Pakistan Journal of Engineering and Technology
ISSN (Print): 2664-2042; ISSN (Online): 2664-2050
Received: 2023-07-17
Accepted: 2023-11-30
https://doi.org/10.51846/vol6iss3pp44-50

PAKJET
Pakistan Journal of Engineering & Technology

Article

Machine Learning-Based Gait Phase Detection for Semi-Active Prosthetic Knee

Muhammad Awais Khan *, Muhammad Usman Qadir, and Izhar Ul Haq

Department of Mechatronics Engineering, University of Engineering and Technology, Peshawar, Pakistan
* Correspondence: Muhammad Awais Khan (mawaiskhan0486@gmail.com)

Abstract: The human knee plays a vital role in performing day-to-day activities. For a healthy person, it is easy to perform locomotion activities, but for people with transfemoral amputation, it is a very difficult task. To overcome this issue, prosthetic knees are developed. These prosthetic knees provide the necessary function of the gait cycle. In order to mimic the gait cycle of the human knee, it is crucial to detect different phases in the gait cycle. Mechanical sensors such as force and angle sensors are used to collect kinematic data, and then with a heuristic rule base system, the gait phases are detected. The rule-based system performs well, but as the number of gait phases increases, it is difficult to identify them. This paper proposed machine learning-based gait phase detection. Decision trees, linear discriminant analysis, and support vector machines are applied to the kinematics data obtained from strain gauges and angle encoders. These algorithms are easy to implement on embedded hardware as they use low computational power. The Linear Discriminant analysis has the highest validation accuracy of 95.6% and test accuracy of 95.40%, while both the Support Vector Machine and Decision Trees algorithm have 95.2% validation accuracy. The test accuracy of the Support Vector Machine is 95.10%, and for the Decision Tree, it is 95.05%.

Keywords: Gait Phase Detection, Semi-active Prosthetic Knee, Machine Learning, Decision Trees, Linear Discriminant Analysis, Support Vector Machine

1. Introduction

The main function of a prosthetic knee is to provide comfort and restore mobility for lower-limb amputees. The crucial aspect of these prosthetic knees is to actively recognize phases of the gait cycle. This detection further determines how the knee will behave during walking and different activities that involve knee movement. At present, the most common and well-known technique to detect the gait phase is a rule-based system, in which different rules are devised based on subject knowledge. For this purpose, different threshold levels are set to identify gait phases. However, this approach is limited and can impact prosthetic performance, comfort, and stability. The main motivation behind this study is to provide an alternative and more robust solution to this problem. This research provides an intelligent approach to mitigate the issue of identifying complex patterns, such as machine learning. This study provides a comprehensive overview of the implementation of machine learning techniques on real amputee data.

Amputation is a tragic occurrence that affects millions of individuals worldwide. It involves surgical limb removal. According to statistics, congenital limb defects, illnesses, and trauma are the three main causes of amputation in that order of incidence. There are two main categories of amputation: upper and lower limbs. These are further divided into transfemoral, transradial, Transcarpal, and complete disarticulation of knee, ankle, foot, and shoulder. Since many nations do not keep track of the number of persons who have had limbs amputated, it is impossible to determine the precise number of amputations that have occurred globally. According to figures from 2017, there are 35.3 million lower limb amputees worldwide, totaling 57.7 million amputees [1].

Since both the knee and the ankle joint are lost during a transfemoral amputation, moving around becomes quite challenging [2]. The human knee is essential to maintaining human life. It is in charge of carrying out everyday tasks. The most frequent and necessary daily exercise is walking. It is a complicated procedure that combines the capabilities of several muscles and tendons. People may transfer their bodies from one location to another and retain their equilibrium thanks to this optimal construction. Walking is a simple process for healthy people, but for people having an amputation, it is very tough and difficult.

Prosthetic knees are designed to address this problem. Three categories—Active, Semi-active, and Passive—are used to describe prosthetic knees. The passive prosthetic knee has a straightforward mechanism and no microprocessor or actuator. Due to the fixed damping it offers, the gait cycle is improper. The Active and Semi-active prosthetic knees are microprocessor-based prosthetic knees that use electric motors and dampers to offer variable damping. The Active Prosthetic Knee produces both positive and negative energy while a person is moving, therefore it uses a lot of electricity, which makes it pricey and has a short lifespan. Semi active prosthetic knees, on the other hand, have dampers that dissipate energy and use less electricity. Prosthetic knees that are both active and semi-active offer more benefits than inactive knees [3].

