Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Synthesizing Physiological Stress Facial Expressions for Medical Simulation Mannequins Using AI-Driven Facial Action Units

Ahmad Ridwan Fauzi

ahmadridwanfauzi@telkomuniversity.ac.id

Telkom University

Trini Handayani

Sekolah Tinggi Ilmu Kesehatan

Research Article

Keywords: Facial expression synthesis, Facial Action Coding System, OpenFace, medical simulation, shape keys, projection mapping

Posted Date: August 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10696643/v1

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

Additional Declarations: No competing interests reported.

Synthesizing Physiological Stress Facial Expressions for Medical Simulation Mannequins Using AI-Driven Facial Action Units

Ahmad Ridwan Fauzi[1]*

Trini Handayani[2]

[1]Department of Information System, Faculty of Industrial Engineering,
Telkom University, Jakarta, Indonesia

[2]Department of Nursing, Sekolah Tinggi Kesehatan Cianjur,
Cianjur, Indonesia

Abstract

Realistic facial expressions of physiological stress, such as pain, choking, and coughing, are important for high-fidelity medical simulation, yet existing simulators provide limited dynamic facial animation. This paper presents a technical pipeline for synthesizing stress-related facial expressions on 3D medical mannequins using the Facial Action Coding System (FACS) as a unifying framework. The model-based component generates graded pain expressions from force sensor input using the Prkachin-Solomon formulation, in which the intensities of the core action units (AUs) are derived by uniformly distributing the pain intensity value. The video-based component captures choking and coughing expressions from video recordings of a performer demonstrating these expressions, estimates AU intensity time series with OpenFace 2.0, and drives shape-key deformations of a 3D head model in Blender. A validation study with 16 nursing students assessed the recognizability of the synthesized expressions across four presentation conditions. The highest recognition accuracy (11 of 16 participants, 68.8%) was achieved when the facial animation was presented without audio, while the lowest accuracy (4 of 16, 25.0%) was observed when audio accompanied the dynamic expression, indicating that poorly synchronized audio cues can reduce recognition performance. The pipeline offers an open-source, low-cost approach for integrating dynamic patient affect into simulation-based healthcare training.

Keywords: Facial expression synthesis, Facial Action Coding System, OpenFace, medical simulation, shape keys, projection mapping

1 Introduction

Healthcare simulation has evolved from simple task trainers to sophisticated systems incorporating virtual reality, haptic feedback, and dynamic physiological responses [1]. A critical yet underdeveloped dimension of simulation fidelity is the realistic representation of patient facial expressions. During clinical procedures such as endotracheal suctioning, patients may exhibit a range of involuntary facial responses: pain from tissue contact, choking from oxygen desaturation, and coughing from airway irritation. Each response conveys essential clinical information to the practitioner. Current commercial simulators provide limited or static facial animation, lacking the dynamic and evidence-based expression synthesis needed to create immersive training scenarios. Recent advances in high-fidelity mannequin fabrication, including 3D-printed pediatric airway models [3], case-based virtual reality simulation for facial wound suture skill training [4], and immersive haptic simulation for training nurses in emergency medical procedures [5], have further expanded the technological landscape.

The Facial Action Coding System (FACS), developed by Ekman and Friesen [6], provides a comprehensive taxonomy of all visually distinguishable facial movements through action units (AUs). Each AU corresponds to the contraction of specific facial muscles, enabling any facial expression to be described as a combination of activated AUs with associated intensity values. FACS has been applied extensively in pain research [7, 8], psychology [9], and more recently in computer vision for automated facial behavior analysis [10, 11]. However, its application to medical simulation, particularly for synthesizing choking and coughing expressions, remains largely unexplored. Recent work demonstrated that inverse kinematics can be integrated with FACS-based facial rigs to

*Corresponding author: ahmadridwanfauzi@telkomuniversity.ac.id

achieve physically grounded expression synthesis [12], and real-time frameworks based on Unity Sentis further demonstrate the feasibility of translating detected AUs directly into character deformation [13].

The challenge of synthesizing choking and coughing expressions differs fundamentally from pain. Pain expression follows a relatively stereotyped pattern; Prkachin and Solomon identified core pain-related AUs (brow lowering, cheek raising, lid tightening, nose wrinkling, upper lip raising) whose activation intensity correlates with perceived pain magnitude, enabling model-based generation from sensor input [7]. In contrast, choking and coughing expressions exhibit high inter-individual variability, and, to the best of the authors' knowledge, no standardized FACS-based description of these expressions existed in the literature at the time of this work. This necessitates a fundamentally different approach: video-based motion capture combined with automated AU estimation.

