# Image categorisation II

Jonathan Windle & Robert Levett

# K-Nearest Neighbours

K-NN is a simple classifier.

No pre-training required. Classification on the fly.

Difference distances used:

- Euclidean (I2)
- Manhattan (I1)

Size of K can affect the classification accuracy.

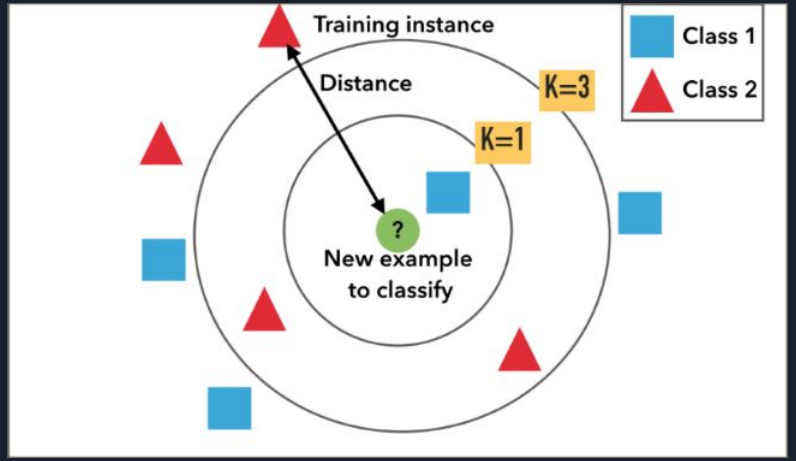



Fig 1. Example of k-NN decision boundaries.
http://blog.xnextcon.com/?p=213

# Support Vector Machines

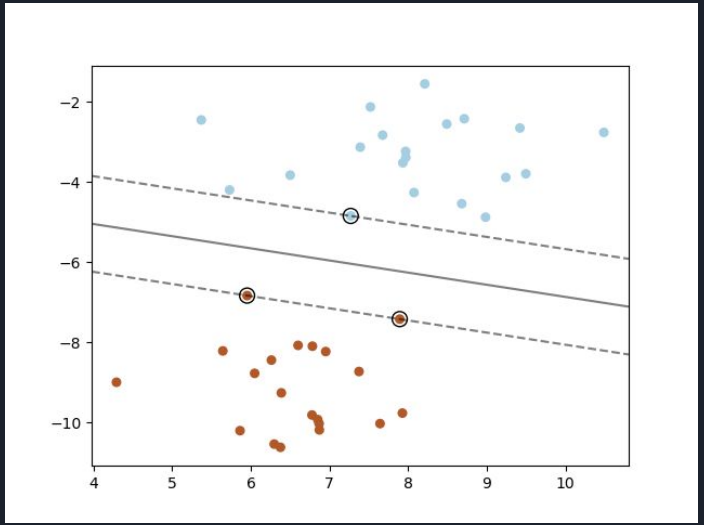



Fig 2. Example of SVM boundaries
https://scikit-learn.org/stable/auto_examples/svm/plot_separating_hyperplane.html

Requires pre-training.

Finds the closest instances from each class (based on a distance metric).

Draws a line of best fit perpendicular to the line drawn between the other classes.

Is one vs all; so finds the distances from every point to every other point, then draws the decision line perpendicular to the two closest points of each class.

The further from the margin, the higher the classification confidence.

# CNNs/Artificial Neural Networks

CNNs: Extract features using convolution and feeds into ANN

Applies weights to input and neurons contain activation function

Uses backpropagation to change weights and therefore activations

Transfer Learning

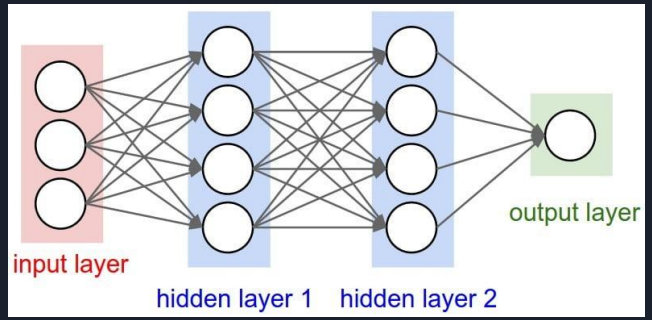



Fig.3 Artificial Neural Network example.
https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/

# Vocabulary/Bag of Words

Formed by extracting SIFT features for an image.

Repeated across entire image training set until vocabulary is fully formed and sift descriptors for each feature point have been extracted.

K-means clustering is then used to form a vocabulary that represents the most common key points within an image, and the vocabulary words are defined as the center of the learned clusters.

The number of clusters corresponds to the size of the vocabulary.

This gives a visual vocabulary, that can be used to form histograms when an image features are detected and matched to the vocabulary.

# Spatial Information

When L = 0, it is just a standard bag of features.

The image is split into $2^{L*2}$ sections at each level.

Increasing the resolution too far becomes redundant as the weight of each feature becomes much lower.

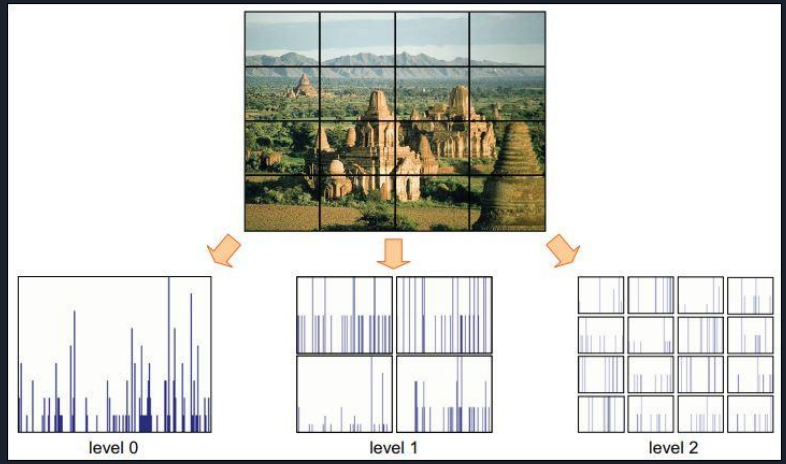



Fig.5 Example of constructing a three level pyramid.

Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories, Lazebnik, Schmid, Ponce.

Spatial features can help define an image category. I.e. if ‘sky’ features found in the upper sections, and green ‘tree’ features found in lower section, it’s likely a forest.

Colour information is calculated for each channel then concatenated together.

