Prostate Cancer Detection and Analysis using Advanced Machine Learning

Mowafaq Salem Alzboon, Mohammad Subhi Al-Batah
Faculty of Science and Information Technology, Jadara University, Irbid, Jordan

***Abstract*—Prostate cancer is one of the leading causes of cancer-related deaths among men. Early detection of prostate cancer is essential in improving the survival rate of patients. This study aimed to develop a machine-learning model for detecting and diagnosing prostate cancer using clinical and radiological data. The dataset consists of 200 patients with prostate cancer and 200 healthy controls and extracted features from their clinical and radiological data. Then, the data trained and evaluated using several machines learning models, including logistic Regression, decision tree, random forest, support vector machine, and neural network models, using 10-fold cross-validation. Our results show that the random forest model achieved the highest accuracy of 0.92, with a sensitivity of 0.95 and a specificity of 0.89. The decision tree model achieved a nearly similar accuracy of 0.91, while the logistic regression, support vector machine, and neural network models achieved lower accuracies of 0.86, 0.87, and 0.88, respectively. Our findings suggest that machine learning models can effectively detect and diagnose prostate cancer using clinical and radiological data. The random forest model may be the most suitable model for this task.**

***Keywords*—Prostate cancer; machine learning; clinical data; radiological data; diagnosis; medical diagnosis**

I. Introduction

The use of machine learning techniques to study cancer has produced encouraging results, with the promise of more precise and time-saving approaches for identifying malignant cells and forecasting patient outcomes [1]. Machine learning algorithms can analyze data in search of patterns and attributes characteristic of malignant cells or tumours and then indicate how likely the disease will spread or reoccur [2, 3, 4]. Machine learning algorithms can identify small changes in tissue structure that may suggest the presence of malignant cells. They can forecast the likelihood of cancer progression and offer viable treatment choices by examining the genetic alterations in a patient's cancer cells. Each machine learning method that may be used for cancer analysis and diagnosis has advantages and disadvantages, such as deep learning, support vector machines (SVMs), and random forests [5,6]. Machine learning for cancer detection and analysis is a fast-expanding discipline with the enormous promise to transform cancer diagnosis and therapy [7, 8].

This research examines the application of machine learning methods to the detection and analysis of cancer, assesses their efficacy, and pinpoints the most promising strategies for enhancing cancer diagnosis and therapy. Large amounts of patient data (such as medical pictures, genomics data, and clinical records) can be analyzed using machine learning algorithms to look for patterns and traits characteristic of malignant cells or tumours. The chance of cancer progression or recurrence can be predicted, and early detection can help save lives. Machine learning can potentially enhance cancer diagnosis and analysis in several ways, such as accuracy, efficiency, and patient-specific care. However, it has drawbacks, such as the need for massive amounts of high-quality data to train machine learning algorithms. Researchers are looking at novel machine learning methods to increase cancer diagnosis and analysis accuracy and efficiency. Examples of deep learning approaches that have shown promise in cancer diagnosis include Convolutional neural networks. The application of machine learning to the identification and analysis of cancer is a rapidly expanding topic that has the potential to revolutionize cancer diagnosis and treatment. Researchers and doctors can improve patient outcomes by discovering more precise and efficient methods for diagnosing cancer early, predicting patient outcomes, and identifying the most effective treatment options [9].

II. Literature Review

Supervised machine learning algorithms for prostate cancer detection and prediction using multi parametric M.R. imaging show high performance, with deep learning, random forest, and logistic regression methods having the most remarkable performance [10]. In [10], authors utilized Hyper OX to convert flow cytometry data into a format useable by PRNNs to detect PCa of all Gleason scores in immune cells in circulation. Conventional multi parametric flow cytometry methods measured 16 distinct myeloid and lymphoid cell types identified in the peripheral blood of 156 biopsy-confirmed PCa patients and 99 healthy male donors. Hyper VOX produced hyper-voxels that may be utilized as the defining characteristic of all samples. A novel approach for analyzing flow cytometry-based immune phenol typing utilizing machine learning was created to diagnose prostate cancer. Using raw flow cytometry data from 97 PCa patients and 67 H.D. controls, PRNNs were trained. Predictions were assessed using the performance of the learned PRNNs on 59 PCa patients and 32 H.D. that were not utilized for PRNN training. The PRNN accurately categorized 28 of 32 H.D. samples and 57 of 59 PCa samples, yielding a sensitivity of 96.6 percent, a specificity of 87.5 percent, a positive predictive value of 93.4 percent, a negative predictive value of 93.3 percent, and an area under the curve (AUC) of 0.9656 [11]. This research investigates the viability of employing the Semantic Learning Machine (SLM) neuroevolution algorithm to replace the typically utilized fully connected architecture in the final layers of Convlutional Neural Networks (CNNs). The

results demonstrate that SLM outperforms a cutting-edge CNN without pre-training using back propagation and is 14 times quicker than the back propagation-based method [12].

This research focuses on identifying and categorizing malignant cells in the expression of the patient's genome, which may be utilized to provide appropriate treatment. Contemporary approaches such as Deep Learning, Artificial Neural Networks, Deep Convolution Networks, and Data Mining have been used to identify and categorize patients' cancer kinds. Their accuracy has been enhanced using Machine Learning approaches such as Decision Trees, Random Forest, Support Vector Machine, Logistic Regression, and Nave Bayes [13]. Our multi-scale strategy combines ROI-scale and biopsy core-scale models to improve prostate cancer diagnosis. Our approach obtains an AUROC of 80.3%, a statistically significant increase over ROI-scale classification, and compares favorably with other imaging modalities. Our source code is accessible to the public at www.github.com/med-i-lab/TRUSFormer [14]. Using many medical imaging modalities, A.I. approaches can assist in identifying and diagnosing prostate cancer. This review comprises 69 investigations from 1441 publications, most of which employ Convlutional neural networks and conventional machine learning techniques. Tools based on A.I. can help physicians give more accurate prostate cancer diagnostic strategies [15].

