STROMAL TUMOR PATTERN CLASSIFICATION USING SVM LEARNING WITH GLCM FEATURE EXTRACTION

K. KOWSALYA[1], Mrs. M. KAVITHA[2],
[1]M.Phil Research Scholar, [2]Assistant Professor,
PG & Research Department of Computer Science,
Tiruppur Kumaran College For Women, Tirupur.

Abstract: Predicting malignant potential is one of the most critical components of a computer-aided diagnosis (CAD) system for Gastro Intestinal Stromal Tumors (GISTs). These tumors have been studied only on the basis of subjective Computed Tomography (CT) findings. Object detection plays a major role in many areas like medical imaging, aerial surveillance, optimal manipulation and analysis, surgical microscopes, etc. The objective of this paper is to develop a model for brain tumors detection and classification i.e., to classify whether the tumor is cancerous or non-cancerous using SVM algorithm. Earlier many have detected using GLCM feature extraction which works on Empirical Risk Minimization. We are using Support Vector Machine algorithm that works on structural risk minimization to classify the images. The SVM algorithm is applied to medical images for the tumor extraction, and a Simulink model is developed for the tumor classification function. This paper presents a prototype for SVM-based object detection, which classifies the images and evaluates whether the classified image is cancerous or non-cancerous. Our proposed classification techniques based on Support Vector Machines (SVM) and histogram based image segmentation are proposed and applied to brain image classification. Here feature extraction from Gastro tumor images will be carried out by gray scale, symmetrical and texture features. This intelligent system improves accuracy rate and reduces error rate of Gastro intestine tumor classification using SVM.

Index Terms - Gastro Intestinal Stromal Tumor, Benign and Malignant images, GLCM Feature Extraction, SVM classifiers.

I. INTRODUCTION

A Gastro Intestinal Stromal Tumor (GIST) is a type of tumor that occurs in the gastrointestinal tract, most commonly in the stomach or small intestine. The tumors are thought to grow from specialized cells found in the gastrointestinal tract called interstitial cells of Cajal (ICCs) or precursors to these cells. GISTs are usually found in adults between ages 40 and 70; rarely, children and young adults develop these tumors. The tumors can be cancerous (malignant) or noncancerous (benign).Small tumors may cause no signs or symptoms. However, some people with GISTs may experience pain or swelling in the abdomen, nausea, vomiting, loss of appetite, or weight loss. Sometimes, tumors cause bleeding, which may lead to low red blood cell counts and, consequently, weakness and tiredness. Bleeding into the intestinal tract may cause black and tarry stools, and bleeding into the throat or stomach may cause vomiting of blood. Gastro Intestinal Stromal Tumors (GISTs) are soft-tissue sarcomas that can be located in any part of the digestive system. Their most common sites are the stomach and small intestine. GISTs start in specialized nerve cells located in the walls of your digestive system. These cells are part of the autonomic processes as movement of food through the intestines, gives rise to a GIST. Determinant clinical characteristics and subjective CT features were assessed to separately construct a corresponding model. The models showing statistical significance in a multivariable logistic regression analysis were used to develop a nomogram. The diagnostic performance of these models was evaluated using ROC curves. Further calibration of the nomogram was evaluated by calibration curves. The breast stromal micro environment is a pivotal factor in breast cancer development, growth and metastases. Although pathologists often detect morphologic changes in stroma by light microscopy, visual classification of such changes is subjective and non-quantitative, limiting its diagnostic utility. Furthermore, without being trained specifically using ductal carcinoma in situ as an outcome, the algorithm detected tumor-associated stroma in greater amounts and at larger distances from grade 3 versus grade 1 ductal carcinoma Collectively, these results suggest that algorithms based on deep convolutional neural networks that evaluate only stroma may prove useful to classify breast biopsies and aid in understanding and evaluating the biology of breast lesions.

II. LITRATURE REVIEW

B. P. RUBIN says Gastrointestinal stromal tumors are the most common mesenchymal neoplasm of the gastrointestinal tract and are highly resistant to conventional chemotherapy and radiotherapy. Such tumors usually have activating mutations in either KIT (75–80%) or PDGFRA (5–10%), two closely related receptor tyrosine kinases. These mutations lead to ligand-independent activation and signal transduction mediated by constitutively activated KIT or PDGFRA. Targeting these activated proteins with imatinib mesylate, a small-molecule kinase inhibitor, has proven useful in the treatment of recurrent or metastatic gastrointestinal stromal tumors' and is now being tested as an adjuvant or neo adjuvant [1]

S. YAMAGUCHI says Small Submucosal Tumors (SMT) without symptoms are frequently found by endoscopic and radiological examinations. To find proper diagnostic measures and therapeutic indications for histologically undiagnosed SMT, we reviewed published articles in PubMed between 1990 and March 2013 using the key words 'submucosal tumor' and the name of a specific disease. SMT is observed in a wide range of gastrointestinal (GI) diseases and conditions, including compression by extra-GI organs and lesions, congenital tumors, inflammation, and benign as well as malignant neoplastic lesions. [2]

BEHAM says the gastrointestinal stromal tumor (GIST) is the most common mesenchymal tumor of the intestinal tract, known to be refractory to conventional chemotherapy or radiation. Still, the only curative option for GIST is given after complete surgical removal even in a metastatic setting, but recurrence is common, and the risk can be defined by surgical factors like incomplete resection, intraperitoneal rupture, or bleeding and tumor associated factors like tumor size, mitotic index, or localization. Conclusion consequently, adjuvant therapy with imatinib mesylate or other tyrosine kinase inhibitors is recommended for high-risk patients after complete resection. For unresectable and advanced GIST, a partial response or stable disease can be achieved in about 80% of patients with imatinib mesylate [3]

HYUN SEOK LEE says Gastro Intestinal (GI) Sub Epithelial Tumors (SETs) are detected incidentally by endoscopy. GI SETs are of various types, with varied prognoses ranging from benign to aggressive and a malignant potential. Gastro Intestinal stromal tumors (GISTs) are the most commonly identified mesenchymal tumors in the stomach, and 10% to 30% of GISTs have a malignant potential. 3,4 Although several features on EUS are specific for the prediction of the malignant potential of GISTs, the diagnostic accuracy of EUS findings varied widely, even with tissue sampling by EUS guided fine-needle aspiration, EUS guided Tru-Cut biopsy, and recently, EUS-guided fine-needle biopsy. Moreover, compared with computed tomography, EUS with contrast enhancement showed limited efficacy in characterizing lesions [4]

NISHIDA says although gastrointestinal stromal tumors (GISTs) are a rare type of cancer, they are the commonest sarcoma in the gastrointestinal tract. Molecularly targeted therapy, such as imatinib therapy, has revolutionized the treatment of advanced GIST and facilitates scientific research on GIST. Nevertheless, surgery remains a mainstay of treatment to obtain a permanent cure for GIST even in the era of targeted therapy. Many GIST guidelines have been published to guide the diagnosis and treatment of the disease. We review current versions of GIST guidelines [5]

LIU.S says Seventy-eight patients with histopathologically confirmed GISTs underwent preoperative CT. Texture analysis was performed on unenhanced and contrast-enhanced CT images, respectively. Fourteen CT texture parameters were obtained and compared among GISTs at different malignancy risks with one-way analysis of variance or independent-samples Kruskale Wallis test. Correlations between CT texture parameters and malignancy risk were analysed with Spearman's correlation test. Diagnostic performance of CT texture parameters in differentiating GISTs at low/very low malignancy risk was tested with receiver operating characteristic (ROC) analysis [6]

