Vol. 2 Issue IV, April 2014

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

A Semantic Concept Detection for Video Based on Regularized Extreme Learning Machine

Julin Rose Jacob

Dept of Computer Science
Karunya University
Coimbatore,India.

Abstract: This paper explains the concept detection based on the proposed method regularized extreme learning machine. In regularized extreme learning machine deals with the modification of the extreme learning machine to solve the missing data problems. Semantic concept detection is an important step in concept-based semantic video retrieval, which can be regarded as an intermediate descriptor to bridge the semantic gap. Support Vector Machines (SVM) and ELM (Extreme Learning Machine) is most existing methods. However, there are several drawbacks of using SVM, such as the high computational cost and large number of parameters to be optimized. The drawback facing by using ELM is some parameters are needed to be tuned manually. This consumes time for classification process. Instead of these disadvantages we use a proposed ELM called Regularized extreme learning machine (RELM) is used to detect semantic concept of videos. It uses a cascade of L1 penalty (LARS) and L2 penalty (Tikhonov regularization) on ELM (TROP-ELM) to regularize the matrix computation.

I.INTRODUCTION

In imaging science, image processing is any form of signal processing for which the input is an image, such as photography or a video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. A technique that used for developing powerful retrieval or filtering systems for multimedia data is semantic concept detection. Semantic concept detection is also known as high-level feature extraction. It can be processed by any supervised learning algorithm. Learning algorithms often assumes that the positive/negative data distribution is balanced. Semantic concept detection also refers to the task for assigning an input video sequence one or multiple labels indicating the presence of one or multiple semantic concepts in the video sequence. The existing concept detection methods are SVM and ELM as concept classifiers. There are several drawbacks of using SVM, such as high computational cost and large number of parameter to be used. The drawback facing by using ELM is some parameters are needed to be tuned manually. The semantic concept detection based on SVM classifier enhances practical performance and the accuracy of semantic concept detection still to be improved. The main objective of this paper is to detect the presence of semantic concepts in video shots using regularized extreme learning machine.In proposed work, Regularized extreme learning machine (RELM) is used to detect semantic concept of videos. The proposed method, which uses the advanced modification of the original extreme learning machine with a new method to solve the missing data problem. This method uses a regularized ELM algorithm[5], which uses a cascade of two regularization penalties: first a L1 penalty to rank the neurons of the hidden layer, followed by a L2 penalty on the regression weights(regression between hidden layer and output layer).Experimental result of proposed work provides better performance when compare with the existing work. The advantages are fast computational speed, no parameter needs to be tuned and it appears more stable and reliable generalization performance by the two penalties.

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

II.RELATED WORKS

In related works explains the existing methods of this paper. There are mainly two existing methods. They are Support Vector Machine (SVM) and Extreme Learning Machine (ELM). In support vector machine is used to analyse the semantic concept detection. SVM algorithm is the algorithm that used for the semantic concept detection. SVM [13] is the concept classifiers. In concept classifier is based on the visual features are manually labeled with video shots and also indicates the probability of a target concept that present in the video shots. SVM[13] are supervised learning models that are associated with learning algorithms. For improving the semantic concept detection the SVM detector is used. SVM detector is used for finding the prior probability of concepts in the video or the image.By using the SVM classifier we can used to train each SVM classifiers and train each features. There are different steps used for improving the detection score. First step is the detecting the detection score from the each individual SVM detectors. Then incorporates the concept and employ a probability prediction rule to estimate the detection. Finally apply a weighted linear combination to aggregate probabilities into final detection score. To produce final detection score,$V_i$by combining all the estimates [2][13].

$$V_i = \sum_{j=1}^{n} W_{ij} P_{Ej}(Y_i = 1)$$

Here W_{ij}denotes the weight of the estimate of the j^{th} detector forE_i. By using the TRECVID [13] data sets we can evaluate the SVM detectors. But SVM have several disadvantages they are high computational cost and large number of parameters need to be used.

Another related work is Extreme Learning Machine (ELM) explains that it is a single hidden layer feedforward networks that focused on input sets and training samples. ELM[13] generates the hidden layer output matrix. For past decades the application of neural networks are slower than the other techniques. It is slower because of two main reasons. The one reason behind this is slow gradient-learning algorithm and the second reason behind this is parameters are tuned by these algorithms. By avoiding these here we propose the algorithm is extreme learning machine for single hidden layer feedforward neural networks. By using this ELM algorithm can produce good performance and can learn faster than other algorithms. In single hidden layer feedforward network contain N hidden neurons. By using ELM we can determine the hidden nodes and find the output matrix. ELM is also good for multi-categories problem. To improve the accuracy of the semantic concept detection ELM is used based on the multi-modality classifier combination frameworks. It explains three different steps. First step explains that extracting the visual features like color, edge and texture. Besides the classical ELM methods ELM classifier is used to extracting each feature for training the datasets. In this step also used One-Against-All (OAA) [4][13] method is used. In second step explains prediction results and the probability fusion methods. The final results explain the probability of the prediction results in the concept of the video shots. By evaluating the ELM by using the mean average precision (MAP).In ELM the MAP ranges from 0.3 to0.5. By using the ELM there are some disadvantages in semantic concept detection. There are some missing values occurring and their some missing data problem occurs. By these disadvantages here proposed a method called Regularized Extreme Learning Machine(RELM).