Detecting the gait cycle is the most important component of prosthetic knees. The human gait cycle is periodic; it starts from the heel strike (HS) and completes the cycle, then repeats itself [4]. The gait cycle has two major phases: Stance and Swing. Stance makes up around 60% of a gait cycle. The stance phase occurs when the foot is in contact with the ground. The swing phase, which begins when one foot leaves the ground and lasts until it strikes the ground again, makes up around 40% of the remaining gait cycle. The HS shows the start of stance phase, while toe off (TO) shows the start of the swing phase. A typical walking cycle has two supports for single limbs and two supports for double limbs. When only one foot is on the ground it is called single support and when feet of both legs are in contact with the ground then it is called double support [5].

There are several sub-phases within the gait phase. Although the number of gait phases employed by different writers varied, we have grouped the gait phases into 5 sub-phases in this work. Two for the swing and three for the stance. Initial stance/foot contact, flat foot, heel off, initial and terminal swing are the sub phases. Regarding the basic principles of gait, a gait cycle starts with heel strikes and ends with heel strike.

One foot's heel touches the ground during the initial phase. On this foot, the burden will soon be transmitted. After heel strike, the body weight is supported entirely by one foot in the second phase of the stance, known as flat foot or mid stance. The third stage is heel off, in which the toes are in contact to the ground and while the heel is not. Stability is necessary for all stance stages. 60% of the gait cycle is made up of the initial stance, the flat foot, and the heel off. Toe off initiates the initial swing phase. The heel is in air, and the toe is about to enter the air during this phase. Mid stance is also called single support because the whole body weight is on single foot. The opening swing lasts until the mid-swing, at which point the terminal stance starts. It continues until the heel strike, at which point the subsequent cycle begins. Each swing phase accounts for 20% of the gait cycle. Fig. 1 depicts the gait phases.

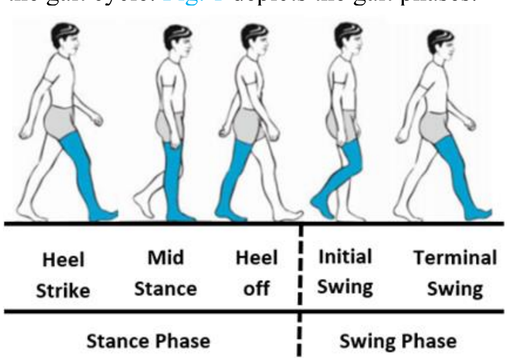

Figure 1. Gait phases in gait cycle.

Nearly all studies employed the data of healthy volunteers for the detection of the gait phase. There is a difference between data on sound leg and amputated leg. The sound leg data is used to validate the prosthetic knee. Measurements based on IMU signals, IMU and force sensor combinations and EMG signals are utilized to determine gait phases.

Mo et al. [6] proposed a method for identifying HC (heel contact) and TO occurrences that employed three IMUs mounted on the foot, shank, and pelvis to assess the peaks of acceleration impulses from distinct sensor locations. The initial contact identification was done by calculating the mean absolute difference.

A method for recognizing transitions between gait phases was demonstrated by Gorsic et al. [7]. The algorithm received information from two pressure insoles and seven IMUs attached to various body segments. The results demonstrated extremely accurate identification, with an overall online detection rate of 97% over four phases. The detection performance was equivalent when employing the HMM (hidden Markov model) technique, despite the process being less involved without learning dataset and model training. Based on deep CNN, in [8] a pressure, gyro, and acceleration sensor arrays were used to create a smart insole. With this method, the primary categorization scheme for the numerous gait characteristics was found in the stance and swing phases.

By incorporating eight EMG sensors positioned beneath the feet of an exoskeleton, Joshi et al. [9] offered an LDA-based control system that could recognize eight gait stages. The complexity of the signal processing technique was increased by the mean value, variance, waveform time, and slope sign change. The EMG-based method was less promising because of the challenging process for data collection and digital signal processing; moisture causes sensitivity issues that gather between the skin and sensors, and the positioning of the sensors beneath the subject's epidermis.

There are two types of methods that are frequently used to identify the human gait cycle. A heuristic rule basis makes up the first, while machine learning makes up the second. Thresholds are found through the heuristic rule base system. The many types of value rules known as threshold algorithms are used to specify certain elements of gait phases or occurrences.

