



IMAGE AUTO-ORIENTATION

ADVANCED MACHINE LEARNING – A.A. 2020/2021

MATTEO CAMPIRONI - 801850

SERENA DI MAGGIO - 821063

IL PROBLEMA

Ruotare correttamente un'immagine è un compito semplice per un essere umano, ma che si rivela una vera e propria sfida per un computer.

La maggior parte delle applicazioni nel mondo della computer vision richiede che prima di essere processate queste abbiano un corretto orientamento.



OBIETTIVO

Ruotare le immagini è però un lavoro tedioso, che richiede tempo e ad alto rischio di errore.

L'obiettivo è quindi sfruttare le reti neurali convoluzionali (CNN) per orientare automaticamente le immagini di angoli multipli di 90° .

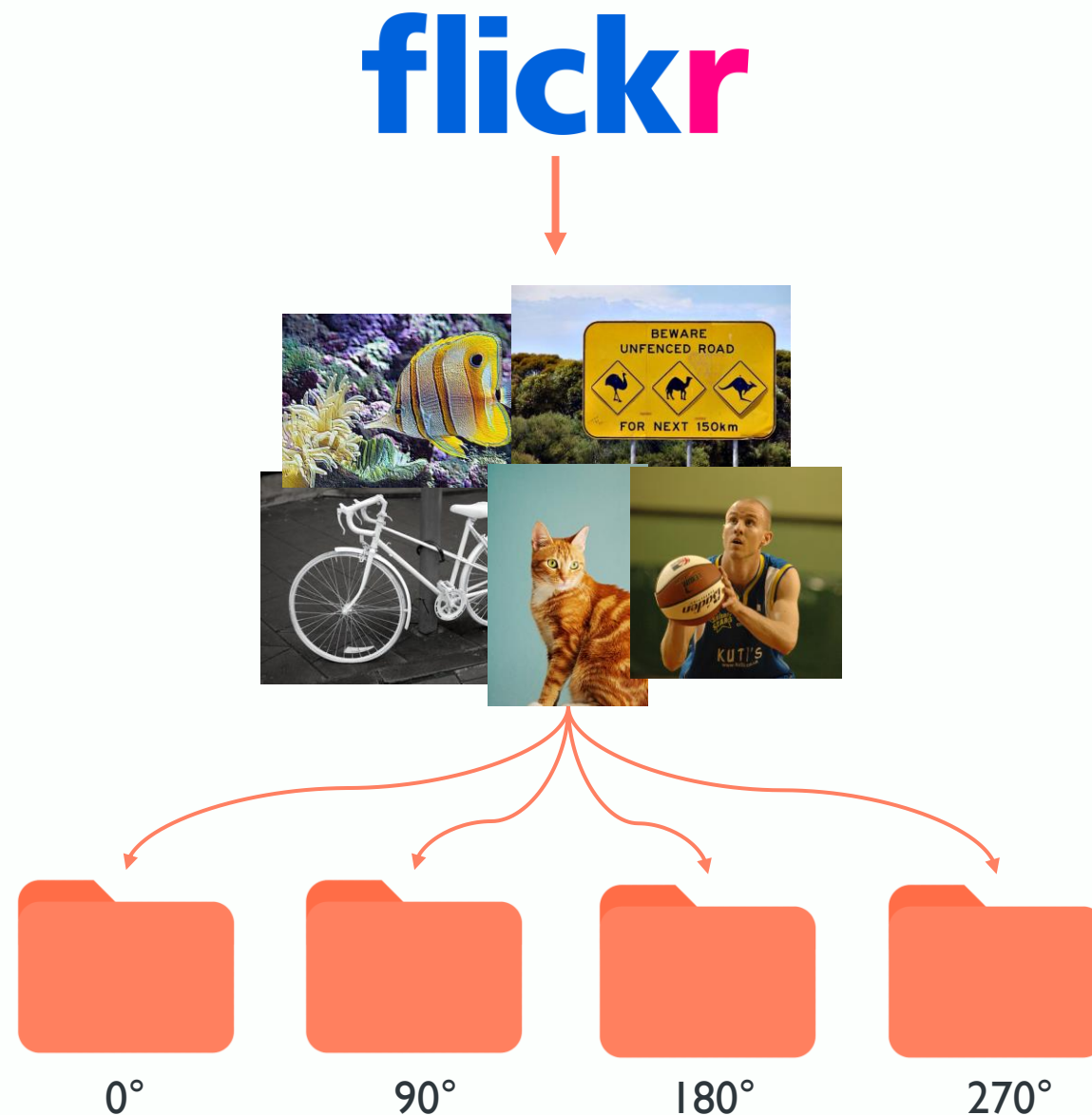


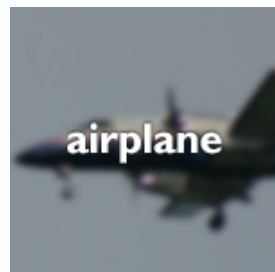
IL DATASET

Il dataset è stato costruito da zero utilizzando le API fornite da Flickr.

