Journal of Human Centered Technology Vol. 1: No. 2 (2022) 33 – 43

Journal of Human Centered Technology

E-ISSN: 2821-3467
https://humentech.utm.my

Cluster Analysis of Biomechanical Gait Data and Pain Score as a Potential Classification of Severity in Knee Osteoarthritis

Husnir Nasyuha Abdul Halim[1], Aizreena Azaman[1,2*], Izwyn Zulkapri[1,2]

[1]School of Biomedical Engineering and Health Sciences, Faculty of Engineering, Universiti Teknologi Malaysia, 81310, Johor Bahru, Johor, Malaysia

[2]Sport Innovation and Technology Centre (SITC), Institute of Human-Centered Engineering (iHumEn), Faculty of Engineering, Universiti Teknologi Malaysia, 81310, Johor Bahru, Johor

*Corresponding Author aizreena@utm.my

Received 31 March 2022; Accepted 18 June 2022; Available online 06 August 2022
https://doi.org/10.11113/humentech.v1n2.19

Abstract:
Osteoarthritis (OA) is the most common type of arthritis affecting approximately 240 million people globally, with increasing prevalence with age. The knee is the most prevalent joint affected by OA and it causes physical disability and decreased motor function which consequently affects the activity of daily living including mobility. Pain is the main symptom that is characterized in OA, which is measured using self-rated scales or questionnaires to determine several aspects of the pain including the intensity, frequency, and pattern. Quantifying pain is a standard clinical practice to diagnose and monitor symptomatic OA, however, its application for severity assessment is not well explored. To date, the severity assessment of knee OA is only by radiographic severity assessment that does not necessarily reflect the symptomatic OA. In this study, gait analysis was performed on symptomatic knee OA patients. Distinctive gait kinematic features were extracted using principal component analysis (PCA). Pain score and the gait features including spatiotemporal and kinematics were used for clustering analysis. Two clustering algorithms, K-means and K-medoids were conducted to cluster samples with similar features to assess knee OA characterization. The clustering solutions were evaluated based on three measures which are the Davies Bouldin index, Calinzki Harabasz index, and Silhouette index. This study discovered that majority of the datasets which is 5 out of 9 datasets had the best performance (fulfill at least 2 out of 3 performance index criteria) when the number of clusters, *k* is 4 and using the k-means algorithm. These clustering models can be used in the future as the labeling class of symptomatic knee OA that is based on pain and gait characteristics of knee OA. Future studies are suggested to test other pain assessment scores, include other gait features such as kinetic and muscle activity features, and employ various types of feature selection methods to improve the clustering performance.

Keywords: Knee osteoarthritis; Gait, Pain, Clustering analysis; Rehabilitation

1. Introduction

Among various types of musculoskeletal disorders, osteoarthritis (OA) is one of the most common and the knee is the most reported joint affected [1]. Knee OA is commonly associated with aging as the prevalence increases with increasing age groups [2]. OA is a degenerative and progressive joint disease that causes metabolic, structural, and mechanical changes to every part of the joint including articular cartilage, subchondral bone, ligaments, and tissues surrounding the joint [1, 3]. The changes mainly cause pain at the joint, therefore being the must-have clinical criterion to diagnose OA according to established international standards such as the American College of Rheumatology (ACR) and European League Against Rheumatism (EULAR) [4, 5].

As knee OA is a progressive disease, severity assessment of OA is crucial for disease management. Although knee OA can be described as either radiographic OA; OA that is diagnosed by anatomical and structural changes of radiographic observations, or symptomatic OA; OA that is diagnosed based on clinical symptoms [6, 7], the current severity classification that is available is only through radiographic classification which most widely used is Kellgren/Lawrence (K&L) grading that classifies severity into four severity grades [8].A systematic literature review reported discordances between the main symptom of OA, knee pain with radiographic OA in which the proportion of patients with knee pain that have radiographic OA ranged between 15% to 76% while 15% to 81% radiographic OA patients had experienced knee pain [7]. A study found that pain experienced is associated differently with K/L radiographic severity between grade 3 and 4 [9]. Another study reported pain absence of up to 31.2% among severe radiographic knee OA (K/L grade 4) populations [10]. Knee OA management includes nonpharmacologic treatment such as exercise, weight loss, mechanical interventions such as orthosis and the pharmacological approach such as nonsteroidal anti-inflammatory drug (NSAID) before proceeding with surgical modalities [11-13].

As mentioned before, the discordance between structural progression and pain exists, therefore an additional severity assessment based on symptom is needed. Moreover, the role of radiographic severity assessment in clinical perspective is more towards diagnosis compared to monitoring purpose [14]. Beside repeated X-ray procedure will expose patient to unnecessary radiation, the structural changes observed from radiographic images is more relevant to be assessed during late stage of severity for clinician to make the decision on knee replacement surgery. Recent recommendation of knee OA management is treatment plan that is individualized, evidence-based, and multimodal that target to relieve symptoms, delay progression, and improve joint function [15, 16] in which certainly not solely depended on structural progression.

