Demo: Probability Maps for Modeling Relative Location

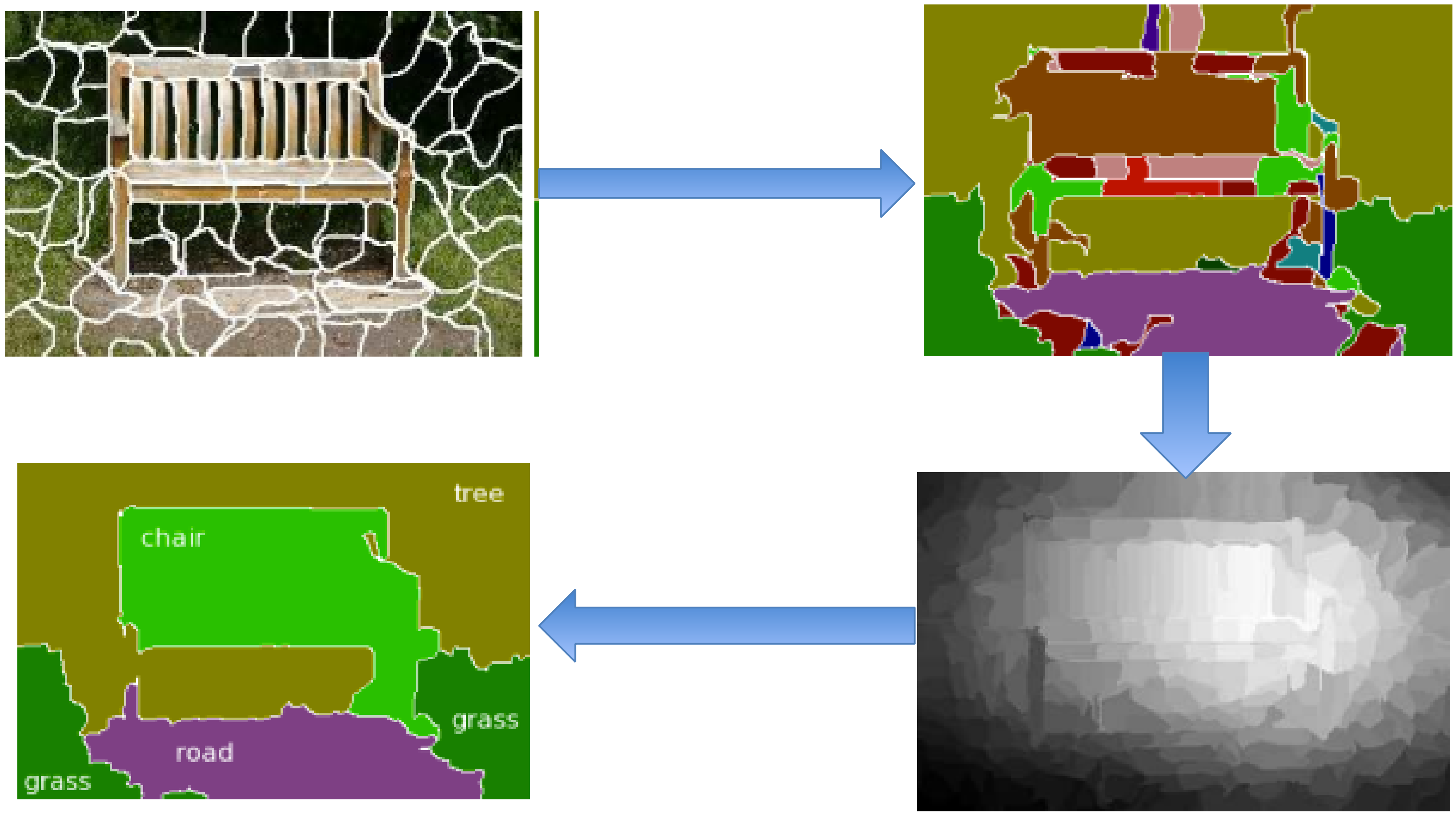
David J. Garcia

October 5, 2012 / CS395T Demo

High Level Idea: Relative Location vs. Absolute Location

- Global spatial information can aid in multi-class segmentation.
- Instead of absolute location, relative location is used in the construction of "learned relative location maps."
- Code from Gould [1].
- Data/Images from MSRC [2].
- -[1] http://ai.stanford.edu/~sgould/svl/
- -[2] http://ai.stanford.edu/~sgould/svl/media/msrcsegexample.tar.gz

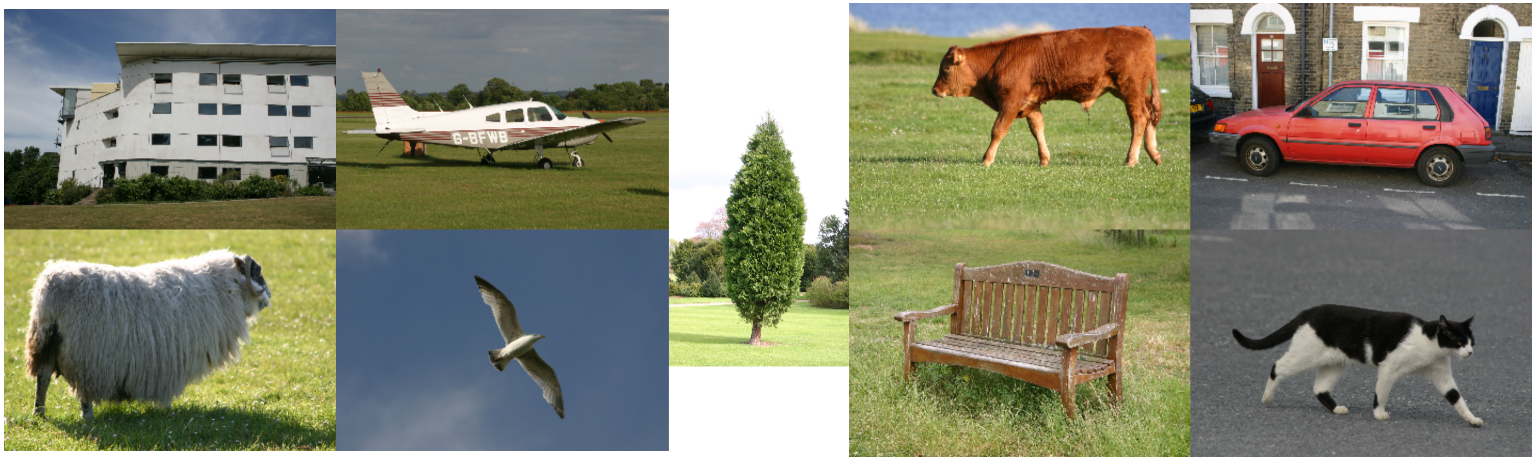
Pipeline: From Superpixels to Classified Segments



