

TFG informe inicial

Raquel

October 23, 2017

Contents

1	Data	1
1.1	Initial Data	1
1.2	Generated Data	1
2	Análisis	3
2.1	Intra sinset	3
2.2	Distribución total de las features	5
2.3	Distribución de las features por layer	6
2.4	Features per image	9
2.5	Distribución de las features por synset	11
2.6	Images per feature	12

1 Data

1.1 Initial Data

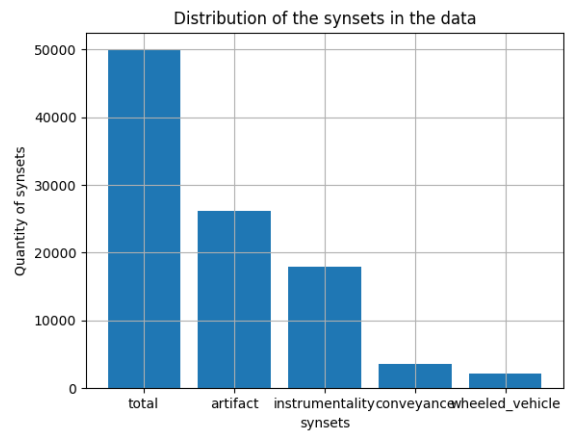
- Embedding matrix of size (50000, 12416), con 62080000 features.
- **labels** Labels vector of size 50k which every label is in numeric format (0, 999)
- **synsets = synset0 synset1 synset2 ...** The set of synsets that we will analyze:
synsets = [living_things, mammal, dog, hunting_dogs]
- **categories = {-1 0 1 }** The possible values of the features.

1.2 Generated Data

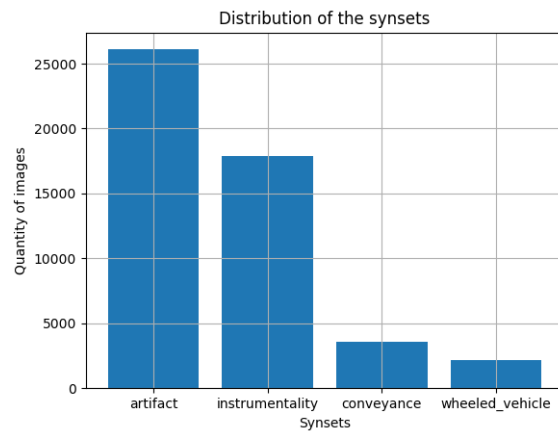
- **synset_index_hyponim.txt** A list with all the hyponims of every synset.
- **synset_index.txt** For each synset a list with the index of the elements of the hyponim list in the embedding.

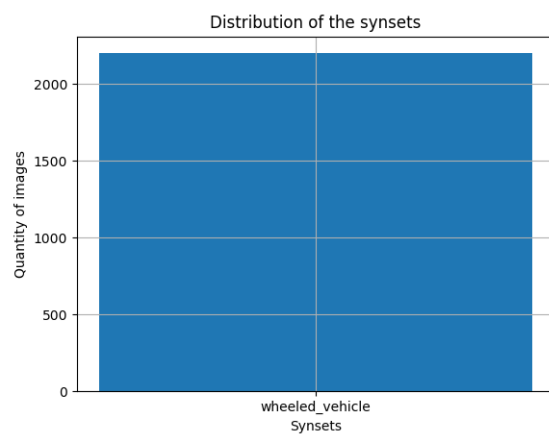
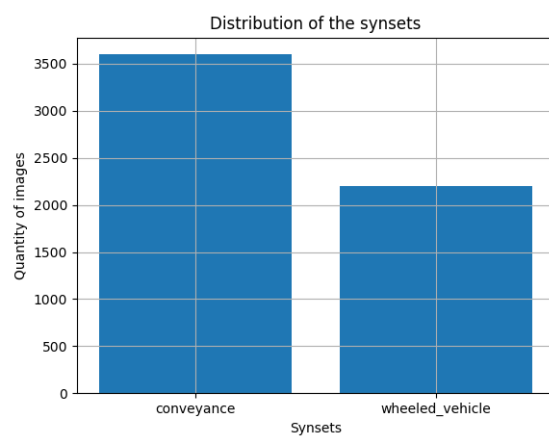
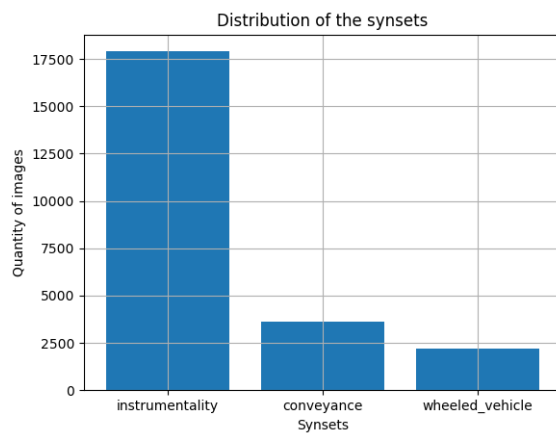
- Un diccionario con la cantidad de imágenes que tiene cada feature para cada category.
- `features_per_image[synsets].pkl` dfd
- `features_per_layer[synsets].pkl` dsf
- `images_per_feature_per_synset[synsets].pkl` dsfs
- `intra_synset[synsets].pkl` dsfs