To evaluate the main symptom of knee OA, knee pain, there are various types of outcome measure that includes pain as a subscale for example the most widely used, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). The extension version of WOMAC is Knee Injury and Osteoarthritis Outcome Score (KOOS) which focused on grading the frequency and intensity of pain. The most recent pain assessment questionnaire developed is Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP) which grade different types of pain experienced. However, pain is deemed as a multifactorial phenomenon, subjective to individual therefore quantifying severity solely using pain scores is insufficient. Gait or walking pattern is one of the changes that can be observed and measured quantitatively among knee OA patients as a result of pain experienced. Commonly reported gait changes among knee OA in previous studies include decreased walking speed, cadence, step and stride length, joint range of motion; and increased step and stride time, step width, and muscle activation [17-20].

Many studies reported association between these knee OA gait characteristics with knee pain for example study by Nebel et al. that found correlation of pain measured with Arthritis Impact Measurement Scales (AIMS) and WOMAC with speed, stride length, knee range of motion and peaks vertical force [21]. Another study reported significant correlation between knee adduction moment at different gait phases with pain measured using Japanese Knee Osteoarthritis Measure (JKOM) [22]. A study by Marriot et al. investigated association of gait with pain using both univariate and multivariate analysis and discover correlation of knee varus angle, knee flexion and extension angle with pain [23]. Other than that, kinetics features which are knee adduction impulse moment, peak knee flexion moment, peak knee extension moment were the features included in final model of multivariate analysis with change of pain as the dependent variable [23]. These examples proved the importance of relationship between main symptom of knee OA, knee pain and gait features for knee OA characteristics, however, pain and gait-based severity classification for knee OA is still not available.

The application of machine learning for knee OA in previous studies mainly aimed to improve radiographic classification or to distinguish between knee OA and able-bodied. The studies that utilized biomechanics or gait features mostly utilized supervised learning algorithms. For examples support vector machine (SVM) is used in studies to assess knee OA using spatiotemporal parameters [24], and combination of several spatiotemporal parameters and knee and ankle kinematic parameters during stair ascent [25]. Other than that, decision tree was performed on ground reaction force (GRF) and combination of spatiotemporal parameters, kinematic and clinical characteristics [26, 27]. Another ML algorithm used was neural network to predict K/L severity and pain level using hip, knee and ankle kinetics and kinematics features [28] and random forest based on 3D GRFs [29]. The application of these supervised ML requires pre-labelled data to characterize severity which mostly based on the radiographic severity. To date, only a few studies used unsupervised learning for instance study by Elbaz et al. using k-means clustering to cluster severity based on spatiotemporal parameters [30], and studies that used kinematic features to characterize gait profiles among later stages of knee OA [31] and among total knee arthroplasty candidates [32]. These few studies indicate that unsupervised learning can be used to characterize and cluster knee OA without depending on pre-labeled information.

However, there is still lack of studies that apply unsupervised learning such as clustering, that attempts to cluster severity based on both pain and different types of gait features. Using unsupervised learning, the gait pattern and new information on the characterization of the disease severity could be discovered, independent of radiographic severity grading. Among various unsupervised learning, partitioning clustering such as k-means and k-medoid are commonly used to cluster a dataset into separate clusters. Therefore, in this study, pain and gait features were collected among knee OA patients and clustering analyses were performed using pain score and the correlated gait features and to analyze the performance of 9 subsets of features that comprised of 54 different clustering solutions on the knee OA gait dataset using three clustering performance indexes.

2. Materials and Method

The overall study workflow is presented as the flowchart in Figure 1:

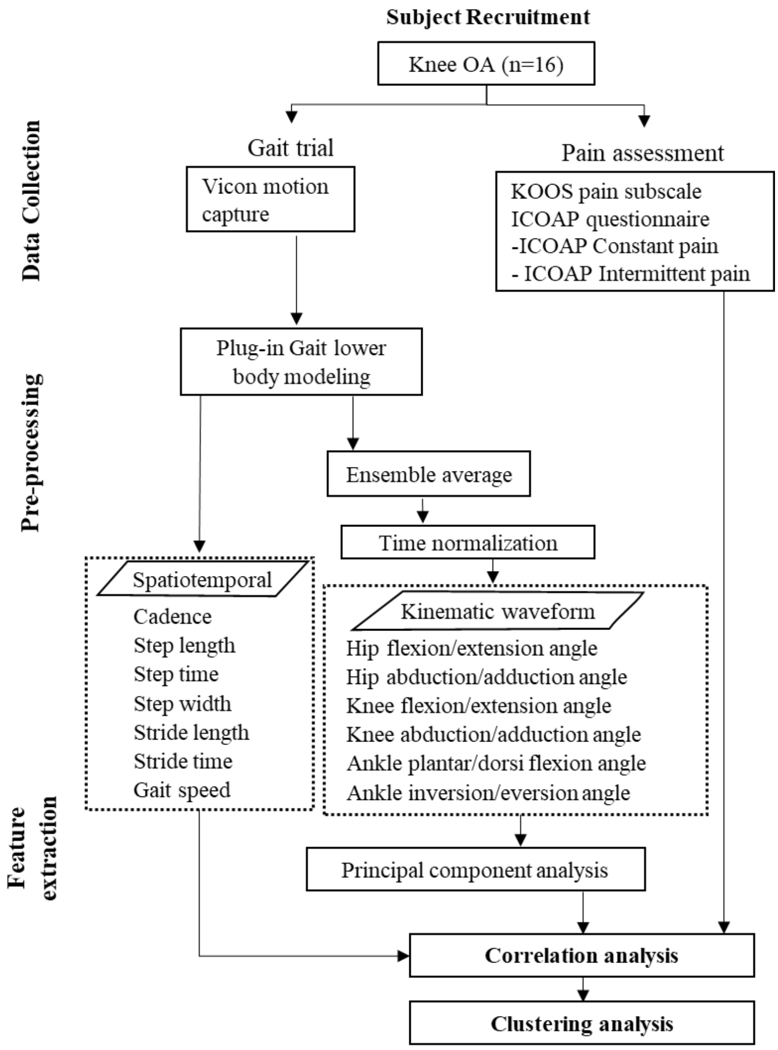

Figure 1. Overall study flowchart

2.1 Subject recruitment and data collection

Sixteen subjects were recruited with the following inclusion criteria: 1) diagnosed as knee OA patient with the presence of knee pain on either side of leg 2) aged ranging from 40-65 years old 3) able to walk without any mobility aid such as cane, walker, and crutches 4) less than 5 visits to physiotherapy sessions and with exclusion criteria: 1) have a history of lower limbs surgery 2) diagnosed with other neuromusculoskeletal disease/disorders 3) wearing any electronic medical devices. The subjects were all recruited from local physiotherapy clinics. There were 16 (12 females) knee OA patients recruited with age: 56.6 (SD =7.67) years old, height: 1.58 (SD =0.07) m, and weight 69.41 (SD = 12.81) kg.

2.2 Data collection

Prior to data collection, the procedure of data collection and informed consent form were presented to subjects and if the subjects consented, the subjects signed the form. The procedures of data collection and informed consent form were approved by the Malaysian Medical Research and Ethics Committee (MREC), Ministry of Health Malaysia with ID number NMRR-19-2586-46320(IIR). The data collection for each subject started with the recording of demographic (age, gender) and anthropometry (height, weight, leg length, knee width, ankle width) details.

Next, subjects were instructed to wear fit and tight outfits for the placement of 16 reflective markers at specified locations by Plug-in Gait for lower body model [33]. Once prepared, static capture of each subject was taken using 5 MX-Cameras (Vicon, Oxford, UK) while subjects standing at the center of volume to build subject's model on Vicon Nexus software. After that, subjects will be asked to walk straight at self-preferred speed along a 4.5m walkway without any footwear while marker data were recorded using the cameras. Lastly, subjects were instructed to answer Knee injury and Osteoarthritis Outcome Score (KOOS) and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP) questionnaires to assess their knee pain.

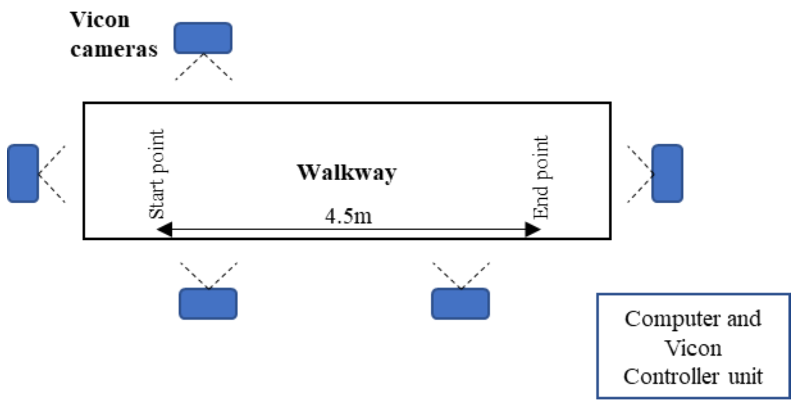

Figure 2. Experimental setup in gait analysis laboratory

Figure 3. Reflective marker and Vicon camera

2.3 Data processing

For unilateral knee OA patients, the data analyzed were only from the affected leg whereas, for bilateral knee OA patients, the data were analyzed from both legs. Each sample of data was comprised of the average of at least two gait cycles of each trial, in which selected gait cycles were the ones that heel strike and toe-off were detected. The pre-processing workflow include marker reconstructing and labeling, gap filling, and running pipeline in the Vicon Nexus software. The process generated lower limb kinematics (angle of hip, knee, and ankle on sagittal and frontal plane) and spatiotemporal data for the gait trials. The kinematic waveforms were time-normalized to produce 101 data points for each sample.