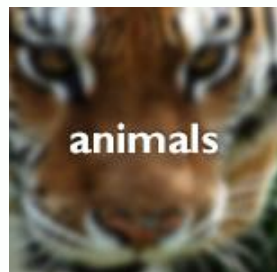
Sono stati selezionati **27 tag** tra i più popolari di sempre e della settimana in modo da ottenere un dataset generico.

Le immagini sono state ruotate artificialmente in modo da ottenere un dataset equilibrato composto da 151.640 fotografie.





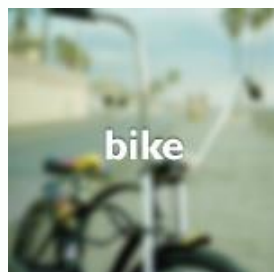
airplane



animals



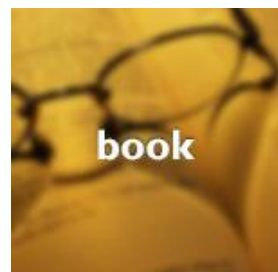
beach



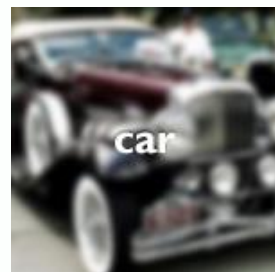
bike



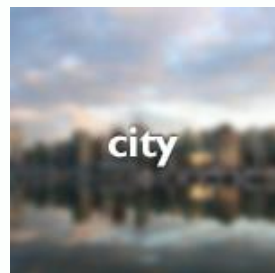
boat



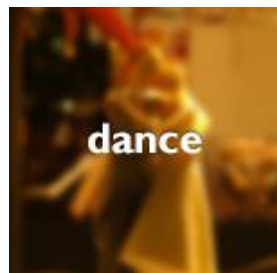
book



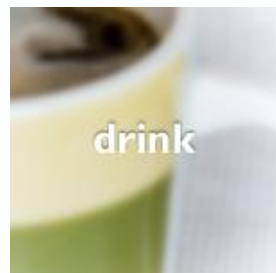
car



city



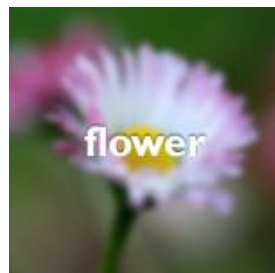
dance



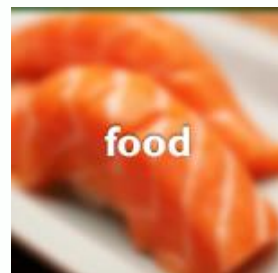
drink



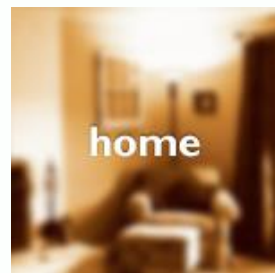
fish



flower



food



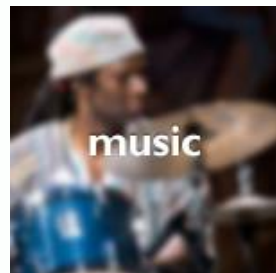
home



