



HAL
open science

Adaptation de domaine par une méthode d'apprentissage profond parcimonieuse. Application à la segmentation d'images biomédicales.

Alexandre Stenger, Luc Vedrenne, Patrick Schultz, Sylvain Faisan, Benoît Naegel,
Étienne Baudrier

► To cite this version:

Alexandre Stenger, Luc Vedrenne, Patrick Schultz, Sylvain Faisan, Benoît Naegel, et al.. Adaptation de domaine par une méthode d'apprentissage profond parcimonieuse. Application à la segmentation d'images biomédicales.. ORASIS 2023, Laboratoire LIS, UMR 7020, May 2023, Carqueiranne, France. <hal-04219289>

HAL Id: hal-04219289

<https://hal.science/hal-04219289v1>

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

Adaptation de domaine par une méthode d'apprentissage profond parcimonieuse. Application à la segmentation d'images biomédicales.

Alexandre Stenger^{1 ‡} Luc Vedrenne^{1 ‡} Patrick Schultz² Sylvain Faisan¹
Étienne Baudrier¹ Benoît Naegel¹

¹ ICube, UMR 7357 CNRS, Université de Strasbourg, France

² IGMBC, UMR 7104 CNRS, U1258 INSERM, Université de Strasbourg, France

[‡] alstenger@unistra.fr - vedrenne@unistra.fr

Résumé

La microscopie électronique permet aux biologistes de visualiser des coupes de cellules à l'échelle nanométrique dans lesquelles les structures biologiques peuvent être analysées quantitativement. Cette analyse requiert l'annotation des contours par un expert, ce qui est très coûteux en temps. L'automatisation de l'annotation (segmentation) représente donc un enjeu important. On constate qu'un réseau de neurones arrive à bien généraliser la segmentation d'une base d'images à une autre si elles sont acquises selon le même protocole. Cependant les performances de généralisation peuvent chuter si le protocole d'acquisition diffère entre les deux bases car cela peut beaucoup modifier les caractéristiques des images. Pour résoudre ce problème, nous proposons une méthode d'adaptation de domaine non supervisée (UDA) s'appuyant sur la normalisation par lots (batch normalization ou encore BatchNorm). Notre méthode a été expérimentée dans le contexte de la segmentation de mitochondries et a donné de meilleurs résultats que les méthodes de l'état de l'art. De plus, elle utilise peu de ressources de calcul, ce qui permet une utilisation presque en temps réel. Enfin, nous proposons une métrique basée sur la distance de Wasserstein pour évaluer l'effet de cette normalisation. Nous montrons empiriquement que cette mesure est log-linéairement corrélée avec la baisse de performances de la segmentation adaptée et qu'elle met en lumière le gain obtenu par notre méthode d'adaptation.

Mots Clef

Adaptation de domaine, apprentissage profond, segmentation, microscopie électronique, FIB-SEM.

Abstract

Electron microscopy imaging techniques allow biologists to obtain nanoscale slice volumes of cells in which biological structures can be statistically analyzed under the condition that the tedious task of segmentation can be automated. If a learned segmentation has good results on images obtained under the same acquisition process, variations in

this acquisition can drastically affect the final image's features and can lead to a drop of segmentation performances. To tackle this problem, we propose an Unsupervised Domain Adaptation (UDA) framework based on Batch Normalization. Our method tested on mitochondria segmentation shows better results than state-of-the-art methods and uses few computation resources, allowing a near real-time use. Furthermore, we propose a metric based on the Wasserstein distance to evaluate the effect of this normalization. We empirically show that this measure is log-linearly correlated with the drop of performances of the adapted segmentation and sheds light on the gain obtained by our UDA framework.

Keywords

Domain Adaptation, Deep Learning, Segmentation, electron microscopy, FIB-SEM.

1 Introduction

Dans le contexte de la microscopie électronique à balayage avec faisceau d'ions focalisés (FIB-SEM), un réseau de neurones dédié à la segmentation sémantique est capable de prédire le masque de segmentation lorsqu'il est entraîné sur des cellules imagées avec un même processus d'acquisition [2, 3]. Toutefois, de grands décalages de domaine peuvent apparaître entre deux images acquises selon des protocoles différents. Par exemple, on peut voir sur la Fig. 1(a) une fixation chimique et sur la Fig. 1(c) une cryofixation à côté de leur masque de segmentation mitochondriale respectif Fig. 1(b)(d). Ainsi, entre les processus de fixation chimique (Fig. 1(a)) et de cryofixation (Fig. 1(c)), les mitochondries changent fortement d'apparence en termes de forme, de contraste et de texture, ce qui entraîne une baisse de performance lorsqu'on tente de généraliser d'une base à l'autre. Cette situation conduit soit à faire appel à des annotations manuelles qui sont très coûteuses en temps, soit à adapter le modèle d'une manière ou d'une autre.

