

23/02/2022

Rapport étude technique

Analyse et traitement d'images hyperspectrales pour
la classification de grain de blé par Deep Learning
sous Python



Quentin Dupuis, Kieran Mcnamara
IMT MINES ALÈS

Table des matières

Table des matières	1
Table des figures	2
Introduction	3
I. Remarques préliminaires	4
II. Détection des contours	5
A. Passage en réflectance	5
B. Seuillage de l'image	6
C. Labélisation des grains	7
III. Extraction des grains	8
A. Création de l'image du grain	8
B. Détermination de la bande la plus brillante	8
IV. Apprentissage profond	9
A. Définition du réseau	9
B. Choix des jeux de données	10
C. Premiers résultats	10
V. Compréhension et amélioration	12
A. Comparaison surface des grains	12
B. Comparaison réflectance des grains	13
C. Remaniement des jeux de données et résultats	14
D. Complexité temporelle	16
VI. Perspectives	18
Conclusion	19
Sources	20

Table des figures

Figure 1. Image au format RGB d'une image contenant les grains de la variété 1 (gauche) grossissement d'un grain sur l'image (droite)	3
Figure 2. Comparaison spectralon luminosité et réflectance	6
Figure 3. Image binaire des grains	6
Figure 4. Résultats de la détection pour des grains séparés	7
Figure 5. Résultats de la détection pour des grains regroupés	7
Figure 6. À gauche, isolation de la zone de détection autour d'un grain. À droite, application du masque binaire sur la zone de détection	8
Figure 7. Visualisation d'un grain pour une bande spectrale sur 20	9
Figure 8. Structure du réseau de neurones convolutionnel	9
Figure 9. Évolution de la précision et de la valeur de la perte au cours des époques pour 21 bandes	10
Figure 10. Tableau comparatif des précisions d'entraînement et de test en fonction du nombre de bandes choisies pour les grains	11
Figure 11. Luminosité moyenne du spectralon en fonction de la bande spectrale	12
Figure 12. Répartition de la taille des grains pour chaque image	12
Figure 13. Boxplot des réflectances des grains des différentes images pour la bande 150 .	13
Figure 14. Précision de l'entraînement et de la validation en retirant l'image avec une réflectance élevée	14
Figure 15. Précision de l'entraînement et de la validation en utilisant l'image de réflectance élevée en validation et en ajoutant 2 images de 2021 pour l'apprentissage	15
Figure 16. Scores obtenus pour un entraînement sur 20 époques	16
Figure 17. Temps pris pour extraire les caractéristiques	17

Introduction

Dans le cadre de nos études à l'IMT Mines Alès, nous avons effectué une Étude Technique en collaboration avec L'INRAE (Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement). L'objectif de cette étude est de montrer ce que peuvent apporter des algorithmes de deep learning dans la classification de variétés de grains de blé dans des images hyper-spectrales. L'usage d'image hyperspectrale est justifié par le fait que les 2 variétés que nous avons à traiter sont très similaires à l'œil nu. De même, le recours au deep learning est justifié par le fait que les algorithmes de classification classiques ne donnent pas de très bons résultats.

Nous avons travaillé avec des images hyper-spectrales, de résolution 2048 * 9100 pixels. Chaque image contient 216 bandes spectrales (les images RGB classiques n'en contiennent que 3), chaque bande spectrale correspond à une unique longueur d'onde entre 400 et 1000 nm. Dans une image on retrouve 3 éléments : un spectralon en forme de demi-cercle qui renvoie la lumière de façon uniforme, une bande de papier millimétré et des dizaines de grains de blé, appartenant tous à la même variété (Figure 1).

Il nous a été fourni au total 8 images hyper-spectrales de ce type : 4 provenant de la récolte de 2020 et 4 de l'année 2021. L'objectif était de pouvoir prédire les variétés de l'année 2021 en entraînant notre modèle uniquement sur l'année 2021. Pour chaque année, nous avons 2 images contenant uniquement des grains de blé de la variété 1 et 2 images contenant uniquement la variété 8. Notre étude s'est restreinte à la discrimination entre ces 2 classes.

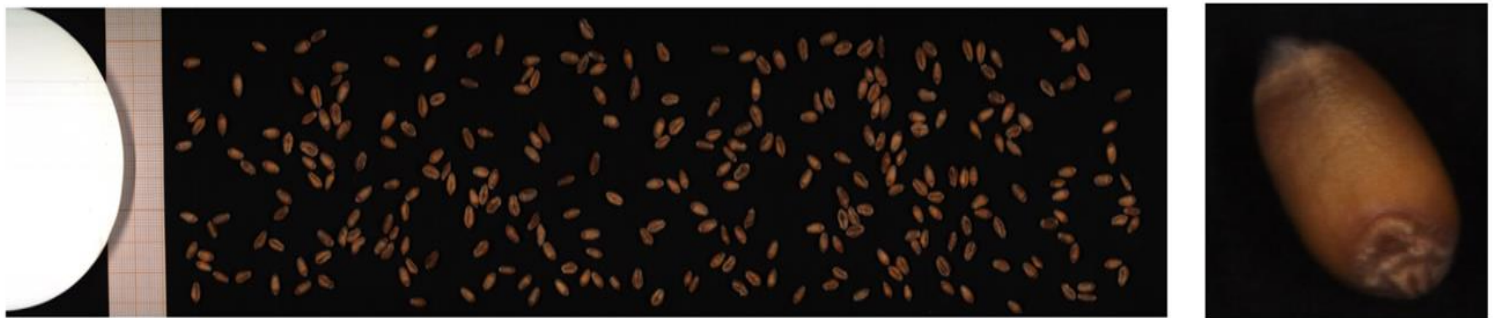


Figure 1. Image au format RGB d'une image contenant les grains de la variété 1 (gauche) grossissement d'un grain sur l'image (droite)

I. Remarques préliminaires

Lors de la prise en main des méthodes pour lire une image hyperspectrale, un premier problème est apparu, celui du chargement. En effet, ces images faisant près de 8GB, une mémoire vive suffisante est nécessaire afin de pouvoir naviguer dans celles-ci et de faire des opérations. Cependant, les configurations utilisées n'étaient pas dotées d'une capacité suffisante pour cela. Il a donc fallu trouver une solution au chargement de l'image complète dans la RAM. Une remarque similaire peut être faite sur le disque qui sert au stockage. Plus le disque aura un transfert de données rapide, plus les lectures des images se feront rapidement. Dans le cas idéal, un SSD est préféré à un disque dur classique mais il faut que celui-ci puisse avoir un stockage suffisant, ce qui peut être coûteux.

Le langage utilisé pour traiter les images hyperspectrales est Python, notamment via l'utilisation de la librairie Spectral qui a été développée pour ce type d'application. Elle permet la lecture en image complète mais aussi en bande voire sous-image. Bien que facile d'utilisation, cette librairie admet des temps d'exécution conséquent, ne permettant qu'un affichage pour une bande donnée de l'image en environ 1 min. De même, certaines fonctions ralentissent lorsque trop de bande sont extraites, notamment lors de l'extraction d'une partie de l'image où il est préférable d'extraire bande par bande plutôt que d'utiliser la fonction prédéfinie pour extraire une sous-image. En effet, des tests ont été réalisés sur un nombre désiré de bande. Plus ce nombre est bas, plus le traitement est rapide. Le nombre de bande maximale à avoir été testé est 108 et le temps d'exécution se compte en heures (1h30 pour l'extraction des grains de toutes les images). Il est donc difficile d'obtenir un affichage en temps réel de l'image. À la vue de ces éléments, il est préférable de lire le moins possible les images originales en stockant le plus possible les résultats des différentes opérations si possible.