# SIFT features

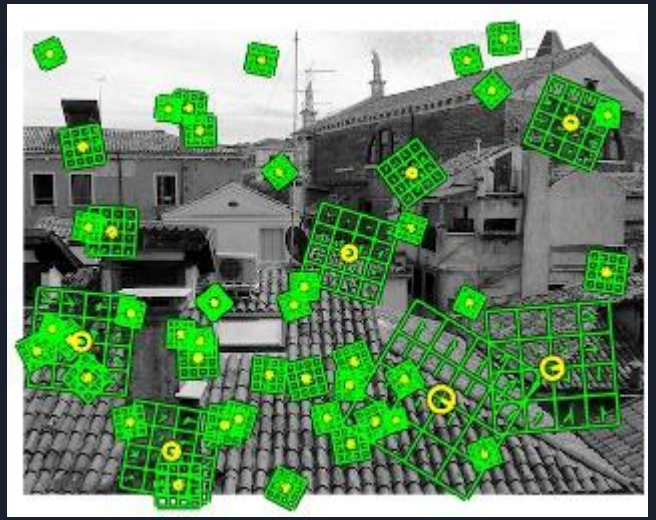


Fig.4 Example of SIFT feature selection.
http://www.vlfeat.org/overview/sift.html

Used to find points of interest from an image.

Yellow dots are points of interest, green panes are SIFT features.

SIFT features are scale & rotation invariant, so similar features in different images should be detected.

Gaussian smoothing and the Difference of Gaussians (DoG) are calculated.

Edge points are removed.

A histogram is formed by quantising the orientation into n bins.

# Bag of SIFT features

Extract features for each image.

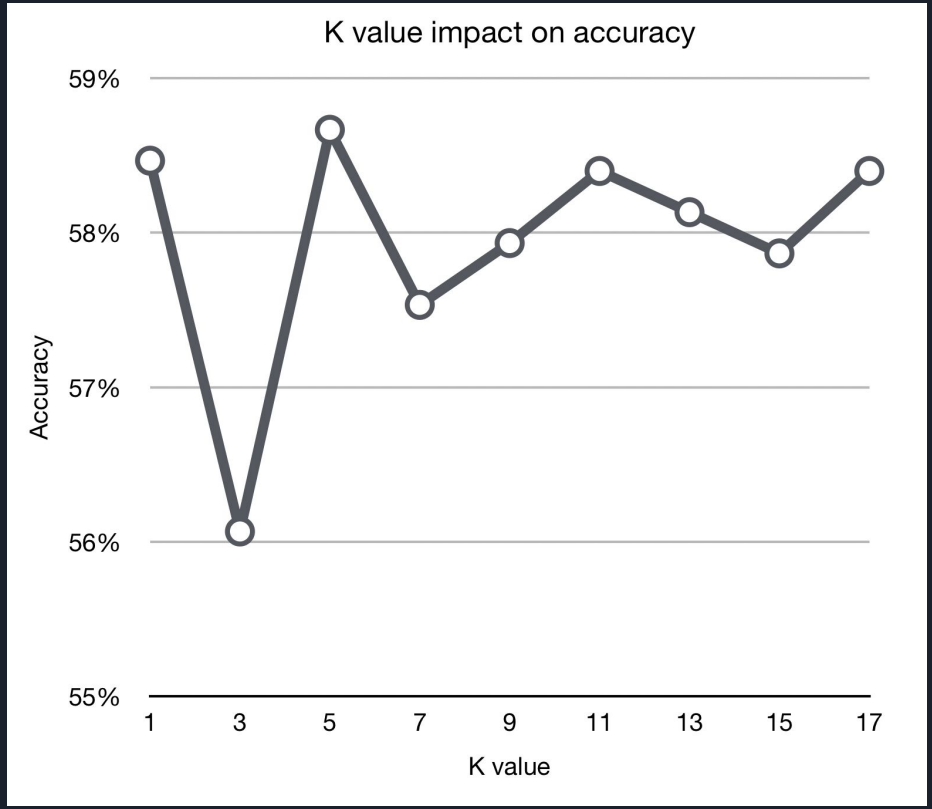
Get the SIFT descriptor for each feature point.

Match the feature descriptors with the previously built vocabulary, and increment the most appropriate feature bin.

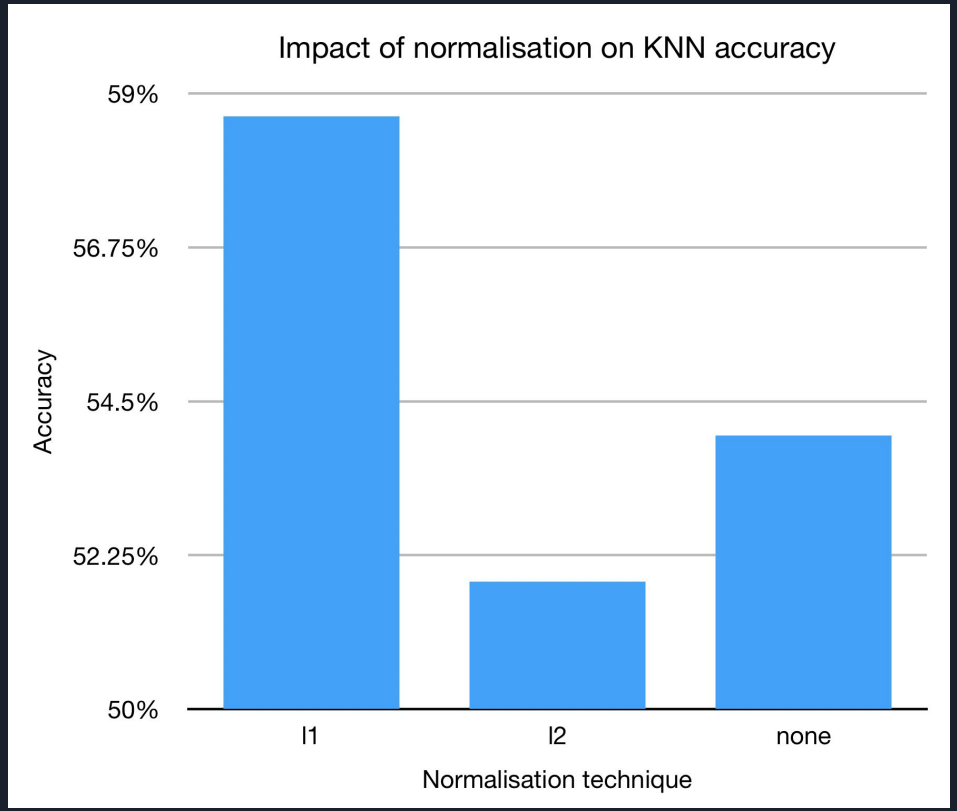
This is done by finding the closest feature in the vocabulary, then incrementing a histogram at that point.

