Architectural Decision Document

Garbage classifier - Capstone project

• Data quality assessment

The dataset used in the project contains 2527 RGB images which are all 512x384. They have been checked and labeled by hand. The 6 different classes contain different number of elements and this imbalance must be compensated.

• Feature engineering

To obtain better performance with our model, we augmented the dataset. Due to the relatively small size of the dataset, the image augmentation can provide images which are new to the model and thus prevent overfitting and increase the generality of the model. The augmentation consist of: rotations, flips, zooms, shears, and shifts. Furthermore the images are normalised such as their values are bounded between 0 and 1.

• Framework

The framework of choice for this project is Keras with TensorFlow as backend and Scikit-learn to perform metrics evaluation. Keras is a powerful, open and user friendly environment. It allows to load, preprocess, augment, and serve images very easily. With its sequential API, it's also very easy to build models by stacking different pre-defined layers on top of each other. Sometimes it can be limiting because it's API doesn't allow access to some stages of the workflow, but in this case the library performed well and we never felt limited by it. Scikit-learn has been chosen because it contains many powerful and useful functions, especially during the test phase. In this project we use it to calculate the confusion matrix and the F1 score.

• Algorithms

The projects compares the results of two different algorithms: one from Machine Learning and one as part of a Deep Neural Network model. The first one is a Logistic Regression and by only considering individual pixel values it’s not supposed to perform so well on image classification. Nevertheless it can still serve as a reference base line to evaluate the performance of the other model. The second model is a Convolutional Neural Network which consists of 4 2D Convolutional layers and 3 Fully Connected layers. Furthermore, MaxPooling and Dropout layers are present to help with overfitting. The Convolutional way is the preferred when dealing with image classification problem, because the convolutional layers can continuously abstract spatial features from the image regardless their absolute position or rotation in the image.

• Performance Indicators

The metrics used to evaluate the classification performance of the model are accuracy, confusion matrix and F1 score. The accuracy is the simplest and has been chosen because it reflects the ability of the model to recognise the correct class of a sample. The confusion matrix was chosen because it shows how the model perform for each class, for example if there is a class that the model struggles to identify there is room for further feature engineering specific for that class. Finally the F1 score calculates a weighted average between precision and recall. It evaluates the balance between the exactness and the completeness of the predictions, and it’s a good overall indicator of the performance of the model.

• Deployment

The use case of the model is to embed it in the control process of a waste separation machine. Once the model is trained it doesn’t require much computing power nor hard drive space to function. The model is about 14 MB in size and can be run by any modern CPU. To make the system work correctly, a segmentation model should be added, in order to segment the video feed from the cameras and to provide still images with one item each to the classifier. Furthermore it is suggested an hardware solution to physically separate the waste, controlled by the software section, providing a great increase in the efficacy of the system.