Pain Level Modeling of Intensive Care Unit patients with Machine Learning Methods: An Effective Congeneric Clustering-based Approach

Ruijie Fang
University of California, Davis
rjfang@ucdavis.edu

Ruoyu Zhang
University of California, Davis
ryuzhang@ucdavis.edu

Sayed M. Hosseini
Johns Hopkins Hospital
se.mohammad.ho@gmail.com

Mahya Faghih
Johns Hopkins Hospital
mfaghih2@jhu.edu

Soheil Rafatirad
University of California, Los Angeles
Soheil.rafatirad@gmail.com

Setareh Rafatirad
University of California, Davis
srafatirad@ucdavis.edu

Houman Homayoun
University of California, Davis
hhomayoun@ucdavis.edu

ABSTRACT

Recent studies showed that machine learning can assist to better evaluate the induced pain level(s) on healthy individuals in a controlled environment. However, the role of these methods in clinical settings remained unclear and there is an unmet need to develop machine learning assisted tools in pain. The aim of this paper is to develop an automatic pain level assessment model based on patients' physiologic measures from the Medical Information Mart for Intensive Care (MIMIC III). There were two study phases; 1) pilot study, tested three existing machine learning methods proposed recently for healthy individuals. However, these yield poor performances in MIMIC III patients. 2) group study, a novel congeneric clustering method which divided patients into eleven categories and trained a dedicated model for each one. The clustering effectiveness of the proposed congeneric clustering method by showing the highest classification accuracy of 82.86%.

ACM Reference Format:
Ruijie Fang, Ruoyu Zhang, Sayed M. Hosseini, Mahya Faghih, Soheil Rafatirad, Setareh Rafatirad, and Houman Homayoun. 2022. Pain Level Modeling of Intensive Care Unit patients with Machine Learning Methods: An Effective Congeneric Clustering-based Approach. In *2022 4th International Conference on Intelligent Medicine and Image Processing (IMIP 2022), March 18–21, 2022, Tianjin, China.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3524086.3524100

1 INTRODUCTION

More than 1 in 5 adults (50.2 million people) are suffering from pain on most days or every day, which lead to a total value of loss of productivity at $296 million annually, reported by National Health Interview Survey (NHIS) [1]. Pain level assessment is the key to optimal pain treatment. Suboptimal pain management may lead to mental and physical comorbidities, as well as economic losses and low quality of life [2] [4]. More importantly, the use of prescription opioids for the treatment of chronic non-cancer pain is associated with a substantial risk for abuse, dependence and overdose [3]. The development of an objective test for guiding non-opioid therapy in patients with chronic pain could help decrease the risk of opioid-related harms. In addition, pain is a complex and dynamic phenomenon that changes over time. Therefore, a continuous pain level assessment tool is required for optimal pain management. Patient's self-reported tools such as visual analog scale (VAS) and numerical rating scale (NRS) have been widely used in pain level assessment. [5] [6]. Self-report methods require patients to report pain intensity under the instructions of caregivers. However, these methods may be unreliable in certain populations and conditions. Self-reported measures may not be possible for patients with cognitive, developmental or behavioral impairments. For example, a patient may not be able to communicate and report pain verbally, in writing or using facial gestures(e.g., blinking as yes or no) [29]. Under these circumstances, the observation and evaluation of professional and experienced caregivers may be needed. Furthermore, the real-time monitoring of dynamic pain patterns which is not an achievable and subjective aspect of self-report measures can be affected by patients' underlying comorbidities (e.g., depression).

To overcome these shortcomings, various automatic pain measurements have been developed in the last decades. In general, two types of data sources have been used to measure pain: specific behavioral(e.g., facial video) and physiological signals. Lucey *et al.* [7] utilize Active Appearance Model and Support Vector Machine (SVM) to monitor the presence of pain on the UNBC-McMaster Shoulder Pain Dataset. Additionally, Werner *et al.* [8] created the BioVid dataset and apply SVM on the extracted details on facial expression(s) and head pose(s) to classify pain level(s). However, behavioral signals like facial video are difficult to measure and pattern recognition for behavioral signals is computing-intensive.