The histogram can then be used to classify an image.

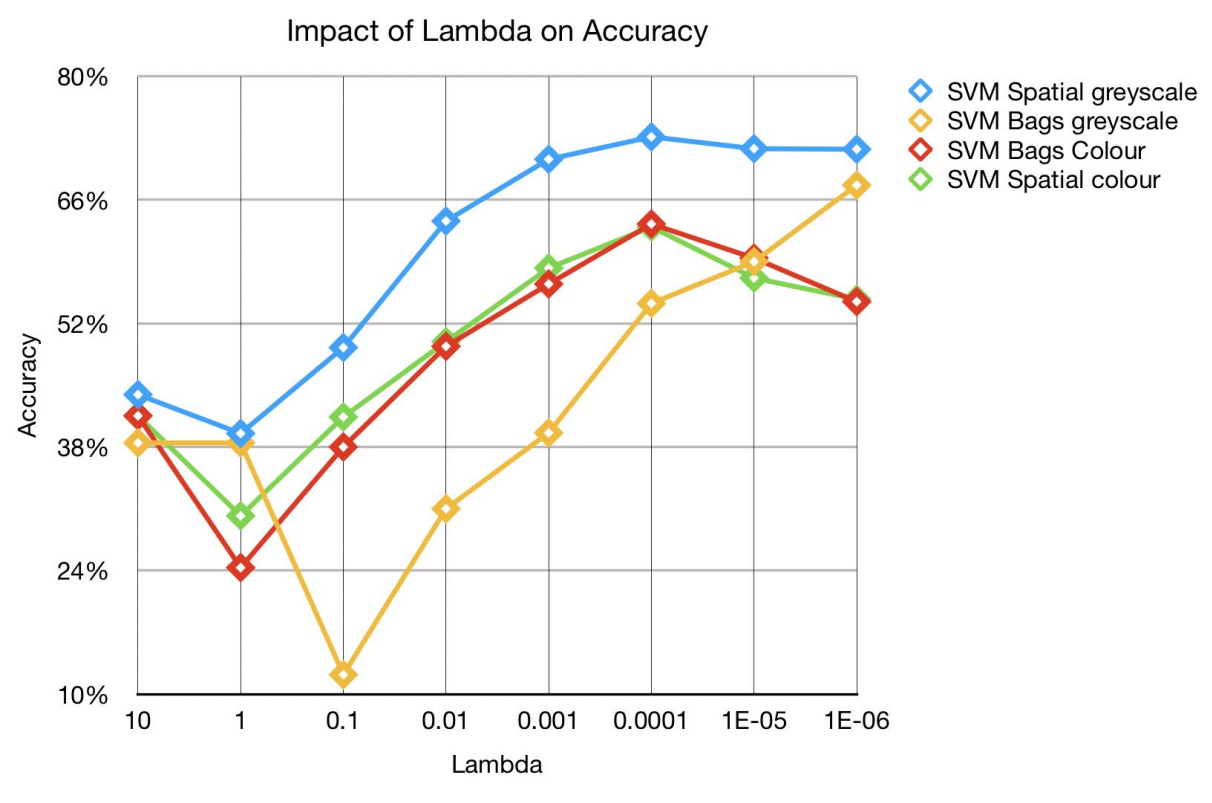
# K Value impact



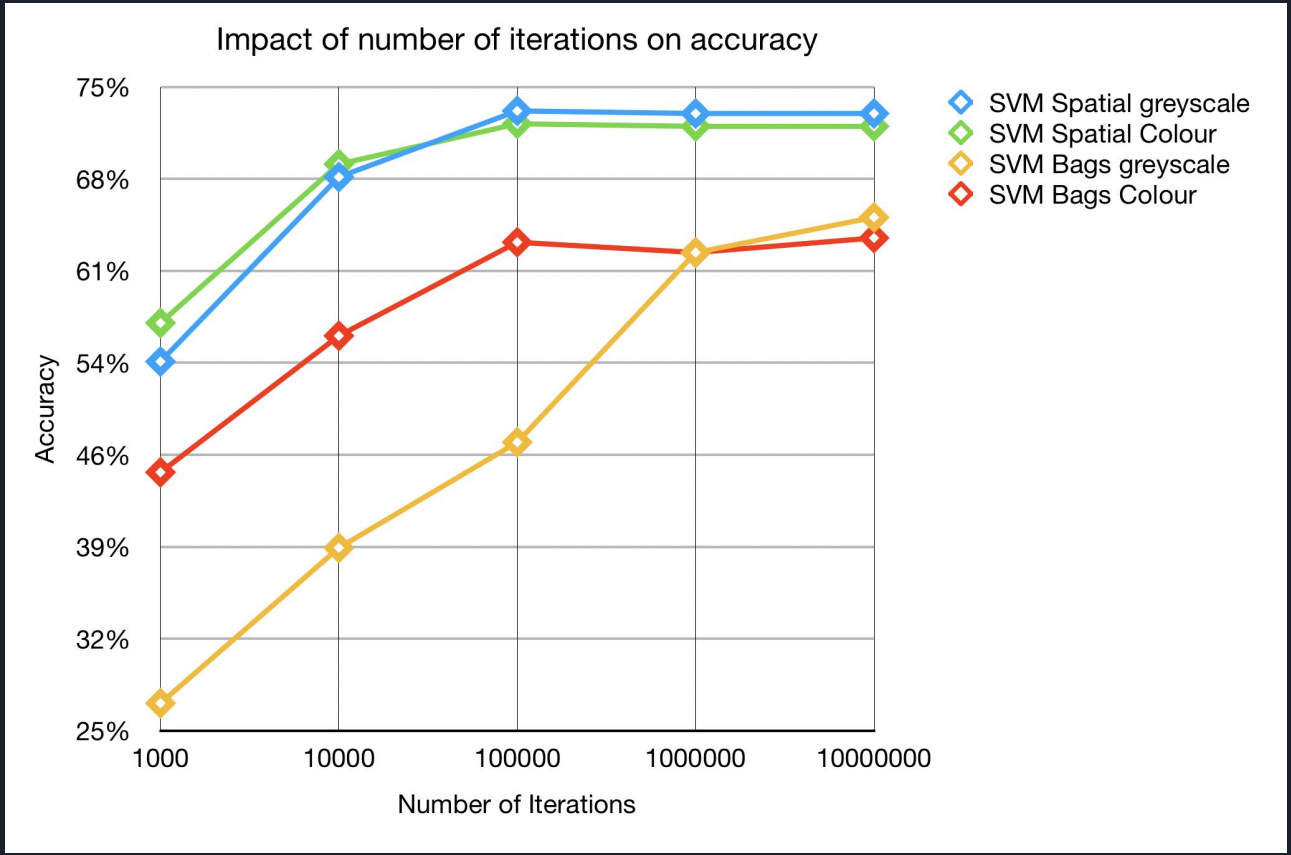
# Normalisation technique



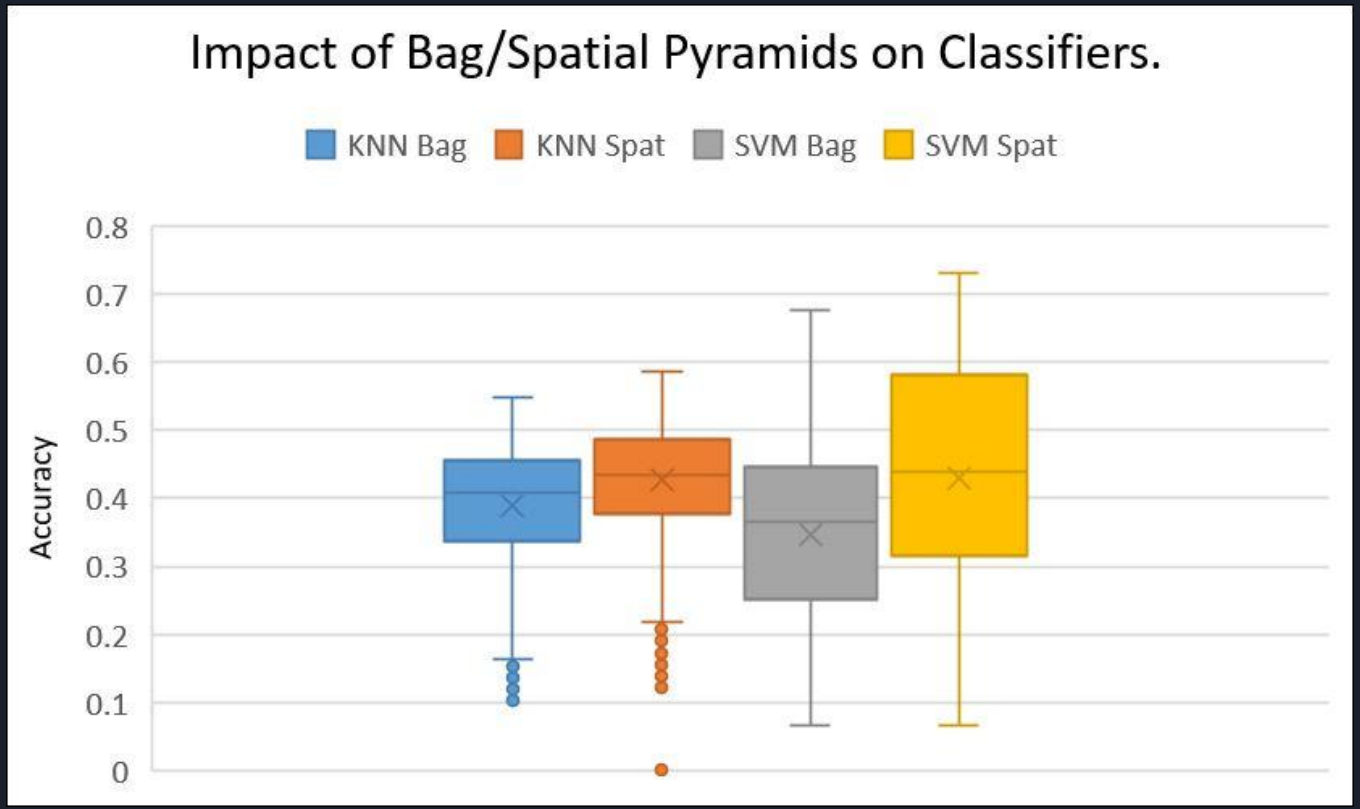
# Impact of Lambda



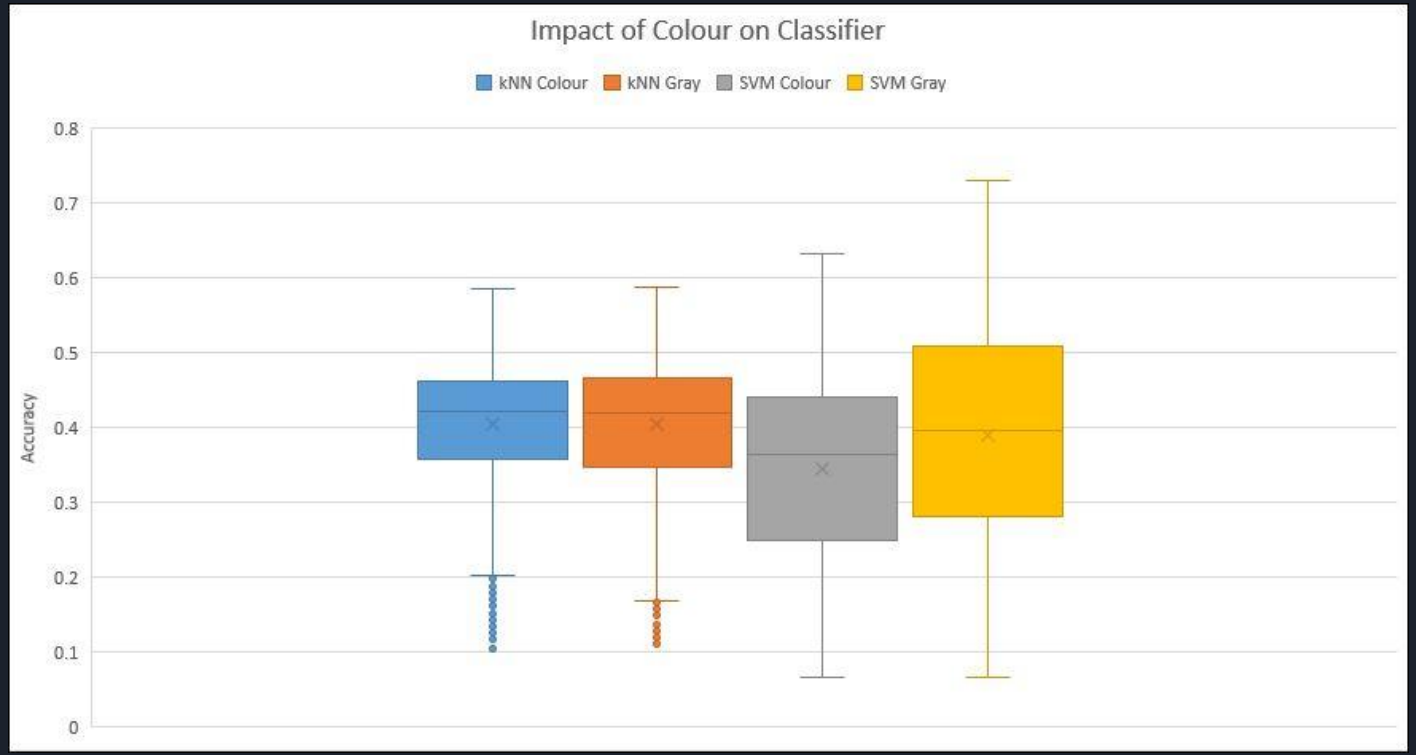
# Impact of Number of iterations



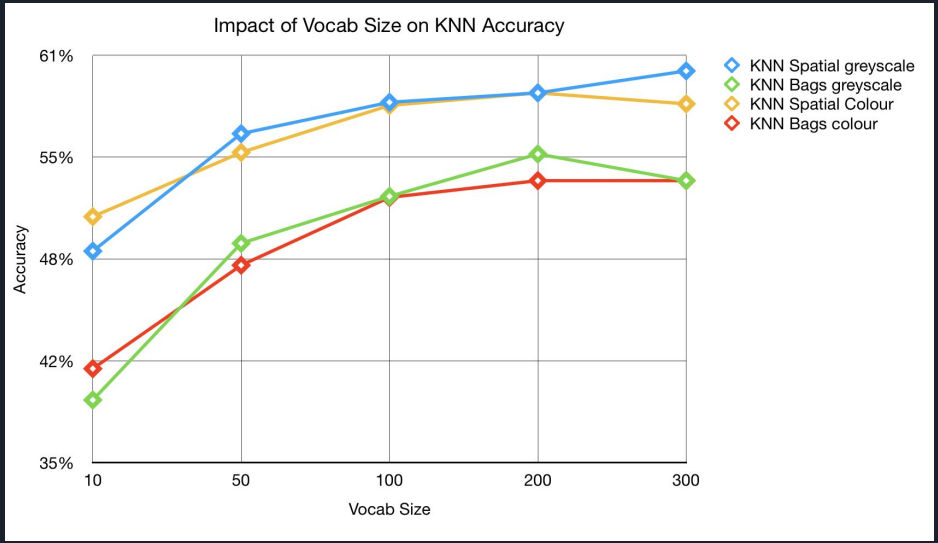
# Impact of Bag / Spatial on Classifiers.

