Automated Subjective Perception of a Driver's Pain Level Based on Their Facial Expression

F. Hadi[1], O. Fukuda[2], W. LYeoh[3], H. Okumura[4], Y. Rodiah[5], Herlina[6], A. Prasetyo[7]
Graduate School of Science and Engineering, Saga University, Saga, Japan[1, 2, 3, 4]
Electrical Engineering Department, University of Bengkulu, Bengkulu, Indonesia[5, 6]
Faculty of Medicine, Diponegoro University, Semarang, Indonesia[7]

***Abstract*—One factor that has a positive correlation with the risk of traffic accidents is the pain experienced by drivers. This pain is sometimes expressed facially by the driver and can be subjectively perceived by others. By observing the facial expression of drivers, it can estimate the pain experienced at that point in time and intervene to prevent some accidents. A method to automatically estimate the pain level expressed by a driver using their facial expression will be proposed in this study. The model is trained by a convolution neural network based on a public dataset of facial expressions at various pain levels. This model is then used to automatically classify the pain level perceived using only the facial expressions of drivers. The result of the automated classification is then compared to ratings of subjective feelings of the driver's pain evaluated by a medical doctor. The experiment results showed that the model classified the pain level expressed facially by the drivers matched that of the classification by the medical doctor at a rate of 80%.**

Keywords—Pain; driver; convolution neural network; facial expression

I. Introduction

Neck and back pain are part of chronic pain that can cause discomfort, be annoying, and interfere with a person's ability to concentrate. If a person is driving and experiencing pain, it can cause collisions [1]. In fact, pain in drivers is one of the main contributing factors to traffic accidents [2]. Furthermore, pain impairs a person's ability for quick thinking when operating a vehicle [3]. Certain medications can also impair cognitive performance and cause pain [4]. Driving over extended periods of time increases the chance of pain, and psychological elements linked to pain, such as stress and anxiety, can also influence driver behaviors and raise the risk of accidents in addition to physical discomfort [5].

The definition of pain, according to the International Association for the Study of Pain (IASP), is an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage[6]. Pain is generally classified into two main types: acute pain dan chronic pain. Acute pain typically triggered by injury or disease, while chronic pain persists beyond the usual healing period and may not have a clear underlying cause[7]. In addition, the relationship between chronic pain and psychological disorders is complex and often bidirectional. For example, individuals with chronic neck pain are at higher risk for mood and anxiety disorders. This suggests that chronic pain may lead to psychological distress, which in turn may worsen pain perception[8].

Assessment of pain is a crucial component of pain management since it enables the determination of the intensity and consequences of pain on the patient's quality of life. Multiple instruments, including the Visual Analogue Scale (VAS) and the Numeric Rating Scale (NRS), are frequently employed to quantify pain intensity by relying on self-reported information provided by patients [9]. VAS was popularized by Aitken and Zealley in the late 1960s, who focused its function on psychological assessment [10]. They showed that VAS could effectively measure subjective feelings, which laid the foundation for its subsequent application in clinical pain measurement. In its development, VAS continued to be validated through several follow-up studies [11]. Moreover, the VAS has been used into more extensive research projects, including those involving cancer patients [12], emergency rooms [13], and measuring pain levels in fibromyalgia patients correlating pain intensity with functional disability and psychological symptoms [14].

The VAS measurement method is generally using a straight line, usually 10 cm long, with the starting point representing "no pain" and the end point indicating "maximum imaginable pain"[15]. One of the main limitations of VAS is that it is highly dependent on subjective assessment of the patient. Each individual may interpret the scale differently. This can be influenced by various psychological factors, such as anxiety and mood when the assessment is carried out. The results will be inconsistent and therefore do not reflect the actual pain[16]. Another critical limitation is that it may not provide adequate accuracy especially over a narrow scale range resulting in reduced sensitivity to detect small changes in pain levels[17].

Several studies on facial expression-based face detection have been developed. One of the leading methods is the Facial Action Coding System (FACS), which categorizes facial movements into certain Action Units (AU). This method has been used to distinguish between real and fake pain expressions in children undergoing dental treatment[18]. However, the reliability of FACS in assessing pain in various populations is still a concern, especially since this technique requires trained personnel to code facial expressions accurately [19]. The combination of facial electromyography (EMG) with electroencephalography (EEG) has also been used to explore the relationship between brain activity and facial expressions when experiencing pain. Although the results of the study showed a correlation between facial muscle movements and brain electrical activity during pain stimuli, this system is considered more complex and variability in the result[20]. In addition, the role of observers is often distorted in interpreting facial expressions due to the influence of previous exposure to pain assessment[21]. This suggests that the context in which facial expressions are viewed can significantly affect pain recognition, indicating the need for standardized training for health care providers to reduce bias.

Since the use of Visual Analog Scale for pain assessment that has been widely used in various studies still has limitations, the use of image processing techniques and the application of machine learning in classifying features is expected to provide better results. Standardization by experts can also avoid subjectivity of assessment so that it can be applied in a wider field, including pain detection in drivers to reduce traffic accidents.

This study aims to automatically estimate the pain level expressed through a driver's facial expression. This may be a viable non-invasive approach to detecting whether a driver is experiencing pain. Additionally, a model capable of detecting a wider range of pain levels would be more beneficial than one that merely differentiates between two conditions.

