

Analyse semi-automatique d'images aériennes pour la détection de panneaux solaires photovoltaïques

Curier R. Lyana, Ziemons H., de Jong T., Iren, D., Bromuri S.¹

Résumé

Le principal objectif de l'étude est d'évaluer l'efficacité de modèles d'apprentissage profond pour l'identification de panneaux solaires dans des images aériennes. À cette fin, divers modèles d'apprentissage par transfert sont évalués pour identifier la stratégie la plus efficace. Cette contribution fait seulement usage d'architectures d'apprentissage profond basées sur des filtres convolutifs (CNN, Convolutional neural networks) avec pré-apprentissage. Les CNNs peuvent être utilisés avec succès pour la classification de panneaux solaires dans des images aériennes. Le modèle le plus performant était la version pré-entraînée du réseau Xception où toutes les 126 couches étaient configurées pour apprendre. Lors des tests et de l'apprentissage, l'exactitude était de plus de 90%. Toutefois, nous avons aussi découvert que la généralisation est un problème, puisqu'il ne semble pas possible de transférer l'apprentissage d'Xception à d'autres régions géographiques. Toutefois, la performance du modèle VGG16, lors des expériences de généralisation, suggère qu'il est possible d'appliquer avec succès un modèle entraîné avec les images d'une autre région.

Mots clés : mégadonnées, transition énergétique, apprentissage automatique, panneau solaire

1. Introduction

1.1 Contexte

Ce projet est réalisé dans le cadre de la transition énergétique en cours en Europe. Cette transition vise à créer des systèmes énergétiques urbains, qui émettent moins de carbone et utilisent moins d'énergie, et est étroitement associée aux Objectifs de développement durable 7 et 11. Aux Pays-Bas, les statistiques actuelles sur l'énergie solaire sont basées sur des enquêtes portant sur l'importation de panneaux solaires et la production solaire globale est basée sur l'estimation de la capacité installée et la capacité de production par unité. Cette méthodologie fournit seulement des chiffres nationaux sur une base annuelle et a une incertitude estimée de 20 pourcent, tandis que la demande d'information est au niveau régional et sur des périodes de temps plus courtes. Les images aériennes de satellite à haute résolution fournissent une abondance d'informations statistiques régionales, toutefois, l'analyse manuelle des images est laborieuse et longue. Une méthode semi-automatique rapide avec un faible taux d'erreur est donc d'un grand intérêt. Dans ce projet, Statistics Netherlands (CBS) de concert avec la Open University et en collaboration avec d'autres agences nationales de statistiques en Belgique et en Allemagne mettent leur expertise en commun pour fournir un aperçu complet et détaillé des installations actuelles de panneaux solaires en Flandres, en Westphalie au nord du Rhin, et dans le Limbourg. Le processus d'extraction de l'emplacement des panneaux solaires à partir d'images sera automatisé pour produire des cartes de la distribution spatiale des panneaux solaires ainsi que des statistiques régionales. En pratique, plusieurs approches d'apprentissage automatique, telles que la forêt aléatoire (RF, Random Forest), la machine à vecteurs de support (SVM, Support Vector Machine) et les réseaux neuronaux convolutifs profonds (c.-à-d. l'apprentissage profond (DL, Deep Learning)), sont à l'étude et les registres disponibles, tels que les informations concernant les déclarations de taxe sur la valeur ajoutée (VAT, Value Added Tax) des propriétaires de panneaux solaires et l'information provenant des fournisseurs d'énergie, sont utilisés pour entraîner et valider l'algorithme. Dans cet article, les premiers résultats des expériences basées sur la méthodologie d'apprentissage profond sont présentés pour l'état du Nord Rhin Westphalie (NRW, DE). Les résultats présentés dans ce document ont fait l'objet de discussions approfondies dans [6], auquel nous référons le lecteur pour des informations plus détaillées.