Meng et al. [10] used IMU inputs to calculate the knee and tibia angles to detect seven gait phases with the best detection result of 100% reliability. Boutaayamoue et al. [11] suggested a technique that accurately recognized four events within 10 ms. For each of the two systems, four sensors were required, and they had to be fastened to the leg components. The author done validation of gait events by thresholding, indicating crucial delays. He combined the heuristics and the zero-crossing technique, for example, records the delay at an average of 100 ms to compute HS (heel strike) and TO.

One of the most employed techniques for classifying gait phases in both offline and online data is machine learning (ML) algorithms. Different gait phase identification methods have been developed by using CNN, DLNN, NN models, and HMM, among other machine-learning approaches. For instance, a number of uses for the ML subfield of HMM have been proposed. Regardless of whether it is being utilized for online or offline detection, this approach reliably detects four event stages [12].

Using a uniaxial gyro mounted on foot in order to monitor the foot's angular velocity, Mannini et al. [13, 14] used HMM to identify the 4 gait phases of initial stance (heel strike), mid stance (flat foot), terminal stance (heel off) and initial swing (toe off). The author identified the gait phases in real-time. Although ML algorithms show promising results as compared to threshold data but the study on real time data of prosthetic knee is missing. In this study we will apply machine learning algorithms on real time data collected from semi active prosthetic knee. The kinematic data is collected from prosthetic knee and then imported in excel sheet. Further signal processing is applied and finally ML algorithms are applied on that data.

2. Methodology

There are three input parameters that were used in this study: Knee angle, Knee angular velocity and Knee force data. This kinematics data was obtained from incremental encoder and strain gauges mounted on shank of the prosthetic knee. The following are the steps applied:

- Data preparation
- Noise reduction
- Implementation of machine learning algorithm

The labelled data was provided by Advanced Robotics & Automation Lab, University of Engineering and Technology Peshawar. This sensors data was collected from a Magnetorheological damper based semi-active prosthetic knee. The raw data contained two variables: knee angle and force readings. The knee angular velocity was calculated from angle data. The data was smoothened with digital low pass filter. After data pre-processing, Statistical features was selected and applied. Finally, there were three ML models were applied: DT, LDA and SVM. Fig. 2 shows data of knee angle, knee angular velocity and knee force.

3. Feature Selection

A statistical-based feature selection method was used in this study. It involves a relationship of input and output by using statistics. There are two kinds of features time and frequency domain features. The 4 distinct features in the time domain were applied on data. These were Mean, Range, Variance and Standard Deviation [15].