Ainsi, les images hyperspectrales sont utilisées pour l'extraction des grains et ceux-ci sont sauvegardés directement sur disque. Cela permet notamment au réseau de neurones d'avoir un accès rapide aux données tout en n'occupant que peu de mémoire vive puisque les images des grains vont être traitées par regroupement. D'autres informations ont également été stockées pour permettre une meilleure extraction des grains mais aussi une caractérisation de ceux-ci.

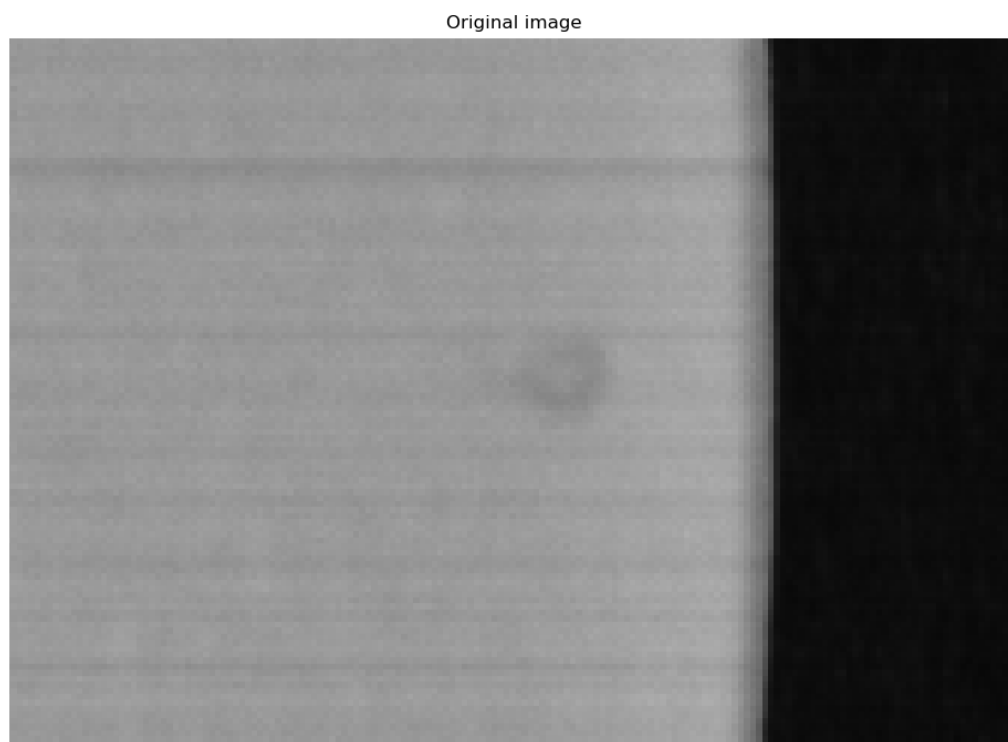
II. Détection des contours

A. Passage en réflectance

L'extraction des grains passe par plusieurs étapes. D'abord, il faut trouver dans quelle bande (longueur d'onde) se placer pour permettre la meilleure détection possible. Ensuite, pour la bande choisie, il faut étalonner l'image en passant des valeurs de luminosité en valeurs de réflectance, c'est-à-dire, utiliser le spectralon présent à l'un des extrémités de l'image. Enfin, les grains sont détectés à travers une méthode de seuillage.

Puisque la position des grains ne varie pas à travers les bandes, la détermination des contours peut se faire sur l'une d'entre elles. En considérant le traitement qui sera fait par la suite, c'est-à-dire le seuillage, la bande qui possède la luminosité la plus élevée semble être le meilleur choix. Pour cela, la valeur moyenne du spectralon est calculée pour chaque bande et la valeur maximale ainsi que le numéro de la bande associée sont retournées.

Une fois cette sélection effectuée, l'image est transformée de valeur de luminosité en valeur de réflectance. La réflectance est ici calculée par colonne et consiste en la division de chaque pixel par la moyenne de luminosité du spectralon si celui-ci fait plus de 50 pixels de longueur. De ce fait, les extrémités de l'image sont coupées lorsqu'il y a trop peu d'informations liées au spectralon. Cette opération permet de lisser les disparités de l'image lié à la capture, comme observable sur l'image (Figure 2).



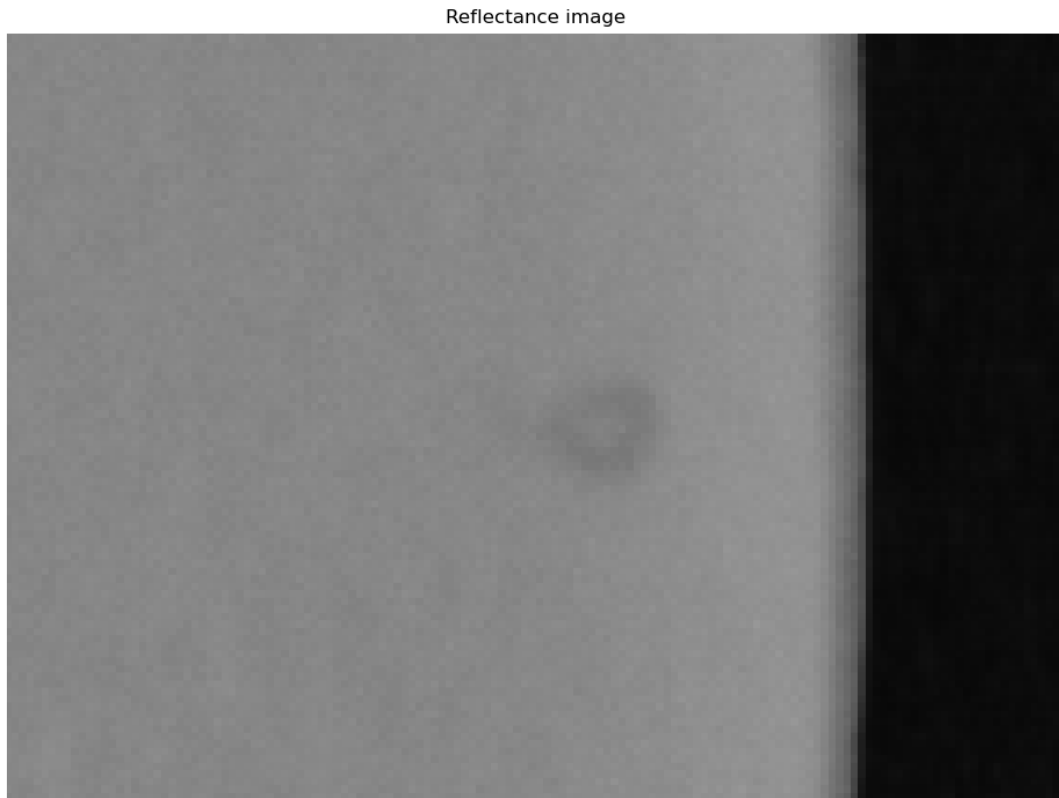


Figure 2. Comparaison spectralon luminosité et réflectance

B. Seuillage de l'image

Il est maintenant possible d'extraire les grains dans les meilleures conditions possibles. Puisque les images sont en nuances de gris, l'idée principale est de réaliser un seuillage. Si la valeur du pixel est inférieure à un certain seuil, celle-ci est mise à 0. Inversement, si elle est supérieure, elle est mise à 1. Grâce à cette méthode, l'ensemble des grains sont trouvés est mis en valeur (Figure 3).

