

Image Classifier: Status Update

- Pixel Dawgs

Data Collection

- Data set collected from Flickr
- Contained text files which had image IDs, tags, and their URLs
- Parsed files and downloaded ~300,000 images
- Associated each image with its given tags and stored that separately as another text file (image IDs stored as indices and tags stored as values)

Data Analysis

- The current test for each image is that each image is not from an unknown set of images
- Generated confusion matrix for each image
- Used histogram back projection to track "probable" images in each category
- This generated a resulting image showing the probability of each pixel belonging to the target histogram

Future Work

- Image Segmentation (offering an image into separate parts for improvement)
- Improved algorithm to allow clustering with the image recognition system
- Increase of assigning tags to the relevant image, instead of using tags as individual tags
- More assignment of tags to the image will need to be done manually

Data Collection

- Data set collected from Flickr
- Contained text files which had image IDs, tags, and their URLs
- Parsed files and downloaded ~300,000 images
- Associated each image with its given tags and stored that separately as another text file (image IDs stored as indices and tags stored as values)

Data Analysis

- The source text file was reformatted so that each tag could have an associated list of images
- Computed composite histograms for each image
- Used histogram back projection to "back project" images onto composite histograms
- This generated a resulting image showing the probability of each pixel belonging to the tagged histogram

Future Work

- Image Segmentation (splitting an image into layers) needs to be implemented
- Selected algorithm is K-Means clustering with histogram recombination analysis
- Instead of assigning tags to the whole image, we will assign tags to individual layers
- Initial assignment of tags to the layers will need to be done manually