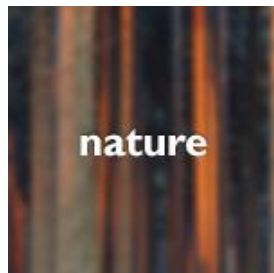
insect



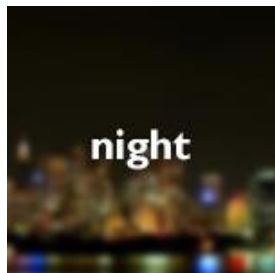
landscape



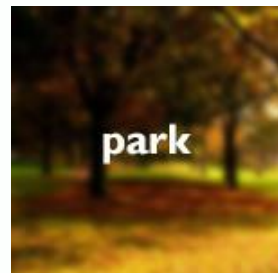
music



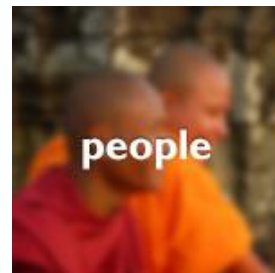
nature



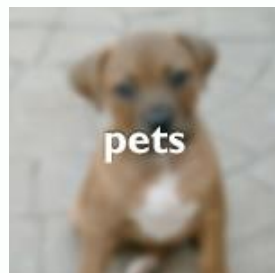
night



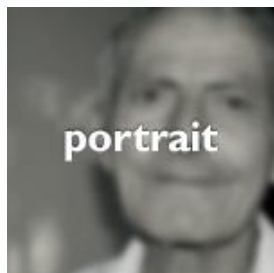
park



people



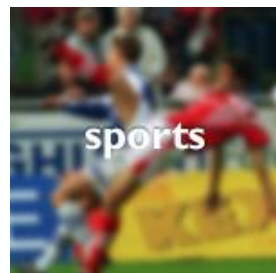
pets



portrait



road sign



sports



street



sunset

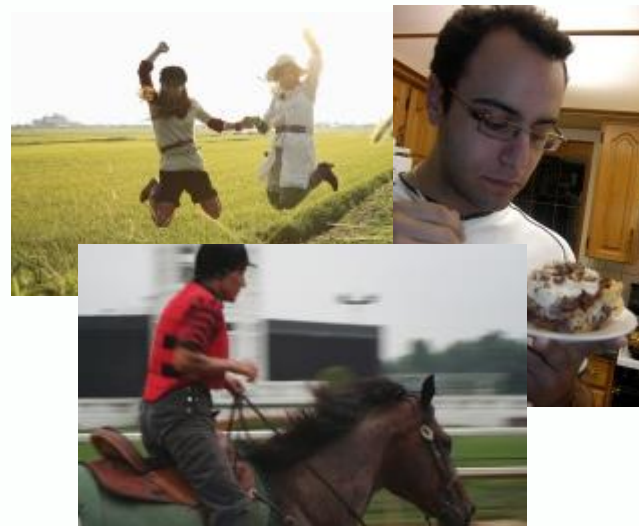
ULTERIORI DATASET

Per testare le performance dei modelli su immagini diverse da quelle che si possono trovare su flickr sono stati selezionati altri famosi dataset.

SUN2012



PASCAL VOC 2012



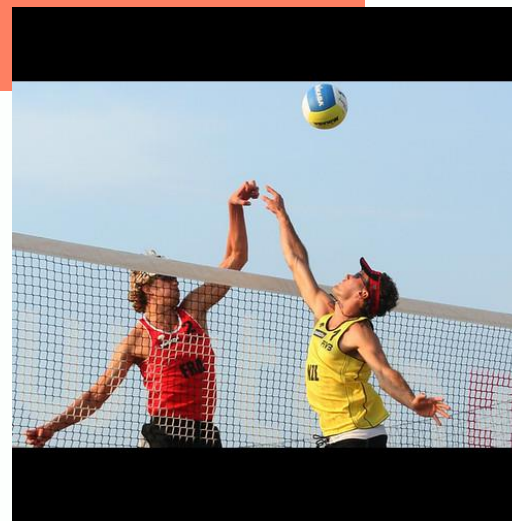
INRIA HOLIDAYS



IMAGE PADDING

Il dataset è composto da immagini di varie risoluzioni mentre le CNN utilizzate lavorano con input di 224x224 pixel.

