

Reconhecendo as emoções dos alunos com base em Análise de Expressão Facial

Xiaojie Li, Renli Yue, Weikuan Jia, Hong Wang*, *Escola Yuanjie
Zheng de Ciência da Informação e Engenharia
Universidade Normal de
Shandong*

Jinan, China e-mail: xli162011@163.com, wanghong106@163.com

Resumo—No campo educacional, a análise emocional dos alunos tem atraído grande atenção há várias décadas, uma vez que está intimamente associada à eficácia da aprendizagem e aos resultados educacionais. É de grande ajuda que os instrutores avaliem o estado de envolvimento dos alunos com base no monitoramento oportuno dos estados afetivos dos alunos. É impraticável avaliar as emoções dos alunos observando-os numa sala de aula grande. A prevalência dos modos de aprendizagem online trouxe novos desafios para a estimativa das respostas emocionais dos alunos. Assim, é desejável o reconhecimento automatizado dos estados emocionais dos alunos durante o processo de aprendizagem. A análise da expressão facial é uma das formas eficazes de identificar automaticamente avaliar os estados emocionais dos alunos nas atividades educativas.

Neste artigo, é proposto um algoritmo automatizado de reconhecimento de emoções utilizando informações de expressões faciais. Especificamente, a região facial é detectada pela primeira vez a partir da imagem de entrada usando redes convolucionais em cascata multitarefa (MTCNN). Em seguida, um modelo de aprendizagem profunda contendo rede backbone e função de perda de entropia cruzada é empregado para distinguir diferentes tipos de emoções faciais. Os resultados experimentais em bancos de dados públicos de emoções indicam que a estrutura proposta pode realizar o reconhecimento de emoções de forma eficaz.

Palavras-chave—Reconhecimento de emoções dos alunos, expressões faciais, eficácia da aprendizagem, aprendizagem profunda, rede neural convolucional

I. INTRODUÇÃO

As emoções dos alunos nas atividades educacionais, como palestras e exames, são conhecidas como um fator importante que afeta a qualidade da aprendizagem e, portanto, são uma grande preocupação dos instrutores. Os instrutores podem, consequentemente, ajustar o conteúdo instrucional e mudar a estratégia de ensino com base na compreensão das respostas afetivas dos alunos no processo de ensino, o que poderia eventualmente ajudar os alunos a aumentar o seu ganho de aprendizagem. Geralmente, na aprendizagem convencional em sala de aula, o estado emocional dos alunos nas atividades educativas pode ser avaliado através da observação dos seus comportamentos, incluindo gestos, posturas de cabeça e expressões faciais. No entanto, é problemático observar com precisão muitos alunos em salas de aula grandes. Além disso, o modo de aprendizagem online tem desempenhado um papel fundamental no sistema educativo desde o início da pandemia da COVID-19. É ainda mais desafiador obter feedback emocional dos alunos nos modos de aprendizagem on-line, onde os alunos podem estar sentados em frente às telas sem se concentrarem na instrução. Assim, é desejável lidar com os desafios de acompanhar o estado emocional dos alunos.

Na última década, o reconhecimento de emoções baseado em expressões faciais tem atraído grande atenção, uma vez que as expressões faciais podem efetivamente refletir o estado emocional dos alunos e são consideradas indicadores críticos de envolvimento e nível de compreensão [1-4].

Os métodos convencionais de avaliação são insuficientes para medir as respostas emocionais dos alunos em todos os contextos, especialmente para os modos de aprendizagem online, onde apenas as partes faciais dos alunos são observáveis. Assim, são necessários métodos automáticos de reconhecimento de emoções. O sucesso dos métodos de aprendizado de máquina permitiu aplicações promissoras deles no reconhecimento automático de emoções dos alunos com base na análise de expressões faciais [22-25]. Essas técnicas analisam vários sinais faciais de emoções humanas usando processamento de imagens e métodos de aprendizado de máquina. O reconhecimento automático de emoções pode facilitar o ajuste oportuno do curso tanto em salas de aula convencionais quanto em ambientes de aprendizagem digital, ajudando os instrutores a compreender melhor o status de aprendizagem dos alunos.

