Poselets: Body Part Detectors Trained Using 3D Human Pose Annotations

LUBOMIR BOURDEV AND JITENDRA MALIK
Outline

H3D dataset
Pipeline
Analysis of Poselets fired
Selective parts – torso, legs and face
Other cases – Clutter, Rotation and Occlusion
Analysis of Hough Transform
Conclusion
Outline

SVM 1
SVM 2
SVM 3
SVM n

Hough Transform

Localized object
H3D dataset

Keypoint Annotations

3D Pose

Region Labels
Given an image, use SVM's trained for ~300 poselets to get poselet activations
Poselet Activation Clusters

Using the H3D training set we fit the transformation from the poselet location to the object. Cluster the hypothesis using KL divergence.
Poselet Activations

Run each poselet detector at every position and scale of the input image, collect all hits and use mean shift to cluster nearby hits.
Find peaks in Hough space by clustering the cast votes using agglomerative clustering and compute the sum over the poselets within each cluster.
Object Hits

All the clusters in terms of image patches
Poselets

Poselet Activations for the best match

Poselet Activations for the last matches
Experiment Setup

Available code - takes an image and draws the bounding box on the subject

Uses a pretrained model for poselets which is used to fire on images and generate hypothesis from 3-D space to 2-D space

Uses a pretrained model for weights of different poselets which is used to combine the probability of object location corresponding to the poselet
Test Cases

Good localization examples
Different poselets which are activated
Change in subject conditions
Training Data and Analysis of Hough transform space
What works

Good quality of bounds on the subject
High score – support from a good number of poselets
Poselets corresponding to head and whole body
Different scales
Part poselets

Poselets when only certain part of body is seen in the image

Poselets corresponding to the part should contribute the most towards the score
Face Poselets
Torso Poselets
Lower body Poselets

Best match

Examples of the selected poselet
Lower body Poselets

Second Best match
Image Conditions

Look at the performance of poselets in presence of different image conditions like Clutter, Rotation and occlusion
Clutter

Good detection in presence of clutter. Poselets corresponding to lower body and the whole body contribute the most in localization.
Extreme Rotation

Best match – incorrect localization
Poselets corresponding to face fired on this

False positives – Decent localization but votes from incorrect poselets
Occlusion

Tenth match with score = 0.42
Highest Match = 0.82
Analysis of Hough Transform

Look at the peaks generated in the Hough space

Each peak corresponds to an image patch localizing the object

Votes from poselets for the image patch vote for the plausible object location

Votes in Hough space clustered using agglomerative clustering
Analysis of Hough transform
Poselet activation with highest score = 0.18
Poselet activations which would lead to good localizations with score \( \sim 0 \).
Limited Training Data?
Score = 1.10

Though the score of best match is low, none of the poselets fired are on the subject. Instead objects are detected in the background.
Training Data

~1500 annotated images

Many images have people upright or facing the camera

The limitations in previous slides can be solved by adding more training data for different postures where poselets other than face, whole body and legs are fired

Difficult to generate annotated data?
Conclusion

Current methods like R-CNN perform exceptionally well for person category compared to poselets

If we take into account the amount of training data used then poselets fares well

However from experiments though the image patch obtained is of considerable quality the poselet activations corresponding to the patch is not right in terms of the structure, scale and orientation in many cases
References

Poselets: Body Part Detectors Trained Using 3D Human Pose Annotations - Lubomir Bourdev and Jitendra Malik

Rich feature hierarchies for accurate object detection and semantic segmentation - Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik