Advanced predictive models for thyroid disease comorbidities using machine learning and deep learning: a comprehensive review

Mohammed Yacoob B. A., Jayashree J.
School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology (VIT University), Vellore, India

Article Info

Article history:

Received Jul 8, 2024
Revised Nov 19, 2024
Accepted Dec 15, 2024

Keywords:

Comorbidities
Deep learning
Endocrine disorders
Machine learning
PCOS
Predictive models

ABSTRACT

With advances in machine learning (ML) and deep learning (DL), the future of thyroid disease diagnosis and prognosis looks very bright. The integration of various data such as imaging and medical record data has increased the accuracy of the model. Advanced DL models such as convolutional neural network (CNN) and recurrent neural network (RNN) further improved disease detection in precision medicine. However, some of the major disadvantages of effective clinical integration include unbalanced samples, unclear sampling, having to communicate in different populations, decreased physician confidence due to the vagueness of current models therefore, and few studies available to identify thyroid comorbidities such as polycystic ovary syndrome (PCOS) and thyroid eye disease (TED) in a variety of different populations to develop the line. It is important to focus future research activities on model definition and validation an improving and thus the diagnosis and prognosis of thyroid comorbidities is of utmost importance. What this will bring is ML and DL, an opportunity to make very significant improvements in the diagnosis, treatment, and management of thyroid diseases, thereby improving patient outcomes and health care by seeking crystals as a group they work interdisciplinary to collaborate in developing flexible solutions, sharing knowledge, and responding to these stated deficiencies.

Corresponding Author:

Jayashree J.
School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology (VIT University)
Vellore, Tamil Nadu, India
Email: jayashree.j@vit.ac.in

1. INTRODUCTION

About 200 million people worldwide suffer from thyroid disease, iodine deficiency affects more than 40% of the world population and is a major cause of thyroid problems. In India, thyroid disease affects about 10% of the population, including women aged 17 to 54 years, are more affected [1], [2]. Because iodine deficiency impairs thyroid hormone production, it contributes to significant increases in morbidity and mortality worldwide [3]. Triiodothyronine (T3) and thyroxine (T4) regulate metabolism and affect bodily functions including heart rate and temperature. Hormone abnormalities caused by an underactive thyroid can lead to symptoms such as high blood pressure, poor circulation, and high cholesterol [4]. At least 1 in 8 women have an underactive thyroid, making it more common in women. Bauer *et al.* [5] hypothyroidism is severely underdiagnosed in 20% of postmenopausal women, making it more difficult to diagnose psychosocial issues caused by thyroid dysfunction [6].

Both hypothyroidism and hyperthyroidism can lead to thyroid eye disease (TED), which affects eye health and has symptoms such as macular degeneration and vision loss. They can also lead to disorders such as polycystic ovary syndrome (PCOS) characterized by insulin resistance and hormonal imbalances [7], [8]. Figure 1 shows the architecture of thyroid prediction system.

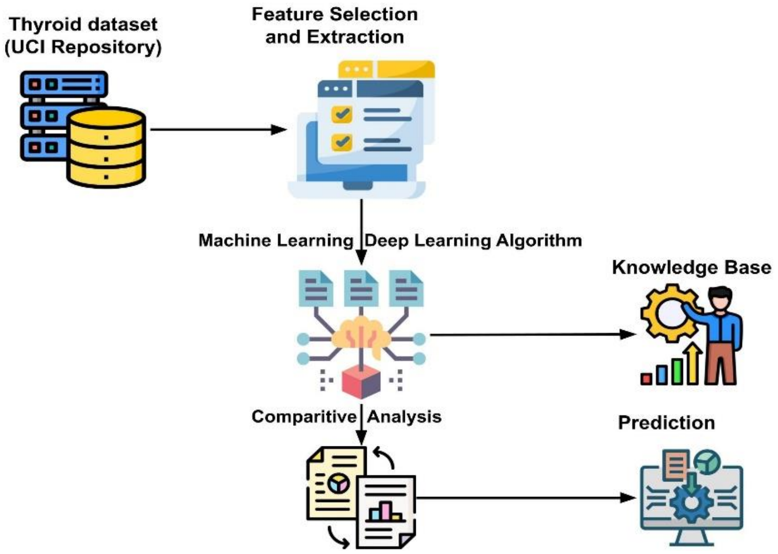

Figure 1. Architecture of thyroid prediction system

Various diseases such as eye boldness, cancer, heart problems, infertility, and PCOS can be caused by any thyroid problem. PCOS is a hormonal disorder diagnosed early mostly, so they will heal quickly and avoid any side effects [9]. Another thing to note is the overall increase in the prevalence of thyroid dysfunction worldwide, which is 8 times more common in women than in men [10]. Both sex and age play an important role in susceptibility to thyroid disease and the fact that risk increases with age makes this age group more susceptible to thyroid disease [11]. These individuals have an increased risk of thyroid cancer even if they have a prior medical history of thyroid conditions [12]. The field of thyroid disease detection has seen, among other things, promising developments in health informatics, especially in machine learning (ML) algorithms [13]. To improve the accuracy of the study, researchers used two learning techniques including deep-contractile neural networks (DCNN), and supervised learning to process medical images [14]. These enhancement techniques aim to reduce the potential for bias in the data and increase the accuracy of the diagnosis of thyroid disease [15]-[17]. Table 1 shows the features of thyroid tests.

Table 1. Thyroid test features

S. No.	Feature	Level
1	Free T3 (FT3)	2.2-4.0 pg/mL
2	T3 triiodothyronine tests	100-200 ng/dL
3	Free T4 (FT1)	0.9-1.7 ng/dL
4	Test T4 (FT4)	Low refers to hypothyroidism
5	Thyroid-stimulating hormone (TSH)	It is produced by the pituitary to regulate thyroid hormone (blood TSH levels determined by laboratory tests)

ML and DL models have been developed to accurately diagnose multidimensional and globally diverse thyroid problems. The use of subjective clinical symptoms and laboratory results in current diagnostic procedures can lead to errors or delays. Utilizing large patient databases with symptomatic and medical information, these systems aim to increase diagnostic accuracy. Issues of data imbalance, noise, and interpretability of the model require robust solutions. To increase acceptance and improve patient outcomes and health care efficiency through the use of sophisticated computerized thyroid disease diagnosis, general coverage is required in population and health care.