Neste artigo, é proposta uma estrutura de reconhecimento de emoções de estudantes em dois estágios, baseada na análise de expressões faciais. Primeiramente, as imagens faciais são detectadas e alinhadas usando MTCNN. Então, para extração e classificação de características, as imagens faciais alinhadas são alimentadas em um modelo de aprendizado profundo, no qual ResNet-18 é usado como rede backbone e a perda de entropia cruzada é empregada como função de perda.

O resto do artigo está organizado da seguinte forma. A Seção II fornece uma breve introdução à análise de expressões faciais e aos métodos de reconhecimento de emoções dos alunos. A Seção III apresenta o sistema proposto. Resultados experimentais e discussões são apresentados na Seção IV. A Seção V fornece algumas observações finais.

II. TRABALHO RELATADO

Nesta seção são apresentados os trabalhos anteriores relacionados ao sistema proposto, ou seja, reconhecimento de expressões faciais e reconhecimento de emoções dos alunos usando informações de expressões faciais.

A. Análise de expressão facial

A análise de expressão facial oferece uma variedade de aplicações em muitos campos. A análise da expressão facial é uma tarefa desafiadora devido à iluminação variável, ao fundo complexo, às diferenças sutis e à anotação difícil. Para resolver essas questões, um bom número de estudos de pesquisa foram propostos e alcançaram vários graus de sucesso.

*Autor correspondente (wanghong106@163.com)

Normalmente, um pipeline de análise de expressão facial inclui vários módulos, ou seja, detecção de rosto (ou alinhamento de rosto), extração de região de interesse (ROI), extração de recursos e um classificador. Especificamente, para uma determinada imagem, as regiões faciais são localizadas e as ROIs são extraídas. Em seguida, as características são extraídas das regiões faciais e alimentadas em um classificador para determinar a categoria da expressão facial. Métodos convencionais baseados em recursos feitos à mão e baseados em aprendizagem profunda são empregados para análise de expressão facial. No estágio inicial, características artesanais que descrevem informações de textura e forma são extraídas de regiões faciais com base no conhecimento prévio dos pesquisadores [5-8]. Esses recursos artesanais incluem, entre outros, LBP [9], Gabor [10], HOG [11] e NMF [12]. Nos últimos anos, o sucesso das redes neurais profundas na visão computacional encorajou a mudança de métodos artesanais baseados em recursos para métodos de aprendizagem profunda [13-19].

B. Reconhecimento emocional dos alunos

No campo educacional, o reconhecimento emocional tem atraído grande atenção nas últimas duas décadas, uma vez que desempenha um papel fundamental na decisão dos estados de aprendizagem dos alunos. Os pesquisadores propuseram muitos sistemas automáticos de reconhecimento de emoções para aplicações educacionais usando métodos de aprendizado de máquina.

Sun et al. propuseram um método de reconhecimento de expressões faciais em tempo real baseado em máquina de vetores de suporte para estimar o estado afetivo dos alunos no moderno sistema de ensino à distância [20]. Com esse sistema foram considerados sete tipos de emoções, ou seja, felicidade, raiva, tristeza, surpresa, nojo, medo e neutro. Tang et al. apresentou um protótipo de sistema para analisar automaticamente as expressões dos alunos, no qual a Sequência de Histograma de Padrão Binário Uniform Local Gabor (ULGBPHS) foi extraída como características e o K-vizinho mais próximo (KNN) foi empregado como classificador [21]. Savva et al. propuseram uma aplicação web para analisar as emoções dos alunos usando algoritmos de aprendizado de máquina [22]. Bosch et al. utilizou visão computacional e técnicas de aprendizado de máquina para detectar as respostas afetivas dos alunos. O método proposto foi avaliado em um conjunto de dados coletados em um laboratório de informática de uma escola, onde os alunos estavam envolvidos em um jogo educacional desenvolvido para ensinar princípios fundamentais da física newtoniana [23]. Usando o padrão móvel dos olhos e da cabeça, Krithika et al. projetou um sistema para identificar e monitorar as emoções dos alunos em um ambiente de e-learning e melhorar o ambiente de e-learning [24]. Lasri et al. propuseram um método de análise de emoções faciais baseado em redes neurais convolucionais (CNN). O método proposto demonstrou ser eficaz no banco de dados FER 2013 com sete tipos de expressões [25]. Feng et al. desenvolveu um método de reconhecimento automático de emoções baseado em uma rede neural convolucional orientada a aspectos (A-CNN) e um classificador usando o mecanismo de memória de longo curto prazo com atenção [26].

