공학박사 학위논문

Machine learning based gait data analysis for objective evaluation of knee osteoarthritis

퇴행성 슬 관절염의 객관적 평가를 위한 기계학습 기반의 보행 데이터 분석 연구

2020 년 8 월

서울대학교 대학원
협동과정 바이오엔지니어링 전공
권 순 빈

Machine learning based gait data analysis for objective evaluation of knee osteoarthritis

지도교수 김 희 찬

이 논문을 공학박사 학위논문으로 제출함
2020 년 8 월

서울대학교 대학원
협동과정 바이오엔지니어링 전공
권 순 빈

권순빈의 공학박사 학위논문을 인준함
2020 년 8 월

위 원 장	이 정 찬	(인)
부위원장	김 희 찬	(인)
위 원	윤 형 진	(인)
위 원	노 두 현	(인)
위 원	구 윤 서	(인)

Ph. D. Dissertation

Machine learning based gait data analysis for objective evaluation of knee osteoarthritis

BY

Soon Bin Kwon

August 2020

INTERDISCIPLINARY PROGRAM IN BIOENGINEERING

THE GRADUATE SCHOOL

SEOUL NATIONAL UNIVERSITY

Machine learning based gait data analysis for objective evaluation of knee osteoarthritis

BY
Soon Bin Kwon

INTERDISCIPLINARY PROGRAM IN BIOENGINEERING
THE GRADUATE SCHOOL
SEOUL NATIONAL UNIVERSITY

THIS DISSERTATION IS APPROVED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

AUGUST 2020

Approved by Thesis Committee:

Professor ________________ Chairman
Jung Chan Lee, Ph. D.

Professor ________________ Vice chairman
Hee Chan Kim, Ph. D.

Professor ________________ Member
Hyung-Jin Yoon, M.D., Ph. D.

Professor ________________ Member
Du Hyun Ro, M.D.

Professor ________________ Member
Yunseo Ku, Ph. D

ABSTRACT

Machine learning based gait data analysis for objective evaluation of knee osteoarthritis

Soon Bin Kwon
Interdisciplinary Program in Bioengineering
The Graduate School
Seoul National University

Osteoarthritis (OA) is a disease that affects above 30% of the elderly population aged 60 years and older. Western Ontario and McMaster Osteoarthritis (WOMAC) and radiographic−based Kellgren–Lawrence (KL) grade methods are currently used to evaluate the severity of knee osteoarthritis (KOA). However, the WOMAC is a subjective method which cannot be performed to certain patients, and is not suitable for tracking changes in severity over time. KL grade requires highly trained experts and is a time consuming process. This dissertation hypothesized that objective and biomechanical gait data can supplement unmet needs of current gold standard. It was hypothesized that specific features from gait data would reflect the severity of KOA. Therefore, this study aims to identify key gait features associated with the severity of KOA and provide a new objective and explainable evaluation method for KOA based on gait analysis. Features were extracted from the gait signal and an automated severity evaluation model was designed based on machine learning technique for WOMAC severity evaluation model. To develop an automated severity evaluation algorithm for KL grade, features were extracted from the plain radiography image using deep learning network, and machine learning was applied to select features from the gait data. Both image and gait features were used to develop a machine learning algorithm for KL grade evaluation. The evaluation algorithm for WOMAC and KL grade showed a correlation of 0.741 and an accuracy of 75.2% with gold standard method, respectively. This dissertation proposed a new evaluation method for KOA and showed the clinical utility of the gait data application that was limited in clinical practice due to the complexity of the signal.

Keywords: Gait analysis, Machine learning, Deep learning, Knee osteoarthritis

Student number: 2016－26470

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	Two−dimensional
3D	Three−dimensional
ANOVA	Analysis of variance
AUC	Area under the curve
BLOKS	Boston Leeds Osteoarthritis Knee Score
CNN	Convolution neural net
IMU	Inertia measurement unit
IR	Infrared
IRB	Institutional Review Board
JSN	Joint space narrowing
KAM	Knee adduction moment
KL	Kellgren−Lawrence
KOA	Knee osteoarthritis
KOGS	Knee Osteoarthritis Grading System
MRI	Magnetic resonance image
NCA	Neighborhood component analysis
OA	Osteoarthritis
OCT	Optical coherence tomography
PROMs	Patient reported outcome measures
ROC	Receiver operating characteristic

SF−36 Short Form 36

SVM Support vector machine

TKA Total knee arthroplasty

US Ultrasound

WOMAC Western Ontario and McMaster Universities Osteoarthritis Index

WORMS Whole−Organ Magnetic Resonance Imagine Score

1. INTRODUCTION

1.1. KNEE OSTEOARTHRITIS

Osteoarthritis (OA) is one of the most prevalent musculoskeletal disease and is established as a public health problem [1-3]. It has been reported that 70% to 85% of the elderly population over 55 are affiliated with OA and 10% of the world's elderly population have severe case of OA [4-6]. The economic burden associated with osteoarthritis (OA) is high with 1-2% of the gross national product spent on OA-related healthcare [7, 8]. With the aging of the global population, the number of patients who suffer from knee osteoarthritis (KOA) is expected to increase [9].

OA is defined as a clinical condition of joints characterized by focal areas of degeneration of the articular cartilage with reactive formation of new bone at the articular margins [10]. This clinical condition of OA is irreversible. Depending on the underlying cause of this clinical condition, OA is classified into two types: primary and secondary. Primary OA is more common of the two. Even though the primary OA is related with age and genetic factor, there is no definite identifiable underlying cause. Secondary OA has identified causes of the disease, including injury, repeated surgery, and obesity [11]. Also, unlike the primary OA, secondary OA may occur at any age.

Even though OA can affect all joints, it is most common in hand, hip and knee. Among the three joints, the knee joint is a complex synovial joint and involves multiple structures to perform its function. The degeneration of articular cartilage and reformation of bone causes the typical symptoms of OA. The typical symptoms of KOA include pain, stiffness, which worsen in accordance with an increase in the disease progression [12, 13]. These symptoms occurring at the knee joint could significantly reduce the gait function and impair functional independence of individuals.

1.2. SEVERITY EVALUATION OF KNEE OSTEOARTHRITIS

1.2.1. Symptomatic Severity evaluation

The typical symptoms of knee osteoarthritis (KOA) include pain, stiffness and decreased joint motion. Even though there are objective methods to measure the decrease in joint motion, it is difficult to measure pain and stiffness objectively. The assessment of pain and stiffness are often performed with patient reported outcome measures (PROMs).

PROMs are patient self-complete questionnaires to assess patient's symptoms and clinical status [14]. PROMs can be conducted in a written format or verbally and can be translated into multiple language. PROMs often consist of questions to measure patient's symptom and how the symptoms affect the functionality in daily life activity. There are many disease-specific PROMs suggested to assess symptomatic and functional severity of various disease. Several PROMs are used in clinical practice for KOA such as, the Oxford Knee Score, the Short Form 36 (SF-36), and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).

Among the number of PROMs suggested for the assessment of KOA, the WOMAC is most widely used to determine the symptomatic severity of KOA [15]. There are total of 24 questions in WOMAC consist of three subscales: pain, stiffness, and physical function with 5, 2 and 17 questions, respectively. The WOMAC index has been widely used in clinical studies as well as clinical practice [16-18].

1.2.2. Structural Severity evaluation

The assessment of structural severity evaluation of KOA is performed based on medical image modalities; magnetic resonance image (MRI), ultrasound (US), optical coherence tomography (OCT) and radiography. Among these modalities MRI and radiography have established guidelines system to evaluate the KOA in clinical practice [19]. The evaluation guides for MRI includes, the Knee Osteoarthritis Grading System (KOGS) [20], the Whole-Organ Magnetic Resonance Imagine Score (WORMS) [21], and the Boston Leeds Osteoarthritis Knee Score (BLOKS) [22]. Even though MRI provides detailed information of the knee joint, radiography is the most accessible and remains as a gold standard for structural severity evaluation of KOA.

Among various grading system suggested for radiography modalities [23-25], the Kellgren-Lawrence (KL) grading system is current gold standard and most widely used in clinical practice. The KL grading system classifies KOA into five grades, ranging from 0-4, where Grade 0 indicates healthy subjects with no KOA symptoms, and Grade 4 indicates the most severe cases suitable for total knee arthroplasty (TKA). The KL grade is determined by observing the presence of joint space narrowing, osteophytes, bone deformity, and sclerosis from radiographic images. For the accurate evaluation of the KL grades, two experts are required to independently conduct radiographic evaluations without considering other data. If the evaluation of the two experts are contradictory, the results are discussed to reach a conclusion.

1.3. UNMET CLINICAL NEEDS

Even though gold standard for both symptomatic and structural severity evaluation are cost and time efficient method to assess symptoms and function of patients, there are limitation with the current gold standards. WOMAC has been validated in many previous studies to be efficient method in multiple language in various format [26-28], however, it is limited to the drawbacks of PROMs. WOMAC, just like other PROMs, is an unexplainable method. [29, 30].

Limitations also exist for KL grade system. Although the KL grading system is widely implemented in clinical applications for the severity evaluation of KOA, it is time consuming and requires highly trained experts, generally with fellowship training experiences in arthroplasty or radiography [31]. For the accurate evaluation of the KL grades, two experts are required to independently conduct radiographic evaluations without considering other data. If the evaluation of the two experts are contradictory, the results are discussed to reach a conclusion, which requires fair amount of time of the experts. [31, 32].

1.4. GAIT ANALYSIS AND KOA

Modern gait analysis is an effective technique for the analysis of the biomechanical information of lower joints, and it provides a temporal signal of each joint and additional gait information such as cadence, stride length, and step width. The main purpose of gait analysis is for quantified assessment of human locomotion.

The gait data is obtained with multiple infrared cameras and force plates placed on the ground. The gait cycle is composed of two main phase, stance and swing phase, which can further be categorized into initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid swing, and terminal swing. Each phases of gait cycle are shown in Figure 1.1. The obtainable biomechanical parameters from the gait analysis includes, hip flexion angle, hip adduction angle, hip rotation angle, hip extension moment, hip abduction moment, hip rotation moment, hip power, knee flexion angle, knee varus angle, knee extension moment, knee adduction moment, knee rotation moment, knee power, ankle flexion angle, ankle plantar-flexion moment, ankle varus moment, ankle power, pelvic tilt angle, pelvic obliquity angle, pelvic rotation angle, foot progression angle, foot progression moment, and tibia torsion angle. The typical signal of healthy subjects of the gait parameter is shown in Figure 1.2.

One of typical symptoms of KOA is decrease in gait function. The spatiotemporal gait parameters, such as cadence and speed, decreases for the KOA patients. Based on these observation, Thorp.et.al [33] have suggested both functional and structural distinction between KL grade. Other previous studies [13, 34, 35] also have analyzed and have shown the relationship between these gait parameters and the severity of KOA. It was shown that some traditional features, such as range of motion of the knee joint, had significant difference between OA and non-OA subjects. Even though these results suggests possible application of gait data, the complexity of the gait data remains as a barrier for application in clinical practice [36].

