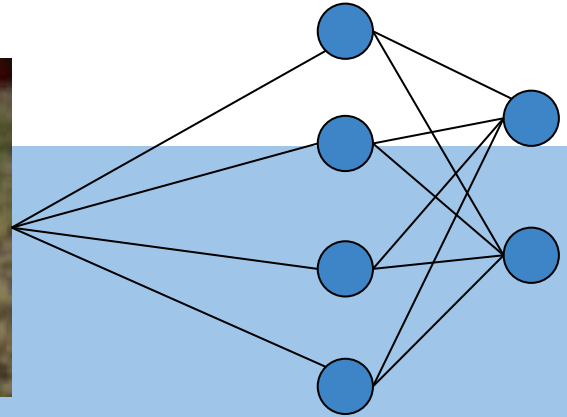


Classez des images à l'aide d'algorithmes de Deep Learning

par Ana Bernal

mentor: Samir Tanfous

Avril 2023



Repo github pour ce projet:

github.com/ana-bernal/P6_img-classif

Programme

- 1 Rappel mission
- 2 Exploration et aperçu des données
- 3 Tests pour classification: *from scratch* et *transfer learning*
Meilleur modèle + inférence
- 4 Conclusions

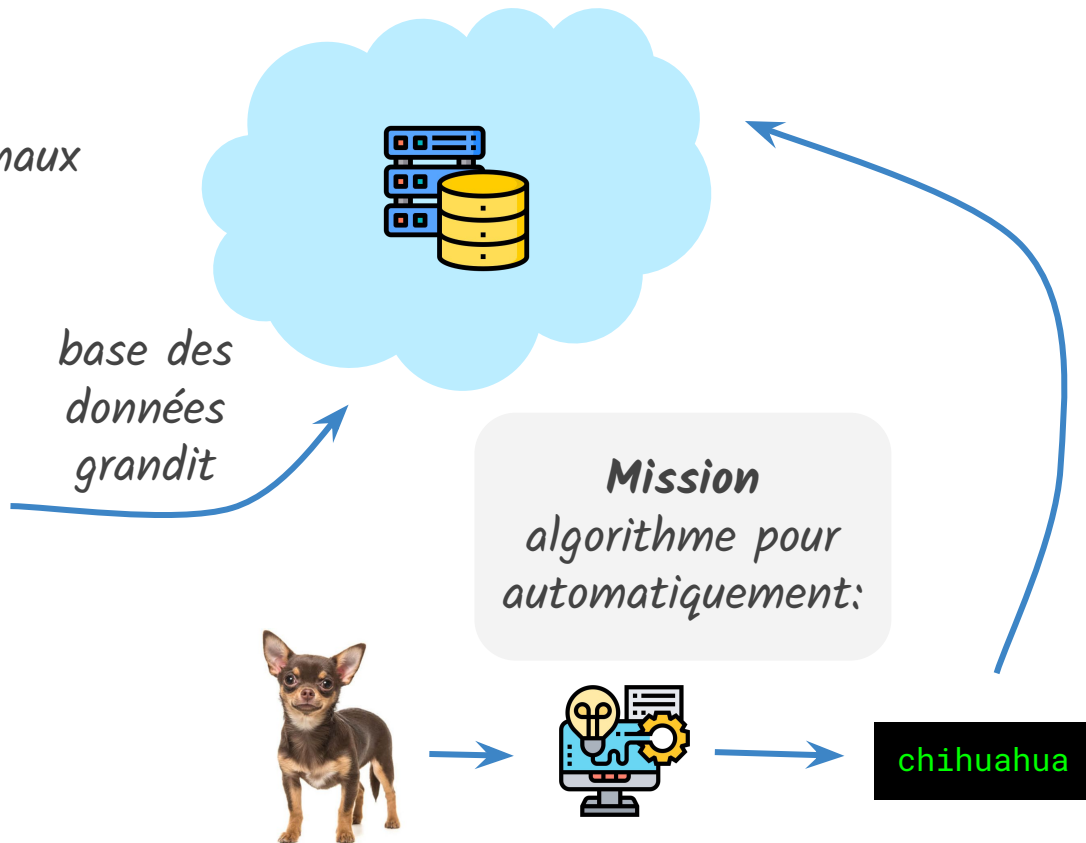
1

Rappel mission

Rappel mission

1

Association de protection d'animaux



2

Exploration des données

Aperçu des données

2

Source des données



Stanford Dogs Dataset

Aditya Khosla Nityananda Jayadevaprakash Bangpeng Yao Li Fei-Fei

Stanford University

The Stanford Dogs dataset contains images of 120 breeds of dogs from around the world. This dataset has been built using images and annotation from ImageNet for the task of fine-grained image categorization. Contents of this dataset:

Quelques statistiques

Nombre total d'images	20581
Nombre de classes (races de chien)	120



Sample of 10 breed names:

```
['Chihuahua',  
'Japanese_spaniel',  
'Maltese_dog',  
'Pekinese',  
'Shih',  
'Blenheim_spaniel',  
'papillon',  
'toy_terrier',  
'Rhodesian_ridgeback',  
'Afghan_hound']
```

Aperçu des données

2

Source des données



Stanford Dogs Dataset

Aditya Khosla Nityananda Jayadevaprakash Bangpeng Yao Li Fei-Fei

Stanford University

The Stanford Dogs dataset contains images of 120 breeds of dogs from around the world. This dataset has been built using images and annotation from ImageNet for the task of fine-grained image categorization. Contents of this dataset:

Quelques statistiques

Nombre total d'images	20581
Nombre de classes (races de chien)	120



Pour classification après	
Nombre total d'images	$2537 = 2030 + 507$
Nombre de classes (races de chien)	15

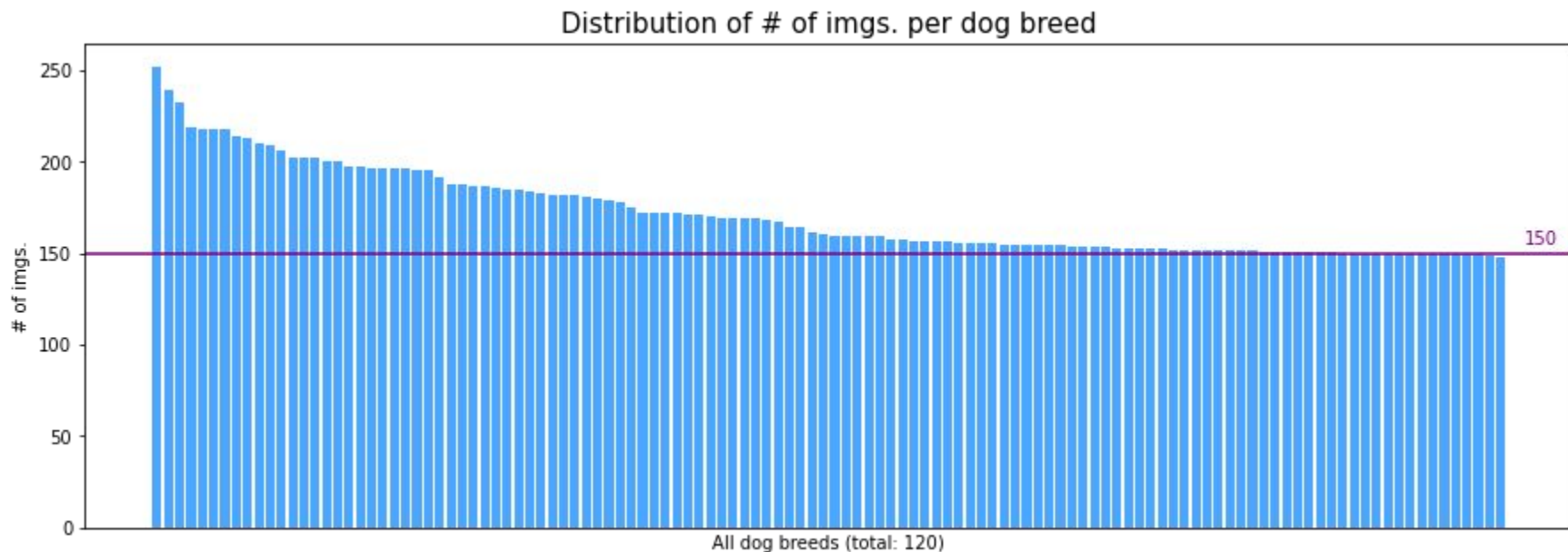
Sample of 10 breed names:

```
['Chihuahua',  
'Japanese_spaniel',  
'Maltese_dog',  
'Pekinese',  
'Shih',  
'Blenheim_spaniel',  
'papillon',  
'toy_terrier',  
'Rhodesian_ridgeback',  
'Afghan_hound']
```

