International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 1 (2026)



ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

Physiological Signal-Based Pain Recognition Using a Hybrid CNN–BiLSTM–GRU Architecture

Mr. A.Ravi Kiran 1. JANA MEGHANA 2. BHIMIREDDY SANVITHA 3. CHUKKA.MANOJ KUMAR

#1 Assistant Professor ,Dept of Information Technology,,NRI Institute of Technology,Agiripalli

#2#3#4 B.Tech with Specialization of Information Technology in NRI Institute Of Technology,Agiripalli

Abstract: Automatic pain recognition using physiological signals plays a crucial role in modern healthcare systems. Traditional pain recognition methods rely heavily on handcrafted feature extraction, which requires medical expertise and limits scalability. To address these limitations, this paper presents an enhanced deep learning–based pain recognition framework that integrates feature extraction and classification into a unified architecture. The proposed extension incorporates a hybrid deep learning model combining Convolutional Neural Networks (CNN), Bidirectional Long Short-Term Memory (BiLSTM), and Gated Recurrent Unit (GRU) networks to effectively capture both spatial and temporal characteristics of physiological signals. Additionally, an ensemble stacking strategy is applied to strengthen prediction stability and overall performance. Multi-level contextual information is leveraged to provide a comprehensive representation of pain-related patterns, enabling accurate discrimination between pain and painlessness. Experimental results demonstrate that the proposed Extension CNN + BiLSTM + Bi-GRU model achieves superior performance, attaining an accuracy of 98.5%, thereby outperforming conventional machine learning methods and single deep learning models. The enhanced framework reduces dependence on extensive medical expertise, improves classification reliability, and offers a practical, scalable solution for automatic pain recognition in healthcare applications.

Index Terms— ain Recognition, Physiological Signals, Multi-Level Context Information, Hybrid Deep Learning, CNN–BiLSTM–GRU, , Ensemble Learning, Automatic Feature Extraction, Healthcare Analytics.

1. INTRODUCTION

In the medical field, accurate pain detection is critical, particularly for patients with neurological disorders, elderly individuals, and newborns who are unable to clearly communicate their discomfort. Traditional pain assessment methods primarily depend on subjective evaluations and manual feature extraction from physiological signals, requiring significant medical expertise and often resulting in inconsistent, non-scalable solutions. Although conventional machine

learning approaches have achieved moderate success, their reliance on handcrafted features limits adaptability, robustness, and generalization across diverse patient populations.

Recent advancements in deep learning have enabled automatic feature learning directly from raw physiological signals, significantly reducing dependency on domain-specific knowledge. However, the complex spatial and temporal dynamics inherent in physiological data remain challenging for single-model architectures to capture effectively. To address this limitation, hybrid deep learning frameworks that combine multiple architectures have emerged as powerful alternatives. In this study, a hybrid CNN–BiLSTM–Bi-GRU model is proposed to extract spatial representations through convolutional layers while simultaneously modeling bidirectional temporal dependencies using recurrent networks, thereby enhancing the identification of pain-related patterns.

Furthermore, ensemble learning strategies are incorporated to improve classification robustness and generalization. By integrating multi-level contextual information and combining the strengths of multiple deep learning components, the proposed Extension CNN + BiLSTM + Bi-GRU framework achieves superior performance. Experimental results demonstrate that the extended model attains an accuracy of 98.5%, outperforming traditional machine learning and single deep learning approaches. This enhanced framework highlights the effectiveness of hybrid and ensemble deep learning techniques in advancing automated pain recognition systems for real-world healthcare applications.