Mean: this feature can be described as,

$$M = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

Range: it is given by,

$$Range\ (x) = Max(x) - Min(x) \tag{2}$$

Variance: VAR is given by,

$$VAR = \frac{1}{1-N}\sum_{i=1}^{N} x_i^2 \tag{3}$$

Standard Deviation: SD is given by,

$$S.D = \sqrt{(}\frac{1}{1-N}\sum_{i=1}^{N} x_i^2 \tag{4}$$

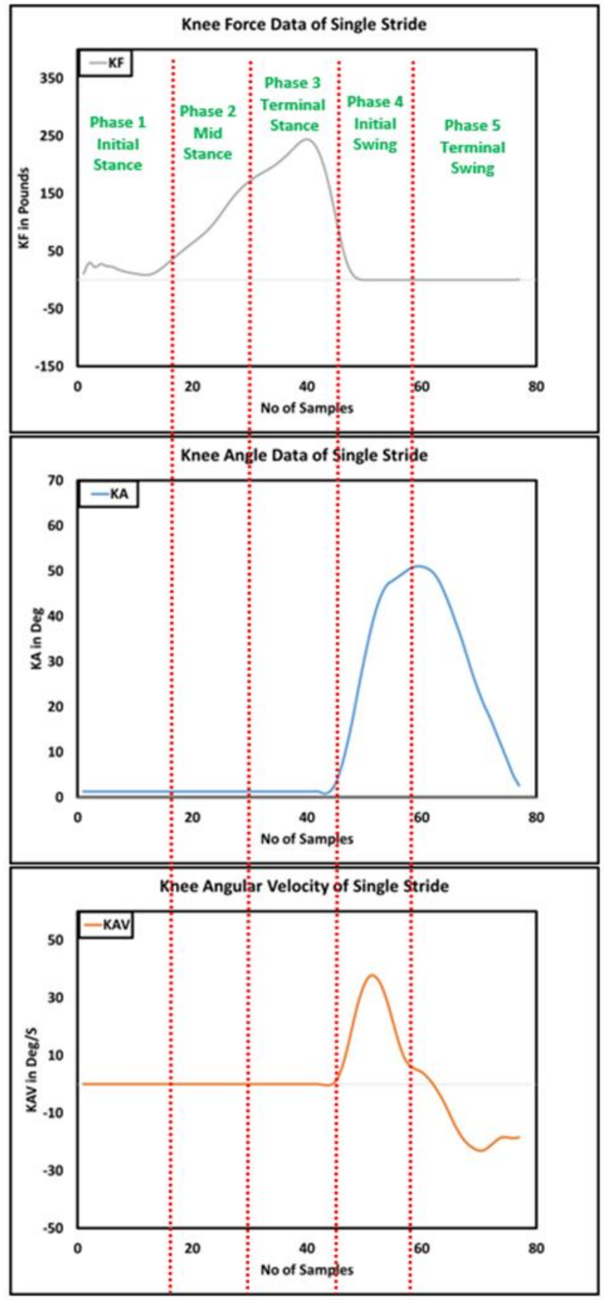

Figure 2. Semi-active prosthetic knee sensors data.

Segmentation, feature extraction, and algorithms were applied to MATLAB by MathWorks. The first step is data segmentation. Segment is data in fixed time slot. This segment is then used to estimate signal feature. The segment should not be very small because it will lead to bias and variance, and it should not be very large as it will lead to high computational load and will fail to perform in real-time operation. Segmentation was done using a non-overlapping windowing technique. There were 50 samples selected for each gait phase. From these samples first 5 and last 5 were not selected and middle 40 readings were selected as, data in the middle of the segment is stable. There were 3 segments (40 samples angle vector, 40 samples angular velocity vector and 40 samples force vector). The four features were applied to each segment separately and a total of 12 features were obtained (4 from each segment). After feature extraction, these 12 features were input to machine learning models.

4. Machine Learning Algorithms

4.1. Decision Trees (DT)

The main method used for classification and prediction is decision trees. Decision tree learning, which focuses on categorization rules that are displayed as decision trees inferred from a collection of disordered and irregular cases, is a typical inductive technique based on examples. It performs top-down recursive attribute comparisons across internal decision tree nodes, assesses downstream branches in light of multiple node characteristics, and draws a conclusion from the decision tree's leaf nodes. The whole tree is governed by a set of disjunctive expression rules, and each node, from the root to the leaf, is governed by a conjunctive rule [16].

4.2. Linear Discriminant Analysis (LDA)

LDA may be employed to identify various patterns in addition to the two categories into which it is frequently used to divide them. The LDA assumes that the classes are linearly separable results in the production or creation of the multi-LDA function, which represents many hyperplanes in the space of feature. If there are two classes, the LDA creates a single hyperplane and projects the data onto it in a way that minimizes the distance between them [17].

4.3. Support Vector Machine (SVM)

It is a technique that can be used to solve both classification and regression issues. The SVM algorithm uses a training dataset separated into different classes to find a hyperplane that provides the greatest margin between the data that belong to distinct classes. Hyperplane has the largest margin. SVM only uses the objects (samples) on the margin's edges rather than dividing them based on variations in class means (known as support vectors). The separating hyperplane is supported (defined) by the vectors (data points) closest to the margin, giving the SVM method its name [18].

5. Result and Discussion

The 3d scatter plots between features extracted from three input parameters for 5 distinct gait phases in complete gait/locomotion cycle (initial stance, mid stance, terminal stance, initial swing, and terminal swing) are shown below. With the use of a scatter plot, you can investigate the connections between two or three variables for a given collection of data. In these graphs the three axes correspond to three input parameters. Each feature that Figs. 3 to 6 is represented in separate graphs.

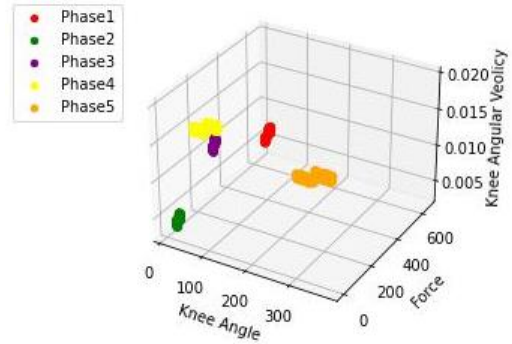

Figure 3. Variance.