II. Methodology

This research is divided into two main stages. The first stage is a classification model trained using a public dataset of facial expressions shown at different levels of pain. For the second stage, the model is used to estimate the level of pain expressed facially by the driver. Figure 1 illustrates the stages of this research process.

A. *You Only Look Once (YOLO) for Pain Classification*

The You Only Look Once (YOLO) is a popular detection algorithm developed by Redmon J, in 2012[22]. The YOLO algorithm process is to divide the facial image into S x S meshes. Each grid is responsible for predicting the target where the actual box will fall in the center of the grid. The total bounding box is generated from the meshes. Each bounding box contains five parameters: Target center point coordinates, target width and height dimensions (x, y, w, h), and confidence. The S x S edge predicts the category probability of the target on that edge. The prediction bounding box confidence and category probability are then multiplied to obtain the category score for each prediction box[23].

In this study, we use YOLOv5 which offers important benefits in speed and computational efficiency[24], [25] lightweight architecture[26], and capability to effectively handle dynamic and diverse facial expressions [27]. Therefore, it is very appropriate for application in real-time scenarios as it enables rapid detection of driver pain, thereby preventing traffic accidents.

YOLOv5 uses a single-stage detection methodology, where the image is processed in a single iteration across the entire network with a convolutional neural network (CNN) approach that simultaneously predicts bounding boxes and class probabilities [28]. It consists of backbone, neck, and head segment components. The main function of the backbone is to extract features, which are then aggregated by the neck to provide predictions at multiple scales. Meanwhile, the head produces the final detection findings [29].

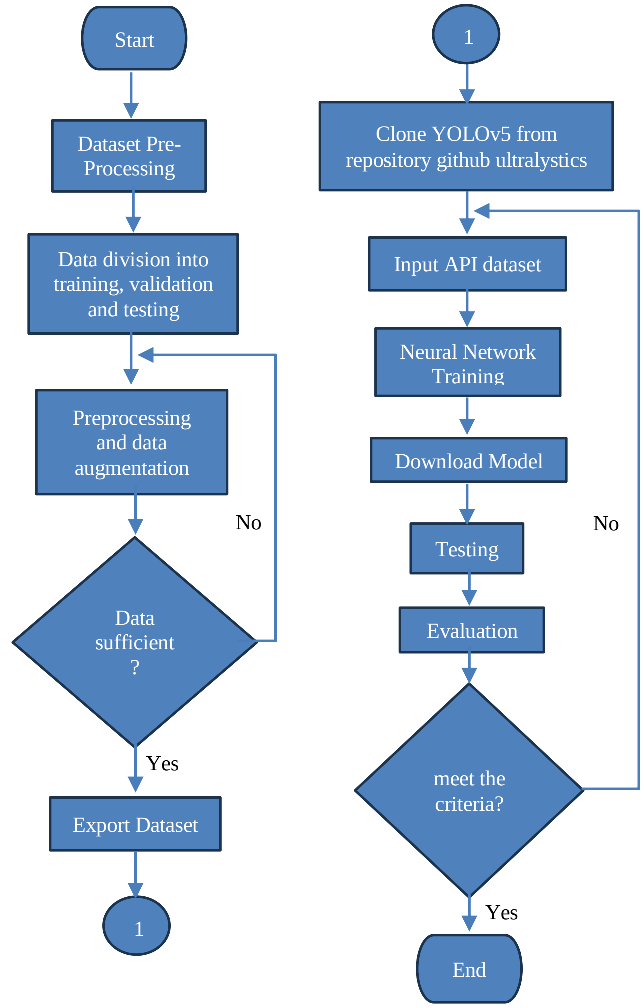


Fig. 1. Research workflow.

A significant improvement in YOLOv5 is the incorporation of anchor boxes, which are pre-defined bounding box shapes that improve the model's ability to reliably estimate object positions. This methodology uses a combination of methods, including data augmentation, to improve the robustness of the model by artificially increasing the size of the training dataset [30].

B. *Datasets Pre-Processing*

The dataset used in this study is from the Multimodal Intensity Pain (MIntPAIN) database[1] with 3079 images divided into 11 classes (0 - 10). Each class will be divided into three classifications based on its class range, namely mild, moderate, and severe. Fig. 2 shows an example dataset [31] with its VAS values. The image is intentionally blurred due to agreements with the dataset provider and privacy reason. The data pre-processing stage is conducted by annotating the area or pixels in the image as a region or area in the bounding box that will be used for model calculations as training or validation. In the YOLO method, annotation is carried out starting by drawing a

[1] (http://www.vap.aau.dk/mintpaon-database

bounding box on each object in the image which then stores the description of the bounding box or object class in a text file database containing the class, x coordinates, y coordinates, width, and height respectively. Data in each VAS value class is divided into 70% for training, 20% for validation, and 10% for testing. The next stage is to carry out the auto orient and resize process. Auto orient aims to adjust the position of the object in the image to ensure that the main object in the image is in the right position. While the resize, the process is to change the image size to the same, 640 x 640 pixels. The purpose of this stage is to equalize the image size because the data obtained can have variations in orientation, both in portrait and landscape formats. By resizing the images, it can ensure that all images have uniform dimensions for data processing and its use in model training.