III.PROPOSED APPROACH

In the proposed approach explains the Regularized extreme learning machine. There are some disadvantages in the existing methods .They are high computational cost, missing values occur. T o remove these disadvantages by adding regularized extreme learning machine. Regularaized extreme learning machine is used for detects the semantic concepts in the video.This proposed method is used as the advanced modification of the extreme learning machine. It is used to solve the missing data problem. In regularized extreme learning machines[5] uses a cascade of two regularization penalities: first a L1 penalty to rank the neurons of the hidden layer, followed by a L2 penalty on the regression weights(regression between hidden layer and output layer).

There are three features are extracted.They are color,edge and texture that are extracted by Grid Color Moment (GCM), Edged Direction Histogram (EDH), and Gabor Filters (GBR) respectively. Finally the proposed Regularized extreme learning machine (RELM) is used as a Multi-modality classifier. Experimental result of proposed work provides better

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

performance when compare with the existing work. The advantages are fast computational speed, no parameter needs to be tuned, and it appears more stable and reliable generalization performance by the two penalties. In remaining sections explains five different modules. The lists of modules are visual feature extraction, multi-modality classifier combination framework based on regularized elm[5] L1 penalty: LASSO andL2 penalty: Tikhonov regularization, Probability fusion method of multi-modality, inference of contextual correlation and performance evaluation. The main idea lying in ELM is the random weights of a single hidden layer feedfoward neural network (SLFN). The essence of ELM is that the hidden layer of SLFNs need not be tuned. Compared with those traditional computational intelligence techniques, ELM provides better generalization performance at a much faster learning speed and with least human intervention[6-9]. TROP-ELM is in order to handle missing data. The goal of using TROP-ELM is to take all the advantages of ELM like speed, and at the same time needs to be robust and more reliable. That is why we need the double regularization[10]. Here there is a flowchart explains the architecture of the regularized extreme learning machine in semantic concept detection. In the below flow chart explains the architecture of the RELM in semantic concept detection. In this explains that first giving as an input video. Then the video is segmented into different frames. After converting the video into frames then extracting the three features from the frames.They are color , edge and texture.These features are extracted by by Grid Color Moment (GCM), Edged Direction Histogram (EDH), and Gabor Filters (GBR) respectively. After extracting the features the applies RELM which is used as the multi-modality classifiers based on L1 penalty (LASSO) and L2 penalty(Tikhonov regularization). A regularized ELM algorithm, which uses a cascade of two regularization penalties, first a L1 penalty to rank the neurons of the hidden layer, followed by a L2 penalty on the regression weights (regression between hidden layer and output layer).This module introduces this algorithm briefly. In below explains the two regularization penalties. Pairwise distance estimation [5]efficiently estimates the expectation of the squared Euclidean distance between observations in data sets with missing data. Therefore, in general,it can be embedded into any distance-based method,like k nearest neighbors, support vector machine(SVM),multi-dimensional scaling (MDS), etc., to solve missing data problem.

L1 penalty:An important part in ELM is to minimize the training error $\|H\beta - y\|$ which is an ordinary regression problem. One technique to solve this is called Lasso, for 'least absolute shrinkage and selection operator[11]. Lasso solution minimizes the residual sum of squares, subject to the sum of the absolute value of the coefficients being less than a constant,that'swhyitisalsocalled'L1 penalty'. The general form which Lasso works on is[5]

$$min_{\lambda,\omega} \left(\sum_{i=1}^{N}(y_i - x_i\ \omega)^2 + \lambda \sum_{j=1}^{p} |\omega_j|\right)$$

Because of the nature of the constant, Lasso tends to produce some coefficients that are exactly 0 and hence give interpretable models. The shrinkage is controlled by parameter l. The smaller l is, the more oj coefficients are zeros and hence less variables are retained in the final model.

L2 penalty: Tikhonov regularization, named for Andrey Tychonoff, is the most commonly used method of regularization[12]. In statistics, the method is also known as ridge regression. The general form of Tikhonov regularization is to minimize[5]

$$min_{\lambda,\omega} \left(\sum_{i=1}^{N}(y_i - x_i\ \omega)^2 + \lambda \sum_{j=1}^{p} \omega j^2\right)$$

The idea behind of Tikhonov regularization[5] is at the heart of the ''bias-variance tradeoff'' issue, thanks to it, the Tikhonov regularization achieves better performance than the traditional OLS solution. Moreover, it outperforms the Lasso solution in cases that the variables are correlated. One advantage of the Tikhonov regularization is that it tends to identify/isolate groups of variables, enabling further interpretability. Since learning in the presence of missing data is pervasive problems in machine learning and statistical data analysis, we propose to extend ELM, particular TROP-ELM [10] in order to handle missing data. The goal of using TROP-ELM is to take all the advantages of ELM like speed, and at the same time, the method needs to be robust and more reliable. That is why we need the double regularization.