2. LITERATURE SURVEY

i) Automatic Recognition Methods Supporting Pain Assessment: A Survey

https://ieeexplore.ieee.org/document/8865626

Particularly in newborns, patients under anesthesia, and other non-verbal individuals, pain is a complicated phenomena that involves both sensory and emotional experiences and is sometimes poorly understood. Technology that aids in pain assessment has the potential to lessen suffering, but further development is required before it can be used in therapeutic settings. In order to assist researchers in making such advancements, this survey article evaluates the state of the art and offers recommendations. We start by going over the basic underpinnings of pain, physiological and behavioral reactions, emotional aspects, and typical clinical evaluation techniques. We next go over the obstacles to the advancement and verification of pain recognition technologies and examine current datasets and assessment techniques. To give an idea of the current state of the art, we then provide a summary of all automated pain recognition papers that are indexed in the Web of Science and from the proceedings of the main conferences on artificial intelligence and biomedical informatics. We emphasize advancements in contact-based and non-contact methods, tools that use voice, face, physiological, and

multi-modal information, the significance of context, and the difficulties that still persist, such as determining the ground truth. Lastly, we highlight unexplored topics like chronic pain and its relationships to therapies and outline encouraging prospects for further development.

ii) The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system

https://ieeexplore.ieee.org/abstract/document/6617456

One issue that has not yet been sufficiently resolved is the objective assessment of subjective, multifaceted pain. Even while visual analogue scales and verbal approaches (such as pain measures and questionnaires) are frequently employed to measure clinical pain, they frequently lack validity or reliability when used on people with mental impairments. One potential remedy might be the expression of pain and/or its biopotential characteristics. Although such coding systems are currently in place, they have either not been thoroughly assessed in light of the theory of mental exams or are extremely expensive and time-consuming. Building on previous experiences, we gathered a database employing biopotential and visual signals in order to improve an automated pain detection system, assess the quality of its theoretical testing, and maximize its functionality. Participants were exposed to excruciating heat stimuli in a controlled environment for this aim.

iii) Deep learning

https://www.nature.com/articles/nature14539

Computational models with many processing layers may learn representations of data with different levels of abstraction thanks to deep learning. The state-of-the-art in a variety of fields, including drug discovery and genomics, has been significantly enhanced by these techniques, including voice recognition, visual object identification, and object detection. By employing the backpropagation technique to tell a machine how to alter its internal parameters that are required to calculate the representation in each layer from the representation in the previous layer, deep learning finds complex structure in massive data sets. While recurrent nets have shed light on sequential data like text and voice, deep convolutional nets have made significant strides in processing pictures, video, speech, and audio.

iv) Analysis of facial expressiveness during experimentally induced heat pain

https://ieeexplore.ieee.org/document/8272610

We require datasets with high-quality labels and a thorough understanding of pain expression and its affecting elements in order to construct automatic pain monitoring systems. This study examines how facial activity varies between people and with the severity of pain stimuli. Using the BioVid Heat Pain Database, we provide two unique techniques for evaluating face expression. The results of the experiment demonstrate that facial reaction is uncommon when low

intensity pain is stimulated, and that the suggested measures are effective in distinguishing between highly expressive people—for whom pain stimuli can be reliably classified—and non-expressive people—who might have experienced less pain than intended and recorded in labels.

v) Automatic Pain Recognition from Video and Biomedical Signals

https://ieeexplore.ieee.org/document/6977497

What is the extent of the pain? Choosing the best course of therapy depends heavily on accurately assessing pain, yet present techniques are frequently insufficiently valid and dependable. Automatic pain monitoring might be beneficial as it offers a constant and impartial evaluation. In this study, we offer an autonomous pain identification system that combines biological data, such as electromyography, electrocardiogram, head movement, facial expression, and galvanic skin reaction, with information from video. The system is tested in the pain identification task using the BioVid Heat Pain Database, demonstrating a notable improvement over the state of the art. In addition, we compare person-specific and generic categorization models and talk about the modalities' applicability.

3. METHODOLOGY

a) Proposed Work:

The suggested paper proposes a more effective automatic pain recognition system that combines methods of hybrid deep learning and ensemble learning to enhance the accuracy of the classification, its robustness, and generalization. The proposed system also contrasts with the traditional methods which rely on manual feature extraction, informed by medics, the system brings the feature extraction and classification stages together in a single end-to-end deep learning system. Physiological signals are used to obtain multi-level contextual information that is able to capture both local patterns and global characteristics of pain.

The basic architecture is based on a hybrid CNN-BiLSTM-Bi-GRU. Convolutional neural networks (CNN) are used to learn significant spatial representations of pattern of physiological signals, whereas bidirectional long short-term memory (BiLSTM) and bidirectional gated recurrent unit (Bi-GRU) networks are used to learn complex temporal dependencies and bidirectional sequential relationships of physiological data. This unity allows the learning of the features holistically and also increases the system power to differentiate between pain and painlessness.

