ASME Journal of Engineering and Science in Medical Diagnostics and Therapy
Online journal at:
https://asmedigitalcollection.asme.org/medicaldiagnostics

Abdullahi Olapojoye[1]
Department of Mechanical Engineering,
University of Texas at Dallas,
Richardson, TX 75080
e-mail: abdullahi.olapojoye@utdallas.edu

Abhishek Singh
Department of Mechanical Engineering,
University of Texas at Dallas,
Richardson, TX 75080
e-mail: aps180000@utdallas.edu

Eri Nishi
Department of Electrical Engineering,
Setsunan University,
Neyagawa, Osaka Prefecture 572-8508, Japan
e-mail: eri.nishi@ele.setsunan.ac.jp

Baowei Fei
Department of Bioengineering,
University of Texas at Dallas,
Richardson, TX 75080;
Department of Radiology,
UT Southwestern Medical Center,
Dallas, TX 75390
e-mail: bfei@utdallas.edu

Aria Nostratinia
Department of Electrical and Computer Engineering,
University of Texas at Dallas,
Richardson, TX 75080
e-mail: aria@utdallas.edu

Fatemeh Hassanipour
Department of Mechanical Engineering,
University of Texas at Dallas,
Richardson, TX 75080
e-mail: fatemeh@utdallas.edu

Infants Sucking Pattern Identification Using Machine-Learned Computational Modeling

Breastfeeding involves a complex coordination of swallowing, breathing, and sucking, with the infant's sucking proficiency being crucial for adequate nutrient intake. However, real-time assessment of milk intake is difficult, and issues with sucking often become apparent after the infant shows signs of nutrient deficiency. Traditional assessments by clinicians rely on the expertise and subjective judgment of healthcare professionals, enabling personalized evaluations. In this research, we introduce a novel approach to identifying sucking patterns by leveraging data collected from infants during breastfeeding sessions. This method utilizes artificial nipple-based sensors to capture the tongue forces exerted by infants, generating valuable clinical data. In the analysis of the collected time-series data, we applied machine-learned computational modeling (MLCM) algorithms to extract pertinent features and identify distinctive sucking patterns. The best-performing model demonstrated an accuracy of 90%, an 80% recall score, a perfect 100% precision score, a 0.90 f1-score, and an area under the curve (AUC) of 0.80. The proposed classification system has the potential to serve as a reliable decision-support tool for clinicians, offering valuable insights into infants' sucking behaviors. By integrating machine learning (ML)-based computational modeling into clinical practice, we aim to enhance the early identification of unhealthy sucking patterns, allowing for timely interventions and pro-active healthcare management.
[DOI: 10.1115/1.4066459]

Keywords: infants sucking pattern, clinical diagnostics, biomechanics, machine learning, electro-mechanical sensors, neural networks

Downloaded from http://asmedigitalcollection.asme.org/medicaldiagnostics/article-pdf/8/3/031003/7384220/jesmdt_008_03_031003.pdf by University Of Texas At Dallas user on 23 February 2026

1 Introduction

Breastfeeding is a complex and dynamic process that requires a finely tuned coordination of several key actions: sucking, swallowing, and breathing. Newborns come equipped with these related reflexes, enabling them to initiate feeding shortly after birth. The complex sucking mechanism involves the elevation of the jaw, pressing the tongue against the nipple's front tip, and a subsequent peristaltic action that effectively extracts milk. This roller-like tongue movement, akin to peristalsis, facilitates the efficient expression of milk from the nipple [1]. However, premature and low birth weight infants, or those with oral motor dysfunction, may face feeding challenges due to underdeveloped primitive reflexes. Understanding the feeding behavior of infants becomes paramount to address such challenges.

Various scientific methodologies, including ultrasonic tomography and intra-oral video cameras, have been deployed in research endeavors aimed at comprehending the nuanced peristaltic-like actions of the tongue during feeding [2–9]. This exploration not only enhances our understanding of infant feeding patterns but also paves the way for tailored interventions to support the unique needs of premature and low birth weight infants in their breastfeeding capabilities.

Machine learning (ML) is a subfield of artificial intelligence that enables learning from and making predictions based on data. Advances in computational power and algorithm design have led to efficient training of these models from large quantities of data, and have improved their accuracy. In biology and medicine, ML can be

[1]Corresponding author.

Contributed by the Applied Mechanics Division Technical Committee on Dynamics & Control of Structures & Systems (AMD-DCSS) of ASME for publication in the Journal of Engineering and Science in Medical Diagnostics and Therapy. Manuscript received August 9, 2024; final manuscript received August 28, 2024; published online September 30, 2024. Assoc. Editor: Lulu Wang.

Copyright © 2025 by ASME

applied to analyze complex biological data and enhance the accuracy and speed of medical diagnoses. Vast amounts of information can be processed in an efficient and reproducible manner, including genetic sequences, medical images, and patient records [10–13].

The integration of machine-learned computational modeling (MLCM) techniques is becoming increasingly prevalent in the realm of healthcare [14–16], with studies focusing on assessing and intervening in various aspects of infant and children care [17–20]. Quality improvement initiatives in healthcare are supported by the predictive capabilities of MLCM and the explanatory insights offered by explainable artificial intelligence techniques [21].

The application of MLCM in the realm of infant care is evolving. Elgersma et al. [22] employed ML to identify factors influencing human milk feeding and direct breastfeeding for infants with single ventricle congenital heart disease, demonstrating the versatility of the technique in addressing complex healthcare dynamics. Additionally, ML techniques have been successfully utilized to identify various factors affecting breastfeeding behaviors, both positively and negatively, leveraging insights derived from social media data [23].

In alignment with recent advancements, our project focuses on utilizing MLCM to develop a robust classification system for infant sucking patterns. In this area, MLCM techniques offer a strong impact in assessing neonatal health, feeding issues, and early developmental problems. In particular, MLCM algorithms can analyze large datasets of infant suckling sounds, pressures, and rhythms to identify patterns that are indicative of healthy versus problematic suckling behaviors. These patterns might include the frequency, strength, and consistency of suckling, which are crucial for diagnosing issues such as poor latch or insufficient milk transfer.

With enough training data, MLCM algorithms can potentially attain previously unavailable levels of accuracy. Further, MLCM models can detect subtle abnormalities in suckling patterns that might be missed by human observation, leading to early diagnosis of potential feeding disorders or neurological issues that affect feeding. MLCM-powered devices can provide real-time feedback to parents and healthcare providers about an infant's suckling efficiency, allowing for immediate interventions if necessary. This is particularly useful in neonatal intensive care units or at home for premature infants.

