



4º Congresso Brasileiro
de Ciência e Saberes
Multidisciplinares
**tudo é
ciência**
11º Encontro de Extensão
Universitária do UniFOA

**23 a 25
de outubro**

Submissões abertas até 07/09

Análise de expressões faciais como ferramenta de apoio pedagógico em aulas remotas: uma abordagem baseada em deep learning

Vitor Amadeu Souza¹; 0009-0002-1857-6799

1 – UniFOA, Centro Universitário de Volta Redonda, Volta Redonda, RJ.
vitor.amadeu@foa.org.br

Resumo: O ensino remoto emergencial implementado durante a pandemia de COVID-19 trouxe desafios significativos para a comunicação não-verbal entre professores e estudantes. Este estudo propõe o uso de tecnologias de reconhecimento automático de expressões faciais baseadas em deep learning como ferramenta auxiliar para professores em aulas síncronas online. Através da implementação da biblioteca DeepFace, foram analisadas sete categorias emocionais básicas (alegria, tristeza, raiva, surpresa, medo, desgosto e neutralidade) em imagens teste, obtendo-se uma acurácia de 100% na classificação das emoções testadas. A metodologia envolveu o desenvolvimento de um sistema automatizado capaz de detectar e classificar expressões faciais em tempo real, proporcionando feedback imediato sobre o engajamento emocional dos estudantes. Os resultados demonstram não apenas a viabilidade técnica da solução proposta, mas também sua alta precisão em condições controladas. Conclui-se que a análise automática de expressões faciais pode contribuir significativamente para a melhoria da qualidade do ensino remoto, oferecendo aos educadores insights valiosos sobre o estado emocional e o nível de engajamento de seus estudantes, desde que implementada de forma ética e transparente.

Palavras-chave: Ensino remoto; Reconhecimento de emoções; Deep learning; Tecnologia educacional; Expressões faciais.



INTRODUÇÃO

A pandemia de COVID-19 transformou drasticamente o cenário educacional mundial, forçando instituições de ensino a adotarem rapidamente modalidades de ensino remoto emergencial (Hodges *et al.*, 2020). Esta transição abrupta revelou limitações significativas na comunicação não-verbal entre professores e estudantes, um componente fundamental do processo de ensino-aprendizagem tradicionalmente presencial (Silva; Guerino; Valentim, 2022). No ambiente presencial, os educadores frequentemente utilizam pistas visuais, como expressões faciais e linguagem corporal dos estudantes, para avaliar o nível de compreensão, engajamento e bem-estar emocional da turma (Tytler et al. (2017).

A perda dessa capacidade de leitura emocional no ambiente virtual representa um desafio pedagógico considerável, uma vez que Anderson e Dron (2011) argumentam que a interação efetiva entre professor e estudante é um dos pilares fundamentais do sucesso educacional, independentemente da modalidade de ensino. No contexto remoto, a limitação das telas, problemas de conectividade e a tendência dos estudantes de desabilitarem suas câmeras criam barreiras adicionais para essa interação (Castelli; Sarvary, 2021).

Simultaneamente, os avanços em inteligência artificial e visão computacional têm proporcionado ferramentas cada vez mais sofisticadas para o reconhecimento automático de expressões faciais e estados emocionais (Pantic; Rothkrantz, 2003). Tecnologias baseadas em deep learning, particularmente redes neurais convolucionais, demonstraram capacidades notáveis na classificação de emoções humanas a partir de imagens faciais (Mollahosseini *et al.*, 2017). Essas tecnologias apresentam potencial significativo para aplicação em contextos educacionais, oferecendo aos professores ferramentas objetivas para monitorar o estado emocional e o engajamento de seus estudantes.

O reconhecimento de emoções em contextos educacionais não é um conceito inteiramente novo, pois pesquisas anteriores exploraram o uso de sensores fisiológicos e análise comportamental para medir estados afetivos de aprendizes (D'mello; Graesser, 2012). No entanto, a aplicação específica de tecnologias de reconhecimento facial em ambientes de ensino remoto permanece relativamente inexplorada, especialmente considerando as particularidades técnicas e éticas envolvidas.