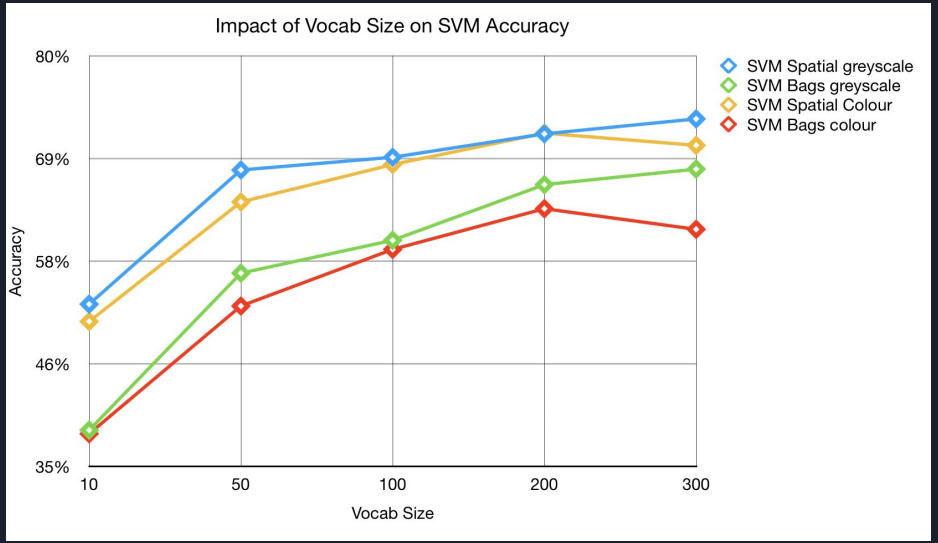


# Impact of Colour on Classifiers.

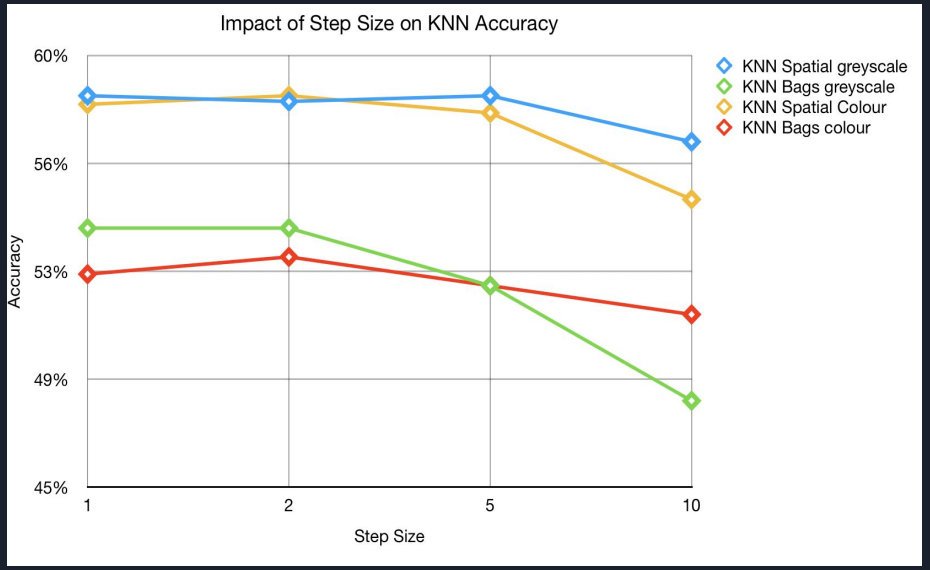


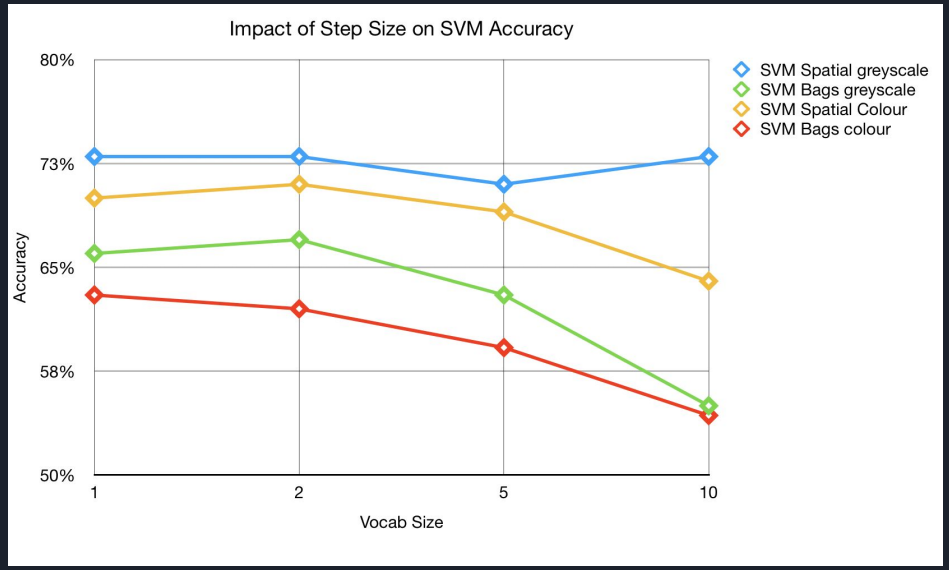
# Impact of Vocab sizes



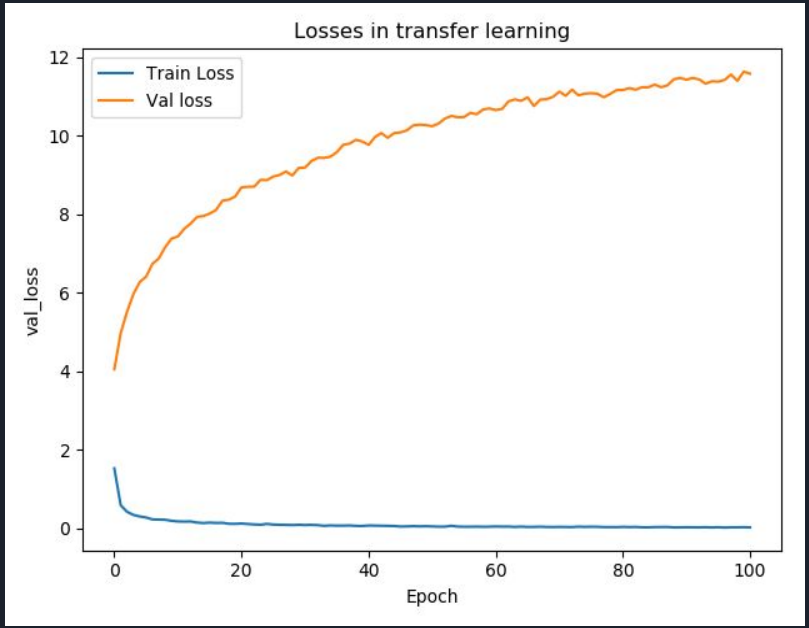



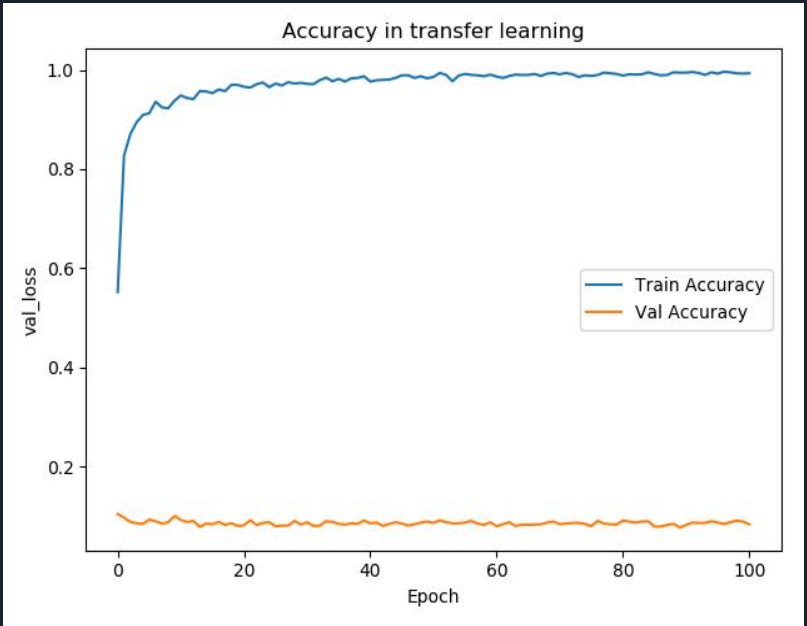
# Step Sizes


Impact of Step Size on KNN Accuracy
