IMAGE AUTO-ORIENTATION

ADVANCED MACHINE LEARNING – A.A. 2020/2021

MATTEO CAMPIRONI - 801850
SERENA DI MAGGIO - 821063

PROBLEM STATEMENT

Correctly rotating an image is a simple task for a human being, but it turns out to be a real challenge for a computer.

Most applications in the world of computer vision require images to be properly oriented before they can be processed.

THE GOAL

However rotating images is a tedious, time-consuming and error-prone activity.

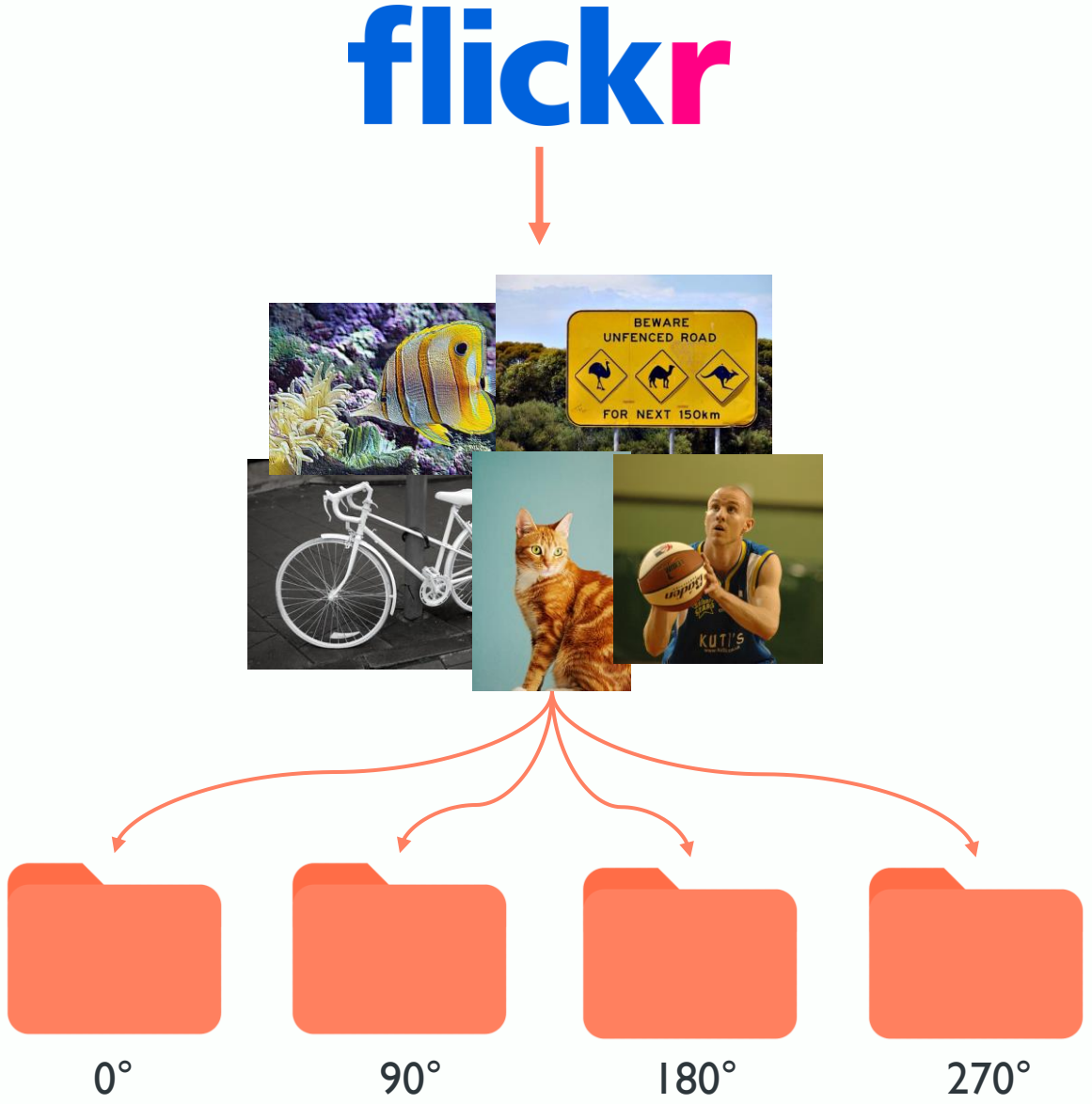
The goal is therefore to exploit convolutional neural networks (CNNs) to automatically orient the images by multiple angles of 90°.

