

Projets de vision artificielle

3A option Intelligence Artificielle

Bordeaux INP – ENSEIRB-MATMECA – ENSC

2020–2021

Michaël Clément

michael.clement@enseirb-matmeca.fr

1 Présentation du projet

Vous devez former des groupes de **3 à 5 personnes**. Ensuite, chaque groupe choisit un projet parmi ceux proposés ci-après. Les sujets sont volontairement assez ouverts et malléables, ceci afin de laisser aux groupes la liberté d'expérimenter et d'explorer différentes approches. Une partie intégrante du projet est de faire une étude des méthodes existantes pour le sujet (modèles, jeux de données, évaluations, etc.) et potentiellement de proposer des idées et variantes originales. Je pourrai vous donner des idées et des conseils au fur et à mesure de votre avancement. Si vous êtes bloqués ou avez des questions, n'hésitez pas à me contacter.

Le rendu de chaque groupe sera composé de :

- **L'ensemble du code réalisé** : fichiers Python et/ou Jupyter Notebooks utilisés pour le projet (définition des modèles, scripts d'apprentissage et de test, résultats des expériences, etc.).
- **Un rapport de maximum 8 pages**, au format d'un article scientifique, suivant globalement la structure suivante :
 - Introduction présentant le contexte du projet et le problème étudié.
 - Description de la ou des méthodes utilisées (algorithmes, modèles, etc.).
 - Présentation du protocole expérimental (jeu de données, paramètres, etc.) et des résultats quantitatifs et qualitatifs obtenus.
 - Discussion et conclusion : synthèse de ce que vous avez appris, quelles étaient les difficultés, comment votre projet pourrait être amélioré, etc.
 - Références bibliographiques.
- **Un fichier README** contenant une description du projet, l'organisation des fichiers, et toute instruction supplémentaire nécessaire pour l'exécution du code.

Les projets seront à rendre début janvier 2021. Si la situation sanitaire le permet, la dernière séance pourra servir à faire des présentations de vos travaux. L'évaluation prendra en compte l'ensemble des aspects explorés dans le projet : investissement au fil des séances, qualité du code produit, pertinence et originalité des méthodes proposées, rigueur et qualité des résultats expérimentaux, compétences rédactionnelles et de présentation.

Remarque : pour réaliser le projet, vous pouvez vous inspirer d'implémentations déjà disponibles en ligne. Cependant, vous êtes invités autant que possible à développer vos propres programmes. N'oubliez pas que l'objectif du projet est avant tout pédagogique, et que vous devez comprendre ce que vous utilisez. Dans tous les cas, en cas de réutilisation de codes ou de résultats, vous devez impérativement faire références aux auteurs initiaux.

2 Sujets proposés

2.1 Description textuelle d'images

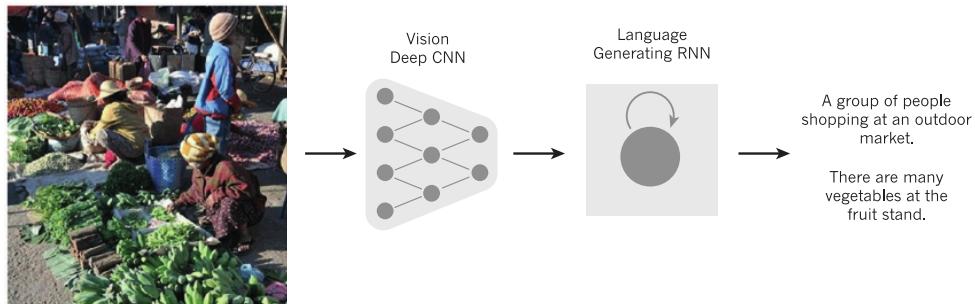


Figure 1 – Description textuelle d'images (*image captioning*) [1].