Aperçu des données

2

Distribution du # d'images par race de chien



Aperçu des données

2

Aperçu de quelques images et leurs labels

Maltese_dog



Bernese_mountain_dog



Rhodesian_ridgeback



kuvasz



Brittany_spaniel



Bernese_mountain_dog



schipperke



Kerry_blue_terrier



Weimaraner



Norwich_terrier



EntleBucher



dumber



3

Tests de classification

Classification

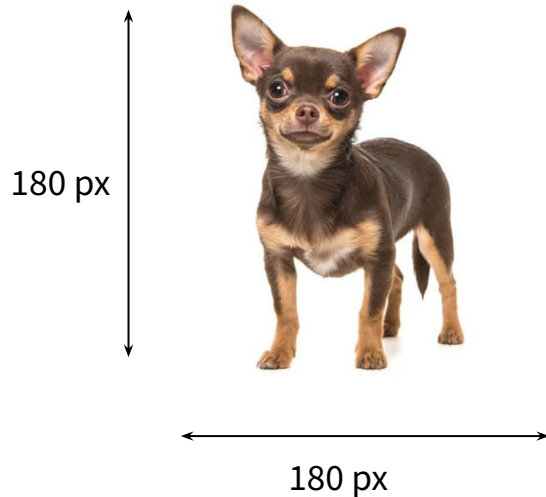
3

- Classification d'images → label.
- Un seul label possible.
- Réseaux des neurones

Classification

3

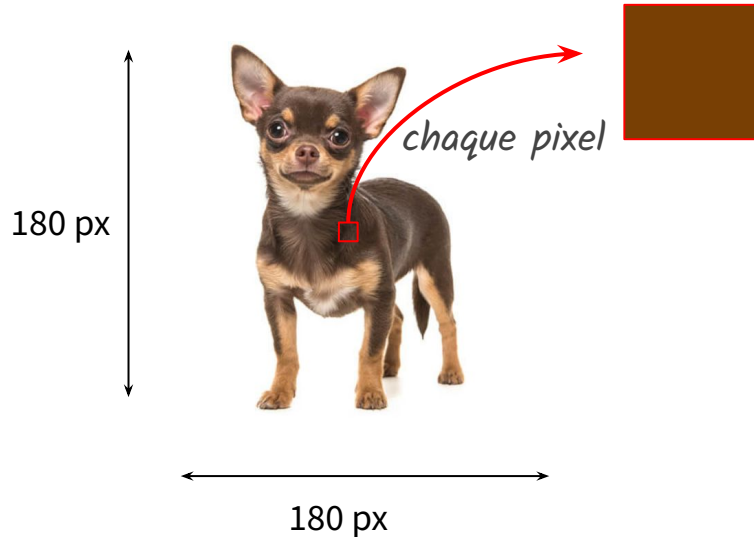
- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



Classification

3

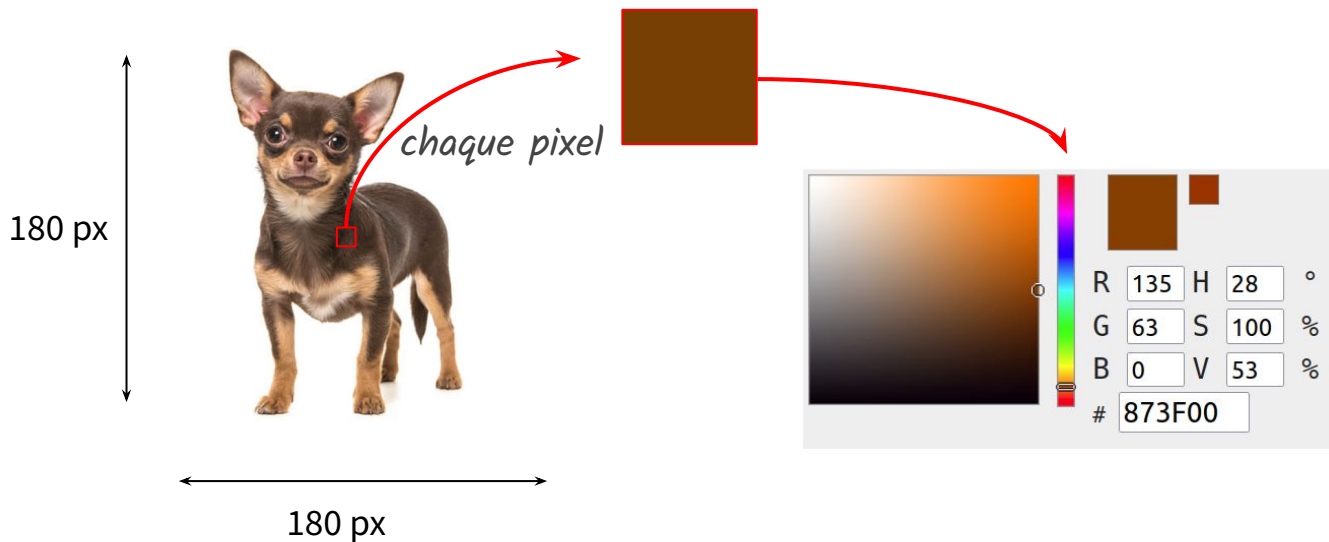
- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



Classification

3

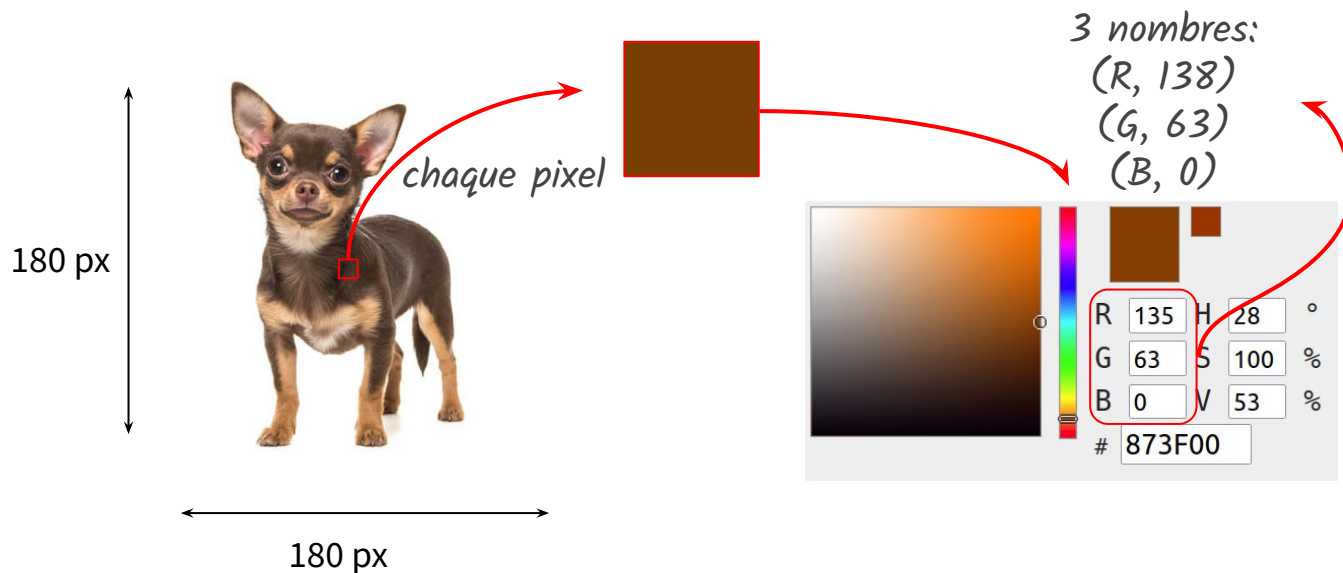
- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