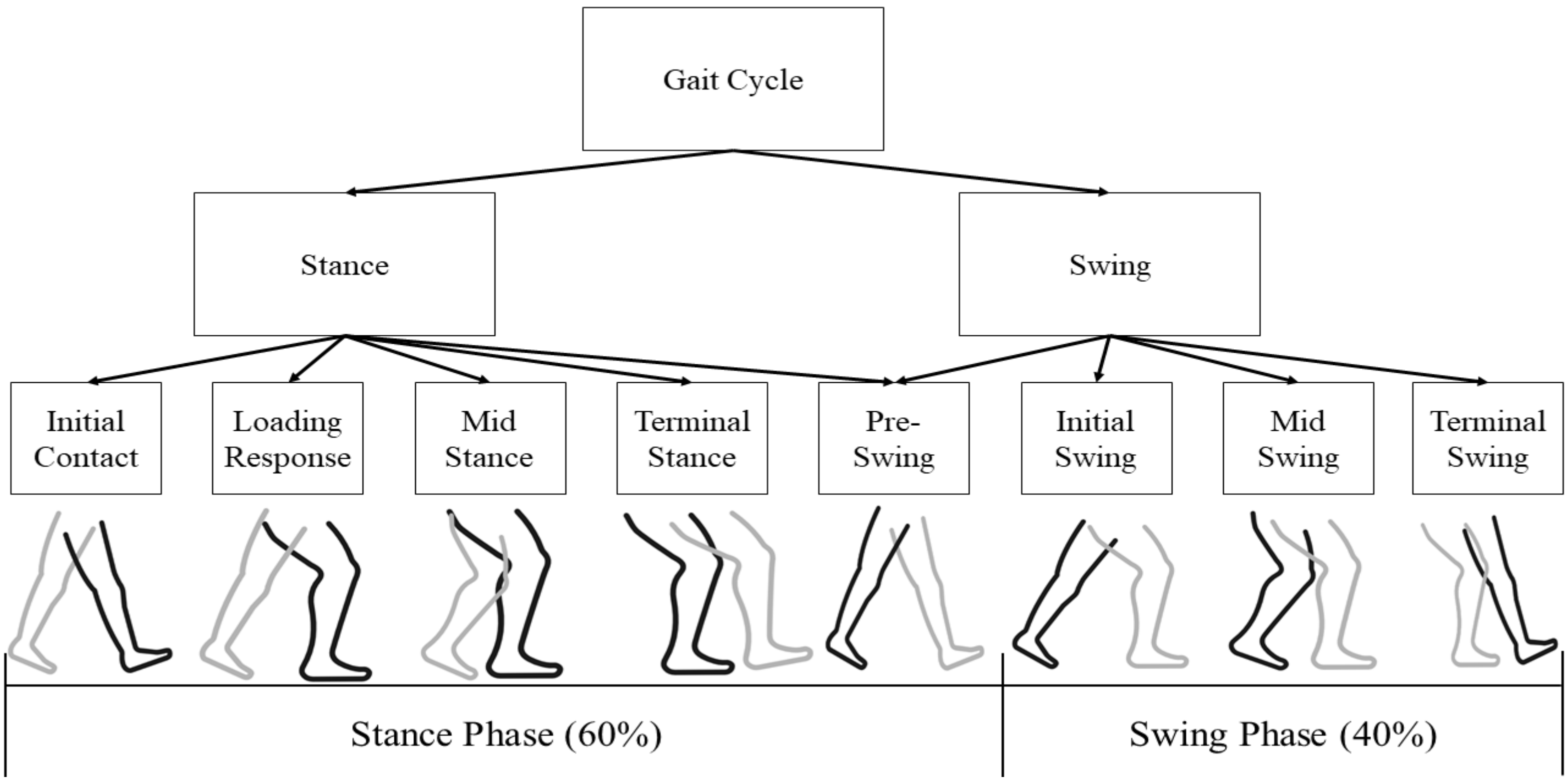

Figure 1.1 Gait cycle and gait motion for each phase during the cycle

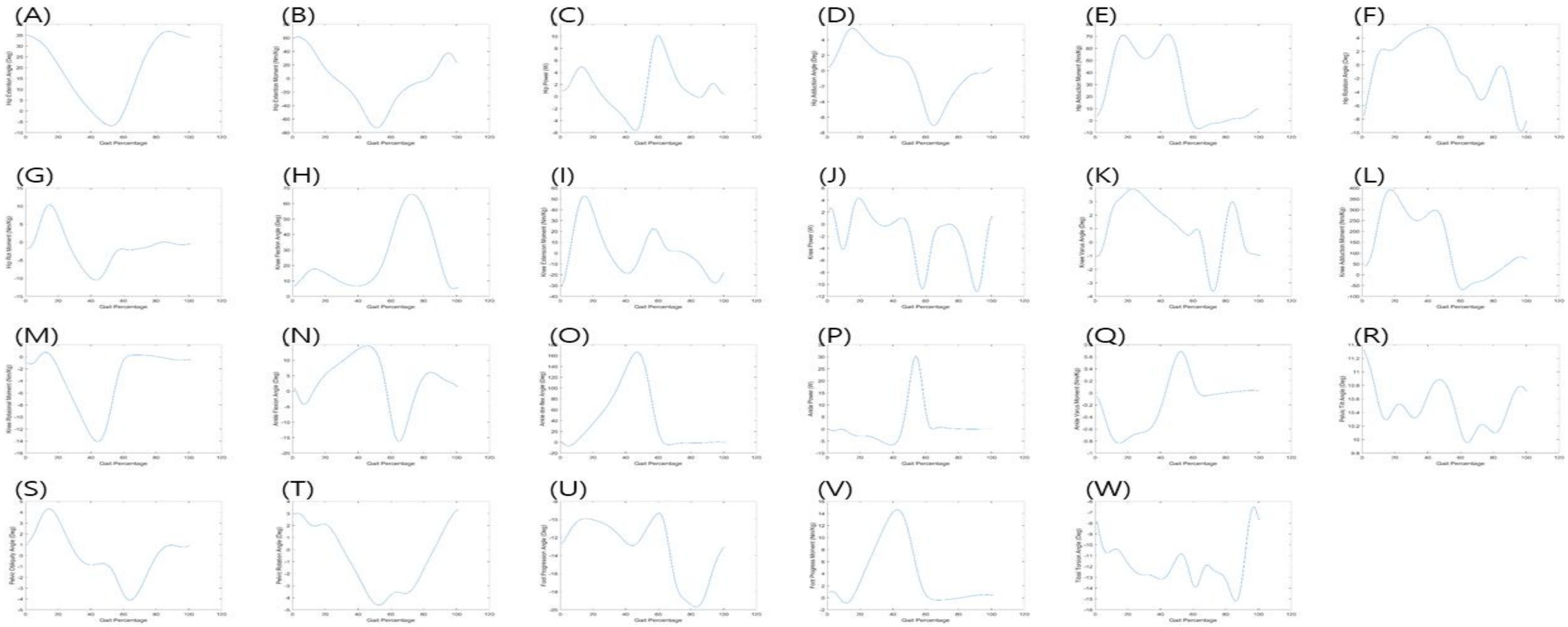

Figure 1.2 Gait parameters of control during the gait cycle. Starting from the (A) hip flexion angle (B) hip adduction angle (C) hip rotation angle (D) hip extension moment (E) hip abduction moment (F) hip rotation moment (G) hip power (H) knee flexion angle (I) knee varus angle (J) knee extension moment (K) knee adduction moment (L) knee rotation moment (M) knee power (N) ankle flexion angle (O) ankle plantar−flexion moment (P) ankle varus moment (Q) ankle power (R) pelvic tilt angle (S) pelvic obliquity angle (Y) pelvic rotation angle (U) foot progression angle (V) foot progression moment (W) tibia torsion angle

1.5. THESIS OBJECTIVES

KOA is irreversible and cannot be cured. Therefore, prevention and management of the disease is important. Patients who suffers from KOA gets TKA or drug treatment, therapy, reduces the symptoms with appropriate exercises, and management of the risk factors. For precise treatment of the disease, accurate and objective severity evaluation is important. However, there are limitation with current gold standards which led to the objectives of this dissertation. This dissertation hypothesized that gait data is closely related with both symptomatic and structural KOA and can supplement the unmet clinical needs of current gold standard severity evaluation methods. It was hypothesized that specific gait features would reflect the severity of the symptoms. Further, radiographic image and gait data were assumed to be complementary with each other. Thus, the objective of this dissertation is to suggest the application of gait data in clinical practice by developing objective and automated severity evaluation of KOA. The following aims were investigated to achieve the objective.

(1) Identify gait features related with symptomatic severity of KOA and develop an objective severity evaluation algorithm for symptomatic severity of KOA.

(2) Apply features from radiographic image and gait data to develop an automatic severity evaluation algorithm for the structural severity of KOA.

2. SYMPTOMATIC SEVERITY OF KNEE OSTEOARTHRITIS

2.1. INTRODUCTION

For many patients, relief from the symptoms of KOA is important to maintain functional independence. Thus, KOA therapy is aimed at reducing pain and improving gait function. The results from pain and gait disorder assessments are used to develop treatment plans, determine the effectiveness of treatment, and inform disease prognosis. However, there is no objective gold standard method exists for the assessment of symptomatic severity of KOA. The current gold standard for the assessment of symptomatic severity of KOA relies on a subjective questionnaire method, WOMAC. A Korean version of WOMAC is shown in Figure 2.1. As shown in the figure, there are total of 24 questions, and the score ranges from 0 to 96.

Previous studies have suggested gait analysis as an alternative tool for measuring patient disabilities since symptomatic dysfunction can be evaluated objectively using gait data [29, 30, 37]. This study anticipates that the analyses between WOMAC and gait analysis data can provide potentially objective measures of symptoms and provide insight regarding the relationship between symptomatic severity evaluation of disease and gait quality. Accordingly, the WOMAC estimation models based on the gait analysis features would provide an objective assessment of symptomatic severity of KOA.

This cross-sectional study analyzed the relationship between gait data and the WOMAC scores of KOA patients. The WOMAC indices of KOA patients without cognitive impairment, depression and who were willing to answer accurately, were included to avoid longitudinal bias and other possible inaccuracies. This study hypothesized that the WOMAC index and its three subscales would closely relate to KOA patients' gait function and that specific features would change with disease progression. Overall, the aim of this chapter is to identify the key features associated with the WOMAC index and its three subscales, and to apply machine learning algorithms to objectively evaluate the symptomatic severity of KOA.

< WOMAC >

1. 통증에 관한 질문입니다. 해당되는 항목에 v표시 해 주세요.

	통증이 없다 (0)	약간의 통증이 있다 (1)	중간정도의 통증이 있다 (2)	심한 통증이 있다 (3)	매우 극심한 통증이 있다 (4)
1) 평지걷기					
2) 계단 오르내리기					
3) 밤에 잘 때					
4) 앉거나 누워서 쉬고 있을 때					
5) 서 있을 때					

2. 다음은 무릎이 얼마나 뻑뻑한지 (굳어있는지)에 관한 질문입니다. 해당되는 항목에 v표시 해 주세요.

	뻑뻑하지 않다 (0)	약간 뻑뻑하다 (1)	중간정도의 뻑뻑함이 있다 (2)	심하게 뻑뻑하다 (3)	매우 극심하게 뻑뻑하다 (4)
6) 아침에 자고 일어났을 때					
7) 일과 중 앉거나 눕거나 쉬고 난 후 움직이려 할 때					

3. 기능에 관한 질문입니다. 다음의 활동들에 대하여 얼마나 어려움을 느끼십니까?