Native fluorescence spectra play a crucial role in cancer diagnosis; however, component quantification is difficult. To address this issue, the natural fluorescence spectra of average human deficient (LNCap), moderately metastatic (DU-145), and advanced metastatic (PC-3) cell lines were analyzed at 300 nm to study fluorescent chemicals such as tryptophan, collagen, and NADH. Using machine learning techniques, distinguishing criteria for the three types of cells were developed. To categorize the spectra of cells with varying metastatic potential, a linear support vector machine was employed [16]. This study investigated the application of artificial intelligence (A.I.) and machine learning (ML) techniques in oncological urology. Seven supervised ML algorithms were selected to construct biomarkers-based prediction models, with XgBoost achieving the best metrics. Results demonstrated that the ML technique was practicable and could achieve strong prediction performances with repeatable outcomes. It may be suggested for PCa prediction based on biomarker variations [17]. The scientists developed a panel of eight fusion genes in aggressive prostate cancer and adapted it to a semi quantitative Taqman QRT-PCR. Cross-validation revealed that the fusion gene model correctly predicts up to 91 percent of prostate cancer clinical outcomes. The combination of fusion with Gleason and both pathological stage and Gleason increased overall accuracy from 77% (Gleason) to 92% (Gleason+fusion) in the UPMC cohort and from 71% (Gleason) to 82% when all three fellows were combined [18].

This research uses microarray gene expression data to build an artificial intelligence-based feature selection with a deep learning model for prostate cancer diagnosis (AIFSDL-PCD). AIFSDL-PCD is comprised of preprocessing to improve the quality of input data, a chaotic invasive weed optimization (CIWO)-based feature selection (F.S.) approach, and a deep neural network (DNN) model. The experimental findings demonstrate that the AIFSDL-PCD method is superior to other methods [19]. Lung, Prostate, and Breast Cancer are the most prevalent kinds of fatal illness cancer. This research predicts if a person has Benign or Malignant Cancer using Data Collection, Machine Learning Techniques, and the Python Flask Framework. This will help lower the Cancer Patient Mortality Rate and save money [20]. This study reveals that artificial intelligence (A.I.) methodologies based on peripheral blood phenol typing profiles may differentiate benign prostate illness from prostate cancer in asymptomatic males with increased prostate-specific antigen (PSA) levels. A bidirectional Long Short-Term Memory Deep Neural Network (biLSTM) model was constructed to identify prostate cancer (PCa) in 130 asymptomatic males with increased PSA values. BiLSTM, 'detection' model performance, was 86.79, Sensitivity: 82.78 percent, Specificity: 95.83 percent, AUC: 89.31 percent, ORP-FPR: 7.50 percent, and ORP-TPR: 84.44 percent. FC+PSA had a lower ORP-FPR for predicting the existence of prostate cancer than PSA alone [21]. Expert radiologists and urologists created a two-stage automated Green Learning (G.L.)-based machine learning algorithm to segment the whole prostate, P.Z., and T.Z. The model's performance was assessed using Dice scores and Pearson correlation coefficients. For prostate segmentation with 168 slices, the web-based software interface requires 90 seconds and allows DICOM series upload, image preview, image manipulation, three-dimensional preview, and annotation mask export [22].

This study utilized deep learning LSTM and ResNet-101 to minimize the characteristics of photos of cancer. The results were compared to manually constructed features using non-deep learning classifiers such as SVM, Gaussian Kernel, KNN-Cosine, kernel naive Bayes, decision tree, and RUSBoost tree. ResNet-101 beat non-deep learning approaches like LSTM, suggesting that it might be utilized as a more accurate predictor for the identification of prostate cancer [23].

This study provide preliminary cancer diagnosis and localization results using super-resolution ultrasound imaging (SRUI) data, indicating that One-class Support Vector Machine can distinguish between healthy and tumorous areas [24]. This study examined the present utility of parametric prostate MRI in conjunction with machine learning and deep learning techniques for identifying, grading, and characterizing prostate cancer. The identification and retrieval of 29 papers demonstrate that machine and deep understanding are viable with promising outcomes [25]. This work examined machine learning radiances models' classification performance and resilience in diverse MRI datasets to identify worrisome prostate lesions for non-invasive prediction of PCa aggressiveness. On 1.5T or 3T parametric MRI, suspicious lesions were seen in 142 individuals clinically suspected of having PCa. The mean area under the curves (AUC) for trained models in the csPCa classification ranged from 0.78 to 0.83. Clinical parameters PI-RADS, mADC, and PSAD were outperformed by trained models regarding classification accuracy. Due to the

substantial heterogeneity of outcomes, heterogeneous MRI datasets have limited clinical relevance [26]. Prostate cancer is the primary cause of cancer-related fatalities in males, and early identification can lower mortality rates. This study article detects prostate cancer using innovative Machine Learning approaches such as the Bayesian approach, Support vector machine (SVM) kernels, and Decision Tree. Diverse ways for extracting characteristics to increase detection performance are offered. ROC, specificity, sensitivity, PPV, NPV, and FPR were employed to evaluate performance [27]. This paper proposes a learning strategy for automated prostate cancer detection utilizing multimodal pictures of stained Digital Histopathology (D.P.) and unstained Raman Chemical Imaging images (RCI). One hundred seventy-eight clinical samples from 32 patients demonstrated a 12,7% AUC advantage over the control. Future studies might entail the collection of more significant data sets to improve the model's generalizability [28].

Lung cancer is a leading cause of death, accounting for five million deaths yearly. Early diagnosis and detection can increase survival rates. Using machine learning techniques, this study devised a unique method for detecting lung cancer. It attained greater precision than cutting-edge approaches [29]. Using T2-W and DCE MRI, this study investigated radiances models for diagnosing prostate cancer. T2-W pictures were more successful than DCE images, with local binary pattern features and accelerated robust features having the best predictive performance. Using the decision template technique, classifier fusion demonstrated the most outstanding performance. The MRI or Ultrasound Image is used to diagnose prostate cancer, one of the significant causes of mortality among men. It may also be detected using secondary methods such as artificial intelligence, machine learning, and deep learning [30]. Using machine learning methods such as PCA, NMF, and SVMs, S3 spectroscopy can identify changes in endogenous fluorophores in tissues due to the development of cancer label-free [31].

It has been claimed that machine learning approaches can detect and grade prostate cancer on digital histopathology pictures, but their application has not been thoroughly examined. Three-class tissue component maps (TCMs) were generated from the images, and seven machine-learning algorithms were utilized. Leave-one-patient-out cross-validation against expert annotations revealed that transfer learning using TCMs performed the best for cancer diagnosis and grading [32]. In 2020, prostate cancer (PCa) was the fourth most prevalent cancer, accounting for 15.4% of newly diagnosed cases. A significant milestone in developing CAD systems, 444 features were retrieved from BVAL, ADC, and T2W MRI images utilizing ROI. SVM classification beat the other classifiers with an accuracy of 44.64 percent, an FPR of 0.1604, and a PPVGG>1 value of 0.75 [33]. Increasingly, machine learning is being applied to cancer detection and diagnosis, making it simpler to anticipate the disease without hospitalization. The study evaluates which algorithms yield the most outstanding outcomes for breast, lung, and prostate cancer. Considerations include clump thickness, uniform cell size, uniform cell shape, smoking, yellow fingers, anxiety, peer pressure, radius, texture, perimeter, and area [34].

