

Early detection and classification of bone marrow changes in lumbar vertebrae using machine learning techniques

Yasir Hussein Shakir¹, Tiong Siah Kiong^{1,2}, Chai Phing Chen^{1,3}

¹College of Graduate Studies, Universiti Tenaga Nasional (UNITEN), Kajang, Malaysia

²Institute of Sustainable Energy, Universiti Tenaga Nasional (UNITEN), Kajang, Malaysia

³Department Electrical and Electronics Engineering, Universiti Tenaga Nasional (UNITEN), Kajang, Malaysia

Article Info

Article history:

Received Jun 8, 2024

Revised Feb 14, 2025

Accepted Mar 15, 2025

Keywords:

Bone marrow prediction

Dimensionality reduction

Feature extraction

Healthcare

Machine learning

Magnetic resonance imaging

ABSTRACT

Bone marrow changes in lumbar vertebrae (BMCLVB) have emerged as a significant correlation of chronic low back pain (CLBP) severity, especially in patients with comorbid conditions like HIV, osteoporosis, and cancer. Identifying these correlations not only aids governments and health insurance providers but also facilitates early treatment for those at risk. However, challenges lurk due to the unavailability and quality of healthcare data. The collaboration between data science and artificial intelligence, particularly machine learning (ML), has propelled biomedical research forward. So far, accessing and processing hospital and clinical data remains a hurdle. In doing so it aims to provide an opportunity for early intervention and treatment. In addition, the goal of the current study was to overcome data shortcomings using advanced ML techniques to unlock complex magnetic resonance imaging (MRI) features. We believe that extending the dataset with that obtained from an Iraqi hospital will not only assist in diagnosing BMCLVB but also fill the gap between data science and healthcare. Above all, the upgrade is intended to empower biomedical research and increase the chances of successful patient treatment.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Yasir Hussein Shakir

College of Graduate Studies (COGS), Universiti Tenaga Nasional (UNITEN)

Kajang, Malaysia

Email: yasserhessein19855@gmail.com

1. INTRODUCTION

The study investigates the early detector and classifier bone marrow changes in lumbar vertebrae (BMCLVB) using magnetic resonance imaging (MRI) data and addressing a key clinical challenge in diagnosing lumbar vertebrae. By improving classification techniques, the research looks to improve the diagnostic accuracy and outcome of patients. This study also addresses an important gap on the literature (BMCLVB) particularly in the application of machine learning (ML) to classification, a field currently constrained by the lack of publicly available datasets. BMCLVB was related to chronic low back pain (CLBP), and it was associated with increased disease severity in those who have HIV and osteoporosis therefore, BMCLVB may serve as an imaging biomarker for individuals with different skeletal, metabolic, and hematological illnesses as a surrogate in therapy response [1], [2]. It is possible for a fracture involving the verbal behavior (VB) or virtual assistant (VA) to cause nerve pressure when the space between the vertebral bodies is reduced and the intervertebral discs (IVD) may slip or herniate. As a result, bulging or slipped discs put stress on local nerves, producing tingling sensation in the legs and, in the occurrence of nerve injury, lower body helplessness [3], [4]. The bone marrow is only in charge of oxygen transport, coagulation, and immunity. It comprises hematopoietic cells (megakaryocytes, lymphoid, erythrocytes and

myeloid cells, and pluripotent mesenchymal stem cells) that interact with bone marrow adipocytes in a well-organized signal work supplied by bony trabeculae bordered by a fibrous reticulum [3].

MRI gives a non-invasive view into the bone marrow by mapping local, regional, or systemic alterations in marrow water position due to its unequalled sensitivity to the presence of fat and the relatively low effect of mineralized tissue on MRI signal [5], [6]. The patient receives an appropriate surgical therapy treatment, such as decompression, to relieve pressure on the nerves, consequently improving the patient's quality of life (QoL) [7]. The diagnostic process of persistent low back pain takes extensive effort, yet imaging tests like X-rays, MRI scans, and CT scans help the clinician to obtain specific spinal ailment information and relate this data to pain symptoms [8], [9]. Currently, they are two methods exist today for spinal surgical and radiology diagnosis either through human expertise or technical computer-based image analysis approaches. The diagnostic report provides information about bone marrow changes in lumbar vertebral and also confirms the diagnosis while the spinal surgeons analyze patients through clinical means based on displayed symptoms. The assessment of spinal alignment and spinal deformation along with spinal balance stands as a crucial aspect because scientific evidence proves that variations in spinal posture directly impact lumbar structure stability as well as pressure distribution [10]. Over time, researchers have discovered a link between changed spinal alignment and lumbago [11].

Medical imaging has advanced significantly in recent years, with MRI being a key component in the diagnosis and evaluation of many illnesses. The categorization of BMCLVB is crucial but difficult assignment among areas of interest in medical imaging. There is a clear vacuum in the body of literature despite significance of this field. Although lumbar vertebral imaging and classification have been subject of several research investigations, thorough work on BMCLVB classification is conspicuously lacking. Furthermore, construction and assessment of models for this particular job are made more difficult by the lack of a specialized, publicly accessible database. The establishment of this new dataset represents a substantial advancement in the field of BMCLVB using deep learning techniques. The main novelty and contributions of our research are as follows.

This study aims at the early detection and classification of BMCLVB in order to enhance the diagnosis and classify of the disease. Besides creating a new MRI dataset and data for BMCLVB classification, this work also proposes an novel, automated approach for BMCLVB classification. The strategy includes MRI-based data acquisition with histogram of oriented gradients (HOG) in feature extraction and further reduction with principal component analysis (PCA), linear discriminant analysis (LDA) and independent component analysis (ICA) and also used multiple classifiers are employed to enhance classification accuracy with the goal of establishing BMCLVB as a robust accessible biomarker in clinical settings.

In the present paper, the content is divided into six sections. Section 2 will discuss prior studies in disease prediction or classification and prior work, section 3 gives a brief technical background of the methods used in deep learning and medical images of several datasets. The explanation about the proposed and the experimental outcome is shown in section 4, followed by the section 5, which discusses the critical analyze. Section 6 looks to the future and outlines further research directions for this project.

2. PREVIOUS WORKS

In this part, given the lack of prior research tackling the issue, we will present an overview of the relevant literature on disease classification and prediction in broad terms. Literature can be structured in multiple ways. We can categorize them based on the disease they address, such as COVID-19, tumor brain, or based on the type of method used, for instance, traditional ML-based and deep learning-based.

2.1. Traditional machine learning-based

ML algorithms have been broadly connected over distinctive spaces, just like the classification and detection therapeutics and analysis. Al-Dhief *et al.* [12], proposed classifying dysphonia disease in the pathology detection application by using the mel frequency cepstral coefficient and the naive Bayes method as a classifier class from the normal and abnormal class by the Saarbrücken voice database. By training models on a dataset that is publicly accessible, the researchers investigate how effective these algorithms are with and without the use of data augmentation techniques. Additionally, Uysal and Ozturk [13] suggested the use of ML for diagnosing Alzheimer's disease (AD), including support vector machine (SVM), K-nearest neighbors (KNN), logistic regression (LR), decision tree (DT), gram-negative bacteria (GNB), and reasonable and non-discriminatory (RAND) have been utilized in tasks involving regression and classification. have been using classification and regression tasks. Multiple biomarkers used together is how most researchers agree is the best use and treatment for the diagnosis of conditions, rather than the single evaluation of one process alone. The study also reinforces that both gender and atrophy level heavily influence diagnostic judgments. Using (KNN) provides a maximum accuracy of 98% for the diagnostic group when using brain region, age, and gender volumes. However, despite this novelty, there is still a lack of a deeper understanding of the connection between AD and cognitive declines in the left hemisphere of the brain.

Gudadhe *et al.* [14] used A novel ML-based method for categorizing cerebral bleeding in computed tomography images is examined, and progress in the categorization of intracranial hemorrhage in CT scan images is shown in this work. They combine texture-based grey-level co-occurrence matrix (GLCM) features with transform-based features, like discrete wavelet transform and discrete cosine transform, by utilizing ML classifiers, and these characteristics are combined, they provide joint feature sets, which improve classification accuracy over separate approaches. This study offering a noteworthy accomplishment and compares the efficacy of this integrated feature method to traditional procedures and the combined feature set of wavelets transform with GLCM achieves the best accuracy of 87.22% when apply the random forest (RF) classifier.

Castro *et al.* [15] invented a technique to improve feature selection in AD stage classification to increase accuracy. To prepare and segment structural and functional brain MRI data, a non-human connectome project preamble approach is used and multiple ML classifiers, a collection of critical network parameters that change significantly throughout a specific progressive connection is collected. The average accuracy of the binary classification for Hadamard classifier (HC) against early mild cognitive impairment (EMCI), HC vs late MCI (LMCI), HC versus AD, 86.0% versus EMCI versus LMCI, 87.0% versus LMCI versus AD, and 88.5% versus EMCI versus AD is shown below. This study compares DT, KNN, and ensemble techniques. SVM routinely surpasses these approaches in terms of accuracy.

Like this work, El-Sappagh *et al.* [16] evaluated how five well-known ML algorithms namely LR, DT, KNN, RF, and SVM estimated the development of AD. The study uses different types of cost-efficient time-series characteristics that combine patient demographics along with medications taken and assessments of cognitive performance and existing medical conditions. These models execute four classification duties which consist of ND and prodromal MCI and subjective MCI and AD. Rephrase the RF model achieved high accuracy of 90.51% among ML algorithms in this study. The research demonstrates that high predictive value exists for every model which incorporates comorbidity and medication variables and other early-stage features to perform better than RF models. Dutta and Bandyopadhyay [17] detected a brain cancer classifier to tackle this and issued their proposal advocating for the use of ensemble-based ML algorithms. They introduce the NG-Boost algorithm, coupled with a (5-fold) stratified cross validation approach, as the primary classifier model for automated tumor detection, and contrasts ML techniques such as RF, Glioblastoma (GB), enhanced TumorNet (ET), and adaptive boosting (AdaB). It is worth noting, NG-Boost yields the most hopeful result, achieving an accuracy of 98.5%.

Khan *et al.* [18] discussed the issues involved in diagnosing migraine headaches, a common and complex neurovascular disease. They convincingly argue that the subjective measurements of pain intensity do not provide an accurate diagnosis, and hence the new methods of new measurements were imperative. ML, in turn, more specifically the ML models, have been proven to be quite successful in several other fields. Thus, their application to healthcare is also demanding, considering the shortage of resources. They predict and classify the many kinds of migraines using state-of-the-art ML techniques, such as SVM, KNN, RF, disturbance storm time index (DST), and deep neural networks (DNN). These are trained on public datasets, both with and without data augmentation, to enhance migraine diagnosis. The results show improvement in the range of 99.66% accuracy of DNN, 94.60% of SVM, 97.10% of KNN, 88.20% of DST, and 98.50% of RF. The prove the transformational power of ML in migraine diagnostics.

Karamti *et al.* [19] review focuses primarily on the problem of cervical cancer, particularly its burden and the impact of cervical cancer on women in the developing world. However, missing values in existing datasets for automated detection often compromise ML model performance. In this work, the authors suggest an automated system for forecasting cervical cancer and deal aptly with missing values by using synthetic minority oversampling technique (SMOTE) features to reach a high rate of accuracy. To deal with missing values, the system employs an ensemble voting classifier with three ML models using KNN imputer and SMOTE up-sampled features. When trained on KNN-imputed SMOTE features, the proposed model gains a remarkable 99.99% for the four metrics of both accuracy, precision, recall, and F1 score. Experimental results demonstrate the proposed model is more effective as compared to a few exists ML algorithms under different scenarios.

Tang *et al.* [20] investigated the use of quantitative characteristics taken from chest CT scans to categorize COVID-19 patients. They used a collection of CT scans from COVID-19 patients to extract variables such as texture, shape, and intensity. ML models were trained and evaluated for COVID-19 classification and high accuracy in discriminating between patients. Using 3-fold cross-validation, the RF model performed well, achieving true positive and negative rates of 0.933 and 0.745, accuracy of 0.875, and an area under the curve (AUC) of 0.91. These results highlight the efficiency of using quantitative CT image features for COVID-19 classification.

2.2. Deep leaning technical-based

Deep learning approaches, such as convolutional neural networks (CNN), are commonly utilized to identify and classify medical images. Ramos *et al.* [21] provide a special computer-aided detection (CAD) system for the automatic detection of breast cancer. Image segmentation, manual feature extraction (e.g.,

HOG, uniform local binary patterns (uLBP)), and deep feature extraction using Dense-Net 201 are all included in the system. A hybrid technique combines deep and handmade characteristics with statistical criteria and used genetic algorithms (GA) and mutually beneficial selection are used to choose the most important traits. The performed well on mammography and ultrasound datasets (accuracy 97.6%, precision 98%, recall 98%, F1-score 98%, and index balanced accuracy (IBA) 95%). The suggested CAD outperforms current approaches and shows promise for clinical applications.

The research work done by Javed *et al.* [22] specifically in identification and classification of lung focal opacity (2023), DL techniques like DNN and stack autoencoder classifiers have been defined for the identification and classification of lung focal opacity. This is indicative of the fact that deep learning algorithms are indeed useful in the field of medical image analysis. Moreover, facial images were employed when it comes to pain recognition and pain level differentiation utilizing ML. Munadi *et al.* [23] propose enhancing the DL models for diagnosing TB by increasing image quality. The following works were discussed: Contrast limited adaptive histogram equalization (CLAHE), high-frequency emphasis filtering (HEF), and unsharp masking (UM). Res-Net and Efficient-Net transfer learning models used enhanced images for more efficient classification compared to traditional images. Using the Shenzhen dataset to test the model, accuracy and the AUC metrics were as high as 89.92% and 94.8% and 7%, respectively, endorsing image enhancement as a tool to build a better, more powerful system for diagnosing TB to get better detection.

Sharma *et al.* [24] present and examine diverse CNN architectures designed for feature extraction and classification in the realm of pneumonia detection. Additionally, they explored how dataset size affects CNN performance via training models on both original and augmented datasets. Our findings, which revealed varying testing accuracies among different model configurations, assure the potential of CNNs to enhance pneumonia detection. Specifically, they achieved magnificent accuracy of 90%. Among authors, Sungheetha and Sharma [25] has the task of diagnosing the diabetes related symptoms to detect signs which are visible in blood vessels of eyes. Their technique involved using a deep feature CNN, which was specifically designed for detecting these markers. Taking an example, the model proved to work with accurate to an impressive 97% in diagnosing diabetic conditions. This study proves that deep learning is effective in medical diagnostics while also attempting to reveal the capabilities of de that make it possible to identify signs of illnesses by analyzing their visual patterns.

Devulapalli *et al.* [26] employed Gabor multiscale texture features in conjunction with a pre-trained Googlenet model as a feature extractor. From the extensive image collection, the final feature vector will be used to extract the required picture data. It has outperformed state-of-the-art techniques with an accuracy of 91%. This was done by Shankar *et al.* [27], who made a significant attempt at identifying COVID-19 by analyzing chest X-ray pictures. Their approach involves the following steps: pre-processing, which involves Weiner filtering, feature extraction with GLCM, generalized low-rank models (GLRM), and local binary patterns (LBP), and the final classification with an artificial neural networks (ANN). While feature selection was performed with great care to select the optimal feature subsets, this model showed potential to improve the classification efficiency between infected samples and healthy ones. This work also reveals the versatility of DL in solving medicinal diagnosis problems, especially the emergent health challenges in the world.

The research consisting of the detection of coronary artery stenosis was also carried out by Glessgen *et al* [28]. A retrospective study with 217 emergency room patients evaluated aortic CTA images for visual tally by two readers. The deep learning model (DLM) provided an independent classification of coronary arteries and the images as either high-quality (HQ) or low quality (LQ); accuracy: sensitivity, 92.9%; specificity, 95.7%; and net present value (NPV), 99.4% for patients with HQ images. Although it may have slightly lower sensitivity in the patient group with the lesser definition of the LQ images, the diagnostic efficiency did not change. Notably, the DLM in this regard exactly targeted all ACS patients and ICA-confirmed stenoses while highlighting the modality's possible strengths, primarily, including the delivery of an accessible non-parametric very high NPV for the substantial coronary stenosis in anticancer peptide (ACP) patients. Gottipalla and Yalla [29] further investigated a new novel hybrid ML approach to addressing dimensionality problem. This reliable algorithm combines the Bees Algorithm and ICA for optimal performance. As a next step, after establishing the feature reduction, various classifiers like KNN, SVM, LR, and RF. The aspects were inspected to evaluate the efficiency of these classifiers accuracy, memory consumption, precision, recall values, and F1 score.

3. PROPOSE APPROACHES

The aim of our study is to collect enough MRI-based bone marrow change data to be used for ML model development. It could also be utilized as a complementary assist for radiologists or clinicians when differentiating to make intelligent decisions and wiser. We will work with our dataset (BMCLVB), and the experimental process comprises several steps, the proposed model, and experimental methods utilized as shown in Figure 1.

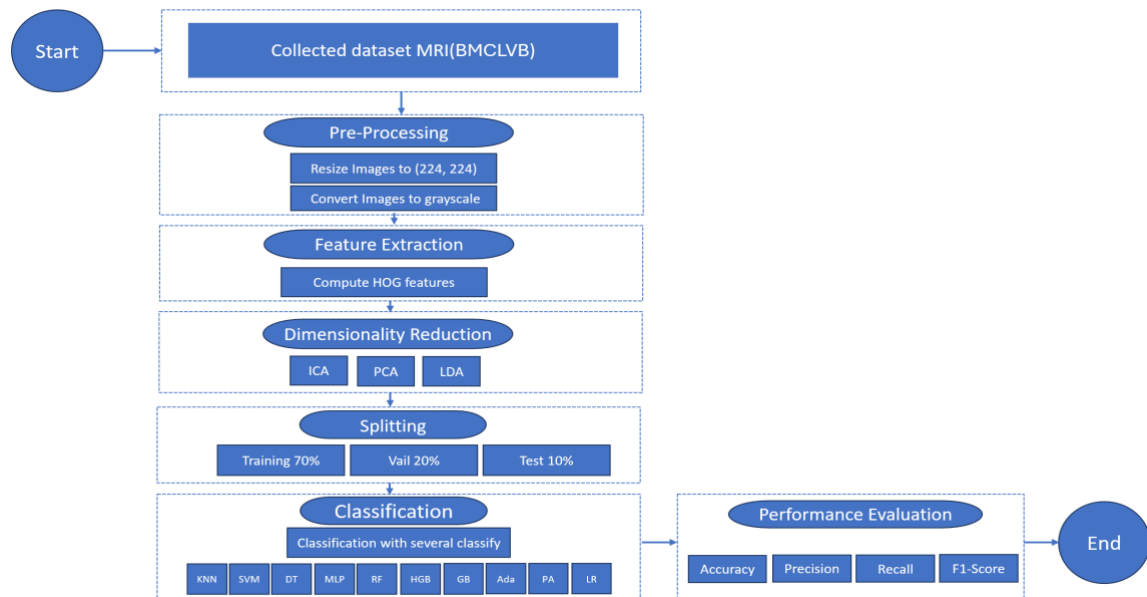


Figure 1. Flowchart of the proposed model and experimental methods used

3.1. Collected dataset MRI bone marrow changes in lumbar vertebrae

BMCLV dataset, comprising 1018 MRI scans from 134 BMCLVB patients who tested for bone marrow changes and were referred to the training-therapeutic Al-Kafeel super specialty hospital (KSSH), Iraq, from April 10, 2022, to September 2, 2023. The imaging protocol consisted of T1-weighted, T2-weighted, and fat-suppressed sequences with a slice thickness of 5 mm and a matrix size of 256×256 pixels. The images were stored in DICOM format and anonymized before analysis, and the ages of the patients ranged from 18 to 80 years old. The present study was conducted according to the guidelines provided by the institution's review board and the declaration of Helsinki principles regarding ethical conduct of research involving human subjects. According to the findings of this study, BMCLVB diagnosis is based on clinical signs and symptoms, blood analysis, MRIscans, and x-ray bone inspection. The usual intervention entails long span dosage of antibiotic or antifungal chemicals and, at times, surgery to drain out abscess or even remove affected bone.

A patient is identified as positive for BMCLVB if they meet one or more the following criteria: i) displaying suspicious clinical symptoms (such as fatigue, fever, back pain, neurological issues, and leg pain) in conjunction with positive MRI findings and ii) exhibiting suspicious laboratory findings (such as dysplastic changes like megaloblastic rubricates, neutrophil hyper segmentation, or giant platelets, which may indicate myelodysplastic syndrome or certain myeloid leukemias) along with positive MRI results.

Three expert radiologists performed readings to determine whether the images displayed normal or abnormal BMCLVB findings. We based our assessment on the more probable diagnosis identified by the agreement between two radiologists when their individual diagnoses varied from their peers. Images with visible lesions from other diseases or known masses together with those of poor quality and substandard radiation were eliminated from the study.

3.2. Pre-processing

The radiology images exported had sizes of approximately 527 KB, and the formats were in DICOM with a pixel resolution of 512×512 . The images were therefore cropped in JPG format to enhance patient anonymity and to match most programming standard libraries. Even where the file size was reduced to 92 KB.

3.2.1. Resize images

At this phase, BM in different resizes and properties is normalized. Specifically, the images are resized to 224×224 . In this process, in the next step, all image features are provided that have the same size and property in feature extraction, facilitating consistent analysis.

3.2.2. Convert images to grayscale

Grayscale conversion is one of the most basic or initial image processing and is often adopted as the first filter in image processing. This is versatile for images through reducing the amount of color information but maintaining structural features through the conversion element called grayscale. This helps to simplify

because it aids in the elimination of complex computational activity that can mask radical patterns and texture [30]. The output image is black and white while the intensity varies from 0 to 255 (least and greatest values respectively) which means it can easily be analyzed later.

3.3. Feature extraction using histogram of oriented gradients

Proper for the objective of our study, we applied HOG as a feature extraction on the textural and shapes of MRI images. It calculates gradient orientations in the localized regions of an image that then identifies the edges and corners of an image [31]. The image is usually converted to grayscale and feature extraction occurs through HOG which stands for HOG. HOG parameters consist of Orientations:9, Pixels per cell:(8, 8), cells per block:(2, 2), Transform square root: Yes, Block normalization method: L2 - Hys. The equation for calculating HOG features is as (1):

$$HOG(x, y) = x = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (1)$$

3.4. Dimensionality reduction techniques

Reduction in dimensions is a process where the overall volume of available features (or dimensions) is reduced while making sure that the data's important characteristics are retained. This process is crucial in dealing with high-dimensional data, which can suffer from issues such as the curse of dimensionality, increased computational complexity and overfitting. Include this process is crucial for dealing with high-dimensional data, which can suffer from issues such as the curse of dimensionality, increased computational complexity, and overfitting and include:

3.4.1. Principal component analysis

To decrease the number the data while maximizing the variance, to reach a limit, we used PCA. In essence, PCA can assist us in identifying more significant ways of explaining our data. It helps reducing the complexity essential for data representation bearing only minor informational loss.

3.4.2. Linear discriminant analysis

LDA was used as the second form of feature extraction under the category of supervised linear methods that is used for dimensionality reduction. To refresh our memory recall, LDA was helpful in a way that it enhanced between class variance to within class variance ratio. This improvement contributed to better class separability in our dataset.

3.4.3. Independent component analysis

We also used another technique called ICA whose objective was to further cut down on the dimensionality of our MRI image dataset to discover the latent statistically independent sources contained herein. ICA was particularly helpful in estimating and separating a non-Gaussian signal. This signal might be included in the sample.

3.5. Dataset split

The split the dataset into three subsets: training, validation, and test sets, where data are divided for learning the model, assessing the model's accuracy based on the unseen data of the validation, and the final assessment of the model on the unseen test data. The data set in this study involved a sample size of one thousand and eighteen. Hence, we partitioned 70% of the data, equal to (712 cases), into the training set for our models' training. The final set of data, which was also the validation set, contained 20% of the subjects (203 cases), and the last, the hold-out set, consisted of 10% of the data (103 cases), which was not exposed to the training and the cross-validation and was used to rate the general performance of the models as shown in Table 1.

Table 1. Split dataset information

| Dataset | Number of cases | Percentage of total dataset (%) |
|------------|-----------------|---------------------------------|
| Training | 712 | 70 |
| Validation | 203 | 20 |
| Test | 103 | 10 |
| Total | 1,018 | 100 |

3.6. Classification

In the classification process, the KNN, SVM, DT, multi-layer perceptron (MLP), RF, histogram (HG), gradient boosting (GB), AdaB, passive aggressive (PA), and LR methods were used to train the

identified features. These techniques fall under traditional methodologies of learning, and the results that can be achieved are highly beneficial in binary and multiclass formations. Even though artificial learning methods have been gradually advancing in recent years, leading to reduced utilization of traditional learning methods, it is pivotal to remember that classical learning methods are just as effective as well. Our results signify that utilizing factors like f1-score, and accuracy akin to our work may demonstrate that traditional learning techniques, such as classical ML, may also provide fruitful results. Additionally, we've discovered that when applying ML, it is possible to take advantage of classical learning methods for those kinds of tasks that require low complexity and rapid results.

3.6.1. K-nearest neighbors

In terms of purpose, KNN is one of the easiest to understand and most powerful algorithms used in classification and identification concerns. It sorts data based onto KNN in a feature space. The output is a class membership where the sample is classified by a voting system and the votes are cast by neighbors. The algorithm used to implement cluster incremental is nonparametric and of the lazy learning type, which implies that the data distribution does not make any assumptions and does not manipulate the training data points to make certain inferences [32].

3.6.2. Support vector machine

One of the most effective and widely used supervised learning algorithms, SVM performs well in both classification and function approximation. Although it is nearly equally applicable to all forms of problems, it is most frequently applied to classification problems. The implementation of the SVM algorithm involves placing each item of the data set on the n-dimensional plane, where n is the number of attributes in the dataset, with the qualities of attributes being the measure of the coordinate axis. Next, the classification step is carried out, which involves identifying the plane that maximally separates the two classes [33].

3.6.3. Decision tree

DT undergo learning algorithms that can only be employed in those problems where the target variable is already defined or given, these are more useful in exploratory type of problems like classification problems. Since it can estimate any functional form, it is suitable for both continuous and categorical input and output variables. This algorithm means splitting the population into two or more homogenous groups in accordance with the most distinguishing feature/input variable [34].

3.6.4. Multi-layer perceptron

MLP contains multiple perceptrons because it is composed of more than one perceptron. The true computational power of MLP exists between its input and output layers in various hidden layers which along with one hidden layer. These, along with one hidden layer, enable any continuous function approximation [35].

3.6.5. Random forest

RF is an ensemble learning method for classification, regression, and other tasks that work by constructing a multitude of DT at training time. It outputs the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct the DT habit of overfitting to their training set [36].

3.6.6. Histogram-based gradient boosting

Histogram-based gradient boosting (HGB) is a ML approach for classification and regression that extends the gradient boosting machine architecture. It constructs an extra model in a forward stage-wise approach, allows for the optimization of any differentiable loss functions. Additionally, it uses histograms to accelerate the learning process [37].

3.6.7. Gradient boosting

GB serves as a ML technique that enables regression and classification by constructing prediction models through multiple weak prediction models made of DT. The modeling process happens stage-by-stage while the method enables optimization of any differentiable loss function. This approach creates generalized results [38].

3.6.8. Adaptive boosting

The ML meta-algorithm AdaB operates under the acronym AdaBoost. Yoav Freund along with Robert Schapire developed this algorithm. The method works along with multiple kinds of learning algorithms to enhance their performance results. Several weak learning algorithms produce their output which gets combined into weighted values forming the boosted classifier output [39].

3.6.9. Passive aggressive

PA algorithms are online learning algorithms that remain passive for proper classification results. However, they become aggressive in the case of a mistake, updating and modifying. It's appropriate for large-scale learning and may be utilized for binary or multi-class classification [40].

3.6.10. Logistic regression

LR is an analysis of statistics employed to analyze the validity of a given data when one or more explanatory factors influence the variable of interest. A metric, which is an ordered quantitative variable, represents the result to meet a dichotomous variable, in which there are only two options. It [41] has been applied in numerous fields such as social sciences, marketing, as well as in the medical field amongst others.

4. RESULTS

During the experiments, the system was configured with an Intel Core (i7-9700 CPU), two GPUs (T4×2), and 16 GB of RAM. All the processes described by the flow were done using Kaggle. The classification outcome was evaluated to compare the results of the classifiers with the actual results using accuracy, precision, recall, and F1 score. Metrics used in this view include the following that is described by (2) to (5). The equations demonstrate the calculation of these values: where higher value for each of these metrics is indicative of better model performance. Here, TP represents true positive, which is the TN signaling normal sample successfully categorized by the model, and TN signifies true negative, which is the BMCLVB sample accurately classified by the model; FP is for false positive, which is the BMCLVB sample misclassified as normal by the model; FN stands for false negative, which is the normal sample incorrectly classified as BMCLVB by the model. It is important to note that in our work, we think that BMCLVB samples are negative samples, while normal samples are positive samples.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5)$$

The HOG feature extraction method achieved its highest success percentage rates from Table 2. The marks the best performance rates with boldface. HOG feature extraction with LR classifier generated the highest performing results according to table data. The patient classification precision reached 89.36% while recall performed at 89.36% and the F1 score achieved 89.36% according to the recorded table data. The detection rate for normal as well as abnormal cases the accuracy reaches 90.29%.

Table 2. Classification results with only HOG

| Algorithm HOG | Accuracy | Precision | Recall | F1 score |
|--------------------|----------|-----------|--------|----------|
| KNN HOG | 0.8058 | 0.7288 | 0.9149 | 0.8113 |
| SVM HOG | 0.8932 | 0.9091 | 0.8511 | 0.8791 |
| DT HOG | 0.6408 | 0.6087 | 0.5957 | 0.6022 |
| MLP HOG | 0.8058 | 0.9655 | 0.5957 | 0.7368 |
| RF HOG | 0.8340 | 0.8947 | 0.7234 | 0.7990 |
| Hist gradient HOG | 0.8350 | 0.8409 | 0.7872 | 0.8132 |
| Gradient boost HOG | 0.7660 | 0.7556 | 0.7234 | 0.7391 |
| AdaBoost HOG | 0.6505 | 0.6122 | 0.6383 | 0.6250 |
| PA HOG | 0.8835 | 0.8431 | 0.9149 | 0.8772 |
| LR HOG | 0.9029 | 0.8936 | 0.8936 | 0.8936 |

Table 3 of HOG with PCA features indicates the highest hit percentages and where the best metrics are noted with bold marks. The results indicate that feature extraction using HOG in combination with the PCA reduction along with passive-aggressive classification produces maximum success rates. The precision rate in patient identification reached 91.07% while recall achieved 92.73% and F1 score amounted to 89.36% according to the table analysis. The accuracy 91.89% of all correct detections between normal and abnormal specimens.

