

# Classify Elderly Pain Severity from Open-Source Automatically Video Clip Facial Action Units Analysis. A Study from the Integrated Pain Artificial Intelligence Network (I-PAIN) Data Repository.

Patama Gomutbutra (✉ [Patthama.g@cmu.ac.th](mailto:Patthama.g@cmu.ac.th))

Chiang Mai University

Adisak Kittisares

Chiang Mai University

Atigorn Sanguansri

The College of Arts, Media and Technology, Chiang Mai University

Noppon Choosri

The College of Arts, Media and Technology, Chiang Mai University

Passakorn Sawaddiruk

Chiang Mai University

Puriwat Fakfum

Chiang Mai University

Peerasak Lertrakarnnon

Chiang Mai University

Sompob Saralamba

Mahidol University

---

## Research Article

**Keywords:** Facial Action Coding System, Chronic pain, Elderly, Dementia, Asian

**Posted Date:** November 1st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1010233/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** It is increasingly interesting to monitor pain severity in elderly individuals by applying machine learning models. In previous studies, OpenFace© - a well-known automated facial analysis algorithm, was used to detect facial action units (FAUs) that initially need long hours of human coding. However, OpenFace© developed from the dataset that dominant young Caucasians who were illicit pain in the lab. Therefore, this study aims to evaluate the accuracy and feasibility of the model using data from OpenFace© to classify pain severity in elderly Asian patients in clinical settings.

**Methods:** Data from 255 Thai individuals with chronic pain were collected at Chiang Mai Medical School Hospital. The phone camera recorded faces for 10 seconds at a 1-meter distance briefly after the patients provided self-rating pain severity. For those unable to self-rate, the video was recorded just after the move, which illicit pain. The trained assistant rated each video clip for the Pain Assessment in Advanced Dementia (PAINAD). The classification of pain severity was mild, moderate, or severe. OpenFace© process video clip into 18 FAUs. Five classification models were used, including logistic regression, multilayer perception, naïve Bayes, decision tree, k-nearest neighbors (KNN), and support vector machine (SVM).

**Results:** Among the models that included only FAU described in the literature (FAUs 4, 6, 7, 9, 10, 25, 26, 27 and 45), multilayer perception yielded the highest accuracy of 50%. Among the machine learning selection features, the SVM model for FAU 1, 2, 4, 7, 9, 10, 12, 20, 25, 45, and gender yielded the best accuracy of 58%.

**Conclusion:** Our open-source automatic video clip facial action unit analysis experiment was not robust for classifying elderly pain. Retraining facial action unit detection algorithms, enhancing frame selection strategies, and adding pain-related functions may improve the accuracy and feasibility of the model.

## Introduction

Pain severity is crucial information for pain management decision-making. However, the accuracy in the assessment of older patients is still challenging. While self-reported pain ratings are the gold standard, elderly patients have limited cognitive and physical functions, making it more difficult to assess pain. In addition, the COVID-19 situation and shortage of caregivers make hospital visits more complicated. Therefore, home-based pain management becomes essential for these patients. Integration of machine learning that can provide automated and continuous pain monitoring at home could be the ideal solution.

However, there are potentially complicated objective pain measurements, such as fMRI and heart rate variability. These are not feasible for the actual clinical setting. Facial expression is an accessible, informative feature associated with pain severity, but several limitations inhibited real-life application. The original facial activity measurement, the Facial Action Coding System (FACS), [1] needs manual coding, which is both timely and costly. There were attempts to develop an automated facial analysis algorithm to overcome this limitation. One of the well-known algorithms is OpenFace©[2]; anyway, it was

trained by a dataset dominated by a young, healthy Caucasian person. The widely used UNBC McMaster Pain Dataset has conducted illicit pain in shoulder pain patients while being videotaped in a laboratory setting [3]. Therefore, 'wrinkles' in the elderly [4], cognitive impairment[5], and gender- or ethnic-related fairness of skin[6] might cause the representative bias of the model.

Furthermore, in the regular clinical environment, chronic pain patients usually have persistent pain accompanied by spontaneous pain without physical pain stimuli. While taking everyday life, the facial angle and lighting are hard to set in perfect position as a laboratory setting. These factors may impact the performance of the model.