L'adaptation de domaine (AD) a suscité un grand intérêt avec le développement de l'apprentissage profond. Dans

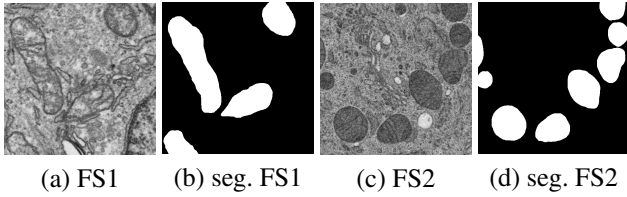


FIGURE 1 – Exemple de patches 256×256 de nos ensembles FS1 et FS2 (a,c) tels que décrits en partie 4.1, avec les vérités terrains des mitochondries (b,d).

le contexte de la segmentation sémantique, la plupart des travaux se concentrent encore sur l’adaptation des jeux de données Cityscapes, GTA5 et SYNTHIA [14]. L’application de ces mêmes méthodes aux images biomédicales s’avère très difficile, en raison de la petite taille des ensembles de données, du manque de structure claire et de la forte variation interne de ce type d’image. En ce qui concerne l’adaptation de domaine pour la segmentation cellulaire, *Haq et al.* [1] ont introduit des méthodes adversariales (basées sur les GAN) pour l’adaptation de domaines entre les données histologiques, et *Li et al.* [21] ont utilisé la même méthode. *Roels et al.* ont étudié l’architecture Y-Net pour la segmentation cellulaire en microscopie électronique [20], tandis que *Franco et al.* ont introduit l’Attention Y-Net [19]. Ces réseaux sont difficiles à entraîner, prennent du temps et consomment beaucoup d’énergie. D’un autre côté, il est possible de faire de l’adaptation de domaine en recalculant les statistiques des couches BatchNorm. Cette méthode, rapide et facile à utiliser, a attiré notre attention. De plus, cette couche BatchNorm est autant capable de réduire le décalage interne des covariables [17] que le décalage inter-domaine [18]. Nous proposons un cadre d’adaptation basé sur le recalcul des BatchNorm et montrons que dans certains cas, il fonctionne même avec de petits ensembles de données, ce qui le rend adapté aux images biomédicales. Nous introduisons également une métrique pour quantifier la qualité de l’estimation statistique, et montrons qu’elle est un bon indicateur des performances après adaptation. En résumé, cette méthode est rapide, nécessite peu de ressources de calcul, ne nécessite pas de nouvel apprentissage du réseau et peut être utilisée par les biologistes eux-mêmes.

2 État de l’Art

Les réseaux de neurones convolutionnels (CNN) avec BatchNorm (BN) ont été introduits par *Ioffe et al.* [17] en 2015. L’idée d’utiliser la BatchNorm dans un cadre d’adaptation de domaine est apparue 3 ans plus tard avec AdaBN [18]. Ce travail a servi de base à de multiples développements : DSBN [16] crée une nouvelle couche de CNN qui possède des sous-couches de BatchNorm spécifiques au domaine. UBNA [15] améliore ce travail en pondérant le calcul des statistiques des BatchNorm en fonction de leur profondeur dans le réseau. Dans notre travail, nous utili-

sons simplement l’idée originale d’AdaBN.

Nisar et al. [11] et *Stacke et al.* [12] ont proposé des mesures très similaires pour quantifier le changement de domaine. Cependant, ils font tous deux l’hypothèse qu’une telle métrique doit être symétrique, ce qui ne correspond pas aux performances observées. De plus, afin de calculer la distance de Wasserstein, généralement très coûteuse, ils se basent sur un cas unidimensionnel. Ils prennent la distribution des moyennes des activations d’un réseau de neurones, ce qui suppose l’hypothèse – non validée à notre connaissance – que ces activations présentent une très faible variance. *Schneider et al.* [7] ont étudié l’effet de la normalisation par lots, et ont exploité la forme fermée de la distance de Wasserstein entre deux lois gaussiennes, qui s’adapte au contexte des activations des réseaux de neurones. Ils ont introduit une normalisation de la distance par la distribution source pour prendre en compte l’asymétrie de la tâche. Cependant, Leur travail se concentre sur la robustesse au bruit pour la classification.