2. LITERATURE REVIEW

Tyagi *et al.* [18] suggests that ML predicted thyroid disease by focusing on the organ controlled by the thyroid gland. The two main types of thyroid problems are hyperthyroidism and hypothyroidism. Healthcare organization data has been processed previously, mostly ML for disease detection. Data were provided by the University of California Irvine repository. The study compares support vector machine (SVM) and K-nearest neighbor (KNN) algorithms for thyroid disease prediction.

Unnikrishnan and Menon [19] examined the prevalence of thyroid disorders in India, where an estimated 42 million people are affected. The report, which discusses five key conditions—namely hypothyroidism, hyperthyroidism, Goiter, Hashimoto's thyroiditis, and thyroid cancer-shows the regional variance brought about by iodine deficiency. Further, it discusses the efforts to achieve a standardized reference range of thyroid hormone for pregnant women and children as part of improving diagnosis and treatment.

Rehman *et al.* [20] evaluate KNN algorithms based on the detection of thyroid disease from the full KEEL repository dataset and hospital dataset from Pakistan. In their work, three KNN variations with other clinical features such as pulse rate, BMI, and blood pressure were tested to achieve the accuracy of detecting the disease. The findings indicate that the combination of these functions like Euclidean and Cosine distance with their selection chi-square based has enhanced the classification performance and therefore draw attention to feature selection importance during thyroid diseases prediction.

Chaganti *et al.* [21] presents a unique ML approach for thyroid disease prediction, which focuses on the early detection and classification of conditions such as cancer, hyperthyroidism, and hypothyroidism. The study uses various algorithms that include logistic regression (LR), random forest (RF), convolutional neural networks (CNNs), and SVM, evaluates performance using metrics such as accuracy, precision, recall, and F1 as well as scores that reflect the potential to improve study accuracy and patient outcomes.

This study [22] has proposed a ML algorithm to classify thyroid diseases. Given the challenges of using large health data, this article focused on ML methods. Using the data collected from Iraqis in the study, some tissues indicate hyperthyroidism, thyroid dysfunction, and hypothyroidism. ML techniques such as SVM, RF, KNN, Naïve Bayes (NB), and multilevel perceptron were used to identify thyroid groups. Comparative effectiveness studies were also conducted. However, the classification accuracy needs to be increased.

K-nearest algorithm for screening thyroid disorders [23]. This study uses real-time data from Pakistan hospital and KEEL dataset repository data set to test three selection methods: KNN feature selection, chi-square-based feature selection, and uncertain feature selection. It measures things like heart rate, BMI, and blood pressure. Among the tested methods, chi-square-based feature selection proved to be very accurate and successful, especially for recently added variables.

Thyroid disease prediction using different ML algorithms has been analyzed in [24], this work used a unique set of data from the UCI archive to test the accuracy of decision trees (DT) , KNN, and LR in predicting prediction if thyroid problems are, it exceeds DT and LR at least with an accuracy of 96%. A ML model was used by [25] to distinguish patients with hyperthyroidism from hypothyroidism. The results of this study are expressed as an accuracy chart: prediction accuracy for hypothyroidism was 90.9% and for hyperthyroidism was 93.8%.

The extreme gradient boosting (XGBoost) model was used by [26] to predict thyroid status in the knowledge discovery dataset from UC Irvin based on the observed model in China. Tiwari *et al.* [27] introduced an epidemiological model to predict the occurrence of COVID-19 cases in India. The diagnosis of thyroid disease was studied in [28]. Expert algorithms were used to analyze symptoms and make predictions about the disease [29]. Simplified swarm optimization (SSO) was used to analyze the function of the thyroid gland. The SSO was developed by the authors on a dataset obtained from the UCI [30].

Learning vector quantization techniques, radial basis functions, and backpropagation [31] used artificial neural networks (ANN) to predict thyroid disease. In contrast [32] ANN and SVM were tested for thyroid disease detection, in [33] along with diabetes and malignancy used similarity measures to identify patterns associated with thyroid problems. Several ML techniques were used to classify thyroid detection with high accuracy: logistic regression scored 98.7% in disease classification, and modified fuzzy hyper line segment clustering neural network (MFHLSCNN) showed improved clustering skills over multilayer perceptron (MLP) and linear discriminant analysis (LDA) hypothyroid disease with an accuracy score of 99.62% in prediction [34].

3. DATASET OF THYROID DISEASE

Datasets for thyroid disease prediction by ML and DL methods include data sets including textual image-based datasets from various sources. The groups provide a comprehensive view of thyroid disease, in which complex patterns and relationships that predict thyroid disease status occur. It helps to develop recognizable patterns [19], [21]. Table 2 shows sample dataset features and sample count.

Table 2. UCI thyroid dataset

Features	Sample count
31	9172

Features and structural abnormalities of the thyroid gland have been revealed in image-based data from research databases and medical libraries, such as computer tomography (CT) and magnetic resonance imaging (MRI), and ultrasound ML and DL models need such images to remove these to better identify the recognizable symptoms of thyroid diseases. Combining text-based and image-based data to create a comprehensive collection that accurately reflects the complexity of thyroid disease will help develop reliable predictive models. The UCI repository provides text-based data and ultrasound images for disorders of hyperthyroidism and hypothyroidism for these datasets of 9,172 samples each with 31 unique characteristics enable feature importance analysis for optimal classification of thyroid diseases [35], as shown in Table 3 [21], [36].

