A Study of the N-D-K Scalability Problem in Large-Scale Image Classification

Carlos E. del-Castillo-Negrete
Yale University, carlos.del-castillo-negrete@yale.edu

Sreenivas R. Sukumar
Oak Ridge National Lab, sukumarsr@ornl.gov

A Study of the N-D-K Scalability Problem in Large-Scale Image Classification

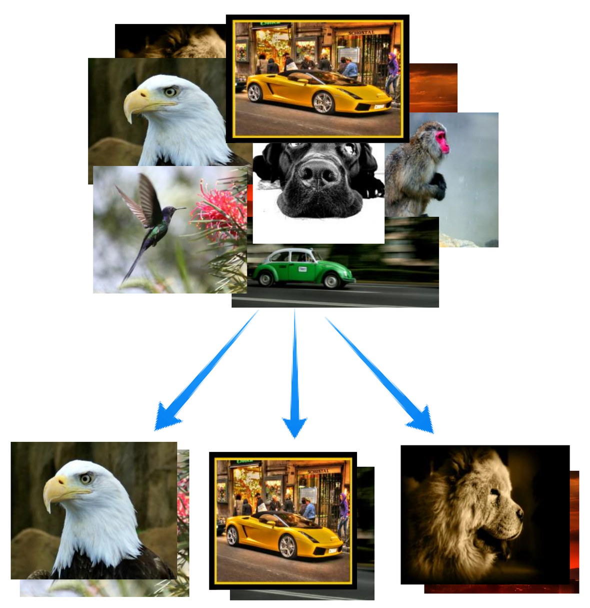
Carlos del-Castillo-Negrete, Yale University

Mentor: Sreenivas R. Sukumar

Image Classification

Image classification is a problem of central importance to computer vision.

Success has been achieved at small scales, but the main challenge remains to build system to rival human visual system? Algorithms must scale up!



N-D-K Scalability Challenges:

N = Number of images used to train model.
D = Dimensionality of feature vector space
K= Number of classes in training problem

•N → Messy and huge data sets. How can we leverage them for maximum efficiency?
•D → Wealth of different feature vectors for images. Which one(s) to use? How can different information be incorporated for a better model?
•K → Classification challenge increases dramatically with number of classes. Semantically close classes (e.g., cat vs. dog) can be very difficult to distinguish.

Machine Learning Approach to Image Classification

•Several machine learning algorithms exist for general image classification.
•We used Linear Support Vector Machines (SVMs) that provide a powerful method for image classification, especially on big data sets.
•SVM can classify large data sets in both N (number of data points) and D (dimensionality of feature space) relatively fast.
•For scale up in K, SVM divides a multi-class problem into several instances of binary classification.
•Several one vs. all models trained to distinguish a given class from the rest.

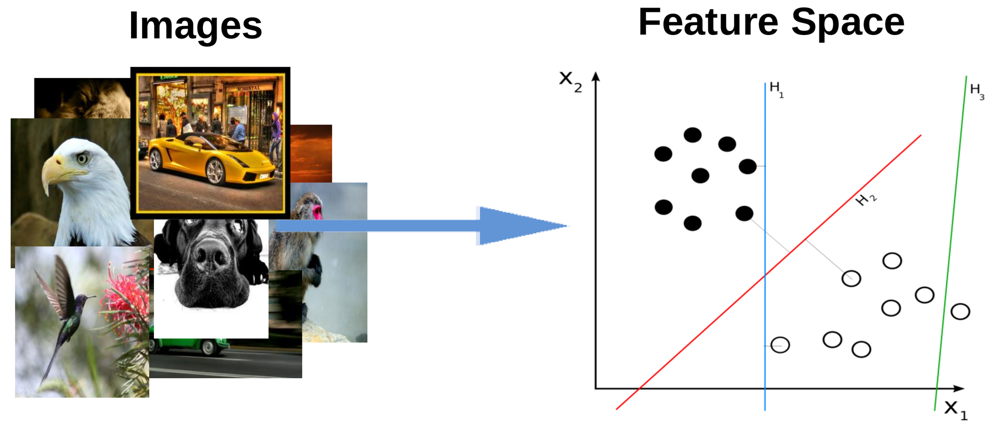


Figure: SVM Model as applied to image classification. Images are transformed into a feature space using pattern recognition algorithms. Each image is represented by a point in this feature hyperspace. The SVM model finds the optimal hyperplane separating two given classes of the training set.

N-D-K Study using SVMs and NUS-WIDE Image Data Set

Scaling with number of classes (K)
•The SVM classification algorithm scales poorly with K.
•Performance of SVM drops significantly as the number of classes increases [Fig.1].
•Performance depends heavily on feature vectors. BOW has the best accuracy and one of the lowest runtimes while WT has one of the worst accuracies and the highest runtime.
•Unbalanced nature of data set skews classification of SVM towards the most dominant classes [Fig.2].
•SVM struggles most with classes that are semantically close [Fig.2].

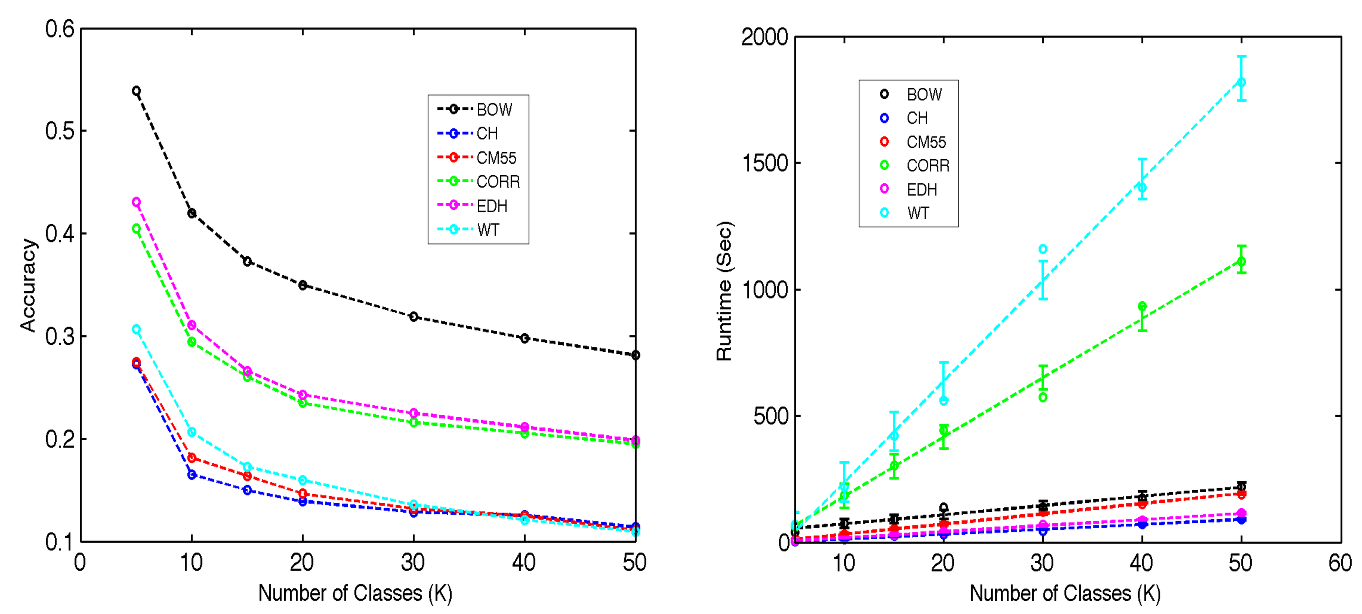


Figure 1: Accuracy (left) and Runtime (right) vs. Number of classes K. SVM trained with 130,000 images for each K.

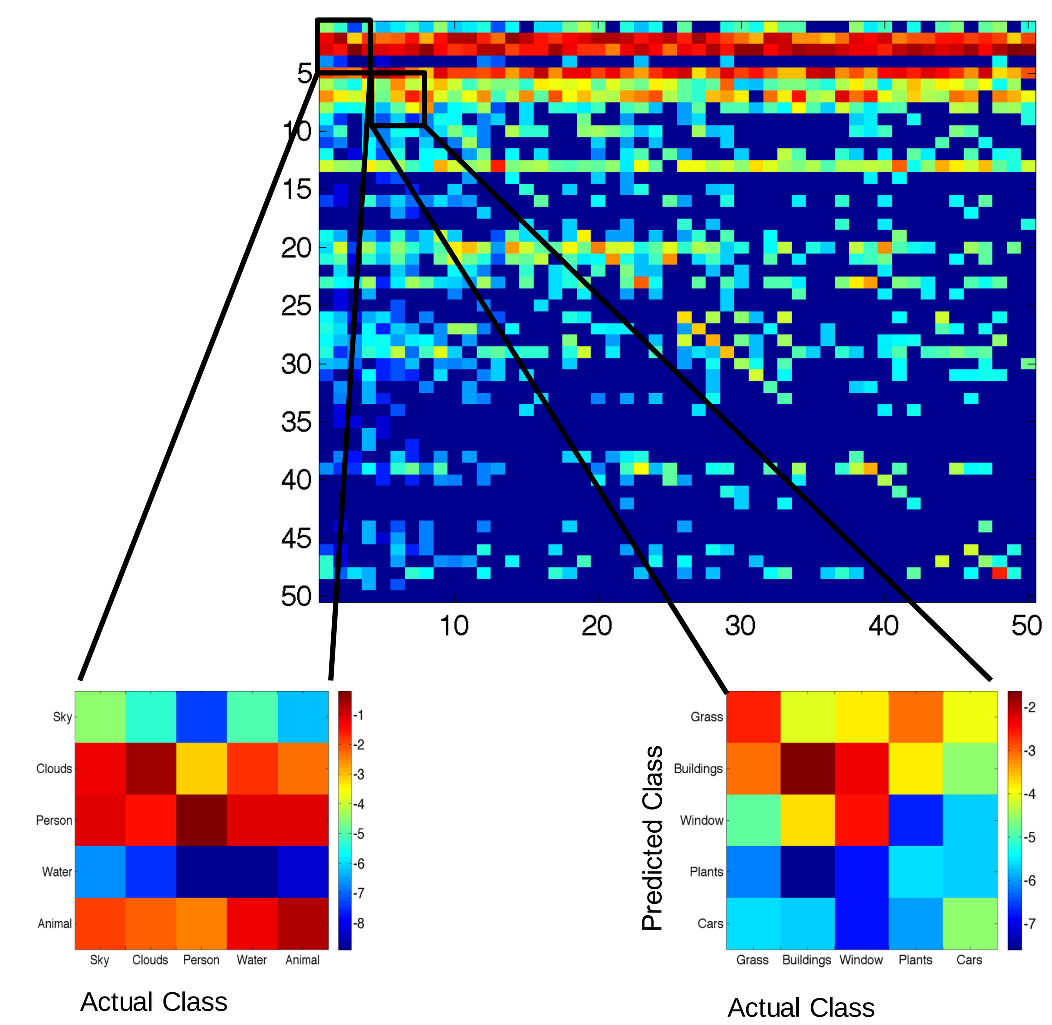


Figure 2: Confusion Matrix showing the performance of linear SVM using BOW feature vectors. Trained model with 130,000 images. Bottom Left: first 5X5 portion of matrix. Pixel (i,j) represents number of times class j was predicted as class i. Note how most populated classes (e.g., person/animal) dominate the misclassification. Bottom Right: Next 5X5 portion of matrix along diagonal. Note how semantically close classes (e.g., grass/plant) have high misclassification rates.

Scaling with number of images (N) in training set [Fig.3]
•The SVM model scales well with the size of the training set.
•A steep learning curve is observed in the accuracy vs. N plots for all features vectors.
•Learning saturates fast, and for N>10^4 accuracy plateaus.
•The overall shape of the learning curve is independent of K
• Interesting crossover is observed in the BOW learning curve.

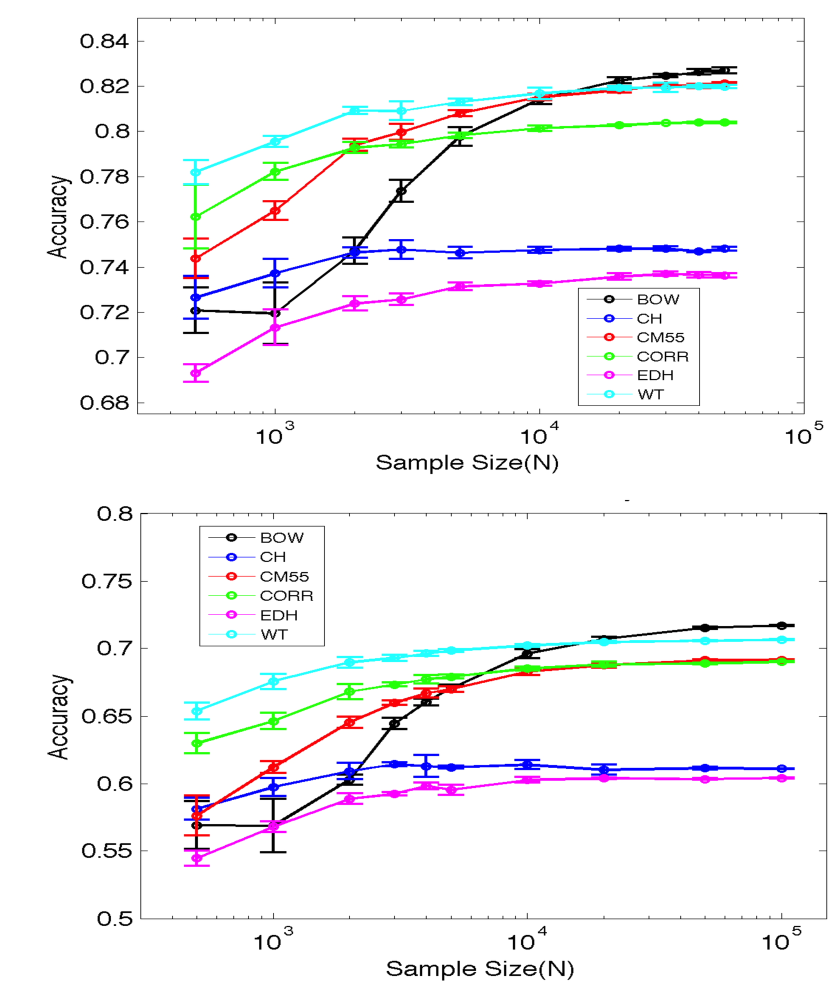


Figure 3: Accuracy vs. sample size (N) for K=2 (top) and K=3 (bottom) classification problems. Results of SVM trained using different feature vectors.

Scaling with dimension (D) of feature vector [Fig.4-5]
•D scaling studied by training SVM with concatenated feature vectors.
•Concatenation increases run-time consistently
•While in general concatenation tends to increases accuracy [Fig.4], in some cases [Fig.5] concatenation has a negligible effect.
•Potential for increased accuracy by combining feature vector information seems promising.
•However, a more targeted approach beyond blind concatenation should be studied.