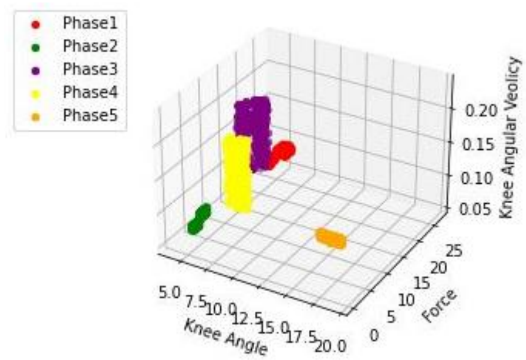

Figure 4. SD.

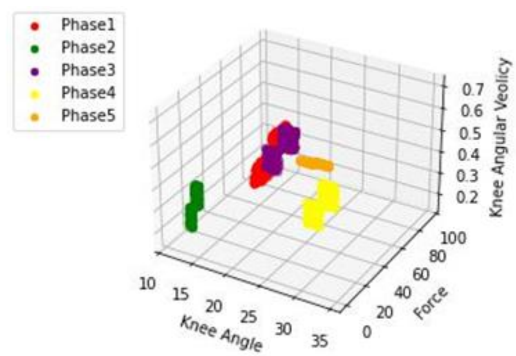

Figure 5. Range.

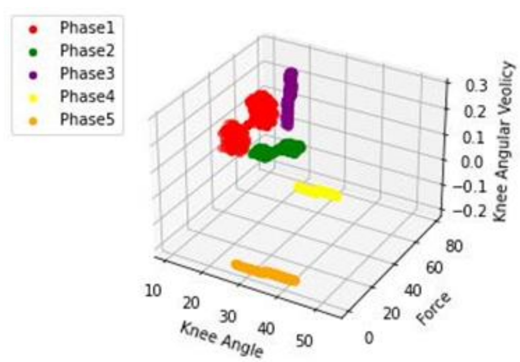

Figure 6. Mean.

Statistic and ML Toolbox of MATLAB is used to generate classification results. Table 1 shows the model parameters.

Table 1. Model parameters.

Model	Parameter	Value
SVM	Kernel Function	Linear
	Kernel Scale	Automatic
	Box Constraint Level	1
	Multiclass Method	one vs. One
	Standardized Data	yes
LDA	Covariance Structure	Full
DT	Max Splits	100
	Split Creation Criterion	Gini's Index
	Surrogate Decision Split	off

Table 2 provides a comprehensive evaluation of three machine learning techniques, specifically Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), and Decision Tree (DT). The study comprises various performance metrics, which suggest that LDA exhibits greater performance than other methodologies. The LDA model has exceptional performance in terms of precision (95.76%), recall (95.58%), and F1 Score (95.67%), indicating its robust classification capabilities. Additionally, the LDA model demonstrates a superior level of validation accuracy (95.60%) and test accuracy (95.40%), hence highlighting its consistent and accurate performance. The performance of the SVM model is noteworthy, as it achieves a high precision rate of 95.70%, a recall rate of 95.18%, and an F1 Score of 95.44%.

Furthermore, it attains a noteworthy validation accuracy of 95.20% and a test accuracy of 95.10%. Despite displaying somewhat lower precision and F1 Score in comparison to LDA and SVM, DT still maintains a competitive degree of accuracy. More precisely, the achieved results include a precision rate of 93.73%, a recall rate of 95.18%, and an F1 Score of 94.45%. Furthermore, it exhibits exceptional performance with regard to validation accuracy, achieving a score of 95.20%, as well as test accuracy, with a score of 95.05%. In summary, it can be observed that LDA and SVM demonstrate robust classification abilities, whilst DT showcases a notable level of accuracy.

Table 2. Accuracy of algorithms.

Algorithms	Precision	Recall	F1 Score	Validation Accuracy	Test Accuracy
SVM	0.9570	0.9518	0.9544	0.9520	0.9510
LDA	0.9576	0.9558	0.9567	0.9560	0.9540
DT	0.9373	0.9518	0.9445	0.9520	0.9505

The model accuracy of all algorithms with their ROC curve is shown in the Figures below.

