

# Classification automatique d'images de propergol solide en combustion par utilisation de réseaux de neurones convolutifs

MATTHIEU NUGUE<sup>1</sup>, ROBIN W. DEVILLERS<sup>1</sup>, ADRIEN CHAN HON TONG<sup>2</sup>,  
GUY LE BESNERAIS<sup>2</sup>, JULIEN PICHILLOU<sup>3</sup>

<sup>1</sup> ONERA, Département Multi-Physique pour l'Energétique

<sup>2</sup> ONERA, Département Traitement de l'Information et Systèmes  
Chemin de la Hunière, BP 80100, 91123 Palaiseau Cedex, France

<sup>3</sup> CNES, Direction des Lanceurs

52 rue Jacques Hillairet, 75612 Paris Cedex, France

<sup>1</sup>matthieu.nugue@onera.fr, robin.devillers@onera.fr,

<sup>2</sup>adrien.chan\_hon\_tong@onera.fr, guy.le\_besnerais@onera.fr

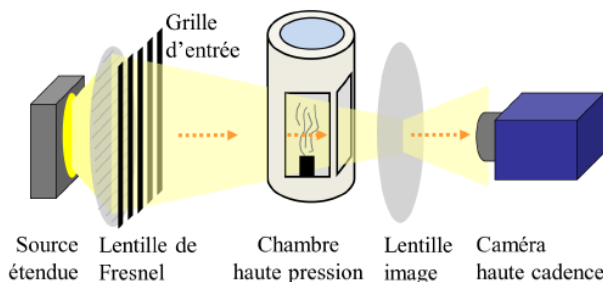
<sup>3</sup>julien.pichillou@cnes.fr

**Résumé** – Dans ce papier, différentes méthodes de classification pixel par pixel sont comparées pour des images expérimentales de propergol solide en combustion. Nous proposons une nouvelle méthode utilisant des réseaux de neurones convolutifs qui permet de classer correctement une plus grande variété d'objets et nous étudions sa sensibilité au nombre de classes.

**Abstract** – In this paper, various pixelwise classification methods are evaluated on experimental images of solid-propellant combustion. A new method using a convolutive neural network is proposed, which allows to classify specific objects that were difficult to detect with other segmentation methods. The sensitivity of the new method with respect to the number of classes used for network training is also investigated.

## 1 Introduction

Les propergols solides sont largement utilisés pour des applications spatiales et militaires, par exemple pour la propulsion de la fusée Ariane. La présence dans certaines compositions de particules d'aluminium, servant à augmenter la performance du moteur, conduit à des écoulements diphasiques au sein de celui-ci. Cependant, la combustion de l'aluminium peut contribuer aux instabilités de fonctionnement du moteur (e.g. instabilité thermo-acoustique [1]). Connaître les caractéristiques des particules d'aluminium en combustion (notamment leur taille et leur vitesse initiale) est aujourd'hui un enjeu important dans les études de nouveaux moteurs à propergols solides. Dans ce cadre, l'ONERA utilise l'ombroscopie focalisée (montage illustré sur la Figure 1) pour visualiser à haute cadence des particules d'aluminium quittant la surface d'un échantillon de propergol en combustion.



**Figure 1 : Schéma de principe du dispositif d'ombroscopie permettant de visualiser un échantillon de propergol solide en combustion.**

Des analyses d'images ont été réalisées sur un cas simple (propergols comportant des particules inertes), et

ont permis de mesurer la taille des objets quittant la surface du propergol [2]. Cependant, les objets présents sur les images de cas réalistes (propergol comportant des particules d'aluminium) ont des apparences diverses et il s'agit de classer ces objets avant d'en extraire les caractéristiques. Dans cette étude, nous proposons une méthode de classification pixel à pixel (on parle aussi de segmentation sémantique) utilisant des réseaux de neurones convolutifs et nous comparons ses résultats aux méthodes d'analyse utilisées par le passé.