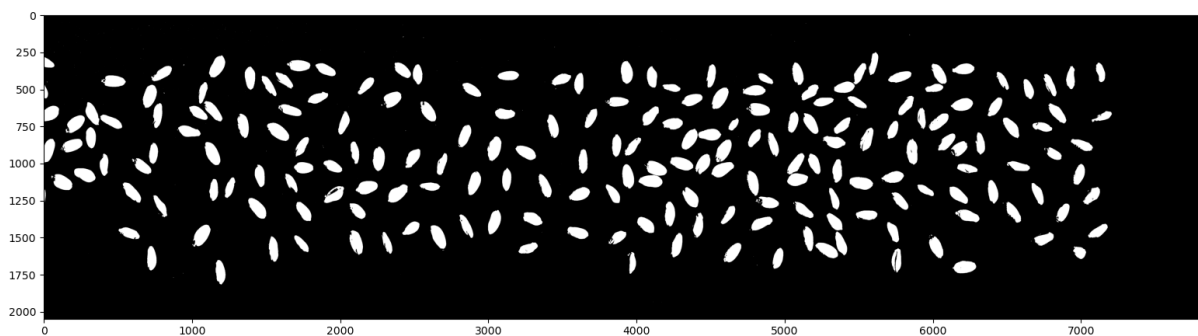


Figure 3. Image binaire des grains

C. Labélisation des grains

La difficulté réside alors dans le fait d'isoler chaque grain par rapport aux autres. Une labellisation directe est possible en connectant deux pixels voisins ayant la même valeur. Cependant, si lors du seuillage, une connexion s'est créée entre deux grains, du fait de leur proximité, ils seront considérés comme une même entité. Une première idée pour prendre en compte ce problème est de réaliser un deuxième seuillage pour les labels qui ont une surface élevée, indice du regroupement de plusieurs grains. La valeur du seuil sera plus élevée pour permettre d'obtenir une meilleure distinction au détriment des bords des grains qui vont être rongés. Une deuxième idée est d'utiliser des techniques de séparation comme « Watershed » ou du clustering. À la vue des résultats suffisants de la première approche, de sa rapidité et du temps imparti, les méthodes plus complexes n'ont pas été implémentées.

Grâce à la labellisation, des propriétés peuvent être récupérées directement comme la boîte englobante (« Bounding box » en anglais) et l'image convexe. Cette image permet d'obtenir de manière directe un filtre pour extraire uniquement les grains et rendre l'arrière-plan noir. De cette manière, la position des grains est obtenue dans l'image originale et permettra par la suite d'être utilisé pour les extraire à travers les bandes tout en appliquant le filtre pour ne ressortir que le grain.

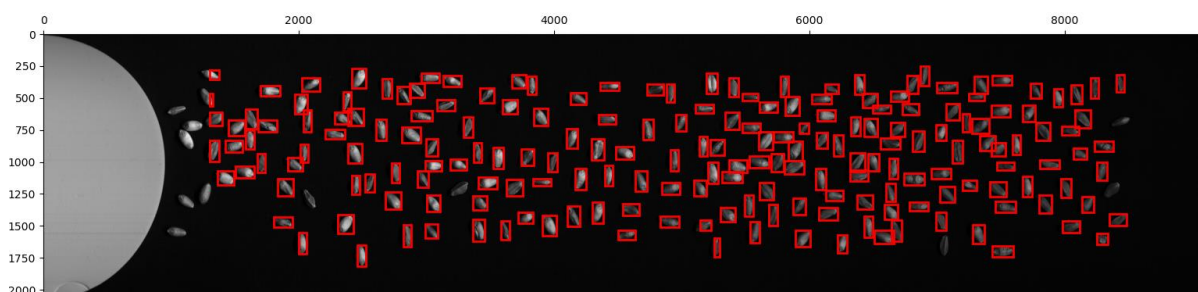


Figure 4. Résultats de la détection pour des grains séparés

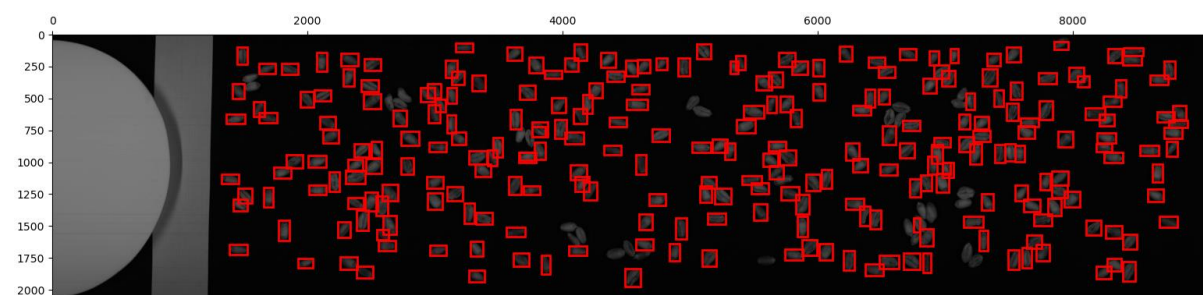


Figure 5. Résultats de la détection pour des grains regroupés

Cette méthode d'extraction fonctionne bien dans le cas général, mais admet quelques difficultés dans les scénarios décrits précédemment. Si plusieurs grains sont proches (Figure 5), ils seront difficilement séparés. Dans le cas où les grains sont dispersés, très peu d'erreurs sont produites pour la détection mais certains grains trop sombres par rapport aux autres peuvent ne pas être considérés (Figure 4).

III. Extraction des grains

A. Création de l'image du grain

Une fois que nous avons effectué cette détection, nous récupérons l'ensemble des zones de détection sous forme de liste. Chaque élément de cette liste contient une autre liste qui contient la position en x et y du coin supérieur gauche de la détection ainsi que sa largeur et sa hauteur. Puisque le grain se situe à la même position sur chaque bande spectrale, nous avons pu rogner chaque grain sur chaque bande afin de construire une nouvelle image hyperspectrale contenant qu'un seul grain.