OpenFace© [2] is the deep neuronal network algorithm that reports 18 out of the original 46 AUs. It cannot discriminate AU24-27, which uses the same muscle as AU23, and AU43 (eye closure) from AU45. The OpenFace outputs values for each AU, indicating the algorithm's confidence in the AU being present. The per-frame label representing an integer value from 0 to 5 is estimated from either a classification or a regression. The data from computer vision detect FAUs into time series of continuous values. Previous studies applied two methods for achieving point estimation from time series for each AU. The first method uses mean measurements for different AUs[7], while the second method uses time series by calculating the area under the curve of the pulse of FAUs[8]. The accuracy of OpenFace© compared to human FACS coding is 90% for constrained images and 80% for everyday life images. However, the validity of OpenFace© in detecting a face in elderly individuals and dementia is still problematic. One study using manual code FACs found that OpenFace has a precision of AU4 of only 54% and AU43 of 70.4%[5]. Meanwhile, the Delaware database project applies OpenFace© in an ole population younger than 30 years with a higher proportion of nonCaucasian ethnicity and found that the precision of AU4 is 98% for AU4 and 73% for AU45[9]

Defining pain-related facial action units from computer vision automated facial coding is the most challenging area. The benchmark study on pain-related facial action units by Prkachin et al. uses picture frames of 129 people with shoulder pain and rates facial pain during illicit acute pain by motion[3]. They adopted Ekman's coding and utilized four certified human coders who provided 1-5 levels of pain intensity (1 = trace to 5 = maximum) for each picture frame. They summarized six AUs significantly associated with pain, including brow lowering (AU4), orbital tightening (AU6 and AU7), levator contraction (AU9 and AU10), and eye closure (AU43). The prediction model was proposed by the sum of AU4, AU6, or AU7, whichever is higher, AU9 or AU10, whichever is higher, and AU43 for each AU (Pain = AU4 + AU6 or AU7 + AU9 or AU10 + AU43)[7]. A recent systematic review summarized the AUs consistently reported as related to pain; these comprised AU4, 6, 7, 9, 10, 25, 26, 27, and 43[10].

With these overall knowledge gaps, it is still questionable whether the machine learning model from the currently available automated facial analysis can be used as decision support. Therefore, this study aims to evaluate the accuracy of the model using data from OpenFace© to classify pain severity in Asian elderly individuals who receive chronic pain care.

# Methods

## Study Design

This is a prospective registry building and analysis of facial expressions in elderly Thai individuals with chronic pain. The research project was approved by the Chiang Mai University's Institutional Review Board (CMU IRB no 05429), and the informed consent for participation was obtained from the patients themselves or the designated caregivers for both study participation and publication of images in an online open-access publication. All research was performed in accordance with the IRB regulations.

## Study population

Cases were collected from the pain clinic, internal medicine ward, orthopedic ward, and nursing home institute of Chiang Mai University Hospital during May 2018-December 2019. Patients aged more than 55 years old have a history of chronic (more than three months) and ongoing pain during the assessment. Patients or caregivers able to communicate in the Thai language were included in this study.

## Data collection

Research assistants enrolled participants on weekdays between 9 AM and 4 PM. The data collection form consisted of demographic questions. The cognitive status of each patient was evaluated using the Minimal Mental Status (MMSE) Thai 2002[11]. The facial expression data were recorded by phone camera for 10 seconds at a 1-meter distance. A patient who could communicate was asked to give pain information just before starting the video clip recording. The pain information included pain site, quality, and self-rating severity by the visual analog scale and Wong-Baker Face Scale. We recorded video noncommunicable pain during bed bathing, moving, or measuring blood pressure while assuming illicit pain behavior. Both video clips of communicable and uncommunicable patient research assistance-trained PAINAD were assigned to rate the video clip. The Pain Assessment In Dementia (PAINAD) is a simple score based on five pain behavior observation domains, including breathing, negative vocalization, facial expression, body language, and consolability[12]

## Importing and preparing the dataset

The videos from the mobile camera feed were then manually trimmed to remove noise (e.g., frames where patients were talking). OpenFace© automatically coded the video clips into 18 facial action coding system-based AUs. The data are kept as an I-PAIN data repository for further research.

As the target class variable of this study, pain severity was defined by a self-rating, WBS. Mild if 0 and 2, moderate for 4 and 6, and severe for 8 and 10. For uncommunicable patients, their pain levels were defined using the PAINAD rating by a trained research assistant. One-hot encoding [13] was applied as the premodeling process. Mild for 0-1, moderate 2-4, and severe if five or more.[14]. The facial action unit-time series data of each action unit that represented the facial movement over time from each patient generated by OpenFace© were transformed into two forms: 1) the average movement intensity [15] and 2) the area under the curve (AUC) around the maximum peak [8]. The AUC of each action unit was

calculated using the data from 22 frames (0.03 seconds per frame) around the maximum peak. These two data sets were then explored if the action units were related to the pain severity levels and could be used as the attributes to classify the pain intensity. **Figure 1** outlines the entire data processing.