¹ Statistics Netherlands and Business Intelligence and Smart Services Institute, The Netherlands (rll.curier@cbs.nl);

2. Jeux de données

Les résultats présentés dans cet article portent seulement sur la convenance des images aériennes. Deux jeux de données sont employés:

1. Images aériennes du Nord Rhin Westafalie de OpenGeodata . NRW [4] est disponible dans le format 20x20cm. Les images disponibles de quatre villes différentes ayant un nombre élevé de panneaux solaires et différents paysages urbains furent choisies, Aachen, Munster, Dusseldorf et Hopsten, et un jeu de données équilibré de 4 665 images annotées fut créé. Pour ce faire, les images disponibles dans le format 30 000x30 000 pixels furent téléchargées et annotées pour chacune de ces villes. À partir de ces images, 2 332 panneaux solaires furent identifiés and découpés en morceaux de 75x75 pixels, pour réduire le volume de traitement. Pour obtenir des images négatives, plusieurs images furent utilisées comme tuiles dans des découpages de 75x75 pixels de façon aléatoire et des découpages sans panneaux furent sélectionnés.
2. Images aériennes des données de Californie [1]. C'est le jeu de données le plus utilisé dans les articles de recherche sur la détection des panneaux solaires [1]. Ces données ont été utilisées comme référence dans l'expérience 2 et 3 pour comparer les résultats du meilleur modèle de la première phase d'expérimentation et évaluer à quel point elle peut être généralisée. Ces données contiennent des annotations sur les panneaux solaires provenant de 601 images de quatre villes californiennes, à une résolution d'au moins 30 cm² par pixel. À partir de ceci, un jeu de données équilibré de 39 722 images annotées de 75x75 pixels fut créé pour l'apprentissage et les tests.

3. Méthodologie

3.1 Apprentissage profond

L'apprentissage profond [3] essaie d'imiter l'activité des couches de neurones dans le néocortex, la partie ridée qui constitue 80% du cerveau où la pensée se produit. Le logiciel apprend véritablement à reconnaître des patrons dans les représentations digitales de sons, d'images et d'autres types de données. L'idée principale est déjà vieille de plusieurs décennies, et elle a conduit à plusieurs échecs et percées. Toutefois, à cause de l'amélioration des formules mathématiques et d'ordinateurs toujours plus puissants, les informaticiens peuvent plus que jamais modéliser plus de couches de neurones virtuels. Dans cette expérimentation, les architectures antérieures seront étudiées. Ces architectures sont disponibles dans la librairie Keras avec des poids pré-entraînés.

3.2 Apprentissage par transfert

L'apprentissage sur de grands jeux de données est à la fois coûteux en calculs et long avec les réseaux neuronaux basés sur l'état actuel d'avancement des connaissances. Par conséquent, dans notre cas, entraîner un CNN à partir de rien n'est pas une option idéale et l'apprentissage par transfert est préférable. Le principal avantage de l'apprentissage par transfert est de rapidement créer des modèles puissants. L'apprentissage par transfert consiste à employer des modèles pré-entraînés pour classer des images sur lesquelles ils n'ont jamais été entraînés et qui pourraient même utiliser des catégories différentes. L'apprentissage par transfert peut se faire avec diverses approches:

- Le nouveau jeu de données est plutôt petit et très différent du jeu de données original. Toutes les couches entièrement connectées du modèle pré-entraîné, sauf la dernière, sont considérées et utilisées en tant qu'extracteurs de caractéristiques comme dans les approches traditionnelles de vision par ordinateur. Un classificateur par apprentissage automatique (par ex. Un SVM) est entraîné au sommet des couches avec les catégories du nouveau jeu de données.
- Le nouveau jeu de données est très différent du jeu original mais large. Cette approche ajuste les poids du modèle pré-entraîné, soit pour toutes les couches ou juste pour les dernières, tout en gardant les autres couches gelées.
- Enfin, quelques unes des couches du réseaux pré-entraîné sont utilisées pour un nouveau modèle, suivi par les blocs de construction convolutifs typiques, qui seraient entraînés à partir de rien.

