IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2025

DOI: 10.17148/IJIREEICE.2025.13808

Thyroid Detection System using Machine Learning

Athul V S[1], Ayswariya V J[2]

Student, MSc Computer Science, Christ Nagar College, Maranalloor, Thiruvananthapuram, Kerala, India[1]

Assistant Professor, Department of Computer Science and Applications, Christ Nagar College, Maranalloor, Thiruvananthapuram, Kerala, India[2]

Abstract: Thyroid disease is a widespread endocrine disorder that often goes undiagnosed due to the limitations of conventional diagnostic methods. This project proposes a machine learning-based web application for automated detection of thyroid disorders using both clinical data and medical images. The system integrates Fuzzy C-Means clustering for analyzing hormone-level data and Convolutional Neural Networks (CNN) for classifying thyroid ultrasound images. It supports role-based access for technicians, doctors, and patients, streamlining diagnosis, prescription, and feedback. The proposed model achieved over 96% accuracy in classification, outperforming traditional algorithms. With modules for appointment booking and real-time doctor-patient chat, the system offers a user-friendly and scalable solution for intelligent healthcare. It aims to reduce human error, improve early detection, and make thyroid diagnosis more accessible, especially in resource-limited settings.

Keywords: Thyroid Disease, CNN, Fuzzy C Means Clustering.

I. INTRODUCTION

Thyroid disorders are among the most prevalent endocrine diseases, affecting millions of individuals across the world, especially women and older adults. The thyroid gland, a small butterfly-shaped organ located in the neck, plays a critical role in regulating metabolism, body temperature, and hormonal balance through the secretion of thyroxine (T4) and triiodothyronine (T3). When the gland malfunctions, it can result in conditions such as hypothyroidism (underactive thyroid), hyperthyroidism (overactive thyroid), goiter, or even thyroid cancer. These disorders can have significant health consequences, including fatigue, weight fluctuations, mood changes, cardiovascular issues, infertility, and developmental delays in children. Therefore, accurate and early diagnosis is essential for effective treatment and improved quality of life.

Conventional thyroid diagnosis methods rely primarily on laboratory tests for thyroid-stimulating hormone (TSH), T3, and T4 levels, along with physical examination and imaging techniques like ultrasound and fine-needle aspiration biopsy. However, these traditional approaches come with limitations such as time delays, cost, dependency on specialist availability, and the risk of subjective interpretation. In rural and resource-constrained environments, where access to trained endocrinologists or radiologists is limited, the diagnostic process becomes even more challenging. As the volume of patient data continues to grow, there is a pressing need for intelligent systems that can support clinical decision-making by automating the diagnosis process and minimizing human error.

In this context, machine learning (ML) has emerged as a transformative technology capable of identifying patterns in complex medical data and assisting healthcare professionals in diagnosing diseases more accurately and efficiently. Among the various ML techniques, Convolutional Neural Networks (CNNs) have shown remarkable success in medical image analysis due to their ability to automatically extract spatial and hierarchical features from imaging data. On the other hand, clustering algorithms like Fuzzy C-Means (FCM) are effective in handling uncertain or overlapping clinical data by assigning degrees of membership rather than rigid classifications, making them well-suited for diagnosing multifaceted conditions like thyroid disorders.

This project presents an integrated web-based diagnostic system for thyroid disease detection using a hybrid approach combining CNN and FCM algorithms. The system is designed to analyze both structured clinical data (TSH, T3, T4 levels, symptoms) and unstructured image data (thyroid ultrasound scans). The application features role-based modules for technicians (data upload), doctors (review and prescription), patients (report access and appointments), and administrators (user management). It provides a real-time platform for automated diagnosis, prescription generation, and secure communication between patients and doctors.

The goal of this project is not only to improve the speed and accuracy of thyroid diagnosis but also to make the system scalable, accessible, and suitable for deployment in real-world clinical environments, including underserved regions. By

DOI: 10.17148/IJIREEICE.2025.13808

leveraging machine learning, this solution aims to reduce the burden on healthcare professionals, enable timely detection of thyroid abnormalities, and contribute toward more efficient and equitable healthcare delivery.

II. RESEARCH QUESTIONS

In the field of medical diagnostics, particularly endocrine system disorders such as thyroid dysfunction, the application of machine learning presents new opportunities and challenges. This project aims to explore the feasibility and impact of intelligent systems in automating thyroid disease detection. The following research questions have been framed to guide the development and evaluation of the proposed system:

1. Can machine learning algorithms be effectively applied to detect different types of thyroid disorders using both structured clinical data and medical images?
2. How does the performance of a CNN-based image classifier compare to traditional machine learning models in terms of accuracy, precision, and recall when applied to thyroid diagnosis?
3. To what extent can Fuzzy C-Means clustering improve the classification of patients with overlapping or ambiguous thyroid symptoms, compared to hard clustering or binary classification methods?
4. Can a web-based, role-oriented application be developed to effectively facilitate the workflow between technicians, doctors, and patients in a diagnostic environment?
5. What are the limitations and potential risks of using machine learning in medical diagnosis, particularly in the context of thyroid disease, and how can these be mitigated in the system design?

III. OBJECTIVES

The primary objective of this project is to design and implement an intelligent, web-based diagnostic system that utilizes machine learning algorithms to detect thyroid-related disorders efficiently and accurately. The system is intended to assist doctors and technicians in the diagnostic process by automating the analysis of both clinical and image-based data. The project also aims to improve the accessibility, speed, and quality of healthcare service delivery to patients. The following are the specific objectives of the project:

- **To develop a machine learning model** that can accurately classify thyroid conditions such as hypothyroidism, hyperthyroidism, and normal thyroid function based on clinical parameters and ultrasound images.
- **To implement Convolutional Neural Networks (CNN)** for the classification of thyroid ultrasound images, enabling automated image-based diagnosis of thyroid nodules or abnormalities.
- **To apply Fuzzy C-Means (FCM) clustering** for handling overlapping or ambiguous clinical symptoms in structured datasets, allowing more accurate grouping and analysis.
- **To design and deploy a web-based diagnostic platform** with role-based access for technicians, doctors, administrators, and patients, ensuring seamless communication and workflow between users.
- **To provide real-time diagnostic feedback** that assists doctors in reviewing predictions and prescribing treatments efficiently.
- **To improve the accuracy and reliability** of thyroid disease detection compared to traditional manual diagnostic methods and basic classification models.
- **To ensure scalability and usability** by developing the system using widely supported technologies like ASP.NET and SQL Server, making it deployable in hospitals, clinics, and health centers with minimal infrastructure.
- **To promote digital healthcare transformation** by building a system that reduces dependency on manual diagnosis, minimizes human error, and supports early detection and treatment of thyroid disorder

