

Synthèse bibliographique sur les méthodes d'apprentissage pour la détection de spots en microscopie de fluorescence

Victor Piriou / Encadrant : Thibault Lagache (Unité d'Analyse d'Images Biologiques, Institut Pasteur, Paris)

27 octobre 2022

Table des matières

1	Introduction	2
2	Synthèse bibliographique	3
2.1	Méthodes traditionnelles	3
2.1.1	Vue d'ensemble	3
2.1.2	Conclusion	3
2.2	Méthodes plus récentes basées sur l'apprentissage profond	3
2.2.1	Méthodes spécifiquement dédiées à la détection de spots en imagerie de fluorescence	3
2.2.1.1	DetNet	3
2.2.1.2	deepBlink	4
2.2.1.3	DeepSpot	5
2.2.1.4	Conclusion	6
2.2.2	Méthodes plus générales dédiées à la segmentation d'objets en imagerie biologique	6
2.2.2.1	StarDist et SplineDist : deux méthodes basées sur la même philosophie	6
2.2.2.2	StarDist	7
2.2.2.3	SplineDist	7
2.2.2.4	Conclusion	7
2.3	Conclusion	8

1 Introduction

En imagerie biologique, la microscopie de fluorescence permet de marquer des objets avec des protéines fluorescentes afin de les rendre visibles à l'image. La détection de ces objets est une étape nécessaire pour permettre une analyse quantitative de leur dynamique.

En raison du phénomène de diffraction et des imperfections du système optique, lorsque ces objets sont plus petits que la résolution du microscope, ils apparaissent à l'image sous formes de petites taches brillantes, floues et sans contour clairement définissables. Ce flou est caractérisé par la fonction d'étalement du point (ou "PSF" pour *Point Spread Function*)

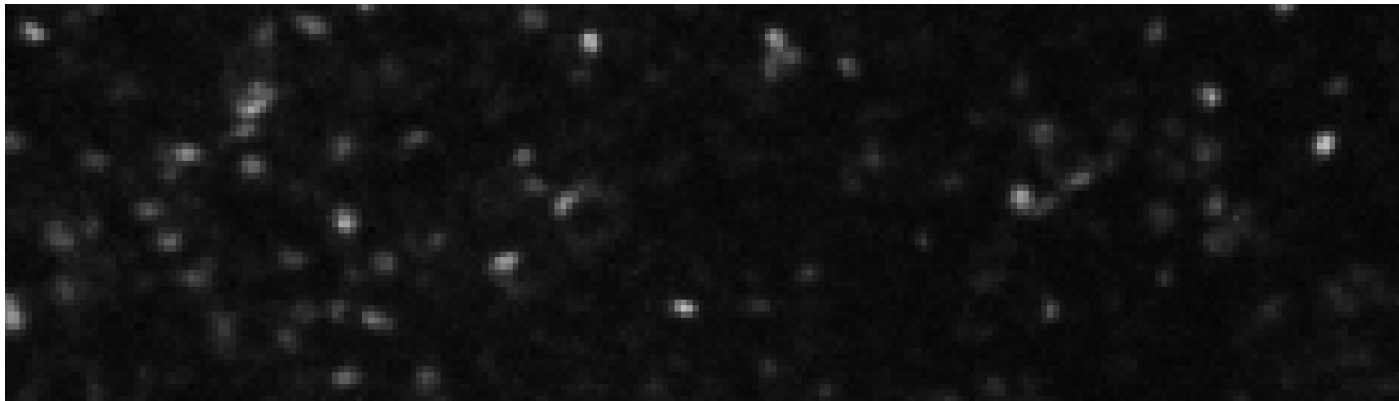


Fig. (1) Morceau d'une des images fournies pour la réalisation de ce projet

tion) du système, qui est l'image d'une source ponctuelle de lumière et qui peut être approximée avec un haut degré de précision par une gaussienne [1].

Outre la résolution spatiale limitée par la diffraction, une autre source majeure d'aberrations introduites dans le processus d'imagerie est le bruit photonique intrinsèque, qui résulte de la nature aléatoire de l'émission des photons. Le bruit des photons (bruit de Poisson), qui est indépendant de l'électronique du détecteur (qui ajoute un bruit gaussien), ne peut être réduit qu'en augmentant l'intensité lumineuse ou le temps d'exposition. Toutefois, augmenter l'intensité lumineuse afin d'améliorer la qualité de l'image augmente aussi l'altération photochimique des marqueurs fluorescents, menant à l'atténuation de leur fluorescence, un processus appelé photoblanchiment.

En conséquence, il est souvent difficile, même pour les biologistes experts, de distinguer les objets du bruit ou des structures de fond non pertinentes dont l'intensité peut excéder celle des spots. De plus, ce travail manuel est extrêmement fastidieux, prend du temps et dépend de l'observateur, ce qui le rend non-reproductible.