Bhadana et al. propuseram um método de reconhecimento de emoções faciais em tempo real usando classificação em cascata HAAR para detecção de rosto e CNN para classificação de emoções [27].

Considerando que as emoções humanas básicas são diferentes das emoções dos alunos no processo de aprendizagem, os pesquisadores criaram especificamente uma categoria de emoções acadêmicas. Pang et al. construiu um conjunto de dados contendo quatro emoções acadêmicas pertinentes, ou seja, feliz, confuso, entediado e neutro [28]. Tanto um método convencional usando recursos artesanais e um profundo

abordagem de aprendizagem são testadas no conjunto de dados proposto como avaliações de base.

Além disso, a análise de expressões faciais foi empregada para medir os níveis de compreensão e envolvimento, uma vez que eram fundamentais para avaliar a eficácia da aprendizagem dos alunos.

Para reconhecer a qualidade da compreensão e do envolvimento ativo dos alunos, Jain e Sah propuseram um método de reconhecimento de emoções utilizando informações da fala e das expressões faciais [29]. Shi et al. usou expressão facial para identificar a confusão acadêmica no aprendizado on-line com métodos de aprendizado de máquina [30]. Yang et al. propuseram um método de reconhecimento de emoções para identificar os níveis de compreensão dos alunos em ambientes virtuais de aprendizagem [31]. Whitehill et al. propuseram um método de reconhecimento automático de engajamento baseado nas expressões faciais dos alunos [4]. Os recursos Gabor e a máquina de vetores de suporte (SVM) foram empregados para treinar o classificador binário para detecção de engajamento.

III. O QUADRO PROPOSTO

Para reconhecer automaticamente as emoções dos alunos no processo de aprendizagem, um sistema de reconhecimento de emoções é proposto neste artigo. Nesta seção é apresentada a introdução ao framework proposto.

A. Visão geral do sistema

Conforme mostrado na Figura 1, o sistema proposto inclui dois estágios, ou seja, um módulo de detecção e alinhamento de rosto para detecção e alinhamento de região facial, e uma rede neural profunda para a ponta treinada para extração de características e atribuição de rótulos.

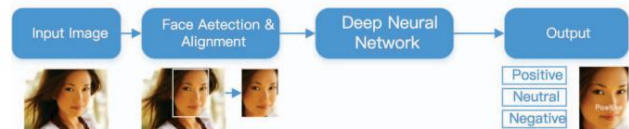


Figura 1. Ilustração do framework proposto.

Dado um conjunto de imagens de entrada, o MTCNN [32] é aplicado para detectar e alinhar regiões faciais. Depois disso, as imagens faciais são redimensionadas e alimentadas na rede neural profunda, mostrada na Figura 2.

Para as imagens faciais alinhadas, uma rede backbone é aplicada primeiro para extrair características profundas para processamento posterior. Devido à subjetividade dos anotadores e à ambiguidade entre as diferentes emoções faciais, as incertezas são inevitáveis no processo de anotação. Os rótulos inconsistentes e incorretos causados por essas incertezas levam a uma degradação significativa do desempenho dos métodos de reconhecimento de emoções faciais baseados em dados. Para suprimir essas incertezas, o módulo de ponderação de importância de autoatenção e o módulo de reetiquetagem introduzidos em [33] são integrados à estrutura proposta. Para ser mais específico, o esquema de ponderação de importância da autoatenção, que consiste em uma camada linear totalmente conectada (FC) e uma função de ativação Sigmóide, é usado para gerar pesos de atenção. Os pesos gerados são posteriormente aplicados à função de perda e ao módulo de reetiquetagem.

O módulo de reetiquetagem foi projetado para corrigir amostras incorretamente rotuladas com base nas probabilidades softmax. Para cada imagem, se a probabilidade máxima prevista for maior que a

probabilidade de seu rótulo anotado original com um determinado limite, a imagem será atribuída ao novo rótulo previsto. Assim, as amostras mal rotuladas podem ser corrigidas.

