ABSTRACT

Pain is a highly personal and stressful subjective experience linked to damage to the tissue. At present, the management of pain continues to remain ambiguous and disappointing at hospitals. In particular, the management of patients' postoperative pain has become a major medical and nursing challenge. Hospitals have taken initiatives to measure pain using self-report measures such as the Visual Analogue Scale (VAS) and the Numeric Pain Intensity Scale (NPIS). But these methods are inaccurate and subjective as it depends on the patient's input. Therefore, there is a need for an objective, quantitative method to monitor pain continuously. Thus, this work presents the various data-driven approaches to automatically measure and monitor postoperative patient's pain severity levels continuously.

This work utilizes minimal raw data, i.e., two physiological signal data and one behavioral data, to determine pain. Therefore, this research work reduces the constraints imposed by multimodal signal processing and also helps to establish the field of wearable technologies. The physiological signals used for the study are Electrocardiogram (ECG) and Electro-Dermal activity (EDA), and the behavioral data used for the study are the facial expressions of the individuals. Evidence from several cohort studies has shown that physiological signals such as ECG and EDA signals and the facial expression data of individuals are the best sources of the presence of acute pain in adults (especially in postoperative patients).

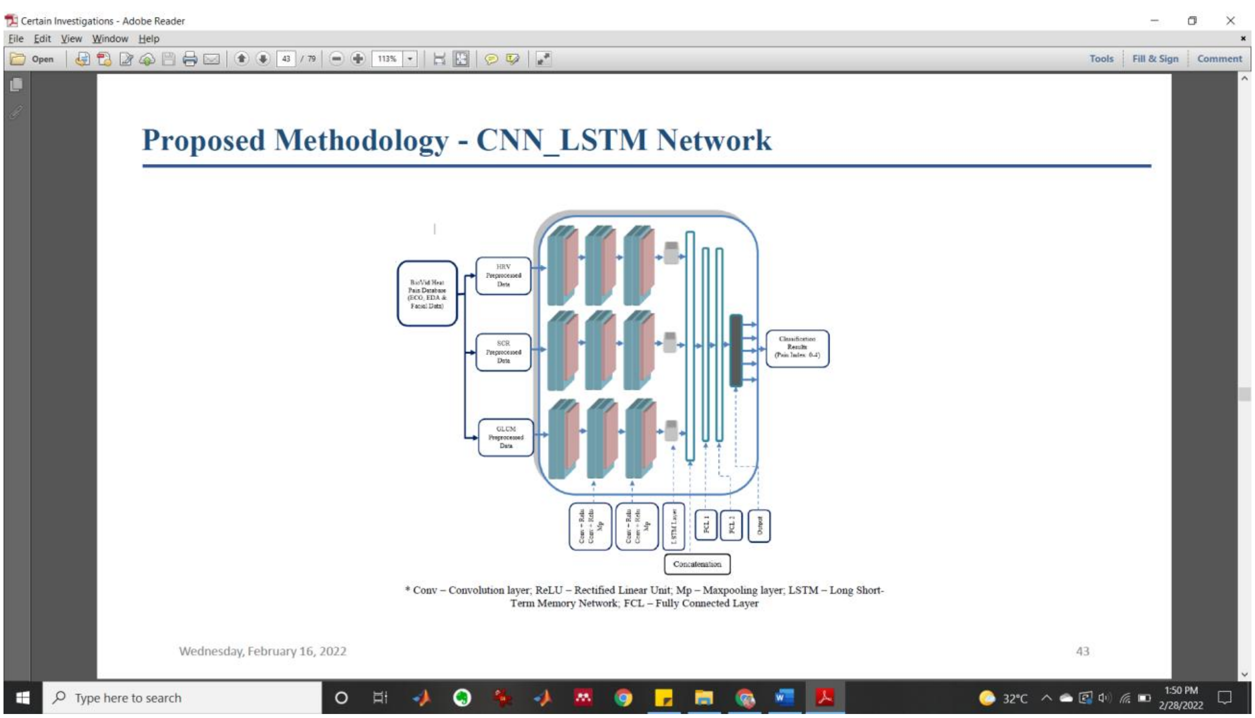
Work 1:

This research work imparts physiological and behavioral data to the different data-driven approaches to evaluate the pain severity levels. The first research work modeled is the supervised ML and DL-based classification models that use the features extracted from the source data such as ECG, EDA,and facial expression data. The notable features used for this study are the use of the Heart Rate Variability (HRV) features of the ECG signal, the phasic component features of the EDA signal, and the Gray Level Co-Occurrence Matrix (GLCM) features of the facial expression data.

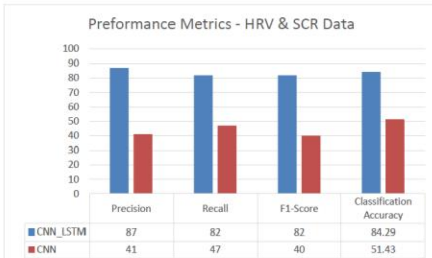
A filter-based method, i.e., one-way ANOVA, is applied to the data to select the best pain-associated features. Thus, the features are selected based on statistically significant values ($P < 0.05$) for the classification. Finally, a classification task implementation helps to classify five different levels of pain (Pain Index: 0-4, namely No pain as BL1, Mild pain as PA1, Moderate pain as PA2, Severe pain as PA3, Intolerable

pain as PA4) using supervised ML algorithms such as Neural Network (NN), Support Vector Machine (SVM), and Random Forest (RF) and DL algorithms like a hybrid Convolutional Neural Network Long Short-Term Memory Network (CNN_LSTM). The algorithm's performance is tested using the following metrics: classification accuracy, recall, precision, f1-score, and confusion matrix. This work utilizes the BioVid Heat Pain database (BVHP DB) and the UNBC-McMaster Shoulder Pain Expression Archive database (UNBC DB).

