Part-based Object Detection using RealAdaBoost and ANN

Manuel Montoya Catalá

Computer Vision

Carlos III University of Madrid

June, 2015

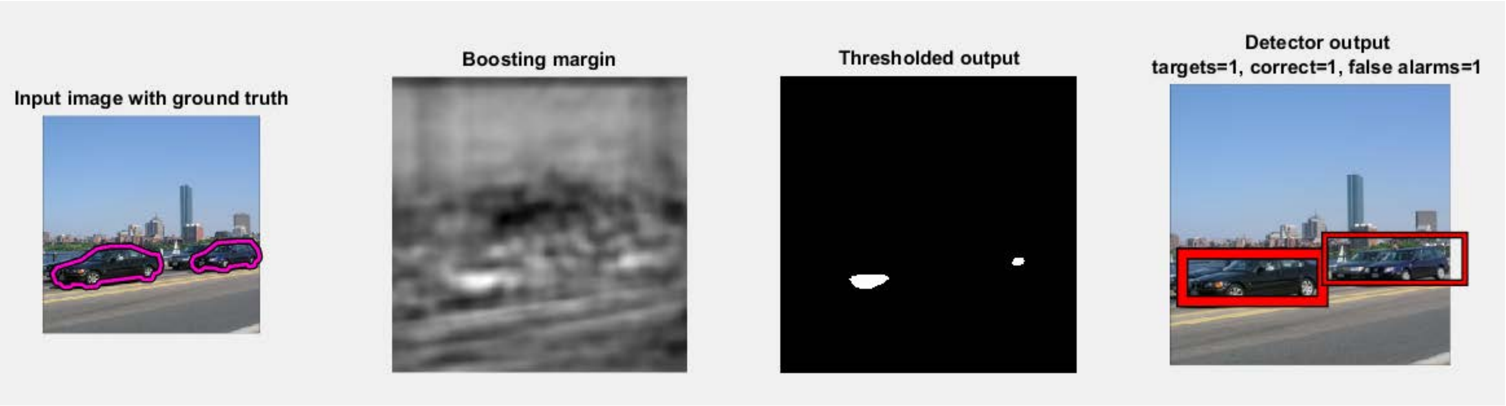
Introduction

In this project we propose a part-based system for car detection.

- Using a Star model for the relationship among the parts.
- RealAdaBoost + Neural Nets as Classifier.
- NN trained with the Extreme Learning Machine algorithm.
- Results evaluated using the Precision-Recall curve.

Project is based in the Lab Session 8 – Object Detection of this subject.

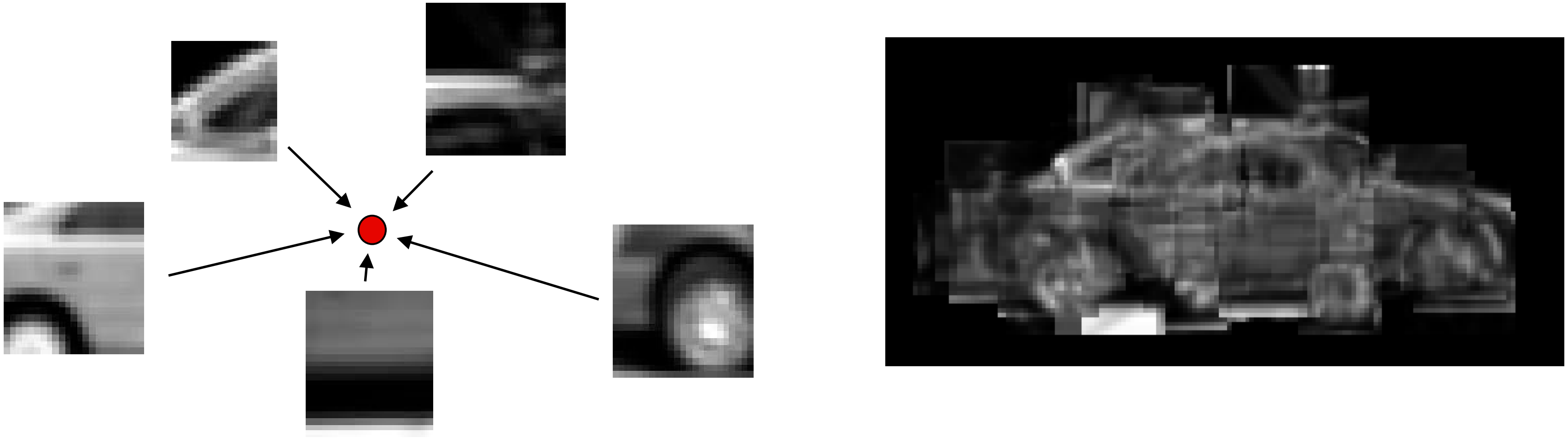
- Improvement on the Detector.