Classification

3

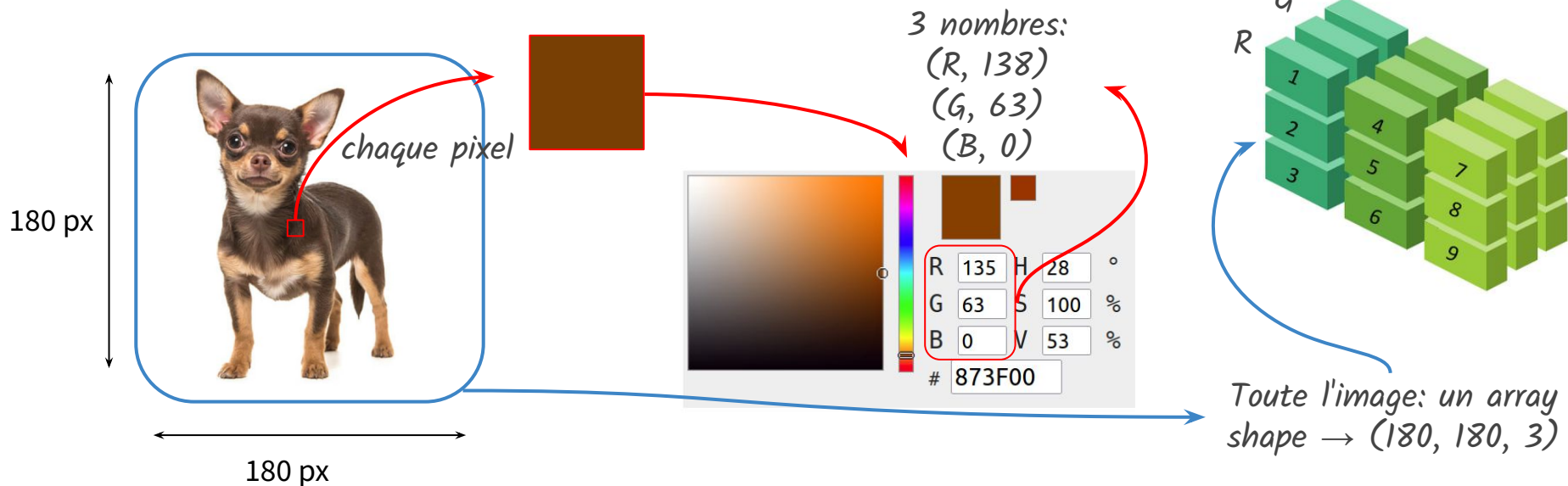
- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



Classification

3

- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



Classification

3

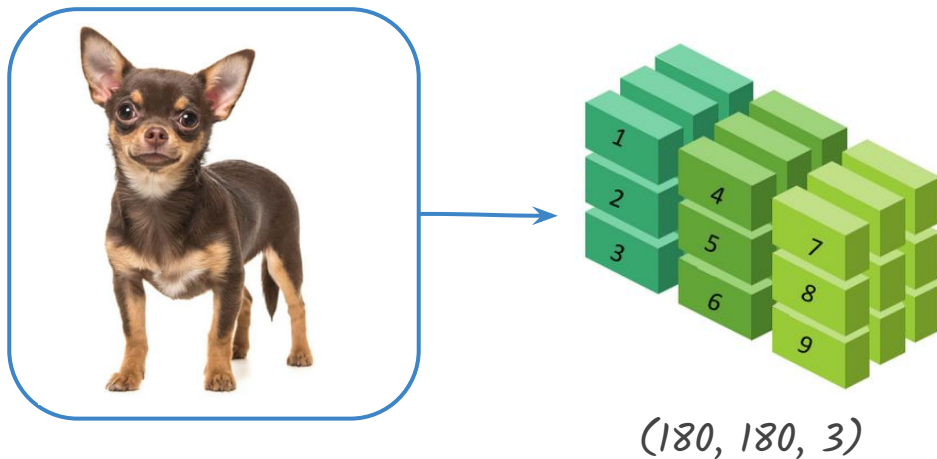
- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



Classification

3

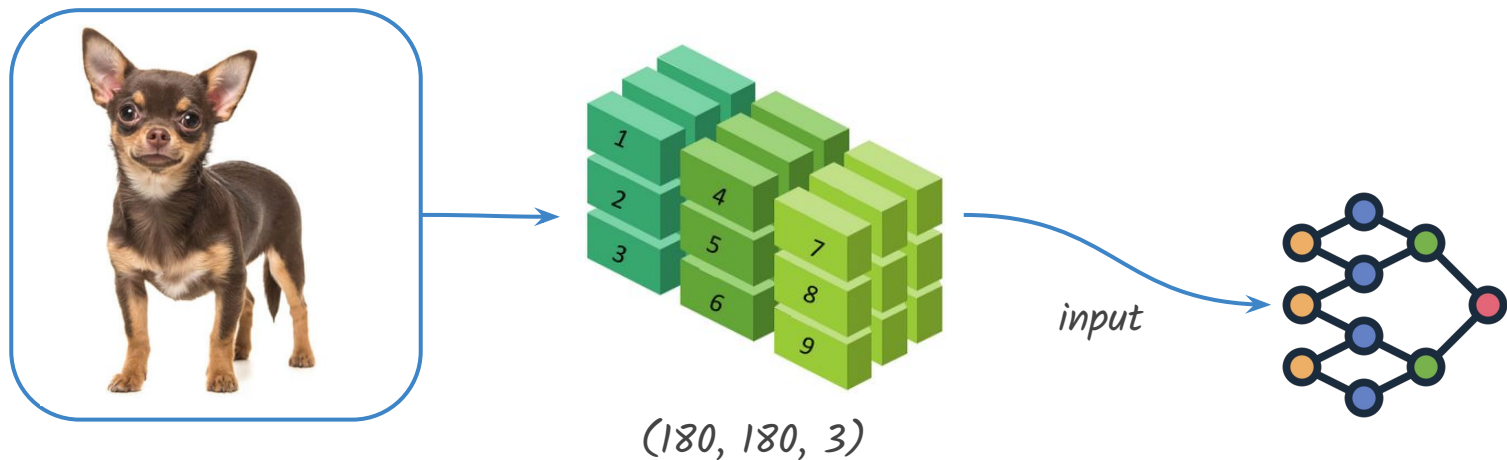
- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



Classification

3

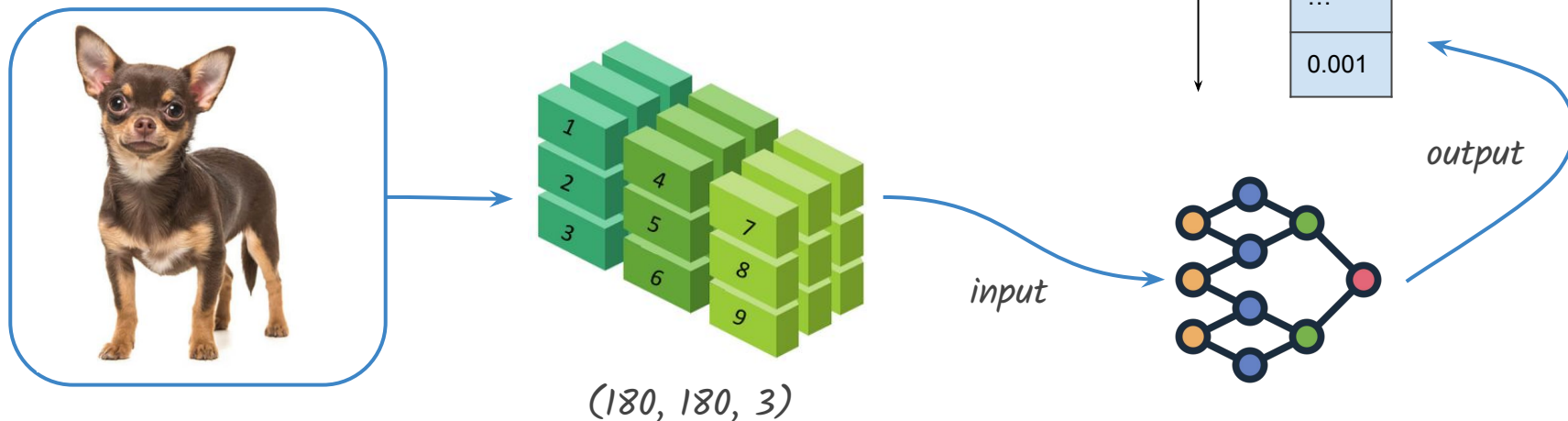
- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