2.4 Feature extraction and computation

Kinematic features were extracted using principal component analysis (PCA). PCA is a dimensionality reduction technique that transforms a multidimensional dataset into uncorrelated variables called principal components (PCs) [34]. From the covariance matrix of the dataset, the PCs which are called eigenvectors were computed in hierarchical order from highest to lowest variance. The loadings of PCs determine the magnitude of each PC and its orthogonal projection called PC scores are its corresponding measure for each sample [34]. In this study, PC scores were used as a feature and the interpretation of the features was done using the representative extremes method, a method that compares high and low percentile samples as described in previous studies that utilized PCA for gait features extraction [35]. The retained PCs were the PCs that explained 90% of the variance. The spatiotemporal features computed were walking speed, cadence, step and stride length, step and stride time, and step width.

2.5 Statistical analysis

All statistical analyses were performed using SPSS (SPSS v26, IBM Corp, USA). Normality of features was tested using the Shapiro-Wilk test with a p-value ≤ 0.05. To select gait features that will be input for clustering analysis of each dataset, Spearman's Rank order correlation ($p<0.05$ (2-tailed)) was used to assess the association between the KOOS, ICOAP constant, and ICOAP intermittent pain scores with gait features extracted, in which selected features together with pain score were utilized as features for clustering.

2.6 Clustering analysis

In this study, unsupervised learning utilized was clustering. The clustering analyses were carried out in MATLAB (v.2021a, MathWorks, USA). The clustering analyses were done using two types of partitioning clustering algorithms, k-means, and k-medoids. Both k-means and k-medoids clustering purpose are to partition the dataset into clusters, in which every cluster is represented by a point, and to have the other cluster members with minimal distance to the point. For k-means the point is centroids, mean of the cluster members while for k-medoids, the center or the representative is medoid, an actual data point of the cluster [36]. The number of clusters tested were 2 to 4, with 20 repetitions for each number of clusters. There were three datasets of gait features tested with each pain score: 1) kinematic and spatiotemporal, 2) spatiotemporal only, and 3) kinematic only.

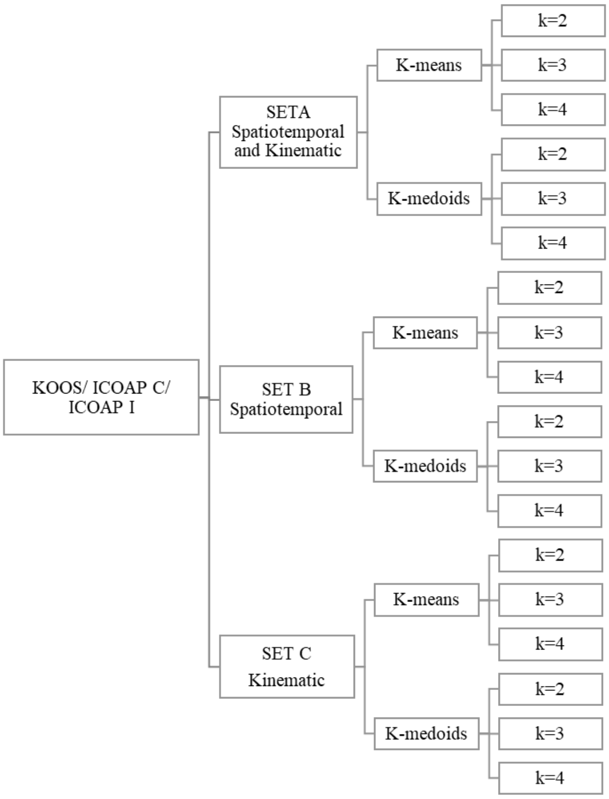

Figure 4. Clustering analysis

The clustering performance was assessed using three types of indexes: a) Davies Bouldin index, b) Calinski Harabasz index, and c) Silhouette index. The indexes are defined as in Table 1 and the clustering solutions were analyzed based on the index interpretation. The best clustering solution has atleast 2 of the criteria: the lowest value of Davies Bouldin index, and the highest Calinski Harabasz and Silhouette index.

Table 1. Clustering performance index

Clustering performance index	Definition	Interpretation of index on cluster performance
Davies Bouldin Index [37]	The ratio of the intra-cluster dispersion with separation of clusters	The lower the index, indicate the lesser the similarity between clusters
Calinski Harabasz Index [38]	The ratio of the variance of the sums of squares of the distances of individual objects to their cluster center with the sum of squares of the distance between the cluster centers.	The higher the index, the higher dispersion
Silhouette index [39]	The difference between the average distance between all clusters with an the average distance of individual cluster members divided by the maximum distance	The index is between -1 to 1, clusters towards -1 indicate clusters are assigned wrongly, toward 0 indicate clusters' distances is insignificant and toward 1 indicate that clusters are clearly distinguished

3. Results and Discussion

For each sample, there were 14 features extracted from frontal and sagittal kinematics of the ankle (5 features), hip (4 features), and knee (5 features) using PCA. All 14 kinematic features extracted, and 7 spatiotemporal features computed, were tested for normality using the Shapiro-Wilk test and revealed to be all not normally distributed ($p<0.05$). The Spearman's Rank order correlation analysis revealed the following significant correlation between pain score and gait features for selection of gait data subsets for each pain score.

