INJ INTERNATIONAL NEUROUROLOGY JOURNAL

Statistical Feature Analysis and Machine Learning based Classification of Lower Back Pain Using Thermal Images

P. Praveen*[1], Mallikarjunaswamy M.S.[1] & Chandrashekara S[2]

[1]Dept of Electronics & Instrumentation, Sri Jayachamarajendra College of Engg., JSS Science and Technology University, Mysuru, India

[2]ChanRe Rheumatology & Immunology Center & Research, Bengaluru, India

Abstract

Nearly one-quarter of the world's population suffers from low back pain (LBP). LBP can originate from several different areas of the body, including the nerves, spinal cord, discs, bones, and tendons of the lumbar spine. Getting a proper diagnosis of LBP at an early stage is the initial step toward a speedy and complete recovery. Though much effort and money have been invested in LBP research approaches, successful diagnosis remains an important objective, and LBP remains to be a major reason for concern in primary healthcare. The inability of conventional medical images to identify the LBP was one of the reasons for the above problem. This research intends to provide a five-step process for automatic LBP detection using thermal images. At first, the thermal images of both healthy and LBP individuals are taken. Second, the images are analysed by employing Grey Level Co-occurrence Matrix (GLCM) and Grey Level Run-Length Matrix (GRLM) techniques to get a total of 18 features. Third, using the Particle Swarm Optimisation (PSO) method, the most crucial of the 18 features is determined. Fourth, ML model training involves using both the raw feature data and the optimized feature data. Finally, the performance of the ML model is assessed using both data before and after feature selection to determine the optimal method for LBP detection.

Keywords: Thermal Images, Low Back Pain, Feature Extraction, Optimization, Machine Learning.

Introduction

LBP-related disability is a major public health problem worldwide. Adults aged 10-24 and 50-74 lose an average of an entire year of health due to LBP every year [1]. Therefore, healthcare providers treating patients with LBP need more knowledge about the nuances of this musculoskeletal disorder. Fear of mobility and a reduction in everyday and social activities are prominent symptoms of LBP [2]. The prevalence of LBP peaks in people's thirties and rises steadily through their fifties and sixties before levelling off in people's seventies. There is a significant public health and economic burden associated with LBP since it affects the working population and is the leading cause of missing workdays, especially among those whose jobs require more strenuous physical activity [3]. The estimated one-year recurrence rate for LBP is between 24 and 80 percent [4], making preventative an attractive option. Ninety percent of those who suffer from LBP are classed as having non-specific LBP [5] which implies that no definite cause can be discovered and therapy instead focuses on pain reduction and its effects. Complications of LBP can be caused by several structures in the spinal column. Less than 5% of those seeking primary care for LBP have a significant systemic illness. Although clinicians may classify patients with back pain as having a muscle spasm, sacroiliac discomfort, or trigger point, there is no consensus on the reliability of these labels. It is not always possible to pinpoint structural abnormalities in the spine as the root cause of a patient's LBP. Despite the patient's continued complaints, imaging tests including X-rays, Computed Tomography (CT), and Magnetic Resonance Imaging (MRI) frequently show nothing wrong. In these instances, it is crucial to know the root of the discomfort and how to effectively treat it. Symptoms, clinical testing, and state-of-the-art diagnostic tools like CT and MRI are all useful in detecting LBP [6] when the pain is caused by structural abnormalities in the spine. A common misconception is that these methods are too costly or time-consuming to be practical in daily life. A standard outcome measure is essential for quantifying the issue, even though clinical testing is necessary for diagnosing the cause of back pain. While many methods have been developed to evaluate this phenomenon, infrared thermal imaging has emerged as the industry standard.

Thermography, or thermal imaging, has been widely employed in the industry over the past few decades for a wide range of objectives, including detection, monitoring, and prediction across a wide range of disciplines, from engineering to medical and biological studies. One of the diagnostic parameters utilized in medicine is a patient's internal body temperature. The fact that a person's core temperature may be an indicator of health problems is also generally established. The therapies used in physiotherapy are quite successful. However, it is tough to assess efficacy because it is so dependent on patient symptoms. In light of these findings, it is postulated that infrared thermal imaging can serve as a reliable method of evaluating LBP in its preliminary stages. So, LBP auto-identification with thermal images of the individual were employ. Some of the works on pain prediction using ML and Deep Learning (DL) are detailed below.

Rim et.al. [7] predicted the degree of lumbar radiculopathy in different groups using digital infrared thermographic imaging (DITI) and a machine-learning (ML) technique. The DITI data included both people with radiculopathy from herniated lumbar discs and people who did not have disc problems as controls. To test the accuracy of the model, 1,000 patients were split evenly between 73% training data and 37% test data. Applying the ML approach to a pain-severity categorization based on thermographic pictures would help doctors better treat lumbosacral radiculopathy and evaluate the efficacy of their treatments for the ailment. The research [8] aimed to create a Natural Language Processing (NLP) system that could spot X-ray, CT, and MRI evidence of LBP. An NLP system using a rule-based approach was trained and evaluated using medical imaging records from LBP patients. The radiology department's free-text reports can be analysed by an NLP system. The unstructured clinical data originally included in radiology reports has been converted into a more usable format. Clinical data is used systematically in data-driven investigations of LBP. With the groundwork laid by this study, future clinical studies can proceed toward their goal of autonomously extracting clinical data from free-text radiological reports. In one of the research articles [9], the researchers detail their strategies for gathering curated imaging data within legally permissible parameters. In-depth instructions on data collection and tagging are provided. Next, detail how ML could be applied to the gathered data to create an objective imaging biomarker. The final objective is to make it possible to use validated ML models trained on imaging data as objective biomarkers for the clinical diagnosis of lumbar radiculopathy and the guidance of related treatment protocols.