Table 3. Types of data sample attributes

S. No.	Attributes	Definition	Type
1	Sex	Sex patient identifies	(String)
2	Age	Patient age	(Integer)
3	Thyroxine	The patient is taking thyroxine	(Boolean)
4	Query on thyroxine	The patient is taking thyroxine	(Boolean)
5	On antithyroid meds	The patient takes antithyroid medications	(Boolean)
6	Sick	The patient is ill	(Boolean)
7	Pregnant	The expectant patient	(Boolean)
8	Query hypothyroid	The patient thinks they may have hypothyroidism.	(Boolean)
9	Query hyperthyroid	The patient thinks they may have hyperthyroid.	(Boolean)
10	TSH measured	Measured TSH in the blood	(Boolean)
11	T3 measured	Measured T3 in the blood	(Boolean)
12	TT4 measured	Measured TT4 in the blood	(Boolean)
13	T4U measured	Measured T4U in the blood	(Boolean)
14	FTI measured	Measured FTI in the blood	(Boolean)
15	TBG measured	Measured TBG in the blood	(Boolean)

Table 4 [21] presents an imbalanced class distribution for the thyroid health data set, with multiple samples not classified in any one class. As outlined in the method, feature selection and preprocessing were performed to balance the dataset for analysis. Patients without thyroid disease make up the majority of the “no condition” group, although severe patients with non-thyroid symptoms have had altered thyroid levels due to chronic conditions [10], [37], [38].

Table 4. Classification contained within the dataset

S. No.	Class	Definition	Details
1	-	Absence of condition	The patient has no normal thyroid disease/report.
2	A	Increased protein interaction	Protein binding facilitates mass assembly [39].
3	B	Under-replaced	Levothyroxine medications for hypothyroidism usually work well by replenishing the deficiency of thyroxine hormone [38].
4	C	Coexisting non-thyroid conditions	The term “non-thyroid disorder” refers to chronic diseases or hormonal changes that have nothing to do with pancreatic thyroid dysfunction [40].
5	D	Over-replaced	Reduce allocation errors by using low TSH and T4 abnormalities to identify patients with inadequate replacement and overreplacement to replace T4-TSH deficiency [9].
6	E	Discordant assay results	Early detection of test interruptions helps to avoid inappropriate patient care caused by abnormal thyroid function tests [41].
7	F	Hyperthyroid	Sleep issues, restlessness, tremors, skin and hair changes, palpitations, rapid heartbeat, and muscle weakness are all symptoms of hyperthyroidism [9], [42].
8	G	Primary hypothyroid	Damage to the thyroid gland, usually caused by the immune system or after radioiodine, radiation, or surgery, is the primary cause of hypothyroidism [42].
9	H	By replacement therapy	Thyroid hormone treatment, also known as hypothyroidism, is often recommended when the thyroid does not produce enough hormones. Sometimes it could also help control goiter growth [9], [43].
10	I	Under-replaced	Rare: used to assess the extent of thyroid goiter [43].
11	J	Compensated hypothyroid	In subclinical hypothyroidism, serum TSH levels are elevated, also known as compensatory hypothyroidism, while serum-free thyroxine (FT4) levels are normal [44].

Out of the 9,172 patient records in the dataset, 6,771 are classified as normal, indicating no thyroid condition. The new records included 346 patients with compensated hypothyroidism, 233 patients with primary hypothyroidism, 456 patients with non-thyroid disease, and 359 patients with compensated high-protein-binding hypothyroidism. Sample counts and classification categories for the data set are summarized in Table 5. For balance, 400 samples were randomly selected from 6,771 normal records, and new classes were switched to ensure that at least each class contained 200 samples. Table 5 provides information on the balanced data, while Table 6 provides sample information. The diagnosis of hypothyroidism is made based on laboratory blood tests, which require careful evaluation of basic parameters and hormone levels possible misdiagnosis and inappropriate treatment due to small statistical changes in thyroid hormone levels to increase the accuracy [21].

Table 5. Balanced dataset for the classification of thyroid diseases

S. No.	Class	Count after preprocessing	Final
1	Normal	6,771	400
2	Increased binding protein	346	346
3	Present non-thyroidal ailment	436	436
4	Primary hypothyroid	233	233
5	Compensated hypothyroid	359	359

Table 6. Count of images and diseases

S. No.	Classification of a disease	Number of images
1	Normal thyroid	88
2	Hyperthyroid	77
3	Hypothyroid	110
4	Thyroid nodule	146
5	Thyroiditis	74
6	Thyroid cancer	99

3.1. Data collection

Many academic databases like Kaggle, UCI, PubMed, IEEE Xplore, and Google Scholar were searched to assemble materials for this survey article. The data sets were acquired from reputable open data portals, government databases, and institutional repositories emphasizing the diversity of subject categories, historical periods, and geographical scope of the data. The selection criteria adopted in ensuring the inclusion of relevant and high-quality datasets that would enable a comprehensive understanding of the subject matter were quite stringent [1], [2], [19], [23].

3.2. Data preprocessing

Preparing unprocessed data for use in a ML model is known as data preprocessing. When running a ML project, it is common to encounter very dirty or unstructured raw data. Real-world data often contain noise and missing values, making them unsuitable for direct incorporation into ML models. Consequently, data preprocessing is essential for data cleaning and transformation, qualifying ML models for their maximum accuracy and efficiency [45].

4. ALGORITHMS IN DEEP LEARNING AND MACHINE LEARNING

Algorithms for computer learning and DL are needed, eliminating the need for explicit design. These algorithms can use hierarchical mental models to interpret design information [46]. DL techniques based on neural networks improve data classification and prediction accuracy, which is crucial for reliable disease diagnosis in modern medicine. The health service produces a lot of information in addition to medical examinations and patient records. Medical data sets are systematically analyzed using ML techniques, which offer analytical methods [47].

Accurate modeling and disease model classification are achieved through DL algorithms, which facilitate research objectives and treatment hypotheses ML enables automated data analysis and prediction, and DL provides improved it has data acquisition, storage, and computing power [48]. Since unstructured data makes it difficult to reveal important risk detection characteristics, DL makes feature engineering work automatically by going through multiple layers, thus highlighting the importance of the user's involvement. Figure 2 shows a comparison between ML and DL architectures. The entire figure shows the difference between a simple and a deep neural network. Figure 2(a) shows a simple neural network used by ML where feature extraction is separate from classification. Figure 3(b) shows a deep neural network that integrates

feature extraction and classification to create a highly automated approach. Figures 2(a) and 2(b) show the three classifications of layers in deep neural networks as follows:
i) The function of the input layer is to store input data.
ii) The number of hidden layers is variable, but its functionality is critical to the network.
iii) Output layer, which handles data classification and prediction.