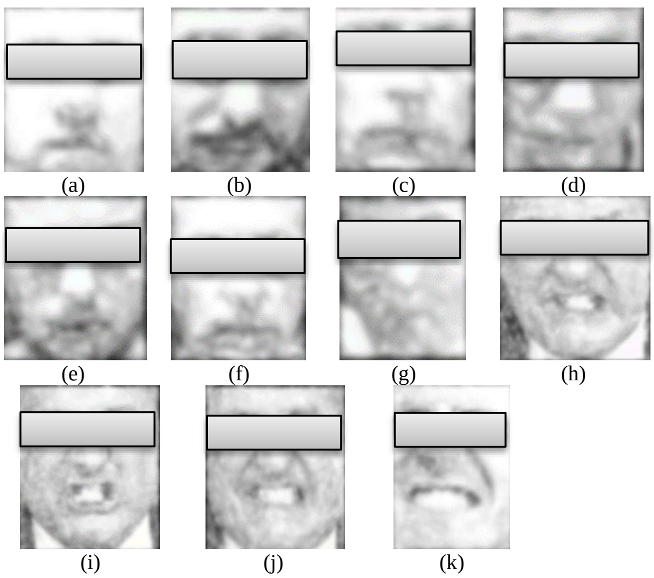


Fig. 2. VAS value dataset sample (a) 0, (b) 1, (c) 2, (d) 3, (e) 4, (f) 5, (g) 6, (h) 7, (i) 8, (j) 9 (k)10.

C. Model Training

The model-building stage is carried out on Google Colab. The first step in this process is training the YOLOv5 model available in the ultralytics github repository. The previously prepared dataset is cloned and entered using the roboflow API with the Model training configuration shown in Table I.

TABLE I. MODEL TRAINNING CONFIGURATION

Image Size	640x640 pixels
Parameters	Batch Size = 16
	Epoch = 300
Hyperparameter	default
IoU	0.5 and 0.75

The 0.5 threshold is often used as a baseline to evaluate model performance, while the 0.75 threshold is utilized for more stringent evaluation [32]. The model is not only accurate but also robust across a wide range of datasets and conditions [33]and models evaluated at the 0.5 IoU threshold performed significantly differently compared to models evaluated at lower thresholds[34].

D. Model Evaluation

The evaluations conducted in this study include precision, recall, mean Average Precision (mAP), and Intersection over Union (IoU). This process is carried out to assess the effectiveness of the object identification algorithm.

Precision measures the ratio of true positives (TP) correctly identified from all positive predictions, while recall evaluates the ratio of positive examples correctly identified from all object examples. These metrics are very important for assessing the precision of the model in object identification while avoiding excessive false positives[35]. F1-score is a metric that considers the trade-off between precision and recall, offering a single number that represents the overall performance of the model[36]. The mathematical representation for recall, precision, and F1-score are given in Eq.(1), Eq.(2), and Eq.(3).

$$Recall = \frac{TP}{TP+FN} = \frac{TP}{All\ Ground\ Truth} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} = \frac{TP}{All\ Prediction} \quad (2)$$

$$F1 - Score = \frac{2 \times (Recall+Precision)}{(Recall+Precision)} \quad (3)$$

Mean Average Precision (mAP) is an important quantification in the evaluation of YOLOv5. This value is calculated by taking the average of the precision scores at various levels of recall. This value is often calculated at a certain IoU threshold, such as mAP@0.5, to assess the model's capacity to predict object locations accurately. Studies have shown that YOLOv5 can achieve high mean precision (mAP) scores, thus validating its efficacy in various detection tasks. [37], [38]. The formula for mAP is given in Eq.(4).

$$\text{mAP} = \frac{1}{n}\sum_{i=1}^{n} APi \quad (4)$$

where,

n: Number of data AP

AP: Average Precision

If the evaluated model does not meet the desired value, which is above 0.8, then the model is re-created by adding a new image dataset to improve the parameter values to match the desired so that the model is considered suitable for use.

To evaluate the overlap between predicted bounding boxes and ground truth boxes, the intersection over Union (IoU) metric is a key measure. A higher intersection over Union (IoU) coefficient indicates superior accuracy in localization. If the IoU value is greater than the threshold value of 0.5 (the value assumed to increase the accuracy of detected objects), then the results are acceptable[39] [40]. IoU can be calculated using the formula, as shown in Eq.(5).

$$IoU = \frac{Intersection\ Area}{Union\ Area} \quad (5)$$

E. Automatically Estimating Pain Level of Drivers from Their Facial Expression

We involved 15 males online taxi drivers who had been driving for 5 to 6 hours as research subject. During driving, many VAS values will be read according to the participant's facial expression because the driver has previously driven for 5-

6 hours. Therefore, for the limit, the VAS value chosen is the highest VAS value with an observation period of no more than 2 hours.

For the research environment setting, the camera is mounted in front of the driver, which is focused on the driver's head and is arranged in such a way that it does not obstruct the driver's view as shown in Fig. 3. The data taken from the camera is video which is then converted into image and then processed to detect VAS value.