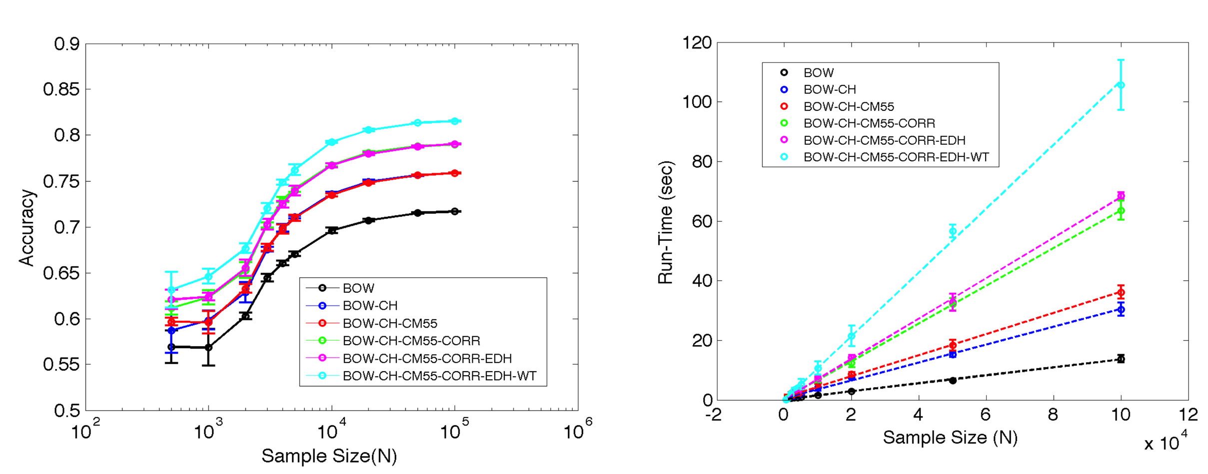


Figure 4: Accuracy (left) and Runtime (right) vs. Sample Size (N) for linear SVM trained using different concatenated feature vectors with K=3.

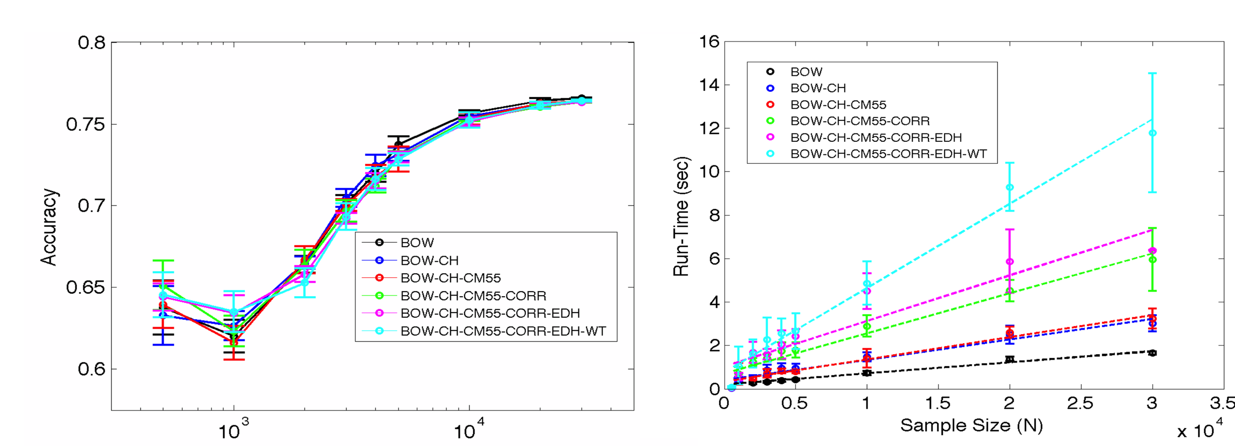


Figure 5: Accuracy (left) and Runtime (right) vs. Sample Size (N) for linear SVM trained using different concatenated feature vectors with K=2.

NUS-WIDE Image Data Set [1]

• 269,648 images mined from Flicker.com
• Associated with each image, six different feature vectors quantifying image content.
• A list of 81 ground truth concepts associated with each image.