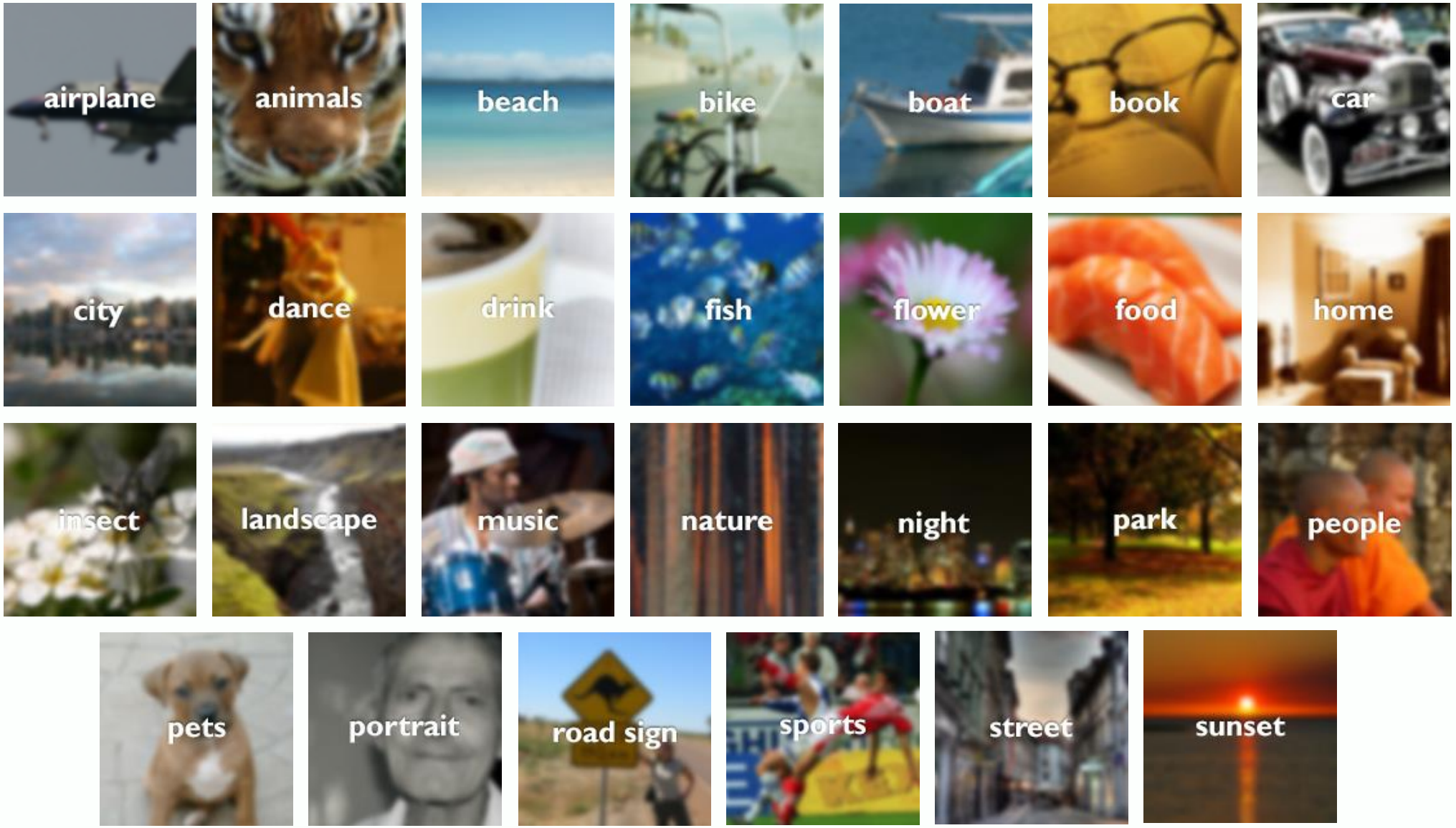
THE DATASET

The dataset is built from scratch using the APIs provided by Flickr.