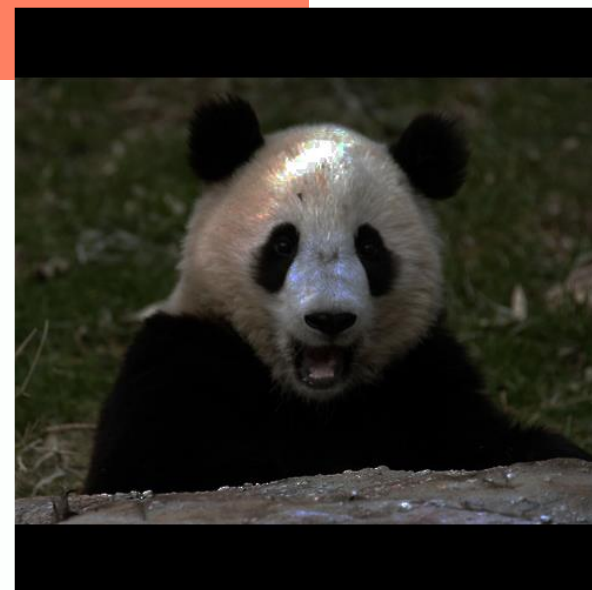
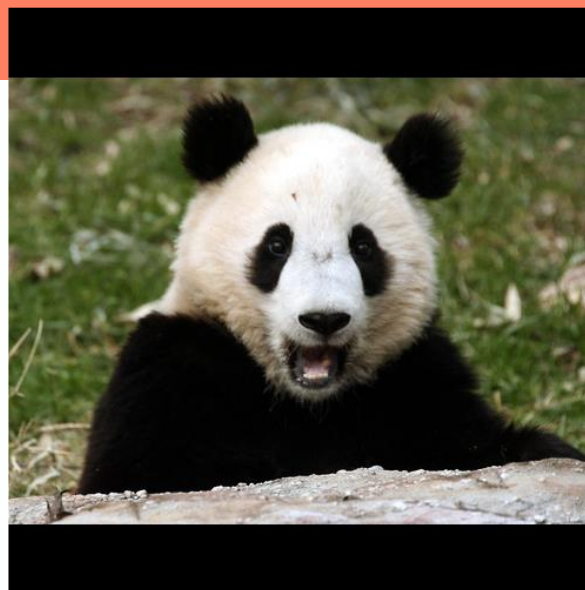
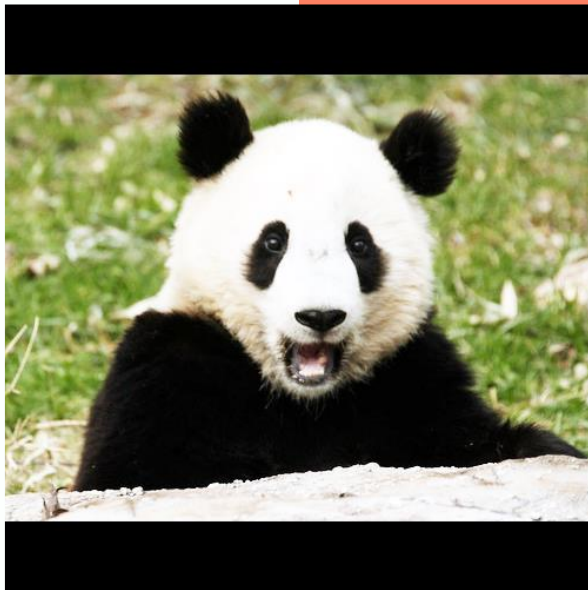
Per questo motivo si è scelto di effettuare un padding delle immagini con pixel neri, in modo che l'output fosse quadrato.



DATA AUGMENTATION

Sfruttando *ImageDataGenerator* in Keras è stata utilizzata la tecnica della data augmentation in tempo reale durante il training.

Sono stati scelti *brightness_range* e *channel_shift_range* come parametri.



OVERVIEW

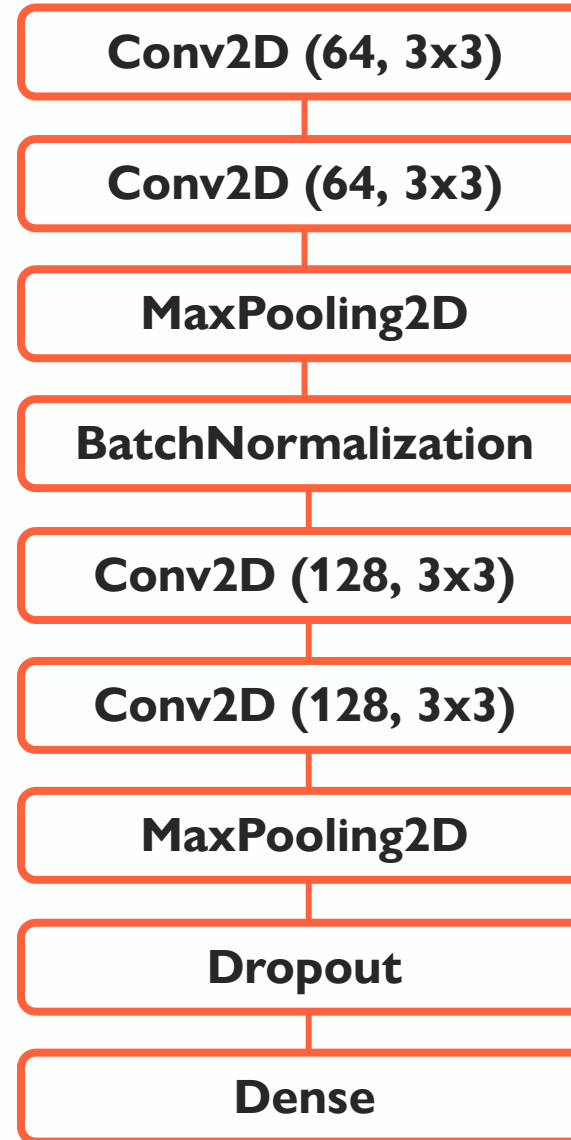


CUSTOM MODEL

Il primo modello proposto è stato costruito da zero.

