PROSTATE CANCER DETECTION USING GREEDY SEARCH FOR LESIONS IN MAGNETIC RESONANCE IMAGES (MRI): A NOVEL TECHNIQUE

GUNDOORI MUBARAK¹, RAJABHUSHANAM C²

¹Research Scholar, ²Professor

¹Bharath Institute of Science and Technology,
²Bharath Institute of Higher Education and Research, Selaiyur, Chennai, Tamil Nadu

Email: ¹mubarak.gvic@gmail.com, ²rajabhushanamc.cse@bharathuniv.ac.in

ABSTRACT

Prostate cancer is one of the leading causes of cancer death in males. The death rate from prostate cancer can be drastically lowered with well-thought-out plans for treatment and maintenance. Cancer diagnosis by magnetic resonance imaging is a great first step in treatment (MRI). New research shows that MRI can also be used to categorise prostate cancer. There may be hereditary cancer implications for men and their families because genetic abnormalities can greatly raise the risk for prostate cancer (PC), and may be associated with aggressive disease and poorer outcomes. MRI is a useful tool for diagnosing the pathological conditions associated with PC (MRI). Swarm intelligence is the basis for Particle Swarm Optimization (PSO), an algorithm used to solve optimization problems in search spaces and to analyse and forecast social behaviour under the assumption of the existence of goals. In order to facilitate PC detection in MRIs, the authors herein employ a greedy technique for hybridising PSO optimization. The Adaboost classifier is another option that has been proposed. In order to better diagnose prostate cancer, this study suggests using an internal learning technique.

Keywords: Prostate cancer diagnosis, MRI imaging, object recognition, inner learning, and prostate segmentation

1. INTRODUCTION

According to recent worldwide cancer data [1,] lung cancer accounts for 14.5% of male cases, with prostate cancer coming in second at 13.5%. In the biography, a needle biopsy was performed to get a case sample, and then H&E staining and microscopy were used to provide a standard diagnosis of prostate cancer. In order to study the morphological structure of the tissues of the Glandular structures are maintained by epithelial cells in low-grade tumours but are lost in high-grade tumours. Microarray tissue analysis of men with prostate cancer has identified five distinct growth patterns, numbered 1 through 5, that correlate to distinct cell tissue morphologies associated with improved survival rates. Normal tissue and poor prognosis are essentially indistinguishable. Pathological sections' growth patterns are often divided into primary structure and secondary structure, with each representing a different percentage of the whole. Totaling the scores for the major and secondary structures yields a total score; this value is used to categorise patients into one of many prognosis categories, with lower scores (below 6) often indicating a more favourable outcome.

Machine learning models, regression trees, and Probabilistic classifiers are frequently used in the Demonstrating significant instantaneous performance assessment method for prostate cancer to extract this same functionality establishment and identify the available. It is discussed how analytical paradigms like Bayesian classifiers (which identify likelihood functions as logics premised on situations parsing rather than hard sampling rate) and supervised learning classification technique and multiple regressions like random forests (which use random data segments and are naturally suitable for binary classification concerns) can be used to address these issues. To generalise feature classification, any of the aforementioned techniques can be applied. Using ResNet18 as a basis, reference [4] classifies neural networks as either "classification" or "generation" networks.

A classification model is used by the discriminative network to achieve its objectives. The presence of glandular structures is initially identified using textural cues and morphometric information acquired from glandular units [5]. The power spectrum displays the temporal evolution of an image's textures [6].
A large discrepancy was seen between the Gleason scores reported by the various classifiers. Convolutional neural networks and other deep learning-based algorithms can execute feature learning and classification operations in a unified framework once training data sets reach a certain size [7], allowing for improved results without excessive reliance on manual annotation. More and more clinical settings are using computer auxiliary diagnostic systems for a variety of medical diagnosis and therapy [8, 9] as deep learning and computer vision continue to develop. Using a dataset of 120k images as training data, the inceptionv3 classification system performed at a level on par with that of human dermatologists [10]. Finally, a deep network classifier was trained using hundreds of images of histocells to provide referral and diagnostic predictions [2]. Therefore, both cancer's presence and the affected tissue's Gleason score have been verified.