Our approach leverages artificial nipple-based sensors to capture the tongue forces exerted by infants during breastfeeding sessions. These sensors collect precise clinical data, which is then used to develop a system capable of identifying and classifying sucking patterns in real-time. By providing objective, data-driven insights, our system aims to enhance the accuracy and consistency of infant sucking assessments, ultimately improving the diagnosis and management of feeding challenges in infants, particularly those who are premature, have low birth weight, or exhibit oral motor dysfunction.

2 Data Extraction and Preparation

For the purpose of data collection and better understanding of the complex dynamics of infant feeding, a measurement sensor device embedded in artificial nipple designed to capture the peristaltic-like movement of an infant's tongue during sucking. This device is equipped with two sensors attached to a baby bottle nipple, offering a perspective into the forces generated during the sucking activity. The force sensor integrated into the artificial nipple uses a cantilever structure, utilizing a thin stainless-steel plate as a beam. A strain gauge, strategically affixed to its surface, facilitates precise force measurement. The two sensor channels are positioned at the tip and base of the nipple to measure forces at the root and tip of the tongue, respectively. Channel 1 (ch. 1) captures forces at the root, while Channel 2 (ch. 2) measures forces at the tip of the tongue as depicted in Fig. 1(*a*).

The experimental protocols used in this study were reviewed and received ethical clearance by the "Research Ethics Committee for

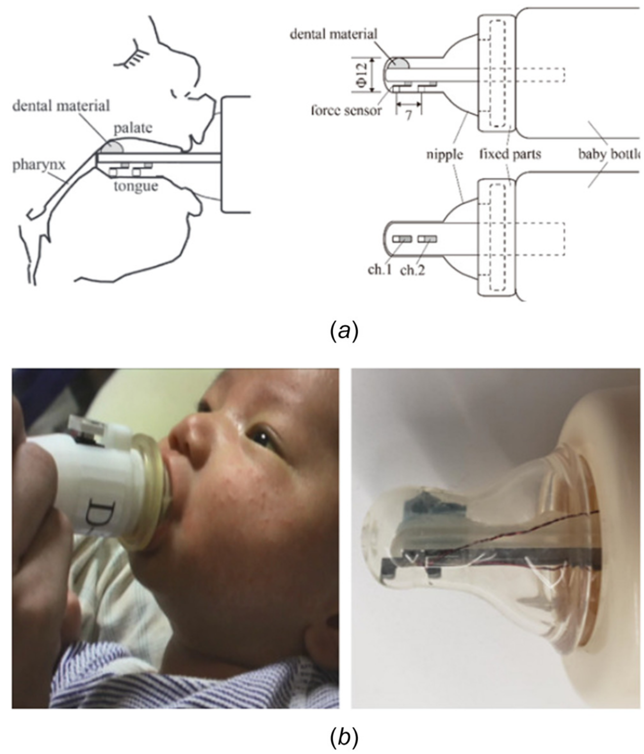

Fig. 1 Device and experiment setup: (*a*) schematic diagram of the artificial nipple with integrated force sensor and (*b*) the experimental setup demonstrating sensor device positioning on an infant

Human Subjects" at Setsunan University (approval number 2015–006) [24,25]. Before commencing the measurements, the aims and methods of the study were thoroughly explained to the infants' parents. Written informed consents were acquired from the parents to confirm their understanding of and consent to their infants' participation in the study. During the experiments, the examiner cradled each infant and carefully positioned the sensor device in their oral cavity, mimicking the typical posture observed during breastfeeding, as illustrated in Fig. 1(*b*).

To translate these force measurements into usable data, the output signal undergoes A/D conversion through a bridge circuit and an amplifier before being transmitted to a personal computer via universal serial bus. Operating at a sampling frequency of 100 Hz with a quantization resolution of 12 bits, this system ensures detailed and accurate representation of the peristaltic-like tongue movements. Figure 2 describes the workflow used in this study. The collected data reveal force waveforms generated by tongue movements that closely resemble sine waves, incorporating an offset component. By identifying peak values on a period, we can effectively quantify the maximum force exerted during the peristaltic movement of the tongue. This information, derived from the force sensors embedded in the artificial nipple, opens avenues for MLCM algorithms to analyze and classify these waveforms, contributing to a deeper understanding of infant feeding dynamics and potentially aiding in interventions for feeding challenges. Figure 3 represents a force profile of a healthy sucking, characterized by regular, rhythmic patterns and an unhealthy sucking, with irregular and less defined patterns, as measured by sensors on channels 1 and 2.

The data preparation process commenced by addressing several key aspects. Initially, a low-pass filter was applied to reduce noise and enhance the important signal components in the force data. This was followed by systematically cleaning the data to handle missing values, outliers, redundant, and duplicate entries. Subsequently, input and output variables were separated; the force data measured in Newtons and time in seconds were identified as crucial inputs, while the output involved classifying the sucking patterns as healthy or unhealthy. Label encoding was employed to convert categorical

Downloaded from http://asmedigitalcollection.asme.org/medicaldiagnostics/article-pdf/8/3/031003/7384220/jesmdt_008_03_031003.pdf by University Of Texas At Dallas user on 23 February 2026

