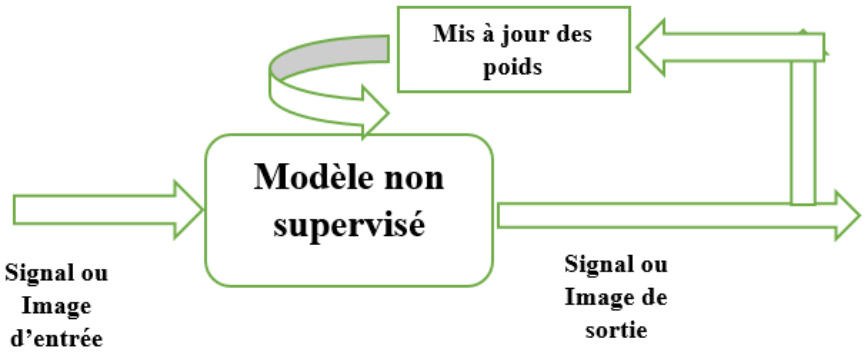


Figure 2. 4. Schéma de principe de l'apprentissage non supervisé [Wit14]

Tableau 2. 2. Quelques domaines d'applications de l'apprentissage non supervisé

Algorithmes	Applications
ACP	Compression des données, neuroscience
K plus proches voisins	Regroupements des documents
Auto encodeurs	Détection des contours

2.3. Apprentissage semi supervisé

Il s'agit d'une combinaison d'apprentissage supervisé et d'apprentissage non supervisé. Cette combinaison est parfois appréciée vu la facilité relative d'avoir des données non labélisées. On note parmi les approches connues en apprentissage semi supervisé les fonctions radiales de base [Kon07], l'empilement de Machine restreinte de Boltzman [Hin06], [Hra15], [Nzh18], [P3]. Cette dernière permet à l'utilisateur d'extraire les caractéristiques et ensuite appliquer les techniques de la rétro- propagation et du réglage fin sur le réseau. Elle sera détaillée par la suite.

2.4. Apprentissage par renforcement

Ce type d'apprentissage permet de réaliser l'apprentissage en fonction des règles et contraintes de l'environnement. Le système d'apprentissage appelé, agent dans ce contexte, à partir de l'environnement, sélectionne et accomplit des actions et obtient en retour des corrections et améliorations. Il apprend par lui-même la meilleure stratégie pour avoir autant de corrections que possible.

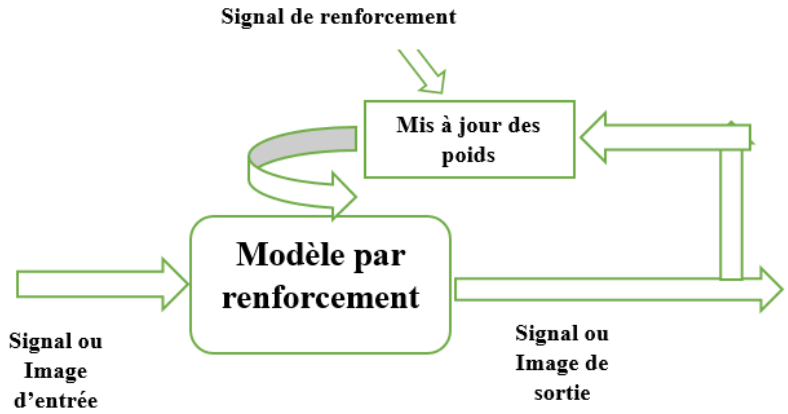


Figure 2. 5. Schéma de principe de l'apprentissage par renforcement [Zen20]

III. Détails sur l'extraction et le nettoyage des données

1. Extraction de données

Le prétraitement des données dans les systèmes d'apprentissage automatique implique à la fois l'extraction des données et des caractéristiques. Le prétraitement des images consiste à convertir des données brutes en données préparées. Par ailleurs, l'étape de l'extraction des

caractéristiques traite ensuite ces données préparées de manière à créer des attributs utiles pour l'apprentissage automatique.

2. Données brutes et données transformées

Les données à traiter par les systèmes d'apprentissage peuvent être brutes (dans un lac de données) ou transformées (dans un entrepôt de données). Les données brutes sont des données sous leur forme source non traitées pour l'apprentissage automatique. Les données transformées d'un entrepôt de données peuvent avoir été transformées en se basant sur leur forme d'origine pour être utilisées à des fins d'analyse. En plus, les données envoyées à partir de systèmes de diffusion en continu qui seront utilisées par des modèles d'apprentissage automatique pour la prédiction sont considérées comme des données brutes.

3. Prétraitement et nettoyage des données

Le prétraitement des données comprend diverses opérations. Chaque opération vise à aider le système d'apprentissage automatique à élaborer de meilleurs résultats de classification et prédiction. Le nettoyage de données consiste généralement à la suppression ou correction des données contenant des valeurs corrompues ou non valides au sein des données brutes, ou la suppression des données pour lesquels il manque un grand nombre d'informations. L'amélioration de la qualité des caractéristiques peut être réalisé à travers une égalisation d'histogramme, une normalisation, interpolation des valeurs manquantes, le découpage des anomalies et l'ajustement des valeurs ayant des distributions asymétriques et la simple suppression de caractéristiques si nombreuses valeurs sont manquantes.

IV. Méthodes d'extraction des régions d'intérêts à appliquer

Il y a trois étapes principales dans les techniques de détection de fumée dans les images : le prétraitement de l'image, l'extraction de caractéristiques et la détection de fumée. Parmi ces étapes, l'extraction des caractéristiques est la partie de base. Les méthodes traditionnelles dépendent de la sélection manuelle des caractéristiques, ce qui dépend des connaissances professionnelles de l'opérateur. Bien que les chercheurs développent de nombreuses études sur les caractéristiques des images de fumée et de flammes, seules les caractéristiques simples, telles que la couleur, les bords et les textures simples, ont été proposées. Cependant, en raison de la complexité des types et des scènes de fumée et d'incendie, ainsi que des nombreux événements d'interférence dans les enregistrements pratiques, parfois les caractéristiques extraites ne permettent pas une bonne distinction du feu et de la fumée, ce

qui entraîne une mauvaise précision et une faible capacité de caractérisation. Les techniques d'extraction des régions d'intérêts utilisées dans notre travail sont résumées par les étapes du tableau 2.3.

1. Extraction de région d'intérêt contenant la fumée par Mask-RCNN [Gir14]

La détection d'objets est un domaine de recherche qui permet de classer et localiser des régions d'une image ou d'un flux d'images vidéo. Cette tâche nécessite la combinaison de la classification d'image et la localisation d'objets. Le principe de la détection d'objets est comme suivi: dans une image donnée, on cherche les régions de celle-ci qui pourraient contenir l'objet désiré, puis on l'extrait et on la classe à l'aide d'un modèle de classification. Par conséquent, pour avoir une bonne méthode de détection d'objets, il est nécessaire d'avoir un algorithme robuste de détection de régions et un bon algorithme de classification d'images.

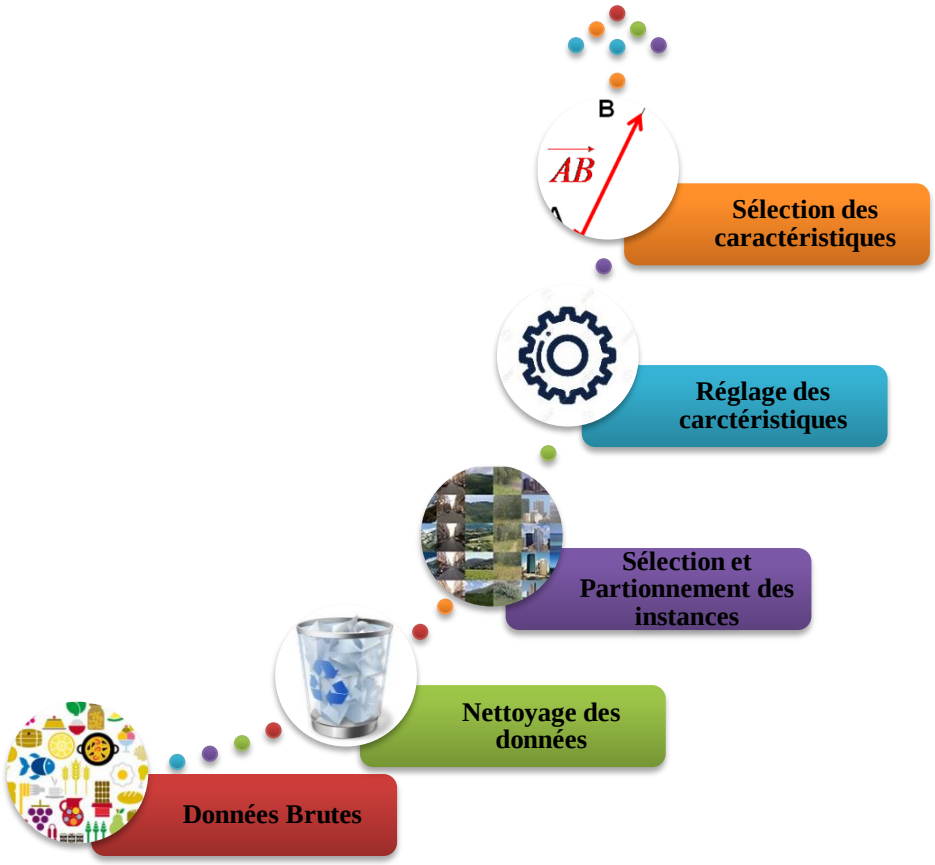


Figure 2. 6. Etapes de préparation de données pour les modèles à apprentissage automatique

Tableau 2. 3. Résumé des étapes proposées pour l'extraction des régions d'intérêt par Mask-RCNN et Yolo et classification par DBN

Algorithme d'apprentissage proposé en utilisant Mask RCNN ou Yolo pour la détection des régions d'intérêts et la classification par DBN 1. Procédure d'apprentissage par Mask RCNN ou Yolo et la classification par DBN 2. Etape 1 : Nettoyage des images d'entrées pour l'apprentissage 3. Etape 2 : Extraction des régions d'intérêts par Mask RCNN ou Yolo 4. Etape 3 : Classification du vecteur caractéristique extrait de l'étape 2 par DBN

Les réseaux de neurones convolutifs et l'apprentissage profond sont devenus les méthodes adéquates pour résoudre ces problèmes de détection et classification. Le recherche en détection d'objets est naturellement intégré les modèles de classification d'image, ce qui a permis de créer des méthodes tels que SSD (Single Shot Detector) [Yun20] et R-CNN, qui sont couplés avec un réseau pré-calculé pour la classification d'image. Pour ce genre de tâche le principe reste globalement le même, simplement, à la fin de notre réseau est ajouté une étape de création de masque capable de dire si oui ou non un pixel fait partie de l'objet détecté.

Les algorithmes R-CNN [Gir15], Fast R-CNN [Yun20] qui peuvent servir à la détection sont basés sur le CNN (Convolution Neural Network). Ils sont plus lents que les détecteurs à coup unique, mais donnent des résultats plus précis. Dans notre travail, on a utilisé une méthode de la même famille mais plus efficace : l'algorithme Mask RCNN [Gir15] dont l'architecture est exposée dans la Figure 2.7.

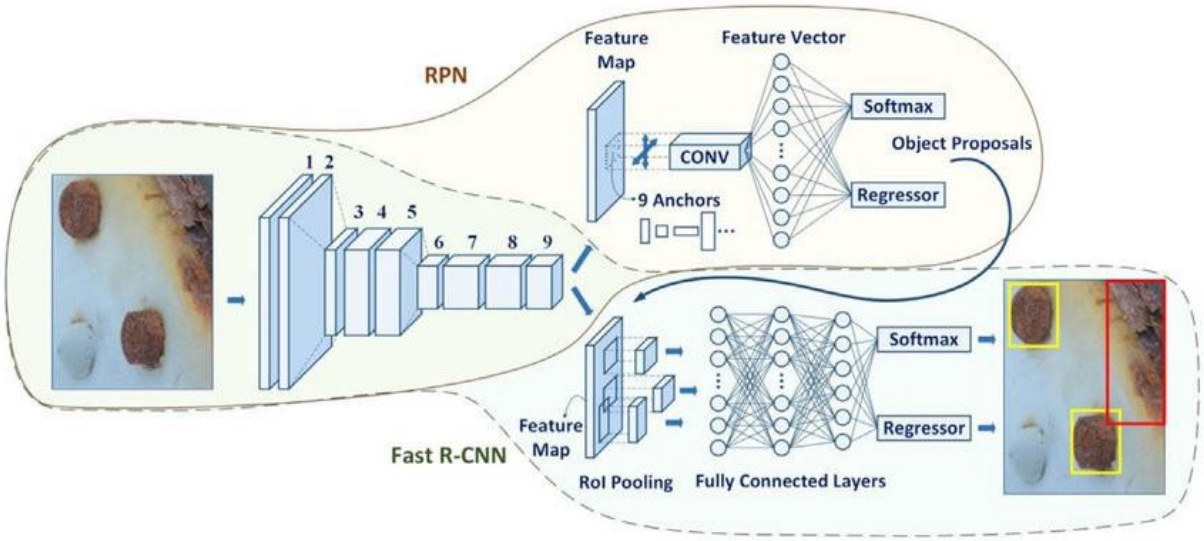


Figure 2. 7. Architecture de Mask-RCNN

Pour la détection des régions de fumée, nous avons utilisé le framework Tensor flow, le réseau Mask-RCNN ResNet101 dont l'apprentissage a été effectué avec la base de donnée COCO [Tsu14] qui permet de détecter plusieurs objets de différentes natures. C'est un ensemble de plus de 330000 images contenant des dizaines de types d'objets différents.