Recent advances in computer vision have produced tools capable of automated facial behavior analysis from video input. OpenFace, an open-source toolkit developed for academic research, estimates 17 facial AU intensities, head pose, eye gaze, and facial landmarks from single-camera video [14, 15]. These capabilities create an opportunity to bridge the gap between video data and 3D facial animation: AU intensities extracted from recordings of coughing and choking expressions can drive expression synthesis on a 3D head model through rigging and shape-key deformation. State-of-the-art AU estimation methods, including self-supervised facial expression parsing based on AU patterns [16], deep learning approaches for AU intensity estimation [18] and facial expression recognition [19], EMG-based AU recognition [20], and the release of OpenFace 3.0 [17], continue to improve the fidelity of such pipelines. Blend-shape based expression synthesis has also been shown to implement facial expressions efficiently on 3D characters [21].

This paper presents a dual-approach pipeline for synthesizing physiological stress expressions on 3D medical mannequins. The model-based component generates pain expressions from sensor-derived intensity values using the Prkachin-Solomon FACS formulation [7]. The video-based component captures choking and coughing expressions through OpenFace analysis of video recordings. Both pathways converge on a unified 3D rigging framework in which AU intensities drive shape-key deformations. The recognizability of the synthesized expressions is evaluated through a validation study with 16 nursing students, demonstrating that FACS-driven expression generation produces visually distinguishable patient states. The contribution of this work is a complete, open-source-compatible pipeline that enables simulation developers to synthesize dynamic patient affect on low-cost hardware, complementing recent advances in immersive simulation environments for healthcare training [5]. The practical value of the approach is supported by a companion evaluation in which dynamic facial feedback improved situational awareness and procedural performance during respiratory care training [22].

2 Methods

2.1 Study design

The study employed a developmental research design comprising two phases: (1) a technical development phase in which the FACS-based expression synthesis pipeline was designed, implemented, and validated through iterative testing; and (2) an observational evaluation phase in which the recognizability of the generated expressions was assessed by human participants. The study was conducted at Sekolah Tinggi Kesehatan Cianjur between January and August 2025. The research protocol, including the participant evaluation procedures, was reviewed and approved by the Institutional Review Board of Telkom University. All participants provided written informed consent prior to participation.

2.2 FACS framework and 3D model preparation

The FACS taxonomy served as the foundational framework for all expression synthesis. Among the 44 single AUs defined by FACS, 21 AUs relevant to the target expressions (pain, choking, and coughing) were selected for implementation on the 3D head model. A commercially available 3D character model was imported into Blender, an open-source 3D modeling and animation software. Facial rigging was performed by placing skeletal control points (bones) at positions corresponding to the anatomical origins of facial muscles as described in the FACS manual [6]. Each bone was assigned a weight map defining its region of influence on the model mesh. Shape keys were then created for each AU by deforming the mesh to match the visual appearance illustrated in the FACS reference diagrams. The intensity of each shape key, denoted $SK_i \in [0,1]$, controls the degree of deformation of the corresponding AU_i, where 0 represents the neutral state and 1 represents maximum activation [26].

2.3 Model-based pain expression synthesis

For pain expression, a mathematical model was developed based on the work of Prkachin and Solomon [7], who identified four core pain-related AUs, brow lowering (AU4), cheek raising (AU6), lid tightening (AU7), and nose wrinkling (AU9), with upper lip raising (AU10) as an additional indicator. Lucey et al. formalized the relationship between pain intensity P and the core AU intensities as follows [8]:

$$P = \mathrm{AU}_4 + \max(\mathrm{AU}_6, \mathrm{AU}_7) + \max(\mathrm{AU}_9, \mathrm{AU}_{10}) \tag{1}$$

To generate pain expressions from an arbitrary pain intensity value, the inverse mapping was implemented by uniformly distributing the input pain intensity P across the core AUs. For a given pain intensity $P \in [0,1]$, each core AU intensity was computed as

$$\mathrm{AU}_i = \frac{P}{5}, \quad \text{for } i \in \{4,6,7,9,10\} \tag{2}$$

This uniform distribution model assumes a linear relationship between perceived pain and AU activation, consistent with the findings of Prkachin and Solomon [7].

