**How to Cite**:
Panwar, S. A., Munot, M. V., Mane, P. B., Deshpande, P. S., & Kanwade, A. B. (2022). Novel hybrid approach based on combination of textural features and clinical parameters for reliable prediction of thyroid malignancy. *International Journal of Health Sciences*, *6*(S7), 4011–4032.
https://doi.org/10.53730/ijhs.v6nS7.12707

# Novel hybrid approach based on combination of textural features and clinical parameters for reliable prediction of thyroid malignancy

**Dr. Sarika A. Panwar**
Department of Electronics and Telecommunication Engineering, AISSMS's Institute of Information Technology, Pune, India, PIN: 411001
Email: vrushalimendre@gmail.com

**Dr. Mousami V. Munot**
Department of Electronics and Telecommunication Engineering, SCTR's Pune Institute of Computer Technology, Pune, India, PIN: 411043

**Dr. Pradeep B. Mane**
Department of Electronics and Telecommunication Engineering, AISSMS's Institute of Information Technology, Pune, India, PIN: 411001

**Dr. Pallavi S. Deshpande**
Department of Electronics and Telecommunication Engineering, Bharati Vidyapeeth Deemed to be University College of Engineering, Pune, India, PIN: 411043

**Dr. Archana B. Kanwade**
Department of Electronics and Telecommunication Engineering, Marathawada Mitramandal College of Engineering, Pune, India

***Abstract***—The dramatic increase in thyroid cancer, particularly among the younger population demands development of an automated decision support system for timely and reliable prognosis of the disease so as to facilitate improved chances of recovery in the subjects. While numerous methods are already reported by the researchers for the detection of Thyroid malignancy, the most crucial parameter of Thyroid Malignancy Index (TMI) has received very less attention. TMI is of paramount importance in diagnosis and treatment of the patients having malignant thyroid and its consideration is therefore inevitable. This research aims to develop an automated and a reliable decision support system for detection of thyroid malignancy. The proposed hybrid approach incorporates a novel combination of texture features and clinically observable parameters to initially identify a malignant thyroid tumor using support vector machines and


*Manuscript submitted: 9 May 2022, Manuscript revised: 18 July 2022, Accepted for publication: 27 August 2022*

further predicts its TMI value, thus exhibiting a performance like a trained radiologist. Publically available database comprising of 99 cases and 134 ultrasound images are used to validate the superiority of the proposed approach. Apt consideration and reliable prediction of the TMI in this research makes the designed approach novel and marks a mega leap towards its practical deployment in the clinical environment.

***Keywords*---**Thyroid Malignancy Index (TMI), Decision Support System (DSS)

**Introduction**

The occurrence of thyroid cancer is escalating globally. Thyroid cancer is the seventh most frequent cancer in women. In total it is probably 2,200 deaths (1,050 men and 1,150 women) pertaining to this disease might cause this year. From the year 2009 to 2020 it is observed that the annual death rate has increased by a half percent. The number of thyroid cancer patients is rising across India, predominantly in the younger generation (specific age group < 45) [1]. Early Prognosis significantly raises the probability of recovery in the case of thyroid cancer. This research area, therefore, expects some significant contribution in identifying the biomarkers for modelling them using machine learning algorithms.

Thyroid diseases are caused due to malfunctioning of the thyroid gland and its associated hormones. The problem of Thyroid dysfunction can be categorized into two groups: 1st group reports dysfunction of the thyroid called thyroiditis and 2nd engross tumors of the thyroid. In general, these disorders are very common [2]. If doctors suspect thyroid disease, the blood test is suggested that can provide important information about several hormones such as TSH (Thyroid Stimulating Hormone), T3 (Triiodothyronine), T4 (thyroxin), TSI (Thyroid Stimulating Immunoglobulin) related to thyroid function.

Thyroid tumors or nodules refer to the anomalous growth in thyroid cells and further creates a lump in the thyroid gland. A patient with thyroid nodules has growths in the thyroid, which may be solid or fluid-filled. Fluid-filled nodules are cystic in nature. The majority of thyroid nodules are benign (non-cancerous). Thyroid nodule treatment is given on the basis of the decision of being it to be benign, or malignant. A thyroid nodule is found malignant (cancerous) infraction of cases. Symptoms are not seen in most of the thyroid nodules. While patients are carrying out routine check-ups in that process quite a lot of times accidentally thyroid nodules get revealed. It is also observed that many patients themselves locate thyroid nodules by observing a lump in the neck. Although thyroid function tests are a significant reason for detecting thyroid nodules. The presence of Thyroid nodules is responsible for the secretion of the excess quantity of thyroid hormone leads to further hyperthyroidism in the patient. However, thyroid nodules, even including cancerous ones, are in fact non-functioning, resulting in diagnostic tests like TSH showing normal results. Complaints like pain in the neck, jaw, or ear are even not often reported by patients with thyroid nodules. When a nodule grows big enough it starts compressing the windpipe or esophagus, and it starts giving difficulty in inhalation, swallowing, or might give a "scratch feeling in the throat". Very few times hoarseness is observed in the patients which happens when the nodule attack the nerve which is responsible for the control of the vocal cords, however, this is typically linked to thyroid cancer [2].

With the help of the Thyroid imaging technique, the images of the thyroid gland for clinical intentions are created. This ability to achieve information about the thyroid gland has many useful clinical applications. Thyroid imaging includes Magnetic Resonance Imaging (MRI), Ultrasound Imaging, and Computed Tomography (CT) [3]. Ultrasound imaging happens to be a non-invasive and low-priced imaging practice that offers the intra-nodular construction of the thyroid gland. It did not cause radioactive hazards to patients and was also done in a small acquisition time. Furthermore, this is a low-cost procedure as compared to peer imaging procedures like CT scan and MRIs. Advanced technology has made it possible to have high-frequency probes. these probes, provide excellent resolution in USG image acquisition. Reliability in clinical diagnosis always relies on basically the quality of USG images and of course the skill and expertise possessed by the radiologist who understands and interprets the images. USG images are simply affected by echo perturbations and speckle noise which interfere with the correct diagnosis [3].

Thyroid disease diagnosis using a decision support system through a blood test and ultrasonography image is a long-standing problem. In the last 25 years, substantial research has been carried out to develop a Computer-Aided Diagnosis (CAD) system for thyroid disease diagnosis. Developing an assistive automated system built with human expert knowledge is still a challenging issue. This section presents a comprehensive survey of the developments and current trends in the field of thyroid disease diagnosis using DSS. The survey details the overall advancements in CAD systems for thyroid disease diagnosis and also briefs about the previously reported techniques and algorithms to enhance the system.