On a aussi utilisé les outils suivants :

-Linux Ubuntu 18.04.

-Python.

-Tensorflow object detection (Open source)

-Librairie de Open CV : utilisé pour gérer la gestion du flux vidéo.

2. Présentation du détecteur des objets en temps réel : YOLO [Red16]

Les détecteurs d'objets « à coup unique » « Single Shot Detectors » détectent les frontières des régions et les classifient en même temps. Comme il s'agit d'un processus en une seule étape, il est beaucoup plus rapide que les détecteurs régionaux. Il convient de noter que ce type de détecteurs sont peu performants lorsqu'il s'agit d'objets de petite taille. Les algorithmes populaires qui utilisent cette méthodologie sont YOLO [Red16] (You Only Look Once) et SSD [Yun20] (Single Shot multibox Detector).

Sa grande force est sa rapidité : il peut être appliqué en temps réel (jusqu'à 45 image / seconde). Yolo est plus rapide que des R-CNN, car il découpe l'image en petits blocs et

génère des tenseurs pour chaque bloc. Le principe du YOLO est de faire parcourir l'image une seule fois par un réseau de neurones profonds (d'où le nom de You Only Look Once), par opposition aux méthodes dites de régions (notamment utilisées par les modèles basés sur R-CNN).

Le schéma de principe de l'algorithme Yolo est affiché dans la Figure 2.8.

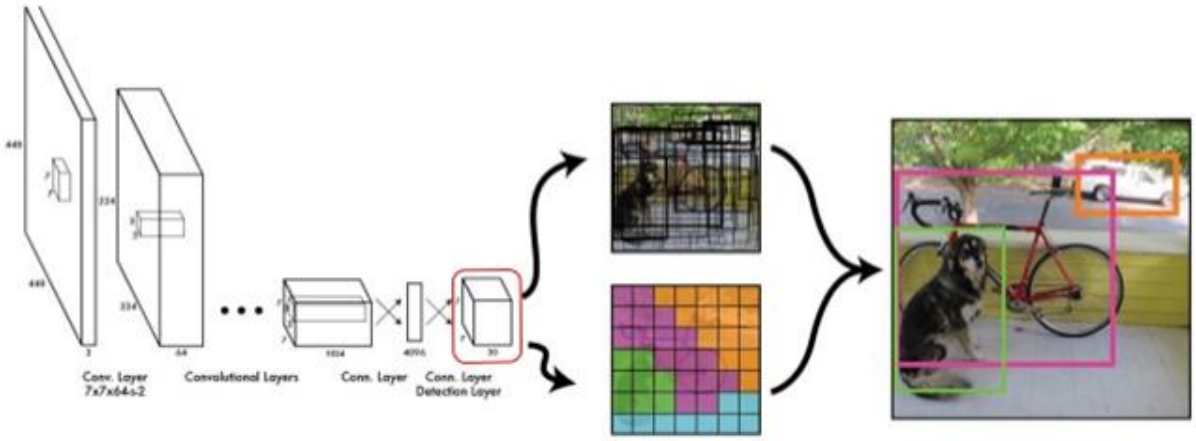


Figure 2. 8. Schéma de principe de l'algorithme Yolo (You Only Lok Once) [Red16]

La technique de détection de cette méthode peut se décomposer en deux grands composantes :

-Première composante : Le réseau de YOLO est composé d'un premier CNN de 53 couches appelé Darknet53. Il a été pré-entrainé sur la base de données ImageNet.

L'idée principale est d'appliquer une méthode d'apprentissage par transfert « tranfert learning » qui permet de s'appuyer sur un modèle pré-entraîné avec une tâche semblable à la nôtre, pour pouvoir réaliser un apprentissage qui convergera plus rapidement [Red16]. Les images sont divisées en une grille de cellules et la cellule chargée de détecter l'objet est celle qui contient le centre de l'objet.

-Deuxième composante : Un second réseau de 53 couches très proche du premier et empilé. Lors de l'apprentissage du modèle, les poids de ce second réseau seront ajustés avec les données de la base d'apprentissage, contrairement aux poids du premier réseau qui seront fixés (principe du transfert learning).

La version 3 qu'on a utilisé du modèle YOLO est caractérisée par l'augmentation du nombre de couches du réseau et l'implémentation de trois échelles des « boxes » pour détecter des objets de tailles petites, ce qui permet de détecter plusieurs « boxes » qui se chevauchent

pour un même objet. Ceci est réalisé avec l'application de la technique de « Non Max Suppression » pour ne garder que les régions les plus significatives.

On note à la fin que la limite de ces détecteurs de zones d'intérêt est la nécessité d'avoir un volume très important de la base de données et des ressources et aussi parfois le manque de précision.

V. Structure du réseau d'apprentissage profond utilisé le DBN (Deep Belief Network)

Il n'existe pas un meilleur algorithme d'apprentissage automatique efficace pour toutes les applications en classification d'images. Avec un volume de données assez important, comme le cas des vidéos de contrôle des feux dans les forêts, et grâce aux avancées en matière de stockage des données et de traitement en temps réel, on peut se permettre d'explorer des outils d'apprentissage automatique de plus en plus complexes tels que les réseaux de neurones DBN.

1. Présentation du réseau DBN : (Réseaux de conviction profonde)

Le Deep Belief Network [Hin06] [LeC15] est un modèle probabiliste génératif composé de plusieurs couches cachées. Chaque couche est associée à des corrélations d'ordre élevé. L'approche générative consiste à modéliser la distribution conjointe des variables observées en utilisant généralement un critère basé sur la log-vraisemblance. Cette approche ne nécessite pas de données étiquetées. Le DBN utilisé pour l'apprentissage de nos données est composé d'une couche visible et de deux couches cachées. En effet, le DBN est une machine Boltzmann restreinte (RBM : Restricted Boltzmann Machine) qui apprend les caractéristiques (phase de prétraitement) et les propage ensuite pour affiner les réseaux. Après la phase de prétraitement, le réseau DBN agit comme un perceptron multicouches (Multi-layer Perceptron). Enfin, après avoir ajouté une couche de sortie, on peut insérer une régression logistique comme classifieur. Dans ce qui suit, nous présentons les différentes étapes de l'apprentissage, du prétraitement et de l'extraction des caractéristiques.
Le RBM [Hin06], [Hra15], [Nzh18], est un algorithme adapté à la classification, la réduction des dimensions et à l'apprentissage des caractéristiques. Il peut être considéré comme le bloc de structure de base des réseaux à apprentissage profonds. Le RBM est généralement utilisé

comme extracteur de caractéristiques et classificateur non linéaire autonome, ainsi que pour la régression logistique [Efr75] et les SVM [Chi02], [Gub09], [Mar10].

2. Choix des paramètres du réseau proposé

Chaque nœud visible prend une caractéristique de bas niveau d'un élément de l'ensemble de données avec lequel l'apprentissage sera fait. Les désignations et les choix des hyper-paramètres du réseau utilisé sont détaillés dans le Tableau 2.4.

Tableau 2. 4. Choix et désignation des hyper-paramètres

Hyper paramètre	Désignation
$v_i = \{1, 2, \ldots, 100\}$	Valeur de nœuds i dans la couche visible
$h_j = \{1, 2, 3\}$	Valeur de nœuds j dans les couches cachées
$b = \{b_1, b_2, b_3, \ldots .. b_{100}\}$ $c = \{c_1, c_2, c_3\}$	b_i, c_j: biais associé à la i^{eme} couche visible et la $j^{ème}$ couche cachée respectivement.
w_{ij}	Le poids entre la $i^{ème}$ couche visible et la $j^{ème}$ couche cachée
σ_i^2	La covariance de $i^{ème}$ distribution
batch	64

3. Minimisation de la fonction coût et optimisation des poids [Hin02],[Hin06]

L'énergie $E(v,h)$ utilisée pour affiner le modèle du réseau DBN est définie par :

$$E(v,h) = \sum_{i=1}^{100} \frac{(v_i - b_i)^2}{2\sigma_i^2} - \sum_{j=1}^{3} c_j h_j - \sum_{i=1}^{100}\sum_{j=1}^{3} \frac{v_i}{\sigma_i} h_j w_{ij} \quad (2.1)$$

Les modèles probabilistes basés sur la minimisation de l'énergie définissent une distribution de probabilité à travers la fonction énergétique à minimiser donnée par :

$$P(v) = P(v|h) = \sum_h p(v|h) = \frac{1}{Z}\sum_h e^{-E(v,h)} \quad (2.2)$$

$$\text{Où } Z = \sum_{v,h} e^{-E(v,h)} \quad (2.3)$$

L'apprentissage consiste à mettre à jour les paramètres du réseau (poids et biais) avec la méthode du gradient. Elle nous permet de réduire aussi la divergence de Kulback Leibler

[Car05], [Hin02]. La fonction de perte à minimiser dépend de la fonction énergétique et elle est définie avec :

$$\ln L(\theta|v) = \ln p(\theta|v) = \ln \frac{1}{Z}\sum_{h} e^{-E(v,h)} = \ln \sum_{h} e^{-E(v,h)} - \ln \sum_{v,h} e^{-E(v,h)} \tag{2.4}$$

Où θ est un vecteur qui contient les paramètres du modèle W, b et c.

Le terme dérivé de "gradient", a des parties positives et négatives :

$$\frac{\partial}{\partial\theta}\ln L(\theta|v) = \ln p(\theta|v) = \ln \frac{1}{Z}\sum_{h} e^{-E(v,h)} = \frac{\partial}{\partial\theta}(\ln \sum_{h} e^{-E(v,h)}) - \frac{\partial}{\partial\theta}(\ln \sum_{v,h} e^{-E(v,h)}) \tag{2.5}$$

Les parties positive et négative ne se réfèrent pas à leurs signes mais à la variation de la densité de probabilité définie par le modèle. Le premier terme augmente la probabilité d'obtenir des données d'apprentissage (en réduisant l'énergie correspondante), tandis que le second diminue la probabilité que des échantillons soient générés par le modèle.

Ainsi, on obtient l'équation suivante pour l'apprentissage [Hin02]

$$\frac{\partial \ln p(v)}{\partial w_{ij}} = \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{reconstructed} \tag{2.6}$$

Où i = 1,..,100, j = 1,2,3, et les symboles $\langle\ \rangle_{data}$ and $\langle\ \rangle_{reconstructed}$ sont utilisés pour représenter les valeurs à estimer des données et du modèle reconstruit, respectivement.

$$\frac{\partial \ln p(v)}{\partial w_{ij}} = \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{reconstructed} \tag{2.7}$$

Cela conduit à une règle d'apprentissage très simple pour effectuer la plus forte ascension stochastique dans la probabilité de log des données d'apprentissage :

$$\Delta w_{ij} = \left(\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{reconstructed}\right).Lr \tag{2.8}$$

Le terme multiplicatif Lr est le Learning rate (Coefficient d'apprentissage) qui sera fixé expérimentalement par la suite et qui dépendra essentiellement de comportement de la fonction coût à minimiser lors de l'apprentissage. Il influe sur la rapidité et la stabilité de l'apprentissage.

Nous pouvons commencer avec un état aléatoire dans l'une des couches, puis effectuer un échantillonnage de Gibbs en alternance [Hua97]. Toutes les unités d'une couche sont mises à jour en parallèle en fonction de l'état actuel des unités de l'autre couche. Cette étape est identique à la génération de données dans les réseaux d'apprentissage profond.

L'échantillonnage de Gibbs réduit la divergence de Kullback Leibler [Car05], [Hin02]. Les unités visibles sont échantillonnées simultanément avec des valeurs de la couche cachée (et vice versa).

Les valeurs de *v*, *h* sont calculées avec :

$$\begin{cases} h^{(n+1)} \simeq sigm(W'v^{(n)} + c) \\ v^{(n+1)} \simeq sigm(W'h^{(n+1)} + b) \end{cases} \quad (2.9)$$

$$\begin{cases} P(h^{(n+1))} \left| v^{(n)} \right.) = sigm(W'h^{(n+1)} + b) \\ P(v^{(n+1)} \left| h^{(n)} \right.) = sigm(W'v^{(n)} + c) \end{cases} \quad (2.10)$$

Où *sigm* est la fonction sigmoïde.

4. Comportement du réseau comme un perceptron multicouche après le prétraitement

La deuxième couche RBM est formée à partir du poids et de biais de la première couche RBM, les unités cachées du deuxième niveau RBM ont tendance à présenter des poids positifs élevés similaires à ceux de la première couche [Hin02]. En fait, la deuxième couche RBM extrait des caractéristiques du niveau supérieur. Après la phase de prétraitement utilisant la machine restreinte de Boltzmann, le réseau se comporte comme un perceptron multicouche. Le principe du perceptron multicouche proposée et développé par Franck Rosemblatt [Kho98] est présenté par la Figure 2.9.

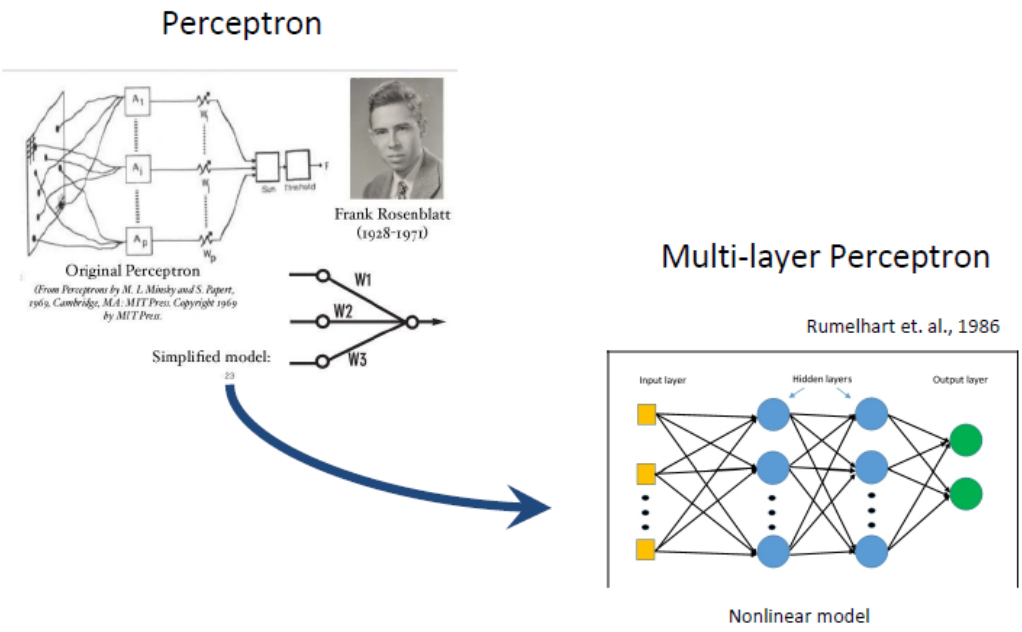


Figure 2. 9. Rappel du Schéma de principe du Perceptron multicouches [Kho98]

Le perceptron multicouche [Kho98] est un réseau neuronal artificiel à anticipation composé d'au moins trois couches. Chaque nœud formant une couche représente un neurone utilisant une fonction d'activation non linéaire. Pour l'apprentissage, le perceptron multicouche utilise la méthode classique de la rétro-propagation [Cha06], [Hra15] comme outil d'apprentissage supervisé. Après l'apprentissage préalable, nous ajoutons une couche de sortie à partir des paramètres de la dernière couche cachée. Nous affinons les poids du modèle préformé en continuant la rétro propagation. Cette ancienne technique est encore utilisée pour former de grands réseaux d'apprentissage en profondeur. Il est possible d'adapter avec précision toutes les couches du réseau d'apprentissage profond, mais on peut aussi conserver certaines des couches précédentes fixes (en raison de problèmes de configuration excessive) et adapter uniquement une partie du réseau de niveau supérieur.

Enfin, nous allons utiliser la régression logistique en tant que classificateur pour classifier les données qui se trouve dans la dernière couche cachée.

5. Etape de classification avec la régression logistique

La régression logistique [Effr75] est basée sur l'estimation conditionnelle avec la méthode de maximum de vraisemblance. Nous choisissons les paramètres w qui maximisent la probabilité de log de *y* dans les données d'apprentissage en fonction des observations *v*.

L'idée est de prédire la probabilité que les nœuds de la couche d'entrée appartiennent à une certaine classe.

$$P(y=1|v) = f_{w_{ij}}(v) = \text{logistic}(w_{ij}.v) = \frac{1}{1+e^{(-w_{ij}v)}} \tag{2.11}$$

où i=1,…,100, j=1,...,3

Pour l'apprentissage, on définit la fonction coût suivante :

$$Loss(y_t, f_{w_{ij}}(v)) = -y_t \ln(f_{w_{ij}}(v)) - (1-y_t).\ln(1-f_{w_{ij}}(v)) \tag{2.12}$$

Pour l'apprentissage de la couche visible, les poids sont adaptés avec :

$$\begin{aligned} w_{ij} &\rightarrow w_{ij} - \alpha \frac{\partial}{\partial w_{ij}} Loss(y\ , f_w(v)) \\ w_{ij} &\rightarrow w_{ij} + \alpha(y - f_{w_{ij}}(v)) \end{aligned} \tag{2.13}$$

L'architecture proposée du réseau DBN utilisé dans nos expériences de détection de fumée est représentée dans la Figure 2.10.

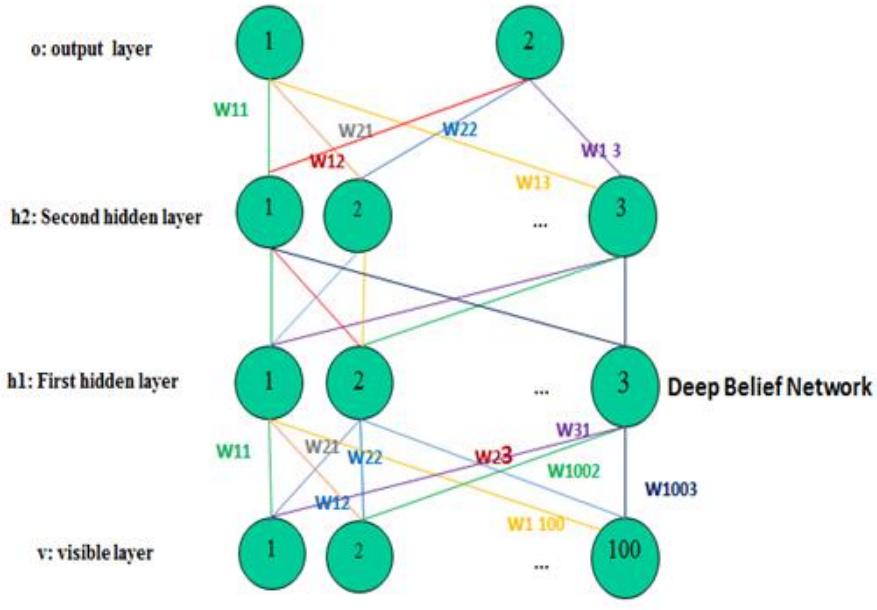


Figure 2. 10. Architecture utilisée du réseau DBN

6. Phase de classification en utilisant l'algorithme d'estimation du moment adaptatif ADAM [Des18], [Zou19]

Le choix de l'algorithme d'optimisation pour n'importe quel modèle d'apprentissage approfondi peut faire la différence entre les performances et la rapidité des résultats obtenus. L'algorithme d'optimisation d'ADAM (Algorithme d'estimation du moment adaptatif) (Adaptive Moment Estimation Algorithm) [Des18], [Zou19] est une extension de la méthode de la descente du gradient stochastique (SGD) (la descente stochastique du gradient, Stochastic Gradient Decent) [LeC98] qui a récemment connu une utilisation plus large dans les applications d'apprentissage profond, la vision par ordinateur et le traitement du langage naturel.

ADAM est un algorithme d'optimisation qui peut être utilisé à la place de la procédure classique de descente de gradient stochastique pour mettre à jour les poids du réseau en fonction des données d'apprentissage. Les avantages intéressants de l'utilisation d'ADAM sur des problèmes d'optimisation non convexes sont comme suit :

-Facilité de la mise en œuvre.

-Efficacité de calcul.

-Mémoire requise pas très grande

-Bonne adaptation aux problèmes de Big Data.

-Adaptation au contexte non stationnaires.

Les hyper-paramètres ont une interprétation intuitive et nécessitent généralement peu de réglages. En effet, l'algorithme ADAM combine les avantages de deux versions étendues de l'algorithme de la descente du gradient stochastique : l'algorithme du gradient adaptatif [Lou98], et l'algorithme de la propagation de la moyenne quadratique. Ce dernier algorithme utilise des coefficients d'apprentissage par paramètre qui sont adaptés en fonction de la moyenne des amplitudes récentes des gradients pour le poids (par exemple, la vitesse à laquelle il change). Cela signifie que l'algorithme fonctionne bien pour les problèmes en ligne, dans un contexte bruité ou non stationnaire.

Par contre, au lieu d'adapter les coefficients d'apprentissage en fonction du premier moment uniquement, comme dans l'algorithme de la moyenne quadratique, ADAM utilise également la moyenne des seconds moments des gradients dans cette adaptation.

Malgré la grande popularité d'Adam, nous avons remarqué que parfois, il ne converge pas vers une solution optimale dans le cas de grandes valeurs du coefficient d'apprentissage.

Pour des raisons de comparaison, on a aussi utilisé dans cette phase l'algorithme classique de SGD, (la descente stochastique du gradient, Stochastic Gradient Decent) [LeC98]. Les variations de la fonction coût d'apprentissage et de la précision lors de l'apprentissage sont présentées sur les Figure 2.11 et 2.12 et montrent une supériorité de l'optimiseurs ADAM par rapport à l'algorithme classique SGD (une fonction coût moindre et une précision meilleure lors de l'apprentissage, surtout au début de l'apprentissage). C'est la raison pour laquelle nous avons ensuite choisi d'utiliser ADAM dans nos expériences présentés dans le chapitre suivant.

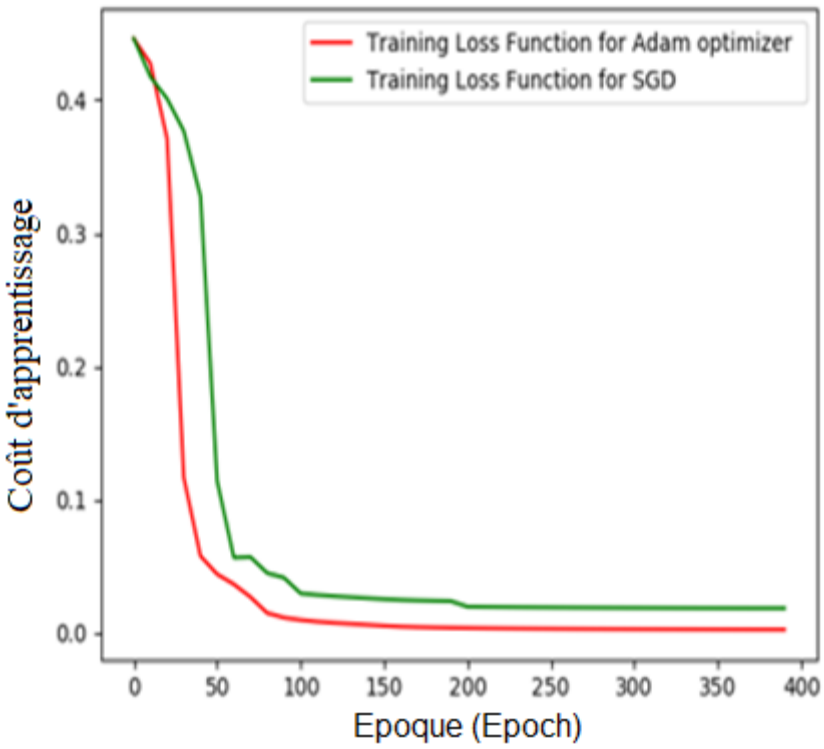


Figure 2. 11. Comparaison de la variation de la fonction coût pendant l'apprentissage en utilisant le SGD et l'optimiseur ADAM

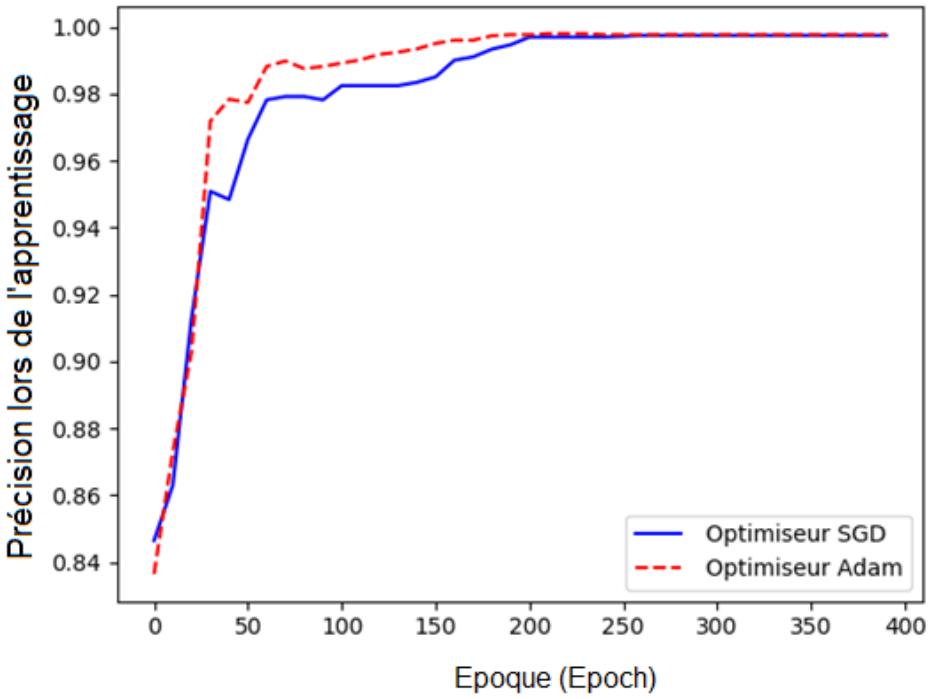


Figure 2. 12. Comparaison de variation de la fonction de précision pendant l'apprentissage en utilisant le SGD et l'optimiseur ADAM »

VI. Conclusion

Dans ce chapitre, on a commencé par présenter les différents systèmes d'apprentissage automatique, et aussi les méthodes d'extraction des régions d'intérêt, qu'on a appliqué à la localisation de la fumée. Nous avons ensuite exposé la méthodologie proposée basée sur la classification par DBN. Par ailleurs, nous avons fixé les hyper-paramètres du réseau trouvés avec différentes expériences de simulations et de comparaison pour optimiser le choix du nombre de nœuds, nombre de couches, et des coefficients d'apprentissage. Enfin, nous avons présenté deux types de systèmes de classification : SGD et l'optimiseur ADAM, et nous les avons comparés. On a aussi étudié leur influence sur la fonction coût et le taux de détection. Les résultats des expériences simulations et les comparaisons entre la méthode proposée et les méthodes traitées dans la littérature fera l'objet du chapitre suivant.

