



HAL
open science

SEGMENTATION SÉMANTIQUE DE FEU DE FORÊT A L'AIDE D'UN UNET-RESNET50

Mohamed Tlig, Moez Bouchouicha, Mounir Sayadi, Eric Moreau

► **To cite this version:**

Mohamed Tlig, Moez Bouchouicha, Mounir Sayadi, Eric Moreau. SEGMENTATION SÉMANTIQUE DE FEU DE FORÊT A L'AIDE D'UN UNET-RESNET50. ORASIS 2023, Laboratoire LIS, UMR 7020, May 2023, Carqueiranne, France. ⟨hal-04219356⟩

HAL Id: hal-04219356

<https://hal.science/hal-04219356v1>

Submitted on 27 Sep 2023

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.



HAL Authorization

SEGMENTATION SÉMANTIQUE DE FEU DE FORÊT A L'AIDE D'UN UNET-RESNET50

Mohamed Tlig^{1,2}, Moez Bouchouicha², Mounir Sayadi¹ and Eric Moreau ²

- 1 Member of SIME Laboratory, ENSIT University of Tunis, Tunis 1008, Tunisia; mohamed-tlig@etud.univ-tln.fr (M.T.); mounirsayadi@yahoo.fr (M.S.)
- 2 Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France ; moez.bouchouicha@univ-tln.fr (M.B.); eric.moreau@univ-tln.fr (E.M.)

Résumé

Les feux de forêts représentent un risque important pour plusieurs pays dans le monde. Ils provoquent souvent des pertes humaines, économiques et écologiques. Étant donné que seules les images visuelles ou infrarouges (IR) ne peuvent fournir des informations précises et grâce aux avancées technologiques des systèmes de détection de feu basés sur la vidéo, la combinaison des images visuelles et infrarouges peut jouer un rôle essentiel dans la détection des feux. Il est donc précieux de fusionner ces deux types d'images, en combinant les avantages des informations détaillées sur la texture dans les images visibles et les informations sur le rayonnement thermique dans les images infrarouges. Dans cet article, nous sélectionnons deux techniques de fusion d'images de pointe pour les évaluer dans le cadre de la tâche spécifique de la fusion d'images de feux et nous comparons les performances des méthodes de fusion en se basant sur des métriques sélectionnées. Ensuite, nous utilisons le UNet-ResNet50 pré-entraîné pour l'étape de segmentation du feu. Enfin, les résultats expérimentaux montrent que l'étape de fusion proposée avant l'étape de segmentation sémantique conduit à de meilleurs résultats par rapport aux images visibles uniquement.

Mots Clef

Image visible, image infrarouge, feu de forêt, segmentation sémantique, critères d'évaluation.

Abstract

The forest fire represents a significant risk for many countries around the world. In this paper, first, we select two state-of-the-art image fusion techniques to evaluate for the specific task of fire image fusion and compare the performance of infrared-visible fusion methods based on selected metrics of

fusion performance. Second, we use the UNet with pre-trained ResNet50 for the fire segmentation stage.

Finally, experimental results show that the proposed fusion stage before the semantic segmentation stage leads to better results compared to visible images only.

Keyword

Visible image, infrared image, fire forest, semantic segmentation, evaluation criteria.

1. Introduction

Au cours des dernières années, le changement climatique et les facteurs d'origine humaines ont eu un impact significatif sur l'environnement [1]. Certains de ces facteurs comprennent les vagues de chaleur, les ouragans, la poussière, les sécheresses, les inondations, les tempêtes et les feux de forêts [2,3]. Les feux de forêts sont l'un des principaux éléments responsables de dommages de vies humaines, de pertes économiques et de catastrophes environnementales majeures qui prendraient plusieurs décennies à surmonter. Selon les informations publiées par l'Université du Maryland, publiées sur Global Forest Watch, les tropiques ont perdu 12,2 millions d'hectares de couvert arboré en 2020 [4]. Les émissions de carbone résultant de cette perte de forêt primaire (2,64 Gt de dioxyde de carbone (CO₂)) équivalent aux émissions annuelles de 570 millions de véhicules, soit plus du double des véhicules en circulation aux États-Unis [5]. La surveillance et la détection des feux de forêt sont devenues une solution essentielle pour la protection de la vie humaine et la conservation des ressources naturelles, qui suscite un intérêt croissant dans le monde entier. Diverses techniques de détection des feux sont apparues aujourd'hui pour aider à surmonter les effets catastrophiques de ce phénomène. Traditionnellement, des observateurs humains

ou des tours de guet étaient utilisés pour la détection des feux de forêt [6].