3 Méthodes

Dans une configuration habituelle en apprentissage supervisé pour la segmentation, nous considérons un ensemble d’apprentissage $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n$ avec des images $\mathbf{x}_i \in \mathbb{R}^{h \times w}$ et des masques de segmentation binaires $\mathbf{y}_i \in \{0, 1\}^{h \times w}$. Nous désignons par \mathcal{P} la distribution conjointe sous-jacente des données et par $\hat{\mathcal{P}} = \frac{1}{n} \sum_i \delta_{\mathbf{x}_i, \mathbf{y}_i}$ son estimation empirique. Notre objectif est d’apprendre une fonction $f : \mathbb{R}^{h \times w} \mapsto \{0, 1\}^{h \times w}$ prédisant un masque à partir d’une image, par la minimisation du risque empirique $\hat{\mathcal{R}}(f) = \mathcal{R}_{\hat{\mathcal{P}}}(f) = \mathbb{E}_{\mathbf{x}, \mathbf{y} \sim \hat{\mathcal{P}}} [L(\mathbf{y}, f(\mathbf{x}))]$, où L mesure l’écart entre l’étiquette réelle et l’étiquette prédite. Dans la configuration d’adaptation de domaine non supervisée, deux domaines \mathcal{D}^s (source) et \mathcal{D}^t (cible) représentés par leurs distributions \mathcal{P}^s et \mathcal{P}^t sont considérés. L’objectif est d’apprendre une fonction f qui prédit sur \mathcal{P}^t en utilisant les images et les étiquettes de \mathcal{P}^s mais uniquement les images de \mathcal{P}^t .

3.1 Rappel : la normalisation par lots

Dans un réseau, la normalisation par lots est effectuée dans des couches intermédiaires appelées *couches de BatchNorm*. Soit X l’entrée et Y la sortie d’une couche BatchNorm. Dans un schéma d’inférence classique d’apprentissage profond, la couche BatchNorm calcule

$$Y = \gamma \times \frac{X - \hat{\mu}}{\sqrt{\hat{\sigma}^2 + \epsilon}} + \beta \quad (1)$$

où γ et β sont des paramètres apprenables tandis que $\hat{\mu}$ et $\hat{\sigma}^2$ sont la moyenne et la variance estimées sur l’ensemble d’apprentissage. Notez que $\hat{\mu}$ et $\hat{\sigma}^2$ sont généralement fixés dans un cadre d’inférence classique. Le paragraphe suivant présente l’utilisation de la BatchNorm en adaptation de domaine.

3.2 Normalisation par lots pour une adaptation de domaine en temps réel

Soit un domaine source \mathcal{D}_S représenté par sa distribution source \mathcal{P}^s et n échantillons (x_s, y_s) de paires image-étiquette de \mathcal{D}_S . De la même manière, nous avons \mathcal{D}_T représenté par sa distribution source \mathcal{P}^t et des échantillons d'image x_t provenant de \mathcal{D}_T (dans le cadre de l'UDA, nous n'avons pas accès aux étiquettes cibles).

Nous proposons une procédure en deux étapes :

Première étape : Nous entraînons un U-Net [5] de manière supervisée sur le domaine source en utilisant des paires image-étiquette source $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n, (\mathbf{x}_i, \mathbf{y}_i) \sim \hat{\mathcal{P}}^s$.

Seconde étape : Nous gelons toutes les couches du U-Net sauf les BatchNorms. Pour chaque couche BatchNorm, nous recalculons leurs statistiques $\hat{\mu}$ et $\hat{\sigma}^2$ en utilisant tous les échantillons cibles de la manière suivante. Soit B le nombre de lots de données. Soit $\hat{\mu}_s$ et $\hat{\sigma}_s^2$ les statistiques calculées sur le domaine source tel que défini dans la formule (1). Pour chaque lot (indexé par j) et à chaque couche BatchNorm, soit \mathbf{z}_t^j la sortie de la couche convolutive précédente, et $\hat{\mu}_t^j, (\hat{\sigma}_t^j)^2$ les statistiques estimées à l'étape j sur le domaine cible. Le recalcul par une couche BatchNorm est décrit dans l'Algorithme 1.