2 Análisis

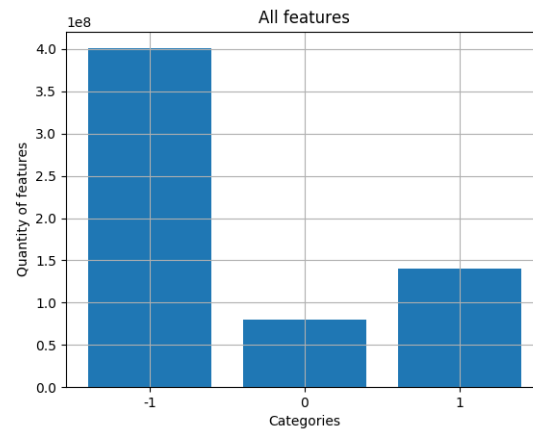


2.1 Intra sinset

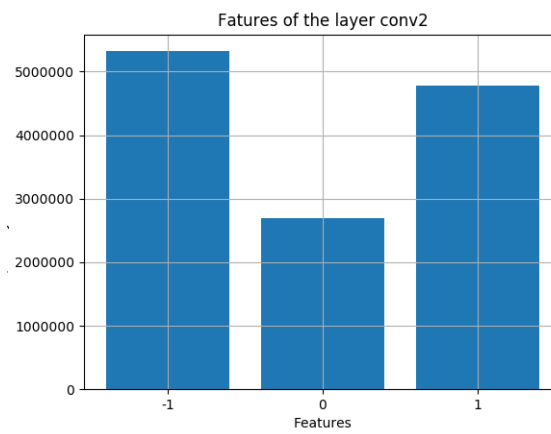
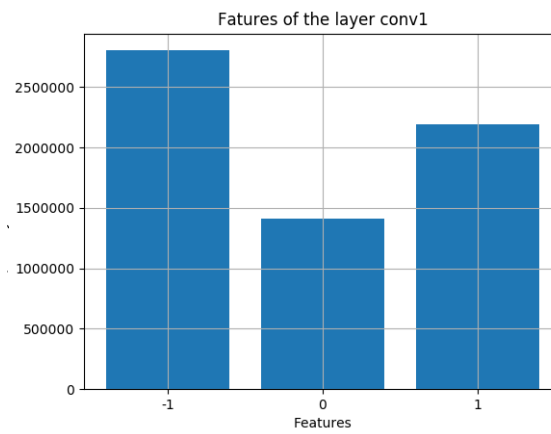


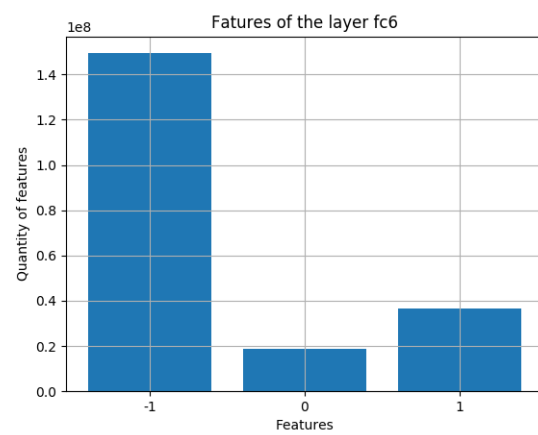
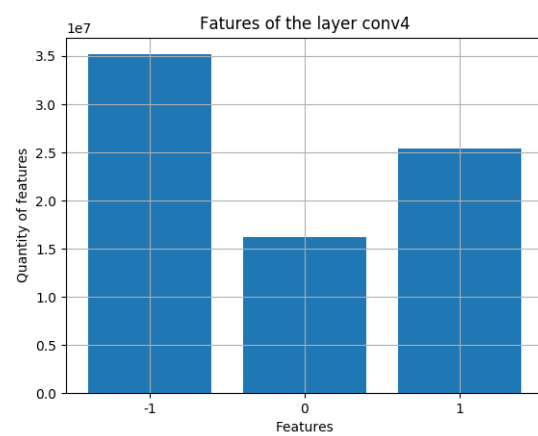
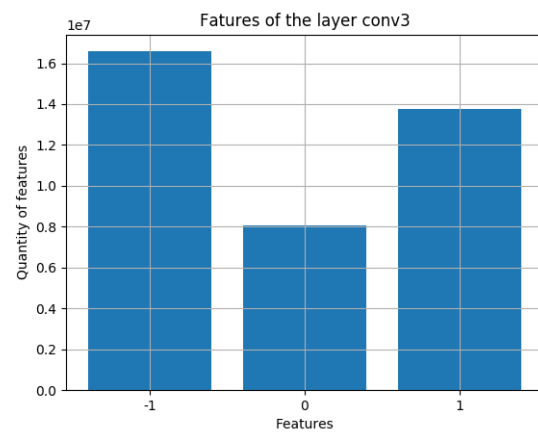