Table 2. Pain score and its correlated gait features ($p<0.05$ (two-tailed))

KOOS pain subscale	ICOAP Constant pain (ICOAP C)	ICOAP intermittent pain (ICOAP I)
Cadence	Cadence	Stride Length
Step Length	Step Length	3PC1
Step Time	Step Time	3PC2
Step Width	Step Width	4PC1 (hip abduction/adduction overall magnitude)
Stride Length	Stride Length	5PC2
Stride Time	Stride Time	
Walking Speed	Walking Speed	
1PC2(ankle flexion phase shift)	3PC1	
3PC1 (hip flexion/extension overall magnitude)	3PC2	
3PC2 (hip flexion/extension range of motion)	5PC2	
5PC2 (knee flexion/extension range of motion)	6PC1 (knee abduction/adduction overall magnitude)	

The clustering performance for each of the datasets were listed in Table 3 - 5. The bold results indicate that the index values were the best within the sets. Based on the clustering performance index results, most of the datasets clustering solutions was with at least two criteria of best performance is when using k-means with the number of *k* being 4. There were some datasets in which the best cluster solutions were with *k*=4 but both k-means and k-medoids perform equally which are KOOS dataset B and all ICOAP I datasets. For sets A (both spatiotemporal and kinematic features) and sets C (kinematic only), clustering using features with KOOS performed the best according to Davies Bouldin and Calinski Harabasz index. On the other hand, for sets B (spatiotemporal features only), clustering using features with ICOAP I performed the best based on all indexes.

Table 3. Clustering performance for datasets A (DB: Davies Bouldin Index, CH: Calinski Harabasz Index, S: Silhouette index)

Performance Index				DB	CH	S
SET A Spatiotemporal and kinematic	KOOS	K-means	k=2	1.1037	38.8612	0.3375
			k=3	0.9332	37.6967	0.3613
			k=4	**0.7748**	**47.3938**	**0.4118**
		K-medoids	k=2	1.1023	38.8363	0.3350
			k=3	0.9490	37.0514	0.3650
			k=4	0.7947	46.4054	0.4074
	ICOAP C	K-means	k=2	0.9866	27.3459	0.3563
			k=3	1.0414	36.1244	0.3848
			k=4	**0.8832**	**37.5386**	**0.4109**
		K-medoids	k=2	1.4767	21.2518	0.2721
			k=3	0.9172	29.0343	0.3526
			k=4	0.9486	35.5815	0.3864
	ICOAP I	K-means	k=2	1.1697	33.7243	0.3767
			k=3	0.8754	36.8219	0.4283
			k=4	**0.7955**	**42.2393**	**0.4842**
		K-medoids	k=2	1.2332	28.5044	0.3394
			k=3	0.9472	28.1132	0.4040
			k=4	**0.7955**	**42.2393**	**0.4842**

Table 4. Clustering performance for sets B (DB: Davies Bouldin Index, CH: Calinski Harabasz Index, S: Silhouette index)

Performance Index				DB	CH	S
SET B Spatiotemporal	KOOS	K-means	k=2	0.8268	57.3372	0.4805
			k=3	0.6399	82.4894	0.5280
			k=4	**0.5736**	**89.6149**	**0.5569**
		K-medoids	k=2	0.8268	57.3372	0.4805
			k=3	0.6354	79.7919	0.5298
			k=4	**0.5736**	**89.6149**	**0.5569**
	ICOAP C	K-means	k=2	0.7964	76.5135	**0.4888**
			k=3	0.8101	74.3832	0.4468
			k=4	**0.7646**	**77.6586**	0.4767
		K-medoids	k=2	0.8449	69.3085	0.4403
			k=3	0.8092	71.0572	0.4418
			k=4	0.8294	67.7985	0.4365
	ICOAP I	K-means	k=2	0.5473	147.3120	0.6097
			k=3	0.4657	193.2303	0.6224
			k=4	**0.3845**	**292.8713**	**0.6697**
		K-medoids	k=2	0.5404	136.9348	0.5885
			k=3	0.4657	193.230	0.6224
			k=4	**0.3845**	**292.8713**	**0.6697**

The results highlighted the importance of feature selection step before performing clustering. There are various types of feature selection algorithms which can be divided into three techniques, including filter-based, wrapper-based, and embedded techniques [40]. Filter-based technique is performed by univariate or multivariate statistical tests, which one of the methods is correlation analysis [41]. The advantages of using filter method is the low computational cost, easier implementation, and more generalizable compare to other techniques [42]. In this study, the gait features to be included were selected using correlation analyses with pain scores because pain is the most important characteristic of knee OA, and the aim of this study is to characterize knee OA based on the symptom. However, improvement of this study can be done by considering the multicollinearity issue between the gait features selected in the datasets. As the features were characteristics of the knee OA gait, correlation among the features themselves may had occurred and caused redundancy which affecting the performance of the clustering.