해당되는 항목에 v표시 해 주세요.

	전혀 힘들지 않다 (0)	약간 힘들다 (1)	중간정도로 힘들다 (2)	매우 힘들다 (3)	아주 극심하게 힘들다 (4)
8) 계단 내려오기					
9) 계단 올라가기					
10) 앉은 상태에서 일어나기					
11) 서있기					
12) 바닥을 향해 무릎을 구부리거나 물건을 줍기					
13) 평지 걷기					
14) 차에 오르내리기					
15) 장보기					
16) 양말이나 스타킹 신기					
17) 침대에서 일어나기					
18) 양말이나 스타킹 벗기					
19) 침대에 눕기					
20) 욕조에 들어가기,나오기					
21) 앉기					
22) 변기에 앉기,일어나기					
23) 힘든 집안일 하기(무거운 물건 옮기기, 마루닦기)					
24) 쉬운 집안일 하기(요리, 먼지제거)					

Figure 2.1 Korean version of WOMAC index

2.2. METHODS

2.2.1. Participants

This study was approved by Institutional Review Board (IRB) of Seoul National University Hospital (IRB no. 1810-004-974) and were performed in accordance with relevant guidelines and regulations. Written informed consent was obtained from all participants. This study was performed using the gait lab database. The database consists of gait reports of KOA patients with various degrees of knee pain and healthy volunteers without any knee pain from 2013 to 2017. This study excluded subjects based on the following criteria: (1) missing some data for both legs; (2) aged < 20 years; (3) spine disease, hip, or ankle arthritis on x-ray; (4) inflammatory or traumatic arthritis of the knee; (5) any prior bone surgery in the lower extremities; and

(6) cognitive impairment or depression. After the exclusion, the result of first WOMAC evaluation of each subject before any surgical procedure were selected. A total of 375 subjects were included in the study. Table 1 summarizes the participants' demographic characteristics and symptomatic severity

Table 1 Subject characteristics of symptomatic KOA patients

Feature	Mild (n=140)	Moderate (n=182)	Severe (n=53)	p-value
Age	62.6(9.1)	63.7(10.2)	63.3(10.2)	0.101
WOMAC	18.9(11.9)	48.5(6.8)	71.7(10.3)	<0.0001
Physical Function	13.8(8.9)	35.4(5.6)	52.8(7.8)	<0.0001
Pain	3.4(3.2)	9.1(2.3)	13.2(3.7)	<0.0001
Stiffness	1.7(1.7)	4(4)	5.7(5.7)	<0.0001

2.2.2. Gait Data Collection

All gait analysis data, including kinetic, kinematic and spatial-temporal data, were collected at the Human Motion Analysis Laboratory of Seoul National University Hospital following OrthoTrack 6.6 Reference Manual [38] with daily quality check to maintain the error within 1mm. All data collection process was performed by an operator with 20 years of experience. The subjects has a few minutes to warm up to acclimate to the setting and reflective markers were placed on the subjects based on the Helen Hayes arrangement. After placing the markers, an operator asked the subjects to walk along a 9 m track. Motion data were collected using twelve charge-coupled device cameras with a three-dimensional optical motion capture system (Motion Analysis Corp., Santa Rosa, CA, USA) at a sampling frequency of 120 Hz. Two floor-embedded force plates were used to obtain the kinetic data. An average of five or six trials of the 9 m walk of the kinetic and kinematic data for each joint were used in this study.

2.2.3. Statistical Analysis and WOMAC Estimation Model

All data analyses and classification were performed using MATLAB 2019b (MathWorks, Natick, MA). The gait analysis data were used to extract kinetic and kinematic data of hip, pelvic area, knee and ankle. These features included, but were not limited to, area under the curve, maximum value of swing phase, and minimum value of the curve. An additional 16 gait characteristics (i.e., velocity and cadence) were also selected as classification model features. Only the right leg was included to avoid statistical dependency from multiple observations of single individuals [39].

To statically analyze the relationship between the WOMAC score and gait features, the severity of WOMAC was classified into three classless: mild, moderate, and severe. Each WOMAC questions are answered into 5 different answers: none (0), mild (1), moderate (2), severe (3) and extreme (4). To divide the WOMAC score into three different severities, 1.5, the midpoint between mild to moderate, and 2.5, the midpoint between moderate to severe, were chosen as the decision point and were multiplied by the number of WOMAC questionnaires (24). The cut points adopted in this study is similar with other cut points used in previous studies [40, 41]. These studies applied 4 and 6 as cut points in scale of 0–10, which is 1.6 and 2.4 in scale of each WOMAC questionnaire (0–4). Accordingly, WOMAC scores below 36 was classified as mild, scores between 36 and 60 were classified as moderate, and the scores above 60 were classified as severe. The WOMAC subscales were divided into three classes using the same procedure.

A one-way analysis of variance (ANOVA) with a significance level of 0.0001 was performed. A student t-test was used to analyze class differences between each severity groups for features with significant difference as the result of ANOVA. For a multiple-comparison correction, a new alpha value of 0.00003 was used as significance level according to Bonferroni correction [42]. Features that were significant for all three comparisons between each classes were selected as key features.

Two different machine learning algorithm, random forest and support vector machine, were used to build regression model to estimate WOMAC score using the selected features as input. A random forest algorithm is an ensemble learning method constructed with multiple decision trees. Support vector machine is also widely used machine learning algorithm including the medical practice [43]. To resolve dataset imbalances, this study performed a down-sampling method.

The hold-out method was used for model validation. Seventy percent of the data were randomly selected to train the model and the other thirty percent of data were used for validation. The model was analyzed by observing correlation.

2.3. RESULTS

A total of 1083 features (of 23 gait parameters) were extracted from the gait analysis dataset and 44 features (12 hip, 1 pelvic, 17 knee, 9 ankle, 1 foot, and 4 spatiotemporal) were selected according to ANOVA and t−test results. The gait parameter features included hip rotation moment, hip flexion angle, hip adduction angle, hip power, pelvic obliquity angle, knee extension moment, knee flexion angle, knee power, knee varus angle, ankle plantarflexion moment, ankle power, foot progression angle, total speed, duration of single limb support phase (% of gait cycle), timing of initial double limb support (% of gait cycle), and timing of weight acceptance (% of gait cycle). Physical function was significantly related to all features, except for hip power. Pain differed significantly in relation to hip adduction angle, hip power, knee power, knee varus angle, ankle plantarflexion moment, and ankle power. Stiffness was significantly different in relation to hip rotation moment, hip adduction angle, knee flexion angle, and knee varus angle.

Table 2 summarizes the key WOMAC features and statistical results. The representative mean values of parameters for each group were divided according to WOMAC score (Figure 2.2). The correlation between actual and estimated WOMAC score was 0.725 and the 0.741, respectively for support vector machine and random forest (Figure 2.3).

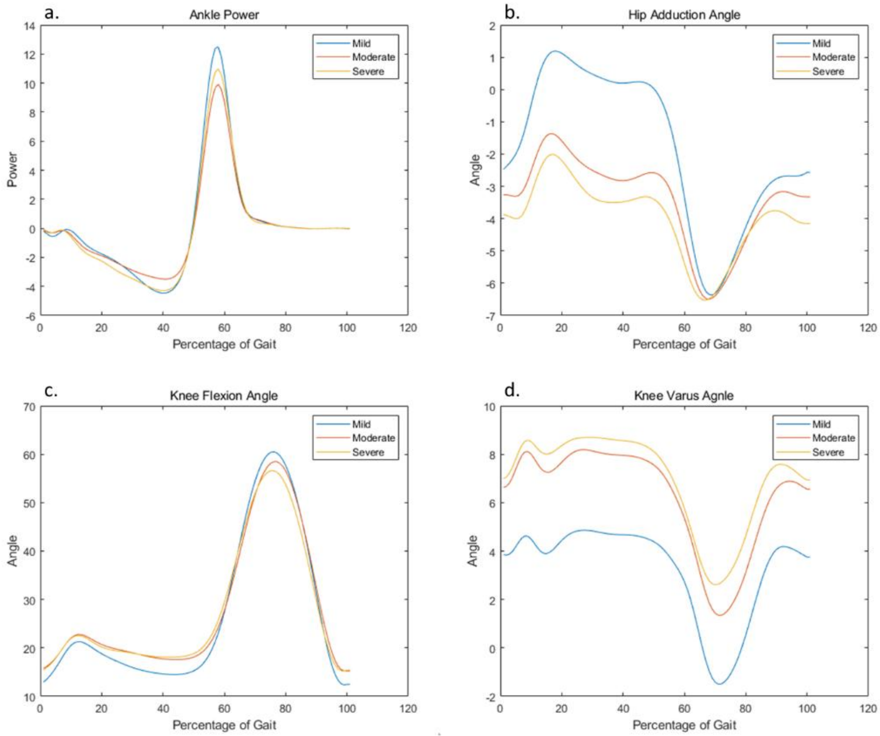

Figure 2.2 Mean values of representative gait parameters for each symptomatic severity of KOA where features were extracted from the a) ankle power, b) hip adduction angle, c) knee flexion angle, and d) knee varus angle.

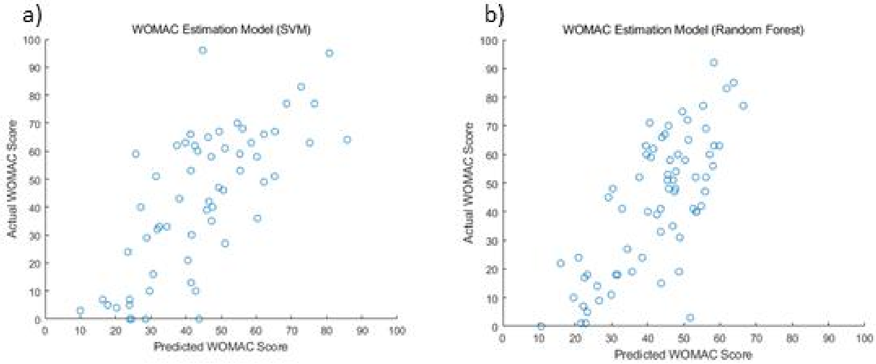

Figure 2.3 Regression result for WOMAC results using a) support vector machine b) the random forest algorithm and identified key features.