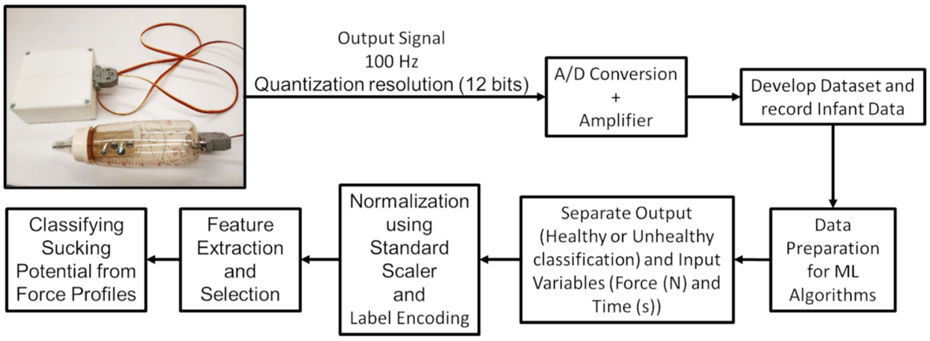

Fig. 2 Sucking pattern classification workflow

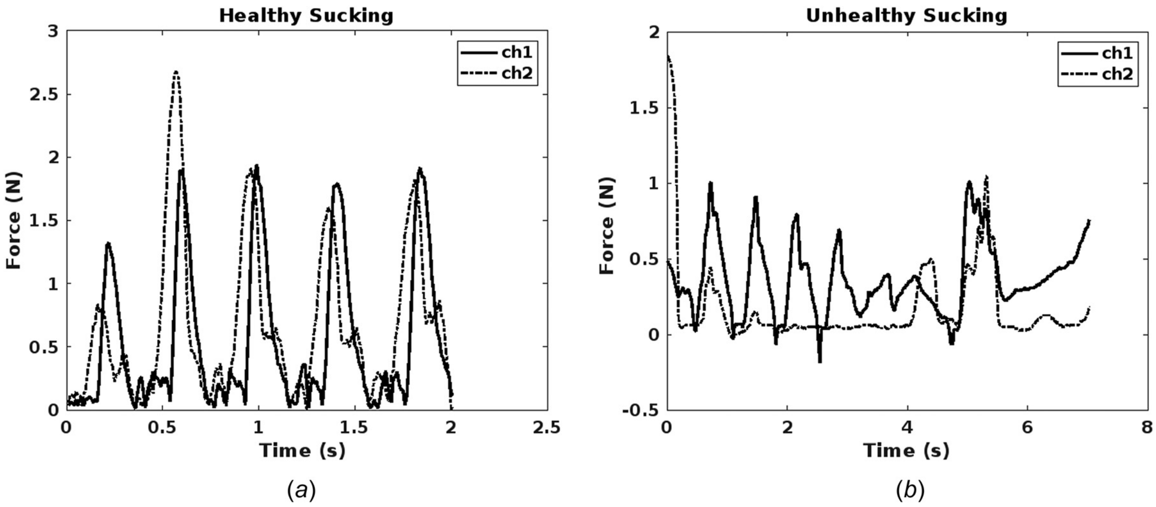

Fig. 3 Comparative force-time plots for infant sucking patterns as measured by sensors on channels 1 (solid line) and 2 (dashed line): (*a*) healthy sucking, characterized by regular, rhythmic patterns and (*b*) unhealthy sucking, which show irregular and less defined patterns

variables into a format compatible with the MLCM model. The dataset was then divided into training and test sets to ensure model generalization, comprising a total of 40 subjects, with 18 infants categorized under unhealthy sucking and 22 under healthy sucking. This categorization laid the groundwork for subsequent analyses. Finally, normalization was performed using the Standard Scaler to ensure uniform scaling of features. This meticulous data preparation process ensures that the models receives clean, standardized, and properly formatted input, setting the stage for robust and reliable analysis of infant sucking patterns.

2.1 Determination of Ground Truths. The methods used to establish the clinical ground truths in our dataset, essential for validating the accuracy of our computational models in predicting infant sucking patterns, are outlined here. Criteria such as a birth weight under 2500 g, weight under 3000 g at 30 days postbirth, gestational age shorter than 35 weeks, and the presence of palatal anomalies were identified as potential markers for developmental challenges. These parameters are important as they can be early indicators of developmental issues and were verified and comprehensive evaluated by qualified healthcare professionals [24].

In the specific context of this study, which focuses on analyzing sucking potential, we narrowed our analysis to the mechanical force profiles generated during the sucking process. While factors like birth weight, gestational age, and palatal issues could influence an infant's sucking ability, the exact impact of these variables remains uncertain. There are maternal factors like nipple shape, mammary gland count, and milk secretion, that may also influence the efficacy of breastfeeding [26]. To eliminate any potential bias in our computational analysis and ensure the reliability of our findings, we excluded these broader developmental factors. This targeted approach allows us to focus exclusively on the mechanical aspects of sucking behavior, providing a clear, objective basis for our study of infant sucking patterns.

3 Machine-Learned Computational Model Development

In this study, four computational models were developed. Two classical ML and two deep learning models were developed to classify infants sucking behavior based on time series force data. Detailed descriptions of the algorithms are presented in this section.

3.1 Support Vector Classifier. A support vector classifier (SVC), also known as a support vector machine (SVM), is a supervised ML algorithm used for classification and regression tasks. The primary goal of an SVM is to find a hyperplane that best separates different classes in the feature space. In the case of a binary classification problem, this hyperplane aims to maximize the margin between the two classes while minimizing the classification error. The margin is the distance between the hyperplane and the nearest data point (support vector) from either class. SVM can handle nonlinear decision boundaries by transforming the input features into a higher-dimensional space. The optimization problem involves finding the optimal weights (coefficients) for each feature and the bias term. The objective function includes a regularization term, which penalizes the misclassification of data points. The tradeoff

Downloaded from http://asmedigitalcollection.asme.org/medicaldiagnostics/article-pdf/8/3/031003/7384220/jesmdt_008_03_031003.pdf by University Of Texas At Dallas user on 23 February 2026

Fig. 4 SVM classification surface for infant sucking patterns. Classification is performed on the force data, which has two kinds of subjects represented by triangle and square, respectively.

parameter (C) controls the balance between maximizing the margin and minimizing the misclassification.

Support vector machine include key parameters such as the kernel choice and the regularization parameter (C), which can be finely tuned to enhance the model's efficacy on a particular time series dataset. Given their proficiency in managing both linear and nonlinear decision boundaries, Support Vector Classifiers are highly effective for time series classification. This capability underpinned our decision to employ them in analyzing sensor data collected during our breastfeeding experiment. Their successful application in numerous sensor-based studies further validates this choice [27,28].

In the present study, the implementation of the SVM is based on the strategy employed by Ref. [29]. Three kernels were considered: Radial basis function (rbf), linear and polynomial kernels. The C parameter was also tuned until the best performance was achieved. The classification surface for the SVM classifier with rbf kernel is depicted in Fig. 4.

3.2 Gradient Boosting Classifier. Gradient Boosting (GB) is an ensemble learning technique that builds a predictive model in a stage-wise fashion, optimizing for errors made by the previous models in the ensemble. Specifically, the GB Classifier is designed for classification tasks. The primary idea behind GB involves creating a strong predictive model by combining the outputs of several weak models, typically decision trees. At each stage, a new weak model is trained to correct the errors made by the combined ensemble of existing models. The learning process focuses on minimizing a cost function, such as the logistic loss for classification problems. The loss function quantifies how well the model is performing.

Applying GB to time series data involves adapting the algorithm to handle the temporal nature of the data. Some GB libraries allow for handling time dependencies explicitly. GB can capture complex nonlinear relationships in time series data, making it suitable for a wide range of tasks. The iterative nature of GB allows it to be robust to outliers and noise in the data. It also automatically performs feature selection by assigning higher importance to features that contribute more to reducing the loss function. The ensemble nature of Gradient Boosting helps in reducing overfitting, which is particularly important in time series modeling.