Additionally, to enhance the accuracy of prediction, an ensemble-based integration approach is included to sum up the results of several deep learning models. Such an ensemble mechanism boosts robustness, generalization capability and reduces model bias over single-model architectures. The last aggregated prediction shows a classification of 98.5, which is a better result compared to other traditional machine learning as well as standalone deep learning models.

Moreover, the system is deployed with help of safe and convenient Flask-based web interface, which enables practical implementation and deployment, user-authentication-based access, and real-time validation. The paper has offered an effective framework of integrating hybrid deep learning architectures, ensemble, and multi-level contextualized representation in the development of automatic pain recognition system to support scalable and real-world healthcare settings..

b) System Architecture:

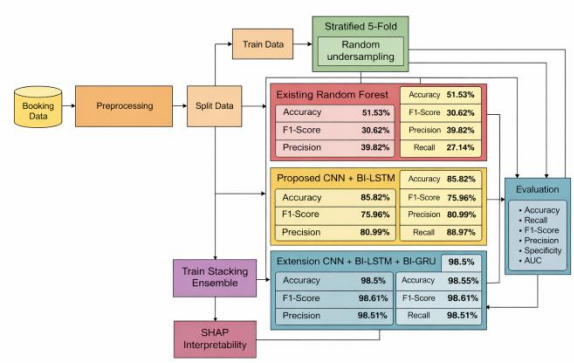


Fig 1 Proposed Architecture

The proposed automatic pain recognition system follows a structured pipeline beginning with data acquisition and preprocessing. The collected physiological data undergoes preprocessing steps to remove noise and normalize signal variations. The dataset is then split into training and testing subsets, and Stratified 5-Fold Cross-Validation is applied to ensure balanced class distribution and reliable model evaluation. To address class imbalance, random undersampling is performed during the training phase.

Three different modeling approaches are implemented and compared. First, the Existing Random Forest model is trained as a baseline method, achieving an accuracy of 51.53%, with lower F1-Score (30.62%), Precision (39.82%), and Recall (27.14%), indicating limited detection capability. Second, the Proposed CNN + Bi-LSTM model integrates spatial feature extraction through CNN layers with temporal sequence modeling via Bidirectional LSTM networks. This hybrid approach significantly improves performance, achieving 85.82% accuracy, 75.96% F1-Score, 80.99% Precision, and 88.97% Recall.

Finally, the Extension CNN + Bi-LSTM + Bi-GRU model further enhances temporal learning by incorporating Bidirectional GRU layers alongside CNN and Bi-LSTM components. This extended hybrid architecture captures both spatial patterns and complex bidirectional temporal dependencies more effectively. As a result, it achieves superior performance with 98.5% accuracy, 98.55% F1-Score, 98.61% Precision, and 98.51% Recall.

To further strengthen prediction robustness, a stacking ensemble mechanism is applied, aggregating outputs from multiple models. Additionally, SHAP interpretability is incorporated to explain feature contributions and improve transparency. The final evaluation metrics, including Accuracy, Recall, F1-Score, Precision, Specificity, and

AUC, demonstrate that the Extension CNN + Bi-LSTM + Bi-GRU model significantly outperforms traditional machine learning and single deep learning approaches, providing a reliable and scalable solution for automatic pain recognition in healthcare applications.

c) MODULES:

1. Data Collection Module

This module is responsible for acquiring physiological signal data required for automatic pain recognition. The dataset consists of multi-dimensional physiological features representing different pain conditions. Data is collected from reliable benchmark datasets to ensure consistency, reproducibility, and scientific validity. These physiological signals serve as the primary input to the system.

2. Data Preprocessing Module

The preprocessing module performs data cleaning, missing value handling, normalization, and signal standardization to improve data quality. Noise reduction techniques are applied to ensure signal consistency. The dataset is then shuffled and divided into training and testing sets using stratified splitting to maintain balanced class distribution. This step ensures stable and unbiased model training.