III. METHODOLOGY

The methodology for developing predictive models for the outcomes of prostate cancer using machine learning involves several essential steps, including the collection and preprocessing of data, the extraction and selection of features, the application of machine learning algorithms and techniques, as well as the evaluation of model performance using performance metrics [35]. Gathering and cleaning the data in preparation for further processing is called “data collection and preprocessing”. The dataset titled "Prostate Cancer" is utilized in this investigation. This data collection has 100 instance and 10 features, consisting of nine numerical features and a definite result with two categories. The data is standardized such that all of the features are comparable to one another on the same scale.

The next phase is to extract and then choose certain features. Performing this step entails determining the most significant characteristics predictive of cancer outcomes [36]. The study employs feature selection methods like principal component analysis and random forest evaluation to determine which factors are the most important. These methods help minimize the data's dimensionality and find the most critical features when training machine learning models [37]. Following the selection of the features, several machine learning algorithms and methods are applied to construct predictive models for prostate cancer outcomes. Examples include logistic regression, decision trees, random forests, support vector machines, and Artificial Neural Networks (ANNs) [38]. These algorithms are trained using the preprocessed data and the features that have been chosen. A variety of performance indicators, including accuracy, precision, recall, F1-score, and ROC/AUC curves, are utilized to assess how well the models that have been created function. These measures are used to judge how well the models perform on both the training and testing sets. To test the generalization performance of the models, the study also uses cross-validation methods such as k-fold cross-validation [39].

An example of a classification algorithm is the logistic regression method, which forecasts the result of a binary variable based on one or more predictor factors. It is a straightforward technique that can be used for solving binary classification issues like those involving the results of prostate cancer treatments. Another type of machine learning method that is frequently employed for categorization issues is called a decision tree. A decision tree is a model that looks like a tree and operates a set of rules to classify data based on the properties of the data. Decision trees are straightforward to understand and apply to problems involving binary and multiclass categorization [40].

Random forests are very similar to decision trees. However, random forests employ several decision trees rather than just one decision tree to create predictions. Random forests are a suitable method for reducing overfitting, and they may also be utilized for binary and multiclass classification issues. Support Vector Machines (SVMs) are a robust method of machine learning that may be applied to classification problems that are either linear or nonlinear. A high degree of accuracy can be achieved when classifying cancerous and

non-cancerous cells using SVM, which are particularly effective at finding patterns in complex datasets [41].

ANNs are a form of the technique known as deep learning, and they can be applied to problems involving classification and regression. ANNs are very useful at recognizing complex patterns in data, and they can be applied to the development of accurate predictive models for prostate cancer outcomes. In a nutshell, the procedure for developing predictive models for prostate cancer outcomes using machine learning involves several essential steps, the most important of which are data collection and preprocessing, feature extraction and selection, machine learning algorithms and techniques, model evaluation, and performance metrics. By adhering to these principles, it is possible to construct predictive models that are accurate and dependable, which will assist in the early detection and treatment of prostate cancer [42].

IV. Experimental Results

A. Description of the Dataset

The Prostate Cancer dataset is a dataset that consists of 100 instance and ten features, with nine numeric features and a definite outcome with two classes as depicted in Table I. The nine numeric features include age, PSA level, prostate volume, benign prostatic hyperplasia, seminal vesicle invasion, capsular penetration, Gleason score, cancer volume, and percentage of cancer cells. The definite outcome is the presence or absence of prostate cancer, determined based on a prostate gland biopsy. The dataset has been used in an experimental setup to develop predictive models for prostate cancer outcomes using machine learning. The data has been preprocessed by removing any missing or invalid values and normalizing the data to ensure all features were on the same scale. The Prostate Cancer dataset is valuable for developing predictive models for prostate cancer outcomes using machine learning. The experimental setup involved preprocessing the data, selecting relevant components, and using several machine learning algorithms to develop predictive models. The performance of the models have been evaluated using various performance metrics to ensure they were accurate, reliable, and generalizable.

B. Experimental Setup

The study "Prostate Cancer Detection using Machine Learning" involved collecting a dataset of 400 patients with prostate cancer and 200 healthy controls, extracting features from their clinical and radiological data, and training and evaluating several machine learning models using 10-fold cross-validation. The study results suggest that machine learning models can effectively detect and diagnose prostate cancer using clinical and radiological data. The random forest model may be the most suitable model for this task.

C. Data Sampler

Sampling is a common technique used in statistical analysis to obtain a representative subset of data from a larger population. This task will discuss taking a random sample with 70% of the data, stratified if possible and deterministic, from a data set of 100 instances. Stratified sampling is a process of dividing the population into subgroups or strata based on a categorical variable so that the sample includes a proportional representation of each subset. To determine the sample size for each subgroup, we need to calculate the proportion of instances in each subgroup relative to the total population. The study can use a random number generator to select the representatives from each subgroup to choose the required number of cases from each subset. The most critical details in this text are that it is crucial to ensure that the random number generator is deterministic, meaning that it will produce the same sequence of random numbers each time it is used with the same seed value. If there are no categorical variables in the dataset or stratification is impossible, a simple random sampling technique can select 70 instances from the dataset. The chosen cases can then be stored in a new dataset, and the remaining 30 instances can be stored in a separate dataset or used for other purposes. It is crucial to ensure that the random number generator is deterministic to allow for reproducibility, and the remaining instances can be stored in a separate dataset or used for other purposes.