Chapitre III : Méthode proposée pour la détection de fumée avec le DBN et résultats expérimentaux

Chapitre III : Méthode proposée pour la détection de fumée avec le DBN et résultats expérimentaux

I. Introduction et présentation des données

Dans ce dernier chapitre, nous présentons les résultats expérimentaux que nous avons faites sur la détection et la localisation de la fumée avec la technique présentée dans le chapitre précédent basée sur le système d'apprentissage DBN, que nous appelons aussi modèle. Nous commençons par présenter la base de données utilisée. Ensuite on présente les simulations réalisées pour fixer les paramètres du réseau de détection utilisé. Nous présentons enfin une comparaison entre les résultats de plusieurs méthodes de détection de fumée traitées dans la littérature et notre méthode développée pour présenter la vraie valeur ajoutée dans la détection de fumée.

Les algorithmes proposés ont été exécutés sur un processeur PC Intel Core (TM) i5, 1.80GHz, RAM 4 GB. Le programme proposé a été réalisé en utilisant Python et la librairie open source Open CV. L'ensemble des données utilisées est disponible à l'adresse suivante : http://signal.ee.bilkent.edu.tr/VisiFire/Demo/SampleClips et http://www.openvisor.org.

Tout d'abord, nous avons divisé les vidéos en trames d'images de taille 320 × 240 pixels. L'ensemble des images a été divisé en deux parties : un premier ensemble pour l'apprentissage et un deuxième ensemble pour la validation. Afin de tester en contrôler les performances du système de détection e fumée proposé, nous avons évalué les critères de performances classiques, qui seront détaillés au cours de ce chapitre.

Pour l'apprentissage du modèle, nous avons utilisé un ensemble de données de dix vidéos contenant de la fumée et six vidéos sans fumée. A partir de ces vidéos, on a extrait un total de 20000 images. Nous avons utilisé 70 % des images extraites pour l'apprentissage, et le reste ont été utilisées pour la validation et le test. Le prétraitement que nous proposons d'appliquer à la base de données a permis de réduire la complexité de la méthode de détection.

Plusieurs exemples d'images extraites de la base de donnée sont présentés dans les Figures 3.1, 3.2 et 3.3, dans les cas où il n'y a que de la fumée, présence de fumée avec des nuages et du brouillard et dans las de non présence de fumée.



Figure 3. 1. Exemples d'images de la base de données contenant de la fumée



Figure 3. 2. Exemples d'images de la base de données contenant de la fumée, des nuages, des personnes en mouvement, du brouillard etc.



Figure 3. 3. Exemples d'images de la base de données sans fumée

II. Détails de la méthode proposée pour de la détection de fumée

1. Résumé de la technique proposée :

Une des idées principales que nous proposons dans notre travail est de sélectionner les images contenant un mouvement en utilisant l'algorithme GMM des mélanges de gaussiennes [Yan14], [Mck00], [Sta00]. avant de commencer la détection de la fumée. Cette technique, qui a été détaillé dans le chapitre II de ce rapport, nous permet de garder uniquement les zones de l'images concernées par un déplacement ou un mouvement : c'est la dynamique temporelle de la fumée qui est prise en compte. Des exemples de résultat de l'application du GMM pour l'extraction du mouvement de la fumée sont illustrés sur la Figure 3.6. Par la suite, nous calculons la différence entre les composantes R, V, B de chaque trame. Pour les trames où R, V, B sont proches, nous conservons les valeurs pour calculer le rapport énergétique de ces trames.
A ceci s'ajoute le calcul du rapport entre les énergies de chaque trame et de la trame de référence correspondante, avant l'apparition de la fumée. Si les contours sont flous, le rapport d'énergie est inférieur à 1. Nous conservons donc la trame de la fumée et nous effectuons l'extraction de la région d'intérêt.

Ces régions d'intérêt seront ensuite insérées dans un vecteur de caractéristiques qui viendra alimenter le réseau DBN. Les organigrammes des Figures 3.4 et 3.5 donnent un aperçu sur la méthode que nous proposons dans le cadre de ce travail.

Pour autant que nous sachions, le seul travail qui a utilisé le DBN comme méthode de détection de la fumée est celui de Pundir [Pun17]. La différence entre notre approche proposée et la méthode développée dans [Pun17] est que nous ne nous intéressons qu'aux caractéristiques de couleur, de mouvement et d'énergie. Ces caractéristiques sont insérées dans le classificateur DBN contrairement à la méthode susmentionnée [Pun17] basée sur l'analyse du mouvement, de la couleur et de la texture.

Pour cette raison, nous proposons dans ce chapitre de montrer l'importance d'utiliser l'énergie comme une caractéristique pertinente de la fumée dans le cadre de la classification en comparant les performances des deux méthodes.

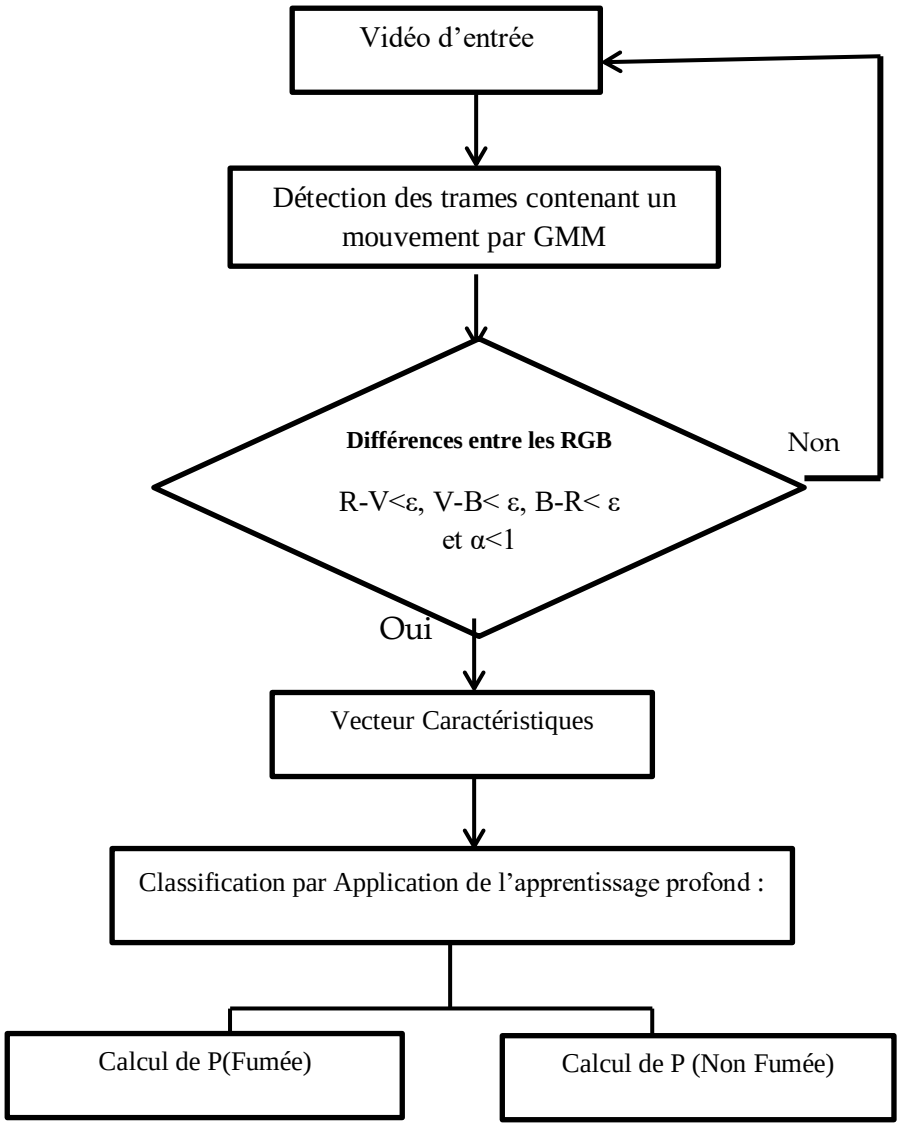


Figure 3.4. Organigramme de la technique proposée

En outre, la méthode proposée présente une amélioration de la performance de la détection de fumée en la comparant avec celle de Pundir et al [Pun17] en jouant sur la combinaison des caractéristiques tels que le mouvement, la couleur et l'énergie pour extraire le vecteur des caractéristiques et limiter les régions d'intérêts en même temps. Les anciens travaux ne mentionnent pas l'emplacement réel de la fumée et ne précisent pas les régions d'intérêt.

La contribution importante consiste à l'idée de limiter les régions d'intérêt avant de les introduire dans le DBN et de calculer les critères de détection avec des ressources matérielles abordables et un temps de calcul minimal.

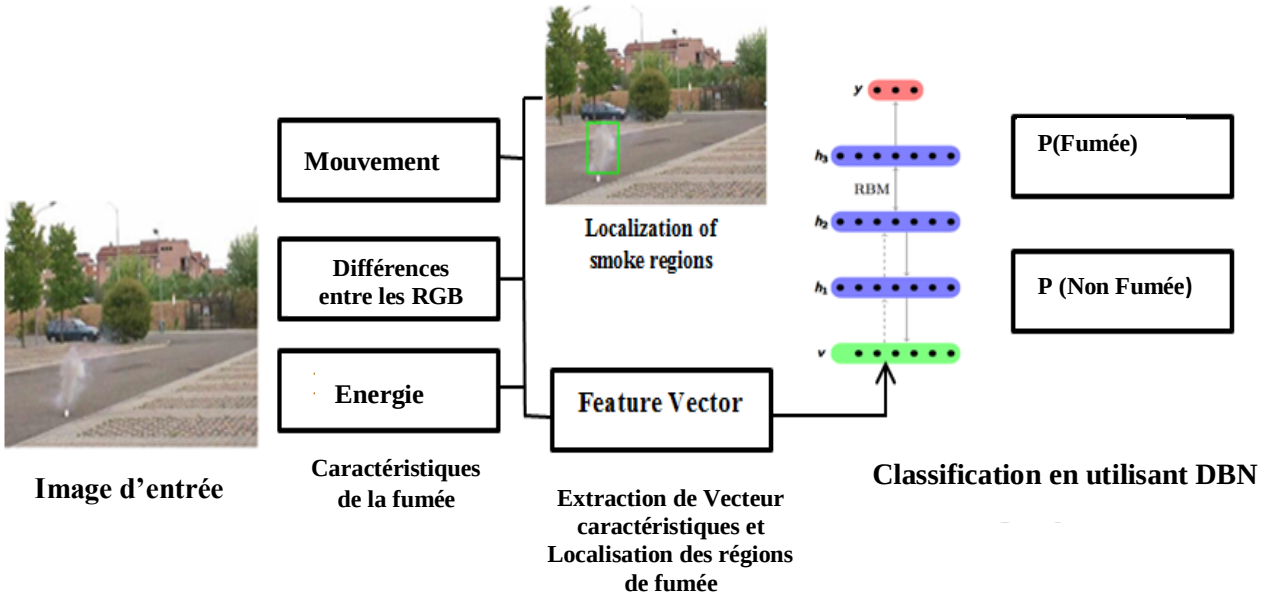


Figure 3. 5. Schéma de récapitulation de la méthode proposée pour la détection de la fumée

2. Etape du prétraitement et de localisation des fumées dans les images

Les images contenant de la fumée sont caractérisées par des composants couleurs R, G, B. L'idée est de fixer expérimentalement un seuil pour détecter les zones où les différences des composantes : R - G, R - B et G - B ont des valeurs proches de zéro. Après avoir appliqué les tests des différences des RGB aux trames contenant des mouvements élevés en utilisant le modèle des mélanges de gaussiennes, comme schématisé dans la Figure 3.2, nous calculons le rapport de l'énergie entre la trame actuelle et l'énergie de la trame d'arrière-plan avant l'apparition de la fumée. Les contours des régions de fumée deviennent donc flous. Par conséquent, l'énergie des trames contenant les régions de fumée diminue. Ainsi, si le rapport entre l'énergie de la région de l'image considérée et l'image d'arrière-plan est inférieur à 1, la probabilité que ce soit de la fumée est importante, et la région d'intérêt pour localiser la fumée sera détectée. On note à ce niveau que le calcul de l'énergie est effectué sur les sous bandes issues de la transformation par ondelettes discrètes que nous avons appliquées aux trames traitées, comme détaillé dans les équation (1.4) et (1.5) du chapitre II de ce rapport. Cette approche nous a permis de localiser facilement la fumée dans les images vidéo. La région de fumée sera mentionnée dans le résultat par un rectangle vert, comme illustré avec des exemples de résultats dans la Figure 3.4.