De nos jours, les chercheurs se sont tournés vers une solution encourageante pour un système de surveillance des incendies comme les caméras optiques, les systèmes satellitaires et les réseaux de capteurs sans fil [6,7]. Les techniques de détection d'incendie basées sur la vision peuvent être divisées en systèmes de détection d'incendie visibles ou infrarouges en fonction de la gamme spectrale des caméras utilisées [8]. Cependant, l'utilisation de caméras visibles pendant la nuit donne une mauvaise qualité d'image. Généralement, il y a un risque de fausse alarme incendie dans certaines situations où il y a une forte probabilité d'un fond enfumé qui pourrait cacher les flammes et où les feuilles de couleur rouge deviennent prédominantes dans la forêt. De plus, la fumée est transparente pour les caméras infrarouges et les objets de couleur rouge ne présentant aucune menace pour la détection des incendies car ils ne sont pas la source de chaleur [9]

Par conséquent, nous pouvons obtenir l'image infrarouge dans des conditions de faible luminosité ou d'absence de lumière [10]. Une autre solution importante consiste à fusionner les images infrarouges et visibles en intégrant les caractéristiques des images infrarouge et visibles [6]. Ainsi, une nouvelle approche consistant à fusionner des images infrarouges et visibles peut être plus bénéfique pour réduire les fausses alarmes incendie et extraire de nombreuses informations liées aux feux de forêt. Plusieurs chercheurs se sont concentrés sur le développement d'approches de fusion pour les images visibles et infrarouge [11,12].

La fusion des images visibles et infrarouges peut être bénéfique pour améliorer la précision, la robustesse et la fiabilité des systèmes de détection des feux. Pour évaluer l'applicabilité de certaines approches de la fusion infrarouge-visible à la fusion d'images de feux de forêt, nous avons choisi deux méthodes de pointe et les avons évaluées sur la tâche particulière de la fusion d'images de feux. Les deux techniques mentionnées sont les suivantes : la méthode proposée par Hui Li, et al [11] présente une inclusion plus équilibrée des informations texturales et thermiques et elle est capable d'extraire de manière robuste les caractéristiques saillantes des données corrompues. Une autre méthode, introduite par Zhiqiang Zhou et al. [12], permet de bien transférer les informations importantes d'image infrarouge dans l'image fusionnée. De plus, elle préserve simultanément les détails et le paysage de fond dans l'image visible d'entrée.

Nous évaluons et comparons ces méthodes sur la base de données « Corsican Fire » [13] en utilisant les métriques les plus utilisées pour évaluer la qualité des images fusionnées. Selon la revue de la littérature sur la fusion d'images visibles et infrarouges, nous avons trouvé d'autres métriques utilisées dans certains articles mais ces métriques sont les plus utilisées [14,15]. Pour la tâche de segmentation, les réseaux d'encodeur-décodeur comme FCN [16], UNet [17], etc. sont généralement préférés aux autres approches.

Les premières couches de convolution de l'encodeur apprennent à détecter des caractéristiques de bas niveau, telles que des bords, des coins, des textures, etc. Ensuite, les

couches de convolution suivantes apprennent à combiner ces caractéristiques de bas niveau pour détecter des motifs de plus en plus complexes, tels que des formes géométriques, des objets, des parties d'objets, etc. Les couches de convolution plus profondes apprennent finalement à capturer des caractéristiques sémantiques de haut niveau, telles que des concepts abstraits comme des relations spatiales ou des actions.