L'idea era quella di provare a sviluppare una rete neurale con un numero ristretto di layer.

I risultati ottenuti si sono rivelati buoni, ma non soddisfacenti.



train accuracy

73.53%

val. accuracy

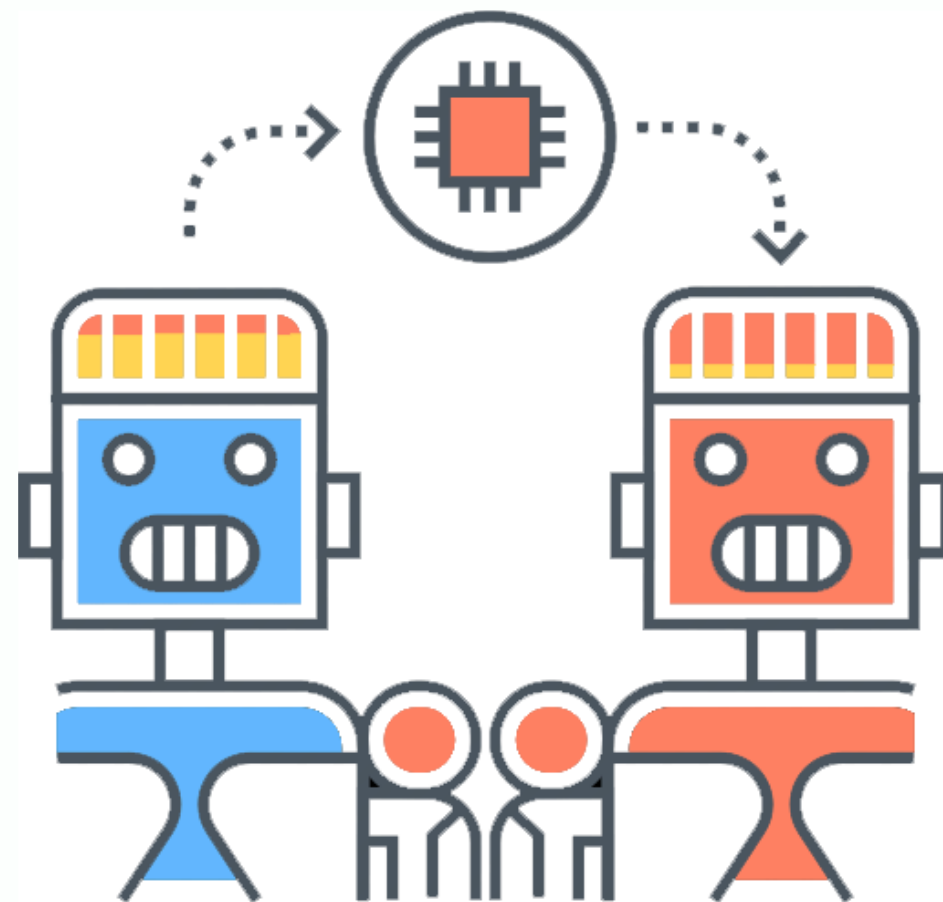
73.25%

TRANSFER LEARNING

Tecnica di Machine Learning in cui un modello allenato su un task è sfruttato come punto di partenza per un secondo task.

Sono stati scelti 3 modelli da cui partire:

- VGG16
- MobileNetV2
- DenseNet201



VGG16

È stata sostituita l'ultima parte dell'architettura con i seguenti layer:

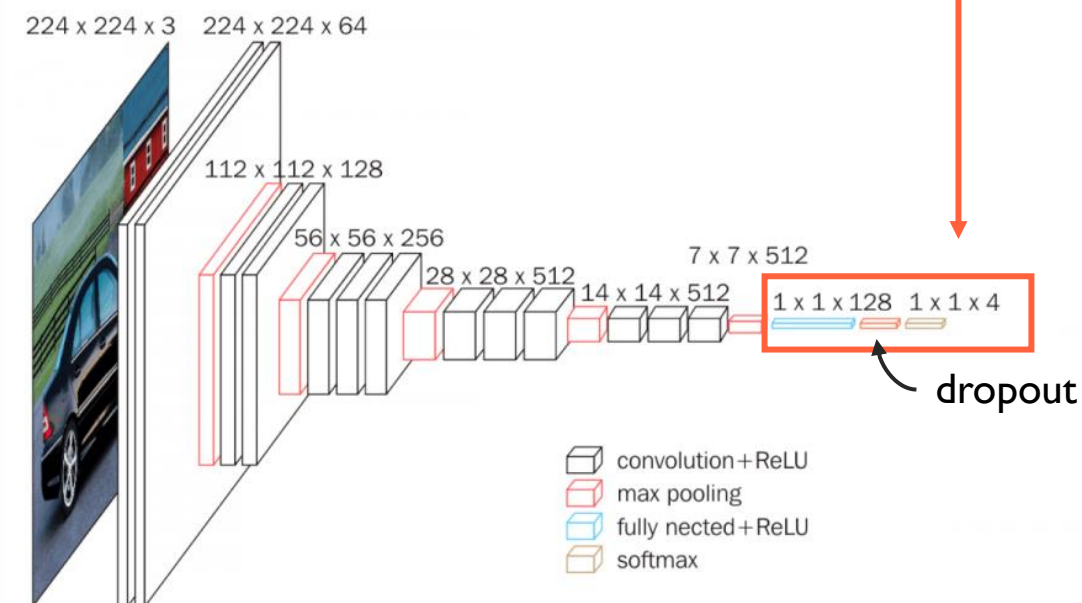
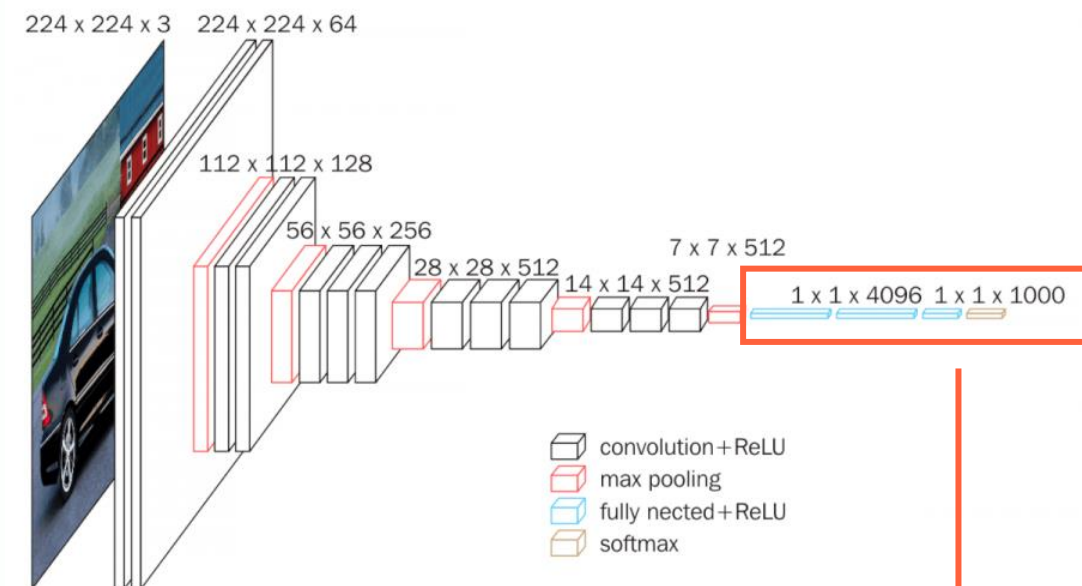
- dense (128, ReLU)
- dropout (0.7)
- dense (4, Softmax)

train accuracy

93.42%

val. accuracy

92.43%



MOBILENETV2

Come prima, è stata sostituita l'ultima parte dell'architettura con i seguenti layer:

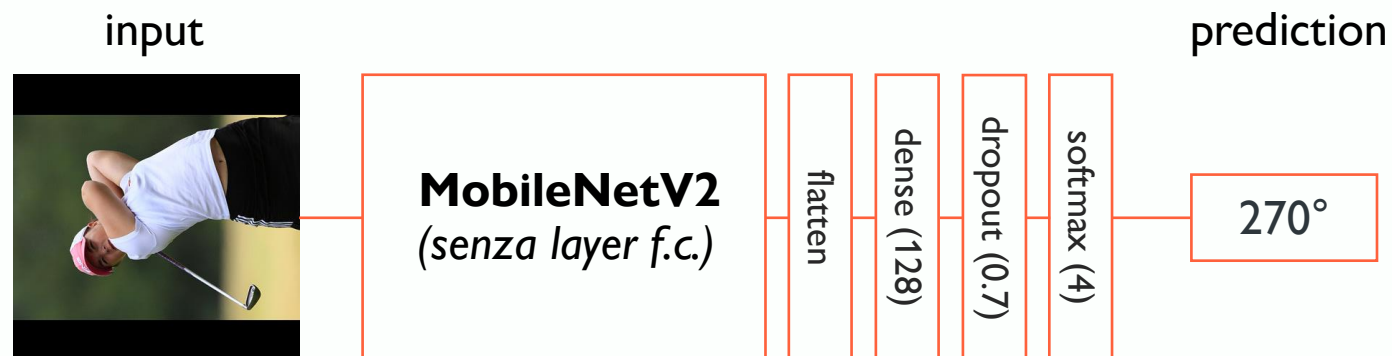
- dense (128, ReLU)
- dropout (0.7)
- dense (4, Softmax)

train accuracy

97.56%

val. accuracy

93.32%



DENSENET201

La struttura è rimasta invariata.

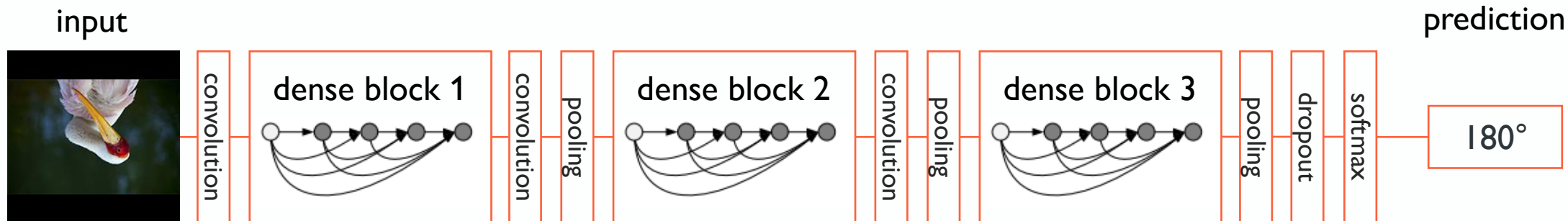
È stato aggiunto un layer di dropout e l'ultimo è stato sostituito da un softmax.