Figure 1. Flow to identify the pain intensity-related facial action unit

Demographic data, such as age, gender, and dementia, were identified as missing values. One case was deleted due to a lack of information about MMSE. Age was categorized into four classes: less than 60, 61 to 70, 71 to 80, and more than 80. Gender was coded 0 for female and 1 for male. Dementia was classified by MMSE cut point, score 18 or below for those who only completed a lower education and 22 for those who completed higher education.

## **Descriptive analysis**

R studio version 1.3 [16] and MATLAB version 7.0 [17] were used for statistical analysis and creation of figures. Demographic data were summarized into percentages and means, medians, and standard deviations. The correlation analysis between each AU grouping based on a correlation of 0.3 compared to the group of facial anatomical movements. The correlation of each FAU to pain severity was explored using one-way ANOVA with a defined statistical p-value < 0.05.

## **Classification models and models evaluation**

WEKA software [18] was used for the data mining process. A total of 255 samples were split into training and test datasets at 70:30 (180:75). The synthetic minority oversampling technique (SMOTE) was applied for the imbalanced data. Tenfold cross-validation was applied for the attribute selection process. The models were built using five commonly used classification machine learning techniques, including generalized linear model, multilayer perception, which is a subtype of the artificial neural network, J48 decision tree, naïve Bayes, k-nearest neighbors (KNN) – in this study, the optimized K number = 10, and a sequential minimal optimization support vector machine. In the model evaluation, ten iterations of 10-fold cross-validation were used in each data set, and the overall classification accuracy percentage was compared between the models.

## **Results**

### ***Descriptive analysis***

A total of 255 Thai communicable elderly participants were assessed. The mean age was 67.72 years (SD 10.93, range 60-93). About half (55%) of the patients were male. Among the groups, 47.1% were diagnosed with cancer and 23% with dementia. Approximately 10% were bed-bound. Moderate pain was classified as 55.4%, severe pain as 24.4%, and mild pain as 20.2%. The most common site for pain is the back (33.8%). The most common quality of pain was lancinating or 'sharp shooting' pain (40.2%). The demographic details of the patients are portrayed in **Table 1**.

Table 1

Demographic data (N=255)

Characteristic	N (SD or %)
Age (mean (SD))	67.72 (10.39)
Sex = Male (%)	140 (54.9)
Education (%)	
No or less than four years	27 (10.8)
Less than a college degree	112 (45.0)
College degree or above	110 (44.2)
Religious (%)	
Buddhism	250 (98.0)
Christian	2 ( 0.8)
Islam	3 ( 1.2)
Physical status (%)	
Walking	162 (63.5)
Wheelchair	66 (25.9)
Bed bounded	27 (10.6)
Underlying disease	
Cancer (%)	114 (47.1)
Diabetes (%)	54 (21.5)
Depression (%)	31 (12.4)
Dementia (%)	60 (23.5)
MMSE score(%)	
0-10 : severe dementia	17 ( 6.9)
11-18 : moderate dementia	37(15.0)
19-25 : mild dementia	101 (41.1)
26-30 : no dementia	91 (37.0)
Pain severity (%)	
mild	52 (20.2)
moderate	141 (55.4)
severe	57 (22.4)

Site of pain	
Head (%)	16 ( 4.1)
Face (%)	14 ( 3.6)
Chest (%)	30 (14.9)
Abdomen (%)	44 (17.4)
Back (%)	80 (33.8)
Hip (%)	11 ( 5.1)
knee (%)	31 (11.8)
Foot (%)	24 (9.8)
Quality of pain	
Burning (%)	63 (23.8)
Troubling (%)	73 (28.4)
Lancinating (%)	112 (41.2)
Dull (%)	18 ( 8.2)
Sore (%)	40 (14.5)
Spastic (%)	52 (20.7)
Paresthesia/difficult to explain (%)	19 ( 7.7)

The network plot of each AU was examined. The results indicated that they were not independent, and when grouped by a correlation of 0.3 or more, they appeared close to the grouping of facial anatomical movement, as shown in **Figure 2**. The closely related AUs, called coexisting AUs, are grouped. It should be noted that the area under the curve approach provided a more accurate grouping than the average approach (e.g., AU1, AU2, and AU5 are anatomically closely related, and they were well acknowledged as coexisting from a previous study[19]).