## I. LITERATURE SURVEY

As thyroid disease diagnosis is a very important diagnosis problem, many researchers started working to build up a computer-aided diagnosis system to assist clinicians in diagnosis. A.S.Varde et.al. [4]in 1991 developed a clinical laboratory proficient system for the diagnosis of thyroid dysfunction. It was a rule-based expert system that considered the clinical findings and the results of applicable laboratory tests along with the patient's medical history. Okidaet. al. [5] in the year 2000 developed the expert system to support the diagnosis of thyroid nodules. It was also a rule-based system that is integrated to a database, that uses production rules, certainty factors qualitative reasoning for the computational representation of knowledge. When machine learning techniques emerged, machine learning algorithms were designed and used to analyze medical datasets. During the last 30 years, several machine learning algorithms have been designed and developed for the diagnosis of thyroid diseases.

Radiologists in Ultrasound society set up a team of specialised persons combining from a wide range of medical subject field to gather together to a consensus on the administration of thyroid nodules recognized with thyroid ultrasonography, with a typical spotlight on deciding which nodules should be subjected to US-guided fine-needle aspiration and which of the thyroid nodules are not needed for subjecting to fine-needle aspiration. The panel meeting was held in Washington, DC, October 26–27, 2004, and formed a consensus report [6]. According to the consensus statement given by Frates et al. [6], FNA is strongly recommended in a solid solitary nodule (having a maximum diameter greater than or equal to 1cm) with microcalcifications or coarse calcifications. The consensus statement further clarifies that FNA can be considered for mixed solid and cystic or almost entirely cystic nodules with the solid mural components. If substantial growth is observed since prior ultrasound examination, ultrasound-guided FNA needs to be considered. FNA is assumed to be superfluous in diffusely large glands with numerous nodules of analogous US appearance without intervening parenchyma. The existence of any calcification inside the nodule rise the probability of malignancy. In peculiar, microcalcifications in a preponderantly solid nodules are related with close to threefold addition in cancer risk, and coarse calcifications are related with a twofold rise, when compared with preponderantly solid nodules without calcifications

The study done by Iannuccilli et al. [7] points that intrinsic microcalcifications is simply statistically reliable criterion on which increased suspicion for malignancy in thyroid nodules. The result obtained in the study points the need for biopsy in deciding further workup. All nodules that show the presence of intrinsic microcalcifications should undergo biopsy, principally in case of calcifications that bear a resemblance to a form of a snowstorm in sonography.

Hoang et al. in 2007 [8] also reports the valuable characteristics for identifying malignant thyroid nodules. The features like microcalcifications, local invasion, lymph node metastases, a nodule that is taller one than the wide one, and specifically decreased echogenicity. Some of the remaining features like the absence of a halo, poorly defined irregular margins, solid composition, and vascularity, are not very significant but perhaps useful subsidiary signs.

Literature reports the use of sono-graphic features for detecting thyroid malignancy. The lacuna of this type of analysis exhibits inter-and intra-observer inconsistency within the inferences drawn for the results, this happens very commonly even by a veteran medical professional. Heterogeneity nature is seen in most of the thyroid nodules with a variety of inner components which ultimately creates confusion in the radiologists and physicians. The situation leads or motivates the researchers to think and devise the automation of the system. With the automation, we will be able to get superior quality in images and also better support by the algorithmic approach in the interpretation of the same.

Table 1 presents the methodology adapted by researchers for the development of CAD system for thyroid disease diagnosis.

Despite the scrupulous efforts to develop DSS for automated diagnosis of thyroid disorders, it has limited success when compared to trained clinicians in clinical service demand. Some of the reasons for poor performance are lack of utilization of expertise and experience and lack of ability to understand the clinical parameters of thyroid malignancy. Great efforts to develop the techniques in line with automatic pattern recognition which will aid the diagnosis of thyroid gland-related diseases have been made so far. DSS for automated diagnosis of thyroid disorders is an attempt to minimize human efforts in the process of diagnosis. Hence the system must have both, expert cognition and experience of the trained clinician and/or radiologists with the clinical acceptability perspective. This makes the development of the DSS for thyroid disorders hard and thought-provoking and hence needs additive research. Few significant findings, demands or challenges, and issues in the processes of development about DSS for thyroid diseases are discussed below: The features which describe the texture of thyroid nodule image can be extracted. However, there can be other information available in the USG image of nodule which gives an idea about the malignancy of the nodule. Various approaches are reported in literature for feature extraction. Most of the researcher used GLCM based feature extraction method [9,20,23,24,30]. Keramidas et al. [25] used Fuzzy Local Binary Pattern [FLBP]. Acharya et al. [23] used DWT and GLCM features, but these features don't consider the very important clinical parameter (Only observable parameters) that is nodule's solid or cystic nature which decides the malignancy of thyroid nodule

1. The algorithms fail in measuring the solidity percentage which increases the likelihood of malignancy. It is necessary to evaluate and confirm the thyroid malignancy based on solidity and calcifications (microcalcifications and coarse calcification).

2. According to the consensus statement given by Frates et al [6] the likelihood of malignancy is more with the existence of calcification in the nodule. Particularly, microcalcifications in a solid nodule are related to a threefold rise in cancer threat whereas a twofold increase is related to coarse calcifications, as compared with mainly solid nodules lacking in calcification. Solid nodules are also susceptible to malignancy. Statistically speaking, the bigger criteria for taking a doubt in malignancy within thyroid nodules is happened to be a solid nodule (predominantly or purely solid nodule) having some calcifications

3. Acharya et al. [29] formulated the TMI based on texture and wavelet features. Still, it doesn't consider the clinically observable feature to formulate TMI.

While numerous methods are already reported by the researchers for the detection of Thyroid malignancy, the most important aspect of TMI prediction has received less attention by the researchers. This index is of paramount importance in planning the line of treatment for patients having malignant thyroid.

This research aims to design and develop an automated system for the early and reliable detection of thyroid cancer. The proposed automated DSS uses a novel combination of texture features and clinically observable parameters to predict the TMI thus exhibiting a performance like a trained radiologist.

Comprehensive research is carried out aiming at the generation of several textural, wavelet as well as clinical features with the implementation of a powerful SVM classifier and fuzzy system for the prediction of the TMI factor in a thyroid nodule. The extracted features are distinct and can be incorporated in the overall diagnostic procedure as explained in the next section.

## II. METHODOLOGY

A new approach has been developed towards thyroid nodule malignancy detection. The proposed approach is broadly divided in two parts viz. part 1 and part 2.

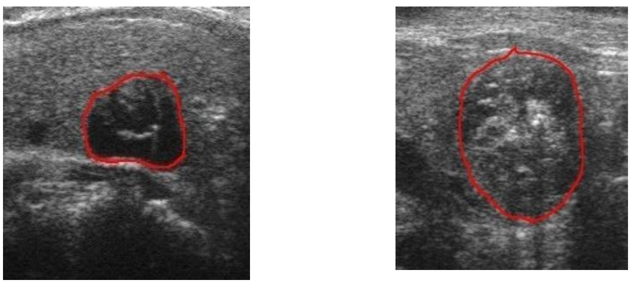
Part I: Thyroid Malignancy Detection

Part II: Thyroid Malignancy Index (TMI) value prediction.

