Medical Image Classification Based on Curriculam Learning

S.Preethi Saroj, P.Balasubramanie, J.Venkatesh

***Abstract:** With the emergence of large medical images and exceptional growth of diagnostic methods, categorizing them into respective class has always been a dominant topic in computer vision. Though the system seems ubiquitous, achieving higher accuracy rates for classification is critical. Semi-Supervised Learning (SSL) is better than supervised learning as it eliminates labeling all images thus reducing computational cost and time. Existing methods suffer from classification accuracy due to the presence of outliers in critical images. This paper is an attempt to apply SSL through Multi-Modal Curriculum Learning (MMCL) strategy over medical images. Through this, medical images can be categorized into normal and abnormal images. Experimental results demonstrate good accuracy for classification.*

***Keywords:** Medical imaging, Semi-Supervised Learning, Multi-modal Curriculum, Pyramid Histogram of Gradients*

I. INTRODUCTION

With the advent of technology and growth in size, medical image database has given rise to massive amount of images. It is necessary to categorize the images into appropriate classes. This multi-class classification problem addresses two great challenges. First, though having adequate images for training improves classification accuracy, obtaining labels is time consuming process. Second, as the data comes from multiple sources, integrating heterogeneous features is difficult. With supervised image classification the available labelled images are insufficient to train since it has to deal with dramatic growth of images. Furthermore, it involves more time in labelling the images. In medical applications, images are obtained from various sources hence supervised learning cannot be adopted. The need to reduce the label assigning time and classification accuracy is the motivation towards choosing Semi-Supervised Learning(SSL). SSL is capable of learning from limited labelled image and classifies unlabeled data.

Classifying images with limited label information and also eliminating higher computational time has become major area of interest in image processing domain. Gigantic quantities of images grow exponentially day-by day increasing image databases. But all the images appended are not labelled. So semi-supervised learning is appropriate for medical applications. This requires only limited labelled images and with itself it can start the learning process.

Revised Manuscript Received on June 22, 2019.

S.Preethi Saroj, Research Scholar, Dept. of CSE, Kongu Engineering College,Erode

P.Balasubramanie, Professor, Dept. of CSE, Kongu Engineering College,Erode

J.Venkatesh, Associate Professor, Dept. of Management Studies, Anna Universtiy, Coimbatore

Though unlabeled images do not contain explicit labels, they give distribution information of entire dataset helping accurate classification. Original input images taken into account for image classification task may contain unnecessary parts. This burdens the size of database and may slowdown the process. While analyzing the data, unwanted features may also be considered. To overcome this limitation image segmentation is employed. It simplifies the representation of image, making it easier for analysis and for locating boundaries, curves or lines. Feature extraction extracts possible image characteristics from the image.

It reduces the amount of resources required to describe large set of data. Medical image may contain varying range of features due to its orientation, scale, intensity or patter of pixel distribution. Extracted features form feature vector and set of all feature vector constitute the feature subspace. PHOG feature descriptor counts for occurrence of gradient orientation in localized region of image. This descriptor extracts features from specified region of interest.

Images are obtained from multiple sources. To process this multi-modal data, extracted feature vectors have to be concatenated into a long vector. But forming and managing this vector is difficult and it may affect the performance. Hence multi-modal learning is selected that takes only correlative features and fuses them. Curriculum learning aims at presenting the image information in a organized fashion. It starts with a small task , learns its easier aspects and then gradually increments the difficulty level. It involves adaptive value of starting, where experimental results show that classification learning does not depend on knowledge entropy but on initiating with limited set having limited information and expanding it gradually as the learning progresses.

Classification accuracy is highly affected by the presence of outliers as they misclassify the images. But through repetitive iterations and feedback this problem is eliminated. Pyramid Histogram of Oriented Gradients (PHOG) feature descriptor is chosen that extracts features from labelled and unlabeled images. MMCL analyses the features relation and builds the curriculum. While integrating the curriculum the labels are propagated from labelled set to unlabeled set. Thus the integrated curriculum helps in classifying the unlabeled images.

