DOI: 10.54392/irjmt2517

Asian Research Association

INTERNATIONAL RESEARCH JOURNAL OF
MULTIDISCIPLINARY TECHNOVATION

Acute Pain Recognition using an Ensemble Learning Methods: Evaluation of Performance and Comparison

Manisha S. Patil [a, *], Hitendra D. Patil [a]

[a] Department of Computer Science Engineering, Shri Shivaji Vidya Prasarak Sanstha's Bapusaheb Shivajirao Deore College of Engineering, Dhule (MS), India

* Corresponding Author Email: manishavpatil2007@gmail.com

DOI: https://doi.org/10.54392/irjmt2517

Received: 06-07-2024; Revised: 09-01-2025; Accepted: 15-01-2025; Published: 22-01-2025

Abstract: Accurate assessment and classification of acute pain are critical for optimal therapy, particularly in healthcare environments in which early intervention might prevent chronic pain development. Conventional pain recognition approaches mostly depend on the self-reported information, which can be subjective by psychological factors and communication problems, especially in nonverbal organizations. Recent advancements in technology have provided new opportunities for pain recognition using facial images and biomedical signals such as electromyography (EMG). In this work, we proposed an ensemble learning-based model that combines both face images and EMG data for acute pain classification, and the CNN ShuffleNet V2 approach is used for feature extraction. Our objective for pain classification is to correct classification for pain intensity levels from T0 to T4 (no pain vs. pain). We proposed ensemble learning-based techniques like TabNet, LightGBM, Hidden Markov, and Gaussian Process for acute pain classification. We used many kinds of approaches to improve prediction performance, which created a comprehensive framework for pain classification and insights into the physiological and psychological responses to acute pain. Our analysis of results also indicates that the ensemble approach definitely surpasses previous approaches whereby TabNet model accuracy came to be 97.8%. Also, this model has great F1 score of 97.6%, as well as recall at 97.3%, while on kappa score, it goes up to 92.4%, indicating great dependability. These results present a good optimism that our ensemble learning technique could change the face of pain assessment procedures and therefore patient care in acute pain treatment.

Keywords: Pain, Feature, Classification, Facial, Physiological Signal, EMG

1. Introduction

Acute pain is a clinically complex condition that may severely affect an individual's function and quality of life [1]. Specifically in healthcare, where immediate treatments can potentially prevent chronic pain, accurate identification and characterization of acute pain are essential for effective management and therapy [2]. Traditional methods of pain assessment usually involve self-reported measures, which are not objective and depend on many factors, including the mental state of a person and communication problems, especially in nonverbal societies [3]. Advances in technology have led to new ways of assessing pain based on face imaging and physiological information [4-6]. Thereby, facial expressions may have proved strong indicators of emotional states-in this case, pain, while statistics allow for the analysis of it via machine learning approaches [7]. Additional data on muscle activity pertinent to pain responses can also be taken from physiological signals like EMG [8, 9].

The present research illustrates the potential of our proposed approach to transform pain evaluation processes in health care institutions by offering an extensive performance evaluation and comparison of it versus current methods. The final objective of this research is to further enhance care for patients and acute pain treatment results by integrating the increasing amount of information on automatic pain recognition systems. The following sections discuss relevant studies and present our results of our performance study, and compare how different ensemble learning methods classify acute pain.

The method for pain recognition, which uses facial expression images and ensemble learning, was proposed by [4]. This method uses ShuffleNet V2 for feature extraction and class activation map techniques. Base methods include CatBoost and XGBoost, with a support vector machine model as a meta-model. Another method is the three-stream hybrid deep neural network (EDLM), which was proposed by [5] and which extracts the features from the facial images with pain classification. The VGGFace and principal component analysis are used in this EDLM model to extract features from the Multimodal Intensity Pain database. The models have also outperformed the competitors in the multi-level pain detection database through achieving better accuracy of the feature classification.

2. Related work

In such a study that utilized ensemble approaches to classify pain intensity, a trend was shown for how ensemble learning methods can be improved further to enhance classification evaluation, by integrating the most effective features from several iterations. A series of investigations have also looked at the integration of facial and physiological data, pointing toward a need for multimodal approaches to accurately assess pain intensity. In the study by [10], for example, EDA data has been used to classify the level of pain using a hybrid model that combines deep learning and traditional machine learning approaches. This is in line with the objectives of our study, which aims to enhance the accuracy of acute pain detection using the combination of both face images and EMG signals.

This section seems at significant research that has improved techniques to evaluate pain, with a focus on using physiological signals such as electromyography (EMG) and facial expressions [11]. Much study has been given in recent years to the classification of acute pain using facial images and physiological signals, especially since the advent of advanced machine learning techniques. Based on relevant research that has informed our understanding and the enhancement of pain assessment methods, the following section discusses the employment of physiological signals, namely, electromyography (EMG), and facial expressions. Indeed, work by [12], for example, demonstrated effectively how deep feature extraction in images of the face can be applied for automatic pain classification, meaning that such models can aid doctors in classifying pain indirectly. According to their findings, the integration of multiple physiological signals may improve the efficiency of pain classification techniques, especially where patients may not be able to communicate their level of pain appropriately. Most studies done in this area have relied greatly on the Biovid Part A dataset, which has allowed for an enormous amount of data to be used to train pain classification models. The dataset is appropriate for the development of machine learning algorithms for pain assessment because it includes different physiological signals along with the corresponding pain levels [13]. In addition to facial emotions, physiological signals, particularly electromyography, or EMG have been used to provide more information for the classification of pain. A number of studies has shown that the integration of EMG signals with facial expressions could improve the accuracy of pain recognition [14].

As indicated by the study in which pain intensity classifications through ensemble methods were used, the advancements in the ensemble learning method indicate that there may be an opportunity for improvement of assessment classification by combining various models. Again, further study showed a multimodal combination between physiological data and facial data to enable efficient pain evaluation. The study of [10] utilized a hybrid model that includes deep learning and conventional approaches to machine learning in an attempt to classify the intensity of pain using EDA signals. That is compatible with our research direction of improving acute pain detection by using face images and EMG signals.

Our research expands the mentioned foundations by using the Biovid Part A dataset for very detailed performance evaluation and comparison with an ensemble learning-based approach in order to predict the classification of acute pain. Lots of studies that examined the combination of facial and physiological data point at the significance of multimodal approaches toward precise evaluation of the condition called pain. For example, the study applied deep learning and traditional approaches of machine learning to create a hybrid model that quantifies the degree of pain based on electrodermal activity (EDA) biomarkers [15]. That proves the technique; classification performance of pain may be improved by combining multiple biomedical indicators with facial emotions [16].

In conclusion, the present state of research underlines the importance of combining physiological signals and facial images to recognize pain, with methods of ensemble learning providing an opportunity for improved efficacy. In order to improve on these frameworks, our research introduces an ensemble learning-based approach to predict acute pain classification using the Biovid Part A dataset for extensive performance analysis and comparison

3. Methodology

3.1 Data Sets

The most significant dataset used in the study is the BioVid Heat Pain Database, Part A. The supervisor of the BioVid investigation squad, Philipp Werner, allowed the access to the data. The study involved four distinct stages of independently validated thermal pain stimulation, labeled as T1, T2, T3, and T4. Pain levels were categorized into five classes, with 20 participants representing every class, resultant in an overall of 8,600 samples recorded for each signal. Each pain intensity measurement spanned a duration of 5.5 seconds [17].

3.2 Feature Extraction

In this work, we construct a convolutional network based framework using five CNN layers and batch normalization [18-21]. We utilize the CNN ShuffleNet V2 algorithm's parameters for learning the structure. The Figure 1 illustrates the systematic feature extraction method. ShuffleNet's concentration on variable minimization and computational efficiency sets it apart from localized patterns. Bunch convolution connects every set of input streams simultaneously [22]. The following minimizes the total amount of elements and streamlines the estimation procedure. To add or aggregate a map of feature elements, we implemented element-wise total or aggregate feature fusion of face images and EMG biomedical signal [23]. The inexpensive approach integrates information obtained from multiple perspectives. To demonstrate the significant movements, we employ a class activation map. Following that, the extraction of features model incorporates a fully linked layer and a reshape function to generate a two-dimensional array [24].

In this investigation, we use an incremental generalization approach that integrates estimates from numerous base models to improve the accuracy of predictions. We use a variety of classifiers as a meta-model and produce results that incorporate predictions from the base models. Train the CatBoost and XGBoost algorithms to generate a prediction. The predicted results are utilized as an attribute for training the various classifiers [25, 26]. The various classifiers train by identifying the significant variations in the features and separating them into different categories. The proposed pain intensity level recognition is provided in Figure 2.

CatBoost, compared to existing gradient boosting algorithms, does not require prior treatment, encoding, or one-hot encoding for working with data that is categorical. Therefore, categorical information might be analyzed using minimum computing power. CatBoost offers excellent processing speed with parallelism, which makes it perfect for large datasets and elegant algorithms. The method employs sequential enhancing as well as stochastic variations to reduce overfitting and promote applicability. CatBoost tackles missing data without relying on approximation [27]. Throughout training, heedless trees manage the values that are absent. CatBoost trains quicker and more efficiently than other gradient-boosting structures. We utilize the Hyperband optimization approach to improve the CatBoost algorithm's effectiveness in recognizing pain. XGBoost's objective functionality includes regularization approaches such as Least Absolute Shrinkage and Selection Operator (LASSO) [28]. It eliminates overfitting as well as improves model adaptability. XGBoost's development allows for efficient analysis of face images and EMG signals across several processors [29].

In XGBoost, builders can create unique objectives and functions for explicit difficulties. This capability allows researchers to utilize hyperband optimization to increase the XGBoost model's efficiency. To handle missing data during the training stage, XGBoost was created. It can automatically manage missing values based on data. We apply regularization during tree construction to manage complexity, improving the model's applicability [30].The integrated cross-validation in XGBoost reduces overfitting and helps monitor model performance during training [31].

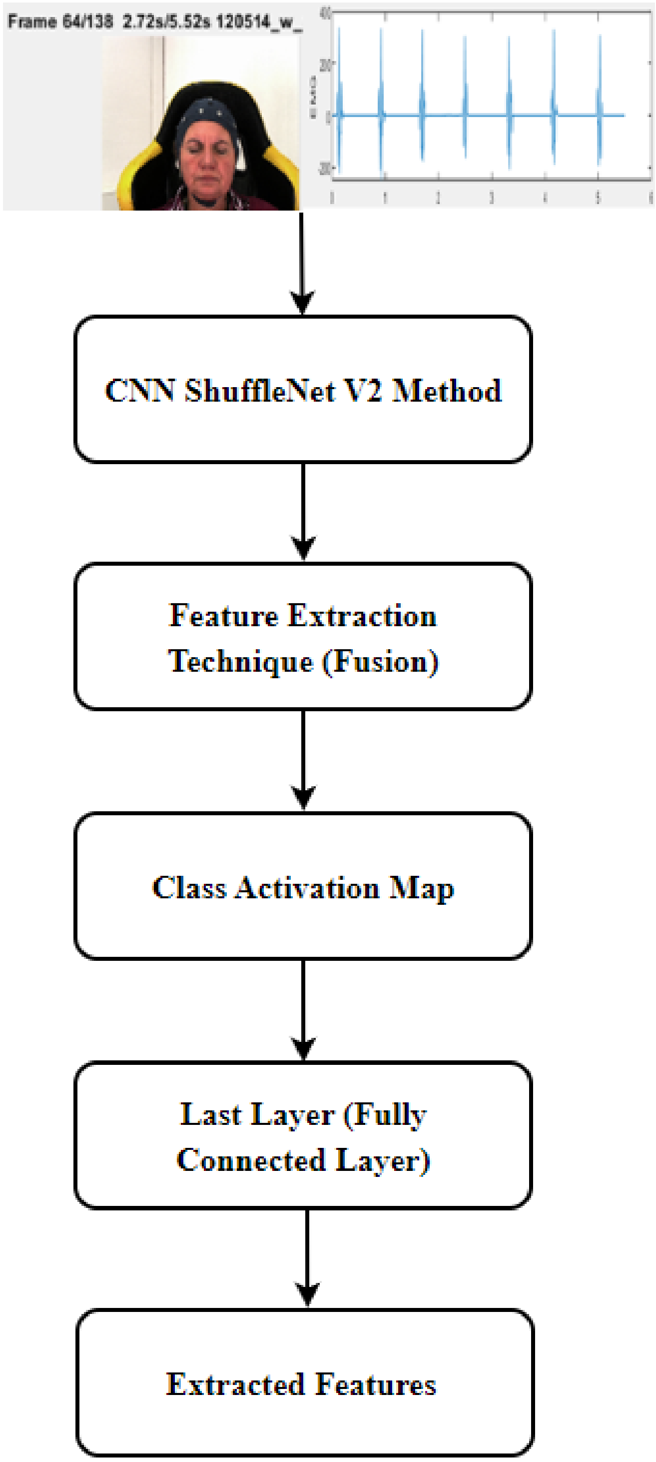

Figure 1. Proposed feature extraction method from face and physiological signal (EMG)

3.3 Classification Model

The proposed pain intensity level recognition using different classifier shown in Figure 2.

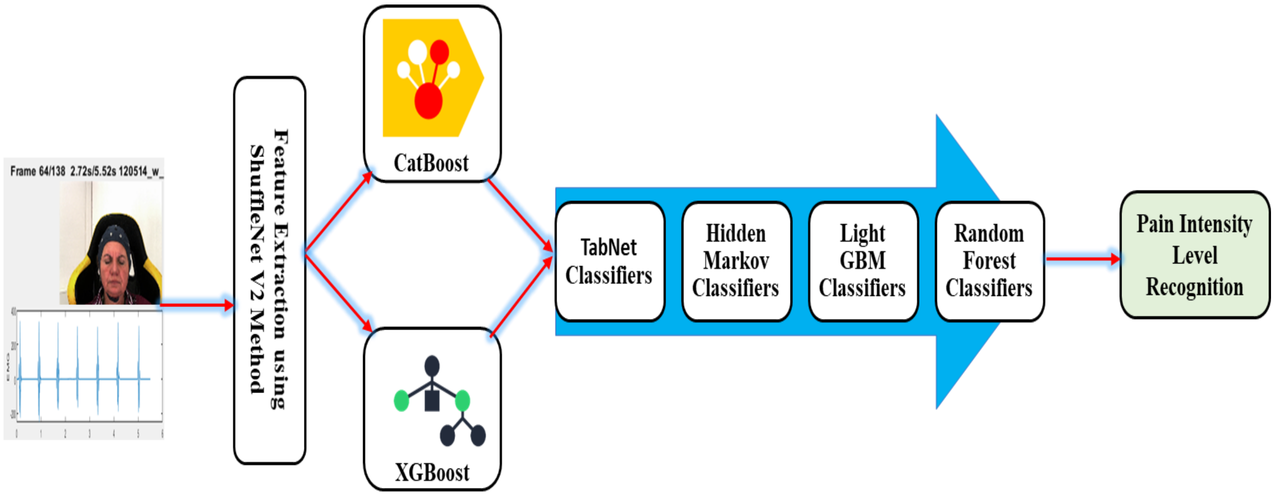

Figure 2. Proposed framework for pain level classification using various classifier

3.3.1 TabNet Model

The primary objective is to use facial and physiological signals to distinguish between pain (B0) and its absence (P1, P2, P3, P4). We create four binary classifications task for this. We extract features from face and EMG signal has been combined. Based on fused feature we use decision tree (DT) classifiers for our analysis [31]. The incorporation of deep neural networks into DTs is known as TabNet [32]. The TabNet design includes many transformer blocks to learn pertinent characteristics and a Batch Normalization layer to filter the raw input [33]. Additionally, it uses learnable masks to select which feature to evaluate at each decision step and a sequential attention mechanism. Because the learning capacity is allocated to the most salient features, this trait facilitates effective learning. Based on their performance, we may finally choose the optimal set of signals, features, and models [34-36].

3.3.2 Light Gradient Boosting Machine (LightGBM) Model

LightGBM is a type of gradient-boosting method used for predictive modelling in supervised learning tasks, such as regression, classification, and ranking. .Microsoft created the open-source library known as LightGBM. The way LightGBM constructs trees sets it apart from more conventional gradient boosting methods like XGBoost and GBM [37]. It reduces learning time and memory usage by employing a gradient-based one-sided sampling technique, selecting only the most crucial data points for tree construction. LightGBM keeps adding decision trees to the ensemble iteratively until an interruption requirement is satisfied, like achieving the lowest improvement in the validation set error or the maximum number of trees [38].LightGBM was used for this investigation because of its strong predictive ability. It constructs trees that progressively fix the prior trees prediction mistakes as an ensemble approach. Its support for parallel computing and its fast convergence time during training with very big datasets give it an additional computational edge over ensemble algorithms like random forest [39]. LightGBM has been shown to have the potential to outperform CART and SVM by at least 16 times and outperform the extreme gradient boost by 26 times [40].

3.3.3 Hidden Markov Model

The hidden Markov model (HMM) is a type of probabilistic model that represents time series data as a sequence of K states. Each state is linked to a unique probability distribution, often called an observation model, which belongs to a specific family of probability distributions (such as Gaussian). Using a data-driven approach, HMM inference estimates the state variables, the transition probability matrix (i.e., the probabilities of switching between states or remaining in the same state), initial state probabilities (i.e., the likelihood of each state at the beginning of a sequence), and the probability of each state being active at each time point [41-43]. Here, two HMM variants with distinct observation models the HMM-MAR and the HMM-TDE were investigated [44].

We changed the following in our analyses:

- The relevant model hyperparameters for the HMMTDE include the lag structure, defined by the width L and the inter-lag steps S, as well as the order P for the HMM-MAR.
- The Dirichlet distribution concentration parameter, 2 here as d, parameterizes the prior probability of staying in the similar state as opposite to changing states.
- K is the number of states.

3.3.4 Gaussian Process Classification Model

A Gaussian Process (GP) is the extension of the multidimensional Gaussian distribution to infinite numerous dimensions, where each finite number has a multimodal Gaussian distribution [45]. A GP's mean and covariance functions may be used to characterize it in the same way that a Gaussian distribution's mean vector and covariance matrix does. GP frameworks may be seen from two viewpoints; "weight space" and "function space". GPR is a Bayesian extension of ordinary linear regression, where estimates are derived from a weight vector and Gaussian noise. In Gaussian Process mathematical models, a zero mean Gaussian Process prior is applied to the weights, and the posterior distribution is determined using Bayes' theorem [46].

Gaussian Process Classification (GPC) builds upon Gaussian Process Regression (GPR) by applying a Gaussian Process prior to an unconstrained latent function and then computing its posterior distribution. In this context, instead of observing the function directly, we employ a sigmoidal response function to map it to the unit interval. This study utilizes the cumulative Gaussian density $\Phi\ (x)$, commonly referred to as probit likelihood. The response function converts an unbounded regression problem into a classification task with outputs limited to the unit interval, ensuring an appropriate probabilistic interpretation [47]. Developing GPC projections involves two steps. To make a probabilistic forecast, we first compute the dormant variable's distribution at the check point, followed by its anticipation. Class probabilities are calculated by integrating the latent function across the full distribution at the test data point, as opposed to SVM's point predictions [48].

3.4 Evaluation metrics for pain classification models

Assessing the effectiveness of a pain classification model involves utilizing a range of metrics to assess its effectiveness. Here's a breakdown of key metrics, including their formulas.

3.4.1 Accuracy

Accuracy evaluates the overall correctness of the model's predictions and can be defined as:

$$\text{Accuracy} = \frac{\text{TP + TN}}{\text{TP + FP + TN + FN}} \tag{1}$$

A high accuracy score suggests that the model is correctly predicting a significant number of instances [49].

3.4.2 Precision

Precision determines the amount of presented outcomes that are extremely beneficial. Precision express as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP + FP}} \tag{2}$$

The high level of precision indicate that in the event the algorithm predicts a favourable situation, there is an excellent possibility that its prediction is accurate [50].

3.4.3 Recall

Recall is an indicator of the algorithm's capability to accurately recognize all true positive instances. It shows the number of actual instances which have been accurately predicted. Recall may be expressed as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP + FN}} \tag{3}$$

Excellent recall show that the algorithm accurately recognizes a large number of positive instances. This is especially important whenever the implications of false negatives are significant [49].

3.4.4 F1-Score

The F1-score is the harmonics mean of accuracy and recall, giving a balanced assessment, and it is particularly useful for distorted datasets. The F1 score can be expressed as:

$$\text{F1-Score} = 2 * \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{4}$$

A high F1-score redirects an effective balance between precision and recall [49, 50].

3.4.5 Kappa

Kappa evaluates the level of consistency among an algorithm's predictions and actual labels, considering into account the possibility of chance matching. Kappa can be expressed as follows:

$$\text{Kappa} = \frac{\text{Observed Accuracy} - \text{Expected Accuracy}}{1 - \text{Expected Accuracy}} \tag{5}$$

It is important that you choose the most appropriate metric(s) depending on the specific intends and characteristics associated with your pain classification problem. As an example, whenever missing a pain assessment actually undesirable, emphasize recollection. If avoiding unnecessary treatments due to incorrect results is critical, prioritise accuracy. The F1-score achieves an appropriate proportion provided both are important [51].

4. Results and Discussion

Table 1 summarizes the results of a five-fold cross-validation analysis on Part A of the Biovid dataset, which used the Leave-One-Subject-Out cross-validation strategy. The classification job relies on the TabNet algorithm for separating between T0 and T4. Table 1 shows the performance characteristics for each of the five folds. The algorithm's prediction accuracy varies from 94.2% to 97.8% across folds, having Fold 5 providing the highest accuracy.