2. RELATED WORK

Although non-CNN-based prostate cancer screening approaches have apparent drawbacks, they have made some progress. Klein, C et al. [10] proposed a completely automated computer-assisted prostate cancer screening approach in particular. This method categorises voxels of various sizes using the Hessian-based blob detection algorithm. On the multi-parameter MR picture, the candidate detection frame is then defined using histogram analysis. Finally, the acquired feature set is classified as benign or malignant using a two-stage classification approach. Chung, M et al. [11] employed the log-likelihood function to distinguish the prostate in diffusion-weighted MR images. The local strain caused by the client's inhalation and area motion is then addressed using a non-rigid matching approach. The prostate was segmented using a level set model, then normalised and optimised using an upgraded Gauss-Markov random field picture model. Finally, using a stacked non-negative constraint technique, Prostate tumours are classified as cancerous using a deep learning ANN model network. Using a radiation-driven provisional random utility methodology, Leyh-Bannurah, Set al. [12] suggested a different technique for autonomously identifying bladder cancer.

To better diagnose prostate cancer, they used a relationship seen between geographic and physicochemical properties of volumetric coordinates, as well as a full set of vertex shader quantitative radiological factors [13]. Shvero, Aet al. [14] devised a method for automating the diagnosis of different bladder illnesses. Create a harmful risk map for all verticles in the prostate, then locate the troublesome region for bladder cancer localization using a prospective categorization technique. The BoB MIL approach was developed by Morote, J et al. [15], and each of the increased bags (or parent bags) contains numerous subordinate occurrence containers. The BoB MIL technique was used to overcome the problem of diagnosing prostate cancer using magnetic resonance imaging data.

Keeping in mind the distinction between prostate cancer lesion and non-lesion zones. After segmenting the prostate and collecting various Tonyali, Set al. [16] Based on demographic data in the prostate area, used an image similarity measure to diagnose tumours. MRI series with Interpolation techniques, gradient, and calculated mass transfer were used in this study, Carina et al. retrieved image brightness, slope, contour motion, and length information. These attributes are then utilised to diagnose tumours using an extra discriminant analysis model. Banno, Tet al. [17] suggested a layer group learning system for radical prostatectomy that can simultaneously create a medical framework and extract explainable clinical criteria.

Due to this rapid expansion of deep learning, many CNN-based algorithms have improved prostate cancer diagnosis findings in recent years. Klein, C et al. [18], for example, used a novel form of deep convolutional neural network to develop a machine prostate diagnostic accuracy and diagnostic approach by van Riel, L. A et al [19]. During a multi-parameter magnetic resonance imaging session, the prostate is discovered first. Then go to CNN and look for photographs of guys with prostate cancer to get a broad picture of the disease. Finally, the aggressiveness of each local lesion is determined using SVM classifiers and multi-modal CNN features. A multi-modal CNN-based automatic prostate cancer screening system was proposed by Porcaro, A. Bet al. [20]. Chun, Fet al. [21] proposed a multi-modal CNN-based automatic prostate cancer detection system. Abedi, Aet al. [22] proposed a prostate cancer detection system with multiple sub-networks: a bladder clinical diagnostic adult neurogenesis warping network and a testicular cancer detection twin convolutional neural network. Abedi, A. et al. [23] enhanced the ability of a CNN to detect prostate malignant cells and divide cells correctly. They used focus loss to balance the malignant and
non-cancerous regions, and developed selective dense artificial potential sectors for thread to improve lesion identification by optimising the segmented lesion mask [23]. Cantiello, et al [25] 2.5D's approach can segment the prostate while also diagnosing prostate cancer is a type of cancer that affects. They employ deep learning algorithms to find more elements in the picture T2W graphical, then reconfigure the attributes to the source to provide each frame the predicted value [26].

Porcaro, Aet al. [27] integrated two parallel convolutional networks to develop a combined training neural framework for ADC and T2W images. They also introduced a 3D sliding window technique that keeps the 2D domain's complexity while using the 3D situational geographical data supplied by the MRI dataset. Yoo, C et al. [28] created Focal Net, a one-of-a-kind multi-class CNN for detecting prostate cells and providing score-based aggressiveness prediction. To identify colorectal cancer, Zhao, Zet al. [29] employed feature extraction packet prioritization to construct a picture matrix from an image. Elkoushy, M. A.et al. [30] combined the extended short attention span network with latent structure to diagnose diseases despite the presence of hand-made features using deep learning. This method uses a modular grading scale to eliminate unneeded tissue slices, then feeds the image block into a proprietary Neuron to create a map of the world that can indicate colorectal cancer. A new approach for detecting and subdividing testicular cancerous cells was described, that combines morphological and occurrence branching and employs focus components to effectively blend internally and externally picture information.