Figure 1 and 8 shows the flow diagram for Part-I and Part-II respectively.

### *A. Part I: Thyroid Malignancy Detection*

The flow diagram of part I is presented in Figure 1. The thyroid malignancy detection scheme comprises at first the ROI separation then feature extraction and finally malignancy detection. 99 cases of thyroid nodule are referred from the publically accessible Computer Imaging and Medical Applications (CIM) Laboratory database [36]. The said thyroid ultrasound images digital database is an open-access resource for the research community. This project was supported by Universidad National de Colombia, CIM@LAB and IDIME ( Instituto de Diagnostico Medico). The most important intention of the Dataset is to promote the activity of developing algorithms for CAD systems. Further, these algorithms will facilitate in-depth analysis of thyroid nodules. Contrarily the database is used as a practicing means for radiologists. The database includes 99 cases and 136 images. Figure 2 depicts some of the images from the database, used for this research.



(a) (b)

Figure 2: Samples from the CIM lab database used (a) Benign nodule (b) Malignant Nodule

1) *ROI Separation*: The first component of computer aided detection is Region of Interest (ROI) separation. Segmentation of thyroid nodule region is important for extracting the features of thyroid nodule part.

Echogenicity is very important for the segmentation of thyroid nodule. Echogenicity can be assessed with respect to the normal thyroid parenchyma and can be classified as hypoechoic, hyperechoic and isoechoic nodule.

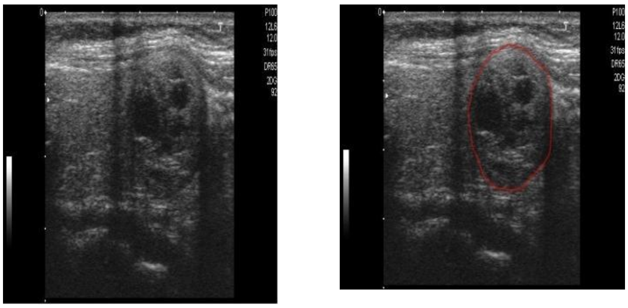
- Hypoechoic nodule shows a relatively hypoechoic pattern with respect to the regular thyroid parenchyma.

- Isoechoic nodule shows an isoechoic pattern with respect to the normal thyroid parenchyma.

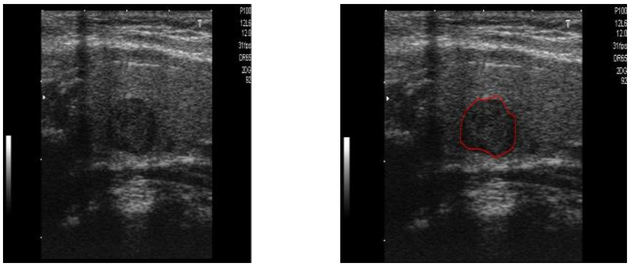
- Hyperechoic nodule shows a relatively higher echogenicity in regard to the normal thyroid parenchyma.

Figure 3 depict the examples of hypoechoic, hyperechoic and isoechoic nodules. Figure 3 (a) and (b) depicts the isoechoic nodule with boundary highlighted by expert. Figure 3 (c) and (d) shows the hypoechoic nodule with nodule boundary marked in (d). Figure 3 (e) and (f) shows the hyperechoic nodule. It is observed from Figure 3, that if the nodule is isoechoic or hyperechoic, then edges or boundaries are not well defined. It is very difficult to segment or separate the nodule in case of isoechoic or hyperechoic nodules. Most of the thyroid nodule boundary detection algorithms proposed in past [21, 37, 38, 39] work well on hypoechoic nodules.

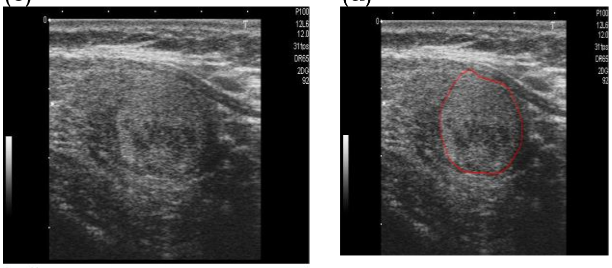
In the database of CIM laboratory [36], 136 images of 99 cases are provided with the boundary highlighted by expert radiologists. Considering this as benefit for segmentation, nodule highlighted images in which thyroid boundary is highlighted with red line (shown in Figure 3 (b), (d) and (f)), can be used for ROI separation. The ROI is separated by detecting the red boundary surrounding to thyroid nodule. Figure 4 depicts the separated nodule from background thyroid parenchyma.



(a) (b)

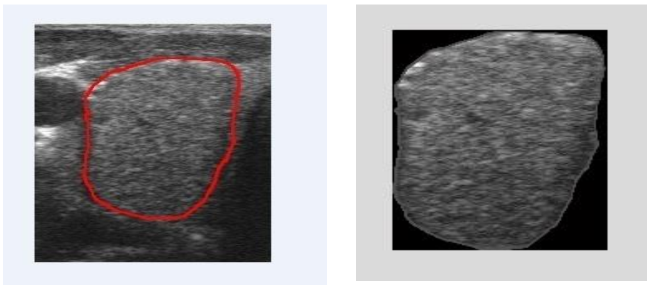


(c) (d)



(e) (f)

Figure 3: Examples of thyroid nodule echogenicity



(a) (b)

Figure 4 : Output of ROI separation algorithm (a) Original Image (b) ROI separated image

2) *Feature Extraction and Classification:* Feature extraction is divided into two modules as follows:

Discrete Wavelet Transform: Figure 5 shows the feature extraction using DWT. Thyroid nodule USG image f(i, j) is given as an input to Low Pass Filter (LPF) and High Pass Filter (HPF). The output of LPF and HPF are down sampled by 2. The down arrow in the rectangle shows the down sampling operation. The output of down sampling is further given to LPF and HPF to get LL, HL, LH, and HH coefficients.

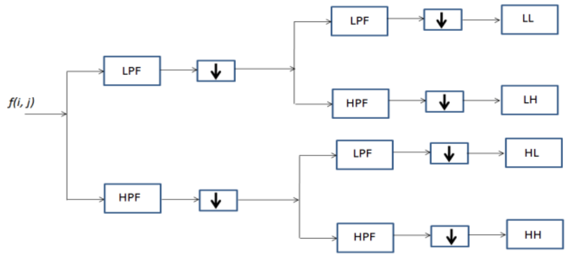
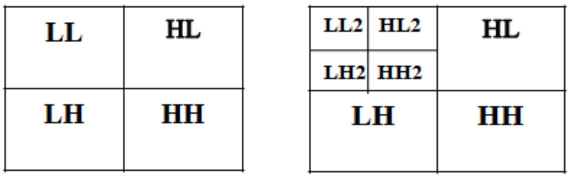



Figure 5: Feature extraction using DWT

Figure 6 (a) shows the first level of DWT decomposition. Decomposition is further performed on the LL coefficient to get coarser scale of wavelet coefficients. Decomposition is stopped at second level because further decomposition didn't give significant results. Figure 6(b) shows second level of DWT decomposition. This DWT feature extraction method uses Daubechies (Db) 8 as the mother wavelet.