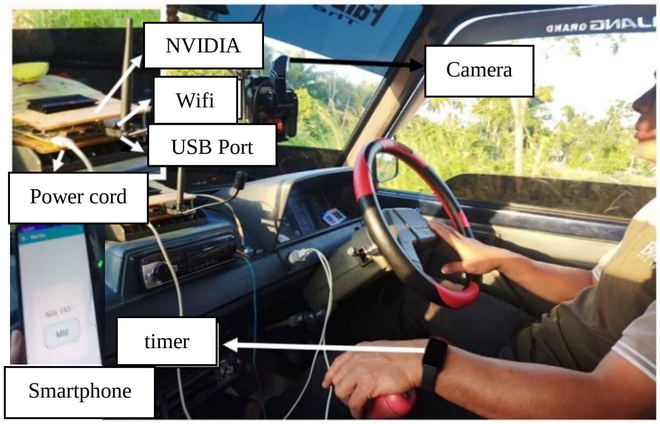


Fig. 3. A view of the experimental scenario.

III. Result and Discussion

A. YOLOv5 Model Training Result

In YOLO model training with 300 epochs and a batch size of 16, the results are shown in the confusion matrix as shown in Fig. 4. In addition, the recall, precision, F1-score, and mAP values are also calculated to measure the performance of the model. The all-normalization technique is applied in this study by dividing each element in the confusion matrix by the total number of instances in the data set. This procedure transforms each cell into a representation of the overall proportion of the data set rather than simply showing the absolute number. This is relevant to this study, where the classes correspond to different levels of pain (ranging from 0 to 10), all normalization can give an overall understanding of how well the model performs across all levels.

Based on the confusion matrix results, the prediction scores obtained can provide an understanding of the extent to which the YOLO model is able to classify each class with accuracy. The value class 8 gets the highest prediction score of 0.96. This shows that the model is very good at identifying and predicting objects included in these classes. The high prediction score in class 8 can be caused by several factors, such as clear and consistent representation in the training data, sufficient variation in the objects representing this class, and optimal hyperparameter selection. Meanwhile, the three classes of VAS values with values 2, 3, and 4 have almost the same and relatively low confusion matrix values of 0.40; 0.63; and 0.42 respectively. This is influenced by several factors, including the expression for these values has almost the same expression because pain with VAS values 0 to 4 has not changed much of the expression. Therefore, if the Interval between the classes is relatively small, difficulty in distinguishing between the classes can occur. In addition, the subjectivity of the interpretation of VAS values by individuals can cause similar perceptions of different levels and the limitations of the samples used in data collection also affect this. As for the result, the value of the train data is as shown in Fig. 5.

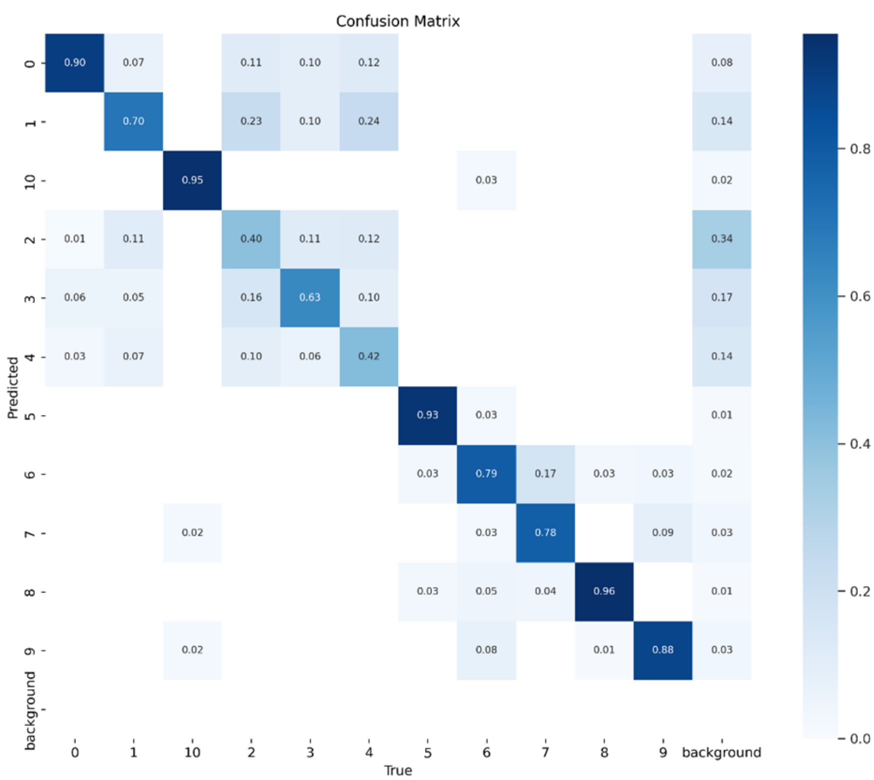


Fig. 4. Confusion matrix result.

From these results, numerous figures illustrate variations in the performance of the YOLO model throughout the training process. The graph illustrates that box loss, object loss, and class loss exhibit a general level of consistency, with minor variations over the training phase. This observation suggests that the model exhibits consistency in its learning. Nevertheless, the recall graph exhibits instability resulting from imbalances in class distribution, less than ideal model parameters and configurations, and inadequate quality of the dataset. To enhance recall stability, it is imperative to examine the distribution of samples in each class, fine-tune model parameters, and enhance dataset quality by collecting more representative data. Furthermore, to enhance the legibility of the model training outcomes, they are shown in Table II. The training results obtained an average value of mAP at 0.5 of 0.82673 and mAP at 0.95 of 0.67048 which were produced with epoch 300. The training process with default hyperparameters was stable at epoch 300 so that the model was said to have been fulfilled and could be used.