A review [10] shows that automatic pain assessment can benefit greatly from multimodal techniques, especially in clinical situations, and that substantial gains can be shown when temporal exploitation of modalities is added. It hints at improved DL structures and techniques. It also offers tips for creating objective and understandable results by implementing rigorous evaluation processes and interpretation approaches. This research also delves into the shortcomings of currently available pain datasets in terms of their ability to back up the development, validation, and deployment of DL models as decision-support instruments in practice. The author [11] used an ML decision model informed by clinical biochemistry to make predictions about joint pain (the dependent variable). About 650 people sought out orthopaedic care for conditions such as joint swelling and myalgia. Used supervised learning to train, test, and cross-validate a decision tree. Age, gender, uric acid, and C-reactive protein were used in the model's assessment. Diagnoses of joint discomfort were given to 44% of patients. When the decision tree model was trained and evaluated, it produced good results. Arthralgia was strongly associated with uric acid levels. Early detection of joint pain with ML can help avoid more serious orthopaedic issues in the future. In the publication [12], the authors provide a DL system that can extract and classify features from physiological data without the need for expert medical knowledge. The authors suggest using multiple dimensions of context to differentiate painful from nonpainful physiological signals. Using information from both the Emopain 2021 and Part A of the BioVid Heat Pain databases, it shows that multi-level context information is superior to uni-level context information. The suggested method capitalizes on the superior performance of DL over more conventional methods by applying it to physiological inputs.

Materials and Methods

The theoretical concept of important algorithms used in this research to implement automatic detection of LBP are discussed in this section.

Feature Extraction

Feature Extraction (FE) is the procedure of examining an image's texture [13]. The findings advance the understanding of texture and object behaviour detection. The steps taken to extract statistical features using the GLCM and run length features using the GRLM are outlined below.

ISSN:2093-4777 | E-ISSN:2093-6931
DOI: 10.5123/inj.2023.4.inj53
Vol. 27 Iss. 4 (2023)

GLCM: The GLCM technique's grey-level co-occurrence data are utilized to extract 'texture features' and maintain a relationship between pixels. This method is based on the 'conditional probability density functions $'p\ (i, j \mid d,)'$ and chosen directions of $'S = 0°, 45°, 90°, 135°'$, etc., as well as distances d ranging from 1 to 5. The function $'p\ (i, j \mid d,)'$ [14] denotes the probability that two grey level pixels $'i'$ and $'j'$ are situated at the same inter-sample distance 'd' and in the same direction $'S'$. GLCM is concerned with entropy, correlation, energy, variance, dissimilarity, contrast, average, homogeneity, and cluster shade.

GLRM: A matrix represents the geometric characteristics of GRLM. It returns a number that represents the average pixel brightness along the Run length-specified axis. It's a two-dimensional item [15]. The $'j'$ number of components and the $'i'$ intensity in the provided directions reflect each constituent in this scenario. The 'run-length matrix' calculates the frequency of a run for each grayscale value. Then it examines whether three consecutive pixels have the same intensity value, then four, and so on. The length of a run is represented by the number of individual pixels in it. GLRM collects features like long runs, short runs, run ratio, run length, low grey level runs, high grey level runs, entropy, short runs at low grey levels, and short runs at high grey levels.

Feature Selection

The section on Feature Selection discusses the optimizers for successful feature selection. This article introduces the Particle swarm optimization (PSO) algorithm and delves into its inner workings. Feature selection must be changed into a more trustworthy and acceptable form before the classifier can classify the LBP category.

PSO is an optimization method that employs a random selection process, similar to that of flocks of birds or schools of fish. Each bird in the swarm modifies its hunting model based on what it has learned as part of the swarm's collaborative effort to find food [16]. The PSO algorithm draws inspiration from evolutionary algorithms as well as swarm artificial life systems. The constituent particles of the swarm, known as "Birds," freely roam the search space's numerous dimensions. Each particle moves at its own pace and establishes its place along the way. Each particle is updated, affecting the aggregate population. Because the swarm arrangement is self-motivated, particles gather at the maximum value of the target function. PSO comprises the following steps:

Step 1: The PSO's correct searching capability comes from the inclusion of variables for each particle. The collision between particles is ignored in the weighting factors c_1, c_2. To avoid a collision, the range in which the particle's i, velocity V, and random number R are protected must be updated.

$$\vec{V}_i = W\vec{V}_i + c_1 R_1\left(\vec{P}_{i,best} - \vec{P}_i\right) + c_2 R_2\left(\vec{g}_{i,best} - \vec{P}_i\right) \quad \text{-----(1)}$$

Where,

$\vec{V}_i$ → Particle i's velocity
$\vec{P}_{i,best}$ → Particle i's best position
$\vec{g}_{i,best}$ → Particle i's best location
W → Control parameter
R_1, R_2 → Random numbers ϵ 0,1
c_1, c_2 → Learning factor.

Step 2: Position updates - Particle positions are updated due to the interval between iterations, as shown in the equation below.

$$\vec{P}_i = \vec{P}_i + \vec{V}_i \qquad \text{-----(2)}$$

Verify that $\vec{P}_i$, is within acceptable limits after a refresh.

Step 3: Memory update: Use the below equations to modify P$\vec{P}_{i,best}$ and $\vec{g}_{i,best}$,

$$\vec{P}_{i,best} = \vec{P}_i, if\ f\left(\vec{P}_i\right) > f(\vec{P}_{i,best}) \qquad \text{-----(3)}$$

$$\vec{g}_{i,best} = \vec{g}_i, if\ f(\vec{g}_i) > f(\vec{g}_{i,best}) \qquad \text{-----(4)}$$

Step 4: Destination Verification: The procedure repeats steps 2 and 3 an infinite number of times, or until a user-specified number of end states have been reached. The estimation and the result are given.

In PSO algorithms, the fitness values are ignored. When the population size is very large, this provides a significant computational benefit compared to alternative approaches. The calculations for velocity and position use arithmetic operations on real numbers.

INJ INTERNATIONAL NEUROUROLOGY JOURNAL

Classification

The LR classifier, the NB classifier, and the XGBoost algorithm have all been tested and compared. List each classifier's supporting theory and mathematical equation given below.

Logistic regression: When it comes to assessing LBP, LR is one of the most reliable statistical methods accessible. This model uses a binary dependent variable and allows the independent variables to be measured on a variety of scales, including ordinal, nominal, and ratio. Nonlinearity describes the connection between the dependent and the independent variable [17]. This connection is expected to be more complicated than the linear models [18], despite the common misconception that LR is a particular form of a generalized linear model. Since the dependent variable is a binary one (with and without LBP), the conditional distribution is a Bernoulli rather than a Gaussian. The following equation illustrates the connection between occurrence and its dependence on numerous variables in LR analysis.

$$p = \frac{1}{1+e^{-z}} \qquad \text{-----(5)}$$

Where,

p → Probability of an LBP occurring ϵ 0,1,

z → Linear combination of several variables associated with LBP. To analyze data using logistic regression, an equation of the form:

$$z = b_0 + b_1x_1 + b_2x_2 + \cdots b_nx_n \qquad \text{-----(6)}$$

Where,

b_0 → Intercept of the model,

b_i → Slope coefficients ϵ $(i = 0, 1, 2, \ldots, n)$,

x_i → Independent variables ϵ $(i = 0, 1, 2, \ldots, n)$.

As the linear model, a logistic regression of LBP presence or absence (on the independent variables) is constructed.

Naïve Bayes: Using a probability model established in terms of Bayes's theory, Bayesian Classification [19] determines the likelihood that a new observation fits into a given category. This method posits that classification may be approximated by computing the posterior probability and conditional probability density function, which are derived from a large amount of training data characterized by several characteristics, and then evaluates the prior probability of each category based on this data. It was possible to derive an equation for the posterior probability:

$$P\left(\frac{C_j}{X}\right) = \frac{P\left(\frac{X}{C_j}\right) * P(C_j)}{P(X)} \qquad \text{-----(7)}$$

Where,

X → Unknown observation

$P(X)$ → Prior probability of X.

C_j → Category

$P\left(\frac{C_j}{X}\right)$ → Posteriori probability (Probability X belongs to the category C_j)

$P\left(\frac{X}{C_j}\right)$ → Probability, given C_j, an unknown observation belongs to this category,

$P(C_j)$ → Prior probability X to be observed in C_j,

Naive Bayes is a straightforward approach to Bayesian classification that is used in situations when all of the variables used to characterize the training data are uncorrelated and equally important to the classification task [20]. The following equation can be used to determine the probability $P\left(\frac{x_j}{C_j}\right)$ under the conditional independence presumptions,

$$P\left(\frac{X}{C_j}\right) = \prod_{i=1}^{k} P\left(\frac{x_j}{C_j}\right) \qquad \text{-----(8)}$$

The numerical attribute values within each category follow a normal distribution concerning standard deviation and mean, so that the conditional probability can be computed as follows:

$$P\left(\frac{x_j}{c_j}\right) = \frac{1}{\sqrt{2\Pi\delta}} e^{\frac{-(x_j-\mu)^2}{2\delta^2}} \quad \text{-----(9)}$$

Where,

μ → Mean

δ → Standard deviation.

A learning algorithm will often sift through a set of possible hypotheses in search of the most likely one. To find the greatest posterior hypothesis, apply the following equation. The Boolean result of the analysis is y_j, which indicates whether or not the prediction is for LBP given an incidence of k variables linked with LBP. The forecast is made for the category with the highest posterior probability using the formula below.

$$y_j = argmax\ P\ (y_j) \prod_{i=1}^{k} P\left(\frac{x_j}{y_j}\right) \quad \text{-----(10)}$$

XGBoost: The XGBoost is a unified ML technique for handling supervised learning issues; it uses a gradient ascending framework and is tree-based for classification. "Ensemble learning" is the process of forecasting a dataset using a large number of relatively weak classifiers and then combining their anticipated results using a specified technique [21]. It enhances the standard gradient boosting decision tree (GBDT) technique's efficiency, accuracy, and scalability. In contrast to GBDT, XGBoost employs regularisation in the Loss Function (LF) to determine the goal:

$$J(\theta) = L(\theta) + \Omega(\theta) \quad \text{-----(11)}$$

Where

$$L(\theta) = (\hat{y}_j, y_j) \quad \text{-----(12)}$$

$$\Omega(\theta) = \gamma T + \frac{1}{2}\lambda||w||_2 \quad \text{-----(13)}$$

As shown in (11), the objective function includes two components: $L(\theta)$and $\Omega(\theta)$

Where

θ → A set of parameters that can be obtained from the given data.

$L(\theta)$ → Differentiable convex LF (Difference between the actual and the forecast outcome)

The logistic and the mean square LF are two popular LF. The LF utilized in this paper is mean square [22].

Ω(θ) → Regularisation is used to penalize method complexity.

T → Total number of nodes in the tree.

R → Learning rate ϵ 0,1.

λ → regularisation parameter

w → Fraction of leaves.

Because (11) takes a function as a parameter and thus cannot be optimized using the classic Euclid technique, XGBoost constructs the regression tree and adds an optimization element at each iteration. As an outcome in the t^{th} iteration, the objective function is stated in the following way:

$$J^{(t)} = \sum_j^N l\left(y_j, \hat{y}_j^{(t-1)} + f_t(x_j)\right) + \Omega(f_t) \quad \text{-----(14)}$$

Additionally, XGBoost allows for parallel processing. It optimizes the method by choosing the optimal split point and processing elements in parallel during enumeration.

Results

The research on LBP detection by thermal images are discussed in this section from data collection to model evaluation.

Data and its processing

For this research, the data are collected by own. Figure 1 shows a person's sample with and without LBP.

DOI: 10.5123/inj.2023.4.inj53
ISSN:2093-4777 | E-ISSN:2093-6931
Vol. 27 Iss. 4 (2023)

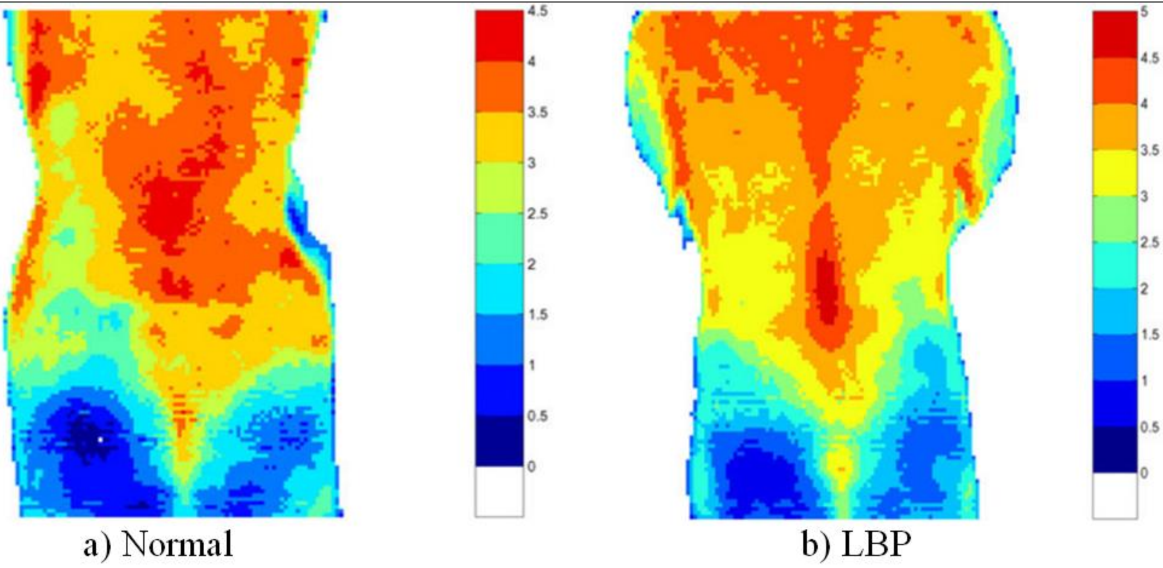


Fig 1. LBP sample images

The experiments were conducted in the morning, in a climate-controlled room with temperatures between 23^{0}C and 24^{0}C and humidity between 45 and 55% with no exposure to natural light. Participants waited 5-10 minutes before the commencement of the experiment so that they could adapt to the temperature of the room. The next step was to have them sit in a chair with their backs uncovered. In this research, images of the subjects' backs were taken using a Fluke PTi 120 thermal imaging camera. The selected method is a portable and low-cost thermal imaging device that can detect signals between 8 and 14 micrometres in wavelength. In addition, it has a temperature resolution of 0.1^{0}C and can detect temperatures from -20^{0}C to 400^{0}C. totally out of 400 images collected, 200 images are LBP and the remaining 200 are non-LBP images. And the images are partitioned for training and testing as 80% and 20%.

Analysis of Features

When the input data set of the algorithm becomes too large, it should be translated into a more manageable dimension. The procedure of converting an input image into a standard set of features is known as FE. The FE approach was used on the segmented images to convert pixel groupings into numerical data. The values of all nine extracted features using GLCM on both LBP and non-LBP are given in Figure 2.a to 2.i. In all the below figures, the feature value of LBP is represented by an orange plot, and the normal is denoted by a blue plot. In the figure, the number of samples is represented on the x-axis, and feature values are given on the y-axis.

DOI: 10.5123/inj.2023.4.inj53
ISSN:2093-4777 | E-ISSN:2093-6931
Vol. 27 Iss. 4 (2023)

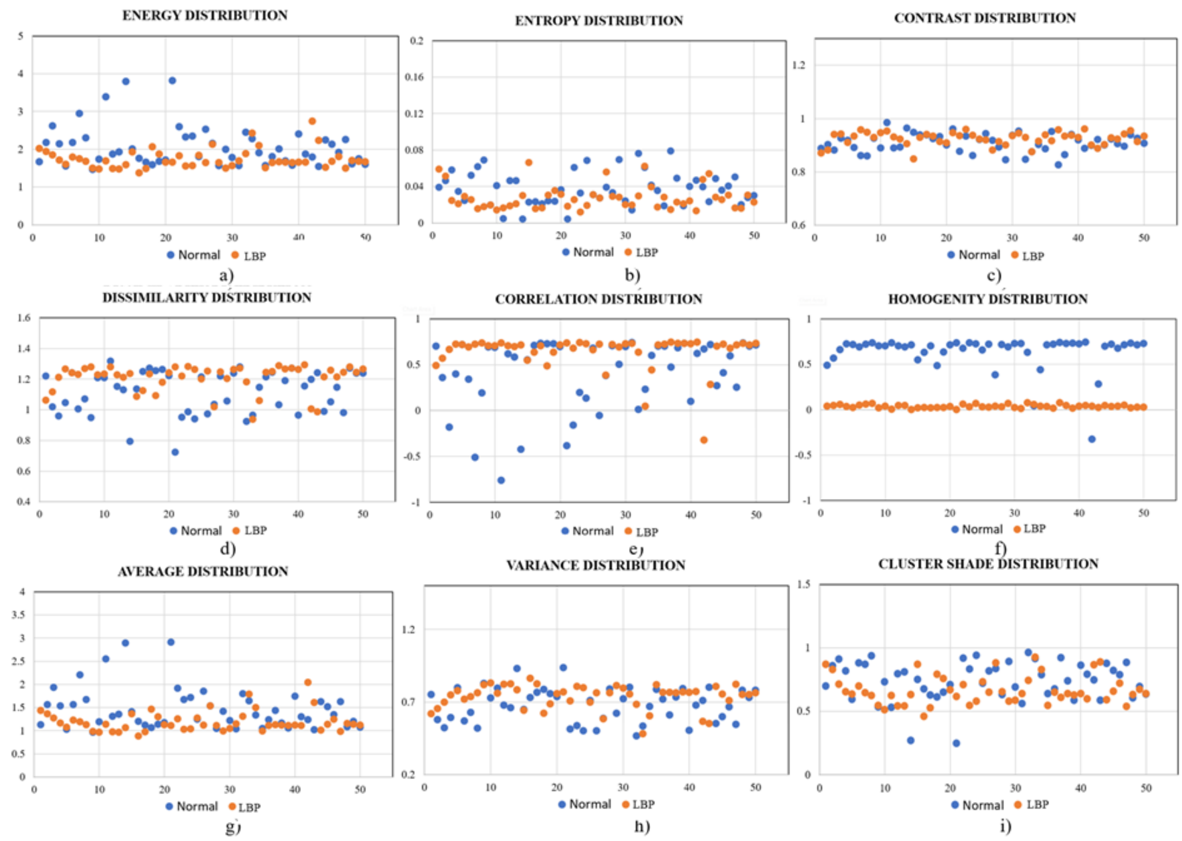
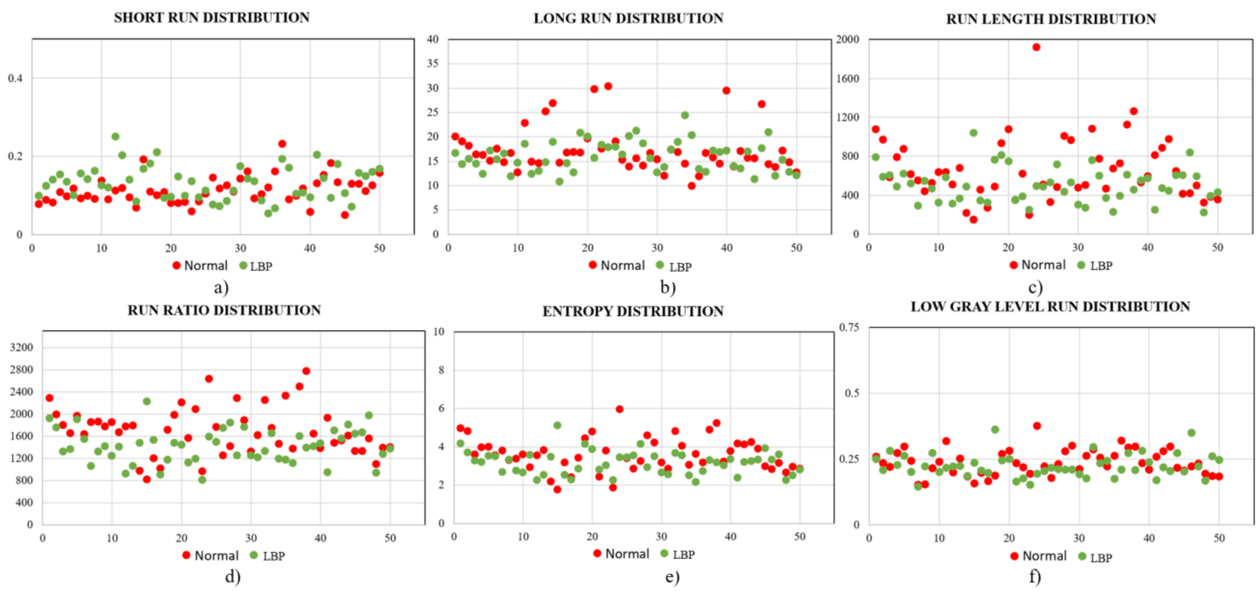