Classification

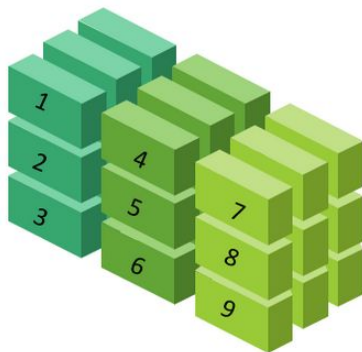
3

- Classification d'images → label
- Un seul label possible
- Réseaux des neurones



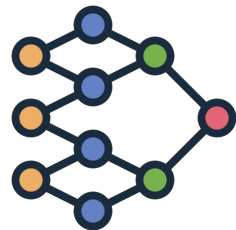
Classification

- Classification d'images → label
- Un seul label possible
- Réseaux des neurones

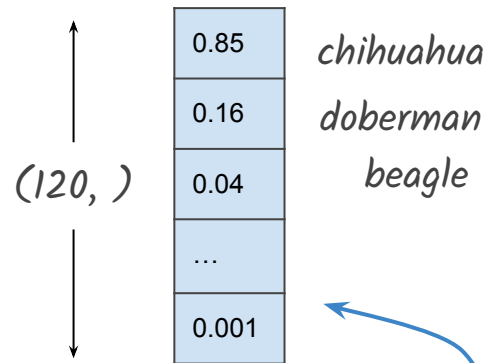


$(180, 180, 3)$

input



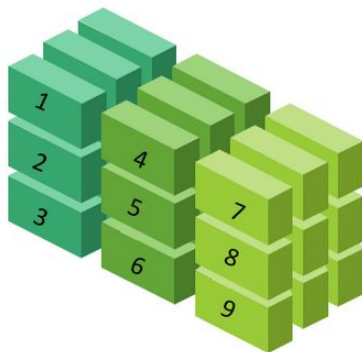
output



3

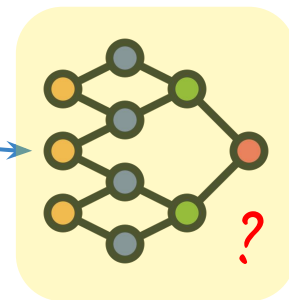
Classification

- Classification d'images → label
- Un seul label possible
- Réseaux des neurones