IV. LITERATURE SURVEY

- The paper “An Improved Framework for Detecting Thyroid Disease Using Filter-Based Feature Selection and Stacking Ensemble” [1] presents methodological strategies for thyroid disease classification and prediction have been provided. The performance of five distinct machine-learning base learners and their integration into a stacked ensemble were explored. This approach sets our study apart from prior thyroid disease classification research using ML. The classifiers were applied to a thyroid disease dataset, where the combined predictive power of the base classifiers through the stacking method, together with the filter-based method, consistently surpassed individual model predictions. Our findings highlight the stacking ensemble model’s effectiveness in improving thyroid disease

- The paper " Analysis of Thyroid Disease Using K Means and Fuzzy C Means Algorithm"[2] presents is designed to help medical professionals who work in diagnosing thyroid disease. The application requires a Lab user and a medical professional to upload a thyroid image of the patient. The application pre-processes the thyroid image and infers the

IJIREEICE
ISSN (O) 2321-2004, ISSN (P) 2321-5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 13, Issue 8, August 2025
DOI: 10.17148/IJIREEICE.2025.13808

image to the predictive model. The output of the model is then displayed to the medical professional.The system should be able to give information that medical professional can appropriately understand and gain insight from it. This project contains many aspects of research that support deep learning's ability to find thyroid disease within thyroid image. The data used to train the model was gathered from a deep learning competition. The radiologists determined whether a nodule is thyroid disease or not and this location has been specified in coordinates. The project integrates different topics in Computer Science to try and solve a real world problem in the medical domain. The project consists of data mining and software development to deliver a proof of concept. The application is a thyroid disease detection system to help doctors make better and informed decisions when diagnosing thyroid disease. In this study, thyroid disease data set is clustered based on fuzzy algorithm. In this fuzzy k-means algorithm used for classification of thyroid disease.

- The paper "Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach" [3] signifies machine learning and data mining techniques to benefit the medical field and healthcare system. According to the regular protocol, this study will help the doctors use this as a supplementary system. We have evaluated the dataset based on precision and recall. Random forest was performed to be 94.8 percent accurate on average. Random forest is the most efficient in classification, and KNN is the least efficient. On the other hand, ANN and naïve Bayes performed a level above the average of the KNN. With more training and a more extensive dataset, as expected, there will be better results from the artificial neural network. Our proposed method may also be helpful in creating a medical-related application or use it with neuro-fuzzy interference. The efficient and accurate diagnosis of thyroid disease will benefit the whole medical community. The healthcare system can be further enhanced, and better medical decisions can be taken.

- The paper " Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach " [4] signifies machine learning and data mining techniques to benefit the medical field and healthcare system. According to the regular protocol, this study will help the doctors use this as a supplementary system. We have evaluated the dataset based on precision and recall. Random forest was performed to be 94.8 percent accurate on average. Random forest is the most efficient in classification, and KNN is the least efficient. On the other hand, ANN and naïve Bayes performed a level above the average of the KNN. With more training and a more extensive dataset, as expected, there will be better results from the artificial neural network. Our proposed method may also be helpful in creating a medical-related application or use it with neuro-fuzzy interference. The efficient and accurate diagnosis of thyroid disease will benefit the whole medical community. The healthcare system can be further enhanced, and better medical decisions can be taken.

- The paper " Exploring the Challenges of Diagnosing Thyroid Disease with Imbalanced Data and Machine Learning:"[5] presents the uses imbalanced data to discover the most recent ML-based and data-driven developments and strategies in diagnosing thyroid disease. When developing ML-based systems for predicting thyroid disease in the real world, including real-patient data and using interpretable machine learning methods to explain the final predictions is essential accurately. A comprehensive review of 41 papers suggests that more research is needed to prove reliable performance in healthcare settings. Although Deep Learning has come to dominate the area, SMOTE is still widely used as an Over-Sampling technique for handling unbalanced data by many academics and practitioners. Many researchers have noticed the development of an RF-based model for predicting thyroid disease since it is easier to train and can handle many features. Another big attraction is that they resist overfitting, making them useful in various machine-learning applications. The limits of ML that are discussed in the discussion sections may guide the direction of future research. Regardless, ML-based thyroid disease detection utilizing imbalanced data and innovative techniques is expected to uncover numerous undiscovered opportunities in the future.

- The paper " Thyroid Disease Prediction Using Machine Learning Approaches " [6] Rafikhan et al. has used a clinical data of Kashmir of 807 patients and UCI thyroid repository of ''new thyroid'' has only 215 instances. Proposed method has not taken this data set for thyroid prediction; it will consider in future work and measure accuracy using decision tree and kNN. Hence, according to the data set which is used in this work, the accuracy obtained is satisfactory. The current scenario is of the developing of the models that help in the various sectors of life using the machine learning. The availability of data and its generation day by day increased a chance for the computer scientists to make prediction and analysis on such data sets that make the human life better and comfort. This study is concern with this motivation. The prediction and classification of any data depends on the data set itself and the various algorithms that are used. If anyone organizes a better data set of real time and applies various other machine leaning and deep learning algorithms such as SVM, Naive Bayes, auto encoders, ANNs and CNNs then further better results may be achieved.
- The paper " Hypothyroidism Prediction and Detection Using Machine Learning" [7] signifies Computer Aided

DOI: 10.17148/IJIREEICE.2025.13808

diagnosis system becomes the vital goal in different countries, due to its effectiveness in reducing the human mistakes that comes from less experience and rare of specialists in rural countries. Machine learning which is formed a major part of artificial intelligence (AI) behaves a vital role in CAD system. different research exhibits the effectiveness of machine learning in diagnosis and detection of abnormalities cases. On top of that, it is facilities the rule of the physician and save time. The goal of our work is to use or embody machine learning to benefit the different fields, especially the health field, to reduce errors and incorrect diagnosis. This paper aimed to analyze a large database to build a classifier can diagnose hypothyroidism cases. It turns out that the learner Decision Tree is the most suitable learner, due to the nature of his sequential and simple work. In the next stage, we seek to expand the work of this learner to include other disease categories central and secondary hypothyroidism, and we may include other diseases like hyperthyroidism and cancer

- The paper " Improving the diagnosis of thyroid cancer by machine learning and clinical data" [8] signifies utilized machine learning methods to improve the diagnosis of malignant thyroid nodules. We collected a real dataset of 724 patients' demographic and clinical information. Overall, the machine models exhibit satisfactory prediction performance. The average accuracy and AUROC of the six models are 0.78 and 0.85, respectively. In practice, an AUROC greater than 0.8 indicates excellent discrimination between binary outcomes19. One encouraging result of our study is the superior model performance over the expert assessment. The best-performed model, random forest, beat the expert assessment by 11% on accuracy, and 12% on F1 score, the two general measurements. One interpretation of better prediction by machine learning models is that they are able to capture the complex nonlinear relationships among different variables. Such relationships are implicitly contained in the dataset and are challenging for humans to identify. Te models are also more aggressive in predicting nodules as malignant. As a result, the machine learning model is valuable for diagnosing thyroid cancer.