We harnessed the power of the GB Classifier to adeptly manage nonlinear relationships and automatically select features, resulting in the development of accurate models that effectively capture complex temporal patterns within our data. Our model development involved the exploration of numerous hyperparameter combinations. Among these, the best model performance was attained with the following hyperparameters: Learning Rate: 0.01, Maximum Depth of Trees: 3, Minimum Samples for a Leaf: 1, Minimum Samples for a Split: 5, Number of Trees: 50.

3.3 Convolutional Neural Network. One-dimensional convolutional neural network (1D-CNNs), is a type of neural network architecture specifically designed for processing one-dimensional sequences of data. While traditional convolutional neural network (CNNs) are commonly used for image-related tasks, 1D-CNNs are well-suited for tasks involving sequential data, such as time series analysis and natural language processing.

The hierarchical feature learning of 1D CNNs allows them to capture patterns in both short and long sequences, making them versatile for different time series datasets. 1D-CNNs are also capable of learning from relatively small datasets, making them suitable for applications where labeled time series data is limited. The shared weights and hierarchical feature learning help in reducing overfitting, especially in scenarios with limited training data. 1D CNNs have proven to be effective and efficient for time series classification tasks. Their ability to automatically extract relevant features, capture local and global patterns, and handle varying sequence lengths makes them a valuable tool for a wide range of applications, including signal processing, health monitoring, and financial forecasting.

Convolutional neural network model architecture was designed using the Keras library in TensorFlow [30], an open source ML platform, to develop an appropriate 1D-CNN model. We applied the developed 1D-CNN model to our complex data, and obtained a very good result after several tuning of parameters. Our model is made up of many layers: three convolution layers that apply convolutional operations to input sequences, allowing the network to learn local patterns and features; three pooling layers for downsampling the output of the convolution layers in order to reduce the dimensionality of the learned features while retaining their essential information; ReLU (Rectified Linear Unit) [31] activation function was used to introduce nonlinearity to the model, enabling it to capture complex patterns in the data; Flatten layer for reshaping the output into a one-dimensional vector, preparing it for fully connected layers; Fully Connected Layers with a sigmoid [32] activation function for processing the flattened features, combining them to make classifications. We also employed ADAM optimizer [33] to train the neural networks because of its robustness and proven ability in neural network training. The loss function used by the model is "binary cross entropy." The purpose of this loss function is to penalize the model more when it confidently predicts the wrong class. The model's architecture is shown in Fig. 5(*a*).

3.4 Long Short-Term Memory. Long short-term memory (LSTM) is a type of recurrent neural networks architecture designed to address the challenges of learning long-range dependencies in sequential data. Traditional recurrent neural networks often struggle to capture and retain information over extended time intervals due to issues like vanishing gradients. LSTMs, introduced to mitigate these problems, feature memory cells and gating mechanisms that enable them to selectively retain or discard information, making them well-suited for processing sequential data.

Our decision to utilize LSTM networks stems from the time-sensitive nature of the sensor data. LSTMs are particularly adept at handling data from sensors due to several reasons: Sensor data frequently display temporal dependencies, meaning current readings are influenced by previous ones. This characteristic makes LSTMs an ideal choice for processing and analyzing such data. LSTMs, designed for sequential data, can capture and learn these dependencies effectively; Sensors may provide information with long-term dependencies, and LSTMs excel at handling such scenarios. The memory cells and gating mechanisms in LSTMs enable the model to selectively store and retrieve relevant information over extended periods; Sensors may produce data at irregular intervals, and LSTMs can naturally adapt to variable-length input sequences, making them suitable for real-world

Downloaded from http://asmedigitalcollection.asme.org/medicaldiagnostics/article-pdf/8/3/031003/7384220/jesmdt_008_03_031003.pdf by University Of Texas At Dallas user on 23 February 2026

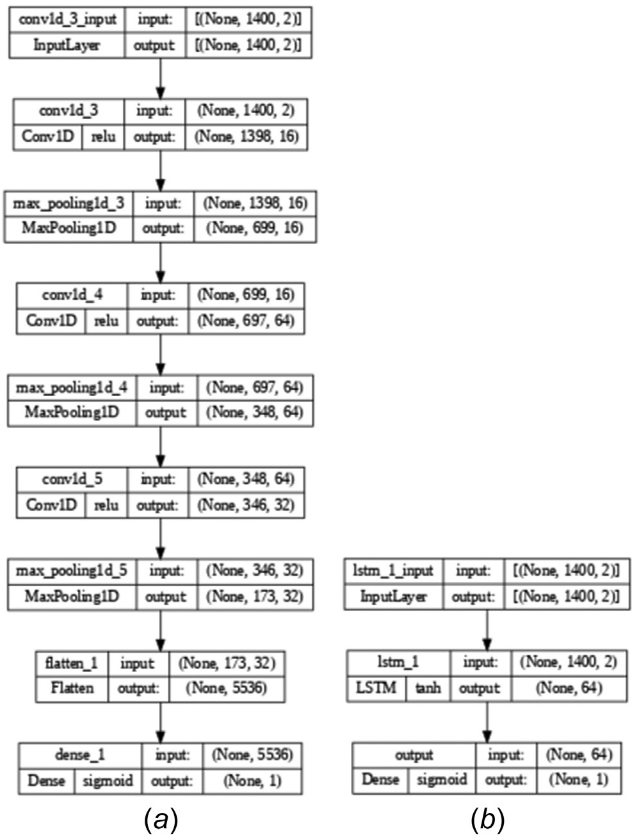

Fig. 5 Architectural diagrams of DL models used in the study: (*a*) 1D-CNN model architecture, detailing the sequence of convolutional and max-pooling layers leading to a dense output layer and (*b*) LSTM model architecture, which starts with an input layer followed by an LSTM layer and culminates in a dense layer with a sigmoid activation function

scenarios where data collection is not strictly periodic; LSTMs can automatically learn hierarchical features from sensor data [34], capturing both short-term fluctuations and long-term patterns. This relieves the need for manual feature engineering, a common requirement in traditional signal processing methods. Sensor data is often subject to noise and variations. LSTMs are robust to noisy inputs and can learn to filter out irrelevant information, enhancing their performance in noisy environments.

The first layer of our model is the input layer, which consists of force data on the subject's mother's nipple during breastfeeding. The second layer consists of LSTM node with hyperbolic tangent activation function. Inputs from the LSTM layer are fed into a fully connected dense layer with sigmoid activation that gives the probability of the event (healthy or unhealthy sucking). The architecture is as shown in Fig. 5(*b*). We also applied binary cross entropy as the loss function because of its proven effectiveness in binary classification.