IMIP 2022, March 18–21, 2022, Tianjin, China

ACM ISBN 978-1-4503-9620-2/22/03.
https://doi.org/10.1145/3524086.3524100

On the other hand, physiologic measures showed promising results. Werner *et al.* [9] included physiological signals to form a multi-modal database(i.e the BioVid database). They utilized random forest and SVM to evaluate changes in pain level on electrocardiogram (ECG), galvanic skin response (GSR) and electromyography (EMG) as well as facial video. Sascha Gruss *et al.* [10] extract 42 different features from ECG, EMG and GSR in the pattern learning perspective and found features from amplitude and similarity group and were derived from facial electromyography were selected the most. Werner et al. [16] [17] first conducted pain level assessment on multimodalities on both sensors and pain stimulation on the new XITE(Experimentally Induced Thermal and Electrical) database and verified the possibility to assess both tonic and phasic pain. Automatic pain level assessment is still in the developing stage and only a few systematic databases are available [28]. Further-more, more modalities need to be included in automatic pain level assessment. For example, respiratory modalities such as respiratory rate, SpO2 and pupil size can be a potential pain indicator [13] [24] but there have never been research that combined respiratory features and pupil size together.

To overcome the lack of studies on patients and modalities, we conducted a series of machine learning-based studies on Medical Information Mart for Intensive Care III (MIMIC-III) dataset. Our study consist of two stages: pilot study and group study. The pilot study reproduces existing methods and group study cluster congeneric patients into the same groups based on their diagnosis and 11 groups are formed. This congeneric clustering is based on the International Classification of Diseases, Ninth Revision (ICD-9) [11], where if varying ICD-9 with different organ involvement were available for a patient, manual classification of patient's primary diagnoses were extracted through review of patient's characteristics and database unstructured free texts by two independent reviewers (MF, SMH). Any disagreements between the two reviewers were resolved through discussion.

The contributions of this paper can be summarized as follows:

- Introduce new modalities into automatic pain level assessment studies including respiratory rate, SpO2 and pupil size which are selected carefully out of many modalities presented exist in our studied dataset.
- Conduct automatic pain level recognition experiments on MIMIC-III database which collects ICU patients' data
- Cluster patients into 11 categories based on the diagnosis to obtain the similarities between congeneric samples and improve the pain level recognition rate.

2 METHODS

In this section we give an overview of the MIMIC-III database and describe our assessment approach. Pain level assessment has three stages: preprocessing, pilot study and group study. In preprocessing stage, unreasonable data and outliers were removed from the dataset and EasyEnsemble [12] algorithm was involved to resample the dataset. When data is ready to use, pilot study was conducted to find the optimal features and machine learning algorithm by performing existing machine learning study on all training data. Results from pilot study also provided references for the verification of our proposed congeneric clustering-based method. In group study stage, we clustered patients into 11 categories and performed machine learning study on each group. Cross-validation was involved in both pilot study stage and group study stage.

2.1 Dataset

Medical Information Mart for Intensive Care III (MIMIC-III) database [20] is a large, available database comprising de-identified health-related data associated with over sixty thousand patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012. The database includes information such as demographics, vital sign measurements and diagnosis notes. We extracted physiological measurement data (Heart Rate, Respiratory Rate, Pupil Size, SpO2), demographic data (Gender, Age), diagnostic data (DRG Code) and pain data (Pain Level) to form a sub-dataset. In total, 63,829 data samples from 11,428 patients were included in this study.

Since data in the MIMIC-III database is recorded by caregivers manually, physiological measures data is recorded discretely with random time intervals range from 5 minutes to 1 hour. Therefore, there's no need to perform feature extraction on the database because of the independence of each data sample so that we leave the original data as features to conduct machine learning experiments. The database is labeled by pain level data that BL0 (baseline 0) refers to no pain and PL1 to PL3 refers to pain level 1 to pain level 3 with increasing pain intensity.

The MIMIC-III database is a comprehensive database containing a lot of information unrelated to our study. Therefore, an exhaustive data cleaning is required to gather pain level assessment-related information from the database. First of all, in the physiological measures data section as well as pain data section, data is recorded with timestamp. In each timestamp data point, we check if there's data we need missed and remove the entire timestamp data if so. By doing this, we clean the database to a sub-database with only information necessary to our study and has a significantly reduced size. After that, we perform a series of data preprocessing techniques to make sure the valuable features of the data have been extracted. Then, strict data cleaning and outlier detection are implemented so that garbage data is guaranteed to be cleaned up. To begin with, we came up with a unreasonable range of vital signs based on previous research [14] [15] and removed these data. The removing principles are:

