Automated Groceries Storage: Fridge or Non-Fridge decision

April 22, 2018

1 Definition

1.1 Project Overview

This project is an approximation of detecting if a product should be stored in the fridge or not.

Nowadays, time is limited, human have a lot of amount of work to do and thinking in system that could help in everyday tasks is the next step in the human evolution.

Some initiatives in Computer Vision and Deep Learning for detecting products were proposed. See "The Freiburg Groceries DataSet" Philipp Jund at el. [4] . Those are focussed in helping big retailers in product storage.

This project uses the "Precios Claros" Dataset. It's a public database that store information of retailers products including price, category, picture, etc. Only the images will be used and the objective is based in the product shape determine if it should be stored in the fridge or not.

1.2 Problem Statement

This project proposes a way to detect if a grocery product should be stored on the fridge or not based on the shape. There are no other works in the field about this problem.

The solution proposed uses Convolutional Neural Networks. Two approaches were considered: using transfer learning and an original CNN from the scratch.

The expectation is that the model build from the scratch could beat in accuracy the pre-trained model.

1.3 Metrics

At training time, the accuracy will be used to measure the success of the network. Also, the loss and validation loss progress will be checked to provided an accurate model without underfitting or overfitting conditions.

The final score will be given by the **F1-score** over the test dataset.

The F1-Score will be calculated using this formula:

$$F1\_Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The Precision will indicate the percentage of products we are storing correctly in the fridge. The formula will be:

$$Precision = \frac{PredictedFridge|Fridge}{PredictedFridge|Fridge + PredictedFridge|Non\_\_Fridge}$$

The Recall is going to be definined around the fridge products. We are interested in correctly store them in the fridge and non of them should be out of it. That would lead in the wasting of the product.

$$Recall = \frac{PredictedFridge|Fridge}{PredictedFridge|Fridge + PredictedNon\_\_Fridge|Fridge}$$

2 Data Exploration

2.1 Exploratory Visualization

The Precios Claros Dataset is composed by 4685 images of grocery products. All the images have only 1 product and they have white background. The image is clear. Each image is RGB and the dimension are 240x240.

All the products were classified in fridge or non-fridge according the storage requirements.

Some products could be either. For example, a closed soda could be stored in the fridge or not. The criteria in this case is to store it in the fridge.

Here are some sample images:

Figure 1: Precios Claros Dataset

The statistics of the clases is the following:

Class Id	Class Name	Products Qty	Percentage over Total
0	fridge	683	0.15%
1	non_fridge	4002	0.85%
Total		4685	100.00%

Figure 2: Precios Claros Classes Distribution

Notice that most of the products are non-fridge products.

2.2 Algorithms and Techniques

This project is solved using Deep Learning. The techiques applied are:

- Data Augmentation
- Transfer Learning
- Convolutional Neural Networks
- Dropout
- Hyperparameters tuning
- Optimizer selection

2.2.1 Data Augmentation

Data Augmentation is a techinique broadly used to prevent overfitting and to increment the training data. The idea is to increment the data by processing the images in the dataset like scales, rotations, etc.

For example, if you apply data augmentation over the Precios Claros dataset we can have:

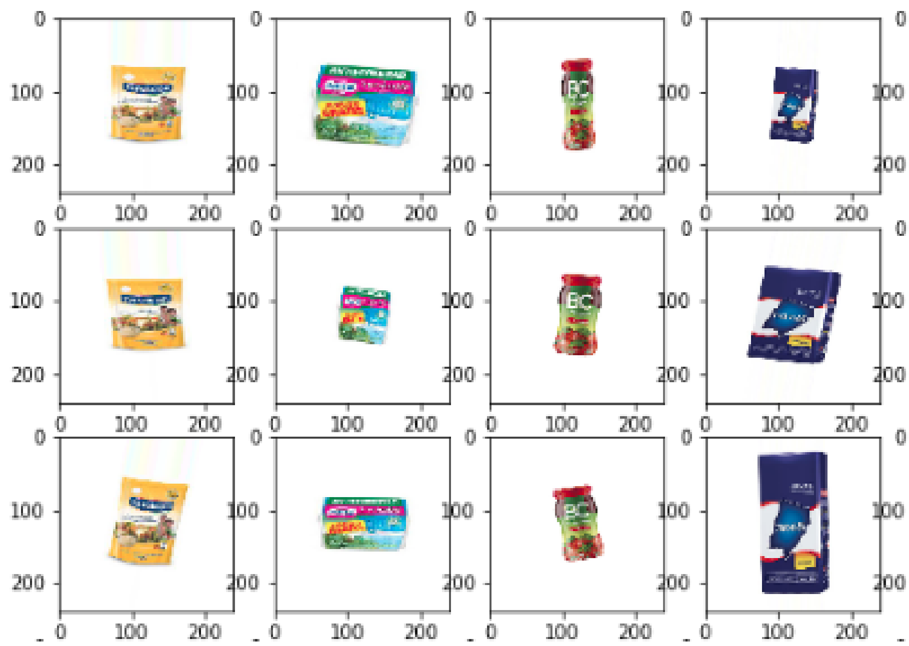

Figure 3: Data Augmentation

2.2.2 Transfer learning

Transfer learning is a very useful techinique in deep learning that let us use the previous trained model in our data.

Models that were trained for weeks and on multiple GPU could be used over new data to make a prediction with amazing results.

2.2.3 Convolutional Neural Networks

CNN are a kind of Neural Networks mainly used in Computer Vision. It consist in applying convolution operations over an image pixels to create filters that extract different features/characteristic of an image. For example, borders, lines, etc.

2.2.4 Dropout

Consist in randomly disable some of the units in a Neural Network layer. It uses a probability to decide if the node (unit) needs to be disabled.

This will let the NN to not stick to a fixed architecture and don't get stuck in characteristics that can be shared by some nodes.

2.2.5 Hyperparameters tuning

The accuracy of a neural network will depends on different parameters:

- Learning Rate
- Number of Layers
- Batch Size
- Number of Units/kernel/strides in the convolutional layer
- Dropout probability in the Dropout layers

2.2.6 Optimizer selection

Optimizer play a fundamental role in learning. The classic Stochastic Gradient Descent algorithm had many improvements. One of them was the adding of momentum, then a Adam, Adamax, and Adagrad new techniques.

2.3 Benchmark

The Benchmark model is a VGG16 transfer learning model.

The network structure is defined:

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
global_average_pooling2d_2 ( (None, 512)               0
_________________________________________________________________
dense_8 (Dense)              (None, 2)                 1026
=================================================================
Total params: 1,026
Trainable params: 1,026
Non-trainable params: 0
_________________________________________________________________
```

Figure 4: VGG16 Model

The accuracy of the model is: 51.56%

	fridge prediction	no-fridge prediction	total by category
fridge	17	45	62
non-fridge	17	49	66
Total by prediction	34	94	

How many of the products predicted to go to the fridge actually needs to go in the fridge? Precision: 50.00%

Of the fridge items, how many were correctly predicted Recall: 27.42%

F1-Score: 0.35

3 Methodology

3.1 Data Preprocessing

The Precios Claros dataset provides 4685 clear images. As it was discussed before Data Augmentation will be used to increase the performance of the Neural Networks. Usually, training over more data allow us to get better results.

Also, for using Transfer Learning, the images should be resized (VGG16 uses 224x224x3 images) and the bottleneck extracted.

The other preparation required is the separation between train, valid, and test set.

The complete pre-processing is summarized in this graph:

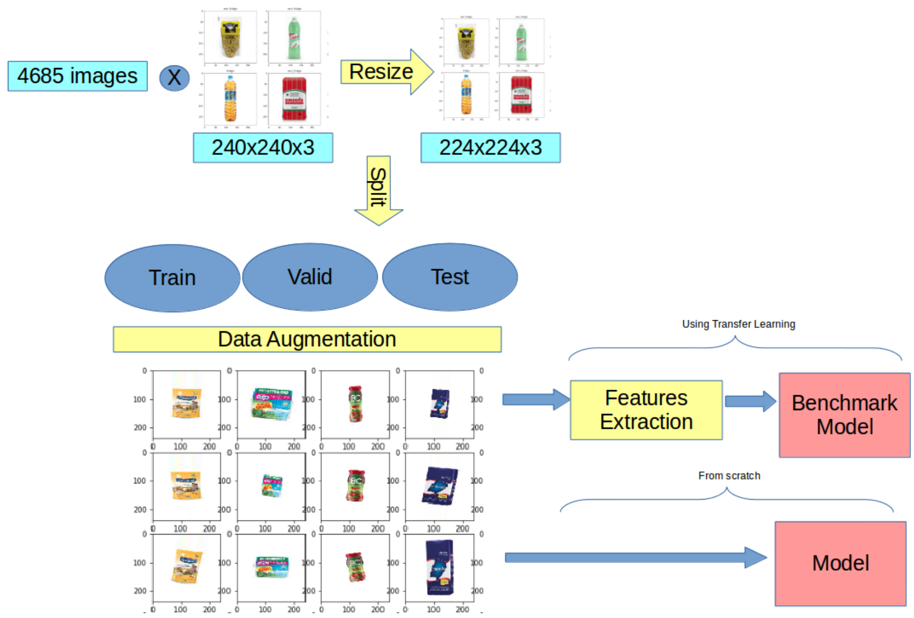

Figure 5: Data preparation flow

3.2 Implementation

The main issue was the data aumentation implementation. The idea was to generate all the images in memory. The flow method finally allowed me to do that.

For improving benchmark results, a CNN from the scratch was implemented. The technique uses a Convolutional Neural Network followed by a Max Pooling layer.

Dropout layers were used to prevent overfitting.

Layer (type)	Output Shape	Param #
conv2d_26 (Conv2D)	(None, 223, 223, 16)	208
max_pooling2d_26 (MaxPooling	(None, 111, 111, 16)	0
conv2d_27 (Conv2D)	(None, 110, 110, 32)	2080
max_pooling2d_27 (MaxPooling	(None, 55, 55, 32)	0
conv2d_28 (Conv2D)	(None, 54, 54, 64)	8256
max_pooling2d_28 (MaxPooling	(None, 27, 27, 64)	0
dropout_21 (Dropout)	(None, 27, 27, 64)	0
conv2d_29 (Conv2D)	(None, 26, 26, 128)	32896
max_pooling2d_29 (MaxPooling	(None, 13, 13, 128)	0
flatten_7 (Flatten)	(None, 21632)	0
dense_21 (Dense)	(None, 256)	5538048
dropout_22 (Dropout)	(None, 256)	0
dense_22 (Dense)	(None, 128)	32896
dropout_23 (Dropout)	(None, 128)	0
dense_23 (Dense)	(None, 64)	8256
dropout_24 (Dropout)	(None, 64)	0
dense_24 (Dense)	(None, 2)	130

Figure 6: Data preparation flow

3.3 Refinement

The new network was trained for 300 epochs using Model Checkpoint to save the best model weights.

The learning rate 0.0001 in Adamax optimizer. Others learning rate, 0.05, 0.01, 0.001 were tried but the best results were in a slow learning.

The batch size that best worked was 64. Other values like 32, 128 were tried without outperform the selected parameter value.

