Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 246 (2024) 3977–3989

Procedia Computer Science

www.elsevier.com/locate/procedia

28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2024)

SpineSighter: An AI-Driven Approach for Automatic Classification of Spinal Function from Video

Zebang Liu[a,*], Yulia Hicks[a], Liba Sheeran[b,c]

[a]School of Engineering, Cardiff University, Cadiff CF24 3AA, United Kingdom
[b]Cardiff School of Healthcare, Cardiff University, Cardiff CF24 0AB, United Kingdom
[c]Biomechanics and Bioengineering Research Centre Versus Arthritis, Cardiff University, Cardiff CF24 0AB, United Kingdom

Abstract

Low Back Pain (LBP) is a prevalent musculoskeletal disorder affecting over 80% of the population over their lifetime and is a leading cause of disability globally. The most frequent type, non-specific LBP (NSLBP) does not have a clearly identifiable pathology cause. Current clinical guidelines advocate for tailored management and self-care approaches for NSLBP. The effectiveness of these personalised management plans significantly depends on accurate and on-going assessment of the patient's spinal function. This presents considerable challenges for both clinicians and patients.
This study introduces "SpineSighter", an artificial intelligence (AI) model developed to tailor management of NSLBP by categorising patients based on their spinal function either into High Function (HF) and Low Function (LF) subsets. Utilising standard video recordings and computer vision technology, SpineSighter analyses motion features such as angular displacement, velocity, and acceleration during repeated forward flexion tests. The model showed high accuracy in classifying spinal function, achieving an accuracy of 95.13%, sensitivity of 93.81%, specificity of 96.00%, and an F1 score of 0.9442. This innovative use of AI highlights the importance of velocity as a critical indicator of spinal functional differences, opening new avenues for personalised clinical management, self-care and recovery strategies of NSLBP.

Peer-review under responsibility of the scientific committee of the 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems

Keywords: Non-specific Low Back Pain (NSLBP); Spinal Functions Classification; Artificial Intelligence (AI); Human Posture Estimation (HPE); Motion Features Analysis.

1. Introduction

Low Back Pain (LBP) affects more than 80% of individuals over their lifetime, marking it as a highly prevalent musculoskeletal disorder and a significant global public health issue [1]. It notably impacts occupational productivity and brings a considerable economic burden worldwide, particularly among those aged 40 to 69 [2]. The economic impact of LBP exceeds one billion USD in many countries, and it stands as a leading cause of disability globally [1, 2], highlighting its importance as a critical public health concern.

* Corresponding author. Tel.: +44(0)7731696509;
E-mail address: Liuz91@cardiff.ac.uk

1877-0509 © 2024 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems
10.1016/j.procs.2024.09.172

Investigations into the aetiology of LBP reveal that nearly 90% of cases do not show pathoanatomical or radiological abnormalities, thus classified as non-specific LBP (NSLBP) [2]. NSLBP, marked by pain and impaired motor function without a clear anatomical or pathological cause [2], prompts clinical guidelines to advocate for treatments that include education, reassurance, non-pharmacological interventions, self-management practices, as well as exercise customised to individual needs [3].

However, formulating personalised treatment plans for NSLBP is challenging due to its complex nature and the diverse physical alterations and disability levels of patients [2]. Variability in spinal function among NSLBP patients is intricately linked to NSLBP [5, 6], necessitating rehabilitation recommendations that consider the specific spinal posture and movement alterations of each subgroup [6, 7, 8], highlighting the importance of accurate spinal functions classification [3, 6, 8].

Clinicians often use digital images and videos to assess posture and movement patterns in NSLBP [7, 15]. Artificial Intelligence (AI), with its robust capabilities, has made significant inroads across various medical fields [9], helping in utilising this technology for LBP management. Most efforts focus on deriving diagnostic insights from medical images, while others utilise inertial measurement units (IMU) [10], E-skin [11], or 3D motion capture system [6] to gather data for NSLBP analysis, though these methods present substantial challenges for patient daily use. To our knowledge, no research has yet explored using AI to classify spinal functions based on videos.

Human Pose Estimation (HPE), a critical AI challenge, analyses motion patterns from video data [14], showcasing the potential of applying AI to easily accessible standard video recordings without the need for elaborate setups. The efficacy of HPE in extracting video-based motion features for NSLBP classification has been proven [15]. However, to the best of our knowledge, there have been no studies on using AI to extract spinal movement from videos for classifying spinal function during daily tasks such as forward bending.

Feature importance estimation is vital for interpreting complex machine learning models. Perturbation-based methods are popular for their simplicity and applicability across different models. By altering a feature's values and observing changes in model performance, these methods highlight key features and support the advancement of explainable artificial intelligence (XAI), promoting transparency and trust in machine learning [16].

This study looked to develop an AI model using computer vision technology to automatically classify spinal function in NSLBP patients through standard video recordings of their routine physical examination. This method aimed to use movement pattern differences in forward flexion movements to offer a more objective, scalable, and economically practical clinical solution. Our vision is that integrating HPE into an AI framework will significantly enhance the convenience and accuracy of classifying spinal function in NSLBP patients, offering an opportunity for highly personalised self-management and patient monitoring to assess intervention effectiveness.

