AI-Based Prototype For Training In Surgical Suturing Procedures

Oscar Eduardo Gualdrón G [1], Juan Jose Ardila V[2], Marisol Maestre D. [3]

[1] GISM Research Group, Faculty of Engineering and Architecture, Electronic Engineering, University of Pamplona, Km 1 vía Bucaramanga, 543050, Pamplona, Colombia.
[2] Agroinnova Research Group, Ernesto Schiefelbein Foundation, Bogotá, Colombia.
[3] Gramy Research Group, Faculty of Economic and Business Sciences, Economics program, University of Pamplona, Km 1 vía Bucaramanga, 543050, Pamplona, Colombia.

ABSTRACT
Surgical suturing is a fundamental procedure in clinical practice, used for tissue approximation and wound healing. Although its teaching is a routine part of medical education, it faces challenges such as the limited availability of controlled scenarios, the risks of practicing on real patients, and the subjective evaluation of skills. These limitations emphasize the need for technological tools that enable safe and objective training.
This article presents the development of an artificial intelligence-based prototype for surgical suturing training, integrating kinematic sensors for data acquisition and a supervised classification model using support vector machines (SVM) to distinguish between expert and novice users. The system provides a realistic and safe practice environment, enhancing the teaching-learning process in surgery, reducing clinical risks, and strengthening the objective evaluation of surgical skills.
KEYWORDS: Artificial intelligence, support vector machines, wearable device, surgical suture, surgical simulation, medical training, clinical skills assessment.

1. INTRODUCTION

Surgical suturing is one of the most relevant procedures in medical practice, as it enables tissue approximation and promotes proper wound healing. This technique is fundamental in general surgery and various specialties, since its correct execution reduces postoperative complications and fosters a faster patient recovery [1]. Nevertheless, the learning process of suturing presents multiple challenges, particularly in the context of training new healthcare professionals.
Traditionally, training has been carried out on real patients, which entails risks associated with clinical safety, in addition to a strong dependence on direct instructor supervision. Recent studies have shown that nearly 20% of medical residents do not meet the basic standards in surgical skills, which is related to the reduced availability of practical hours and the subjectivity in assessment processes [2], [3]. This scenario highlights the need to implement safer, more objective, and reproducible strategies in the teaching of suturing procedures.
The incorporation of simulation technologies has emerged as an effective alternative to strengthen learning in controlled environments, minimizing risks and promoting the

progressive acquisition of clinical skills [4]. Such tools allow for systematic repetition of procedures, immediate feedback, and the possibility of establishing objective performance metrics, thereby contributing to the training of both undergraduate students and surgical residents [5].
In recent years, artificial intelligence (AI) has gained relevance in the analysis of medical skills due to its ability to process large volumes of data and recognize patterns associated with human performance. Supervised classification algorithms, such as support vector machines (SVM), have been successfully applied to the identification of surgical gestures and the differentiation of expertise levels between experts and trainees [6], [7]. These techniques provide a solid framework for the implementation of training systems that reduce reliance on subjective evaluations.

In this context, the development of an AI-based prototype for training in surgical suturing, specifically focused on the simple interrupted stitch, is proposed. The system combines kinematic sensors for acceleration, tilt, and angular velocity with signal acquisition and analysis software, allowing for the objective characterization of the executed movements. Subsequently, through a classification model, the distinction between expert users and trainees is established, thus providing a safe, realistic practice environment with quantitative feedback [8].
The purpose of this research is to contribute to the improvement of the teaching-learning process in surgery, strengthening medical education and reducing the risks associated with practice on real patients. In this way, the prototype is presented as an innovative tool that integrates clinical simulation with intelligent data analysis, in line with current trends in medical education and bioengineering [9].

2. MATERIALS AND METHODS

The study is classified as quasi-experimental [10], with a comparative approach. The main objective was to evaluate the effectiveness of a system based on sensors and artificial intelligence (AI) algorithms to classify performance in the suturing procedure between two groups of participants: experts and novices.
For this purpose, data were collected using a wearable wristband equipped with acceleration, tilt, and angular velocity sensors, which were employed to capture time series [11] during the execution of the surgical suturing procedure.
The recorded signals were subsequently processed and used to train an artificial intelligence algorithm, with the aim of classifying the participants' performance according to their level of expertise. The analysis focused on comparing the classifier's results in distinguishing between both groups.

2.1 Prototype Design

The prototype developed for surgical suturing training was oriented toward the emulation of the simple interrupted stitch, considered the basic wound closure technique. The system is composed of:

- A suturing prosthesis, designed to reproduce the anatomical conditions of the skin and soft tissues.
- A signal acquisition module integrating kinematic sensors.
- Specialized software responsible for processing and analyzing the collected data.

The prosthesis was made of polymeric materials that simulate the resistance and elasticity of the skin, allowing for needle insertion and suture thread traction. To ensure stability during practice procedures, it was mounted on a rigid base.

2.1.1 Suturing Prosthesis.

For this study, prostheses provided by the Canadian company ProxSIMity Inc. were used. Figure 1 shows the model employed, designed with polymeric materials that reproduce the texture and resistance of soft tissues, allowing repeated insertion of the surgical needle and traction of the suture thread.

The prosthesis features an external surface that simulates the skin and a lower support that provides stability during the procedure, preventing displacements that could affect signal recording. This configuration ensures homogeneous training conditions and structural resistance against repeated use.

Furthermore, the modularity of the model facilitates its replacement after multiple trials, offering a reproducible, safe, and suitable environment for the systematic practice of the suturing technique..

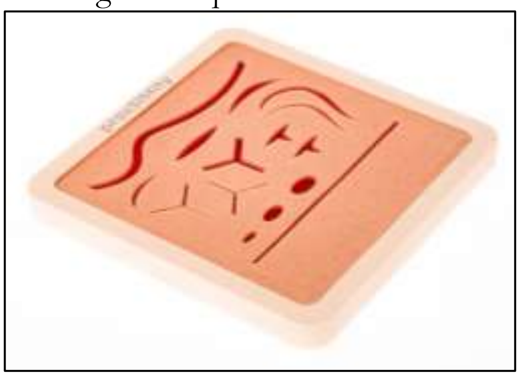

Figure 1. Suturing prosthesis. **Source:** ProxSIMity.