Table 2 List of gait features related with symptomatic severity of KOA

Joint	Parameter	Feature	PF	PN	SF
Hip	Rotation Moment	Standard Deviation	O		O
	Flexion Angle	Lower Bound of Autocorrelation	O		
		Bandwidth Frequency Bounds	O		
	Adduction Angle	Area Under the Curve during Stance Phase	O	O	O
		Standard Deviation of Absolute Value	O		
	Power	Minimum Value during Mid−Stance	O		
		Maximum Value during Terminal Stance	O		
		Area Under the Curve	O		
		Maximum − Minimum	O	O	
		Distance between Stance and Swing Phase using Dynamic Time Wrapping	O		
		Maximum Value during Mid−Swing	O		
		Minimum Value during Terminal Swing			
Pelvic	Obliquity Angle	Minimum value during Terminal Stance to Pre−Swing	O		

	Extension	Kurtosis	O		
	Moment	Peak2RMS	O		
		Variance	O		O
	Flexion Angle	Standard Deviation	O		O
		Maximum − Minimum	O		O
		Area Under the Curve of Power Spectral Density	O		O
	Power	Maximum Value during Terminal Swing	O		
		Maximum Value during Mid−Stance	O	O	O
Knee		Maximum Value during Terminal Stance	O		O
		Area Under the Curve of Stance Phase	O	O	O
		Area Under the Curve	O	O	O
		Root Mean Square (RMS)	O		O
	Varus Angle	Peak2RMS	O	O	O
		Mid−reference level	O	O	O
		Area Under the Curve of Power Spectrum	O		O
		Maximum Value during Terminal Swing	O	O	O
		Minimum Value during Loading Response	O	O	O
Ankle	Plantarflexion	Minimum Value during Loading Response	O	O	
	Moment	Maximum value during initial Swing	O		

		Maximum − Minimum	O
		Kurtosis	O
		Peak2RMS	O
	Power	Maximum − Minimum	O
		Lower Bound of Autocorrelation	O
		Occupied Bandwidth	O
		Bandwidth frequency bound	O
Foot	Progression Angle	Average of Absolute Value	O
Spatiotemporal		Total Speed	O
		Duration of single limb support phase	O
		Timing of initial double limb support	O
		Timing of weight acceptance	O

PF=Physical Function; PN=Pain; SF=Stiffness.

Table 3 Example of gait features showing difference among patients with same WOMAC score

Feature	Subject 1	Subject 2	Subject 3
Total WOMAC Score	52	52	52
WOMAC-Physical Function	35	35	34
WOMAC- Pain	11	11	14
WOMAC - Stiffness	6	6	4
Hip Power Maximum - Minimum	10.55	11.51	9.84
Knee Varus Angle Area Under the Curve of Stance Phase	8.11	12.64	1.26
Knee Varus Angle Area Under the Curve	8.36	12.59	0.99

Knee Varus Angle Peak2RMS	458	681.23	21.19
Knee Varus Angle Mid-reference level	7.11	9.38	1.51
Knee Varus Angle Maximum value during Terminal Swing	1.56	1.35	2.68
Knee Varus Angle Maximum during Loading Response	1.47	3.24	0.02
Ankle Plantarflexion Moment Minimum Value during Loading Response	11.12	10.17	4.05
Total WOMAC Score	52	52	52

2.4. DISCUSSION

While previous studies [30, 44, 45] have reported the relationship between spatiotemporal gait features, such as speed and stride length, and WOMAC indices of KOA or hip OA patients, this is the first study to analyze the relationship between kinetic and kinematic gait parameters and the WOMAC indices. Biomechanical intervention is recognized an alternative method to control pain and improve physical function [46]. Gait analysis provides meaningful KOA biomechanical information, but its complexity has limited its clinical applicability [36, 47]. Here, this study statistically analyzed key gait cycle features and identified critical KOA biomechanical information. In addition, this study built machine learning estimation models for the WOMAC index based on the identified features. While PROM methods are cheap, easy and quick, they are not applicable to patients who are unable or unwilling to perform the task. Despite the ability of gait analysis to provide valuable information about KOA biomechanical properties, a standardized method is not available for clinical use. The estimation model provides objective and reliable symptomatic results and has utility as a consistent method for evaluating gait analysis data. Finally, this study has identified key features based on both conventional methods and novel engineering methods. Conventional features, such as peak and minimum gait data values, are limited to load or motion at a single time point during the gait cycle and do not contain information over the gait cycle [48]. This study has developed methods that include information over the entire gait cycle, such as area under the curve, root mean square (RMS) and power spectrum. This study also conducted detailed feature analysis during gait cycle sub-phases: loading response, mid-stance, terminal-stance and pre-swing of stance phase, and initial swing mid-swing and terminal swing of swing phase.

This study identified well-known joint parameters that are specific to KOA patients and function in gait performance (listed in Table 2). Ankle dorsiflexion moment, for example, is an ankle joint movement involved in supination and pronation and three-dimensional ankle joint motions [49]. Previous studies have shown that knee varus angle changes are closely related to KOA [50, 51]. Lo and colleagues reported an association between knee varus angle and knee pain during weight bearing activities, most likely due to narrowing of the medial joint space, opening of the lateral space or increased lateral soft tissue pretension. This study found that hip, knee and ankle joint power, the product of torque and angular velocity, differed significantly according to WOMAC severity. Similarly, Segal et al. [52] reported joint power differences between symptomatic KAO patients and high-functioning controls. Ro et al. [53] have reported the difference between maximum and minimum value of both hip flexion angle and hip adduction angle were smaller in KOA patients compared with control group. Weidow et al. have reported that [54] the maximum value of hip rotation moment significantly differed between symptomatic and asymptomatic group. KOA patients was reported to have significantly lower knee flexion range of motion in swing and stance phase during gait cycle, which is in agreement with our findings [55]. McCarthy et al [56] claimed that knee extension moment is also an important gait characteristic to analyze the relationship between KOA and gait data. Bechard et al [57] reported that toe-out angle of foot progression angle was significantly smaller in patients with KOA and pelvic obliquity angle was reported to be correlated with symptoms of KOA.

In this study, physical function was influenced by the greatest number of features (43 from 14 parameters), indicating that WOMAC is a comprehensive score that incorporates the movement of many joints. This is reasonable given that KOA also effects the kinetic and kinematics of hip and ankle joints. Thus, to improve the physical function of patients, it is important to train not just the knee joint but also other KOA-affected joints [58]. The results of the study could provide guidelines for KOA exercise and rehabilitation (Table 2). Pain and stiffness were most related to knee-specific parameters. This pattern is demonstrated by tibiofemoral OA, which is a fairly common form of OA related to varus alignment. Tibiofemoral OA patients report higher pain levels than patellofemoral OA patients. Knee extension moment was not significantly related to pain. However, the WOMAC pain questionnaire only included one stair−related question, which may have influenced this result. In addition, the questionnaire also lacked questions related to knee adduction moment. Stiffness showed a significant relationship with knee flexion angle, a sagittal plane parameter. This is notable because the main movement of the knee, extension and flexion, is included in the sagittal plane. Also as shown in Table 3, gait features showed clear difference among patients with same WOMAC score but with different subscale scores.

A limitation of this study was that it was validated internally; to validate the model for overfitting it should be subjected to external validation. In addition, the features identified in this study were not applied to actual rehabilitation.

Future studies should apply the key features to patient rehabilitation and determine the therapeutic effects.

2.5. CONCLUSION

In conclusion, this study has identified features associated with the WOMAC its subscales and built estimation models for the WOMAC index. The features have been extracted using a feature engineering technique and statistically selected and validated. The estimation models were generated by traditional linear regression and random forest regression models. The estimation model and list of key features represents an objective and alternative option for KOA symptom severity evaluation.

3. STRUCTURAL SEVERITY OF KNEE OSTEOARTHRITIS

3.1. INTRODUCTION

For a brief review, the current gold standard for the structural severity evaluation of KOA is the Kellgren–Lawrence (KL) grading system [23]. An example of a knee radiographic image of each KL grade is shown in Figure 3.1. As shown in the figure, even though the difference in KL grade 0 and 4 is distinct, it is difficult to observe the difference between each consecutive KL grades. Thus, highly trained experts are required to make radiographic assessment. However, the accurate severity evaluation process of KL grade is time consuming process for the experts. An automated severity evaluation system can decrease the time–consumption, thus allowing for the clinicians to direct their attention toward clinical findings.

Accordingly, deep learning models were proposed in previous studies for automated KOA severity evaluation using radiographic imaging [31, 59, 60]. Deep learning is an effective technique for the analysis and classification of images, which is widely applied in various fields such as the medical field, and demonstrates excellent performances. Besides, the deep learning approach did not demonstrate a satisfactory performance when it was applied to the classification of KOA based on radiographic images. Although deep learning demonstrated a satisfactory performance for binary classification between OA and non-OA cases, with an area under the curve (AUC) of 0.92, the overall results of the multi-class severity evaluation of KOA were unsatisfactory, with an accuracy [31]; when compared with the accurate results of the deep learning approaches in other fields of application [61-64]. The low performance of deep learning can be attributed to the fact that a fair amount of radiography images is not perfectly adequate for KL grade classification. Even though it is important to obtain clear visualization of all compartments of knee joints for suitable radiography image for both tibiofemoral OA and patellofemoral OA classification, a considerable number of images do not satisfy all conditions. Previous studies [31, 60] have trained and validated the models with large dataset, each with 8892 and 5960 knees, the multi-class accuracy of the models were 57.4% and 66.7%, respectively.

Due to the degradation of gait functions in accordance with the progression of the disease, the relationship between gait analysis and the severity of KOA based on radiographic images has received significant research attention [13, 35, 48]. Kean *et al.* [65] reported that the knee adduction moment (KAM) impulse was positively correlated with the severity of disease.

Spatiotemporal gait parameters, including speed, cadence, and duration of the stance phase, differ significantly in KOA patients and asymptomatic groups [66]. However, most previous studies concentrated on specific joints, such as KAM, and features were separated into trends, unsuited for clinical application. To make gait data more applicable clinically, this study first sought to extract as many features from various joints as possible and then to identify key features by Kellgren–Lawrence (KL) grade.

This cross–sectional study analyzed the gait data of subjects ranging from no KOA to end–stage KOA. This study hypothesized that the patients' gait function would decrease gradually with the stage of KOA and specific features would change with the progression. Also, the radiographic images and gait data were hypothesized to contain critical and complementary information with respect to the knee joint and the severity of the disease; thus, the use of both data improves the classification accuracy of the model. Overall, the aim of this chapter was to find key features that can be used as mechanical biomarkers of KOA progression and features were extracted from radiographic images using a deep learning network, and then a classification model was developed for KOA based on image features and gait features identified as KOA-related features.

KL grade 0	KL grade 1	KL grade 2	KL grade 3	KL grade 4
• No symptoms of KOA	• Doubtful JSN • Possible osteophytic lipping	• Possible JSN • Definite osteophytes	• Definite JSN • Moderate osteophytes • Some sclerosis • Possible bone-end deformity	• Marked JSN • Large osteophytes • Severe sclerosis • Definite bone ends deformity

Figure 3.1 Example of representative knee radiographic image of each KL grade from KL 0 to 4 and explanation of how the severity evaluation was made for each grade.

3.2. METHODS

3.2.1. Participants

This study was approved by our Institutional Review Board (IRB no. 1810−004−974). Written informed consent was obtained from all participants This study was performed with the database of our gait lab. The database consists of the gait reports of the various degrees of knee OA patients and healthy volunteers from 2012 to 2017. The inclusion criteria of the database were healthy volunteers or knee OA patients who decided to participate the gait analysis and x ray analysis. The subjects' medical records were obtained, and all participants underwent a physical examination and standing, knee−extended position, full−limb radiography of the knee. This study excluded subjects based on the following criteria: (1) patients who lacked some data for both legs; (2) patients aged or < 20 years; (3) spine disease, hip, or ankle arthritis on x-ray; (4) inflammatory or traumatic arthritis of the knee ; (5) any prior bone surgery in the lower extremities ; and (6) All participants with equal KL grades for both knees. Consequently, 227 unilateral subjects with KOA participated in this study for gait feature identification. The degree of KOA was determined using the KL grading system. Table 4 summarizes the participants' demographic characteristics and walking speed. For the classification model, participants with criteria (6) were not excluded and each limb was counted as separate data, and a total of 728 limbs were included in this study.