## 2 Images étudiées et méthodes

La comparaison des méthodes a été réalisée sur des images issues d'une campagne d'essai réalisée en 2016. L'acquisition des images se fait avec une caméra CMOS de marque Photron, à des cadences dépassant le kHz. Les images, fournies par un logiciel livré avec la caméra, sont au format tiff 12 bit, en niveau de gris. Leur résolution varie entre 256 x 512 px et 1024 x 1024 px suivant la cadence adoptée.

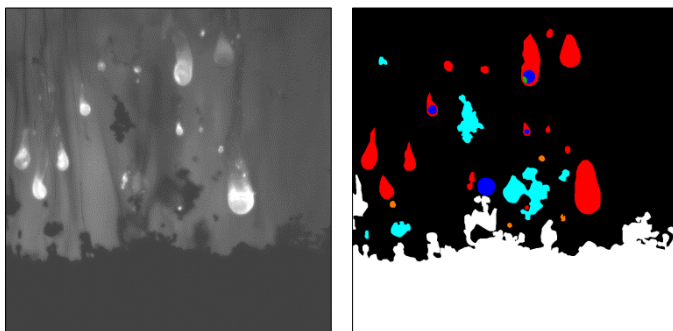
### 2.1 Présentation des images expérimentales et des objets à classer

La Figure 2 présente, sur l'image de gauche, un exemple d'image expérimentale, et sur l'image de droite, la vérité terrain de la classification obtenue par annotation manuelle. Les différentes classes élémentaires de pixels qui nous intéressent sont listées dans le Tab 1 ci-dessous avec la couleur de leur représentation dans la vérité terrain des Figure 2 et Figure 3.

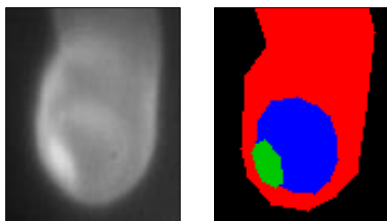
**Tab 1 : Liste des classes élémentaires présentes sur les images ainsi que leurs abréviations et leurs couleurs de représentation dans la vérité terrain**

Nom	Abréviation	Couleur
Fond de l'image	FI	Noir
Propergol	PR	Blanc
Agglomérats non allumés	AN	Cyan
Particules non allumées	PN	Orange
Flamme	FL	Rouge
Gouttes en combustion	GC	Bleu
Lobe d'alumine	LA	Vert

Précisons ces classes. Les particules d'aluminium quittent la surface du propergol en étant entraînées vers le haut de l'image par l'écoulement des gaz de combustion. Ces particules ont des apparences très différentes si elles sont en combustion (une flamme entoure la goutte d'aluminium et un lobe d'alumine très lumineux est présent sur la surface, voir la Figure 3), ou bien si elles ne sont pas encore allumées (apparence sombre). De plus, les particules peuvent s'agréger lorsqu'elles sont encore sur la surface du propergol pour former des agglomérats de formes variables (quelques agglomérats apparaissent en cyan sur l'image de vérité terrain de la Figure 2). Lorsque la combustion d'un agglomérat est initiée, son apparence change rapidement vers celle d'une goutte en combustion. La diversité des classes ainsi que le changement d'apparence des objets au cours du temps rendent complexe l'analyse automatique de ces images.



**Figure 2 : Image expérimentale (à gauche) et vérité terrain de la classification obtenue par annotation manuelle (à droite)**



**Figure 3 : Zoom sur une goutte d'aluminium en combustion**

## 2.2 Méthodes existantes

Deux méthodes existantes vont servir de référence par rapport au traitement proposé, fondé sur les réseaux de neurones convolutifs. Ces deux méthodes ne permettent pas de classer les pixels sur toutes les classes élémentaires présentées dans le Tab 1, mais seulement sur 4 classes, dont certaines regroupent plusieurs classes élémentaires. La deuxième colonne du Tab 2 indique l'affectation des 7 classes élémentaires pour les 4 classes pouvant être obtenues par les deux méthodes existantes.

