Object Recognition: HOG/SVM Classifiers

Artificial Intelligence @ Allegheny College

Janyl Jumadinova

October 2, 2018
Observations are classified into two or more classes, represented by a response variable $Y$ taking values $1, 2, ..., K$.

We have a feature vector $X = (X_1, X_2, ..., X_p)$, and we hope to build a classification rule $C(X)$ to assign a class label to an individual with feature $X$.

We have a sample of pairs $(y_i, x_i), i = 1, ..., N$. Note that each of the $x_i$ are vectors.
Object Detection/Recognition

Goal:
Find an object of a pre-defined class in a static image or video frame.
Goal:
Find an object of a pre-defined class in a static image or video frame.

Approach:
- Extract certain image features, such as edges, color regions, textures, contours, etc.
- Use some heuristics to find configurations and/or combinations of those features specific to the object of interest.
Process of Object Detection/ Recognition

1. Input image
2. Preprocessing
3. Features: HAAR, HOG, SIFT, SURF
4. Learning Algorithm: SVM, Random Forests, ANN
5. Label Assignment: Cat or Background
Histogram of Oriented Gradients (HoG) Feature Descriptor
Histogram of Oriented Gradients (HoG)
Histogram of Oriented Gradients (HoG)

From Dalal and Triggs paper
Linear classifiers

Find linear function to separate positive and negative examples

\[ x_i \text{ positive: } x_i \cdot w + b \geq 0 \]
\[ x_i \text{ negative: } x_i \cdot w + b < 0 \]
Support Vector Machines (SVMs)

- Discriminative classifier based on optimal separating line (for 2D case)
- Maximize the margin between the positive and negative training examples
Support Vector Machines (SVMs)

- Want line that maximizes the margin.
The **precision** is the ratio \( \frac{tp}{(tp+fp)} \) where \( tp \) is the number of true positives and \( fp \) the number of false positives.

- The precision is intuitively the ability of the classifier not to label as positive a sample that is negative.
Evaluation

- The **precision** is the ratio $\frac{tp}{(tp+fp)}$ where $tp$ is the number of true positives and $fp$ the number of false positives.
  - The precision is intuitively the ability of the classifier not to label as positive a sample that is negative.

- The **recall** is the ratio $\frac{tp}{(tp+fn)}$.
  - The recall is intuitively the ability of the classifier to find all the positive samples.
The F-beta score can be interpreted as a weighted harmonic mean of the precision and recall, where an F-beta score reaches its best value at 1 and worst score at 0.
- The F-beta score weights recall more than precision by a factor of beta. $\beta = 1.0$ means recall and precision are equally important.
Evaluation

- The **F-beta score** can be interpreted as a weighted harmonic mean of the precision and recall, where an F-beta score reaches its best value at 1 and worst score at 0.
- The F-beta score weights recall more than precision by a factor of beta. \( \beta = 1.0 \) means recall and precision are equally important.
- The **support** is the number of occurrences of each class in the correct target values.
Support Vector Machines (SVMs):
- works for linearly separable and linearly inseparable data; works well in a highly dimensional space (text classification)
- inefficient to train; probably not applicable to most industry scale applications

Random Forest:
- handle high dimensional spaces well, as well as the large number of training data; has been shown to outperform others
No Free Lunch Theorem:

Wolpert (1996) showed that in a noise-free scenario where the loss function is the misclassification rate, if one is interested in off-training-set error, then there are no a priori distinctions between learning algorithms. On average, they are all equivalent.
**No Free Lunch Theorem:**
Wolpert (1996) showed that in a noise-free scenario where the loss function is the misclassification rate, if one is interested in off-training-set error, then there are no a priori distinctions between learning algorithms. On average, they are all equivalent.

**Occam’s Razor principle:**
Use the least complicated algorithm that can address your needs and only go for something more complicated if strictly necessary.

“Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?”
1. Represent each example with a single, fixed HoG template

2. Learn a single [linear] SVM as a detector