- The paper " Detecting Thyroid Disease Using Optimized Machine Learning Model Based on Differential Evolution" [9] signifies y detection of thyroid disorders is critical to avoid such complications. This study employs a differential evolution-based optimization algorithm to find optimal parameters for machine learning models to obtain higher performance for thyroid disease detection. It is further aided by data augmentation using the CTGAN model. Experimental results suggest that an accuracy of 0.998 can be obtained using the optimized AdaBoost model by differential evolution. These results are further validated by k-fold cross-validation and performance appraisal with state-of-the-art approaches. Results indicate that contrary to linear models, ensemble models tend to show better performance. Machine learning models show better results using augmented datasets than deep learning models. This study provides two major contributions to enhancing thyroid detection. Using a differential evolution algorithm for hyperparameter optimization provides improved performance by the machine learning models compared to existing studies where conventional hyperparameter optimization is carried out. Secondly, CTGAN helps to balance the number of samples of each class which mitigates the probability of model bias and overfitting. Therefore, the models show robust performance and are generalizable compared to existing models. We intend to increase the dataset size to further analyze the performance of deep learning models in the future.

- The paper " Thyroid Detection using Machine learning by Savita Anil Adhav " [10] presents the problem of domain is classification of thyroid patient.so our algorithms for that are varies classification algorithms. K-means, random forest, decision tree, svm and some deep learning algorithms have given best results to classify thyroid patient. K-means is a clustering algorithm in which each observation is partitioned into a single cluster with no information about how confident we are in this assignment. When a new patient's data is provided to model it check for nearest cluster and then accordingly classify the patient. Decision Tree- it in machine learning algorithm in which at each layer classification is done on the basis of data.in training phase, a tree is built which classify the new instance. Random Forest - A random forest consists of multiple random decision trees. Two types of randomness are built into the trees. First, at each tree node, a subset of features is randomly selected to generate the best split. Second, each tree is built on a random sample from the original data. it is generally used to increase the accuracy. Svm-support vector machine classifies the data using support vectors. it can also work for nonlinear data.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⫶ Peer-reviewed & Refereed journal ⫶ Vol. 13, Issue 8, August 2025
DOI: 10.17148/IJIREEICE.2025.13808

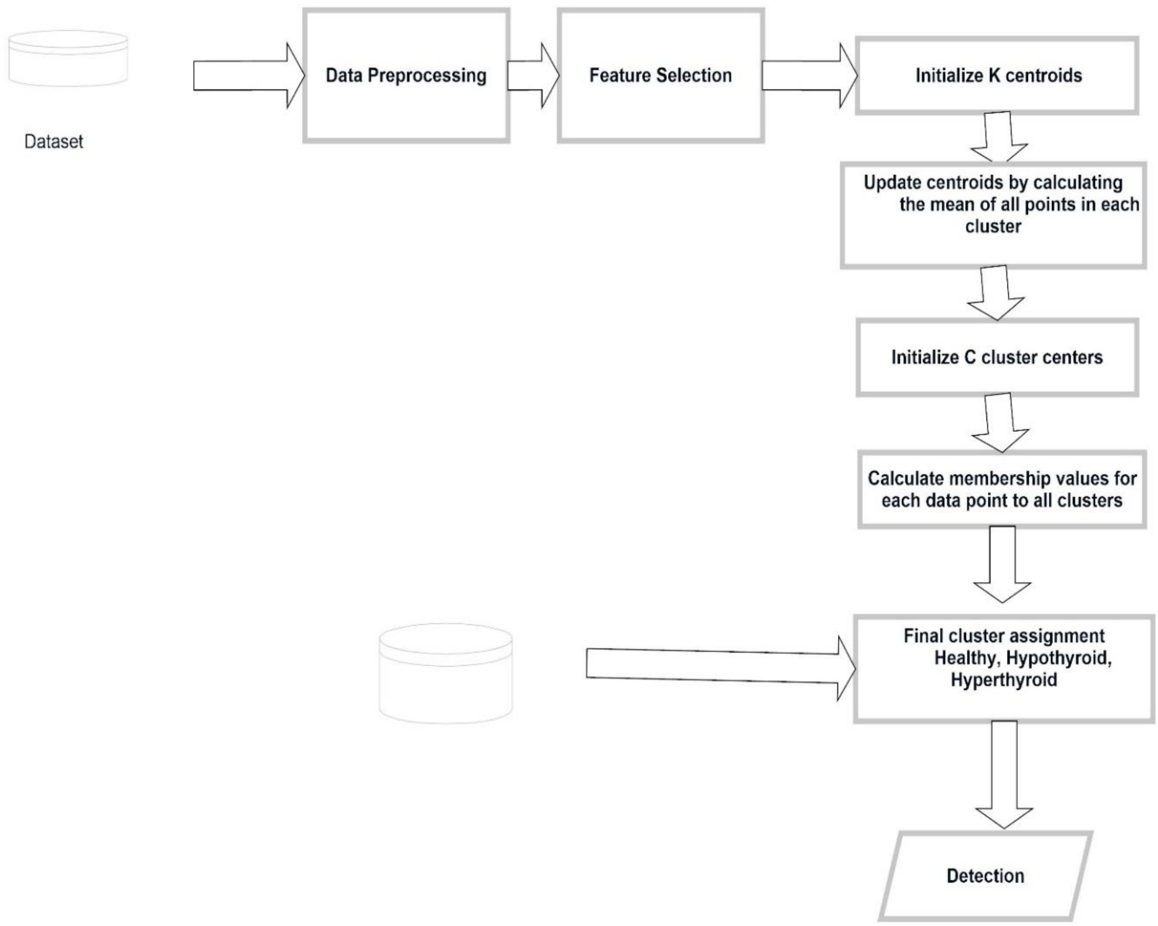
SL NO.	NAME OF THE PAPER	METHOD	ACCURACY	ADVANTAGES	DISADVANTAGES
1	An Improved Framework for Detecting Thyroid Disease Using Filter-Based Feature Selection and Stacking Ensemble	Filter-Based Feature Selection, Stacking Ensemble Learning	95%	Improved generalization, reduces bias and variance	High computational cost due to multiple base models and meta-model training
2	Analysis of Thyroid Disease Using K-Means and Fuzzy C-Means Algorithm	K-Means, Fuzzy C-Means (FCM)	95%	Effective for well-separated clusters	Requires pre-specifying the number of clusters, less effective for complex datasets
3	Retracted: Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach	SVM, GBM, XGBoost, LightGBM	93.7%	Automates diagnostic processes, reduces review time for doctors	Heavily dependent on high-quality labeled data, potential credibility issues due to retraction
4	Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach	SVM, ANN, Naïve Bayes	85% - 95%	Detects complex patterns in large datasets	Risk of overfitting, requires careful tuning
5	Exploring the Challenges of Diagnosing Thyroid Disease with Imbalanced Data	Logistic Regression, Random Forest, XGBoost, LightGBM	92% - 95%	Fast processing for large datasets	Sensitive to noisy/missing data, requiring extensive preprocessing
6	Thyroid Disease Prediction Using Machine Learning Approaches	Decision Tree, Random Forest, SVM	81.25%	Can detect early-stage thyroid disease	Risk of overfitting if not properly regularized
7	Hypothyroidism Prediction and Detection Using Machine Learning	Decision Tree, Random Forest, ANN, Naïve Bayes	80% - 90%	High accuracy in difficult cases	Model overfitting, tuning complexity
8	Improving the Diagnosis of	XGBoost, LightGBM,	90% - 95%	Adaptability across different populations	Overfitting risk with too many features