On remarque cependant quelques artefacts qui peuvent survenir avec cette méthode comme le bruit de fond ou d'autre morceau de grain présent sur l'image (Figure 6). Ces artefacts peuvent perturber le traitement de l'image par le réseau de neurones, c'est pourquoi il nous a fallu extraire et appliquer un masque sur notre image de grain initiale afin de sélectionner uniquement les pixels qui se situent à l'intérieur du grain (Figure 6)

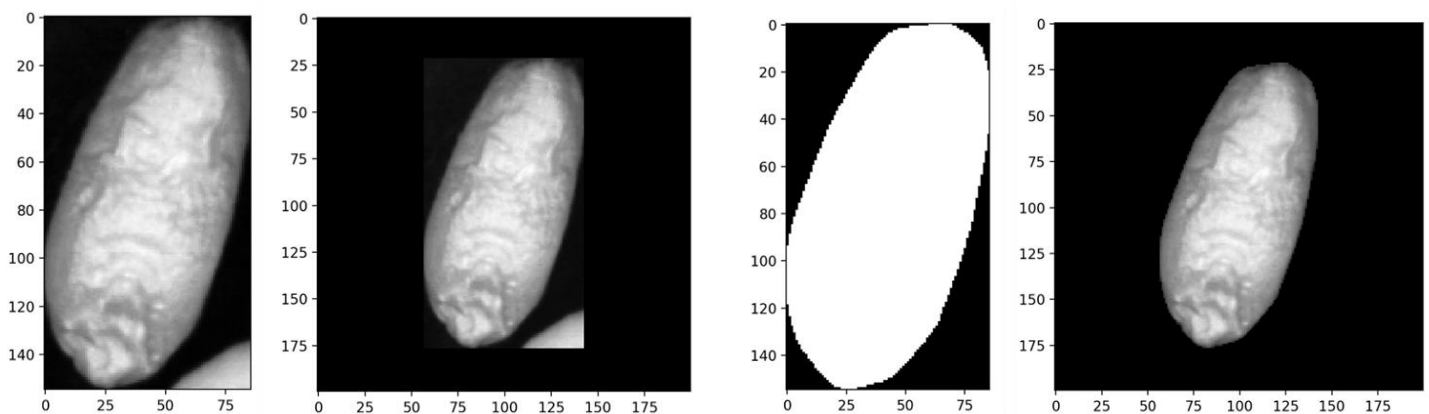


Figure 6. À gauche, isolation de la zone de détection autour d'un grain. À droite, application du masque binaire sur la zone de détection

B. Détermination de la bande la plus brillante

Lorsque l'on observe toutes les bandes spectrales d'un grain de blé (Figure 7) on remarque que les bandes aux extrémités sont plus bruitées que celles au centre. Cela s'explique en grande partie par le spectre d'émission de la lampe utilisée lors de l'acquisition des images. Celui-ci est beaucoup plus sombre aux extrémités donc le bruit est plus important. Nous avons donc calculé pour chaque image la bande la plus lumineuse afin de réaliser l'extraction dans les meilleures conditions.

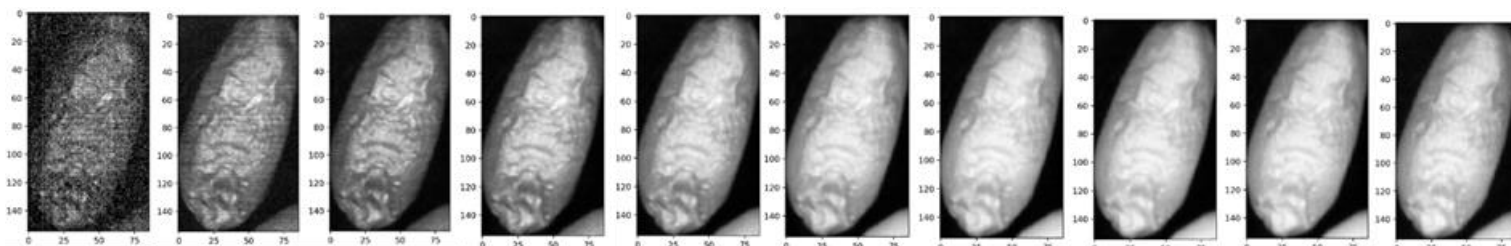


Figure 7. Visualisation d'un grain pour une bande spectrale sur 20

IV. Apprentissage profond

A. Définition du réseau

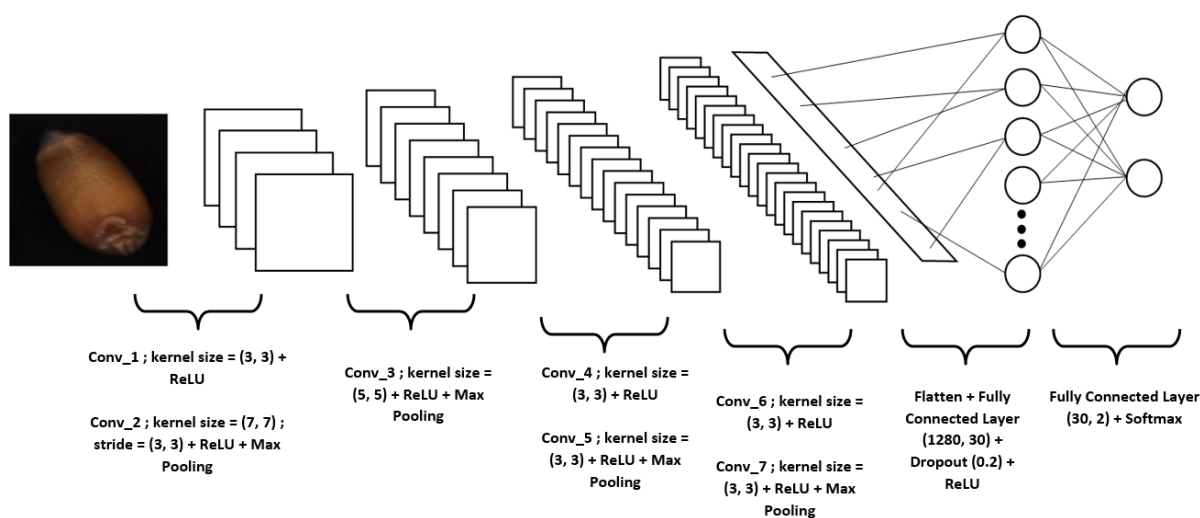


Figure 8. Structure du réseau de neurones convolutionnel

Le réseau que nous utilisons ici (Figure 8) est un réseau de neurones convolutionnel, nous avons fait le choix de ce type d'architecture car il est adapté pour le traitement d'image. Ce réseau prend en entrée une image d'une résolution de 200x200 pixels avec n bandes (n peut prendre n'importe quelle valeur entre 1 et 216). Il contient 7 couches de convolution, 4 "max pooling" et 3 "fully connected layers" [1]. En sortie nous avons un vecteur de 2 éléments : un pour chaque classe, auquel on vient appliquer la fonction Softmax pour obtenir une densité de probabilité.

Lorsque l'on effectue notre boucle d'entraînement, nous fournissons un lot de 12 images avant de faire notre descente de gradient. Cette méthode est couramment utilisée afin d'améliorer les performances de l'entraînement.

B. Choix des jeux de données

Le principale avantage de travailler par traitement par lots (ou « Batch ») dans le cas présent est de ne pas charger toute l'image dans la mémoire mais plutôt d'aller lire uniquement un nombre limité de grains. Cependant, cela nécessite de nombreuses allocations et désallocations et donc une certaine lenteur de l'algorithme. Pour pallier ce problème, un SSD peut être utilisé pour accélérer les lectures. À titre d'exemple, pour des images où 108 bandes ont été extraites, chacune d'entre elles fait près de 34 Mo. Si 200 détections ont eu lieu, cela représente 6.8 Go de données qui doivent transiter par la mémoire pour chaque image. Pour un entraînement sur 5 images, chaque époque requiert donc 34 Go. Ce phénomène explique pourquoi les boucles d'entraînement prennent du temps, non pas à cause du calcul mais à cause de la lecture des données.