II. RELATED WORK

Image segmentation aids to separate the desired foreground object from its background eliminating unwanted boundaries of the image. When foreground and background have similar features it is difficult to segment the image. To overcome this issue [11] proposed interactive image segmentation that incorporates simple user interaction into image segmentation in a supervised or semi-supervised manner. It gets prior information from user that forms a prime part in guiding image segmentation. The simple inputs can cause chaos that can decline object boundary preservation. Adaptive Constraint Propagation (ACP) overcomes this limitation by employing kernel matrix that adaptively propagates the information about the entire image without affecting its coherence. Traditional watershed algorithm over segments the image. [8] Handled this problem by hybrid gradient and self-adaptive marker extraction. Gradient of the image is calculated using colour space histogram and gradient function is improved through information entropy theory. From the low frequency region of image gradients, adaptive marker approach extracts the local minima. These weak edges can be effectively enhanced. Thus without manual setting of parameters, image segmentation can be carried out in an efficient way.

A major problem in feature extraction is that medical image dataset are not uniform because of its orientation, scale or pattern of pixel distribution. [12] Developed generalized grey scale and rotation invariant operator for detecting uniform patterns in images. Edges and corners are primitive micro-features that correspond to most frequent uniform binary patterns and the underlying distribution in given by histogram. LBP first divides the images into cells. For each pixel in cell, LBP compares it with surrounding 8 neighbours either in clockwise or counter-clockwise. Histogram over the cell is computed based on whether the pixel is greater or smaller than the 8 neighbour cells. After normalizing the histograms, LBP integrates all histograms resulting in feature vector of the image. [1] uses pyramid feature descriptor that represents the image shape and its spatial layout along with its pyramid kernel. It also generalizes the spatial pyramid and learns its level weighting parameters. This critically improves classification performance. It represents the image by its local shape and spatial layout of the shape. Histogram of edge orientations within specified region is quantized into K-bins representing local shape. Spatial layout is obtained by tilting the image into regions at various resolutions. Histogram of Oriented Gradients when represented in multiple resolution levels collectively forms PHOG.

Semi-supervised learning method (SSL) learns in the presence of both labelled and unlabeled data. The goal of semi-supervised learning is to understand the labelled and unlabeled data thereby improving learner behaviour. SSL is widely used in machine learning and data mining since it can easily use available unlabeled data to improve the classification task when label data are scarce or expensive. [4] Classifies images with available labelled and unlabeled data using ensemble of prototypes and then learn a discriminative feature representation of an unlabeled image by calculating its projected values from previous samples. [2] Proposes a scalable graph-based algorithm and its complexity varies linearly with image dataset size. It also analyzes the results for high efficient approximation in semi-supervised learning. It adopts graph-based semi-supervised learning. With unlabeled data it forms a graph and edges are taken as nxn matrix. The matrix is diagonalized to get Eigen vectors. Eigen values of the graph solves semi-supervised problem in a reduced dimensional space.

The goal of image classification is to decide whether an image belongs to certain class or not. [6] uses the tags associated with labelled and unlabeled images to improve classification accuracy through semi-supervised learning. Tags of the image might not be directly related to them. A strong classifier learns from labelled images. It takes both image labels and tags as features. Multiple Kernel Learning framework (MKL) framework combines kernel with label and kernel with encoded tag. This result aids in predicting the labels of unlabelled training images. Both labelled data and output of classifier on unlabeled data are used to learn another classifier that takes only visual features as input. Testing samples are examined with the final classifier.

Matrix Completion (MC) is another method for classification. It effectively deals with missing data and noise in feature and label space. Besides its advantages, MC cannot handle image classification that has heterogeneous features. One solution to this problem is to concatenate all features into a long vector. But this may lead to over-fitting problem and results in very long matrix increasing the computing time. To overcome this problem, [10] combines MC based classification outputs of different views and developed new framework namely multi-view matrix completion. This can handle wide range of features. Two-fold cross validation is performed on labelled set to calculate the coefficients. The method divides labelled training samples into two sets. Assumption is made such that one set is unknown and the labelled set information is propagated to the other. Thus combination co-efficient of different view are learned resulting in optimized precision for the multi-label classification.

