# Graph-based Active Learning for Semi-supervised Classification of SAR Data


Jason Setiadi

School of Statistics - University of Minnesota Twin Cities


University of Minnesota — Driven to Discover®

College of Liberal Arts — University of Minnesota — Driven to Discover

## Abstract


**We present a novel method for classification of Synthetic Aperture Radar (SAR) data by combining ideas from graph-based learning and neural network methods within an active learning framework. Graph-based methods in machine learning are based on a similarity graph constructed from the data. When the data consists of raw images composed of scenes, extraneous information can make the classification task more difficult. In recent years, neural network methods have been shown to provide a promising framework for extracting patterns from SAR images. These methods, however, require ample training data to avoid overfitting. At the same time, such training data are often unavailable for applications of interest, such as automatic target recognition (ATR) and SAR data. We use a Convolutional Neural Network Variational Autoencoder (CNNVAE) to embed SAR data into a feature space, and then construct a similarity graph from the embedded data and apply graph-based semi-supervised learning techniques. The CNNVAE feature embedding and graph construction requires no labeled data, which reduces overfitting and improves the generalization performance of graph learning at low label rates. Furthermore, the method easily incorporates a human-in-the-loop for active learning in the data-labeling process. We present promising results and compare them to other standard machine learning methods on the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset for ATR with small amounts of labeled data**


## Introduction

Synthetic Aperture Radar (SAR) utilizes the movement of an antenna over a distance from the target to capture finer resolution images than standard radar. There are both phase and amplitude components of the signal that in combination can provide more detailed information about the objects in the scene. Automatic target recognition (ATR) of SAR data seeks to classify the objects of interest in such SAR images. Hand-labeling images by human eye is an impractical and time-consuming task for large datasets. This makes SAR very amenable to automated machine learning methods.

Supervised machine learning algorithms, such as deep learning, rely on an abundance of labeled data to learn from. In many applications, labeling data can be quite costly, while unlabeled data is ubiquitous and easy to obtain. Semi-supervised learning (SSL) methods use both labeled and unlabeled data in the learning task and aim to achieve good quality results with far less labeled data than fully supervised methods. A common way to use the unlabeled data is through the construction of a similarity graph, which effectively leverages relations between unlabeled data points for dimension reduction and classification tasks.

A successful application of graph-based learning to image classification hinges on constructing a high-quality graph that accurately encodes the similarities between data points while ignoring or suppressing differences between images that are due to spurious noise or image acquisition artifacts. Since the raw pixel values are sensitive to noise, contrast, lighting, or small shifts or rotations that commonly corrupt image data, it is important to apply a feature transformation to the images before constructing a graph. Several recent papers have successfully used variational autoencoders (VAEs) for unsupervised feature extraction in hyperspectral imagery, SAR imagery, and for constructing similarity graphs in graph-based learning. VAE feature learning is an unsupervised method that retains the power and flexibility of deep supervised learning, making it ideal for problems with limited amounts of data.

In addition to constructing a high quality graph and considering the amount of labeled data available to a classifier, the choice of training (labeled) points can significantly affect classifier performance. Active learning is a branch of machine learning that judiciously selects a limited number of unlabeled data points to query for labels, with the aim of maximally improving the underlying SSL classifier's performance. In applications like ATR in SAR imagery, the chosen active learning query points are labeled by an oracle, or human in the loop, such as a domain expert. These query points are selected by optimizing an acquisition function over the discrete set of data points available in the unlabeled pool of data. Active learning can greatly increase the performance of classifiers at very low label rates, and minimize the cost of labeling data with a human-in-the-loop.

In this work we present a novel pipeline for combining graph-based semi-supervised learning and VAE-based feature extraction methods within an active learning framework to improve ATR in SAR imagery data. In particular, we use a convolutional variational autoencoder (CNNVAE) to learn feature representations of SAR images. The CNNVAE feature embedding is completely unsupervised (i.e., requires no labeled data), so the method is compatible with active learning at low label rates. We compare our method with various acquisition functions against other standard machine learning methods on the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset for ATR with small amounts of labeled data. Our main results show that our active learning method for SAR data can outperform state of the art SAR classification methods while using only a fraction of the labeled data used in existing approaches.

## Methodology

The Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset contains 6,874 images of 10 types of military vehicles. A Sandia X-band radar operating at 9.60GHz with a bandwidth of 0.591GHz was used to collect the data. We follow a standard training and testing split according to the angle at which the SAR data was collected; namely, the training data was obtained at an angle of 15° while the testing data was obtained at 17°. Given this pre-defined train and test split of the data, we accordingly restrict the possible set of active learning query points at any iteration to belong to the training set, and we only test the accuracy on the testing set.