L'objectif de ce sujet est de développer une méthode permettant de générer une description en langage naturel du contenu d'une image (voir la Figure 1). Ce problème est souvent désigné sous le nom de *image captioning*. Pour cela, une approche communément employée consiste à coupler un réseau convolutif (calculant des caractéristiques visuelles) avec un réseau récurrent (par exemple *LSTM*) permettant de manipuler des données textuelles. Pour ce sujet, vous pourrez notamment vous inspirer des travaux suivants : [1] et [2].

2.2 Colorisation d'images



Figure 2 – Colorisation d'images [3].

L'objectif de ce sujet est de développer une méthode permettant de coloriser des images initialement à niveaux de gris (voir la Figure 2). Le principe consiste généralement à utiliser un réseau de neurones afin de prédire les canaux manquants correspondant aux couleurs. Plusieurs variantes peuvent être envisagées : colorisation à partir d'image(s) de référence, interaction utilisateur avec marqueurs manuels, etc. Pour ce sujet, vous pourrez notamment vous inspirer des travaux suivantes : [3] et [4].

2.3 Transfert de style

L'objectif de ce sujet est de développer une méthode permettant d'appliquer un style visuel à une image donnée (voir la Figure 3). Pour cela, une approche consiste à utiliser un réseau convolutif pré-entraîné pour la classification. À partir de ce réseau, il s'agit de modifier progressivement l'image de test de façon à ce que les *feature maps* obtenues aux différentes couches

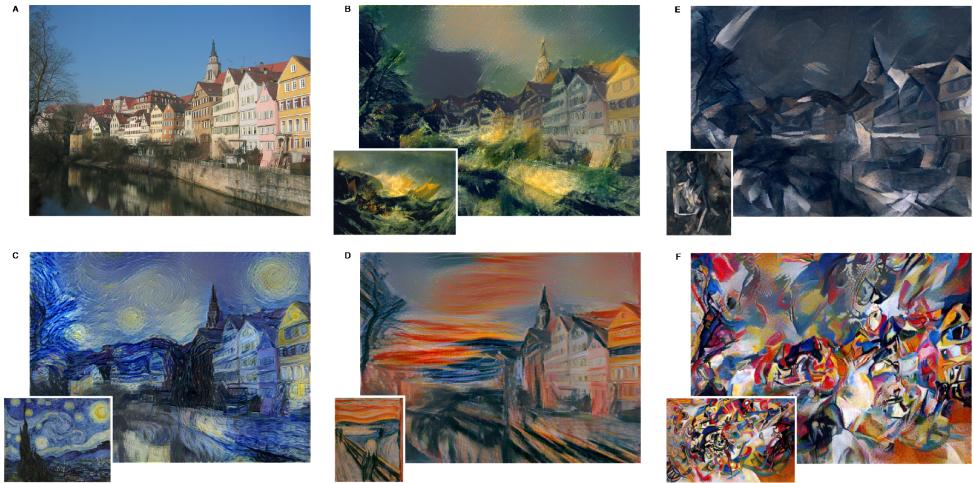


Figure 3 – Transfert de style [5].

du réseau correspondent à celles des images de référence et de style souhaitées. Pour ce sujet, vous pourrez notamment vous inspirer des travaux suivants : [5] et [6].

2.4 Manipulation de dessins

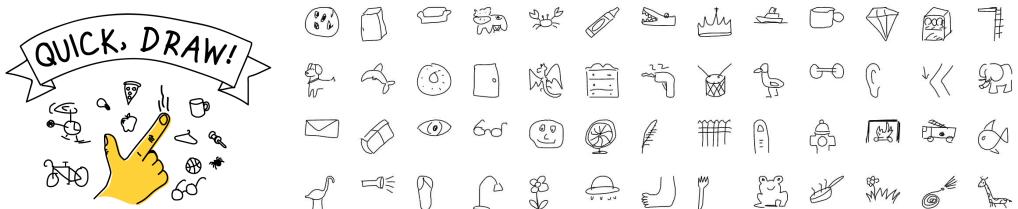


Figure 4 – Application *Quick Draw!* pour la reconnaissance de dessins.