C. ZHOU says Purpose to determine the predictive CT imaging features for risk stratifications in patients with primary Gastro Intestinal Stromal Tumors (GISTs). Materials and methods One hundred and twenty-nine patients with histologically confirmed primary GISTs (diameter >2 cm) were enrolled. CT imaging features were reviewed. Tumour risk stratifications were determined according to the 2008 NIH criteria where GISTs were classified into four categories according to the tumour size, location, mitosis count, and tumour rupture. The association between risk stratifications and CT features was analyzed using univariate analysis, followed by multinomial logistic regression and Receiver Operating Characteristic (ROC) curve analysis. [7]

PHILIPPE LAMBIN Imaging is an important technology in medical science and is used in clinical practice to aid decision making. The role of medical imaging, however, is swiftly evolving from being primarily a diagnostic tool to also include a central role in the context of personalized precision medicine. In radiomics, digitally encrypted medical images that hold information related to tumour path physiology are transformed into mineable high-dimensional data. The process used in radiomics involves the identification of vast arrays of quantitative features within digital images, storage of such data in federated databases (that is, a system in which several independent databases function as a single entity) and the subsequent mining of the data for knowledge extraction and application. [8]

M. VALLIÈRES says this study aims at developing a joint FDG-PET and MRI texture-based model for the early evaluation of lung metastasis risk in Soft-Tissue Sarcomas (STSs). We investigate if the creation of new composite textures from the combination of FDG-PET and MR imaging information could better identify aggressive tumors. The incorporation of features into multivariable models was performed using logistic regression. The multivariable modeling strategy involved imbalance-adjusted bootstrap resampling in the following four steps leading to final prediction model construction: feature set reduction; feature selection; prediction performance estimation; and computation of model coefficients. [9]

M. ZHOU says To create a radiogenomic map linking Computed Tomographic (CT) image features and gene expression profiles generated by RNA sequencing for patients with Non–Small Cell Lung Cancer (NSCLC). Materials and Methods: A cohort of 113 patients with NSCLC diagnosed between April 2008 and September 2014 who had preoperative CT data and tumor tissue available was studied. Next, total RNA was extracted from the tissue and analyzed with RNA sequencing technology. Ten highly coexpressed gene clusters, termed metagenes, were identified, validated in publicly available gene-expression cohorts, and correlated with prognosis. [10]

III. RELATED WORK

In existing system the comprehensive survey of existing tumor enhancement and Classification techniques. Each method is classified, analyzed, and compared against other approaches. To examine the accuracy of the tumor enhancement and Classification techniques, the sensitivity and specificity of the approaches is presented and compared where applicable. Finally, this research provides taxonomy for the available approaches and highlights the best available enhancement and Classification methods. It only categorized tumor Classification techniques into mass detection using a single view and mass detection using multiple views. The mass detection using single view Classification in turn is divided into four categories: model-based methods, region-based methods, contour-based methods, and clustering methods. To the best of our knowledge, this study is the first time to use radiomics model and CNNs for GISTs classification. Analysis and comparison between the two methods are conducted in detail.

IV. PROPSOSED METHODOLOGY

The proposed system Support Vector Machine with GLCM matrix function is chosen in order to distinguish the interior area from other organs in the MR image dataset. Then modified gradient magnitude region growing algorithm is applied, in which gradient magnitude is computed by Median Filter and employed as the definition of homogeneity criterion. This implementation allowed stable boundary detection when the gradient suffers from intersection variations and gaps. By analyzing the gradient magnitude, the sufficient contrast present on the boundary region that increases the accuracy of Classification. The first stage is to determine the input image labels and the number of pixels in each label. The second stage is to determine the output requested region to get total number of pixels accessed. Segmented areas are automatically calculated and to get desired tumor area per slice.