The original SAR images are of various sizes, and so the magnitude and phase images were all center-cropped to 88 × 88 pixels. A vast majority (> 99.8%) of the pixel values in the magnitude images are within the range [0, 1]. The pixels outside this range appear to be noise, and so we clipped the magnitude images to the range [0, 1]. We then converted each magnitude and phase image-pair into a 3-channel image by taking the first channel to be the magnitude image, and the second two channels to be the real and imaginary parts of the SAR image. This transformation ensures all image channels have pixel values in the range [0, 1], which is necessary for the loss function in variational autoencoders.

We now describe our novel data pipeline for applying graph-based active learning to SAR data. We utilize neural network architectures called variational autoencoders (VAE) to learn latent representations of the SAR images from which we construct our similarity graph to apply graph-based active learning. Variational Autoencoders (VAE) transform the input data to (usually) a lower-dimensional space via the use of an "encoder" structure from which a "decoder" structure is designed to reconstruct the input data. The encoder and decoder neural architectures we use involve convolutional layers, hence we refer to our VAE architecture as a CNNVAE (Convolutional Neural Network Variational AutoEncoder).

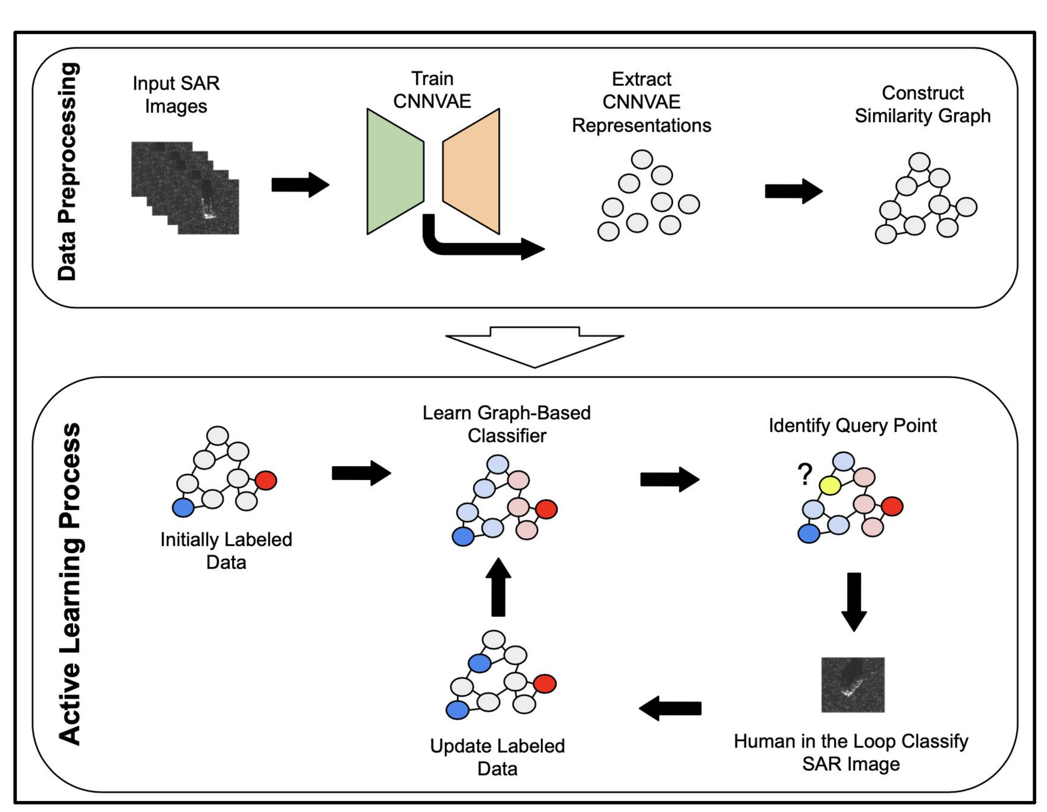
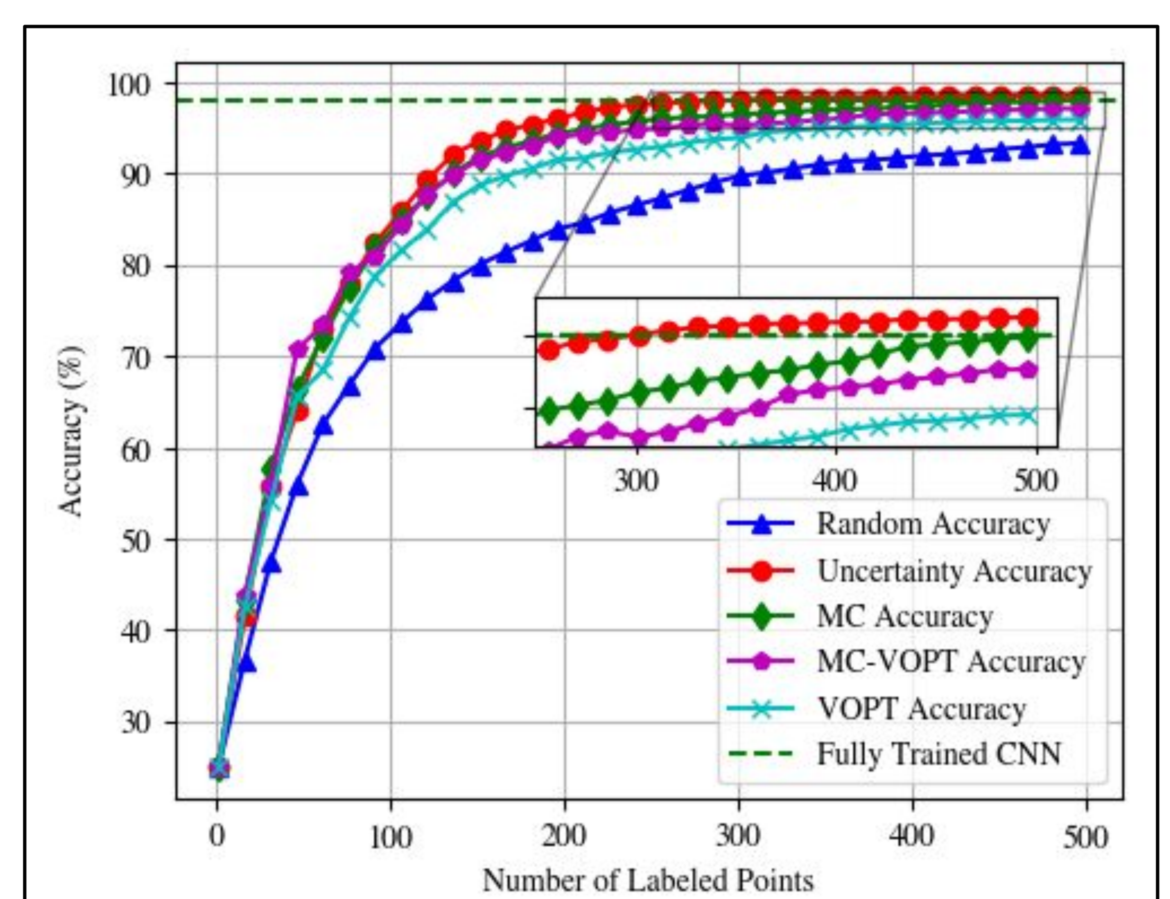