	Concept	Images
1	sky	74190
2	clouds	54087
3	person	51577
4	water	35264
5	animal	33887

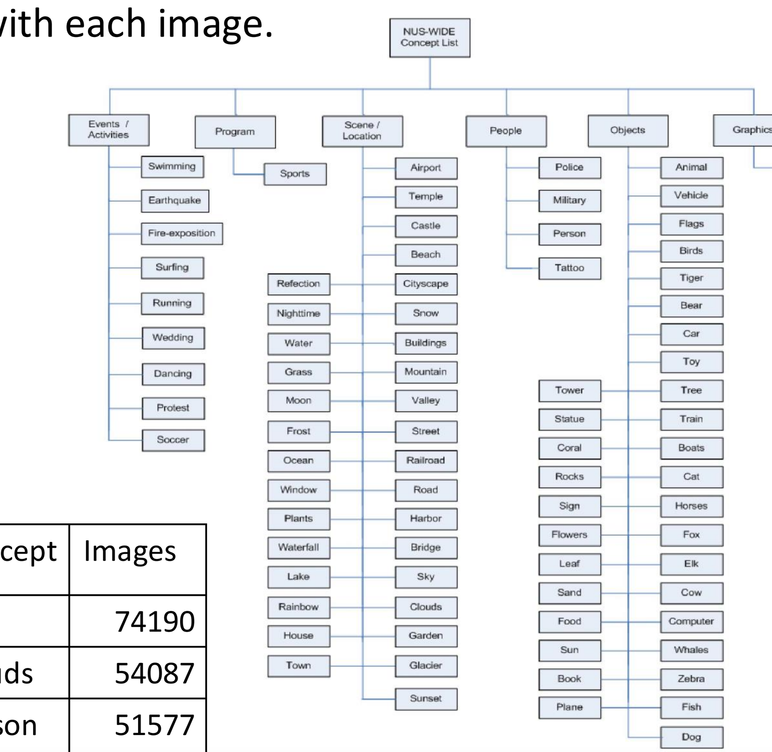


Figure 6: Top: Taxonomy of NUS-WIDE Concepts.
Left: Dominant classes

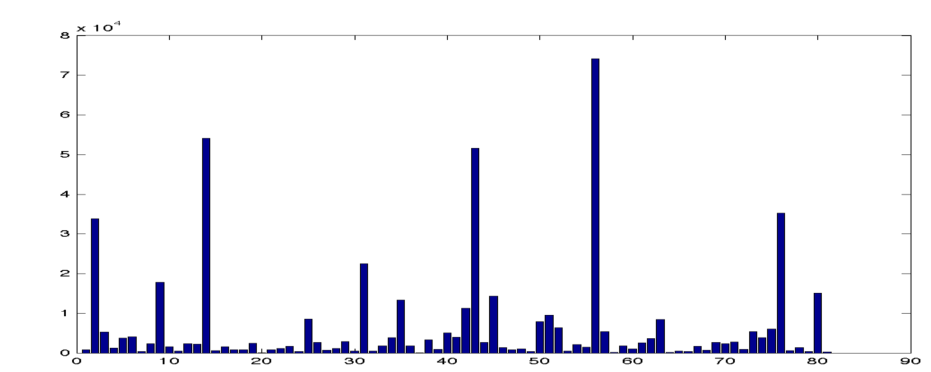


Figure 7: Histogram of number of images associated with each of the 81 concepts.

Software Implementation

A Library to Study N-D-K Scalability

• Developed a C-library of functions to study the scalability of LIBLINEAR [2] implementation of SVM classifiers on large image data sets.
• Built library on-top of LIBLINEAR package and provide functions to convert existing data.
• Used the whole NUS-WIDE data set to sample and construct different instances of image classification problems.
• Measured statistics on the data sets sampled and recorded performance of the LIBLINEAR SVM algorithm when applied to different image classification problems.

References

1. Tat-Seng Chua, et.al.. "NUS-WIDE: A Real-World Web Image Database from National University of Singapore" (2009). http://lms.comp.nus.edu.sg/research/NUS-WIDE.htm
2. Fan, et. Al., LIBLINEAR: A library for large linear classification J. of Machine Learning Research (2008). http://www.csie.ntu.edu.tw/~cjlin/liblinear/.

Yale University
ORAU
OAK RIDGE INSTITUTE FOR SCIENCE AND EDUCATION Managed by ORAU for DOE
OAK RIDGE National Laboratory
UT-BATTELLE
U.S. DEPARTMENT OF ENERGY Office of Science