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

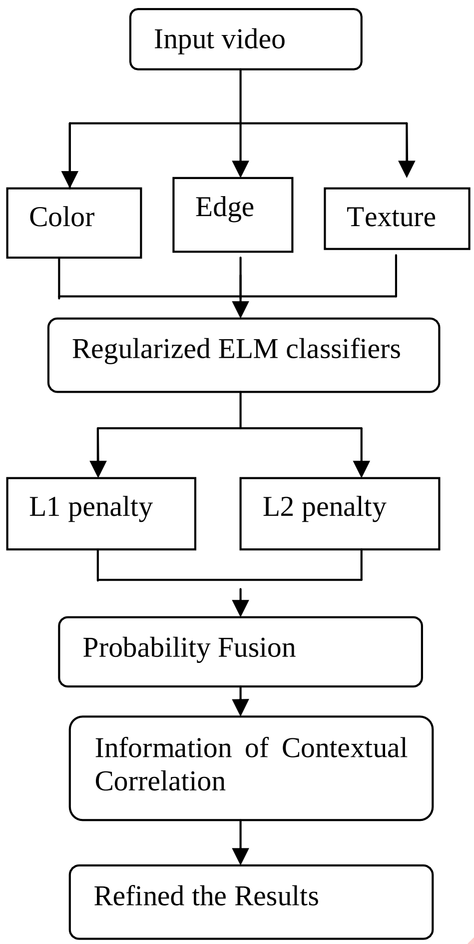

Fig1: Architecture diagram of RELM

IV. COMPARISION

By comparing the existing methods and the proposed methods we can easily identify that which the method acquires high performance in MAP improvements calculations. The following table contains the MAP improvements of the concept detections in various techniques. The methods are SVM[2], ELM[1] and RELM[5].

No	Methods	MAP improvements
1	SVM	0.167
2	ELM	0.179
3	RELM	0.2

V.CONCLUSION

In this proposed method RELM for detecting semantic concept of videos, Initially visual features such as color, edge, and texture can be extracted by Grid Color Moment (GCM), Edged Direction Histogram (EDH), and Gabor Filters (GBR) respectively. Then Multi-modality classifier combination framework based on Regularized extreme learning machine (RELM) is used multi-categories classification problem. This algorithm, which uses a cascade of two regularization penalties: first a L1 penalty to rank the neurons of the hidden layer, followed by a L2 penalty on the regression weights. This regularizes the matrix computations and hence makes the MSE computation more reliable, and on the other hand, it estimates the expected pair wise distances directly on incomplete data so that it offers the ELM a solution to solve the missing data issues.

REFERENCES

[1] Bo Lu,Guoren Wang,Ye Yuan,Dong Han,Semantic concept detection for video based on extreme learning machine,Neurocomputing 102(2013)176-183.

[2] Yusuf Aytar, O.Bilal Orhan and Mubarak Shah, "improving semantic concept detection and retrieval using contextual estimates"

[3] A. Yanagawa, W. Hsu, S.-F. Chang, Brief Descriptions of Visual Features for Baseline TRECVID Concept Detectors. Columbia University ADVENT Technical Report 219-2006-5, 2006.

[4] H.J. Rong, G.B. Huang, Y.-S. Ong, Extreme learning machine for multi- categories classificaiton applications, in: Proceedings of IEEE International Joint Conference on Neural Networks, 2008, pp. 1709–1713.

[5] Qi Yu, YoanMiche, EmilEirola , MarkvanHeeswijk , EricSe´ verin , AmauryLendasse, Regularized extreme learning machine for regression with missing data,in Neurocomputing 102 (2013) 45–51

[6] G.-B. Huang, D.H. Wang, Y. Lan, Extreme learning machines: a survey, Int. J. Mach.

ISSN: 2321-9653

Learn. Cybernet. 2 (2) (2011) 2107–2122.

[7] G.B. Huang, L. Chen, Enhanced random search based incremental extreme learning machine, Neurocomputing 71 (2008) 3460–3468.

[8] G.B. Huang, L. Chen, C.K. Siew, Universal approximation using incremental constructive feedforward networks with random hidden nodes, IEEE Trans. Neural Networks 17 (4) (2006) 879–892.

[9] G. B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: theory and applications, in: 2004 International Joint conference on Neural Networks (2004).

[10] Y. Miche, A. Sorjamaa, P. Bas, O. Simula, C. Jutten, A. Lendasse, Op-elm: optimally-pruned extreme learning machine, IEEE Trans. Neural Networks 21 (2010) 158–162.

[11] R. Tibshirani, Regression shrinkage and selection via the Lasso, J. R. Statist. Soc. 58 (1) (1996) 267–288.

[12] A.E. Hoerl, R.W. Kennard, Ridge regression: biased estimation for nonortho- gonal problems, Technometrics 12 (1) (1970) 55–67.

[13] Julin Rose Jacob,Survey on semantic concept detection, Vol. 1 (02), 2012, ISSN 2278 - 0181