- Heart rate < 50 or > 150
- Respiratory rate < 10 or > 61
- SpO2 < 40

Data falls in the above principles is regarded as clinically unreasonable and the entire timestamped data containing unreasonable data was be removed. Then, an isolation forest outlier detection algorithm was employed and 2.34% of outlier data points were discarded. Afterward, resample algorithm was performed to balance the dataset to avoid bias when training machine learning model. We chose EasyEnsemble as the resample algorithm which took the same number of samples from the majority class as the minority class randomly, then combined them with the minority class as the training set to train one weak machine learning model and repeat the process several times until we take use of all majority data to obtain several weak machine learning models and constitute all weak models to be a single strong model. With the EasyEnsemble algorithm, we made effective use of every data point to maximize the extraction of useful information from the dataset. After resampling, the dataset is standardized to eliminate the influence of different scales and units in features. Finally, the dataset is imported to a cross-validation (CV) process. Cross-validation, as the name suggests, is to repeatedly use the data, segment the sample data, and combine them into different training sets and test sets. The training set is used to train the model, and the test set is used to evaluate the quality of the model. On this basis, we can get several (here we set it to 5) different training and test sets. A sample in a training set may become a test sample next time. CV can prevent the model from overfitting.

2.2 Pain Level Assessment

We divided pain level assessment into two stages: pilot study stage and group study stage. In pilot study, we intend to verify the effectiveness of existing machine learning methods on patients' data while in group study stage we cluster patients into 11 different groups and perform machine learning recognition on each group to try to upgrade the recognition accuracy.

2.2.1 Pilot study. Different algorithms have their own advantages and disadvantages. Diverse results may come from the same task with different machine learning algorithm [18]. In addition, modalities used are crucial to the success of machine learning classification task [21]. In the MIMIC-III database we have multiple available modalities including Heart Rate, Respiratory Rate, SpO2 and Pupil Size. Among all these modalities, Heart Rate has been proved to be effective in automatic pain level assessment by previous publications [8] [22]. On the other hand, Respiratory Rate, SpO2 and Pupil Size have seldom been used in automatic pain level assessment studies [25]. To explore the performances of potential modalities and machine learning algorithms, we designed the pilot study stage. In pilot study, three different prevalent machine learning methods have been implemented on the entire dataset: SVM, Adaboost and Neural Networks, also four different modality combinations are tested: all four features, remove Respiratory Rate, remove SpO2 and remove Pupil Size.

SVM is a generalized linear classifier that classifies data based on support vectors. Kernel tricks like RBF kernel make SVM a non-linear classifier substantially. SVM suits small sample learning because its final decision only focuses on support vectors. However, such a mechanism requires delicate preprocessing and tuning before applying SVM to maximize the potential of SVM.

AdaBoost is an iterative algorithm. Its core idea is to train different classifiers (weak classifiers) for the same training set and then combine them to form a stronger final classifier (strong classifier). We choose RandomForest (RF) as weak classifiers of Adaboost. RF is widely used in automatic pain recognition studies other than SVM. Compared with SVM, it has better interpretability and robustness.

Neural network is one of the most popular machine learning algorithms [26]. It is often used in the fields of natural language processing and computer vision. When the amount of data is large enough, neural network often has the best performance. But it requires more computation and is also unexplainable.

We also apply GridSearchCV algorithm in each machine learning experiment to dig the potential of each algorithm. We use 5-folds cross-validation to train and validate our models.

The result as shown in table 1 and figure 1 illustrates the classification accuracy under different modalities combinations and different classification tasks (BL0 v.s. PL1 to PL3). Among performances from three different machine learning algorithms, Adaboost with all features involved gives the highest average accuracy (63.71%) in three classification tasks. Therefore, Adaboost was chosen as the machine learning algorithm used in the following work and all features were involved. Using classification accuracies as baseline, accuracy drop/increase is calculated for three different modality combinations.

According to table 2, all three modalities yielded negative accuracy changes and the same results can be found from statistical correlation analysis as shown in figure 2. From table 2 we conclude that SpO2 and pupil size contribute significantly to the assessment accuracy and all three modalities have positive influence on the assessment accuracy which validate their effectiveness in automatic pain level assessment study.