Figure 3. 6. Exemples de résultats de l'application du GMM pour l'extraction du mouvement de la fumée

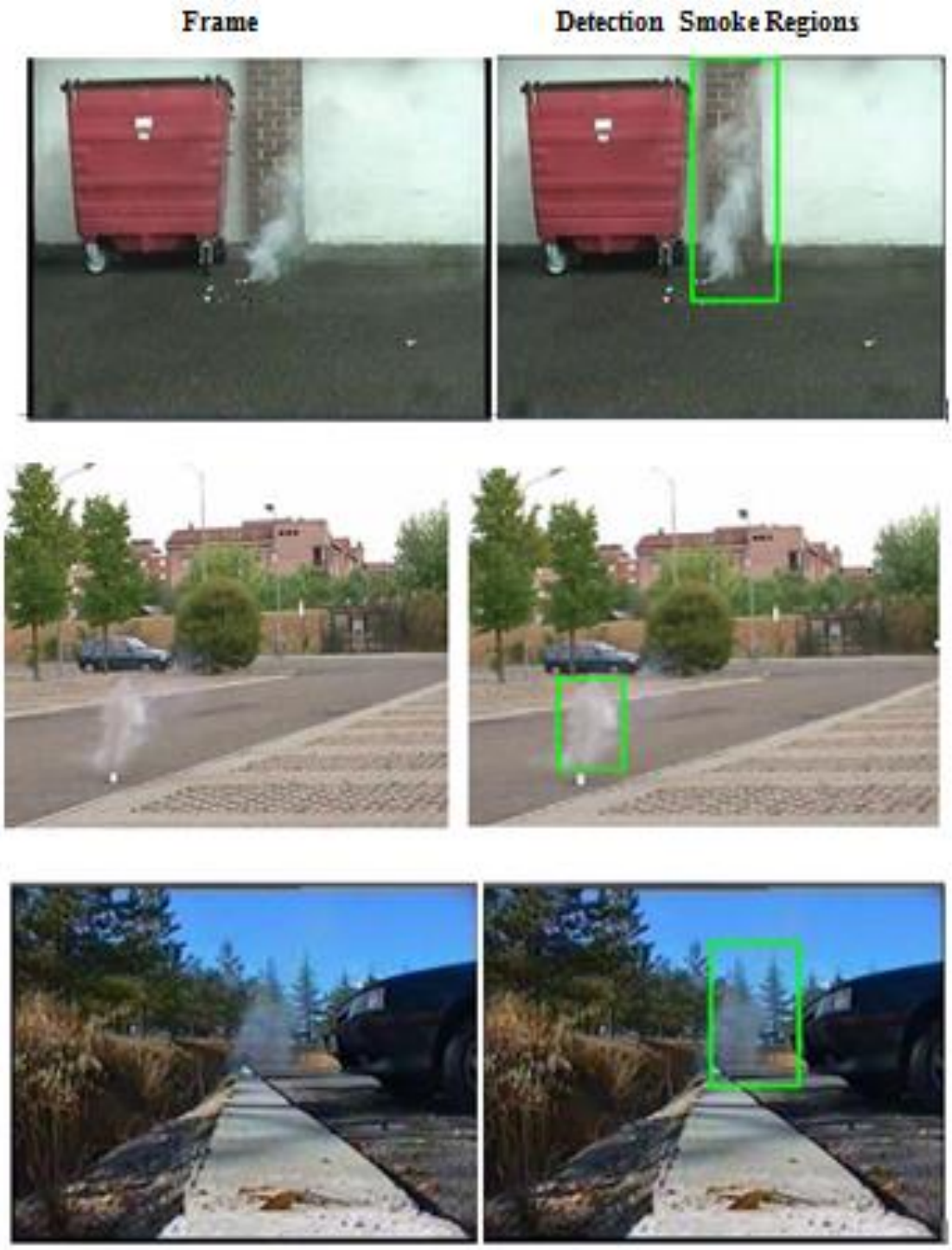


Figure 3. 7. Exemples de résultats obtenus de localisation de fumée

3. Expériences réalisées pour le choix des paramètres du réseau DBN :

Nous présentons dans cette section les simulations que nous avons réalisées pour pouvoir bien choisir les paramètres du réseau DBN utilisé [P1]. En effet, le bon choix des hyper paramètres du réseau nous permettra d'améliorer les performances du système de

classification et détection de fumée. Il s'agit donc d'optimiser le choix des paramètres suivants : le coefficient d'apprentissage *Lr* (Learning Rate), le nombre de couches cachées et le nombre maximal d'itérations de pré-apprentissage.

Nous avons testé plusieurs valeurs du coefficient d'apprentissage et nous avons opté pour un coefficient d'apprentissage *Lr*=0.001 où le coût de prétraitement diminue d'une façon stable et plus rapidement par rapport aux autres valeurs testées de *Lr*. Un exemple du coût de prétraitement en fonction du nombre d'époques pour 4 valeurs de *Lr* est illustré sur la Figure 3.8.

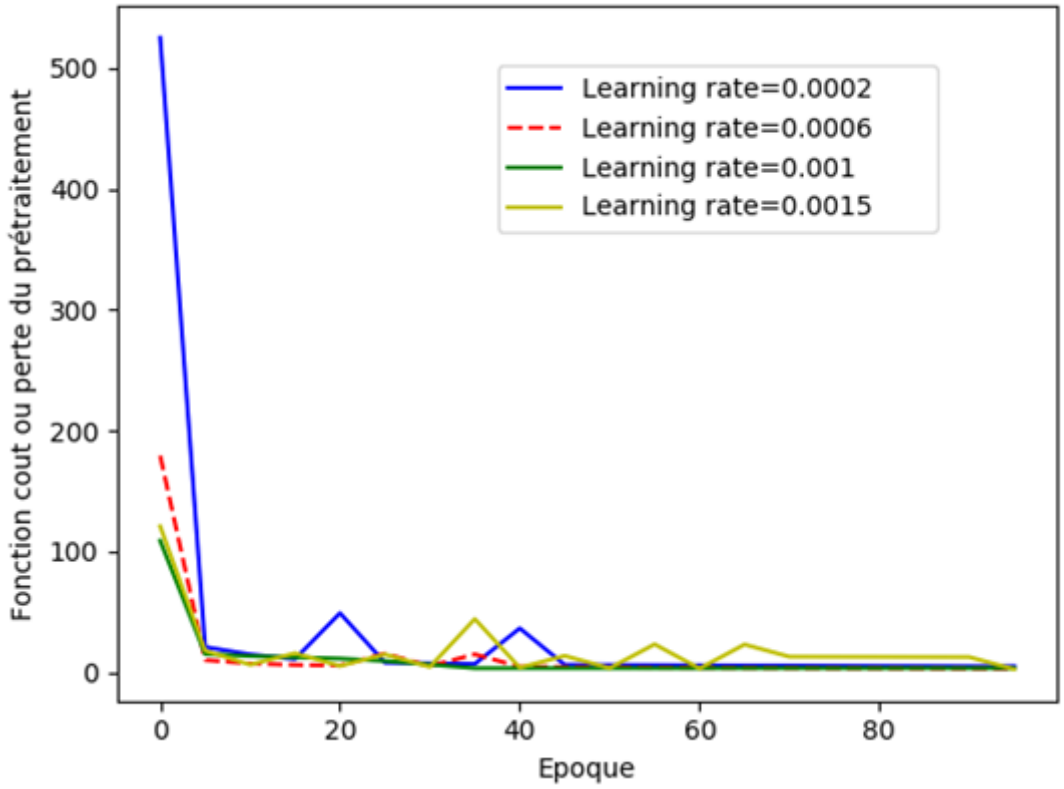


Figure 3. 8. Exemple de l'influence de la variation du coefficient d'apprentissage *Lr* sur la fonction coût du prétraitement

On note à ce niveau que la différence entre l'époque (epoch) et l'itération est que l'époque décrit le nombre de fois que l'algorithme d'apprentissage reçoit tout l'ensemble de données tandis que l'itération décrit le nombre de fois qu'un lot particulier de données passe à travers algorithmes. En plus, le nombre d'époques de prétraitement est l'un des paramètres du réseau.

Après avoir choisi le coefficient d'apprentissage, nous devons fixer le nombre de couches cachées. Le nombre de couches cachées est l'un des paramètres les plus importants dans le

réseau. Il n'y a pas de nombre spécifique optimal des couches cachées mais on peut estimer expérimentalement ce nombre en testant la rapidité de convergence de la fonction coût du prétraitement pour différents nombres de couche cachées. On a présenté dans la Figure 3.1 la variation de la convergence de la fonction coût du prétraitement par rapport au nombre d'époques, pour 3 valeurs du nombre de couche cachée, noté h : h=2, h=3 et h=4. Les résultats montrent que les cas où h=2 et h=3 donnent les meilleurs résultats de rapidité de convergence de la fonction coût.

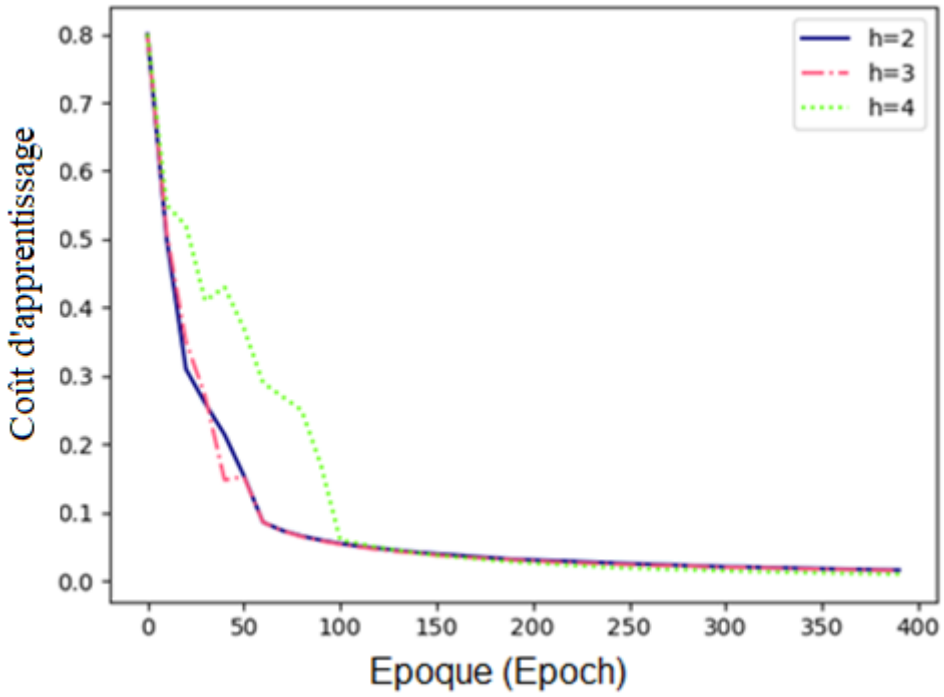


Figure 3. 9. Influence de la variation du nombre de couche cachées *h* sur la convergence de la fonction coût d'apprentissage

En plus, pour optimiser la phase d'apprentissage du réseau, on doit aussi minimiser le temps de prétraitement, noté *Tp*, et le temps de la fin de l'apprentissage, noté *Tf*, tout en garantissant la plus haute valeur possible du taux de détection *Dr* (detection rate), qu'on va définir en détail dans l'équation (3.1).

Pour ceci, on a présenté sur la Tableau 3.1 les résultats de nos expériences sur l'influence du nombre de couches cachées (h=2, 3 ou 4) sur le temps d'apprentissage *Tp*, le temps de réglage fin *Tf* et le Taux de détection *Dr*.

D'après les résultats obtenus dans la Figure 3.9 et le Tableau 3.1, on peut conclure que l'utilisation de deux couches cachées (h=2) seulement donne les valeurs les plus faibles de *Tp* et *Tf* tout en garantissant le taux de détection *Dr* le plus élevé.

Tableau 3. 1. Influence du nombre de couches cachées sur le temps d'apprentissage *Tp* , le temps de fin de l'apprentissage *Tf* et le Taux de détection *Dr*

Nombre de couches cachées	h=2	h=3	h=4
Tp	0.511 s	1.27 s	4.34 s
Tf	0.528 s	12.43 s	29.18 s
Taux de détection *Dr*	96%	90%	87%

En plus, nous remarquons, d'après la figure 3.9, que la fonction coût de prétraitement commence à se stabiliser à partir de 400 époques. Par conséquent, nous avons choisi de réaliser le prétraitement du réseau avec pendant seulement 400 époques (l'ensemble de données sera divisé en 400 sous-ensembles, et les poids seront mis à jour 400 fois au maximum).

La fonction coût peut être évaluée pour chaque époque lors de la phase d'apprentissage et de la phase de validation. Elle donne des indications sur l'efficacité du système d'apprentissage et sur son niveau d'entrainement ce qui influe sur sa capacité de généralisation. La fonction coût d'apprentissage est définie par la somme des erreurs commises pour chaque exemple dans les ensembles d'apprentissage ou de validation.
La variation de la fonction coût d'apprentissage et de validation en fonction de l'époque est illustrée à la Figure 3.10.