The pain intensity P was derived from a force sensor embedded in the simulated trachea. The force sensor measured contact force as a three-dimensional vector $F = (f_x, f_y, f_z)$. The magnitude of force change between consecutive readings was computed using the Euclidean norm of the force delta vector,

$$\|\Delta F\| = \sqrt{\Delta f_x^2 + \Delta f_y^2 + \Delta f_z^2} \tag{3}$$

and mapped to pain intensity through a linear transformation calibrated such that a force delta of 0.1 N corresponded to maximum pain intensity ($P = 1.0$):

$$P = \frac{\|\Delta F\|}{0.1} \tag{4}$$

This calibration was determined through empirical observation during system development, following the procedure described in the original implementation of the simulator.

2.4 Video-based expression capture

For choking and coughing expressions, where no standardized FACS-based mathematical model exists, a video-based motion capture approach was employed. A healthy adult volunteer was asked to deliberately perform coughing and choking expressions, and the performance was recorded using a 720p webcam under consistent lighting conditions. The video recordings were processed using OpenFace 2.0, an open-source facial behavior analysis toolkit that performs facial landmark detection, head pose estimation, and AU intensity recognition [15]. OpenFace 2.0 estimates 17 AU intensities on a continuous scale from 0 (absent) to 5 (maximum intensity), outputting time-series data in comma-separated value format for each video frame [15].

The raw AU intensity values output by OpenFace (range 0 to 5) were normalized to the shape-key intensity range $[0,1]$ using the linear mapping

$$SK_i = \frac{\mathrm{AU}_i}{5} \tag{5}$$

where AU_i is the intensity estimated by OpenFace for the i-th AU. The normalized time-series data were then imported into Blender using a custom Python script that assigned each frame's AU values to the corresponding shape keys as keyframe animation data. This process produced temporally coherent 3D facial animations that reproduced the expression dynamics captured in the video recordings.

2.5 Expression state transitions

To enable smooth transitions between expression states, a state-based transition system was implemented. Four patient states were defined based on the clinical triggers encountered during endotracheal suctioning: neutral when the patient is not triggered; pain when force sensor contact is detected; choking when oxygen saturation falls below a threshold; and coughing as a probabilistic trigger during tube insertion. State transitions were governed by external sensor inputs and triggered pre-authored animation sequences. Transition blending was applied over a short interval (0.5 to 1.5 seconds depending on the clinical context) to avoid abrupt expression changes, ensuring that the 3D head model exhibited plausible dynamics rather than instantaneous expression switching.

2.6 Expression validation test

An expression validation test was conducted in August 2025 to evaluate whether observers could correctly identify the synthesized facial expressions. Sixteen nursing students of Sekolah Tinggi Kesehatan Cianjur (mean age 22.4 years; 9 males, 7 females) were recruited. Each participant provided written informed consent.

Participants were seated approximately one meter from the simulator display and observed the synthesized expressions in four randomized trials. Each trial presented one of four patient states, neutral, full pain, half pain (two intensity levels of the pain expression), or choking, under controlled audiovisual conditions. Trial 1 presented the static projected face together with sound cues. Trial 2 presented the static projected face without sound. Trial 3 presented the dynamic facial expression together with sound. Trial 4 presented the dynamic facial expression without sound. After each trial, participants were asked to identify which patient state was displayed from the four possible options. Following all four trials, participants completed a brief questionnaire providing qualitative feedback on expression realism and suggestions for improvement.

2.7 Statistical analysis

The primary outcome measure was the number of participants correctly identifying the displayed expression in each trial. Recognition accuracy was calculated as the proportion of correct responses per trial. Response distributions were summarized per trial using descriptive statistics. Given the single-trial-per-condition design, no inferential statistical tests or inter-rater reliability analyses were performed.

3 Results

3.1 Expression synthesis pipeline output

The dual-approach pipeline successfully generated three distinct facial expressions on the 3D head model. The model-based pain expression responded to force sensor input with graded intensity: at low pain intensity, subtle brow lowering and lip tightening were visible, whereas at the maximum pain intensity, all five core AUs (brow lowering, cheek raising, lid tightening, nose wrinkling, and upper lip raising) showed pronounced activation, producing an unambiguous pain grimace (Fig. 1). The video-based pipeline reproduced the recorded choking and coughing expressions on the 3D head model (Fig. 2), with AU intensity profiles extracted from the video recordings driving the shape-key animation. The complete generation pipeline, from video capture through OpenFace analysis to 3D keyframe animation, is illustrated in Fig. 3.