Este estudo visa preencher essa lacuna investigando a viabilidade e as implicações do uso de sistemas automáticos de reconhecimento de expressões faciais como ferramenta de apoio pedagógico em aulas remotas. A pesquisa fundamenta-se na hipótese de que a análise automática de expressões faciais pode fornecer aos educadores informações valiosas sobre o estado emocional coletivo de seus estudantes, permitindo ajustes pedagógicos em tempo real e contribuindo para uma experiência de aprendizagem mais efetiva.

O objetivo geral desta pesquisa é desenvolver e avaliar um sistema de análise automática de expressões faciais aplicado ao contexto educacional remoto, enquanto os objetivos específicos incluem implementar um sistema de reconhecimento de emoções utilizando a biblioteca DeepFace, testar a acurácia do sistema na detecção das sete emoções básicas de Ekman relacionadas ao uso de tecnologias de reconhecimento facial em ambientes educacionais.

MÉTODOS

A fundamentação teórica desta pesquisa baseia-se na teoria das emoções básicas de Paul Ekman, que identifica sete expressões faciais universais: alegria, tristeza, raiva, surpresa, medo, desgosto e neutralidade (Ekman, 1992). Essa categorização tem sido amplamente utilizada em estudos de reconhecimento automático de emoções e fornece uma base sólida para aplicações computacionais. Adicionalmente, o estudo incorpora princípios da teoria do aprendizado afetivo, que reconhece a importância dos estados emocionais no processo de aprendizagem (Pekrun *et al.*, 2002), uma vez que pesquisas demonstram que emoções positivas como alegria e surpresa podem facilitar a aprendizagem, enquanto emoções negativas como medo e tristeza podem criar barreiras cognitivas (Tyng *et al.*, 2017).

O sistema desenvolvido utiliza a biblioteca DeepFace, uma ferramenta de código aberto que integra diversos modelos pré-treinados de reconhecimento facial e análise de emoções (Serengil; Ozpinar, 2020). A escolha desta biblioteca justifica-se por sua facilidade de implementação, acurácia comprovada e suporte a múltiplos backends de detecção facial, incluindo RetinaFace, MTCNN e OpenCV. A arquitetura do sistema compreende cinco componentes principais: o módulo de aquisição de imagens, responsável pelo download e



armazenamento local das imagens de teste representando diferentes estados emocionais; o módulo de pré-processamento, que realiza a normalização das imagens, incluindo redimensionamento e ajustes de contraste quando necessário; o módulo de detecção facial, que utiliza o algoritmo RetinaFace para localizar e segmentar faces nas imagens de entrada; o módulo de análise emocional, que aplica redes neurais convolucionais pré-treinadas para classificar as expressões faciais detectadas; e o módulo de visualização, que apresenta os resultados através de interfaces gráficas utilizando matplotlib.

Para avaliar a eficácia do sistema proposto, foi criado um dataset de teste contendo imagens representativas das sete emoções básicas. As imagens foram selecionadas de bancos de imagens livres (Freepik), seguindo critérios específicos de qualidade: resolução mínima de 640x480 pixels, iluminação adequada, face claramente visível e expressão emocional não ambígua. O processo de coleta automatizada utilizou requisições HTTP para download das imagens, que foram organizadas em uma estrutura de diretórios categorizados por emoção, abordagem que permite a replicabilidade do estudo e facilita a expansão do dataset para pesquisas futuras.

O código-fonte está disponível para download através do link: <https://github.com/vitor-souza-ime/facial>.

RESULTADOS E DISCUSSÃO

Os resultados obtidos demonstram a viabilidade técnica do sistema proposto para identificação automática de expressões faciais, com o algoritmo baseado em DeepFace alcançando uma acurácia de 100% na classificação das sete emoções básicas testadas. Esta performance indica que, em condições controladas com imagens de alta qualidade e expressões bem definidas, a tecnologia de reconhecimento automático de emoções apresenta capacidade de classificação equivalente à percepção humana especializada. O Quadro 1 demonstra os resultados esperados e alcançados para cada predição feita.