60%
56%
53%
49%
45%
Accuracy
1
2
5
10
Step Size
KNN Spatial greyscale
KNN Bags greyscale
KNN Spatial Colour
KNN Bags colour

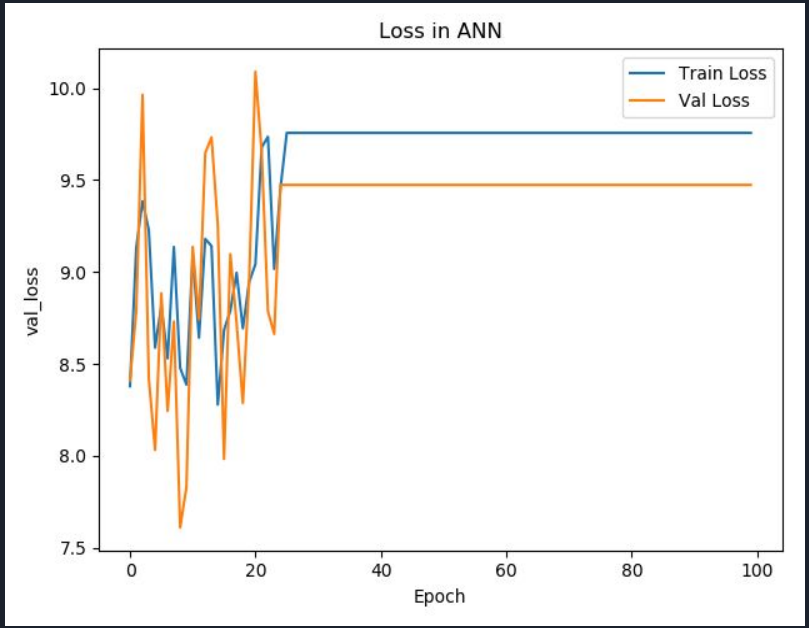

Impact of Step Size on SVM Accuracy
80%
73%
65%
58%
50%
Accuracy
1
2
5
10
Vocab Size
SVM Spatial greyscale
SVM Bags greyscale
SVM Spatial Colour
SVM Bags colour