3. METHODOLOGY

In the previous few decades, there have been significant advancements in neuro imaging techniques, including technical innovations in data gathering and the development of innovative models. PS is the most neurodegenerative disorder among neurodegenerative illnesses. Modern MRI techniques are used to evaluate volume deficits in regional volume deficits associated with structural brain abnormalities connected to the development and progression of PS. A curvelet transform feature extraction approach was utilised to extract features from the PANDA dataset. The IPCA is used to reduce dimensionality through the usage of IPCA-bagging and IPCA-boosting classifiers.

3.1 Prostate cancer grade Assessment (PANDA) Dataset

The Data that had been used here has been obtained from a database of PANDA (https://www.kaggle.com/c/prostate-cancer). In the year 2013, PANDA was commenced by National Institute on Aging (NIA), and the National Institute of Biomedical Imaging and Bioengineering (NIBIB), with that of the Food and Drug Administration (FDA), with certain non-profit organizations and also certain private pharmaceutical companies as a $60 million, in a 5-year public–private partnership. PANDA's major goal is to use MRI, positron emission tomography, and several types of biological markers, as well as clinical and neuropsychological tests, to track the evolution of MCI and early PS. If the early onset and progression of the disease condition is identified at an early stage, it might help the researchers/clinicians to start with specific treatment through which the time and cost of clinical experiments can be controlled.

3.2. The Proposed Method

In prostate cancer, feature extraction is crucial. IPCA is an incremental variation of the well-known unsupervised IPCA dimensionality reduction technique. To begin, IPCA is a statistical approach for analysing n-dimensional data. IPCA determines main dimensions by observing correspondence between different dimensions and determining which dimensions have the most fluctuation in the data. The IPCA basis dimensions or vectors are in the direction of the training vectors' greatest variance. These basis vectors are known as eigenvectors, and they can be considered as eigen pictures because they are further specified in the image space.

The average image is the first eigenimage, and the remainder of the eigenimages are variations on that average image. Each of the eigen images might be considered a feature. Because an image of prostate cancer can be concisely described by a feature vector with a few items, the Eigen image technique is useful. This is a template matching method that outperforms attribute-based methods. This approach, however, has certain drawbacks, including susceptibility to image circumstances such as background noise, lighting, object occlusion, scaling, and shifting of images. Additionally, the standard IPCA technique, preparatory to addressing the problem, computes eigen vectors and eigen values for a sample covariance matrix obtained from a well-known
given image data matrix, which must also be accessible.

A covariance matrix \( \mathbf{C} \) is displayed by the eigenvector \( \mathbf{a} \) and the eigenvalue \( \mathbf{CM} \) in Equation (1):

\[
\mathbf{C} = \mathbf{a} \mathbf{x} \mathbf{a}^T
\]  

(1)

By replacing (1) an unknown \( \mathbf{CM} \) with a sample covariance matrix and \( z = x \), we get (2) the equation below:

\[
\sum_{k=1}^{n} (\mathbf{x}^T \mathbf{a})^k z^{n-k}
\]

(2)

Where \( z(n) \) represents the n-th z-step following n image inputs.

Since \( \|x\| \) and \( x = z/\|z\| \), \( x \) I has been fixed to \( z(1) \). By approximating \( z(i) \) based on the prior value of the \( z \). By plugging in Eq. (2), we get Eq. (10).

\[
\sum_{k=0}^{n} z(1)^k z^{n-k}
\]

(4)

The eigenvectors are recomputed and re-listed in descending order, along with the corresponding eigen value, whenever a new image is presented. Discriminant vectors are derived from characteristics extracted from images of the face, ear, and hand veins using this method. Then, these feature vectors are concatenated which is readily available for the next stage.

Kennedy and Eberhart [13] introduced PSO in 1995 as a population-based evolutionary algorithm. The social behaviour of a swarm is outlined in accordance with sociopsychological principles. At the outset of the evolution process, we define a population consisting of individual candidates as some arbitrary speculations of the problem and its candidate solutions. Following this, an iterative process should be established to further refine these potential answers. After repeatedly assessing the viability of candidate solutions, the particles recall the spots where they were most successful. By "local best," we mean the optimal answer proposed by a given candidate. Each individual particle broadcasts this data to its

Figure 1. Proposed architecture Diagram

As a result, a different method is required, which led to the use of incremental PCA in this study. IPCA is a recursive algorithm that constantly updates the eigen vectors when new images are introduced.

Figure 2. Proposed system Diagram
neighbours. Further, it is made clear where in the neighbourhood people have had success. These successful outcomes direct the steps taken while traversing the search space. At the end of a trial, the population tends to converge on a globally optimal solution. PSO is a more recent evolution algorithm than GAs and genetic programmes. Yet PSO can be used to efficiently optimise a wide range of parameter values. The PSO has been implemented in numerous optimization and recognition tasks with positive results [14].