A success percentage analysis of our data through HOG-LDA combination appears in Table 4. Among the results the most significant values stand out through bold formatting. According to the research

findings HOG features together with LDA dimensionality reduction achieve maximum success rates through KNN classification. The ability to differentiate patients precision achieved 76.59% success rate followed by 76.59% recall, and 76.59% F1 score according to the table examination. The accuracy which measures correct detection rate stands at 78.64% for all normal and abnormal cases.

Table 5 when using HOG with ICA. Each entry in the displayed table indicates results as the highest point before different metric in boldface. The table demonstrates that maximum success occurs when HOG features extract data while applying ICA dimensionality reduction combined with AdaBoost classifier model. Examination of the table shows precision 46.15% stands as the highest result while recall 38.20% and F1 score 41.80% are also included. All Normal and abnormal cases had accuracy rate of 51.46% when correctly detected.

Table 3. Classification results with HOG and PCA

| Algorithm HOG and PCA | Accuracy | Precision | Recall | F1 score |
|------------------------------------|----------|-----------|--------|----------|
| KNN HOG and PCA | 0.8058 | 0.7288 | 0.9149 | 0.8113 |
| SVM HOG and PCA | 0.8252 | 0.7636 | 0.8936 | 0.8235 |
| PA HOG and PCA | 0.9126 | 0.9107 | 0.9273 | 0.9189 |
| DT HOG and PCA | 0.6990 | 0.6071 | 0.7907 | 0.6876 |
| MLP HOG and PCA | 0.9126 | 0.8958 | 0.9149 | 0.9053 |
| RF HOG and PCA | 0.6893 | 0.9412 | 0.3404 | 0.5000 |
| Hist gradient boosting HOG and PCA | 0.8738 | 0.7931 | 0.9787 | 0.8762 |
| GB HOG and PCA | 0.8252 | 0.7544 | 0.9149 | 0.8269 |
| AdaBoost HOG and PCA | 0.7087 | 0.6269 | 0.8936 | 0.7368 |
| LR HOG and PCA | 0.9029 | 0.8936 | 0.8936 | 0.8936 |

Table 4. Classification results with HOG and LDA

| Algorithm | Accuracy | Precision | Recall | F1 score |
|---------------------------------|----------|-----------|--------|----------|
| KNN HOG and LDA | 0.7864 | 0.7659 | 0.7659 | 0.7659 |
| SVM HOG and LDA | 0.7767 | 0.7609 | 0.7447 | 0.7527 |
| PA HOG and LDA | 0.7573 | 0.7200 | 0.7659 | 0.7427 |
| DT HOG and LDA | 0.6408 | 0.5926 | 0.6809 | 0.6333 |
| MLP HOG and LDA | 0.7660 | 0.7447 | 0.7447 | 0.7447 |
| RF HOG and LDA | 0.6408 | 0.5926 | 0.6809 | 0.6333 |
| Hist gradient boost HOG and LDA | 0.7767 | 0.7727 | 0.7234 | 0.7477 |
| GB HOG and LDA | 0.7573 | 0.7292 | 0.7447 | 0.7368 |
| AdaBoost HOG and LDA | 0.7660 | 0.7347 | 0.7659 | 0.7500 |
| LR HOG and LDA | 0.7669 | 0.7346 | 0.7659 | 0.7499 |

Table 5. Classification results with HOG and ICA

| Algorithm | Accuracy | Precision | Recall | F1 score |
|---------------------------------|----------|-----------|--------|----------|
| KNN HOG and LDA | 0.4563 | 0.3953 | 0.3617 | 0.3775 |
| SVM HOG and LDA | 0.4757 | 0.4653 | 1.0 | 0.6347 |
| MLP HOG and LDA | 0.4466 | 0.4242 | 0.5957 | 0.4966 |
| PA HOG and LDA | 0.4369 | 0.3778 | 0.3617 | 0.3696 |
| DT HOG and LDA | 0.5049 | 0.45 | 0.382 | 0.413 |
| RF HOG and LDA | 0.5144 | 0.4612 | 0.382 | 0.4181 |
| Hist gradient boost HOG and LDA | 0.4757 | 0.4478 | 0.6383 | 0.5279 |
| GB HOG and LDA | 0.5049 | 0.4773 | 0.8936 | 0.6211 |
| AdaBoost HOG and LDA | 0.5146 | 0.4615 | 0.382 | 0.418 |
| LR HOG and LDA | 0.4175 | 0.3673 | 0.382 | 0.3746 |

5. DISCUSSIONS

The main objective of this research project is to develop a diagnostic system which detects changes in BMCLVB during the clinical process. The study together with its results demonstrate enough evidence to accomplish the determined goal. The system combines ML with feature extraction through HOG and dimensionality reduction via PCA, LDA, and ICA techniques as shown by Figure 2.

Table 6 shows the results obtained from our data set and the comparison of ML performance across feature extraction methods and dimensionality reduction. The utilization of passive-aggressive learning with HOG for feature extraction and PCA for dimensionality reduction yielded notable results. As our investigation shows, this approach provides an extraordinary accuracy of 91.26%, which signifies that it is effective when used for the evaluation of changes in the bone marrow. Besides, if we look at the precision part, it reads 91.07%, indicating that this model will greatly reduce false positives, which evidences the problem, that is, to improve the diagnosis and treatment of patients. Equally importantly, the high recall rate of 92.73%. The F1 score rate is 91.89%, and it increases confidence about the model in terms of a good balance between the

false-positive rate and the false-negative rate. In clinical practice, there is often a need for a perfect storm scenario where sensitivity and specificity are both optimal, which is achieved here. Therefore, the study provides evidence on how the passive-aggressive learning approach alongside HOG and PCA can be used to estimate the change in bone marrow in the lumbar vertebrae from MRI data. By artificial intelligence and ML, this paper presents a robust approach to enhance diagnostic accuracy and support clinicians during treatment planning. However, future studies are needed to extend and corroborate this methodology in different clinical contexts to check the extent of its cross-sectional and interventional translation potential.

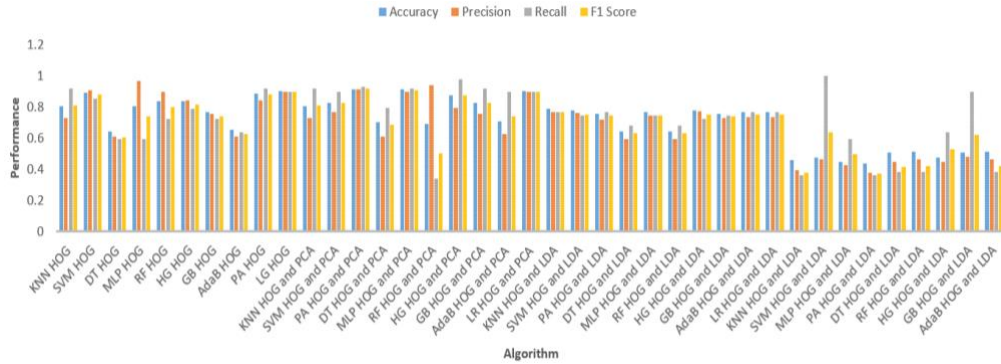


Figure 2. Comparison of ML algorithm performance across feature extraction methods and dimensionality reduction