2.2.2 Group study. We hypothesis that pain of congeneric patients has the same feature and therefore, performing machine learning study on dataset that only contains congeneric patients' data will come up with better results than general study regardless of patient's diagnosis. To validate our hypothesis, we deployed Adaboost algorithm which we proved to be the optimum and included all features (Heart Rate, Respiratory Rate, Pupil Size, SpO2, Gender and Age) in the machine learning process. Then, 11 ICD-9 based clusters were chosen to cluster patients according to their diagnosis as shown in table 3. ICD-9 refered to International Classification of Diseases, Ninth Revision which was published in 1978 [11]. 11 different Adaboost models were trained on each cluster data following the same process in pilot study. Table 4 and figure 3 show the classification results.

The classification results show that models for each cluster have higher accuracy than the baseline accuracy from pilot study (63.7%, last row of table 4). Figure 4 displays the Receiver Operating Characteristic (ROC) results of the congeneric clustering method. However, we doubt that the increase of accuracy may be due to the reduction of sample points. The same phenomenon can be seen from the highest clusters: Endocrinology and Rheumatology clusters have the highest accuracy over 80% (82.86% and 81.75%) and they have relatively smaller sample sizes (1679 and 541). To validate the increment of accuracy from the resemblance of congeneric patients rather than the sample size, we analyze the Pearson correlation coefficient between classification accuracy and sample size. The calculated p-value equals 0.054 which is higher than 0.05 so that we are prone to decline the hypothesis that classification accuracy correlates with sample size. We maintain the previous view: by clustering congeneric patients into groups, features from group tend to be similar and therefore, easier for machine learning models to classify. Finally, the classification accuracy increase.

Table 1: Classification accuracy from different algorithms and modalities

Algorithm	Used modalities	BL0 v.s. PL1 (%)	BL0 v.s. PL2 (%)	BL0 v.s. PL3 (%)
SVM	All features	48.82	52.34	58.55
	Remove RR*	49.31	53.29	58.08
	Remove SpO2	51.66	52.59	53.64
	Remove PS**	50.18	51.09	55.2
Adaboost	All features	52.87	58.52	**63.71**
	Remove RR	51.78	58.69	**61.91**
	Remove SpO2	51.14	56.07	**58.75**
	Remove PS	50.99	57.25	**59.21**
NN	All features	53.04	55.8	60.16
	Remove RR	51.78	55.17	61.89
	Remove SpO2	53.53	52.45	58.64
	Remove PS	52.04	53.28	59.3

*RR refers to Respiratory Rate

**PS refers to Pupil Size

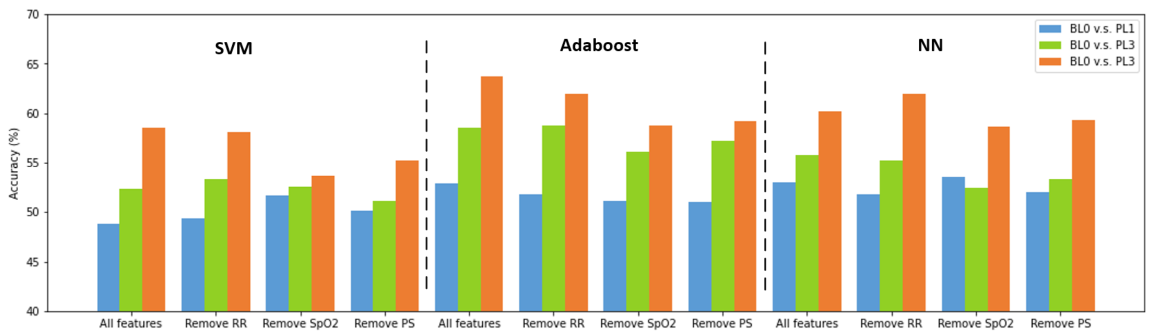

Figure 1: Classification accuracy with different machine learning algorithms and combinations of modalities. Accuracy displayed in percent (%). RR refers to Respiratory Rate, PS refers to Pupil Size

Table 2: Average accuracy changes under different modalities

	BL0 v.s. PL1 (%)	BL0 v.s. PL2 (%)	BL0 v.s. PL3 (%)	average
Remove RR*	-1.14	0.33	-0.25	-0.36
Remove SpO2	1.15	-3.23	-6.23	-2.77
Remove PS**	-0.88	-3.02	-4.74	-2.88

*RR refers to Respiratory Rate

**PS refers to Pupil Size

Table 3: Group list and corresponding sample size

Cluster	Cardio	Endo	Gastro	Hem	Infec	Neuro	Ortho	Pulm	Renal	Rheum	Others
Sample size	9006	679	6398	1921	4582	9133	2442	4361	1263	541	23513