1.1. Classifying level of spinal function impairment

In the absence of a pathological cause, numerous methods classify NSLBP often focusing on functional and motor performance as a critical factor [4, 7]. These classification systems identify specific patterns that correlate with specific subgroups, each requiring tailored intervention strategies. Assessments of spinal posture and movement are crucial for subgroup classification [6, 7, 17], helping to distinguish clinical subgroups by distinct postures (e.g., sitting, standing) and movements (e.g., forward and backward bend, sit-to-stand, squat). These subgroups display unique spinal and pelvic kinematics associated with varying functional limitations such as difficulties sitting or standing. Additionally, the level of functional limitation varies within each patient, reflecting the fluctuating nature of NSLBP, which alternates between periods of remissions and exacerbation.

When classifying the level of spinal function impairment, movements of patients with a high level of function (HF) are likely to appear as smooth and confident but may degrade with repetitions. In contrast, those with low function (LF) are more likely to show hesitant, slow, uncoordinated, movements or abrupt stops when changing speeds, showing movement disruption [8]. Real time differentiation between low and high function may allow immediate customisation of education, exercise, and other management strategies. This responsive approach would encourage full patient adherence and engagement, thus enhancing both the effectiveness and efficiency of interventions for LBP.

1.2. LBP research technology

To advance the understanding and treatment of NSLBP, many studies have concentrated on objective methods of classifying NSLBP [4]. Yet, the utilisation of Artificial Intelligence (AI) to extract motion information from standard videos for spinal functions classification still is relatively unexplored. Traditional approaches for gathering human movement data have relied on wearable devices and common motion capture systems such as Vicon™ [6]. Although the motion capture using Vicon™ [6] offers precision, collecting and processing its voluminous data needs specialised ability and substantial computing power and resources. E-skin [11], an emerging LBP data collection tool, had also attracted people's attention. However, the requirements of the laboratory environment and its high cost limited its daily clinical use. Inertial measurement units (IMUs) [10] offer a promising clinical tool for data collection and classification in NSLBP. However, their complex data processing requirements currently limit their practical use in clinical settings.

1.3. Artificial Intelligence (AI) application to spinal function classification

Integrating computer vision into LBP classification marks a major advance in the clinical application of AI. Hartley et al. [14] used Human Posture Estimation for motion analysis but did not consider the classification of spinal function considering the motion characteristics (velocity and acceleration) as potential indicators of level of spinal function.

Much of the computer vision research related to LBP has been focused on biomedical imaging [19]. While imaging from professional medical equipment can provide intricate details about the spine, this approach overlooks the multifaceted aetiology of LBP often lacking patho-anatomical or radiological abnormalities [2, 4]. The associated costs, potential radiation exposure, and their inability to record dynamic movements further constrained the applicability of imaging particularly to inform management of NSLBP. In contrast, our study utilises data from a standard video-camera combined with AI-driven analysis to significantly reduce invasiveness and patient risks while offering a more comprehensive view of motion dynamics [9, 18].

The presented approach combines state-of-the-art computer vision technologies with sophisticated machine learning algorithms, establishing a new standard for classifying spinal function. The following section offers a concise overview of HPEs (Human Pose Estimation), delineating their function and significance within the framework of this research.

1.4. Automatic Human Pose Estimation

Human Pose Estimation (HPE) is a critical area in computer vision focused on identifying human figures in visual media and simplifying their poses. HPE models such as OpenPose [20], AlphaPose [21], among others, have made significant contributions.

HigherHRNet [22] stands out in this field for its superior ability to capture the dynamics of posture, especially in bending movements. It employs a bottom-up method to accurately locate 17 key anatomical points in images or videos. In our study, HigherHRNet [22] was utilised to precisely detect essential points which are the neck, hip, and ankle in patient videos, as key points to accurately assess spinal function. This targeted approach facilitates the extraction of motion features essential for our analysis.

2. Method

This section outlines our method, encompassing data collection, clinical expert classification, and the AI classification model for HPE. It covers feature extraction and organization, alongside the model's training and evaluation. The study focused on automatically classifying spinal function in NSLBP patients by primarily analysing their performance during forward flexion movements, which is most affected functionally.

2.1. Dataset

A video dataset [15] was used for model training and validation. It included 83 individuals (n=47 females, mean age 44.7 years [SD=11.8, Range 22-76 years old]; height 170 cm [SD=9.9cm, Range 153cm-188cm]; mass 81.3kg [SD=16.7kg, Range 46kg-123kg]) with LBP lasting over 3 months, independently categorised into High Function (HF; n=48) or Low Function (LF; n=35) by two expert physiotherapists [4, 6, 7, 15], with specifics of the assessment described previously [15].

Scaled videos of forward flexion movements, documented over ten repetitions and proven effective for classification were adjusted to 768 pixels on the shortest side and played at 30 frames per second. Patients' faces were blurred for anonymity using a self-written script. All following procedures was performed on the processed dataset. Videos were categorised into two tags based on expert function classification and stored in separate folders. Only authorized researchers have access to dataset.

Table 1 - Sample size of dataset

Classification	Number
High Function (HF)	48
Low Function (LF)	35
Total	83

(a)

(b)

Fig. 1. (a) The movement of a patient with high function (HF) during a 10-repetition forward flexion test was recorded by a camera. (b) The movement of a patient with low function (LF) during the same test was recorded.

2.2. Motion Features Extraction

Extracting motion features followed the method detailed previously [15], extending it to new velocity-based and acceleration-based features and applying it to the new task of HF/LF classification.

