

CNN pour la reconnaissance d'artiste en peinture

Charles Combelles (charles.combelles@student.ecp.fr)

CentraleSupélec, 20 février 2021

Abstract

Il est parfois difficile pour les historiens de l'art d'identifier avec certitude l'auteur d'un tableau. De nombreuses techniques ont été développées, allant de l'observation visuelle à l'analyse chimique des pigments. L'objectif de ce document est de tenter une autre approche : celle du deep learning. Loin d'avoir la prétention d'égaliser l'expertise des spécialistes ou de distinguer un vrai d'un faux, nous tenterons simplement d'entraîner un modèle capable de prédire qui est l'auteur d'une toile.

1. Introduction

1.1 Données

Le jeu de données [1] utilisé comporte 8446 œuvres réalisées par 50 artistes célèbres. Les artistes varient du plus prolifique, Vincent van Gogh avec 877 œuvres, au moins prolifique, Jackson Pollock, avec 24 œuvres. 15% des œuvres de chaque artiste sont mises de côté en tant qu'ensemble de test.

1.2 Méthodes envisagées

Pour obtenir les résultats souhaités, deux méthodes ont été envisagées.

- Créer un réseau à base de CNN spécialisé pour reconnaître les artistes à partir des œuvres
- Réaliser du *transfer learning* à partir d'un réseau spécialisé dans la

reconnaissance d'images de type ResNet.

Après étude de l'état de l'art, notamment les travaux de [x], et quelques essais infructueux il a été décidé de se concentrer sur le *transfer learning* qui permet d'obtenir de bons résultats avec moins d'entraînement car ResNet a été pré-entraîné. Ce choix a également été motivé par une puissance de calcul limitée à disposition (ordinateur portable).

2. Préprocessing

2.1 Exploration des données

Les données se présentent sous la forme de tableaux, fresques, esquisses d'une grande variété de forme et de style (cf Figure 1). Une première étape est donc de redimensionner toutes les œuvres en un format unique et carré propre à être donné en entrée par le réseau ResNet.



Figure 1 : Différents œuvres du dataset

Le format choisi est (224, 224, 3). Même si cela représente une augmentation significative du temps de calcul, on conserve les canaux RGB car la palette de couleur est un élément important du style d'un artiste.

Une des particularités de ce problème est le peu de données disponible du fait que chaque artiste a produit un nombre limité d'œuvres au cours de sa vie. On utilise différentes méthodes pour remédier à ce problème.

2.2 Elimination des classes avec trop peu d'instances

On retire des données (et donc du modèle) les artistes les moins prolifiques. Cette méthode a pour inconvénient de réduire le domaine d'application du modèle, mais permet à la fois un entraînement plus rapide (moins de classes à différencier) et d'éviter les pertes de précision dues au pur manque de données. Deux versions de cette méthode seront testées : dans un premier temps garder seulement les artistes avec au moins 200 œuvres, puis avec au moins 100 œuvres.

2.3 Data Augmentation

En réalisant de petites transformations aléatoires sur les images, on peut virtuellement démultiplier la taille de l'ensemble d'entraînement. En pratique, on utilisera les transformations suivantes :

- Petites rotations ($<5^\circ$)
- Inversion horizontale
- Modification de la luminosité

Les modifications sont réalisées aléatoirement sur chaque batch, ce qui permet non seulement d'obtenir de meilleurs résultats de prédiction, mais aussi de lutter contre l'overfitting en ayant un ensemble d'entraînement « différent » à chaque epoch.

3. Construction du modèle

3.1 ResNet50

Les Residual Networks ont été introduits par Kaiming He *et al.* [2] comme une manière de régler le problème du *vanishing gradient* dans les réseaux CNN très profonds. L'architecture repose sur le *Residual Block* (gauche sur Figure 2) qui ont la particularité d'avoir une *Skip Connection* [3] qui réinjecte l'entrée en sortie des convolutions.

Ce type de réseau a déjà fait ses preuves dans de nombreuses applications pratiques de traitement d'image tels que la reconnaissance d'objets ou la classification d'images, c'est pourquoi il semble particulièrement adapté au problème présent.