(a) (b)

Figure 6: 2D sub-band DWT (a) First Level of Decomposition (b) Second Level Decomposition

The Normalized energies A2, V2, H2, D2, H1, V1, and D1 are obtained from matrices LL2, HL2, LH2, HH2, LH, HL and HH respectively.

Equations 1 to 7 represent the mathematical model of the DWT based features which play a significant role in the designed system. These equations detail the formation of normalized energies from individual sub-band matrices, which formulate a feature vector F (F1 - F7).

| |
|---|
| $F1 = LL2 = A_2 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} LL2_{ij}\right]^2$ (1) |
| $F2 = HL2 = V_2 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} HL2_{ij}\right]^2$ (2) |
| $F3 = LH2 = H_2 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} LH2_{ij}\right]^2$ (3) |
| $F4 = HH2 = D_2 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} HH2_{ij}\right]^2$ (4) |
| $F5 = LH = H_1 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} LH_{ij}\right]^2$ (5) |
| $F6 = HL = V_1 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} HL_{ij}\right]^2$ (6) |
| $F7 = HH = D_1 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} HH_{ij}\right]^2$ (7) |

Gray-Level Co-occurrence Matrix (GLCM) Features: GLCM features are used to obtain the texture information of nodule GLCM feature vector is explained further.

Contrast gives intensity contrast between a pixel and its neighbor over the whole image.

$F8 = contrast = \sum_{i,j}[i-j]\,2 * p(i,j).$

Correlation correlates a pixel to its neighbor over whole image.

$F9 = correlation = \sum_{i,j}\frac{(i-\mu_i)(j-\mu_j)(p(i,j))}{\sigma_i\sigma_j}$, where μ= mean {p (i,j)}

Energy is the sum of squared elements in the GLCM.

$$F10 = energy = \sum_{i,j} p(i,j)^2$$

Homogeneity is the value that measures closeness of elements in the GLCM matrix.

$$F11 = homogeneity = \sum_{i,j}\frac{p(i,j)}{1+|i-j|}$$

Inverse Difference Moment [IDM] is calculated as follow:

$$F12 = IDM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{1}{1+(i-j)2}p(i,j)$$

F13=Variance= {p( i, j)}

F14 $= sumaverage = \sum_{i=0}^{2G^2} ip_{x+y}(i).$

F15=Entropy= entropy {P}

Where $P = P_x + P_y$

$F16 = ClusterShade = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i+j-\mu_x-\mu_y)^3 p(i,j).$

Where $\mu_x = mean\{p_x\}$

$$\mu_y = mean\{p_y\}$$

$\mu=\frac{1}{2}\sum_{i=0}^{2G2} iH_s(x|\Delta x,\Delta y)$

cluster prominence is defined as:

$F17 = C.Prom. = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i + j - \mu x - \mu y\}^4 p(i,j).$

$F18 = inertia = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - j)^2 p(i,j).$

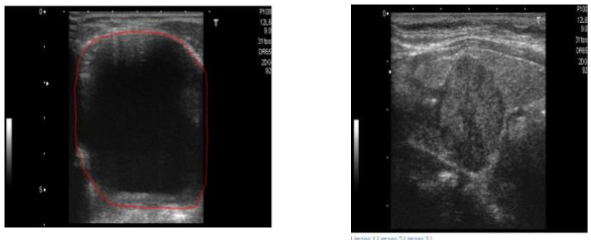
GLCM is popularly used in the literature for efficient statistical feature extraction.

GLCM based algorithm extracts total 11 features (F8- F18).

Clinical features (percentage cystic area and presence of calcifications): ROI separated USG image is considered as an input for extracting two clinical features:

F19=Percentage cystic area.

F20=Presence of calcifications.



(a) (b)

Figure 7: Example of internal content of thyroid nodule (a) purely cystic (b) purely solid

The proposed algorithm for extracting Clinical features is based on a consensus statement [6]. As per the consensus statement, the presence of any calcification (coarse and micro-calcifications) in predominantly solid nodule is related to an around threefold growth in cancer threat as compared with predominantly solid nodules without calcifications. Considering this consensus statement, two very important clinical features are extracted from the thyroid nodule USG image. These two features are used to predict Thyroid Malignancy Index (TMI). The purely solid and purely cystic nodules as seen in the USG image are shown in Figure 7 (a) and (b). The cystic area and presence of calcifications are calculated as follows:

ALGORITHM

1. Calculate the size of nodule (segmented thyroid nodule variable, generated after ROI separation).

2. Remove all small objects from image by creating morphological structuring element.

3. Calculate total nodule area (Nodule A).

4. Calculate cystic area in a nodule (Cystic A).

5. %Cystic area = (cystic A/nodule A) *100.

6. Let calc is the variable containing the average pixel value of nodule.

7. Compare calc with T (where T is the predefined threshold).

8. Define “Presence_of_Calcification” Variable.

9. Presence_of_Calcification= 0 (if calc< T)

   = 1 (if calc>T)

The cystic area (cystic %) in a nodule varies regardless of the size of the nodule. The cystic area is calculated in percentage. The presence of the calcification feature takes only two values: 1 or 0. If the calcification is present in nodule, it will take value 1, and if the calcification is absent it will take 0 value.

As per the consensus statement mentioned in [6], mainly a solid nodule which is probably <25% cystic having microcalcifications shows 31.6% probability of detecting as cancer, when its compared with a principally cystic nodule having probability >75% to be cystic and that to not having calcification, this depicts 1.0% chance of being detected as cancer. Based on this statement, Thyroid Malignancy Index [TMI], can be calculated using the above algorithm.

If the cystic area (%) > 75%, then the nodule is predominantly cystic nodule.

If the cystic area (%) < 25%, then the nodule is a predominantly solid nodule.

If the nodule is purely solid (with a cystic percentage between 0 to 25%), or predominantly solid with calcifications (=1) present in the nodule will be highly susceptible to malignancy. Using this criterion TMI is predicted.

% Cystic Area and presence of calcifications in nodule forms feature vectors F19 and F20respectively.

The feature vector of 20 features is presented in table 2.

TABLE 2.
FEATURE VECTOR TABLE

| **Feature number** | **Feature Names** | **Feature number** | **Feature Names** |
|---|---|---|---|
| F1 | A2 | F11 | Homogeneity |
| F2 | V2 | F12 | IDM |
| F3 | H2 | F13 | Variance |

| | | | |
|---|---|---|---|
| F4 | D2 | F14 | Sum Average |
| F5 | H1 | F15 | Sum Entropy |
| F6 | V1 | F16 | Cluster Prominence |
| F7 | D1 | F17 | Cluster Shade |
| F8 | Contrast | F18 | Inertia |
| F9 | Correlation | F19 | Percentage cystic area |
| F10 | Energy | F20 | Presence of calcifications |

The features mentioned in Table 2 are utilized for malignancy detection (Part I) and TMI prediction (Part II).