The PSO offers a number of advantages, including [15] the following: Since it is based on intelligence, the search can be done at the speed of a particle and can be put to use in both science and technology. Only the hopeful particle, which has survived through many generations, can pass on its knowledge to its progeny. No effort is made to determine overlap or mutation. The PSO calculation is simple and easy to carry out in comparison to other developing computations. When using the correct code number, the PSO verifies the result by itself. The PSO just needs less lines of code than the rest. Additionally, it's really simple to put into action.

The partial optimism and scattering-related issues that PSO can't solve are two of the algorithm's drawbacks. Since this is the case, regulating its velocity and course is less precise. The rules for how particles in an energy field move and how that field is solved are not problems amenable to this method because they require a coordinate system.

3.3 Adaboost

As a result of Adaboost's greedy combining of multiple weak learners, a powerful classifier can be created. It picks the least inaccurate weak classifier and discards the rest. Gradient boosting takes a probabilistic approach to component selection, which allows it to handle difficult problems with many features and learn to generate a sparse classification rule with fewer features. However, this might also be a downside due to the Adaboost's greedy character, which limits optimization to just reducing the error and increasing the margin with regard to the selected feature. The biggest problem is that it tends to group all the weak kids into one big one.

3.4 Proposed IPCA-PSO-Adaboost Algorithm

In order to improve PSO's convergence in the presence of high-dimensional problems, the IPCA PSO algorithm was created.

<table>
<thead>
<tr>
<th>IPCAPSO-Adaboost algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initialization</td>
</tr>
<tr>
<td>(a) Commence swarming with initialization</td>
</tr>
<tr>
<td>(b) Covariance matrix initialization</td>
</tr>
<tr>
<td>(c) Figure out the primary components</td>
</tr>
<tr>
<td>(d) Trace individual particles in a three-dimensional Z-plane</td>
</tr>
<tr>
<td>2. Particles for Fly compacts</td>
</tr>
<tr>
<td>(a) Particles in X-space that are in motion</td>
</tr>
<tr>
<td>(b) Particles in Z-space that are in motion</td>
</tr>
<tr>
<td>i. Transpose the coordinates of the newly discovered Z-space locations into X-space</td>
</tr>
<tr>
<td>3. Create a formula using the weights of 2.a and 2.b.i</td>
</tr>
<tr>
<td>4. Best ever for each other</td>
</tr>
<tr>
<td>(a) Update Pbest</td>
</tr>
<tr>
<td>(b) Updating the p-best covariance:</td>
</tr>
<tr>
<td>i. incorporate the new weighted Pbest position into the covariance matrix.</td>
</tr>
<tr>
<td>ii. Compute the PCA again</td>
</tr>
<tr>
<td>iii. convert the positions and velocities in X space to Z space as of right now</td>
</tr>
<tr>
<td>5. A new standard of excellence in the globe</td>
</tr>
<tr>
<td>(a) Improvements to Gbest</td>
</tr>
<tr>
<td>(b) Covariance matrix updating using the gbest method:</td>
</tr>
<tr>
<td>i. Recalculate the covariance matrix, taking into account the revised weighted Gbest position</td>
</tr>
<tr>
<td>ii. Adjust the major components</td>
</tr>
<tr>
<td>iii. Convert the existing positions and velocities from X space to Z space</td>
</tr>
<tr>
<td>6. Particles Once More Fly</td>
</tr>
<tr>
<td>7. Scoring the Adaboost t Boosting Potential</td>
</tr>
</tbody>
</table>

Implementation of the IPCA-PSO-Ad boost Algorithm

Since the greedy algorithm only chooses solutions that meet the greedy criterion, it may complete its task quickly and easily. However, the greedy solution only comes close to the global solution on rare occasions. The PSO algorithm takes into account the best particles in the population. This is because, under PSO, all of the other particles will gradually shift to take up the best possible position around the particle that has already been chosen. When the particle's position is optimised, the PSO can find a solution to the
problem more quickly. The PSO's starting population has been exposed to the greedy solution in the hopes of improving the system overall. The proposed IPCA-PSO-Adaboost process is shown in Figure 1.

Flowchart 1: The IPCA-PSO-Adobest Boost Proposal

4 RESULTS AND DISCUSSION

Adaboost classifier optimization is performed using a hybrid IPCA-PSO algorithm, with the number of classification trees (from 25 to 300) and tree depth (from 1 to 6) being tuned. The proportion of correct classifications, the proportion of false negatives, and the overall false positive rate are displayed in Table 1 and figures 2–4 for several methods including IPCA–Adaboost, IPCA–PSO–Adaboost, and IPCA–PSOG Adaboost.