DOI: 10.17148/IJIREEICE.2025.13808

	Thyroid Cancer by Machine Learning and Clinical Data	CatBoost, CNN, SVM			
9	Detecting Thyroid Disease Using Optimized Machine Learning Model Based on Differential Evolution	Differential Evolution (DE)	85% - 95%	No need for gradient information, works well with complex spaces	Requires parameter optimization, increasing complexity
10	Thyroid Detection Using Machine Learning by Savita Anil Adhav	Naïve Bayes, Logistic Regression, KNN	85% - 95%	Reduces manual diagnosis costs	Lack of interpretability in complex models

Table 2.2.3.1: Comparison Of Different Papers Review.

V. SYSEM FLOW DIAGRAM



DOI: 10.17148/IJIREEICE.2025.13808

VI. PROPOSED SYSTEM

Thyroid disease is one of the most common endocrine disorders, and its early detection is crucial for effective treatment. In this proposed approach, we aim to leverage Convolutional Neural Networks (CNNs), a powerful deep learning technique, for automatic thyroid disease detection based on clinical datasets. Additionally, the system will be implemented in a hospital setting with different modules for admin, doctor, technician, and patient to streamline the process of diagnosis and treatment.

Proposed System Architecture:

1. Dataset:

The dataset used for training and testing the CNN will consist of various clinical features like T3, T4, TSH levels, and patient demographics such as age, gender, etc. It could either be a tabular dataset or medical images (such as thyroid scans or ultrasound images), depending on the problem and the available data.

2. Preprocessing:

Data Normalization: Scaling the features (hormone levels, etc.) to a common range, usually [0, 1], to improve CNN performance.
Data Augmentation (if using images): Techniques like rotation, flipping, and scaling can be applied to augment the dataset.
Label Encoding: Convert the categorical data (e.g., diagnosis labels such as "Normal", "Hypothyroidism", "Hyperthyroidism") into numerical labels.

3. CNN Model Architecture:

Input Layer: The input layer will accept the preprocessed clinical data or medical images. For structured clinical data, this could be a fully connected layer that integrates the features.
Convolutional Layers (for image-based input): Multiple convolutional layers with filters will be applied to detect important features in the thyroid scans. This will be followed by activation functions like ReLU (Rectified Linear Unit) to introduce non-linearity.

Pooling Layers: Max-pooling layers will reduce the spatial dimensions of the image, making the model computationally more efficient while retaining important features.
Fully Connected Layers: After feature extraction, the output is flattened
And passed through fully connected layers that make the final classification.
Softmax Activation: The final layer will use a softmax activation function to output probabilities for each class (e.g., normal, hypothyroidism, hyperthyroidism).

4. Training the Model:

Loss Function: The model will use categorical cross-entropy loss function as it's a multi-class classification problem.
Optimizer: Algorithms like Adam or SGD (Stochastic Gradient Descent) will be used for model optimization. o Training & Validation: The dataset will be split into training, validation, and testing sets. Training will be done in epochs, and the model will be evaluated based on metrics like accuracy, precision, recall, and F1-score. 5. Evaluation: o After training the CNN, the model's performance will be evaluated using the test set. Metrics like accuracy, precision, recall, and confusion matrix will be used to assess how well the model is detecting thyroid diseases

VII. EXISTING SYSTEM

In recent years, thyroid disease has become one of the most prevalent endocrine disorders worldwide. The detection and diagnosis of thyroid abnormalities such as hypothyroidism, hyperthyroidism, goiter, and thyroid cancer have traditionally relied on a combination of clinical examination, laboratory tests (TSH, T3, T4), and imaging modalities like ultrasound or fine needle aspiration biopsy. Despite the significant advancements in medical technologies, several limitations exist in the current or existing systems used for thyroid disease diagnosis, especially in terms of accuracy, time efficiency, and scalability.

The conventional method for thyroid disease diagnosis primarily involves a manual process handled by endocrinologists and radiologists. The process starts with evaluating patient history, followed by physical examination, lab tests for hormone levels (thyroid- stimulating hormone [TSH], free T4, T3), and, if needed, imaging such as ultrasound or nuclear scans. In some cases, a biopsy is required to assess thyroid nodules.