The box plot comparing FAUs from the average activity approach and the area under the curve approach classified pain as mild, moderate, and severe, as shown in Fig. 3. ANOVA yielded significant differences for the average activities of AU4 ( $p=0.04$ ), AU7 ( $p=0.005$ ), AU10 ( $p=0.03$ ), and AU25 ( $p=0.005$ ) from the results of the area under the curve approach, which identified AU17 ( $p=0.07$ ) and AU23 ( $p=0.0045$ ).

Since a dependent relationship was discovered between each FAU, it may imply that regression may not suit pain severity prediction. However, there was a relationship between the average activity of AU4, AU7, AU10, and AU25 and the area under the curve of AU17 and 23 as well as dementia and sex and pain severity. Therefore, these features were included in the classification model.

## ***Classification models and models evaluation***

There are two sources of features to build pain severity classification models: pain-related AUs consistently found in previous studies (AUs 4, 6, 7, 9, 10, 25, 26, 27, and 45) and selection by machine learning. Features selected using each machine learning method are demonstrated in **Table 2**, and the accuracy of each machine learning model is illustrated in **Table 3**. It can be concluded that machine-selected features provide the best accuracy. The SVM model for average activities of AU1, 2, 4, 7, 9, 10, 12, 20, 25, 45, and gender yielded an accuracy of 58%. The second most accurate model is the KNN model, which measures the area under the curve of AU1, AU2, AU6, AU20, and females, which provided an accuracy of 56.41%. Among the features selected from pain-related FAU in previous studies, the most accurate models are multilayer perception (50%) and KNN (44.87%).

Table 2

The features selected by each machine learning method

Selected features	Average	AUC
Classification		
Logistic regression	Gender, Dementia ,AU01, AU02 AU06,AU09,10 AU10,AU12 AU20, AU23, AU25	Dementia, AU01,AU02,AU04,AU06 AU07, AU09,AU10,AU12,AU14 AU17, AU20
Multilayer perceptron	Gender,Dementia,AU01,AU04,AU05 AU06,AU07,AU09,AU10,AU12 AU14, AU15, AU17,AU20, AU23 AU25,AU26,AU45	Gender,AU05, AU07,AU10,AU12 AU14, AU23, AU25, AU26, AU45
Decision tree	AU01,AU04,AU05,AU06, AU07,AU12,AU14	Gender, AU4,AU12,AU15,AU45
Naïve Bayes	Gender, AU9, AU12	Gender, AU4,AU12,AU15,AU45
KNN (K =10)	Gender, AU7,AU9,AU20, AU25	Gender, AU1, AU2, AU6, AU20
SVM	Gender,AU01,AU02,AU04,AU07 AU09,AU10,AU12,AU20,AU25,AU45	Gender, AU20, AU45

Table 3

The accuracy of pain intensity prediction models

Feature selection	From previous studies[10]		Machine learning	
	Average <sup>1</sup>	AUC <sup>2</sup>	Average <sup>5</sup>	AUC <sup>6</sup>
Classification				
Logistic regression	30.76	38.46	41.02%	15.38
Multilayer perceptron	50	23.07	39.74%	33.33
Decision tree	38.46	32.05	50.00%	44.87
Naïve Bayes	21.79	25.64	55.40%	24.36
KNN (K=10)	28.2	44.87	23.96%	56.41
SVM	30.76	29.48	58.78%	14.12
Note:				
t Accuracy of classification into three levels of pain severity: mild, moderate, severe				
1 AU4,AU6,AU7,AU9,AU10, AU 25, AU26, AU27 and AU45				
2 AUC of AU6,AU7,AU9,AU10, AU 25, AU26, AU27 and AU45				
5 Difference based on classification technique see in Table 2				
6. Difference based on classification technique see in Table 2				

The confusion matrix between pain classified mild, moderate, severe by either WBS or PAINAD (actual severity) with the SVM and average value from Open Face© model found that the misclassification of moderate pain is higher than mild or severe pain. The ROC areas of mild, moderate, and severe pain were 0.514, 0.408, and 0.496, and the F statistics were 0.651, 0.333, and 0.560, respectively, as described in **Table 4.**

Table 4

Confusional matrix between actual severity and model classification for the task classify pain severity.

