

Détection de colonies de micro-organismes dans des images haute résolution

Aurélien Launay^{1,2}

Guillaume Perrin²

Ernest Hirsch¹

¹ICube, 300 Bd Sébastien Brant, 67400 Illkirch-Graffenstaden

²bioMerieux, 376 chemin de l'Orme, 69280 Marcy L'Etoile

a.launay@unistra.fr

1 Résumé

La détection et l'énumération automatique de colonies de micro-organismes observées dans des boîtes de Pétri est un enjeu majeur en microbiologie (détection de contaminations en milieu industriel par exemple). De nombreuses études [3] [6] [9] [10] permettent, par analyse d'images, d'automatiser cette détection en fin d'incubation. Cependant ces méthodes (par exemple segmentation automatique des colonies), ne permettent de gagner du temps qu'à la lecture finale. Des instruments spécifiques (*i.e.* EvisightTM Compact [7]) permettent un suivi en temps réel, par prise régulière d'images, pendant l'incubation. Avec le développement des méthodes d'apprentissage profond, les performances de détection d'objets dans les images ont été considérablement améliorées au cours des dernières années. L'étude par apprentissage profond menée dans [6] montre un gain important par rapport aux méthodes traditionnelles d'analyse d'image. La combinaison du dispositif d'acquisition utilisé et de méthodes d'apprentissage telles que les réseaux de neurones convolutifs, permettent d'obtenir des résultats encourageant aussi bien pour la détection que pour l'énumération des colonies des micro-organismes observées dans des boîtes de Pétri.

2 État de l'art

La culture de colonies de micro-organismes dans des boîtes de Pétri est utilisée afin d'étudier le niveau de contamination d'un échantillon (environnement, clinique, etc.). Cette contamination est évaluée en fonction du nombre de colonies qui se développent dans la boîte pendant son incubation [3]. La numération des colonies dans les boîtes de Pétri à la fin de l'incubation, bien que relativement aisée pour un expert humain, s'avère longue et fastidieuse. Des études ont été menées afin d'automatiser ce comptage. La méthode privilégiée est celle de la prise d'images de boîtes en fin d'incubation, puis du comptage de colonies après segmentation des objets d'intérêt. Les outils algorithmiques varient du seuillage manuel [10], pour lequel la définition du seuil nécessite la maîtrise de l'ensemble des paramètres d'acquisition de l'image, à des méthodes de seuillage automatique, plus complexes [9]. Des techniques de morphologie mathématique sont également utilisées en complément des méthodes de seuillage, augmentant le nombre de paramètres à fixer. C'est cette succession de seuillages et d'opérations morphologiques qui est utilisée au sein de l'instrument d'acquisition [7]. Celui-ci permet d'incuber des boîtes et d'en acquérir des images tout au long de l'incubation, permettant ainsi un suivi en temps réel de la croissance microbienne. Une autre méthode, récemment utilisée dans [6], consiste à utiliser l'image en fin d'incubation, et à compter les colonies à l'aide de réseaux d'apprentissage profond, à savoir des réseaux de neurones convolutifs. Cette méthode d'apprentissage utilise comme base de données d'entraînement des sous-images (d'une boîte de Pétri) contenant de zéro à six colonies. Dans cette application, le réseau, une fois entraîné, compte le nombre de colonies présentes dans une fenêtre, fenêtre qui parcourt l'ensemble de l'image. Notre contribution cherche à combiner les images de l'instrument d'acquisition et les techniques d'apprentissage profond pour construire un réseau de neurones capable de détecter et énumérer les colonies observées dans une boîte de Pétri durant son incubation. Cette approche vise à éviter de nombreux artefacts dus à la présence du couvercle de la boîte de Pétri (occlusion partielle des colonies, sérigraphies, etc.), mais également à être utilisable tout au long de l'incubation.

3 Méthode et données

Les données utilisées sont des images haute résolution de boîtes de Pétri, prises tout au long de l'incubation, à intervalle de temps régulier (toutes les 30min). La taille de ces images varie de 6000*6000 pixels à 9000*9000 pixels selon les boîtes, avec une résolution de 2400dpi. Les données à disposition représentent 132 images (ou séquences d'images) pour deux espèces de micro-organismes différentes, ainsi que des boîtes sans colonies. Les espèces sont réparties comme suit (en pourcentage du nombre d'images) : *Candida albicans* (47%), *Staphylococcus aureus* (27%), boîtes sans colonies (25%).

Plusieurs difficultés sont à prendre en compte dans les images :

- la présence de petites colonies regroupées, difficiles à détecter et surtout à compter ;
- la présence d'artefacts (*e. g.* gouttes d'eau, poussières) ;
- des occlusions dues au couvercle des boîtes (les colonies apparaissent déformées, ou en partie dissimulées), ces couvercles étant conservés durant toute l'incubation afin d'éviter tout risque de contamination.