Figure above shows our pipeline for processing the SAR data. We first train the CNNVAE to learn lower-dimensional embeddings of the SAR imagery data, and then use the learned embeddings to construct a similarity graph on the inputs. This similarity graph is then used for inferring the labels (classifications) of the unlabeled data from the small amount of labeled data that we possess initially. We then apply graph-based active learning acquisition functions in order to sequentially select unlabeled points for a human in the loop then label and add to the labeled data.

Once we have constructed a similarity graph from the CNNVAE representations of the SAR images, we then apply Laplace learning to infer labels for the unlabeled data from a small set of initially labeled data. For our experiments on the MSTAR dataset, we choose this initial set by selecting uniformly at random a single image from each class.

We now discuss the graph-based acquisition functions for applying in the active learning iterations on the MSTAR dataset. We apply Random Sampling, Uncertainty Sampling, VOpt, Model Change (MC), and a novel acquisition function MCVOpt. Random sampling selects query points by sampling uniformly at random from the unlabeled set at each iteration. Uncertainty Sampling selects query points that the current graph-based classifier is the most uncertain about at each iteration. The VOpt acquisition function calculates the amount that the trace of the covariance matrix decreases as a result of adding an unlabeled point to the current labeled set. Model Change (MC) is a recently proposed graph-based acquisition function that quantifies the amount by which the underlying graph-based model would change if an unlabeled point were added to the current labeled set with the label predicted by the current model. Lastly, MCVOpt combines MC and VOpt's acquisition functions into one.