En général, les premières couches d'un CNN détecte les caractéristiques plus générales, comme les filtres de Gabor ou les gouttes de couleur tandis que les neurones des couches hautes sont spécialisés dans des tâches particulières et apprennent des caractéristiques de plus en plus complexes ou composites. Ainsi, placer la coupure, entre les couches qui apprennent et celles qui sont gelées, à ces emplacements peut conduire à réduire la performance [5]. Dans cette étude, la coupure, entre les couches qui apprennent et celles qui sont gelées, fut effectuée approximativement après deux tiers des couches du réseau.

3.3 Description du modèle

Tel que mentionné à la Section 3.2, les modèles pré-entraînés existants sont étudiés et aussi entraînés sur les images aériennes des jeux de données.

Ici, quatre modèles furent étudiées, notamment InceptionV3, InceptionResNetV2, DenseNet et Xception. Toutes ces architectures ont été créées au cours des trois dernières années. Elles sont disponibles dans la librairie Keras avec des poids pré-entraînés à partir de ILSVC [2] et ensemble, ils représentent l'amélioration récente la plus importante des CNN, en tant que modules de démarrage et couches résiduelles. Néanmoins, il y a des différences importantes dans les architectures de leurs réseaux, comme par exemple leurs nombres de paramètres et de couches (voir la table 1 ci-dessous). Pour davantage d'informations, voir [6].

3.4 Configuration d'apprentissage

Dans cette étude, les modèles InceptionV3, InceptionResNetV2, DenseNet et Xception, avec les poids du ILSVC [2] comme point de départ, seront entraînés sur nos nouveaux jeux de données. Les modèles pré-entraînés furent adaptés pour le cas particulier de la détection des panneaux solaires, la couche extrante des modèles est modifiée par addition d'une moyenne globale par regroupement pour réduire la dimension, une couche totalement connectée de 512 noeuds et finalement une couche sigmoïde fut ajoutée.

3.5 Évaluation

Dans cet article, les trois expériences effectuées pour évaluer les modèles sont présentées. L'exactitude et le score F1 sont les 2 mesures qui furent choisies pour l'évaluation. L'exactitude mesure la proportion globale de bonnes prédictions parmi le nombre total de prédictions, c.-à-d. le pourcentage d'images qui sont correctement classées (positives ou négatives). Dans le cas d'une classification binaire, le score F1 est une mesure de l'exactitude du test et est utile quand vous voulez rechercher l'équilibre entre la précision Error! Bookmark not defined. et le rappel Error! Bookmark not defined.. Le score F1 est une moyenne harmonique de la précision et du rappel, où un score F1 atteint sa meilleure valeur à 1 (précision et rappel parfaits) et sa plus basse valeur à 0.

4. Premiers résultats

Cet article présente les premiers résultats pour l'application des cinq modèles de CNN: VGG16, InceptionV3, InceptionResNetV2, DenseNet et Xception. Ces modèles sont publiquement disponibles dans la librairie Keras. Les modèles sont ajustés sur des photographies aériennes de paysages urbains différents, en utilisant les registres existants pour ajuster et valider les algorithmes développés.

Le point de départ pour les modèles est l'utilisation des poids de la Compétition ImageNet de reconnaissance visuelle à grande échelle (ILSVRC, ImageNet Large Scale Visual Recognition Competition) [2]. Dans cette compétition, différents algorithmes sont testés en ce qui concerne leur performance de détection d'objets et la classification de différentes images. La détection des panneaux solaires ne fait pas partie de cette compétition.