Le jeu de données à considérer est composé de 8 images réparties sur les années 2020 et 2021. Pour chaque année, deux images sont fournies pour les variétés 1 et 8. Comme l'objectif initial est de pouvoir prédire les variétés d'une année sur l'autre, les images de l'année 2020 ont été utilisées pour la partie entraînement du modèle et l'année 2021 pour la validation et le test. À noter qu'il n'y a pas de mélange de grains de différentes images, les 3 parties précédentes sont bien indépendantes entre elles.

C. Premiers résultats

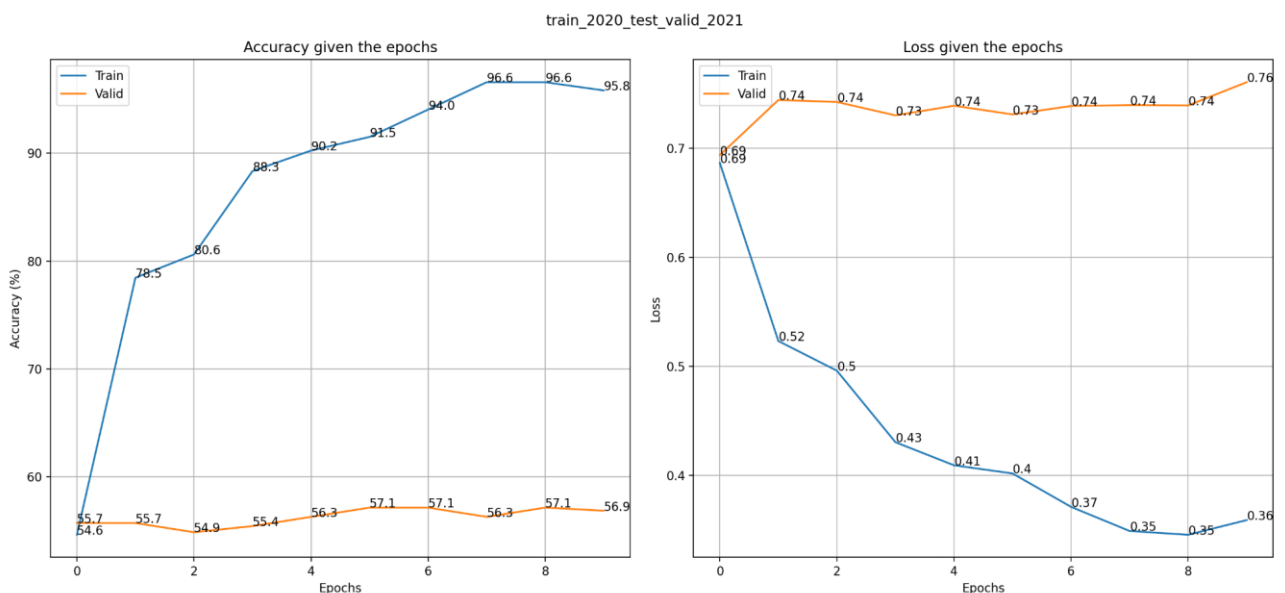


Figure 9. Évolution de la précision et de la valeur de la perte au cours des époques pour 21 bandes

Un entraînement a été réalisé pour des images de grains où 21 bandes ont été extraites. À travers les résultats (Figure 9), plusieurs remarques peuvent être faites. D'abord, l'apprentissage sur l'année 2020 se fait bien et une précision de plus de 95% est atteinte, le réseau de neurones apprend des grains. Cependant, celui-ci prédit mal sur l'année 2021 avec une précision qui se situe aux alentours de 57% pour la validation. À noter que pour cette validation, il y a 55.4% de grains de la variété 8. Après vérification, la variété 8 est bel et bien majoritairement prédite. Sur le jeu de test, une précision de 59.2% est atteinte mais la proportion de la variété 8 est de 58%, la prédiction d'une seule classe se retrouve comme pour la validation.

Une première idée pour améliorer ces résultats est de comparer l'impact du nombre de bandes choisies. Pour cela, des tests ont été réalisés dans les mêmes conditions, c'est-à-dire entraînement sur 2020 et validation et test sur 2021 pour un nombre de 2, 10 et 108 bandes, en plus de l'essai précédent.

Nombre de bandes	Précision après 10 itérations Entraînement	Précision sur le jeu de Test
2	89.5%	56.5%
10	91.0%	56.7%
21	95.8%	59.2%
108	95.6%	55.4%

Figure 10. Tableau comparatif des précisions d'entraînement et de test en fonction du nombre de bandes choisies pour les grains

Cette expérience met en évidence que le nombre de bandes, pour un nombre de 10 itérations, influe peu sur la précision sur le jeu de test (Figure 10). Pour une valeur de 108, la précision sur le jeu d'entraînement est proche de la valeur pour 21 bandes mais la précision sur le test est plus faible. Cela s'explique par le faible nombre d'itérations alors qu'il y a plus d'informations fournies par les images de grains. Une augmentation de ce nombre permettrait une meilleure comparaison puisque plus le nombre de bandes est élevé plus la convergence vers un minimum pour la fonction de perte est lente.

V. Compréhension et amélioration

A. Comparaison surface des grains

Pour interpréter ces résultats nous avons mesuré la luminosité moyenne du spectralon pour chaque bande dans nos images d'entraînement et de test. Les résultats sont présentés dans la Figure 11 (orange & bleu pour 2021). On remarque que cette luminosité varie du simple au double ce qui traduit le fait que les conditions d'acquisition des images n'étaient pas les mêmes entre chaque prise ce qui complique la classification par la suite.

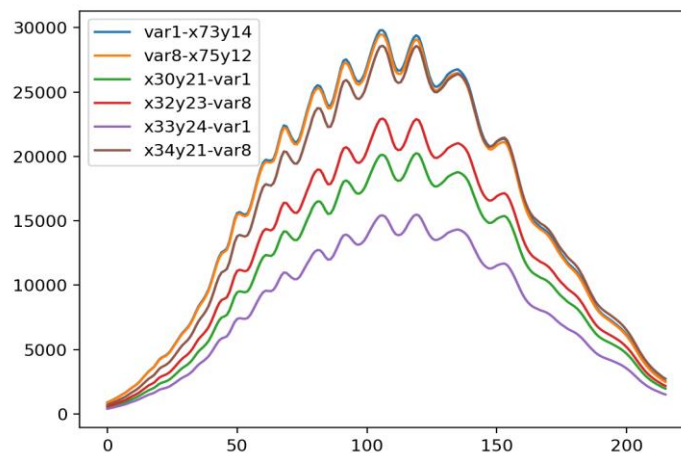


Figure 11. Luminosité moyenne du spectralon en fonction de la bande spectrale

Sur le graphique nous pouvons voir la répartition de la taille des grains pour chaque image (taille calculée en nombre de pixel dans le masque). Cette répartition apporte peu d'information sur la variété cependant nous remarquons que les grains de l'année 2021 (les 4 à gauche) sont en moyenne plus gros que ceux de 2020 (les 4 à droite). Cela peut s'expliquer par le fait que la luminosité est plus élevée dans les images de 2021 (Figure 12) ce qui fait que les masques contiennent plus de pixels, notamment sur les bords des grains.