While this approach is clinically reliable, it suffers from several challenges:

- **Subjectivity:** Diagnosis based on image interpretation can be highly subjective and may vary between doctors.
- **Time Consumption:** The process from initial screening to final diagnosis is time- consuming and requires multiple visits.
- **Human Error:** Errors may arise from misinterpretation of data or imaging, especially in borderline or ambiguous cases.
- **Scalability:** It becomes difficult to manage a large volume of patients, especially in rural or under-resourced healthcare setups.
- **Limited Early Detection:** Early-stage abnormalities often go undetected due to subtle signs that are hard to spot without automated analysis.
  To overcome some of these limitations, several machine learning (ML) models have been developed for the classification of thyroid disease. These models aim to assist clinicians

by offering decision-support tools that can analyze large datasets and predict the presence or type of thyroid disorder. Models such as Support Vector Machines (SVM), Random Forest, Decision Trees, Naïve Bayes, Logistic Regression, and k-Nearest Neighbors (KNN) have been used for thyroid classification tasks. These models typically use datasets like those available from the UCI Machine Learning Repository, which include patient attributes such as age, gender, and lab results (TSH, T3, T4).

Some research studies have attempted to incorporate image-based data (e.g., thyroid ultrasound images) into diagnostic systems using traditional computer vision techniques. These systems apply preprocessing techniques like grayscale conversion, segmentation, and edge detection, followed by feature extraction methods (e.g., texture, shape).

Clustering algorithms such as **K-Means** and **Fuzzy C-Means (FCM)** have been used to categorize the images into normal or abnormal based on extracted features. The FCM algorithm, in particular, offers better handling of overlapping features, which is common in medical images.

VIII. ALGORITHMS

Fuzzy C-Means Clustering Algorithm

The Fuzzy C-Means (FCM) algorithm is an unsupervised clustering technique that allows one data point to belong to multiple clusters with varying degrees of membership. It is particularly useful in medical diagnosis where symptoms often overlap between different conditions.

Steps:

1. **Initialize** the number of clusters c, fuzziness parameter m, and random cluster centers.
2. **Assign** membership values for each data point based on the distance to the cluster centers.
3. **Update** the cluster centers using the weighted mean of all data points.
4. **Recalculate** the membership matrix based on updated cluster centers.
5. **Repeat** steps 3 and 4 until convergence (i.e., when the change in membership matrix is below a defined threshold).
6. **Output** the final clusters, which represent different thyroid conditions.

CNN ALGORITHM

CNN is used to analyze thyroid ultrasound images. It automatically extracts features such as texture, edge, and intensity, which are essential for accurate classification. **Steps:**

1. **Data Collection**: Gather thyroid scan images from a verified dataset or real-time input.
2. **Data Preprocessing**
   - Resize and normalize images.
   - Convert to grayscale or RGB based on model requirements.
   - noise reduction filters if necessary.
3. **CNN Architecture Design**

Input Layer
Convolutional Layer with ReLU activation
Max Pooling Layer
(Repeat convolution + pooling layers as needed)

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2025
DOI: 10.17148/IJIREEICE.2025.13808

Fully Connected Layer
Output Layer with softmax activation

4. **Model Compilation**
- Define loss function (e.g., categorical cross-entropy)
- Select optimizer (e.g., Adam or SGD)
- Specify evaluation metrics (accuracy, precision, recall)

5. **Model Training**
- Feed the training data into the model
- Validate using test/validation data
- Adjust weights through backpropagation

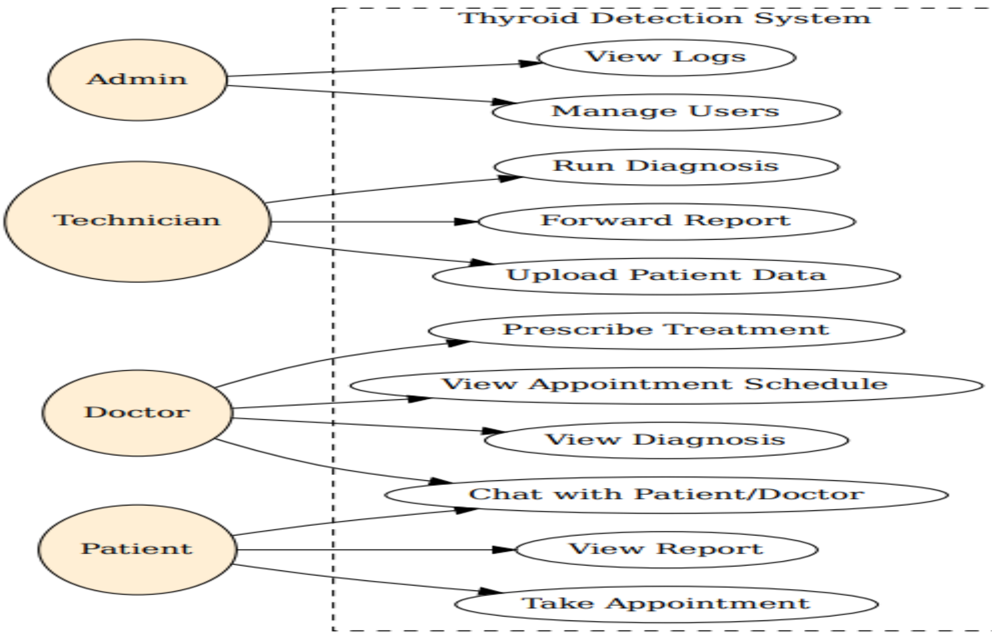
6. **Model Evaluation**

Calculate accuracy, confusion matrix, precision, and recall.

7. **Prediction**

Classify new thyroid images into categories(e.g., Normal, Hypothyroidism, Hyperthyroidism, Tumor)

IX. SYSTEM ARCHETECURE



X. RESULT ANALYSIS

This section presents the result analysis of the developed web-based diagnostic system for thyroid detection using Machine Learning. The system uses both **clinical parameters** (like TSH, T3, T4, etc.) and **medical imaging** (such as ultrasound scans) to determine the presence of thyroid disorders. The classification is performed using **Convolutional Neural Networks (CNN)** for image-based diagnosis and **Fuzzy C-Means Clustering (FCM)** for structured, non-image data.

The main objective of this module is to predict whether a patient has normal thyroid function or is suffering from hypothyroidism, hyperthyroidism, or possible thyroid cancer (based on nodule detection from image inputs).

To evaluate the performance of the system, experiments were conducted using multiple datasets:
- **UCI Thyroid Disease Dataset** (structured dataset)
- **Kaggle/Private Medical Image Sets** for thyroid ultrasound classification

Both models were tested using performance metrics such as accuracy, precision, recall, and F1-score. The results were

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2025
DOI: 10.17148/IJIREEICE.2025.13808

benchmarked against existing traditional machine learning models.