Curriculum learning aims at presenting the image information in a organized fashion. It starts with a small task learn its easier aspects and then gradually increments the difficulty level. [9] proposes self-paced approach for visual category discovery. It focuses on easiest objects first and progressively expands to include complex objects. At each iteration, easiness of each sub-window having large unlabeled images are re-estimated. Then single prominent cluster among the easiest is selected. As system gradually accumulates models, new one learns from the early models. Experimental result shows that starting with easier instances and then gradually including complexity results in better learning.

Label propagation plays a prime part in supervised learning that aims in classifying massive number of unlabeled images with relatively few labelled images. Experiments are carried out to study how label information is propagated in graph and ways to improve the speed. Existing algorithms treat unlabeled images equally. Transmission of labels from labelled to unlabeled images happen with the help of neighbourhood matrix. But this method does not hold good in presence of outliers. So [5]

Treats unlabeled samples at various levels of difficulty. Itproposes iterative label propagation algorithm where it involves teaching and learning process. Here the learner propagates the labels to unlabeled examples assigned by the teacher. Learner's feedback to adjust the subsequent changes are incorporated by the teacher. Teaching-to Learn and Learning to Teach (TLLT) algorithm improves the accuracy of label propagation making it robust while tuning parameters.

III. METHODOLOGY

This works aims to classify medical image using semi-supervised learning. First image segmentation is done to identify major parts that are required for learning. Image segmentation makes images easier for analyzing. Then the segmented image features are extracted. Various feature descriptors are available namely Local Binary Pattern(LBP), Speeded Up Robust Features(SURF), Pyramid Histogram of Oriented Gradients(PHOG) are available. PHOG is selected for this work since it is simple, fast and suits for larger datasets. We can specify the region of interest in images. PHOG counts occurrences of gradient orientation in that specified region. Computed gradient ensures the normalization in images. It also involves bin orientation. Histogram levels, number of bins and angle need to be assigned for the descriptor. Angle can have value 180 or 360 depending on whether gradient is unsigned or signed respectively. Here the gradients are unsigned and hence takes the value 180. PHOG computes histogram matrix with same size of the image where (i,j) position contains the histogram value for pixel at position (i,j) and another gradient matrix where the position (i,j) contains gradient value for pixel at position(i,j).

Adjacency matrix represents the similarity between features in terms of features. Laplacian matrix gives difference between degree matrix and adjacency matrix. The parameter $\varkappa 2$ controls the sharpness of Gaussian distribution. Alternating Direction of Multipliers (ADM) algorithm is adopted which splits the images into multiple quadrants. Conditional entropy is a measure that quantifies the amount of information needed to describe the outcome of random variable given that value of another random variable is known. Here conditional entropy is small for simple curriculum for the given label set indicating its reliability. Whether an image belongs to a class or not is provided by average commute matrix.

In each iteration, labels are propagated from labelled set to unlabeled set. But only distinct images can be selected eliminating the duplicates. For this purpose binary selection matrix is used and it ensures each image is selected only once. To improve the performance, single modal curriculum is expanded to multi-modal curriculum. For that purpose, results of multiple single curriculum results are collected and integrated into a single curriculum. Resultant curriculum becomes optimal after performing standard alternating minimization. Two major constraints are carried out in the process. First, the images with simple features are selected. Second, all simple feature sets formed has to agree with each other. Forbinus norm is computed between single curriculum and optimal curriculum to find the sum of absolute squares.