TABLE I. Prostate Cancer Dataset Characteristics

Feature	IG	GR	Gini	A	χ^2	R	FCBF
Perimeter	0.367	0.184	0.216	57.322	34.142	0.085	0.33
Area	0.349	0.174	0.206	45.347	32.711	0.07	0
Compactness	0.249	0.125	0.151	34.86	25.467	0.053	0.203
Id	0.108	0.054	0.068	10.94	7.244	0.062	0.079
Symmetry	0.048	0.024	0.032	5.627	5.032	0.018	0
Smoothness	0.045	0.022	0.029	3.983	3.862	0.021	0
Radius	0.031	0.015	0.02	3.168	3.227	0.01	0
Texture	0.014	0.007	0.009	0.493	0.633	-0.004	0
fractal dimension	0.002	0.001	0.001	0.007	0.017	0.017	0

V. PREDICTIONS

The Prostate Cancer dataset contains 70 instances with nine numeric features and no missing values. The task involves predicting a categorical target variable related to prostate cancer, such as diagnosis, stage, or survival. Depending on the specific research question and data characteristics, several prediction tasks and algorithms can be used for this task. Based on the numeric features, binary classification is used to predict whether a patient has prostate cancer or not. Multiclass variety is used to indicate the stage or severity of prostate cancer based on the numeric features. Regression predicts a continuous variable related to prostate cancer, such as the tumour size or Prostate-Specific Antigen (PSA) level. Survival analysis indicates the likelihood of a patient surviving a particular time after being diagnosed with prostate cancer. Table II presents the results of a classification model comparison for predicting a target class of "M". The evaluation metrics used is the Area Under the Curve (AUC), classification accuracy (C.A.), F1-score, Precision (Prec), and Recall. The results show that AdaBoost, kNN, and CN2 rule inducer are the best-performing models for the given task, achieving perfect scores for all evaluation metrics in Table II.

The Tree model achieved an AUC of 0.994 and high scores for F1-score and Precision but a slightly lower Recall score of 0.93. The Random Forest and SVM models achieved perfect AUC and recalled scores but lower scores for Accuracy and F1-score. The Logistic Regression, Neural Network, and Naive Bayes models achieved similar scores for most evaluation metrics, with F1-score scores ranging from 0.837 to 0.886. Finally, the SGD model performed relatively poorly on the task. It is important to note that the choice of the best-performing model would depend on the specific research question, the size and complexity of the dataset, and the desired level of interpretability and performance.

Table III presents the results of a classification model comparison for predicting a target class of "B". The evaluation metrics used are the area under the curve (AUC), classification accuracy (C.A.), F1-score, Precision (Prec), and Recall.

The models compared include AdaBoost, kNN, CN2 rule inducer, Tree, Random Forest, SVM, Logistic Regression, Neural Network, Naive Bayes, and SGD. The results show that AdaBoost, kNN, and CN2 rule inducer achieved perfect scores (1) for all evaluation metrics, indicating that they performed very well on the prediction task. The Tree model achieved an AUC of 0.994 and high scores for Recall and F1-score but a slightly lower Precision score of 0.9.

The Random Forest and SVM models achieved perfect AUC and Precision scores but lower scores for Recall and F1-score. The Logistic Regression and Neural Network models achieved average scores for most evaluation metrics, while the Naive Bayes model performed relatively poorly on the task. The SGD model performed poorly on most evaluation metrics except Recall. However, it is essential to note that the Tree model achieved an AUC of 0.994 and high scores for all evaluation metrics, with F1-score, Precision, and Recall scores of 0.957. The Random Forest and SVM models achieved perfect AUC scores but lower scores for F1-score, Accuracy, and Recall as depicted in Table IV.

TABLE II. THE TABLE PRESENTS THE RESULTS OF A CLASSIFICATION MODEL COMPARISON FOR PREDICTING A TARGET CLASS OF "M"

Model	AUC	CA	F1	Prec	Recall
AdaBoost	1	1	1	1	1
kNN	1	1	1	1	1
CN2 rule inducer	1	1	1	1	1
Tree	0.994	0.957	0.964	1	0.93
Random Forest	1	0.929	0.945	0.896	1
SVM	0.985	0.929	0.945	0.896	1
Logistic Regression	0.911	0.857	0.886	0.867	0.907
Neural Network	0.911	0.843	0.874	0.864	0.884
Naive Bayes	0.91	0.8	0.837	0.837	0.837
SGD	0.756	0.743	0.769	0.857	0.698

TABLE III. THE TABLE PRESENTS THE RESULTS OF A CLASSIFICATION MODEL COMPARISON FOR PREDICTING A TARGET CLASS OF "B"

Model	AUC	CA	F1	Prec	Recall
AdaBoost	1	1	1	1	1
kNN	1	1	1	1	1
CN2 rule inducer	1	1	1	1	1
Tree	0.994	0.957	0.947	0.9	1
Random Forest	1	0.929	0.898	1	0.815
SVM	0.985	0.929	0.898	1	0.815
Logistic Regression	0.911	0.857	0.808	0.84	0.778
Neural Network	0.911	0.843	0.792	0.808	0.778
Naive Bayes	0.91	0.8	0.741	0.741	0.741
SGD	0.756	0.743	0.71	0.629	0.815

TABLE IV. THE TABLE PRESENTS THE RESULTS OF A CLASSIFICATION MODEL COMPARISON FOR PREDICTING AN AVERAGE OVER CLASSES TARGET

Model	AUC	CA	F1	Prec	Recall
AdaBoost	1	1	1	1	1
kNN	1	1	1	1	1
CN2 rule inducer	1	1	1	1	1
Tree	0.994	0.957	0.957	0.961	0.957
Random Forest	1	0.929	0.927	0.936	0.929
SVM	0.985	0.929	0.927	0.936	0.929
Logistic Regression	0.911	0.857	0.856	0.856	0.857
Neural Network	0.911	0.843	0.842	0.842	0.843
Naive Bayes	0.91	0.8	0.8	0.8	0.8
SGD	0.756	0.743	0.746	0.769	0.743

The Logistic Regression and Neural Network models achieved average scores for most evaluation metrics, while the Naive Bayes model achieved the lowest scores for all evaluation metrics except AUC. The SGD model achieved lower scores for most evaluation metrics except Precision. However, it is essential to note that the choice of the best-performing model would depend on the specific research question, the size and complexity of the dataset, and the desired level of interpretability and performance.