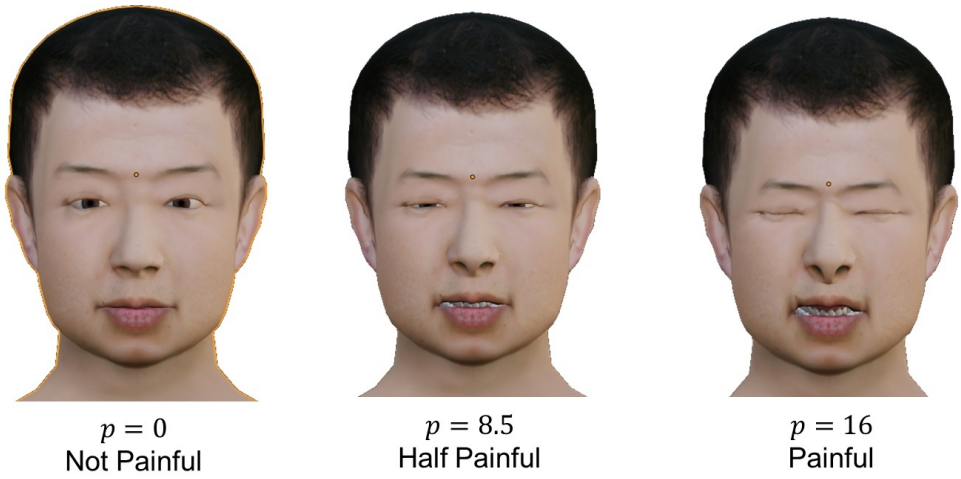


Fig. 1 Facial expression of pain generated on the 3D head model at increasing pain intensities, from the neutral state to the maximum pain intensity, showing progressive activation of the core pain-related AUs



Fig. 2 Video frames of a performer demonstrating physiological stress expressions and the corresponding expressions synthesized on the 3D head model, illustrating the video-to-model transfer achieved by the pipeline

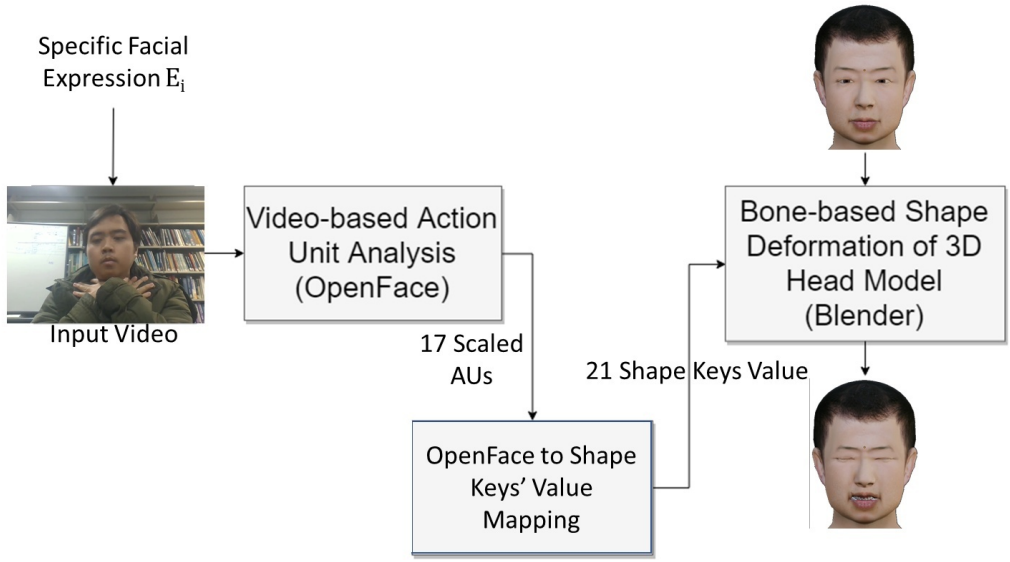


Fig. 3 Pipeline of facial expression generation: input video, video-based AU analysis with OpenFace, mapping of the 17 scaled AUs to 21 shape-key values, and bone-based shape deformation of the 3D head model in Blender

Of the 21 AUs implemented as shape keys, 17 could be driven by the OpenFace output. Four AUs, lip pucker, tongue show, and two additional oral AUs, could not be animated through the video-based pipeline because OpenFace 2.0 does not estimate these AUs. The 17 available AUs proved sufficient to produce recognizable choking and coughing expressions, particularly when combined with the temporal dynamics captured from the video recordings (Table 1). The transition system produced smooth expression changes without perceptible discontinuities between states.