Motion features were derived by tracking the position of three key points (neck, hip and ankle) throughout the video recording using HighHRNet [22]. This process involved calculating the Euclidean distance (in pixels) between the hip and neck (hn), ankle and hip (ah), and ankle and neck (an), as shown in Fig. 2. These measurements are used to understand the motion dynamics of an object.

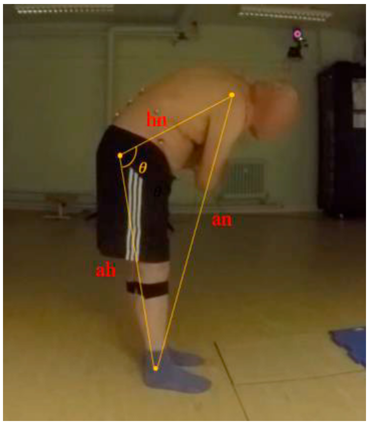


Fig. 2. The Euclidean distances in pixels between hip to neck (hn), ankle to hip (ah) and ankle to neck (an) were calculated.

In the dataset [15], individuals were side facing the camera while performing a forward flexion test. The HPE model was used to detect and connect the neck, hip and ankle keypoints. The angle between the line connecting the hip and neck and the line connecting the hip and ankle was defined as the forward bending angle, calculated as follows:

$$\theta = \frac{\cos^{-1}(hn^2 + ah^2 - an^2)}{2 \cdot (hn \cdot ah)} \tag{1}$$

where

$\cos^{-1}$ - Inverse function of cosine function, hn - The line from hip to neck, ah - The line from ankle to hip, an - The line from hip to neck, θ - The angle between the lines, hn and ah.

After obtaining the waveform diagram of the angle changing with time, the velocity and acceleration of the patient's forward flexion can be obtained by calculating the first-order derivative and the second-order derivative with respect to time, respectively using the following formulas:

$$V = \frac{d\theta}{dt} \tag{2}$$

where

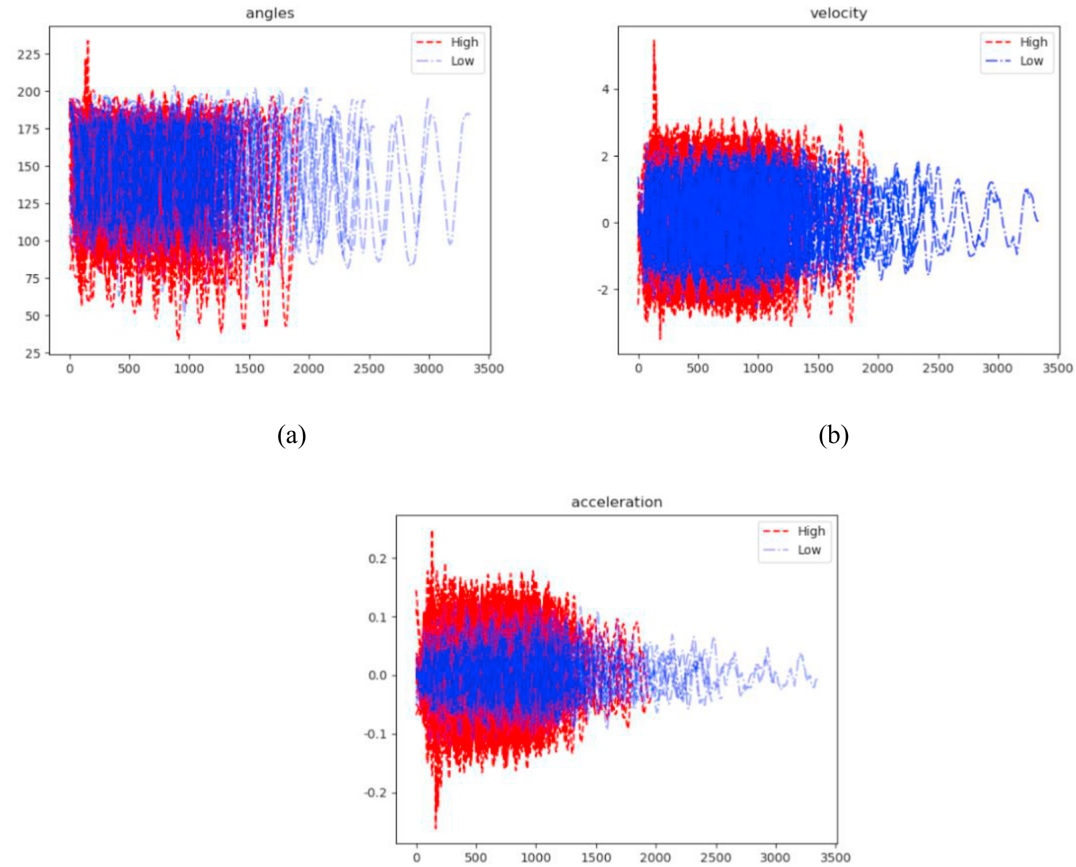
V - velocity, $d\theta$ - Differentiation of angle (θ), dt - Differentiation of time (t).

$$a = \frac{dV}{dt} = \frac{d^2\theta}{dt^2} \tag{3}$$

where

a - acceleration, dV - Differentiation of velocity, dt - Differentiation of time (t), $d^2\theta$ - Second differential of angle, dt^2 - Second differential of time (t).