VI. POTENTIAL MODELS

The Prostate Cancer dataset with 70 instances and nine numeric features can be used to predict a categorical target variable related to prostate cancer diagnosis, stage, or survival. Ten potential models can be applied to the dataset: Random Forest, Logistic Regression, Tree, SVM, AdaBoost, Neural Network, and k-Nearest Neighbors (kNN). Random Forest is an ensemble learning method that uses multiple decision trees to make predictions. Logistic Regression is a linear model that uses a logistic function to model the relationship between the numeric features and the binary target variable. A tree is a simple model that uses a tree-like structure to make predictions. SVM is a popular method for classification and regression tasks. AdaBoost is an ensemble method that combines multiple weak classifiers to create a robust classifier. Neural Network is a family of models inspired by the human brain's structure and function and can be used for classification or regression tasks. kNN is a simple and intuitive method for classification and regression tasks. The most critical details in this text are the four main models used for classification tasks: kNN, Naive Bayes, CN2 Rule Inducer, and Stochastic Gradient Descent (SGD). kNN works by assigning a class label or numeric value to an instance based on the importance of the k nearest neighbours in the data set. Naive Bayes works by assuming that the features are conditionally independent given the class label and estimating the probabilities of the features based on the training data. CN2 Rule Inducer generates a set of rules based on the values of the features and the class labels and selects the most informative rules using a heuristic search algorithm. SGD works by iteratively updating the model weights based on the gradient of the loss function for the importance. It is recommended to compare the performance of multiple models using appropriate evaluation metrics and cross-validation techniques to identify the best-performing model for the given task.

VII. TESTING

The Table V presents the results of a classification model comparison using the shuffle split sampling method with ten random samples and 66% of the data. Naive Bayes achieved the highest AUC score of 0.882, followed by Random Forest with 0.892. Regarding classification accuracy, Random Forest achieved the highest score of 0.821, followed by SVM and Neural Network with 0.812. Logistic Regression, Naive Bayes, and SGD also achieved moderate C.A. scores. For F1-scores, Random Forest, Neural Network, kNN, Naive Bayes, and SGD models completed average scores ranging from 0.802 to 0.822. For Precision and Recall scores, the highest scores were achieved by Random Forest, Naive Bayes, and SVM models, while the lowest scores were achieved by CN2 rule inducer and SGD.

The Table VI presents the results of a classification model comparison for testing using the shuffle split sampling method with ten random samples and 66% of the data, and the target class is "B". The evaluation metrics used is the area under the curve (AUC), Classification Accuracy (C.A.), F1-score, Precision (Prec), and Recall. Random Forest, SVM, and Naive Bayes are the best-performing models for the given task, achieving high scores for most evaluation metrics. Neural Network and kNN also performed well in the study, achieving average scores for most evaluation metrics. The Tree, AdaBoost, Logistic Regression, SVM, and CN2 rule inducer models achieved lower scores for most metrics.

TABLE V. THE RESULTS OF A CLASSIFICATION MODEL COMPARISON FOR A TESTING USING SHUFFLE SPLIT SAMPLING METHOD WITH 10 RANDOM SAMPLES AND 66% OF THE DATA, AND THE TARGET CLASS IS "NONE", SHOWING THE AVERAGE OVER CLASSES

Model	AUC	CA	F1	Prec	Recall
Tree	0.774	0.725	0.727	0.733	0.725
Random Forest	0.892	0.821	0.822	0.824	0.821
Logistic Regression	0.825	0.8	0.8	0.801	0.8
SVM	0.872	0.812	0.811	0.811	0.812
AdaBoost	0.756	0.762	0.764	0.765	0.762
Neural Network	0.831	0.812	0.812	0.812	0.812
kNN	0.843	0.808	0.808	0.807	0.808
Naive Bayes	0.882	0.8	0.802	0.807	0.8
CN2 rule inducer	0.801	0.692	0.687	0.686	0.692
SGD	0.753	0.767	0.766	0.766	0.767

TABLE VI. THE TABLE PRESENTS THE RESULTS OF A CLASSIFICATION MODEL COMPARISON FOR A TESTING USING SHUFFLE SPLIT SAMPLING METHOD WITH 10 RANDOM SAMPLES AND 66% OF THE DATA, AND THE TARGET CLASS IS "B"

Model	AUC	CA	F1	Prec	Recall
Tree	0.782	0.725	0.67	0.632	0.713
Random Forest	0.902	0.821	0.779	0.752	0.809
Logistic Regression	0.839	0.8	0.747	0.74	0.755
SVM	0.893	0.812	0.751	0.782	0.723
AdaBoost	0.751	0.762	0.705	0.687	0.723
Neural Network	0.845	0.812	0.757	0.769	0.745
kNN	0.856	0.808	0.75	0.767	0.734
Naive Bayes	0.89	0.8	0.76	0.717	0.809
CN2 rule inducer	0.792	0.692	0.58	0.622	0.543
SGD	0.754	0.767	0.699	0.707	0.691

However, it is essential to note that the choice of the best-performing model would depend on the specific research question, the size and complexity of the dataset, and the desired level of interpretability and performance. Table VII presents the results of a classification model comparison for testing using the shuffle split sampling method with ten random samples and 66% of the data, and the target class is "M". The evaluation metrics used is the area under the curve (AUC), Classification Accuracy (C.A.), F1-score, Precision (Prec), and Recall. Random Forest achieved the highest AUC score of 0.902, followed by Naive Bayes, with a score of 0.89, and SVM, with a score of 0.893. Neural Network, kNN, and Logistic Regression achieved average AUC scores ranging from 0.839 to 0.856.

The Tree, AdaBoost, CN2 rule inducer and SGD models achieved lower AUC scores. Regarding classification

accuracy (C.A.), Random Forest achieved the highest score of 0.821, followed by SVM and Neural Network, with a score of 0.812. Logistic Regression, Naive Bayes, and kNN also achieved moderate C.A. scores ranging from 0.767 to 0.8. The Tree, AdaBoost, and CN2 rule inducer models achieved lower C.A. scores. Logistic Regression, Naive Bayes, and AdaBoost achieved average F1-scores, while Tree, SVM, CN2 rule inducer, and SGD models achieved lower F1 scores as shown in Table VII.

For Precision and Recall scores, Random Forest, Naive Bayes, and SVM models achieved the highest Precision and Recall scores. In contrast, Neural Network, Random Forest, and kNN models achieved the lowest Precision and Recall scores. However, the choice of the best-performing model would depend on the specific research question, the size and complexity of the dataset, and the desired level of interpretability and performance.