Part-based Models

Objects are represented as a set of parts and their spacial relationship.

- A part is any element of an object that can be reliably detected.
- Parts were chosen randomly from the grountruth of 8 images from the dataset.



Car model

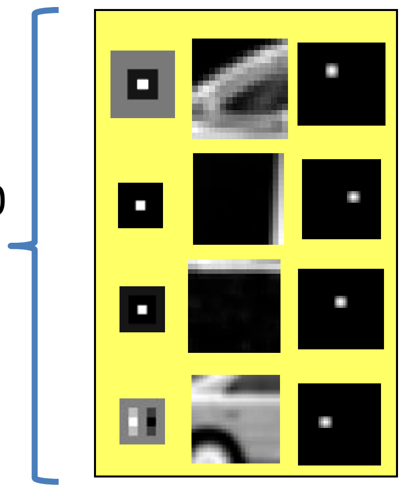
Part-based Models

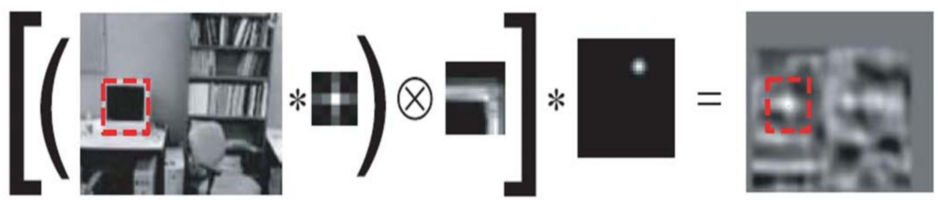
Each part of the model is transformed into a feature. The car-model is composed by 640 features.

Each Features composed by a triplet:

- Filter
- Patch
- Location

640



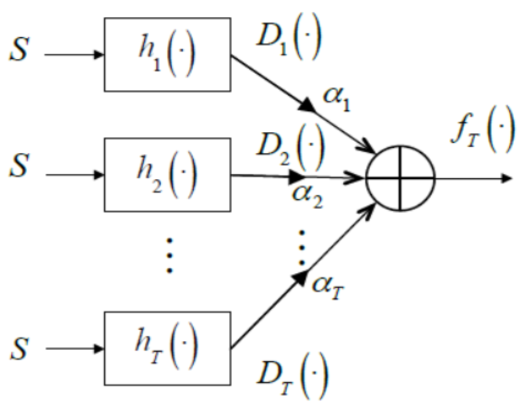


$$h_i(I, x, y) = [|I * f_i| \otimes P_i] * g_i$$

RealAdaBoost

Ensemble method for creating a strong classifier as a linear combination of Weak Classifiers.

- Each Weak Learner is trained with a different distribution of samples $\bar{D}_t$

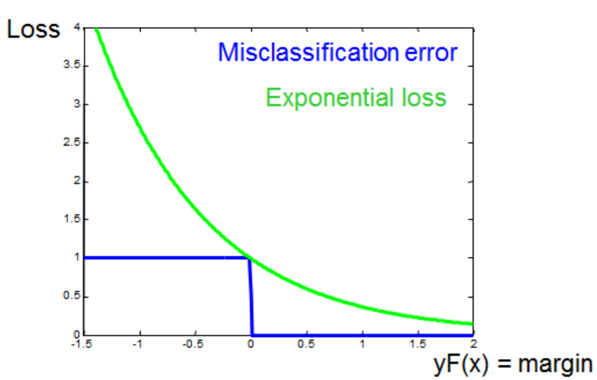


$$H(\bar{x}_i) = sign\left(\sum_{t=1}^{T} \alpha_t\, h_t(\bar{x}_i)\right)$$

$$\bar{D}_t = [D_t(1), D_t(2), \cdots, D_t(M)]$$

Update Rule:

$$D_{t+1}(i) = \frac{D_t(i) e^{-\alpha_t\, h_t(\bar{x}_i) y_i}}{Z_t}$$



Loss function:

$$E = \sum_{i=1}^{M} e^{-y_i f(\bar{x}_i)}$$

RealAdaBoost

As a result of the Updating Rule, the error probability is bounded by:

$$P_e \leq \prod_{t=1}^{T} Z_t \qquad\qquad Z_t = \sum_{i=1}^{M} D_t(i) e^{-\alpha_t \, h_t(\bar{x}_i) y_i}$$

Normalization constant

Once the Weak Learner $h_t(\cdot)$ is trained:

- The value of α_t should minimize Z_t.