Depuis plusieurs décennies, des méthodes de détection automatique de spots sont développées. Cependant, même les méthodes les plus récentes rapportent soit trop de faux positifs - corrompant ainsi l'analyse par la présence d'objets inexistants - soit moins de spots qu'il n'y en a en réalité. Cela pose aussi problème en vidéomicroscopie, où les spots doivent être suivis dans le temps pour étudier la dynamique des objets correspondants. Dans les algorithmes de suivi courants, qui consistent le plus souvent en deux étapes distinctes et consécutives de détection (spatiale) et de liaison (temporelle), la performance du détecteur est cruciale. En effet, une mauvaise détection peut par exemple entraîner la procédure de liaison à trouver des trajectoires aberrantes dans lesquelles des objets pourtant correctement détectés se retrouvent liées à de fausses détections d'une trame voisine. Un autre type d'erreur est la terminaison précoce de trajectoires due au fait qu'aucun spot similaire n'ait été détecté dans les trames suivantes. Les performances imparfaites des méthodes d'apprentissage actuelles ne sont pas surprenantes étant donné que les spots d'un même jeu de données ont tendance à beaucoup se ressembler, tandis que leur apparence peut beaucoup varier d'un jeu de données à un autre, en raison des différents composants biologiques suivis, des différentes conditions expérimentales et d'imagerie. Toutes ces difficultés motivent la recherche de nouvelles méthodes de détection automatique de spots fluorescents.

Ce document présente les méthodes de détection de spots, triées dans l'ordre chronologique. Il contient non seulement une synthèse bibliographique sur les méthodes d'apprentissage mais fournit également une vue d'ensemble d'autres méthodes plus anciennes pouvant être une source d'inspiration.

2 Synthèse bibliographique

2.1 Méthodes traditionnelles

2.1.1 Vue d'ensemble

En 2009, Ihor Smal *et al.* ont publié une comparaison quantitative des méthodes de détection de spots [2] les plus utilisées en microscopie de fluorescence. Elles étaient basées sur l'analyse multi-échelles basée sur les ondelettes [3] [4], la morphologie mathématique [5] [6] [7] et l'apprentissage automatique [8] [2]. Chacune des méthodes comparées dans cet article s'applique en trois étapes et la différence entre chacune de ces méthodes est la façon dont chaque étape est mise en oeuvre. Ces trois étapes sont :

1. **Débruitage de l'image originale.**
2. **Rehaussement de l'intensité des spots et atténuation de l'intensité des structures de l'arrière-plan.** Les méthodes [5] (Bright *et al.*, 1987) [6] (Joss *et al.*, 1991) sont basées sur la transformation « chapeau haut-de-forme » (*top-hat filter*) qui permet de conserver uniquement les éléments de l'image qui sont à la fois plus petits que l'élément structurant et plus clairs que leurs voisins.

Une des méthodes [4] (Olivo-Marin, 2002) basée sur les ondelettes exploite le fait que les particules sont représentées par un petit nombre de coefficients d'ondelettes qui sont corrélés entre les échelles et qu'au contraire, les coefficients dus au bruit sont distribués de manière aléatoire et ne sont pas propagés entre les échelles.

L'utilisation d'une des méthodes [8] (Jiang *et al.*, 2007) d'apprentissage implique le glissement sur l'image d'une fenêtre dont l'intensité du pixel situé en son centre est ensuite remplacée par le résultat de la classification d'une variante du méta-algorithme AdaBoost. La sortie finale de ce dernier est une somme pondérée des sorties d'autres classifieurs faibles. Dans cet article les classifieurs faibles qui ont été utilisés sont des caractéristiques pseudo-Haar, sensibles à la directionnalité des motifs de l'image. Cependant, le nombre total de caractéristiques pseudo-Harr est très grand, bien plus grand que le nombre de pixels. Afin d'assurer une classification rapide, une grande majorité des caractéristiques disponibles doit être exclue afin de se concentrer sur un petit ensemble de caractéristiques critiques. Par conséquent, la version d'AdaBoost qui a été choisie contraint le classifieur faible de sorte à ce qu'il dépende d'une seule caractéristique. Ainsi, la sélection de chaque nouveau classifieur faible correspond à la sélection d'une seule caractéristique.

La méthode d'apprentissage introduite dans cette étude comparative [2] consiste à convoluer l'image avec un filtre linéaire qui est appris après avoir effectué une analyse discriminante de Fisher sur deux ensembles de vecteurs : un qui contient des morceaux aplatis d'images qui représentent des particules, et un autre qui contient des morceaux aplatis d'images qui ne correspondent pas à des particules. La meilleure projection est celle qui maximise à la fois la variance inter-classe et minimise la variance intra-classe.

3. **Seuillage de l'image traitée afin d'obtenir une image binaire.** Dans la plupart des cas, parce que le débruitage n'est pas parfait, les groupes de pixels non nuls dans l'image binaire qui appartiennent au même point ne sont pas connectés ou contiennent des pixels nuls erronés à l'intérieur du groupe. Afin de résoudre ce problème, l'opération de fermeture de la morphologie mathématique est fréquemment utilisée comme étape de post-traitement.

Ces méthodes ont été évaluées sur des images synthétiques et il a été montré que ce sont les deux méthodes supervisées qui obtiennent la meilleure proportion de vrais positifs. Par ailleurs, à partir des expérimentations sur des images réelles, il a été confirmé que la performance réelle des méthodes de détection dépend de l'application.

