International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Enhancing Image Diagnosis by the Implementation of Transfer Classifiers

Vinay Agrawal, Ashutosh Shankhdhar, Arushi Mangla

Abstract— ***Images generated from a variety of sources and foundations today can pose difficulty for a user to interpret similarity in them or analyze them for further use because of their segmentation policies. This unconventionality can generate many errors, because of which the previously used traditional methodologies such as supervised learning techniques less resourceful, which requires huge quantity of labelled training data which mirrors the desired target data. This paper thus puts forward the mechanism of an alternative technique i.e. transfer learning to be used in image diagnosis so that efficiency and accuracy among images can be achieved. This type of mechanism deals with variation in the desired and actual data used for training and the outlier sensitivity, which ultimately enhances the predictions by giving better results in various areas, thus leaving the traditional methodologies behind. The following analysis further discusses about three types of transfer classifiers which can be applied using only small volume of training data sets and their contrast with the traditional method which requires huge quantities of training data having attributes with slight changes. The three different separators were compared amongst them and also together from the traditional methodology being used for a very common application used in our daily life. Also, commonly occurring problems such as the outlier sensitivity problem were taken into consideration and measures were taken to recognise and improvise them. On further research it was observed that the performance of transfer learning exceeds that of the conventional supervised learning approaches being used for small amount of characteristic training data provided reducing the stratification errors to a great extent.***

Index Terms—Transfer Learning, Biased SVM, Adaptive SVM, Heat Maps, Machine Learning

I. INTRODUCTION

Image segmentation today plays a significant role in most of the industries around the globe whether it be medical field or space technologies. As industries' requirements have it today, various automated segmentation techniques have been developed which in contrast to manual segmentation are more efficient and effective.

Voxel wise categorization by supervised learning is the basis of most of the automated segmentation techniques today. Supervised Learning uses a linear or non-linear separator for classifying the target data into a specific category according to its characteristic. It focuses on creating a separator by extracting the attributes of both the target and training dataset provided. The classification of the target data into the desired category then takes place with the help of the separator. It tries to do this by creating a maximum margin hyperplane which puts the new data points in its decision boundaries and performs classification on them with the help of its support vectors and the maximum margin line created.

To increase the efficiency of supervised learning mechanisms, many different methods have been proposed earlier, such as incorporating rotation invariance into SVM, which are the virtual support vector SVM and the tangent vector SVM [2]. Along with that brain tumor classification [3] has also been improved via this mechanism by using a hybrid technique of SVM and fuzzy c-means. For increasing utility, the given training data must be prototypical of the desired target data. Though it might be possible that in case of medical sectors, only limited quantity of training data be given because of various factors which can create disturbances in the actual result.

This is where we aim to introduce the concept of Transfer Learning, which allows the use of smaller amounts of training data for image segmentation. In addition to this, it can also be used to minimize the classification errors in video classification as has been introduced by Dai Chunni [1] by the use of textural attributes, and colours of the video. The foundation of this mechanism lies in the fact that it tries to achieve maximum efficiency by using the existing models and using their related behaviour to create a new model. It also uses the advantage of those mechanisms to extract information during the initial training phase, so that maximum utilisation can be achieved with minimum efforts.

II. BACKGROUND

One of the machine learning techniques that has emerged in the recent times has been Transfer Learning which by the use of small amount of training data obtained from different sources increase the efficiency of the given problem and accept datasets where the training and target labels may not be the same. This gives a great advantage to the user because now one can use different labelling attributes, constraints, and characteristics of both the actual and desired datasets. The mechanism in which we use same labelling function, attributes and distribution can be mentioned as data which has been collected from the same origin. Our aim in Transfer Learning is to create such a mechanism which focuses and learns from the already given data which comes from different origins or in cases where the data may be stored in cloud services securely and may be required later for analysing.

Manuscript published on November 30, 2019.
* Correspondence Author
Vinay Agrawal*, Computer Engineering and Applications, GLA University, Mathura, India.
Ashutosh Shankhdhar, Computer Engineering and Applications, GLA University, Mathura, India.
Arushi Mangla, Computer Engineering and Applications, GLA University, Mathura, India.