- - [1] S. Gould, J. Rodgers, D. Cohen, G. Elidan, and D. Koller, “Multi-Class Segmentation with Relative Location Prior,” IJCV, 2007

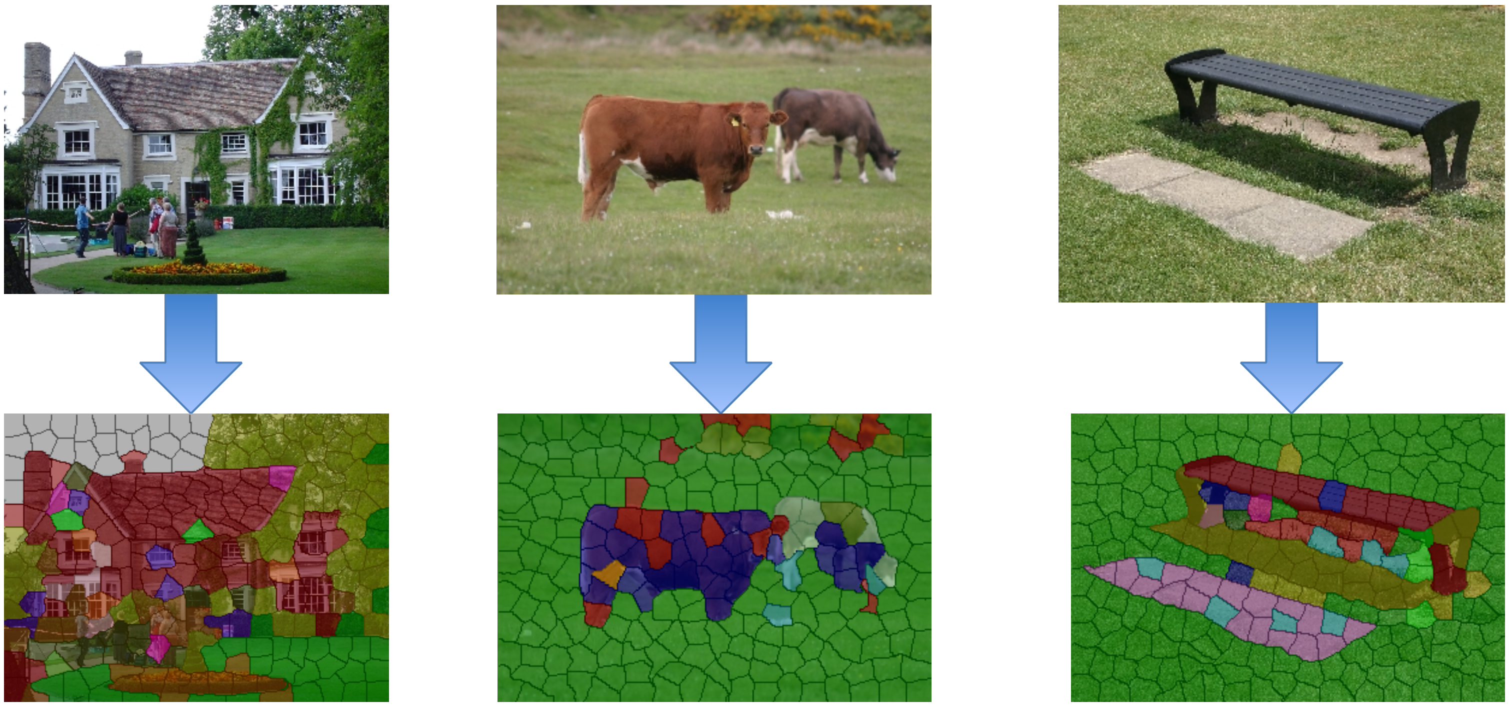
Example Results from Testing…

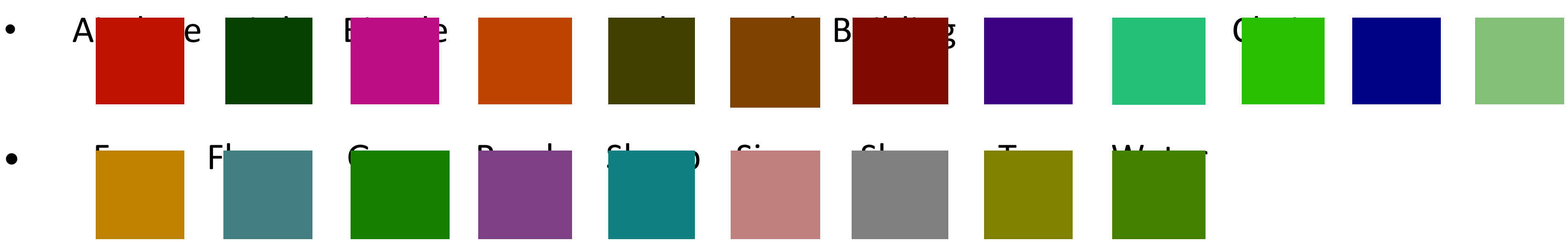
- The categories are: building, grass, tree, cow, sheep, sky, airplane, water, face, car, bicycle, flower, sign, bird, book, chair, road, cat, dog, body, and boat.
- The models were trained with a random sampling of half of the total number of images. Testing results are from the other half.
- Here are some examples input images from the dataset:



Example Results from Testing…

- Here are a few results from six categories (using the Author's settings):





Segmentation for the Experiments

- In their work, the authors used mean shift segmentation (not specified in the paper). The strategy replaces each pixel value with an average of pixel values around it for some r-neighborhood; neighborhood pixel are discriminated based upon an additional threshold d.

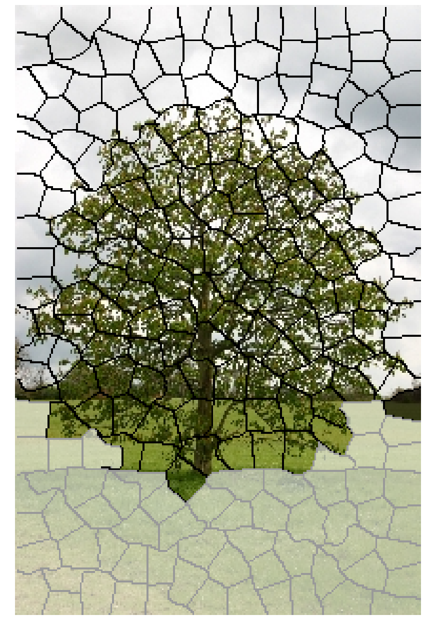
- Unfortunately, it was not known what parameters were used in the segmentation. So, a variety of segmentations will be tested with the vlfeat quickseg function. As an illustration, consider these two segmentations:

- Notice that the right image's lettering is more visible. The d threshold in the left was set to 20, and in the right, it was set to 10.

Segmentation Size Experiment

- The number of segments, in an image, is dependent upon both the segmentation settings and the image content. Each of these images has a different number of segments.

(b) Superpixel over-segmentation

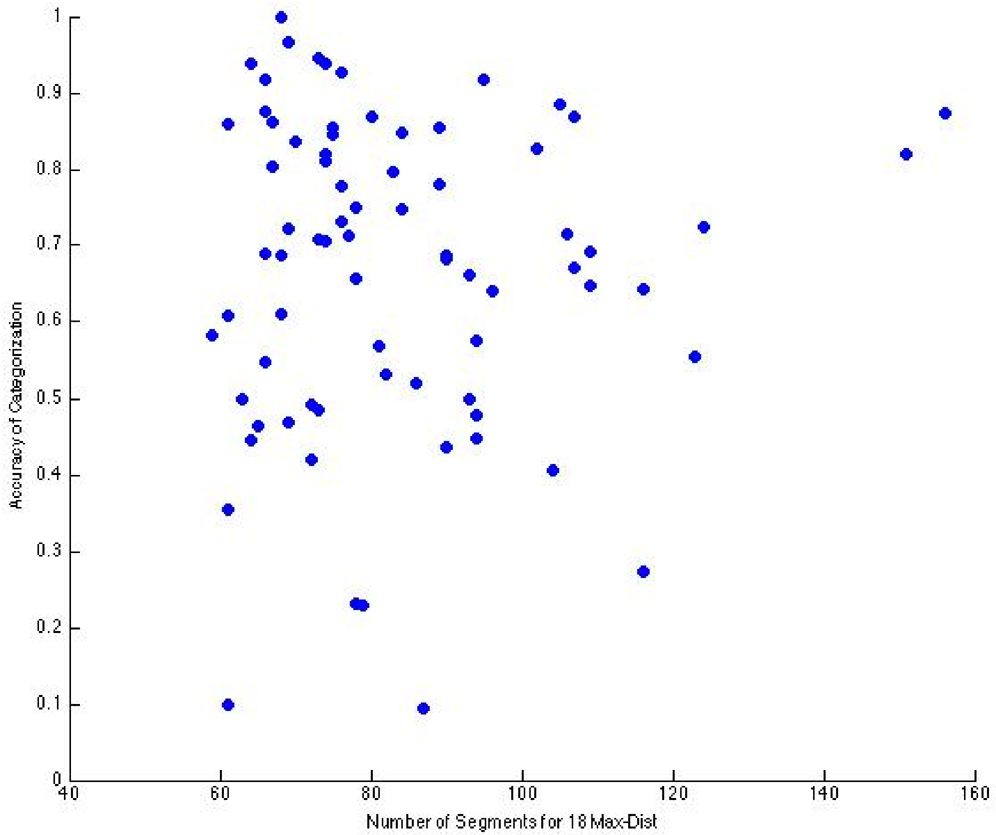


- It might be an interesting experiment to observe the method's performance with respect to the number of segments.

- - [1] S. Gould, J. Rodgers, D. Cohen, G. Elidan, and D. Koller, "Multi-Class Segmentation with Relative Location Prior," IJCV, 2007

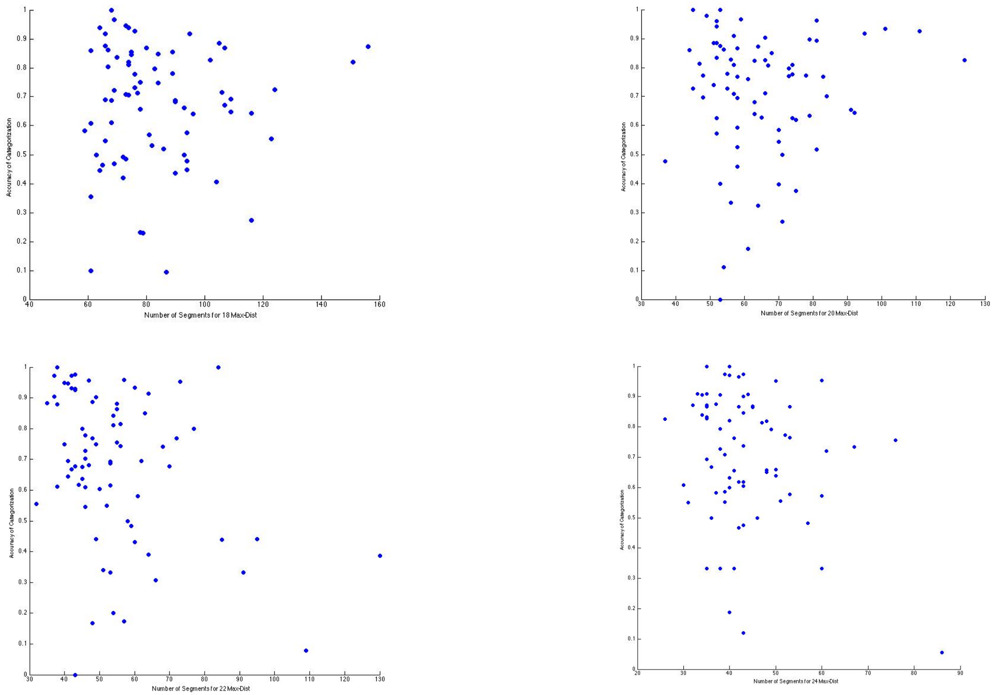
Segmentation Size Experiment

- The experiment was performed with a variety of segmentation settings. This plot was formed from a random sampling of 75 test-points at Max-Dist=18.



Segmentation Size Experiment

- The experiment was done for Max-Dist sizes: 18, 20, 22, and 24:



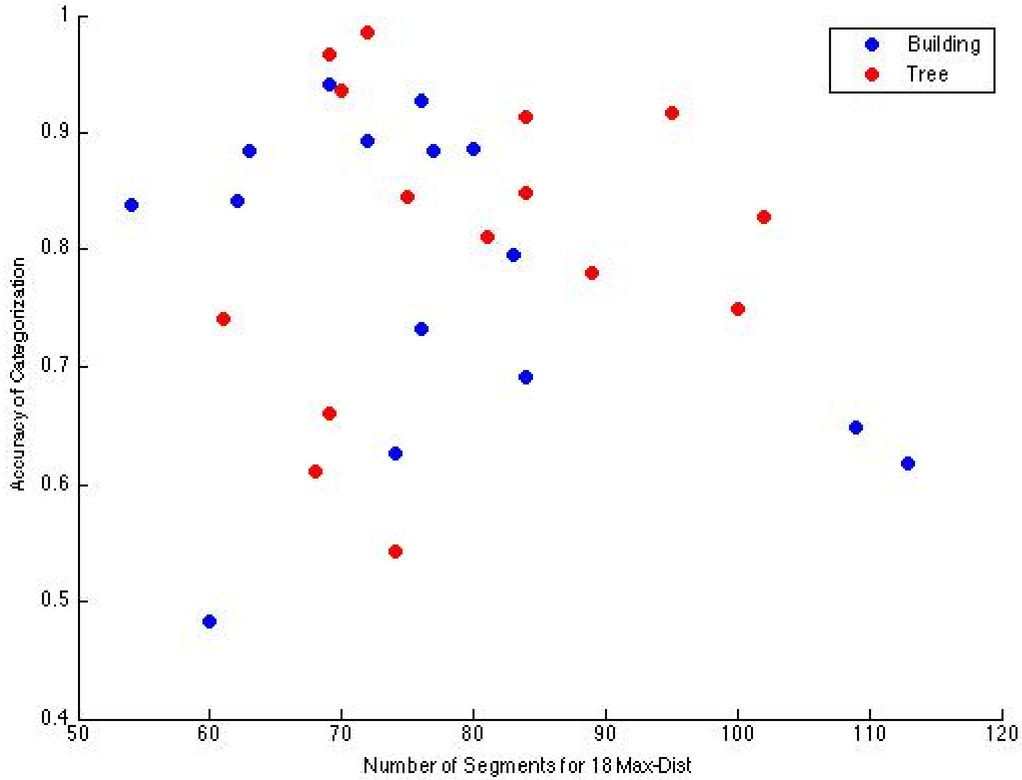
- The range of segmentation-numbers that score relatively well shift to the left with every larger value for Max-Dist. This is not really surprising since the input drives the output.

2D vs 3D Experiment

- Another experiment was done to compare how classification compared against objects that make “more” sense in a 2D context versus objects that might need the aid of a 3D context.
- As a test case, test results from “planes” and “bodies” were compared with “trees” and “buildings”.
- Planes are typically displaced well away from ANY objects for safety reasons. Any objects in the image frame will most likely be “behind” the plane.
- People, especially when “posing” in pictures, might tend to be “in front” of interesting objects. (i.e. tourists, family photos in front of the oceans, buildings…etc).
- Pictures of Buildings, on the other hand, might have plenty of 2D context (buildings next to other buildings as well as streets). Likewise, trees are likely to be next to other trees and above the ground (few trees grow out of the street).
-  And  VS.  And 

2D vs 3D Experiment

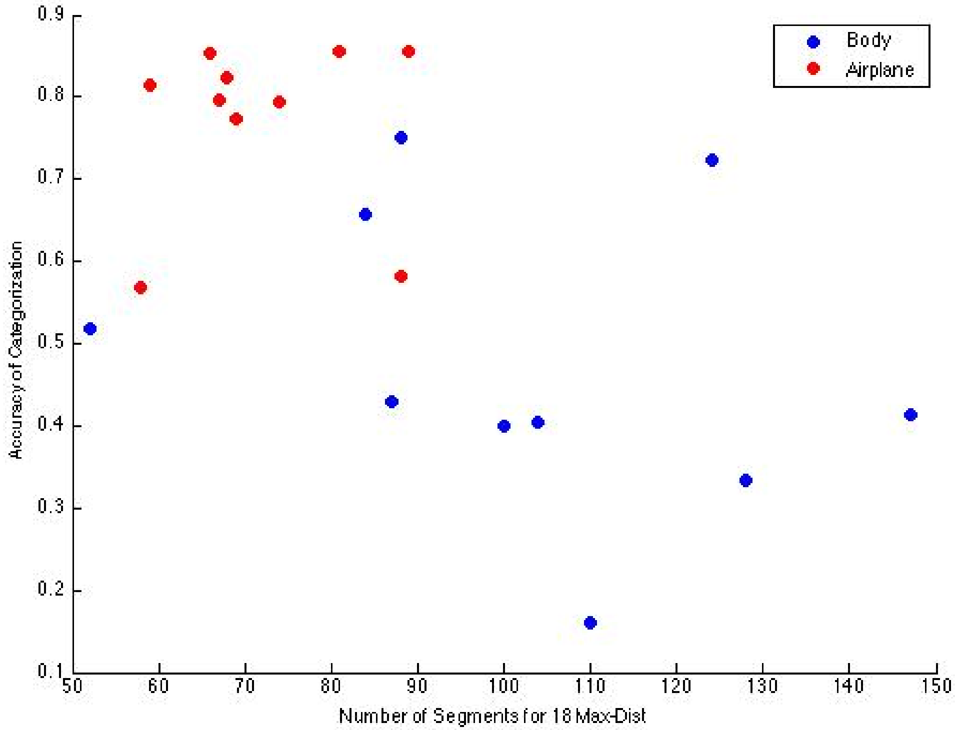
- “Tree” and “Building” are compared:



- This result is expected. Both sets of points are mixed with each other WRT accuracy.

2D vs 3D Experiment

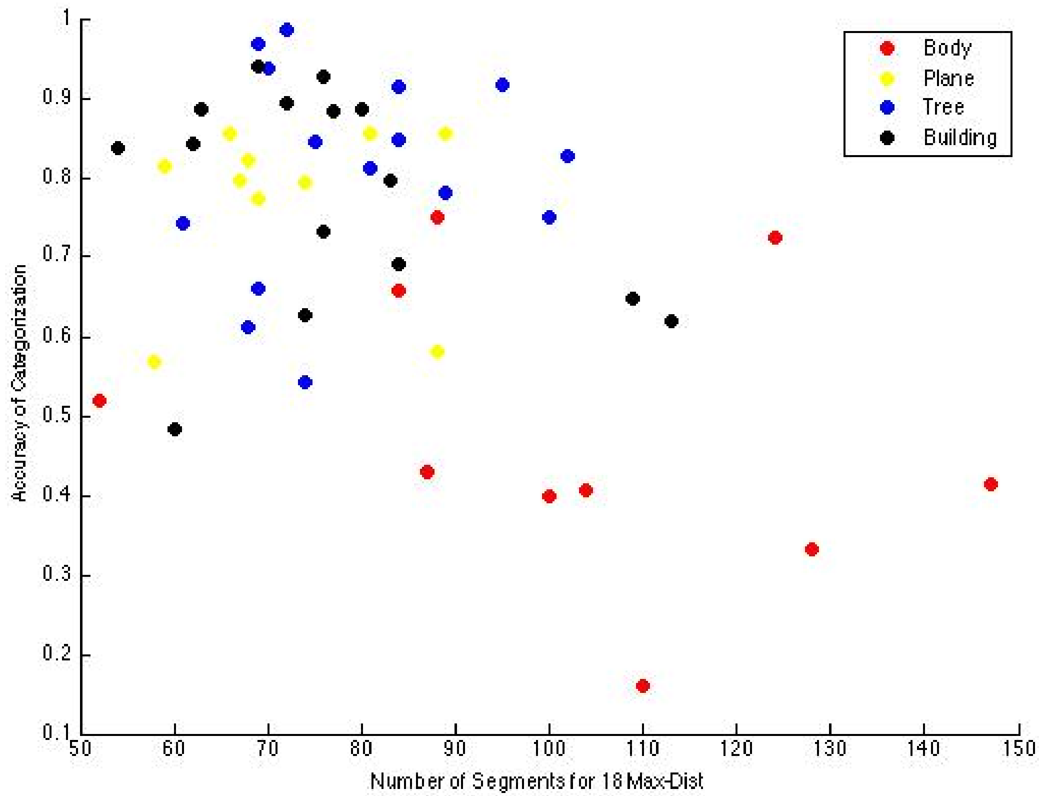
- "Body" and "Plane" are compared:



- With the exception of 3 Body-points, the two sets are linearly separable at just above 0.5 accuracy. Perhaps there is another relationship that should be tested.

2D vs 3D Experiment

- Here are all four test-sets overlaid in one scatter plot.