En utilisant ces caractéristiques sémantiques apprises, le décodeur peut reconstruire l'image d'origine à partir de la représentation de bas niveau en utilisant des couches de déconvolution. Les connexions sautées (skip connections) entre les couches de convolution de l'encodeur et les couches de déconvolution du décodeur permettent de combiner les informations de bas niveau avec les informations de haut niveau pour obtenir une meilleure reconstruction de l'image. L'ensemble de ce processus d'apprentissage hiérarchique de caractéristiques sémantiques allant du bas niveau au haut niveau est ce qui rend les réseaux d'encodeur-décodeur très performants dans de nombreuses tâches de vision par ordinateur, y compris la segmentation d'image, la détection d'objets, la reconnaissance d'actions et bien plus encore

L'encodeur peut être remplacé par un réseau pré-entraîné tel que VGG16 [18], VGG19 [18], etc. Ces réseaux pré-entraînés sont déjà entraînés sur le jeu de données ImageNet [19] et possèdent les capacités d'extraction de caractéristiques nécessaires. Avec le succès du réseau résiduel [20], ResNet50 est l'une des architectures les plus utilisées pour toute tâche d'apprentissage par transfert. La structure de cet article est la suivante : la section deux présente les deux méthodes comparées, les critères d'évaluation, la base de données utilisés, les résultats obtenus des critères d'évaluation calculés à partir du modèle de fusion sélectionné. La section trois fournit les résultats des expériences de segmentation sémantique des feux utilisant UNET-RESNET50. Enfin, la section 4 présente la conclusion et les travaux futurs.

2. Méthodes de fusion, critères d'évaluation et données

Dans cette section, nous présenterons dans un premier temps les deux méthodes retenues : la fusion d'images par Latent Low-Rank Representation (LatLRR) proposée par Hui et al. [11]. Dans cette méthode, la décomposition multi-échelle des images sources améliore la robustesse au bruit et aux valeurs aberrantes. La méthode de fusion d'images infrarouges et visibles « guided filter-based context enhancement » (GFCE) donnée dans [12], Elle permet d'injecter les caractéristiques spectrales infrarouges les plus pertinentes dans l'image visible. Ceci est réalisé en préservant ou en améliorant correctement les repères perceptuels importants du paysage d'arrière-plan et les détails de l'image visible. Dans un second temps, nous introduisons l'ensemble des critères d'évaluation

considérés. Enfin, la base de données des feux de Corse présentée par Toulouse et al [13] est introduite.

2.1 Méthode de fusion d'images évaluées

2.1.1 Technique basée sur LatLRR

Cette méthode, proposée par Hui et al. [11]. Dans ce travail, les auteurs proposent une méthode de fusion d'images basée sur LatLRR qui est simple et robuste.

En utilisant cette méthode, la fusion d'images visibles et infrarouges peut donc être réalisée avec une meilleure qualité d'image et une meilleure efficacité de calcul. De plus, cette technique peut être adaptée à différents types d'images et peut être utilisée dans diverses applications de vision par ordinateur, telles que la surveillance de la sécurité et la reconnaissance de forme.

2.1.2 Technique basée sur GFCE

Cette méthode, est proposée par Zhiqiang Zhou et al [12]. Dans ce travail, les auteurs proposent un nouvel algorithme d'amélioration du contexte de vision nocturne par la fusion d'images infrarouges et visibles avec le filtre guidé. Premièrement, une méthode d'amélioration adaptative est développée. Ensuite, une décomposition multi-échelle hybride basée sur le filtre guidé est introduite pour injecter les informations de l'image infrarouge dans l'image visible par une approche de fusion multi-échelle. Enfin, en comparant la saillance perceptive des informations de l'image infrarouge et visible, une méthode de sélection des paramètres de régularisation basée sur la perception est utilisée pour déterminer la quantité relative des caractéristiques spectrales infrarouges injectées.

2.2 Critère d'évaluation

L'évaluation objective consiste à créer un modèle mathématique correspondant en simulant le système de perception visuelle humain et à comparer les performances de chaque méthode de fusion à l'aide de données statistiques. Nous pouvons classer les méthodes d'évaluation de la fusion d'images en deux types : la méthode basée sur les caractéristiques et la méthode basée sur les statistiques [21]. La méthode basée sur les caractéristiques utilise des mesures de qualité des caractéristiques de l'image telles que l'information mutuelle et l'entropie croisée. La méthode basée sur les statistiques utilise des mesures de qualité standard telles que l'entropie, le SNR et l'écart type. Les indicateurs d'évaluation objectifs sont énumérés dans le tableau 1 ci-dessous.