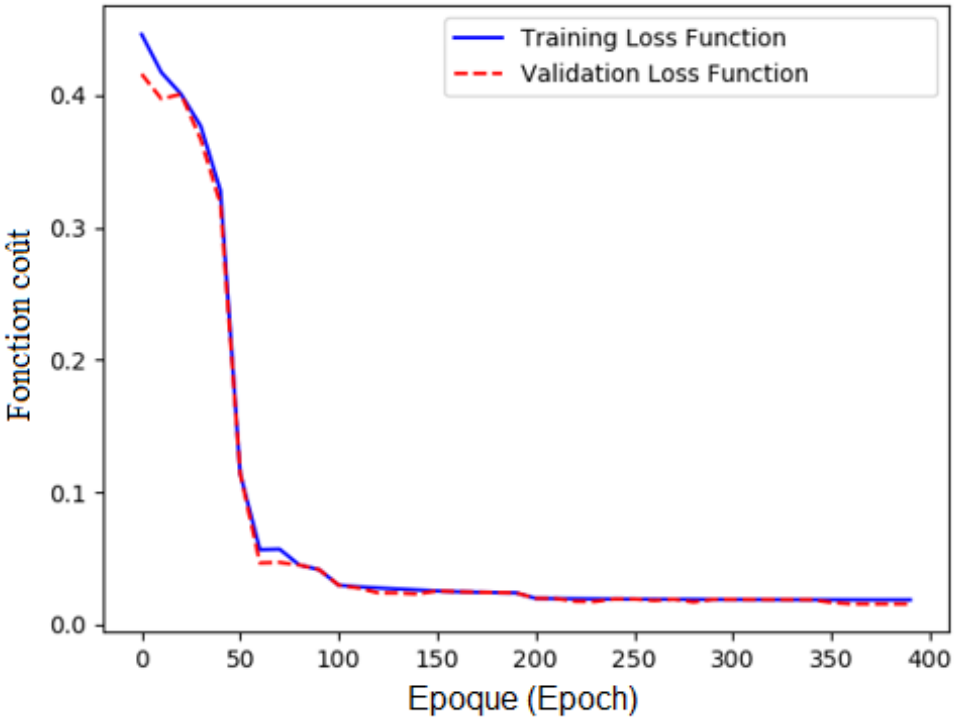


Figure 3. 10. Variation de la fonction coût d'apprentissage et de validation en fonction de l'époque

A la fin de cette partie, on présente la variation du Taux de détection (défini dans (3.1)) en fonction des itérations d'apprentissage, en utilisant les valeurs optimisées des paramètres du réseau et de l'apprentissage c. à. d. : 2 couches cachées, un coefficient d'apprentissage *Lr*=0.001 et l'algorithme d'optimisation d'ADAM.

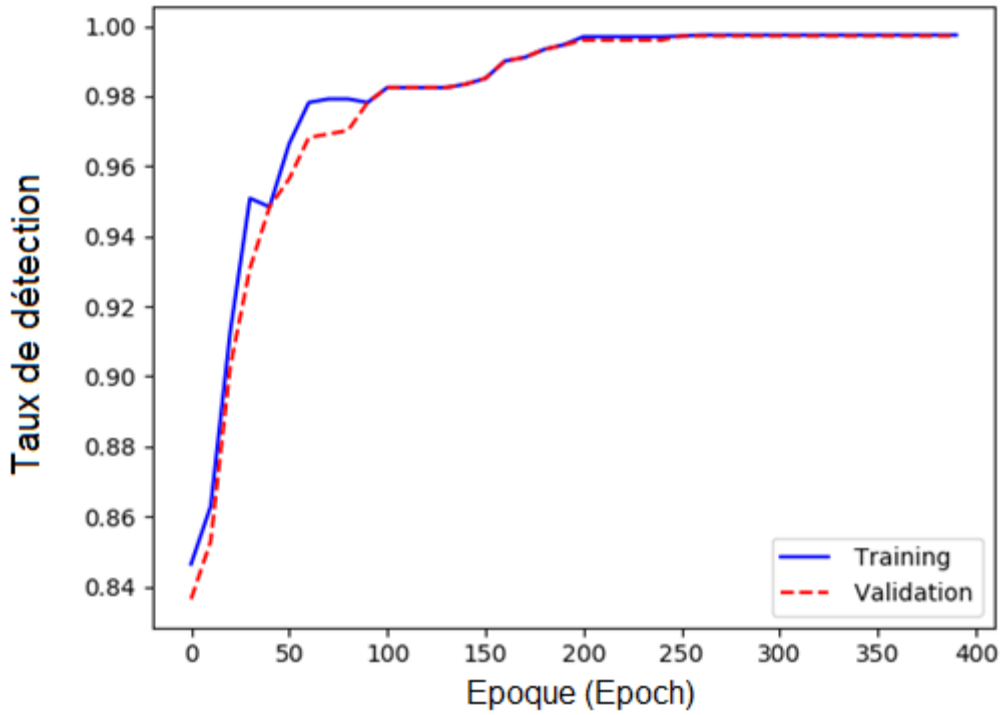


Figure 3. 4. Variation du taux de détection en fonction des itérations de l'apprentissage en phase d'apprentissage et de validation

III. Comparaison avec d'autres méthodes de détection de fumée

1. Présentation des résultats obtenus :

Afin de valider la méthode de détection de fumée proposée, on présente dans cette section des résultats expérimentaux de comparaison appliqués sur un ensemble de données (vidéos contenant de la fumée). Notre méthode proposée pour la détection de fumée (en utilisant les valeurs optimisées des paramètres du réseau et de l'apprentissage) a été comparée, sur la même base de données, à celle proposée dans Wang Yuanbin [Yua16] qui a été détaillé dans l'étude bibliographique du chapitre I de ce rapport. En effet, dans [Yua16], les caractéristiques statiques et dynamiques ont été combinés et insérés dans un vecteur multidimensionnel qui représente l'entrée du classificateur SVM pour classifier les régions de fumée et de non fumée.

L'expérience de détection de fumée a été réalisé sur un ensemble de 12 vidéos contenant de la fumée, du brouillard ou des nuages. Afin de diversifier la base de données pour bien tester l'efficacité et la robustesse des méthodes testées, les vidéos testées sont choisies comme suit

: Les six premières vidéos testées (de la vidéo 1 à la vidéo 6) contiennent toutes de la fumée. Les trois vidéos suivantes (de la vidéo 7 à la vidéo 9) contiennent des nuages, du brouillard et des personnes en mouvement. Les trois dernières vidéos testées (de la vidéo 10 à la vidéo 12) ne contiennent pas de fumée.

Dans le Tableau 3.2, nous avons présenté les résultats de cette comparaison en donnant, pour les 12 vidéos testées, les valeurs des 6 critères classiques de performance qu'on va détailler dans le paragraphe suivant : Le taux de détection, la précision, le rappel, le F1Score, et « l'intersection sur l'union » : IoU. On a aussi comparé le temps d'exécution de l'apprentissage Tf.

Le résultat de chaque trame testée est compté parmi un des quatre résultats suivants :

- **VP (Vrai Positif)** : cas où la détection est positive, sachant que la valeur réelle est effectivement positive. L'image contient de la fumée, et il s'agit de la présence de fumée réelle.
- **VN (Vrai Négatif)** : Cas où la détection est négative, et la valeur réelle est effectivement négative. L'image ne contient pas de la fumée, et effectivement il n'y a pas de fumée dans l'image.
- **FP (Faux Positif)** : Cas où la détection est positive, mais la valeur réelle est négative. Le test indique qu'il y a une fumée dans l'image, mais réellement il n'y a pas de fumée.
- **FN (Faux Négatif)** : les cas où la détection est négative, mais la valeur réelle est positive. Le test indique qu'il n'y a de fumée dans l'image, mais réellement il y a de la fumée.

En fonction de ces quatre quantités, on peut définir les valeurs des 4 critères classiques de performance de la classification comme suit :

$$\text{Taux de détection} = \frac{VP + VN}{VP + VN + FP + FN} \tag{3.1}$$

$$\text{Précision} = \frac{VP}{VP + FP} \tag{3.2}$$

$$\text{Rappel} = \frac{VP}{VP + FN} \tag{3.3}$$

$$\text{F1 Score} = 2.\frac{\text{Précision.Rappel}}{\text{Précision+Rappel}} \tag{3.4}$$

Tableau 3. 2. Comparaison des métriques de classification par la méthode proposée par Wang Yuanbin [Yua16] (M1) et notre méthode (M2)

Vidéo testée	Nb de trames testées	VP		VN		FP		FN		Taux de Détection %		Précision %		Rappel %		F1 Score		Tf	*IoU*
		M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M2	M2
Vidéo 1	275	210	225	14	23	26	14	28	17	82	90	89	94	88	93	0.88	0.93	0.6	0.85
Vidéo 2	483	439	457	15	7	23	4	28	19	94	96	95	99	94	96	0.94	0.97	0.68	0.92
Vidéo 3	123	104	109	6	3	6	3	7	6	89	91	94	97	93	94	0.93	0.95	0.43	0.87
Vidéo 4	888	784	829	32	5	32	16	68	34	92	94	96	98	92	96	0.94	0.97	0.82	0.91
Vidéo 5	808	731	762	20	14	30	7	38	29	93	96	96	99	95	96	0.95	0.97	0.76	0.94
Vidéo 6	466	408	435	12	8	21	4	31	20	90	95	95	99	93	95.5	0.94	0.97	0.65	0.93
Vidéo 7	**626**	**536**	**560**	**27**	**10**	**59**	**48**	**11**	**5**	**90**	**91**	**90**	**92**	**98**	**99**	**0.93**	**0.95**	**0.71**	**0.88**
Vidéo 8	**214**	**160**	**174**	**30**	**19**	**17**	**16**	**8**	**5**	**89**	**90**	**90**	**91.5**	**95**	**97**	**0.92**	**0.94**	**0.44**	**0.87**
Vidéo 9	**507**	**400**	**430**	**36**	**21**	**54**	**45**	**21**	**18**	**86**	**90**	**88**	**91**	**95**	**96**	**0.91**	**0.93**	**0.68**	**0.89**
Vidéo 10	304	0	0	291	298	15	6	-	-	95	98	0	0	0	0	-	-	0.48	-
Vidéo 11	109	0	0	102	104	7	5	-	-	93	95	0	0	0	0	-	-	0.31	-
Vidéo 12	84	0	0	79	81	5	3	-	-	94	96	0	0	0	0	-	-	0.29	-

En plus des critères de classification présentés, un autre critère pour tester les performances de localisation de la fumée a été aussi étudié. Dans les résultats de détection de la fumée, il n'est pas toujours certain que les coordonnées spatiales du cadre qui délimite ma zone détectée de la fumée correspondent exactement aux coordonnées du cadre réel de la « vérité terrain ».

Pour cette raison, nous avons utilisé une métrique d'évaluation qui quantifie la différence entre les deux cadres en fonction du chevauchement mutuel. C'est le critère IoU : l'intersection sur l'union [Rah16], dont un exemple simple est donnée sur la figure 3.12.

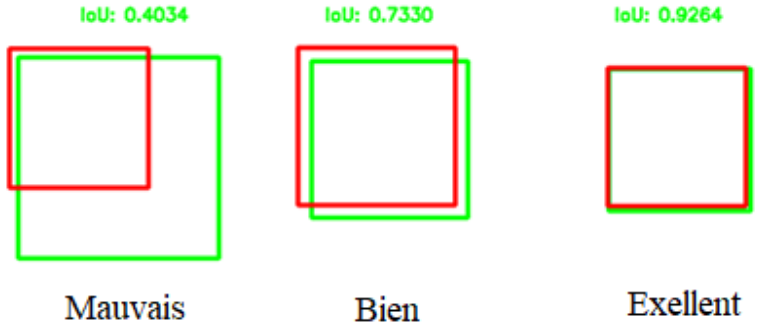


Figure 3. 5. Exemple d'évaluation de localisation par la métrique IoU

Il s'agit du rapport de la surface de la zone de chevauchement entre le cadre détecté de la fumée (par la méthode de détection) (Spred) et la vérité terrain (Strue) divisée par la surface de la zone de leur Union, défini comme suit :

$$IoU = \frac{Strue \cap Spred}{Strue \cup Spred} \tag{3.5}$$

L'IoU varie de 0% s'il n'y a aucun chevauchement et à 100 % si les deux cadres ont parfaitement la même position. Les cadres des zones détectées par la méthode de détection qui se chevauchent fortement avec les cadres de vérité terrain ont des scores IoU plus élevés que celles avec moins de chevauchement. Cela fait de l'intersection sur Union une métrique efficace pour évaluer les détecteurs d'objets personnalisés. Les valeurs trouvées du critère IoU correspondants à la méthode de détection proposée sont donné sur la dernière colonne du tableau des résultats de la Figure 3.2. Ceux de la méthode de [Yua16] n'a pas été donnée car ils n'ont pas réalisé la localisation de la fumée avec des cadres dans leurs travaux.

2. Interprétation des résultats du Tableau 3.2:

Les résultats expérimentaux montrent que, pour la plut part des tests réalisés, la méthode que nous proposons permet d'obtenir un taux de détection, une précision, un Rappel et un F1 Score supérieur à ceux correspondants au résultats de Yuanbin [Yua16]. Elle a généralement le plus grand nombre de VP et le plus petit nombre de VN.