For model evaluation using the IoU value, The IoU value is varied to 0.5 and 0.75 which aims to determine the effect of the given IoU. The evaluation results on the model are shown in Fig. 6. The f1-score graph results have almost the same shape, this means that the variation of the IoU value does not have a significant effect on the f1-score value because the precision and recall values do not change much when the IoU value is varied in this model. From the graph, it can be concluded that the recommended confidence value to use is in the range of 0.2 to 0.5.

TABLE II. MODEL TRAINING RESULT

Type of Data	***Box Loss***	***Object Loss***	***Class Loss***
Training	0.013476	0.0049682	0.0095172
Validation	0.012259	0.003167	0.014258
Metrics	Precision	Recall	F1-Score
	0.73679	0.82046	0.789075908
	mAP		
	@0.5	@0.5:0.95	
	0.82673	0.67048	

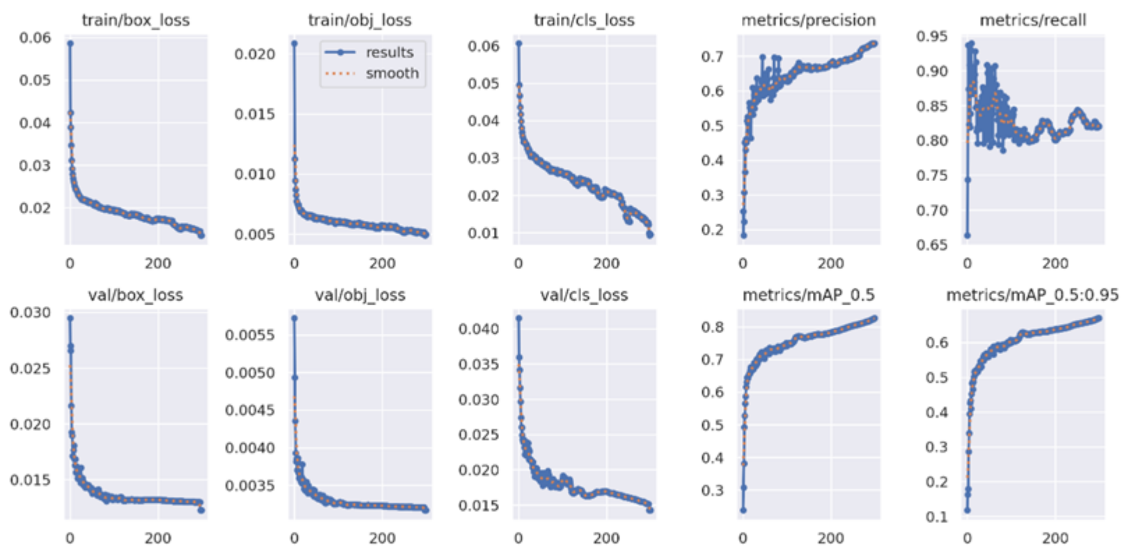
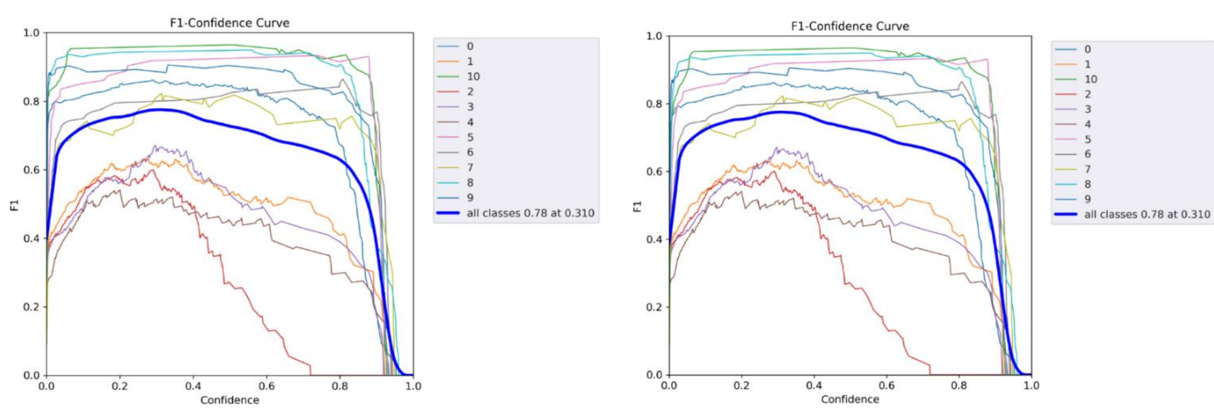


Fig. 5. Model training result.



(a) (b)

Fig. 6. F1-score (a) IoU 0.5, (b)).75.

B. Evaluation with Real World Drivers

The estimation results from the model are compared with ratings from a medical doctor. The classification results are divided into three levels, namely “mild” with a VAS value range of 0-3, “moderate” with a VAS value of 4-7, and “severe” with a VAS value of 8-10. The results from the 15 drivers recorded are shown in Table III.