Table 1 Summary of Results

<table>
<thead>
<tr>
<th></th>
<th>Prostate image-IPCA-Adaboost</th>
<th>Prostate image-IPCA - PSO-Adaboost</th>
<th>Prostate image-IPCA – PSOG Adaboost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy of Classification</td>
<td>85.53</td>
<td>91.49</td>
<td>93.45</td>
</tr>
<tr>
<td>Normal True Positive Rate</td>
<td>91.43</td>
<td>95.43</td>
<td>96.57</td>
</tr>
<tr>
<td>The MCI True Positive Rate</td>
<td>73.75</td>
<td>82.5</td>
<td>86.25</td>
</tr>
<tr>
<td>True Positive Rate - Prostate</td>
<td>57.5</td>
<td>75.35</td>
<td>63.18</td>
</tr>
<tr>
<td>True Negative Rate – Normal</td>
<td>81.19</td>
<td>88.07</td>
<td>88.63</td>
</tr>
<tr>
<td>True Negative Rate – MCI</td>
<td>93.97</td>
<td>96.55</td>
<td>97.37</td>
</tr>
<tr>
<td>True Negative Rate - Prostate</td>
<td>93.35</td>
<td>96.62</td>
<td>98.07</td>
</tr>
</tbody>
</table>

Figure 2: Accuracy of Classification with Principal Component Analysis and PSOG Adaboost

Figure 2 shows that compared to IPCA-Adaboost (8.79%) and IPCA-PSO-Adaboost (2.1%), the classification accuracy of IPCA-PSOG Adaboost is significantly greater.
Figure 3 shows that compared to IPCA-Adaboost, which has a false-positive rate of 5.47% in the typical population, IPCA-PSOG Adaboost has a true-positive rate of 1.19%. In the instance of MCI, the IPCA-PSOG Adaboost has a higher true positive rate than the PCA-Adaboost (15.63%) and the IPCA-PSO-Adaboost (4.44%), whereas in the case of Prostate, the IPCA-PSOG Adaboost has a higher true positive rate than the PCA-Adaboost (32.73%) and the PCA-Adaboost (6.45%).

Figure 4 shows that the genuine negative rate for IPCA-PSOG Adaboost is 8.73% greater than that of IPCA-Adaboost in the typical population, whereas the true negative rate for IPCA-PSO-Adaboost is just 0.6% higher. The true negative rate for IPCA-PSOG Adaboost is 3.6% higher than that of PCA-Adaboost and 0.85% higher than that of IPCA-PSO-Adaboost for MCI, and 4.93% higher than that of IPCA-Adaboost and 1.49 percentage points higher than that of IPCA-Adaboost for Prostate.

5. CONCLUSION
Prostate cancer is the most common kind of cancer, accounting for 60%-80% of all occurrences; it is a fatal, progressive neurological disease for which there is now no treatment. Different research groups have proposed a variety of detection approaches using MRI. As well as being efficient, the greedy algorithm is also quite easy to implement. In this paper, a greedy algorithm that combines elements of PSO and a hybrid of the two is proposed. The research also recommends an efficient IPCA-PSOG-Adaboost hybrid. PSOG Adaboost outperforms Adaboost and PSO-Adaboost in terms of classification accuracy by 8.35% and 2.14%, respectively. IPCA-Adaboost improves classification accuracy by 8.79%, but IPCA-PSO-Adaboost only improves it by 2.1%. IPCA-PSOG Adaboost has a true positive rate of 1.19 percent, whereas IPCA-Adaboost has a rate of 5.47 percent in the general population. IPCA-PSOG Adaboost's actual positive rate for PCA-Adaboost is 32.7 percent higher than IPCA-6.45 Adaboost's percent, and IPCA-true Adaboost's positive rate for MCI is 15.63 percent higher than IPCA-and Adaboost's 4.44 percent higher than IPCA-PSO-Adaboost. Under normal circumstances, the IPCA-PSOG Adaboost has a true negative rate that is 0.6% greater than that of the PCA-Adaboost. The true negative rate of the IPCA-PSOG Adaboost is 3.6% higher than that of the PCA-Adaboost and 0.85% higher than that of the IPCA-PSO-Adaboost for MCI, and 4.93% higher than that of the IPCA-Adaboost and 1.49% higher than that of the IPCA-Adaboost for Prostate.

REFERENCES