Figure 1 presents a flow diagram for part I. A feature vector of 20 features is used to classify the thyroid nodule into two classes namely: malignant and benign. as this is a two-class problem, SVM is used for classification. SVM is proved to be a powerful classifier for two-class problems.

a non-linear decision hyperplane is created by SVM classifier. The non-linearly separable data can be properly classified using this hyperplane. The kernel function is applied on all data samples to map the fresh non-linear samples into a higher-dimensional space where they can be easily separable separable. Therefore, the SVM performance is analysed for various kernel functions namely: Linear, RBF, MLP, and Polynomial kernel with degrees 1, 2, 3.

SVM model is designed using various kernel functions as the features are linearly non-separable. The performance analysis of the SVM classifier is done on the basis of ROC analysis. Area Under ROC Curve [AUC] can be a more significant performance measure and also it represents the likelihood that arbitrary pairs of high risk, as well as low-risk thyroid nodules, are acceptably classified. Sensitivity [SN] depends only on measurement of malignant nodule and specificity [SP] only on measurement of the benign nodule.

The performance of SVM for various kernel functions is depicted in Table 3. It is observed from Table 3, that the highest classification accuracy is obtained if a polynomial kernel with 3rd degree is used. AUC is 0.98 with a specificity of 0.97 and a sensitivity of 0.88. The likelihood ratio estimates the capability or power of every classifier for providing higher degree of confidence in assuring about a positive diagnosis. The likelihood ratio is also highest for SVM with a polynomial of 3rd degree. The total count of support vectors for the SVM algorithm alongwith a polynomial of 3rd degree is 55 which coincides with the maximum classification accuracy. The total number of support vectors ranges from 2.7% to 5% of the total number of data points. The decision hyperplane of SVM classifier creates decision hyperplane which is determined solitarily by count of number of support vectors. Having a low number of support vectors, is a denotation of SVM low complexity and its differentiation capacity. Figure 7 depicts the ROC curve for SVM classifier with polynomial of 3rd degree.

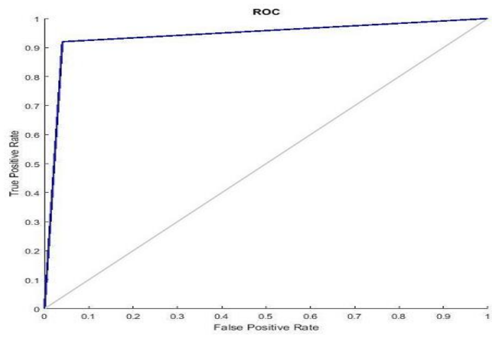



Figure 7: ROC curve for SVM classifier with polynomial kernel of 3rd degree

*B. Part II: TMI Prediction*

Thyroid Malignancy Index [TMI] is predicted based on two clinical parameters namely: cystic % and presence of calcifications. For each nodule cystic percentage and presence of calcifications is calculated. These two features are given as an input to a fuzzy-based classifier. The details of fuzzy-based classifier are as follows:

Fuzzy Classifier for TMI:

The block diagram to find TMI is shown in Figure 9. Cystic percentage and presence of calcifications are the two inputs given to the fuzzy classifier. Membership functions are designed with the help of Mamdani model.

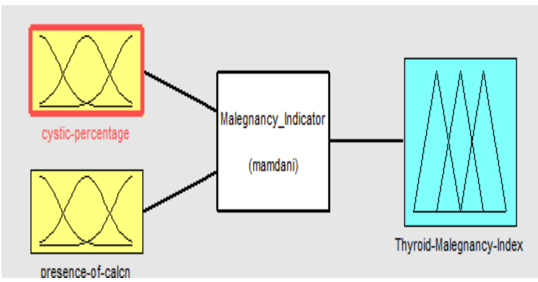



Figure 9: Fuzzy classifier for TMI

Figure 10 depicts superlative performance of the designed fuzzy based DSS in terms of predicting TMI using cystic percentage and presence of calcification. The yellow portion showcases highest TMI when cystic percentage is least and dense calcification is present. Whereas lowest TMI index is obvious when cystic percentage is high and calcification is hardly present (region in blue color). Figure 10 depicts the 3D view of output for all ranges of inputs.

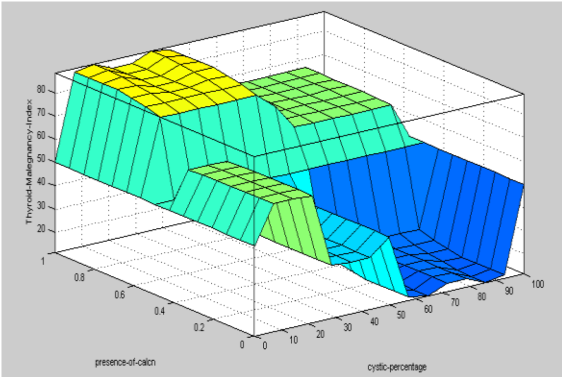



Figure 10: TMI Prediction

## III. DISCUSSION AND CONCLUSIONS

Automated detection of thyroid malignancy is a crucial problem attracting attention of many researchers. Most of the researchers have focused on the use of texture feature extraction for malignancy detection. Although high accuracy is obtained using these features in the reported literature, they do not necessarily support the prediction of TMI which is the most important parameter in thyroid malignancy detection. This research presents a decision support system which serves as a facilitating tool for the clinicians/ radiologist in reliable detection of thyroid malignancy. Both, manual approach and human expertise are imbibed in the designed automated system to ensure accuracy and reliability of diagnosis. In this research, the quantification of texture characteristics, wavelet features, and only observable features (cystic percentage and presence of calcifications) has directed towards extensive feature generation that facilitates in encoding the available information of every thyroid nodule. Initial phase of the designed system (part I) detects malignancy of tumor where, SVM classifier has exploited the relation between clinical feature, texture features, and wavelet features towards an accurate differentiation between benign and malignant thyroid nodule. Later, in part 2, the fuzzy-based system uses clinically observable parameters to predict TMI. The presence of calcifications and cystic percentage in the nodule has been proved as a significant characteristic that predicts thyroid malignancy index. The fuzzy system proved to be the best method for the prediction of TMI and the overall performance of the designed system matches that of a human expert (radiologist). Apt consideration and reliable prediction of the TMI in this research makes this study novel and marks a mega leap towards practical deployment of the proposed DSS.

## References

[1] https://www.cancer.org/cancer/thyroid-cancer/about/key-statistics.html

[2] Salvatore D, Davies TF, Schlumberger MJ, et al. Thyroid physiology and diagnostic evaluation of patients with thyroid disorders. In: Melmed S, Polonsky KS, Larsen PR, et al., eds. *Williams Textbook of Endocrinology*. 12th ed. Philadelphia, PA: Elsevier Saunders; 2011:chap 11

[3] Chaudhary V, Bano S. Imaging of the thyroid: Recent advances. Indian J Endocrinol Metab.2012; 16:371–6.