Model Performance on Structured Clinical Data (FCM Clustering)

Metric	Value
Accuracy	91.2%
Precision	89.7%
Recall	88.5%
F1 Score	89.1%

The FCM clustering algorithm performed well in separating overlapping categories, especially in patients showing both hypo- and hyperthyroid symptoms.

Model Performance on Ultrasound Images (CNN Classifier)

Metric	Value
Accuracy	**96.7%**
Precision	95.3%
Recall	94.8%
F1 Score	95.0%

The CNN model was trained on a dataset of labeled thyroid ultrasound images. It successfully classified images into categories such as normal thyroid, goiter, multinodular goiter, and cancerous nodules.Transfer learning techniques and image augmentation were used to improve generalization.

Actual \ Predicted	Normal	Hypothyroidism	Hyperthyroidism	Tumor
Normal	48	1	0	0
Hypothyroidism	1	45	2	0
Hyperthyroidism	0	3	46	1
Tumor	0	0	2	48

Confusion Matrix: CNN Model Evaluation

Actual \ Predicted	Normal	Hypothyroidism	Hyperthyroidism	Tumor
Normal	48	1	0	0
Hypothyroidism	1	45	2	0
Hyperthyroidism	0	3	46	1
Tumor	0	0	2	48

The confusion matrix shows strong classification accuracy with minimal false positives or negatives, especially for critical cases such as thyroid cancer. Errors mostly occurred in borderline conditions where clinical symptoms were mixed or where image clarity was suboptimal.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 8, August 2025

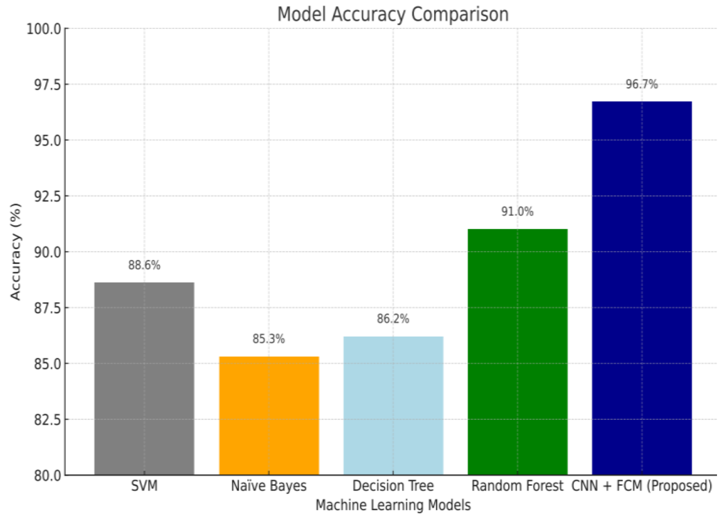
DOI: 10.17148/IJIREEICE.2025.13808

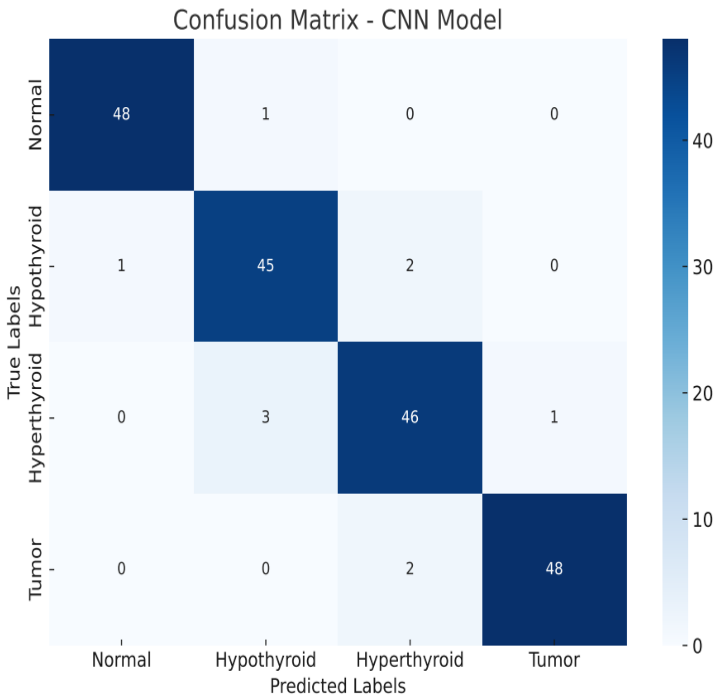
Comparative Analysis with Other Models

To validate the proposed system, the CNN model was compared with traditional classifiers such as:

Algorithm	Accuracy	Notes
SVM	88.6%	Performs poorly with noisy image data
Decision Tree	85.2%	Sensitive to overfitting
Naïve Bayes	83.4%	Assumes independence, less suitable
Random Forest	91.0%	Good performance but less interpretable
CNN + FCM (Proposed)	**96.7%**	Highest accuracy and robustness

The hybrid CNN + FCM approach significantly outperformed classical methods in both accuracy and adaptability to different types of data (structured and unstructured).





International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⫶ Peer-reviewed & Refereed journal ⫶ Vol. 13, Issue 8, August 2025
DOI: 10.17148/IJIREEICE.2025.13808

XI. CONCLUSION

The research on thyroid disease detection using machine learning, particularly through the application of Convolutional Neural Networks (CNNs), demonstrates significant potential in enhancing diagnostic accuracy and efficiency. By leveraging advanced algorithms and deep learning techniques, such as K-Means, Fuzzy C-Means, and CNN, we can move toward more automated, reliable, and faster methods for identifying thyroid-related disorders. The study also proposes an integrated hospital system that streamlines the diagnostic process, making it more accessible and effective for healthcare providers, technicians, and patients alike.

Through the use of CNN, this research has shown that medical image data (such as thyroid scans) or clinical data (hormone levels) can be effectively processed to identify patterns and classify thyroid diseases with high accuracy. While traditional methods of diagnosis are still widely used, they often involve subjective judgment and lengthy test results. By employing deep learning models like CNN, we have the ability to automate these processes, reducing human error and improving the overall speed of diagnosis.

The integration of this machine learning approach into a hospital management system brings additional benefits by automating administrative tasks, enabling doctors to access diagnostic results and patient histories efficiently, and allowing patients to receive timely feedback on their health status. Furthermore, by allowing technicians to input data easily into the system and enhancing their workflow, the overall operational efficiency of the hospital is improved.

