Fruits and Vegetables Images Classification

Carlo Mugno carlo.mugno@outlook.it
Claudio Ragaglia claudioragaglia@gmail.com
Alessio Migliorino migliorinoalessio@gmail.com

1. Model Description

We describe here the **Neural Network** we have built to classify images from a fruits and vegetables dataset.

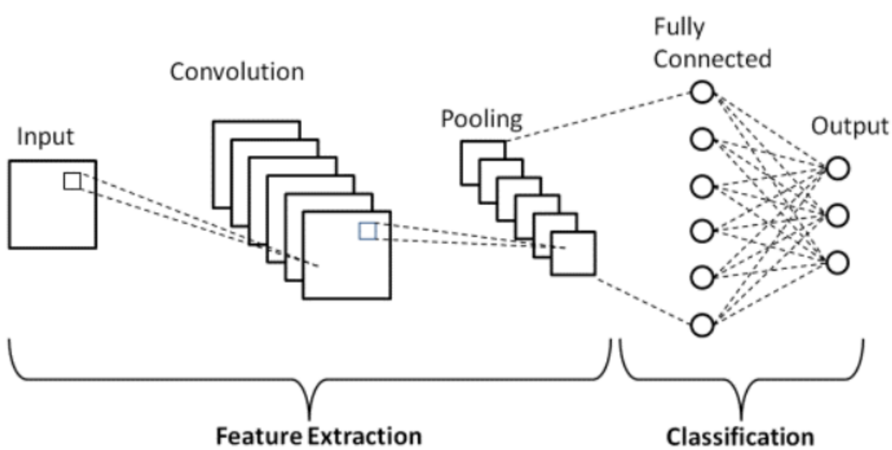

Figure 1: A simple representation of a general neural network model.

In the specific case it presents an initial input layer followed by five convolutional layers. Each layer has an activation function (the Rectified Linear Unit one, known as ReLu) and a max pooling layer, always with kernel size equal to 2 and stride equal to 2. The input channels of each layer is always the output channels of the previous layer. The first layer has an output channels of 64, a kernel size of 3x3 with padding 1 and stride 1. It is the only layer without the max pooling. The second layer has an output channels of 128, a kernel size of 3x3 with padding 1 and stride 1. The third layer has an output channels of 256, a kernel size of 3x3 with padding 1 and stride 1. The fourth layer has an output channels of 512, a kernel size of 3x3 with padding 1 and stride 1. The fifth layer has an output channels of 512, a kernel size of 3x3 with padding 1 and stride 1. The output channels of 512 will not be the input channels of the fully connected layer. It has to be computed (2048) and after we can proceed to the final classifier layer.