### 2.2.1 Méthode par seuillage fixe

La détection des objets présents dans les images a initialement été réalisée à l'aide de la librairie Ni-Vision de Labview. La méthode, décrite dans la référence [3], consiste à utiliser deux filtres successifs (lissage puis laplacien) afin d'accentuer les contours. Un seuillage fixe est ensuite appliqué à l'image pour en extraire des régions représentées sous forme de cercle équivalent. Chaque région est ensuite classée en utilisant sa taille et son niveau d'intensité lumineuse. Cette méthode offrait de bons résultats pour la détection d'objets sombres (particules n'étant pas en combustion). Cependant, le grand nombre de paramètres à ajuster manuellement (de 15 à 30), et le temps de traitement élevé sous Labview limitait le volume des statistiques disponibles.

### 2.2.2 Méthode à double détecteur

L'autre méthode de référence combine deux détecteurs. Le premier détecteur permet d'extraire la région correspondant au propergol, par une technique de contours actifs [4]. L'initialisation de ce contour sur la première image est faite en sélectionnant la partie inférieure comme étant le propergol, puis par la suite en utilisant le contour estimé pour l'image précédente. Le deuxième détecteur utilise la méthode MSER (Maximally Stable Extremal Regions) [5], et permet de détecter les différents objets dans l'écoulement au-dessus du propergol. L'emploi de MSER a permis d'améliorer la robustesse des détections par rapport à la méthode à seuils fixes, comme présenté dans [2] pour des particules inertes. Une classification est réalisée en définissant un tri sur des intervalles de validité pour 9 paramètres descriptifs, incluant le niveau de gris de l'objet détecté, son contraste et sa circularité.

## 2.3 Méthode VGG + SVM

Les méthodes décrites dans la Section 2.2 présentent plusieurs limitations, notamment elles ne permettent pas de distinguer une goutte d'aluminium en combustion de la flamme qui l'entoure. Pour y remédier l'utilisation de réseaux de neurones convolutifs est explorée. A ce stade, la base d'apprentissage annotée est limitée à 6 images 1024 x 1024 pixels sous-échantillonnées en 256 x 256 pixels par contrainte de place mémoire, et on adopte une approche associant réseaux de neurones et classification linéaire par SVM (Support Vector Machine), tel que montré dans [6]. Le réseau de neurones VGG16 (Visual Geometry Group) présenté dans [7], dont le modèle est issu de l'entraînement sur la base de données ImageNet, est utilisé pour extraire les caractéristiques des images expérimentales. Plus précisément, notre approche consiste à concaténer les sorties de 2 couches de niveau d'abstraction différent de VGG16 afin d'extraire à la fois une information spatiale fine (couche conv1\_2), et une information « purifiée » (couche conv5\_1). Ce choix rappelle la structure d'échange entre niveaux de U-Net [8]. Les paramètres ainsi obtenus servent à entraîner un modèle de prédiction SVM, issu de la référence [9], qui fournit la classification pixel par pixel.