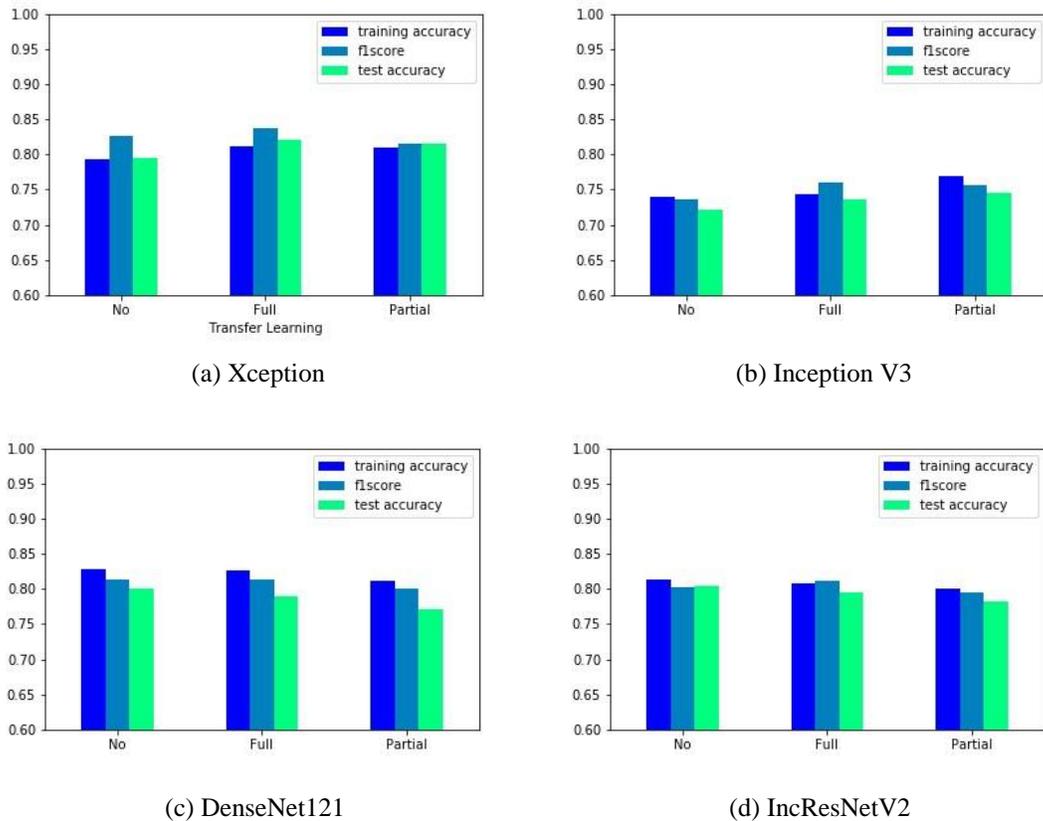
4.1 Expérience 1 – Apprentissage par transfert

Dans cette expérience, nous étudions dans quelle mesure les modèles qui ont déjà été ajustés pour d'autres applications peuvent aussi être utilisés pour la détection des panneaux solaires. À cette fin, les modèles InceptionV3, InceptionResNetV2, DenseNet et Xception sont considérés pour différentes options d'apprentissage par transfert.

Les résultats sont présentés dans la Figure 4.1-1, ci-dessous. Le modèle le plus performant est la version pré-ajustée du réseau Xception, où toutes les 126 couches sont configurées pour être ajustables, Figure 2(a). Une exactitude de 81,08% avec le jeu d'apprentissage et de 82.01% avec le jeux de test, avec un score F de 83.8%. Ces résultats sont en accord avec les résultats observés dans le ILSVRC [2] où le modèle Xception surpassa DenseNet121 et InceptionV3. À la différence de la compétition renommée, avec le jeu de données NRW Xception surpassa aussi InceptionResNetV2. Il faut aussi noter que les trois versions du modèle InceptionV3 ont la pire performance et que le modèle InceptionResNetV2 ajusté avec une approche a une performance légèrement meilleure que le modèle DenseNet ajusté avec la même approche.

Figure 4.1-1

Expérience 1: Aperçu des résultats pour chaque modèle de base et pour l'apprentissage par transfert.