Quadro 1 - Resultado das predições



Expected: Happy Detected: Happy 	Expected: Sad Detected: Sad 	Expected: Angry Detected: Angry 
Expected: Surprise Detected: Surprise 	Expected: Neutral Detected: Neutral 	Expected: Fear Detected: Fear 
	Expected: Disgust Detected: Disgust 	

Fonte: Freepik.

A literatura científica sugere que sistemas de reconhecimento de emoções em condições reais tipicamente alcançam acurácias entre 70% e 85% (Happy; Routray, 2015), o que ainda representa um nível de precisão útil para aplicações educacionais. As implicações pedagógicas dos resultados obtidos são substanciais para o contexto educacional remoto. A capacidade demonstrada de identificar automaticamente estados emocionais pode fornecer aos professores informações valiosas sobre o engajamento da turma e a efetividade de suas estratégias pedagógicas (Picard *et al.*, 2004). Em cenários práticos de ensino remoto, um sistema com essa precisão poderia alertar professores sobre momentos de



4º Congresso Brasileiro
de Ciência e Saberes
Multidisciplinares
**tudo é
ciência**
11º Encontro de Extensão
Universitária do UNIFOA

**23 a 25
de outubro**

👉 Submissões abertas até 07/09

confusão generalizada, detectada através de expressões de surpresa ou neutralidade predominantes, ou desengajamento, identificado por expressões de tédio ou tristeza.

Essa informação permitiria ajustes pedagógicos em tempo real, como mudanças no ritmo da aula, introdução de atividades interativas ou pausas para esclarecimentos (Kapila Vani; Jayashree, 2025). A tecnologia também oferece potencial para personalização do ensino baseada em dados emocionais agregados, permitindo que professores identifiquem tópicos que consistentemente geram confusão ou desinteresse, possibilitando refinamentos no conteúdo e nas estratégias de apresentação para aulas futuras (Calvo; D'Mello, 2010).

CONCLUSÕES

Este estudo demonstrou a viabilidade técnica e o potencial pedagógico da análise automática de expressões faciais como ferramenta de apoio ao ensino remoto. A acurácia de 100% obtida nos testes realizados com a biblioteca DeepFace representa um marco significativo, indicando que a tecnologia de reconhecimento emocional baseada em deep learning atingiu um nível de maturidade suficiente para aplicações práticas em ambientes educacionais controlados.

Estes resultados, embora obtidos em condições experimentais ideais, sugerem que a implementação dessa tecnologia em cenários reais de ensino remoto pode fornecer aos educadores informações valiosas e confiáveis sobre o estado emocional de seus estudantes. A capacidade demonstrada de identificar automaticamente as sete emoções básicas com precisão abre possibilidades concretas para o desenvolvimento de sistemas pedagógicos adaptativos que respondam em tempo real ao engajamento e bem-estar emocional dos aprendizes.

O potencial transformador dessa tecnologia para a melhoria da qualidade educacional é significativo, mas sua realização plena depende da colaboração contínua entre pesquisadores, educadores e formuladores de políticas para garantir implementações responsáveis e eficazes que verdadeiramente servem aos interesses dos estudantes e da sociedade como um todo.



REFERÊNCIAS

ANDERSON, T.; DRON, J. Three generations of distance education pedagogy. *The International Review of Research in Open and Distributed Learning*, v. 12, n. 3, p. 80-97, 2011. Disponível em: <https://www.erudit.org/en/journals/irrodl/2011-v12-n3-irrodl05132/1067616ar/abstract/>. Acesso em: 12 mai. 2025.

KAPILA VANI, R. K.; JAYASHREE, P. Multimodal emotion recognition system for e-learning platform. *Education and Information Technologies*, [S.l.], p. 1-32, 2025. DOI: <https://doi.org/10.1007/s10639-024-13279-6>. Acesso em: 12 abr. 2025.

CALVO, R. A.; D'MELLO, S. Affect detection: an interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, v. 1, n. 1, p. 18-37, 2010. DOI: <https://doi.org/10.1109/T-AFFC.2010.1>. Acesso em: 20 abr. 2025.