Table 4 Subject characteristics for structural KOA patients

	KL 0	KL 1	KL 2	KL 3	KL 4	p-value
age	28.7(5.8)	64.6(4.1)	67.3(7.4)	69.2(7)	69.3(6.6)	<0.01
height	167.5(8.7)	156.3(7.9)	156.7(8)	154.3(7)	153.8(7.5)	<0.01
weight	63.2(12.3)	59.4(9.1)	62.8(10.8)	63.4(9.6)	63.9(10.2)	0.014
BMI	22.4(2.9)	24.3(3)	25.5(3.3)	26.6(3.2)	27(3.5)	<0.01
TS	124.1(7.6)	106.9(15.3)	88.4(23.1)	84(21.2)	78.9(20.1)	<0.01
Gender Ratio	95:92	46:48	85:22	97:22	97:152	N/A

3.2.2. Gait Data Collection

All gait analysis data, including kinetic, kinematic and spatial−temporal, were collected at the Human Motion Analysis Laboratory of Seoul National University Hospital. The subjects were asked to walk for a few minutes to get used to the setting. After warming up, an operator with 19 years of experience placed reflective markers on the subjects based on the Helen Hayes set. The subjects were asked to walk along a 9−m track. Motion data were collected using twelve charge−coupled device cameras with a three−dimensional optical motion capture system (Motion Analysis Corp., Santa Rosa, CA, USA) at a sampling frequency of 120 Hz. The kinetic data was obtained with two force plates which is embedded in the floor. The kinetic and kinematic data for each joint were averaged after five or six trials of the 9-m walk and then used as study data.

3.2.3. Radiographic Assessment

All the full limb radiographic images used in this study were digitally obtained using an image archiving and communication system (Maroview 5.4; INFINITT Healthcare, Seoul, Korea). The radiographic evaluations were conducted independently, in accordance with the KL grading system, by two experts with fellowship training experience in arthroplasty. The two experts did not consider the other information related to the subjects. When the evaluation results were contradictory, the grading was discussed to reach a conclusion. The inter-observer reliability of the radiographic assessments was satisfactory (intra-class correlation coefficient, 0.93).

3.2.4. Feature Extraction and Classification

All data analyses and classifications were performed using MATLAB 2019b (MathWorks, Natick, MA). In all, 149 features were extracted from the kinetic and kinematic data for the hip, knee, and ankle in the gait analysis data. The representative gait parameters are shown in Figure 3.2. These features are extracted by calculating the area under the curve, maximum value during swing phase, Kurtis, area of absolute value of the curve, and other characteristic of the kinetic and kinematic curves. An additional 16 gait characteristics were selected as features for the classification model, such as velocity and cadence. All features were extracted only from the right leg. This study analyzed only the right leg to remove statistical dependency from multiple observation from single individuals [39]. Neighborhood component feature selection [67] was performed to reduce the number of features. Neighborhood component analysis (NCA) feature selection performs regularization to obtain feature weights by minimizing the error for leave-one-out classification. Leave-one-out classification is conducted using one sample as a test set and the remaining data as a training set. This process is repeated until every sample has been used as a test set. Features with approximately zero feature weight were excluded. The remaining features were identified as key features and were used for the classification model. The student's *t*-test was performed to identify differences in the features among KL grades.

The Inception-ResNet-v2, which is a pre-trained convolutional neural network based on the ImageNet database, was used for two purpose: 1) To extract the features from radiographic images; 2) Classification model to compare our proposed SVM model. The Inception-ResNet-v2 combines the Inception architecture with a residual connection for the acceleration of the training and the improvement of the accuracy of the network [68]. It consists of 164 layers with image input dimensions of 299 × 299, and is considered as an effective and efficient network. The full limb radiographic images obtained from subjects were cropped around the knee and resized into dimensions of 299 × 299. The features were extracted after the final average pooling layer before the Softmax layer. A NCA was then conducted after the feature extraction using a deep learning network. For the selected gait and image features, the Student's

t−test was conducted to determine the significantly different features between two different KL grades ($p < 0.0001$). This analysis provides information on the features that significantly influence the classification of KL grades.

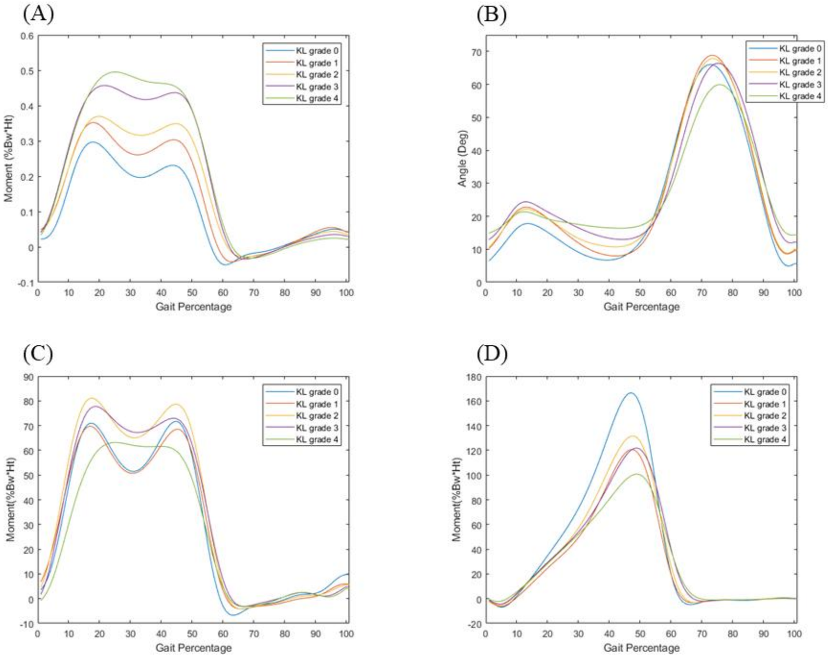

Figure 3.2 Mean values of representative gait parameters for each KL grade where features were extracted from the a) knee abduction moment, b) knee flexion angle, c) hip abduction moment, and d) ankle dorsiflexion moment. All three moments were normalized using weight × height.

A support vector machine (SVM) with the cubic kernel function was used for final classification. A hold-out validation was conducted to train and test the model. The data were split into train and test sets with a ratio of 7:3. The training set was used to train the SVM model, and the remainder of the unknown data was used for the test. A diagram of the study flowchart is shown in Figure. 3.3.

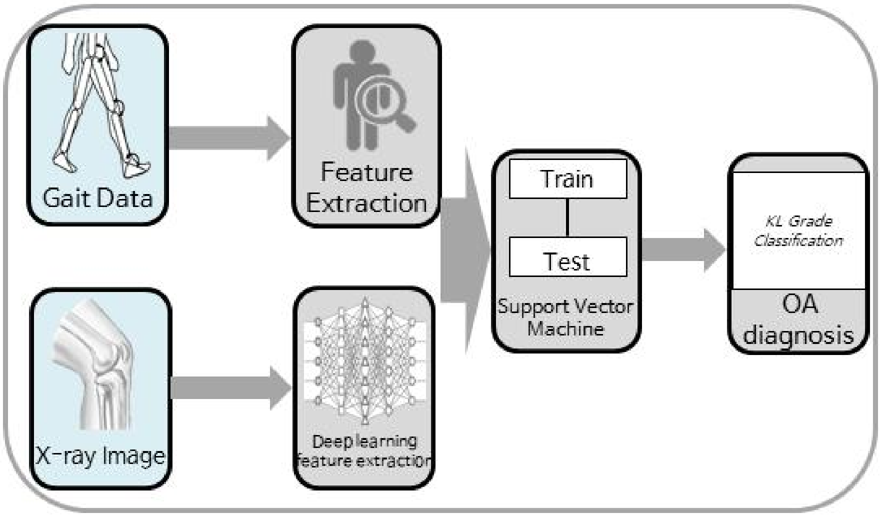

Figure 3.3 Flow diagram of deep learning approach with X-ray image input

For a comparison of the proposed SVM model, using both gait and image features as input, with deep learning methods using the same dataset as input, a deep learning network was trained and validated using the Inception-ResNet-v2. An Adam optimizer and a cross-entropy loss function were used with a mini-batch size of 32. A learning rate of $1e-4$ was set to stochastically optimize the network. Radiographic images cropped used feature extraction was used as input to train and validate the model. A hold-out validation method was used to test the model by splitting dataset into train and test sets with a ratio of 7:3. The total number of iteration was 900. The two models were examined by calculating the AUC of the receiver operating characteristic (ROC), sensitivity, precision, and F1-score. The sensitivity, precision, and F1-score were calculated using Equations 1-3, respectively:

$$Sensitivity = \frac{True\ \ Positive}{(True\ \ Positive + False\ \ Negative)} \tag{1}$$

$$Precision = \frac{True\ \ Negatives}{(True\ \ Negative + False\ \ Positive} \tag{2}$$

$$F1\ \ Score = 2 * \frac{Precision * Sensitivity}{(Precision + Sensitivity)} \tag{3}$$

The occlusion map [69], which indicates the relative significance of the 2D area for classification, was also observed to validate if the deep learning model uses the appropriate location for decision-making. The occlusion map is generated by covering a small portion of image with a mask, which moves across the image. The changes in the probability score for each mask location were measured and used to identify the relative significant area from an image.

3.3. RESULTS

3.3.1. Features Analysis

Among the 165 features extracted from the gait analysis data, 20 features (9 from the knee, 7 from the hip, 2 from the ankle joint, and 2 spatiotemporal parameters) remained after the NCA feature selection and were selected as the final features. The selected features are shown in Table 5 and the NCA result of gait features in shown in Figure 3.4. The gait parameters included as key features were: knee extension moment, knee abduction moment, knee rotational moment, knee flexion angle, hip abduction moment, hip extension moment, hip extension angle, ankle dorsiflexion moment, cadence and stride length. Among the 1,536 features extracted from a deep learning model, 65 features were selected.