Table 1 Summary of the action unit (AU) implementation on the 3D head model

Category	Number	Description
Targeted AUs	21	AUs relevant to pain, choking, and coughing, implemented as shape keys
Driven by OpenFace 2.0	17	AUs estimable from video recordings and animated
Not animatable	4	Lip pucker, tongue show, and two additional oral AUs

3.2 Expression recognition test

All 16 participants completed the four trials of the expression validation test. Recognition accuracy varied across trial conditions (Fig. 4 and Table 2). The highest accuracy was achieved in Trial 4 (dynamic expression without sound), in which 11 of 16 participants (68.8%) correctly identified the displayed expression. Trial 3 (dynamic expression with sound) yielded the lowest accuracy, with 4 of 16 participants (25.0%) responding correctly. Trials 1 and 2 (static projected face with and without sound) both resulted in 4 of 16 correct responses (25.0%), consistent with chance-level performance.

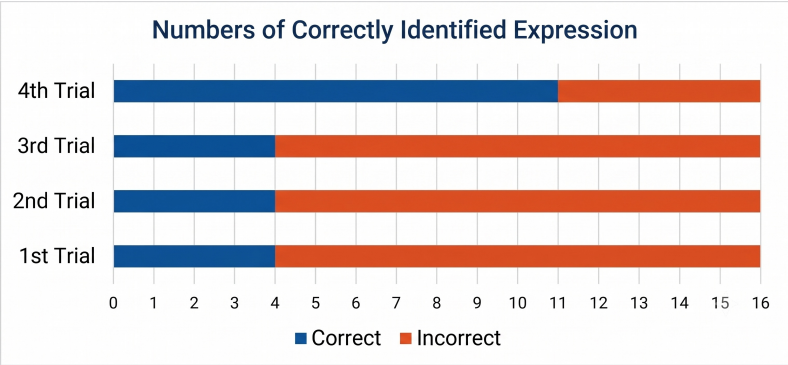


Fig. 4 Number of participants correctly identifying the displayed expression across the four trials (N = 16). Trial 1: static projection with sound; Trial 2: static projection only; Trial 3: dynamic expression with sound; Trial 4: dynamic expression only

Table 2 Recognition results per trial (N = 16 nursing students)

Trial	Condition	Correct	Incorrect	Accuracy
1	Static projection with sound	4	12	25.0%
2	Static projection only	4	12	25.0%
3	Dynamic expression with sound	4	12	25.0%
4	Dynamic expression only	11	5	68.8%

Qualitative feedback from the post-test questionnaire indicated that sound-visual asynchrony was a primary source of confusion. Several participants noted that the audio cues, coughing and choking sounds, did not always align temporally with the visual facial movements, leading to uncertainty about the patient's state. In addition, response patterns in the fourth trial suggested that some participants confused the choking expression with pain, consistent with the overlapping AU activations of the two states, which both involve brow lowering, eye closure, and mouth opening.

4 Discussion

This study demonstrates the feasibility of synthesizing physiological stress facial expressions on 3D medical mannequins using FACS as a unifying framework. The dual-approach pipeline combines model-based pain generation with video-based capture of unstructured expressions, addressing a gap in current simulation technology in which dynamic facial expression remains largely absent despite its recognized importance for clinical realism [24]. Simulation realism is understood as a multi-dimensional construct spanning physical, semantic, and phenomenal aspects [2], and facial expression fidelity contributes directly to the phenomenal realism experienced by learners.

The expression recognition test revealed an important finding: the addition of audio cues paradoxically reduced recognition accuracy when synchronization was suboptimal. The highest accuracy (68.8%) occurred in Trial 4, in which the dynamic expression was presented in isolation, whereas the presence of sound reduced accuracy to chance level (Trial 3, 25.0%). The chance-level performance in Trials 1 and 2 confirms that the static projected face alone, with or without sound, did not convey the patient state, and that the dynamic facial expression was the informative factor. The result is consistent with the participant feedback collected in this study, in which subjects reported that poorly synchronized sound cues confused their perception of the patient's state. For simulation designers, this implies that visual fidelity should be prioritized and that audio integration must achieve tight temporal synchronization to avoid degrading the user experience. The findings also align with emerging evidence that expression-aware virtual reality systems can support simulation-based training when facial cues are reliably rendered [25].