3.5 Hyperparameters Optimization. Hyperparameter optimization is the process of selecting the best set of hyperparameters for a ML model. Hyperparameters are configuration settings that are not learned from the data but are set prior to the training process. Examples of hyperparameters include the learning rate, the number of hidden layers in a neural network, and the regularization term.

The goal of hyperparameter tuning is to find the hyperparameter values that result in the best model performance on a validation set or through cross-validation. The process typically involves searching through a predefined hyperparameter space, evaluating different combinations, and selecting the set that optimizes a chosen performance metric. For all our models, we defined an

Table 1 Five-fold cross-validation experiment results

	Classification rate (%)		
Model	Training	CV	Testing
SVC ("lin")	100.0	42.5	50.0
SVC ("rbf")	76.7	60.0	70.4
SVC ("poly")	66.7	60.0	60.0
GBC	100.0	75.6	80.0
LSTM	80.0	83.3	90.0
1D-CNN	90.0	90.0	90.0

hyperparameter space, selected the scikit-learn's "GridSearchCV" as a search method and, five-fold cross-validation (CV) as the validation scheme to achieve our optimization objective.

4 Results

Thirty subjects' data are used for training and cross-validation while ten subjects' data are used for testing. Summarized in Table 1 is the five-fold cross validation experiment results. Neural networks models outperformed classical MLCM models. In terms of accuracy, the CNN model with around 90% for training, cross-validation and testing, performed better than all other models. It was closely followed by the LSTM model with around 80%, 83%, and 90% for training, cross-validation and testing, respectively. The excellent performance of CNNs and LSTM models in finding complex relationships in our data may be attributed to their ability to automatically extract relevant features, capture local patterns and dependencies, and handle noise inherent in experimental data.

The best performing classical MLCM algorithm in this study in terms of accuracy is the GB Classifier. After hyper-parameters tuning, it achieved around 100%, 76%, and 80% accuracy for training, cross-validation and testing, respectively. The robustness of the model to handle noisy features and explore nonlinear relationships in time series data could be responsible for its better performance. It outperforms the best Support Vectors Classifier with the radial basis function (rbf) kernel which achieved around 70% accuracy.

4.1 Further Performance Evaluation. Accuracy is not always the best metric for evaluating the performance of a classification model. To this end, our analysis focused on three metrics: recall, precision, and the F1-score, as outlined in Table 2.

A confusion matrix is commonly used in MLCM to assess how well a model is performing in terms of classifying instances into different categories. It provides a tabular summary of the predictions made by a model compared to the actual ground truth. The confusion matrices for the models are shown in Fig. 6. The confusion matrices for each model reveal that the 1D CNN (see Fig. 6(*a*)) and LSTM models exhibited the highest accuracy, each correctly identifying 90% of the cases. These models demonstrated a strong ability to discern unhealthy patterns with minimal false negatives, indicating only one unhealthy infant misclassified as healthy. This high level of accuracy and low rate of false negatives suggest that both 1D CNN and LSTM are reliable for use in clinical settings where accurately detecting unhealthy conditions is crucial. In contrast, the SVC

Table 2 Results from performance evaluation

Model	Recall (%)	Precision (%)	F1-score
SVC ("poly")	20.0	100.0	0.52
SVC ("lin")	60.0	50.0	0.49
SVC ("rbf")	40.0	100.0	0.67
1D-CNN	80.0	100.0	0.90
GBC	60.0	100.0	0.79
LSTM	80.0	100.0	0.90

Downloaded from http://asmedigitalcollection.asme.org/medicaldiagnostics/article-pdf/8/3/031003/7384220/jesmdt_008_03_031003.pdf by University Of Texas At Dallas user on 23 February 2026