2.1.2 Conclusion

Les méthodes les plus anciennes [5] [6] reposaient sur la paramétrisation manuelle d'opérateurs mathématiques représentant la forme des spots de façon idéalisée. Ce n'est pas le cas des méthodes plus récentes [4] [8] [2] mais elles nécessitent toujours une optimisation indépendante de chacune de ces trois étapes ainsi qu'un seuillage global de l'image. L'utilisation d'un seuil global est problématique puisqu'en raison de plusieurs facteurs tels que le photoblanchiment, les statistiques des images peuvent varier dans le temps (*e.g.* les images d'une séquence vidéo ou celles acquises lors de différentes expérimentations espacées dans le temps). Cela contraint donc à l'optimisation d'un seuil différent pour chacune des images en fonction de leur plage d'intensité. Pour pallier ce problème, une méthode [9] (Basset *et al.*, 2015) de détection incluant un seuil localement adapté et automatiquement déduit des statistiques d'intensités locales a toutefois été proposée.

2.2 Méthodes plus récentes basées sur l'apprentissage profond

2.2.1 Méthodes spécifiquement dédiées à la détection de spots en imagerie de fluorescence

2.2.1.1 DetNet DetNet (Wollman *et al.*, 2019) [10] est un réseau de neurones basé sur un *deconvolution network* [11] prédisant, pour chacun des pixels de l'image originale, la probabilité qu'il corresponde à une particule. Il a été adapté de sorte à ce que le nombre de paramètres soit fortement réduit (17000 paramètres au lieu des 1.1 million de paramètres du réseau duquel il a été inspiré). Plus précisément, les nombres de *feature maps* et d'opérations de *pooling* ont été réduits, et

les convolutions transposées de la partie décodeur ont été remplacées par des interpolations bilinéaires. Son faible nombre de paramètres lui permet d'être entraîné avec peu de données. Il a été évalué sur deux jeux de données synthétiques et

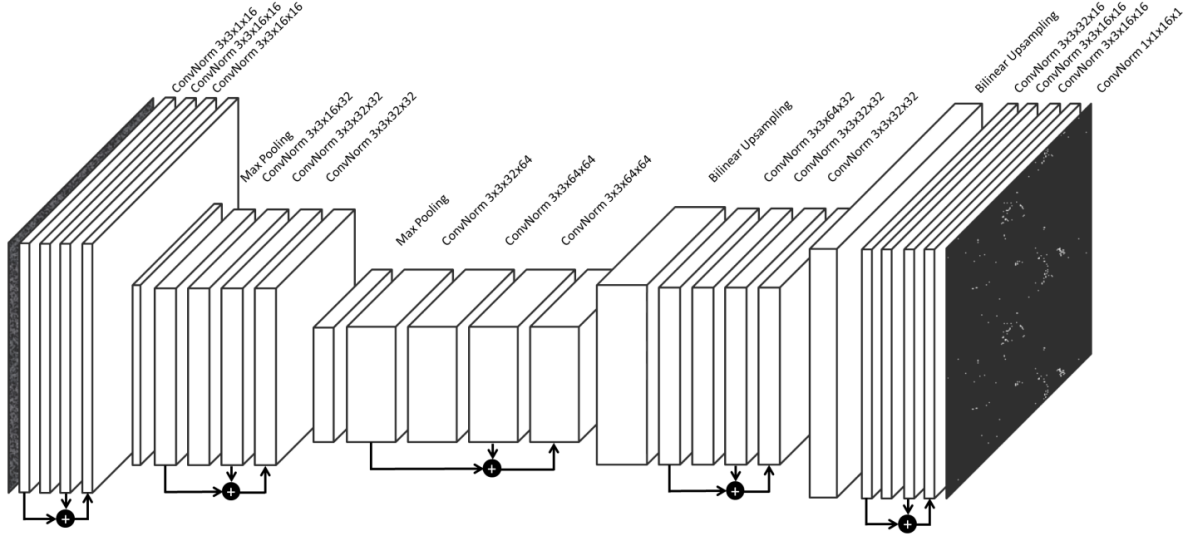


Fig. (2) **Architecture de DetNet.** Sur cet illustration, les flèches noires correspondent à des « connections résiduelles » qui consistent à ajouter, à la sortie de deux couches convolutionnelles consécutives, l'image (ou la *feature map*) d'entrée. Leur but est de réduire le problème de *vanishing gradient* dû à la profondeur du réseau et donc faciliter son apprentissage.

a par exemple été comparé avec la transformation *h-dome* [7]. Dans la plupart des scénarios (qui diffèrent selon la taille, la forme des particules et le rapport signal sur bruit), DetNet est la méthode la plus performante. Ces performances ont été évaluées par le score F1 qui est la moyenne harmonique de la précision et du rappel. Par ailleurs, les performances de localisation ont été évaluées par la racine de l'erreur quadratique moyenne.

La potentielle difficulté à distinguer les particules qui se recouvrent n'a pas été évoquée dans cette article.