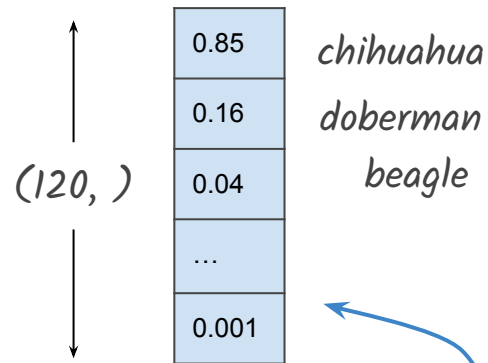


$(180, 180, 3)$

input

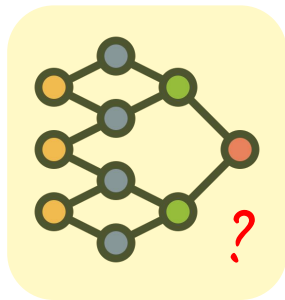


output



3

Classification



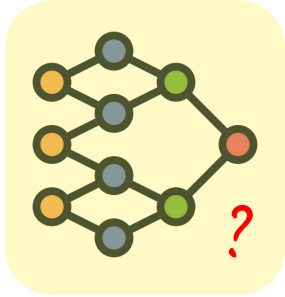
Convolutional Neural Network

Parenthèse:
couche convolution

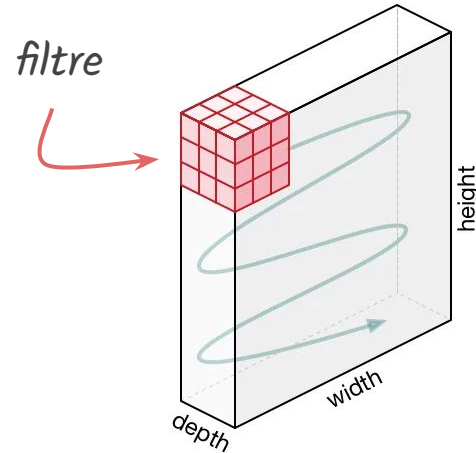
Classification

Convolutional Neural Network

3



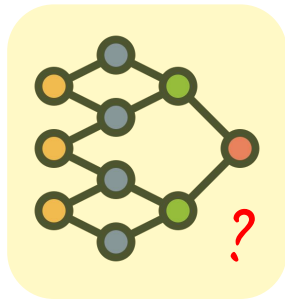
Parenthèse:
couche convolution



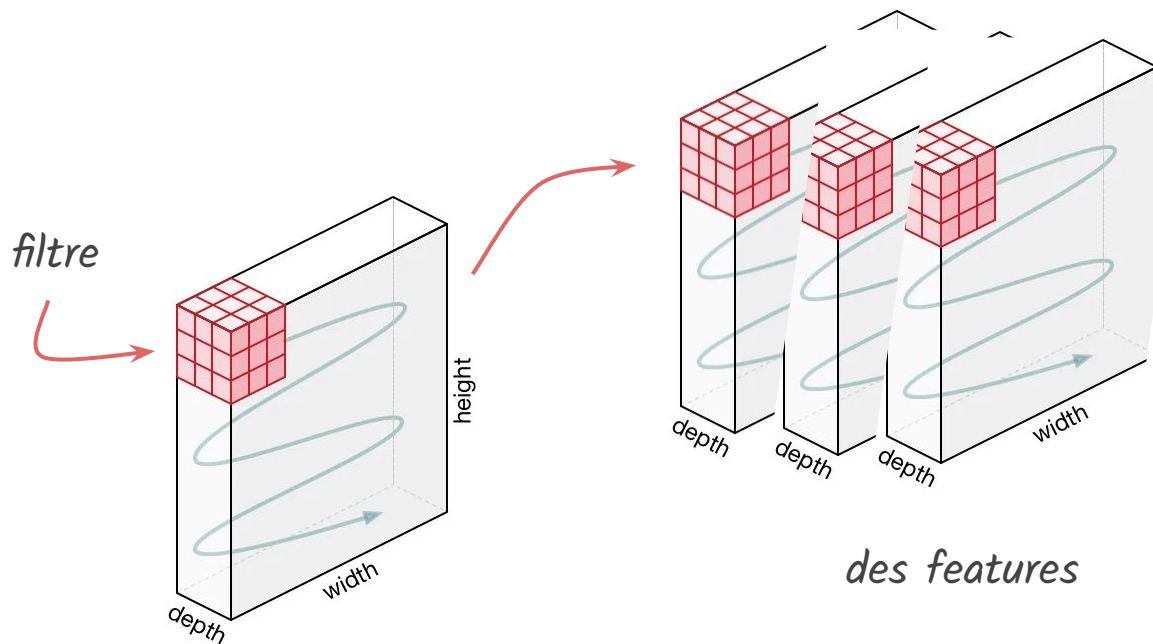
Classification

Convolutional Neural Network

3



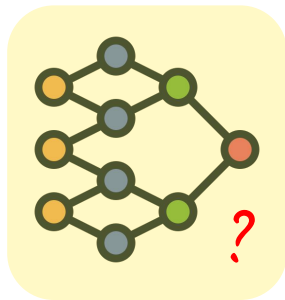
Parenthèse:
couche convolution



Classification

Convolutional Neural Network

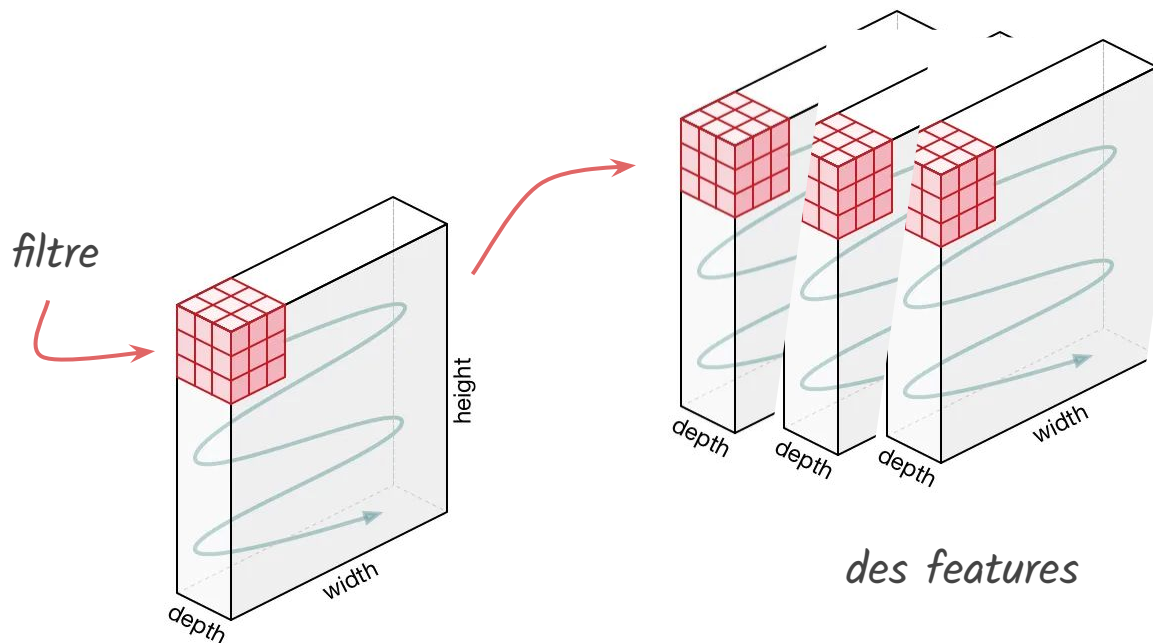
3



Parentèse:
couche convolution

Beaucoup d'hyper
paramètres

filtres, taille filtres
stride
couches conv.
...



Classification

3

Démarche

Classification

3

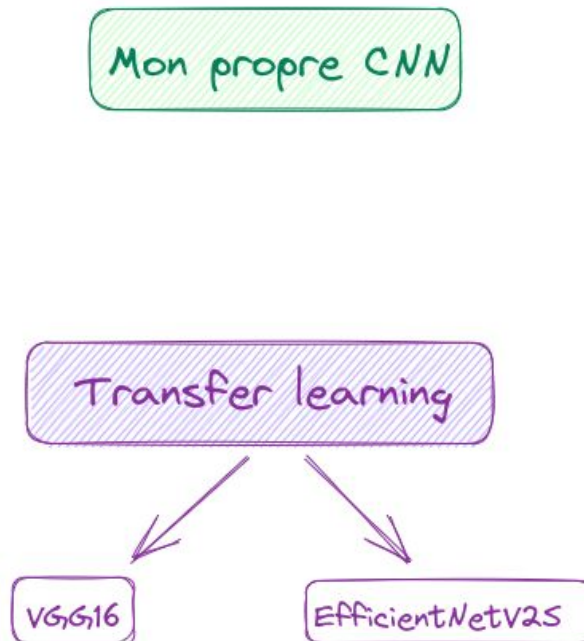
Démarche

Mon propre CNN

Classification

3

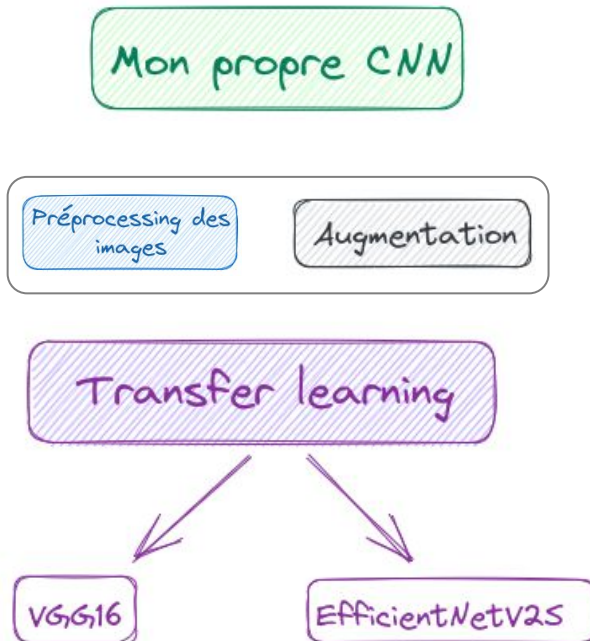
Démarche



Classification

3

Démarche



Classification

3

Préprocessing des
images

Augmentation

Classification

3

Préprocessing des
images

Augmentation

VGG16

The images are converted from RGB to BGR, then each color channel is zero-centered with respect to the ImageNet dataset, without scaling.

EfficientNetV2S

Included in the model

Classification

3

Préprocessing des
images

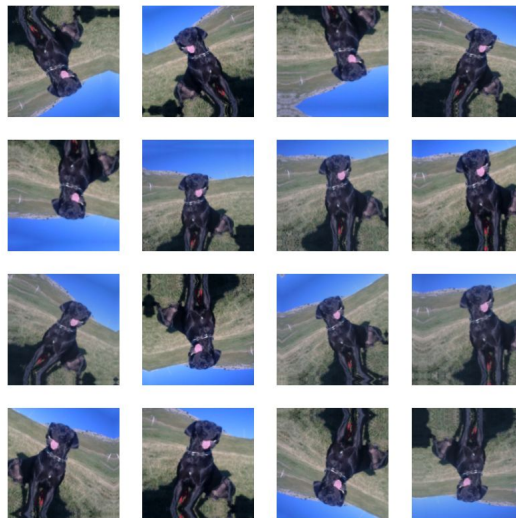
Augmentation

VGG16

The images are converted from RGB to BGR, then each color channel is zero-centered with respect to the ImageNet dataset, without scaling.