This data might be confidential in many times, and thus requirement of a secure transfer mechanism is evident in such cases. Ashutosh Shankhdhar has also talked about such a method [6] in which how a successful secure transmission can take place between two users by a lock-key encryption technique which can prove to be very useful in the coming age, since reliable transmission of data is highly needed now. The data once securely transferred, will needed to be trained. Though our requirement initially used to be of some same distribution data, but now the labelling functions may not necessarily be similar to the target data. This makes our method used different than the traditional methodologies in which it was believed that the training and test data should come from the same origin.

This paper deals with the methodology of inductive transfer learning, where both the actual and desired datasets have different labelling attributes, characteristics and functions used. We have taken an assumption that only a small quantity of the actual data, that is the training data is used, which is called the same disposition training data, and which uses information from the already available large amount of data obtained from different origins other than the target data, known as different disposition training data. This method, also, assumes that even though there might be changes in the labelling function of training and target sources, they are similar in some fashions so that some additional insights can be found in the different-distribution sources that are lacking in the same- distribution training data.

In the next section we have discussed the three classifiers that use the above mentioned type of training data sets. They are based on SVM (support vector machine) classification, where two of the three classifiers use sample weighting.

III. METHODS

The efficiency of three transfer classifiers is compared to the normally used SVM classifier in the further sections. A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. The hyperplane is a decision boundary which classifies a new data point and our aim is to maximise this decision boundary with the help of support vectors so that we can minimize the number of wrong classifications or reduce the number of new data points that lie within the boundary. The hyperplane is created by converting the problem into a linear function. It can be done as –

$$f(x) = B_o + S\,(A_i * (x_i, y_i))$$

where x_i is the training sample and y_i is the collection of support vectors. The coefficients B_o and A_i can be calculated from the training dataset provided. In cases, where the data might not be linearly separable we can use the kernel trick to create a soft-margin by plotting the data into a higher dimension to make it linearly separable. This has also been shown by Saruar Alam and Moonsoo Kang [5] in which they have used applied Principal Component Analysis (PCA) and kernel SVM to improve the accuracy of diagnosis methods being used and also achieved high sensitivity and specificity. Moreover, if we study further, we can also try to find a suitable kernel for the classifier, so that better separation can be done by the use of weighted feature selection [11].

A. Biased SVM

Biased SVM works by assigning a weight, greater than or equal to zero, to each point in the dataset to show its importance in making efficient classification for a new data point. Primary differences between the biased SVM and the normal SVM have also been discussed [8] like the upper bounds difference in the Lagrange multipliers in a dual problem. If a point helps then it is given more weightage, whereas outliers and likewise data points are given low weightage. Weights when assigned in the SVM Objective Function, results in –

$$Minimum\ ½\ |a||^2 + Constant \sum_{i=1}^{N} W_i E_i$$

One alternative to achieve transfer learning is by assigning a weight to S_t data points of one, and a weight of G_w to S_d data points. This will result in –

$$Minimum\ ½\ |a||^2 + Constant\ G_w \sum_{i:\ xi\ Ɛ\ S_d} E_i + Constant \sum_{i:\ xi\ Ɛ\ S_t} E_i$$

B. Transfer Adaboost

Transfer Adaboost mechanism works iteratively in a loop and adjusts the weights assigned to the data points according to how much value they are adding to the classifier. After a majority poll of the final resultant data points the classifier is achieved.

It works a little like the Biased SVM by assigning a weight Z_i to each training sample. After each loop, the weights are revised so that they all sum up to one, and that is how a classifier is looked as qualified. The next loops for assigning weights work by –

$$Z_i^{\,r+1} = \begin{cases} Zi^t\ £^{\,(1/2)|f\,t\,(xi)-yi|}, & for\ (x_i, y_i)\ Ɛ\ S_d \\ Zi^t\ £^{\,-(1/2)|f\,t\,(xi)-yi|}, & for\ (x_i, y_i)\ Ɛ\ S_t \end{cases}$$

where, £ is for calculating how the weights will be assigned after each loop. The weights of S_d samples will be reduced by £ which are not helping in the classification and the weights of S_t will be increased by £.