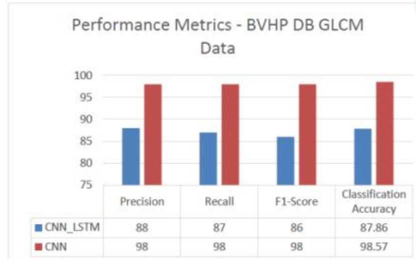
Although the feature-based supervised ML and DL models achieve high accuracy in classifying various pain levels, the hand-engineered features are its main drawback. In healthcare systems, DL algorithms that do not use predefined features provide several benefits. The ability to extract features without requiring medical professionals to comprehend the health issue fully is their most significant benefit. The present work aims to develop models to achieve good classification results for untrimmed continuous physiological data using a feature learning approach. This approach resolves ML challenges such as feature selection and the instances related to the small datasets. This study had grouped into three modules. The first module is to create a unimodal pain recognition system; then, the next module is to create amultimodal system. The final module is to create a gender-based multimodal system. The proposed three modules are tested on the BVHP physiological Database.



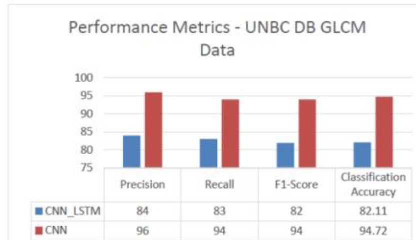
Performance Measure



	Precision	Recall	F1-Score	Classification Accuracy
CNN_LSTM	87	82	82	84.29
CNN	41	47	40	51.43

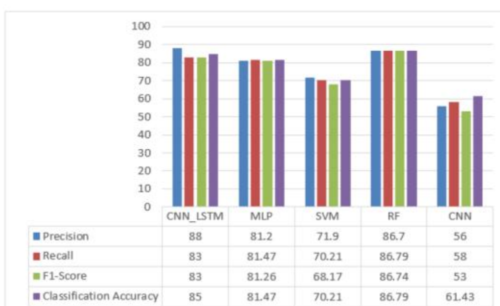


	Precision	Recall	F1-Score	Classification Accuracy
CNN_LSTM	88	87	86	87.86
CNN	98	98	98	98.57



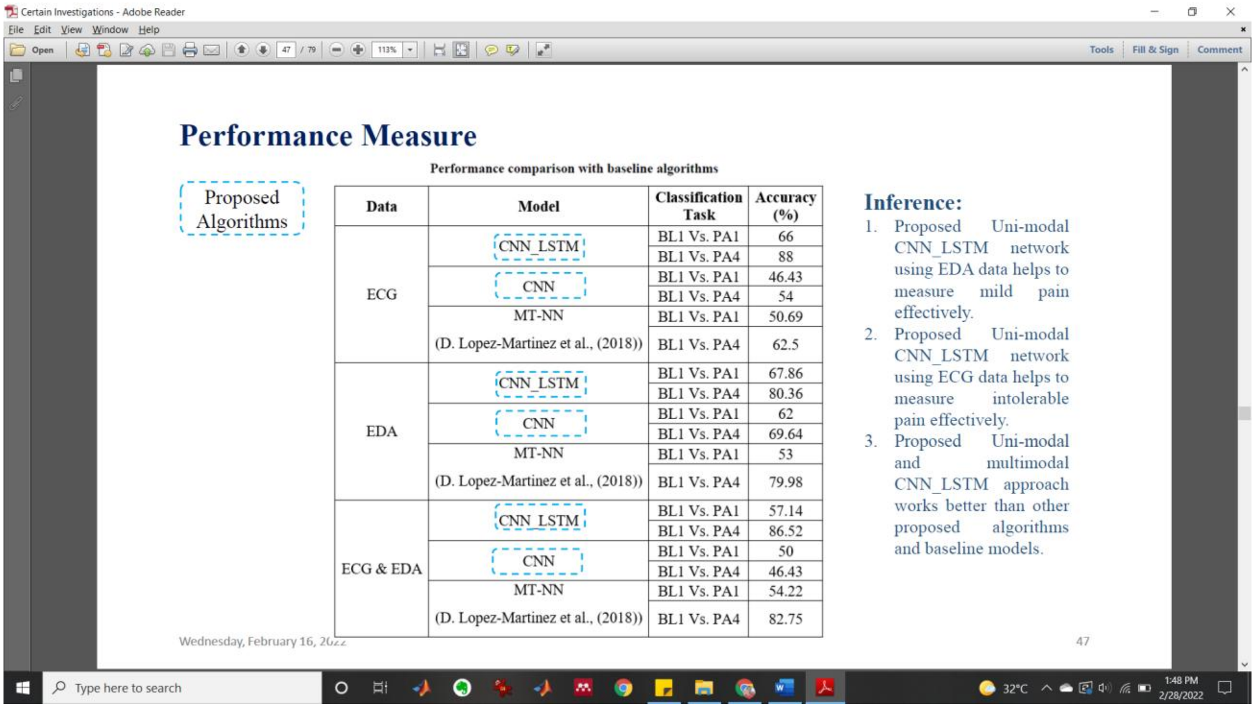
	Precision	Recall	F1-Score	Classification Accuracy
CNN_LSTM	84	83	82	82.11
CNN	96	94	94	94.72

Performance Measure



	CNN_LSTM	MLP	SVM	RF	CNN
Precision	88	81.2	71.9	86.7	56
Recall	83	81.47	70.21	86.79	58
F1-Score	83	81.26	68.17	86.74	53
Classification Accuracy	85	81.47	70.21	86.79	61.43

Performance metrics for the multiclass classification task using HRV, SCR, and GLCM features