27 of the most popular tags of all time and of the week were selected in order to obtain a generic dataset.

The images were artificially rotated in order to obtain a balanced dataset, consisting of 151,640 photographs.




airplane
animals
beach
bike
boat
book
car
city
dance
drink
fish
flower
food
home
insect
landscape
music
nature
night
park
people
pets
portrait
road sign
sports
street
sunset

ADDITIONAL DATASETS

Other famous datasets were selected in order to test the performance of the models on images that are different from those that can be found on flickr.

SUN2012

PASCAL VOC 2012

INRIA HOLIDAYS

IMAGE PADDING

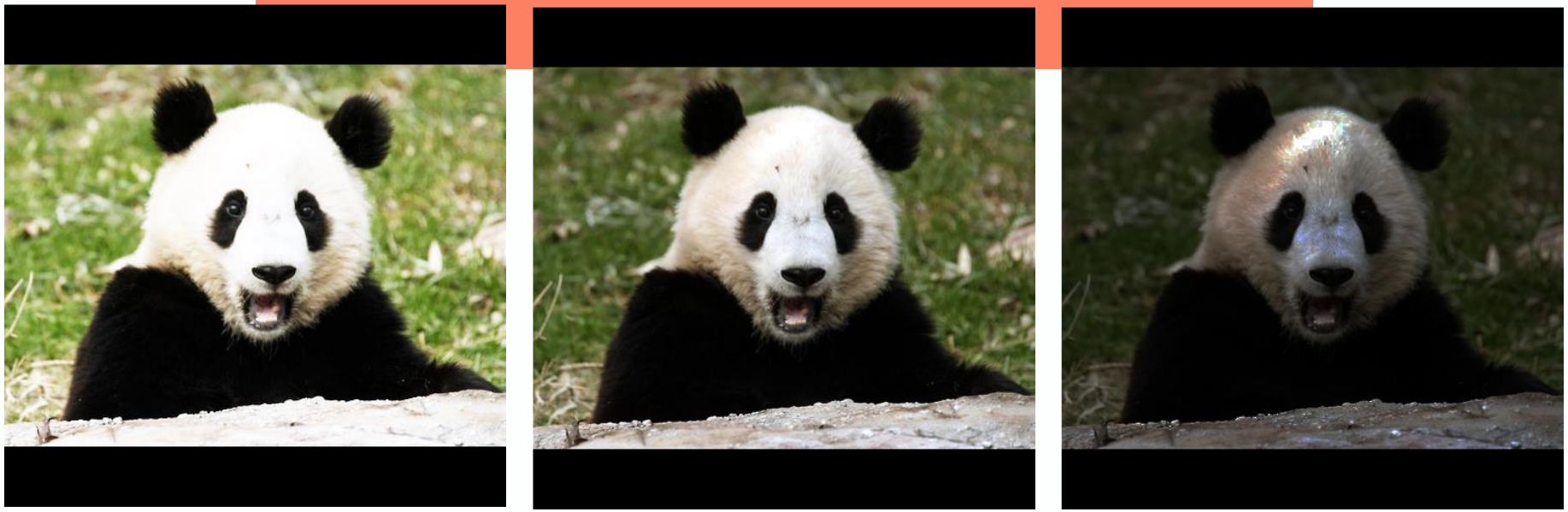
The dataset is composed of images of various resolutions, while the CNNs used for the task work with inputs of 224x224 pixels.

For this reason, it has been decided to pad the pictures with black pixel, in order for the output to be a square.