Model classify	Mild	Moderate	Severe	ROC area	F measure
Actual severity					
mild	85	3	10	0.496	0.651
moderate	46	23	30	0.408	0.333
severe	13	32	54	0.514	0.560

## Discussion

The main findings for chronic pain in elderly Asian individuals are that the classification pain severity (mild, moderate, severe) model developed from OpenFace shows no robustness. The best model is SVM for average activities of AU1, 2, 4, 7, 9, 10, 12, 20, 25, 45, and gender yielded the best accuracy, at 58%. There are three possible explanations and solutions to improve the accuracy of the model. First, the limitation of OpenFace© in detecting facial action units in painful elderly individuals. Second, the correlation between facial expression and perceived pain severity in chronic pain patients may not be very high. Third, the point estimation method may not be optimized.

We defined the pain severity-related facial action unit in two ways. The AUs described 'consistently pain-related from the systematic review[10] and machine learning selection. We consistently overlap AUs from these two methods, including AU4, AU7, AU10, and AU45, already described in human coding studies[3]. However, it is noted that the machine learning technique also chooses AU1, AU2, and AU20, which are defined as fear expressions, and AU9, 10, 17 have been described as related to the disgusting face. This may support the 'total pain' theory that chronic pain comprises complex emotional components. [26]

This machine learning also showed the contribution of gender and dementia but not increasing age to the model. This finding is consistent with previous studies that indicated that gender influences pain severity expression in terms of more expressive[8] and possibly fairer skin, which influences model accuracy in facial landmark detection[6]. Previous studies have shown that elderly individuals with dementia tend to show more activity in the mouth area than in the upper part of the face[21].

The confusion matrix indicates that higher misclassification in moderate pain than obvious mild or severe pain. This nature of the pain classification model has been described in A previous study on the accuracy of OpenFace© to classify pain severity from UNBC McMaster.[15]. Therefore, it may not be feasible to adopt the current automated pain severity classification model for critical decision-making, such as adjusting the dosage of opioid analgesics. On the other hand, it may be more suitable to augment gross triage tasks, such as supporting evidence of self-rating severity.

## Strengths and limitations

To our current knowledge, this study is one of few extensive facial recognition studies conducted in the Asian elderly population. This study is also conducted in a natural setting in which stakeholders benefit from the solution. The information from our research may fill the representative bias of the current model. It also directly answers that currently available open-source facial analysis software and classification models still need reasonable improvement accuracy to classify elderly pain.

The major limitation of this study is that we did not yet have the validity of the facial landmark from OpenFace© compared to the standard human coding. This was caused by limited personal resources that were available for training at that time. This defect may be questionable if the reported AUs in this study are the same anatomical landmark as those described in the standard FACs.

The second limitation is the strategy to estimate value from the time-event series produced by OpenFace®. From a previous study, the video clip of UNBC McMaster has 80% off frames and shows no pain.[22] Therefore, the average method may influence 'no pain' frames vary person by person. The area under the curve approach, in which the activity also correlates well with facial muscle anatomical movement. may mitigate this effect. Unfortunately, our experiment shows no advantage of this approach. The minor limitation is that the data collection process lags time after the patient's self-report and starting video. Although it is brief, this may decrease the association between the patient-rated and the face in the video clip.

## **Implication and suggestion for further research**

The next step for this project is to collaborate with a computer engineer to improve the algorithm. The trained human coders will validate the facial landmarks of the new algorithm. We also create a collecting data application that silently captures patients' faces while sliding the pain scale and may apply the appearance action model to select a 'possible pain' frame. We may focus on the most consistent pain-related AUs, such as AU4, AU7, AU10, and AU45, to develop real-time sensing and predict extreme pain severity (e.g., no and severe pain) to check if the patient pain self-rating is valid. This will help gather more information along with clinical usability. The training data collection set examined the clinical outcome of using an automated pain severity monitoring platform to support decision-making regarding pain medication.

## **Conclusion**

Our open-source automatic video clip facial action unit analysis experiment was not robust to classifying elderly pain. Retraining facial action unit algorithms, enhancing frame selection strategy, and adding pain-related functions may improve the accuracy and feasibility of the model.

## **Declarations**

## **Data accessibility**

The CSV file can be downloaded at <https://w2.med.cmu.ac.th/agingcare/indexen.html>

## **Acknowledgments**

Thank you, Ms.Napassakorn Sanpaw, for the data collection process. This work was supported by a grant received from The Thai Society of Neurology, 2019. The authors have no conflicting financial interests.