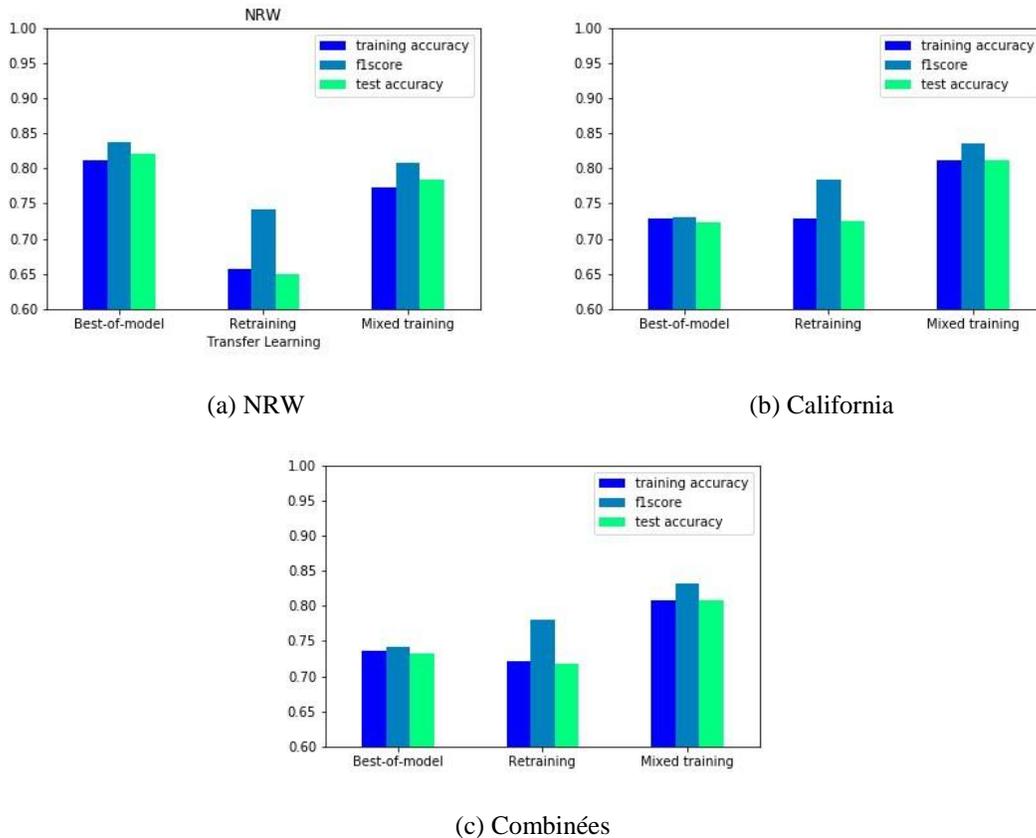
4.2 Expérience 2 – Jeu de référence et généralisation

Le but de cette expérience est d'étudier la performance de modèles pré-ajustés lorsqu'ils sont appliqués à des photographies aériennes d'autres régions. À cette fin, le modèle le plus performant de l'expérience 1: Xception avec apprentissage par transfert total ajusté sur les données du Nord Rhin Westfalie est utilisé de la sorte sur les photographies aériennes de Californie. Deuxièmement, le meilleur modèle mentionné plus haut est utilisé comme référence, ajusté sur le jeu de donnée de Californie et ensuite évalué sur les jeux de données du NRW et de Californie. Finalement, le meilleur modèle susmentionné est réinitialisé avec les poids pré-ajustés ILSVRC et ensuite ajusté sur un jeu d'apprentissage combiné avec des images du NRW et de Californie. La performance du modèle, présentée à la Figure 4.2-2, a clairement baissé avec l'application au jeu de données Californie – sans ajustement supplémentaire. En réajustant le modèle complètement, les résultats s'améliorent, mais de façon surprenante le meilleur résultat est obtenu si le modèle est seulement partiellement réajusté (quelques couches). Le résultat est le même pour l'application au NRW. En cas de réajustement partiel ou total, et spécialement dans ce

dernier cas, le modèle a une moins bonne performance, possiblement à cause du surapprentissage.

Figure 4.2-2

Expérience 2: Aperçu des résultats pour le modèle Xception avec les données du NRW, de Californie et combinés, respectivement.