DATA AUGMENTATION

Using *ImageDataGenerator* from Keras, random transformations to input samples were applied during network training.

brightness_range and *channel_shift_range* have been chosen as parameters.



OVERVIEW

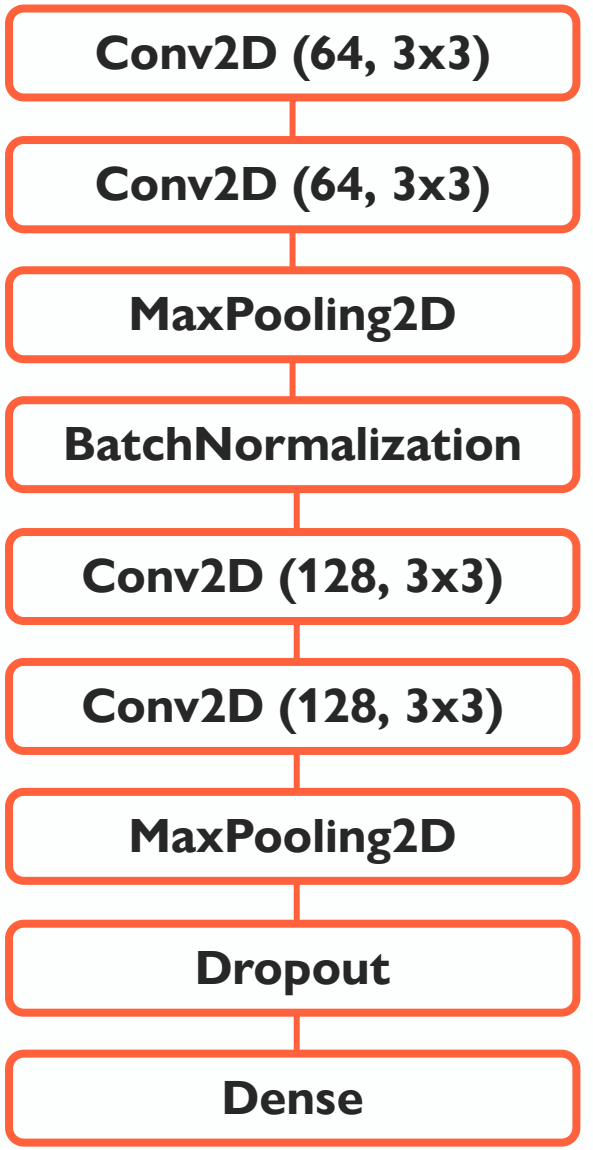

CUSTOM MODEL
VGG16
MOBILENETV2
DENSENET201
TRANSFER LEARNING

CUSTOM MODEL

The first model proposed for image orientation detection has been built from scratch.

The idea was to develop a simple neural network, with a restricted number of layers.

The results were good, but not completely satisfying.