C. Adaptive SVM

Adaptive SVM stands out from the other two classifiers because it uses a seprate SVM classifier by training it on the T_d samples, and this classifier can also be used to normalize that SVM classifer which was trained using the T_s samples.

The SVM classifier is adjusted to the Ts samples via training the $\mu F(x)$ function, which uses $F^d(x)$ so that we get the desired SVM classifier F(x).

This can also be shown mathematically as –

$$F(x) - F^d(x) + \mu\ F(x)$$
$$- F^d(x) + a^T x + a_o$$

The a and a_o constraints are calculated from S_t by adjusting

$$Minimum\ ½\ |a||^2 + Constant^S \sum_{i=1}^{N} E_i$$
$$y_i F^d(x_i) + y_i(a^T x_i + a_o) \geq 1 - E_i$$
$$E_i \geq 0$$
$$¥\ (x_i, y_i)\ Ɛ\ S_t$$

Apart from also being used with the kernel, one major improvement of Adaptive SVM is that the SVM classifier which is trained on S_d samples needs to be trained only once, thus reducing the computations of the separator.

It also requires lesser memory consumption as comapred to the othe separators since all the data points need not be entered together.

Other than these three, there are some other variants too which can bring great improvements in their respective fields. Some of which have been discussed are the fuzzy SVM and the power SVM [4] which respectively train the orginal SVM by taking multiple instances of the exemplars and by modifying the constraints instead of the objective function. Zhang also proposed the R-SVM [7], also known as recursive SVM which uses recursive feature selection algorithm and can perform great on noisy data, making it applicable in datasets where there is a need to select multiclass features.

IV. RELATED WORK

By using the three SVM's discussed in the previous section along with the traditional method being used to perform segementaion on voxelwise classification of data acquired from mutliple sources, performance between the classifiers and also among them can be observed and proper diagnoisis of images can be achieved. This can be used in a variety of areas such as weather forecast, medical sector, aeronautical sector, satellite communication.

The effectiveness of this can be tested on one such area where it can be used to perform analogy on images for better derivations are heat maps. Heat maps generated form various sources can often be disturbing as they give different training distributions, functions and characteristics of the training data. This can create variations in the test data for a traditional SVM to work upon them. This can be solved with the help of Transfer Learning in which sample data obtained form various sources and scanners can be used even if they have different or small distributions as compared to the test data.

We have taken the data of different heat maps generated from various sites over a geographical area of 3000 km^2 and perfomed. We separated the dataset into two classes and used one dataset for training the classifier on different distributions and functions. The training dataset also included outliers to see how our classifier performs on misclassified data points. Using the traditional SVM methodology, we created a hyperplane and created the surrounding support vectors and denoted them with a circle. Also, we can also apply another technique here to reduce the number of points which may not serve as the support vectors [9], as has been discussed and experimented by Jie Meng. The area between the two dotted lines is the decision area and our initial aim is to maximise this area so that efficiency can be increased. The hyperplane boundary will initially be disoriented because of the outliers but the biased SVM and adaptive SVM performed greatly in spite of this. Also on removing the outliers, it was observed that the traditional SVM and biased SVM had almost the same observations.

From this, it can be interpreted that the traditional SVM and the our studied methods perform almost similarly when the dataset do not contain irrelevant data and further study to find out the reason behind this is going on. Still in spite of this, the adaptive SVM and the biased SVM has outperformed the traditional SVM greatly which sometimes generates errors because of the different imaging techniques being used.

The number of test damples in our data set were 1600 (1000 training data points and 600 testing data points). The value for our constant was evaluated via cross validation. All the contraints were set and a linear kernel was used to perform the experiment.