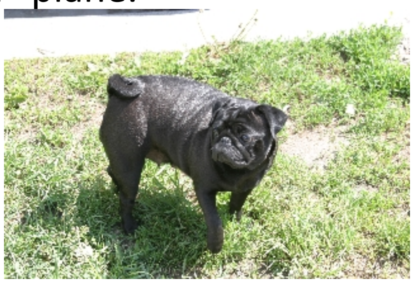


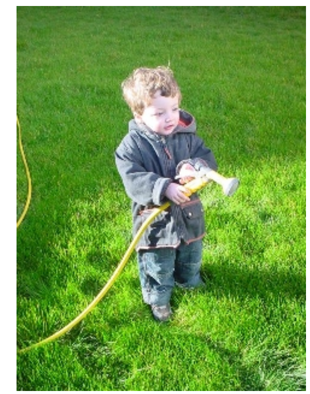
- Again, with the exception of three body-points and one building-point, the body-points are linearly separable around 0.5 accuracy. “Plane”, “Tree”, and “Building” might have something in common.

Deformable Parts Experiment

- In the dataset, there are several categories that can be considered deformable (i.e. dog, cat, body…etc). There are also numerous non-deformable categories (i.e. tree, building, plane…etc).

- The categories “dog”, “cat”, and “body” are compared against “tree”, “building” , and “plane.”

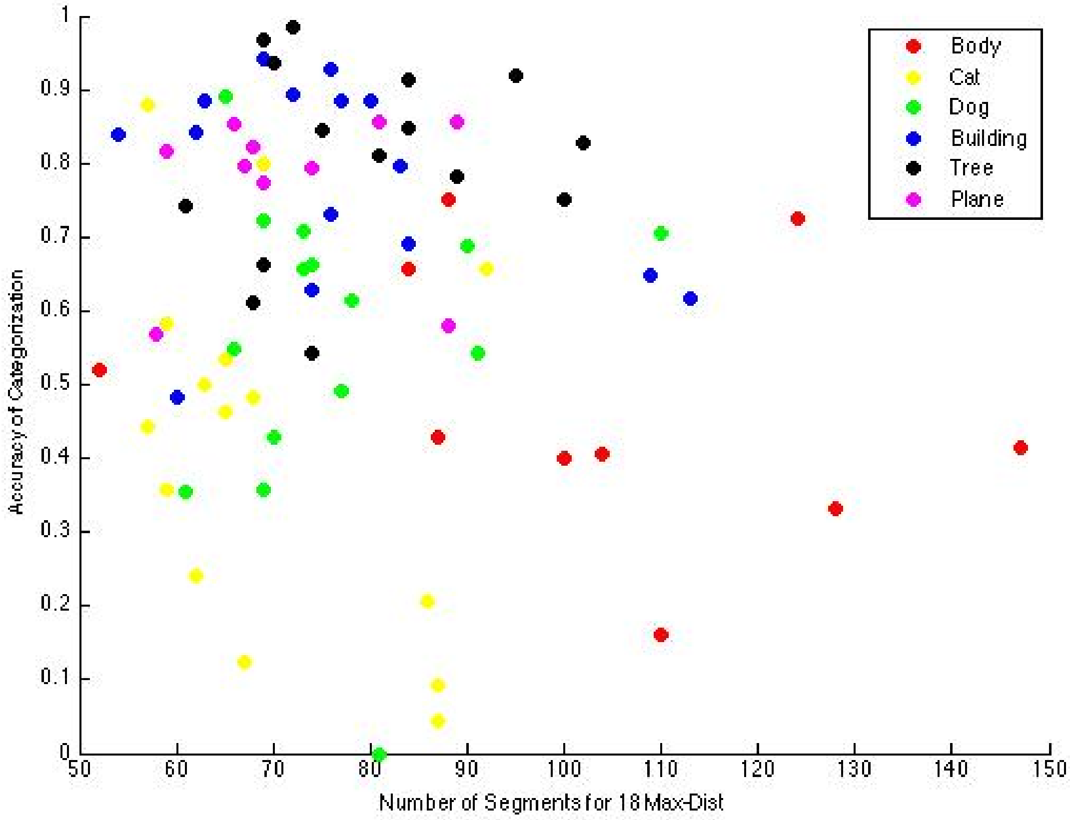






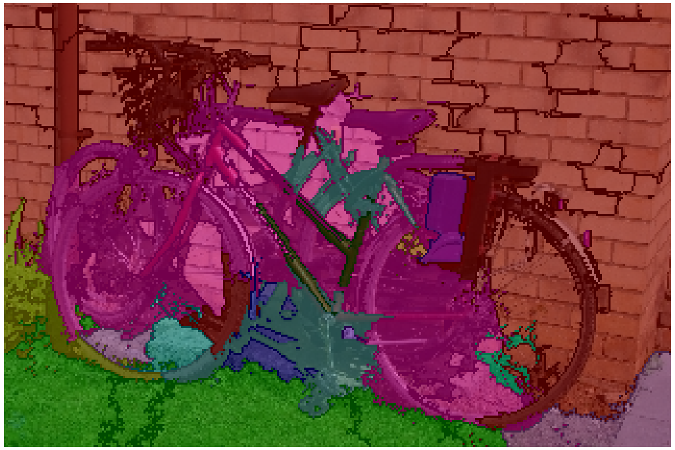

Deformable Parts Experiment

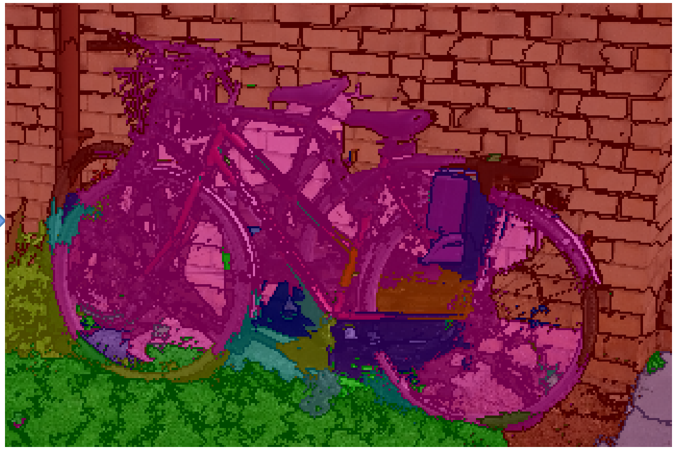
- This bifurcation might be a little better. The “Cat” class has almost as much trouble as the “Body class. The dog class seems a bit better than either.