train accuracy
73.53%

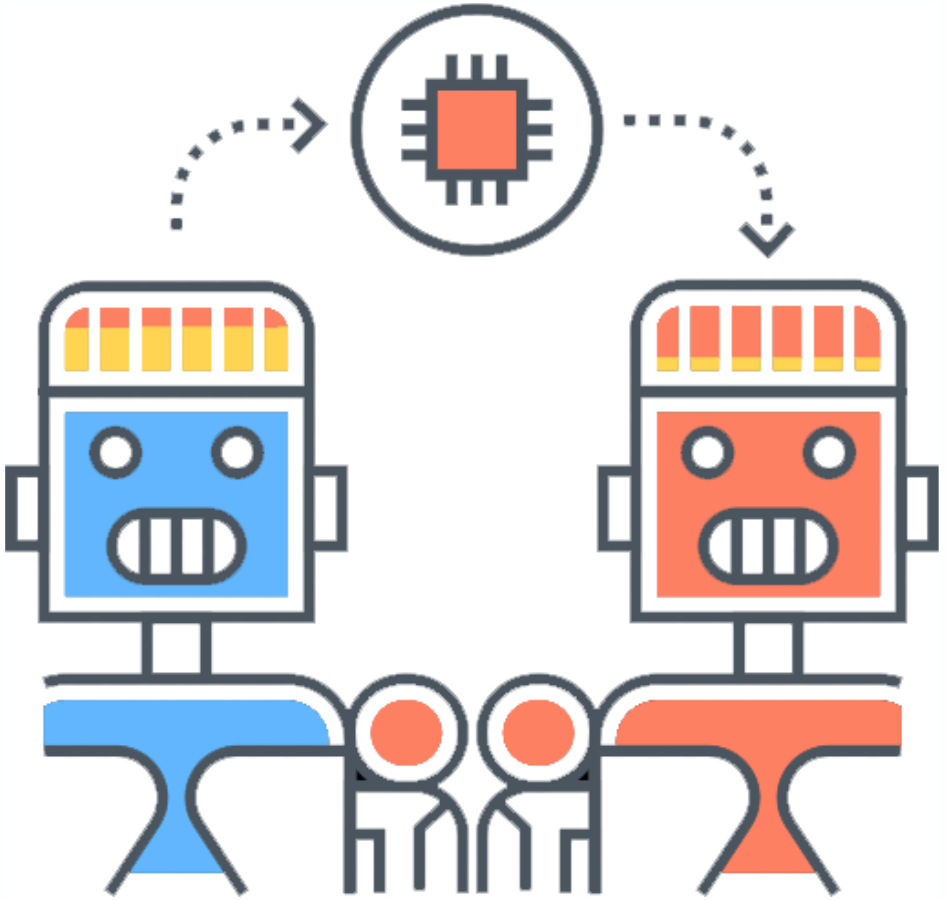
val. accuracy
73.25%

TRANSFER LEARNING

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task.

Three models were selected:

- VGG16
- MobileNetV2
- DenseNet201

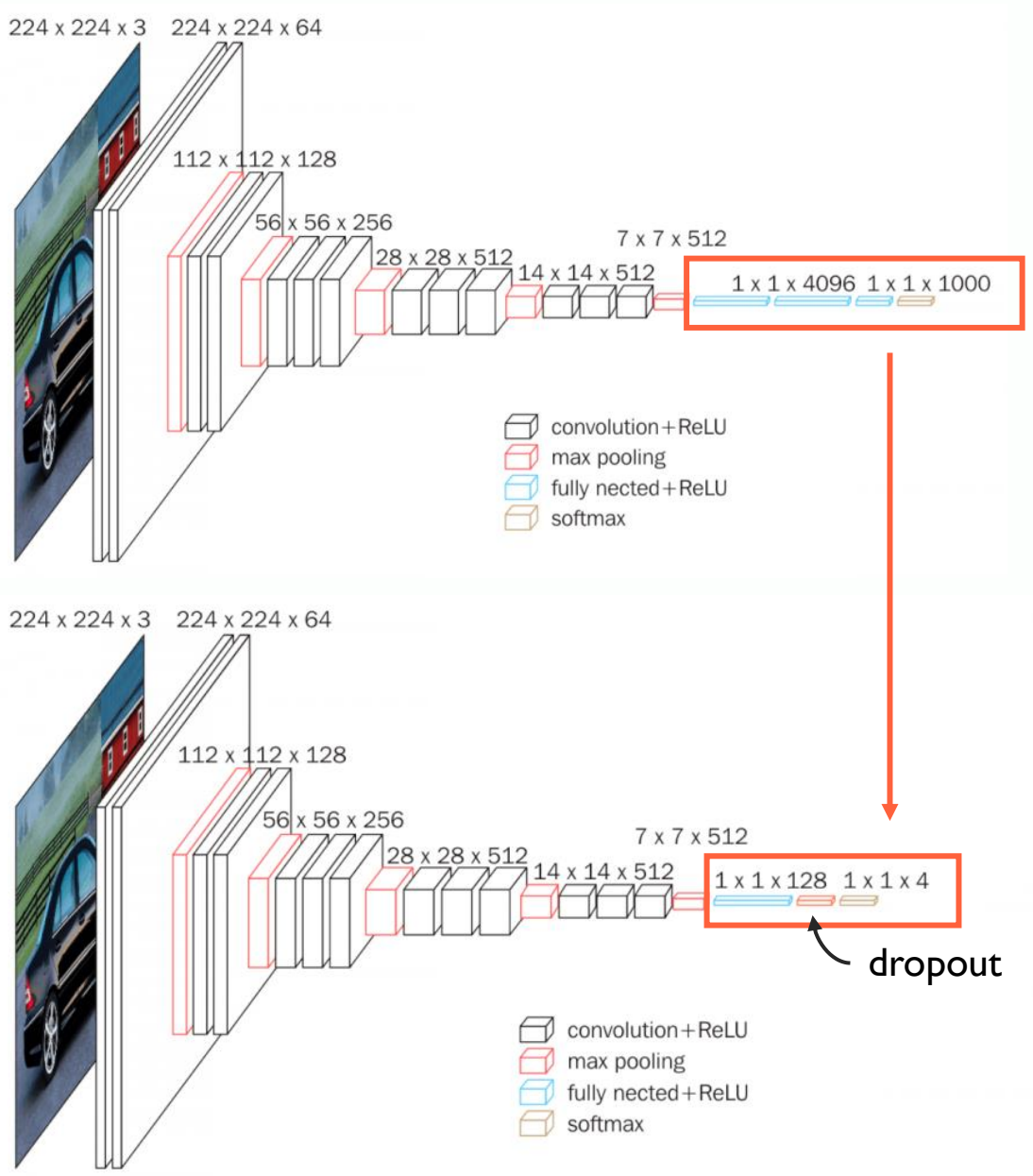


VGG16

Only the last part of the architecture has been replaced by the following layers:

- dense (128, ReLU)
- dropout (0.7)
- dense (4, Softmax)

train accuracy	val. accuracy
93.42%	92.43%

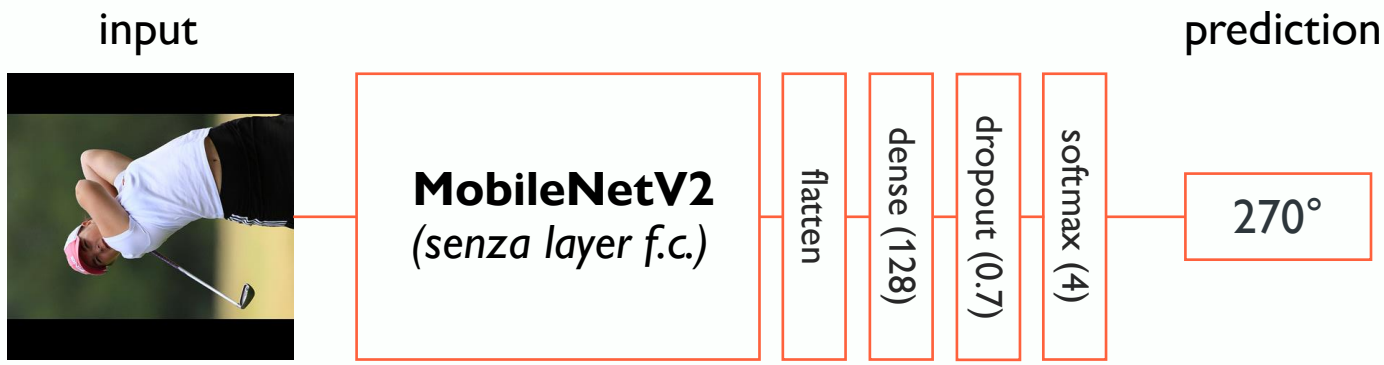


MOBILENETV2

Just like before, only the last part of the architecture has been replaced by the following layers:

- dense (128, ReLU)
- dropout (0.7)
- dense (4, Softmax)

train accuracy	val. accuracy
97.56%	93.32%

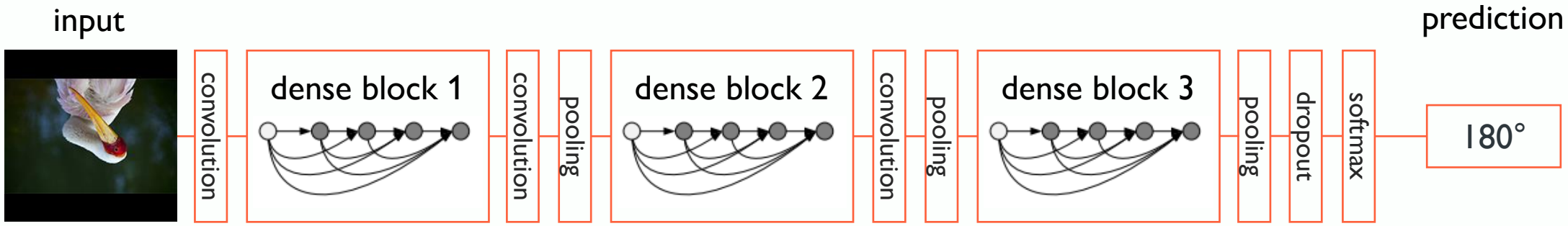


DENSENET201

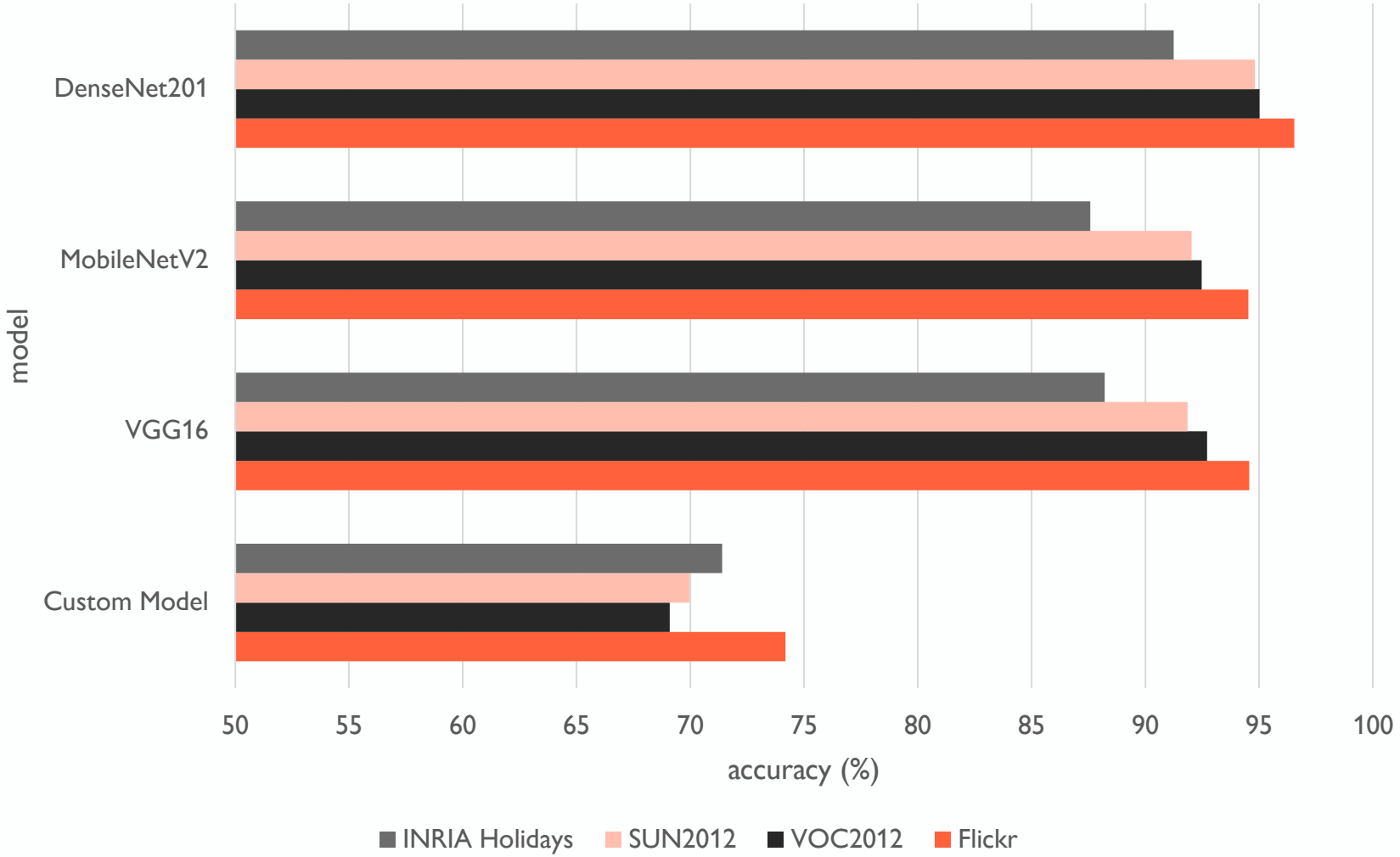
The network structure was left unchanged.

A dropout layer has been added and the last one has been replaced with a softmax.

train accuracy	val. accuracy
96.64%	95.02%

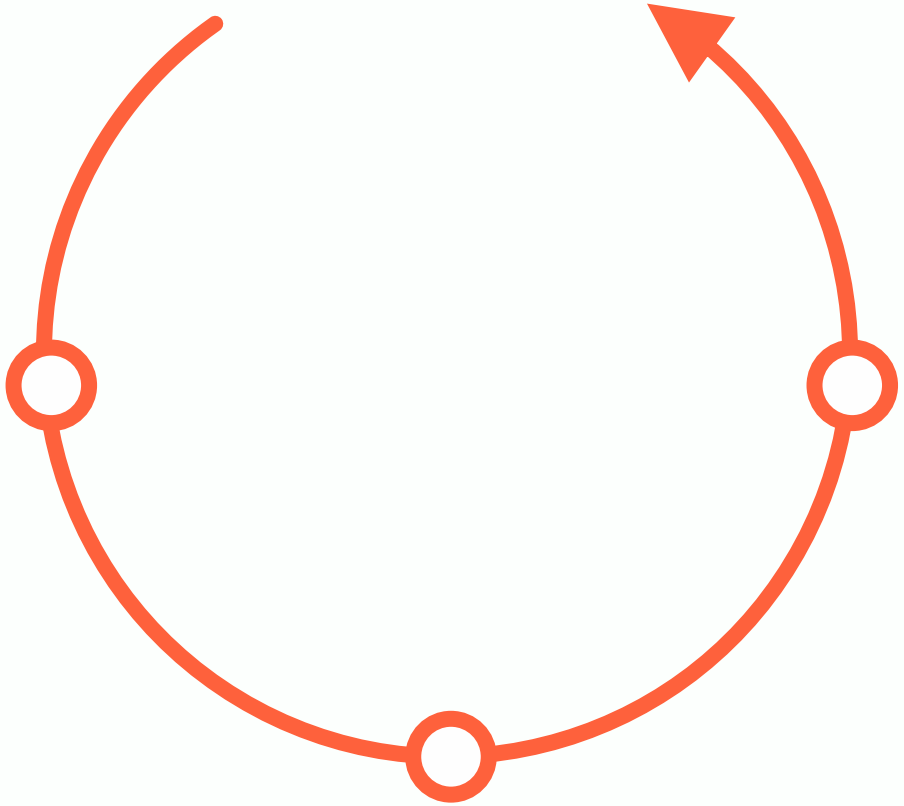


PERFORMANCE ON MULTIPLE TEST SET



FUTURE IMPROVEMENTS

Bigger dataset with different categories



Model optimization

External libraries for data augmentation

THANK YOU FOR YOUR ATTENTION