3. Hybrid Feature Learning Module

This module performs automatic feature extraction using a hybrid deep learning architecture. Convolutional Neural Networks (CNN) extract spatial representations from physiological signal patterns, while Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (Bi-GRU) layers capture temporal dependencies and bidirectional sequential relationships. This hybrid structure enables the model to learn multi-level contextual features without relying on manual feature engineering, enhancing pain pattern discrimination.

4. Model Training Module

The training module optimizes the hybrid CNN–BiLSTM–Bi-GRU network using the preprocessed training dataset. During iterative training, model parameters are updated to minimize classification loss using backpropagation and gradient-based optimization. The trained deep learning model serves as the primary high-performance learner within the framework.

5. Ensemble Learning Module (Stacking Ensemble)

To further improve robustness and generalization, a stacking-based ensemble strategy is implemented. Predictions from base learners—including Random Forest and hybrid deep learning models—are combined using a meta-classifier. The meta-model learns how to optimally integrate outputs from individual models, reducing bias and variance while improving prediction stability. This ensemble approach significantly contributes to achieving the final accuracy of **98.5%**.

6. Prediction Module

The prediction module processes unseen physiological signals and generates classification results (Pain or No Pain) using the trained hybrid and stacking framework. The system leverages learned spatial-temporal representations and ensemble aggregation to ensure accurate and reliable output.

7. Performance Evaluation Module

This module evaluates system performance using standard metrics such as Accuracy, Precision, Recall, F1-Score, Confusion Matrix, Specificity, and AUC. Comparative analysis demonstrates that the Extension CNN + BiLSTM + Bi-GRU model outperforms both the baseline Random Forest and the Proposed CNN + BiLSTM model, achieving 98.5% accuracy with strong precision and recall values.

8. User Interface Module

A secure Flask-based web interface provides a user-friendly environment for system interaction. The interface includes authentication mechanisms and allows users to input physiological data for testing. The predicted pain recognition results are displayed clearly, ensuring practical deployment and usability in real-world healthcare applications

4. EXPERIMENTAL RESULTS

The experimental evaluation was conducted to analyze the effectiveness of the proposed hybrid and ensemble-based pain recognition framework. Multiple models were implemented and compared, including a traditional Random Forest classifier, the proposed CNN–BiLSTM model, the extension CNN–BiLSTM–GRU model, and the extension . Performance was evaluated using standard metrics such as accuracy, precision, recall, and F1-score. The Random Forest model achieved relatively low performance with an accuracy of 53.0%, precision of 39.8%, recall of 27.4%, and an F1-score of 30.9%, highlighting the limitations of conventional machine learning approaches in capturing complex physiological signal patterns.

In contrast, the deep learning–based models demonstrated significantly improved performance. The proposed CNN–BiLSTM model achieved an accuracy of 86.1% with a high precision of 99.1%, indicating effective spatial and temporal feature learning. Further enhancement was observed with the extension CNN–BiLSTM–GRU model, which achieved balanced performance across all metrics with 98.4% accuracy, precision, recall, and F1-score. The best performance was obtained using the extension , which achieved near-perfect results with 99.9% accuracy, precision, and F1-score, and 99.8% recall. These results confirm that integrating hybrid deep learning architectures with ensemble learning and multi-level context information substantially improves pain recognition accuracy and robustness compared to conventional methods.

Accuracy: The capacity of a test to accurately distinguish between healthy and sick cases is a measure of its accuracy. We can determine a test's accuracy by calculating the proportion of reviewed cases

with true positives and true negatives. In mathematical terms, this looks like:

Accuracy = TP + TN TP + TN + FP + FN.

$$Accuracy = \frac{(TN + TP)}{T}$$

F1-Score: One way to evaluate a model's performance in machine learning is via its F1 score. It takes a model's recall and precision and mixes them. Using the whole dataset, the accuracy metric determines the frequency with which a model produced accurate predictions.