The moving average of a 30-frame window (equivalent to one second of video data) was used to enhance the smoothness of the angle, velocity, acceleration waveforms. The calculated statistical features on the smoothed angular waveform are shown in Table 2. The waveforms of all patients processed by this filter are shown in Fig. 3 shown.



(a) (b)

(c)

Fig. 3. Visualization of data of the same type of feature for 83 patients on High/Low. (a) angle waveform; (b) angular velocity waveform; (c) angular acceleration waveform.

These statistical features were used to distinguish HF and LF within NSLBP patients. Minimal, Maximal, Standard Deviation, Variance, Range are some conventional statistical features. The spine bending Repetition Time (RT) and Repetition Time Variance (RTV) were derived from the frame count between these minimum points on the angle plot [15]. Depth Variance (DV) provided insight into how consistently patients achieve a specific flexion depth and the extent to which flexion depth fluctuates during repetitions. Stability Time (ST) was obtained by dividing the waveform into two halves, determining the angular range of each half, and then subtracting one half from the other. This feature can provide insight into whether patients increase their maximum bending angle over multiple repetitions. More details on calculating these features refer to the study by Hartley et al. [15].

Table 2 – Motion statistical features used for classification. All features were obtained from angle, velocity and acceleration waveforms.

1. Minimal (full spine flexion angle)	**6. Stability Time (ST)**
2. Maximal (full spine standing angle)	**7. Depth Variance (DV)**
3. Standard Deviation (SD)	**8. Repetition Time (RT)**
4. Variance	**9. Repetition Time Variance (RTV)**
5. Range	

2.3. High Function (HF) and Low Function (LF) Classification Model

Considering that changes in movement patterns can reflect differences in spinal function, the changes in movement speed and acceleration during the forward bending test in patients with NSLBP informed the extraction of speed-related features [23].

A feedforward neural network was used to perform prediction of spinal function classification using only motion features, with parameter values extracted from labeled instances. The architecture and process of the model are shown in Fig. 4. The model was composed of three identical branches that accept different inputs, processed by the same batch normalization layer, linear layer, and a LeakyReLU activation layer [24] before the final linear layer. Each linear layer consists of the 9 neural nodes, the same number as statistical features. Two identical self-attention mechanism modules [25] were placed before and after the final linear layer.

The attention mechanism in each branch utilized a 1D convolution layer to generate a query from the feature dimension of the input. This query undergoes a softmax operation to form an attention map, which is then element-wise multiplied with the original input to yield an output that emphasizes the most informative features. This focused processing helps in effectively distinguishing between two subgroups of spinal function.

For each motion feature, the model accepted the 9 statistical features detailed in Table 2 to classify the spinal function and generate the prediction result of LF or HF (0 or 1), respectively.

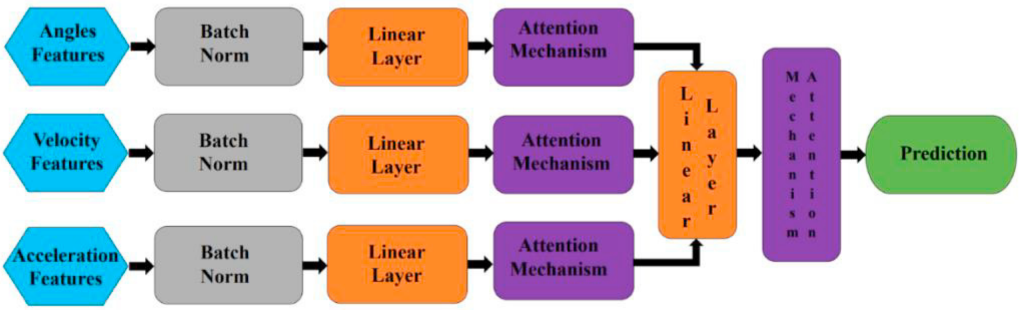


Fig. 4. Feedforward neural network for spine function classification.

2.4. Evaluation

In evaluating AI model, four key metrics are important: accuracy, sensitivity, specificity, and F1 score [26]. Sensitivity represents the model's ability to accurately identify all positive class (LF) individuals, while specificity measures the model's ability to correctly identify all negative class (HF) individuals. These metrics are, using true positive detections (TP), true negative detections (TN), false positive detections (FP) and false negative detections (FN). Their calculation methods are outlined as follows:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$Specificity = \frac{TN}{FP+TN} \tag{6}$$

Accuracy reflects the model's correct classification rate, while sensitivity and specificity gauge its capability to accurately identify specific categories. The F1 score, a harmonic mean of precision and recall, assesses the balance between identifying true positives and the model's overall accuracy in pinpointing the correct class, offering a comprehensive measure of performance. Its range fluctuates between 0 and 1, and the higher the value, the better the performance of the model.

$$F1\ Score = \frac{2TP}{2TP+FP+TN} \tag{7}$$

where:

TP - True positive detections, FP - False positive detections, TN – True negative detections.

In the evaluation, 5-Fold cross-validation, a technique to maximize data use in model training while ensuring model reliability and generalizability, was used [27]. Data was divided into five approximately equal parts; four were used for training and one was used for testing, rotating until each part has been used as a test set.

A perturbation-based method was employed for assessing the significance of various inputs within our artificial intelligence model. This approach is predicated on the premise that marked fluctuations in performance metrics are indicative of substantial feature importance. It involved systematically altering input features and observing the effects on model performance. Specifically, we perturbed each feature by nullifying its value and then assessed the resulting changes in model accuracy or loss. This procedure was executed for each feature in isolation, culminating in a hierarchical ranking that highlights the contribution of each feature to the model's decision-making process.