Fig. 2. GLCM Features

For both LBP and non-LBP, the values of all nine features derived using GLRM are shown in figures 3.a through 3.i. The LBP feature value is depicted in green, while the normal distribution is shown in red, in each of the following figures. The y-axis of the graphic depicts feature values, while the x-axis shows the number of samples.



INJ INTERNATIONAL NEUROLOGY JOURNAL

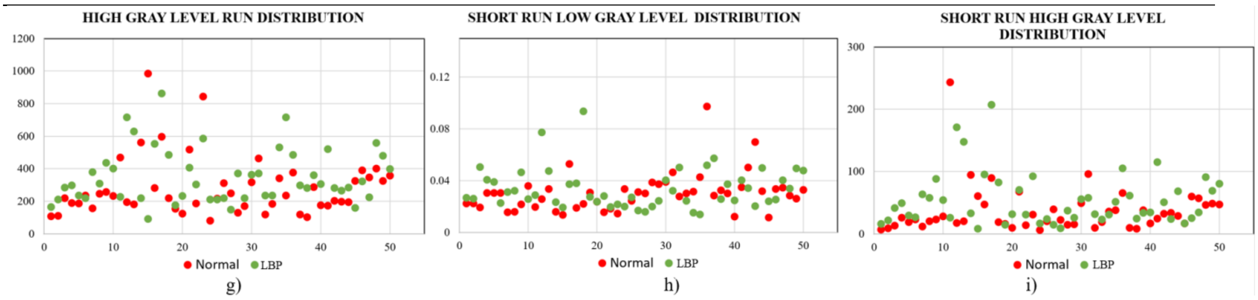


Fig. 3. GLRM Features

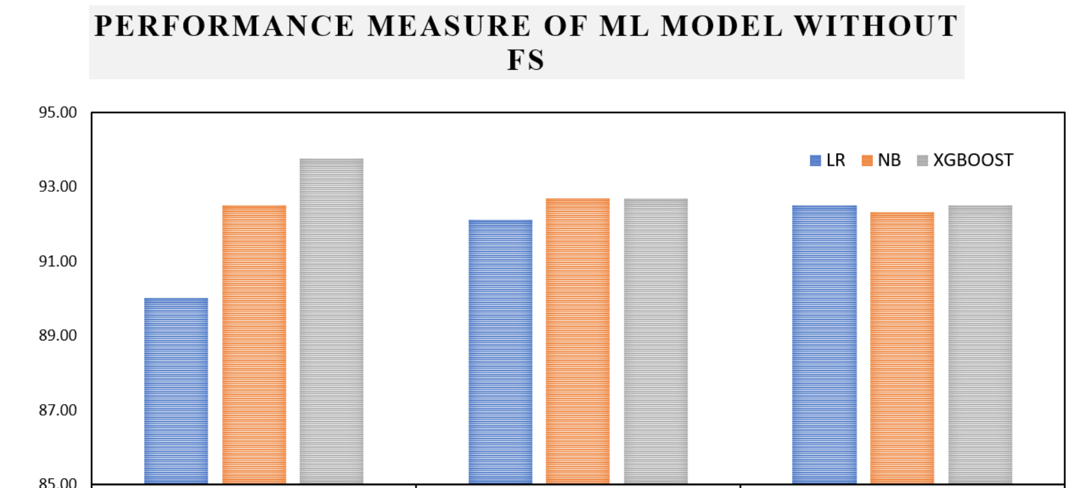
ML Model comparison

The ML-based automatic identification of LBP using thermal images is detailed in this section. First, the data were collected from thermal cameras, and the features were extracted from the data using GLCM and GLRM approaches. A total of 18 features are generated, GLCM gives 9 features and GLRM gives nine features. From the 18 features, important features are selected using the optimization technique called PSO. The features of low back thermal images before and after plying PSO are given to the ML model for training. Next, the ML model is tested using the 20% data.