## **References**

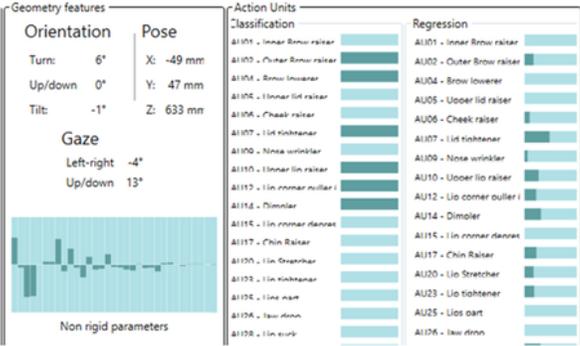
1. Facial Action Coding System. In: Paul Ekman Group [Internet]. [cited 28 Jun 2020]. Available: <https://www.paulekman.com/facial-action-coding-system/>
2. Lötsch J, Utsch A. Machine learning in pain research. *Pain*. 2018;159: 623–630. doi:10.1097/j.pain.0000000000001118
3. Prkachin KM, Solomon PE. The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain: *Pain*. 2008;139: 267–274. doi:10.1016/j.pain.2008.04.010
4. Kunz M, Seuss D, Hassan T, Garbas JU, Siebers M, Schmid U, et al. Problems of video-based pain detection in patients with dementia: a road map to an interdisciplinary solution. *BMC Geriatr*. 2017;17: 33, s12877-017-0427–2. doi:10.1186/s12877-017-0427-2
5. Taati B, Zhao S, Ashraf AB, Asgarian A, Browne ME, Prkachin KM, et al. Algorithmic Bias in Clinical Populations—Evaluating and Improving Facial Analysis Technology in Older Adults With Dementia. *IEEE Access*. 2019;7: 25527–25534. doi:10.1109/ACCESS.2019.2900022
6. Buolamwini J, Gebru T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: Friedler SA, Wilson C, editors. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. New York, NY, USA: PMLR; 2018. pp. 77–91. Available: <http://proceedings.mlr.press/v81/buolamwini18a.html>
7. Lucey P, Cohn J, Lucey S, Matthews I, Sridharan S, Prkachin KM. Automatically Detecting Pain Using Facial Actions. *Int Conf Affect Comput Intell Interact Workshop Proc ACII Conf*. 2009;2009: 1–8. doi:10.1109/ACII.2009.5349321
8. Haines N, Southward MW, Cheavens JS, Beauchaine T, Ahn W-Y. Using computer-vision and machine learning to automate facial coding of positive and negative affect intensity. Hinojosa JA, editor. *PLOS ONE*. 2019;14: e0211735. doi:10.1371/journal.pone.0211735
9. Mende-Siedlecki P, Qu-Lee J, Lin J, Drain A, Goharзад A. The Delaware Pain Database: a set of painful expressions and corresponding norming data. *PAIN Rep*. 2020;5: e853. doi:10.1097/PR9.0000000000000853
10. Kunz M, Meixner D, Lautenbacher S. Facial muscle movements encoding pain—a systematic review: *PAIN*. 2019;160: 535–549. doi:10.1097/j.pain.0000000000001424
11. Boonkerd P, Assantachai P, Senanarong W. *Clinical practice guideline for dementia (in Thai) Guideline for Dementia*. Bangkok: Neuroscience Institute; 2003.
12. Warden V, Hurley AC, Volicer L. Development and psychometric evaluation of the Pain Assessment in Advanced Dementia (PAINAD) scale. *J Am Med Dir Assoc*. 2003;4: 9–15. doi:10.1097/01.JAM.0000043422.31640.F7
13. Ramasubramanian K, Moolayil J. *Applied supervised learning with R: use machine learning libraries of R to build models that solve business problems and predict future trends*. Birmingham: Packt; 2019.
14. Gomutbutra P, Sawasdirak P, Chankrajang S, Kittisars A, Lertrakarnnon P. The Correlation between Subjective Self-pain Rating, Heart Rate Variability and Objective Facial Expression in Thai Elders With and Without Dementia. *Thai J Neurol*. 2020;36: 10–18.