Figure 4 : Poids associés aux classes

index		name	paintings	weight
0	8	Vincent_van_Gogh	877	0.123678
1	30	Edgar_Degas	702	0.098999
2	13	Pablo_Picasso	439	0.061909
3	15	Pierre-Auguste_Renoir	336	0.047384
4	19	Albrecht_Durer	328	0.046256
5	46	Paul_Gauguin	311	0.043858
6	16	Francisco_Goya	291	0.041038
7	31	Rembrandt	262	0.036948
8	20	Alfred_Sisley	259	0.036525
9	32	Titian	255	0.035961
10	22	Marc_Chagall	239	0.033705
11	4	Rene_Magritte	194	0.027359
12	0	Amedeo_Modigliani	193	0.027218
13	42	Paul_Klee	188	0.026512
14	28	Henri_Matisse	186	0.026230
15	45	Andy_Warhol	181	0.025525
16	12	Mikhail_Vrubel	171	0.024115
17	24	Sandro_Botticelli	164	0.023128
18	26	Leonardo_da_Vinci	143	0.020166
19	14	Peter_Paul_Rubens	141	0.019884
20	5	Salvador_Dali	139	0.019602
21	10	Hieronymus_Bosch	137	0.019320
22	21	Pieter_Bruegel	134	0.018897
23	27	Diego_Velazquez	128	0.018051
24	11	Kazimir_Malevich	126	0.017769
25	17	Frida_Kahlo	120	0.016923
26	23	Giotto_di_Bondone	119	0.016782
27	9	Gustav_Klimt	117	0.016500
28	47	Raphael	109	0.015372
29	44	Joan_Miro	102	0.014384

5. Résultats

On compare les performances du modèle pour vérifier l'efficacité des différents ajustements mis en place.

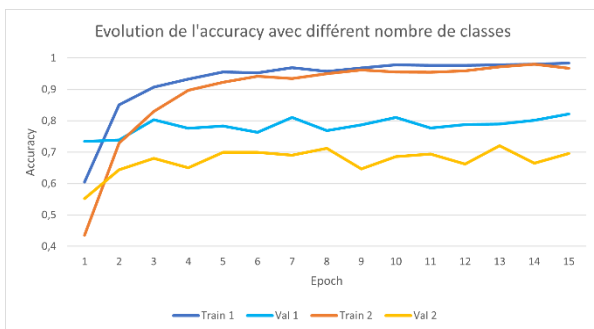


Figure 5 : Comparaison des performances avec 11 (série 1) et 30 (série 2) classes.

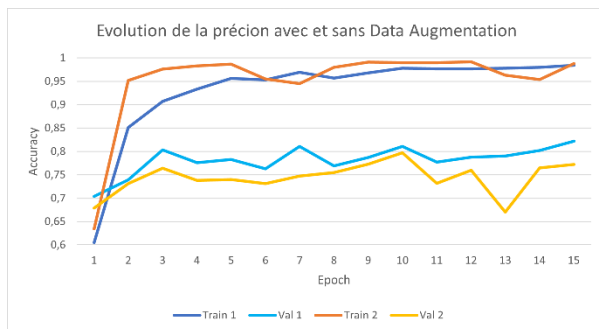


Figure 6 : Comparaison des performances avec (1) et sans (2) data augmentation (test avec 11 classes)

Comme on s'y attendait, on a à la fois un entraînement plus rapide et de meilleures performances en sélectionnant uniquement les classes avec plus de 200 instances. De même, l'utilisation de la data augmentation améliore la précision du modèle sur l'ensemble de validation. La data augmentation semble bien éviter l'overfitting (validation accuracy plus stable).