The confusion between choking and pain expressions suggests that these two states share overlapping AU activations, both involving brow lowering, eye closure, and mouth opening, and may require additional contextual cues for differentiation; in the FACS framework, similar AU combinations can represent different affective states depending on context and temporal dynamics. Similarly, recent work on 3D character expression using blend shapes has shown that shape-key configurations can implement facial expressions efficiently across different characters, supporting the generalizability of the shape-key approach adopted here [21].

The use of OpenFace for AU estimation enabled a data-driven approach to expression synthesis that would be difficult to achieve through manual FACS coding alone. Recent advances in automated facial analysis, including self-supervised AU-based expression parsing [16] and deep learning approaches for AU intensity estimation and facial expression recognition [18, 19], suggest that the fidelity of the pipeline could be further improved through integration with more recent AU estimation models. The release of OpenFace 3.0 with enhanced multitask capabilities and improved landmark detection [17] may also address the limitation of 17 estimable AUs identified in this study. EMG-based AU recognition offers a complementary sensing modality for real-time applications [20].

Compared with existing simulation approaches, the FACS-based pipeline offers several advantages. Unlike commercial simulators that rely on pre-scripted facial animations, this system generates expressions dynamically in response to real-time sensor input. The rigging and shape-key framework is transferable to any 3D head model, enabling reuse across different mannequin platforms. Furthermore, the video-based capture pathway provides a mechanism for expanding the expression repertoire: any facial expression that can be performed and recorded can, in principle, be synthesized on the 3D model through the OpenFace-to-shape-key pipeline. The approach complements immersive simulation environments for healthcare training, such as the haptic nursing simulation environment described by Gutiérrez-Fernández et al. [5], by adding realistic patient affect to the visual channel of the simulation. A companion evaluation of the feedback system in respiratory care training reported that dynamic facial feedback improved procedural performance, reducing task completion time among nursing students by 13.0% [22].

4.1 Limitations

Several limitations should be acknowledged. First, the validation test participants were nursing students who may have limited experience with clinical situational awareness and may not fully represent the perceptual expectations of experienced clinical practitioners. Future studies should include experienced clinicians as evaluators. Second, the choking and coughing recordings were performed by a healthy volunteer rather than recorded from clinical episodes. Although deliberately performed expressions are standard practice in facial expression research, recordings of actual patients experiencing these symptoms would further increase physiological validity; this extension requires clinical collaboration and appropriate ethics approval. Third, the coughing expression was synthesized but not included in the recognition test, which covered the neutral, full pain, half pain, and choking states. Fourth, the single-trial-per-condition design prevented the computation of inter-rater reliability and limited the analysis to descriptive statistics. Fifth, the study did not evaluate how the technology contributes to simulation-based learning or learner performance, and no comparison with commercial mannequin facial animation systems was performed. Future work should address these limitations through more diverse evaluator samples, multi-performer video capture, comparison with established simulation platforms, integration of advanced AU estimation models, and evaluation of learning outcomes in simulation curricula.

5 Conclusion

This paper presented a dual-approach pipeline for synthesizing physiological stress facial expressions on 3D medical mannequins using the Facial Action Coding System as a unifying framework. The model-based component generates graded pain expressions from force sensor input using the Prkachin-Solomon FACS formulation, while the video-based component captures and reproduces choking and coughing expressions through OpenFace-based AU analysis and 3D shape-key animation. An expression validation test with 16 nursing students demonstrated that the synthesized expressions are visually recognizable, with the highest accuracy (68.8%) achieved when the facial animation was presented without potentially desynchronized audio cues. The pipeline offers a scalable, open-source-compatible framework for enhancing the visual realism of medical training simulators, with potential applications in emergency medicine, nursing education, and surgical simulation.

Statements and Declarations

Funding. The authors did not receive support from any organization for the submitted work. No funding was received to assist with the preparation of this manuscript.

Competing interests. The authors have no relevant financial or non-financial interests to disclose.

Ethics approval. The research protocol was reviewed and approved by the Institutional Review Board of Telkom University.

Consent to participate. All participants provided written informed consent prior to participation.