Algorithm 1 Recalcul de la BatchNorm dans le cadre de l'UDA

- 1: $\hat{\mu}_t^0 \leftarrow \hat{\mu}_s$
 - 2: $(\hat{\sigma}_t^0)^2 \leftarrow (\hat{\sigma}_s)^2$
 - 3: **for** $i \in \{1, \dots, B\}$ **do**
 - 4: $\mu_t^j \leftarrow \mathbb{E}[\mathbf{z}_t^j]$
 - 5: $(\sigma_t^j)^2 \leftarrow \text{Var}[\mathbf{z}_t^j]$
 - 6: $\hat{\mu}_t^j \leftarrow (1 - \alpha)\hat{\mu}_t^{j-1} + \alpha\mu_t^j$
 - 7: $(\hat{\sigma}_t^j)^2 \leftarrow (1 - \alpha)(\hat{\sigma}_t^{j-1})^2 + \alpha(\sigma_t^j)^2$
 - 8: **end for**
 - 9: $\hat{\mu}_t \leftarrow \hat{\mu}_t^B$
 - 10: $(\hat{\sigma}_t)^2 \leftarrow (\hat{\sigma}_t^B)^2$
-

Comme nous le verrons dans la partie 4, cette procédure permet l'adaptation de domaine. Le paragraphe suivant propose des éléments pour comprendre ces résultats.

3.3 Interpréter l'effet de la normalisation par lots

Soit $p_s, p_t : \{0, 1\}^{h \times w} \times \mathbb{R}^{h \times w} \mapsto \mathbb{R}^+$ les densités de probabilité des distributions source et cible respectivement. Nous avons $p_s(\mathbf{y}, \mathbf{x}) = p_s(\mathbf{y}|\mathbf{x})p_s(\mathbf{x})$ et $p_t(\mathbf{y}, \mathbf{x}) = p_t(\mathbf{y}|\mathbf{x})p_t(\mathbf{x})$. En suivant l'exemple de [10], nous définissons le décalage des covariables comme les deux distributions conjointes ayant des probabilités conditionnelles identiques mais des marginales différentes, c'est-à-dire $p_s(\mathbf{y}|\mathbf{x}) = p_t(\mathbf{y}|\mathbf{x})$ et $p_s(\mathbf{x}) \neq p_t(\mathbf{x})$. Soit $\mathbb{E}_{\{s,t\}}$ et $\mathbb{V}_{\{s,t\}}$ les moyennes et variances réelles des distributions source et cible, de sorte que $\tilde{f}_{\{s,t\}} = (f(\mathbf{x}) - \mathbb{E}_{\{s,t\}}[f(\mathbf{x})]) (\mathbb{V}_{\{s,t\}}[f(\mathbf{x})])^{-1/2}$ sont les distri-

butions normalisées des activations d'un réseau de neurones. Soit \mathbb{E}_r et \mathbb{V}_r les moyennes et variances réelles des distributions source (avec $r = s$) et cible (avec $r = t$), de sorte que $\tilde{f}_r = (f(\mathbf{x}) - \mathbb{E}_r[f(\mathbf{x})]) (\mathbb{V}_r[f(\mathbf{x})])^{-1/2}$ avec $r \in \{s, t\}$ les distributions normalisées des activations d'un réseau de neurones. Si les distributions d'activation de la source et de la cible sont des distributions gaussiennes, on a $p(\tilde{f}_s(x)|x)p_s(x) \approx p(\tilde{f}_t(x)|x)p_t(x)$. et le décalage des covariables est compensé. Notre méthode vise à approximer une telle compensation en estimant les deux premiers moments de la distribution cible. Nous affirmons que l'erreur d'estimation entre les domaines source et cible est presque entièrement due à la qualité de cette estimation, sous l'hypothèse gaussienne, qui est généralement considérée comme vérifiée dans le cadre des activations des réseaux de neurones. Cette qualité peut être mesurée par le carré de la 2-distance de Wasserstein entre les moments réels (μ_t, σ_t^2) et ceux estimés par le réseau adapté $(\hat{\mu}_t, \hat{\sigma}_t^2)$, qui a une forme fermée suivante pour les distributions gaussiennes [8] :

$$\mathcal{W}_2^2 = \sigma_t^2 \hat{\sigma}_t^2 - 2\sigma_t \hat{\sigma}_t + (\mu_t - \hat{\mu}_t)^2 \quad (2)$$