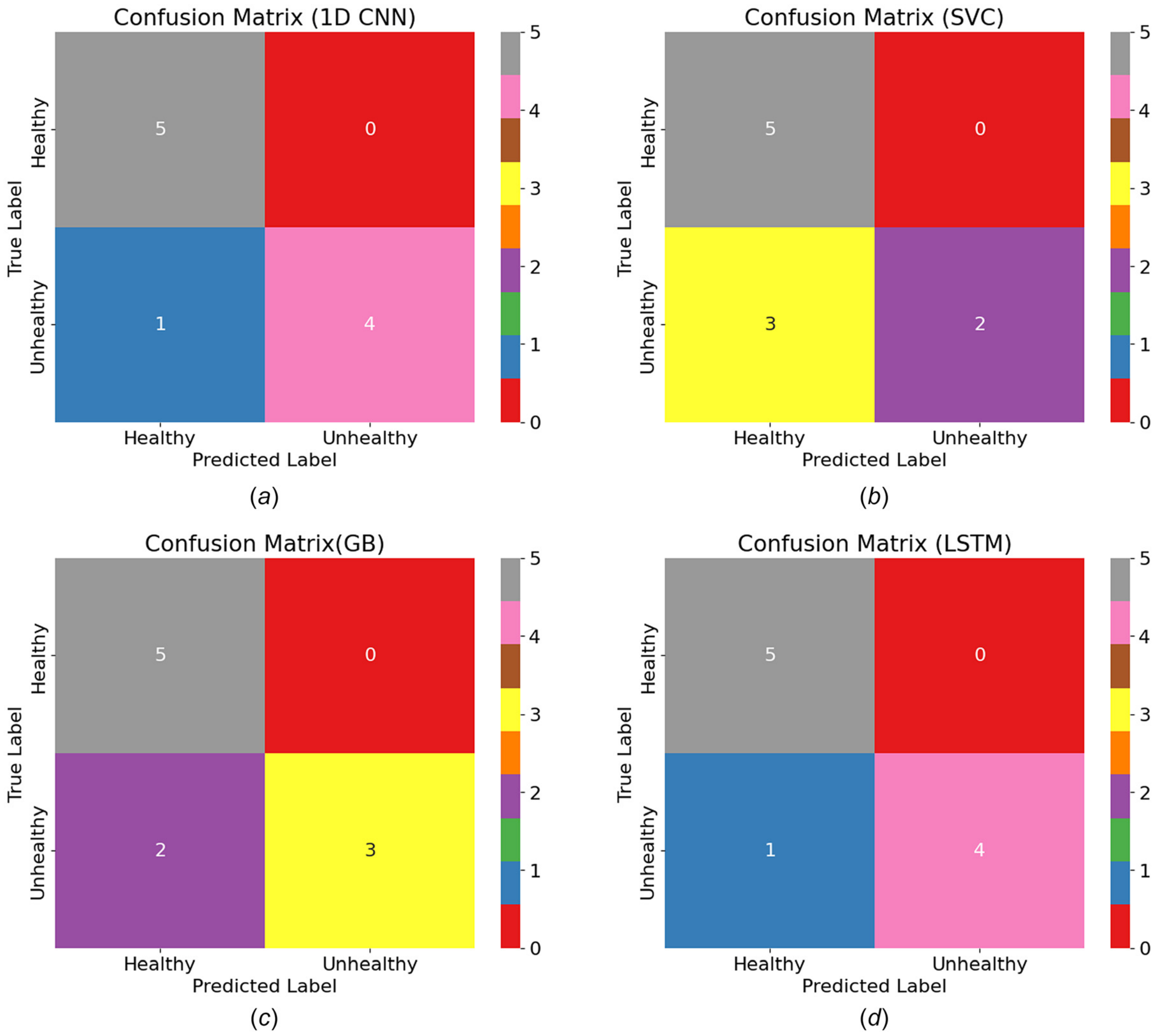

Fig. 6 Confusion matrices of models for classifying infant sucking patterns: (*a*) 1D-CNN model, (*b*) support vector classifier (SVC), (*c*) gradient boosting (GB) classifier, and (*d*) long short-term memory (LSTM) model. Each matrix displays the number of true positive, true negative, false positive, and false negative predictions.

(Fig. 6(*b*)) and GB Classifier (Fig. 6(*c*)) models showed lower accuracies of 70% and 80%, respectively, with a higher incidence of false negatives, which could pose risks of underdiagnosing unhealthy conditions.

The 1D CNN and LSTM (Fig. 6(*d*)) models outperformed others across other metrics, both achieving an impressive 80% recall and 100% precision, leading to the highest F1 scores of 0.90. This indicates their robustness in correctly identifying unhealthy sucking patterns while minimizing false positives. The GBC model also performed well, with a recall of 60% and precision at 100%, resulting in a commendable F1 score of 0.79. In contrast, the SVC models displayed varied results; the "rbf" kernel showed a moderate recall of 40% but a perfect precision, culminating in an F1 score of 0.67, while the "poly" kernel had a notably low recall of 20%, although with perfect precision, and the "linear" kernel demonstrated a balanced recall of 60% but lower precision of 50%, reflecting in their respective F1 scores of 0.52 and 0.49.

These results underscore the effectiveness of 1D CNN and LSTM in ensuring accurate diagnosis and reducing the likelihood of false diagnoses, which is crucial in clinical applications where early detection and accurate categorization of health conditions can significantly influence outcomes. The variability in performance among the SVC models suggests that while certain kernels may yield high precision, their practical application could be limited by lower recall, highlighting a tradeoff between identifying most unhealthy cases and maintaining accuracy in predictions.

We also used the receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) to evaluate the performance of our models. The AUC serves as a measure of the models' ability to distinguish between healthy and unhealthy sucking patterns, with a value closer to 1.0 indicating higher diagnostic accuracy. The ROC curve for the 1D CNN model shows an AUC of 0.80, indicating good model performance in distinguishing between the classes (see Fig. 7(*a*)). This curve, staying significantly above the diagonal line of no-discrimination, suggests that the 1D CNN is reliable in classifying the sucking patterns with a reasonable balance between sensitivity and specificity.

The SVC exhibits an AUC of 0.84, which is the highest among the four models. The ROC curve shows several steps, reflecting distinct thresholds where the sensitivity of the model increases at specific false positive rates (see Fig. 7(*b*)). This model demonstrates an excellent ability to differentiate between the two sucking patterns, possibly due to effective feature separation enabled by the kernel method used. Like the 1D CNN, the GB Classifier achieves an AUC of 0.80. Its ROC curve is a smooth curve approaching the top-left corner, indicating that it also maintains a good balance between detecting true positives and minimizing false positives, although it may not reach the sensitivity peaks of the SVC (see Fig. 7(*c*)). The

Downloaded from http://asmedigitalcollection.asme.org/medicaldiagnostics/article-pdf/8/3/031003/7384220/jesmdt_008_03_031003.pdf by University Of Texas At Dallas user on 23 February 2026