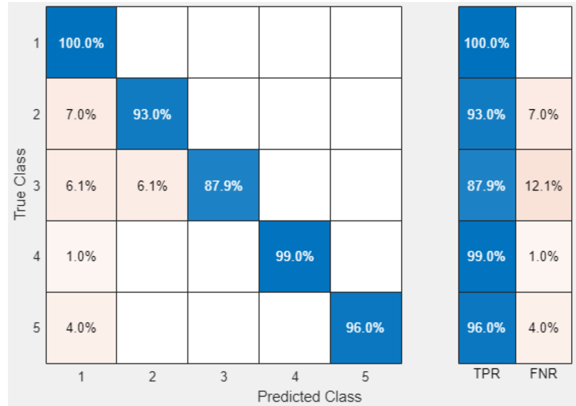

Figure 7. LDA.

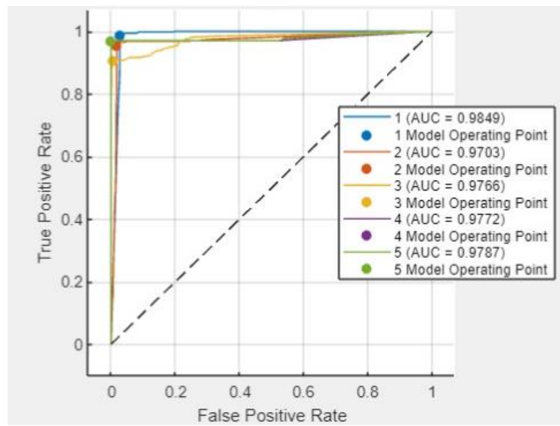

Figure 8. ROC curve for LDA.

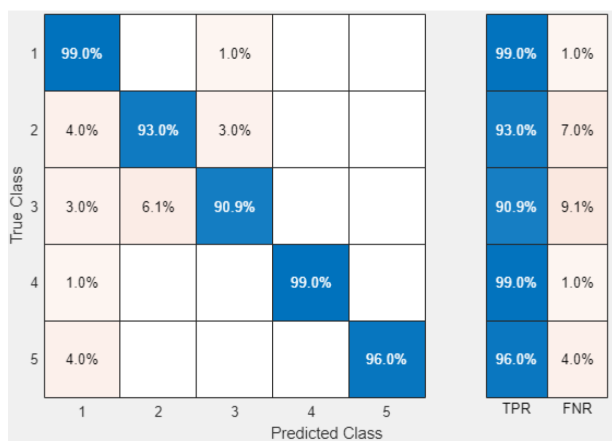

Figure 9. SVM.

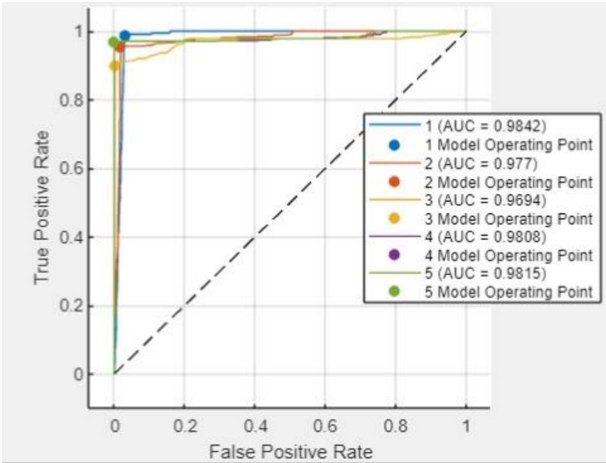

Figure 10. ROC curve for SVM.

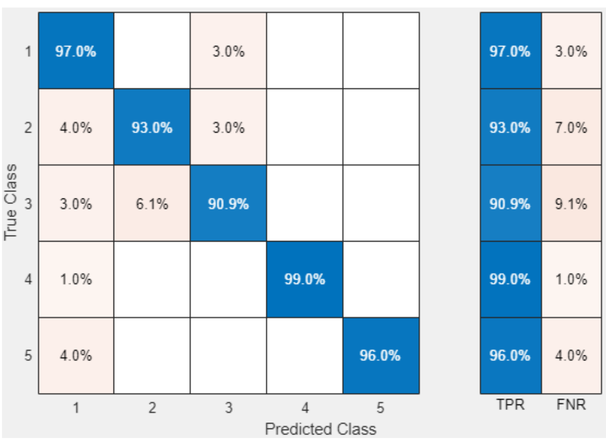

Figure 11. DT.