[4] A. S. Varde, K.L.Massey, H.C.Wood, “ A Clinical Laboratory Expert System for the Diagnosis of Thyroid Dysfunction”, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol.13, No.3, 1991

[5] Okida, Sergio, F. M. de Azevedo, and J. L. B. Marques. "Expert system to support the diagnosis of thyroid nodules." Engineering in Medicine and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE. Vol. 3. IEEE, 2000.

[6] Frates, Mary C., et al. "Management of Thyroid Nodules Detected at US: Society of Radiologists in Ultrasound Consensus Conference Statement 1."Radiology 237.3 (2005): 794-800.

[7] Iannuccilli, Jason D., John J. Cronan, and Jack M. Monchik. "Risk for malignancy of thyroid nodules as assessed by sonographic criteria the need for biopsy." Journal of Ultrasound in Medicine 23.11 (2004): 1455-1464.

[8] Hoang, Jenny K., et al. "US Features of Thyroid Malignancy: Pearls and Pitfalls 1." Radiographics 27.3 (2007): 847-860.

[9] Smutek, Daniel, et al. "Image texture analysis of sonograms in chronic inflammations of thyroid gland." Ultrasound in medicine & biology 29.11 (2003): 1531-1543.

[10] Tsantis, Stavros, et al. "Development of a support vector machine-based image analysis system for assessing the thyroid nodule malignancy risk on ultrasound." Ultrasound in medicine & biology 31.11 (2005): 1451-1459.

[11] Savelonas, Michalis A., et al. "Computational characterization of thyroid tissue in the radon domain." Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07). IEEE, 2007.

[12] Keramidas, Eystratios G., et al. "Thyroid texture representation via noise resistant image features." Computer-Based Medical Systems, 2008. CBMS'08. 21st IEEE International Symposium on. IEEE, 2008.

[13] Chang, Chuan-Yu, Ming-Feng Tsai, and Shao-Jer Chen. "Classification of the thyroid nodules using support vector machines." Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on. IEEE, 2008.

[14] Tsantis, Stavros, et al. "Morphological and wavelet features towards sonographic thyroid nodules evaluation." Computerized Medical Imaging and Graphics 33.2 (2009): 91-99.

[15] Kodaz, Halife, et al. "Medical application of information gain based artificial immune recognition system (AIRS): Diagnosis of thyroid disease." Expert Systems with Applications 36.2 (2009): 3086-3092

[16] Chang, Chuan-Yu, Shao-Jer Chen, and Ming-Fong Tsai. "Application of support-vector-machine-based method for feature selection and classification of thyroid nodules in ultrasound images." Pattern recognition 43.10 (2010): 3494-3506.

[17] Ma, Jieming, et al. "Differential diagnosis of thyroid nodules with ultrasound elastography based on support vector machines." 2010 IEEE International Ultrasonics Symposium. IEEE, 2010

[18] Iakovidis, Dimitris K., Eystratios G. Keramidas, and Dimitris Maroulis. "Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns." Artificial Intelligence in Medicine 50.1 (2010): 33-41.

[19] Dogantekin, Esin, AkifDogantekin, and DeryaAvci. "An automatic diagnosis system based on thyroid gland: ADSTG." Expert Systems with Applications 37.9 (2010): 6368-6372.

[20] Dogantekin, Esin, AkifDogantekin, and DeryaAvci. "An expert system based on Generalized Discriminant Analysis and Wavelet Support Vector Machine for diagnosis of thyroid diseases." Expert Systems with Applications 38.1 (2011): 146-150.

[21] Ding, Jianrui, et al. "Quantitative measurement for thyroid cancer characterization based on elastography." Journal of Ultrasound in Medicine 30.9 (2011): 1259-1266. A Thyroid Nodule Detection System for Analysis of Ultrasound Images and Videos.

[22] Legakis, Ioannis, et al. "Computer-based nodule malignancy risk assessment in thyroid ultrasound images." International Journal of Computers and Applications 33.1 (2011): 29-35.

[23] Acharya, U. R., Faust, O., Sree, S. V., Molinari, F., Garberoglio, R., Suri, J. S. Cost-effective and non-invasive automated benign and malignant thyroid lesion classification in 3D contrast-enhanced ultrasound using combination of wavelets and textures: A class of thyroscan algorithms. Technol Cancer Res Treat 10, 371-380 (2011).

[24] Nikita Singh and Alka Jindal "A Segmentation and Comparison of Classification Method for Thyroid Ultrasound Images." International Journal of Computer Applications (0975-8887) Volume 50-No 11,.July 2011

[25] Keramidas, Eystratios G., Dimitris Maroulis, and Dimitris K. Iakovidis. "TND: a thyroid nodule detection system for analysis of ultrasound images and videos." Journal of medical systems 36.3 (2012): 1271-1281.

[26] Acharya, U. Rajendra, et al. "Non-invasive automated 3D thyroid lesion classification in ultrasound: a class of ThyroScan™ systems." Ultrasonics 52.4 (2012): 508-520.

[27] Chen, Hui-Ling, et al. "A three-stage expert system based on support vector machines for thyroid disease diagnosis." Journal of medical systems 36.3 (2012): 1953-1963.

[28] Liu, Da-You, et al. "Design of an enhanced fuzzy k-nearest neighbor classifier based computer aided diagnostic system for thyroid disease." Journal of medical systems 36.5 (2012): 3243-3254.

[29] Acharya, U. Rajendra, et al. "ThyroScreen system: high resolution ultrasound thyroid image characterization into benign and malignant classes using novel combination of texture and discrete wavelet transform." Computer methods and programs in biomedicine 107.2 (2012): 233-241.

[30] Acharya, U. Rajendra, et al. "Computer-Aided Diagnostic System for Detection of Hashimoto Thyroiditis on Ultrasound Images From a Polish Population." Journal of Ultrasound in medicine 33.2 (2014): 245-253.

[31] Acharya, U. R., Chowriappa, P., Fujita, H., Bhat, S., Dua, S., Koh, J. E., ... & Ng, K. H. (2016). Thyroid lesion classification in 242 patient population using Gabor transform features from high resolution ultrasound images. Knowledge-Based Systems, 107, 235-245.

[32] Choi, Young Jun, et al. "A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of thyroid nodules on ultrasound: initial clinical assessment." Thyroid 27.4 (2017): 546-552.