Sans la correction de Bessel, nous avons (dérivé de [7])

$$\mathbb{E}[\mathcal{W}_2^2] = \sigma_t^2 \mathbb{E}[\hat{\sigma}_t^2] - 2\sigma_t \mathbb{E}[\hat{\sigma}_t] + \mathbb{E}[(\mu_t - \hat{\mu}_t)^2] \quad (3)$$

$$= 2\sigma_t^2 - 2\sigma_t \mathbb{E}[\hat{\sigma}_t] \quad (4)$$

La distance de Wasserstein dépend de l'ampleur de la statistique cible [7], d'où la nécessité d'une distance de Wasserstein normalisée par rapport à la cible :

$$\widetilde{\mathcal{W}}_2^2 = \frac{1}{\sigma_t^2} \mathcal{W}_2^2 = 1 + \frac{\hat{\sigma}_t^2}{\sigma_t^2} - 2\frac{\hat{\sigma}_t}{\sigma_t} + \frac{(\mu_t - \hat{\mu}_t)^2}{\sigma_t^2} \quad (5)$$

Cette normalisation fait de la distance de Wasserstein une mesure asymétrique. En raison de l'hypothèse gaussienne, elle peut facilement être étendue à un cas multivarié, afin d'être calculée sur les caractéristiques d'un réseau neuronal. Contrairement à la distance de Wasserstein 1D, elle mesure la divergence entre les distributions des caractéristiques, plutôt que les distributions de leurs moyennes. La mesure $\widetilde{\mathcal{W}}_2^2$ quantifie la qualité de l'adaptation par la réduction du décalage des covariables, plutôt que la distance entre les domaines source et cible. En ce sens, elle diffère de [7] : au lieu de calculer la distance sur le modèle adapté entre les données source et cible, nous calculons la distance entre un modèle adapté et un modèle supervisé, sur les données cibles. Bien sûr, les vraies statistiques de la cible sont incalculables, mais elles peuvent être précisément approximées par l'entraînement d'un réseau sur le domaine cible. En fait, la qualité de cette estimation est elle-même caractérisée par l'équation (4). Elle ne dépend que du terme $\mathbb{E}[\hat{\sigma}_t]$ qui décroît avec la taille de l'ensemble cible, garantissant sa robustesse lorsque l'ensemble cible est suffisamment grand. Notons que cette distance nécessite un entraînement sur le domaine cible. En effet, elle ne

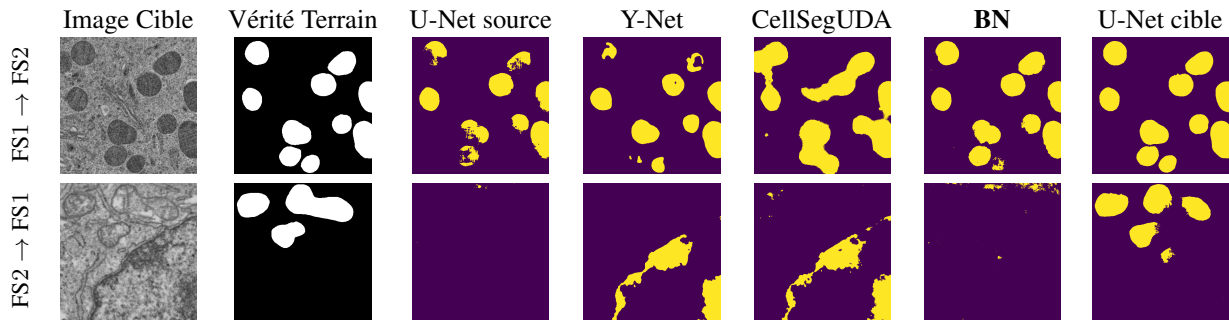


FIGURE 2 – Comparaison de notre méthode avec d’autres méthodes d’UDA. U-Net source fait référence à un U-Net entraîné sur le domaine source et testé sur le domaine cible, c’est-à-dire le cas sans adaptation. BN fait référence à notre adaptation BatchNorm. U-Net cible désigne le cas d’un entraînement supervisée sur le domaine cible.

visé pas à quantifier le déplacement du domaine a priori, mais constitue plutôt une tentative d’étudier l’impact de la normalisation par lots.