Category Performance Across Segmentation Size Experiment

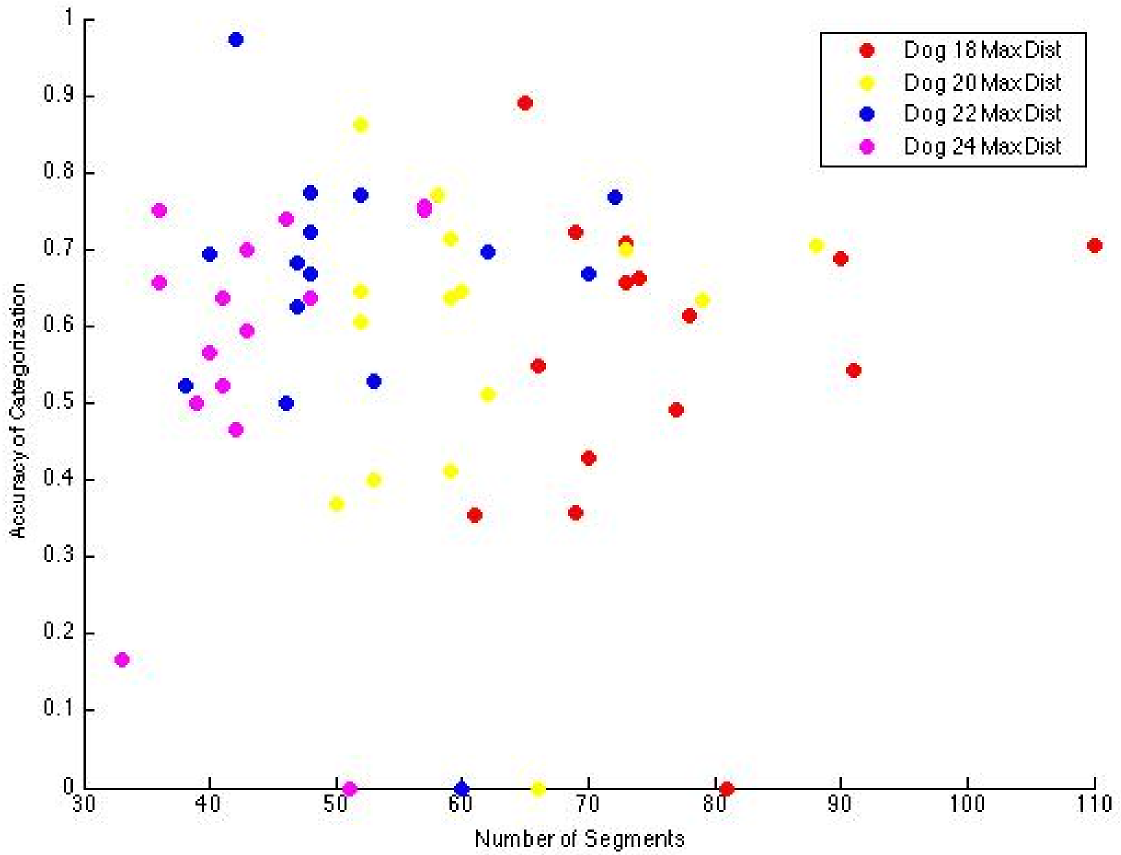
- How does a particular category perform across different segmentation sizes? This experiment was done over “Dog”, “Cat”, and Bicycle. “Bicycle” is unlike either “Cat” and “Dog” in that it’s structure is less contiguous. It’s possible that this type of structure is more amenable to a smaller segmentation preprocess.





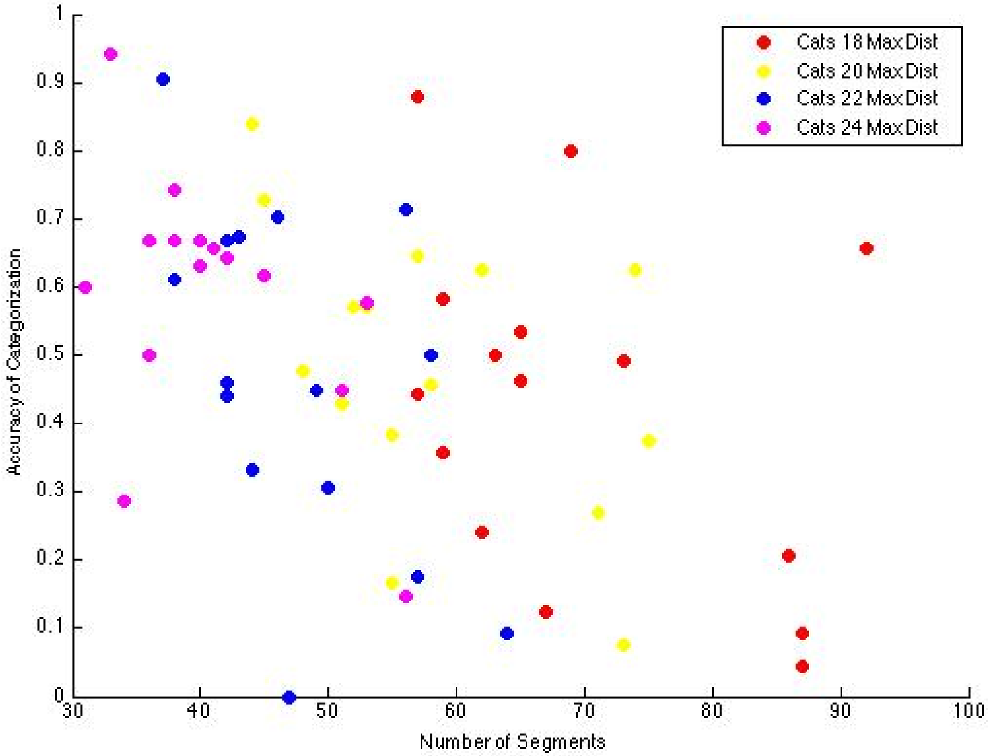
Category Performance Across Segmentation Size Experiment