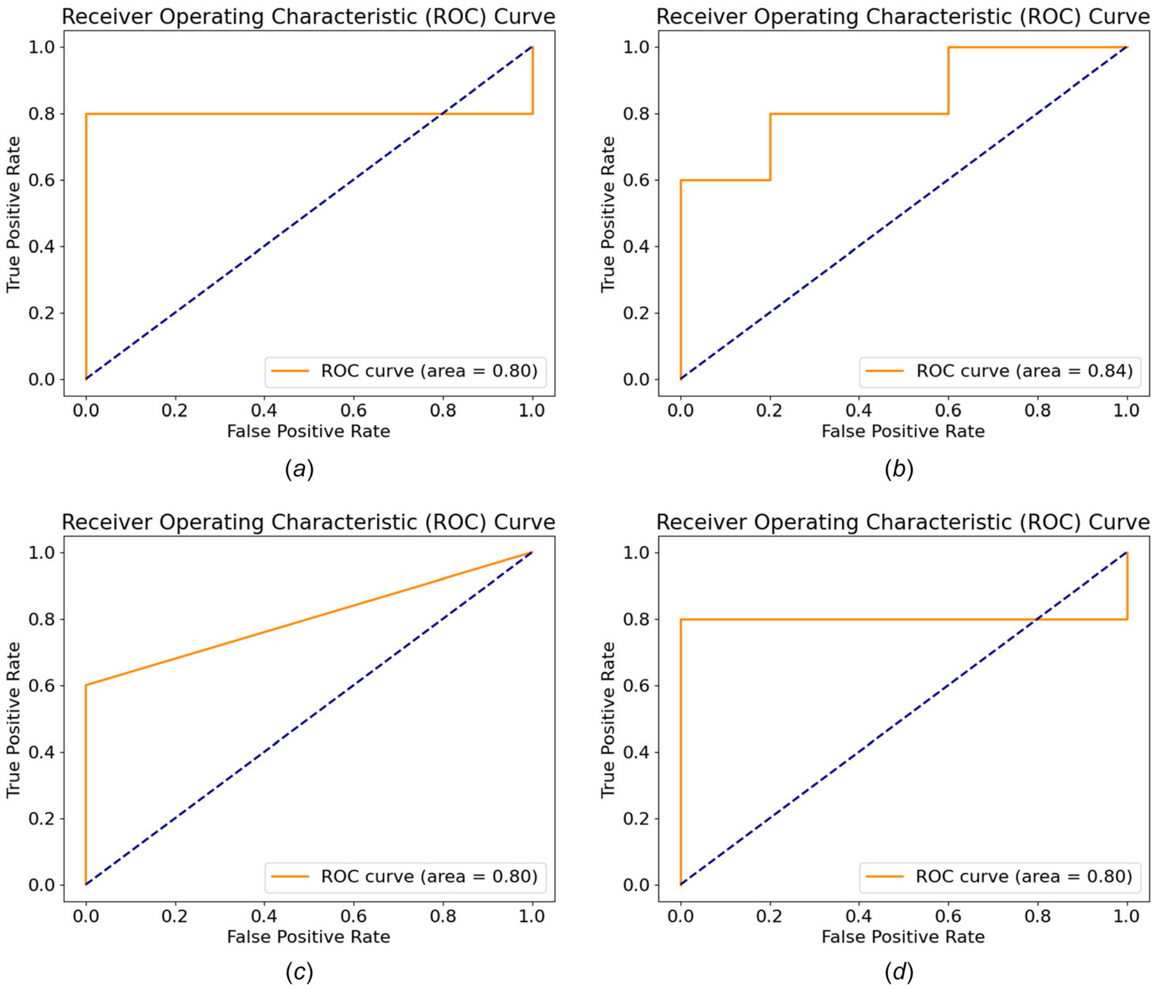

Fig. 7 ROC curves for different models. This figure displays the ROC curves for four models: (*a*) 1D-CNN, (*b*) SV classifier, (*c*) GB classifier, and (*d*) LSTM. Each curve plots the true positive rate against the false positive rate, with the area under the curve (AUC) indicated, demonstrating each model's effectiveness in distinguishing between healthy and unhealthy sucking patterns.

LSTM model's ROC curve also displays an AUC of 0.80, which suggests that it performs comparably to the 1D CNN and GB Classifier in terms of overall accuracy in distinguishing healthy from unhealthy patterns (see Fig. 7(*d*)). The smoothness of the curve suggests consistent performance across different thresholds.

The variance in performance between the training and testing phases for some models, particularly the SVC with a linear kernel, underscores the problem of overfitting, suggesting that while some models excel on training datasets, they falter on unseen data. In contrast, the GBC and SVC models with "rbf" and "poly" kernels show better, though not optimal, generalization capabilities. Future research could explore combining these models or integrating additional physiological data to further improve diagnostic accuracy and reliability.

The findings suggest that leveraging these models, healthcare providers could detect sucking issues earlier than traditional methods allow, potentially leading to better health outcomes for infants.

5 Conclusion

Our study introduces an advanced identification system for analyzing infant sucking patterns, utilizing artificial nipple-based sensors that measure the tongue forces exerted by infants. This technology generates crucial clinical data that is instrumental in understanding newborn feeding behaviors. We have rigorously evaluated MLCM models for classifying these sucking patterns, with a particular focus on the performance of 1D-CNN and LSTM networks. These models have shown excellent performance in terms of consistency during training, testing, and cross-validation phases, as well as superior ROC curve analysis.

The robustness of the 1D-CNN and LSTM models is evidenced by their high accuracy, recall, precision, and F1 scores. Such metrics underscore their reliability and potential for widespread clinical application. Furthermore, this study not only confirms the effectiveness of sophisticated computational modeling techniques in pediatric healthcare but also paves the way for their practical deployment. This represents a significant advancement in medical diagnostics and child welfare, promising to enhance the standard of care and intervention strategies in neonatal settings.

Acknowledgment

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Funding Data

- National Science Foundation (Grant Nos. 2121075 and 1454334; Funder ID: 10.13039/100000001).
- Japan Society for the Promotion of Science KAKENHI (Grant Nos. JP20K14756 and JP24K07510; Funder ID: 10.13039/501100001691).

Downloaded from http://asmedigitalcollection.asme.org/medicaldiagnostics/article-pdf/8/3/031003/7384220/jesmdt_008_03_031003.pdf by University Of Texas At Dallas user on 23 February 2026

Data Availability Statement

The algorithms and datasets used in this study are available here.

References

[1] Woolridge, M. W., 1986, "The 'Anatomy' of Infant Sucking," Midwifery, **2**(4), pp. 164–171.

[2] McClellan, H. L., Sakalidis, V. S., Hepworth, A. R., Hartmann, P. E., and Geddes, D. T., 2010, "Validation of Nipple Diameter and Tongue Movement Measurements With B-Mode Ultrasound During Breastfeeding," Ultrasound Med. Biol., **36**(11), pp. 1797–1807.

[3] Burton, P., Deng, J., McDonald, D., and Fewtrell, M. S., 2013, "Real-Time 3D Ultrasound Imaging of Infant Tongue Movements During Breast-Feeding," Early Hum. Dev., **89**(9), pp. 635–641.

[4] Geddes, D. T., Kent, J. C., Mitoulas, L. R., and Hartmann, P. E., 2008, "Tongue Movement and Intra-Oral Vacuum in Breastfeeding Infants," Early Hum. Dev., **84**(7), pp. 471–477.

[5] Stone, M., and Shawker, T. H., 1986, "An Ultrasound Examination of Tongue Movement During Swallowing," Dysphagia, **1**(2), pp. 78–83.

[6] Geddes, D. T., and Sakalidis, V. S., 2016, "Ultrasound Imaging of Breastfeeding–a Window to the Inside: Methodology, Normal Appearances, and Application," J. Hum. Lactation, **32**(2), pp. 340–349.

[7] Alatalo, D., Jiang, L., Geddes, D., and Hassanipour, F., 2020, "Nipple Deformation and Peripheral Pressure on the Areola During Breastfeeding," ASME J. Biomech. Eng., **142**(1), p. 011004.

[8] Jiang, L., Varnousfaderani, N. S., Hallac, R., Kane, A., Tadesse, Y., and Hassanipour, F., 2022, "Bio-Inspired Breastfeeding Simulator Integrated With Software Interface and Feedback Controls," IEEE Rob. Autom. Lett., **7**(2), pp. 1118–1125.

[9] Elad, D., Kozlovsky, P., Blum, O., Laine, A. F., Po, M. J., Botzer, E., Dollberg, S., Zelicovich, M., and Ben Sira, L., 2014, "Biomechanics of Milk Extraction During Breast-Feeding," Proc. Natl. Acad. Sci., **111**(14), pp. 5230–5235.