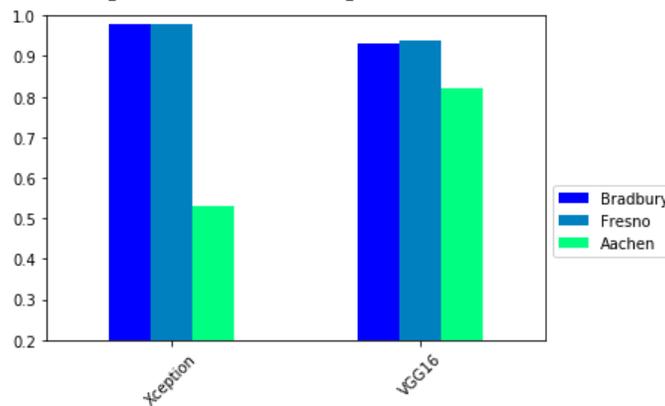


4.3 Expérience 3 – Généralisation en validation croisée

Dans la troisième et dernière expérience une validation croisée est effectuée entre Xception, le meilleur modèle susmentionné, et VGG16, une architecture plus traditionnelle. Les modèles sont ajustés à Bradbury (Californie) et appliqués à Aachen (Nord Rhin Westfalie) et Fresno (Californie). Dans les résultats sont présentés à la Figure 4.3-3, le modèle Xception surpasse légèrement le modèle VGG16 à Bradbury et Fresno. En utilisant Xception sur le jeu de données de Aachen, l'exactitude et la précision se détériorent rapidement. Cependant, la meilleure performance du modèle VGG16 sur Aachen indique que ce dernier généralise mieux que Xception. Cette baisse de performance pourrait s'expliquer par le surapprentissage de Xception sur les données.

Figure 4.3-3

Expérience 3 : Aperçu du score F pour les modèles Xception et VGG16 sur Bradbury, Fresno et Aachen.



5. 5. Conclusion et prochaines étapes

Les premiers résultats de cette étude montrent que les modèles pré-entraînés existants peuvent être utilisés avec succès pour classer les panneaux solaires à partir de photographies aériennes. Les modèles les plus performants, Xception et VGG16, atteignent une précision d'apprentissage et de test et un score F1 de plus de 90%. L'expérience 1 montre que le choix d'un modèle de base est d'une importance primordiale et que ce choix ne peut pas toujours être déduit de la littérature. De plus, la généralisation à de nouveaux jeux de données a été étudiée, les expériences 2 et 3, ce qui, selon les auteurs, n'a pas été fait auparavant. Les résultats indiquent que l'application réussie d'un modèle entraîné sur des images d'une zone différente est possible, sans perte extrême d'informations. Les performances et la généralisation d'un tel réseau peuvent être considérablement augmentées, non pas en réentraînant le modèle sur les nouvelles données, mais en prenant le modèle uniquement avec ses poids pré-entraînés de l'ILSVRC et en répétant la procédure d'entraînement avec un jeu de données mixte.

Remerciements

Cette recherche est menée dans le cadre de l'action ESS 'Fusionner la géostatistique et l'information géospatiale dans les États membres' (convention de subvention n°: 08143.2017.001-2017.408) et d'un investissement CBS pour le développement d'un algorithme d'apprentissage profond.

Bibliographie

- [1] Kyle Bradbury, Raghav Saboo, Timothy L. Johnson, Jordan M. Malof, Arjun Devarajan, Wuming Zhang, Leslie M. Collins, and Richard G. Newell. Distributed solar photovoltaic array location and extent dataset for remote sensing object identification. *Scientific Data*, 3:160106, December 2016.
- [2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.
- [3] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [4] Bezirksregierung Köln Geobasis NRW. Digitale orthophotos - paketiert: nach gemeinden [data set]. Retrieved June 14, 2018, from <https://www.opengeodata.nrw.de/produkte/geobasis/dop/dop/>.

- [5] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? *CoRR*, abs/1411.1792, 2014.
- [6] Hannah. Ziemons. Chasing the sun: A deep learning approach to solar panel classification in aerial images. Master's thesis, Maastricht University School of Business and Economics, 2018.