Consent for publication. All participants provided written informed consent for participation. The performer depicted in Fig. 2 provided written consent for the publication of the images, and all other human participant data are presented in aggregated form.
Data availability. The data supporting the findings of this study are available from the corresponding author on reasonable request. Video recordings are not publicly shared to protect participant privacy.
Author contributions. Conceptualization: Ahmad Ridwan Fauzi; Methodology: Ahmad Ridwan Fauzi, Trini Handayani; Software: Ahmad Ridwan Fauzi; Validation: Ahmad Ridwan Fauzi, Trini Handayani; Formal analysis: Ahmad Ridwan Fauzi; Investigation: Ahmad Ridwan Fauzi; Resources: Trini Handayani; Writing, original draft: Ahmad Ridwan Fauzi; Writing, review and editing: Ahmad Ridwan Fauzi, Trini Handayani; Supervision: Trini Handayani.
Use of artificial intelligence. The OpenFace 2.0 toolkit used in this study employs deep neural networks for facial landmark detection and AU intensity estimation. During the preparation of this work, the authors used an AI-based writing assistant to help proofread the manuscript text, to identify relevant references, and to synthesize the reference list into the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgments

The authors gratefully acknowledge the Faculty of Industrial Engineering, Telkom University, for providing research facilities, and the nursing students of Sekolah Tinggi Kesehatan Cianjur who volunteered for the expression validation test.

Biographies



Ahmad Ridwan Fauzi received the bachelor's degree in informatics engineering from Institut Teknologi Sepuluh Nopember (ITS), Indonesia, the M.Sc. degree in information science and technology, and the M.B.A. degree from EDHEC Business School, France. He worked as a senior software engineer at Xfers (now Payfazz) and previously taught business at Universitas Suryakancana. He is currently a lecturer in the Department of Information System, Faculty of Industrial Engineering, Telkom University, Jakarta, Indonesia. His research interests include medical simulation, computer graphics, and applied artificial intelligence.



Trini Handayani received the medical doctor degree from Universitas Sebelas Maret, Indonesia, and the doctoral degree in law from Universitas Islam Bandung. She is currently a lecturer in the Department of Nursing at Sekolah Tinggi Kesehatan (STIKes) Cianjur, where she teaches nursing education, and also teaches law at Universitas Suryakancana. Her research interests include simulation-based learning, nursing education, and health law.

References

[1] Motola, I., Devine, L.A., Chung, H.S., Sullivan, J.E., Issenberg, S.B.: Simulation in healthcare education: a best evidence practical guide. AMEE Guide No. 82. Med. Teach. 35(10), e1511–e1530 (2013). https://doi.org/10.3109/0142159X.2013.818632

[2] Dieckmann, P., Gaba, D., Rall, M.: Deepening the theoretical foundations of patient simulation as social practice. Simul. Healthc. 2(3), 183–193 (2007). https://doi.org/10.1097/SIH.0b013e3180f637f5

[3] Mencarelli, M., Serio, P., Puggelli, L., Baggi, R., Avenali, S.: 3D-printed high-fidelity pediatric mannequin for rigid and flexible bronchoscopy training on difficult airways. Eng. Rep. 8(5), e70786 (2026). https://doi.org/10.1002/eng2.70786

[4] Zhang, Y., Nie, S., Dai, Y., Yan, Y.: Case-based virtual reality simulation for penetrating facial wound suture skill training in medical students: design and pilot study. BMC Med. Educ. 25(1) (2025). https://doi.org/10.1186/s12909-025-08151-1

[5] Gutiérrez-Fernández, A., Fernández-Llamas, C., Vázquez-Casares, A.M., Mauriz, E., Riego-del-Castillo, V., John, N.W.: Immersive haptic simulation for training nurses in emergency medical procedures. Vis. Comput. 40, 7527–7537 (2024). https://doi.org/10.1007/s00371-023-03227-9

[6] Ekman, P., Friesen, W.V.: Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto (1978)

[7] Prkachin, K.M., Solomon, P.E.: The structure, reliability and validity of pain expression: evidence from patients with shoulder pain. Pain 139(2), 267–274 (2008). https://doi.org/10.1016/j.pain.2008.04.010

[8] Lucey, P., Cohn, J.F., Prkachin, K.M., Solomon, P.E., Matthews, I.: Painful data: the UNBC-McMaster shoulder pain expression archive database. In: 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG), pp. 57–64. IEEE, Santa Barbara (2011). https://doi.org/10.1109/FG.2011.5771462

[9] Ekman, P.: An argument for basic emotions. Cogn. Emot. 6(3-4), 169–200 (1992). https://doi.org/10.1080/02699939208411068