Discussions

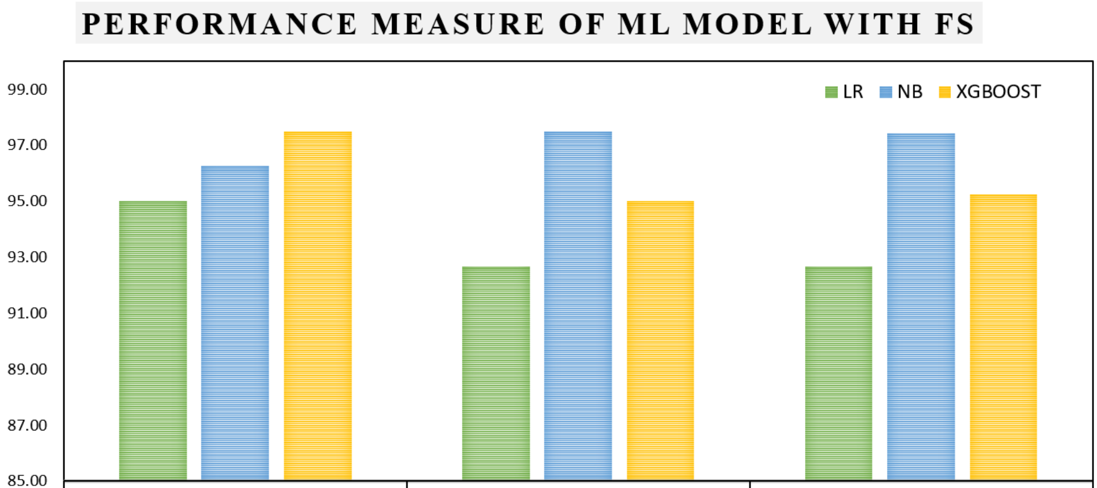
The outcome of the ML model using data before applying the optimization approach is given in Figure 4. For evaluating the ML model, metrics like accuracy, recall, and precision are employed. The XGBoost generates the highest accuracy of 93.75% when compared with the other models such as LR and NB. In this case, the accuracy produced by LR and NB is 90% and 92.5%. The maximum value of recall is 92.68%, which is attained by two models namely XGBoost and NB. The lowest value of 92.11% is the result of LR. Similar to recall, the highest precision is also attained by two models such as LR and XGBoost, whose value is 92.5%. The NB produces minimum precision of 92.31%.

Figure 5 displays the results of an ML model using the optimized feature data. The ML model is assessed using the same criteria as in the previous case. When compared to other models like LR and NB, XGBoost produces the greatest accuracy of 97.5%. Here, LR and NB yield accuracies of 95% and 96.25%, respectively. NB's recall is 97.50%, which is the highest achievable value. XGBoost produces 95%, and LR produces 92.6% in the recall. The NB and LR methods achieve the highest (97.44%) and lowest (92.68%) levels of precision, respectively. The XGBoost algorithm then provides a precision of 95.24%



	ACCURACY	RECALL	PRECISION
LR	90.00	92.11	92.50
NB	92.50	92.68	92.31
XGBOOST	93.75	92.68	92.50

Fig. 4. ML model performance on LBP identification before FS



	ACCURACY	RECALL	PRECISION
LR	95.00	92.68	92.68
NB	96.25	97.50	97.44
XGBOOST	97.50	95.00	95.24

Fig. 5. ML model performance on LBP identification after FS

Conclusion

LBP is not a disease, but rather a set of symptoms whose origins are unclear in the majority of instances, although risk factors have been established. LBP is painful and has huge societal and economic effects. LBP is problematic for surgical and medical therapy options and a leading cause of short-term impairment. In this research, it is shown that the thermal imaging can be used to correctly identify LBP. Thermal images have their features retrieved and selected using optimization methods. Predictions of LBP are made using ML models like LR, NB, and XGBoost. The ML model gives better results on data after feature selection. Three ML models are used and evaluated on feature selection data. The results suggest that NB outperforms the other two ML approaches in terms of recall (97.5%) and precision (97.44%); but, when accuracy is utilized as the evaluation parameter, the XGBoost algorithm demonstrates greater performance (97.5%). These findings point to the promising future of using ML approaches for LBP prediction, and the possibility of selecting the most appropriate algorithm in light of their concerns.