2.5. Implementation details

The proposed model was trained for classifying HF/LF based on motion features using a single NVIDIA 3060Ti GPU. The hyperparameters of model were optimised: learning rate (lr) at 0.1, batch size set to 256, over 300 epochs, with momentum at 0.9, weight decay of 0.01, dropout rate at 0.5, utilising stochastic gradient descent (SGD) as the optimiser and BCEWithLogitsLoss for the loss function. To ensure robustness, each fold underwent 10 training runs, with the iteration yielding the optimal performance and minimal loss being retained [27]. This rigorous training regimen was complemented by exhaustive experimentation with various feature combinations, aiming to enhance the model's accuracy and generalisability. The decrease in loss and increase in accuracy during training and testing are shown in the left and right sides of Fig. 5 respectively.

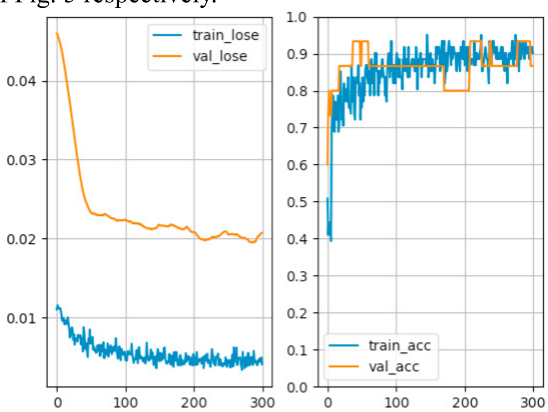


Fig. 5. Loss and accuracy changes during training and testing.

3. Results

3.1. Classification Performance

In this section, the classification performance results obtained using classification models exploiting different features are presented in Table 3. The results were averaged across all metrics for 5-fold cross-validation.

Employing 'Velocity' as a feature yields superior accuracy, sensitivity, specificity, and F1 scores compared to other individual features or their combinations. Conversely, the traditionally emphasized feature 'Angle' shows a relatively poor performance in accurately identifying HF, leading to lower overall accuracy and F1 score. This suggests that 'Angle' may have a limited or even negative impact on classification, a finding that is consistent across other feature combinations that incorporate 'Angle.' Notably, the utilisation of all three features—'Angle,' 'Velocity,' and 'Acceleration'—does not outperform the use of 'Velocity' and 'Acceleration' individually or their combination (refer to Table 3).

Table 3 - Performance of spinal functions classification using 9 motion features.

Feature	Accuracy	Sensitivity (LF)	Specificity (HF)	F1 Score
Angle	87.76%	91.43%	85.33%	0.8961
Velocity	**95.13%**	**93.81%**	**96.00%**	**0.9442**
Acceleration	91.38%	90.95%	91.78%	0.9124
Angle + Velocity	87.93%	76.77%	95.78%	0.9016
Angle + Acceleration	91.51%	87.78%	94.00%	0.8991
Velocity + Acceleration	94.12%	91.43%	96.00%	0.9200
All three features	92.44%	88.45%	93.00%	0.8965

in which "Sensitivity" denotes accuracy of correctly assigning low function (LF), and "Specificity" denotes accuracy of correctly assigning high function (HF). "All three features" denotes 9 statistical features each of the three motion features - Angle, Velocity and Acceleration, being used together.

3.2. Importance of Features

The nine statistical features calculated for the three motion features and their importance to the model were shown in Table 4 and Fig. 4.

In the normalised importance analysis of motion features, the minimum angle feature showed the highest importance score (0.3691), suggesting that extreme angle variations are pivotal for model predictions. Following closely were the maximum angle value (Max) with a score of 0.1825 and the standard deviation (SD) at 0.1768, indicating the model's sensitivity to angular variability and response. The velocity's repetition time (RT) scored the highest in importance (0.4682), highlighting its critical role in differentiating categories based on instantaneous speed changes. Additionally, the maximum value (Max) and range (Range) of velocity also showed significant importance, with scores of 0.3007 and 0.0598, respectively. For acceleration features, the maximum value (Max) was most significant (0.2613), while the minimum value (Min) and standard deviation (SD) were also noteworthy, with scores of 0.2200 and 0.1910, respectively, highlighting the importance of peak and trough patterns in acceleration for classification. These results reveal that extreme values and variability within motion features are crucial indicators for class distinction in the model.

Table 4 - Importance of Statical Features.

Statistical Features / Motion Features	Minimal	Maximal	SD	Variance	Range	ST	DV	RT	RTV
Angle	**0.3691**	0.1825	0.1768	-0.0135	0.0681	0.0239	0.0702	0.1458	-0.0230
Velocity	0.0910	0.3007	0.0540	0.0040	0.0598	0.0090	0.0131	**0.4682**	0.0002
Acceleration	0.2200	**0.2613**	0.1910	0.0073	0.0684	0.0198	0.0004	0.2460	-0.0142

The accompanying bar chart illustrates the normalised importance of the statistical features for angle, velocity, and acceleration. It highlights that the minimum angle possesses the predominant weight in model predictions, while for velocity, the response time (RT) emerges as the most distinguishing factor. The chart further visualises that the maximum value of acceleration holding considerable significance in its influence on the model's classification capability.