Table 6. Comparison of ML algorithm performance across feature extraction methods and dimensionality reduction

| Algorithm and FE+DR | Accuracy | Precision | Recall | F1 Score |
|---------------------------|----------|-----------|--------|----------|
| KNN HOG | 0.8058 | 0.7288 | 0.9149 | 0.8113 |
| SVM HOG | 0.8932 | 0.9091 | 0.8511 | 0.8791 |
| DT HOG | 0.6408 | 0.6087 | 0.5957 | 0.6022 |
| MLP HOG | 0.8058 | 0.9655 | 0.5957 | 0.7368 |
| RF HOG | 0.8340 | 0.8947 | 0.7234 | 0.7990 |
| Hist gradient HOG | 0.8350 | 0.8409 | 0.7872 | 0.8132 |
| GB HOG | 0.7660 | 0.7556 | 0.7234 | 0.7391 |
| AdaBoost HOG | 0.6505 | 0.6122 | 0.6383 | 0.6250 |
| PA HOG | 0.8835 | 0.8431 | 0.9149 | 0.8772 |
| LR HOG | 0.9029 | 0.8936 | 0.8936 | 0.8936 |
| KNN HOG and PCA | 0.8058 | 0.7288 | 0.9149 | 0.8113 |
| SVM HOG and PCA | 0.8252 | 0.7636 | 0.8936 | 0.8235 |
| PA HOG and PCA | 0.9126 | 0.9107 | 0.9273 | 0.9189 |
| DT HOG and PCA | 0.6990 | 0.6071 | 0.7907 | 0.6876 |
| MLP HOG and PCA | 0.9126 | 0.8958 | 0.9149 | 0.9053 |
| RF HOG and PCA | 0.6893 | 0.9412 | 0.3404 | 0.5000 |
| Hist gradient HOG and PCA | 0.8738 | 0.7931 | 0.9787 | 0.8762 |
| GB HOG and PCA | 0.8252 | 0.7544 | 0.9149 | 0.8269 |
| AdaBoost HOG and PCA | 0.7087 | 0.6269 | 0.8936 | 0.7368 |
| LR HOG and PCA | 0.9029 | 0.8936 | 0.8936 | 0.8936 |
| KNN HOG and LDA | 0.7864 | 0.7659 | 0.7659 | 0.7659 |
| SVM HOG and LDA | 0.7767 | 0.7609 | 0.7447 | 0.7527 |
| PA HOG and LDA | 0.7573 | 0.7200 | 0.7659 | 0.7427 |
| DT HOG and LDA | 0.6408 | 0.5926 | 0.6809 | 0.6333 |
| MLP HOG and LDA | 0.7660 | 0.7447 | 0.7447 | 0.7447 |
| RF HOG and LDA | 0.6408 | 0.5926 | 0.6809 | 0.6333 |
| Hist gradient HOG and LDA | 0.7767 | 0.7727 | 0.7234 | 0.7477 |
| GB HOG and LDA | 0.7573 | 0.7292 | 0.7447 | 0.7368 |
| AdaBoost HOG and LDA | 0.7660 | 0.7347 | 0.7659 | 0.7500 |
| LR HOG and LDA | 0.7669 | 0.7346 | 0.7659 | 0.7499 |
| KNN HOG and LDA | 0.4563 | 0.3953 | 0.3617 | 0.3775 |
| SVM HOG and LDA | 0.4757 | 0.4653 | 1.0 | 0.6347 |
| MLP HOG and LDA | 0.4466 | 0.4242 | 0.5957 | 0.4966 |
| PA HOG and LDA | 0.4369 | 0.3778 | 0.3617 | 0.3696 |
| DT HOG and LDA | 0.5049 | 0.4500 | 0.3820 | 0.4130 |
| RF HOG and LDA | 0.5144 | 0.4612 | 0.3820 | 0.4181 |
| Hist gradient HOG and LDA | 0.4757 | 0.4478 | 0.6383 | 0.5279 |
| GB HOG and LDA | 0.5049 | 0.4773 | 0.8936 | 0.6211 |
| AdaBoost HOG and LDA | 0.5146 | 0.4615 | 0.3820 | 0.4180 |

Another striking issue in the same is the average detection rate of spinal cord injuries by MRI is 75–80% and in a study published in the Journal of Neurotrauma, researchers highlighted the pivotal role of MRI scans in the assessment acute spinal cord injury (SCI). This underscores the significance of MRI in the timely and accurate diagnosis of SCI [42], [43]. Another important aspect of the study contributing to the literature is the duration of diagnosis. Our research achieves goal through reducing the time needed for radiology specialists during their work diagnosis. We have emphasized in various sections of our study that our core objective involves creating programs to help healthcare providers make diagnoses. Laboratory results demonstrate that the intensity of CLBP shows direct correlations with BMCLVB particularly in patients dealing with cancer and osteoporosis together with HIV. Stressful and tiring work conditions naturally lead radiologists to miss some issues during diagnosis. According to Table 7 we obtained these results in our research.

Table 7. The comparison of detection performance with the previous worker study using different ML and DL algorithms

| Author | Year | Method | Dataset | Performance (%) | Application |
|---------------------------------|------|--|----------------------------|--------------------------------|--|
| Lewandrowski <i>et al.</i> [44] | 2020 | NN and CNN | MRI T1, T2 | Accuracy=85.2 | Reporting in routine spine |
| Pan <i>et al.</i> [45] | 2021 | Faster Region-based CNN | lumbar MR Images T2 | Accuracy=88.8 | Detect other lumbar abnormalities and cervical spondylosis |
| Tavana <i>et al.</i> [46] | 2023 | 1-Pre-trained Xception+SVM 2- Pre-trained MobileNetV2+SVM | X-ray images | 1-Accuracy=96 2-Accuracy=96 | Degenerative lumbar spinal stenosis |
| Jiang <i>et al.</i> [47] | 2024 | XGBoost, LR, RF AdaBoost, GNB, MLP, SVM, and KNN | DXA scanning and lumbar CT | XGBoost Accuracy=84.70 | Screw loosening after degenerative lumbar fusion surgery |
| Our study | 2024 | Passive aggressive HOG and PCA | MRI BMCLVB | Accuracy=91.26 | Detection and classification of BMCLVB |

6. CONCLUSION AND FUTURE WORKS

Our study on the classifier of BMCLVB using MRI data highlights several key findings and implications for clinical practice effective classification with ML. Our investigation demonstrates the feasibility of utilizing ML algorithms, particularly passive-aggressive learning with HOG and PCA, for accurate classification of bone marrow changes. The achieved accuracy, precision, recall, and F1 score indicate the potential of these techniques in enhancing diagnostic capabilities and streamlining clinical decision-making processes. Accurate classification of BMCLVB is crucial for diagnosing various pathologies, including degenerative disc disease, infections, and neoplastic conditions. This research proposes a solid way for increasing diagnosis accuracy and aiding physicians during treatment planning through the use of artificial intelligence methods and ML. The robust performance of the proposed methodology suggests its suitability for integration into clinical workflows, providing clinicians with valuable insights derived from MRI data. The implementation of such automated classification systems has the possibility to enhance efficiency, reduce errors in diagnosis, and improve medical practice's outcomes for patients. Further validation and refinement of the developed classification model are warranted, including validation on independent datasets and exploration of additional feature extraction techniques. Longitudinal studies are needed to assess the model's performance in tracking disease progression and response to treatment over time. Our study contributes to the expanding research on ML applications in medical imaging analysis, demonstrating its potential to augment diagnostic capabilities and improve patient care. Collaboration between clinicians, radiologists, and ML experts is essential for translating research findings into clinically actionable tools and practices. In summary, our study underscores the promise of ML approaches in advancing the field of medicinal imaging analysis, with implications for improved diagnosis, treatment planning, and patient results in lumbar vertebrae pathology. Continued study and collaboration are crucial for realizing the full potential of these techniques in clinical practice. Many potential way into future research were identified, including experiments with different deep learning methods.