2.2.1.2 deepBlink deepBlink (Eichenberger *et al.*, 2021) [12] est un réseau de neurones basé sur un réseau U-Net [13] prédisant, pour chacune des cellules d'une grille (plus petite que l'image originale), la probabilité que la zone de l'image correspondante contienne une particule, ainsi que la localisation exacte de la particule. La taille des cellules impactent

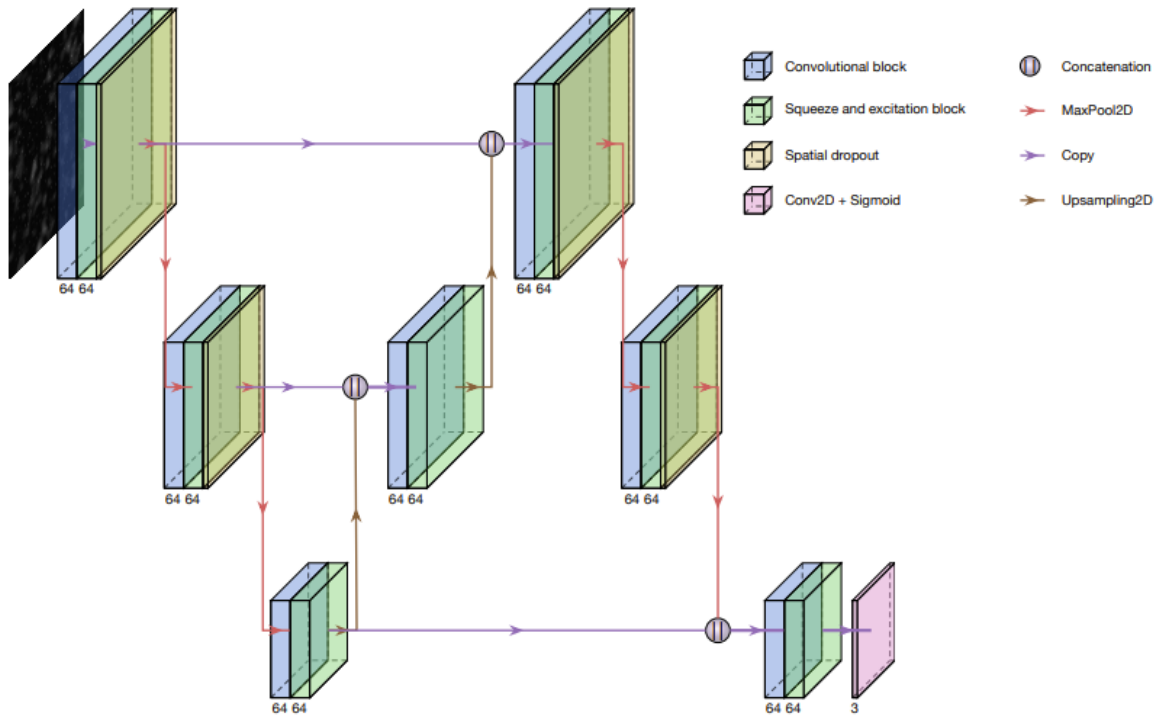


Fig. (3) **Architecture de deepBlink.** Le réseau U-Net, qui est reconnaissable par sa forme en "V", est situé à gauche. Les 2 couches de droite servent à réduire la dimension spatiale de l'image originale afin de générer la grille contenant les cellules. Les *Squeeze-and-Excitation blocks* [14] sont des petits réseaux de neurones qui calculent, à partir des sorties des neurones qui les précèdent, une pondération pour chaque canal.

fortement les performances du réseau. Si elle est inférieure à la taille optimale, la proportion de faux positifs sera artificiellement augmentée. Cela pose problème puisque les réseaux de neurones apprennent aussi la répartition des classes. Si la taille est supérieure à la taille optimale, certaines cellules contiendront plusieurs particules alors qu'à chacune d'entre elles est associée une seule détection potentielle.

Il a été montré que les performances de deepBlink surpassent celles de DetNet sur 2 jeux de données réels et 4 jeux de données synthétiques. Ces performances ont été évaluées par :

- les mêmes formules utilisées dans l'article de DetNet ;
- une nouvelle *metric*, introduite dans l'article, qui combine les deux formules précédemment citées.

Par ailleurs, dans le but de déterminer l'importance de chaque composant du réseau, des expériences d'ablation ont été réalisées où les composantes individuelles ont été retirées afin de mesurer la diminution de performance associée. Cela a montré que l'usage du coefficient de Dice dans la fonction de coût est le facteur améliorant le plus les performances. Comme cela est expliqué dans l'article, ce n'est pas surprenant puisque, dans une autre étude, [15] la supériorité du coefficient de Dice a été montré dans les cas où les classes du jeu de données sont déséquilibrées.

2.2.1.3 DeepSpot DeepSpot (Bouilhol *et al.*, 2022) [16] est un réseau dont le but est de générer une image que l'on peut ensuite ajouter à l'image originale afin de rehausser l'intensité de ses particules d'ARN et de les rendre plus facilement détectables.

Les performances de détection de spots en aval a été évaluée sur 20 jeux de données simulées et 3 jeux de données réelles à l'aide d'un détecteur (Olivo-Marin, 2002) [4] basé sur les ondelettes et mis à disposition dans la plateforme Icy [17]. Il est important de noter que le réseau est entraîné de sorte à ce que l'intensité des spots de l'image originale soit rehaussée sans élargir les spots, ce qui est primordial puisque des spots proches peuvent être susceptibles de fusionner si on les élargit.

Les nombres de faux négatifs et de faux positifs ont été déterminés à l'aide de l'algorithme *k-d tree* [18] qui a l'avantage de mémoriser les concordances ambiguës qui peuvent arriver lorsque la localisation d'une détection est très proche de celles deux particules réelles.