References

1. Abbafati, Cristiana, D. B. Machado, B. Cislaghi, O. M. Salman, Marina Karanikolos, M. McKee, K. M. Abbas et al. "Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019." The Lancet 396, no. 10258 (2020): 1204-1222.
2. André, Maria, and Mari Lundberg. "Thoughts on Pain, Physical Activity, and Body in Patients with Recurrent Low Back Pain and Fear: An Interview Study." Physical Therapy 102, no. 2 (2022): pzab275.
3. Serranheira, F., M. Sousa-Uva, F. Heranz, F. Kovacs, and A. Sousa-Uva. "Low Back Pain (LBP), work and absenteeism." Work 65, no. 2 (2020): 463-469.
4. Hoy, Damian, Paul Brooks, Fiona Blyth, and Rachelle Buchbinder. "The epidemiology of low back pain." Best practice & research Clinical rheumatology 24, no. 6 (2010): 769-781.
5. Bardin, Lynn D., Peter King, and Chris G. Maher. "Diagnostic triage for low back pain: a practical approach for primary care." Medical journal of Australia 206, no. 6 (2017): 268-273.
6. Samuel, Stephen Rajan, CG Shashi Kumar, and G. Arun Maiya. "Infrared thermal imaging as an outcome measure in low back pain." Research Journal of Pharmaceutical, Biological and Chemical Sciences 7, no. 1 (2016): 1806-1810.

7. Rim, Jinu, Seungjun Ryu, Hyunjun Jang, Hoyeol Zhang, and Yongeun Cho. "Machine Learning-Based Pain Severity Classification of Lumbosacral Radiculopathy Using Infrared Thermal Imaging." Applied Sciences 13, no. 6 (2023): 3541.
8. Kim, Yeji, Chanyoung Song, Gyuseon Song, Sol Bi Kim, Hyun-Wook Han, and Inbo Han. "Using Natural Language Processing to Identify Low Back Pain in Imaging Reports." Applied Sciences 12, no. 24 (2022): 12521.
9. Gaonkar, Bilwaj, Kirstin Cook, Bryan Yoo, Banafsheh Salehi, and Luke Macyszyn. "Imaging Biomarker Development for Lower Back Pain Using Machine Learning: How Image Analysis Can Help Back Pain." Biomedical Engineering Technologies: Volume 1 (2022): 623-640.
10. Gkikas, Stefanos, and Manolis Tsiknakis. "Automatic assessment of pain based on deep learning methods: A systematic review." Computer Methods and Programs in Biomedicine 231 (2023): 107365.
11. Shahid, Saman, and Aatir Javaid. "Application of Machine Learning Decision Tree in Diagnosing Joint Pain." In 2022 Medical Technologies Congress (TIPTEKNO), pp. 1-4. IEEE, 2022.
12. Phan, Kim Ngan, Ngumimi Karen Iyortsuun, Sudarshan Pant, Hyung-Jeong Yang, and Soo-Hyung Kim. "Pain Recognition with Physiological Signals Using Multi-Level Context Information." IEEE Access 11 (2023): 20114-20127.
13. George, Shini, and V. Srividhya. "A REVIEW ON TECHNIQUES FOR FEATURE EXTRACTION IN OPINION MINING." Journal of Pharmaceutical Negative Results (2022): 9356-9361.
14. Arora, Kratika, and Ashwani Kumar Aggarwal. "Approaches for image database retrieval based on color, texture, and shape features." In Handbook of research on advanced concepts in real-time image and video processing, pp. 28-50. IGI Global, 2018.
15. Vani Kumari, S., and K. Usha Rani. "Analysis on various feature extraction methods for medical image classification." In Advances in Computational and Bio-Engineering: Proceeding of the International Conference on Computational and Bio Engineering, 2019, Volume 2, pp. 19-31. Springer International Publishing, 2020.
16. Tran, Binh, Bing Xue, and Mengjie Zhang. "Overview of particle swarm optimisation for feature selection in classification." In Simulated Evolution and Learning: 10th International Conference, SEAL 2014, Dunedin, New Zealand, December 15-18, 2014. Proceedings 10, pp. 605-617. Springer International Publishing, 2014.
17. Cheeseman, Peter C., and John C. Stutz. "Bayesian classification (AutocFlass): theory and results." Advances in knowledge discovery and data mining 180 (1996): 153-180.
18. Domingos, Pedro, and Michael Pazzani. "Beyond independence: Conditions for the optimality of the simple bayesian classi er." In Proc. 13th Intl. Conf. Machine Learning, pp. 105-112. 1996.
19. Boateng, Ernest Yeboah, and Daniel A. Abaye. "A review of the logistic regression model with emphasis on medical research." Journal of data analysis and information processing 7, no. 4 (2019): 190-207.
20. LaValley, Michael P. "Logistic regression." Circulation 117, no. 18 (2008): 2395-2399.
21. Chen, Sijie, Wenjing Zhou, Jinghui Tu, Jian Li, Bo Wang, Xiaofei Mo, Geng Tian, Kebo Lv, and Zhijian Huang. "A novel XGBoost method to infer the primary lesion of 20 solid tumour types from Gene expression data." Frontiers in Genetics 12 (2021): 632761.
22. Zhang, Dan, Hua-Dong Chen, Hasan Zulfiqar, Shi-Shi Yuan, Qin-Lai Huang, Zhao-Yue Zhang, and Ke-Jun Deng. "iBLP: an XGBoost-based predictor for identifying bioluminescent proteins." Computational and Mathematical Methods in Medicine 2021 (2021): 1-15.