SMAI Team Project

Team 18

by,
Krishna Tulsyan
Romil Aggarwal

Project Name

Scene Classification using Weak Supervision

Basic Ideas

Datasets

Low Level Representation

Semantic Theme Representation

Theme Vector Generation

Scene Classification

Collection of Datasets

Corel Stock Photo

We use the images from the Corel Stock photo Dataset.

15-scene Category

15 scene categories are picked from MIT 67 Dataset.

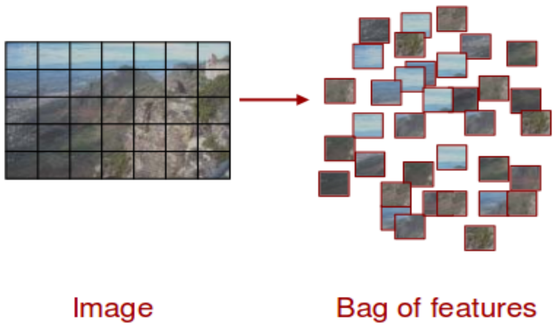
Caltech 101

This dataset contains 101 object categories. We will run this model on this dataset to compare its performance with other object classification algorithms.

Low-Level Representation

Image Representation

At the low level, we represent the images as bags of 8 X 8 vectors of DCT coefficients sampled on a uniform grid. Only 36 (out of 64) top left coefficients are kept.



Scene Categories

Each image ‘Ii’ is labeled with a “Ground Truth” ‘Si’.
The image label ‘Si’ is considered to be an observation from a Semantic scene category ‘S’ defined on {S1, S2, ... , Sk}. This framework of low-level representation is similar to the “Bag-of-Visterms” representation.

Semantic Theme Representation

Concept of Themes

To represent image by semantic themes the dataset is augmented by a vocabulary L = {t1, t2, t3, … , tL} of semantic themes 'ti'. For example an image of "street" can have themes such as "road", "trees", "cars", "people" etc.

Theme Representation

Each image 'Ii' is labeled with a "Ground Truth" 'Si'. The image label 'Si' is considered to be an observation from a Semantic scene category S defined on {S1, S2, ... , Sk}.

Casual Annotation

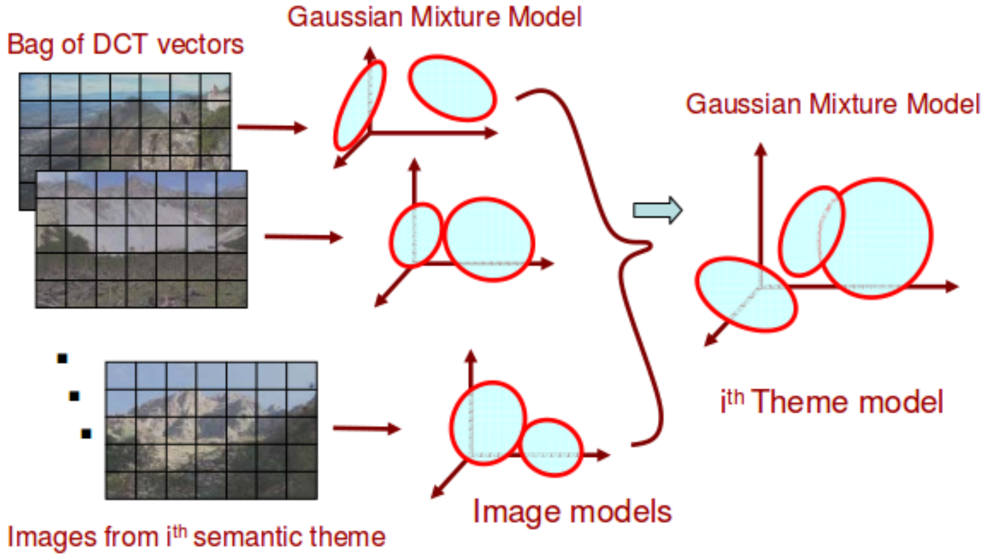
Usually theme annotations are absent in the training dataset. In such cases semantic theme categories serve as a proxy for the theme vocabulary. This is casual annotation and for this paper casual annotation is used.