Performance Measure

Proposed Algorithms

Performance comparison with baseline algorithms

Data	Model	Classification Task	Accuracy (%)
ECG	CNN_LSTM	BL1 Vs. PA1	66
		BL1 Vs. PA4	88
	CNN	BL1 Vs. PA1	46.43
		BL1 Vs. PA4	54
	MT-NN (D. Lopez-Martinez et al., (2018))	BL1 Vs. PA1	50.69
		BL1 Vs. PA4	62.5
EDA	CNN_LSTM	BL1 Vs. PA1	67.86
		BL1 Vs. PA4	80.36
	CNN	BL1 Vs. PA1	62
		BL1 Vs. PA4	69.64
	MT-NN (D. Lopez-Martinez et al., (2018))	BL1 Vs. PA1	53
		BL1 Vs. PA4	79.98
ECG & EDA	CNN_LSTM	BL1 Vs. PA1	57.14
		BL1 Vs. PA4	86.52
	CNN	BL1 Vs. PA1	50
		BL1 Vs. PA4	46.43
	MT-NN (D. Lopez-Martinez et al., (2018))	BL1 Vs. PA1	54.22
		BL1 Vs. PA4	82.75

Inference:

1. Proposed Uni-modal CNN_LSTM network using EDA data helps to measure mild pain effectively.
2. Proposed Uni-modal CNN_LSTM network using ECG data helps to measure intolerable pain effectively.
3. Proposed Uni-modal and multimodal CNN_LSTM approach works better than other proposed algorithms and baseline models.

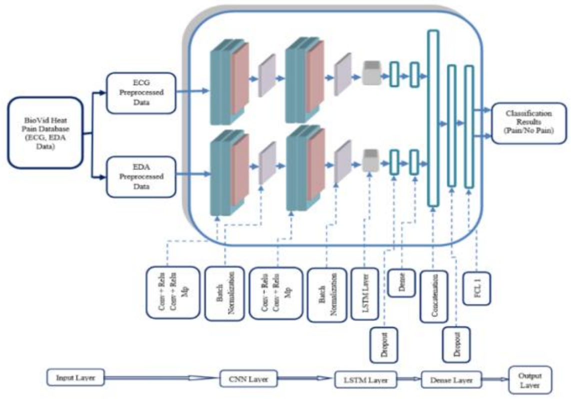
Wednesday, February 16, 2022 47

Work 2:

This research work is further extended by developing a pain recognition system using a multimodal approach. Many researchers believe that simply determining the presence of pain is too rough for estimating pain in practice. Therefore, to measure the actual level of pain intensity (Pain Index: 0-4) in each patient, a unimodal and multimodal classification approach is implemented. This work divides into three parts. The first part is establishing a multimodal system using physiological data (i.e., ECG & EDA); the second module is to develop a gender-based multimodal system using physiological data (i.e., ECG & EDA). And the third work is to establish a unimodal system using facial expression data. The performance of all the research work gets tested on the BVHP DB.

All the above proposed models of this study had achieved good classification results as that of the state-of-the-art by producing considerable improvement in the classification accuracy.

A hybrid CNN-LSTM network



* Conv – Convolution layer; ReLU – Rectified Linear Unit; Mp – Maxpooling layer; LSTM – Long Short-Term Memory Network; FCL – Fully Connected Layer

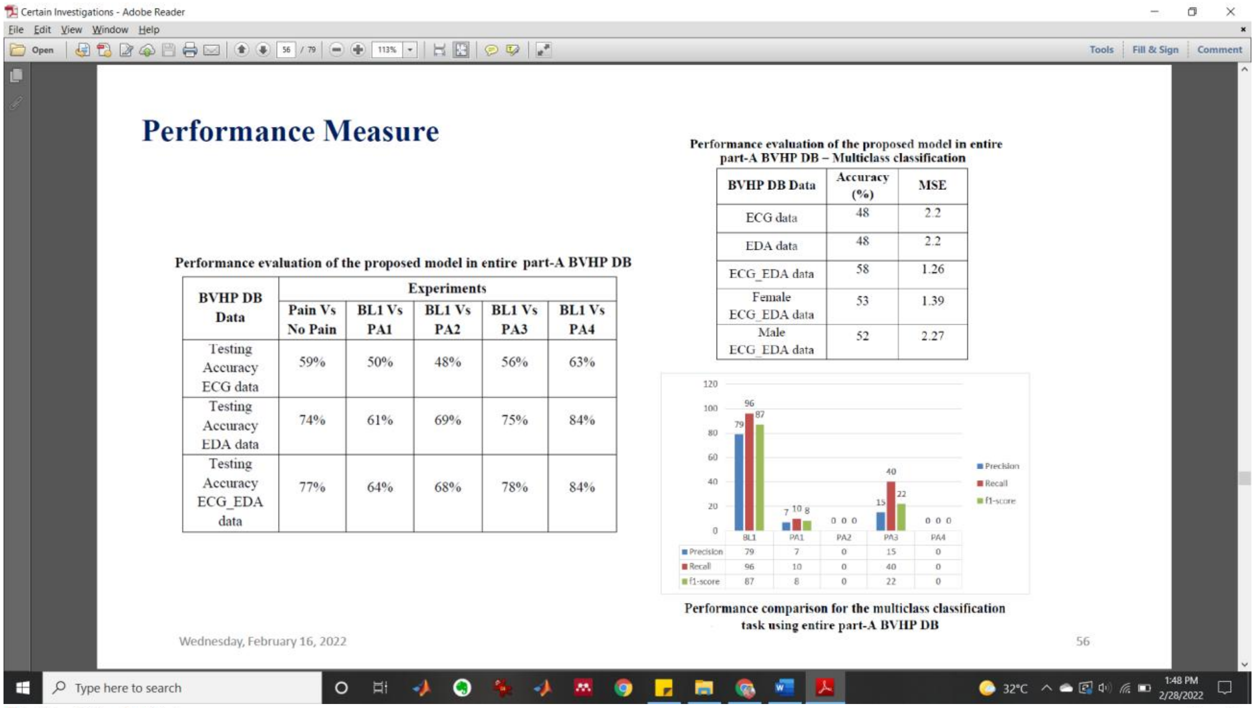
Performance Measure

Performance evaluation of the proposed model in gender- based BVHP DB

BVHP DB Data	Experiments				
	Pain Vs No Pain	BL1 Vs PA1	BL1 Vs PA2	BL1 Vs PA3	BL1 Vs PA4
Testing Accuracy Male ECG data	52 %	49 %	48 %	48 %	56 %
Testing Accuracy Female ECG data	57 %	48 %	53 %	56 %	64 %
Testing Accuracy Male EDA data	74 %	62 %	66 %	74 %	82 %
Testing Accuracy Female EDA data	74 %	61 %	68 %	74 %	89 %
Testing Accuracy Male ECG_EDA data	73 %	62 %	68 %	76 %	84 %
Testing Accuracy Female ECG_EDA data	73 %	59 %	66 %	76 %	87 %

Inference:

1. Measuring intolerable pain in female is comparatively easier than male
2. EDA serves as the best source of measuring acute pain in female as well as in male
3. Multimodal CNN_LSTM approach helps to detect the presence of pain effectively and also helps to measure intolerable pain.