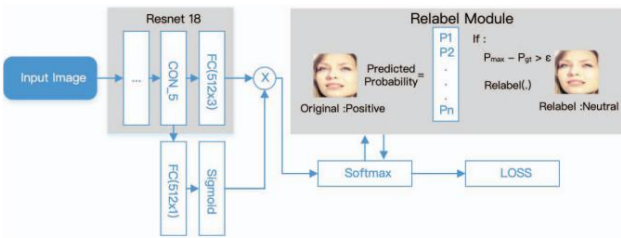


Figura 2. Ilustração da rede neural profunda.

B. Implementação

A estrutura profunda proposta é implementada com a caixa de ferramentas Pytorch. Tanto o VGG quanto o Resnet-18/50 são redes backbone amplamente utilizadas. Neste artigo, o Resnet-18 é usado como rede backbone para extração profunda de recursos. Ele foi pré-treinado no conjunto de dados de reconhecimento facial MS-Celeb-1M e as saídas de sua última camada de pooling foram extraídas como recursos. Todas as imagens faciais alinhadas foram redimensionadas para 224x224 pixels antes de serem alimentadas na rede backbone.

4. RESULTADOS EXPERIMENTAIS

Nesta seção, são apresentados os conjuntos de dados utilizados nos experimentos e os resultados da avaliação.

A. Bancos de dados

Atualmente, não há nenhum conjunto de dados disponível publicamente projetado especificamente para pesquisas sobre o reconhecimento de emoções dos alunos. Os conjuntos de dados existentes são construídos para analisar as emoções humanas básicas, que incluem principalmente felicidade, raiva, tristeza, surpresa, nojo, medo, desprezo e neutralidade. Como as emoções humanas básicas são diferentes das emoções dos alunos no processo de aprendizagem, não é apropriado utilizar diretamente esses conjuntos de dados para avaliar o processo de aprendizagem dos alunos [34]. Assim, inspirados no trabalho [29], os movimentos básicos humanos são ainda reclassificados em três tipos de emoções, ou seja, positivas, neutras e negativas. Uma emoção positiva pode ocorrer quando os alunos sentem que podem facilmente acompanhar os materiais de aprendizagem, enquanto a emoção negativa indica as respostas dos alunos aos desafios do processo de aprendizagem. A emoção neutra pode ocorrer quando nenhuma emoção positiva ou negativa óbvia é detectada. Para ser mais específico, a felicidade e a surpresa são consideradas emoções positivas, enquanto a tristeza, a raiva, o nojo e o medo são considerados emoções negativas. A lista de categorias de emoções é apresentada na Tabela 1.

TABELA I. LISTA DE CATEGORIA DE EMOÇÃO

Emoções humanas básicas	Tipo de emoções
Felicidade	Emoções positivas
Surpresa	
Tristeza	
Raiva	Emoções negativas
Nojo	
Temer	
Neutro	Emoção Neutra

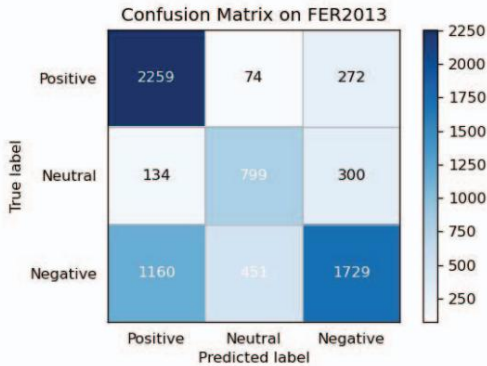
Para avaliar a eficácia da estrutura proposta, dois conjuntos de dados emocionais, ou seja, FER2013 [35] e RAF-DB [13]

são empregados neste artigo. O RAF-DB compreende cerca de 30.000 imagens faciais rotuladas por 40 anotadores com emoções básicas (subconjunto de rótulo único) ou emoções compostas (subconjunto de duas guias). Em nosso experimento, apenas o subconjunto de rótulo único com sete tipos de emoções básicas (ou seja, neutro, felicidade, surpresa, tristeza, raiva, nojo, medo) foi utilizado, resultando em 12.271 imagens para treinamento e 3.068 imagens para teste. O FER2013 foi coletado usando a API de pesquisa de imagens do Google e lançado nos Desafios ICML 2013 com as mesmas sete emoções faciais básicas do RAF-DB. Contém imagens faciais registradas em escala de cinza com tamanho de 48x48 pixels em um número total de 28.709 imagens de treinamento, 3.589 imagens de validação e 3.589 imagens de teste.