Semantic Theme Representation

Forming Theme Models

Now for all the images belonging to the same theme, the gaussian mixture model is estimated. Now using EM we estimate the Gaussian mixture model for the complete theme.

Hence we obtain $P_{X|T}$. Using this we will later estimate $P_{T|X}$.



Generating Theme Vector

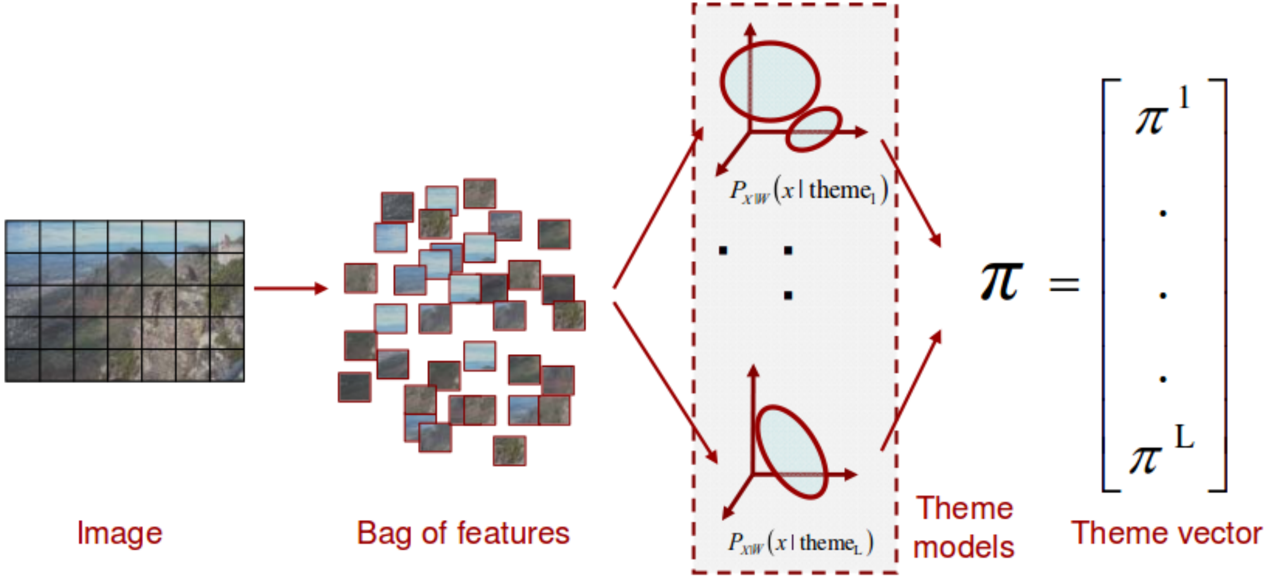
Why Not use Annotations?

Annotations seem to be good parameter for using as a feature vector. But we don't use it because usually it's only available during training and if an image is not annotated with a particular theme then it doesn't necessarily mean that it doesn't have that theme.

Procedure

We will use theme vector as the feature vector of an image for learning. Feature vector 'Π' = $\{\pi_1, \pi_2 \ldots\ldots, \pi_L\}$ is a L-dimensional vector. Here 'π_i' represents the probability that the image feature vector is drawn from the i^{th} theme. To calculate this we use Bayes rule as we already know $P_{X|T}$ and need $P_{T|X}$.

Generating Theme Vector



Here is the pictorial representation of the procedure.

Scene Classification

Procedure

Till the previous step we learned the theme vector (π) for each image.

Now we will use a multi-class SVM using one-vs-all strategy with Gaussian kernel for classification, with the parameters obtained by 3-fold cross validation.

Deliverables

Create a model that is able to classify images based on its semantic theme.

Also provide the theme vectors for each image which captures the different semantic meanings of the image.

Timeline

6/10/2015
Understand the paper and creating the bag of DCT vectors.

20/10/2015
Learning gaussian mixture model using EM algorithm.

27/10/2015
Computing the theme vectors for each image.

6/11/2015
Use SVM to classify images and deliver the final deliverables.

Thank You

Any Questions... ?