Performance Measure

Performance evaluation of the proposed model in entire part-A BVHP DB

BVHP DB Data	Experiments				
	Pain Vs No Pain	BL1 Vs PA1	BL1 Vs PA2	BL1 Vs PA3	BL1 Vs PA4
Testing Accuracy ECG data	59%	50%	48%	56%	63%
Testing Accuracy EDA data	74%	61%	69%	75%	84%
Testing Accuracy ECG_EDA data	77%	64%	68%	78%	84%

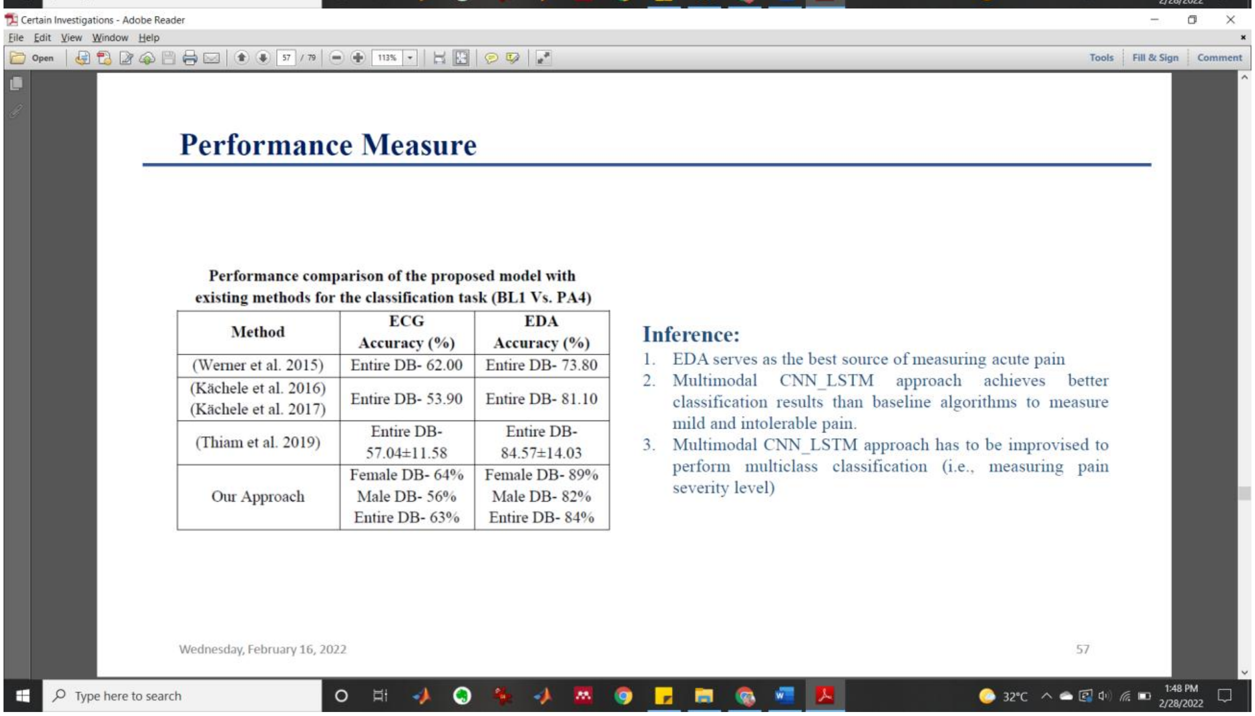
Performance evaluation of the proposed model in entire part-A BVHP DB – Multiclass classification

BVHP DB Data	Accuracy (%)	MSE
ECG data	48	2.2
EDA data	48	2.2
ECG_EDA data	58	1.26
Female ECG_EDA data	53	1.39
Male ECG_EDA data	52	2.27

	BL1	PA1	PA2	PA3	PA4
Precision	79	7	0	15	0
Recall	96	10	0	40	0
f1-score	87	8	0	22	0

Performance comparison for the multiclass classification task using entire part-A BVHP DB

Wednesday, February 16, 2022 56



Performance Measure

Performance comparison of the proposed model with existing methods for the classification task (BL1 Vs. PA4)

Method	ECG Accuracy (%)	EDA Accuracy (%)
(Werner et al. 2015)	Entire DB- 62.00	Entire DB- 73.80
(Kächele et al. 2016) (Kächele et al. 2017)	Entire DB- 53.90	Entire DB- 81.10
(Thiam et al. 2019)	Entire DB- 57.04±11.58	Entire DB- 84.57±14.03
Our Approach	Female DB- 64% Male DB- 56% Entire DB- 63%	Female DB- 89% Male DB- 82% Entire DB- 84%

Inference:

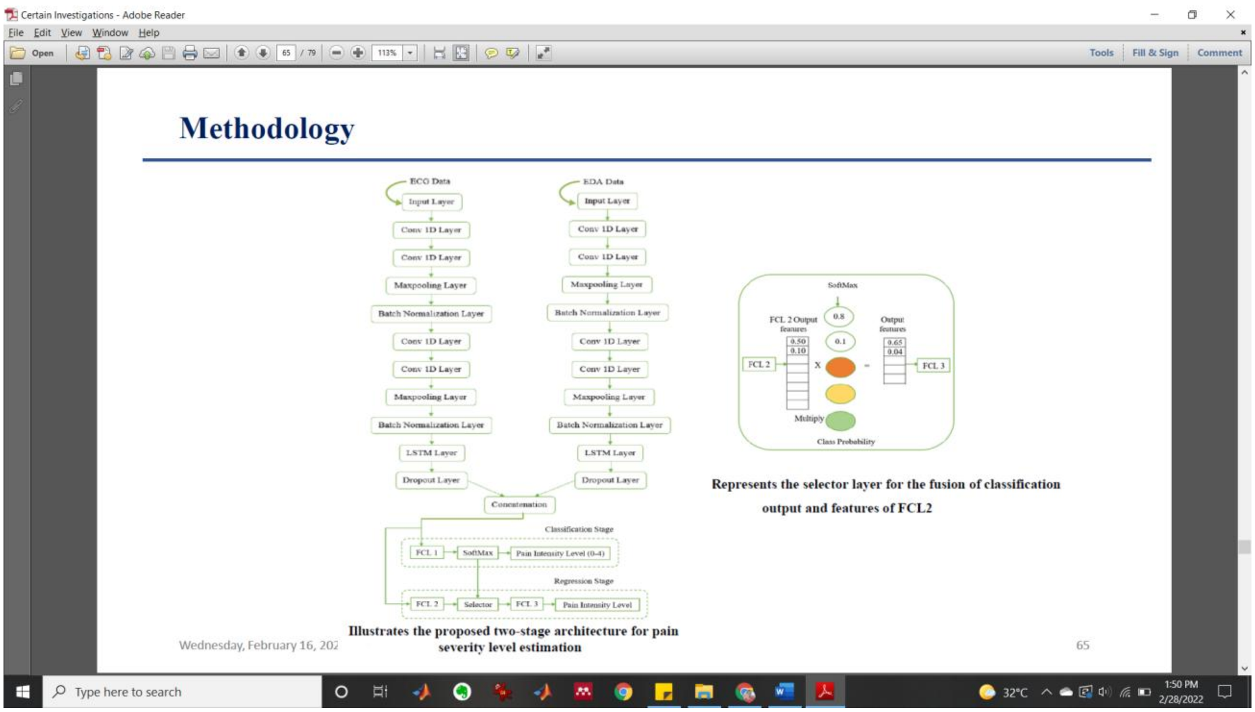
1. EDA serves as the best source of measuring acute pain
2. Multimodal CNN_LSTM approach achieves better classification results than baseline algorithms to measure mild and intolerable pain.
3. Multimodal CNN_LSTM approach has to be improvised to perform multiclass classification (i.e., measuring pain severity level)

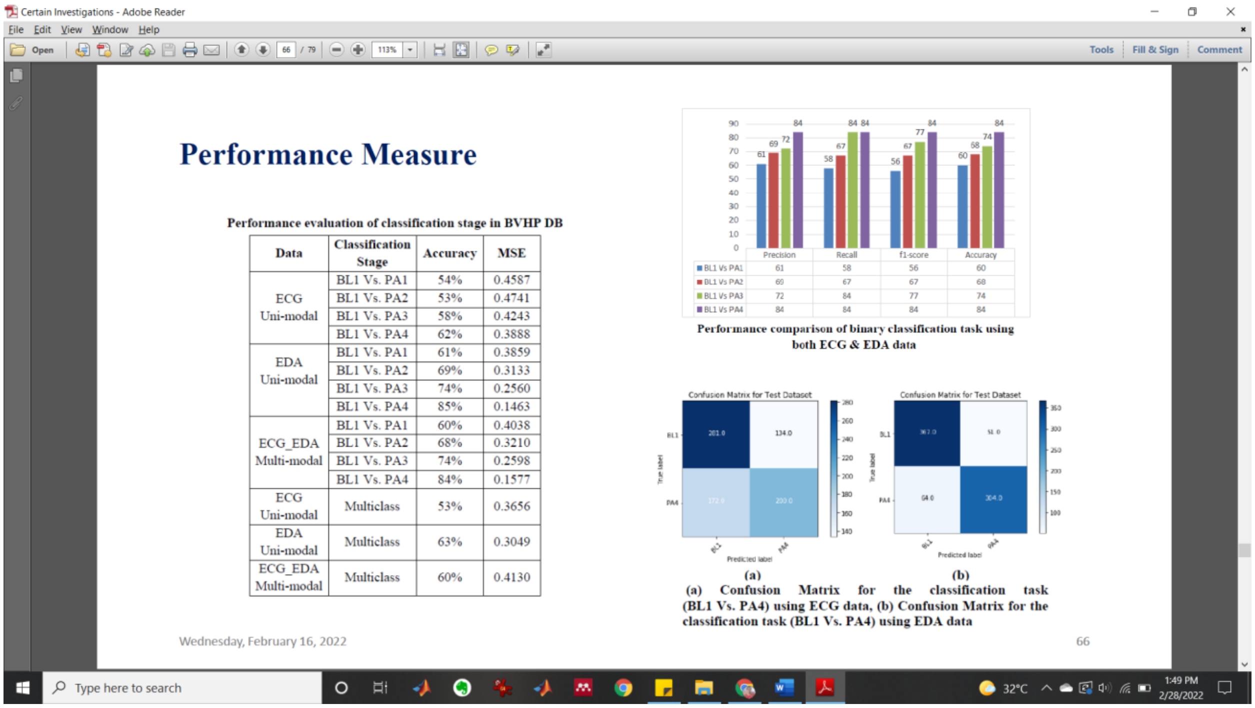
Wednesday, February 16, 2022 57

Work 3:

The next novel research contribution is developing the two-stage approach to estimate pain level using physiological signals (i.e., ECG & EDA) and the class separation method. An algorithm's capability to identify pain from indeterminate length streaming sequences is expected to be a complex problem. This task gets achieved by using a DL-based joint classification and regression framework. This method is the improvised model of the previous work, which had accomplished in two stages. The first stage is to classify the given pre-processed physiological signals into five different pain levels. Then, the second stage is to establish a regression model based on the training samples from each class of the first stage.