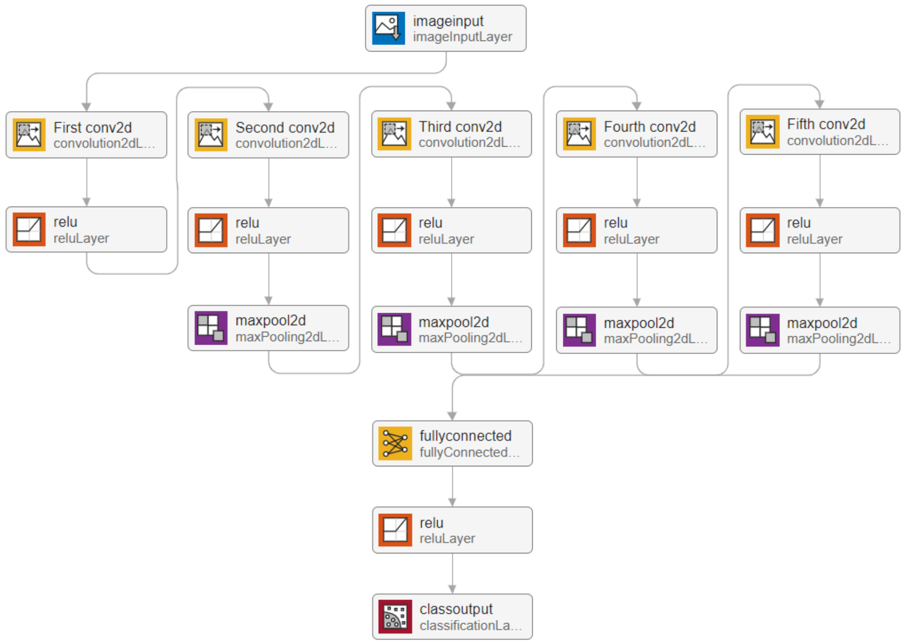

Figure 2: A representation of our neural network model in MatLab

2. Dataset

The dataset was found in a GitHub public repository: https://github.com/Horea94/Fruit-Images-Dataset. It consists in 90380 fruits and vegetables images. They all have a 100x100 pixels image size and they are divided in 131 classes. The following fruits and vegetables are included: Apples (different varieties: Crimson Snow, Golden, Golden-Red, Granny Smith, Pink Lady, Red, Red Delicious), Apricot, Avocado, Avocado ripe, Banana (Yellow, Red, Lady Finger), Beetroot Red, Blueberry, Cactus fruit, Cantaloupe (2 varieties), Carambula, Cauliflower, Cherry (different varieties, Rainier), Cherry Wax (Yellow, Red, Black), Chestnut, Clementine, Cocos, Corn (with husk), Cucumber (ripened), Dates, Eggplant, Fig, Ginger Root, Granadilla, Grape (Blue, Pink, White (different varieties)), Grapefruit (Pink, White), Guava, Hazelnut, Huckleberry, Kiwi, Kaki, Kohlrabi, Kumsquats, Lemon (normal, Meyer), Lime, Lychee, Mandarine, Mango (Green, Red), Mangostan, Maracuja, Melon Piel de Sapo, Mulberry, Nectarine (Regular, Flat), Nut (Forest, Pecan), Onion (Red, White), Orange, Papaya, Passion fruit, Peach (different varieties), Pepino, Pear (different varieties, Abate, Forelle, Kaiser, Monster, Red, Stone, Williams), Pepper (Red, Green, Orange, Yellow), Physalis (normal, with Husk), Pineapple (normal, Mini), Pitahaya Red, Plum (different varieties), Pomegranate, Pomelo Sweetie, Potato (Red, Sweet, White), Quince, Rambutan, Raspberry, Redcurrant, Salak, Strawberry (normal, Wedge), Tamarillo, Tangelo, Tomato (different varieties, Maroon, Cherry Red, Yellow, not ripened, Heart), Walnut, Watermelon. The folder already contains training and test sets. Then we decided to extract from the train set a 10 percent of validation set. The three final datasets have these dimensions: training set with 60923 images, validation set with 6769 images and test set with 22688 images.

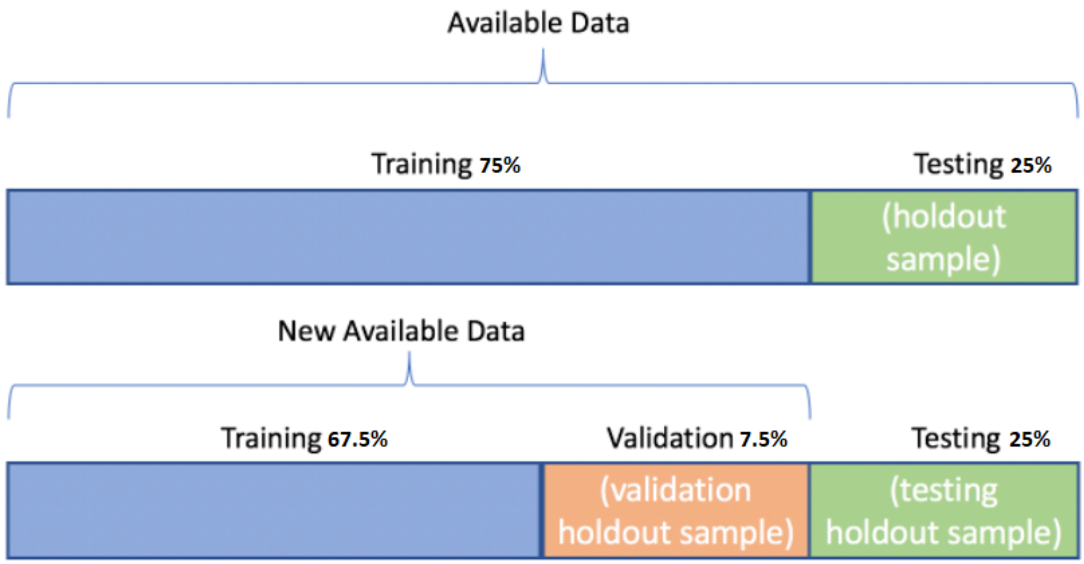

Figure 3: A representation of dataset split.

3. Training procedure

This section is about the training options we decided to use for the model. Before starting the neural network, previously presented, we divided the images of each data loaders (containing train, validation and test) in batches of 32 element size. This means that for the data loader of the train we get 60923/32 batches, for the validation one 6769/32 batches and for the test one 22688/32 batches. This technique is useful because we can analyze the images together in batches. We also used an optimizer with a learning rate (which is a tuning parameter in an optimization algorithm that determines the step size at each iteration searching a minimum of a loss function) equal to 0.005 and we defined the cross entropy criteria for the loss, commonly used for classification problems. We processed the model in 10 epochs, without a stopping criteria: reached the tenth epoch the algorithm is stopped. At each epoch algorithm returns several values; the three most important are Validation Accuracy, Test Accuracy and Validation Loss. We usually choose the Test Accuracy correspondent to the highest Validation Accuracy and the lowest Validation Loss; often (but not always) the best result is at the tenth epoch with an high Test Accuracy value.

4. Experimental Results

Several experiments were conducted to test the effectiveness of the proposed Neural Network. Model was trained as described in the previous section. Table 1 shows test results for the proposed architectures as well as of ablation studies (i.e., different variants of the final architecture when adding or removing layers). From the below table we can see that every time a new layer is added, the test accuracy decreases, except for the fifth layers. There, we can find a little increase compared to the previous one.

Model	Accuracy
− + Layer 1	94.37%
− + Layer 2	94.43%
− + Layer 3	92.53%
− + Layer 4	90.01%
− + Layer 5	91.41%

Table 1: At each layer corresponds a value of test accuracy with the lowest validation loss and the highest validation accuracy.

For the final model (5 layers) we also created a confusion matrix, but we prefer to leave it only in the code section because it was too difficult put an image of a matrix with 131 classes. Usually the rows correspond to the predicted class and the columns correspond to the true class. The diagonal cells correspond to observations that are correctly classified. The off-diagonal cells correspond to incorrectly classified observations. Almost all classes get a high percentage of being correctly classified, with probability values between 80 and 100 percent. The lowest percentage corresponds to Plum 3 class and it is 35 percent. Also Pear class has one of the lowest percentage, equal to 45.