Augmented Lagrangian Multiplier tries to eliminate constraints in sub-problem by introducing penalty function. This speeds up the convergence rate and computational time. Greedy approach converts optimal curriculum matrix into binary values. When overall optimal curriculum is found, labels are propagated from labelled set to unlabeled set. Suppose in nth iteration, nth label propagation happens and in subsequent iteration (n+1)th curriculum is framed and the process continues through which image classification manages to attain higher accuracy.

IV. RESULTS

The algorithm is evaluated for correct classification of unlabeled images. These metrics helps in analyzing the efficiency of algorithm and guide towards finding right direction for fine tuning system parameters so as to obtain optimized results.

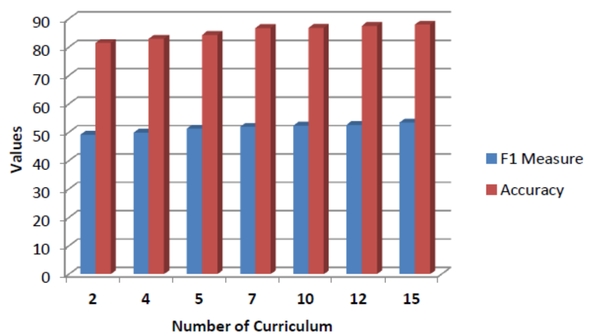
F1 measure is best suited metric for finding out the error ratio between actual and predicted labels. Precision is the number of true positives and false positives i.e. number of positive predictions under the positive class values predicted. On the other hand, recall is the number of positive predictions divided by the number of positive class values in the test data. Thus precision is measure of classification exactness and recall is measure of classification completeness. F1 measure portraits the balance between precision and recall. Correct Classification Rate(CCR) gives the accuracy i.e. extent to which images are correctly classified.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

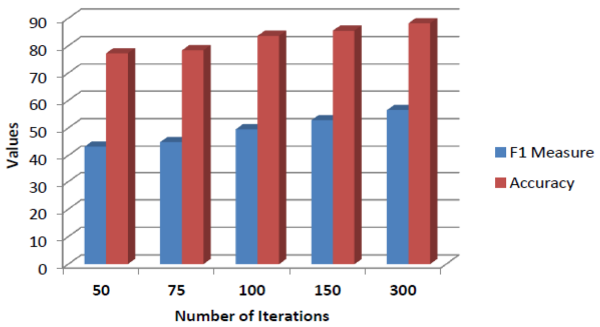
$$F1\ Measure = \frac{2.Precision.Recall}{Precision + Recall}$$

$$Accuracy = \frac{TP + TN}{P + N}$$



Through curriculum, unlabeled images are logically classified. Curriculum learning helps to reorganize the learning sequence. By this unlabeled images can be reliably labelled as they start from simple labelled data.

For each curriculum, binary selection matrix is computed to avoid duplicate instances. This binary selection matrix is calculated for each curriculum and each of it holds different features. Increasing the curriculum helps to better organize unlabeled data. Fig. 4.1 shows that by rising the number of curriculums, better results are obtained. All the experiments are carried out for 100 iterations.



Influence of number of iterations

Iterations play a prime part in boosting up performance in learning process. Through iterations only, classification process is made efficient. In iteration, label propagation is carried and label matrix is computed. By this improvement in iterations better label matrix is got. Experiments are carried out by fixing the number of curriculum as 5. Fig. 4.2 depicts the performance of the algorithm by varying the number of iterations. When number of iterations is increased, F1 measure and accuracy gives steady state improvement. But when the iterations is carried out for larger times it takes more time to converge.

V. CONCLUSION

With this approach, features of labelled and unlabeled images are exploited based on which curriculum is framed. This work involves the feature extraction using the Pyramid Histogram of Oriented Gradients. Iterative propagation helps to achieve better accuracy in image classification. This work also employs curriculum learning using supervised classification learning also. From the test results, multi-modal curriculum learning provides better accuracy rates compared to other standard algorithms. Thus with large-scale image datasets, image classification is successfully carried out with scarce labelled data.