15. Sikka K. Facial Expression Analysis for Estimating Pain in Clinical Settings. Proceedings of the 16th International Conference on Multimodal Interaction. Istanbul Turkey: ACM; 2014. pp. 349–353. doi:10.1145/2663204.2666282
16. R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
17. MATLAB. version 7.10.0 (R2010a). Natick, Massachusetts: The MathWorks Inc.; 2010.
18. Eibe Frank, Mark A. Hall, and Ian H. The WEKA Workbench. Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques.” 4th ed. Morgan Kaufmann,; 2016.
19. Peng G, Wang S. Weakly Supervised Facial Action Unit Recognition Through Adversarial Training. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT: IEEE; 2018. pp. 2188–2196. doi:10.1109/CVPR.2018.00233
20. Corrigendum: Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. Psychol Sci Public Interest. 2019;20: 165–166. doi:10.1177/1529100619889954
21. Lautenbacher S, Kunz M. Facial Pain Expression in Dementia: A Review of the Experimental and Clinical Evidence. Curr Alzheimer Res. 2017;14: 501–505. doi:10.2174/1567205013666160603010455
22. Lucey P, Cohn JF, Matthews I, Lucey S, Sridharan S, Howlett J, et al. Automatically Detecting Pain in Video Through Facial Action Units. IEEE Trans Syst Man Cybern Part B Cybern. 2011;41: 664–674. doi:10.1109/TSMCB.2010.2082525

## Figures



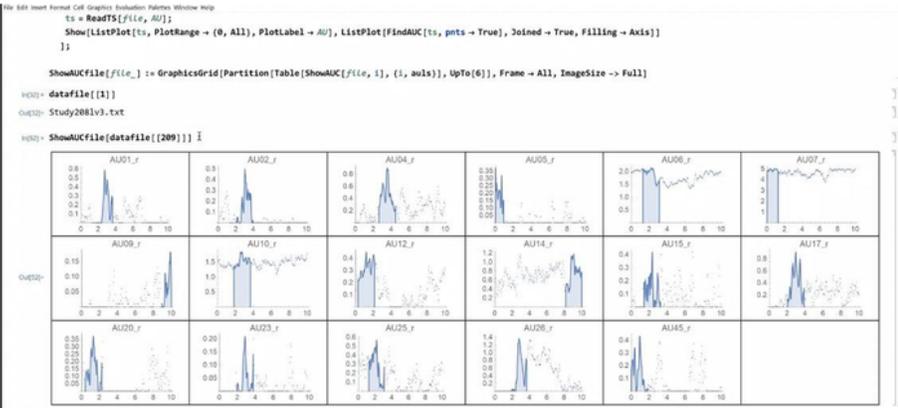
The phone camera video recorded patients for 10 seconds at 1 meter distance.



The video clips were automated coding by OpenFace© into 18 Facial action coding system based AUs

ts	AU1	AU2	AU4	AU6	AU7	AU9	AU10	AU12	AU14	AU15	AU17	AU23	AU25	AU18
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0

The OpenFace© outputs values for each AU indicating the algorithm's confidence in the AU being present. The per-frame label representing an integer value from 0 to 5 estimated from either a classification or a regression. The rationale is that the lower the intensity is, the harder classifying



The Area Under the Curve (AUC) of 22 seconds before and after the peak of each AU evidence time-series was determined as a representative –‘point estimation’ of facial action unit movement