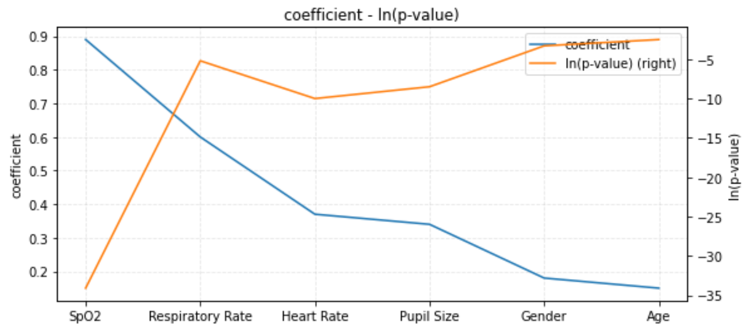

Figure 2: Correlation coefficient and p-value between modalities and pain level. Logarithm of P-value is displayed rather than absolute value since p-value ranges from 10^{-16} to 10^{-1}.

Table 4: Classification results from group study

Cluster	Number ofsamples	BL0 v.s. PL1 (%)	BL0 v.s. PL2 (%)	BL0 v.s. PL3 (%)
Cardiovascular	9006	53.73	56.15	66.52
Endocrinology	1679	66.91	73.81	**82.86**
Gastroenterology	6398	58.62	66.55	74.82
Hematology	1921	61.59	69.26	78.77
Infectious Diseases	4582	56.41	62.52	71.22
Neurological disorders	9133	56.94	64.98	72.65
Orthopedics	2442	58.48	64.51	71.98
Pulmonary	4361	58.96	64.64	73.24
Renal Diseases	1263	62.31	70.09	77.85
Rheumatology	541	64.24	70.71	**81.75**
Others	23513	55.57	62.59	70.23
General study	63829	52.87	58.52	63.71

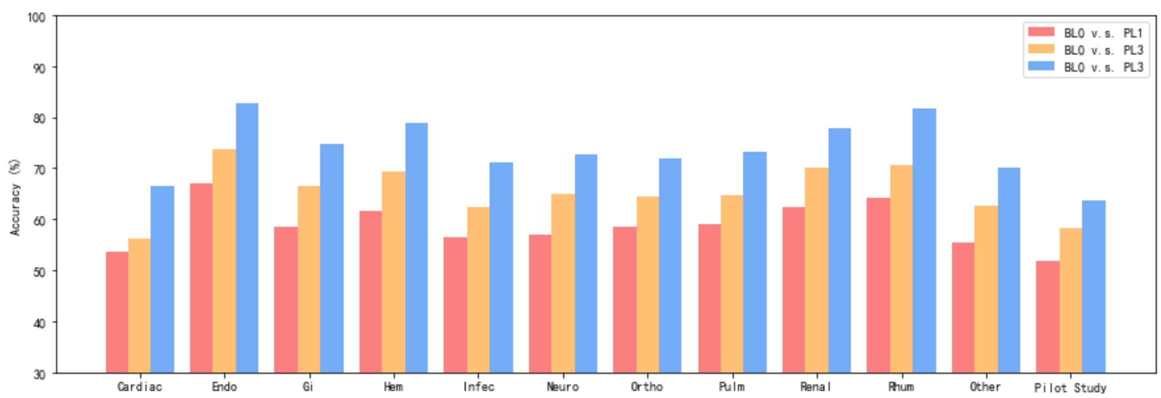

Figure 3: Classification accuracy of group study

3 DISCUSSION

Previous studies proved strong evidence that induced pain in healthy volunteers can be recognized by physiological signals. However, poor performances were yielded by using existing automatic pain level assessment methods from our pilot study. In table 5 we compared previous studies and our work from various aspects. From the table, our method has better performance on clinical patients. The conversion from healthy volunteers to patients as research subjects is crucial in automatic pain study because first, experimental pain tests on healthy volunteers are restricted in time with short pain stimulus in seconds or several minutes and the choice of time window length in HRV analysis is limited [8][22]. For example, in