[10] Martinez, B., Valstar, M.F., Jiang, B., Pantic, M.: Automatic analysis of facial actions: a survey. IEEE Trans. Affect. Comput. 10(3), 325–347 (2019). https://doi.org/10.1109/TAFFC.2017.2731763

[11] Zhi, R., Liu, M., Zhang, D.: A comprehensive survey on automatic facial action unit analysis. Vis. Comput. 36(5), 1067–1093 (2020). https://doi.org/10.1007/s00371-019-01707-5

[12] Wang, B., Jing, L., Shin, J.: Integrating inverse kinematics and the Facial Action Coding System for physically grounded facial expression synthesis. Electronics 15(3), 558 (2026). https://doi.org/10.3390/electronics15030558

[13] Luo, Y., Wu, F.: Real-time character animation generation via deep learning-based facial expression recognition using Unity Sentis. In: Proceedings of the 2025 6th International Conference on Computer Information and Big Data Applications (CIBDA), pp. 83–86 (2025). https://doi.org/10.1145/3746709.3746725

[14] Baltrušaitis, T., Robinson, P., Morency, L.-P.: OpenFace: an open source facial behavior analysis toolkit. In: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1–10. IEEE (2016). https://doi.org/10.1109/WACV.2016.7477553

[15] Baltrušaitis, T., Zadeh, A., Lim, Y.C., Morency, L.-P.: OpenFace 2.0: facial behavior analysis toolkit. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), pp. 59–66. IEEE, Xi'an (2018). https://doi.org/10.1109/FG.2018.00019

[16] Zou, K., Jin, L., Xu, X.: Self-supervised facial expression parsing: unveiling global patterns through facial action units. Vis. Comput. 41, 8139–8152 (2025). https://doi.org/10.1007/s00371-025-03860-6

[17] Hu, J., Mathur, L., Liang, P.P., Morency, L.-P.: OpenFace 3.0: a lightweight multitask system for comprehensive facial behavior analysis. In: 2025 IEEE 19th International Conference on Automatic Face and Gesture Recognition (FG), IEEE (2025). https://doi.org/10.1109/fg61629.2025.11099277

[18] Chen, Y., Zhang, J., Wang, T., Li, X., Yang, M.-H.: Trend-aware supervision: on learning invariance for semi-supervised facial action unit intensity estimation. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38 (2024). https://doi.org/10.1609/aaai.v38i1.27803

[19] Solis-Arrazola, M.A., Sanchez-Yañez, R.E., Garcia-Capulin, C.H.: Enhancing image-based facial expression recognition through muscle activation-based facial feature extraction. Comput. Vis. Image Underst. 240, 103927 (2024). https://doi.org/10.1016/j.cviu.2024.103927

[20] Veldanda, A., Liu, H., Koschke, R.: Can electromyography alone reveal facial action units? A pilot EMG-based action unit recognition study with real-time validation. In: Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies, pp. 142–151 (2024). https://doi.org/10.5220/0012399100003657

[21] Kim, K.M.: A study on the facial expressions of 3D characters: with a focus on using the blend shapes. J. Basic Des. Art 26(2), 29–41 (2025). https://doi.org/10.47294/ksbda.26.2.3

[22] Fauzi, A.R., Handayani, T., Sazkia, A.C., Ardiansyah, M.: Dynamic facial feedback in simulation: effects on situational awareness during respiratory care training. Clin. Simul. Nurs. 118, 102049 (2026). https://doi.org/10.1016/j.ecns.2026.102049

[23] Sellberg, C.: Caring for the mannequin: social interactions in simulation-based nursing education. SSRN preprint (2025). https://doi.org/10.2139/ssrn.5203039

[24] Schwarz, P., Spanknebel, S., Immel, D., Müller, R., Wagner, B.: Automated capture and transfer of human facial expressions to humanoid robots for realistic patient simulation. Front. Robot. AI (2026). https://doi.org/10.3389/frobt.2026.1798227

[25] Alekhine, J., Fabroyir, H.: Facial expression-aware virtual reality for simulation training using deep learning. In: 2025 15th International Conference on Information & Communication Technology and System (ICTS), pp. 1–6 (2025). https://doi.org/10.1109/icts67612.2025.11369582

[26] Blender Foundation: Blender, open-source 3D creation suite. https://www.blender.org (2017). Accessed 10 Aug 2026