Tableau 1. Statistiques de quelques critères d'évaluation et références. "+" signifie qu'une grande valeur indique de bonnes performances, tandis que "-" signifie qu'une petite valeur indique de bonnes performances.

Criteres d'évaluation	Reference	+/-
Entropy (EN)	[22]	+
Correlation coefficient (CC)	[23]	+
Structural similarity index measure (SSIM)	[24]	+
Standard deviation (SD)	[25]	+
Mean Squared Error (MSE)	[26]	-
Peak signal-to noise ratio (Psnr)	[27]	+

2.3 Base de données

Dans ce travail, nous utilisons les paires d'images visibles et infrarouges de la base de données de Corse. Diverses expériences ont été menées pour analyser les performances de l'approche de fusion proposée. Pour évaluer le potentiel de l'approche présentée, nous avons pris une base de données de 640 paires d'images provenant de la base de données de Corse présentée par Toulouse et al [13]. Cette base est constituée d'un ensemble de triplets d'images. Pour chaque image dans le domaine visible, nous disposons de sa vérité terrain, établie manuellement par un expert, et de l'image acquise simultanément dans l'infrarouge. La figure 1 montre deux triplets d'images issus de la base de données.

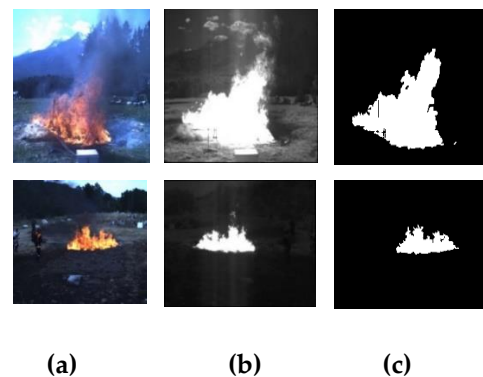


Figure 1 : Exemples d'images (a) image prise dans le spectre visible ; (b) image prise dans le spectre infrarouge ; (c) vérité terrain

2.4 Résultats expérimentaux et discussions

Dans cette section, nous analysons la performance de l'approche de fusion sélectionnée dans plusieurs expériences. La totalité de la base de données de 640 paires d'images sont tirée de la base de données de Corse afin d'évaluer le potentiel de l'approche présentée. Deux d'entre elles sont également

représentées sur la figure 2 (a) et (b). Les images fusionnées sont présentées dans les Figures. 2(c), 2(d), d'où il ressort clairement que les images de feux fusionnées sont approximativement similaires aux images visibles de la source en incorporant les caractéristiques des images infrarouges et visibles. Une comparaison détaillée des résultats de fusion obtenus par les méthodes sélectionnées. Pour évaluer les résultats des différentes approches de fusion, les indices de performance sont comparés.

Pour calculer la performance quantitative des approches de fusion, 640 paires d'images de feu visibles et infrarouge sont sélectionnées pour présenter l'analyse comparative des résultats. Les images sources visibles et infrarouges sont présentées dans la Figure. 2(a) et (b), respectivement. Les images fusionnées correspondantes obtenues par les deux méthodes sont présentées dans les figures 2(c)-(d), respectivement. La performance objective des différentes méthodes de fusion en utilisant les six métriques est listée dans le tableau 2. Pour chaque métrique, le score moyen de chaque méthode sur la base de données est calculé. La valeur indiquée en gras indique le meilleur score parmi toutes les méthodes. La méthode basée sur LatLRR obtient la première place pour les métriques CC, PSNR, SSIM et MSE avec des avantages évidents. Pour les deux autres métriques STD et ENT, la méthode basée sur LatLRR obtient la deuxième place, mais l'écart est relativement faible par rapport à sa supériorité sur les quatre autres métriques. Dans l'ensemble, la méthode basée sur LatLRR réalise la meilleure performance objective parmi les deux méthodes proposées.