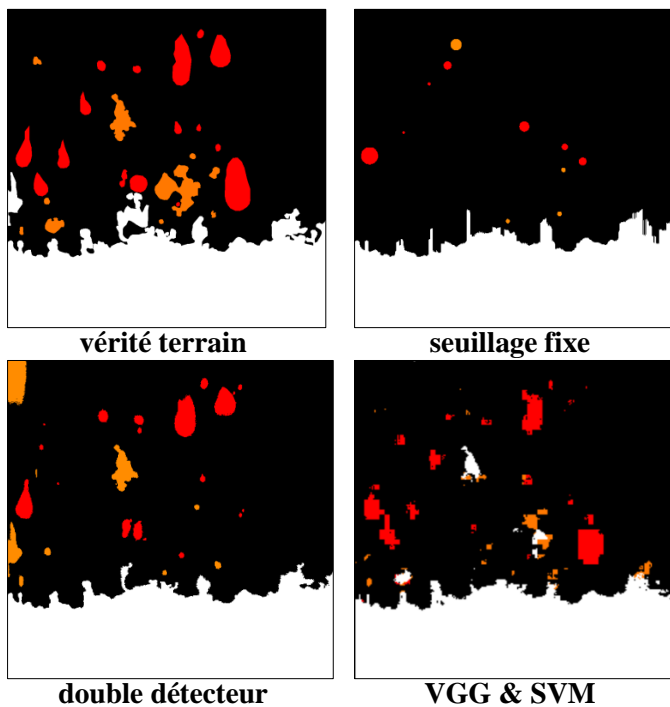
### 3 Résultats

La comparaison entre les 3 méthodes se fait sur 10 images annotées et considère les 4 classes présentées dans Tab 2. La performance est évaluée par calcul du IoU (Intersect over Union) par classe. L'image de prédiction obtenue par les différentes méthodes est par ailleurs comparée qualitativement à la vérité terrain présentée dans la Section 2.1. Enfin, la Section 3.2 présente une analyse de l'impact du nombre de classes choisi pour entraîner la méthode VGG + SVM.

**Tab 2 : Affectation des classes élémentaires en fonction du nombre total de classes considérées**

Nombre de classes :	3	4	5	7
Fond de l'image	FI	FI	FI	FI
Propergol	PR	PR	PR	PR
Agglomérats non allumés	+ AN	AN	+ AN	AN
Particules non allumées	+ PN	+ PN	+ PN	PN
Flamme	FL	FL	FL	FL
Gouttes en combustion	+ GC	+ GC	GC	GC
Lobe d'alumine	+ LA	+ LA	LA	LA

#### 3.1 Comparaison des méthodes sur 4 classes



**Figure 4 : Comparaison de la vérité terrain avec une image prédite par les trois méthodes**

La Figure 4 compare la vérité terrain avec la prédiction des trois méthodes sur les 4 classes : « fond de l'image » (FI), « propergol » (PR), « objets non allumés » (AN + PN) et « objets clairs » (FL + GC + LA). Les détections produites par la méthode à seuillage fixe sont très lacunaires, seules les petites particules non allumées et quelques flammes sont localisées ; la détection de la surface manque également de finesse. La méthode par double détecteur permet de définir un contour de surface propergol plus fin, et améliore nettement la détection des objets au-dessus du propergol. On note en revanche des fausses alarmes liées à des fumées. La méthode VGG + SVM détecte un grand nombre d'objets présents sur l'image, mais la définition de leurs contours manque

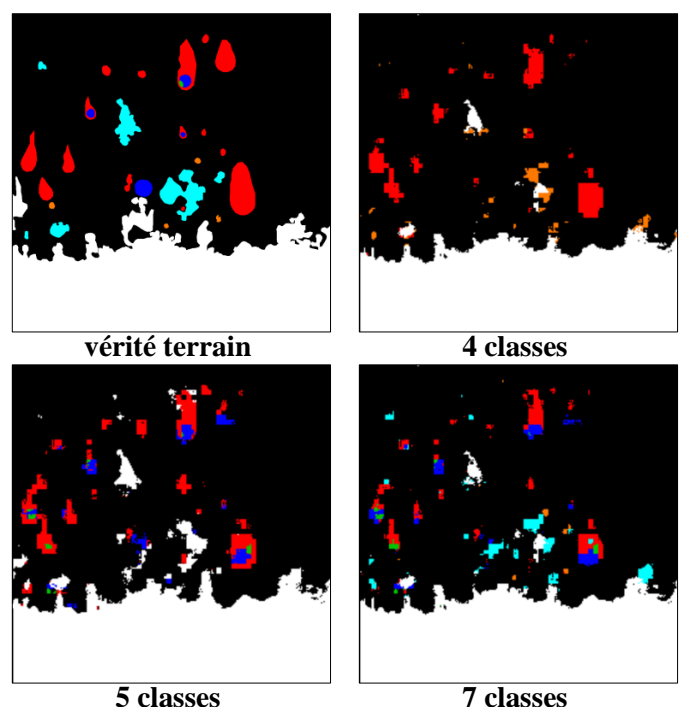
de finesse. Il est intéressant de remarquer la présence de pixels de la classe propergol parmi des objets non allumés : de fait, ces deux classes présentent des intensités très similaires, et c'est le contexte qui permet de les distinguer, ce qui indique que le réseau est encore trop fixé sur des caractéristiques locales.