EfficientNetV2S

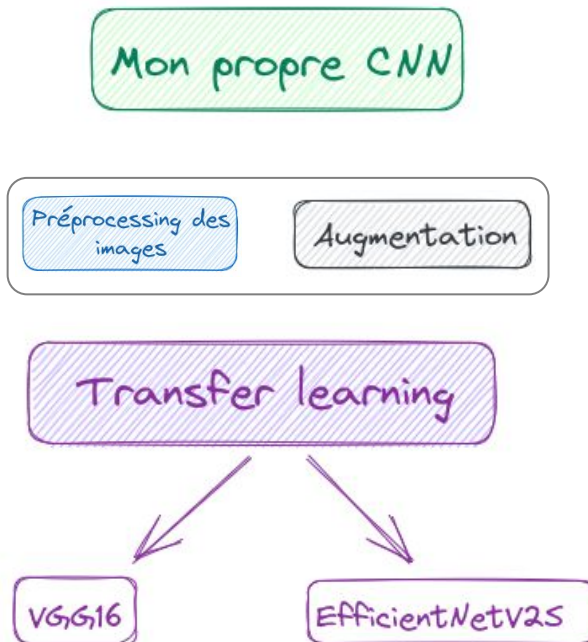
Included in the model



Classification

3

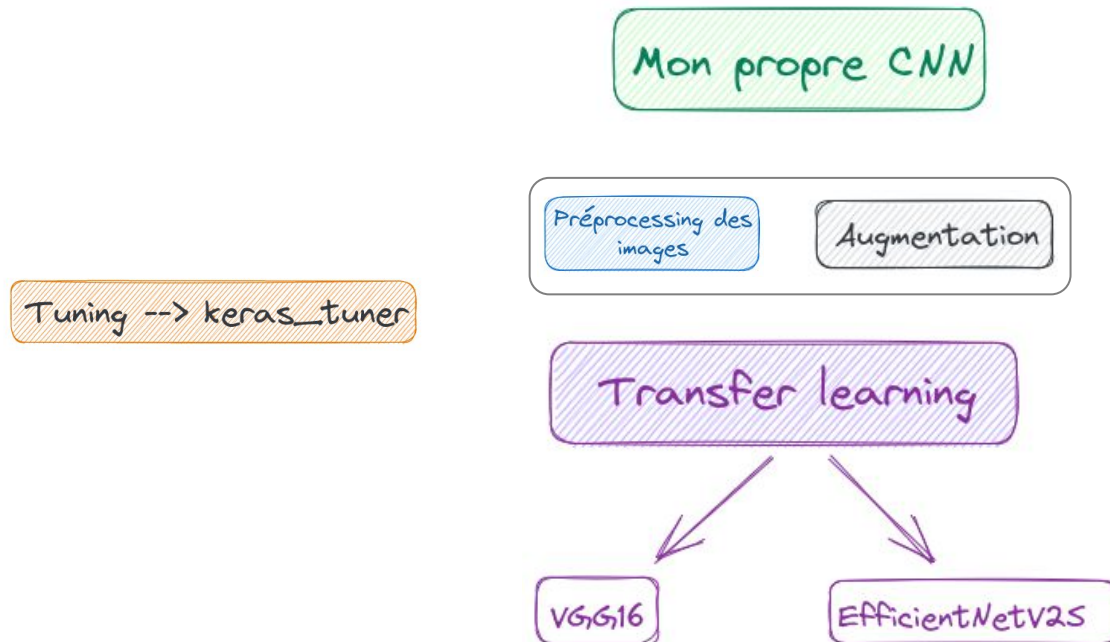
Démarche



Classification

3

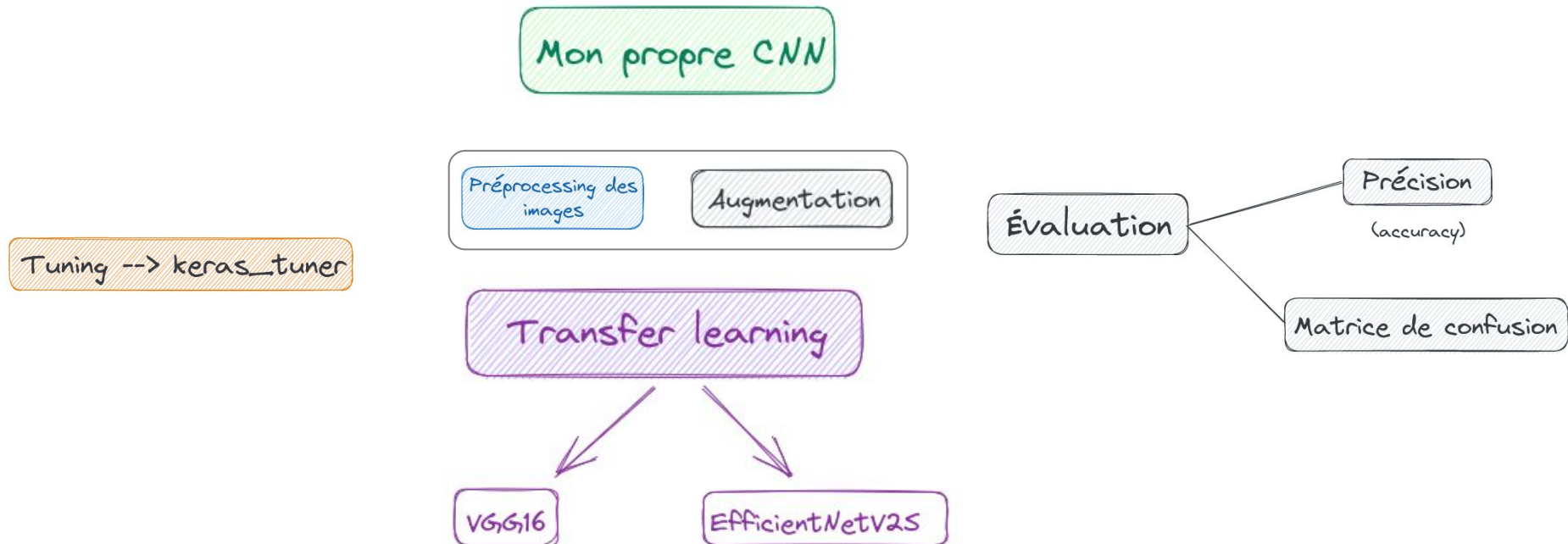
Démarche



Classification

3

Démarche



Classification

3

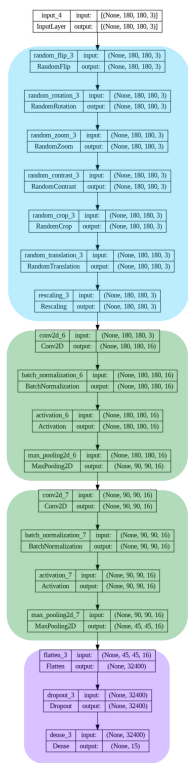
Architectures

Classification



Classification

CNN - S



augmentation

Bloc convolution:

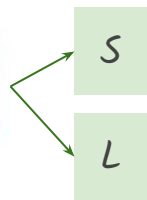
Conv2D
batch norm
ReLU
MaxPooling

bloc convolution

Bloc classification:

flatten
dropout
fully connected + softmax

Mon propre CNN



CNN - L



augmentation

bloc convolution

bloc convolution

bloc convolution

bloc convolution

bloc convolution

Bloc classification:

flatten
fully connected
dropout
fully connected
dropout
fully connected + softmax

3

Classification

3

CNN - S

params = 491 167

augmentation

Bloc convolution:

Conv2D

batch norm

ReLU

MaxPooling, *strides = (2,2)*

bloc convolution

Bloc classification:

flatten

dropout rate = *0.3*

fully connected + softmax

Mon propre CNN

S

L

CNN - L

params = 1 270 143

augmentation

K = 64
F = 10

bloc convolution

K = 64
F = 5

bloc convolution

K = 32
F = 5

bloc convolution

K = 32
F = 10

bloc convolution

K = 16
F = 10

bloc convolution

Bloc classification:

flatten

fully connected, *u = 1023*

dropout = *0.2*

fully connected, *u = 512*

dropout = *0.5*

fully connected + softmax

Classification

CNN - S

Mon propre CNN

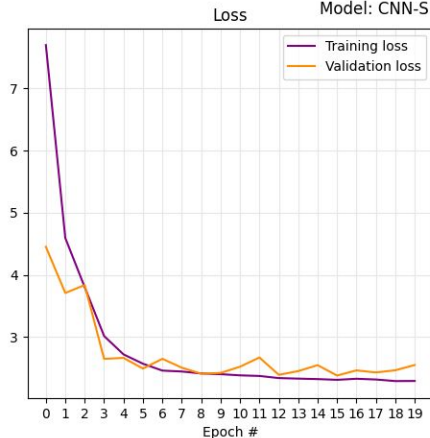
S

L

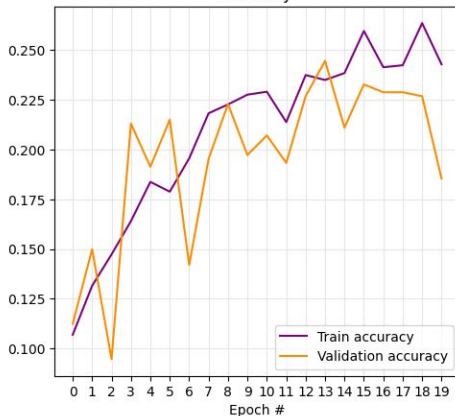
CNN - L

3

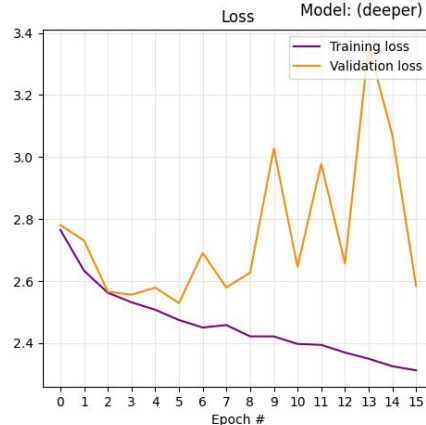
Model performance
Model: CNN-S from scratch (tuned)