Table 5 List of gait features related with structural severity of KOA

Joint	Parameter	Feature	KL 0	KL 1	KL 2	KL 3	KL 4
Knee	Abduction Moment	Area of the absolute value	131.95(67.29)^{1,2,3,4}	159.03(50.85)^{0,3,4}	176.11(80.12)^{0,3,4}	217.71(96.29)^{0,1,2}	227.56(105.93)^{0,1,2}
	Rotational Moment	Area during the stance phase	−28.89(14.16)^{3,4}	−27.52(12.37)^{3,4}	−32.59(17.38)^{3,4}	−43.24(24.3)^{0,1,2}	−44.25(25.24)^{0,1,2}
		Area of the absolute value	34.87(13.57)^{3,4}	31.73(11.01)^{3,4}	35.14(16.83)^{3,4}	45.14(23.36)^{0,1,2}	46.31(24.06)^{0,1,2}
	Flexion Angle	Maximum value during the late phase	17.87(5.46)^{1,2,3,4}	22.75(5.38)^{0}	22.65(5.81)^{0}	22.65(6.88)^{0}	22.7(9.98)^{0}
		Area of the absolute	746.6(195.64)^{1,2,3,4}	1001.29(260.6)^{0,2,3,4}	1181.12(370.59)^{0,1}	1215.46(434.1)^{0,1}	1319.04(643.92)^{0,1}

		value during the stance phase					
		RMS	747.82(193.53)^{1,2,3,4}	1004.04(254.4)^{0,2,3,4}	1182.74(366.98)^{0,1}	1217.82(429.08)^{0,1}	1347.81(586.75)^{0,1}
		Mean	2456.06(318.88)^{1,2,3,4}	2740.38(373.73)^{0}	2748.77(413.98)^{0}	2723.97(545.14)^{0}	2831.74(753.14)^{0}
		Kurtosis	418.56(68.6)^{2,3,4}	397.76(79.73)^{2,3,4}	321.93(102.7)^{0,1,3,4}	270.92(99.42)^{0,1,2}	235.18(143.73)^{0,1,2}
		Lower boundary of autocorrelation	62.1(4.95)^{2,3,4}	61.39(5.85)^{2,3,4}	55.87(8.2)^{0,1,3,4}	50.45(9.57)^{0,1,2,4}	45.53(13.51)^{0,1,2,3}
		Area under the curve	259.31(81.26)^{1,2,3}	300.41(69.67)^{0,4}	312.54(108.92)^{0,4}	302.46(116.57)^{0,4}	253.61(114.71)^{1,2,3}
Hip	Abduction Moment	Area of the absolute value during	260.3(82.77)^{1,2,3}	298.62(68.76)^{0,4}	310.33(107.3)^{0,4}	303.29(114.78)^{0,4}	255.1(113.41)^{1,2,3}

	the stance phase					
	Area of the absolute value	261.35(80.45)^{1,2}	301.16(70.07)^{0}	315.08(108.46)^{0,4}	305.65(116.3)^{0,4}	260.99(111.01)^{2,3}
	RMS	276.24(84.27)^{1,2,3}	313.48(72.8)^{0}	326.76(111.15)^{0,4}	316.89(118.3)^{0,4}	272.52(113.01)^{2,3}
Rotational Moment	Area under the curve	−4.06(11.55)^{1,2,3,4}	14.41(13.84)^{0}	16.5(19.41)^{0}	15.44(23.91)^{0}	17.01(27.12)^{0}
	Area of the absolute value during the stance phase	−6.61(12.42)^{1,2,3,4}	11.5(13.78)^{0}	14.48(19.36)^{0}	14.2(24.12)^{0}	16.06(26.71)^{0}
Extension Moment	Area of the absolute value during the stance phase	−42.54(66.77)^{4}	−52.78(81.97)^{4}	−20.92(95.27)	−25.38(97.62)	4.9(95.89)^{0,1}

	Area of the absolute value	200.38(64.15)^{1,2}	171.32(53.08)^{0}	167.09(59.19)^{0}	161.22(70.37)	163.09(69)
	RMS	264.04(83.91)^{1,2,3,4}	214.53(65.79)^{0,3,4}	207.91(71.17)^{0}	196.99(80.54)^{0,1}	203.7(82.13)^{0,1}
Extension Angle	Area under the curve	697.99(331.89)^{1,2,3,4}	852.54(376.38)^{0,4}	1057.45(545.43)^{0}	1055.03(447.11)^{0}	1241.66(582.13)^{0,1}
	Area of the absolute value during the stance phase	1803.47(562.39)^{1,2,3,4}	2016.5(586.61)^{0,4}	2173.77(740.04)^{0}	2145.6(641.98)^{0}	2375.54(804)^{0,1}
	Area of the absolute value	927.25(174.42)^{1,2,3,4}	1047.72(227.79)^{0}	1184.9(419.96)^{0}	1164.55(332.68)^{0}	1316.83(488.89)^{0}
	RMS	2050.86(371.88)^{1,2,3,4}	2221.47(425.62)^{0}	2305.05(604.01)^{0}	2260.31(521.51)^{0}	2454.9(701.22)^{0}
	STD	251(41.27)^{2,3,4}	255.34(62.95)^{2,3,4}	208.15(72.01)^{0,1,4}	205(66.52)^{0,1}	181.84(65.51)^{0,1,2}

Ankle	Dorsiflexion Moment	Upper boundary of autocorrelation	44.81(3.64)^{2,3,4}	45.89(5.59)^{2,3,4}	41.53(7.28)^{0,1}	41.66(6.4)^{0,1}	40.02(7.41)^{0,1}
		Area of the absolute value during the stance phase	347.33(106.16)^{1,2,3,4}	224.32(79.11)^{0}	255.43(98.49)^{0}	257.85(110.66)^{0}	257.13(114.13)^{0}
		RMS	368.43(109.38)^{1,2,3,4}	255.86(72.06)^{0}	277.03(98.73)^{0}	275.1(111.32)^{0}	271.62(111.4)^{0}
		STD	26.24(14.86)^{1,2,3,4}	12.34(6.83)^{0}	13.78(8.48)^{0}	13.3(8.67)^{0}	11.81(7.73)^{0}
Spatial-temporal		Total Speed	124.12(7.58)^{1,2,3,4}	106.9(15.27)^{0,2,3,4}	88.38(23.14)^{0,1,4}	83.99(21.19)^{0,1}	78.86(20.09)^{0,1,2}
		Total Stride Length	131.19(7.92)^{1,2,3,4}	113.5(12.24)^{0,2,3,4}	98.81(20.74)^{0,1,4}	96.35(18.42)^{0,1}	91.48(16.98)^{0,1,2}
		Step Width	65.47(4.48)^{1,2,3,4}	56.67(6.51)^{0,2,3,4}	49.54(10.76)^{0,1,}	48.2(9.35)^{0,1}	45.7(8.63)^{0,1,2}

4

[0] Significantly different with KL grade 0 ;[1] Significantly different with KL grade 1; [2] Significantly different with KL grade 2 ;[3] Significantly different with KL grade 3 ;[4] Significantly different with KL grade 4;

$^{*}p<0.001$

Figure 3.4 NCA result for gait features

Figure 3.5 represents the NCA result of radiographic image features. Table 6 presents the numbers of significantly different image and gait features based on comparisons between two KL grade groups. In 5 out of 10 binary comparisons, more than 30% of a total of 20 gait features were significantly different. However, less than 30% of image features were significantly different in the two cases of the 6 comparisons (KL1 vs. KL3 and KL2 vs. KL4). With respect to the image features, more than 30% of total 65 image features were significantly different in 6 out of 10 binary comparisons. However, no significantly different gait features were observed in the three cases of those 6 comparisons (KL1 vs. KL2, KL 2 vs. KL3 and KL3 vs. KL4). With respect to the radiographic image and gait features, more than 30% were significant, namely, KL0 vs. KL3, KL0 vs. KL4, and KL1 vs. KL4.

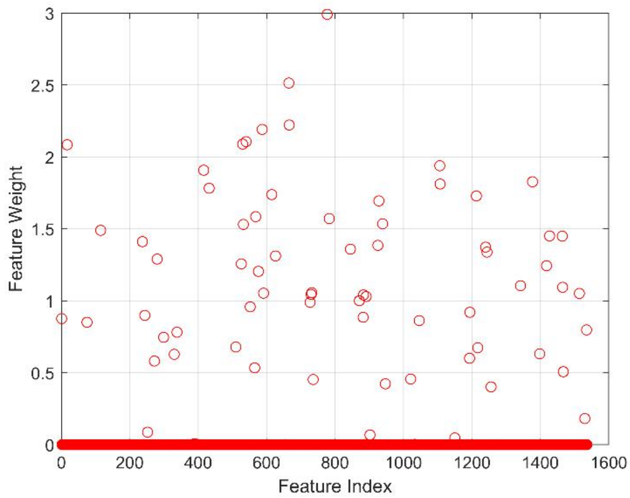

Figure 3.5 NCA result for radiographic image features

Table 6 List of gait features related with structural severity of KOA

Groups compared	# image features significantly different	# gait features significantly different
KL0 – KL1	11(17%)	4(20%)
KL0 – KL2	11(17%)	5(25%)
KL0 – KL3	20(31%)	11(55%)
KL0 – KL4	36(55%)	17(85%)
KL1 – KL2	46(70%)	2(10%)
KL1 – KL3	4(6%)	10(50%)
KL1 – KL4	28(43%)	14(70%)
KL2 – KL3	42(65%)	3(15%)
KL2 – KL4	10(15%)	7(35%)
KL3 – KL4	26(40%)	2(10%)

3.3.2. Deep Learning Approach Based on Radiographic Images

Figure. 3.6 presents a cropped radiographic image of the knee used as an input, in addition to an occlusion map of the model. Table 7 presents the confusion matrix of the validation result of the model and the AUC results, which were 0.93, 0.80, 0.85, 0.78, and 0.97 for KL Grades 0–4, respectively. The sensitivity, precision, and F1-score of the model were 0.55, 0.60, and 0.55, respectively. Figure. 3.7 (A) presents an ROC curve for the results obtained using the deep learning approach based on radiographic images.

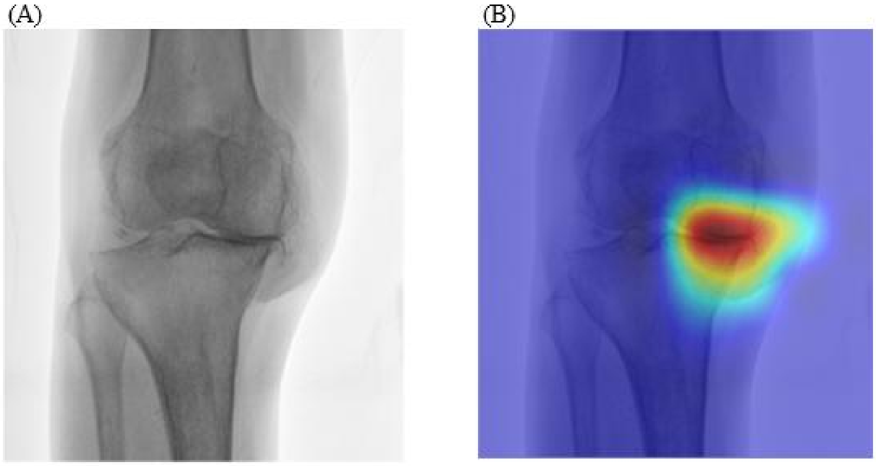

Figure 3.6 a) Cropped radiographic image of knee, used as an input to the deep learning network for radiographic feature extraction and KOA classification model to compare with proposed algorithm. b) An occlusion map indicating the relative spatial significance of features of a 2D image obtained as output of the deep learning classification model. The red area is the significant region, as indicated by the deep learning network, which is consistent with the ROI for KOA severity evaluation.

Table 7 Confusion matrix for deep learning approach with X-ray input

	KL 0	KL 1	KL 2	KL 3	KL 4
KL 0	60	6	0	1	0
KL 1	27	18	3	0	2
KL 2	3	5	5	1	3
KL 3	0	1	3	8	18
KL 4	0	0	0	3	48

3.3.3. Proposed Model Based on Gait Data and Radiographic Images

Table 8 presents the confusion matrix and AUC for the SVM model with total of 85 gait and radiographic image features. The AUC results for KL Grades 0–4 were 0.93, 0.82, 0.83, 0.88, and 0.97, respectively. The sensitivity, precision, and F1-score of the model were 0.70 0.76, and 0.71, respectively. Figure. 3.7 (B) presents the ROC of the classification result.