Table 1. A five-fold cross-validation analysis performed on Biovid (Part A) dataset for the classification task T0 vs. T4 using the TabNet algorithm

Folds	Accuracy %	Precision %	Recall %	F1-score %	Kappa %
1	94.2	94.5	93.3	94.1	89.1
2	95.6	95.5	94.5	94.6	89.6
3	95.8	96.6	95.6	95.8	90.8
4	96.7	97.8	96.8	97.1	91.9
5	97.8	97.3	97.3	97.6	92.4

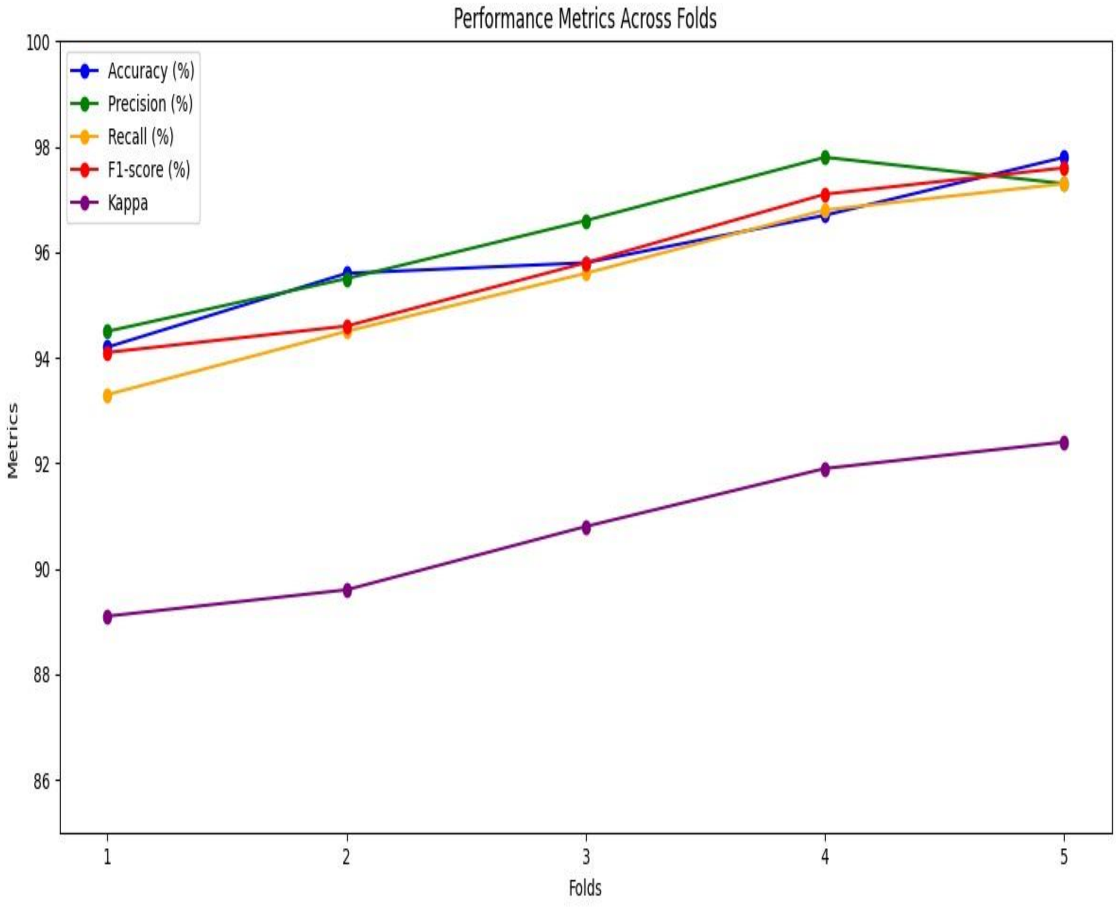

Figure 3. Graphical representation of performance metrics for classification task T0 vs. T4 using the TabNet model

Table 2. A five-fold cross-validation analysis has been performed on the BVDB (Part A) in a LOSO for the classification task T0 vs. T4 using the LightGBM model.

Folds	Accuracy %	Precision %	Recall %	F1-score %	Kappa %
1	92.5	91.8	92.2	92.2	87.2
2	93.4	92.3	93.5	92.8	87.6
3	94.5	93.8	94.3	92.9	88.8
4	95.2	94.8	94.9	93.2	89.6
5	95.8	95.3	95.1	93.8	90.3

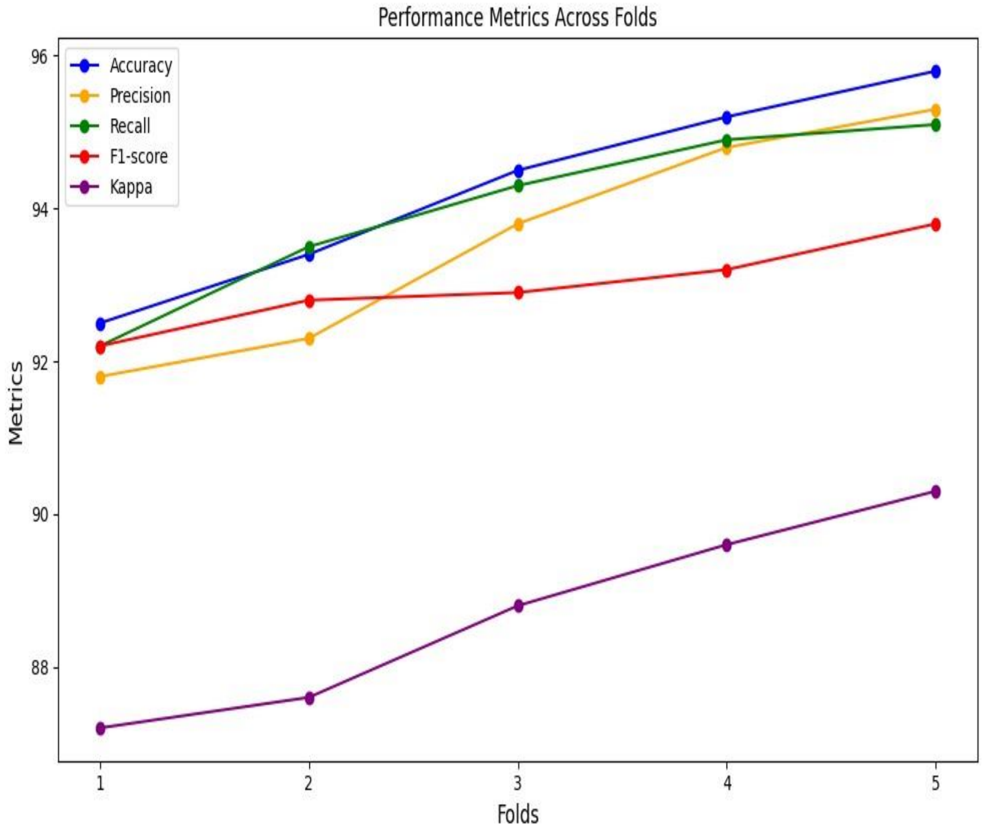

Figure 4. Graphical representation of performance metrics for classification task T0 vs. T4 using the LightGBM model