In order to apply graph-based learning, we build a k-nearest neighbor similarity graph on the latent CNN and CNNVAE features. When we refer to the latent CNN features of either network, we are referring to the output of the initial convolutional layers, once flattened into a vector. For the supervised CNN, the latent CNN features have dimension 6400, while for the CNNVAE, the latent features have dimension 64 × 11 × 11 = 7744. We use an approximate k-nearest neighbor search to efficiently perform the nearest neighbor search in high dimensions. Instead of the usual Euclidean distance, we use the angular metric for the k-nearest neighbor search, which is equivalent to normalization all the feature vector to unit norm and using the Euclidean distance. After performing a k-nearest neighbor search, we construct a self-tuning similarity graph with edge weights. We used k=20 in all experiments and symmetrized the weight matrix.




## Results

We now present our results from applying our novel graph-based active learning pipeline to the ATR task on the MSTAR dataset. We use a CNNVAE to learn a 32-dimensional embedding of the original 88 x 88 images. We then construct a similarity graph from these lower dimensional embeddings, from which we then iteratively select new query points to add to the labeled data; we use Laplace learning as the underlying graph-based classifier for accuracy evaluation.

Given the similarity graph constructed according to the pipeline, we perform experiments as follows. A single point per class (i.e., 10 in total) is selected uniformly at random to comprise an initially labeled set. This set is then used as the start for choosing 500 query points for evaluating each acquisition function in the active learning process. After each point is selected and added to the labeled data with its corresponding ground-truth label, the resulting accuracy in the updated Laplace learning model's classifier is recorded. For each acquisition function's set of active learning iterations, an accuracy curve is produced by averaging the accuracies over all 10 iterations. These accuracy curves represent how useful the corresponding acquisition functions queries were for improving the underlying classifier's performance in the trials. We use hyperparameter values of gamma = 0.5 and m = 300 (spectral truncation cutoff) for the acquisition functions MC, MC-VOPT, and VOPT which utilize the spectral truncated Gaussian Regression model for acquisition function evaluation.

Since we have ground truth for MSTAR, the active learning queries use this information to provide the new labels. Each trial used a different random seed to produce the single, initially-labeled point per class. The green dotted line depicts the accuracy for the state-of-the-art CNN method for ATR on the MSTAR dataset that was trained on the entire training set (i.e., with 3,671 labeled points).

With just 300 points labeled during the active learning process, our graph-based semi-supervised classifier achieves the state-of-the-art CNN classifier accuracy! That is, with less than 10% of the labeled data that are used by modern CNN classifier architectures, we can achieve state-of-the-art classification accuracy on the MSTAR dataset. This is convincing evidence that our proposed graph-based semi-supervised classification and active learning pipeline is a data-efficient and accurate methodology for application to SAR imagery.

By comparing the relative shapes of these accuracy curves, one can measure the relative utility of the given acquisition functions. It is desired that the acquisition function yields the greatest initial increases in accuracy while also ending in the highest overall accuracy. From the accuracy plot, we observe that each of the shown graph-based acquisition functions are superior to random sampling in the earlier stages of the active learning process. They each provide similar initial increases in accuracy for roughly the first 200 choices. Thereafter, while the VOpt, MC, and MCVOpt acquisition functions' corresponding accuracy curves level off at an overall accuracy of around 95%, Uncertainty Sampling continues to steadily improve to an overall accuracy of just over 97%. These results suggest that Uncertainty Sampling is the superior choice of acquisition function for graph-based active learning on the MSTAR dataset.

## Conclusions

**We present a novel machine learning pipeline for active learning with graph-based classification applied to SAR imagery data. With tests of ATR in the MSTAR dataset, our method proves to be very useful in leveraging small amounts of labeled data to classify the images. Further directions include testing on other SAR imagery datasets as well as improving the graph construction with other unsupervised representation learning methods.**

## Acknowledgements


**I would like to thank my faculty mentor Dr. Jeff Calder from the School of Mathematics, UMTC for guiding me throughout this research. I would also like to thank Dr. Bertozzi and Kevin Miller for their mentorship throughout the UCLA Summer CAM REU. Finally I would like to thank my collaborators Jack Mauro, Xoaquin Baca, and Zhan Shi during the REU.**