Several input, output, and hidden layers that make up the DL process are shown. The more sophisticated characteristics that each layer acquires enable deeper insights to be gained from the data. They increase the layer depth and thicken the materials. On the other hand, Figure 3 summarizes the evolution of AI and is divided into Figure 3(a): performance timeline of ML (1980) and DL (2006). Figure 3(b): hierarchical representation (AI → ML → DL → CNN). Figures 3(a)-3(c): effect of data volume on ML/DL performance. (Figures 3(a) to 3(c)) DL outperforms classical ML in processing large data sets and greater accuracy while generating more sophisticated neural networks with characteristics collected from human data uses, which limits scalability and speed [49].

While DL algorithms have undoubtedly improved performance and attracted interest in predictive applications in many industries, including healthcare and medicine, ML needs to support performance and accuracy standards same rigor. The results of the study highlight the importance of ML, because of its consistent ability to produce more accurate and precise results [47].

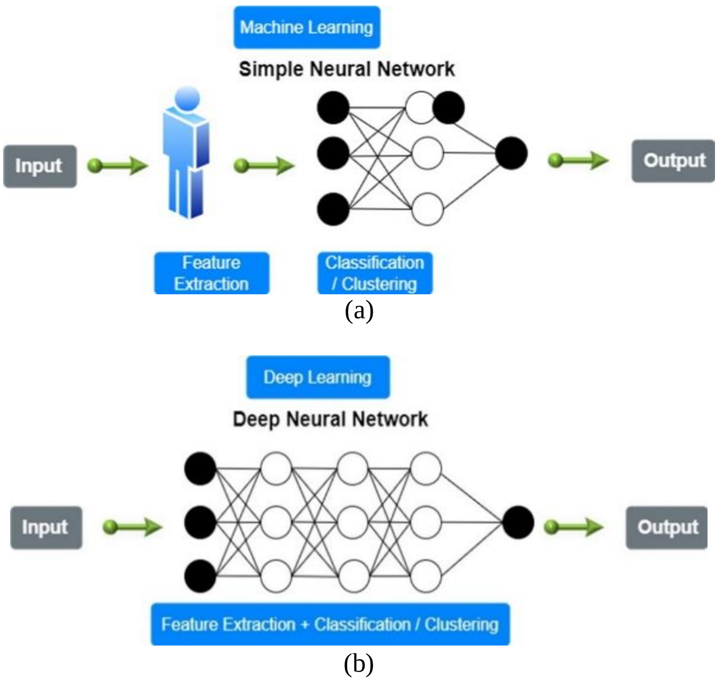

Figure 2. ML and DL (a) different layers and work process ML (simple neural network) and (b) different layers and work process DL (deep neural network)

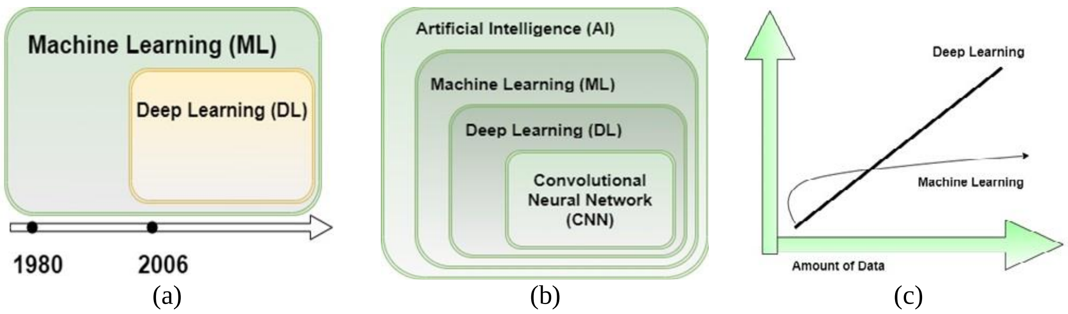

Figure 3. Evolution of AI: (a) performance timeline of ML (1980) and DL, (b) hierarchical representation (AI → ML → DL → CNN), and (c) effect of data volume on ML/DL performance

An overview of ML methods, association algorithms, and widely used DL algorithms and architectures are discussed in Tables 7 and 8 [21]. DL models are an integrative subset of ML. This example involves data processing through multilevel networks, where the results of the previous level are used as signals for the results of the next level [50]. The availability of large label datasets, more powerful parallel computing hardware (e.g., graphics processing units), and faster algorithms have all contributed to the rapid implementation of DL [51]. DL techniques can solve the challenges that the AI community has been dealing with for years. This is because using a significant number of layers improves the quality of the universal approximation and facilitates the identification of connected elements from data at different levels of hierarchy and abstraction [52]. In radiology, CNNs are considered highly sophisticated DL algorithms, which are used for computer vision tasks including detection, classification, and segmentation [53], [54]. CNN relies on 4 middle concepts-local connections, shared weights, pooling, and the incorporation of multiple layers-all of which drastically beautify the precision and effectiveness of the general machine [52]. The classification task or prediction task uses fully connected layers with iterative operation and pooling [54], [55]. Several layer structures are available, and several deep neural network architectures such as GoogleNet [56], AlexNet [57], VGGNet [58], and ResNet [59] have been successfully used in image analysis.