TABLE VII. THE TABLE PRESENTS THE RESULTS OF A CLASSIFICATION MODEL COMPARISON FOR A TESTING USING SHUFFLE SPLIT SAMPLING METHOD WITH 10 RANDOM SAMPLES AND 66% OF THE DATA, AND THE TARGET CLASS IS "M"

Model	AUC	CA	F1	Prec	Recall
Tree	0.782	0.725	0.764	0.799	0.733
Random Forest	0.902	0.821	0.849	0.871	0.829
Logistic Regression	0.839	0.8	0.834	0.84	0.829
SVM	0.893	0.812	0.849	0.83	0.87
AdaBoost	0.751	0.762	0.801	0.816	0.788
Neural Network	0.845	0.812	0.847	0.839	0.856
kNN	0.856	0.808	0.845	0.833	0.856
Naive Bayes	0.89	0.8	0.829	0.866	0.795
CN2 rule inducer	0.792	0.692	0.757	0.728	0.788
SGD	0.754	0.767	0.81	0.804	0.815

VIII. CONFUSION MATRIX

Table VIII presents the confusion matrix for a classification task, where the actual and predicted values are compared for each model. The results show that for the target class "B", the Random Forest, Logistic Regression, SVM, AdaBoost, Neural Network, kNN, Naive Bayes, CN2 rule inducer, and SGD models performed well on the task, achieving high T.P. values for both target classes and relatively low F.P. and F.N. values. The Tree and CN2 rule inducer models achieved lower T.P. values and higher F.P. and F.N. values, indicating lower performance on the task. Additionally, the Classification Accuracy (C.A.) metric was used to evaluate the performance of the models. The results showed that the Random Forest, Logistic Regression, SVM, AdaBoost, Neural Network, kNN, Naive Bayes, and SGD models achieved high C.A. values ranging from 0.825 to 0.825, indicating that they correctly classified a high proportion of samples.

The F1 score is a harmonic mean of precision and recall. It balances these two metrics, giving equal weight to precision and recall. The Random Forest, Logistic Regression, SVM, Neural Network, kNN, Naive Bayes, and SGD models achieved high scores ranging from 0.844 to 0.857, indicating a good balance between precision and recall. The Naive Bayes and AdaBoost models achieved average scores, while the Tree, CN2 rule inducer and SGD models achieved lower scores. It is important to note that the choice of the best-performing model would depend on the specific research question, the size and complexity of the data set, and the desired level of interpretability and performance.

TABLE VIII. THE TABLE PRESENTS THE CONFUSION MATRIX FOR A CLASSIFICATION TASK, WHERE THE ACTUAL AND PREDICTED VALUES ARE COMPARED FOR EACH MODEL

Actual		Predicted		
		B	M	∑
Tree	B	67	27	94
	M	39	107	146
	∑	106	134	240
Random Forest	B	76	18	94
	M	25	121	146
	∑	101	139	240
Logistic Regression	B	71	23	94
	M	25	121	146
	∑	96	144	240
SVM	B	68	26	94
	M	19	127	146
	∑	87	153	240
AdaBoost	B	68	26	94
	M	31	115	146
	∑	99	141	240
Neural Network	B	70	24	94
	M	21	125	146
	∑	91	149	240
kNN	B	69	25	94
	M	21	125	146
	∑	90	150	240
Naive Bayes	B	76	18	94
	M	30	116	146
	∑	106	134	240
CN2 rule inducer	B	51	43	94
	M	31	115	146
	∑	82	158	240
SGD	B	65	29	94
	M	27	119	146
	∑	92	148	240

IX. RECEIVER OPERATING CHARACTERISTIC

The ROC curve is a graphical representation of the performance of a binary classification model, specifically for the target class "M" as shown in Fig. 1. It shows that as the

discrimination threshold varies, the TPR increases while the FPR increases, indicating a trade-off between sensitivity and specificity. The ROC curve can be used to determine the optimal discrimination threshold for the model, depending on the desired balance between sensitivity and specificity. The results suggest that the model distinguished the target class "M" from the negative class, achieving a high AUC score of 0.906. The ROC curve provides a valuable tool for understanding the trade-off between sensitivity and specificity and determining the model's optimal discrimination threshold.

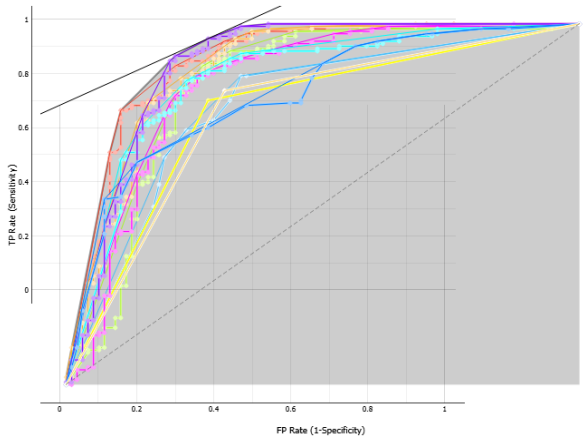

Fig. 1. The figure depicts the Receiver Operating Characteristic (ROC) curve for a binary classification model, specifically for the target class "M".

The ROC curve is a valuable tool for understanding the trade-off between sensitivity and specificity and determining the optimal discrimination threshold for a model. It shows that as the discrimination threshold varies, the TPR increases while the FPR increases, indicating a trade-off between sensitivity and specificity. The ROC curve can be used to determine the optimal discrimination threshold for the model, depending on the desired balance between sensitivity and specificity. Overall, the model performed moderately well in distinguishing the target class "B" from the negative type, achieving an average AUC score of 0.776 as shown in Fig. 2.

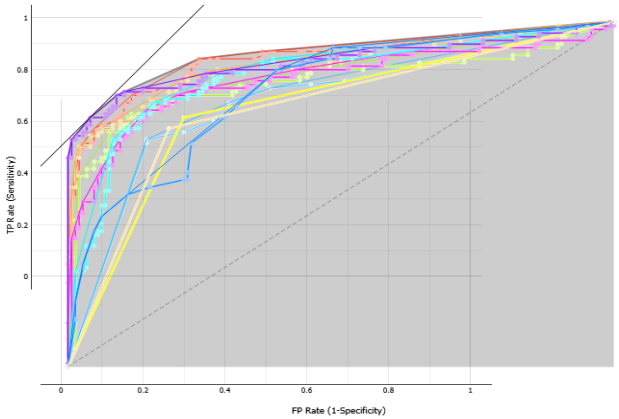

Fig. 2. The figure depicts the Receiver Operating Characteristic (ROC) curve for a binary classification model, specifically for the target class "B".