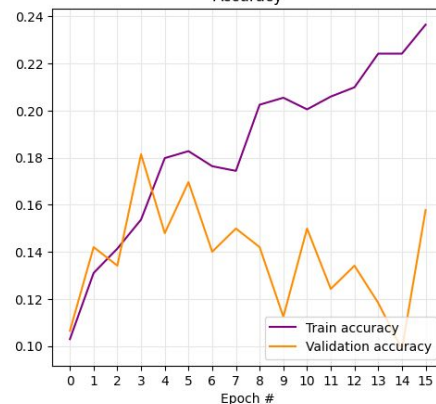
Accuracy



Model performance
Model: (deeper) CNN from scratch, tuned



Accuracy



Classification

CNN - S

Mon propre CNN

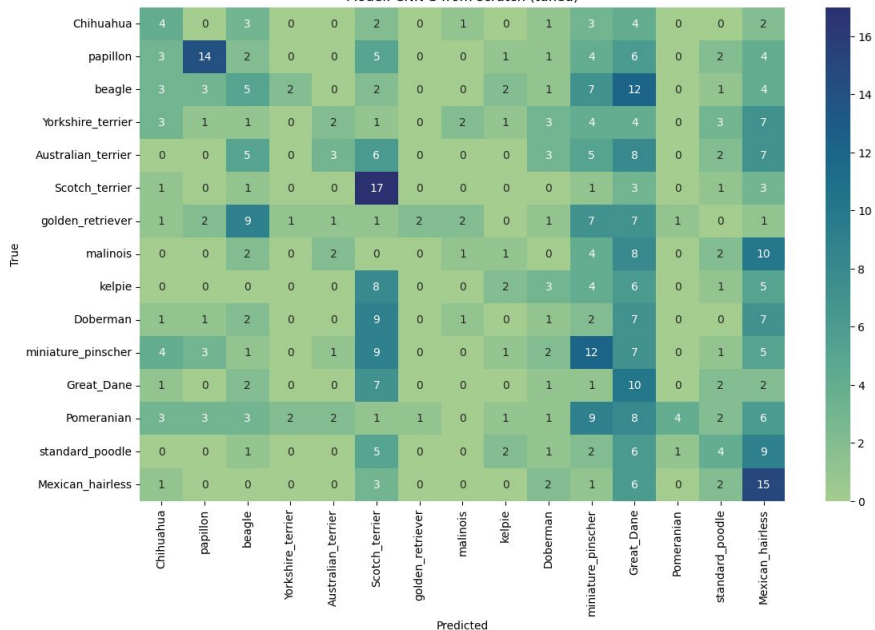
S

L

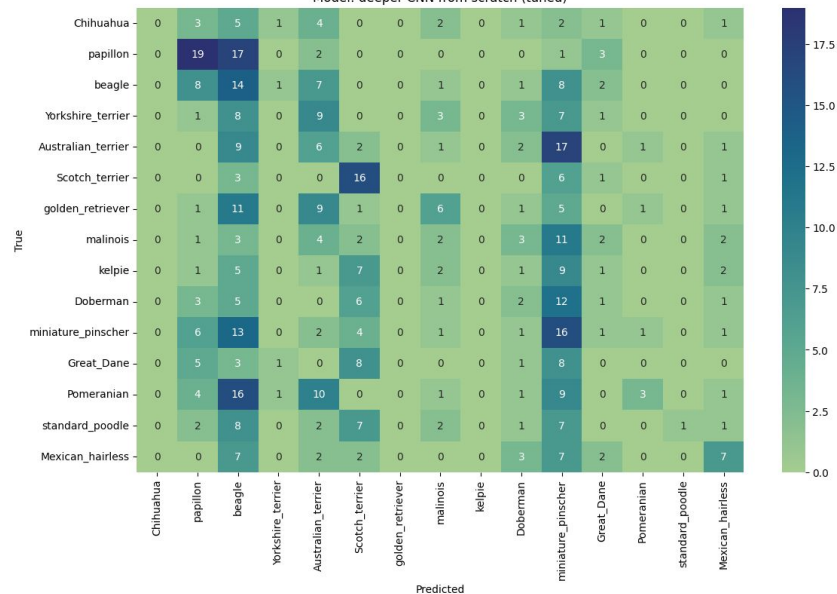
CNN - L

3

Confusion matrix: dog breed classification
Model: CNN-S from scratch (tuned)



Confusion matrix: dog breed classification
Model: deeper CNN from scratch (tuned)

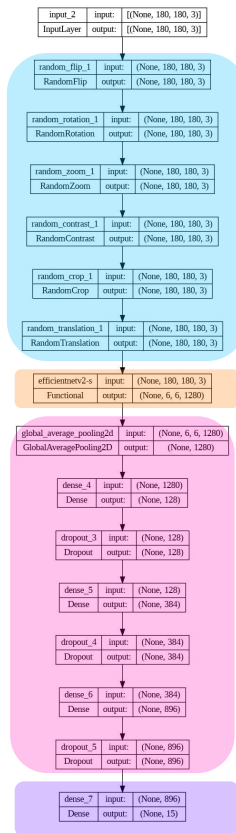


Classification

Transfer learning

3

Architecture



augmentation

VGG16 ou EfficientNetV2S

Bloc FC:

flatten or global average pooling

fully connected + ReLU

dropout

fully connected + ReLU

dropout

fully connected + ReLU

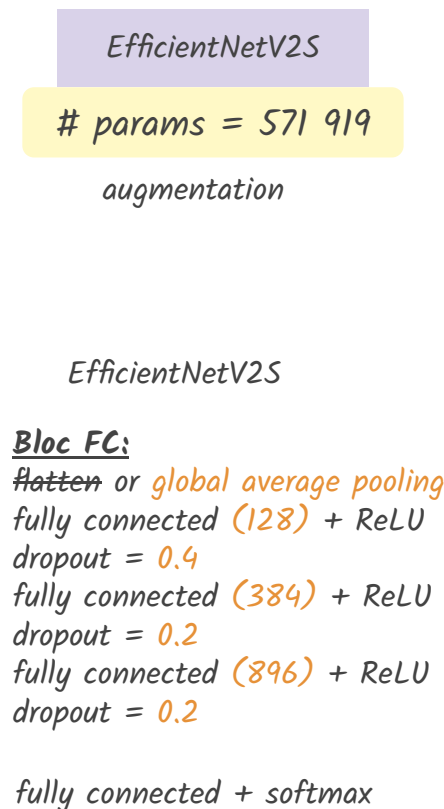
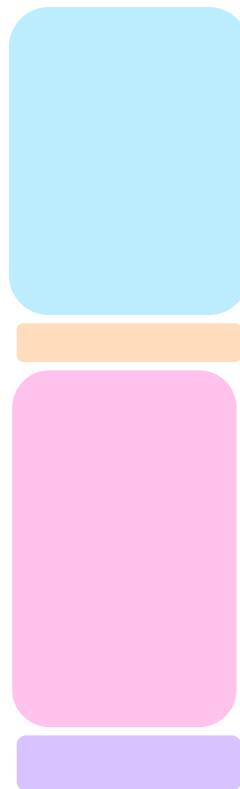
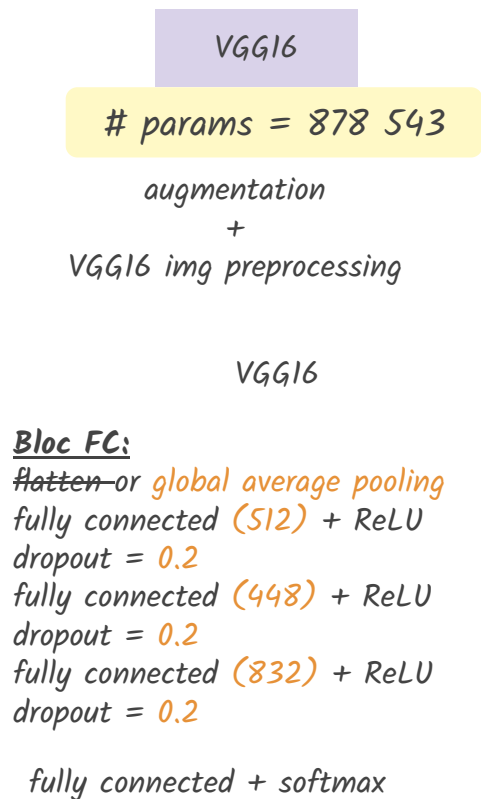
dropout