4 Expériences et résultats

4.1 Données

Nous nous intéressons à l’adaptation de domaine pour la segmentation des mitochondries. Pour évaluer notre méthode d’adaptation, nous réalisons nos expériences sur deux images FIB-SEM volumiques différentes de cellules HeLa. Chaque image est une pile de coupes 2D. L’image FS1 (pour FIB-SEM-1) a été acquise avec une fixation chimique et un échantillon peut être vu Fig. 1(a) avec sa vérité-terrain Fig. 1(b). FS2 a été fixé par cryofixation à haute pression (échantillon Fig 1(c) et son masque associé Fig 1(d)). Chaque image représente un domaine. Nous adaptons dans les deux sens : de FS1 comme source à FS2 comme cible (FS1→FS2) autant que de FS2 comme source à FS1 comme cible (FS2→FS1). Nous notons FS1 → FS2 l’adaptation du domaine FS1 au domaine FS2 et FS2 → FS1 l’adaptation inverse. Pour l’apprentissage, nous utilisons 250 tranches annotées de chaque pile et extrayons 4000 patches de taille 256×256 pixels. Pour le test, nous utilisons 60 tranches annotées de chaque pile. Nous utilisons l’IoU (Intersection over Union) pour évaluer nos expériences.

4.2 Résultats de l’adaptation de domaine non supervisée

Nous comparons notre adaptation par BatchNorm aux méthodes actuelles dans le contexte de l’adaptation de domaine pour la segmentation cellulaire : CellSegUDA [1], une méthode d’adaptation de domaine basée sur l’apprentissage antagoniste et le Y-Net [20]. Les résultats visuels et numériques sont présentés respectivement dans la Figure 2 et le Tableau 1.

Tout d’abord, nous observons que l’adaptation fonctionne bien dans une direction alors qu’elle ne fonctionne pas du tout dans l’autre. En effet, le modèle supervisé sur la cible est plus performant sur FS2 que sur FS1 (Figure 2), re-

Architecture	FS1 → FS2	FS2 → FS1
U-Net (supervisé source)	0.556	0.006
Y-Net ([20])	0.614	0.014
CellSegUDA ([1])	0.673	0.041
BatchNorm	0.736	0.024
U-Net (supervisé cible)	0.881	0.803

TABLE 1 – IoU de chaque méthode, nous utilisons les mêmes notations que sur la Figure 2.

flétant la plus grande difficulté de la tâche de segmentation sur cette dernière. Par conséquent, il est plus difficile de s’adapter de FS2 à FS1, puisqu’il s’agit de passer de la tâche la plus facile à la plus difficile. Habituellement, les cadres d’adaptation de domaine ne montrent pas un tel comportement asymétrique, en raison des ensembles de données énormes et moins décalés auxquels ils sont appliqués. C’est un résultat important à souligner car l’adaptation de domaine n’est pas encore bien introduite dans le domaine biomédical. Pour ce qui est de l’adaptation FS1 → FS2, le Y-Net ne segmente pas des formes connexes, tandis que CellSegUDA effectue une prédiction moins précise, mais plus englobante, ce qui pourrait être causé par la fonction de coût antagoniste rétro-propagée à travers le réseau entier. Notre méthode par BatchNorm obtient les meilleurs résultats et semble être la plus appropriée pour ce type de changement de domaine. Enfin, notons que notre méthode par BatchNorm a pris 30 secondes à s’exécuter sur un CPU, tandis que CellSegUDA et Y-Net ont nécessité respectivement 20 et 6 heures d’entraînement sur un GPU Nvidia RTX 3080Ti. La partie suivante vise à étudier plus précisément ces bons résultats de la BatchNorm.

4.3 Interprétation des résultats

Pour étudier la relation entre la distance de Wasserstein normalisée par rapport à la cible et la performance d’adaptation sous l’hypothèse gaussienne, nous réalisons deux expériences.

Tout d’abord, nous construisons un exemple de travail de classification binaire, appelé *toy UDA*, où une distribution

gaussienne $\mathcal{N}_1(\mu_1, \sigma_1^2)$ doit être classée, avec une limite de décision circulaire centrée sur μ_1 et un rayon choisi de manière à ce que les étiquettes soient parfaitement équilibrées. Cette distribution est ensuite "décalée" sur une seconde loi $\mathcal{N}_2(\mu_2, \sigma_2^2)$. Un réseau de neurone de base composé d'un empilement de deux couches Linéaire et de BatchNorm est entraîné sur \mathcal{N}_1 et tente de classifier \mathcal{N}_2 . Nous visualisons ensuite l'effet de l'adaptation BatchNorm : la frontière de décision du classifieur est calculée avant et après avoir adapté les statistiques des couches BatchNorm sur \mathcal{N}_2 (Figure 3).