Table 8 Confusion matrix for proposed method using gait features and X-ray image as inputs

	KL 0	KL 1	KL 2	KL 3	KL 4
KL 0	60	6	0	1	0
KL 1	27	18	3	0	2
KL 2	3	5	5	1	3
KL 3	0	1	3	8	18
KL 4	0	0	0	3	48

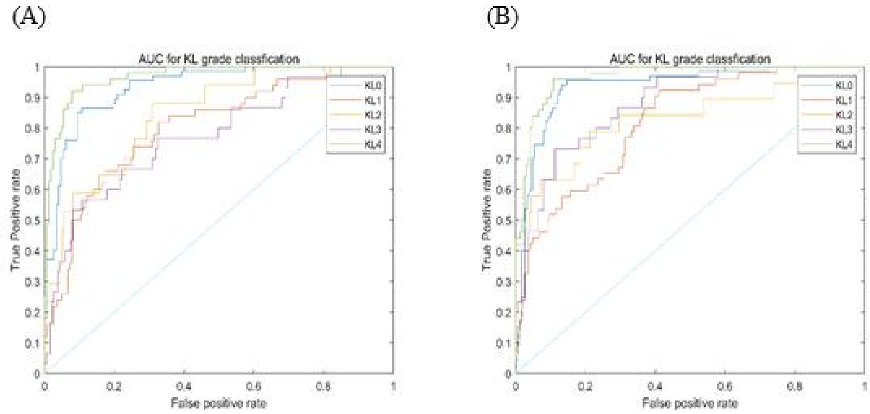

Figure 3.7 ROC curve for (A) deep learning approach using radiography image as input (B) Proposed method using both gait and radiography image as input

3.4. DISCUSSION

This study shows that gait function decreased gradually as the severity of KOA increased by identifying key gait features and demonstrated that the machine learning approach based on gait and radiographic image features can improve the classification performance of KOA at the KL grading scale when compared with the use of only radiographic images. The proposed model is based on gait analysis data and radiographic images of the knee, whereas the models proposed in previous studies utilized only one of the data types. Moreover, it was demonstrated that gait data can be used in clinical applications. Although gait data includes significant information on joints, its application is limited due to its complexity [21]. This paper proposes a method for the application of gait data to the severity evaluation of KOA, in combination with radiographic images. Moreover, the proposed method demonstrated an improved performance in the classification of KOA when compared with the methods proposed in previous studies. In particular, the proposed method demonstrated a relatively high accuracy, AUC, and F1-score, among other metrics.

The parameters listed in Table 5 are well-known joint parameters that have significantly different values at each stage of the disease. Kean *et al.* [65] reported that a feature extracted from the KAM during gait can distinguish individuals with different KL grades. Thorp *et al.* [70] reported that the same feature extracted from the KAM differed significantly among the control (KL grade 0 or 1), mild (KL grade 2), and moderate (KL grade 3) groups. Other parameters in Table 5 have also been reported [71-73]. Weidow *et al.* [74] reported that the maximum hip extension angle was smaller for OA patients compared with controls and that the peak hip flexion moment was also smaller in patients. They also showed that the knee flexion angle decreased in lateral OA patients. Astephen *et al.* [72] reported that the range of peak knee flexion angle differed significantly for severe OA patients and that the peak knee flexion angle during stance phase significantly differed progressively. They also reported that the minimum hip flexion moment during late stance differed significantly for all OA and that the hip flexion angle range significantly differed progressively. The peak and minimum ankle flexion moments were also reported to differ significantly for patients with severe OA [71].

The features listed in Table 5 have different abilities to discriminate the severity of disease. Some differed significantly for all severities, while others differed only in the controls, mild OA, or severe OA. The stride length differed significantly progressively, while maximum value during the late swing phase of hip extension moment differed significantly for KL grades 3 and 4. In comparison, area under the curve of hip abduction moment differed significantly between the most severe groups, KL grades 3 and 4, whereas all features extracted from the hip extension moment differed significantly between the control group.

Features like the peak and minimum values of gait data are reported to be limited to load or motion at an instance during the gait cycle and cannot contain information on the duration [65]. Features with duration information, such as the KAM impulse, are better discriminators than features without duration information, such as the peak KAM [65, 70]. Most of the features extracted in this study, such as the variance, RMS, and area under the curve, contain duration information. The variance and RMS are well-known, widely used parameters that contain overall information on the signal. Area under the curve is an integral of the signal and is also a representative parameter.

As shown in 3.6b, the occlusion map directs the region of interest (ROI), which indicates that the model was suitably trained. Tiulpin et al [31] presented an attention map that accurately determined the ROI and demonstrated a similar classification performance to the proposed deep learning model. However, the classification performance of the deep learning network based on radiographic images was not sufficient for the accurate severity evaluation of KOA using the KL grading system. Therefore, instead of limiting the applicability of the deep learning method as a classifier, deep learning was used to extract features from the radiographic images. Given that the deep learning pointed ROI as relatively important features on spatial location of 2D radiographic image, the features that were extracted using deep learning also contain information on the ROI. The combined gait and radiographic image features demonstrated a superior performance when compared with the deep learning method. The accuracy of the combined model was 75.2%, whereas that of the deep learning method was 64.7%.

As shown in Table 6, the radiographic image and gait data features were found to be complementary. For example, with respect to KL Grades 1–3, half of the 20 gait features were significantly different; whereas there were only four significantly different features among the 65 image features. Moreover, only two significant differences were observed between the gait data features at KL Grades 1 and 2. However, 46 features image features, which constituted more than half of the total number of features, were found to be significantly different. The results support the hypothesis, in that the gait data and radiographic image features are complementary for the distinction of the KL grades. It should be noted that the number of input features in the proposed model is significant less than that of the deep learning network based on radiographic images, which demonstrates a poorer classification performance. With reference to the literature, this was the first study wherein KOA was evaluated based on the KL grading system using both gait and radiographic image features.

3.5. CONCLUSION

In conclusion, the proposed model based on gait data and radiographic images was demonstrated to improve the accuracy of to evaluate the severity of KOA using the KL grading system. The automated classification of KOA using the proposed method can reduce the work of the clinician and improve the reliability of the KL grading system.

5. CONCLUSION

5.1. THESIS SUMMARY AND CONTRIBUTIONS

Modern gait analysis is a powerful technique that provides various biomechanical information of each joints in the lower body. Even though the degradation of gait function is clear along with the progression of KOA, the large volume and complexity of the data was a significant barrier for its application in actual clinical practice. With the development of artificial intelligence, it is now possible to extract as many features as possible from the data, and apply machine learning technique for feature selection to find significant features. Thus, in this this dissertation, various algorithm was applied to gait data for its application in clinical practice.

For application of gait data in clinical practice, automatic severity evaluation algorithm for both symptomatic and structural severity of KOA have developed. Key gait features associated with both severity of KOA have been identified. The feature selection method was performed based on machine learning algorithm and traditional statistical method. The estimation algorithms have successfully estimated the severity of KOA. For WOMAC estimation, the algorithm has shown the correlation of 0.741. The proposed model using radiography image and gait data as input has shown accuracy of 75.2% for discriminating the severity of KOA.

In conclusion, this dissertation has shown that gait data can be applied in clinical practice for KOA severity evaluation. The identified gait features are objective and repeatable, which can cover the limitations of current gold standards for assessment of KOA. The classification algorithm can contribute as an extra tool for KOA evaluation to obtain more detail information of the joint function.

5.2. FUTURE DIRECTIONS

As a next step of the study, it would be meaningful to design a wearable system optimized for knee joint kinematic measurement. The optimized wearable system should be light weighted, efficient with power consumption and convenient for both long and short term measurement. With the data collected from the wearable sensor, validation of the KOA severity evaluation algorithm would be suitable.

While this dissertation focused on the relationship between gait data and current severity of KOA, a study analyzing the relationship between gait data and the progression of disease would also be notable. An appropriately designed long term clinical study for mild KOA patient is needed. The difference between the patients with the disease progression and without the progression can reveal the biomechanical pathology of KOA. The result of this study might prevent the progression of the disease by modifying the gait habit and would provide useful information for rehabilitation for patients with knee joint disease.

BIBLIOGRAPHY

1. Altman, R., et al., *The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand.* 1990. **33**(11): p. 1601-1610.
2. Picavet, H. and J.J.A.o.t.r.d. Hazes, *Prevalence of self reported musculoskeletal diseases is high.* 2003. **62**(7): p. 644-650.
3. Felson, D.T., Y.J.A. Zhang, and R.O.J.o.t.A.C.o. Rheumatology, *An update on the epidemiology of knee and hip osteoarthritis with a view to prevention.* 1998. **41**(8): p. 1343-1355.
4. Woolf, A.D. and B.J.B.o.t.w.h.o. Pfleger, *Burden of major musculoskeletal conditions.* 2003. **81**: p. 646-656.
5. Kee, C.C.J.T.N.c.o.N.A., *Osteoarthritis: manageable scourge of aging.* 2000. **35**(1): p. 199-208.
6. Birchfield, P.C.J.G.N., *Osteoarthritis overview.* 2001. **22**(3): p. 124-131.
7. Vos, T., et al., *Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013.* 2015. **386**(9995): p. 743-800.
8. Murphy, L. and C.G.J.A.T.A.J.o.N. Helmick, *The impact of osteoarthritis in the United States: a population-health perspective.* 2012. **112**(3): p. S13-S19.
9. Kurtz, S., et al., *Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030.* 2007. **89**(4): p. 780-785.
10. Arya, R. and V.J.J. Jain, Indian Academy of Clinical

Medicine, *Osteoarthritis of the knee joint: An overview.* 2013. **14**(2): p. 154-62.

11. Cheng, Y., et al., *Physical activity and self-reported, physician-diagnosed osteoarthritis: is physical activity a risk factor?* 2000. **53**(3): p. 315-322.
12. Kaufman, K.R., et al., *Gait characteristics of patients with knee osteoarthritis.* 2001. **34**(7): p. 907-915.
13. Hall, M., et al., *The knee adduction moment and knee osteoarthritis symptoms: relationships according to radiographic disease severity.* 2017. **25**(1): p. 34-41.
14. Weldring, T. and S.M.J.H.s.i. Smith, *Article commentary: patient-reported outcomes (pros) and patient-reported outcome measures (PROMs).* 2013. **6**: p. HSI. S11093.
15. Bellamy, N., et al., *Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee.* 1988. **15**(12): p. 1833-1840.
16. Houpt, J., et al., *Effect of glucosamine hydrochloride in the treatment of pain of osteoarthritis of the knee.* 1999. **26**(11): p. 2423-2430.
17. Thomas, K., et al., *Home based exercise programme for knee pain and knee osteoarthritis: randomised controlled trial.* 2002. **325**(7367): p. 752.
18. Bachmeier, C.J., et al., *A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery.* 2001. **9**(2): p. 137-146.
19. Braun, H.J. and G.E.J.B. Gold, *Severity evaluation of osteoarthritis: imaging.* 2012. **51**(2): p. 278-288.
20. Oosthuizen, C.R., et al., *The knee osteoarthritis grading*

system for arthroplasty. 2019. **34**(3): p. 450-455.
21. Peterfy, C., et al., *Whole-organ magnetic resonance imaging score (WORMS) of the knee in osteoarthritis.* 2004. **12**(3): p. 177-190.
22. Hunter, D.J., et al., *The development and reliability of a new scoring system for knee osteoarthritis MRI: BLOKS (Boston Leeds Osteoarthritis Knee Score).* 2007.
23. Kellgren, J. and J.J.A.o.t.r.d. Lawrence, *Radiological assessment of osteo-arthrosis.* 1957. **16**(4): p. 494.
24. Altman, R.D., G.J.O. Gold, and cartilage, *Atlas of individual radiographic features in osteoarthritis, revised.* 2007. **15**: p. A1-A56.
25. Oosthuizen, C., et al., *The X-Ray Knee instability and Degenerative Score (X-KIDS) to determine the preference for a partial or a total knee arthroplasty (PKA/TKA).* 2015. **14**(3): p. 61-69.
26. Stucki, G., et al., *Evaluation of a German version of WOMAC (Western Ontario and McMaster Universities) arthrosis index.* 1996. **55**(1): p. 40-49.
27. Ko, T.-S., S.-Y. Kim, and J.-S.J.J.o.K.M.R. Lee, *Reliability and validity of the Korean Western Ontario and McMaster Universities (WOMAC) osteoarthritis index in patients with osteoarthritis of the knee.* 2009. **19**(2): p. 251-260.
28. Bellamy, N., et al., *Osteoarthritis Index delivered by mobile phone (m-WOMAC) is valid, reliable, and responsive.* 2011. **64**(2): p. 182-190.
29. Gandhi, R., et al., *Relationship between self-reported and performance-based tests in a hip and knee joint replacement population.* 2009. **28**(3): p. 253-257.