fully connected + softmax

Classification

Transfer learning

3



Classification

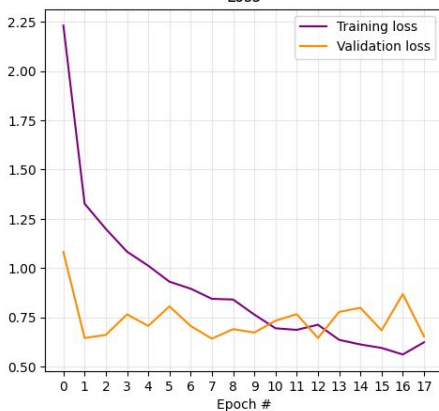
Transfer learning

3

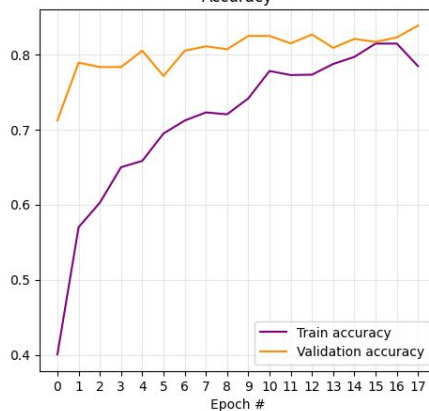
VGG16

EfficientNetV2S

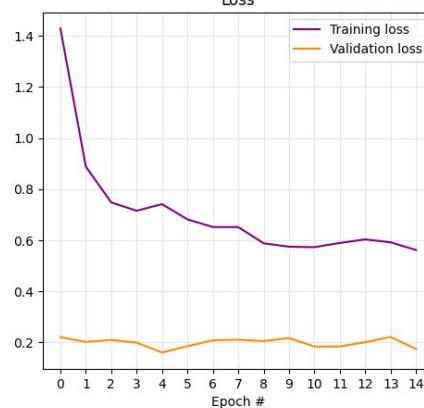
Model performance
Model: VGG16 (transfer, tuned)



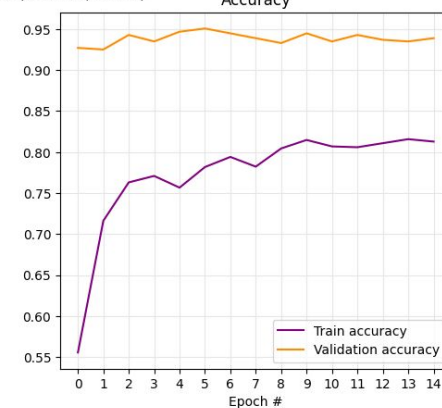
Accuracy



Model performance
Model: EfficientNet (transfer, tuned)



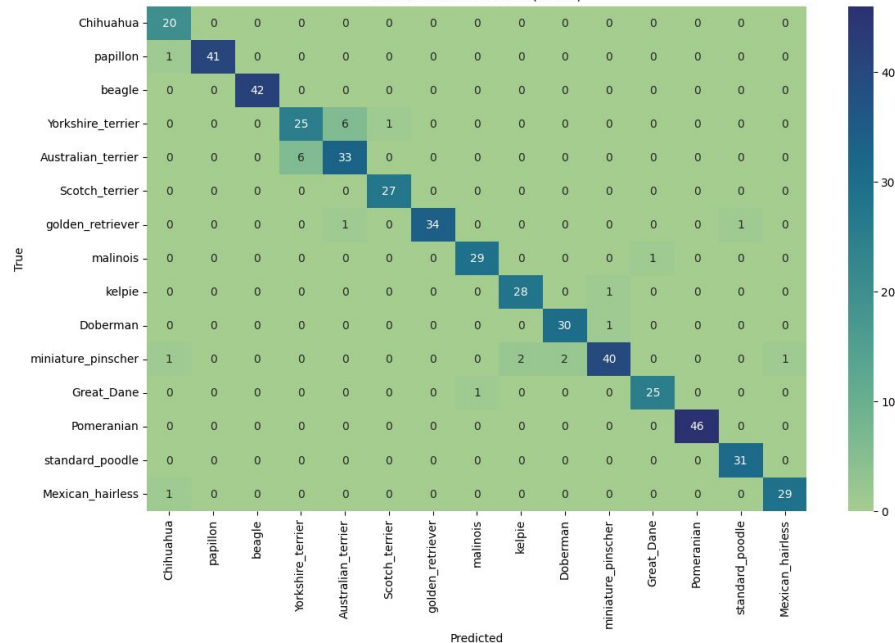
Accuracy



Transfer learning

EfficientNetV2S

Confusion matrix: dog breed classification
Model: VGG16 (tuned)



Classification

3

Modèle choisi



Transfer learning

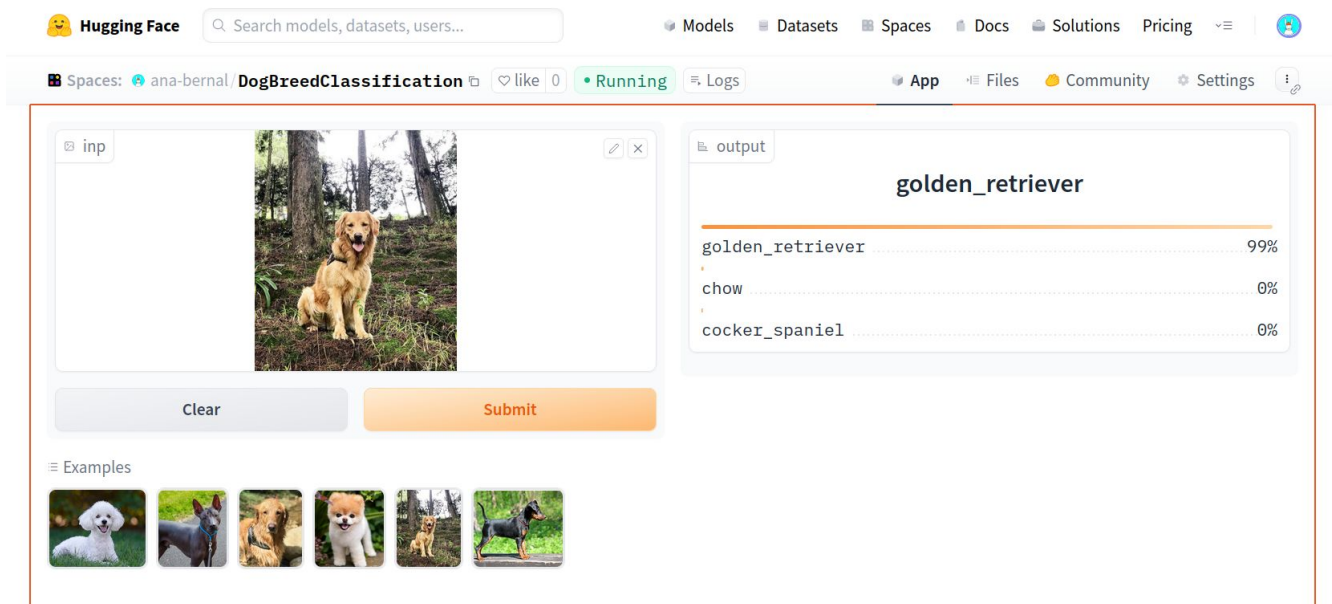
+

EfficientNetV2S

Classification: inférence

3

[Clic ici](#) * pour tester ce modèle en action



The screenshot shows the Hugging Face interface for the 'DogBreedClassification' model by user 'ana-bernal'. The interface includes a search bar, navigation links (Models, Datasets, Spaces, Docs, Solutions, Pricing), and a user profile icon. The main content area displays the model's name, a 'like' count, and a 'Running' status. Below this, there is an input section with a photo of a golden retriever and a 'Submit' button. The output section shows the classification results for the image.

Model	Score
golden_retriever	99%
chow	0%
cocker_spaniel	0%

* huggingface.co/spaces/ana-bernal/DogBreedClassification

4

Conclusions

Conclusions

- EfficientNetV2 → les meilleurs performances sur tous les modèles.
- ImageNet base des données/labels assez efficace déjà pour classification.
- Modèles à disposition aussi très efficaces déjà.

Aller + loin

- 3 datasets: train, test, validation.
- Entraîner tous les paramètres de VGG16 et EfficientNet, ou plus de couches → mieux comprendre les performances avec nos données.
- Appliquer les modèles tels quels avec une autre base des données.
- Forcer le # et tailles des filtres pour modèles from scratch et tuner autres paramètres.
- + d'Epochs.

Merci !