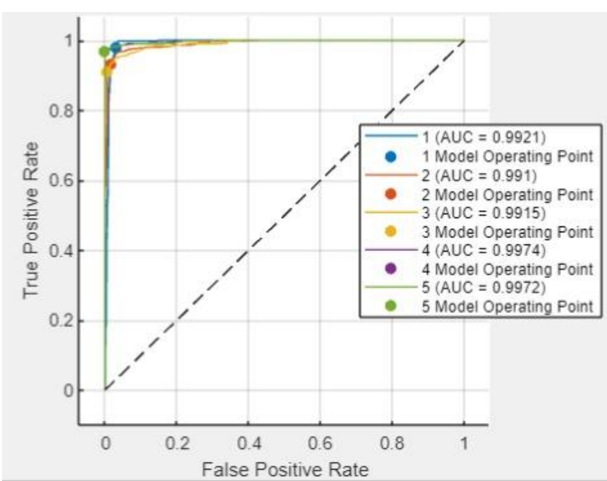

Figure 12. ROC curve for DT.

6. Discussion

In conclusion, this study has successfully investigated a significant issue in the field of prosthetic knees, specifically focusing on individuals who have experienced amputation of the lower leg. The deployment of prosthetic knees plays a crucial role in facilitating the rehabilitation of movement and functionality for those who have suffered amputation. The selection between mechanical and microprocessor-based knees significantly impacts the entire quality of life encountered by individuals in question. Despite the advancements in sensor and actuator systems in microprocessor-based knees, accurately identifying the gait cycle remains a challenging task. As the complexity of gait phases increases, the effectiveness of classic rule-based systems in terms of accuracy decreases. The current study has achieved significant advancements through the utilization of machine learning techniques to examine empirical data collected from individuals who have undergone transfemoral amputations in real-world situations.

The application of machine learning methodologies, including LDA, SVM and DT, in the examination of knee data involving force, angle, and angular velocity, has yielded promising results. The performance of the LDA model was shown to be superior, with a validation accuracy of 95.6% and a test accuracy of 95.40% (see Table 3). The SVM and DT algorithms demonstrated notable performance, attaining validation accuracies of 95.2% and test accuracies of 95.10% and 95.05% correspondingly. The findings of this study indicate that the utilization of machine learning has the potential to significantly improve the precision of gait phase recognition in microprocessor-based prosthetic knees. The enhancement in precision possesses the potential to augment the overall functioning and performance of prosthetic knees.

Table 3. Related studies.

Study	Sensors	Method	Phases/ Events	Accuracy
Zhen et al. (2019) [19]	Three IMUs	LSTM-DNN	2	91.8%
Liu et al. (2016) [20]	Four angular sensors	NN	8	94.5%
Ledoux et al. (2018) [21]	One IMU	THR, LDA, QDA	2	92%
Attal et al. (2018) [22]	Pressure	Hidden Markov Model	6	83.21%
Kim et al. (2020) [23]	One IMU	Time-frequency Analysis	2	97% (TO running events) 99% (Other events)
Our Research	Single IMU, Encoder and Force Sensor	SVM, LDA, DT	5	SVM (95.20%) LDA (95.60%) DT (95.20%)

7. Future Work

To ensure the robustness of our machine learning techniques across different user profiles, it is imperative to undertake additional validation utilizing larger and more diverse datasets of amputee gait data in future research attempts. Moreover, the examination of deep learning methods and the incorporation of these algorithms into prosthetic knee systems possess the capability to augment the precision and flexibility of gait phase recognition in real-time settings. There is a possibility of greatly improving the quality of life for individuals with lower limb amputations by effectively incorporating machine learning into microprocessor-based prosthetic knees. This integration has the potential to offer enhanced movement that is both more natural and useful.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] W. Cao, H. Yu, W. Chen, Q. Meng, and C. J. I. A. Chen, "Design and evaluation of a novel microprocessor-controlled prosthetic knee," vol. 7, pp. 178553-178562, 2019.

[2] M. U. Qadir, M. A. Khan, M. Hussain, I. ul Haq, N. Akhtar, and K. Shah, "Design and Analysis of Knee Joint for Transfemoral Amputees," in *2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*, 2021, pp. 1-5: IEEE.