RealAdaBoost choses the following value of α_t minimizing an approximation function:

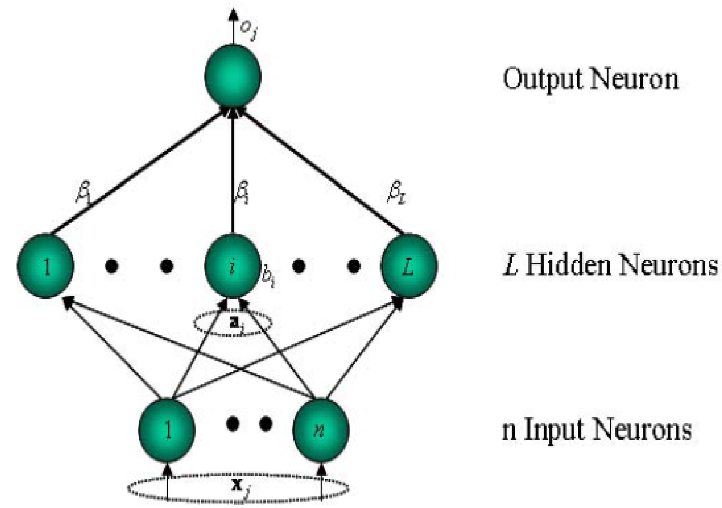
$$\alpha_t = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right) \qquad\qquad r = E\{\text{margin}\}_{\bar{D}_t} = \sum_{i=1}^{M} D_t(i)\, h_t(\bar{x}_i)\, y_i$$

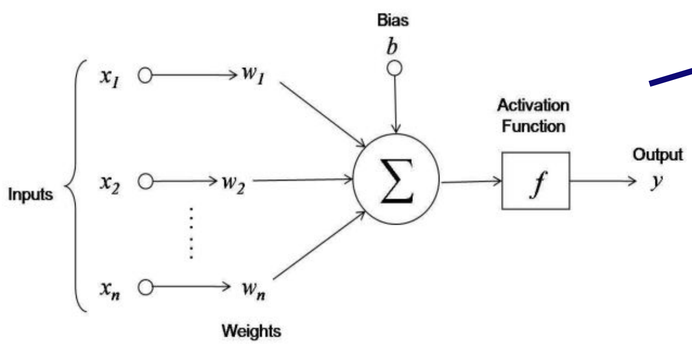
ANN + ELM

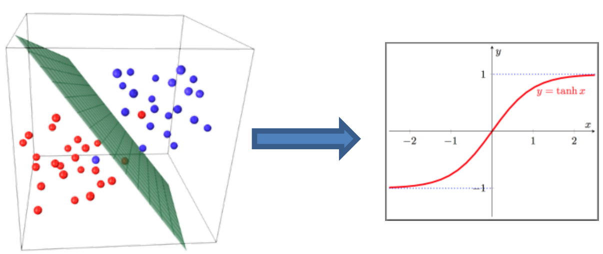
Machine learning system inspired by the architecture of the mammals’ brain.

- Composed by small processing units called neurons.
- The neurons are interconnected and share information with each other.
- Each neuron outputs a transformed linear combiantion of its input.

Neuron can be seen as a hyperplanes.
Their output is the projection of the samples over that hyperplane.







Each neuron has parameters:

$$\{\overline{\mathrm{W}}, \mathrm{b}\}$$

ANN + ELM

Extreme Learning Machine is an algorithm to train an ANN very fast:

- It randomly chooses the parameters of the hidden neurons
- Computes the weights of the output neuron as Least Square Solution o