2.1.2 Description of the Wearable Device

The wearable device used in this research corresponds to a wristband equipped with acceleration, tilt, and angular velocity sensors, as shown in Figure 2. This device, supplied by the Canadian company ProxSIMity Inc., is placed on both wrists of the participants and enables real-time data collection during the execution of the surgical procedure.

Figure 2. Wearable device integrated with sensors. **Source:** ProxSIMity..

Through its sensors, the system records variations in tilt along the X, Y, and Z axes, as well as the acceleration and angular velocity of movements on these axes. This information is transmitted via Bluetooth technology [12] to a computer, where the data are stored in .csv format for further analysis.

For connection management and data acquisition, a specific software was developed in Unity [13]. The interface allows establishing communication with the wristband through Bluetooth, starting the signal collection, pausing the acquisition, and saving the records, as illustrated in Figure 3.

Figure 3. Connection interface functionalities. **Source:** Own elaboration.

2.2 Experimental Procedure

2.2.1 Data Acquisition and Labeling

Specific tests of the surgical suturing procedure were carried out with the aim of collecting comparative data between two groups of participants: experts and trainees. These data were subsequently used for training the artificial intelligence model.

The participants were organized as follows:

- Expert group: physicians, nurses, and faculty members from the medicine, nursing, and related programs at the University of Pamplona.
- Trainee group: students from the medicine, nursing, and related programs of the same institution.

In total, tests were conducted with 10 participants, evenly distributed: 5 experts (surgeons or instructors with suturing experience) and 5 novices (students in training). Each participant performed 5 complete procedures, resulting in a total of 50 samples: 25 from experts and 25 from novices. All tests were stored and labeled, ensuring an adequate class balance for the subsequent AI model training phase.

During data collection, the wristbands equipped with sensors were placed on the participants, and the prosthesis was prepared following a standardized protocol that included the following steps:

a. Material delivery: the participant received the needle holder and forceps required to perform the suture stitch.

b. Start of data capture: once ready, the participant pressed the button on the sensor wristband to activate signal acquisition.

c. Execution of the suture: the participant performed the simple interrupted stitch, finishing with knot tying and thread cutting.

d. End of data capture: upon completing the stitch, the button on the wristband was pressed again to stop data recording.

e. Result verification: it was confirmed that the stitch had been correctly executed according to the expected level of expertise.

f. Sample labeling: finally, the procedure was recorded in the acquisition software and labeled according to the corresponding category (expert or novice).

2.2.2 Exploratory Data Analysis.

The data collected during the tests were organized and processed into a structured dataset using analysis tools such as DataFrames in Python. This dataset included:

- Accelerations: The three components (X, Y, Z) recorded by the accelerometer, reflecting the linear movements of the needle holder during the suturing procedure.
- Angular velocities: The rotational values obtained from the gyroscope on the three axes, describing orientation and rotation changes of the surgical instrument during the procedure.
- Tilt angles: Represented the real-time inclination of the needle holder, allowing analysis of stability and movement control in each sample.

After loading the data, an integrity check was performed to ensure that no missing or inconsistent values were present. Subsequently, time series were plotted for each variable with the objective of visually exploring their behavior.
Figures 4, 5, and 6 present a sample of the signals captured by the prototype's sensors, showing the acceleration graphs on the X, Y, and Z axes, as well as the gyroscope and tilt signals on the same axes.
In each graph, the signals correspond to one of the repetitions performed by each participant during the test. The signals represented in blue correspond to individuals classified as Experts, while the signals in red correspond to individuals classified as Novices.

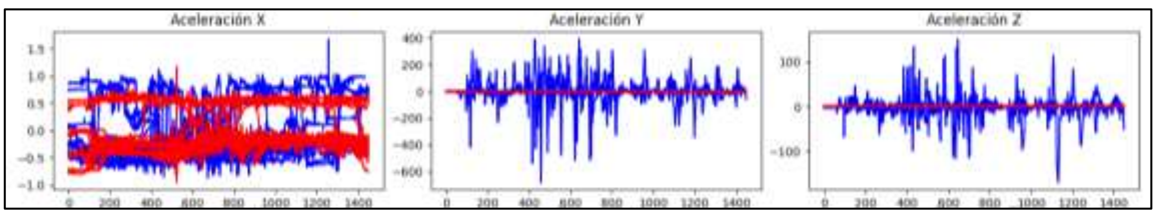

Figure 4. Sample of signals obtained, acceleration sensor. **Source:** Own elaboration.

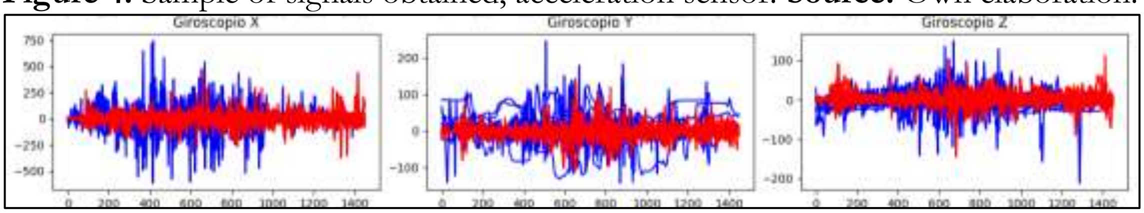

Figure 5. Sample of signals obtained, gyroscope. **Source:** Own elaboration.

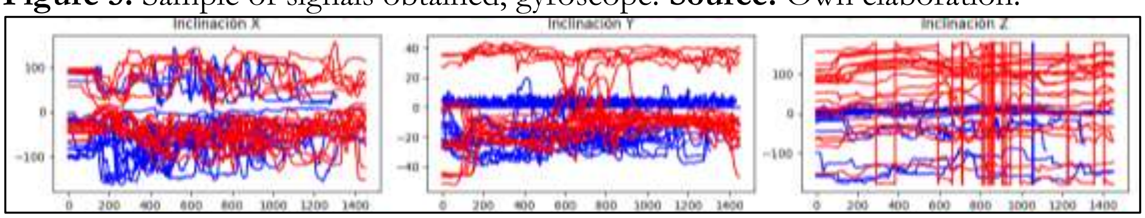

Figure 6. Sample of signals obtained, tilt sensor. **Source:** Own elaboration.