On cherche à entraîner un réseau permettant la détection de *C. albicans* et *S. aureus* (voir figures 2 et 3), deux espèces morphologiquement très proches. Le réseau de neurones doit permettre la détection de colonies dans l'image finale, afin de vérifier que l'on peut atteindre un niveau de reconnaissance proche de l'état de l'art décrit dans [3], faisant état d'une

précision et d'un rappel de 96%. La banque d'images regroupant l'ensemble des souches de *C. albicans* et *S. aureus* représente 3828 colonies.

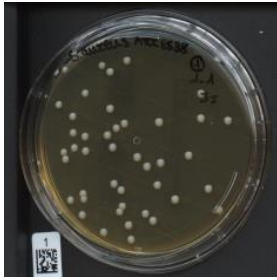


Figure 1: *S. aureus* à 24h.



Figure 2: *C. albicans* à 24h.

Parmi les images de la base de données, 90% sont réservés à l'apprentissage du réseau et 10% au test. Des images de la base d'apprentissage, on extrait des sous-images contenant des exemples de colonies et des exemples de négatifs (pas de colonies présentes dans la sous-image). On obtient ainsi des sous-images semblables à celles observées dans [6], mais on cherche à n'avoir qu'une seule colonie par sous-image au plus. Pour les deux espèces testées, on extrait ainsi 3327 sous-images de colonies et 40000 sous-images de négatifs, qui représentent la base d'apprentissage du réseau. La répartition entre les deux classes est déséquilibrée. Les objets à détecter étant très petits par rapport à la taille de l'image, il est en effet plus aisé d'obtenir des exemples de sous-images sans colonies. Pour pallier cette difficulté et apprendre au mieux les positifs, le nombre de données positives est artificiellement augmenté par une technique d'augmentation de données [11]. On applique aléatoirement à chaque colonie dix translations et éventuellement un retournement. Toutes les colonies sont donc présentes en plusieurs exemplaires différents dans la base de données d'apprentissage. Cette dernière est répartie en une base d'entraînement et une base de validation. Lors de l'apprentissage du modèle, les performances sont estimées sur la base de validation, permettant, si l'écart entre les performances pour les bases d'entraînement et de validation diverge (signe de sur-apprentissage) d'arrêter l'entraînement du réseau de manière optimale.

Les sous-images de la base d'entraînement sont fournies à un réseau de neurones convolutifs dont la structure est celle du réseau VGG16 décrit dans [8]. La base de données restant relativement restreinte, on utilise une technique de transfert d'apprentissage décrite par exemple dans [4]. Les poids du réseau de neurones sont ainsi initialisés avec les poids du réseau pré-entraînés sur une base de données plus importante (ici ImageNet). Les premières couches (ici de 1 à 5), qui extraient les caractéristiques les plus génériques [12], ne sont pas ré-entraînées, et sont donc figées. Les couches convolutives suivantes (de 6 à 13), qui extraient des caractéristiques plus spécifiques aux images [12], sont ré-entraînées avec des sous-images de colonies et de négatifs. Enfin, l'architecture des couches entièrement connectées, dont le rôle est de classer les sous-images en fonction des caractéristiques extraites dans les couches convolutives, est modifiée par rapport à celle du réseau VGG16, afin de répondre au problème de classification posé. Ces couches sont entraînées en même temps que les couches précédentes avec notre base de données. Les hyper-paramètres du modèle sont quant à eux laissés aux valeurs par défaut (« learning rate », etc.).

Deux techniques de normalisation des images sont appliquées pour évaluer l'impact du prétraitement sur les performances:

- sous-images centrées (moyenne nulle) ;
- soustraction de l'image au temps initial à l'image à tester (image au temps final d'incubation). Ce sont les sous-images extraites de cette « image de soustraction » qui sont utilisés pour l'apprentissage. En raison de contraintes mécaniques liées au dispositif d'acquisition, deux images consécutives ne sont pas nécessairement alignées (un décalage de quelques pixels peut être observé). La seconde méthode de normalisation nécessite ainsi l'utilisation de méthodes de recalage [1] entre les deux images, afin d'assurer une soustraction correcte.

4 Résultats et discussions

Les performances de la méthode de classification décrite ci-dessus, pour les deux méthodes de normalisation des images, sont calculées en termes de rappel et de précision.

$$\text{Rappel} = \frac{\text{Vrais Positifs}}{\text{Vrais positifs} + \text{Faux négatifs}} \quad \text{Précision} = \frac{\text{Vrais Positifs}}{\text{Vrais positifs} + \text{Faux positifs}}$$

Une fenêtre glissante parcourt l'ensemble des images de la base de données de test constituée de 10 % de la base de données totale (soit 13 images). Le réseau de neurones est appliqué à chacune de ces fenêtres afin d'obtenir un score de classification. Un algorithme de suppression des non maxima [2] est appliqué pour ne conserver que les sous-images ayant localement les plus hauts scores. Un seuil variable est ensuite appliqué sur les scores conservés afin de déterminer les sous-images positives que l'on cherche à apparier avec les colonies annotées. Pour chaque valeur du seuil, le rappel et la précision sont ainsi calculés. Les courbes de rappel / précision obtenues sont présentées en figure 3.