This system outperforms the state-of-the-art and the previously proposed methods by producing a substantial drop in the MAE, RMSE values. Besides, the proposed system is independent of the subject characteristics; the forecast for new data gives reduced error irrespective of the trained data. The use of patient features such as gender characteristics and physiological data improves the system's predictive ability.





Certain Investigations - Adobe Reader

Performance Measure

Performance evaluation of classification stage in BVHP DB

Data	Classification Stage	Accuracy	MSE
ECG Uni-modal	BL1 Vs. PA1	54%	0.4587
	BL1 Vs. PA2	53%	0.4741
	BL1 Vs. PA3	58%	0.4243
	BL1 Vs. PA4	62%	0.3888
EDA Uni-modal	BL1 Vs. PA1	61%	0.3859
	BL1 Vs. PA2	69%	0.3133
	BL1 Vs. PA3	74%	0.2560
	BL1 Vs. PA4	85%	0.1463
ECG_EDA Multi-modal	BL1 Vs. PA1	60%	0.4038
	BL1 Vs. PA2	68%	0.3210
	BL1 Vs. PA3	74%	0.2598
	BL1 Vs. PA4	84%	0.1577
ECG Uni-modal	Multiclass	53%	0.3656
EDA Uni-modal	Multiclass	63%	0.3049
ECG_EDA Multi-modal	Multiclass	60%	0.4130

	Precision	Recall	f1-score	Accuracy
BL1 Vs PA1	61	58	56	60
BL1 Vs PA2	69	67	67	68
BL1 Vs PA3	72	84	77	74
BL1 Vs PA4	84	84	84	84

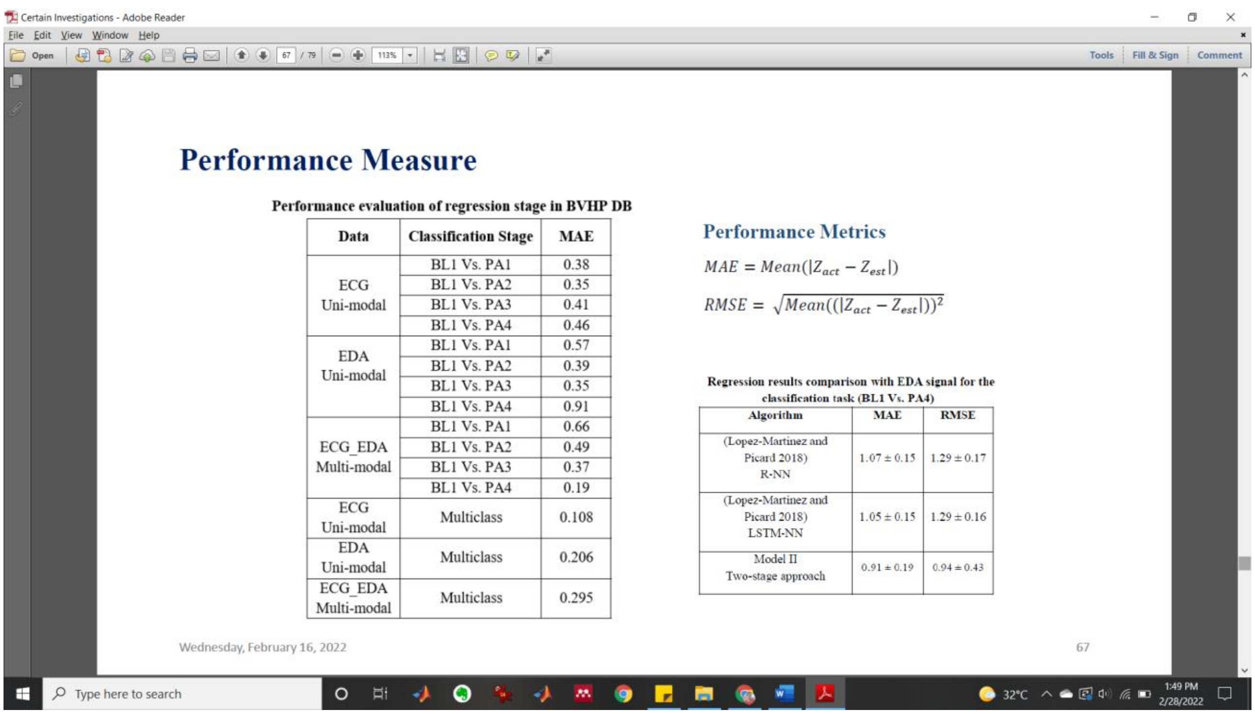
Performance comparison of binary classification task using both ECG & EDA data

(a) (b)

(a) Confusion Matrix for the classification task (BL1 Vs. PA4) using ECG data, (b) Confusion Matrix for the classification task (BL1 Vs. PA4) using EDA data

Wednesday, February 16, 2022

66



Certain Investigations - Adobe Reader

Performance Measure

Performance evaluation of regression stage in BVHP DB

Data	Classification Stage	MAE
ECG Uni-modal	BL1 Vs. PA1	0.38
	BL1 Vs. PA2	0.35
	BL1 Vs. PA3	0.41
	BL1 Vs. PA4	0.46
EDA Uni-modal	BL1 Vs. PA1	0.57
	BL1 Vs. PA2	0.39
	BL1 Vs. PA3	0.35
	BL1 Vs. PA4	0.91
ECG_EDA Multi-modal	BL1 Vs. PA1	0.66
	BL1 Vs. PA2	0.49
	BL1 Vs. PA3	0.37
	BL1 Vs. PA4	0.19
ECG Uni-modal	Multiclass	0.108
EDA Uni-modal	Multiclass	0.206
ECG_EDA Multi-modal	Multiclass	0.295

Performance Metrics

$$MAE = Mean(|Z_{act} - Z_{est}|)$$

$$RMSE = \sqrt{Mean((|Z_{act} - Z_{est}|))^2}$$

Regression results comparison with EDA signal for the classification task (BL1 Vs. PA4)