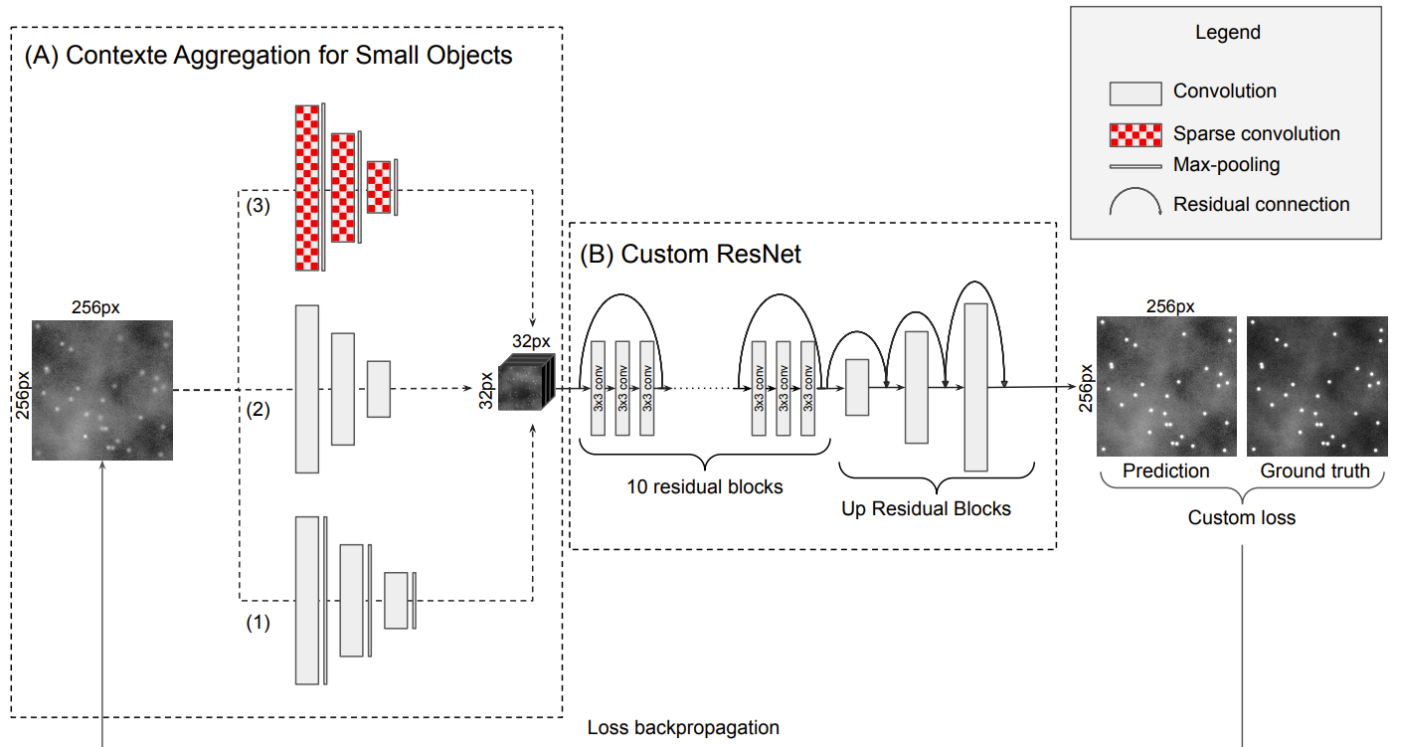


Fig. (4) **Architecture de DeepSpot.** Elle est constituée d'un module *CASO* (Context Aggregation for Small Objects) qui est un réseau neuronal à chemins multiples et peut apprendre des caractéristiques plus complètes et complémentaires qu'un seul chemin. En particulier, l'objectif de la convolution à trous est d'apporter plus de contexte autour des petits spots, tandis que les deux autres chemins agrègent les informations sémantiques des spots brillants et des spots moins brillants pour le premier et le deuxième chemin respectivement. Les résultats des trois chemins d'encodage sont ensuite concaténés pour construire un vecteur de caractéristiques plus long contenant les informations extraites par chaque chemin. Le décodeur ainsi que la fin de la partie de l'encodeur correspondent à un réseau ResNet modifié.

2.2.1.4 Conclusion L'absence d'étape de débruitage et le fait qu'elles se basent toutes sur l'utilisation de réseaux convolutionnels sont les points communs de ces méthodes modernes.

DetNet est le réseau le moins performant malgré le fait qu'il requiert un nombre de paramètres significativement moins élevé que ceux de DeepSpot et deepBlink. La réduction voulue de son nombre de paramètres n'a pas été justifiée et n'est en réalité pas désirable puisqu'elle engendre probablement une baisse des performances. Par ailleurs, un nombre de paramètres élevés pourraient être en partie compensé par l'utilisation de données synthétiques générées à partir de méthodes qui seraient basées sur des réseaux antagonistes génératifs [19].