The graphical record of the time series shows high variability in the data, even for samples within the same category, making it highly complex to identify clear patterns or evident differences between experts and novices through visual inspection.
These characteristics observed in the time-series samples highlight the need to address these issues using more advanced processing and analysis techniques, such as feature extraction and the application of artificial intelligence models, to discern relevant information and build an effective classification model.
It also became evident that the collected samples were not of the same length, since the duration of each procedure varied depending on the skill, experience, and execution pace of each participant. This heterogeneity in signal length represents a major challenge for

the development of artificial intelligence models, as most machine learning algorithms require input data to have a uniform and consistent structure.

This is because AI algorithms—particularly those based on neural networks or supervised learning methods—typically operate under the assumption that input data have fixed dimensionality. When samples have different lengths, additional preprocessing is required to adapt them into a compatible format, which may introduce distortions or loss of relevant information.

Regarding variable-length time series, [14] points out that textbooks dedicated to multivariate time series classification problems rarely address the challenge posed by differences in series length. However, this issue is common in real-world applications. The study proposes three approaches to handle this situation: truncation, data padding, and missing value estimation.

This problem is also addressed in detail in [15], with a particular emphasis on multivariate time series. That study explores various data imputation methods, including deep learning techniques, applied to five healthcare datasets. The authors provide a comprehensive review of recent methods for imputing multivariate time series, focusing on those based on deep learning and published within the last five years.

Finally, this variability in signal duration can mask key patterns that distinguish experts from novices, since differences in execution pace and procedure length may not be directly related to the participant's skill, but rather to other factors such as caution, personal methodology, or even the inherent complexity of the task itself. Therefore, the unequal size of the samples complicates the technical processing of the data and may affect the model's ability to generalize and learn effectively.

2.2.3 Feature Extraction.

After obtaining the DataFrames, feature extraction [16] was performed on both statistical and signal-related parameters for each of the sensors, using the Python library NumPy [17].

This method is widely used in time series analysis [18], particularly in biomedical applications. In these cases, signals are often noisy and complex; therefore, feature extraction helps reduce noise and highlight relevant patterns [19]. This process transforms raw signals into a set of numerical descriptors that summarize their temporal, statistical, or frequency behavior, thereby facilitating the identification of meaningful patterns.

The problem detailed in the previous section regarding variability in signal length was addressed through feature extraction, which allowed for normalization of such differences.

When statistical and time-domain features are extracted, a more homogeneous representation of the data is obtained. This reduces the influence of individual variations and enables classification algorithms to focus on the relevant patterns of the signal. As a result, an improvement in the model's accuracy is expected.

In this case, the extracted features are listed in Table 1:

Tabla 1. Extracted features.

Name	Equation	Meaning or Application
Mean	$\mu = (1/n) \sum x_i$	Average of the values; represents the

		central tendency.
Standard deviation	$\sigma = \sqrt{[(1/n) \sum (x_i - \mu)^2]}$	Measures the spread of the data relative to the mean.
Mean absolute deviation (MAD)	$MAD = (1/n) \sum \lvert x_i - \mu \rvert$	Average absolute deviations from the mean.
Minimum value	min(x)	Lowest value in the signal.
Máximum value	max(x)	Highest value in the signal.
Range (max-min)	max(x) - min(x)	Total range of the signal.
Median	mediana(x)	Central value of the ordered signal; less sensitive to outliers.
Median absolute deviation (MAD_med)	$MAD\_med = (1/n) \sum \lvert x_i - mediana(x) \rvert$	Dispersion with respect to the median.
Interquartile range (IQR)	IQR = Q3 - Q1	Dispersion between the first and third quartile.
Count of negative values	$\sum (x_i < 0)$	Number of values below zero.
Count of positive values	$\sum (x_i > 0)$	Number of values above zero.
Count of values above the mean	$\sum (x_i > \mu)$	Frequency of values greater than the mean.
Number of peaks	peaks(x)	Number of local maxima; useful for analyzing repetitive patterns.
Skewness	skew(x)	Degree of symmetry in the

		signal distribution.
Kurtosis	kurt(x)	Concentration level of values relative to the mean.
Energy	$E = \sum x_i^2$	Measure of the signal's total power.

Source. Own work.

Each of these features was calculated individually for each sensor: acceleration on the X, Y, and Z axes; gyroscope on the X, Y, and Z axes; and tilt on the X, Y, and Z axes.