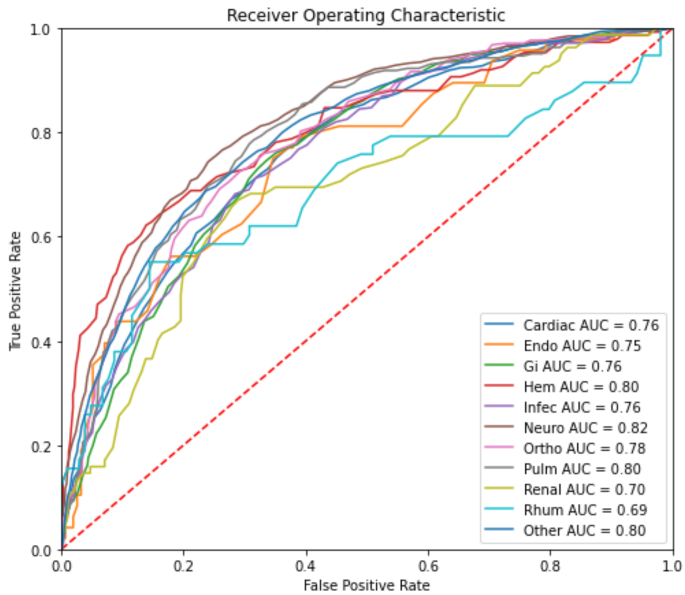

Figure 4: ROC Curve of Congeneric Clustering Method

Table 5: Automatic pain level assessment research comparison

	Algorithm	Modalities	Pain Stimuli	Best accuracy on healthyvolunteers	Best accuracyon patients
Jiang '17 [22]	SVM	ECG	Heat& Elec[1]	AUC = 0.82%	58.55%
Kachele '16 [19]	RF & k-NN	EDA, ECG, EMG	Heat	83.10%	61.89%
Naeini '21 [23]	Adaboost	HRV	Elec	75.68%	63.71%
This Paper	Adaboost	HR[2], RR[3], SpO2, Pupil Size	Patients' Disease	/	82.86%

[1] "Elec" refers to electrical pain stimuli [2] Heart rate [3] Respiratory rate

the BioVid database, thermal pain stimulus was maintained for 4 seconds and then stopped for 8-12 seconds. However, in the actual situation, pain can be constant and last for hours. In this study, we used patient data regardless of the duration of pain so that the conclusion is more generalizable.

In addition, we introduce new modalities to automatic pain level assessment study field, including respiratory rate, SpO2 and pupil size. Respiratory rate and SpO2 are closely related to breath intensity. When human beings are suffering from pain, their breathing will go heavily to inhale more air. Thus, the respiratory rate and SpO2, which reflect breath intensity, can be used to track pain. Besides, pupil size has been proved to be effective in reflecting brain activities [27]. It is reasonable that we hypothesis pain, as one of the brain activities, can be recognized by pupil size and we verified it in this paper. Moreover, group study was deployed after pilot study. We clustered patients into 11 different categories and train machine learning models on each category. Results show that classification accuracy from each group is higher than the baseline accuracy which we got from pilot study. This may be due to the fact that there is a resemblance among patients in same category where pain type, location is similar. Thus, data from homologous has a higher correlation and is easier to be classified by machine learning classifier.

Although we have obtained over 80% classification accuracy in the group study for two categories, there are limitations to our methods; first, the MIMIC-III database is a clinical ICU database and not focused on pain level assessment. Therefore, physiological measures data and pain level data weren't recorded regularly. The time interval between two consecutive data points ranges from 5 minutes to more than 1 hour and the timestamps were imprecise. This particularly raises an issue during the data preprocessing as physiological measures data and pain level data may be mismatched if they share the same timestamp but are not recorded at the same time accurately. Besides, the sample number in the MIMIC-III database is limited. When it comes to group study, there are few samples in some categories, such as Endocrinology which contains only 1679 sample points. Therefore, if a larger-scale database is supplied, the pain level assessment accuracy with our proposed method can go further, especially for categories with small sample sizes like Orthopedics and Renal Diseases.

In addition, the group study method can go further. Group study intends to gather information from cogeneric patients and it can go deeper to personalized. A personalized method might focus on one single patient and it is more focused than group study which may lead to further improved classification accuracy. Kächele *et al.* [19] done a similar study on BioVid database and gained better results than other studies on BioVid [9] [30].

To address these limitations, our future work will focus on 1) the personalized study on the MIMIC-III database and 2) set up a new and systematic database which will be conducted on patients rather than healthy individuals. Also, we will measure more modalities in the new database including well-timestamped ECG signal, PPG signal, SpO2, Pupil Size.