Les performances offertes par deepBlink sont meilleures que celles de DetNet. Intuitivement, cela se comprend par le fait qu'il est basé sur un réseau U-Net, réputé pour sa capacité à classifier d'une manière remarquable chacun des pixels d'une image. Dans un tel réseau, avant chaque sur-échantillonnage (de la partie décodeur du réseau U-Net) une convolution est appliquée sur la concaténation d'informations de contextes (provenant de la sortie de la couche précédente du décodeur) et d'informations spatiales (provenant de la sortie de l'encodeur du U-Net située à la même échelle spatiale). Sans ces connexions¹ entre chaque étage de l'encodeur et du décodeur, l'image de sortie serait uniquement déduite à partir d'informations abstraites issues de l'encodeur et donc dépourvues d'information spatiales dont la perte est due aux opérations de sous-échantillonnages successives de la partie encodeur. Toutefois, le fait que chaque détection potentielle doit être associée à une seule cellule de la grille contraint l'utilisateur de la méthode à devoir trouver la taille optimale des cellules pour chaque type d'objet à détecter.

Bien que DeepSpot offre les meilleures performances, il ne détecte pas directement les spots et nécessite donc l'utilisation d'un autre algorithme pour la détection.

La supériorité des performances de DeepSpot par rapport à celles de DeepBlink s'expliquent probablement par :

- l'encodage plus complet qui est réalisé grâce au module CASO ;
- le seuillage effectué à l'aide d'une méthode robuste basée sur les ondelettes.

2.2.2 Méthodes plus générales dédiées à la segmentation d'objets en imagerie biologique

2.2.2.1 StarDist et SplineDist : deux méthodes basées sur la même philosophie Les réseaux des méthodes StarDist (Schmidt *et al.*, 2018) [20] et SplineDist (Mandal *et al.*, 2021) [21] sont basés sur un réseau U-Net et consistent à prédire, pour chacun des pixels de l'image originale, non seulement la probabilité qu'il y ait un objet mais aussi les valeurs de variables permettant de caractériser le contour de l'objet à segmenter. L'élimination des contours qui ne correspondent à aucun objet réel se fait via l'algorithme de *Non-Maximum Suppression* (NMS). La fonction de coût des réseaux de

Algorithme NMS

Entrée: les seuils s_1 et s_2

Sortie: l'ensemble l des polygones qui n'ont pas été supprimés

$l \leftarrow$ l'ensemble de tous les polygones

supprimer les polygones dont le centre est un pixel ayant une probabilité inférieure à s_1

tant que l n'est pas vide **faire**

$meilleur_p \leftarrow$ le polygone de l associé à la plus grande probabilité

$l \leftarrow l \setminus \{meilleur_p\}$

pour tout $p \in l$ **faire**

$A \leftarrow surface(meilleur_p)$

$B \leftarrow surface(p)$

si $\frac{A \cap B}{A \cup B} > s_2$ **alors**

$l \leftarrow l \setminus \{p\}$

fin si

fin pour

fin tant que

ces deux méthodes encourage la minimisation des aires des surfaces délimitées par les contours prédits ne correspondant à aucun objet réel. Ainsi, lors de l'étape de NMS, les surfaces délimitées par ces contours auront plus de chance d'être supprimées puisque leur intersection avec celles des objets réels sera petite.

La sortie finale du U-Net produit un tenseur T de taille $L \times H \times 128$ où L et H correspondent respectivement à la largeur et à la hauteur de l'image donnée à l'entrée du réseau. Une convolution est appliquée au tenseur T avec un filtre de taille $3 \times 3 \times 128$ afin de produire une matrice M de même taille que celle de l'image originale.

Ensuite, une convolution puis une fonction Sigmoidale sont appliquées à la matrice M afin de générer une carte p prédisant, pour chacun des pixels, la probabilité que ce pixel corresponde à une particule.

1. Elles servent aussi à atténuer le problème de *vanishing gradient*.

En plus de cela, n autres convolutions sont appliquées à la matrice M afin de produire n autres matrices. Chacune de ces matrices contient, pour chacun des pixels, une des n variables servant à caractériser le contour de l'objet.

2.2.2.2 StarDist StarDist modélise le contour de chaque objet du pixel (i, j) par un polygone. Les variables prédites par le réseau sont les R distances $D_{ij} = \left\{ d_{ij}^k \right\}_{k=0, \dots, R-1}$ situées entre le centre et les sommets du polygone. L'angle séparant les distances radiales de deux sommets consécutifs est constant.

La fonction de coût de StarDist est définie par :

$$\mathcal{L}_{\text{StarDist}} \left(Pos, p, \hat{p}, D, \hat{D} \right) = \frac{1}{|\text{Pos}|} \sum_{(i,j) \in \text{Pos}} \mathcal{L}_{\text{ECB}} \left(p_{ij}, \hat{p}_{ij} \right) + \lambda_1 \mathcal{L}_{\text{rayons}} \left(p_{ij}, D_{ij}, \hat{D}_{ij} \right),$$

avec :

$$\mathcal{L}_{\text{rayons}} \left(p_{ij}, D_{ij}, \hat{D}_{ij} \right) = p_{ij} \cdot \mathbb{1}_{p_{ij} > 0} \cdot \frac{1}{R} \sum_{k=0}^{R-1} \left| d_{ij}^k - \hat{d}_{ij}^k \right| + \lambda_2 \cdot \mathbb{1}_{p_{ij} = 0} \cdot \frac{1}{R} \sum_{k=0}^{R-1} \left| \hat{d}_{ij}^k \right|,$$

où :

- Pos est l'ensemble de positions de l'image ;
- ECB désigne l'entropie croisée binaire ;
- $p_{i,j}$ et $\hat{p}_{i,j}$ sont respectivement les probabilités réelles et prédites qu'il y ait un objet situé à la position (i, j) dans l'image ;
- λ_1 et λ_2 sont des hyper-paramètres servant simplement à pondérer les différents termes.