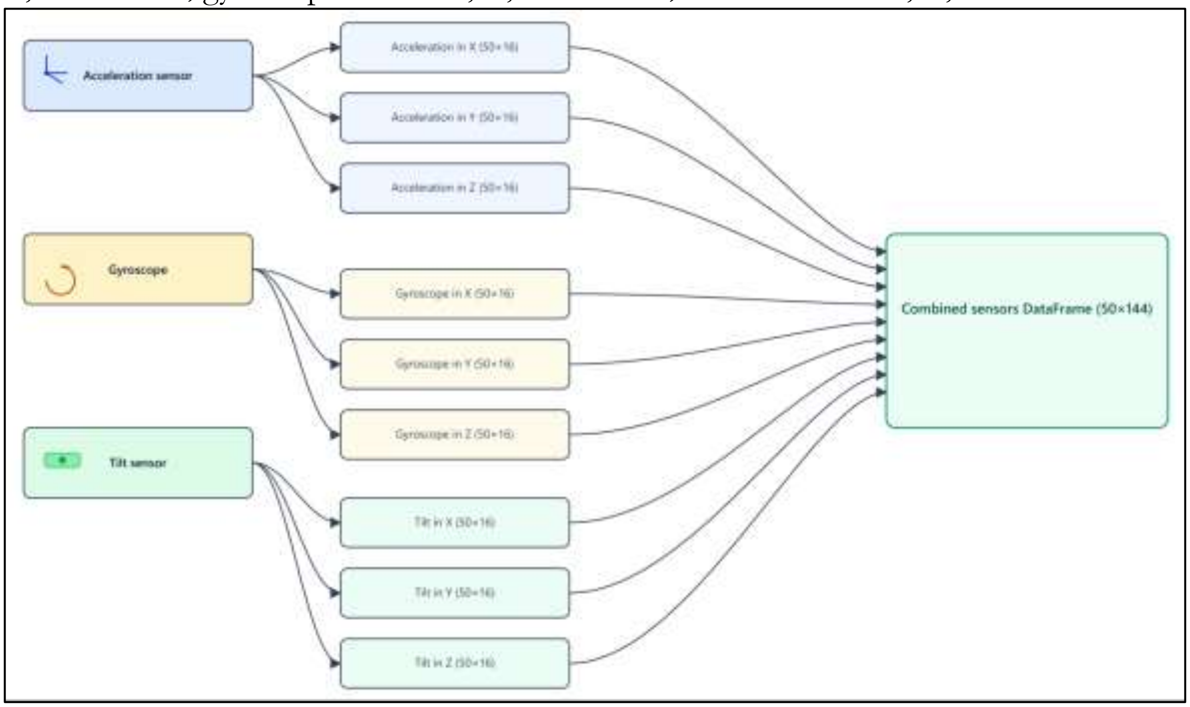

Figure 7. DataFrames resulting from the feature extraction process. **Source:** Own elaboration.

This process produced nine independent DataFrames, one for each sensor, as shown in Figure 7. Each DataFrame has dimensions of 50×16 (50 trials and 16 features). In addition, a consolidated DataFrame was generated, combining all sensors, with dimensions of 50×144.

2.2.4 Application of the Principal Component Analysis (PCA) Technique.

In order to reduce the dimensionality of the data while preserving as much information as possible, Principal Component Analysis (PCA) was applied to each of the DataFrames obtained after feature extraction. The use of PCA facilitates visualization, accelerates the training process of classification models, and helps mitigate issues associated with multicollinearity or redundancy among variables.

The effectiveness of PCA in dimensionality reduction and improvement of classification performance has been widely documented in the scientific literature. For example, in study [20], it was demonstrated that applying PCA to EEG signals for automatic sleep stage detection allowed the models to maintain or even improve classification accuracy while reducing computational load. Similarly, in research [21], the performance of linear dimensionality reduction techniques, including PCA, was evaluated in the classification of cardiac arrhythmias, showing that PCA contributed to improved sensitivity and F-score of the models used.

Through the application of this technique, the original set of 16 features per sensor was projected onto a new lower-dimensional space, composed of two principal components selected according to their ability to preserve the largest possible proportion of the total data variance. This procedure was implemented using the PCA function from Python's sklearn.decomposition module [22].

As a result, nine DataFrames were obtained, consistent with what was previously described, each with dimensios of 50×2 (50 samples and 2 principal components).

Table 2 presents the percentages of variance explained by the first two principal components, obtained through Principal Component Analysis (PCA), for each of the sensors considered. The values corresponding to the first and second components are reported, along with the total accumulated variance percentage explained by both. This information is essential for evaluating the effectiveness of the dimensionality reduction process, as it allows verification that the selected components retain most of the relevant information for the subsequent classification task.

Tabla 2. Percentage of variance explained by the first two principal components, according to the source sensor.

Sensor	Component 1 (%)	Component 2 (%)	Total Variance (%)
Acceleration X	65,64	33,23	98,87
Acceleration Y	74,93	21,61	96,54
Acceleration Z	73,69	24,71	98,4
Gyroscope X	91,62	7,64	99,26
Gyroscope Y	65,78	23,3	89,08
Gyroscope Z	59,62	36,42	96,04
Tilt X	87,32	10,92	98,24
Tilt Y	73,74	15,18	88,92
Tilt Z	98,72	0,92	99,64
Combined sensors	89,38	6	96,23

Source: Own elaboration.

As can be observed in Table 2, in all cases the first two components retained between 88% and 100% of the original variance, justifying their use as a reduced representation of the data without significant information loss.

In conclusion, as shown in Figure 8, a total of 20 DataFrames were generated: ten with the original features extracted from each sensor and ten with their respective reduced representations through PCA. These datasets were used as inputs for the classification models evaluated in this study.

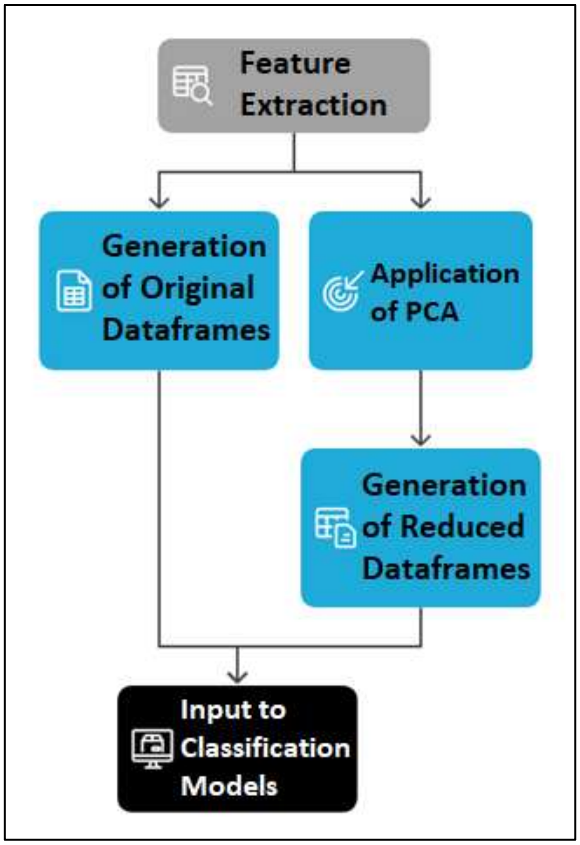

Figura 8. Consolidated DataFrames. **Source:** Own elaboration.