In a word, automatic pain study in this paper reveals the possibility of automatic pain level assessment on patients. However, patient-based pain level assessment study is still very scarce at present. Future research and pain databases need to be developed.

4 CONCLUSION

Congeneric clustering-based automatic pain level assessment on ICU patients' data is deployed in this study. We use the MIMIC-III database for automatic pain level assessment field and adopt three new modalities: respiratory rate, SpO2 and pupil size. By combining all these modalities in our method, we examined existing methods that were proposed for healthy individual level of pain prediction in a pilot study stage. The results show poor prediction accuracy when using the model developed for healthy individuals and apply it to patient data. In response we propose a novel congeneric clustering method. Congeneric patients were clustered into the same group based on their diagnosis (from ICD-9 code) and eleven groups were formed including Cardiovascular, Endocrinology, Gastroenterology, Hematology, Infectious diseases, Neurology, Orthopedics, Pulmonary, Renal Diseases, Rheumatology and Others, and train a machine learning model on each cluster. Finally, we obtained the highest accuracy of 82.86%.

REFERENCES

[1] Yong, R. Jasona,*; Mullins, Peter M.b; Bhattacharyya, Neilc Prevalence of chronic pain among adults in the United States, PAIN: April 02, 2021 - Volume - Issue - doi: 10.1097/j.pain.0000000000002291
[2] Gaskin, Darrell J., and Patrick Richard. "The economic costs of pain in the United States." The Journal of Pain 13, no. 8 (2012): 715-724.
[3] Wilkerson, Richard Gentry, Hong K. Kim, Thomas Andrew Windsor, and Darren P. Mareiniss. "The opioid epidemic in the United States." Emergency Medicine Clinics 34, no. 2 (2016): e1-e23.
[4] Vadivelu, Nalini, Alice M. Kai, Vijay Kodumudi, Richard Zhu, and Roberta Hines. "Pain management of patients with substance abuse in the ambulatory setting." Current pain and headache reports 21, no. 2 (2017): 9.
[5] S. M. G. Zwakhalen, J. P. H. Hamers, H. H. Abu-Saad, and M. P. F. Berger, "Pain in elderly people with severe dementia: A systematic review of behavioural pain level assessment tools," BMC Geriatrics, vol. 6, no. 1, p. 3, Jan. 2006.
[6] A. Williams, H. T. O. Davies, and Y. Chadury, "Simple pain rating scales hide complex idiosyncratic meanings," Pain, vol. 85, no. 3, pp. 457–463, Apr. 2000.
[7] Lucey, Patrick, Jeffrey F. Cohn, Kenneth M. Prkachin, Patricia E. Solomon, Sien Chew, and Iain Matthews. "Painful monitoring: Automatic pain monitoring using the UNBC-McMaster shoulder pain expression archive database." Image and Vision Computing 30, no. 3 (2012): 197-205.
[8] Werner, Philipp, Ayoub Al-Hamadi, Robert Niese, Steffen Walter, Sascha Gruss, and Harald C. Traue. "Towards pain monitoring: Facial expression, head pose, a new database, an automatic system and remaining challenges." In Proceedings of the British Machine Vision Conference, pp. 1-13. 2013.
[9] Werner, Philipp, Ayoub Al-Hamadi, Robert Niese, Steffen Walter, Sascha Gruss, and Harald C. Traue. "Automatic pain recognition from video and biomedical signals." In 2014 22nd International Conference on Pattern Recognition, pp. 4582-4587. IEEE, 2014.
[10] Gruss, Sascha, Roi Treister, Philipp Werner, Harald C. Traue, Stephen Crawcour, Adriano Andrade, and Steffen Walter. "Pain intensity recognition rates via biopotential feature patterns with support vector machines." PloS one 10, no. 10 (2015): e0140330.
[11] World Health Organization. "International classifification of diseases—Ninth revision (ICD-9)." Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire 63, no. 45 (1988): 343-344.
[12] Liu, Tian-Yu. "Easyensemble and feature selection for imbalance data sets." In 2009 international joint conference on bioinformatics, systems biology and intelligent computing, pp. 517-520. IEEE, 2009.
[13] Franck, Linda Sturla, Cindy Smith Greenberg, and Bonnie Stevens. "Pain assessment in infants and children." Pediatric Clinics of North America 47, no. 3 (2000): 487-512.