Table 5. Clustering performance for sets C (DB: Davies Bouldin Index, CH: Calinski Harabasz Index, S: Silhouette index)

Performance Index				DB	CH	S
SET C Kinematic	KOOS	K-means	k=2	1.0920	39.4204	0.3392
			k=3	0.8719	38.9762	0.3913
			k=4	**0.7717**	**48.1573**	0.4095
		K-medoids	k=2	1.0884	39.3915	0.3378
			k=3	0.9352	37.5181	0.3655
			k=4	0.7853	47.6531	**0.4099**
	ICOAP C	K-means	k=2	0.9768	27.8465	0.3616
			k=3	1.0322	36.7121	0.3886
			k=4	**0.8774**	**37.9194**	**0.4116**
		K-medoids	k=2	1.5548	17.2914	0.2656
			k=3	0.9380	29.8135	0.3499
			k=4	0.9369	36.2115	0.3901
	ICOAP I	K-means	k=2	1.1697	33.7244	0.3767
			k=3	0.8754	36.8220	0.4283
			k=4	**0.7955**	**42.2396**	**0.4842**
		K-medoids	k=2	1.2332	28.5045	0.3394
			k=3	0.9472	28.1133	0.4040
			k=4	**0.7955**	**42.2396**	**0.4842**

Other than that, although this study did not highlight the strength of the correlations as the main objective is to select the subset of the features with level of pain as the target groups, this study found that the three types of pain scores that represent different outlook on pain associate with different subsets of gait features. KOOS pain score that assessed pain level with the items focusing on intensity and frequency specific to several movement in daily life activity had correlation with all spatiotemporal features and with kinematic features that represents all three lower limbs joints: hip, ankle, and knee. For ICOAP constant pain score, the correlated spatiotemporal features were also all features measured, however, the kinematic features were slightly different with KOOS as no ankle kinematic features included and replaced with an additional knee kinematic feature. On the other hand, the ICOAP intermittent score had the least number of gait features correlated. This indicate that the choice of pain assessment to associate with gait features is crucial and future study should investigate the potential of employing various other outcome scores that available to assessed knee pain and other symptoms of knee OA.

4. Conclusion

Clustering analysis is unsupervised learning, usually used for unlabeled datasets to explore new information on dataset’s patterns. For this study, using pain scores with different subsets of gait features recorded from knee OA patients, clustering analyses were done and assessed using three types of the clustering performance index. Most of the datasets optimal number of clusters was 4 using k-means algorithm thus for future studies, this clustering model can be used for labelling class of knee OA severity and utilized for training data for classification. In conclusion, it is important to test different datasets with different algorithms and number of cluster k to choose the best clustering solutions generated. Future studies are suggested to include other pain scores available for knee OA outcome score and include other biomechanical features such as kinetic and muscle activity features. Moreover, other types of feature selection methods should be tested to improve the clustering performance.

Acknowledgment

This research was supported by Research University Grant (GUP Tier 2) (.18J11). We thank Ms. Norhidayah Abdullah from Hospital Sultanah Aminah, Johor Bahru, and Ms Yvonne Khor for assistance in subject recruitment. We also thank all members of Motion Analysis Laboratory, Universiti Teknologi Malaysia for their endless support.

References

[1] D.J. Hunter and S. Bierma-Zeinstra, Osteoarthritis, The Lancet, 2019, 393(10182):1745–1759. https://doi.org/10.1016/S0140-6736(19)30417-9

[2] A. Cui, H. Li, D. Wang, J. Zhong, Y. Chen and H. Lu. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies, eClinicalMedicine, 2020, 100587:29–30. https://doi.org/10.1016/j.eclinm.2020.100587

[3] M. Kapoor and N.N. Mahomed, Osteoarthritis: Pathogenesis, diagnosis, available treatments, drug safety, regenerative and precision medicine, Springer International Publishing, 2015. https://link.springer.com/book/10.1007/978-3-319-19560-5

[4] R. Altman, E. Asch, D. Bloch, G. Bole, D. Borenstein, K. Brandt, W. Christy, T.D. Cooke, R. Greenwald, M. Hochberg and D. Howell. Development of criteria for the classification and reporting of osteoarthritis: Classification of osteoarthritis of the knee, Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association, Arthritis & Rheumatism, 1986, 29(8):1039–1049. https://doi.org/10.1002/art.1780290816

[5] W. Zhang, M. Doherty, G. Peat, M.A. Bierma-Zeinstra, N.K. Arden, B. Bresnihan, G. Herrero-Beaumont, S . Kirschner, B.F. Leeb, L.S. Lohmander and B. Mazières. EULAR evidence-based recommendations for the diagnosis of knee OA, Annals of the Rheumatic Diseases, 2010, 69(3):483–489. http://dx.doi.org/10.1136/ard.2009.113100

[6] A.E. Nelson. Osteoarthritis year in review 2017: Clinical, Osteoarthritis and Cartilage, 2018, 26(3):319–325. https://doi.org/10.1016/j.joca.2017.11.014

[7] J. Bedson and P.R. Croft. The discordance between clinical and radiographic knee osteoarthritis: A systematic search and summary of the literature, BMC Musculoskeletal Disorders, 2008, 9(1):116. https://doi.org/10.1186/1471-2474-9-116