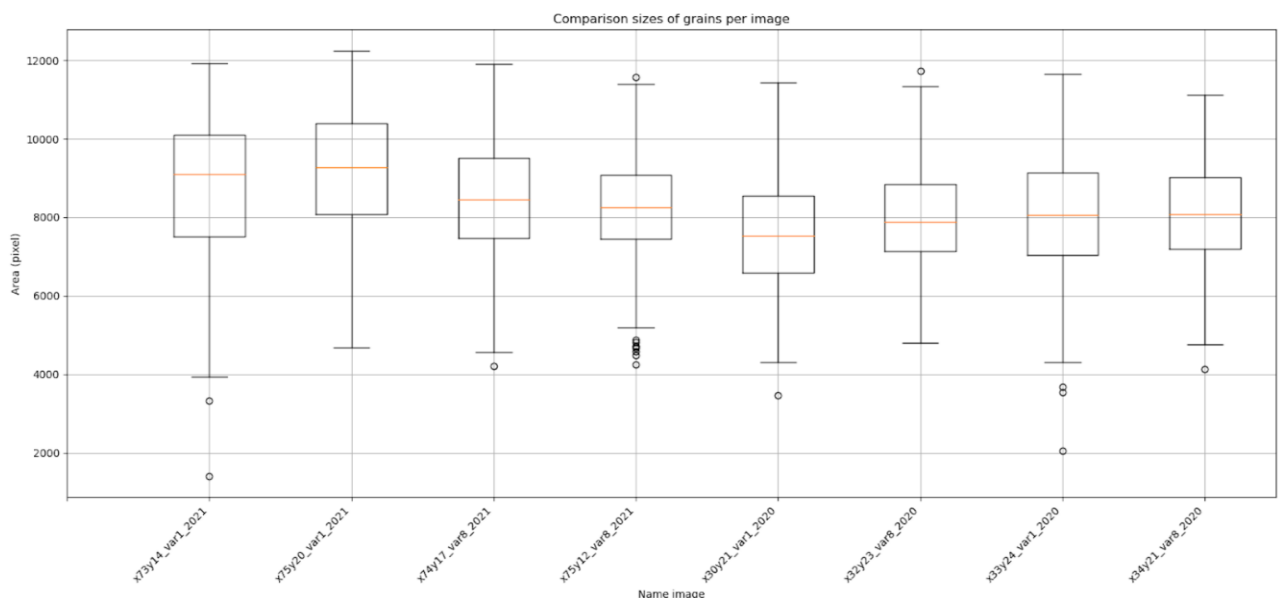


Figure 12. Répartition de la taille des grains pour chaque image

B. Comparaison réflectance des grains

Un autre élément qui peut avoir un impact sur les résultats de l'apprentissage est la réflectance des grains. Si, entre deux images, cette valeur varie de manière significative, le réseau associera des valeurs extrêmes à une variété. Cependant, cette caractéristique provient de l'acquisition de l'image et non du grain en lui-même. Pour vérifier que les images hyperspectrales ont été prises dans les mêmes conditions, en plus de l'étude de leur spectralon, la réflectance des grains peut être analysée. Pour cela, la procédure est la même que pour l'extraction des coordonnées des grains dans l'image. Tout d'abord, la luminance moyenne du spectralon au niveau de chaque colonne pour chaque bande est calculée. Puisque cette opération prend quelques minutes pour chaque image, les résultats sont sauvegardés. Ensuite, pour tous les grains détectés, passage en valeur de réflectance et application du masque à travers toutes les bandes. Enfin, réalisation du calcul de certaines caractéristiques : maximum, minimum, moyenne, médiane et écart-type. Ces valeurs sont également enregistrées.

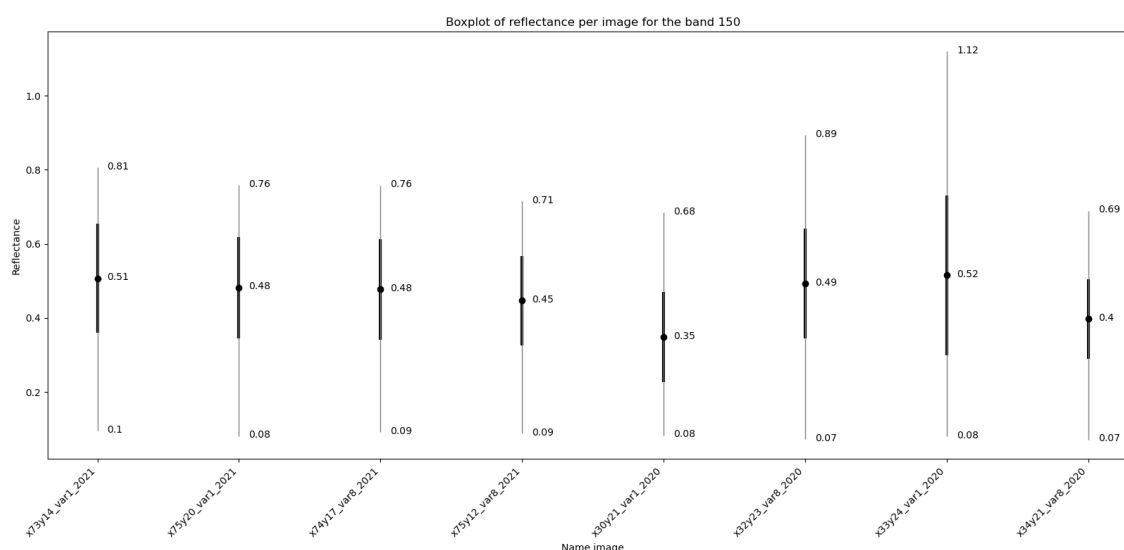


Figure 13. Boxplot des réflectances des grains des différentes images pour la bande 150

Ce boxplot (Figure 13) est obtenu après avoir réalisé une moyenne des minimums, maximums, moyennes et écart-types de chaque grain, à travers toutes les bandes, pour chaque image. Les 4 images de 2021 ont des valeurs ressemblantes avec une moyenne de l'ordre de 0.5 avec un écart-type d'environ 0.1. Cependant, pour l'année 2020, les images présentent de plus fortes disparités. La première d'entre elles offre les valeurs moyennes de réflectance les plus faibles, avec une moyenne à 0.35. Pour la même année, l'avant dernière image est celle avec les valeurs les plus élevées avec une moyenne à 0.52 et un écart-type proche de 0.2. Le fait d'avoir une moyenne des maximums au-dessus de 1 est explicable puisque les grains peuvent renvoyer plus de lumière dans une seule direction alors que le spectralon renvoie la lumière dans toutes les directions. À la vue du caractère unique de cette image, une première idée a été de la mettre de côté lors de l'entraînement pour observer s'il y avait une amélioration des résultats.