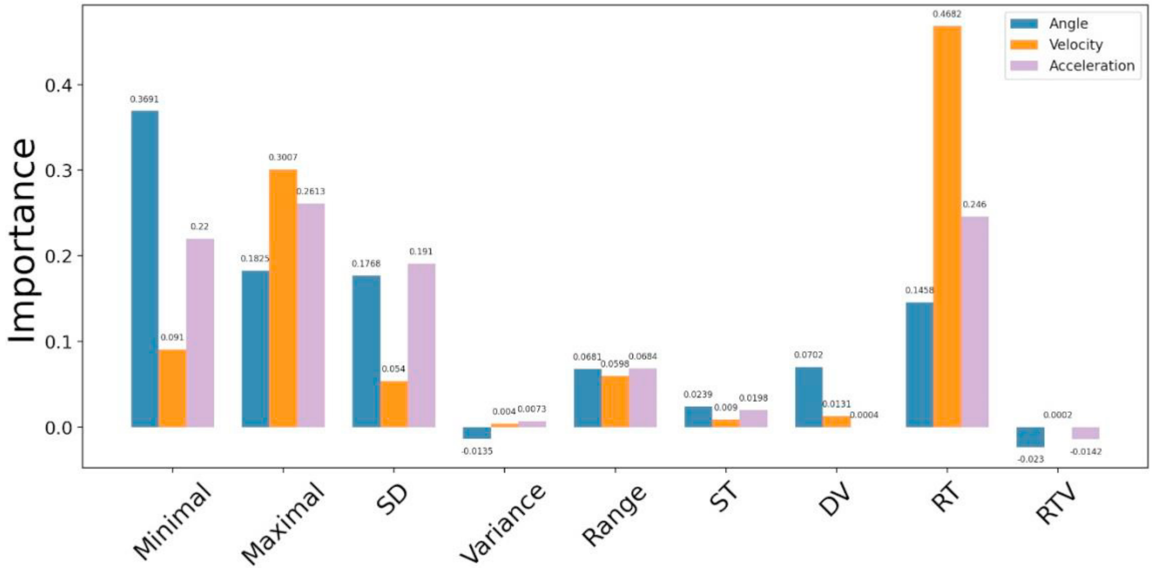


Fig. 4. Importance of Statical Features.

4. Discussion

The results from the SpineSighter classifier highlight its potential to significantly impact clinical practices by accurately differentiating levels of spinal function impairment in NSLBP patients, achieving a peak accuracy of 95.13% and an F1 score of 0.9442. The following sections explore the impact of various motion features in distinguishing levels of function and explore the application of AI models for spinal classification using standard video recordings.

4.1. The impact of different movement features on classification accuracy

Temporal movement features, specifically velocity and acceleration, showed superior accuracy in differentiating levels of spinal function, achieving accuracies of 95.13% and 91.38%, respectively. These measures outperformed angular displacement, which had an accuracy of 87.76%. The superior classification performance of temporal movement features may stem from its comprehensive encapsulation of the continuous dynamic nature of spinal motion, which is crucial for distinguishing different levels of function and spinal impairment [8]. Unlike static angle measurements, temporal features indicating a rate at which spine changes its position over time, provide unique insights into the fluidity and efficiency of spinal motion. These findings are also consistent with biomechanical assessment of lumbar spine, where higher-order kinematics including velocity and acceleration can describe lumbar spine function in patients with LBP in more detail than traditional static measurements [8]. In contrast, study by Williams et al. [28] found that reducing pain through oral analgesia did not significantly affect angular velocity or acceleration in individuals with acute and chronic LBP, challenging the role of pain in altering lumbar kinematics. Clearly, further study of kinematics for varying degrees of spinal function and its interactions with pain is required.

The velocity signatures capture the essence of spinal dysfunction that static angle features may not reflect [29]. Using this feature, the SpineSighter was slightly better at correctly classifying HF than LF as indicated by the measures of sensitivity (96%) and specificity (93.81%), respectively. Velocity's success in accurately classifying high-functioning individuals with LBP may be related to its reflection on movement efficiency and control, which are likely to be preserved in a pain-free spine. Research shows that high-functioning individuals with LBP of pain free controls usually maintain a more efficient and better movement pattern [6, 8, 28]. Further research showed the role of temporal features such velocity and acceleration, in differentiating pathological and non-pathological states of spinal motion [30].

Acceleration, representing the rate of change in velocity of the angular displacement, captures the start and stop of motion but is not in continuous motion like velocity, thereby making it an intermediate classification performer between a static angle and velocity. This was echoed in the biomechanical study by Reeves et al. [31], emphasising the sensitivity of acceleration to changes in dynamic stability of the spine.

Angle displacement was the least accurate in distinguishing between different levels of spinal functionality; however, its relative success in identifying individuals with NSLBP at a 91.43% accuracy rate may stem from its direct measurement of range of motion. This measurement is often limited to patients with LBP and impaired spinal functions [8], aiding in the detection of severe functional limitations.

4.2. Clinical implications

Our SpineSighter classifier goes beyond traditional, subjective assessment methods by providing an objective, efficient alternative for spinal functions analysis, promising substantial improvements in clinical NSLBP management and patient self-care. Clinically, the findings emphasise the superiority of temporal movement features like velocity and acceleration in assessing spinal function using video-recordings, outperforming traditional angular displacement measures. These dynamic metrics offer a more nuanced view of spinal function, suggesting their utility as reliable biomarkers for evaluating spinal health and treatment efficacy.