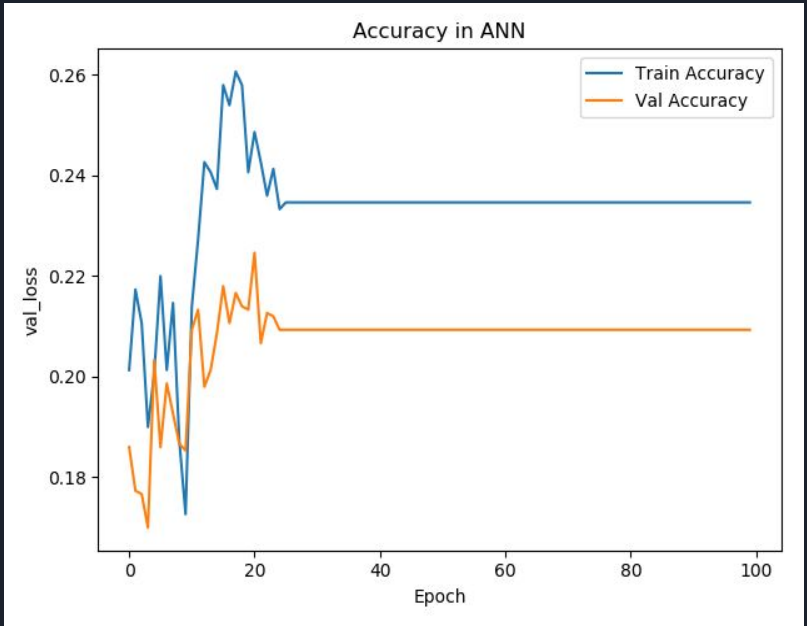
# Transfer Learning results



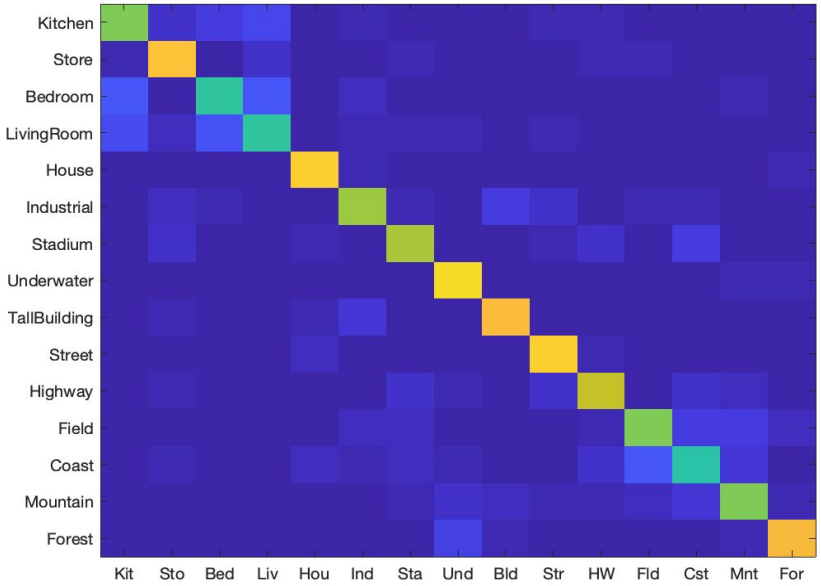
# CNN and Adaboost discussion



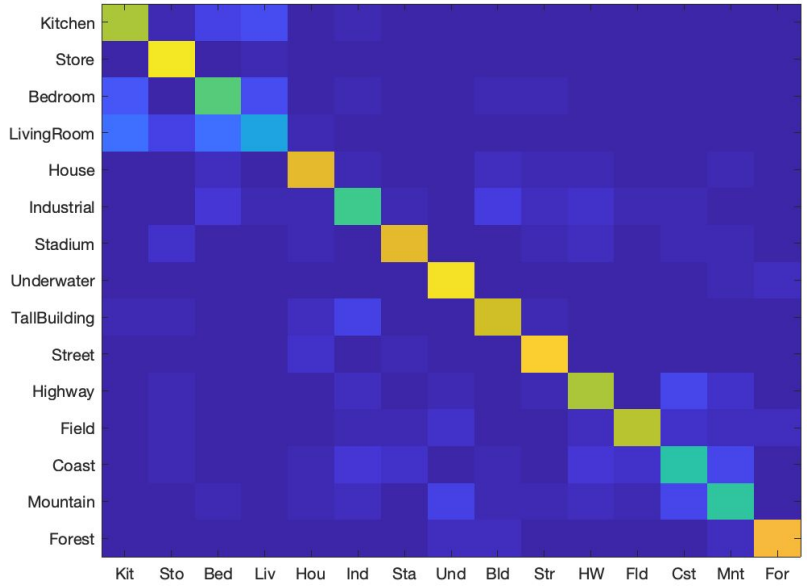



# Confusion - SVM Spatial

73.1% SVM Greyscale Spatial

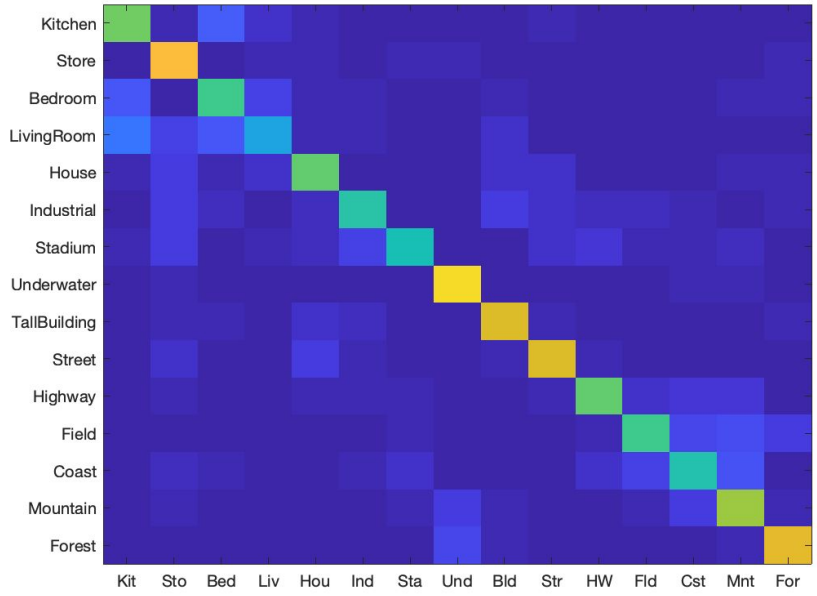



71.5% SVM Colour Spatial

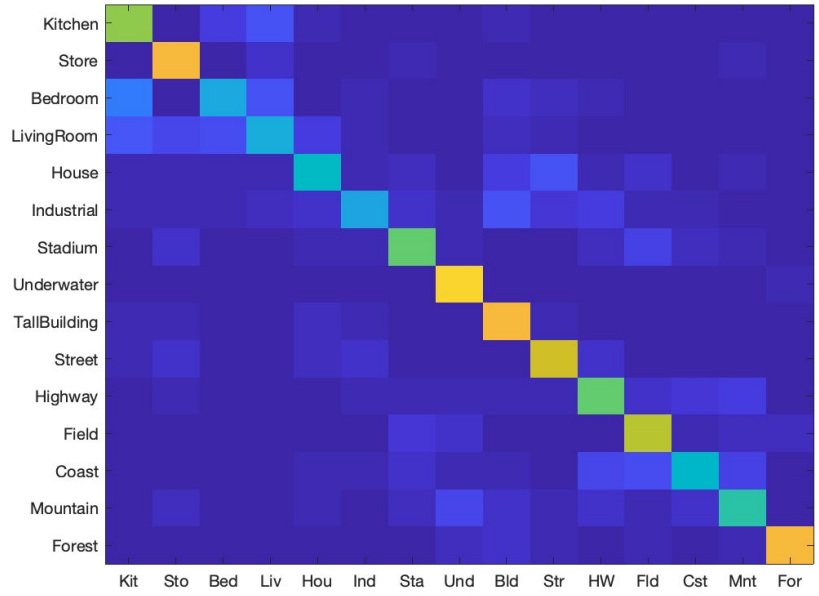