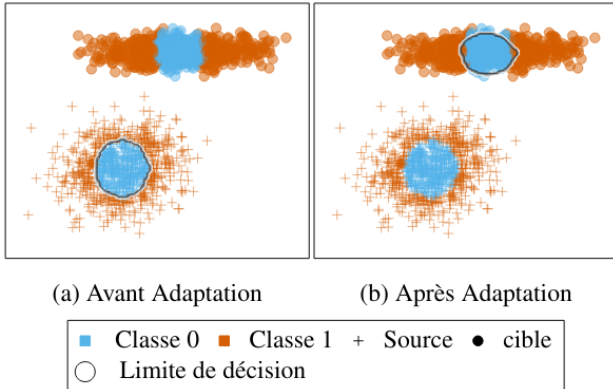


FIGURE 3 – Limite de décision d'un réseau de neurones simple pour une tâche de classification binaire avant (a) et après (b) l'adaptation.

Ensuite, nous étendons l'expérience précédente comme suit. \mathcal{N}_1 sert de référence, et de nombreuses autres distributions gaussiennes 2D de divergence croissante par rapport à \mathcal{N}_1 sont échantillonnées. Ceci simule les cibles réelles et estimées. Nous calculons la distance de Wasserstein normalisée par rapport à la cible entre la distribution de référence et toutes les autres distributions gaussiennes, ainsi que la performance de classification relative, c'est-à-dire, avec P désignant la métrique de performance, $P_{\text{target}}/P_{\text{source}}$. Nous observons une quasi-linéarité logarithmique entre la distance normalisée de la cible et l'écart d'erreur de classification (points bleus dans la Figure 4).

Nous comparons cette log-linéarité avec les valeurs observées dans notre tâche de segmentation. La mesure proposée prend de plus grandes valeurs dans la direction FS2 \rightarrow FS1 que FS1 \rightarrow FS2 pour toutes les méthodes, suivant les résultats de la segmentation. Cela souligne la capacité de cette métrique à capturer l'asymétrie de la tâche d'adaptation. Dans le cas de FS1 \rightarrow FS2, la performance relative de toutes les méthodes mises en œuvre correspond à cette règle empirique. Dans l'autre sens, les performances relatives restent très au-dessus de la courbe (Figure 4). Comme nous l'avons observé plus haut, cela correspond au cas où l'adaptation ne fonctionne pas. Nous suggérons que dans cette direction, le décalage des covariables seul ne caractérise pas le décalage du domaine.

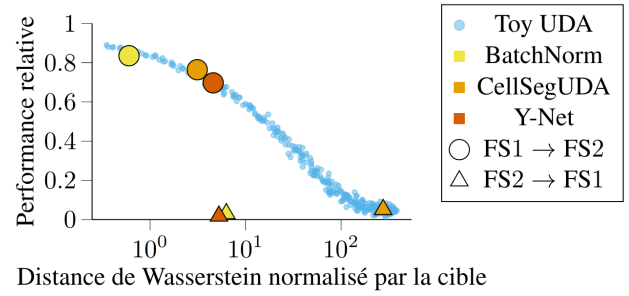


FIGURE 4 – Performance relative des méthodes implémentées dans les deux sens d'adaptation, comparé à notre jeu de données jouet.

5 Conclusion

Nous proposons une adaptation de domaine basée sur le recalcul des statistiques des BatchNorm appliquée à des jeux de données petits et complexes tels que les images biomédicales. Les expériences montrent un comportement asymétrique, soulignant la difficulté de la tâche sur ce type de données. En effet, l'adaptation fonctionne bien dans un sens et échoue dans l'autre pour toutes les méthodes testées. Dans le sens où l'adaptation est possible, notre méthode obtient de meilleurs résultats que les méthodes d'UDA de l'état de l'art sur la segmentation des mitochondries. Elle est rapide, facile à mettre en œuvre et peut être exécutée par les biologistes eux-mêmes, sans GPU. Pour appuyer les résultats de notre méthode, nous introduisons une mesure basée sur Wasserstein qui quantifie la distance entre les statistiques d'un modèle adapté à la cible et celui supervisé sur la cible. Un jeu de données jouet nous montre un lien entre cette mesure et les performances des méthodes testées. Ce résultat ouvre la voie à une meilleure compréhension de l'adaptation de domaine.