4.1 IMAGE PREPROCESSING

Preprocessing images commonly involves removing low frequency, background noise, normalizing the intensity of individual practical images, removing reflections and masking portion of images. Image processing is the technique of enhancing data images prior to computational processing. Following standard preprocessing steps for Gastro intestinal Scan, the corresponding fractal and intensity features are extracted. In the next step, different combinations of feature sets are exploited for tumor Classification and classification. Feature values are then directly fed to the AdaBoost classifier for classification of tumor and non-tumor regions. Manual labeling to tumor regions is performed for supervised classifier training. The trained classifiers are then used to detect the tumor or non tumor segments in unknown Gastro Intestinal Scan.

4.2 GLCM FEATURE EXTRACTION

In this module the Gray Level Co-Occurrence Matrix based Feature extraction is a special form of Dimensionality reduction. When the input data to an Algorithm is too large to be processed and it is suspected to be notoriously redundant (e.g. the same measurement in both feet and meters) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. The below steps to find Feature Extraction for each image sets.

4.3 GASTRO INTESTINAL TUMOR CLASSIFICATION AND CLASSIFICATION FROM NON-TUMOR TISSUE

In CRF kernels functions are used such as graph kernel, polynomial kernel, RBF kernel etc. Among these kernel functions, a Radial Basis Function (RBF) proves to be useful, due to the fact the vectors are nonlinearly mapped to a very high dimension feature space. For tumor/non-tumor tissue segmentation and classification, Scan pixels are considered as samples. These samples are represented by a set of feature values extracted from different Scan modalities. Features from all modalities are fused for tumor segmentation and classification. A modified supervised CRF ensemble of classifier is trained to differentiate tumor from the non-tumor tissues.

4.4 GASTRO INTESTINAL TUMOR CLASSIFICATION USING STRUCTURE PREDICTION

The method proposed for Classification of particular structures of the Gastro intestinal tumor, i.e. whole tumor, tumor core, and active tumor, is evaluated. This method is based on an approach, whose novelty lies in the principled combination of the deep approach together with the local structure prediction in medical image Classification task. There are two type of classification we displaying benign and the other as malignant. A benign tumor is a tumor that does not invade its surrounding tissue or spread around the body. A malignant tumor is a tumor that may invade its surrounding tissue or spread around the body.

4.5 DIGITAL IMAGE PROCESSING

Digital image processing deals with manipulation of digital images through a digital computer. The most common example is Adobe Photoshop. It is one of the widely used applications for processing digital images. Signal processing is a discipline in electrical engineering and in mathematics that deals with analysis and processing of analog and digital signals, and deals with storing, filtering and other operations on signals. These signals include transmission signals, sound or voice signals, image signals, and other signals etc.

V. PERFORMANCE EVAULATION

Our experimental result show Best feature extraction and tissue classification using GLCM-SVM classifier, which does not consider interactions in the labels of adjacent data points. We are going to discuss the result of CNN classification techniques. The comparison of existing classifiers like proposed SVM classifier results are discussed briefly in this section. Comparison will be takes place with of accuracy, precision, F1 Score and Sensitivity with the GIST image pixels.
This section will review our GLCM-SVM, an extension of RBF that uses a Gastro intestinal tumor framework to model interactions in the labels of adjacent data points

$$p(y|x) = \frac{1}{Z} exp\{ \sum_{i\in S} log(O(y_i , \gamma_i(x))) + \sum_{i\in S}\sum_{j\in N_i} V(y_i , y_j , X \} \quad \text{--------- (5.1)}$$

Where $\gamma_i(x)$ computes features from the observations x for location i, $O(y_i, \gamma_i(x))$ s an SVM based Observation-Matching potential, and $V(y_i, y_j, X)$ is the Local- Consistency potential over a pair-wise neighborhood structure, where N_i are the 8 neighbors around location i.