[8] J.H. Kellgren and J.S. Lawrence. Radiological assessment of osteoarthrosis, Annals of the Rheumatic Diseases, 1957, 16(4):494–502. https://doi.org/10.1136%2Fard.16.4.494

[9] S. Muraki, H.Oka, T. Akune, A. Mabuchi, Y. En-yo, M. Yoshida, A. Saika, T. Suzuki, H. Yoshida, H. Ishibashi, S. Yamamoto, K. Nakamura, H. Kawaguchiand and N. Yoshimura, Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: The ROAD study, Osteoarthritis and Cartilage, 2009, 17(9):1137–1143. https://doi.org/10.1016/j.joca.2009.04.005

[10] K.M. Son, J.I. Hong, D.-H. Kim, D.-G. Jang, M.D. Crema and H.A. Kim, Absence of pain in subjects with advanced radiographic knee osteoarthritis, BMC Musculoskeletal Disorders, 2020, 21(1):640. https://doi.org/10.1186/s12891-020-03647-x

[11] K. Sinusas, Osteoarthritis: Diagnosis and treatment. American Family Physician, 2012, 85(1):49–56.

[12] T.E. McAlindon, R.R. Bannuru, M.C. Sullivan 2, N.K. Arden, F. Berenbaum, S.M. Bierma-Zeinstra, G.A. Hawker, Y. Henrotin, D.J. Hunter, H. Kawaguchi, K. Kwoh, S. Lohmander, F. Rannou, E.M. Roos and M. Underwood, OARSI guidelines for the non-surgical management of knee osteoarthritis, Osteoarthritis and Cartilage, 2014, 22(3):363–388. https://doi.org/10.1016/j.joca.2014.01.003

[13] T.L. Vincent and F.E. Watt, Osteoarthritis, Medicine, 2018. 46(3):187–195. https://doi.org/10.1016/j.mpmed.2017.12.009

[14] F.W. Roemer, F. Eckstein, D. Hayashi and A. Guermazi, The role of imaging in osteoarthritis, Best Practice & Research Clinical Rheumatology, 2014, 28(1):31–60. https://doi.org/10.1016/j.berh.2014.02.002

[15] S.S. Yeap, S.R. Abu Amin, H. Baharuddin, K.C. Koh, J.K. Lee, V.K. Mun Lee, N.H. Mohamad Yahaya, C.C. Tai and M.P. Tan, A Malaysian Delphi consensus on managing knee osteoarthritis, BMC Musculoskeletal Disorders, 2021, 22(1):514. https://doi.org/10.1186/s12891-021-04381-8

[16] Astephen Wilson, J.L. and D. Kobsar, Osteoarthritis year in review 2020: Mechanics. Osteoarthritis and Cartilage, 2021, 29(2):161–169. https://doi.org/10.1016/j.joca.2020.12.009

[17] L.D. Duffell, S.J. Jordan, J.P. Cobb and A.H. McGregor. Gait adaptations with aging in healthy participants and people with knee-joint osteoarthritis, Gait & Posture, 2017, 57:246–251. https://doi.org/10.1016/j.gaitpost.2017.06.015

[18] J.A. Zeni and J.S. Higginson. Differences in gait parameters between healthy subjects and persons with moderate and severe knee osteoarthritis: A result of altered walking speed?, Clinical Biomechanics, 2009, 24(4):372–378.

[19] K.R Kaufman, C. Hughes, B.F Morrey, M. Morrey and K.N. An, Gait characteristics of patients with knee osteoarthritis, Journal of Biomechanics, 2001, 34(7):907–915. https://doi.org/10.1016/j.clinbiomech.2009.02.001

[20] J.L. Astephen, K.J. Deluzio, G.E. Caldwell and M.J. Dunbar. Biomechanical changes at the hip, knee, and ankle joints during gait are associated with knee osteoarthritis severity, Journal of Orthopaedic Research, 2008, 26(3):332–341. https://doi.org/10.1002/jor.20496

[21] M.B. Nebel, E.L. Sims, F.J. Keefe, V.B. Kraus, F. Guilak, D.S. Caldwell, J. Pells, R. Queen and D. Schmitt, The relationship of self-reported pain and functional impairment to gait mechanics in overweight and obese persons with knee osteoarthritis, Archives of Physical Medicine and Rehabilitation, 2009, 90(11):1874–1879. https://doi.org/10.1016/j.apmr.2009.07.010

[23] K.A. Marriott, T.B. Birmingham, K.M. Leitch, R. Pinto and J. Robert Giffin, Strong independent associations between gait biomechanics and pain in patients with knee osteoarthritis, Journal of Biomechanics, 2019, 94, 123–129. https://doi.org/10.1016/j.jbiomech.2019.07.015

[24] P. Levinger, D.T. Lai, R.K. Begg, K.E. Webster and J.A. Feller. The application of support vector machines for detecting recovery from knee replacement surgery using spatio-temporal gait parameters, Gait & Posture, 2009, 29(1):91–96. https://doi.org/10.1016/j.gaitpost.2008.07.004