2.2.5 Artificial Intelligence Models.

The objective of the artificial intelligence model was to develop a system capable of classifying users according to their level of experience in performing the surgical suturing procedure, distinguishing between expert individuals and trainees (novices). For this purpose, two models based on supervised learning techniques were implemented.

From the data collected through sensors integrated into the wearable wristband devices, and after the corresponding preprocessing, structuring into DataFrames, and labeling of the samples, a Support Vector Machine (SVM) neural-type model was trained and evaluated.

Support Vector Machines (SVMs) are supervised machine learning algorithms widely used, among other fields, in biomedical and health Big Data analysis for classification and regression tasks. Their ability to handle high-dimensional data and detect complex patterns makes this technique a valuable tool in bioinformatics, clinical medicine, and health data analysis [23].

SVM models [24], [25] work by finding a hyperplane that separates different classes with the maximum margin in a feature space. For non-linear data, kernel functions are used to project the data into a higher-dimensional space where they can be linearly separated. This makes the algorithm effective for solving complex problems in areas such as bioinformatics and pattern recognition.

Figure 10 presents an illustrative example of the application of a Support Vector Machine (SVM) with a non-linear RBF kernel for data classification.

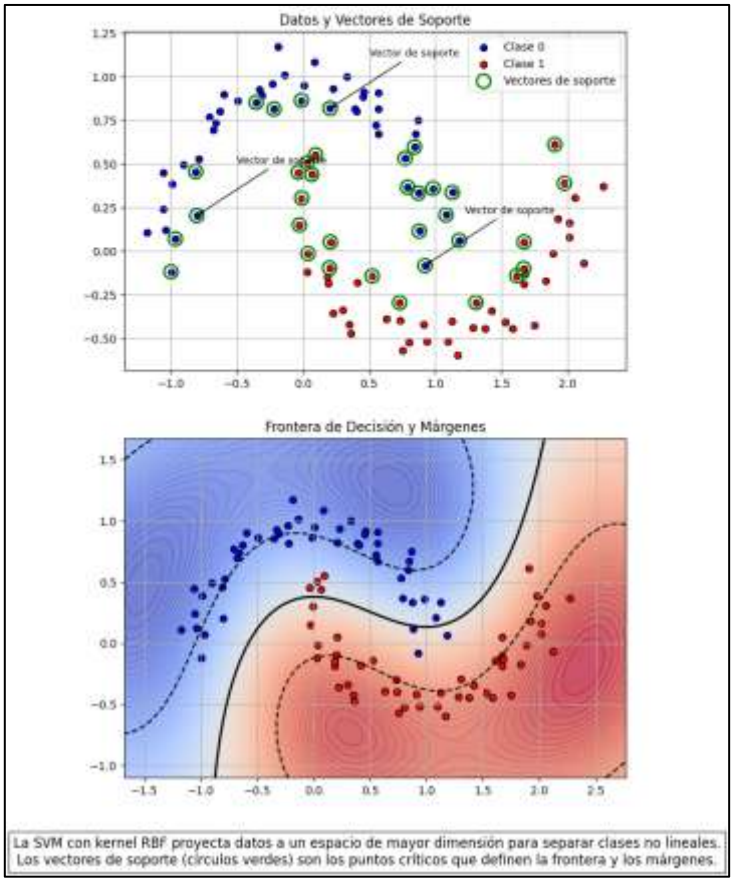

Figure 10. Illustration of the operation of a Support Vector Machine (SVM) with RBF kernel for non-linear classification. **Source:** Own elaboration.

In the first panel, the samples from two different classes are shown, with the support vectors highlighted by green circles. These correspond to the critical data points that determine the position of the decision boundary.

The second panel shows the optimal decision boundary (solid black line) obtained by the model, as well as the margins (dashed black lines) that maximize the separation between classes.

The colored regions represent the areas of the feature space assigned to each class by the classifier. This visualization demonstrates the ability of the SVM to transform the original space using kernel functions, enabling non-linear class separation and highlighting the fundamental role of support vectors in defining the optimal hyperplane.

2.2.5.1 SVM Model Specification.

The classification model implemented in this study corresponds to a Support Vector Machine (SVM), widely used in binary classification problems due to its ability to generate robust decision boundaries even in high-dimensional spaces [26].

The SVM was trained using a radial basis function (RBF) kernel, selected for its capacity to capture non-linear relationships among the features extracted from the sensors. The regularization parameter (C) and the kernel shape parameter (γ) were tuned through k-fold cross-validation ($k = 5$) on the training set, selecting the values that maximized classifier accuracy.

A stratified partition of the data was employed, assigning 70% for training and 30% for validation, while maintaining balance between the "expert" and "novice" classes. All feature vectors were normalized through standardization (z-score), computed from the statistical parameters of the training set. To ensure reproducibility of results, the training process was conducted with a fixed random seed.

The primary performance metric considered was accuracy. Complementary metrics such as precision, recall (sensitivity), F1-score, and confusion matrices on the validation set were also reported to provide a comprehensive evaluation of classifier performance.

The model was implemented in Python using the scikit-learn library, which allowed for standardized hyperparameter tuning, training, and evaluation.

2.2.6 Classifier Model Performance Evaluation.

As shown in Figure 11, the performance of the SVM model was evaluated considering both overall accuracy and complementary metrics that enable a more detailed analysis of the classifier's behavior.

A stratified partition of the data was applied (70% for training and 30% for validation), preserving the balance between the "expert" and "novice" classes. The validation set remained completely independent throughout the process, avoiding bias in hyperparameter selection.

The model was evaluated individually by sensor (acceleration, gyroscope, and tilt) and by each of their axes (X, Y, Z), as well as by combining the three axes (XYZ) for each sensor. This approach allowed identification of which modalities provide greater discriminative power in classifying the participants' level of expertise.