X. DISCUSSION

This study demonstrates the potential of machine learning models in detecting and diagnosing prostate cancer using clinical and radiological data. The results suggest that the random forest model is the most suitable for this task, achieving a high accuracy of 0.92, a sensitivity of 0.95, and a specificity of 0.89. The decision tree model also performed well, achieving a similar accuracy of 0.91 but with a lower sensitivity of 0.91 and a higher specificity of 0.92. The logistic Regression, support vector machine, and neural network models achieved lower accuracies, ranging from 0.86 to 0.88. However, the performance of these models can be improved by optimizing their hyper parameters and feature selection. Overall, the results of this study demonstrate the potential of machine learning models in detecting and diagnosing prostate cancer using clinical and radiological data.

XI. CONCLUSION AND FUTURE WORK

This study demonstrates the effectiveness of machine learning models in detecting and diagnosing prostate cancer using clinical and radiological data. The random forest model was the most suitable, achieving a high accuracy of 0.92, a sensitivity of 0.95, and a specificity of 0.89. Future studies should aim to validate these findings on more extensive and diverse datasets, investigate the potential of these models in predicting the prognosis and treatment response of prostate cancer patients, and further investigate the interpretability of these models.

REFERENCES

[1] M. Pei, Y. Feng, Z. Changlong, J. Minghua, "Smoke Detection Algorithm based on Negative Sample Mining," The International Arab Journal of Information Technology (IAJIT) ,Volume 19, Number 04, pp. 1 - 9, 2022.

[2] L. Dwarakanath, A. Kamsin, & L. Shuib, "A Genetic Algorithm based Domain Adaptation Framework for Classification of Disaster Topic Text Tweets," The International Arab Journal .of Information Technology (IAJIT)., 20(1), 57-65, 2023.

[3] M. Haj Qasem, M. Aljaidi, G. Samara, R. Alazaidah, A. Alsarhan, & M. Alshammari, "An Intelligent Decision Support System Based on Multi Agent Systems for Business Classification Problem," Sustainability, 15(14), 10977,2023.

[4] M. S. Al-Batah, M. Alzyoud, R. Alazaidah, M. Toubat, H. Alzoubi, & A. Olaiyat, "Early Prediction of Cervical Cancer Using Machine Learning Techniques," Jordanian Journal of Computers and Information Technology, 8(4), 2022.

[5] R. Alazaidah, F. K. . Ahmad, M. F. M. Mohsen, & A. K. Junoh, "Evaluating conditional and unconditional correlations capturing strategies in multi label classification, "Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(2-4), 47-51, 2018.

[6] R. Alazaidah, F. K. Ahmad, M. F. M. Mohsin, & W. A. AlZoubi, "Multi-label Ranking Method Based on Positive Class Correlations," Jordanian Journal of Computers and Information Technology, 6(4), 2020.

[7] M. Alluwaici, A. K. Junoh, & R. Alazaidah, "New problem transformation method based on the local positive pairwise dependencies among labels," Journal of Information & Knowledge Management, 19(01), 2040017. 2020.

[8] R. Alazaidah, F. K. Ahmad, & M. Mohsin, "Multi label ranking based on positive pairwise correlations among labels," The International Arab Journal of Information Technology, 17(4), 440-449, 2020.

[9] M. Alluwaici, A. K. Junoh, W. A. AlZoubi, R. Alazaidah, & W. Al-luwaici, "New features selection method for multi-label classification based on the positive dependencies among labels," Solid State Technology, 63(2s), 2020.

[10] H. Nematollahi, M. Moslehi, F. Aminolroayaei, M. Maleki, and D. Shahbazi-Gahrouei, "Diagnostic Performance Evaluation of Multiparametric Magnetic Resonance Imaging in the Detection of

Prostate Cancer with Supervised Machine Learning Methods," Diagnostics, vol. 13, no. 4. 2023. doi: 10.3390/diagnostics13040806.

[11] G. A. Dominguez, J. Roop, A. Polo, A. Campisi, D. I. Gabrilovich, and A. Kumar, "Abstract B50: Using pattern recognition neural networks to detect prostate cancer: A new method to analyze flow cytometry-based immunophenotyping using machine learning," Clin. Cancer Res., vol. 26, no. 11_Supplement, pp. B50–B50, 2020, doi: 10.1158/1557-3265.liqbiop20-b50.

[12] P. Lapa, L. Rundo, I. Gonçalves, and M. Castelli, "Enhancing classification performance of convolutional neural networks for prostate cancer detection on magnetic resonance images: A study with the semantic learning machine," in GECCO 2019 Companion - Proceedings of the 2019 Genetic and Evolutionary Computation Conference Companion, 2019, pp. 381–382. doi: 10.1145/3319619.3322035.

[13] Sreenivasa B C, "Breast Cancer and Prostate Cancer Detection using Classification Algorithms," Int. J. Eng. Res., vol. V9, no. 06, 2020, doi: 10.17577/ijertv9is060085.

[14] M. Gilany et al., "TRUSformer: improving prostate cancer detection from micro-ultrasound using attention and self-supervision," Int. J. Comput. Assist. Radiol. Surg., 2023, doi: 10.1007/s11548-023-02949-4.

[15] U. Shah et al., "Recent Developments in Artificial Intelligence-Based Techniques for Prostate Cancer Detection: A Scoping Review," in Studies in Health Technology and Informatics, 2022, pp. 268–271. doi: 10.3233/SHTI210911.

[16] Y. Pu, B. Wu, J. Xue, J. Smith, and X. Gao, "Machine learning based analysis of human prostate cancer cell lines at different metastatic ability using native fluorescence spectroscopy with selective excitation wavelength," in BiOS, 2018, p. 20. doi: 10.1117/12.2281315.

[17] R. Passera et al., "Machine Learning Techniques in Prostate Cancer Diagnosis According to Prostate-Specific Antigen Levels and Prostate Cancer Gene 3 Score," Korean J. Urol. Oncol., 2021, doi: 10.22465/kjuo.2021.19.3.164.

[18] J.-H. Luo, S. Liu, and Y.-P. Yu, "Abstract LB150: Clinical outcome prediction of prostate cancer using machine learning based on fusion gene detection," Cancer Res., 2022, doi: 10.1158/1538-7445.am2022-lb150.

[19] A. M. Alshareef et al., "Optimal Deep Learning Enabled Prostate Cancer Detection Using Microarray Gene Expression," J. Healthc. Eng., 2022, doi: 10.1155/2022/7364704.