It was also observed that for small number of values the biased SVM outperformed the traditional mechanism greatly. Percenatge error of the different methods were compared to observe how they solve the outlier sensitivity problem. The traditional SVM's performance was mixed-up when the outliers were used in the datatset and its testing error increased. The biased and adaptive SVM's testing error was not so significantly increased and its performance was also not disoriented. Also, adaptive SVM worked the best when the number of samples were small as shown in the graphs and showed the least classification error, except when the number of samples were lesser than 20 than the biased SVM performed better than adaptive SVM. This shows that our proposed methods outweighs the performance of the traditional SVM being used and can give better classification results which further can be used to train efficient classifiers.

For more precision to be achieved in addition to this, selecting and training the boundary areas of a classifier to effectively select the support vectors can also be done, which has been discussed by Qi Mu [13], in which the B-ISVM method has been talked about which is an incremental approach to the normal SVM. This can be used to maximise the maximum decision boundary so that more datapoints can be classified effectively.

All this work has been done taking data from multiple sources so that the viability of this method can be checked and usefulness of this can be easily tested against other images which may be obtained from same souce through their segmentation protocols.

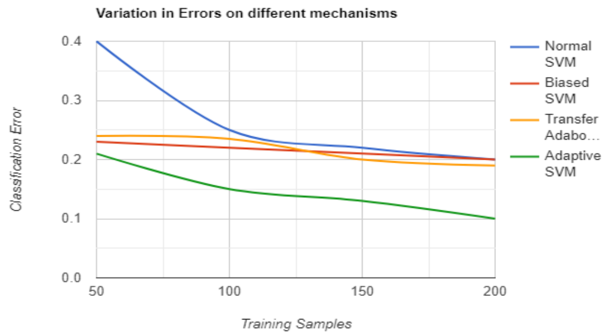

Fig 1. Classification Errors based on different methods

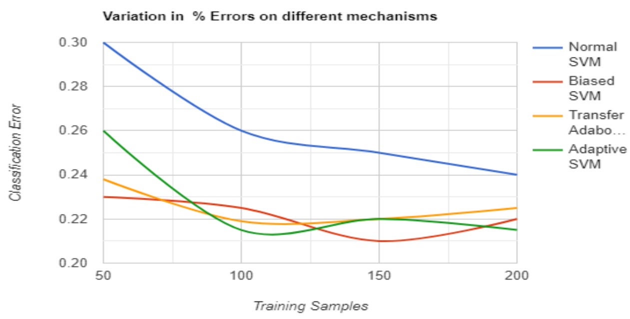

Fig 2. Percentage Errors based on different methods

Retrieval Number: C4060098319/19©BEIESP
DOI:10.35940/ijrte.C4060.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

V. CONCLUSION

This paper initially gives a retrospective view of how the conventional method was used for image segmentation generated from various imaging techniques and how the mechanism was not reliable when it comes to large and sparse data samples. Efficient training methods are then studied which focuses on training a classifier with deviated samples and with numbers of outliers in them, so that efficiency can be increased and distorted images coming from different sources no longer pose difficulty in diagnosing them. We have compared the performances of the three described classifiers used amongst them and along with that we have also paralleled the conventional method being used so that differences in their algorithms can be observed and evaluated along with the results assessed. On evaluation, it was observed that the modern methods performed greatly for the outlier sensitivity problem and the classifications errors were quite significant. The methodologies as described have used robust algorithms to achieve greater efficiency even for a small sample of data and also for data obtained from multiple sources, scanners and sites. This also helps greatly in diagnosing which data points are helping for predicting the target data and which are not.

The three different mechanisms which have been discussed thoroughly in thid paper other than the traditional SVM being used are the Biased SVM, Transfer Adaboost, and Adaptive SVM. These methods when performed resulted in lesser classification errors as compared to the normal SVM and gave better accuracy. It was also observed that when the desired precision was set, the number of training samples required were less by these classifiers which results in great improvement and flexibility. Adaptive SVM is the only classifier among them which does not take the S_d datasets explicitly into account while performing and training a classifier. It performed well on the segmentation of the given images and also outperformed all the other classifiers. The use of these methods can benefit greatly in areas of medical sector, where images often come from different sources and comprises of many outliers. Also, on increasing the number of attributes the proposed methods could bring great enhancement in many cases.