TABLE III. PAIN LEVEL ESTIMATED FROM FACIAL EXPRESSION BY THE MODEL AND A MEDICAL DOCTOR

Subject	VAS Value		Classification	
	Subjective rating by medical doctor	***Proposed Model***	***Subjective rating by medical doctor***	***Proposed Model***
1	7	6	severe	moderate
2	5	6	moderate	moderate
3	6	6	moderate	moderate
4	5	5	moderate	moderate
5	8	8	severe	severe
6	5	6	moderate	moderate
7	1	1	mild	mild
8	5	5	moderate	moderate
9	1	1	mild	mild
10	9	9	severe	severe
11	9	9	severe	severe
12	3	4	mild	moderate
13	8	6	severe	moderate
14	1	1	mild	mild
15	1	1	mild	mild

Of the 15 participants, there were 3 VAS values that deviated from the subjective assessment by the medical doctor so that accuracy can be calculated by subtracting the observed value from the actual value and dividing it by the actual value, then multiplying by 100%, so that:

$$Accuracy = \frac{15-3}{15} = \frac{12}{15} \times 100\% = 80\%$$

With the test results with an error rate of 20%, the following discusses factors that can affect measurement accuracy.

This study was conducted in Indonesia involving male volunteer drivers in Bengkulu City who are predominantly Malay Austronesian ethnicity with an age range of 20-30 years. The results of the study cannot be generalized to be applied to everyone. Older people will have different detection rates. Although the existing datasets are all male, there are still limitations due to demographic, social and cultural diversity. This of course can affect the generalization of the model to different populations. In addition, the availability of datasets from the country of origin is not yet available, which will affect the reliability of the model. Therefore, more accuracy needs to be improved by adding datasets using different ethnic facial variations. Although the results of this study work well on the existing test dataset, if applied to a general context, it still needs to be validated.

Pain is a feeling that can be subjectively felt by someone. It could be that the variation shown in pain at a certain level but is different in giving expression. Variations in facial expressions are a challenge in themselves that can affect the accuracy of this study. Although it has involved external assessments involving medical doctors, it may still be influenced by bias in the dataset. This can also affect the results, especially if applied to different populations. Nevertheless, the proposed model has been proven effective in detecting pain using the Visual Analogue Scale value approach.

Model performance when applied in the real world is influenced by lighting, camera angle, driver movement and if the driver speaks or shows other expressions such as emotions and others. Although research has applied variations in datasets by adding variations in the form of blur, rotation, and brightness enhancement so that the model training process can learn from various data conditions, so that the model can learn from various data conditions, lighting conditions greatly affect the results obtained. Proper camera placement is also a major concern so that data consistency can be maintained. Another thing is if the driver speaks or gives an emotional expression that is very likely to affect the measurement results. These results also may not be generalizable to all conditions due to differences in demographic and cultural factors as discussed above. In addition, variations in pain types (acute and chronic types) should also be considered so that these results may be difficult to apply to different types of pain.

Future research can be further developed by collecting more diverse data sets from different ethnicities which is expected to improve system learning. To improve system validation, this research can be developed by involving biomedical sensors such as heart rate or tone of voice, so that the results are more accurate and can be validated independently. The selectivity and sensitivity of the research can still be improved by applying certain methods so that they can distinguish between real pain and fake pain, without the help of experts in assessing pain. The use of cameras that can work optimally and are not too affected by the environment such as vehicle movement and changes in light intensity can also be considered.

IV. CONCLUSION

In this study we propose to estimate a pain level expressed by facial expressions of drivers by applying VAS value. Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads. The proposed model can process and classify the driver's facial expression based on the Visual Analog Scale (VAS) value scale ranging from 0 to 10. The classification results are divided into three levels, namely "mild" with a VAS value range of 0-3, "moderate" with a VAS value of 4-7, and "severe" with a VAS value of 8-10.

The results of the confusion matrix show that the YOLO model can classify each class accurately. Class value 8 obtained the highest prediction score of 0.96, but the other 2 classes, namely classes 2 and 4, showed accuracy values below 0.5. This can be caused by the relatively small interval between classes or the subjectivity of the interpretation of VAS values so that the same perception is at different levels.

The experiment test results also show that the system can classify the driver's facial expression of pain level with that of classification by the medical doctor at rate of 80%. Thus, the model may be used to detect driver pain and could lead to a reduction the number of traffic accidents in the future.

ETHICS STATEMENT

Our research is based on non-invasive measurement (camera recording from the car dashboard), and we did not intervene in the everyday routine of the experimental participants. The 15 experimental participants were taxi drivers, and their facial expressions during their routine work were recorded and analyzed. All experimental participants gave informed consent before the experiment and written consent was obtained. The pain rating obtained was based solely on the driver's facial expression recorded from a camera, either using a neural-network model or by a medical doctor. We did not obtain the actual subjective pain experience by the participants. Their facial expression may not reflect the actual amount of pain they were experiencing. There is almost no or small possibility of risk or danger arising from the implementation of this research.

In accordance with the local legislation and institutional requirement[2] where this experiment was performed because the study was investigating public behavior and was purely observational (non-invasive and non-interactive), ethical approval was not required.

REFERENCES

[1] A. Aafreen, A. R. Khan, A. Khan, N. K. Maurya, and A. Ahmad, "Prevalence of Neck Pain in Car and Motorcycle Drivers: A Comprehensive Review of Primary, Secondary, and Tertiary Care," JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH, 2023, doi: 10.7860/jcdr/2023/64993.18222.

[2] A. Vaezipour, M. S. Horswill, N. E. Andrews, V. Johnston, P. Delhomme, and O. Oviedo-Trespalacios, "How distracting is chronic pain? The impact of chronic pain on driving behaviour and hazard perception," Accid Anal Prev, vol. 178, 2022, doi: 10.1016/j.aap.2022.106856.

[3] C. Longtin, A. Lacasse, C. E. Cook, M. Tousignant, and Y. Tousignant-Laflamme, "Management of low back pain by primary care physiotherapists using the pain and disability drivers management model: An improver analysis," Musculoskeletal Care, vol. 21, no. 3, 2023, doi: 10.1002/msc.1742.

[4] E. Matsuoka, M. Saji, and K. Kanemoto, "Daytime sleepiness in epilepsy patients with special attention to traffic accidents," Seizure, vol. 69, 2019, doi: 10.1016/j.seizure.2019.04.006.

[5] C. van Vreden et al., "The physical and mental health of Australian truck drivers: a national cross-sectional study," BMC Public Health, vol. 22, no. 1, 2022, doi: 10.1186/s12889-022-12850-5.

[6] S. N. Raja et al., "The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises," Sep. 01, 2020, Lippincott Williams and Wilkins. doi: 10.1097/j.pain.0000000000001939.

[7] M. K. Nicholas, "Time vs mechanism in chronic pain," 2022. doi: 10.1097/j.pain.0000000000002584.

[8] S. Kazeminasab et al., "Neck pain: global epidemiology, trends and risk factors," Dec. 01, 2022, BioMed Central Ltd. doi: 10.1186/s12891-021-04957-4.

[9] E. E. Krebs et al., "Development and Initial Validation of the PEG, a Three-item Scale Assessing Pain Intensity and Interference," J Gen Intern Med, vol. 24, no. 6, 2009, doi: 10.1007/s11606-009-0981-1.

[10] H. M. McCormack, D. J. de L. Horne, and S. Sheather, "Clinical applications of visual analogue scales: A critical review," Psychol Med, vol. 18, no. 4, 1988, doi: 10.1017/S0033291700009934.

[11] C. Maxwell, "Sensitivity and accuracy of the visual analogue scale: a psycho - physical classroom experiment.," Br J Clin Pharmacol, vol. 6, no. 1, 1978, doi: 10.1111/j.1365-2125.1978.tb01676.x.

[12] M. M. Alfonsin, R. Chapon, C. A. B. de Souza, V. K. Genro, M. M. C. Mattia, and J. S. Cunha-Filho, "Correlations among algometry, the visual analogue scale, and the numeric rating scale to assess chronic pelvic pain in women," Eur J Obstet Gynecol Reprod Biol X, vol. 3, 2019, doi: 10.1016/j.eurox.2019.100037.

[13] M. Jiang et al., "Elastic Silicone Occlusive Sheeting Versus Silicone Occlusive Sheeting in the Treatment of Scars: A Randomized Controlled Trial," Dermatol Ther (Heidelb), vol. 12, no. 8, 2022, doi: 10.1007/s13555-022-00763-5.

[14] L. G. R. Fernandes et al., "Correlation between levels of perceived stress and depressive symptoms in the functional disability of patients with fibromyalgia," Rev Assoc Med Bras, vol. 69, no. 11, 2023, doi: 10.1590/1806-9282.20230690.

[15] J. Aceituno-Gómez et al., "Correlation between three assessment pain tools in subacromial pain syndrome," Clin Rehabil, vol. 35, no. 1, pp. 114–118, Jan. 2021, doi: 10.1177/0269215520947596.

[16] Y. S. Park, J. Choi, and S. W. Park, "Blink reflex changes and sensory perception in infraorbital nerve-innervated areas following zygomaticomaxillary complex fractures," Arch Plast Surg, vol. 47, no. 6, pp. 559–566, 2020, doi: 10.5999/aps.2020.01130.

[17] J. Khadka, P. G. Schoneveld, and K. Pesudovs, "Comparing the measurement properties of visual analogue and verbal rating scales," Ophthalmic and Physiological Optics, vol. 42, no. 1, 2022, doi: 10.1111/opo.12917.

[18] M. Alkhouli, Z. Al-Nerabieah, and M. Dashash, "Analyzing the Facial Action Units associated with genuine and fake pain caused by inferior alveolar nerve block in Syrian children: a cross-sectional study," Jun. 21, 2023. doi: 10.21203/rs.3.rs-3044856/v1.

[19] Z. Chen, R. Ansari, and D. J. Wilkie, "Learning Pain from Action Unit Combinations: A Weakly Supervised Approach via Multiple Instance Learning," IEEE Trans Affect Comput, vol. 13, no. 1, pp. 135–146, 2022, doi: 10.1109/TAFFC.2019.2949314.