Le Tab 3 compare la performance des différentes méthodes. Les classes « fond » et « propergol », majoritaires sur les images, sont bien classées pour l'ensemble des méthodes. La méthode par seuillage fixe détecte très peu d'objets d'où le IoU faible pour les classes « objets clairs » et « objets non allumés ». Elle obtient le meilleur score sur la classe « propergol » malgré une estimation de la surface très grossière, dû au score IoU favorisant les résultats de classifications qui englobent complètement la vérité terrain. La méthode à double détecteur obtient le meilleur score sur la classe des « objets clairs » mais le moins bon sur celui de la classe des « objets non allumés », à cause de fausses alarmes plus nombreuses. La méthode VGG + SVM permet de progresser sur les classes « objets non allumés » et « fond » tout en conservant des scores proches de l'optimal sur les autres classes. Le IoU obtenu par cette méthode sur la classe des « objets non allumés » reste bas, car il est encore difficile pour le réseau de différencier le propergol des objets non allumés (qui apparaissent visuellement comme des parties de propergol entraînés dans l'écoulement).

**Tab 3 : Comparaison des IoU pour les trois méthodes**

Classe	Seuillage fixe	Double détecteur	VGG & SVM
Fond	0.944	0.946	<b>0.959</b>
Propergol	<b>0.983</b>	0.981	0.980
Objets clairs	0.215	<b>0.459</b>	0.458
Non allumés	0.200%	0.171	<b>0.235</b>

#### 3.2 Méthode VGG + SVM : Etude de la sensibilité aux ensembles de classe utilisés



**Figure 5 : Comparaison des résultats VGG+SVM en fonction du nombre de classes utilisé**

On s'intéresse ici au résultats de la méthode VGG+SVM pour trois problèmes de classification sur des ensembles de 4, 5 et 7 classes définis dans Tab 2. La Figure 5 représente la vérité terrain comparée aux prédictions de la méthode VGG + SVM. Les pixels de la classe « flamme » (FL) et « objets sombres » (PR + PN + AN) apparaissent toujours correctement classés. Dans les ensembles à 5 et 7 classes, comme cela a déjà été mentionné, il y a une confusion entre les classes « propergol » et « agglomérats non allumés ». Dans ces cas, on remarque aussi une bonne localisation de la classe « gouttes en combustion » à la base d'une flamme (chaque flamme est générée par une goutte d'aluminium en combustion), parfois même lorsque ces éléments ne sont pas dans la vérité terrain pour des raisons d'imprécision de l'annotation manuelle.

**Tab 4 : IoU en fonction du nombre de classes et nombre de pixels servant à l'entraînement**

Classe	4 classes	5 classes	7 classes	nombre de px
Fond	0.959	0.956	0.940	310344
Propergol	0.980	0.975	0.980	64980
Agglomérats non allumés	0.235		0.134	3782
Particules non allumés			0.315	451
Flamme	0.458	0.446	0.386	11774
Gouttes en combustion		0.075	0.061	1581
Lobe d'alumine		0.061	0.039	304