The optimizer selection was delicate. I first started with a stochastic gradient descent and finally, changes to Adamax that it's an improved optimizer.

4 Results

4.1 Model Evaluation and Validation

The accuracy of the model is: 62.32%

	fridge prediction	no-fridge prediction	total by category
fridge	29	40	69
non-fridge	12	57	69
Total by prediction	41	97	

How many of the products predicted to go to the fridge actually needs to go in the fridge? Precision: 70.73%

Of the fridge items, how many were correctly predicted Recall: 42.03%

F1-Score: 0.53

4.2 Justification

The final model is a good aproach to solve the fridge or non-fridge product storage.

In the next graph, we are going to compare accuracy, precision, and recall.

All the values were highly improved by the new model. Recall is important to avoid false negative. It's important that we can correctly store all the fridge product.

The new model works better and the reason could be related to the dropout layers, and the optimizer chosen.

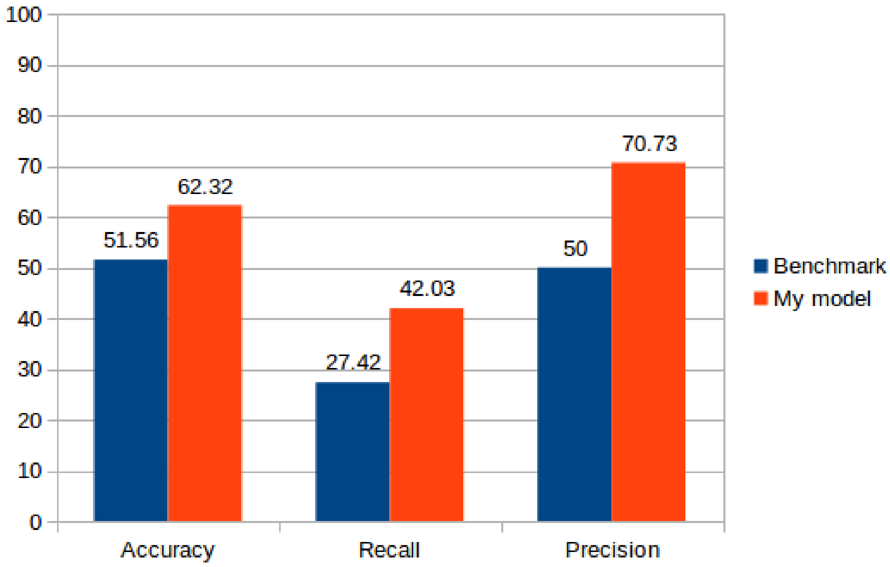

Figure 7: VGG16 Model

Finally, if we compare the F1_score clearly the new model outperforms the benchmark model.

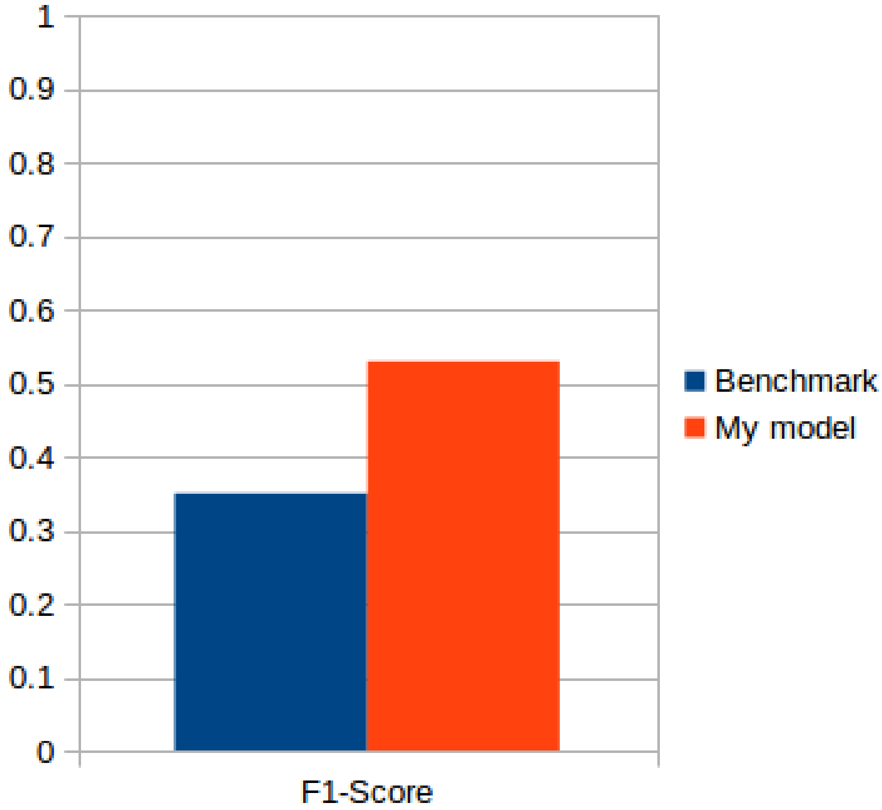

Figure 8: VGG16 Model

5 Conclusion

5.1 Free-Form Visualization

After training, we can review the resutls. They aren't very good.

Figure 9: VGG16 Bechmark Model results - in uppercase the correct label

They are a little bit better if we consider the new proposed model.

Figure 10: Proposed Model results - in uppercase the correct label

5.2 Reflection

The solution proposed is better than the benchmark but no better enough. It's important to realize that a product left out of the bridge could be wasted so it's important to have a high accuracy in fridge products.

The new model does a good work. If we take a look at the precision, the 70% of the products that requieres cold will be correctly stored. However, this is not enough. The 30% of our products will be wasted and that's not good at all.

5.3 Improvement

The problem is hard in its definition. Maybe, just looking the package is not enough information for making a decision.

However, the results are quite interesting. A close analysis over the 30% missclassified fridge products could be performed to develop algorithms that let us increase the accuracy.

Also, some transfer learning advanced model could be used like ResNet50 or Xception and compared to the proposed model.

References

[1] *Super Market Robots.* `https://www.fastcodesign.com/90150368/this-online-supermarkets-robots-put-your-order-together-in-minutes`

[2] *Precios Claros* `https://www.preciosclaros.gob.ar/#!/buscar-productos`

[3] *Freiburg groceries dataset* `http://aisdatasets.informatik.uni-freiburg.de/freiburg_groceries_dataset/`

[4] *Freiburg groceries paper* `https://arxiv.org/abs/1611.05799`

[5] *VGG16 Paper* `https://arxiv.org/abs/1409.1556`