Successful integration of the SpineSighter classifier into daily clinical practice does not require complex solutions to technical and educational aspects. An ordinary mobile phone with a camera function serves as a recording device, capable of capturing detailed patient movements at an appropriate frame rate. Patients do not need to receive in-depth training. They not only need to understand the technical operation of the recording equipment and software, but also need to be informed of the distance from the fixed recording equipment before use and maintain appropriate lighting (conditions that enable the outline of the human body to be clearly seen), which can ensure consistent data quality. The model has no special clothing requirements for patients, and it does not require whether the face is visible. And the video will be encoded before the model processes it to prevent leakage. The classification results can be obtained in a short time. This capability could revolutionise treatment approaches, offering more precise and dynamic assessments compared to static measurements.

4.3. Strengths and Limitations

The study highlights the transformative potential of AI in refining spinal function classification and treatment methodologies for NSLBP. The SpineSighter classifier offers an objective, efficient alternative to traditional subjective assessment methods, promising to enhance clinical management on NSLBP.

However, there are certain limitations. A key challenge is the reliance on video data, which requires high-quality recordings and precise patient positioning for accurate feature extraction. Variabilities in lighting, camera angles, and patient movement can compromise data integrity. Insufficient lighting will bring challenges to the HPE model's accurate positioning of key points on the human body. Changes in camera angles during video recording and the patient's relative movement to the camera will make feature capture unreliable. Although the HigherHRNet used in this study can overcome such barriers to a large extent, it is necessary to be informed of the precautions before actual use.

Additionally, the AI model's effectiveness and its generalisability across diverse patient populations hinge on the training dataset quality. There is a risk of inherent biases if the dataset is not representative of the wider population, which could skew the AI's decision-making processes, favoring certain patient groups over others. This could lead to disparities in the predictive accuracy and efficacy of treatment recommendations derived from the AI model.

Neural network complexity also poses interpretability issues, potentially complicating the understanding of predictive mechanisms.

5.3. Future Work

Future research should address the limitations identified in this study by diversifying the training datasets with high-quality recordings to better represent varied patient demographics, enhancing model generalisability and robustness. Enhancing AI model interpretability is essential to increase clinicians' trust and understanding of AI-based diagnostics. Investigating additional biomechanical metrics, such as joint torques or muscle activation patterns, may enrich the current analysis of movement features. These efforts could lead to more accurate and tailored NSLBP assessment and treatment strategies, significantly advancing the role of AI in enhancing healthcare for people with non-specific pain conditions.

5. Conclusion

The study on the SpineSighter classifier shows its high efficacy in differentiating levels of spinal function in patients with NSLBP, achieving impressive accuracy and F1 scores. This highlights the potential of AI models in clinical settings to advance the understanding and management of spinal function based on motion features captured from standard video recordings. The results show that velocity stands out as a particularly effective feature for this type of classification, surpassing acceleration and angle motion features in terms of accuracy, sensitivity and specificity. The findings show the ability of AI to improve assessment accuracy, facilitate personalised care in clinical settings and offer self-management for people with NSLBP. This work lays an excellent foundation for integrating AI into spinal functional assessment and marks a step forward in advancing the treatment and management of spinal health-related conditions.

Reference

[1] M. van Tulder, B. Koes, and C. Bombardier, ‘Low back pain’, Best Practice & Research Clinical Rheumatology, vol. 16, no. 5, pp. 761–775, Dec. 2002, doi: 10.1053/berh.2002.0267.

[2] Maher C, Underwood M, Buchbinder R. Non-specific low back pain. The Lancet. 2017;389(10070):736-47. doi: https://doi.org/10.1016/S0140-6736(16)30970-9.

[3] NICE. National Institute of Health and Care Excellence, Low back pain and Sciatica over 16s: Assessment and Management (NICE guideline NG59)2016 (last updated 2020). Available from: https://www.nice.org.uk/guidance/ng59.

[4] O’Sullivan P. Diagnosis and classification of chronic low back pain disorders: Maladaptive movement and motor control impairments as underlying mechanism. Manual Therapy. 2005;10(4):242-55. doi: https://doi.org/10.1016/j.math.2005.07.001.

[5] J. R. Beazell, M. Mullins, and T. L. Grindstaff, ‘Lumbar instability: an evolving and challenging concept’, Journal of Manual & Manipulative Therapy, vol. 18, no. 1, pp. 9–14, Mar. 2010, doi: 10.1179/106698110X12595770849443.

[6] R. Hemming, L. Sheeran, R. van Deursen, and V. Sparkes, ‘Non-specific chronic low back pain: differences in spinal kinematics in subgroups during functional tasks’, Eur Spine J, vol. 27, no. 1, pp. 163–170, Jan. 2018, doi: 10.1007/s00586-017-5217-1.

[7] Sheeran L, Sparkes V, Caterson B, Busse-Morris M, van Deursen R. Spinal Position Sense and Trunk Muscle Activity During Sitting and Standing in Nonspecific Chronic Low Back Pain Classification Analysis. Spine. 2012;37(8): E486-E95. doi: Doi 10.1097/Brs.0b013e31823b00ce. PubMed PMID: WOS:000302870700005.