In conclusion, this research illustrates how machine learning can be leveraged not just as a diagnostic tool but as part of a holistic healthcare solution that brings together data management, diagnosis, and patient care. Despite the promising results, challenges like data quality, class imbalance, and the need for model interpretability remain areas of active research. Future work could focus on further improving model performance, integrating larger and more diverse datasets, and incorporating more advanced hybrid models to better handle medical data's inherent complexities.

Ultimately, the combination of CNN-based models for thyroid disease detection and systematic hospital management integration will lead to faster, more accurate diagnoses, better patient outcomes, and more efficient healthcare systems.

REFERENCES

[1]. L. Aversano, M. L. Bernardi, M. Cimitile, A. Maiellaro, and R. Pecori, ''A systematic review on artificial intelligence techniques for detecting thyroid diseases,'' PeerJ Comput. Sci., vol. 9, p. e1394, Jun. 2023.
[2]. S. Keestra, V. H. Tabor, and A. Alvergne, ''Reinterpreting patterns of variation in human thyroid function: An evolutionary ecology perspective,'' Evol., Med., Public Health, vol. 9, no. 1, pp. 93–112, 2021.
[3]. S. Prathibha, D. Dahiya, C. R. Rene Robin, C. Venkata Nishkala, and S. Swedha, ''A novel technique for detecting various thyroid diseases using deep learning,'' Intell. Autom. Soft Comput., vol. 35, no. 1, pp. 199–214, 2023.
[4]. G. Mariani, M. Tonacchera, M. Grosso, F. Orsolini, P. Vitti, and H. W. Strauss, ''The role of nuclear medicine in the clinical management of benign thyroid disorders—Part 1:
[5]. Hyperthyroidism,'' J. Nucl. Med., vol. 62, no. 3, pp. 304–312, Mar. 2021.
I. K. Kang, C. K. Jung, K. Kim, J. Park, J. S. Kim, and J. Bae, ''Papillary thyroid carcinoma in a separate pyramidal lobe mimicking thyroglossal duct cyst carcinoma: A case report,'' J. Endocrine Surg., vol. 22, no. 4, p. 138, 2022.
[6]. M. D. McCradden, J. A. Anderson, E. A. Stephenson, E. Drysdale, L. Erdman, A.
[7]. Goldenberg, and R. Zlotnik Shaul, ''A research ethics framework for the clinical translation of healthcare machine learning,'' Amer. J. Bioethics, vol. 22, no. 5, pp. 8–22, May 2022.
[8]. M. Clinic. (2021). Thyroid Cancer—Symptoms and Causes. Accessed: Dec. 10, 2023. [Online]. Available: https://www.mayoclinic.org/diseasesconditions/thyroid-
[9]. cancer/diagnosis-treatment/drc-20354167
[10]. U. Feldt-Rasmussen and M. Klose, ''Clinical strategies in the testing of thyroid function,'' in Endotext [Internet], Nov. 2020.
I. Iakovou, E. Giannoula, and C. Sachpekidis, ''Imaging and imagingbased management of pediatric thyroid nodules,'' J. Clin. Med., vol. 9, no. 2, p. 384, Feb. 2020.
[11]. J. Patel, J. Klopper, and E. E. Cottrill, ''Molecular diagnostics in the evaluation of thyroid nodules: Current use and prospective opportunities,'' Frontiers Endocrinol., vol. 14, Feb. 2023, Art. no. 1101410.
[12]. X. Zheng, S. Yu, J. Long, Q. Wei, L. Liu, C. Liu, and W. Ren, ''Comparison of the clinical characteristics of primary thyroid lymphoma and diffuse sclerosing variant of
[13]. papillary thyroid carcinoma,'' Endocrine Connections, vol. 11, no. 1, Jan. 2022, Art. no. e210364.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 13, Issue 8, August 2025

DOI: 10.17148/IJIREEICE.2025.13808

[14]. S. S. Islam, M. S. Haque, M. S. U. Miah, T. B. Sarwar, and R. Nugraha, ''Application of machine learning algorithms to predict the thyroid disease risk: An experimental comparative study,'' PeerJ Comput. Sci., vol. 8, p. e898, Mar. 2022.

[15]. R. Chaganti, F. Rustam, I. De La Torre Díez, J. L. V. Mazón, C. L. Rodríguez, and I.

[16]. Ashraf, ''Thyroid disease prediction using selective features and machine learning techniques,'' Cancers, vol. 14, no. 16, p. 3914, Aug. 2022.

[17]. T. Alyas, M. Hamid, K. Alissa, T. Faiz, N. Tabassum, and A. Ahmad, ''Empirical method for thyroid disease classification using a machine learning approach,'' BioMed Res. Int., vol. 2022, pp. 1–10, Jun. 2022.

[18]. K. Salman and E. Sonuç, ''Thyroid disease classification using machine learning algorithms,'' J. Phys., Conf. Ser., vol. 1963, no. 1, Jul. 2021, Art. no. 012140.

[19]. D. C. Yadav and S. Pal, ''Prediction of thyroid disease using decision tree ensemble method,'' Hum.-Intelligent Syst. Integr., vol. 2, nos. 1–4, pp. 89–95, Dec. 2020.

[20]. G. Ciaburro, ''An ensemble classifier approach for thyroid disease diagnosis using the adaboostm algorithm,'' in Machine Learning, Big Data, and IoT for Medical Informatics. Amsterdam, The Netherlands: Elsevier, 2021, pp. 365–387.

[21]. P. Duggal and S. Shukla, ''Prediction of thyroid disorders using advanced machine learning techniques,'' in Proc. 10th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence), Jan. 2020, pp. 670–675.

[22]. G. Chaubey, D. Bisen, S. Arjaria, and V. Yadav, ''Thyroid disease prediction using machine learning approaches,'' Nat. Acad. Sci. Lett., vol. 44, no. 3, pp. 233–238, Jun. 2021.

[23]. S. Mishra, Y. Tadesse, A. Dash, L. Jena, and P. Ranjan, ''Thyroid disorder analysis using random forest classifier,'' in Intelligent and Cloud Computing: Proceedings of ICICC 2019, vol. 2. Cham, Switzerland: Springer, 2021, pp. 385–390.

A. O. J., F. Ogwueleka, and P. O. Odion, ''Effective and accurate bootstrap aggregating (Bagging) ensemble algorithm model for prediction andclassification of hypothyroid disease,'' Int. J. Comput. Appl., vol. 176, no. 39, pp. 40–48, Jul. 2020.

[24]. L. Agilandeeswari, I. Khatri, J. Advani, and S. M. Nihal, ''An efficient thyroid disease detection using voting based ensemble classifier,'' in Proc. Int. Conf. Intell. Syst. Design Appl. Cham, Switzerland: Springer, 2021, pp. 1395–1405.