Dans ce sujet, il s'agit de s'intéresser à l'application *Quick, Draw!* développée par Google, qui est disponible ici : <https://quickdraw.withgoogle.com>. L'application propose également un vaste jeu de données de dessins réalisés par les utilisateurs, et l'objectif du projet serait de développer une méthode exploitant un sous-ensemble de ces données. Les possibilités sont variées : par exemple, la méthode pourrait reconnaître le dessin réalisé par l'utilisateur (comme dans l'application), compléter un dessin déjà commencé [7], ou bien encore rechercher dans une base des images ressemblant à un dessin existant.

2.5 Segmentation d'IRM 3D du cerveau

Ce projet se focalise sur l'imagerie médicale. L'objectif est de développer une méthode permettant de segmenter des régions du cerveau humain à partir d'imagerie à résonance magnétique (IRM) en 3D. Il peut s'agir de segmenter l'ensemble du cerveau, ou bien de se focaliser sur des régions précises comme par exemple l'hippocampe. Pour cela, une architecture de type *U-Net* [9] pourra être utilisée, et des techniques d'augmentation de données [10] pourront être envisagées. Quelques données IRM anonymisées et annotées pourront vous être fournies.

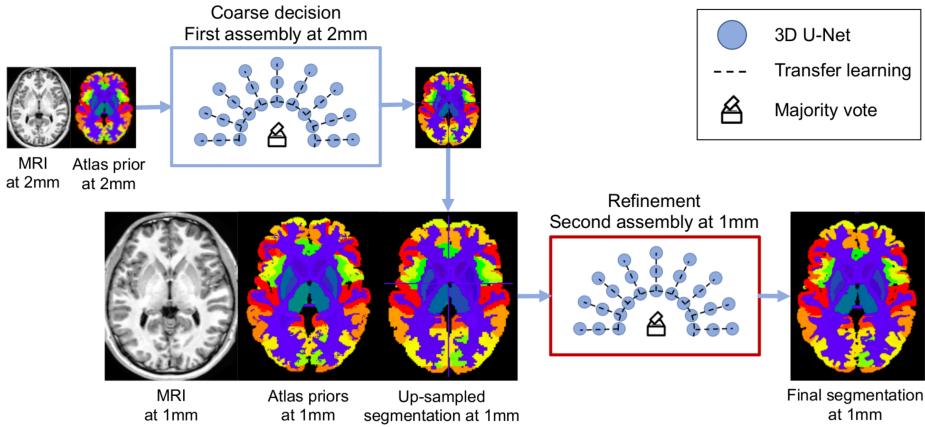


Figure 5 – Segmentation d'IRM 3D du cerveau [8].

Références

- [1] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator", *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [2] A. Karpathy and L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions", *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [3] R. Zhang, P. Isola, and A. A. Efros, "Colorful Image Colorization", *European Conference on Computer Vision (ECCV)*, 2016.
- [4] P. Vitoria, L. Raad, and C. Ballester, "ChromaGAN: Adversarial Picture Colorization with Semantic Class Distribution", *Winter Conference on Applications of Computer Vision (WACV)*, 2020.
- [5] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks", *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [6] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution", *European Conference on Computer Vision (ECCV)*, 2016.
- [7] D. Ha and D. Eck, "A neural representation of sketch drawings", *International Conference on Learning Representations (ICLR)*, 2018.
- [8] P. Coupé, B. Mansencal, M. Clément, R. Giraud, B. D. de Senneville, V.-T. Ta, V. Lepetit, and J. V. Manjon, "AssemblyNet: A novel deep decision-making process for whole brain MRI segmentation", *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2019.
- [9] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation", *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015.
- [10] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond Empirical Risk Minimization", *International Conference on Learning Representations (ICLR)*, 2018.