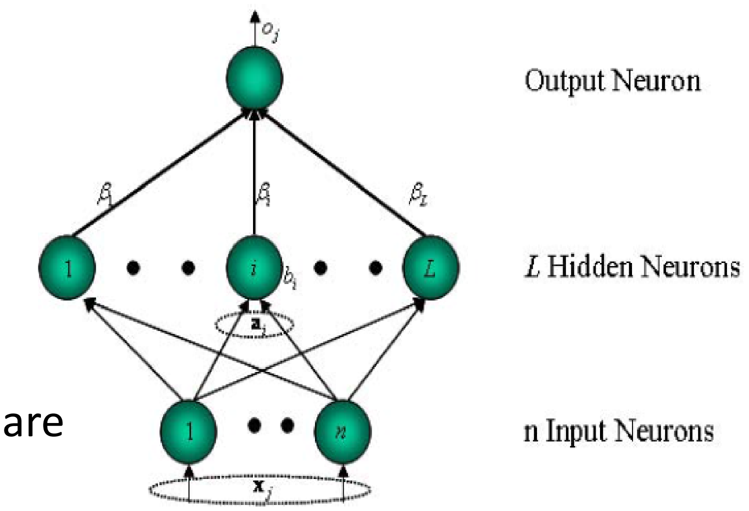
$$E_{MSE} = \sum_{i=1}^{M} D_i \cdot \varepsilon_i^2 = \bar{D} \circ (\bar{\varepsilon}^t \cdot \bar{\varepsilon})$$

$$\bar{\bar{\Lambda}} = \begin{pmatrix} D_0 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & D_M \end{pmatrix} \qquad \bar{D} = \mathrm{diag}(\bar{\bar{\Lambda}})$$

$$\bar{W}_o = (\bar{\bar{H}}^t \bar{\bar{\Lambda}} \bar{\bar{H}})^{-1} \bar{\bar{H}} \bar{\bar{\Lambda}} \cdot \bar{T} = \bar{\bar{H}}^{+} \cdot \bar{T}$$

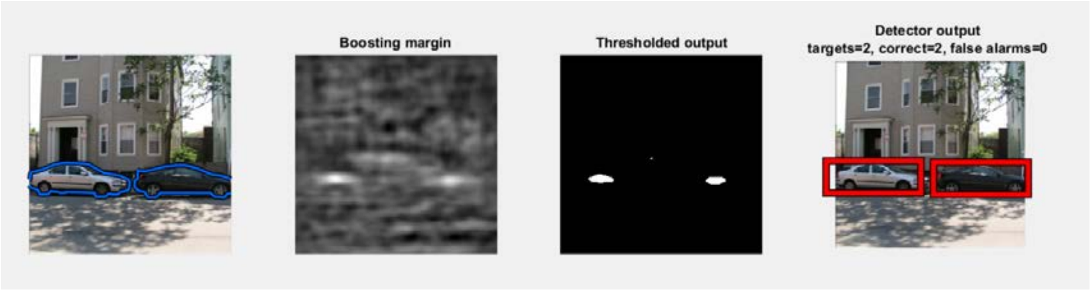
This algorithm can be interpreted as :

- A generation of random orthogonal hyperplanes
- transformed by a sigmoid function
- and linearly combined by $\bar{\beta}$ as the least square solution of the system.

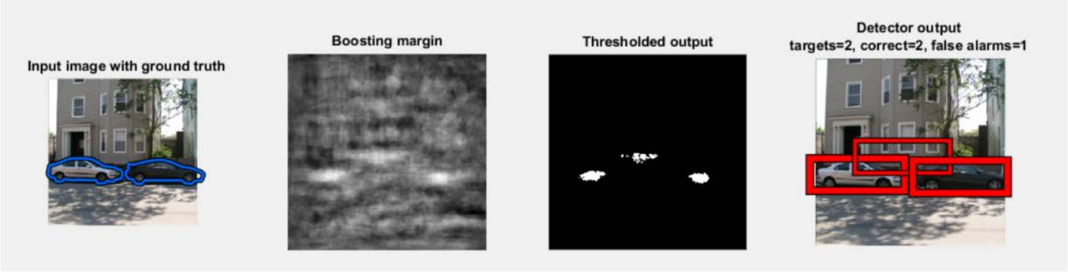


Experiments Setup

The system proposed performs object recognition over the cars repository of the LabelMe Dataset. The dataset is composed of 776 images that contain cars in different scenarios; these cars appear from different points of view including partially occluded positions.



