

Étiquetage de la race de chien depuis une image à l'aide d'un réseau de neurone convolutif

Antoine Gaulin, Othman Mounir, Arthur Pulvéric and Anis Redjda

Polytechnique Montréal

Abstract

Décrire la thèse et le résultat en moins de 200 mots. Le document est de 6 pages maximum. L'article a pour but de présenter un modèle de classification et d'étiquetage d'images de chiens en fonction de leur race.

1 Introduction

1.1 Motivations

Les réseaux de neurones convolutifs permettent de faire de la classification d'image de façon rapide et performante. Ainsi, avec des images de chiens, il est possible de calculer la probabilité qu'un canidé présent dans une image appartienne à une race plutôt qu'une autre. Pour aider la SPCA à se créer un inventaire en ligne de tous ses chiens, et encourager le commerce en ligne, nous voulons concevoir un réseau de neurones qui permet de faire de la classification automatique à partir des images de chien. De cette façon, les clients souhaitant faire l'acquisition d'une race spécifique auront plus de facilité à la retrouver sur leur site internet.

Dans la pratique, un réseau de neurones convolutif classique présente des difficultés pour la classification d'objets présentant de nombreuses caractéristiques communes, ce qui est le cas pour les chiens de la base de données. La Fig. 1 illustre bien ce problème : les chiens présentés ont beaucoup de caractéristiques communes mais n'appartiennent pas à la même race. L'objectif ici est donc d'implémenter un modèle capable d'obtenir des performances acceptables pour la prédiction de races de chien.



Figure 1: De gauche à droite, un Malmute, un Eskimo et un Husky

1.2 Travaux antérieurs

L'article *Using Convolutional Neural Networks to Classify Dog Breeds* [?] est un incontournable. Concrètement, les auteurs proposent deux architectures différentes capables de différencier les petites variations entre les races. Le premier modèle est basé sur l'architecture de LeNet, dont chaque couche de convolution est un filtre avec 2 paramètres : le nombre de pixels qu'elles prennent en entrée et le nombre de canaux en sortie. Le deuxième modèle reprend l'architecture de GoogLeNet, dont chaque couche de convolution est une combinaison de filtres eux-mêmes composés de réseaux de neurones.

Un autre papier intéressant est *Dog Breed Identification*. [?] Il y est présenté un projet qui ayant pour but de prédire les races des chiens à partir d'images. Ce projet utilise des techniques d'apprentissage automatique et de vision par ordinateur. Un réseau de neurones convolutionnel permet dans un premier temps d'identifier les points clés des visages des chiens. Ensuite, des descripteurs SIFT (Scale-Invariant Feature Transform) et des histogrammes de couleur utilisent ces points clés pour extraire des features. Enfin, un classificateur de type machine à vecteur de support (SVM, Support-Vector Machine) prédit la race du chien avec une précision de 52%. Il est à noter que dans 90% des cas, la race du chien à trouver est présente dans les 10 prédictions les plus probables (sur 133 races au total dans la base de données). Ces performances sont supérieures à ce que peuvent faire la plupart des humains.

Dans ce travail, nous faisons le choix d'utiliser un modèle proposé par le premier article. En effet, la simplicité d'architecture et leurs bonnes performances sont totalement adaptées aux contraintes que nous rencontrons. Nous allons ajouter un algorithme qui applique un étiquette sur l'image.

2 Approche théorique

Dans le contexte de la pandémie, nous devons limiter le projet à quelque chose de simple, qui permet toutefois de comprendre le fonctionnement d'un CNN moderne en profondeur. Ainsi, nous allons nous limiter à une seule architecture.

Premièrement, nous avons évalué s'il était possible d'ajouter des étiquettes à une image dans ses métadonnées à l'aide du langage Python. Nous avons trouvé la librairie IPTCInfo qui permet de faire cette tâche aisément.

Ensuite, nous avons cherché un dataset pour entrainer notre modèle. Les images retenues proviennent du *Stanford Dogs Dataset*. [?] Le fait que les clichés sont présents dans des répertoires nommés par la race permet facilement de retrouver l'étiquette pour faire l'entraînement supervisé.

Avant de faire une architecture plus difficile, nous avons commencé par un réseau CNN de base.

3 Discussion

Discussion de nos expériences. Incluant des figures, des tableaux de nos résultats.

Le code source est disponible à cette adresse : <https://github.com/Anteige/INF8225>.

3.1 Exemples de tableau

Scenario	δ (s)	Runtime (ms)
Paris	0.1	13.65
	0.2	0.01
New York	0.1	92.50
Singapore	0.1	33.33
	0.2	23.01

Table 1: Booktabs table

3.2 Exemple d'algorithme

Algorithm 1 Example algorithm

Input: Your algorithm's input

Parameter: Optional list of parameters

Output: Your algorithm's output

```

1: Let  $t = 0$ .
2: while condition do
3:   Do some action.
4:   if conditional then
5:     Perform task A.
6:   else
7:     Perform task B.
8:   end if
9: end while
10: return solution

```

3.3 Exemple de formule

$$\begin{aligned}
 x = & \prod_{i=1}^n \sum_{j=1}^n j_i + \prod_{i=1}^n \sum_{j=1}^n i_j + \prod_{i=1}^n \sum_{j=1}^n j_i + \prod_{i=1}^n \sum_{j=1}^n i_j + \\
 & + \prod_{i=1}^n \sum_{j=1}^n j_i
 \end{aligned} \tag{1}$$

4 Analyse

Une analyse critique de l'approche que vous avez utilisée pour apprendre le sujet que vous avez sélectionné.