[20] C. Ma, C. Wang, D. Zhu, M. Chen, M. Zhang, and J. He, "The Investigation of the Relationship Between Individual Pain Perception, Brain Electrical Activity, and Facial Expression Based on Combined EEG and Facial EMG Analysis," J Pain Res, vol. 18, pp. 21–32, 2025, doi: 10.2147/JPR.S477658.

[21] P. J. Göller, P. Reicherts, S. Lautenbacher, and M. Kunz, "Vicarious facilitation of facial responses to pain," European Journal of Pain (United Kingdom), vol. 28, no. 1, pp. 133–143, Jan. 2024, doi: 10.1002/ejp.2169.

[22] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016. doi: 10.1109/CVPR.2016.91.

[23] R. Lakhotiya, M. Chavan, S. Divate, and S. Pande, "Image Detection and Real Time Object Detection," Int J Res Appl Sci Eng Technol, vol. 11, no. 5, 2023, doi: 10.22214/ijraset.2023.51839.

[24] G. Yang et al., "Face Mask Recognition System with YOLOV5 Based on Image Recognition," in 2020 IEEE 6th International Conference on Computer and Communications, ICCC 2020, 2020. doi: 10.1109/ICCC51575.2020.9345042.

[25] W. Yang, X. Gan, and J. He, "Defect Identification of 316L Stainless Steel in Selective Laser Melting Process Based on Deep Learning," Processes, vol. 12, no. 6, Jun. 2024, doi: 10.3390/pr12061054.

[26] O. M. Lawal, "YOLOv5-LiNet: A lightweight network for fruits instance segmentation," PLoS One, vol. 18, no. 3 March, 2023, doi: 10.1371/journal.pone.0282297.

[2] Guidelines and Ethical Standards National Health Research and Development, Ministry of Health of the Republic of Indonesia, 2021

[27] Y. Xu, “Real-Time Face Expression Recognition Monitoring Using Deep Learning,” in Advances in Transdisciplinary Engineering, IOS Press BV, Mar. 2024, pp. 698–704. doi: 10.3233/ATDE240135.

[28] Y. Ma, “Target tracking and detection based on YOLOv5 algorithm,” Applied and Computational Engineering, vol. 16, no. 1, 2023, doi: 10.54254/2755-2721/16/20230860.

[29] Y. Chang, D. Zhou, Y. Tang, S. Ou, and S. Wang, “An improved deep learning network for image detection and its application in Dendrobii caulis decoction piece,” Sci Rep, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-63398-w.

[30] A. Alotaibi and T. Arif, “Fast and accurate automated intestinal parasites egg detection and classication from images based on YOLOv5 deep convolutional neural network,” 2023, doi: 10.21203/rs.3.rs-2520494/v1.

[31] M. A. Haque et al., “Deep multimodal pain recognition: A database and comparison of spatio-temporal visual modalities,” in Proceedings - 13th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2018, Institute of Electrical and Electronics Engineers Inc., Jun. 2018, pp. 250–257. doi: 10.1109/FG.2018.00044.

[32] M. R. Subhi, E. Rachmawati, and G. Kosala, “Safety Helmet Detection on Field Project Worker Using Detection Transformer,” Journal of Information System Research (JOSH), vol. 4, no. 4, pp. 1316–1323, Jul. 2023, doi: 10.47065/josh.v4i4.3852.

[33] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, “Generalized intersection over union: A metric and a loss for bounding box regression,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, Jun. 2019, pp. 658–666. doi: 10.1109/CVPR.2019.00075.

[34] L. Lam, M. George, S. Gardoll, S. Safieddine, S. Whitburn, and C. Clerbaux, “Tropical Cyclone Detection from the Thermal Infrared Sensor IASI Data Using the Deep Learning Model YOLOv3,” Atmosphere (Basel), vol. 14, no. 2, Feb. 2023, doi: 10.3390/atmos14020215.

[35] S. M. Makhdoomi, C. Khosla, and S. D. Pande, “Real Time Lung Cancer Classification with YOLOv5,” EAI Endorsed Trans Pervasive Health Technol, vol. 9, no. 1, 2023, doi: 10.4108/eetpht.9.3925.

[36] W. Hao, C. Ren, M. Han, L. Zhang, F. Li, and Z. Liu, “Cattle Body Detection Based on YOLOv5-EMA for Precision Livestock Farming,” Animals, vol. 13, no. 22, 2023, doi: 10.3390/ani13223535.

[37] M. M. Rana, M. S. Hossain, M. M. Hossain, and M. D. Haque, “Improved vehicle detection: unveiling the potential of modified YOLOv5,” Discover Applied Sciences, vol. 6, no. 7, Jul. 2024, doi: 10.1007/s42452-024-06029-3.

[38] E. Casas, L. Ramos, E. Bendek, and F. Rivas-Echeverria, “Assessing the Effectiveness of YOLO Architectures for Smoke and Wildfire Detection,” IEEE Access, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3312217.

[39] M. A. Azam et al., “Deep Learning Applied to White Light and Narrow Band Imaging Videolaryngoscopy: Toward Real-Time Laryngeal Cancer Detection,” Laryngoscope, vol. 132, no. 9, 2022, doi: 10.1002/lary.29960.

[40] F. M. A. Mazen, R. A. A. Seoud, and Y. O. Shaker, “Deep Learning for Automatic Defect Detection in PV Modules Using Electroluminescence Images,” IEEE Access, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3284043.