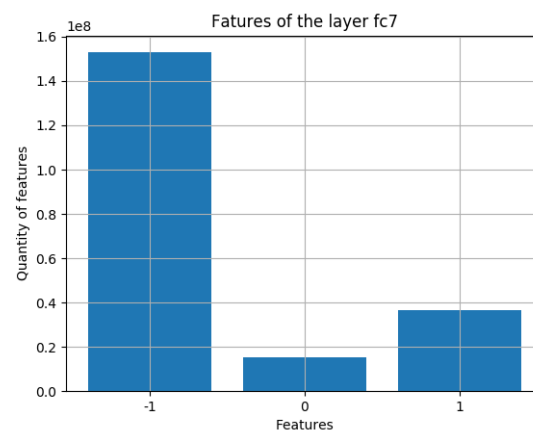
2.2 Distribución total de las features



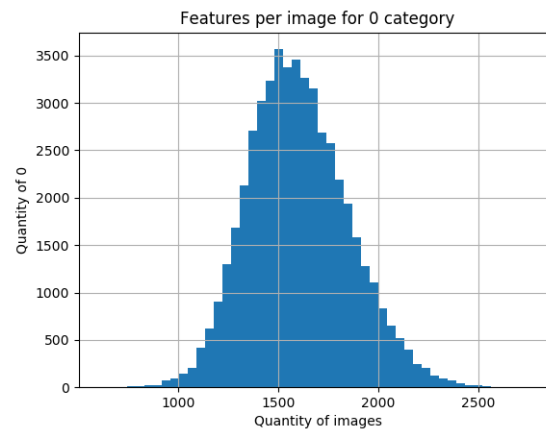
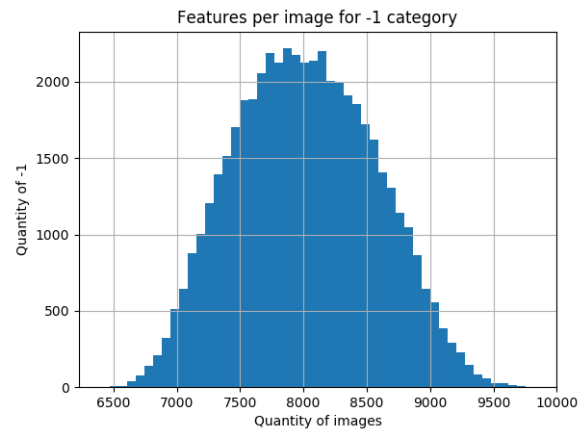
2.3 Distribución de las features por layer

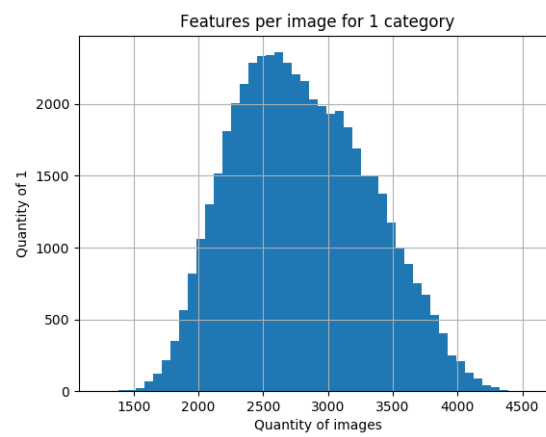




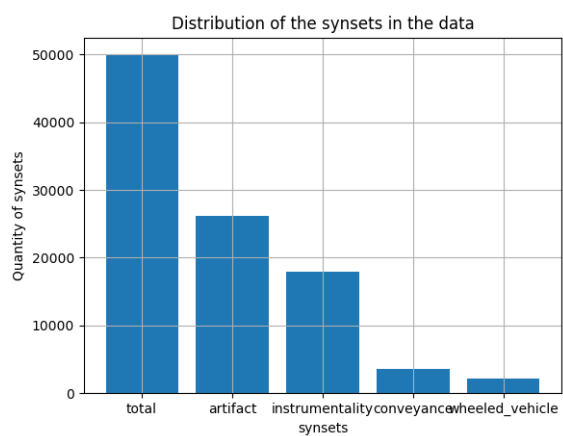


2.4 Features per image





2.5 Distribución de las features por synset



2.6 Images per feature