[33] Koundal, Deepika, Savita Gupta, and Sukhwinder Singh. "Computer aided thyroid nodule detection system using medical ultrasound images." Biomedical Signal Processing and Control 40 (2018): 117-130.
[34] Kumar, S. (2022). A quest for sustainium (sustainability Premium): review of sustainable bonds. Academy of Accounting and Financial Studies Journal, Vol. 26, no.2, pp. 1-18
[35] Nugroho, HanungAdi, et al. "Computer aided diagnosis for thyroid cancer system based on internal and external characteristics." Journal of King Saud University-Computer and Information Sciences (2019).
[36] Shankar, K., et al. "Optimal feature-based multi-kernel SVM approach for thyroid disease classification." The Journal of Supercomputing 76.2 (2020): 1128-1143.
[37] http://cimalab.intec.co/?lang=en
[38] Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2021). Health and treatment of diabetes mellitus. International Journal of Health Sciences, 5(1), i-v. https://doi.org/10.53730/ijhs.v5n1.2864
[39] M. Savelonas, D. Maroulis, D. Iakovidis and S. Karkanis, "Variable Background Active Contour Model for Automatic Detection of Thyroid Nodules in Ultrasound Images ", 2005.
[40] Chan, Tony F., and Luminita A. Vese. "Active contours without edges. "Image processing, IEEE transactions on 10.2 (2001): 266-277.
[41] Keramidas, Eystratios G., et al. "Efficient and effective ultrasound image analysis scheme for thyroid nodule detection." Image Analysis and Recognition. Springer Berlin Heidelberg, 2007. 1052-1060.
[42] Zhang, Xinyu, et al. "Deep convolutional neural networks in thyroid disease detection: A multi-classification comparison by ultrasonography and computed tomography." Computer Methods and Programs in Biomedicine 220 (2022): 106823.

TABLE 3
SVM CLASSIFIER RESULT FOR MALIGNANCY DETECTION

| **Sr. No.** | **Name of Kernel Function** | **Accuracy** | **AUC** | **Specificity** | **Sensitivity** | **Likelihood Ratio** | **Support Vectors** |
|---|---|---|---|---|---|---|---|
| 1 | Polynomial 1 | 71 | 0.95 | 0.46 | 0.91 | 1.685185185 | 74 |
| 2 | Polynomial 2 | 80 | 0.975 | 0.56 | 0.95 | 2.159090909 | 65 |
| 3 | Polynomial 3 | 94 | 0.98 | 0.97 | 0.88 | 29.33333333 | 55 |
| 4 | Linear | 73 | 0.95 | 0.48 | 0.91 | 1.75 | 77 |
| 5 | RBF | 92 | 0.98 | 0.99 | 0.99 | 99 | 99 |
| 6 | MLP | 44 | 0.88 | 0.77 | 0.26 | 1.130434783 | 73 |
| 7 | Quadratic | 78 | 0.98 | 0.54 | 0.96 | 2.086956522 | 61 |

TABLE I
LITERATURE SURVEY SUMMARY

| S.N | Authors | Year | Methodology | Limitations |
|---|---|---|---|---|
| 1 | Smutek et al. [9] | 2003 | Spatial and GLCM features are used for in chronic inflammation of thyroid gland. | Texture analysis to differentiate chronic inflamed thyroid tissue from healthy tissue. |
| 2 | Tsantis et al.[10] | 2005 | Gray tone histogram and GLCM features are fused together to extract the texture information | Clinical features which can imbibe the expert knowledge are not considered. |
| 3 | Savelonas, Michalis A., et al. [11] | 2007 | Radon Transform is used to differentiate directionality patterns in thyroid ultrasonography image. | The classification accuracy achieved is 89.4% with Radon Transform. |
| 4 | Keramidas et al. [12] | 2008 | Fuzzy and LBP approaches are combined. | Clinical features which can imbibe the expert knowledge are not considered. |
| 5 | Chang et al. [13] | 2008 | Specifically shows existence of six kinds of thyroid nodules with remarkably good performance. | Clinical features such as solid or cystic nature has not been considered. |
| 6 | Tsantis et al. [14] | 2009 | Morphological and wavelet features are used for classification. | Thyroid malignancy index is not calculated. |
| 7 | Kodaz, Halife, et al. [15] | 2009 | Information Gain based Artificial Immune Recognition System (AIRS) is proposed for classification of thyroid disease in three classes: Hyperthyroid, Hypothyroid and Normal Thyroid. | Achieved 95.90% accuracy of classification. (Ultrasound Images are not used only blood test reports are considered for classification) |
| 8 | Chang et al. [16] | 2010 | SVM is used for feature selection from 78 texture features extracted from nodular region. | Thyroid Malignancy Index is not calculated. Only malignancy is detected. |
| 9 | Ma, Jieming, et al. [17] | 2010 | Classification based on Ultrasound Elastography using SVM | Focus on classification |
| 10 | Iakovidis et al. [18] | 2010 | Fuzzy Local Binary Pattern and Fuzzy Gray Level Histogram are used to detect malignancy in Ultrasonography images. | Focus is only on classification. Thyroid Malignancy Index (TMI) is not calculated. |
| 11 | Dogantekin et al. [19] | 2010 | Least Square Support Vector Machine is used to build ADSTG (Automatic Diagnosis System based on Thyroid Gland) | The proposed ADSTG (Automatic Diagnosis System based on Thyroid Gland) lack in finding TMI. |
| 12 | Dogantekin et. Al. [20] | 2011 | Wavelet Support Vector Machine based system is proposed for malignancy detection | TMI is not calculated. |

| | | | | |
|---|---|---|---|---|
| 13 | Ding, Jianrui, et al. [21] | 2011 | Statistical and Textural features are used to evaluate thyroid color elastogram | Only solid nodules are considered. Cystic and Mixed solid nodules are skipped from analysis. |
| 14 | Legakis, Ioannis, et al. [22] | 2011 | Textural and shape features of Ultrasonography images are used to delineate thyroid nodule. | Segmentation algorithm is proposed. |
| 15 | Acharya et al. [23] | 2011 | Malignancy is detected using DWT and GLCM with 98% accuracy. | Expert knowledge is not imbibed using clinical malignancy characteristics |
| 16 | Singh et al. [24] | 2011 | GLCM features are used for texture characterization | Accuracy is very low |
| 17 | Keramidas et al. [25] | 2012 | Fuzzy and LBP approaches are combined for thyroid nodule detection. | Could only classify thyroid nodule from thyroid parenchyma |
| 18 | Acharya, U. Rajendra, et al. [26] | 2012 | Features: Fractal Dimension, LBP, Fourier Spectrum Descriptor. TMI calculated using FD, LBP. | Observable clinical parameters are not considered for TMI calculation. |
| 19 | Chen, Hui-Ling, et al. [27] | 2012 | Feature Selection and Classification using Particle Swam Optimization (PSO) and SVM. | Feature selection and parameter optimization is done. |
| 20 | Liu, Da-You, et al. [28] | 2012 | Design of an Fuzzy K-NN classifier | Classifier performance is explored. |
| 21 | Acharya, U. Rajendra, et al. [29] | 2012 | Adaboost algorithm. | Malignancy index is not predicted. |
| 22 | Acharya, U. Rajendra, et al. [30] | 2015 | Textural features using stationary wavelet transform are extracted from Ultrasonography image. Fuzzy classifier is used. | Highest classification accuracy is 88.9% |
| 23 | Acharya, U. Rajendra, et al. [31] | 2016 | Gabor transform is used to extract the features. Accuracy of 94%. | TMI is not considered and predicted. |
| 24 | Choi, Young Jun, et al. [32] | 2017 | The paper presents the diagnostic performance between CAD system and experienced radiologist. | Specificity and correctness of system found lower than the expert radiologist. |
| 25 | Koundal et al. [33] | 2018 | CAD system is proposed. | Malignancy index is not predicted. |
| 26 | Nugroho, HanungAdi, et al. [34] | 2019 | The geometric and textural features are extracted. SVM is used to detect malignancy. | Thyroid malignancy index is not detected. |
| 27 | Shankar, K., et al. [35] | 2020 | Multi- kernel support vector machine is used. Improved gray wolf optimization technique is used for feature selection. | Focus is only on feature section and malignancy detection. TMI prediction is not done. |