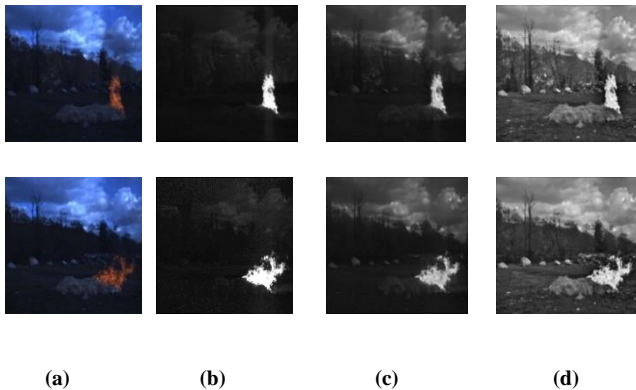


Figure 2. Performances de méthode de fusion (a) images visibles, (b) images infrarouges, (c) méthode LatLRR (d) méthode GFCE

Tableau 2; Mesures de performances de fusion comparatives moyennes calculées à partir du modèle de fusion sélectionné pour toutes les bases de données

Performance Indices	Visible	IR	LatLRR	GFCE
STD	46.73	48.23	42.31	47.18
ENT	6.72	6.12	6.86	7.45
CC	-	-	0.79	0.71
PSNR	-	-	17.96	13.23
SSIM	-	-	0.85	0.60
MSE	-	-	1.29e+03	3.84e+03

3. Méthode de segmentation sémantique et critères d'évaluation

3.1 UNet-ResNet50

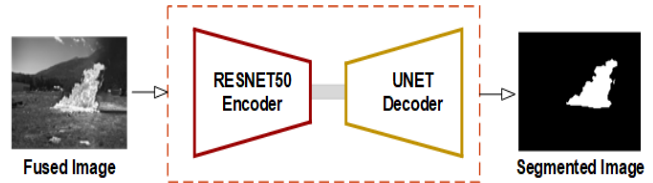


Figure 3 ; L'architecture U-Net-ResNet50

Dans la deuxième étape de la segmentation sémantique du feu, la figure 3 montre un aperçu de l'architecture UNet-ResNet50. Il s'agit d'une architecture basée sur un encodeur-décodeur, où ResNet50 entraîné sur le jeu de données ImageNet [19]. Dans cette section, nous analysons les performances de l'approche de segmentation sélectionnée dans plusieurs expériences. Pour calculer la performance quantitative de l'approche de segmentation. Le modèle UNet-ResNet50 a été choisi pour sa capacité à apprendre et à extraire efficacement les caractéristiques des images, ce qui est crucial pour la segmentation de feux. L'architecture du modèle est une combinaison d'un UNet comme décodeur et ResNet50 comme encodeur, ce qui permet l'extraction de caractéristiques de haut niveau et de bas niveau.

3.2 Critères d'évaluation, résultats expérimentaux et discussions

L'une des questions les plus fréquemment posées après le développement d'une méthode de segmentation sémantique est " comment savoir si notre méthode fonctionne avec une "bonne" précision de segmentation ? ". Pour toutes les méthodes de segmentation d'images, l'évaluation de la métrique est un facteur clé pour évaluer quantitativement la performance de la classification. Dans la littérature, plusieurs métriques d'évaluation ont été proposées [28,29].

Dans cette section, nous analysons la performance de l'approche de segmentation sélectionnée dans plusieurs expériences. Pour calculer la performance quantitative de l'approche de segmentation, 640 paires d'images sont sélectionnées pour présenter l'analyse comparative des résultats. Les images sources visibles et les images de vérité-terrain sont présentées dans la Figure. 3(a) et (b), respectivement. Les images segmentées correspondantes obtenues sont présentées dans les figures 3(c)-(d)-(e), respectivement. Le modèle UNet-ResNet50 a été implémenté en utilisant la bibliothèque Keras avec TensorFlow comme backend. L'entraînement et les tests ont été effectués sur une machine équipée d'un GPU NVIDIA Quadro P620 et de 16 Go de RAM. 80 % des images ont été utilisées pour

l'entraînement, 10 % pour la validation et 10 % pour les tests. Le modèle a été entraîné avec une taille de lot de 16 et pendant 150 époques. L'optimiseur Adam a été utilisé avec un taux d'apprentissage de 0,0001. La performance objective de la méthode de segmentation en utilisant le métrique (IOU) est listée dans le tableau 3. Le score de chaque méthode sur la base de données est calculé. La valeur indiquée en gras indique le meilleur score parmi toutes les méthodes. L'image segmentée de la méthode de fusion LatLRR utilisant U-NET-RESNET50 obtient la première place. Dans l'ensemble, l'image segmentée de la méthode de fusion LatLRR utilisant U-NET-RESNET50 obtient la meilleure performance objective.