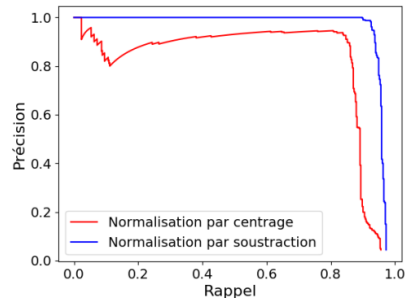


Figure 3: Courbes rappel / précision pour les méthodes de normalisation par centrage et par soustraction.

On utilise la métrique de référence pour la détection d'objet [5], la précision moyenne, qui s'apparente à l'aire sous la courbe (AUC) lorsqu'on travaille en précision/rappel et non en sensibilité/spécificité (1 seule classe d'intérêt). On obtient une précision moyenne de 0.82 pour la méthode de normalisation par centrage, contre 0.96 pour la méthode de normalisation par soustraction. C'est cette dernière que l'on conservera donc pour la suite de l'étude. Les principales erreurs de détection concernent les faux négatifs (rappel plafonnant à 0.97), en particulier lorsque des petites colonies trop proches les unes des autres (figure 4), pour lesquelles l'approche par boîtes englobantes (de taille fixe) présente des limites, sont observées dans les sous-images. Certaines colonies en bord de boîte présentent aussi des scores de classification trop faibles, entraînant également des faux négatifs (figure 5).

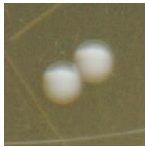


Figure 4 : Sous-image de colonies trop proches.

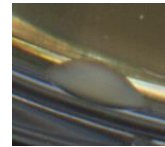


Figure 5 : Sous-image de colonie en bords de boîtes.

Pour résoudre le problème des colonies trop proches, une approche multi-échelle pourrait être envisagée pour adapter la taille des fenêtres à celle des colonies. Pour les colonies en bord de boîte, la prise en compte de l'information de croissance par le réseau de neurones (c'est-à-dire suivre une colonie sur plusieurs pas de temps) devrait permettre d'améliorer les performances. L'étude d'autres micro-organismes tels que *Pseudomonas aeruginosa*, *Aspergillus brasiliensis* et *Bacillus subtilis* (qui ont des morphologies très différentes) est également envisagée par la suite.

Bibliographie

- [1] BETZNER Marlana, PIERQUIN Joseph, KOHLER Sophie *et al.* High Resolution Image Registration for Microcolonies Monitoring on Petri Dishes. *IRBM*, 2017, vol. 38, no 6, p. 325-338.
- [2] BODLA Navaneeth, SINGH Bharat, CHELLAPPA Rama *et al.* Soft-NMS – Improving object detection with one line of code. *Preprint arXiv:1704.04503*, 2017.
- [3] CHIANG Pei-Ju, TSENG Min-Jen, HE Zong-Sian, *et al.* Automated counting of bacterial colonies by image analysis. *Journal of microbiological methods*, 2015, vol. 108, p. 74-82.
- [4] DONAHUE Jeff, JIA Yangqing, VINYALS Oriol *et al.* Decaf : A deep convolutional activation feature for generic visual recognition. Proc. *International conference on machine learning*. Beijing, China, 2014, p. 647-655.
- [5] EVERINGHAM Mark, VAN GOOL Luc, WILLIAMS Christopher K.I. *et al.* The pascal visual object classes (voc) challenge. *International journal of computer vision*, 2010, vol. 88, no 2, p. 303-338.
- [6] FERRARI Alessandro, LOMBARDI Stefano, SIGNORONI Alberto. Bacterial colony counting by convolutional neural networks. Proc. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milano, Italy, 2015, p. 7458-7461.
- [7] GHODBANE Ramzi, ASMAR Shady, BETZNER Marlana *et al.* Rapid diagnosis of tuberculosis by real-time high-resolution imaging of Mycobacterium tuberculosis colonies. *Journal of clinical microbiology*, 2015, vol. 53, no 8, p. 2693-2696.
- [8] KRIZHEVSKY Alex, SUTSKEVER Ilya, HINTON Geoffrey E. Imagenet classification with deep convolutional neural networks. Proc. *25th International Conference on Neural Information Processing Systems (NIPS' 12)*, Lake Tahoe, Nevada, 2012, vol. 1, p. 1097-1105.
- [9] LIU Anan, LIU Zheren, SONG Limin *et al.* Adaptive ideal image reconstruction for bacteria colony detection. *Information Technology and Agricultural Engineering*, 2012, p. 353-360.
- [10] PUTMAN Michael, BURTON Robert, NAHM Moon H. Simplified method to automatically count bacterial colony forming unit. *Journal of Immunological Methods*, 2005, vol. 302, no 1, p. 99-102.
- [11] WONG Sebastien C., GATT Adam, STAMATESCU Victor *et al.* Understanding data augmentation for classification : when to warp?. Proc. *IEEE International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Gold Coast, Australia, 2016. p. 1-6.
- [12] ZEILER Matthew D., FERGUS Rob. Visualizing and understanding convolutional networks. Proc. *European Conference on Computer Vision, Part I, LNCS vol. 8689*, Springer, Zürich, Switzerland, 2014, p. 818-833.