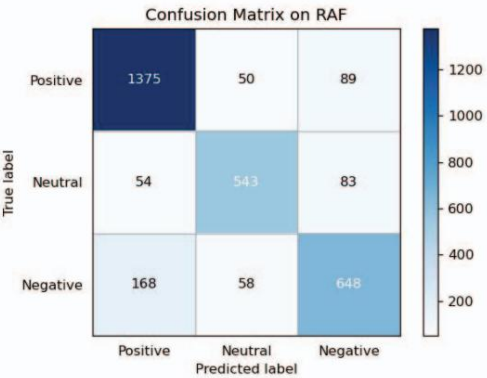
B. Resultados da avaliação

Para avaliar o desempenho do sistema proposto, o modelo é treinado no conjunto de dados de treinamento do RAF-DB e avaliado nos conjuntos de dados de teste de ambos os conjuntos de dados. Os resultados da avaliação estão listados na Tabela 2.

TABELA II. RESULTADOS DA AVALIAÇÃO				
Conjuntos de dados	Precisão			
	Positivo Emoções	Negativo Emoções	Neutro Emoção	Total
FER2013	86,72%	51,77%	64,80%	66,69%
RAF-DB	90,82%	74,14%	79,85%	83,64%



(a)



(b)

Figura 3. As matrizes de confusão nos dois conjuntos de dados. (a) matriz de confusão no FER2013, (b) matriz de confusão no RAF-DB.

Na Tabela 2, pode-se observar que o sistema proposto atinge precisões de reconhecimento de 66,69% e 83,64% no FER2013 e RAF-DB respectivamente, o que indica que o sistema proposto pode distinguir diferentes emoções de forma eficaz. Deve-se notar que o modelo foi treinado usando os dados de treinamento do RAF-DB e aplicado aos dados de teste do FER2013, o que demonstra que o modelo possui, até certo ponto, uma boa capacidade de generalização. Além disso, a Tabela 2 também fornece a precisão do reconhecimento de cada tipo de emoção. Para ambos os conjuntos de dados, as precisões de reconhecimento de emoções positivas são as mais altas, enquanto o desempenho no reconhecimento de emoções negativas é o pior. Mais detalhes sobre o reconhecimento de cada tipo de emoção podem ser vistos na matriz de confusão ilustrada em

Figura 3.

V. CONCLUSÕES

Neste artigo, é proposta uma estrutura de reconhecimento de emoções dos alunos baseada na aprendizagem profunda. Com o sistema proposto, redes convolucionais em cascata multitarefa (MTCNN) são utilizadas para detectar e alinhar as imagens faciais. Resnet-18 é ainda aplicado às imagens faciais como uma rede backbone para extrair recursos profundos. As operações de ponderação de importância e reetiquetagem de autoatenção são integradas para suprimir as incertezas da anotação do rótulo. A perda de entropia cruzada é usada como função de perda. Resultados experimentais em conjuntos de dados públicos demonstram a eficácia do sistema proposto.

RECONHECIMENTO

Este trabalho é apoiado pelo projeto Shandong Undergraduate Teaching Reform Research (chave) (2020124); Fundação de Ciências Naturais da Província de Shandong na China (ZR2020MF076); Foco no Plano de Pesquisa e Desenvolvimento na Província de Shandong (Nº: 2019GNC106115); Fundação Nacional de Ciências da Natureza da China (nº: 62072289); Programa Educacional Superior de Ciência e Tecnologia da Província de Shandong (Nº: J18KA308); Programa Acadêmico Taishan da Província de Shandong da China (Nº: TSHW201502038).

REFERÊNCIAS

[1] Maria Villa, Mikhail Gofman, Sinjini Mitra, Ali Almadan, Anoop Krishnan e Ajita Rattani, "Uma pesquisa de métodos biométricos e de aprendizado de máquina para rastrear a atenção e o envolvimento dos alunos." páginas 948-955.