5.1 OBSERVATION-MATCHING

The Observation-Matching function maps from the observations (features) to class labels. We would like to use SVMs for this potential. However, the decision function in SVMs produces a distance value, not a posterior probability suitable for the DRFs' framework. To convert the output of the decision function to a posterior probability. This efficient method minimizes the risk of over fitting and is formulated as follows:

$$O(y_i = 1 \ \gamma_i(X)) = \frac{1}{1+exp(A X f(\gamma_i(X)) + B} \quad \text{---------- (5.2)}$$

the parameters A and B are estimated from training data represented as pairs where $\langle f(\gamma_i(x)), t_i \rangle$ is the real-valued SVM response (here, distance to the separator), and t_i denotes a related probability that $y_i = 1$, represented as the relaxed probabilities: $t_i = \frac{N+ +1}{N+ +2}$ if $y_i = 1$ $y_i = -1$, where N+ and N− are the number of positive and negative class instances. Using these training instances, we can solve the following optimization problem to estimate parameters A and B:

$$min - \sum_{i=1}^{t}[t_i \ log O\left(t_i, \gamma_i(x)\right) + (1 - O(t_i, \gamma_i(x)))] \quad \text{---------- (5.3)}$$

Platt [15] used a Levenberg-Marquardt approach that tried to set B to guarantee that the Hessian approximation was invertible. However, dealing with the constant directly can cause problems, especially for unconstrained optimization problems [13]. Hence, we employed Newton's method with backtracking line search for simple and robust estimation. To avoid overflows and underflows of exp andlog, we reformulated (3) as

$$min \sum_{i=1}^{t}[t_i \left(A X f(\gamma_i(X)) + B \right) + log(1 + exp(-A X f(\gamma_i(x)) - B))] \quad \text{---------- (5.4)}$$

5.2 PARAMETER ESTIMATION

GLCM-SVMs use a sequential learning approach to parameter estimation. This involves first solving the SVM Quadratic Programming problem (3). The resulting decision function is then converted to a posterior probability using the training data and estimated relaxed probabilities. The Local-Consistency parameters are then estimated from the m training pixels from each of the K training images using pseudo likelihood [12]:

$$\hat{V} = arg\,max \prod_{k=1}^{k}\prod_{i=1}^{m} p\ (y_i^k \mid y^k N_t, X^k \mid, V) \quad \text{---------- (5.5)}$$

We ensure that the log-likelihood is convex by assuming a Gaussian prior over v that is, $p(v|T)$ is a Gaussain distribution with 0 means and $T^2 I$ variance (see [9]). Thus, the local-consistency parameters are estimated using its log likelihood:

$$\hat{v} = argmax_v \sum_{k=1}^{K}\sum_{i=1}^{m}\{ O_i^n + \sum_{j\in N_t} V(y_i^k, y_j^k, X^k) - log(z_i^k) \} - \frac{1}{2T} v^t v \quad \text{---------- (5.6)}$$

where z_i^k is a partition function for each site i in image k, and T is a regularizing constant that ensures the Hessian is not singular. Keeping the Observation- Matching ($O_i^k = Oy_t, \gamma_t(x)$) constant, the optimal Local-Consistency parameters can be found by gradient descent.

We close by noting that the M^3 N [10] framework resembles GLCM-SVMs, as it also incorporates label dependencies and uses a max-margin approach. However, the M3N approach uses a margin that magnifies the difference between the target labels and the best runner-up, while we use the 'traditional' 2-class SVM approach of maximizing the distance from the classes to a separating hyper plane. An efficient approach for training and inference in a special case of M3Ns was presented in [6]. However, the simultaneous learning and the inference strategy used still make computations with this model expensive compared to GLCM-SVMs.

VI. CONCLUSION

Our research brings together two recent trends in the Gastro intestinal tumor Classification literature: model-aware similarity and affinity calculations with GLCM-SVM models with CRF-based evidence terms. In doing so, we make three main contributions. We use super pixel-based appearance models to reduce computational cost, improve spatial smoothness, and solve the data sampling problem for training GLCM-SVM classifiers on Gastro intestinal tumor Classification.

Also, we develop an affinity model that penalizes spatial discontinuity based on model-level constraints learned from the training data. Finally, our structural denoising based on the symmetry axis and continuity characteristics is shown to remove the false positive regions effectively.