StarDist a été évalué sur deux jeux de données synthétiques et un jeu de donnée réel et s'est révélé être un réseau dont les performances dépassent celles de l'état de l'art. Il a également été comparé aux réseaux U-Net.

2.2.2.3 SplineDist SplineDist modélise le contour de chaque objet du pixel (i, j) par une courbe spline paramétrique 2D définie par :

$$\mathbf{s}(t) = \begin{bmatrix} s_x(t) \\ s_y(t) \end{bmatrix} = \sum_{k=0}^{M-1} \mathbf{c}[k] \varphi_M(t - k),$$

où :

- $\mathbf{c}[k] = \left(c_x^k, c_y^k \right) = \left(d^k \cos(\alpha^k), d^k \sin(\alpha^k) \right)$ sont les points de contrôles paramétrisant la courbe ;
- $\varphi_M(t) = \sum_{m \in \mathbb{Z}} \varphi(t - Mm)$ permet d'avoir un contour fermé constitué de B-Splines φ de degré 3 ;
- M est le nombre d'angles et de distances prédites par le réseau.

La fonction de coût de SplineDist est définie par :

$$\mathcal{L}_{\text{SplineDist}} \left(Pos, p, \hat{p}, P, S \right) = \frac{1}{|\text{Pos}|} \sum_{(i,j) \in \text{Pos}} \mathcal{L}_{\text{ECB}} \left(p_{ij}, \hat{p}_{ij} \right) + \lambda_1 \cdot \mathcal{L}_{\text{contour}} \left(p_{ij}, P_{ij}, S_{ij} \right),$$

avec :

$$\mathcal{L}_{\text{contour}} \left(p_{ij}, P_{ij}, S_{ij} \right) = p_{ij} \cdot \frac{1}{N} \sum_{n=0}^{N-1} \left| \mathbf{p}_{ij}^n - \mathbf{s}_{ij}^n \right| + \lambda_2 \cdot \mathbb{1}_{p_{ij} = 0} \cdot \frac{1}{N} \sum_{n=0}^{N-1} \left| \mathbf{s}_{ij}^n \right|,$$

où $S_{ij} = \left\{ \mathbf{s}_{ij}^n \right\}_{n=0, \dots, N-1}$ est l'ensemble des N points obtenus après échantillonnage uniforme de $\mathbf{s}_{ij}(t)$:

$$\mathbf{s}_{ij}^n = \mathbf{s}_{ij}(t) \Big|_{t = \frac{nM}{N}}.$$

N peut être aussi grand que l'on souhaite et ne doit donc pas être nécessairement égal au nombre M de points de contrôle. Il est préférable que N soit supérieur à M puisque plus N est grand, plus le nombre d'ensembles de points de contrôle pouvant aboutir au même contour sera petit et plus la fidélité de la fonction de coût sera élevée.

SplineDist a été évalué sur un jeu de données réelles. Il a été déduit que SplineDist offre des performances similaires à celles de StarDist tout en permettant une plus grande variété de contours possibles.

2.2.2.4 Conclusion StartDist et Spline ont l'avantage d'être des solutions complètes. Mis à part les hyper-paramètres des fonctions de coût, les deux seuls autres hyper-paramètres nécessaires sont les seuils s_1 et s_2 utilisés dans l'algorithme NMS. L'interprétabilité de ces deux hyper-paramètres les rend facilement optimisables. Cette façon de modéliser l'apparence des objets par un contour rend donc ces méthodes plus facilement adaptables à différents jeux de données.

2.3 Conclusion

Parmi les méthodes récentes dédiées à la détection de spots :

- **DeepSpot est la méthode la plus performante.** En plus d'utiliser un réseau convolutionnel classique pour l'encodage de l'apparence des spots, les auteurs de cette méthode ont supposé que l'utilisation de différents réseaux en parallèle permet d'apprendre des caractéristiques complémentaires afin d'encoder de manière plus robuste.
- **SplineDist est la méthode proposant la modélisation des objets la plus polyvalente.** Même si le contour des spots n'est pas clairement définissable, l'approximer par un contour fermé laisse tout de même la possibilité à chacun des pixels d'être le centre d'un objet. Cette modélisation est, par exemple, beaucoup plus flexible que celle de deepBlink qui suppose que chaque cellule de taille fixe contient au maximum le centre d'un seul spot.