[8] van Dieën JH, Reeves NP, Kawchuk G, van Dillen LR, Hodges PW. Motor Control Changes in Low Back Pain: Divergence in Presentations and Mechanisms. J Orthop Sports Phys Ther. 2019 Jun;49(6):370-379. doi: 10.2519/jospt.2019.7917. Epub 2018 Jun 12. PMID: 29895230; PMCID: PMC7393576.

[9] Parashar G, Chaudhary A, Rana A. Systematic Mapping Study of AI/Machine Learning in Healthcare and Future Directions. SN computer science. 2021;2(6):461-. doi: 10.1007/s42979-021-00848-6.

[10] Z. Bacon, Y. Hicks, M. Al-Amri, and L. Sheeran, 'Automatic Low Back Pain Classification Using Inertial Measurement Units: A Preliminary Analysis', Procedia Computer Science, vol. 176, pp. 2822–2831, 2020, doi: 10.1016/j.procs.2020.09.272.

[11] Y. Zhang et al., 'Electronic Skin Wearable Sensors for Detecting Lumbar–Pelvic Movements', Sensors, vol. 20, no. 5, p. 1510, Mar. 2020, doi: 10.3390/s20051510.

[12] D. Robertson, 'Vicon Motion Capture'. Jul. 18, 2013. http://www.vicon.com/

[13] Jin, Heeku Analysis of sitting posture using wearable sensor data and support vector machine model. Med.-Leg. Update 1, 334–338 (2018).

[14] Anthony C, Ankit P, Tasos K, Cotton RJ. Portable in-clinic video-based gait analysis: validation study on prosthetic users. medRxiv. 2022:2022.11.10.22282089. doi: 10.1101/2022.11.10.22282089.

[15] T. Hartley, Y. Hicks, J. Davies, D. Cazzola, and L. Sheeran, 'BACK-to-MOVE: Machine Learning and Computer Vision Model Automating Clinical Classification of Non-Specific Low Back Pain for Personalised Management.', accepted PLOS ONE 2 in production.

[16] Molnar, C. (2020). Interpretable machine learning. Lulu.com.

[17] O'Sullivan P. Clinical instability of the lumbar spine: its pathological basis, diagnosis and conservative management. In Boyling JD, Jull G, eds. Grieve's Modern Manual Therapy, 3rd ed. Amsterdam: Elsevier, 2005:311–22.

[18] M. Vaisy, L. Gizzi, F. Petzke, T. Consmüller, M. Pfingsten, and D. Falla, 'Measurement of Lumbar Spine Functional Movement in Low Back Pain', The Clinical Journal of Pain, vol. 31, no. 10, pp. 876–885, Oct. 2015, doi: 10.1097/AJP.0000000000000190.

[19] D'Antoni F, Russo F, Ambrosio L, Vollero L, Vadalà G, Merone M, Papalia R, Denaro V. Artificial Intelligence and Computer Vision in Low Back Pain: A Systematic Review. Int J Environ Res Public Health. 2021 Oct 17;18(20):10909. doi: 10.3390/ijerph182010909. PMID: 34682647; PMCID: PMC8535895.

[20] Martınez G H. Openpose: Whole-body pose estimation[D]. Carnegie Mellon University, 2019.

[21] Fang H S, Li J, Tang H, et al. Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022.

[22] B. Cheng, B. Xiao, J. Wang, H. Shi, T. S. Huang, and L. Zhang, 'HigherHRNet: Scale-Aware Representation Learning for Bottom-Up Human Pose Estimation'. arXiv, Mar. 12, 2020. Accessed: May 28, 2023. [Online]. Available: http://arxiv.org/abs/1908.10357.

[23] Sihvonen, T., Partanen, J., Hänninen, O., Soimakallio, S. (1997). Electric behavior of low back muscles during lumbar pelvic rhythm in low back pain patients and healthy controls. Archives of Physical Medicine and Rehabilitation, 78(12), 1325-1330.

[24] Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In Proc. icml (Vol. 30, No. 1, p. 3).

[25] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

[26] Powers, D. M. W. (2011). Evaluation: From Precision, Recall and F-Measure to ROC (Receiver Operator Curve), Informedness, Markedness & Correlation. Journal of Machine Learning Technologies, 2(1), 37-63.

[27] Bengio, Y., & Grandvalet, Y. (2004). No unbiased estimator of the variance of k-fold cross-validation. Journal of Machine Learning Research, 5(Sep), 1089-1105.

[28] Williams JM, Haq I, Lee RY. An experimental study investigating the effect of pain relief from oral analgesia on lumbar range of motion, velocity, acceleration and movement irregularity. BMC Musculoskelet Disord. 2014 Sep 16; 15:304. doi: 10.1186/1471-2474-15-304. PMID: 25228332; PMCID: PMC4171543.

[29] Hebert, J.J., Koppenhaver, S.L., Magel, J.S., Fritz, J.M. (2018). The relationship of lumbar multifidus muscle morphology to previous, current, and future low back pain: a 9-year population-based prospective cohort study. Spine, 43(15), 1215-1222.

[30] Mok, N.W., Brauer, S.G., Hodges, P.W. (2004). Hip strategy for balance control in quiet standing is reduced in people with low back pain. Spine, 29(6), E107-E112.

[31] Reeves, N.P., Cholewicki, J., Silfies, S.P. (2009). Muscle activation imbalance and low-back injury in varsity athletes. Journal of Electromyography and Kinesiology, 19(4), 707-715.