Our full system has been thoroughly evaluated on a challenging 20-case GBM and the Bra TS challenge data set and shown to systematically perform on par with the state of the art. The combination of the two tracts of ideas yields better performance, on average, than either alone. In the future, we plan to explore alternative feature and classifier methods, such as classification forests to improve overall performance. we proposed a novel structure that includes a relatively global feature selection with a radiomics model and local feature selection with deep convolution neural networks. This structure also integrates the radiomics and deep convolution features to classify the CT images of GISTs into benign and malignant. Treatment of a brain tumor depends on its size and location. Although benign tumors do not tend to spread, they can cause damage by pressing on areas of the brain if they are not treated early. To avoid manual errors, an automated intelligent classification technique is proposed which caters the need for classification of image.

VII. REFERENCES

1) B. P. Rubin, M. C. Heinrich, and C. L. Corless, "Gastrointestinal stromal tumour," Lancet, vol. 369, no. 9574, pp. 1731-1741, 2007.

2) T. Nishida, N. Kawai, S. Yamaguchi, and Y. Nishida, "Submucosal tumors: comprehensive guide for the diagnosis and therapy of gastrointestinal submucosal tumors," Dig Endosc, vol. 25, no. 5, pp. 479-489, 2013

3) Beham, A. W., Schaefer, I.-M., Schüler, P., Cameron, S., & Michael Ghadimi, B. (2011). Gastrointestinal stromal tumors. International Journal of Colorectal Disease, 27(6), 689–700. doi:10.1007/s00384-011-1353-y

4) Predicting Malignancy Risk in Gastrointestinal Subepithelial Tumors with Contrast-Enhanced Harmonic Endoscopic Ultrasonography Using Perfusion Analysis Software(2019) Hyun Seok Lee, Chang Min Cho, Yong Hwan Kwon, and Su Youn Nam

5) T. Nishida, J. Y. Blay, S. Hirota, Y. Kitagawa, and Y. K. Kang, "The standard diagnosis, treatment, and follow-up of gastrointestinal stromal tumors based on guidelines," Gastric Cancer, vol. 19, no. 1, pp. 3-14, 2016.

6) S. Liu, X. Pan, R. Liu, H. Zheng, L. Chen, W. Guan, H. Wang, Y. Sun, L. Tang, Y. Guan, Y. Ge, J. He, and Z. Zhou, "Texture analysis of CT images in predicting malignancy risk of gastrointestinal stromal tumours," Clin Radiol, vol. 9260, no. 17, pii. 30446-4, 2017

7) C. Zhou, X. Duan, X. Zhang, H. Hu, D. Wang, and J. Shen, "Predictive features of CT for risk stratifications in patients with primary gastrointestinal stromal tumour," Eur Radiol, vol. 26, no. 9, pp. 3086-3093, 2016

8) P. Lambin, R. T. H. Leijenaar, T. M. Deist, J. Peerlings, E. E. C. de Jong, J. van Timmeren, S. Sanduleanu, R. T. H. M. Larue, A. J. G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T. Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F. M. Mottaghy, J. E. Wildberger, and S. Walsh, "Radiomics: the bridge between medical imaging and personalized medicine," Nat Rev Clin Oncol, vol. 14, no. 12, pp. 749-762, 2017

9) M. Vallières, C. R. Freeman, S. R. Skamene, and I. E. Naqa, "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities," Phys Med Biol, vol. 60, no. 14, pp. 5471-5496, 2015

10) M. Zhou, A. Leung, S. Echegaray, A. Gentles, J. B. Shrager, K. C. Jensen, G. J. Berry, S. K. Plevritis, D. L. Rubin, S. Napel, and O. Gevaert, "Non-small cell lung cancer radiogenomics map identifies relationships between molecular and imaging phenotypes with prognostic implications," Radiology, vol. 286, no. 1, pp. 307-315, 2017.