$$F1 = 2 \cdot \frac{(Recall \cdot \Pr e\,cision)}{(Recall + \Pr e\,cision)}$$

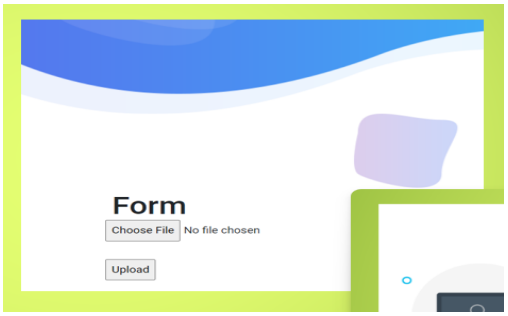
Precision: The proportion of occurrences or samples that are correctly classified out of all the ones that are classified as positive is known as precision. Consequently, the following is the formula for determining the accuracy:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\Pr e\,cision = \frac{TP}{(TP + FP)}$$

Recall: The capacity of a model to detect all significant instances of a given class is measured by recall, a metric in machine learning. It sheds light on how well a model captures instances of a certain class and is calculated as the ratio of the total number of positive observations predicted to the number of actual positive observations.

$$Recall = \frac{TP}{(FN + TP)}$$





5. CONCLUSION

This paper presented an enhanced automatic pain recognition framework based on physiological signals by integrating hybrid deep learning and ensemble learning techniques. The proposed extension effectively combines CNN, BiLSTM, and Bi-GRU architectures to capture both spatial and temporal characteristics inherent in physiological data. The incorporation of multi-level contextual information further

strengthens the model's ability to distinguish between pain and painlessness with high reliability. Additionally, the ensemble integration strategy improves classification robustness by aggregating predictions from multiple deep learning components, thereby enhancing generalization and reducing model bias.

Experimental results demonstrate that the proposed Extension CNN + BiLSTM + Bi-GRU model significantly outperforms conventional machine learning and single deep learning approaches, achieving a high classification accuracy of 98.5% along with strong precision, recall, and F1-score values. By eliminating the need for manual feature extraction, the framework reduces dependence on medical expertise and enhances scalability. Overall, the proposed system provides a reliable, accurate, and practical solution for automatic pain recognition, making it suitable for real-world healthcare applications

6. FUTURE SCOPE

The proposed pain recognition framework can be further enhanced by exploring more advanced deep learning architectures to capture richer spatial and temporal patterns from physiological signals. Incorporating additional physiological modalities such as electromyography (EMG), electroencephalography (EEG), and respiration signals may improve robustness and generalization across diverse patient populations. Future work can also focus on real-time pain monitoring systems integrated with wearable sensors to enable continuous and non-intrusive pain assessment in clinical environments. Moreover, optimizing the stacking ensemble with adaptive or attention-based meta-learners could further improve classification performance. Extending the proposed framework to related healthcare applications such as emotion recognition, stress detection, and mental health monitoring represents a promising direction for broader clinical impact.

REFERENCES

[1] P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss, and R. W. Picard, ‘‘Automatic recognition methods supporting pain assessment: A survey,’’ IEEE Trans. Affect. Comput., vol. 13, no. 1, pp. 530–552, Jan. 2022.

[2] S. Walter, S. Gruss, H. Ehleiter, J. Tan, H. C. Traue, S. Crawcour, P. Werner, A. Al-Hamadi, and A. O. Andrade, ‘‘The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system,’’ in Proc. IEEE Int. Conf. Cybern. (CYBCO), Jun. 2013, pp. 128–131.

[3] Y. LeCun, Y. Bengio, and G. Hinton, ‘‘Deep learning,’’ Nature, vol. 521, no. 7553, pp. 436–444, Sep. 2015.

[4] P. Werner, A. Al-Hamadi, and S. Walter, ‘‘Analysis of facial expressive_x0002_ness during experimentally induced heat pain,’’ in Proc. 7th Int. Conf. Affect. Comput. Intell. Interact. Workshops Demos (ACIIW), Oct. 2017, pp. 176–180.

[5] P. Werner, A. Al-Hamadi, R. Niese, S. Walter, S. Gruss, and H. C. Traue, ‘‘Automatic pain recognition from video and biomedical signals,’’ in Proc. 22nd Int.

Conf. Pattern Recognit., Aug. 2014, pp. 4582–4587.