CASTELLI, F. R.; SARVARY, M. A. Why students do not turn on their video cameras during online classes and an equitable and inclusive plan to encourage them to do so. *Ecology and Evolution*, v. 11, n. 8, p. 3565-3576, 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.7123>. Acesso em: 10 abr. 2025.

D'MELLO, S.; GRAESSER, A. Dynamics of affective states during complex learning. *Learning and Instruction*, v. 22, n. 2, p. 145-157, 2012. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0959475211000806>. Acesso em: 25 mar. 2025.

EKMAN, P. An argument for basic emotions. *Cognition & emotion*, v. 6, n. 3-4, p. 169-200, 1992. DOI: <https://doi.org/10.1080/02699939208411068>. Acesso em: 10 mar. 2025.

HAPPY, S. L.; ROUTRAY, A. Automatic facial expression recognition using features of salient facial patches. *IEEE transactions on Affective Computing*, v. 6, n. 1, p. 1-12, 2015. DOI: 10.1109/TAFFC.2014.2386334. Acesso em: 26 mar. 2025.

HODGES, C. et al. The difference between emergency remote teaching and online learning. *Educause Review*, v. 27, p. 1-12, 2020. Disponível em: <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>. Acesso em: 16 mar. 2025.

SILVA, D. E. dos S.; GUERINO, G. C.; VALENTIM, N. M. C. Emergency Remote Teaching Based on Education 4.0: an Experience Report about the Methodology adopted in the Experimental HCI Class. *Revista Brasileira de Informática na Educação*, [S. l.], v. 30, p. 350-378, 2022. DOI: 10.5753/rbie.2022.2527. Disponível em: <https://journals-sol.sbc.org.br/index.php/rbie/article/view/2527>. Acesso em: 8 abr. 2025.

MOLLAHOSSEINI, A.; HASANI, B.; MAHOOR, M. H. AffectNet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective*



4º Congresso Brasileiro
de Ciência e Saberes
Multidisciplinares
**tudo é
ciência**
11º Encontro de Extensão
Universitária do UNIOA

**23 a 25
de outubro**

👉 Submissões abertas até 07/09

Computing, v. 10, n. 1, p. 18-31, 2017. DOI: 10.1109/TAFFC.2017.2740923. Acesso em: 8 mar. 2025.

PANTIC, M.; ROTHKRANTZ, L. J. Toward an affect-sensitive multimodal human-computer interaction. Proceedings of the IEEE, v. 91, n. 9, p. 1370-1390, 2003. DOI: 10.1109/JPROC.2003.817122. Acesso em: 18 mar. 2025.

PEKRUN, R. et al. Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. Educational psychologist, v. 37, n. 2, p. 91-105, 2002. DOI: https://doi.org/10.1207/S15326985EP3702_4. Acesso em: 25 mar. 2025.

PICARD, R. W. et al. Affective learning—a manifesto. BT technology journal, v. 22, n. 4, p. 253-269, 2004. DOI: <https://doi.org/10.1023/B:BTTJ.0000047603.37042.33>. Acesso em: 18 mar. 2025.

Tytler, R., Murcia, K., Hsiung, CT., Ramseger, J. (2017). Reasoning Through Representations. In: Hackling, M., Ramseger, J., Chen, HL. (eds) Quality Teaching in Primary Science Education. Springer, Cham. https://doi.org/10.1007/978-3-319-44383-6_7. Acesso em: 18 abr. 2025.

SERENGIL, S. I.; OZPINAR, A. LightFace: A hybrid deep face recognition framework. In: 2020 Innovations in Intelligent Systems and Applications Conference (ASYU). IEEE, 2020. p. 23-27. DOI: 10.1109/ASYU50717.2020.9259802. Acesso em: 10 abr. 2025.

TYNG, C. M. et al. The influences of emotion on learning and memory. Frontiers in psychology, v. 8, p. 1454, 2017. DOI: <https://doi.org/10.3389/fpsyg.2017.01454>. Acesso em: 12 abr. 2025.