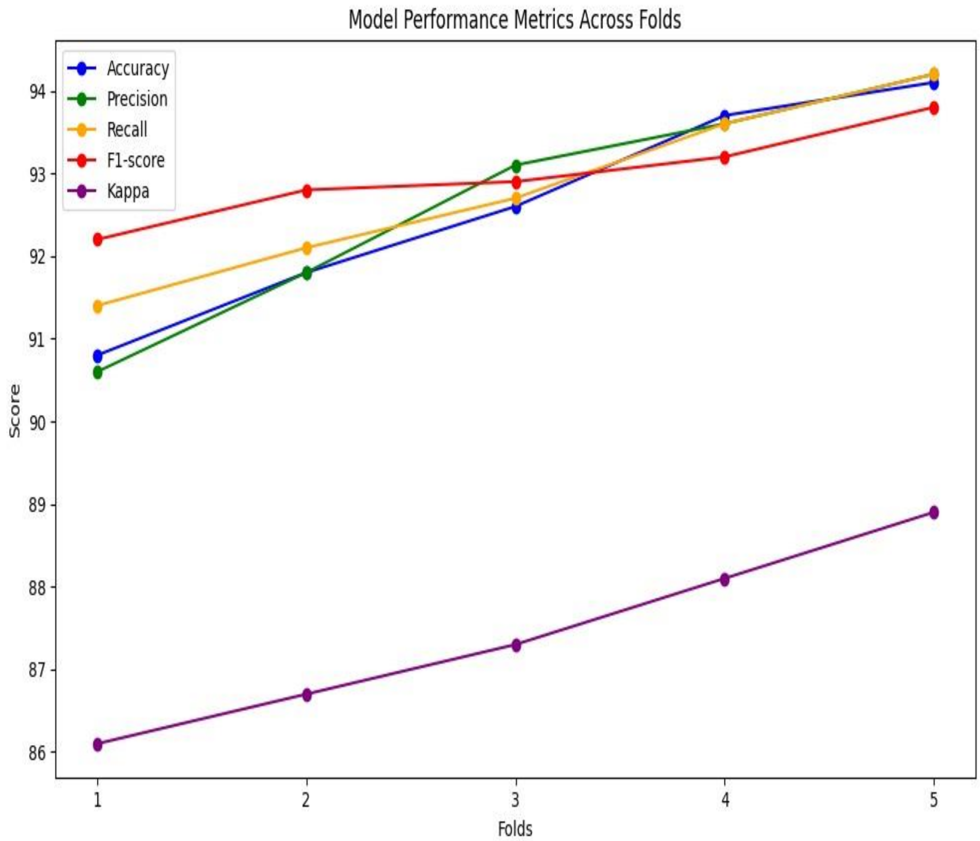

Figure 5. Graphical representation of performance metrics for classification task T0 vs. T4 using the Hidden Markov Model

Table 3. A five-fold cross-validation analysis has been performed on the BVDB (Part A) in a LOSO for the classification task T0 vs. T4 using the Hidden Markov Model

Folds	Accuracy %	Precision %	Recall %	F1-score %
1	90.8	90.6	91.4	92.2
2	91.8	91.8	92.1	92.8
3	92.6	93.1	92.7	92.9
4	93.7	93.6	93.6	93.2
5	94.1	94.2	94.2	93.8

Table 4. A five-fold cross-validation analysis has been performed on the BVDB (Part A) in a LOSO for the classification task T0 vs. T4 using the Gaussian Process classification model

Folds	Accuracy %	Precision %	Recall %	F1-score %
1	90.8	90.6	91.4	92.2
2	91.8	91.8	92.1	92.8
3	92.6	93.1	92.7	92.9
4	93.7	93.6	93.6	93.2
5	94.1	94.2	94.2	93.8

Table 5. Integration performance compared to previous work on the BVDB (Part A) in a LOSO cross-validation setup for classification tasks T0 and T4

Classification Model	Accuracy %	Precision %	Recall %	F1-score %	Kappa %
Susam *et al*. [49]	97.1	95.8	95.7	95.7	90.1
Semwal and Londhe [50]	94.8	94.6	94.5	94.5	91.2
Bargshady *et al*. [51]	95.6	96.1	95.8	95.9	89.7
El Morabit *et al*. [52]	94.1	94.5	94.3	94.4	90.9
Proposed Method (TabNet)	**97.8**	**97.3**	**97.3**	**97.6**	**92.4**

Precision differs from 94.5% to 97.8%, with Fold 4 providing the most accurate results. Recall, or sensitivity, is the proportion of correctly detected positive cases among all real positives and runs from 93.3% to 97.3%, with Fold 5 scoring the highest. The F1-score, indicating the harmonic mean of accuracy and recall, is a balanced statistic for both and runs from 94.1% to 97.6%, with Fold 5 again having the highest result.

Overall, Table 1 demonstrates consistently high performance across all five folds for the TabNet model on the T0 vs. T4 classification task. Fold 5 generally exhibits the best performance across all reported metrics. Graphical representation of performance metrics for classification task T0 vs. T4 using the TabNet model presented in Figure 3.

Table 2 displays the performance of the LightGBM model on this specific task, with accuracy scores across folds ranging from around 92.5% to 95.8%. This variation highlights differences in performance based on the subject excluded for testing. Precision and recall evaluation provided a measure of the framework's capacity to correctly recognize T4 tasks while simultaneously reducing false positives and false negatives. The F1 score optimizes accuracy and recall, and the Kappa statistic measures the algorithm's performance.

To gain a better understanding, compare these results with different approaches (Tables 1, 3, and 4). This comparison demonstrates each algorithm's advantages and disadvantages for the particular task at issue. Figure 4 shows the performance metrics for the T0 vs. T4 classification evaluated with the LightGBM model.

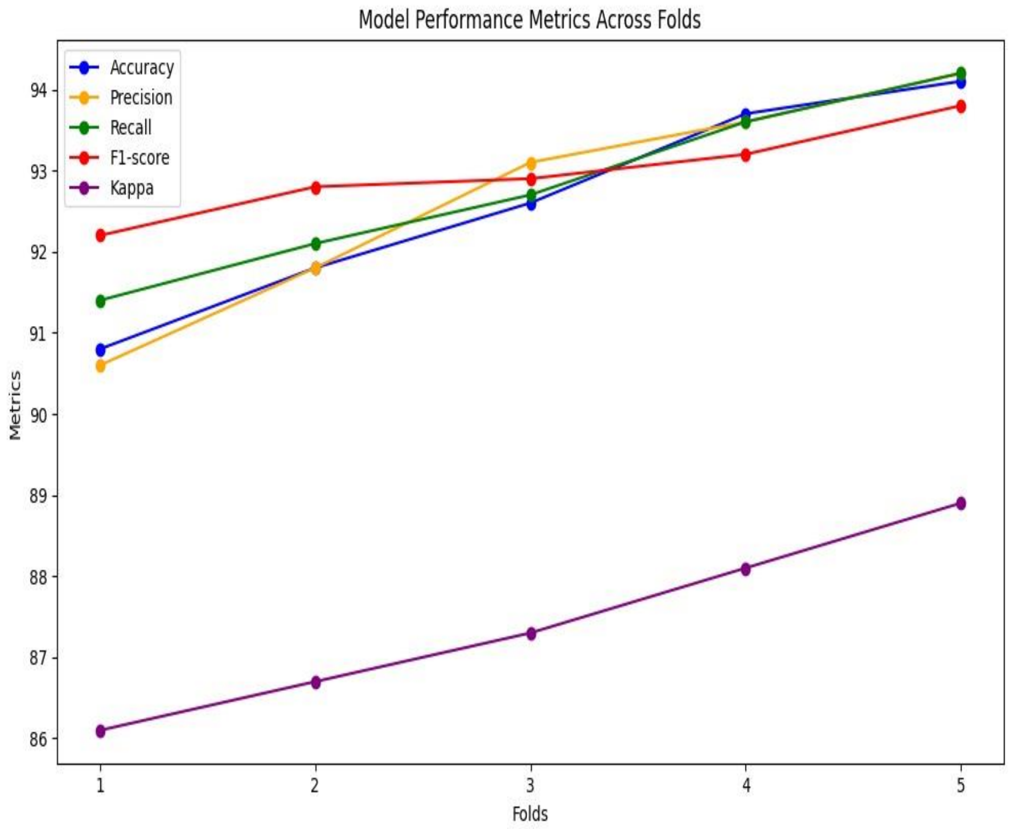

Figure 6. Graphical representation of performance metrics for classification task T0 vs. T4 using the Gaussian Process classification model