On remarque que pour les 6 premiers vidéos testées (de la vidéo 1 à la vidéo 6) qui contiennent de la fumée, on a obtenu des valeurs de détection de fumée assez élevées, allant de 90 à 96 % en utilisant la méthode proposée (M2), tandis que la précision de la méthode (M1) est comprise entre [82-94%]. Pour les trois vidéos (de la vidéo 7 à la vidéo 9) contenant de la fumée, des personnes en mouvement ou des nuages, nous avons constaté que la présence d'objets entraîne une diminution du taux de détection (valeurs ente 89% et 91%). Pour les trois dernières vidéos testées, qui ne contiennent pas de la fumée, les deux méthodes M1 et M2 ne détectent pas la fumée et les valeurs de détection sont alors nulles. Dans ce cas, la précision et le taux de détection dépendent essentiellement des résultats VN et FP en absence de fumée dans les vidéos. La présence de contraintes telles que des personnes en mouvement, du brouillard ou des nuages augmente le nombre de résultats FN.

Par ailleurs, le temps de traitement des vidéos contenant de la fumée varie entre [0,43 s et 0,82 s], mais il est moins (presque la moitié) dans le cas des vidéos qui ne contiennent pas de la fumée.

3. Comparaison entre notre approche et la méthode de Pundir et al [Pun19] basée sur le Deep CNN :

En sachant que la méthode Deep CNN utilisée par Pundir et al [Pun19], et présenté dans le chapitre I de ce rapport, est actuellement une des meilleures méthodes pour la détection de fumée, et même si notre méthode n'arrive pas toujours à améliorer les critères de bonne détection de la méthode de [Pun19], notre originalité est qu'on a proposé une alternative pour la détection de fumée avec une architecture plus simple que celle de [Pun19] (donc plus rapide) et aussi nécessitant moins de ressources.

Comme nous l'avons résumé au chapitre I de ce rapport, nous rappelons que la méthode basée sur Deep CNN permet d'extraire les caractéristiques de la fumée en utilisant deux cadres d'apprentissage profond et tente de classer les régions fumées et non fumées. Le premier cadre d'apprentissage profond utilise un algorithme de super-pixel pour extraire les caractéristiques de la fumée (couleur, texture, désordre, bord tranchant, etc.). Le second

cadre d'apprentissage profond est utilisé pour extraire les caractéristiques de mouvement de la fumée après avoir appliqué le flux optique pour capturer le mouvement chaotique de la fumée. Ces caractéristiques extraites à l'aide du flux optique sont ensuite introduites dans Deep CNN pour extraire les caractéristiques basées sur le mouvement.

Contrairement à cette méthode, dans notre proposition de détection de la fumée basée sur un Deep Belief Network amélioré, nous insérons uniquement le vecteur de caractéristique choisi pour la classification en utilisant le DBN afin d'obtenir directement les probabilités de fumée et de non fumée. La méthode développée par Pundir et al. [Pun19] est basée sur l'extraction des caractéristiques de la fumée par un réseau double CNN après avoir appliqué le super-pixel sur les trames et la méthode de flux optique, et enfin avoir inséré le SVM comme classificateur. Cette méthode nécessite une grande quantité de ressources pour obtenir de bons résultats de classification et le processus d'apprentissage prend beaucoup de temps, surtout si l'ordinateur n'est pas équipé d'un GPU performant. D'autre part, la méthode de détection de la fumée proposée ne nécessite pas autant de données et elle permet aussi de localiser la fumée et de classer les zones de fumée et de non fumée.

4. Robustesse de la méthode proposée en présence de bruit additif :

On a remarqué que la méthode de détection proposée est parfois influencée par la présence de bruit et perturbations sous forme de nuages et de personnes en mouvement, ce qui entraîne une légère diminution du taux de détection. Par conséquent, nous allons tester dans cette dernière section la robustesse de cette méthode en présence de bruit ajouté aux trames d'entrées pendant l'apprentissage, ce qui entraine une dégradation de la qualité des images traitées. En effet, on teste uniquement l'influence du bruit additif aléatoire de haute fréquence (qui peut être réellement la présence de la poussière ou du brouillard) alors que les bruits généralement peuvent être convolutifs de basse fréquence (comme le cas du floue) ou bien multiplicatif (comme le cas de grain de speckle) [Ref04].

En plus, l'ajout d'un bruit aux images d'entrées peut être considéré comme une augmentation de la base de données avec des images bruitées. De cette façon, on peut augmenter à la fois la base de données et tester la robustesse du notre modèle en ajoutant un bruit. On a en plus gardé les mêmes paramètres du réseaux DBN et de l'apprentissage que dans les expériences précédentes.

Les résultats du taux de détection dans le cas d'ajout de bruit additif gaussien de moyenne nulle sont présentés dans le Tableau 3.3. Le niveau de bruit est quantifié par la valeur du rapport signal à bruit (SNR : Signal to noise ratio) [Nee16], calculé avec :

$$SNR = 20 .\log_{10}\left[\frac{\text{var(image à l'entrée)}}{\text{var(bruit ajouté)}}\right] \quad (3.7)$$

Trois niveaux de bruit ont été considérés : faible bruit (SNR = 20 dB), bruit moyen (SNR = 5 dB) et bruit élevé (SNR = 1 dB).

Tableau 3. 3. Variation du taux de détection en fonction de différents niveaux du bruit gaussien ajouté

SNR	Taux de détection
1 dB	87.5%
5 dB	93%
20 dB	94.5%

Enfin, pour étudier la différence entre les techniques de classification utilisées pour la détection de la fumée (SVM, Deep CNN et DBN), mentionnés et détaillés dans le premier chapitre de ce rapport, nous appliquons ces trois techniques sur des vidéos contenant de la fumée avec un bruit gaussien additif de SNR=5 dB (sans déplacement de personnes ni brouillard) et nous calculons les critères classiques : Taux de détection, précision, rappel, F1 Score F1, définis dans les équation (3.1) à (3.4).

Tableau 3. 4. Comparaison entre les différents classificateurs utilisés sans et avec ajout d'un bruit de SNR 5 dB

Condition	Classificateur	Taux de détection	F1 Score	Précision	Rappel
Sans Bruit	SVM	0.93	0.95	1	0.91
	Deep CNN	0.97	0.98	1	0.96
	DBN	0.96	0.97	1	0.95
Avec Bruit Gaussien, SNR 5 dB	SVM	0.91	0.95	0.91	1
	Deep CNN	0.95	0.97	0.94	1
	DBN	0.93	0.96	0.92	1

Les résultats de cette expérience sont présentés dans le Tableau 3.4. Les valeurs obtenues des critères de comparaison montrent que la méthode proposée qui utilise le DBN garantit des performances de classification comparables avec les valeurs trouvées avec le classifieur Deep CNN, et légèrement meilleures que celles trouvé avec le SVM, ce qui confirme la robustesse de la technique de classification adopté, même en présence de brui additif.

On note que malgré que le classifieur Deep CNN est légèrement plus robuste en présence de bruit additif, la méthode basée sur le DBN que nous proposons est plus simple à implémenter, d'architecture moins complexe, et donc plus rapide en implémentation. En plus, c'est une approche semi-supervisée dans laquelle nous exploitons la capacité du DBN à obtenir une représentation hiérarchique de haut niveau des données testées.

IV. Conclusion :

Les avantages de la méthode proposée par rapport aux méthodes classiques de détection de fumée peuvent se résumer en deux points. Le premier à limiter les régions d'intérêts après l'extraction des caractéristiques telles que le mouvement chaotique de la fumée, la couleur et le calcul du rapport énergétique. Toutes ces caractéristiques ont permis d'avoir un vecteur de caractéristiques qui alimente le réseau DBN de détection de fumée.

La méthode proposée assure des métriques intéressantes avec un temps de calcul minimal (taux de détection, score de précision F1, précision, rappel) et la possibilité de localiser les régions de fumée avec une bonne précision en un temps de calcul faible par rapport aux autres méthodes.

Pour autant que nous sachions, le seul travail qui a utilisé le DBN comme méthode de détection de la fumée est celui de Pundir [Pun17]. La différence entre notre approche proposée et la méthode développée dans [Pun17] est que nous ne nous intéressons qu'aux caractéristiques de couleur, de mouvement et d'énergie. On a montré l'importance de l'utilisation de l'énergie comme une caractéristique pertinente de la fumée dans le cadre de la classification.

On note aussi que la différence entre notre méthode et la technique de classification Deep CNN [Pun19] est que, grâce à notre méthode, nous pouvons localiser facilement les régions de fumée avec un vecteur de caractéristiques introduit dans une architecture simple après avoir extrait les caractéristiques de la fumée : mouvement, couleur et analyse énergétique. Cet avantage nous permet évidemment de classer les zones de fumée plus facilement et plus rapidement que la classification Deep CNN. En plus, les modèles basés sur Deep CNN sont

généralement « très profonds », avec de nombreuses couches et inspirés de modèles avec un nombre élevé de classes. Nous pensons que c'est parfois couteux et disproportionné comme approche pour détecter la fumée dans les images ou les vidéos, où il n'y a que deux classes : fumée et non fumée. En plus, Deep CNN nécessite beaucoup de temps de calcul ainsi que beaucoup de données qui ne sont parfois pas facilement disponibles.

Enfin, on a vérifié la robustesse du modèle proposée en ajoutant un bruit gaussien en entrée et aussi en testant l'efficacité de la méthode de la méthode de détection dans le cas de vidéos d'entrées contenant des personnes en mouvement, des nuages et des brouillards.

Conclusion générale et perspectives

Conclusion générale et perspectives

Cette thèse présente nos travaux sur la détection et la classification des régions de fumées des incendies dans des images aériennes. Après avoir implanté et évalué des méthodes classiques de détection et classification de la fumée, nous avons proposé différentes techniques visant à prendre en compte les spécificités de notre problématique. Nous avons développé une approche basée sur l'extraction des régions d'intérêts en utilisant Mask R-CNN et Yolo. Toutefois, ces techniques de détection d'objets présentent des inconvénients tels que la détection non exacte et la nécessité d'avoir une grande base de données de volume assez important. Pour remédier à ces problèmes, nous avons proposé une méthode reposant sur des caractéristiques des vidéos de fumées tels que : le mouvement, la couleur et l'énergie pour faciliter la détection des régions contenant la fumée. Dans un premier temps, nous avons détaillé des contributions présentée dans la littérature sur ce domaine. Dans un second temps, nous avons proposé des alternatives pour améliorer les techniques existantes.

Au cours de notre étude, nous avons conçu et mis en place une base de données permettant d'évaluer des algorithmes de détection sur des images extraites de vidéos. Cette base de données a été fournie avec à la fois des annotations riches, mais également un protocole de test rigoureux et un ensemble de performances d'algorithmes de référence, qui permettent à la communauté de proposer d'autres algorithmes et de les comparer efficacement. Comme enrichissement de cette base, nous avons ajouté un bruit aux entrées pour augmenter les données et renforcer la base de référence.

Une nouvelle technique de détection et de localisation de la fumée a été proposée en combinant différentes caractéristiques de la fumée : couleur, mouvement et énergie.

Dans le premier chapitre, nous avons exposé les méthodes traitées dans la littérature pour la détection de fumée afin d'effectuer une détection précoce de feux. La présentation de la méthodologie de la technique proposée a été présentée et décortiquée dans le deuxième chapitre. Contrairement aux méthodes précédentes, nous avons effectué un prés-traitement des images de notre base de données pour à la localisation des régions de fumée (et à la classification des régions de fumée et des régions sans fumée) à l'aide du DBN comme classifieur. En outre, le système proposé peut être utilisé pour la détection de la fumée en

temps réel. Le temps moyen de traitement des vidéos testées de la base de données varie entre 0,43 à 0,82 s, et dépend essentiellement du nombre total de trames, du nombre de couches cachées et des nœuds du réseau.

L'avantage de la méthode proposée par rapport aux autres méthodes de détection de la fumée est la localisation facile et rapide des régions de fumée, avec un IoU passant de 0,85 à 0,94 pour les vidéos contenant des régions de fumée, par comparaison aux méthodes de détection de la fumée classiques et à la CNN profonde qui nécessitent beaucoup de temps de calcul, et beaucoup plus de données, qui ne sont parfois pas facilement disponibles.

En outre, nous avons travaillé sur une base de données riche en ajoutant aux vidéos, des personnes en mouvement, des nuages, des brouillards et du bruit. Par rapport à la plupart des méthodes de détection de la fumée utilisant CNN, la méthode proposée généré des résultats avec des taux de détection les plus élevés, qui peuvent atteindre 96%, même dans un contexte bruité.

Dans les travaux futurs, nous pouvons proposer l'utilisation d'un double DBN pour améliorer le taux de détection et la localisation de la fumée dans chaque trame. Comme perspectives, nous utiliserons un optimiseur qui combine la rapidité d'exécution et la capacité de généralisation sur de nouvelles données comme le SGD. Une amélioration possible du système développé est l'incorporation d'AdaBound (Adaptive Gradient Methods with Dynamic Bound) pour améliorer substantiellement le taux de détection.

Références

[Bor08] Koerich Borges, P. Vinicius, J. Mayer and E. Izquierdo, Efficient visual fire detection applied for video retrieval, Signal Processing Conference, 16th European, IEEE, 2008.

[Cal08] Simone Calderara, P. Piccinini, R. Cucchiara, "Smoke detection in video surveillance: A mog model in the wavelet domain", 6th International Conference in Computer Vision Systems, ICVS 2008, pp. 119-128.

[Car05] Miguel Carreira-Perpinan," On contrastive divergence learning", Artificial Intelligence and Statistics, Canada, 2005, pp. 1-8.

[Cel07] Turgay 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.