| | | | | |
|---|---|---|---|---|
| 28 | Zhang et al. [40] | 2022 | Deep Convolutional Neural Network is used for thyroid disease classification in multiple classes using USG and CT scan images. | TMI is not detected. |

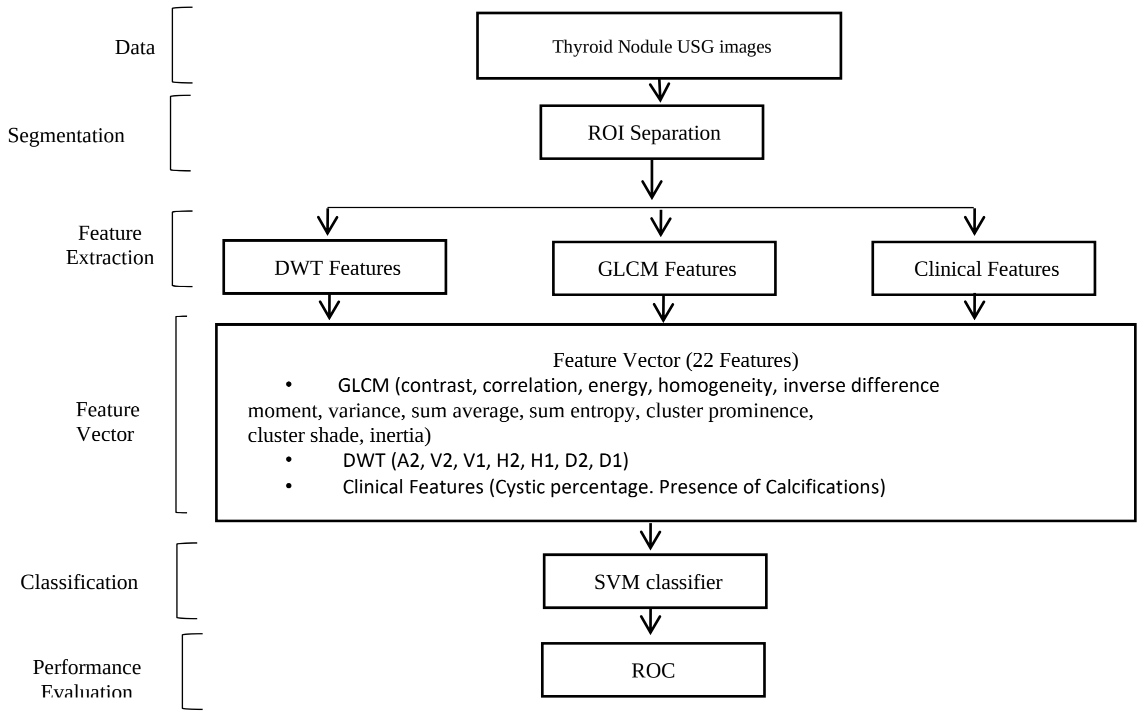



Figure 1: Flow diagram for Part-I (Malignancy Detection)

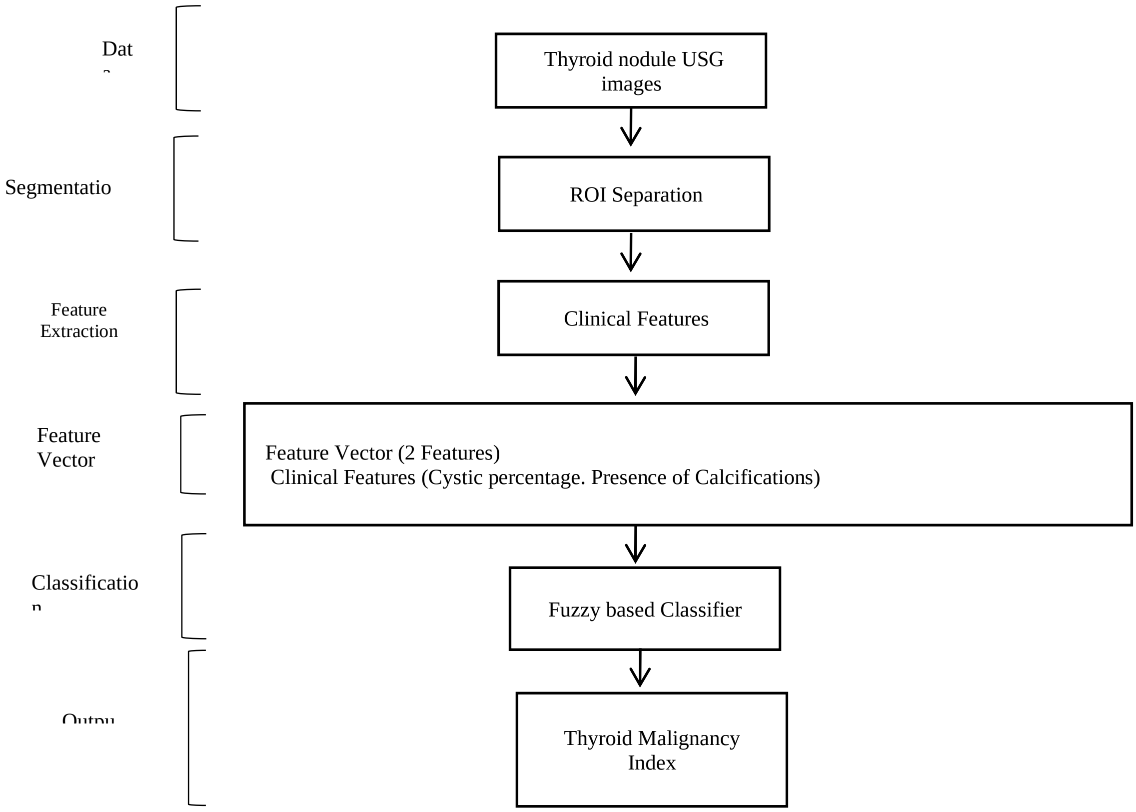



Figure 8: Flow diagram for Part-II (TMI Prediction)