[2] Mohamed Sathik e Sofia G Jonathan, "Efeito das expressões faciais no reconhecimento da compreensão do aluno em ambientes educacionais virtuais", *SpringerPlus*, vol. 2, não. 1, pp. 1-9, 2013.

[3] Joseph Grafsgaard, Joseph B Wiggins, Kristy Elizabeth Boyer, Eric N Wiebe e James Lester, "Reconhecendo automaticamente a expressão facial: Prevendo envolvimento e frustração."

[4] Jacob Whitehill, Zewelanj Serpell, Yi-Ching Lin, Aysha Foster e Javier R Movellan, "Os rostos do engajamento: reconhecimento automático do envolvimento do aluno a partir de expressões faciais", *IEEE Transactions on Affective Computing*, vol. 5, não. 1, pp. 86-98, 2014.

[5] Michael Lyons, Shigeru Akamatsu, Miyuki Kamachi e Jiro Gyoba, "Codificando expressões faciais com wavelets gabor." págs. 200-205.

[6] Yuan Luo, Cai-ming Wu e Yi Zhang, "Reconhecimento de expressão facial baseado no recurso de fusão do PCA

e LBP com SVM", *Optik-International Journal for Light and Electron Optics*, vol. 124, não. 17, pp. 2767-2770, 2013.

[7] Caifeng Shan, Shaogang Gong e Peter W McOwan, "Reconhecimento robusto de expressões faciais usando padrões binários locais." págs. 11-370.

[8] Yuxiao Hu, Zhihong Zeng, Lijun Yin, Xiaozhou Wei, Xi Zhou e Thomas S Huang, "Reconhecimento de expressão facial multivisualização." págs. 1-6.

[9] Matti Pietikäinen, Abdenour Hadid, Guoying Zhao e Timo Ahonen, *Visão computacional usando padrões binários locais*: Springer Science & Business Media, 2011.

[10] Chengjun Liu e Harry Wechsler, "Classificação baseada em recurso Gabor usando o modelo discriminante linear aprimorado de Fisher para reconhecimento facial", *IEEE Transactions on Image Processing*, vol. 11, não. 4, páginas 467-476, 2002.

[11] Navneet Dalal e Bill Triggs, "Histogramas de gradientes orientados para detecção humana." pp. 886-893.

[12] Ruicong Zhi, Markus Flierl, Qiuqi Ruan e W Bastiaan Kleijn, "Fatoração de matriz não negativa esparsa com preservação de gráfico com aplicação ao reconhecimento de expressão facial", *IEEE Transactions on Systems, Man, and Cybernetics, Parte B (Cybernetics)*, vol. 41, não. 1, pp. 38-52, 2010.

[13] Shan Li, Weihong Deng e JunPing Du, "Crowdsourcing confiável e aprendizagem profunda de preservação da localidade para reconhecimento de expressão na natureza." páginas 2852-2861.

[14] Hangyu Li, Nannan Wang, Xinpeng Ding, Xi Yang e Xinbo Gao, "Aprendizagem adaptativa da representação da expressão facial por meio de rótulos CF e destilação", *IEEE Transactions on Image Processing*, vol. 30, pp. 2016-2028, 2021.

[15] Jie Cai, Zibo Meng, Ahmed Shehab Khan, Zhiyuan Li, James O'Reilly e Yan Tong, "Perda de ilha para aprender características discriminativas no reconhecimento de expressões faciais." págs. 302-309.

[16] Yong Li, Jiabei Zeng, Shiguang Shan e Xilin Chen, "Reconhecimento de expressão facial com reconhecimento de oclusão usando cnn com mecanismo de atenção", *IEEE Transactions on Image Processing*, vol. 28, não. 5, pp. 2439-2450, 2018.

[17] Kai Wang, Xiaojiang Peng, Jianfei Yang, Debin Meng e Yu Qiao, "Redes de atenção regional para reconhecimento robusto de expressões faciais de pose e oclusão", *IEEE Transactions on Image Processing*, vol. 29, pp. 4057-4069, 2020.

[18] Yandong Wen, Kaipeng Zhang, Zhifeng Li e Yu Qiao, "Uma abordagem de aprendizagem de características discriminativas para reconhecimento facial profundo." págs. 499-515.