Table 7. Algorithms

S. No.	Algorithms	Features
		Supervised learning
1	LR	Regression is used in the method, where the feature is considered as the independent variable and the class as the dependent variable. A linear classifier does not perform well on nonlinearly separable data.
2	DT	The method achieves objective values through a flowchart-like tree model; This model works well for segmentation but is not suitable for regression or price trend forecasting.
3	NB classifier	The algorithm based on Bayes theorem is a simple probabilistic classification technique that assumes feature independence and uses maximum likelihood estimation for reliable results with small data sets.
4	KNN classifier	Useful for classification and regression, but computationally demanding, this method acts as an instance-based classifier, classifying unknown data according to its similarity to known data.
5	SVM	The algorithm divides the input data into two groups based on several factors using a multidimensional hyperplane using statistical methods, which is suitable for regression and classification applications.
6	RF	The approach that combines decision trees and cluster classification to aggregate findings by averaging the resulting trees performs well but runs the risk of overfitting when dealing with chaotic data.
7	ANN	This method has input, hidden, and output layers. To classify new data with a learning algorithm, the hidden feature mimics neurons and requires a large amount of annotations. This method is effective for nonlinear relationships.
		Unsupervised learning
8	K-means	For large data sets with a predefined number of clusters, the method computes the distance between data points and centroids for classification, comparing between-cluster and inter-cluster differences to optimize the classification.
9	Clustering methods	The algorithm aims to efficiently search clusters or variants for classification, thereby reducing the amount of data required.

Table 8. Deep learning architectures

S. No.	Architecture	Definition
1	AlexNet	One of the basic schemes of high-performance distribution is distinguished by the use of dropouts, data enhancement, and rectified linear unit (ReLU) activation functions.
2	GoogleNet	Develop a basic design that integrates the results of operations applied to the input data; specializing in the distribution of images.
3	VGGNet	It uses a small amount of depth and small pore particles.
4	ResNet	Using skip structures to combine input and processed data, helps the network gain knowledge of residues and complexity.
5	R-CNN	DL methods for analyzing search tasks using bivariate networks with a well-designed model trained for the classification task.
6	YOLO	A flexible and fast one-level interface capable of facilitating real-time discovery and segmentation.

5. RESULTS

A study comparing ML algorithms on thyroid disease prediction evaluated the performance of several ML algorithms [34]. The authors pooled data on 106 patients using DTs, RFs, and LR, among other methods. The RF was more accurate in predicting thyroid disease at 94.7% sensitivity and 94.5% specificity.

"In silico models for predicting interference of molecular origin events associated with thyroid hormone homeostasis" [60]. [61]. A variety of ML methods including RFs, SVM, and neural networks were applied to the training data sets along with traditional statistical methods [62], [63]. Three data balance methods were used to connect the ToxCast database with the scholar's research boom. The models were tested with independent molecular-based data sets. Performance was improved by using multiple neurons, especially for low-input targets. The averages had robust F1 values: 0.81 for thyroid peroxidase inhibition (TPO) and 0.83 for peroxisome proliferator-activated receptor (TR) activation. After parameter changes, the other models also performed well, reaching F1 values of up to 0.77. Recent surveys on the thyroid are shown in Table 9.

Table 9. Recent survey on thyroid disease [9]

Ref.	Sample size	Source of dataset	Models	Class	Metrics	Results
[9]	6,356 samples	Image dataset	CNN-based ResNet and stochastic gradient descent (SGD)	5	Accuracy	Accuracy 94%
[20]	690 samples, 13 attributes	Pakistan district headquarters teaching hospital and KEEL repository datasets	KNN with and without chi-square-based feature selection	3	Accuracy	KNN 98%
[21]	9,172 samples, 31 attributes	UCI cleveland	RF, AdaBoost, GradBoost, long short-term memory (LSTM)	4	F1-score, recall, precision	RF- 99% accuracy and CNN -94% precision
[22]	1,250 samples, 17 attributes	Hospital and lab	SVM, RF, DT, NB, LR, KNN, MLP, LDA	3	Accuracy	DT 90.13%, SVM 92.53%, RF 91.2%, NB 90.67%, LR 91.73%, LDA 83.2% KNN 91.47%, MLP 96.4%
[27]	215 samples with 5 attributes	UCI cleveland	KNN, XGBoost, LR, DT	3	Accuracy	KNN 81.25% and XGBoost 87.5%
[34]	7,200 samples, 21 attributes	UCI cleveland	SVM, MLR, NB, and DT	2	Accuracy	MLR 91.59%, SVM 96.04%, NB 6.31%, DT 99.23%, Sensitivity (94.7%),
[64]	3,162 attributes	UCI cleveland	Extra-trees, CatBoost, light GradBoost, ANN, KNN, RF, DT	4	Accuracy	Accuracy-95.7%
[9], [61]	N/A	ToxCasts	RF, LR, XGB, SVM, ANN	2	F1-score	(TR) RF-81% and (TPO) XGB-83%
[9], [61]	7,247 samples, 21 attributes	UCI cleveland	Grey wolf optimization (GWO), improved GWO (IGWO), hybrid firefly butterfly optimization (HFBO)	3	Accuracy, specificity, sensitivity	Sensitivity (99.2%), accuracy (99.28%), specificity (94.5%)

6 CONCLUSION AND FUTURE RESEARCH DIRECTIONS

The future of thyroid marker detection and prediction seems more promising when robots can learn ML and disseminate this information DL, for example, have been demonstrated that in combination with other data, such as photographs and medical records, can increase model accuracy. Advanced DL models such as CNN and RNN have improved the evaluation of personalized medicine. However, many common issues, such as sample imbalances, unclear models, generalizations across different populations or ethnic groups, and many models that are currently obscure, constrain practitioners to have little confidence in hospitals. For example, there has been little research on the prognosis of common thyroid comorbidities in people with thyroid disease, including PCOS and TED. Thus, to overcome these obstacles, research should be conducted on ways to improve model interpretation capabilities; validation in different population mixtures; a system that integrates multiple data sets is essential for holistic forecasting; and, at the same time, they need to ensure that the ability to detect and predict other thyroid disorders is more effective. To implement ML and DL in the healthcare industry, interdisciplinary teams need to work together to exchange information and develop user-friendly solutions. To further improve the diagnosis, treatment, and management of thyroid disease, these deficiencies need to be corrected for both ML and DL and their associated comorbidities. This will dramatically improve patient outcomes and healthcare efficiency.