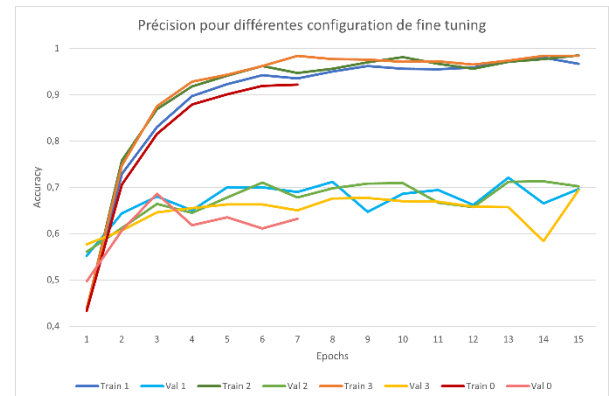


Figure 7 : Performances en réentraînant toutes les couches (0) de ResNet50, les 2 dernières parties (1), la dernière partie (2), le dernier bloc (3) (test avec 30 classes)

Pour ce qui est des choix de fine tuning, les performances semblent moins sensibles à la portion de ResNet50 qui est réentraînée, à partir du moment où on ne part pas dans les extrêmes (réentraînement complet ou quasiment pas de réentraînement, c'est-à-dire séries 0 et 3).

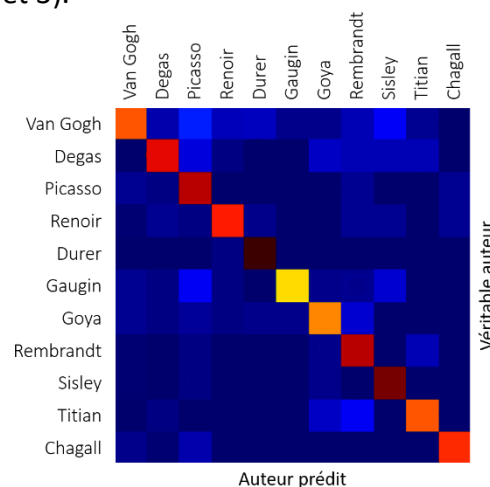


Figure 8 : Matrice de confusion du modèle (sur 11 classes)

L'artiste que le modèle arrive le mieux à reconnaître est Durer, qui a la particularité de n'avoir produit que des esquisses et gravures (donc en niveaux de gris). Les artistes les plus confondus sont Titian et Rembrandt, Gauguin et Picasso, Van Gogh et Picasso, Sisley et Van Gogh.



Rembrandt



Titien



Van Gogh



Sisley

Figure 9 : Comparaison d'artistes

Si Rembrandt et Titien ainsi que Van Gogh et Sisley ont des styles semblables, il ne peut pas en être dit autant de Picasso, Van Gogh et Gauguin. Ces erreurs pourraient être expliquées par la grande variété de style que présentent les œuvres de Picasso et Gauguin.

6. Conclusion

On obtient sur l'ensemble de test des précisions de 81,2% (11 classes) et 69,8% (30 classes). Le modèle ResNet prouve donc encore une fois son efficacité et son adaptabilité à différents problèmes. L'utilisation de la data augmentation et la limitation à des classes ayant un nombre suffisant d'instances sont clés dans l'obtention de bons résultats dans le cadre d'un problème avec peu de données comme celui-ci.

Cependant, le modèle qui a été produit reste plutôt inférieur à un amateur d'art pour ce qui est de reconnaître l'auteur d'un tableau. Notamment, le modèle semble être en difficulté pour reconnaître les artistes dont le style a fortement évolué au cours de leur carrière. Un amateur reconnaîtrait également les peintures de Van Gogh de celles de Sisley assez facilement par son trait de pinceau très particulier.

7. Bibliographie

[1] Kaggle.com. 2021. *Best Artworks of All Time*. <https://www.kaggle.com/ikarus777/best-artworks-of-all-time> (5/01/21)

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition. 2015.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Identity Mappings in Deep Residual Networks. 2016.

[4] Mozejko, M., 2021. *Which layers should I freeze for fine tuning a resnet model on keras?* Stack Overflow. <https://stackoverflow.com/questions/47206714/which-layers-should-i-freeze-for-fine-tuning-a-resnet-model-on-keras> (03/02/21)