30. Lindemann, U., et al., *Gait analysis and WOMAC are complementary in assessing functional outcome in total hip replacement.* 2006. **20**(5): p. 413-420.
31. Tiulpin, A., et al., *Automatic knee osteoarthritis severity evaluation from plain radiographs: A deep learning-based approach.* 2018. **8**(1): p. 1-10.
32. Finnilä, M.A., et al., *Association between subchondral bone structure and osteoarthritis histopathological grade.* 2017. **35**(4): p. 785-792.
33. Thorp, L.E., et al., *Knee joint loading differs in individuals with mild compared with moderate medial knee osteoarthritis.* 2006. **54**(12): p. 3842-3849.
34. Hurwitz, D., et al., *The knee adduction moment during gait in subjects with knee osteoarthritis is more closely correlated with static alignment than radiographic disease severity, toe out angle and pain.* 2002. **20**(1): p. 101-107.
35. Chang, A.H., et al., *External knee adduction and flexion moments during gait and medial tibiofemoral disease progression in knee osteoarthritis.* 2015. **23**(7): p. 1099-1106.
36. Chau, T.J.G. and posture, *A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods.* 2001. **13**(1): p. 49-66.
37. Ornetti, P., et al., *Gait analysis as a quantifiable outcome measure in hip or knee osteoarthritis: a systematic review.* 2010. **77**(5): p. 421-425.
38. *OrthoTrak 6.6 Reference Manual.* 2009.
39. Bryant, D., et al., *How many patients? How many limbs? Analysis of patients or limbs in the orthopaedic literature: a systematic review.* 2006. **88**(1): p. 41-45.

40. Sibille, K.T., et al., *Accelerated aging in adults with knee osteoarthritis pain: consideration for frequency, intensity, time, and total pain sites.* 2017. **2**(3).
41. Kapstad, H., et al., *Cutpoints for mild, moderate and severe pain in patients with osteoarthritis of the hip or knee ready for joint replacement surgery.* 2008. **9**(1): p. 55.
42. Bland, J.M. and D.G.J.B. Altman, *Multiple significance tests: the Bonferroni method.* 1995. **310**(6973): p. 170.
43. Sheng, L., et al. *Diseases classification using support vector machine (SVM)*. in *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP'02.* 2002. IEEE.
44. Blin, O., et al., *Quantitative analysis of walking in patients with knee osteoarthritis: a method of assessing the effectiveness of non-steroidal anti-inflammatory treatment.* 1990. **49**(12): p. 990-993.
45. Solak, A.Ş., B. Kentel, and Y.J.T.J.o.a. Ateş, *Does bilateral total knee arthroplasty affect gait in women?: comparison of gait analyses before and after total knee arthroplasty compared with normal knees.* 2005. **20**(6): p. 745-750.
46. Parween, R., et al., *Methods for evaluating effects of unloader knee braces on joint health: a review.* 2019: p. 1-16.
47. Favre, J. and B.M.J.E.O.R. Jolles, *Gait analysis of patients with knee osteoarthritis highlights a pathological mechanical pathway and provides a basis for therapeutic interventions.* 2016. **1**(10): p. 368-374.
48. Kean, C.O., et al., *Comparison of peak knee adduction moment and knee adduction moment impulse in*

distinguishing between severities of knee osteoarthritis. 2012. **27**(5): p. 520-523.

49. Brockett, C.L., G.J.J.O. Chapman, and trauma, *Biomechanics of the ankle.* 2016. **30**(3): p. 232-238.
50. Gök, H., S. Ergin, and G.J.A.O.S. Yavuzer, *Kinetic and kinematic characteristics of gait in patients with medial knee arthrosis.* 2002. **73**(6): p. 647-652.
51. Lo, G.H., et al., *Associations of varus thrust and alignment with pain in knee osteoarthritis.* 2012. **64**(7): p. 2252-2259.
52. Segal, N.A., et al., *Association of dynamic joint power with functional limitations in older adults with symptomatic knee osteoarthritis.* 2009. **90**(11): p. 1821-1828.
53. Ro, D.H., et al., *Effects of knee osteoarthritis on hip and ankle gait mechanics.* 2019. **2019**.
54. Weidow, J., et al., *Hip and knee joint rotations differ between patients with medial and lateral knee osteoarthritis: gait analysis of 30 patients and 15 controls.* 2006. **24**(9): p. 1890-1899.
55. McCarthy, I., et al., *Analysis of knee flexion characteristics and how they alter with the onset of knee osteoarthritis: a case control study.* 2013. **14**(1): p. 169.
56. Creaby, M.J.O. and Cartilage, *It's not all about the knee adduction moment: the role of the knee flexion moment in medial knee joint loading.* 2015. **23**(7): p. 1038-1040.
57. Bechard, D.J., et al., *Toe-out, lateral trunk lean, and pelvic obliquity during prolonged walking in patients with medial compartment knee osteoarthritis and healthy controls.* 2012. **64**(4): p. 525-532.
58. Bosomworth, N.J.J.C.F.P., *Exercise and knee*

osteoarthritis: benefit or hazard? 2009. **55**(9): p. 871-878.

59. Antony, J., et al. *Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks*. in *International conference on machine learning and data mining in pattern recognition*. 2017. Springer.
60. Antony, J., et al. *Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks*. in *2016 23rd International Conference on Pattern Recognition (ICPR)*. 2016. IEEE.
61. Mansour, R.F.J.B.e.l., *Deep-learning-based automatic computer-aided severity evaluation system for diabetic retinopathy.* 2018. **8**(1): p. 41-57.
62. Choi, J.W., et al., *White blood cell differential count of maturation stages in bone marrow smear using dual-stage convolutional neural networks.* 2017. **12**(12).
63. Cho, S.I., et al., *Dermatologist-level classification of malignant lip diseases using a deep convolutional neural network.* 2019.
64. Dey, D., S. Chaudhuri, and S.J.B.e.l. Munshi, *Obstructive sleep apnoea detection using convolutional neural network based deep learning framework.* 2018. **8**(1): p. 95-100.
65. Kean, C.O., et al., *Comparison of peak knee adduction moment and knee adduction moment impulse in distinguishing between severities of knee osteoarthritis.* Clinical biomechanics, 2012. **27**(5): p. 520-523.
66. Baliunas, A., et al., *Increased knee joint loads during walking are present in subjects with knee osteoarthritis.* Osteoarthritis and cartilage, 2002. **10**(7): p. 573-579.

67. Yang, W., K. Wang, and W. Zuo, *Neighborhood Component Feature Selection for High-Dimensional Data.* Journal of Computers, 2012. **7**(1).
68. Szegedy, C., et al. *Inception-v4, inception-resnet and the impact of residual connections on learning*. in *Thirty-first AAAI conference on artificial intelligence*. 2017.
69. Zeiler, M.D. and R. Fergus. *Visualizing and understanding convolutional networks*. in *European conference on computer vision*. 2014. Springer.
70. Thorp, L.E., et al., *Knee joint loading differs in individuals with mild compared with moderate medial knee osteoarthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2006. **54**(12): p. 3842-3849.
71. Mezghani, N., et al., *Biomechanical analysis to characterize the impact of knee osteoarthritis on hip, knee, and ankle kinematics.* Journal of Biomedical Engineering and Informatics, 2017. **3**(2): p. 36.
72. Astephen, J.L., et al., *Biomechanical changes at the hip, knee, and ankle joints during gait are associated with knee osteoarthritis severity.* Journal of orthopaedic research, 2008. **26**(3): p. 332-341.
73. Deluzio, K. and J. Astephen, *Biomechanical features of gait waveform data associated with knee osteoarthritis: an application of principal component analysis.* Gait & posture, 2007. **25**(1): p. 86-93.
74. Weidow, J., et al., *Hip and knee joint rotations differ between patients with medial and lateral knee osteoarthritis: gait analysis of 30 patients and 15 controls.* Journal of Orthopaedic Research, 2006. **24**(9): p. 1890-

1899.

Abstract in Korean

국문 초록

퇴행성 관절염은 60세 이상의 노인 인구 약 30%에서 발병하는 질병이다. 현재 퇴행성 슬 관절염의 진단은 Western Ontario and McMaster Osteoarthritis (WOMAC) 방법과 방사선 촬영 기반의 Kellgren-Lawrence (KL) grade 방법이 사용되고 있다. 그러나 WOMAC 환자의 주관적인 판단을 토대로 중증도를 정량화하는 방법이어서 일부 환자들에게 적용이 불가능하고, 수술 후의 중증도를 반영하지 못한다는 단점이 있다. KL grade은 고도로 훈련된 전문가를 필요로 하며, 정확한 진단을 위하여서는 많은 시간을 필요로 한다. 반면 보행 신호는 환자의 보행에 따른 객관적인 생체 역학 신호를 제공하며, 보행이 가능한 모든 사람에게 적용이 가능하며, 주기적인 추적 관찰에 용의하다. 따라서 본 연구는 보행 신호를 이용하여 객관적이며, 결과에 대한 생체 역학적

이유를 알 수 있는 퇴행성 슬 관절염의 새로운 분석 방법을 제시함에 있다. 먼저 자동으로 WOMAC 방법을 진단하기 위해 보행신호에서 특징들을 추출하고 기계학습 기법을 이용하여 평가하는 모델을 개발하였다. 또한 KL grade 방법을 평가하기 위해 방사선 영상에서 딥러닝 알고리즘으로 추출한 특징들과 보행신호에서 추출한 특징들을 기계학습 기법을 이용하였다. 제안하는 퇴행성 슬 관절염의 평가 방법은 WOMAC 및 KL grade 방법과 각각 상관관계 0.741, 정확도 75.2%를 보였다. 본 연구는 퇴행성 슬 관절염의 새로운 평가 방법을 제시하였으며, 신호의 복잡성으로 인하여 임상에서 사용되지 못했던 보행 신호의 임상적 활용성을 보여주었다.

핵심어 : 기계학습, 보행 신호, 퇴행성 슬 관절염, 딥러닝