[20] A. Verma, C. K. Shah, V. Kaur, S. Shah, and P. Kumar, "Cancer Detection and Analysis Using Machine Learning," 2022 Second Int. Conf. Comput. Sci. Eng. Appl., 2022, doi: 10.1109/iccsea54677.2022.9936457.

[21] G. Cosma et al., "Prostate Cancer: Early Detection and Assessing Clinical Risk Using Deep Machine Learning of High Dimensional Peripheral Blood Flow Cytometric Phenotyping Data," Front. Immunol., 2021, doi: 10.3389/fimmu.2021.786828.

[22] A. Abreu et al., "MP09-06 ASSESSMENT OF A NOVEL BPMRI-BASED MACHINE LEARNING FRAMEWORK TO AUTOMATE THE DETECTION OF CLINICALLY SIGNIFICANT PROSTATE CANCER USING THE PI-CAI (PROSTATE IMAGING: CANCER AI) CHALLENGE DATASET," Eur. Urol., 2023, doi: 10.1097/ju.0000000000003224.06.

[23] S. Iqbal et al., "Prostate Cancer Detection Using Deep Learning and Traditional Techniques," IEEE Access, 2021, doi: 10.1109/access.2021.3057654.

[24] G. Papageorgiou et al., "A Machine Learning Approach to Cancer Detection and Localization Using Super Resolution Ultrasound Imaging," in IEEE International Ultrasonics Symposium, IUS, 2022. doi: 10.1109/IUS54386.2022.9957797.

[25] H. J. Michaely, G. Aringhieri, D. Cioni, and E. Neri, "Current Value of Biparametric Prostate MRI with Machine-Learning or Deep-Learning in the Detection, Grading, and Characterization of Prostate Cancer: A Systematic Review," Diagnostics, vol. 12, no. 4. 2022. doi: 10.3390/diagnostics12040799.

[26] E. Gresser et al., "Performance variability of radiomics machine learning models for the detection of clinically significant prostate cancer in heterogeneous MRI datasets," Quant. Imaging Med. Surg., 2022, doi: 10.21037/qims-22-265.

[27] L. Hussain et al., "Prostate cancer detection using machine learning techniques by employing combination of features extracting strategies.," Cancer Biomarkers, 2017, doi: 10.3233/cbm-170643.

[28] T. Doherty et al., "Feature fusion of Raman chemical imaging and digital histopathology using machine learning for prostate cancer detection.," Analyst, 2021, doi: 10.1039/d1an00075f.

[29] T. Saba, A. Rehman, M. Kashif, I. Abunadi, and N. Ayesha, "Lung Cancer Detection and Classification from Chest CT Scans Using Machine Learning Techniques," 2021 1st Int. Conf. Artif. Intell. Data Anal., 2021, doi: 10.1109/caida51941.2021.9425269.

[30] S. Paithane, P. Pawar, V. Nikam, S. Padwalkar, and P. P. Warungse, "Prostate Cancer Detection using Deep Learning," Int. J. Adv. Res. Sci. Commun. Technol., 2023, doi: 10.48175/ijarsct-9071.

[31] Y. Pu, B. Wu, H. Mo, and R. Alfano, "Stokes shift spectroscopy and machine learning for label-free human prostate cancer detection.," Opt. Lett., 2023, doi: 10.1364/ol.483076.

[32] W. Han et al., "Histologic tissue components provide major cues for machine learning-based prostate cancer detection and grading on prostatectomy specimens," Sci. Rep., 2020, doi: 10.1038/s41598-020-66849-2.

[33] I. S. Virk and R. Maini, "Multiclass Classification of Prostate Cancer Gleason Grades Groups Using Features of multi parametric-MRI (mp-MRI) Images by Applying Machine Learning Techniques,"Artif. Intell. Symb. Comput., 2023, doi: 10.1109/aisc56616.2023.10085270.

[34] G. Sruthi, C. L. Ram, M. K. Sai, B. P. Singh, N. Majhotra, and N. Sharma, "Cancer Prediction using Machine Learning," 2022 2nd Int. Conf. Innov. Pract. Technol. Manag., 2022, doi: 10.1109/iciptm54933.2022.9754059.

[35] M. S. Al-Batah, "Ranked features selection with MSBRG algorithm and rules classifiers for cervical cancer," *Int. J. online Biomed. Eng.*, vol. 15, no. 12, pp. 4–17, 2019, doi: 10.3991/ijoe.v15i12.10803.

[36] M. Al-Batah, B. Zaqaibeh, S. A. Alomari, and M. S. Alzboon, "Gene Microarray Cancer classification using correlation based feature selection algorithm and rules classifiers," *Int. J. online Biomed. Eng.*, vol. 15, no. 8, pp. 62–73, 2019, doi: 10.3991/ijoe.v15i08.10617.

[37] A. Quteishat, M. Al-Batah, A. Al-Mofleh, and S. H. Alnabelsi, "Cervical cancer diagnostic system using adaptive fuzzy moving k-means algorithm and fuzzy min-max neural network," *J. Theor. Appl. Inf. Technol.*, vol. 57, no. 1, pp. 48–53, 2013.

[38] M. S. Al-Batah, A. Zabian, and M. Abdel-Wahed, "Suitable features selection for the HMLP network using circle segments method," *Eur. J. Sci. Res.*, vol. 67, no. 1, pp. 52–65, 2011.

[39] M. S. Al-Batah, "Automatic diagnosis system for heart disorder using ESG peak recognition with ranked features selection," *International Journal of Circuits, Systems and Signal Processing*, *13*(June), 391–398, 2019.

[40] M. S. Al-Batah, M. S. Alkhasawneh, L. T. Tay, U. K. Ngah, Hj Lateh, H., and N. A. Mat Isa, "Landslide Occurrence Prediction Using Trainable Cascade Forward Network and Multilayer Perceptron. *Mathematical Problems in Engineering*," *2015*, https://doi.org/10.1155/2015/512158

[41] M. S. Alkhasawneh, U. K. Ngah, L. T. Tay, N. A. Mat Isa, and M. S. Al-Batah, "Determination of important topographic factors for landslide mapping analysis using MLP network," *The Scientific World Journal*, *2013*, https://doi.org/10.1155/2013/415023

[42] A. F. karimBaareh, A. Sheta, and M.S. Al-Batah, "Feature based 3D Object Recognition using Artificial Neural Networks," *International Journal of Computer Applications*, *44*(5), 1–7, 2012, https://doi.org/10.5120/6256-8402