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to Al-Kafeel Super Specialty Hospital (KSSH) for their invaluable support in this project. Special thanks to Dr. Ali Kanj, Dr. Hisham Hassan Abd, Dr. Mohamed Abdelreda, as well as the Electronica Department—Mr. Kamal Al-Deen Mahdi Hussein and Mr. Mohammed Abbood Kadhim—for their dedicated efforts and assistance in the collection of the dataset.

FUNDING INFORMATION

The authors state no funding was involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|----------------------|---|---|----|----|----|---|---|---|---|---|----|----|---|----|
| Yasir Hussein Shakir | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | | | ✓ |
| Tiong Sieh Kiong | | | | ✓ | | ✓ | ✓ | | | ✓ | | ✓ | ✓ | |
| Chai Ping Chen | ✓ | | | ✓ | | | ✓ | | | ✓ | | ✓ | ✓ | |

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

The authors state that there are no conflicts of interest.

DATA AVAILABILITY




The data that support the findings of this study are available from the corresponding author, [YHS], upon reasonable request.

REFERENCES




- [1] R. Coleman *et al.*, "Bone markers and their prognostic value in metastatic bone disease: Clinical evidence and future directions," *Cancer Treatment Reviews*, vol. 34, no. 7, pp. 629–639, 2008, doi: 10.1016/j.ctrv.2008.05.001.
- [2] F. Raciborski, R. Gasik, and A. Ktak, "Disorders of the spine. A major health and social problem," *Reumatologia*, vol. 54, no. 4, pp. 196–200, 2016, doi: 10.5114/reum.2016.62474.
- [3] B. Roudsari and J. G. Jarvik, "Lumbar spine MRI for low back pain: Indications and yield," *American Journal of Roentgenology*, vol. 195, no. 3, pp. 550–559, 2010, doi: 10.2214/AJR.10.4367.
- [4] P. Suri, E. J. Boyko, J. Goldberg, C. W. Forsberg, and J. G. Jarvik, "Longitudinal associations between incident lumbar spine MRI findings and chronic low back pain or radicular symptoms: Retrospective analysis of data from the longitudinal assessment of imaging and disability of the back (LAIDBACK)," *BMC Musculoskeletal Disorders*, vol. 15, no. 1, May 2014, doi: 10.1186/1471-2474-15-152.
- [5] M. J. Devlin and C. J. Rosen, "The bone-fat interface: Basic and clinical implications of marrow adiposity," *The Lancet Diabetes and Endocrinology*, vol. 3, no. 2, pp. 141–147, 2015, doi: 10.1016/S2213-8587(14)70007-5.
- [6] D. C. Karampinos *et al.*, "Quantitative MRI and spectroscopy of bone marrow," *Journal of Magnetic Resonance Imaging*, vol. 47, no. 2, pp. 332–353, 2018, doi: 10.1002/jmri.25769.
- [7] K. Hasegawa, M. Okamoto, S. Hatsushikano, H. Shimoda, M. Ono, and K. Watanabe, "Normative values of spino-pelvic sagittal alignment, balance, age, and health-related quality of life in a cohort of healthy adult subjects," *European Spine Journal*, vol. 25, no. 11, pp. 3675–3686, 2016, doi: 10.1007/s00586-016-4702-2.
- [8] D. Rao, G. Scuderi, C. Scuderi, R. Grewal, and S. J. S. Sandhu, "The use of imaging in management of patients with low back pain," *Journal of Clinical Imaging Science*, vol. 8, no. 1, 2018, doi: 10.4103/jcis.jcis_16_18.
- [9] B. P. Hirsch *et al.*, "Applications of SPECT/CT in the evaluation of spinal pathology: a review," *International Journal of Spine Surgery*, vol. 18, no. 1, pp. 9–23, 2024, doi: 10.14444/8552.
- [10] M. Pourahmadi *et al.*, "Spinopelvic alignment and low back pain after total hip arthroplasty: a scoping review," *BMC Musculoskeletal Disorders*, vol. 23, no. 1, 2022, doi: 10.1186/s12891-022-05154-7.
- [11] G. Keorochana *et al.*, "Effect of sagittal alignment on kinematic changes and degree of disc degeneration in the lumbar spine: An analysis using positional MRI," *Spine*, vol. 36, no. 11, pp. 893–898, 2011, doi: 10.1097/BRS.0b013e3181f4d212.
- [12] F. T. Al-Dhief, N. M. A. Latiff, N. N. A. Malik, M. M. Baki, N. Sabri, and M. A. A. Albadr, "Dysphonia detection based on voice signals using naive bayes classifier," in *Conference Proceedings - 2022 IEEE 6th International Symposium on Telecommunication Technologies: Intelligent Connectivity for Sustainable World, ISTT 2022*, 2022, pp. 56–61, doi: 10.1109/ISTT56288.2022.9966535.
- [13] G. Uysal and M. Ozturk, "Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods," *Journal of Neuroscience Methods*, vol. 337, 2020, doi: 10.1016/j.jneumeth.2020.108669.
- [14] S. Gudadhe, A. Thakare, and A. M. Anter, "A novel machine learning-based feature extraction method for classifying intracranial hemorrhage computed tomography images," *Healthcare Analytics*, vol. 3, 2023, doi: 10.1016/j.health.2023.100196.
- [15] A. P.-Castro, E. Fernandez-Blanco, A. Pazos, and C. R. Munteanu, "Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques," *Computers in Biology and Medicine*, vol. 120, 2020, doi: 10.1016/j.compbiomed.2020.103764.
- [16] S. El-Sappagh *et al.*, "Alzheimer's disease progression detection model based on an early fusion of cost-effective multimodal data," *Future Generation Computer Systems*, vol. 115, pp. 680–699, 2021, doi: 10.1016/j.future.2020.10.005.