[10] Libbrecht, M. W., and Noble, W. S., 2015, "Machine Learning Applications in Genetics and Genomics," Nat. Rev. Genet., **16**(6), pp. 321–332.

[11] Erickson, B. J., Korfiatis, P., Akkus, Z., and Kline, T. L., 2017, "Machine Learning for Medical Imaging," Radiographics, **37**(2), pp. 505–515.

[12] Topol, E. J., 2019, "High-Performance Medicine: The Convergence of Human and Artificial Intelligence," Nat. Med., **25**(1), pp. 44–56.

[13] Truong, P., Walsh, E., Scott, V. P., Leff, M., Chen, A., and Friend, J., 2024, "Application of Statistical Analysis and Machine Learning to Identify Infants' Abnormal Suckling Behavior," IEEE J. Transl. Eng. Health Med., **12**, pp. 435–447.

[14] D'Mello, S. K., Tay, L., and Southwell, R., 2022, "Psychological Measurement in the Information Age: Machine-Learned Computational Models," Curr. Dir. Psychol. Sci., **31**(1), pp. 76–87.

[15] D'Mello, S. K., Southwell, R., and Gregg, J., 2020, "Machine-Learned Computational Models Can Enhance the Study of Text and Discourse: A Case Study Using Eye Tracking to Model Reading Comprehension," Discourse Processes, **57**(5–6), pp. 420–440.

[16] Nayyar, A., Gadhavi, L., and Zaman, N., 2021, "Machine Learning in Healthcare: Review, Opportunities and Challenges," Machine Learning and the Internet of Medical Things in Healthcare, pp. 23–45.

[17] Hoodbhoy, Z., Masroor Jeelani, S., Aziz, A., Habib, M. I., Iqbal, B., Akmal, W., Siddiqui, K., Hasan, B., Leeflang, M., and Das, J. K., 2021, "Machine Learning for Child and Adolescent Health: A Systematic Review," Pediatrics, **147**(1), p. e2020011833.

[18] Baker, S., and Kandasamy, Y., 2023, "Machine Learning for Understanding and Predicting Neurodevelopmental Outcomes in Premature Infants: A Systematic Review," Pediatr. Res., **93**(2), pp. 293–299.

[19] Sharifi-Heris, Z., Laitala, J., Airola, A., Rahmani, A. M., and Bender, M., 2022, "Machine Learning Approach for Preterm Birth Prediction Using Health Records: Systematic Review," JMIR Med. Inf., **10**(4), p. e33875.

[20] Li, Y., Mache, M. A., and Todd, T. A., 2020, "Automated Identification of Postural Control for Children With Autism Spectrum Disorder Using a Machine Learning Approach," J. Biomech., **113**, p. 110073.

[21] Oliver-Roig, A., Rico-Juan, J. R., Richart-Martínez, M., and Cabrero-García, J., 2022, "Predicting Exclusive Breastfeeding in Maternity Wards Using Machine Learning Techniques," Comput. Methods Prog. Biomed., **221**, p. 106837.

[22] Elgersma, K. M., Wolfson, J., Fulkerson, J. A., Georgieff, M. K., Looman, W. S., Spatz, D. L., Shah, K. M., Uzark, K., and McKechnie, A. C., 2023, "Predictors of Human Milk Feeding and Direct Breastfeeding for Infants With Single Ventricle Congenital Heart Disease: Machine Learning Analysis of the National Pediatric Cardiology Quality Improvement Collaborative Registry," J. Pediatr., **261**, p. 113562.

[23] Oyebode, O., Lomotey, R., and Orji, R., 2021, "I Tried to Breastfeed but…: Exploring Factors Influencing Breastfeeding Behaviours Based on Tweets Using Machine Learning and Thematic Analysis," IEEE Access, **9**, pp. 61074–61089.

[24] Nishi, E., Okuda, R., Hiraoka, K., and Rikoh, K., 2021, "Measurement of Force Applied by Infant Tongue to Three Types of Artificial Nipple: Investigation of Dynamic Actions in Relation to the Shape and Size of the Artificial Nipple," IEEJ Trans. Electr. Electron. Eng., **16**(8), pp. 1093–1098.

[25] Nishi, E., Nagamatsu, Y., and Niikawa, T., 2016, "Measurement of Force Applied by Infant Tongue to the Nipple During Sucking and Investigation of the Mechanism of Tongue Movement," 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, Aug. 16–20, pp. 2042–2045.

[26] Sriraman, N. K., 2017, "The Nuts and Bolts of Breastfeeding: Anatomy and Physiology of Lactation," Curr. Probl. Pediatr. Adolesc. Health Care, **47**(12), pp. 305–310.

[27] Begg, R., and Kamruzzaman, J., 2005, "A Machine Learning Approach for Automated Recognition of Movement Patterns Using Basic, Kinetic and Kinematic Gait Data," J. Biomech., **38**(3), pp. 401–408.

[28] Halilaj, E., Rajagopal, A., Fiterau, M., Hicks, J. L., Hastie, T. J., and Delp, S. L., 2018, "Machine Learning in Human Movement Biomechanics: Best Practices, Common Pitfalls, and New Opportunities," J. Biomech., **81**, pp. 1–11.

[29] Yoo, J.-H., Hwang, D., and Nixon, M. S., 2005, "Gender Classification in Human Gait Using Support Vector Machine," *Advanced Concepts for Intelligent Vision Systems*, J. Blanc-Talon, W. Philips, D. Popescu, and P. Scheunders, eds., Springer, Berlin, Heidelberg, pp. 138–145.

[30] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., et al., 2015, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," accessed Sept. 18, 2024, https://www.tensorflow.org/

[31] Glorot, X., Bordes, A., and Bengio, Y., 2011, "Deep Sparse Rectifier Neural Networks," Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings, Fort Lauderdale, FL, Apr. 11–13, pp. 315–323.

[32] Goodfellow, I., Bengio, Y., and Courville, A., 2016, *Deep Learning*, MIT Press, Cambridge, MA.

[33] Kingma, D. P., and Ba, J., 2017, "Adam: A Method for Stochastic Optimization," arXiv: 1412.6980.

[34] Hu, B., Dixon, P., Jacobs, J., Dennerlein, J., and Schiffman, J., 2018, "Machine Learning Algorithms Based on Signals From a Single Wearable Inertial Sensor Can Detect Surface-and Age-Related Differences in Walking," J. Biomech., **71**, pp. 37–42.

Downloaded from http://asmedigitalcollection.asme.org/medicaldiagnostics/article-pdf/8/3/031003/7384220/jesmdt_008_03_031003.pdf by University Of Texas At Dallas user on 23 February 2026