- Here is Dog varied over 4 different segmentation inputs:



Category Performance Across Segmentation Size Experiment

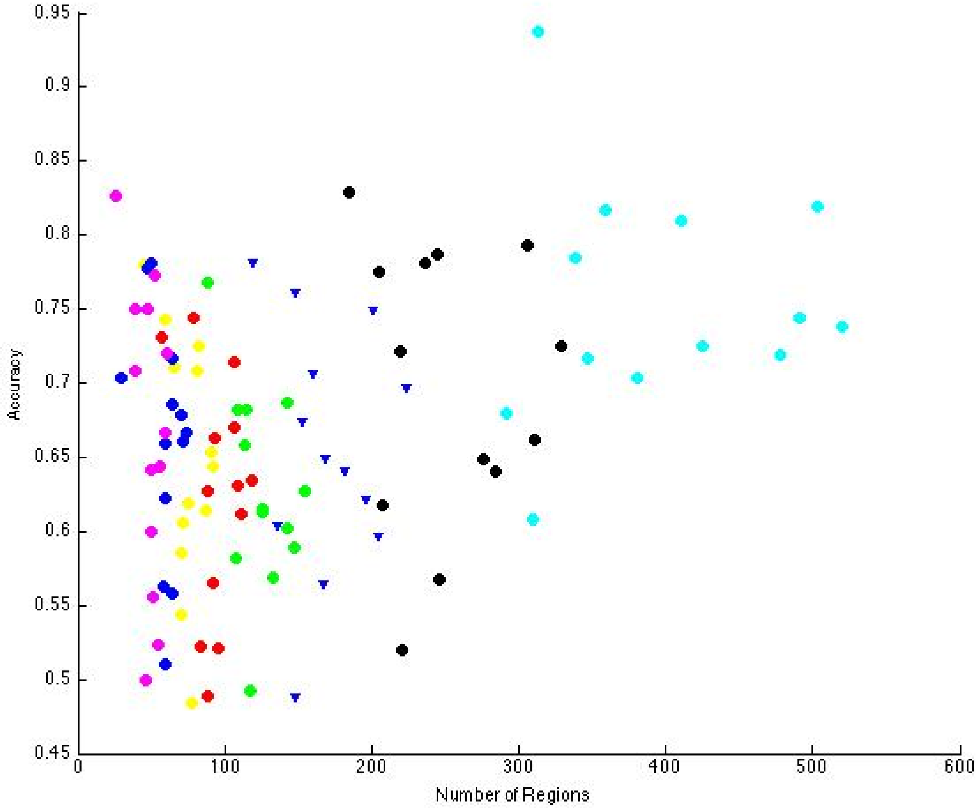
- Here is Cat varied over 4 different segmentation inputs:



Cat seems to have a bit more downward sloping tendency as segments increase in number.

Category Performance Across Segmentation Size Experiment

- The Bicycle shows the opposite trend of Cat:



As the number of segments increases, the data seems to trend to better accuracy. It also seems to vary more with respect to the number of segments.

Category Performance Across Segmentation Size Experiment

- A note on the Standard deviation and Mean of the “Bicycle Points”:

- The points seem to be grouping closer together in the accuracy dimension as they trend toward a greater number of segments.

- Standard deviation of accuracy for the Max-Dist 24 set is 0.1005
- Mean for the Max-Dist 24 set is 0.6661
- Standard deviation of accuracy for the Max-Dist 10 set is 0.0812
- Mean for the Max-Dist 10 set is 0.7541

Perhaps Increasing the number of segments in the preprocessing step leads to increased categorization accuracy over a variety of Bicycle images.

It would be interesting to investigate the false positive rate of bicycles in the other image sets.

The STAIR Vision Library: Good Documentation…No Package Management

- There are some binaries available, but only in 32-bit.   UT labs are 64-bit.
- If using Linux, highly recommend the Ubuntu distribution.
- Use package management for gtk_2.0_dev, opencv.
- If on 64-bit, project needs to be downloaded from: https://stairvision.svn.sourceforge.net/svnroot/stairvision/trunk
- Project needs to be built, but externals may not build correctly…use managed libraries.
- Symbolically link to the opencv libraries from the <install_dir>/externals/opencv/lib directory.
- Run make from <install_dir>
- IMPORTANT: Sometimes build fails silently(i.e. binaries don't work) (due to improperly built libraries in externals directory or bad symlinks)!! Run make clean and rebuild after fixing externals.

The STAIR Vision Library: Additional Help

- Step by step guidance available at:
  - http://ai.stanford.edu/~sgould/svl/howto/svlMultiClassSegmentation.pdf
  - Commands have changed in the latest version. Use the –help option for latest command usage.