- [17] S. Dutta and P. S. K. Bandyopadhyay, "Revealing brain tumor using cross-validated NGBoost classifier," *Sustainability*, vol. 11, no. 1, pp. 1–14, Jul. 2020, doi: 10.21203/rs.3.rs-47048/v1.
- [18] L. Khan, M. Shahreen, A. Qazi, S. J. A. Shah, S. Hussain, and H. T. Chang, "Migraine headache (MH) classification using machine learning methods with data augmentation," *Scientific Reports*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-55874-0.
- [19] H. Karamti *et al.*, "Improving prediction of cervical cancer using KNN imputed SMOTE features and multi-model ensemble learning approach," *Cancers*, vol. 15, no. 17, 2023, doi: 10.3390/cancers15174412.
- [20] Z. Tang *et al.*, "Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images," *arXiv-Electrical Engineering and Systems Science*, pp. 1-18, 2020.
- [21] C. C. -Ramos, O. Garcia-Avila, J. A. Almaraz-Damian, V. Ponomaryov, R. Reyes-Reyes, and S. Sadovnychiy, "Benign and malignant breast tumor classification in ultrasound and mammography images via fusion of deep learning and handcraft features," *Entropy*, vol. 25, no. 7, 2023, doi: 10.3390/e25070991.
- [22] M. A. Javed, H. B. Liaqat, T. Meraj, A. Alotaibi, and M. Alshammari, "Identification and classification of lungs focal opacity using CNN segmentation and optimal feature selection," *Computational Intelligence and Neuroscience*, vol. 2023, no. 1, 2023, doi: 10.1155/2023/6357252.
- [23] K. Munadi, K. Muchtar, N. Maulina, and B. Pradhan, "Image enhancement for tuberculosis detection using deep learning," *IEEE Access*, vol. 8, pp. 217897–217907, 2020, doi: 10.1109/ACCESS.2020.3041867.
- [24] H. Sharma, J. S. Jain, P. Bansal, and S. Gupta, "Feature extraction and classification of chest X-ray images using CNN to detect pneumonia," *Proceedings of the Confluence 2020 - 10th International Conference on Cloud Computing, Data Science and Engineering*, pp. 227–231, 2020, doi: 10.1109/Confluence47617.2020.9057809.
- [25] A. Sungheetha and R. Sharma, "Design an early detection and classification for diabetic retinopathy by deep feature extraction based convolution neural network," *Journal of Trends in Computer Science and Smart Technology*, vol. 3, no. 2, pp. 81–94, 2021, doi: 10.36548/jtcsst.2021.2.002.
- [26] S. Devulapalli, A. Potti, R. Krishnan, and M. S. Khan, "Experimental evaluation of unsupervised image retrieval application using hybrid feature extraction by integrating deep learning and handcrafted techniques," *Materials Today: Proceedings*, vol. 81, no. 2, pp. 983–988, 2021, doi: 10.1016/j.matpr.2021.04.326.
- [27] K. Shankar, E. Perumal, P. Tiwari, M. Shorfuzzaman, and D. Gupta, "Deep learning and evolutionary intelligence with fusion-based feature extraction for detection of COVID-19 from chest X-ray images," *Multimedia Systems*, vol. 28, no. 4, pp. 1175–1187, 2022, doi: 10.1007/s00530-021-00800-x.
- [28] C. G. Glessgen *et al.*, "Artificial intelligence-based opportunistic detection of coronary artery stenosis on aortic computed tomography angiography in emergency department patients with acute chest pain," *European Heart Journal Open*, vol. 3, no. 5, 2023, doi: 10.1093/ehjopen/oead088.
- [29] A. K. Gottipalla and P. Yalla, "Efficacy of machine learning models in lung cancer detection: an emphasis on bees with ICA hybrid feature extraction," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 16s, pp. 31–40, 2024.
- [30] M. M. Rahman, S. Nooruddin, K. M. A. Hasan, and N. K. Dey, "HOG + CNN Net: diagnosing COVID-19 and Pneumonia by deep neural network from chest x-ray images," *SN Computer Science*, vol. 2, no. 5, 2021, doi: 10.1007/s42979-021-00762-x.
- [31] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. 1, pp. 886–893, 2005, doi: 10.1109/CVPR.2005.177.
- [32] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, pp. 986–996, 2003, doi: 10.1007/978-3-540-39964-3_62.
- [33] A. Roy and S. Chakraborty, "Support vector machine in structural reliability analysis: A review," *Reliability Engineering and System Safety*, vol. 233, 2023, doi: 10.1016/j.res.2023.109126.
- [34] B. Kamiński, M. Jakubczyk, and P. Szufel, "A framework for sensitivity analysis of decision trees," *Central European Journal of Operations Research*, vol. 26, no. 1, pp. 135–159, 2018, doi: 10.1007/s10100-017-0479-6.
- [35] E. Bisong, "The multilayer perceptron (MLP)," in *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, Berkeley, CA: Apress, 2019, pp. 401–405, doi: 10.1007/978-1-4842-4470-8_31.
- [36] B. C. Ko, S. H. Kim, and J. Y. Nam, "X-ray image classification using random forests with local wavelet-based CS-local binary patterns," *Journal of Digital Imaging*, vol. 24, no. 6, pp. 1141–1151, 2011, doi: 10.1007/s10278-011-9380-3.
- [37] A. Guryanov, "Histogram-based algorithm for building gradient boosting ensembles of piecewise linear decision trees," *Analysis of Images, Social Networks and Texts*, pp. 39–50, 2019, doi: 10.1007/978-3-030-37334-4_4.
- [38] G. Biau, B. Cadre, and L. Rouvière, "Accelerated gradient boosting," *Machine Learning*, vol. 108, no. 6, pp. 971–992, 2019, doi: 10.1007/s10994-019-05787-1.
- [39] P. Favaro and A. Vedaldi, "AdaBoost," *Computer Vision: A Reference Guide, Springer International Publishing*, pp. 36–40, 2021, doi: 10.1007/978-3-030-63416-2_663.
- [40] L. Gao, W. Zhang, and Q. Tang, "Advanced data mining and applications: 9th International Conference, ADMA 2013, Proceedings," in *Lecture Notes in Computer Science*, Springer Singapore, doi: 10.1007/978-981-96-0850-8.
- [41] E. Bisong, *Building machine learning and deep learning models on google cloud platform: a comprehensive guide for beginners*, California: Apress Berkeley, 2019, doi: 10.1007/978-1-4842-4470-8.
- [42] S. Kadoya, T. Nakamura, S. Kobayashi, and I. Yamamoto, "Magnetic resonance imaging of acute spinal cord injury - Report of three cases," *Neuroradiology*, vol. 29, no. 3, pp. 252–255, 1987, doi: 10.1007/BF00451762.
- [43] C. O. Walsh, S. I. Ziniel, H. K. Delichatsios, and D. S. Ludwig, "Nutrition attitudes and knowledge in medical students after completion of an integrated nutrition curriculum compared to a dedicated nutrition curriculum: A quasi-experimental study," *BMC Medical Education*, vol. 11, no. 1, 2011, doi: 10.1186/1472-6920-11-58.
- [44] K. U. Lewandrowski *et al.*, "Feasibility of deep learning algorithms for reporting in routine spine magnetic resonance imaging," *International Journal of Spine Surgery*, vol. 14, pp. S86–S97, 2020, doi: 10.14444/7131.
- [45] Q. Pan *et al.*, "Automatically diagnosing disk bulge and disk herniation with lumbar magnetic resonance images by using deep convolutional neural networks: Method development study," *JMIR Medical Informatics*, vol. 9, no. 5, 2021, doi: 10.2196/14755.
- [46] P. Tavana, M. Akraminia, A. Koochari, and A. Bagherifard, "Classification of spinal curvature types using radiography images: deep learning versus classical methods," *Artificial Intelligence Review*, vol. 56, no. 11, pp. 13259–13291, 2023, doi: 10.1007/s10462-023-10480-w.
- [47] F. Jiang, X. Li, L. Liu, Z. Xie, X. Wu, and Y. Wang, "Automated machine learning-based model for the prediction of pedicle screw loosening after degenerative lumbar fusion surgery," *BioScience Trends*, vol. 18, no. 1, pp. 83–93, 2024, doi: 10.5582/bst.2023.01327.




BIOGRAPHIES OF AUTHORS

Yasir Hussein Shakir    is a software engineer who obtained his B.Sc. degree in software engineering from Baghdad College of Economic Sciences University in 2014. He further pursued his education and received his M.Sc. degree in computer and communication, specializing in computer programming, from the Faculty of Engineering at the Islamic University of Lebanon (IUL) in 2018. His areas of interest include data mining, image medical processing, medical electronic systems, machine learning, deep learning, and artificial intelligence. Currently, he is a Ph.D. student in the Department of Engineering at Universiti Tenaga Nasional (UNITEN) in Malaysia. He can be contacted at email: yasserhessein19855@gmail.com.



Tiong Siah Kiong    is a senior professor and a senior member of the IEEE, currently serving in the College of Engineering at Universiti Tenaga Nasional (UNITEN). He also holds the position of Director at the Institute of Sustainable Energy (ISE), UNITEN. He received his B.Eng. (Hons), M.Sc., and Ph.D. in electrical and electronic engineering. His research interests include renewable energy, artificial intelligence, data analytics, and communication systems. He is a professional engineer registered with the Board of Engineers Malaysia (BEM) and a member of the Institute of Electrical and Electronics Engineers (IEEE). He can be contacted at email: siahkiong@uniten.edu.my.



Chai Phing Chen    is a senior lecturer in Universiti Tenaga Nasional (UNITEN), Malaysia. She has been associated with technical education for more than ten years and actively participated in various UNITEN projects which involve machine learning. Her project field contains predicting for gas emission from power plant; prediction for short-term wind speed; health condition analysis for HV circuit breaker; fault detection for switchgear; fault detection for transformer; heat waste recovery system via thermoelectric generator and nontechnical losses detection. She can be contacted at email: chenpc@uniten.edu.my.