Furthermore, our future works consists of biomedical image segmentation and computed tomography for diagnosing cancer. It focuses on how the contaminated features disrupt the segmentation and the classifier is not trained properly. In reference to this, applaud worthy work has also been done where datasets may contain duplicate samples or attributes [10] by the use of a modified v-SVM, and also it is being used in neural networks too for restructuring and improving the precision of multisensory translation characteristics [12]. Also, in support to our efforts, work on sparse high-dimensional datasets has been in progress dealing with different misclassification costs among the features so that different distributions among the data can be dealt with and handled efficiently [14]. We think that by the use of these methodologies, various areas can benefit greatly and supervised learning can also be improvised to a great extent. Also, the application of these approaches is still yet to be investigated further so that their maximum potential can be utilized for human use.

REFERENCES

1. Dai Chunni, "SVM Visual Classification Based on Weighted Feature of Genetic Algorithm", 2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA).
2. Yan Yang, Juan Wang, Yongyi Yang, "Improving SVM Classifier with Prior Knowledge in Microcalcification Detection", 19th IEEE International Conference on Image Processing.
3. Parveen, Amritpal Singh, "Detection of Brain Tumor in MRI Images, using Combination of Fuzzy C-Means and SVM", 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN).
4. Ioannis Sarafis, Christos Diou, Theodora Tsikrika, Anastasios Delopoulos, "Weighted SVM from Clickthrough Data for Image Retrieval", 2014 IEEE International Conference on Image Processing (ICIP).
5. Saruar Alam, Moonsoo Kang, Jae-Young Pyun, Goo-Rak Kwon, "Performance of Classification Based on PCA, Linear SVM, and Multi-Kernel SVM", 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN).
6. Ashutosh Shankhdhar, Arushi Mangla, Prateeksha Chaturvedi, "Encoded Data Methodologies for Cloud Computing's Realm Security Enhancement", 2019 International Journal of Recent Technology and Engineering (IJRTE).
7. Qifeng Xu, Xuegong Zhang, "Multiclass Feature Selection Algorithms Base On R-SVM", 2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP).
8. Xulei Yang, Qing Song, Aize Cao, "Weighted Support Vector Machine for Data Classification", 2005 IEEE International Joint Conference on Neural Networks.
9. Shu-Xia Lu, Jie Meng, Gui-En Cao, "Support Vecctor Machine based on a New Reduced Samples Method", 2010 International Conference on Machine Learning and Cybernetics.
10. Yin-Shan Jia, Chuan-Ying Jia, Hong-Wei Qi, "A new NU-Support Vector Machine for training Sets with Duplicate Samples", 2005 International Conference on Machine Learning and Cybernetics.
11. Tinghua Wang, "Improving SVM Classification by Feature Weight Learning", 2010 International Conference on Intelligent Computation Technology and Automation.
12. Iryna Turchenko, Volodymyr Kochan, "Improvement of Identification Accuracy of Multisensor Conversion Characteristic using SVM", Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems.
13. Qi Mu, Yikun Chen, Yongjun Zhang, "Incremental SVM Algorithm to Intrusion Detection Base on Boundary Areas", 2012 International Conference on Systems and Informatics (ICSAI2012).
14. Sauptik Dhar, Vladimir Cherkassky, "Development and Evaluation of Cost-Sensitive Universum-SVM", IEEE Transactions on Cybernetics.

AUTHORS PROFILE

Vinay Agrawal, currently working in GLA University as an Assistant Professor. He has worked on various domains such as Machine Learning, Cloud Security, Data Science and Artificial Intelligence. Currently he is working on improving security among underwater sensor communication, Data Security Dependence etc.

Ashutosh Shankhdhar, currently working in GLA University as an Assistant Professor. His previous works have been in areas such as Data Science, Neural Networks, Cloud Computing Security etc. At present, he is researching more on Cloud Security and Supervised Learning.

Arushi Mangla currently pursuing Bachelors of Technology in Computer Science and Applications with specialization in Data Analytics from GLA University. She has worked on technologies such as Machine Learning, Data Mining, Data Visualization, and is currently doing her research on Data Security, Data Warehouse, Super-marts in developing areas.