C. Remaniement des jeux de données et résultats

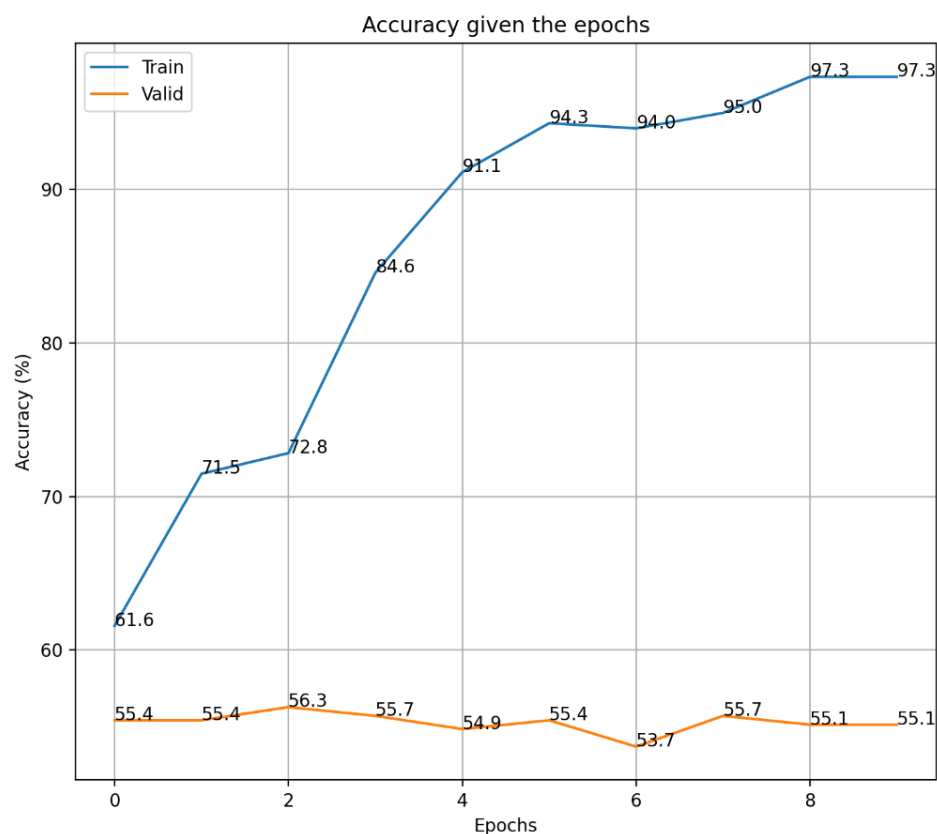


Figure 14. Précision de l'entraînement et de la validation en retirant l'image avec une réflectance élevée

En considérant uniquement 3 images de 2020 pour l'entraînement du modèle et en gardant 2021 pour la validation et le test, la prédiction ne s'est pas améliorée (Figure 14). L'apprentissage se réalise bien sur 2020 mais la variété 8 est toujours majoritairement prédite en validation mais aussi en test, avec des précisions proches de la proportion de celle-ci dans les jeux de données.

Une prédiction d'une année sur l'autre sans avoir accès à des grains de l'année à prédire semble complexe dans une première approche. Pour y remédier, des images de référence peuvent être introduites pour orienter le réseau, provenant de parcelles de monoculture où la variété est déjà connue. Dans ce contexte, deux images, une pour chaque variété, vont être ajoutées à la phase d'entraînement.

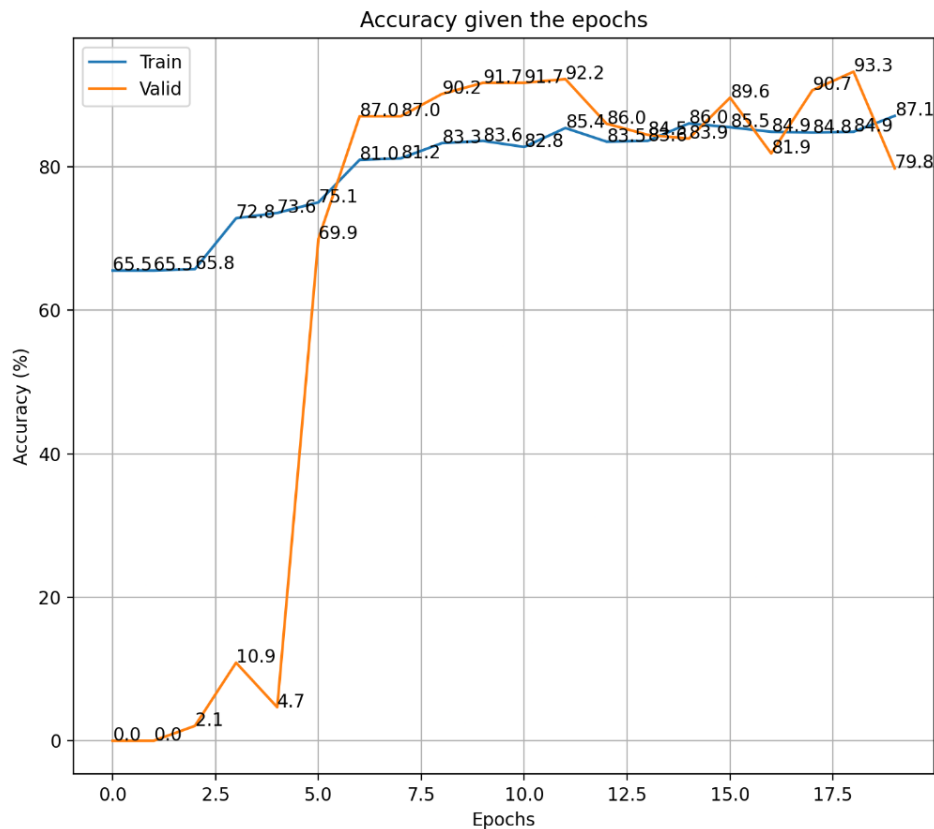


Figure 15. Précision de l'entraînement et de la validation en utilisant l'image de réflectance élevée en validation et en ajoutant 2 images de 2021 pour l'apprentissage

Ainsi, en plaçant l'image de réflectance élevée en validation et les images de 2021 de validation en entraînement, les courbes de précision précédentes sont obtenues (Figure 15). Une première remarque porte sur la variation de la précision pour la validation. En effet, comme attendu avec les valeurs élevées de réflectance, les premières itérations ne permettent pas de prédire la bonne variété. Mais à mesure des époques, l'apprentissage va se réaliser à la fois en prenant en compte les caractéristiques spectrales mais aussi spatiales. De cette façon, la courbe de validation va se stabiliser aux alentours d'une précision de 90%. Concernant celle de l'entraînement, elle croît lentement vers la même valeur de 90%. Le nombre d'itérations réalisées ne permet pas de déterminer si le réseau est en sur-apprentissage, il faudrait itérer plus longtemps pour observer les résultats sur le long terme.

Test :				
Accuracy: 64.7%		Avg loss: 0.662689		
	precision	recall	f1-score	support
0	0.94	0.17	0.29	184
1	0.62	0.99	0.77	255
accuracy			0.65	439
macro avg	0.78	0.58	0.53	439
weighted avg	0.76	0.65	0.56	439

Figure 16. Scores obtenus pour un entraînement sur 20 époques

Néanmoins, l'apprentissage effectué permet d'obtenir des résultats intéressants (Figure 16). Sur le jeu de test, qui a une proportion de 58% de la variété 8, une précision de 64.7% est obtenue. La prédiction prédominante de la deuxième classe se retrouve mais est à nuancer. En effet, les grains de la variété 1, ou classe 0, commence à être trouvés et avec une bonne précision (94%). Malheureusement, seule une trentaine d'entre eux sont prédits sur les 184 (17% de rappel).