Figure 1

Flow to identify the pain intensity-related facial action unit

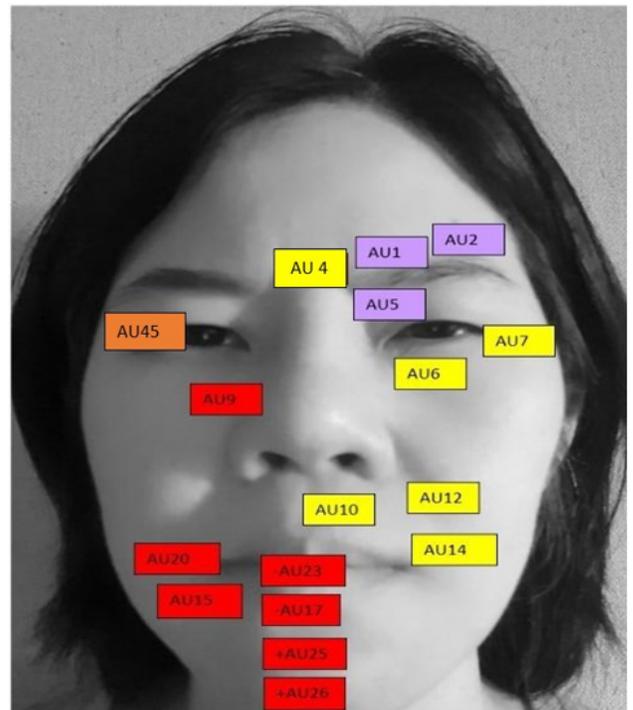
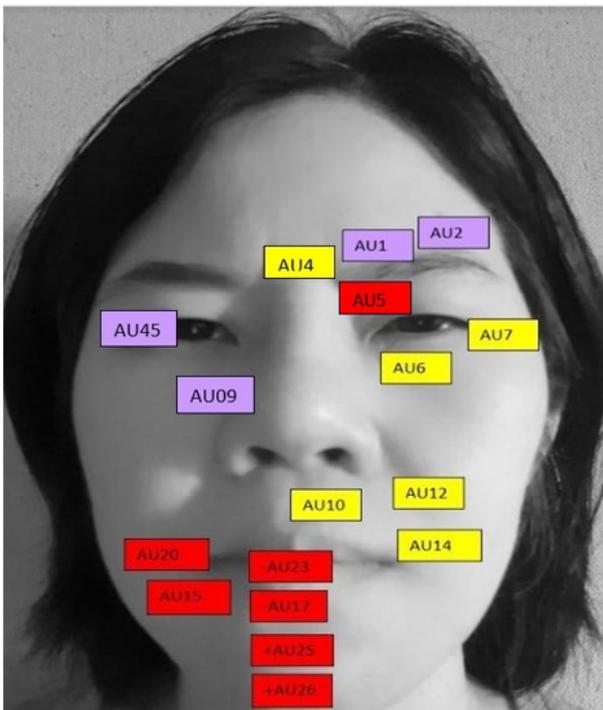
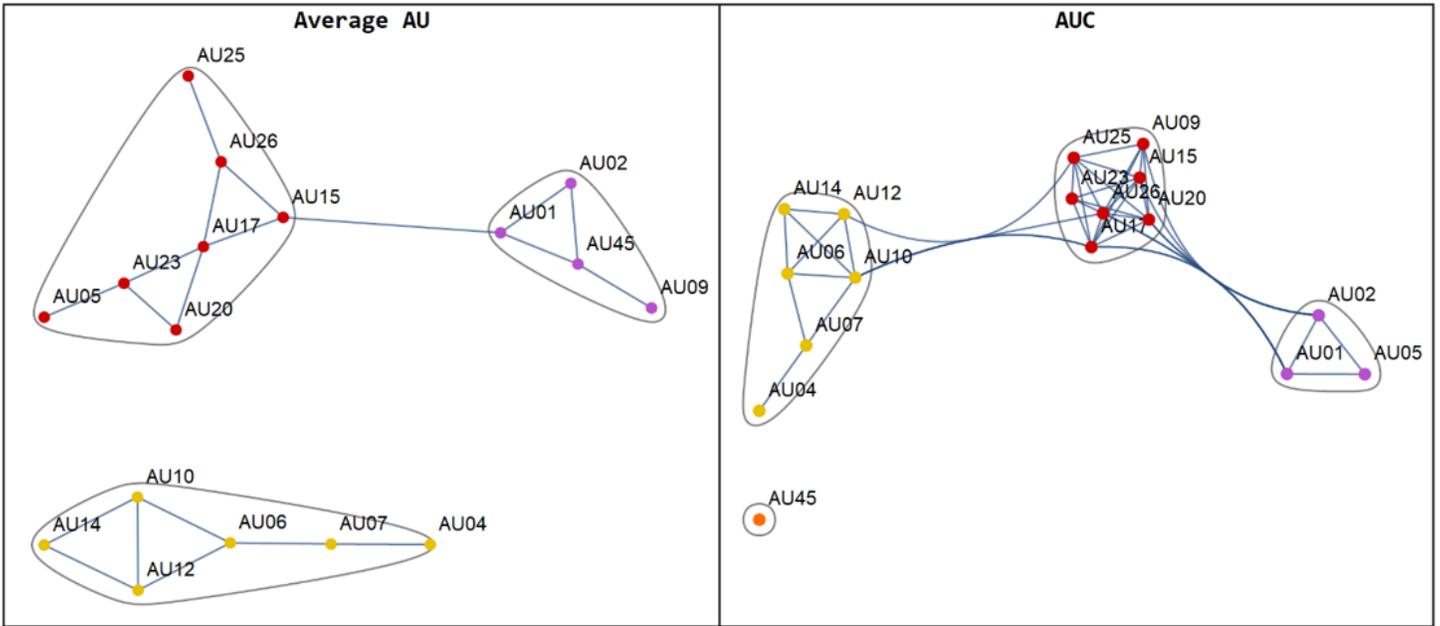


Figure 2

The network plot of the AUs from the average AU (left hand) and the area under the curve: AUC (right hand).



**Figure 3**

The boxplot shows the distribution of intensity for each FAUS categorized into none or mild, moderate, and severe pain.