ACKNOWLEDGEMENTS

We would like to thank our institution, Vellore Institute of Technology, Vellore for their invaluable support during the writing of this article.

FUNDING INFORMATION

The authors declare that no funding was received for this review work.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mohammed Yacoob B. A.	✓	✓				✓		✓	✓		✓			
Jayashree J.				✓	✓					✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY

The data used in this study were obtained from publicly available sources, including the UCI Machine Learning Repository and published literature. Any additional datasets supporting this study are available upon reasonable request to the corresponding author.

REFERENCES

[1] C. Rajeswari, B. Sathiyabhama, S. Devendiran, and K. Manivannan, “Bearing fault diagnosis using wavelet packet transform, hybrid PSO and support vector machine,” *Procedia Engineering*, vol. 97, pp. 1772–1783, 2014, doi: 10.1016/j.proeng.2014.12.329.

[2] M. Hosseinzadeh *et al.*, “A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things,” *Journal of Supercomputing*, vol. 77, no. 4, pp. 3616–3637, Apr. 2021, doi: 10.1007/s11227-020-03404-w.

[3] S. Keestra, V. H. Tabor, and A. Alvergne, “Reinterpreting patterns of variation in human thyroid function,” *Evolution, Medicine and Public Health*, vol. 9, no. 1, pp. 93–112, Feb. 2021, doi: 10.1093/emph/eoaa043.

[4] H. E. Yazdaan *et al.*, “Advances in thyroid function tests: precision diagnostics and clinical implications,” *Cureus*, Nov. 2023, doi: 10.7759/cureus.48961.

[5] M. Bauer, T. Glenn, M. Pilhatsch, A. Pfennig, and P. C. Whybrow, “Gender differences in thyroid system function: relevance to bipolar disorder and its treatment,” *Bipolar Disorders*, vol. 16, no. 1, pp. 58–71, Feb. 2014, doi: 10.1111/bdi.12150.

[6] H. Stre, “Why women are more prone to thyroid problems? | Harley Street ENT Clinic.” https://www.harleystreetent.com/blog/why-women-are-more-prone-to-thyroid-problems.

[7] “Thyroid eye disease,” *Cleveland Clinic*, 2024. https://my.clevelandclinic.org/health/diseases/17558-thyroid-eye-disease (accessed Jun. 28, 2024).

[8] G. Zheng, H. Zhang, and X. Yao, “Research and application of artificial neural network prediction model in deep learning,” in *International Symposium on Robotics, Artificial Intelligence, and Information Engineering (RAIIE 2022)*, Nov. 2022, p. 97, doi: 10.1117/12.2659276.

[9] P. Gupta *et al.*, “Detecting thyroid disease using optimized machine learning model based on differential evolution,” *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 3, Jan. 2024, doi: 10.1007/s44196-023-00388-2.

[10] M. S. Devi, V. D. Kumar, A. Brezulianu, O. Geman, and M. Arif, “A novel blunge calibration intelligent feature classification model for the prediction of hypothyroid disease,” *Sensors*, vol. 23, no. 3, p. 1128, Jan. 2023, doi: 10.3390/s23031128.

[11] J. P. Walsh, “Thyroid function across the lifespan: do age-related changes matter?,” *Endocrinology and Metabolism*, vol. 37, no. 2, pp. 208–219, Apr. 2022, doi: 10.3803/EnM.2022.1463.

[12] X. Zhang, V. C. S. Lee, and J. C. Lee, “Unveiling thyroid disease associations: an exceptionality-based data mining technique,” *Endocrines*, vol. 4, no. 3, pp. 558–572, Jul. 2023, doi: 10.3390/endocrines4030040.

[13] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, “Comparing different supervised machine learning algorithms for disease prediction,” *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, p. 281, Dec. 2019, doi: 10.1186/s12911-019-1004-8.

[14] D. S. Saleh and M. S. Othman, “Exploring the challenges of diagnosing thyroid disease with imbalanced data and machine learning: a systematic literature review,” *Baghdad Science Journal*, vol. 21, no. 3, pp. 1119–1136, Mar. 2024, doi: 10.21123/bsj.2023.8544.

[15] A. Sultana and R. Islam, “Machine learning framework with feature selection approaches for thyroid disease classification and associated risk factors identification,” *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, p. 32, Jun. 2023, doi: 10.1186/s43067-023-00101-5.