ANN + ELM Detection

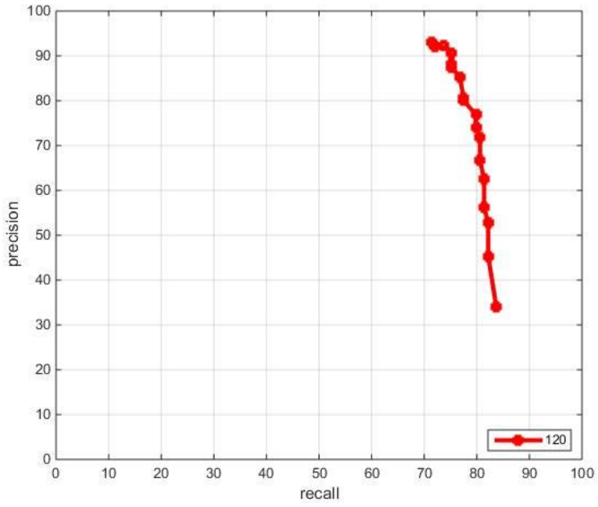


Decision Stump Detection

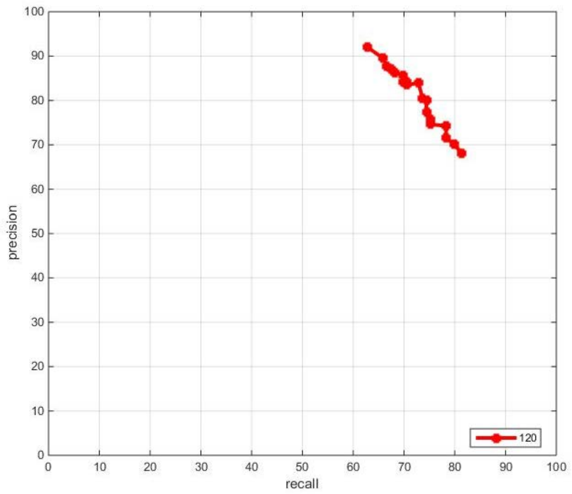
Evaluation

Results were evaluated using the Precision-Recall curve:

$$precision = \frac{Number\ of\ correct\ objects\ detected}{Total\ number\ of\ objects\ detected} \qquad recall = \frac{Number\ of\ objects\ detected}{Total\ number\ of\ correct\ objects}$$



Precision Recall for the Decision Stump



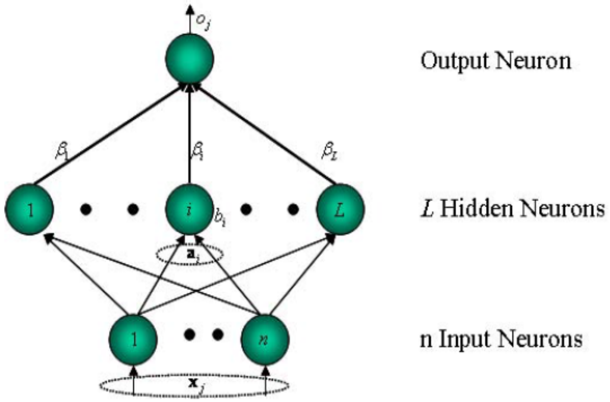
Precision Recall for the ELM-ANN

Conclusions and Future Work

- ANN + ELM + RealAdaBoost is an efficient and fast way to perfrom classification.
- Although boosting is needed to avoid over-fitting, there is no need for deep boosting implementations when using an ANN trained with ELM since a single ANN is good enough for performing a basic classification.
- Having the number of hidden neurons as a parameter increases the adaptability of the classifier, therefore potentially improving the results.

Future Work:

- Train each ANN weak learner over a random subset of the features
- Implement the pre-detection filter that removes small false alarms.
- Perform multi-scale detection.



References

[1] Lab Session 8 – Object Detection from C4.278.12995-1 COMPUTER VISION 14/15-S2.

[2] Antonio Torralba and Bryan Russell , “LabelMe Toolbox: MATLAB Toolbox for the LabelMe Image Database”, 2008 MIT, Computer Science and Artificial Intelligence Laboratory. http://people.csail.mit.edu/torralba/LabelMeToolbox/

[3] Object Detection with Discriminatively Trained Part Based Models. Pedro F. Felzenszwalb, Ross B. Girshick, David McAllester and Deva Ramanan

[4] ICCV 2013 Tutorial on Part-based Models for Recognition. Sydney, Australia.

[5] Extreme learning machine: Theory and applications. Guang-Bin Huang, Qin-Yu Zhu, Chee-Kheong Siew.

[6] Improved Boosting Algorithms Using Confidence-rated Predictions. Robert E. Schapire Yoran Singer

Thank you !!