Tableau 3; Mesure de performance de IOU moyennes calculées pour toutes les bases de données

Performance Indices	Visible	LatLRR	GFCE
IOU	82.10	90.56	90.11

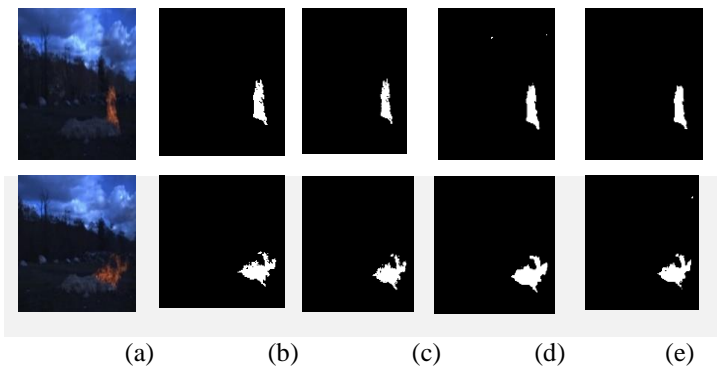


Figure 3. Comparative de performances de segmentation (a) images visibles, (b) vérité terrain, (c) image visible segmentée en utilisant U-NET-RESNET50, (d) image segmentée de la méthode LatLRR en utilisant UNET-RESNET50 (e) image segmentée de la méthode GFCE utilisant UNET-RESNET50

4. Conclusions et travaux futurs

Dans cet article, dans un premier temps, nous avons démontré la faisabilité de (LatLRR) pour la tâche spécifique de la fusion d'images de feu. Cette méthode a affiché les meilleures performances sur la plupart des critères d'évaluation. Dans la deuxième étape, nous avons prouvé que l'image segmentée de la méthode LatLRR en utilisant UNET-RESNET50 affiche la meilleure performance à travers le critère d'évaluation. Comme travail futur, nous visons à développer une segmentation de fumée/feu en temps réel basée sur un modèle d'apprentissage profond pour localiser et segmenter les flammes et la fumée.

5. Bibliographie

- Rossi, J.L. Evolving Risk of Wildfires in Europe: The changing nature of wildfire risk calls for a shift in policy focus from suppression to prevention; UN Office for Disaster Risk Reduction: Brussels, Belgium, 2020.
- European Science & Technology Advisory Group (E-STAG). Evolving Risk of Wildfires in Europe— The Changing Nature of Wildfire Risk Calls for a Shift in Policy Focus from Suppression to Prevention; United Nations for Disaster Risk Reduction—Regional Office for Europe: Brussels, Belgium, 2020. Available online: <https://www.undrr.org/publication/evolving-risk-wildfires-europe-thematic-paper-european-science-technology-advisory> (accessed on 19 June 2021).
- Jolly, M.; Cochrane, P.; Freeborn, Z.; Holden, W.; Brown, T.; Williamson, G.; Bowman, D. Climate-induced variations in global wildfire danger from 1979 to 2013. *Nat. Commun.* 2015, 6, 7537. doi:10.1038/ncomms8537.
- Anshul Gaur, Abhishek Singh, Anuj Kumar, Ashok Kumar, Kamal Kapoor. "Video Flame and Smoke Based Fire Detection Algorithms: A Literature Review", *Fire Technology*, 2020
- Mikaela, W ; Elizabeth , G . Forest Pulse: The Latest on the World'sForests Available online: <https://research.wri.org/gfr/forest-pulse> (accessed on 19 June 2021).
- Alkhatib, A.A.A. A Review on Forest Fire Detection Techniques. *Int. J. Distrib. Sens. Netw.* 2014, 10, 597368,doi:10.1155/2014/597368.
- Li, Pu., Zhao, W.: Image fire detection algorithms based on convolutional neural networks. *Case Stud. Thermal Eng.* 19, 100625 (2020).
- Yuan, C., Zhang, Y., & Liu, Z. (2015). A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. *Canadian Journal of Forest Research*. <https://doi.org/10.1139/cjfr-2014-0347>
- Yuan C, Liu Z, Zhang Y. Fire detection using infrared images for UAV-based forest fire surveillance. In *Fire detection using infrared images for UAV-based forest fire surveillance*. 2017, pp. 567-72
- S. M. Nematidine and D. Gupta, "Nonsampled contourlet domain visible and infrared image fusion framework for fire detection using pulse coupled neural network and spatial fuzzy clustering," *Fire Safety Journal*, vol. 101, pp. 84 – 101, 2018.
- H. Li and X.-J. Wu, "Infrared and visible image fusion using latent low-rank representation," 2018, arXiv:1804.08992. [Online]. Available: <http://arxiv.org/abs/1804.08992>
- Zhou, Z.; Dong, M.; Xie, X.; Gao, Z. Fusion of infrared and visible images for night-vision context enhancement. *Appl. Opt.* 2016, 55, 6480–6490. doi:10.1364/AO.55.006480.
- T. Toulouse, L. Rossi, A. Campana, T. Celik, and M. A. Akhlofi, "Computer vision for wildfire research: An evolving image dataset for processing and analysis," *Fire Safety Journal*, vol. 92, pp. 188 – 194, 2017.
- Piao, J.; Chen, Y.; Shin, H. A New Deep Learning Based Multi-Spectral Image Fusion Method. *Entropy* 2019, 21, 570. <https://doi.org/10.3390/e21060570>
- Xin Jin, Qian Jiang, Shaowen Yao, Dongming Zhou, Rencan Nie, Jinjin Hai, Kangjian He, A survey of infrared and visual image fusion methods, *Infrared Physics & Technology*, Volume