[14] Jose, Anthony D., and D. Collison. "The normal range and determinants of the intrinsic heart rate in man." Cardiovascular research 4, no. 2 (1970): 160-167.
[15] Rodríguez-Molinero, Alejandro, Leire Narvaiza, Jorge Ruiz, and César Gálvez-Barrón. "Normal respiratory rate and peripheral blood oxygen saturation in the elderly population." Journal of the American Geriatrics Society 61, no. 12 (2013): 2238-2240.
[16] Werner, Philipp, Ayoub Al-Hamadi, Sascha Gruss, and Steffen Walter. "Twofold-multimodal pain recognition with the X-ITE pain database." In 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), pp. 290-296. IEEE, 2019.
[17] Gruss, Sascha, Mattis Geiger, Philipp Werner, Oliver Wilhelm, Harald C. Traue, Ayoub Al-Hamadi, and Steffen Walter. "Multi-modal signals for analyzing pain responses to thermal and electrical stimuli." JoVE (Journal of Visualized Experiments) 146 (2019): e59057.
[18] Fatima, Meherwar, and Maruf Pasha. "Survey of machine learning algorithms for disease diagnostic." Journal of Intelligent Learning Systems and Applications 9, no. 01 (2017): 1.
[19] Kächele, Markus, Patrick Thiam, Mohammadreza Amirian, Friedhelm Schwenker, and Günther Palm. "Methods for person-centered continuous pain intensity assessment from bio-physiological channels." IEEE Journal of Selected Topics in Signal Processing 10, no. 5 (2016): 854-864.
[20] Johnson, Alistair EW, Tom J. Pollard, Lu Shen, H. Lehman Li-Wei, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G. Mark. "MIMIC-III, a freely accessible critical care database." Scientific data 3, no. 1 (2016): 1-9.
[21] Campbell, Evan, Angkoon Phinyomark, and Erik Scheme. "Feature extraction and selection for pain recognition using peripheral physiological signals." Frontiers in neuroscience 13 (2019): 437.
[22] Jiang, Mingzhe, Riitta Mieronkoski, Amir M. Rahmani, Nora Hagelberg, Sanna Salanterä, and Pasi Liljeberg. "Ultra-short-term analysis of heart rate variability for real-time acute pain monitoring with wearable electronics." In 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1025-1032. IEEE, 2017.
[23] Naeini, Emad Kasaeyan, Ajan Subramanian, Michael-David Calderon, Kai Zheng, Nikil Dutt, Pasi Liljeberg, Sanna Salantera, Ariana M. Nelson, and Amir M. Rahmani. "Pain Recognition With Electrocardiographic Features in Postoperative Patients: Method Validation Study." Journal of Medical Internet Research 23, no. 5 (2021): e25079.
[24] Erdoğan, Burak, and Hasan Oğul. "Objective Pain Assessment Using Vital Signs." Procedia Computer Science 170 (2020): 947-952.
[25] Werner, Philipp, Daniel Lopez-Martinez, Steffen Walter, Ayoub Al-Hamadi, Sascha Gruss, and Rosalind Picard. "Automatic recognition methods supporting pain assessment: A survey." IEEE Transactions on Affective Computing (2019).
[26] Lopez-Martinez, Daniel, and Rosalind Picard. "Multi-task neural networks for personalized pain recognition from physiological signals." In 2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), pp. 181-184. IEEE, 2017.
[27] Chapman, C. Richard, Shunichi Oka, David H. Bradshaw, Robert C. Jacobson, and Gary W. Donaldson. "Phasic pupil dilation response to noxious stimulation in normal volunteers: relationship to brain evoked potentials and pain report." Psychophysiology 36, no. 1 (1999): 44-52.
[28] Badura, Aleksandra, Aleksandra Masłowska, Andrzej Myśliwiec, and Ewa Piętka. "Multimodal Signal Analysis for Pain Recognition in Physiotherapy Using Wavelet Scattering Transform." Sensors 21, no. 4 (2021): 1311.
[29] Barajas-Montiel, Sandra E., and Carlos A. Reyes-Garcia. "Identifying pain and hunger in infant cry with classifiers ensembles." In International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06), vol. 2, pp. 770-775. IEEE, 2005.
[30] Kächele, Markus, Philipp Werner, Ayoub Al-Hamadi, Günther Palm, Steffen Walter, and Friedhelm Schwenker. "Bio-visual fusion for person-independent recognition of pain intensity." In International Workshop on Multiple Classifier Systems, pp. 220-230. Springer, Cham, 2015.