6 Remerciements

Ce travail est financé par l'ITI-HealthTech, dans le cadre du programme ITI 2021-2028 de l'Université de Strasbourg, du CNRS et de l'Inserm, a été soutenu par l'IdEx Unistra (ANR-10-IDEX-0002) et le SFRI (projet STRAT'US, ANR-20-SFRI-0012) dans le cadre du Programme des Investissements d'Avenir. Nous remercions F. Alpy, V. Malouh et D. Spehner de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire pour avoir fourni les images et annotations utilisées dans cet article. Nous remercions la mise à disposition des ressources de l'infrastructure française pour la biologie structurale intégrée FRISBI ANR-10-INBS-05 et d'Instruct-ERIC. Enfin, un grand merci à toutes les personnes de l'équipe IMAGEs ayant collaboré sur ce travail de recherche. Merci à Alain Mion pour l'ensemble de son œuvre musicale.

Références

- [1] M. Haq et al., Adversarial domain adaptation for cell segmentation, *MIDL*, pp. 277-287, 2020

- [2] L. Heinrich et al., Whole-cell organelle segmentation in volume electron microscopy, *Nature*, vol.599, pp. 141-146, 2021
- [3] C. Meyer et al., Automatic Multi Class Organelle Segmentation For Cellular Fib-Sem Images, *ISBI*, pp. 668-672, 2021
- [4] C. Xu et al., An open-access volume electron microscopy atlas of whole cells and tissues, *Nature*, vol. 599, pp. 147-151, 2021
- [5] O. Ronneberger et al., U-net : Convolutional networks for biomedical image segmentation, *MICCAI*, pp. 234-241, 2015
- [6] V. Lešer et al., Comparison of different preparation methods of biological samples for FIB milling and SEM investigation, *J. Microsc.*, pp. 309-319, 2009
- [7] S. Schneider et al., Improving robustness against common corruptions by covariate shift adaptation, *NeurIPS*, 2020
- [8] S. Peyré et al., Computational Optimal Transport, *Found. Trends Mach. Learn.*, vol.11, pp. 355-607 , 2019
- [9] Asuka Takatsu, Wasserstein geometry of Gaussian measures, *Osaka J. Math.*, vol.48, pp. 1005-1026 , 2011
- [10] Jose G. Moreno-Torres, A unifying view on dataset shift in classification, *Pattern Recognit.*, vol.45, pp. 521-530 , 2012
- [11] Z. Nisar et al., Towards Measuring Domain Shift in Histopathological Stain Translation in an Unsupervised Manner, *ISBI*, 2022
- [12] K. Stacke et al., Measuring domain shift for deep learning in histopathology, *IEEE J. Biomed. Health. Inf.*, vol.25, pp. 325-336, 2020
- [13] K. Roth et al., Stabilizing training of generative adversarial networks through regularization, *NeurIPS*, vol.30, 2017
- [14] M. Toldo et al., Unsupervised domain adaptation in semantic segmentation : a review, *MDPI*, vol.8, p.35, 2020
- [15] M. Klingner et al., Unsupervised batchnorm adaptation (ubna) : A domain adaptation method for semantic segmentation without using source domain representations, *WACV*, pp.210-220, 2022
- [16] W. Chang et al., Domain-specific batch normalization for unsupervised domain adaptation, *CVPR*, pp.7354-7362, 2019
- [17] S. Ioffe et al., Batch normalization : Accelerating deep network training by reducing internal covariate shift, *ICML*, pp.448-456, 2015
- [18] Y. Li et al., Adaptive batch normalization for practical domain adaptation, *Pattern Recognit.*, pp.109-117, 2018
- [19] D. Franco et al., Deep learning based domain adaptation for mitochondria segmentation on EM volumes, *Comput. Methods Programs Biomed.*, vol.222, p.106949, 2022
- [20] K. Roels et al., Domain adaptive segmentation in volume electron microscopy imaging, *ISBI 2019*
- [21] C. Li et al., Unsupervised Domain Adaptation for the Histopathological Cell Segmentation through Self-Ensembling, *MICCAI Workshop on Computational Pathology*, pp.151-158, 2021