CSC345 - OBJECT RECOGNITION REPORT

AYDEN BALLARD – 905438

INTRODUCTION

For this coursework, we are given 10,000 training images, each with a label stating what the object in the image is. The goal of the task is to train a machine to predict the category of the object in an image. The categories of objects to be recognised are: airplane, automobile, bird, cat, deer, dog, frog, horse, ship and truck. We are also given 1000 testing images, these testing images do not contain any labels stating what the object in the image is. The aim is to train the machine using the 10,000 training images, to successfully predict the object in each testing image.

Considering the fact that we are given labels for each of the training data, we will be using supervised learning to help train the machine. In addition to this, I will be trying several different models to train the machine, I will then report and compare each model's success rate, when used to predict the output for the testing images. The models I will be using are: Gaussian Mixture Modelling (GMM), K-Means and Support Vector Machine (SVM). Furthermore, I hope that these tests will help evaluate which machine learning technique is the most fitting for object recognition in images.

METHOD

All image features are computed using the "computeFeatures" function. This function takes an image as a parameter and returns a histogram of the features of that image. The function also returns an image of the histogram itself, this may be used for additional insight on the extracted features of each image.

By using this function, we can loop through the array which stores all 10,000 training images and 1000 testing image, and compute a histogram of the features of each image. The training images features are all stored in a list called trainingFeatures, whereas the testing images features are stored in a list named testingFeatures.

K-MEANS

K-Means clustering is a simple clustering algorithm. "The goal of the K-Means algorithm is to divide M points in N dimensions into K clusters so that the within-cluster sum of squares is minimised". [2] However, in order to implement K-Means clustering, we must first know our value for K. This value, represents the number of clusters that we wish to compute. However, since we are dealing with 10 different categories of objects in images, we know that we wish to compute 10 clusters. This means that our K value is 10.

Once we have determined our value for K. We can randomly choose an initial K centroids and loop through the data, with each iteration of the loop, assigning each data point to its nearest centroid and then updating the centroids based upon the new partitioning of the model. This loop will continue until a stopping criterion is met.

On the other hand, K-Means clustering is not the best method of machine learning for this situation. This is because K-Means clustering is an unsupervised learning technique, whereas for this task we are given the labels for the training images.

GAUSSIAN MIXTURE MODELLING (GMM)

GMM is a model based approach for clustering. It is described as the weighted sum of single Gaussian functions. However, it is not a parametric model because it refers to more than one Gaussian function.

Also, GMM can be used to predict the output for the testing images features, when trained with the training images features. With GMM, we must state the number of dimensions that we are dealing with. Since we have 10 different categories of objects in images, we will have 10 different dimensions. After this, we can start our GMM model with an initial guess of the parameters using K-Means.

Once we have an initial guess of our model, we can compute the posterior probability of each data point (the probability that data point x belongs to component j). After we have computed the posterior probabilities, we can update the parameters and normalise the mixing coefficient. This is repeated until it converges, effectively maximising the posterior probability.

However, similar to K-Means clustering, GMM is also an unsupervised learning technique. This means that if we were to use GMM, the labels given for the training images would be made redundant as GMM does not take labels in to consideration.

SUPPORT VECTOR MACHINE (SVM)

SVM is a supervised machine learning algorithm which can be used for both classification and regression.

"The support vector machine (SVM) is a supervised learning method that generates input-output mapping functions from a set of labelled training data" [3]

A support vector machine tries to find a line which best separates 2 classes, meaning the line with the largest margin. Each data point which falls on the margin, are known as support vectors. In our case, we can plot each of our training features in a 10 dimensional space and then perform classification by finding the hyper-plane which differentiates the classes the best.

SVM then can be evaluated by calculating the positive and negative results of the algorithm, where positive means the number of data samples that belong to its class, and negative meaning the number of data samples that do not belong to its class. We can then use these positive and negative values to compute the true/false positive in addition to the true/false negative. A true positive is a positive sample that has been correctly identified as positive, whereas a false positive is a negative sample that has been incorrectly identified as positive. Afterwards, we can use the true/false positive and true/false negative values to compute the true positive/negativity rate.

The true positive rate is the percentage of correctly predicted positive samples and is calculated by TP/(TP+FN) where TP denotes true positive and FN denotes false negative. Whereas the true negative rate is the the percentage of correctly predicted negative samples and is calculated by TN/(TN+FP) where TN denotes true negative and FP denotes false positive. Finally, we can use these values to compute the overall accuracy of our support vector machine. We can compute the overall accuracy by (TP+TN/P+N).

RESULTS

In this section of the report, I will be discussing the results of the different machine learning algorithms. I will also compare and evaluate the results in order to determine which machine learning algorithm is the best for the given task.

K-MEANS

Below, you can see a table that shows the percentage results of correct guesses of testing images using the K-Means clustering algorithm.

Test	1	2	3	4	5	Average
Accuracy (%)	6.7	8.2	9	9	13.7	9.32

As you can see, the results returned from predictions using the K-Means model varies a fair bit. One reason for this may be that K-Means clustering starts with a random initialisation of cluster centroids, this leads to the variance of results. In addition to this, you can see that K-Means clustering does not produce very good results.

Furthermore, another reason that helps explain the variance in results is that K-Means clustering returns low accuracies is because it is not a supervised learning algorithm, this means that the labels provided for the training images are never utilised.

GAUSSIAN MIXTURE MODELLING (GMM)

Below, you can see the accuracy results of training image predictions using the GMM model.

Test	1	2	3	4	5	Average
Accuracy (%)	11.1	13	13.5	7.6	8	10.64

Similarly, to the results of K-Means model, the GMM accuracy varies as well. The reason for this may be that GMM starts with an initial guess of its parameters using K-Means. This means that each GMM will be different from one-another at the initialisation stage, leading the results to vary.

Furthermore, GMM is a unsupervised learning technique. This means that our labels for the training data is never utilised and causes a lower accuracy result than if we were to use a supervised learning technique.

SUPPORT VECTOR MACHINE (SVM)

The accuracy of correct predictions for the testing data using the SVM model is 42%. This models results does not vary like the results returned from K-Means and GMM. This is because as long as the input data is not changed, the results will be the same as it uses maximising and minimizing functions to compute the classes. Furthermore, the reasoning for SVM having a higher accuracy is because it is supervised learning. This means that we are able to fully utilise the training images labels.

CONCLUSION

In conclusion, it is clear that Support Vector Machine is the superior machine learning technique over K-Means and Gaussian Mixture Modelling for object recognition. I believe that this is because SVM is a supervised learning algorithm. This means that SVM is able to utilise the training data labels when creating the model. In comparison, K-Means and GMM are both unsupervised, meaning that no labels are utilised in these techniques.

In addition to this, it is important to consider the quality and quantity of our input data when evaluating these machine learning techniques. Firstly, we were not given a large sum of training images. If we were given more training images, the accuracy of the predictions made by our machine learning algorithms would be significantly higher. Also, the quality of the training and testing images is not the best as many images contain a lot of noise. This may have effects on the 'computeFeatures' as images with a lot of noise may have outliers in its histogram of features.

REFERENCES

[1] Partridge, M and Calvo, R. (1997). Fast Dimensionality Reduction and Simple PCA. University of Sydney

[2] Hartigan, J. and Wong, M. (n.d.). *A K-Means Clustering Algorithm*. Yale University.

[3] Wang, L. Support Vector Machines: Theory and Applications (2005). Singapore.