Algorithm	MAE	RMSE
(Lopez-Martinez and Picard 2018) R-NN	1.07 ± 0.15	1.29 ± 0.17
(Lopez-Martinez and Picard 2018) LSTM-NN	1.05 ± 0.15	1.29 ± 0.16
Model II Two-stage approach	0.91 ± 0.19	0.94 ± 0.43

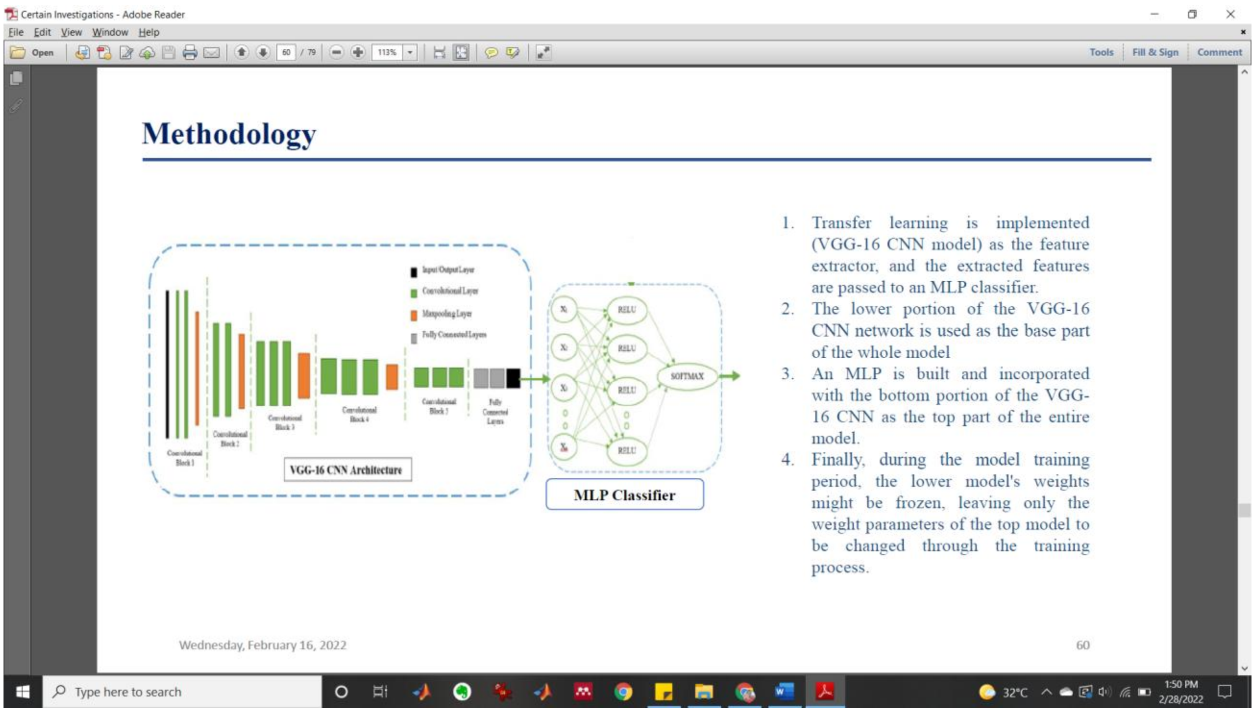
Wednesday, February 16, 2022

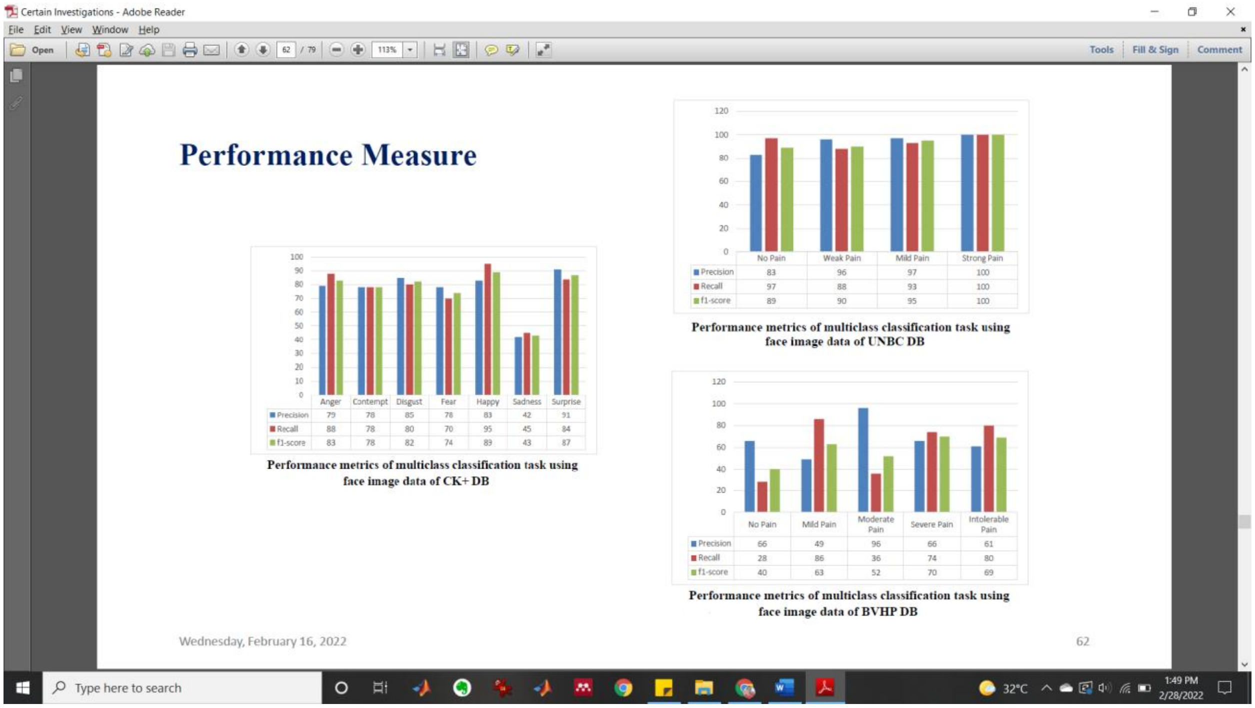
67

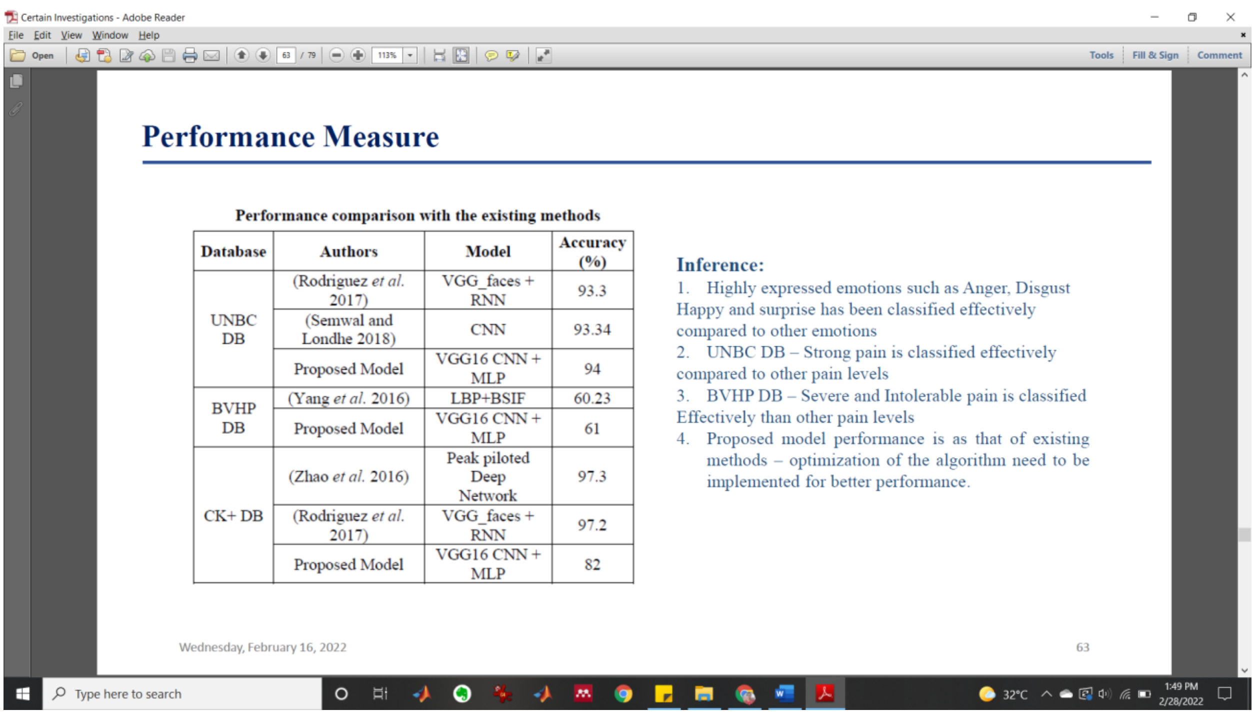
Work 4:

Several studies have shown that pain treatment is primarily dependent on gender. Therefore, gender-based models have been developed for all the proposed research works. Emotional experience plays an important role when measuring pain. This relationship could reinforce or prevent the correct measurement of the severity of pain.

Therefore, to effectively treat acute/chronic pain, subjects' emotional states should be simultaneously measured. In this study, emotions are measured using the same networks which have been proposed to determine pain using facial expression data. This algorithm is tested with the facial expression data of the CK+ dataset to determine various emotions (Emotions: Neutral, Sadness, Surprise, Happiness, Fear, Anger, Contempt, and Disgust).







Performance Measure

Performance comparison with the existing methods

Database	Authors	Model	Accuracy (%)
UNBC DB	(Rodriguez *et al.* 2017)	VGG_faces + RNN	93.3
	(Semwal and Londhe 2018)	CNN	93.34
	Proposed Model	VGG16 CNN + MLP	94
BVHP DB	(Yang *et al.* 2016)	LBP+BSIF	60.23
	Proposed Model	VGG16 CNN + MLP	61
CK+ DB	(Zhao *et al.* 2016)	Peak piloted Deep Network	97.3