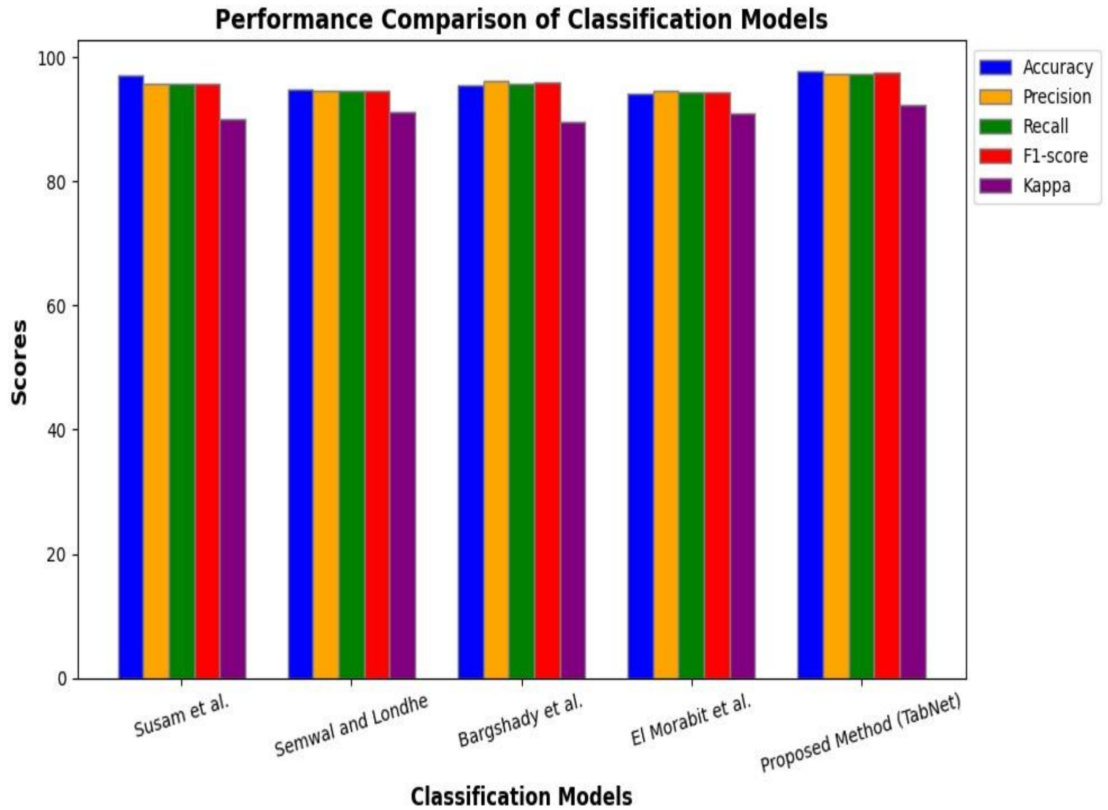

Figure 7. Comparison of existing methods using graphical representation of performance metrics for pain classification task T0 vs. T4

Table 3 presents the performance results of the HMM across various folds, with accuracy ranging from about 90.8% to 94.1%, reflecting some variability based on the subject excluded for testing. Precision, recall, and F1-score provides an additional analysis of the algorithm's performance in classification, emphasis different aspects of accuracy. The Kappa statistic further evaluates the model's effectiveness beyond chance agreement.

A comparison of these results with those in Tables 1, 2, and 4, which employ different models, provides a more thorough evaluation and helps identify the most effective model for this particular task. Figure 5 offers a graphical representation of performance metrics for the T0 vs. T4 classification task using the Hidden Markov Model. The results in Table 4 show the Gaussian Process model's performance across the 5 folds. The accuracy ranges from about 90.5% to 92.8%, indicating some variability depending on the subject held out for testing. Precision, recall, and F1-score provide a a more detailed perspective on the model's performance. Kappa assesses performance beyond chance agreement. Comparing these results with Tables 1, 2, and 3 (using different models) helps determine the most suitable model for this task. Graphical representation of performance metrics for classification task T0 vs. T4 using the Gaussian Process model's shown in Figure 6.

Table 5 displays the combined performance of multiple models of classification for tasks T0 and T4 (pain vs. no pain) on the Biovid Part A dataset in a Leave-One-Subject-Out (LOSO) cross-validation situation. The recommended method has the following metrics: accuracy (97.8%), precision (97.3%), recall (97.3%), F1-score (97.6%), and Kappa (92.4%). Compared with previous approaches, the recommended TabNet method performs significantly with a 97.1% accuracy. Method [49] perform well but falls short in terms of precision, recall, F1-score, and Kappa. Although they performed well models such as [50, 51] failed to satisfy the standards of the recommended approach. Comparison of pain classification using proposed method with existing method shown in Figure 7 with bar graph representation proposed method performs well.

5. Conclusions

The present study provides a distinctive approach to enhance acute pain classification performance via the integration of face images and electromyography (EMG) information. Through using ensemble learning, which incorporates numerous algorithms in order to boost overall performance, this approach not only improves predicted accuracy but also offers understanding on the physiological and emotional elements of acute pain. In the proposed method, we applied a feature extraction strategy based on a convolutional neural network (CNN) architecture, namely the ShuffleNet V2 algorithm, which provides computational simplicity while also extracting meaningful information from images of faces and EMG data. This approach optimizes using five CNN layers and batch normalization, providing a faster and more effective estimated value.

To accurately investigate pain signals, information from both modalities is integrated using an element-wise fusion approach. The proposed approach, operated by TabNet, showed outstanding results, with an accuracy of 97.8%, exceeding preceding models' 97.1%. It also outperformed in other performance criteria, including accuracy (97.3%), recall (97.3%), F1-score (97.6%), and a Kappa score of 92.4, showing its reliability and stability. The results presented show the significant method in automated pain recognition systems, with the potential to improve clinical diagnosis and patient care. This framework will be a very effective solution for pain assessment due to its integrated feature extraction algorithms and remarkable pain classification accuracy.