- 85, 2017, Pages 478-501, ISSN 1350-4495, <https://doi.org/10.1016/j.infrared.2017.07.010>
16. Jonathan Long, Evan Shelhamer, and Trevor Darrell. 2015. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition. 3431–3440.
 17. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention. Springer, 234–241.
 18. Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).
 19. Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, and others. 2015. Imagenet large scale visual recognition challenge. International journal of computer vision 115, 3 (2015), 211–252.
 20. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. 770–778
 21. Y. Chen, "Investigation of image fusion algorithms and performance evaluation for night vision," Ph. D. Dissertation, Electrical engineering, Lehigh University, 2007.
 22. Y. Zhao, G. Fu, H. Wang, and S. Zhang, "The fusion of unmatched infrared and visible images based on generative adversarial networks," Mathematical Problems in Engineering, vol. 2020, p. 3739040, Mar 2020.
 23. Zhan, L.; Zhuang, Y.; Huang, L. Infrared and visible images fusion method based on discrete wavelet transform. J. Comput 2017, 28, 57–71.
 24. Ren, K.; Xu, F. Super-resolution images fusion via compressed sensing and low-rank matrix decomposition. Infrared Phys. Technol. 2015, 68, 61–68.
 25. H. Xu, F. Fan, H. Zhang, Z. Le and J. Huang, "A Deep Model for Multi-Focus Image Fusion Based on Gradients and Connected Regions," in IEEE Access, vol. 8, pp. 26316-26327, 2020, doi: 10.1109/ACCESS.2020.2971137.
 26. Sun, C.; Zhang, C.; Xiong, N. Infrared and Visible Image Fusion Techniques Based on Deep Learning: A Review. Electronics 2020, 9, 2162. <https://doi.org/10.3390/electronics9122162>
 27. Kogan, F.; Fan, A.P.; Gold, G.E. Potential of PET-MRI for imaging of non-oncologic musculoskeletal disease. Quant. Imaging Med. Surg. 2016, 6, 756.
 28. Tlig, L.; Bouchouicha, M.; Tlig, M.; Sayadi, M.; Moreau, E. A Fast Segmentation Method for Fire Forest Images Based on Multiscale Transform and PCA. Sensors 2020, 20, 6429. <https://doi.org/10.3390/s20226429>
 29. Zhao, E.; Liu, Y.; Zhang, J.; Tian, Y. Forest Fire Smoke Recognition Based on Anchor Box Adaptive Generation Method. Electronics 2021, 10, 566 <https://doi.org/10.3390/electronics10050566>