Références

- [1] Bo ZHANG, Josiane ZERUBIA et Jean-Christophe OLIVO-MARIN. "Gaussian approximations of fluorescence microscope point-spread function models". In : *Appl. Opt.* 46.10 (avr. 2007), p. 1819-1829. DOI : 10.1364/AO.46.001819. URL : <https://opg.optica.org/ao/abstract.cfm?URI=ao-46-10-1819>.
- [2] Ihor SMAL et al. "Quantitative Comparison of Spot Detection Methods in Fluorescence Microscopy". In : *IEEE Transactions on Medical Imaging* 29.2 (2010), p. 282-301. DOI : 10.1109/TMI.2009.2025127.
- [3] B. ZHANG et al. "Multiscale Variance-Stabilizing Transform for Mixed-Poisson-Gaussian Processes and its Applications in Bioimaging". In : *2007 IEEE International Conference on Image Processing*. T. 6. 2007, p. VI - 233-VI-236. DOI : 10.1109/ICIP.2007.4379564.
- [4] Jean-Christophe OLIVO-MARIN. "Extraction of spots in biological images using multiscale products". In : *Pattern Recognition* 35.9 (2002), p. 1989-1996. ISSN : 0031-3203. DOI : [https://doi.org/10.1016/S0031-3203\(01\)00127-3](https://doi.org/10.1016/S0031-3203(01)00127-3). URL : <https://www.sciencedirect.com/science/article/pii/S0031320301001273>.
- [5] David S. BRIGHT et Eric B. STEEL. "Two-dimensional top hat filter for extracting spots and spheres from digital images". In : *Journal of Microscopy* 146.2 (1987), p. 191-200. DOI : <https://doi.org/10.1111/j.1365-2818.1987.tb01340.x>. eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2818.1987.tb01340.x>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2818.1987.tb01340.x>.
- [6] G. H. Joss E. J. BREEN et K. L. WILLIAMS. "Locating objects of interest within biological images : The top hat box filter". In : *Journal of Computer-Assisted Microscopy* 3.2 (1991), p. 97-102.
- [7] L. VINCENT. "Morphological grayscale reconstruction in image analysis : applications and efficient algorithms". In : *IEEE Transactions on Image Processing* 2.2 (1993), p. 176-201. DOI : 10.1109/83.217222.
- [8] Shan JIANG et al. "Detection of molecular particles in live cells via machine learning". In : *Cytometry Part A* 71A.8 (2007), p. 563-575. DOI : <https://doi.org/10.1002/cyto.a.20404>. eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cyto.a.20404>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/cyto.a.20404>.
- [9] Antoine BASSET et al. "Adaptive Spot Detection With Optimal Scale Selection in Fluorescence Microscopy Images". In : *IEEE Transactions on Image Processing* 24.11 (2015), p. 4512-4527. DOI : 10.1109/TIP.2015.2450996.
- [10] T. WOLLMANN et al. "Detnet : Deep Neural Network For Particle Detection In Fluorescence Microscopy Images". In : *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. 2019, p. 517-520. DOI : 10.1109/ISBI.2019.8759234.
- [11] Hyeonwoo NOH, Seunghoon HONG et Bohyung HAN. "Learning Deconvolution Network for Semantic Segmentation". In : *CoRR* abs/1505.04366 (2015). arXiv : 1505.04366. URL : <http://arxiv.org/abs/1505.04366>.
- [12] Bastian Th EICHENBERGER et al. "deepBlink : threshold-independent detection and localization of diffraction-limited spots". In : *Nucleic Acids Research* 49.13 (juill. 2021), p. 7292-7297. ISSN : 0305-1048. DOI : 10.1093/nar/gkab546. eprint : <https://academic.oup.com/nar/article-pdf/49/13/7292/39117119/gkab546.pdf>. URL : <https://doi.org/10.1093/nar/gkab546>.
- [13] Olaf RONNEBERGER, Philipp FISCHER et Thomas BROX. *U-Net : Convolutional Networks for Biomedical Image Segmentation*. 2015. DOI : 10.48550/ARXIV.1505.04597. URL : <https://arxiv.org/abs/1505.04597>.
- [14] Jie HU et al. *Squeeze-and-Excitation Networks*. 2017. DOI : 10.48550/ARXIV.1709.01507. URL : <https://arxiv.org/abs/1709.01507>.
- [15] Carole H. SUDRE et al. "Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations". In : *CoRR* abs/1707.03237 (2017). arXiv : 1707.03237. URL : <http://arxiv.org/abs/1707.03237>.
- [16] Emmanuel BOUILHOL et al. "DeepSpot : A deep neural network for RNA spot enhancement in single-molecule fluorescence in-situ hybridization microscopy images". In : *Biological Imaging* 2 (2022), e4. DOI : 10.1017/S2633903X22000034.
- [17] Fabrice de CHAUMONT et al. "Icy : an open bioimage informatics platform for extended reproducible research". In : *Nature Methods* 9.7 (juill. 2012). Number : 7 Publisher : Nature Publishing Group, p. 690-696. ISSN : 1548-7105. DOI : 10.1038/nmeth.2075. URL : <https://www.nature.com/articles/nmeth.2075> (visité le 11/07/2022).

- [18] Jon Louis BENTLEY. “Multidimensional Binary Search Trees Used for Associative Searching”. In : *Commun. ACM* 18.9 (sept. 1975), p. 509-517. ISSN : 0001-0782. DOI : 10.1145/361002.361007. URL : <https://doi.org/10.1145/361002.361007>.
- [19] Ian J. GOODFELLOW et al. *Generative Adversarial Networks*. 2014. DOI : 10.48550/ARXIV.1406.2661. URL : <https://arxiv.org/abs/1406.2661>.
- [20] Uwe SCHMIDT et al. “Cell Detection with Star-Convex Polygons”. In : *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II*. 2018, p. 265-273. DOI : 10.1007/978-3-030-00934-2_30.
- [21] Soham MANDAL et Virginie UHLMANN. “Splinedist : Automated Cell Segmentation With Spline Curves”. In : *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*. 2021, p. 1082-1086. DOI : 10.1109/ISBI48211.2021.9433928.