References

[1] M. Zerriouh, G. De Clifford-Faugère, H.L. Nguena Nguefack, M.G. Pagé, L. Guénette, L. Blais, A. Lacasse, Pain relief and associated factors: a cross-sectional observational web-based study in a Quebec cohort of persons living with chronic pain. Frontiers in Pain Research, 5, (2024) 1306479. https://doi.org/10.3389/fpain.2024.1306479

[2] J.A. Ten Barge, M. Baudat, N.J. Meesters, A. Kindt, E.A. Joosten, I.K. Reiss, G.E. van den Bosch, Biomarkers for assessing pain and pain relief in the neonatal intensive care unit. Frontiers in Pain Research, 5, (2024)1343551. https://doi.org/10.3389/fpain.2024.1343551

[3] L. Mears, J. Mears, The pathophysiology, assessment, and management of acute pain. British Journal of Nursing, 32(2), (2023) 58-65. https://doi.org/10.12968/bjon.2023.32.2.58

[4] [56] A.R. Wahab Sait, A.K. Dutta, Ensemble Learning-Based Pain Intensity Identification Model Using Facial Expressions. Journal of Disability Research, 3(3), (2024) 20240029. https://doi.org/10.57197/JDR-2024-0029

[5] G. Bargshady, X. Zhou, R. C. Deo, J. Soar, F. Whittaker, H. Wang, Ensemble neural network approach detecting pain intensity from facial expressions. Artificial Intelligence in Medicine, 109, (2020) 101954. https://doi.org/10.1016/j.artmed.2020.101954

[6] S. Alphonse, S. Abinaya, N. Kumar, Pain assessment from facial expression images utilizing Statistical Frei-Chen Mask (SFCM)-based features and DenseNet. Journal of Cloud Computing, 13(1), (2024) 142.

https://doi.org/10.1186/s13677-024-00706-9

[7] J. Huo, Y. Yu, W. Lin, A. Hu, C. Wu, Application of AI in multilevel pain assessment using facial images: systematic review and meta-analysis. Journal of Medical Internet Research, 26, (2024) e51250. https://doi.org/10.2196/51250

[8] R. Gutierrez, J. Garcia-Ortiz, W. Villegas-Ch, Multimodal AI techniques for pain detection: integrating facial gesture and paralanguage analysis. Frontiers in Computer Science, 6, (2024) 1424935. https://doi.org/10.3389/fcomp.2024.1424935

[9] S. Gkikas, N.S. Tachos, S. Andreadis, V.C. Pezoulas, D. Zaridis, G. Gkois, A. Matonaki, T.G. Stavropoulos, D.I. Fotiadis, Multimodal automatic assessment of acute pain through facial videos and heart rate signals utilizing transformer-based architectures. Frontiers in Pain Research, 5, (2024) 1372814. https://doi.org/10.3389/fpain.2024.1372814

[10] F. Pouromran, Y. Lin, and S. Kamarthi, Personalized Deep Bi-LSTM RNN based model for pain intensity classification using EDA signal. Sensors, 22(21), (2022) 8087. https://doi.org/10.3390/s22218087

[11] [13] P.D. Barua, N. Baygin, S. Dogan, M. Baygin, N. Arunkumar, H. Fujita, T. Tuncer, R.S. Tan, E. Palmer, M. Mokhzaini Bin Azizan, N.A. Kadri, U.R. Acharya, Automated detection of pain levels using deep feature extraction from shutter blinds-based dynamic-sized horizontal patches with facial images. Scientific Reports, 12(1), (2022) 17297. https://doi.org/10.1038/s41598-022-21380-4

[12] E. Othman, P. Werner, F. Saxen, M.A. Fiedler, A. Al-Hamadi, An Automatic System for Continuous Pain Intensity Monitoring Based on Analyzing Data from Uni-, Bi-, and Multi-Modality. Sensors, 22(13), (2022) 4992. https://doi.org/10.3390/s22134992

[13] P. Thiam, H. Kestler, F. Schwenker, (2020) Multimodal Deep Denoising Convolutional Autoencoders for Pain Intensity Classification based on Physiological Signals. in Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods, Scitepress - Science and Technology Publicationsm 289–296. https://doi.org/10.5220/0008896102890296

[14] M.K. Hasan, G.M.T. Ahsan, S.I. Ahamed, R. Love, R. Salim, Pain Level Detection From Facial Image Captured by Smartphone. Journal of Information Processing, 24(4), (2016) 598-608. https://doi.org/10.2197/ipsjjip.24.598

[15] M. Feighelstein, L. Henze, S. Meller, I. Shimshoni, B. Hermoni, M. Berko, F. Twele, A. Schütter, N. Dorn, S. Kästner, L. Finka, S.P.L. Luna, D.S. Mills, H.A. Volk, A. Zamansky, Explainable automated pain recognition in cats. Scientific reports, 13(1), (2023) 8973. https://doi.org/10.1038/s41598-023-35846-6

[16] R. Fernandez Rojas, N. Hirachan, N. Brown, G. Waddington, L. Murtagh, B. Seymour, R. Goecke, Multimodal physiological sensing for the assessment of acute pain. Frontiers in Pain Research, 4, (2023) 1150264. https://doi.org/10.3389/fpain.2023.1150264

[17] “BioVid Heat Pain Database,” nit. Accessed: Jul. 06, 2024. [Online]. Available: https://www.nit.ovgu.de/-p-1358.html.

[18] N. Ma, X. Zhang, H.T. Zheng, J. Sun, Shufflenet v2: Practical guidelines for efficient cnn architecture design. In Proceedings of the European conference on computer vision (ECCV), (2018) 116-131. https://doi.org/10.1007/978-3-030-01264-9_8

[19] S. Ioffe, C. Szegedy, (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv.

[20] Y. Chen, Z. Zhang, Z. Chen, Y. Zhang, J. Wang, Fine-Grained Classification of Optical Remote Sensing Ship Images Based on Deep Convolution Neural Network. Remote Sensing, 14(18), (2022) 4566. https://doi.org/10.3390/rs14184566

[21] M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning, (2019) 6105-6114.

[22] C. Chatterjee, N. Shah, S. Bhatt, S. Chandi, (2024) A Survey on Multi-modal Emotion Detection Techniques. Research Square. https://doi.org/10.21203/rs.3.rs-3942784/v1

[23] J. Qian, J. Yang, Y. Xu, J. Xie, Z. Lai, B. Zhang, Image decomposition based matrix regression with applications to robust face recognition. Pattern Recognition, 102, (2020) 107204. https://doi.org/10.1016/j.patcog.2020.107204

[24] M.-P. Hosseini, A. Hosseini, K. Ahi, “A Review on Machine Learning for EEG Signal Processing in Bioengineering. IEEE Reviews in Biomedical Engineering, 14, (2021) 204–218. https://doi.org/10.1109/RBME.2020.2969915

[25] A. V. Dorogush, V. Ershov, A. Gulin, (2018) CatBoost: gradient boosting with categorical features support. arXiv.