[Cet13] Çetin Enis, "Video fire detection – Review," Digital Signal Processing, pp: 1827–1843, Juillet 2013.

[Cha06] Tung Chang, C.Chao,"Application of back-propagation networks in debris flow prediction", Eng. Geol. 2006, 85, 270–280.

[Chi02] Hsu Chih.Wei, C.Jen Lin, "A Comparison of Methods for Multiclass Support Vector Machines", IEEE Transactions on Neural Networks, vol. 13, no.2, March 2002, pp. 415 - 425.

[Cir13] Dan 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 2:, pp. 1237–1242, Novembre 2013.

[Des18] Soham De, Mukherjee.A; Ullah. E, "Convergence guarantees for RMS Prop and ADAM in non-convex optimization and their comparison to Nesterov acceleration on autoencoders", ICML Workshop on Modern Trends in Nonconvex Optimization for Machine Learning, Maryland,20 Novembre 2018.

[Efr75] Bradley Efron, "The efficiency of logistic regression compared to Normal Discriminant Analysis", Journal of the American Statistical Association, Vol. 70, No. 352, 1975, pp. 892- 898.

[Fav16] Margarita Favorskaya, A. Pyataeva, et A. Popov, “Spatio-temporal smoke clustering in outdoor scenes based on boosted random forestes”, 20th International Conference on Knowledge based and intelligent information and engineering systems, Grande Bretagne, vol. 93, 2016, pp. 762-771.

[Fel10] Pedro Felzenszwalb, B.Girshick, D.McAllester, D.Ramanan, Object Detection with Discriminatively Trained Part-Based Models, IEEE Transactions on Pattern Analysis and Machine Intelligence,Volume: 32, Issue: 9, Sept. 2010 .

[Fri16] Sébastien Frizzi, R. Kaabi, M. Bouchouicha, F. Fnaiech, E. Moreau :’Convolutional neural network for video fire and smoke detection’, IEEE Transactions IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, Florence,2016,10,(23),pp. 877-882.

[Fri17] Sébastien Frizzi, R.Kaabi, M.Bouchouicha, J-M.Ginoux, F.Fnaiech, E. Moreau, “ Détection de la fumée et du feu par réseau de neurones convolutifs,” Conférence Nationale d’Intelligence Artificielle Année, Caen, France, 2017, pp.45-51.

[Gir15] Ross Girshick, “Fast R-CNN.” Proceedings of the IEEE International Conference on Computer Vision. 2015, pp. 1440–1448.

[Gub09] Jayava Gubbi, S.Marusic, M.Palaniswami,” Smoke detection in video using wavelets and support vector machines,” Fire Safety Journal, Volume 44, Issue 8, November 2009, pp: 1110-1115.

[Hin02] Geoff Hinton,” Training products of experts by minimizing contrastive divergence”, Neural Computation, vol.14, 2002, pp. 1711-1800.

[Hin06] Geoff Hinton, S. Osindero, Y. Whye The. “A fast learning algorithm for deep belief nets”, Neural Computation, vol. 18, no.7, 2006, pp. 1527-1554.

[Hor81] Berthold Horn, K. Schunck, « Determining optical flow », Artificial. Intelligence, 1981, 17, 185–203.

[Hra15] Rafael Hrasko, A.G. C. Pacheco, and R. A. Krohlinga, “Time Series Prediction using Restricted Boltzmann Machines and Backpropagation”, Information Technology and Quantitative Management, Procedia Computer Science, vol. 55; 2015, pp. 990-999.

[Huy18] Yaocong Hu, X. Lu:’Real-time video fire smoke detection by utilizing spatial-temporal ConvNet features’, Multimedia Tools and Applications, 2018,11, pp. 29283–29301.

[Hua97] Chou Huang, “A Gibbs sampling approach to the estimation of linear regression models under daily price limits”, Pacific-Basin Finance Journal, vol.5, no.1, 1997, pp. 39-62.

[Hyu14] Kim Hyuntae, Daehyun. Ryu, et Jangsik. Park, Member, « Smoke detection using GMM and Adaboost, » International Journal of Computer and Communication Engineering, Vol. 3, No. 2, Mars 2014.

[Jen09] John Jensen, J.Im, P. Hardin, R. Jensen,” Chapter 19 | Image Classification,” The SAGE Handbook of Remote Sensing, 2009.

[Kho98] Alireza Khotanzad et C. Chung, “Application of Multi-Layer Perceptron Neural Networks to Vision Problems”, Neural Computing and Applications, vol.7, 1998, pp. 249-259.

[Kon07] Igor Kononenko et Kukar.M,” Artificial Neural Networks,” Machine Learning and Data Mining, 2007, pp: 275-320.

[Lap04] Ivan Laptev, T. Lindeberg Zhu, ”Local Descriptors for Spatio temporal Recognition,” International Workshop on Spatial Coherence for Visual Motion Analysis SCVMA 2004: Spatial Coherence for Visual Motion Analysis, pp 91-103.

[Lec98] Yann Lecun, Bottou.L, Bengio.Y., « Gradient-based learning applied to document recognition », Proc. IEEE 1998, 86, 2278–2324.

[Lec15] Yann Lecun, Y. Bengio, G.Hinton, “Deep Learning”, International Journal of Science Nature, vol. 521, 2015, pp. 436-444.

[Lee07] Byoungmoo Lee, et D. Han, “Real-time fire detection using camera sequence image in tunnel environment”, International Conference on Intelligent Computing, volume 4681, 2007, pp.1209-1220.

[Lie16] Wei Lie, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg. “SSD : Single Shot MultiBox Detector”, ECCV 2016

[Lou98] Stéphane Louvel,”Étude numérique d’un algorithme d’assimilation variationnel avec prise en compte de l’erreur de modèle, et comparaison avec une méthode de filtrage adaptatif”, Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science, Volume 326, Issue 3, Février 1998, pp: 173-179.

[Mar10] Hidenori Maruta, A. Nakamura, and F. Kurokawa, “A novel smoke detection method using support vector machine”, TENCON 2010, 2010 IEEE Region 10 Conference, vol. 431, 2010, pp. 210-215.

[Mck00] Stéphane McKennaa, S.Jabrib, Z. Duricb, A. Rosenfeldc, H.Wechslerb, “Tracking Groups of People”, Computer Vision and Image Understanding Volume 80, Issue 1, Octobre 2000, pp: 42-56.

[Mik04] Kristyan Mikolajczyk, T. TUYTELAARS, C. SCHMID “A Comparison of Affine Region Detectors,” International Journal of Computer Vision, 2004.

[Nan18] Zhang Nan, S. Ding, J. Zhang, and Y.Xue, “An overview on Restricted Boltzmann Machines”, Neurocomputing, vol. 275, Janvier, 2018, pp. 1186-1199.

[Nee16] Arvind Neelakantan, Vilnis. L, Le.Q, V.Sutskever, « Adding Gradient Noise Improves Learning for Very Deep Networks », 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016.

[Pun17] Arun Pundir, B. Raman “Deep belief network for smoke detection”, Fire Technology, 2017, vol. 53, no.6, pp. 1943-1960.

[Pun19] Arun Pundir, B.Raman, “Dual Deep Learning Model for Image Based Smoke Detection”, Fire Technology Issue 6/2019, 2019.

[Ray15] Mark Ryan,” Chapter 1 - Analytics Defined,” Information Security Analytics, Finding Security Insights, Patterns and Anomalies in Big Data 2015, pp: 1-12.

[Rah16] Atiqur Rahman,Yang Wang ,”Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation”, International Symposium on Visual Computing, 2016,pp. 234-244.

[Red16] Joseph Redmon, S.K. Divvala, R.B. Girshick, A. Farhadi,”You only look once: unified, real-time object detection”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016), pp. 779-788.

[Ref04] Réfrégier Paul, “Noise Theory and Application to Physics”, Edition Springer, New York, 2004.

[Say07] Sayadi.M, Tlig.L, Fnaiech.F, « A New Texture Segmentation Method Based on the Fuzzy C-Mean Algorithm and Statistical Features”, Applied Mathematical Sciences, Vol. 1, 2007, no. 60, pp.2999 – 3007.

[Sta00] Chris Staufer, E.Grimson, “Learning patterns of activity using real-time tracking,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, pp: 747 – 757.

[Ter19] Jean Terrail, Réseaux de neurones convolutionnels profonds pour la détection de petits véhicules en imagerie aérienne, Soumis 29 avril 2019, Normandie Université.

[Tli12] Tlig.L, Sayadi.M, Fnaiech.F, "A New Fuzzy segmentation approach based on S-FCM type 2 using LBP-GCO features", in Signal Processing: Image Communication, Elsevier, vol. 27, n. 6, pp. 694-708, July 2012.

[Tor06] Ugur Toreyin, Y. Dedeoglu, A.E. Cetin, “Contour Based Smoke Detection in Video Using Wavelets,” 14th European Signal Processing Conference, EUSIPCO 2006, Florence, Italie, 2006, pp. 1-5.

[Tsu14] Lin Tsung, Michael MaireSerge BelongieJames HaysPietro Perona,” Microsoft COCO: Common Objects in Context,” European Conference on Computer Vision, 2014, pp.740-755.

[Ver12] Steven Verstockt, “A multi-modal video analysis approach for car park fire detection,” Fire Safety Journal 57 (2013) pp, 44–57, Aout2012.

[Vij14] Sudheendra Vijayanarasimhan, K.Grauman, “Large-Scale Live Active Learning: Training Object Detectors with Crawled Data and Crowds, “International Journal of Computer Vision volume 108, pp. 97–114 (2014).

[Vio01] Paul Viola, M.Jones, “Rapid object detection using a boosted cascade of simple features,” Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001.

[Wit14] Petter Wittek, “Unsupervised Learning,” What Quantum Computing Means to Data Mining 2014, pp: 57-62.

[Xio2007] Ziou Xiong, R. Caballero, H. Wang, A.M. Finn, M.A. Lelic, and P.-Y. Peng, “Video-Based Smoke Detection: Possibilities, Techniques, and Challenges”, IFPA Fire Suppression & Detection Research & Applications, East Hartford, 2007, pp. 1-7.

[Yan14] Jianbo Yang, X. Yuan, X. Liao, P. Liul, G. Sapiro, and L. Cari, “Video Compressive Sensing Using Gaussian Mixture Models,” IEEE Transactions on Image Processing, 2014, pp. 4863-4878.

[Yin17] Zhijian Yin, Wan.B, J. A Deep Normalization and Convolutional Neural Network for Image Smoke Detection. IEEE Access 2017, 5, 18429–18438.

[Yua16] Yuan Wang, “Smoke Recognition Based on Machine Vision”, International Symposium on Computer, Consumer and Control, vol. 283, 2016, pp. 668-671.

[Yun20] Li Yundong, Han.DONG, Hongguang.LI, Xueyan.ZHANG, “Multi-block SSD based on small object detection for UAV railway scene surveillance,” Chinese Journal of Aeronautics, Volume 33, Issue 6, June 2020, Pages 1747-1755.

[Zen20] Nianyin Zeng, Li.H, Wang.Z, ”Deep-reinforcement-learning-based images segmentation for quantitative analysis of gold immunochromatographic strip, Neurocomputing, 21 Avril 2020.

[Zha15] Yaqin Zhao, Q.Li, and Z.Gu, “Early smoke detection of forest fire video using CS Adaboost algorithm”, Optik-Int. J. Light Electron Opt., vol. 126, no.19, 2015, pp. 2121-2124.

[Zou19] Fangyu Zou, Shen. L, Jie.Z, Zhang. Liu, « A Sufficient Condition for Convergences of Adam and RMSProp. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 Juin 2019; pp. 11119–11127.

Liste des publications de Rabeb Kaabi

Dans les journaux internationaux

[P1] R. Kaabi, M.Bouchouicha, A. Mouelhi, M.Sayadi, Eric Moreau., "An Efficient Smoke Detection Algorithm Based on Deep Belief Network Classifier Using Energy and Intensity Features,Electronics, vol. 9, 2020, pp. 1390-1411, 27 August 2020.

Dans les conférences internationales

[P2] R. Kaabi, M. Bouchouicha, M. Sayadi, E. Moreau, "Smoke Detection Using GMM and Deep Belief Network", The Fourth International Conference on Advances in Signal Image and Video Processing, Signal, 2019, 6, 6, Athènes, Grèce, pp.1-6.

[P3] R. Kaabi, M. Sayadi, M. Bouchouicha, F. Fnaiech, E. Moreau, and J-M. Ginoux, "Early smoke detection of forest wildfire video using deep belief network", 4th International Conference on Advanced Technologies for Signal and Image Processing, ATSIP 2018, 21-24 March 2018, Sousse, Tunisia.

[P4] R. Kaabi, S. Frizzi, M. Bouchouicha, F. Fnaich, and E. Moreau, "Video Smoke Detection Review", 2017 International Conference on Smart, Monitored and Controlled Cities, SM2C, 17-19 Feb. 2017, Sfax, Tunisia.

[P5] S.Frizzi, R. Kaabi, M. Bouchouicha, F. Fnaiech, E. Moreau, "Convolutional neural network for video fire and smoke detection", IEEE Transactions IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, 2016, 10, (23), pp. 877-882.

[P6] Frizzi.S, R.Kaabi, M.Bouchouicha, J-M.Ginoux, F.Fnaiech, E. Moreau, " Détection de la fumée et du feu par réseau de neurones convolutifs," Conférence Nationale d'Intelligence Artificielle, Caen, France, 2017, pp.45-51.