[19] Nianyin Zeng, Hong Zhang, Baoye Song, Weibo Liu, Yurong Li e Abdullah M Dobaie, "Reconhecimento de expressão facial por meio do aprendizado de codificadores automáticos esparsos profundos," *Neurocomputação*, vol. 273, pp. 643-649, 2018.

[20] Jian-Ming Sun, Xue-Sheng Pei e Shi-Sheng Zhou, "Reconhecimento de emoções faciais no moderno sistema de educação à distância usando SVM." págs. 3545-3548.

[21] Chuangao Tang, Pengfei Xu, Zuying Luo, Guoxing Zhao e Tian Zou, "Análise automática de expressões faciais de alunos em ambientes de ensino." págs. 439-447.

[22] Andreas Savva, Vasso Stylianou, Kyriacos Kyriacou e Florent Domenach, "Reconhecendo expressões faciais de alunos: um aplicativo da web." páginas 1459-1462.

[23] Nigel Bosch, Sidney D'Mello, Ryan Baker, Jaclyn Ocumpaugh, Valerie Shute, Matthew Ventura, Lubin Wang e Weinan Zhao, "Detecção automática de estados afetivos centrados na aprendizagem na natureza." págs. 379-388.

- [24] LB Krithika e Lakshmi Priya GG, "Sistema de reconhecimento de emoções do aluno (SERS) para melhoria do e-learning com base na métrica de concentração do aluno", *Procedia Computer Science*, vol. 85, pp. 767-776, 2016.
- [25] Imane Lasri, Anouar Riad Solh e Mourad El Belkacemi, "Reconhecimento de emoções faciais de alunos usando rede neural convolucional." págs. 1-6.
- [26] Xiang Feng, Yaojia Wei, Xianglin Pan, Longhui Qiu e Yongmei Ma, "Método de classificação e reconhecimento de emoções acadêmicas para ambiente de aprendizagem on-line em larga escala - baseado no método de pipeline de aprendizagem profunda A-CNN e LSTM-ATT", *International Journal of Environmental pesquisa e saúde pública*, vol. 17, não. 6, pp. 1941, 2020.
- [27] Lakshmi Bhadana, PV Lakshmi, D Rama Krishna, G Surya Bharti e Y Vaibhav, "Reconhecimento de emoções faciais em tempo real com rede neural convolucional profunda", *Journal of Critical Reviews*, vol. 7, não. 7500-7507, 2020.
- [28] Jordan Min Han Pang, Tee Connie e Goh Kah Ong Michael, "Reconhecimento de emoções acadêmicas em aulas online." págs. 445-450.
- [29] Ati Jain e Hare Ram Sah, "Feedback do aluno por emoção e reconhecimento de fala por meio de aprendizado profundo." págs. 442-447.
- [30] Zheng Shi, Ya Zhang, Cunling Bian e Weigang Lu, "Reconhecimento automático de confusão acadêmica na aprendizagem online baseada em expressões faciais." págs. 528-532.
- [31] Dongri Yang, Abeer Alsadoon, PW Chandana Prasad, Ashutosh Kumar Singh e Amr Elchouemi, "Um modelo de reconhecimento de emoções baseado no reconhecimento facial em ambiente virtual de aprendizagem", *Procedia Computer Science*, vol. 125, pp. 2-10, 2018.
- [32] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li e Yu Qiao, "Detecção e alinhamento de faces conjuntas usando redes convolucionais em cascata multitarefa", *IEEE Signal Processing Letters*, vol. 23, não. 10, pp. 1499-1503, 2016.
- [33] Kai Wang, Xiaojiang Peng, Jianfei Yang, Shijian Lu e Yu Qiao, "Suprimindo incertezas para reconhecimento de expressões faciais em grande escala." páginas 6897-6906.
- [34] Liping Hu, Fuchang Liu e Ran Fan, "Pesquisa da experiência emocional de aprendizagem dos alunos: baseada na inteligência artificial." págs. 232-235.
- [35] Ian J Goodfellow, Dumitru Erhan, Pierre Luc Carrier, Aaron Courville, Mehdi Mirza, Ben Hamner, Will Cukierski, Yichuan Tang, David Thaler e Dong-Hyun Lee, "Desafios na aprendizagem de representação: um relatório sobre três concursos de aprendizado de máquina." pp. 117-124.