[26] T. Chen, C. Guestrin, (2016) XGBoost: A Scalable Tree Boosting System. in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, USA. https://doi.org/10.1145/2939672.2939785

[27] M. Feurer, F. Hutter, (2019) Hyperparameter Optimization. Automated Machine Learning, Springer International Publishing, 3–33. https://doi.org/10.1007/978-3-030-05318-5_1

[28] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review, 54(3), (2021) 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

[29] M. Dong, L. Yao, X. Wang, B. Benatallah, S. Zhang, Q. Z. Sheng, Gradient Boosted Neural Decision Forest. IEEE Transactions on Services Computing, 16(1), (2023) 330-342. https://doi.org/10.1109/TSC.2021.3133673

[30] X. Xu, X. Wang, Z. Sun, S. Wang, Face recognition technology based on CNN, XGBoost, model fusion and its application for safety management in power system. IOP Conference Series: Earth and Environmental Science, 645(1), (2021) 012054. https://doi.org/10.1088/1755-1315/645/1/012054

[31] V.K. Mishra, A. Rajagopalan, A Novel Extreme Value Theory Based approach to Hyperparameter Optimization. Procedia Computer Science, 218, (2023) 2411–2419. https://doi.org/10.1016/j.procs.2023.01.216

[32] G. Chandrashekar, F. Sahin, A survey on feature selection methods. Computers & Electrical Engineering, 40(1), (2014) 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024

[33] S.Ö. Arik, T. Pfister, TabNet: Attentive Interpretable Tabular Learning. Proceedings of the AAAI Conference on Artificial Intelligence, 35(8), (2021) 6679–6687. https://doi.org/10.1609/aaai.v35i8.16826

[34] S. Kamil Gatfan, A Review on Deep Learning For Electroencephalogram Signal Classification. Journal of Al-Qadisiyah for Computer Science and Mathematics, 16(1), (2024)137–151. https://doi.org/10.29304/jqcsm.2024.16.11453

[35] D.R. Cutler, T.C. Edwards, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, J.J. Lawler, Random forests for classification in ecology. Ecology, 88(11), (2007) 2783-2792. https://doi.org/10.1890/07-0539.1

[36] B.D. Romaissa, O. Mourad, N. Brahim, (2021) Vision-Based Multi-Modal Framework for Action Recognition. 25th International Conference on Pattern Recognition (ICPR), IEEE, Italy. https://doi.org/10.1109/ICPR48806.2021.9412863

[37] M. Hajihosseinlou, A. Maghsoudi, R. Ghezelbash, A Novel Scheme for Mapping of MVT-Type Pb–Zn Prospectivity: LightGBM, a Highly Efficient Gradient Boosting Decision Tree Machine Learning Algorithm. Natural Resources Research, 32(6), (2023) 2417–2438. https://doi.org/10.1007/s11053-023-10249-6

[38] W.Y. Chen, T.C. Wang, R.S. Guo, C. Chen, Integrating Latent Dirichlet Allocation and Gradient Boosting Tree Methodology for Insurance Product Development Recommendation. in 2024 9th International Conference on Big Data Analytics (ICBDA), IEEE, Japan. https://doi.org/10.1109/ICBDA61153.2024.10607213

[39] R. Szczepanek, Daily Streamflow Forecasting in Mountainous Catchment Using XGBoost, LightGBM and CatBoost. Hydrology, 9(12), (2022) 226. https://doi.org/10.3390/hydrology9120226

[40] Y.C. Chu, S.S.S. Chen, K.B. Chen, J.S. Sun, T.K. Shen, L.K. Chen, Enhanced labor pain monitoring using machine learning and ECG waveform analysis for uterine contraction-induced pain. BioData Mining, 17(1), (2024) 32. https://doi.org/10.1186/s13040-024-00383-z

[41] Y. Khalifa, D. Mandic, E. Sejdić, A review of Hidden Markov models and Recurrent Neural Networks for event detection and localization in biomedical signals. Information Fusion, 69, (2021) 52–72. https://doi.org/10.1016/j.inffus.2020.11.008

[42] H. Salam, O. Celiktutan, H. Gunes, M. Chetouani, (2022) Automatic Context-Driven Inference of Engagement in HMI: A Survey. arXiv.

[43] C. Li, A. Pourtaherian, L. Van Onzenoort, W.E.T.A. Ten, P.H.N. De With, Infant Facial Expression Analysis: Towards a Real-Time Video Monitoring System Using R-CNN and HMM. IEEE Journal of Biomedical and Health Informatics, 25(5), (2021) 1429–1440. https://doi.org/10.1109/JBHI.2020.3037031

[44] M. Capecci, M.G. Ceravolo, F. Ferracuti, S. Iarlori, V. Kyrki, A. Monteriu, F. Verdini, A Hidden Semi-Markov Model based approach for rehabilitation exercise assessment. Journal of biomedical informatics, 78, (2018) 1-11. https://doi.org/10.1016/j.jbi.2017.12.012

[45] A. Sumayli, S.M. Alshahrani, Modeling and prediction of biodiesel production by using different artificial intelligence methods: Multi-layer perceptron (MLP), Gradient boosting (GB), and Gaussian process regression (GPR). Arabian Journal of Chemistry, 16(7), (2023) 104801. https://doi.org/10.1016/j.arabjc.2023.104801

[46] A.L. Muyskens, I.R. Goumiri, B.W. Priest, , Schneider, M. D., Armstrong, R. E., Bernstein, J., & Dana, R. (2022). Star–galaxy image separation with computationally efficient gaussian process classification. The Astronomical Journal, 163(4), 148. https://doi.org/10.3847/1538-3881/ac4e93

[47] H. Liu, Y.S. Ong, X. Shen, J. Cai, When Gaussian Process Meets Big Data: A Review of Scalable GPs. IEEE Transactions on Neural Networks and Learning Systems, 31(11), 4405–4423. https://doi.org/10.1109/TNNLS.2019.2957109

[48] S. Grothus, A. Sommer, L. Stahlschmidt, G. Hirschfeld, L. Höfel, R. Linder, J. Wager, (2024). Pediatric chronic pain grading: a revised classification of the severity of pediatric chronic pain. Pain, 165(9), 2087-2097. https://doi.org/10.1097/j.pain.0000000000003226

[49] B.T. Susam, N.T. Riek, M. Akcakaya, X. Xu, V.R. de Sa, H. Nezamfar, D. Diaz, K.D. Craig, M.S. Goodwin, J.S. Huang, (2021) Automated pain assessment in children using electrodermal activity and video data fusion via machine learning. IEEE Transactions on Biomedical Engineering, 69(1), 422-431. https://doi.org/10.1109/TBME.2021.3096137

[50] A. Semwal, N.D. Londhe, Computer aided pain detection and intensity estimation using compact CNN based fusion network. Applied Soft Computing, 112, (2021)107780. https://doi.org/10.1016/j.asoc.2021.107780

[51] G. Bargshady, X. Zhou, R. C. Deo, J. Soar, F. Whittaker, H. Wang, Enhanced deep learning algorithm development to detect pain intensity from facial expression images. Expert systems with applications, 149, (2020) 113305. https://doi.org/10.1016/j.eswa.2020.113305

[52] S. El Morabit, A. Rivenq, M.E. Zighem, A. Hadid, A. Ouahabi, A. Taleb-Ahmed, Automatic pain estimation from facial expressions: A comparative analysis using off-the-shelf CNN architectures. Electronics, 10(16), (2021) 1926. https://doi.org/10.3390/electronics10161926

Authors Contribution Statement

Both the authors equally contributed to the Conceptualization, Methodology, Investigation, Validation, Formal analysis, Data Curation, Writing - Original Draft and Writing - Review & Editing. The final manuscript has been read and approved by all authors.

Acknowledgement

The authors would like to thank Shri Shivaji Vidya Prasarak Sanstha's Bapusaheb Shivajirao Deore College of Engineering, Dhule for the support.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data that support the findings of this study are publically available in at _ https://www.nit.ovgu.de/-p-1358.html [17].

Has this article screened for similarity?

Yes

About the License