

Assignment 3

Belotti Federico - 808708
In team con Scarpellini Gianluca - 807541

24 novembre 2019

1 Obiettivi

L'obiettivo di questo assignment consiste nell'applicare il metodo del *transfer learning* utilizzando una Convolutional Neural Network (CNN) preaddestrata su ImageNet. Le architetture scelte sono VGG16 [4] e ResNet50 [1], mentre il classificatore classico adottato è un SVM lineare. Il dataset utilizzato per testare le features estratte dalle CNN è il sottoinsieme formato dalle classi "dog", "frog", "horse", "ship" e "truck" del CIFAR-10 [2].

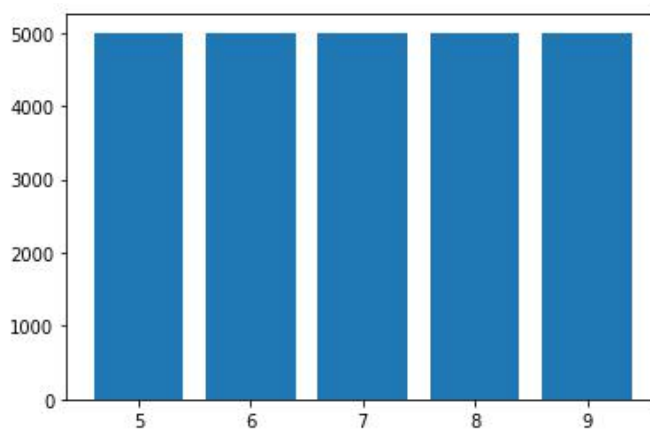


Figura 1: Distribuzione dataset

2 Dataset e preprocessing

Il dataset a disposizione è composto da 25000 immagini, la cui distribuzione è rappresentata in figura 1.

Il dataset è ben bilanciato e le trasformazioni sono applicate in automatico dal metodo `keras.application.{model}.preprocess_input`, dunque dipendenti

dal modello scelto. Dal training set è stata creata randomicamente la parte di validation, composta del 20% dell'insieme di training.

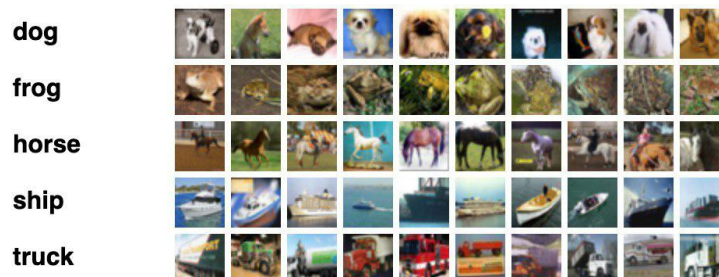


Figura 2: CIFAR-5

3 Dettagli implementativi e risultati

E' stato valutato il transfer learning di due diversi tipi di architetture: VGG16 e ResNet50, alle quali sono stati effettuati dei "tagli" a differenti profondità e ognuno dei sottomodelli ottenuti (uno per ogni taglio) è stato successivamente utilizzato per estrarre un'insieme di features; quest'ultime verranno poi classificate da una classica SVM lineare. I parametri della SVM sono:

- **Kernel:** lineare
- **Penalty:** norma L_2
- **C:** 1
- **Loss:** squared hinge
- **Iterazioni massime:** 1000

La lista completa degli esperimenti è consultabile all'indirizzo https://docs.google.com/document/d/1B8twwLQ0uFo1ZQ_RooeZ_6oudU3WcJUKiScGBLHlvw/edit?usp=sharing.

3.1 VGG16

La VGG16 [4] è un'architettura che adotta banchi di filtri convoluzionali di dimensione 3x3 e impiega un numero maggiore di livelli rispetto alla sua predecessora AlexNet [3]. I risultati ottenuti, in termini di media pesata tra le 5 classi, sono:

Taglio	Split	Precision	Recall	Accuracy	F1
block3_conv1	Validation	0.81	0.81	0.81	0.81
block3_conv1	Test	0.81	0.81	0.81	0.81
block4_conv1	Validation	0.84	0.84	0.84	0.84
block4_conv1	Test	0.84	0.84	0.84	0.84
block5_conv1	Validation	0.82	0.82	0.82	0.82
block5_conv1	Test	0.84	0.84	0.84	0.84

3.2 ResNet50

La peculiarità dell'architettura ResNet è l'utilizzo delle cosiddette *skip-connections*, che sommano l'input di un layer con la sua successiva elaborazione: così facendo vengono a crearsi delle "scorciatoie" che permettono ai gradienti di raggiungere facilmente i layer più vicini all'input (ovvero i più distanti in fase di backpropagation), permettendo quindi di creare architetture più profonde (dunque più rappresentative) [1]. I risultati ottenuti, in termini di media pesata tra le 5 classi, sono:

Taglio	Split	Precision	Recall	Accuracy	F1
conv2_block3	Validation	0.91	0.91	0.91	0.91
conv2_block3	Test	0.90	0.90	0.90	0.90
conv3_block4	Validation	0.92	0.92	0.92	0.92
conv3_block4	Test	0.92	0.92	0.92	0.92
conv4_block6	Validation	0.87	0.87	0.87	0.87
conv4_block6	Test	0.88	0.88	0.88	0.88
conv5_block3	Validation	0.80	0.80	0.80	0.80
conv5_block3	Test	0.81	0.81	0.81	0.81

Riferimenti bibliografici

- [1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.
- [2] Alex Krizhevsky. Learning multiple layers of features from tiny images. Technical report, 2009.
- [3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, pages 1097–1105, USA, 2012. Curran Associates Inc.
- [4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2014.