Summary of the Model Performance Evaluation Process

Feature	Support Vector Machine (SVM)
Data partition	70% training, 30% validation
Main metric	Accuracy
Sensor analysis	Exclude less informative sensors
Optimization	Reduce dimensionality, maintain accuracy

Figure 11. Summary of the model performance evaluation process. **Source:** Own elaboration.

3. RESULTS

3.1 Classifier Model Performance by Sensor

In order to evaluate the performance of the model in the surgical suturing emulation process, the accuracy rates obtained when training and validating the SVM classifier using the data from each sensor individually, as well as from their combinations, were analyzed. Two approaches were considered:

- The original features obtained after the feature extraction process.
- The features reduced through Principal Component Analysis (PCA), with the purpose of observing the effect of dimensionality reduction on the model's performance.

The results obtained are presented in Figure 12, Figure 13, and Figure 14, where the performances achieved in the training and test sets for each sensor and configuration are illustrated.

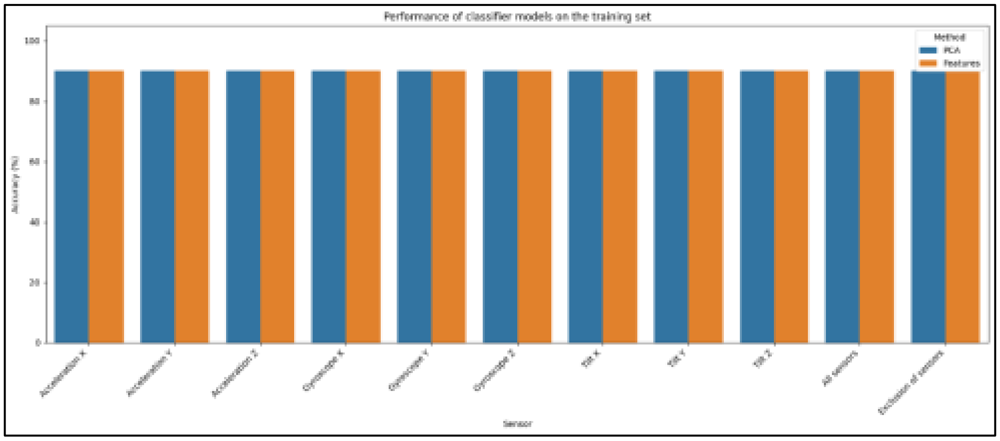

Figure 12. Performance of the classifier models on the training set. **Source:** Own elaboration.

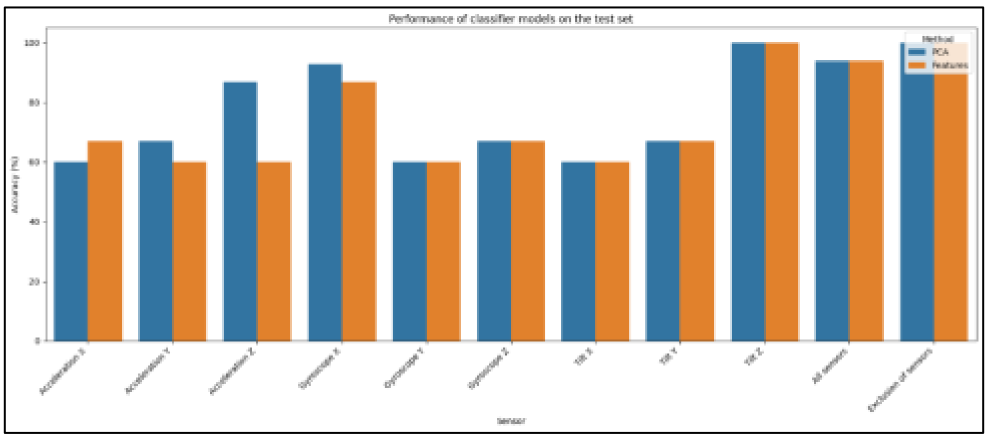

Figure 13. Performance of the classifier models on the test set. **Source:** Own elaboration.

3.2 Combination of Sensors and Selective Exclusion.

When using all sensors in combination, the SVM model achieved 90% accuracy on the training set with both the original features and PCA. In the test set, a precision of 94% was obtained in both configurations, confirming that the fusion of signals provides a more robust view of the procedure and allows the model to maintain high performance in classifying users' level of expertise.

Complementarily, the strategy of selectively excluding less informative sensors was evaluated. In this scenario, the classifier maintained practically identical behavior to that obtained with all sensors: 90% accuracy in training and 100% in testing, both when considering the original features and when applying PCA. This finding suggests that suppressing signals with lower contribution does not compromise the system's predictive capacity and may even simplify the model by reducing dimensionality without sacrificing performance.

In comparative terms, the most outstanding results were achieved with selective exclusion, reaching maximum accuracy in the testing phase. Nonetheless, both sensor combination and the exclusion of low-contribution sensors demonstrated that information integration or reduction should aim at maximizing model generalization without unnecessarily increasing complexity.

3.3 Heatmap of the SVM Model

To intuitively visualize the classifier's performance, heatmaps were created from the confusion matrices generated in the most representative scenarios of the SVM model.

These plots allow for a quick identification of the proportion of correct classifications and errors made by the classifier when distinguishing between expert and novice users.
Procedure. For each sensor configuration (individual, full combination, and selective exclusion), the corresponding confusion matrix was calculated on the validation set. Subsequently, the values were represented in a heatmap using color coding, where correct classifications are concentrated on the main diagonal and errors appear in the off-diagonal cells.