[16] H. Zhao *et al.*, "Diagnosis of thyroid disease using deep convolutional neural network models applied to thyroid scintigraphy images: a multicenter study," *Frontiers in Endocrinology*, vol. 14, Aug. 2023, doi: 10.3389/fendo.2023.1224191.
[17] A. Professor and A. J. SandalinD, "A comparitive study of thyroid disease detection using machine learning-a paper review," *European Chemical Bulletin*, vol. 12, no. 10, pp. 2427–2432, 2023, doi: 10.48047/ecb/2023.12.10.167.
[18] A. Tyagi, R. Mehra, and A. Saxena, "Interactive thyroid disease prediction system using machine learning technique," in *PDGC 2018 - 2018 5th International Conference on Parallel, Distributed and Grid Computing*, Dec. 2018, pp. 689–693, doi: 10.1109/PDGC.2018.8745910.
[19] A. Unnikrishnan and U. Menon, "Thyroid disorders in India: an epidemiological perspective," *Indian Journal of Endocrinology and Metabolism*, vol. 15, no. 6, p. 78, 2011, doi: 10.4103/2230-8210.83329.
[20] H. A. Ur Rehman, C. Y. Lin, and Z. Mushtaq, "Effective K-nearest neighbor algorithms performance analysis of thyroid disease," *Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers,Series A*, vol. 44, no. 1, pp. 77–87, Jan. 2021, doi: 10.1080/02533839.2020.1831967.
[21] R. Chaganti, F. Rustam, I. De La Torre Díez, J. L. V. Mazón, C. L. Rodríguez, and I. Ashraf, "Thyroid disease prediction using selective features and machine learning techniques," *Cancers*, vol. 14, no. 16, p. 3914, Aug. 2022, doi: 10.3390/cancers14163914.
[22] K. Salman and E. Sonuc, "Thyroid disease classification using machine learning algorithms," *Journal of Physics: Conference Series*, vol. 1963, no. 1, p. 012140, Jul. 2021, doi: 10.1088/1742-6596/1963/1/012140.
[23] H. A. Ur Rehman, C. Y. Lin, Z. Mushtaq, and S. F. Su, "Performance analysis of machine learning algorithms for thyroid disease," *Arabian Journal for Science and Engineering*, vol. 46, no. 10, pp. 9437–9449, 2021, doi: 10.1007/s13369-020-05206-x.
[24] G. Chaubey, D. Bisen, S. Arjaria, and V. Yadav, "Thyroid disease prediction using machine learning approaches," *National Academy Science Letters*, vol. 44, no. 3, pp. 233–238, Jun. 2021, doi: 10.1007/s40009-020-00979-z.
[25] M. Hu *et al.*, "Development and preliminary validation of a machine learning system for thyroid dysfunction diagnosis based on routine laboratory tests," *Communications Medicine*, vol. 2, no. 1, p. 9, Jan. 2022, doi: 10.1038/s43856-022-00071-1.
[26] S. Sankar, A. Potti, G. N. Chandrika, and S. Ramasubbareddy, "Thyroid disease prediction using xgboost algorithms," *Journal of Mobile Multimedia*, vol. 18, no. 3, pp. 917–934, Feb. 2022, doi: 10.13052/jmm1550-4646.18322.
[27] S. Tiwari, S. Kumar, and K. Guleria, "Outbreak trends of coronavirus disease-2019 in India: a prediction," *Disaster Medicine and Public Health Preparedness*, vol. 14, no. 5, pp. e33–e38, Oct. 2020, doi: 10.1017/dmp.2020.115.
[28] A. Keleş and A. Keleş, "ESTDD: expert system for thyroid diseases diagnosis," *Expert Systems with Applications*, vol. 34, no. 1, pp. 242–246, Jan. 2008, doi: 10.1016/j.eswa.2006.09.028.
[29] W. C. Yeh, "Novel swarm optimization for mining classification rules on thyroid gland data," *Information Sciences*, vol. 197, pp. 65–76, Aug. 2012, doi: 10.1016/j.ins.2012.02.009.
[30] S. D. Bay, D. Kibler, M. J. Pazzani, and P. Smyth, "The UCI KDD archive of large data sets for data mining research and experimentation," *ACM SIGKDD Explorations Newsletter*, vol. 2, no. 2, pp. 81–85, Dec. 2000, doi: 10.1145/380995.381030.
[31] A. Shukla, R. Tiwari, P. Kaur, and R. R. Janghel, "Diagnosis of thyroid disorders using artificial neural networks," in *2009 IEEE International Advance Computing Conference, IACC 2009*, Mar. 2009, pp. 1016–1020, doi: 10.1109/IADCC.2009.4809154.
[32] S. Shariati and M. M. Haghighi, "Comparison of ANFIS neural network with several other ANNs and support vector machine for diagnosing hepatitis and thyroid diseases," in *2010 International Conference on Computer Information Systems and Industrial Management Applications, CISIM 2010*, Oct. 2010, pp. 596–599, doi: 10.1109/CISIM.2010.5643520.
[33] K. Saastamoinen and J. Ketola, "Medical data classification using logical similarity based measures," in *2006 IEEE Conference on Cybernetics and Intelligent Systems*, Jun. 2006, pp. 1–5, doi: 10.1109/ICCIS.2006.252362.
[34] S. Razia, P. S. Prathyusha, N. V. Krishna, and N. S. Sumana, "A comparative study of machine learning algorithms on thyroid disease prediction," *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 2.8, pp. 315–319, Mar. 2018, doi: 10.14419/ijet.v7i2.8.10432.
[35] G. Mollica *et al.*, "Classification of thyroid diseases using machine learning and bayesian graph algorithms," *IFAC-PapersOnLine*, vol. 55, no. 40, pp. 67–72, 2022, doi: 10.1016/j.ifacol.2023.01.050.
[36] V. R. E. Christo, H. K. Nehemiah, S. K. Sankari, S. Jeyaraj, and A. Kannan, "Classification framework for clinical datasets using synergistic firefly optimization," *IETE Journal of Research*, vol. 69, no. 10, pp. 7051–7070, Oct. 2023, doi: 10.1080/03772063.2021.2007798.
[37] S. M. Wajner and A. L. Maia, "New insights toward the acute non-thyroidal illness syndrome," *Frontiers in Endocrinology*, vol. 3, no. JAN, 2012, doi: 10.3389/fendo.2012.00008.
[38] J. Jonklaas *et al.*, "Guidelines for the treatment of hypothyroidism: prepared by the American thyroid association task force on thyroid hormone replacement," *Thyroid*, vol. 24, no. 12, pp. 1670–1751, Dec. 2014, doi: 10.1089/thy.2014.0028.
[39] P. Bielli, R. Busà, M. P. Paronetto, and C. Sette, "The RNA-binding protein Sam68 is a multifunctional player in human cancer," *Endocrine-Related Cancer*, vol. 18, no. 4, pp. R91–R102, Aug. 2011, doi: 10.1530/ERC-11-0041.
[40] Z. Yasar, C. Kirakli, P. Cimen, Z. Z. Ucar, F. Talay, and G. Tibet, "Is non-thyroidal illness syndrome a predictor for prolonged weaning in intubated chronic obstructive pulmonary disease patients?," *International Journal of Clinical and Experimental Medicine*, vol. 8, no. 6, pp. 10114–10121, 2015.
[41] P. Srichomkwun, N. H. Scherberg, J. Jakšić, and S. Refetoff, "Diagnostic dilemma in discordant thyroid function tests due to thyroid hormone autoantibodies," *AACE Clinical Case Reports*, vol. 3, no. 1, pp. e22–e25, 2017, doi: 10.4158/EP151142.CR.
[42] D. S. Ross *et al.*, "2016 American thyroid association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis," *Thyroid*, vol. 26, no. 10, pp. 1343–1421, Oct. 2016, doi: 10.1089/thy.2016.0229.
[43] B. Biondi and D. S. Cooper, "Thyroid hormone therapy for hypothyroidism," *Endocrine*, vol. 66, no. 1, pp. 18–26, Oct. 2019, doi: 10.1007/s12020-019-02023-7.
[44] M. Salerno, N. Improda, and D. Capalbo, "Management of endocrine disease subclinical hypothyroidism in children," *European Journal of Endocrinology*, vol. 183, no. 2, pp. R13–R28, Aug. 2020, doi: 10.1530/EJE-20-0051.
[45] A. Parashar, A. Parashar, W. Ding, M. Shabaz, and I. Rida, "Data preprocessing and feature selection techniques in gait recognition: a comparative study of machine learning and deep learning approaches," *Pattern Recognition Letters*, vol. 172, pp. 65–73, Aug. 2023, doi: 10.1016/j.patrec.2023.05.021.
[46] K. Chelladurai and N. Sujatha, "A survey on different algorithms used in deep learning process," *E3S Web of Conferences*, vol. 387, p. 05008, May 2023, doi: 10.1051/e3sconf/202338705008.
[47] H. Abdel-Jaber, D. Devassy, A. Al Salam, L. Hidaytallah, and M. El-Amir, "A review of deep learning algorithms and their applications in healthcare," *Algorithms*, vol. 15, no. 2, p. 71, Feb. 2022, doi: 10.3390/a15020071.
[48] M. M. Taye, "Understanding of machine learning with deep learning: architectures, workflow, applications and future directions," *Computers*, vol. 12, no. 5, p. 91, Apr. 2023, doi: 10.3390/computers12050091.