train accuracy

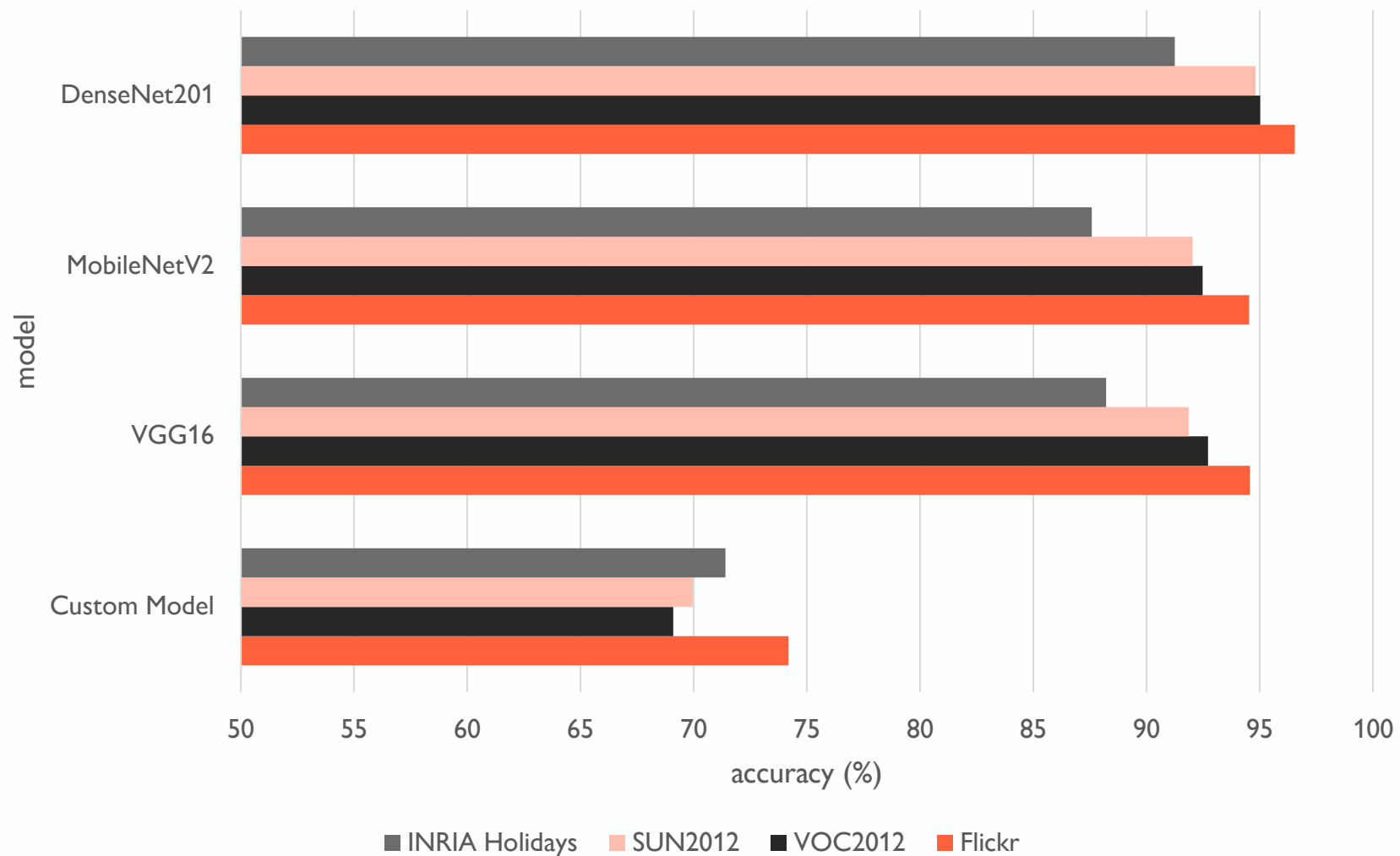
96.64%

val. accuracy

95.02%



PERFORMANCE SUI TEST SET

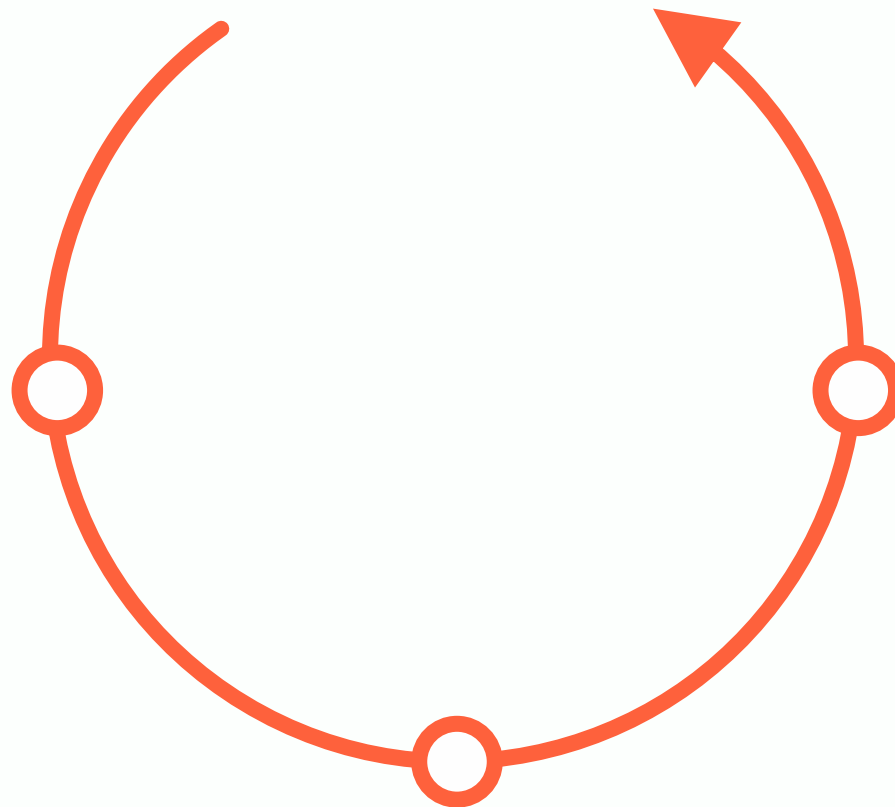


SVILUPPI FUTURI

Dataset più grande e
con maggiore varietà
di immagini

Ottimizzazione dei
modelli

Librerie esterne per
la data augmentation



GRAZIE PER L'ATTENZIONE