The experiments show some limitations in this approach. From the results, we could see that accuracy and F1 Measure with better values is achieved through numerous iterations. But inclusion of numerous iterations has lead to longer computational time. Hence it would have been better if alterations can be made to this step so as to lessen computation time. Further this methodology does not classify images with noisy label cases. Further improvement can be made to handle incorrect labels in noisy images. Performing classification of noisy images involves correcting labels through iterative propagations.

REFERENCES

1. Anna Bosch, Andrew Zisserman and Xavier Munox. (2007), ‘Representing shape with a Spatial Pyramid Kernel’, ACM International Conference on Image and Video retrieval, pp. 401-408.
2. Bengio Y, Louradour J, Collobert R, and Weston J. (2009), ‘Curriculum Learning’, International Conference Machine Learning(ICML), pp. 41-48.
3. Cai X, Nie F, Cai W and Huang H (2013), ‘Heterogeneous image features integration via multi-modal semi-supervised learning modal’, IEEE International Conference on Computer Vision, pp. 1737-1744.
4. Dai D and Van Gool L. (2013), ‘Ensemble projection for Semi-Supervised classification’, IEEE International Conference on Computer Vision, pp. 2072-2079.
5. Gong C, Tao D, Liu W, Liu L and Yang J. (2016), ‘Label Propagation via Teaching-to-Learn and Learning-to-Teach, IEEE Transactions on Neural Networks and Learning Systems, Vol. 28, No. 6, pp. 1452-1465.
6. Guillaumin M, Verbeek J and Schmid C. (2010), ‘Multimodal Semi-supervised learning for image classification’, IEEE Conference on Computer Vision and Pattern Recognition , pp. 902-909.
7. Jun Li, Yu Kong, Handong Zhao, Jian Yang and Yun Fu. (2016), ‘Learning Fast Low-Rank Projection for Image Classification’, IEEE Transaction on Image Processing, Vol. 25, No.10, pp 4803-4815.
8. Li Yuan, QingSong Yu, Chaomin Shen, Wenxin Hu and Zhitong Yang. (2016), ‘New Watershed Segmentation algorithm based on Hybrid Gradient and Self-Adaptive Marker Extraction’, IEEE International Conference on Computer and Communications, pp 624-629.
9. Liping Jing, Chenyang Shen, Liu Yang, Jian Yu, and Michael K. (2017), ‘Multi-Label Classification by Semi-Supervised Singular Value Decomposition’, IEEE Transaction on Image Processing, Vol. 26, No.10, pp 4612-4626.
10. Luo Y, Liu T, Tao D and Xu C. (2015), ‘Multiview matrix Completion for multilabel image classification’. IEEE Transactions on Image Processing, Vol. 24, No. 8, pp. 2355-2368.
11. Meng Jian and Cheolkon Jung. (2016), ‘Interactive image segmentation using Adaptive Constraint Propagation’, IEEE Transaction on Image Processing, Vol. 25, No. 3, pp 1301-1311.
12. Ojala T, Pietikainen M, Maenpaa T. (2002)‘Multiresolution gray-scale and rotation invariant texture classification with local binary patterns’, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, pp 971-987.
13. Qiang Li, Bo Xie, Jane You and Wei Bian. (2016), ‘Correlated Logistic Model With Elastic Net Regularization for Multilabel Image Classification’, IEEE Transaction on Image Processing, Vol. 25, No. 8, pp 3801-3813.
14. Seungryong Kim, RuiCai, Kihong Park, Sunok Kim and KwanghoonSohn. (2017), ‘ Modality-Invariant Image Classification Based on Modality Uniqueness and Dictionary Learning’, IEEE Transaction on Image Processing, Vol. 26, No. 2, pp 884-900.
15. Zhu X and Ghahramani Z. (2007), ‘Learning from labeled and unlabeled data with label Propagation’, School of Computer Science, Carnegie Mellon Univerrsity, Pittsburgh, PA, USA, Tech. Rep. CMU-CALD-02-107