D. Complexité temporelle

D'un point de vue pratique, une difficulté à prendre en compte est la complexité temporelle des différents algorithmes utilisés. Les images originales ont une taille de 8 Go puis les grains vont être extraits et stockés dans une image plus grande. Ainsi, lors des opérations de sélection de la bande la plus brillante ou encore l'extraction du spectralon par bande pour pouvoir passer en valeur de réflectance, il y a une nécessité de stocker le plus d'informations intermédiaires possibles. De la même manière au niveau des grains, la composante principale du réseau en termes de temps d'exécution est le chargement des images, comme évoqué à la partie précédente.

```

New file, 1/7

Extracting features: 100%|██████████| 184/184 [45:01<00:00, 14.68s/it]

New file, 2/7

Extracting features: 100%|██████████| 196/196 [45:18<00:00, 13.87s/it]

New file, 3/7

Extracting features: 100%|██████████| 255/255 [56:35<00:00, 13.32s/it]

New file, 4/7

Extracting features: 100%|██████████| 173/173 [35:45<00:00, 12.40s/it]

New file, 5/7

Extracting features: 100%|██████████| 195/195 [41:46<00:00, 12.86s/it]

New file, 6/7

Extracting features: 100%|██████████| 193/193 [42:06<00:00, 13.09s/it]

New file, 7/7

Extracting features: 100%|██████████| 231/231 [49:45<00:00, 12.92s/it]

Process finished with exit code 0

```

Figure 17. Temps pris pour extraire les caractéristiques

Pour prendre un exemple concret, afin d'obtenir les boxplots de comparaison des réflectances des images, il a fallu pour cela extraire différentes caractéristiques (max, min, moyenne et écart-type). Cette extraction prend environ 50min pour 230 grains (Figure 17) avec des optimisations au niveau des passages en réflectance puisque les valeurs du spectralon ont préalablement été stockées. Ce temps peut être expliqué à travers les différentes parties prises en compte. D'abord, il faut considérer une moyenne de 0.07s de traitement par bande et par grain qui est principalement utilisée par la librairie Spectral pour récupérer la partie du grain dans l'image globale. Ensuite, si on considère 200 grains et les 216 bandes spectrales, on obtient :

$$0.07 * 200 * 216 = 3024s \simeq 50min$$

On retrouve ainsi l'ordre de grandeur du temps pris pour l'extraction et l'obtention d'un fichier de taille (nombre de grains, 216, 5). Ce temps peut être réduit si un disque plus rapide est utilisé comme un SSD.

VI. Perspectives

Les résultats obtenus en prenant en compte 2021 dans l'apprentissage sont encourageants. La variété 8 reste majoritairement prédite mais la seconde commence à être trouvée après une vingtaine d'époques. Plusieurs points peuvent être développés pour affiner la précision. Tout d'abord, si un système plus performant est disponible, l'augmentation du nombre d'itérations et de bandes est envisageable. Ainsi, le jeu de validation sera utilisé pour obtenir le moment pour lequel le sur-apprentissage commence et donc le nombre d'époques optimal à prendre. Les effets du nombre de bande pourront également être plus approfondis, en déterminant le nombre de bandes idéal pour ne pas perdre trop d'information tout en gagnant en temps de calcul. Ensuite, une gestion des jeux de données déséquilibrés favorisera une prédiction mieux répartie entre les classes, ou du moins, chaque classe aurait le même poids dans l'apprentissage. Cela peut être réalisé lors de la répartition des jeux de données mais également à l'intérieur du réseau de neurones où un poids différent peut être attribué à chacune des classes lors du calcul de la perte pour sanctionner plus lourdement une classe mal prédite plutôt que l'autre. De plus, le réseau en lui-même est modifiable en ajoutant des couches ou en remaniant celles déjà existantes.

Un autre aspect à prendre en compte est la détection des grains. La méthode mise en place permet l'extraction de près de 90% des grains lorsque ceux-ci sont rapprochés et plus de 95% lorsque ceux-ci sont dispersés. Pour obtenir une meilleure détection et notamment une meilleure extraction des grains en évitant de ronger les contours via seuillage, d'autres techniques peuvent être mises en place. Le seuillage qui permet de produire une image binaire pour la labellisation peut être remplacé par une détection de contours via l'utilisation de gradient ou de Laplacien par exemple. Par la suite, du regroupement, ou clustering, est envisageable pour isoler chaque grain. Néanmoins, le double seuillage utilisé dans cette étude permet la détection et l'extraction d'un grand nombre de grains avec une faible dégradation des contours.

Enfin, le réseau de neurones ne prend en entrée que les images extraites des grains mais peut également prendre d'autres caractéristiques. Les résultats obtenus n'ont pris en compte que les caractéristiques spatiales et spectrales des grains, d'autres éléments peuvent être introduits comme la position de la parcelle, les conditions météorologiques ou bien la date de récolte. Toutes ces valeurs sont introduites au réseau à la suite des couches de convolution, c'est-à-dire après l'analyse de l'image, au niveau d'un ensemble de couches linéaires servant à la classification.

Conclusion

Cette étude avait pour objectif de mettre en place un apprentissage profond pour évaluer cette approche par rapport aux approches apprentissage machine classiques pour la prédiction de la variété de grains d'une année sur l'autre. Pour cela, des algorithmes de détection et d'extraction de grains ont été définis et automatisés en fournissant une image hyperspectrale en entrée. La difficulté liée à la taille importante des images a été résolue en les décomposant par bandes et en sauvegardant les résultats des calculs intermédiaires mais aussi les images des grains après extraction. Néanmoins, même avec ces optimisations, les temps de calcul restent longs lorsqu'il faut parcourir l'ensemble des bandes spectrales d'une image. Par la suite, le réseau de neurones a été développé, prenant en entrée les grains extraits en valeur de réflectance. Lors de l'entraînement de celui-ci sur 2020 pour prédire 2021, les résultats ne sont pas satisfaisants, une seule variété étant majoritairement prédite. Pour résoudre ce problème, une analyse des images a été menée pour en comprendre l'origine. Au travers de cette analyse, et à l'introduction d'images de référence de 2021 dans l'apprentissage, de nouveaux résultats sont obtenus, plus prometteurs. La variété qui était jusqu'alors peu prédite commence à l'être et une précision de l'ordre de 65% est atteinte sur un jeu de 2021. Les variétés étant difficiles à discriminer, la classification des grains est délicate, que ce soit par une approche classique ou par un réseau de neurones. Dans le cas actuel, ce dernier peut encore être affiné avec l'ensemble des éléments introduits précédemment. Ainsi, l'étude a permis d'introduire une solution par apprentissage profond qui obtient des résultats encourageants et améliorables tout en permettant d'automatiser la détection et l'extraction de grains.

Sources

[1] « Convolutionnal Neural Network », version du 22 février 2022, Wikipédia :
https://en.wikipedia.org/wiki/Convolutional_neural_network