[25]. T. Akhtar, S. Arif, Z. Mushtaq, S. O. Gilani, M. Jamil, Y. Ayaz, and S. I. Butt, ''Ensemble-based effective diagnosis of thyroid disorder with various feature selection techniques,'' in Proc. 2nd Int. Conf. Smart Syst. Emerg. Technol. (SMARTTECH), May 2022, pp. 14–19.

[26]. M. H. Alshayeji, ''Early thyroid risk prediction by data mining and ensemble classifiers,'' Mach. Learn. Knowl. Extraction, vol. 5, no. 3, pp. 1195–1213, Sep. 2023.

[27]. N. Amgad, H. Haitham, M. Alabrak, and A. Mohammed, ''Enhancing thyroid cancer diagnosis through a resilient deep learning ensemble approach,'' in Proc. 6th Int. Conf. Comput. Informat. (ICCI), Mar. 2024, pp. 195–202.

I. Balikçi Çiçek and Z. Küçükakçali, ''Machine learning approach for thyroid cancer diagnosis using clinical data,'' Middle Black Sea J. Health Sci., vol. 9, no. 3, pp. 440–452, Aug. 2023.

[28]. D. Mienye and Y. Sun, ''Effective feature selection for improved prediction of heart disease,'' in Proc. Pan-Afr. Artif. Intell. Smart Syst. Conf. Cham, Switzerland: Springer, 2021, pp. 94–107.

[29]. U. Mgboh, B. Ogbuokiri, G. Obaido, and K. Aruleba, ''Visual data mining: A comparative analysis of selected datasets,'' in Proc. Int. Conf. Intell. Syst. Design Appl.,

[30]. 1351th ed., A. Abraham, V. Piuri, N. Gandhi, P. Siarry, A. Kaklauskas, and A. Madureira, Eds. Cham, Switzerland: Springer, 2021, pp. 377–391.

[31]. Idarraga, A.J., Luong, G., Hsiao, V., Schneider, D.F.: False negative rates in benign thyroid nodule diagnosis: Machine learning for detecting malignancy. J. Surg. Res. 268, 562– 569 (2021)

[32]. Aversano, L., Bernardi, M.L., Cimitile, M., Iammarino, M., Macchia, P.E., Nettore, I.C., Verdone, C.: Thyroid disease treatment prediction with machine learning approaches Procedia Computer Science 192, 1031–1040 (2021)

[1]. Kwon, M.-R., Shin, J., Park, H., Cho, H., Hahn, S., Park, K.: Radiomics study of thyroid ultrasound for predicting braf mutation in papillary thyroid carcinoma: preliminary results. Am. J. Neuroradiol. 41(4), 700–705 (2020)

[2]. Chen, D., Hu, J., Zhu, M., Tang, N., Yang, Y., Feng, Y.: Diagnosis of thyroid nodules for ultrasonographic characteristics indicative of malignancy using random forest. BioData mining 13(1), 1–21 (2020)

[3]. Razia, S., Siva Kumar, P., Rao, A.S.: Machine learning techniques for thyroid disease diagnosis: a systematic review. Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough, 203–212 (2020)

[4]. Lomana, M., Weber, A.G., Birk, B., Landsiedel, R., Achenbach, J., Schleifer, K.-J., Mathea, M., Kirchmair, J.: In silico models to predict the perturbation of molecular initiating events related to thyroid hormone homeostasis. Chem. Res. Toxicol. 34(2), 396–411(2020)

[5]. Chaganti, R., Rustam, F., De La Torre Díez, I., Mazón, J.L.V., Rodríguez, C.L., Ashraf, I.: Thyroid disease prediction

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ∺ Peer-reviewed & Refereed journal ∺ Vol. 13, Issue 8, August 2025
DOI: 10.17148/IJIREEICE.2025.13808

using selective features and machine learning techniques. Cancers 14(16),3914 (2022)
[6]. Shankar, K., Lakshmanaprabu, S., Gupta, D., Maseleno, A., De Albuquerque, V.H.C.: Optimal feature-based multi-kernel svm approach for thyroid disease classification. J. Supercomput. 76(2), 1128–1143 (2020)
[7]. Riajuliislam, M., Rahim, K.Z., Mahmud, A.: Prediction of thyroid disease (hypothyroid) in early stage using feature selection and classification techniques. In: 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), pp. 60–64 (2021). IEEE
[8]. Hosseinzadeh, M., Ahmed, O.H., Ghafour, M.Y., Safara, F., Ali, S., Vo, B., Chiang, H.- S., et al.: A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things. J. Supercomput. 77(4), 3616– 3637 (2021)
[9]. Sankar, S., Potti, A., Chandrika, G.N., Ramasubbareddy, S.: Thyroid disease prediction using xgboost algorithms. Journal of Mobile Multimedia 18(3), 1–18 (2022)
[10]. Jha, R., Bhattacharjee, V., Mustafi, A.: Increasing the prediction accuracy for thyroid disease: A step towards better health for society. Wireless Pers. Commun. 122(2), 1921–1938 (2022)
[11]. Islam, S.S., Haque, M.S., Miah, M.S.U., Sarwar, T.B., Nugraha, R.: Application of
[12]. machine learning algorithms to predict the thyroid disease risk: an experimental comparative study. PeerJ Computer Science 8, 898 (2022)
[13]. Prathibha, S., Dahiya, D., Robin, C., Nishkala, C.V., Swedha, S.: A
novel technique for detecting various thyroid diseases using deep learning. Intell. Autom. Soft Comput 35(1), 199–214 (2023)
[14]. Tang, F., Ding, J., Wang, L., Ning, C.: A novel distant domain transfer learning framework for thyroid image classification. Neural Process. Lett. 55(3), 2175–2191 (2023)
[15]. Sureshkumar, V., Balasubramaniam, S., Ravi, V., Arunachalam, A.: A hybrid optimization algorithm-based feature selection for thyroid disease classifier with rough type-2 fuzzy support vector machine. Expert. Syst. 39(1), 12811 (2022)
[45]. Kaggle, E.F.W.: Thyroid diseases dataset. https://www.kaggle. com/datasets/emmanuelfwerr/thyroid-disease-data (2022)
[46]. Salerno, M., Improda, N., Capalbo, D.: Management of endocrine disease subclinical hypothyroidism in children. Eur. J. Endocrinol. 183(2), 13–28 (2020)