Le tableau Tab 4 représente les IoU obtenus ainsi que le nombre de pixels servant à l'entraînement du modèle par classe. Le faible score obtenu pour les classes « gouttes en combustion » et « lobe d'alumine » peut être expliqué par le faible nombre d'exemples disponibles pour l'entraînement dû à la difficulté d'annoter ces 2 classes : la présence d'une flamme très lumineuse masque souvent la goutte d'aluminium et le lobe d'alumine. Le Tab 5 compare les performances en fonction du nombre de classes utilisées on se base sur les 3 premières classes du Tab 4. Comme observé dans [10], le IoU diminue avec le nombre de classes considérées. La classification portant sur le moins de classes obtient le meilleur score au prix d'une moindre capacité de généralisation.

**Tab 5 : IoU évalué sur trois classes pour trois classifications utilisant trois nombre de classes différents.**

Classe	4 classes	5 classes	7 classes
0 « fond »	<b>0.959</b>	0.957	0.940
1 « objets sombres »	0.969	<b>0.975</b>	0.965
2 « objets clairs »	<b>0.458</b>	0.402	0.359

## 4 Conclusion

Nous avons proposé une méthode VGG+SVM associant un réseau de neurones convolutionnel et un classifieur par SVM pour l'analyse d'images de propergol en combustion. Cette méthode a été comparée

à deux approches de référence [2] [3]. Elle conserve des performances meilleures ou très proches des méthodes de référence, mais permet de raffiner la classification et de détecter des objets difficilement détectables jusqu'à maintenant (y compris pour un œil humain) telles que les gouttes d'aluminium en combustion au sein de la flamme. Ces premiers résultats très encourageants confortent l'idée de continuer les travaux en utilisant des réseaux plus évolués pour la segmentation sémantique, tel que SegNet, ou encore d'approfondir l'apprentissage, en entraînant entièrement le réseau sur une base plus conséquente d'images d'ombroscopie.

## Remerciements

Ces travaux font partie d'une thèse cofinancée par l'ONERA et le CNES et s'inscrivent dans un projet pluriannuel associant l'ONERA et le CNES sur l'analyse d'images d'ombroscopie.

## 5 Bibliographie

- [1] S. Gallier and F. Godfroy (2009). Aluminum Combustion Driven Instabilities in Solid Rocket Motors. In *JPP*, Vol. 25, No. 2.
- [2] M. Nugue, R. W. Devillers, G. Le Besnerais and N. Cesco (2016). Particle detection & size evaluation in solid propellant flames via experimental image analysis to improve two-phase flow simulation in rocket motors. *Space Propulsion Conference*, Rome.
- [3] R. W. Devillers, C. Erades, D. Lambert and J. Belessa (2014). Mesure et suivi de particules, agglomérats et gouttes en combustion au-dessus de la surface d'un propergol en combustion. *14<sup>th</sup> CFTL*, Marseille.
- [4] T. F. Chan and L. A. Vese (2001). Active Contours Without Edges. *IEEE Transactions on Image Processing*, Vol. 10, No 2, pp. 266-277.
- [5] J. Matas, O. Chum, M. Urban and T. Padjla (2002). Robust Wide Baseline Stereo from Maximally Stable Extremal Regions. In *BMVC*, Vol. 22(10), pp. 761-767.
- [6] R. Girshick, J. Donahue, T. Darrell and J. Malik (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *CVPR*.
- [7] K. Simonyan and A. Zisserman (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. In *ICLR*. arXiv: 1409.1556
- [8] O. Ronneberger, P. Fischer and T. Brox (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *MICCAI*, pp. 234-241.
- [9] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang and C.-J. Lin (2008). LIBLINEAR: A Library for Large Linear Classification. *Journal of Machine Learning Research*, Vol. 9, pp. 1871-1874.
- [10] N. Razavi, J. Gall, P. Kohli and L. Van Gool (2012). Latent Hough Transform for Object Detection. In *ECCV*, pp. 312-325.