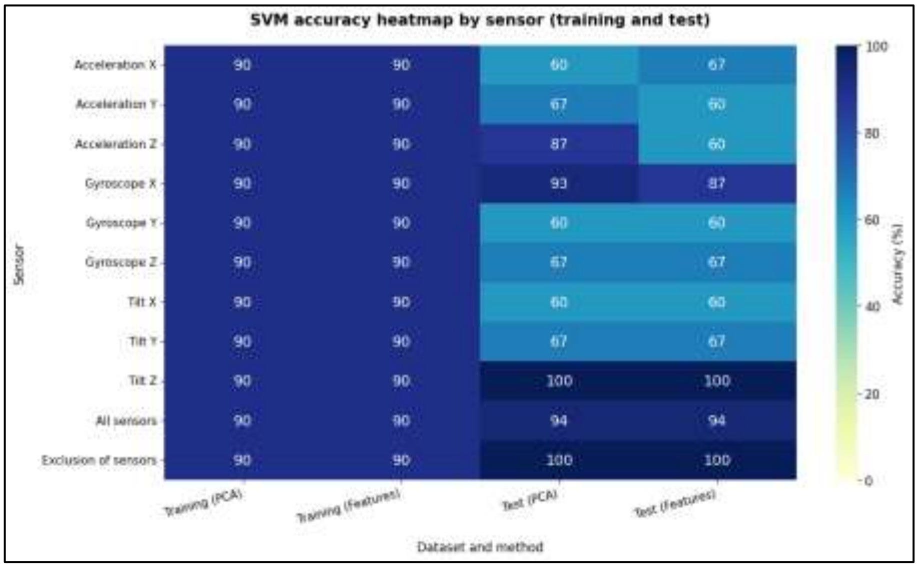

Figure 14. Heatmap of accuracy on the test set. **Source:** Own elaboration.

4. DISCUSSION

The results obtained in this study demonstrate that the SVM model is an effective tool for classifying the level of expertise in performing the suturing procedure. The sensor-by-sensor evaluation showed that, although each modality provides relevant information, there are notable differences in the discriminative capacity of the axes. In particular, gyroscopes and tilt sensors achieved high performance values, while some acceleration axes presented more modest results. This finding is consistent with the literature, where it has been reported that kinematic and postural parameters are often more representative for characterizing fine motor skills [27], [28].
The combination of signals proved to be a highly beneficial approach, as it allowed the integration of information from different modalities and reached 94% accuracy in testing, surpassing the performance obtained with individual sensors. Likewise, the strategy of selectively excluding less informative sensors maintained or even improved performance levels, achieving up to 100% accuracy in testing in some scenarios. This suggests that not all signals contribute significantly to the classification process. This finding highlights the importance of applying feature selection criteria to optimize the balance between accuracy and computational complexity.
The application of PCA enabled dimensionality reduction without sacrificing performance and even improved generalization in certain cases. This observation is key in contexts where the system needs to be implemented on portable devices or in clinical practice environments, since a reduced number of features decreases computational costs and facilitates model scalability.

When compared with other related studies in the field of clinical skills training [29], [30], the results obtained in this work are consistent: machine learning models applied to kinematic data from wearable sensors can differentiate between experts and novices with high accuracy. Moreover, the robustness demonstrated by SVM when integrating or excluding signals suggests that this approach could be extended to other surgical procedures requiring precision and manual coordination, broadening its applicability in medical training.

Overall, the findings reaffirm that the use of artificial intelligence, supported by inertial sensors and data processing techniques, is a promising strategy for strengthening the training of surgical skills in simulation environments. The system's ability to accurately discriminate the level of expertise provides an objective means of feedback that can complement traditional instructor-based evaluation.

5. CONCLUSIONS

The SVM model demonstrated high overall performance, achieving precision levels above 90% in sensor combination scenarios and up to 100% accuracy in testing when applying the strategy of selectively excluding less informative signals.

The combination of sensors showed that the fusion of modalities provides complementary information, enhancing the robustness of the classifier compared to the use of individual sensors. The exclusion of less relevant sensors allowed the model's performance to be maintained while reducing dimensionality, which favors computational efficiency and simplifies practical implementation.

The application of PCA contributed to improved generalization in certain scenarios, showing that reducing data redundancies does not compromise predictive capacity but can even enhance it. The results obtained confirm the feasibility of using wearable sensors and AI algorithms as support tools in clinical skills training, by providing an objective means of feedback on users' performance levels.

The proposed system constitutes an innovative and efficient alternative for training in suturing procedures, with potential to be extended to other contexts of medical education. The use of SVM, together with feature selection and dimensionality reduction strategies, provides a solid approach for discriminating between different levels of expertise, strengthening practical teaching through controlled simulation environments.

6. References

[1] C. C. Chu, "Types and properties of surgical sutures," in Biotextiles as Medical Implants, M. W. King, B. S. Gupta, and R. Guidoin, Eds., Woodhead Publishing Series in Textiles. Cambridge: Woodhead Publishing, 2013, pp. 231–273. doi: 10.1533/9780857095602.2.232.

[2] I. M. Q. Herrera, J. A. M. Paz, and G. E. R. Ortiz, "Análisis de las demandas de responsabilidad médica y su impacto en el sistema hospitalario surcolombiano: el caso de un hospital de tercer nivel," Revista CES Medicina, vol. 28, no. 1, pp. 45–56, 2014.

[3] P. Maziewski, K. Allen, A. Saha, M. G. Arnold, and L. G. Martino, "Impact of surgical training reduction on technical skill acquisition: a multi-center evaluation," Journal of Surgical Education, vol. 77, no. 5, pp. 1090–1097, 2020, doi: 10.1016/j.jsurg.2020.04.015.