[25] T.K. Yoo, S.K. Kim, S.B. Choi, D.Y. Kim and D.W. Kim. Interpretation of movement during stair ascent for predicting severity and prognosis of knee osteoarthritis in elderly women using support vector machine, 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE, 2013, 192–196. https://doi.org/10.1109/EMBC.2013.6609470

[26] N.Ş. Köktaş, N. Yalabik, G. Yavuzer and R.P. Duin. A multi-classifier for grading knee osteoarthritis using gait analysis, Pattern Recognition Letters, 2010, 31(9):898–904. https://doi.org/10.1016/j.patrec.2010.01.003

[27] S.P. Moustakidis, J.B. Theocharis and G. Giakas. A fuzzy decision tree-based SVM classifier for assessing osteoarthritis severity using ground reaction force measurements, Medical Engineering & Physics, 2010, 32(10):1145–1160. https://doi.org/10.1016/j.medengphy.2010.08.006

[28] J. McBride, S. Zhang, M. Wortley, M. Paquette, G. Klipple, E. Byrd, L. Baumgartner and X. Zhao. Neural network analysis of gait biomechanical data for classification of knee osteoarthritis, Proceedings of the 2011 Biomedical Sciences and Engineering Conference: Image Informatics and Analytics in Biomedicine, IEEE, 2011, 1–4. https://doi.org/10.1109/BSEC.2011.5872315

[29] M. Kotti, L.D. Duffell, A.A. Faisal and A.H. McGregor. Detecting knee osteoarthritis and its discriminating parameters using random forests, Medical Engineering & Physics, 2017, 43:19–29. https://doi.org/10.1016/j.medengphy.2017.02.004

[30] A. Elbaz, A. Mor, G. Segal, R. Debi, N. Shazar and A. Herman. Novel classification of knee osteoarthritis severity based on spatiotemporal gait analysis, Osteoarthritis and Cartilage, 2014. 22(3):457–463. https://doi.org/10.1016/j.joca.2013.12.015

[31] G. Leporace, F. Gonzalez, L. Metsavaht, M. Motta, F.P. Carpes, J. Chahla and M. Luzo. Are there different gait profiles in patients with advanced knee osteoarthritis? A machine learning approach, Clinical Biomechanics, 2021, 88:105447. https://doi.org/10.1016/j.clinbiomech.2021.105447

[32] E.T. Petersen, S. Rytter, D. Koppens, J. Dalsgaard, T.B. Hansen, N.E. Larsen, M.S. Andersen and M. Stilling. Patients with knee osteoarthritis can be divided into subgroups based on tibiofemoral joint kinematics of gait – an exploratory and dynamic radiostereometric study, Osteoarthritis and Cartilage, 2022, 30(2):249–259. https://doi.org/10.1016/j.joca.2021.10.011

[33] Lower body modeling with plug-in gait. Nexus 2.5 Vicon Documentation, 2019. https://docs.vicon.com/display/Nexus25/Lower+body+modeling+with+Plug-in+Gait. (Accessed April 15, 2019)

[34] J.E. Jackson, A user's guide to principal components, John Wiley & Sons, 2005, 587.

[35] J.L. Astephen and K.J. Deluzio. A multivariate gait data analysis technique: Application to knee osteoarthritis, Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 2004. 218(4):271–279. https://doi.org/10.1243%2F0954411041560983

[36] L. Kaufman and P.J. Rousseeuw. Finding groups in data: An introduction to cluster analysis, John Wiley & Sons, 2009.

[37] D.L. Davies and D.W. Bouldin. A cluster separation measure, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1979, PAMI-1(2):224–227. https://doi.org/10.1109/TPAMI.1979.4766909

[38] T. Caliński and J. Harabasz. A dendrite method for cluster analysis, Communications in Statistics, 1974, 3(1):1-27. https://doi.org/10.1080/03610927408827101

[39] P.J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics, 1987, 20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7

[40] U.M.Khaire and R. Dhanalakshmi, Stability of feature selection algorithm: A review, Journal of King Saud University - Computer and Information Sciences, 2022, 34(4):1060–1073. https://doi.org/10.1016/j.jksuci.2019.06.012

[41] P.J. Jones, M. Catt, M.J. Davies, C.L. Edwardson, E.M. Mirkes, K. Khunti, T. Yates and A.V. Rowlands, Feature selection for unsupervised machine learning of accelerometer data physical activity clusters – A systematic review, Gait & Posture, 2021, 90:120–128. https://doi.org/10.1016/j.gaitpost.2021.08.007

[42] A. Phinyomark, G. Petri, E. Ibáñez-Marcelo, S.T. Osis and R. Ferber, Analysis of big data in gait biomechanics: Current trends and future directions, Journal of Medical and Biological Engineering, 2018, 38(2):244–260. https://doi.org/10.1007/s40846-017-0297-2

[42] N. Kito, K. Shinkoda, T. Yamasaki, N. Kanemura, M. Anan, N. Okanishi, J. Ozawa and H. Moriyama, Contribution of knee adduction moment impulse to pain and disability in Japanese women with medial knee osteoarthritis, Clinical Biomechanics, 2010, 25(9):914–919. https://doi.org/10.1016/j.clinbiomech.2010.06.008