[49] J. Latif, C. Xiao, A. Imran, and S. Tu, "Medical imaging using machine learning and deep learning algorithms: a review," in *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies, iCoMET 2019*, Jan. 2019, pp. 1–5, doi: 10.1109/ICOMET.2019.8673502.
[50] I. H. Sarker, "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, p. 420, Nov. 2021, doi: 10.1007/s42979-021-00815-1.
[51] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
[52] M. I. Tariq *et al.*, "A review of deep learning security and privacy defensive techniques," *Mobile Information Systems*, vol. 2020, pp. 1–18, Apr. 2020, doi: 10.1155/2020/6535834.
[53] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/s13244-018-0639-9.
[54] A. I. Khan, J. L. Shah, and M. M. Bhat, "CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest x-ray images," *Computer Methods and Programs in Biomedicine*, vol. 196, p. 105581, Nov. 2020, doi: 10.1016/j.cmpb.2020.105581.
[55] B. J. Erickson, P. Korfiatis, T. L. Kline, Z. Akkus, K. Philbrick, and A. D. Weston, "Deep learning in radiology: does one size fit all?," *Journal of the American College of Radiology*, vol. 15, no. 3, pp. 521–526, Mar. 2018, doi: 10.1016/j.jacr.2017.12.027.
[56] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2015, vol. 07-12-June-2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.
[57] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
[58] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.
[59] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2016, vol. 2016-December, pp. 770–778, doi: 10.1109/CVPR.2016.90.
[60] M. Garcia De Lomana *et al.*, "In silico models to predict the perturbation of molecular initiating events related to thyroid hormone homeostasis," *Chemical Research in Toxicology*, vol. 34, no. 2, pp. 396–411, Feb. 2021, doi: 10.1021/acs.chemrestox.0c00304.
[61] V. Sureshkumar, S. Balasubramaniam, V. Ravi, and A. Arunachalam, "A hybrid optimization algorithm-based feature selection for thyroid disease classifier with rough type-2 fuzzy support vector machine," *Expert Systems*, vol. 39, no. 1, Jan. 2022, doi: 10.1111/exsy.12811.
[62] J. Vijayashree and H. P. Sultana, "A machine learning framework for feature selection in heart disease classification using improved particle swarm optimization with support vector machine classifier," *Programming and Computer Software*, vol. 44, no. 6, pp. 388–397, 2018, doi: 10.1134/S0361768818060129.
[63] J. Vijayashree and H. P. Sultana, "Heart disease classification using hybridized Ruzzo-Tompa memetic based deep trained Neocognitron neural network," *Health and Technology*, vol. 10, no. 1, pp. 207–216, Jan. 2020, doi: 10.1007/s12553-018-00292-2.
[64] S. S. Islam, M. S. Haque, M. S. U. Miah, T. Bin Sarwar, and R. Nugraha, "Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study," *PeerJ Computer Science*, vol. 8, p. e898, Mar. 2022, doi: 10.7717/PEERJ-CS.898.

BIOGRAPHIES OF AUTHORS

Mohammed Yacoob B. A. was born in Kanchipuram. Tamil Nadu, India. Received master's degree in computer applications from SCSVMV University, Kanchipuram, Tamil Nadu, in 2010. He is currently pursuing Ph.D. degree with Vellore Institute of Technology, Vellore, Tamil Nadu. After obtaining the master's degree, he began teaching with the MCA Department, C Abdul Hakeem College of Engineering and Technology, as an assistant professor. He taught students about machine and deep learning and assisted them with their research projects, further deepening his research interest, particularly in medicine, machine learning, and deep learning. He is dedicated to finding innovative solutions to real-world problems using advanced technology. His research interests include enhancing medical-related research, especially endocrine disease comorbidities. He can be contacted at email: mohammedyacoob.ba2023@vitstudent.ac.in.

Dr. Jayashree J. associate professor, School of Computer Science and Engineering at VIT University, Vellore, Tamil Nadu, India. She had much teaching and research experience with a good number of publications in reputed international journals and conferences. Research interests include IoT, machine learning, deep learning, and blockchain. She chaired many international conferences delivered keynote lectures and editorial Board member for many journals. She can be contacted at email: jayashree.j@vit.ac.in.