[4] L. Hackney, S. O'Neill, M. O'Donnell, and R. Spence, "A scoping review of assessment methods of competence of general surgical trainees," The Surgeon, vol. 21, no. 1, pp. 60–69, 2023, doi: 10.1016/j.surge.2022.01.009.
[5] D. F. Pepley, A. B. Gordon, M. A. Yovanoff, K. A. Mirkin, S. R. Miller, D. C. Han, and J. Z. Moore, "Training surgical residents with a haptic robotic central venous catheterization simulator," Journal of Surgical Education, vol. 74, no. 6, pp. 1066–1073, 2017, doi: 10.1016/j.jsurg.2017.06.003.
[6] J. Bajwa, U. Munir, A. Nori, and B. Williams, "Artificial intelligence in healthcare: transforming the practice of medicine," Future Healthcare Journal, vol. 8, no. 2, pp. e188–e194, Jul. 2021, doi: 10.7861/fhj.2021-0095.
[7] R. Pedrett, P. Mascagni, G. Beldi, N. Padoy, and J. L. Lavanchy, "Technical skill assessment in minimally invasive surgery using artificial intelligence: a systematic review," medRxiv, 2022. doi: 10.1101/2022.11.08.22282058.
[8] K. Hutchinson, D. Yuh, and P. Vedula, "Challenges in surgical gesture recognition: the JIGSAWS experience," International Journal of Computer Assisted Radiology and Surgery, vol. 14, no. 9, pp. 1645–1655, 2019, doi: 10.1007/s11548-019-02020-w.
[9] A. Satava, "Historical review of surgical simulation—a persnal perspective," World Journal of Surgery, vol. 32, no. 2, pp. 141–148, 2008, doi: 10.1007/s00268-007-9374-y.
[10] Bärnighausen, T. et al., "Quasi-experimental study designs series—paper 4: uses and value," Journal of Clinical Epidemiology, vol. 89, pp. 21-29, 2017.
https://doi.org/10.1016/j.jclinepi.2017.03.012
[11] Hu, C., Sun, Z., Li, C., Zhang, Y., Xing, C., "Survey of Time Series Data Generation in IoT," Sensors, vol. 23, 6976, 2023. https://doi.org/10.3390/s23156976
[12] M. Verma, S. Singh, and B. Kaur, "An Overview of Bluetooth Technology and its Communication Applications," International Journal of Current Engineering and Technology, vol. 5, no. 3, pp. 1588, June 2015. Available:
http://inpressco.com/category/ijcet
[13] Unity Technologies, "Unity Real-Time Development Platform," Unity. [En línea]. Disponible: https://unity.com/. [Accedido: 2-feb-2025].
[14] A. Bier, A. Jastrzębska and P. Olszewski, "Variable-Length Multivariate Time Series Classification Using ROCKET: A Case Study of Incident Detection," IEEE Access, vol. 10, pp. 95701-95715, 2022, doi: 10.1109/ACCESS.2022.3203523.
[15] M. Kazijevs and M. D. Samad, "Deep imputation of missing values in time series health data: A review with benchmarking," Journal of Biomedical Informatics, vol. 144, p. 104440, 2023, ISSN 1532-0464, doi: 10.1016/j.jbi.2023.104440.
[16] S. Krishnan and Y. Athavale, "Trends in biomedical signal feature extraction," Biomedical Signal Processing and Control, vol. 43, pp. 41–63, 2018. doi: 10.1016/j.bspc.2018.02.008.
[17] C. R. Harris et al., "Array programming with NumPy," Nature, vol. 585, pp. 357–362, 2020. doi: 10.1038/s41586-020-2649-2.
[18] Wang, W., Chen, I., Hershkovich, L., Yang, J., Shetty, A., Singh, G., Jiang, Y., Kotla, A., Shang, J., Yerrabelli, R., Roghanizad, A., Shandhi, M., & Dunn, J., "A Systematic Review of Time Series Classification Techniques Used in Biomedical Applications," Sensors (Basel, Switzerland), vol. 22, 2022. https://doi.org/10.3390/s22208016.
[19] Rajoub, B., "Characterization of biomedical signals: Feature engineering and extraction," in Biomedical Signal Processing, Elsevier, 2020, pp. 29–50. https://doi.org/10.1016/b978-0-12-818946-7.00002-0.

[20] A.-M. Tăuțan, A. C. Rossi, R. de Francisco y B. Ionescu, "Dimensionality reduction for EEG-based sleep stage detection: comparison of autoencoders, principal component analysis and factor analysis," Biomed. Eng. / Biomed. Tech., vol. 65, no. 5, pp. 451–461, 2020, doi: 10.1515/bmt-2020-0139.
[21] R. Rajagopal y V. Ranganathan, "Critical Evaluation of Linear Dimensionality Reduction Techniques for Cardiac Arrhythmia Classification," Circuits and Systems, vol. 7, no. 9, pp. 225–234, 2016, doi: 10.4236/cs.2016.79225.
[22] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd ed., Sebastopol, CA, USA: O'Reilly Media, 2019.
[23] K. Taha, "Machine learning in biomedical and health big data: a comprehensive survey with empirical and experimental insights," J. Big Data, vol. 12, p. 61, 2025, doi: 10.1186/s40537-025-01108-7.
[24] T. Azizi, "Comparative Analysis of Statistical, Time–Frequency, and SVM Techniques for Change Detection in Nonlinear Biomedical Signals," Signals, vol. 5, no. 4, pp. 736-755, 2024, doi: 10.3390/signals5040041.
[25] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, y A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," Neurocomputing, vol. 408, pp. 189-215, 2020, doi: 10.1016/j.neucom.2019.10.118.
[26] Y. Bai, "ReLU-Function and Derived Function Review," SHS Web of Conferences, vol. 144, p. 02006, 2022. doi: 10.1051/shsconf/202214402006.
[27] H. A. Almurieb and E. S. Bhaya, "SoftMax Neural Best Approximation," IOP Conference Series: Materials Science and Engineering, vol. 871, no. 1, p. 012040, Jun. 2020, doi: 10.1088/1757-899X/871/1/012040.
[28] A. Mao, M. Mohri, y Y. Zhong, "Cross-Entropy Loss Functions: Theoretical Analysis and Applications," en Proceedings of the 40th International Conference on Machine Learning, PMLR vol. 202, pp. 23803–23828, 2023.
[29] O. Hospodarskyy, V. Martsenyuk, N. Kukharska, A. Hospodarskyy y S. Sverstiuk, "Understanding the Adam Optimization Algorithm in Machine Learning," en Proceedings of the 2nd International Workshop on Computer Information Technologies in Industry 4.0 (CITI'2024), Ternopil, Ukraine, 12–14 Jun. 2024, CEUR Workshop Proceedings, vol. 3742, paper 17, 2024.
[30] K. Thurnhofer-Hemsi, E. López-Rubio, M. A. Molina-Cabello y K. Najarian, "Radial basis function kernel optimization for Support Vector Machine classifiers," arXiv preprint arXiv:2007.08233, Jul. 2020.