

UNIVERSITY OF BAHRAIN
College of Information Technology



**A Hybrid Approach of One-Class
Classification and SVM to classify
patients based on Medical Imaging and
Reports**

A Thesis Submitted in Partial Fulfilment of the Requirements for the Doctor
of Philosophy Degree in Computing and Information Sciences

Submitted By
Fadheela Abbas Hussain

20013739

Supervised By
Dr. Riadh Ksantini
(Associate professor)

University of Bahrain
February 2023

ABSTRACT

Breaches of data can seriously affect healthcare providers' operations and reputations, as they can delay treatment or lead to financial loss. Despite efforts to put in place measures to prevent unauthorized access to patient information, breaches are still happening in the healthcare industry. Due to the increasing number of data breaches, it has become more important that organizations implement effective security measures for their Big Data projects. Classifying the data method tends to produce higher error when compared to other methods due to the large variance directions. One of the most common types of detection technology is one-class classification. This method is very competitive in detecting fake medical images due to the data's unbalanced nature. However, it can also produce higher error when compared to other methods.

One of the most effective ways to improve the accuracy of one-class classification is by implementing incremental covariance-guided support vector machine (iCOSVM) especially with a real time system which is proposed in this thesis. In this part of the system, we first generate fake images using Generative Adversarial Network (GAN) framework. After, extracting features of fake images samples using CNN-VGG-16 and perform further analysis. Then, in order to classify the images, in real time system iCOSVM is introduced. Finally a comparative analysis is performed of iCOSVM and the existing relevant incremental boundary-based methods is performed.

In addition, to increase productivity, an effective unsupervised clustering technique is necessary. It serves as a crucial preprocessing recommended step prior to any supervised machine learning system. Therefore, an Improved Dynamic Autoencoder (IDynAE) model is proposed in this study too as a pre-processing technique for our system.

We evaluate the impact of our proposed IDynAE model which demonstrates that our

method is able to outperform the current generation of detection technology. It has the advantage of being able to provide a superior level of performance over DynAE that are most pertinent deep clustering techniques too. As DynAE compute the conflicted and unconflicted sample using the Euclidian distance of the samples to the centroid which is irrelevant as it is in this case do not consider the topological information of the cluster, whereas the IDynAE compute the conflicted and unconflicted sample based on the topological filter. When we used the topological filter the topology information of the cluster model is preserved, which leads to more accurate estimation of the conflicted and unconflicted samples in order to estimate the with results of 84.36% and 40.97% in terms of ACC and NMI respectively.

Furthermore, in regards to our classification system, the iCOSVM proposed detection system reveals a very accurate and efficient results with 98.914% of the average of Area Under the receiver operating characteristic Curve (AUC) for the seven GANs models proposed in this experiment in the test dataset. Additionally, comparing iCOSVM detection system with other methods (in GAN 7) including batch learning system too, the iCOSVM attains a high accuracy and efficiency comparing to other incremental one class classifications algorithms with AUC (Test) 99.25%. Likewise, it outperforms the batch learning system with AUC (Test) 85.25% minimizing the computational complexity time as well with 2.74 seconds.