# Confusion - SVM Bag of SIFTs

65.3% SVM Greyscale Bags

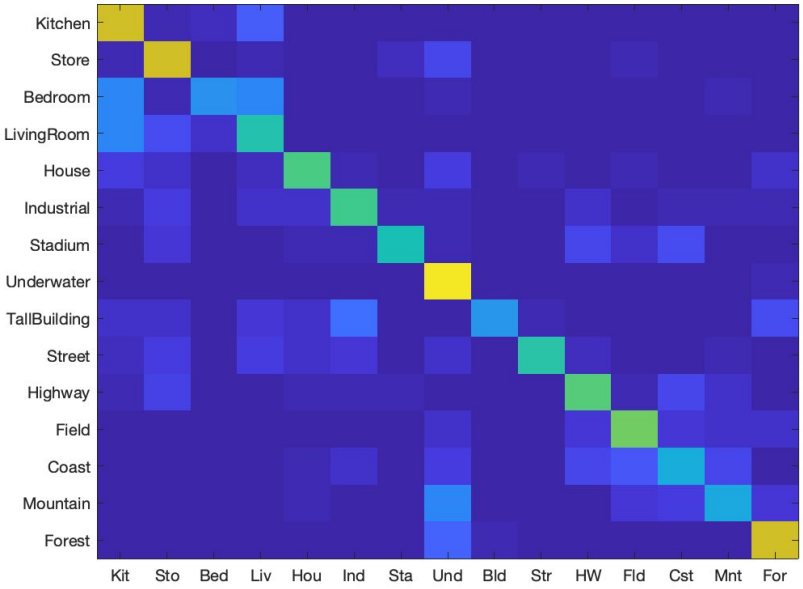



63.1% SVM Colour Bags

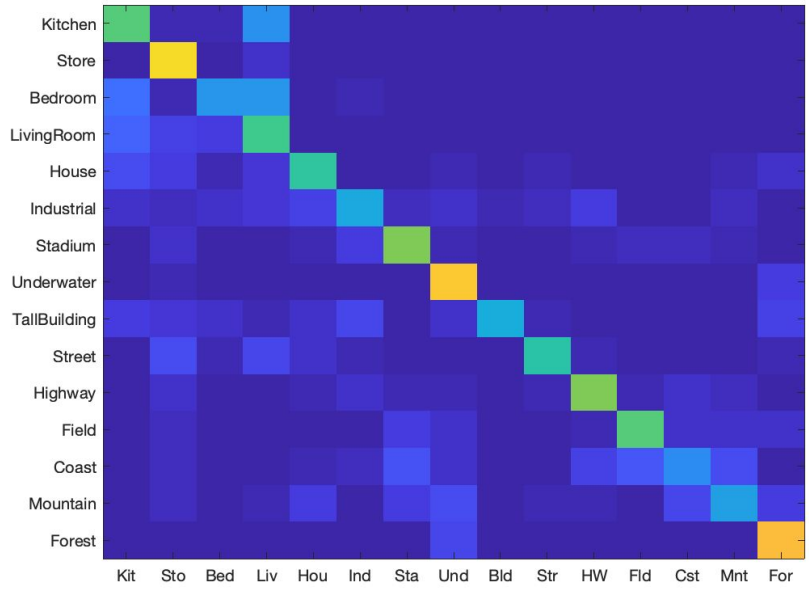


# Confusion - KNN Spatial

58.1% KNN Greyscale Spatial

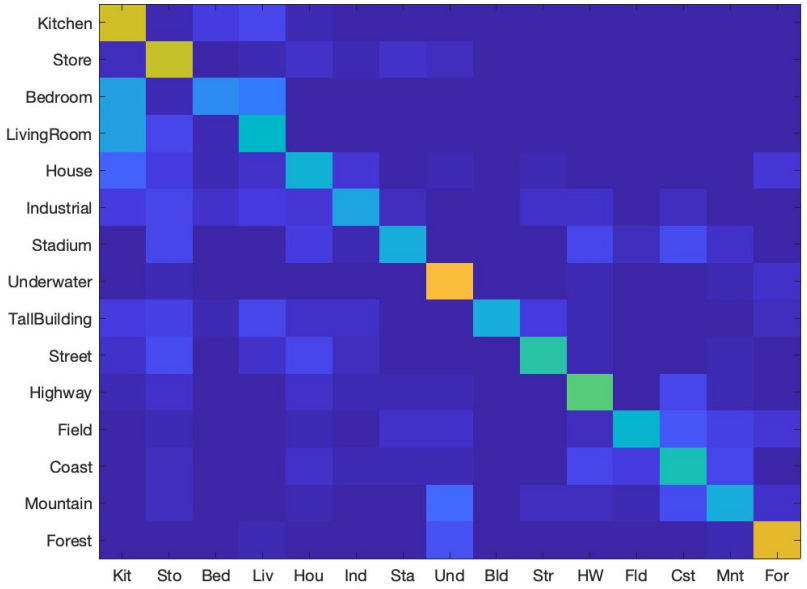



57.8% KNN Colour Spatial



# Confusion - KNN Bag of SIFTs

53.7% KNN Greyscale Bags




50.5% KNN Colour Bags