[3] S. Hoh, J. Chong, and A. C. Etoundi, "Design of a virtual testing platform for assessing prosthetic knee joints," in *2020 5th International Conference on Advanced Robotics and Mechatronics (ICARM)*, 2020, pp. 576-581: IEEE.

[4] H. T. T. Vu, F. Gomez, P. Cherelle, D. Lefeber, A. Nowé, and B. Vanderborght, "ED-FNN: A New Deep Learning Algorithm to Detect Percentage of the Gait Cycle for Powered Prostheses," vol. 18, no. 7, p. 2389, 2018.

[5] K. A. Shorter, J. D. Polk, K. S. Rosengren, and E. T. J. C. B. Hsiao-Wecksler, "A new approach to detecting asymmetries in gait," vol. 23, no. 4, pp. 459-467, 2008.

[6] S. Mo, D. H. J. G. Chow, and posture, "Accuracy of three methods in gait event detection during overground running," vol. 59, pp. 93-98, 2018.

[7] M. Goršič *et al.*, "Online phase detection using wearable sensors for walking with a robotic prosthesis," vol. 14, no. 2, pp. 2776-2794, 2014.

[8] S.-S. Lee, S. T. Choi, and S.-I. J. S. Choi, "Classification of gait type based on deep learning using various sensors with smart insole," vol. 19, no. 8, p. 1757, 2019.

[9] C. D. Joshi, U. Lahiri, and N. V. Thakor, "Classification of gait phases from lower limb EMG: Application to exoskeleton orthosis," in *2013 IEEE Point-of-Care Healthcare Technologies (PHT)*, 2013, pp. 228-231: IEEE.

[10] X. Meng, H. Yu, and M. P. Tham, "Gait phase detection in able-bodied subjects and dementia patients," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 4907-4910: IEEE.

[11] M. Boutaayamou *et al.*, "Development and validation of an accelerometer-based method for quantifying gait events," vol. 37, no. 2, pp. 226-232, 2015.

[12] J. Bae and M. J. M. Tomizuka, "Gait phase analysis based on a Hidden Markov Model," vol. 21, no. 6, pp. 961-970, 2011.

[13] A. Mannini, V. Genovese, A. M. J. I. j. o. b. Sabatini, and h. informatics, "Online decoding of hidden Markov models for gait event detection using foot-mounted gyroscopes," vol. 18, no. 4, pp. 1122-1130, 2013.

[14] A. Mannini and A. M. Sabatini, "A hidden Markov model-based technique for gait segmentation using a foot-mounted gyroscope," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 4369-4373: IEEE.

[15] S. Negi, Y. Kumar, and V. Mishra, "Feature extraction and classification for EMG signals using linear discriminant analysis," in *2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA)(Fall)*, 2016, pp. 1-6: IEEE.

[16] Q.-y. Dai, C.-p. Zhang, H. J. I. J. o. D. T. Wu, and Application, "Research of decision tree classification algorithm in data mining," vol. 9, no. 5, pp. 1-8, 2016.

[17] S. Balakrishnama, A. J. I. f. S. Ganapathiraju, and i. Processing, "Linear discriminant analysis-a brief tutorial," vol. 18, no. 1998, pp. 1-8, 1998.

[18] S. Vishwanathan and M. N. Murty, "SSVM: a simple SVM algorithm," in *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)*, 2002, vol. 3, pp. 2393-2398: IEEE.

[19] T. Zhen, L. Yan, and P. J. A. Yuan, "Walking gait phase detection based on acceleration signals using LSTM-DNN algorithm," vol. 12, no. 12, p. 253, 2019.

[20] D.-X. Liu, X. Wu, W. Du, C. Wang, and T. J. S. Xu, "Gait phase recognition for lower-limb exoskeleton with only joint angular sensors," vol. 16, no. 10, p. 1579, 2016.

[21] E. D. J. I. T. o. B. E. Ledoux, "Inertial sensing for gait event detection and transfemoral prosthesis control strategy," vol. 65, no. 12, pp. 2704-2712, 2018.

[22] F. Attal, Y. Amirat, A. Chibani, and S. J. I. A. T. o. M. Mohammed, "Automatic recognition of gait phases using a multiple-regression hidden Markov model," vol. 23, no. 4, pp. 1597-1607, 2018.

[23] J. Kim, M. N. Bae, K. B. Lee, and S. G. J. E. J. Hong, "Gait event detection algorithm based on smart insoles," vol. 42, no. 1, pp. 46-53, 2020.