	(Rodriguez *et al.* 2017)	VGG_faces + RNN	97.2
	Proposed Model	VGG16 CNN + MLP	82

Inference:

1. Highly expressed emotions such as Anger, Disgust Happy and surprise has been classified effectively compared to other emotions
2. UNBC DB – Strong pain is classified effectively compared to other pain levels
3. BVHP DB – Severe and Intolerable pain is classified Effectively than other pain levels
4. Proposed model performance is as that of existing methods – optimization of the algorithm need to be implemented for better performance.

Wednesday, February 16, 2022 63

Work 5:

The real-time data gathered from the PSGIMSR, Coimbatore, is tested with all the proposed research works. The real-time data acquired for the study are annotated with VAS measure to evaluate the pain severity levels.

Certain Investigations - Adobe Reader

Performance Measure

Performance results of the two-stage method on the real-time data (postoperative patients)

Method	Accuracy (%)	Classification Error	Test Samples
Unimodal - ECG	67	0.33	10
Unimodal - EDA	33	0.66	10
Multimodal – ECG & EDA	63	0.36	10

Performance results of the two-stage method on the real-time data (only normal subjects)

Method	Accuracy (%)	Classification Error	Test Samples
Unimodal - ECG	50	0.5	10
Unimodal - EDA	70	0.3	10
Multimodal – ECG & EDA	40	0.6	10

Inference:

1. MAE and RMSE values are comparatively low when compared to the baseline algorithms
2. For real-time data, the proposed algorithm performance is better for ECG data