

Início (<http://revistas.poli.br/index.php/anais/index>)
/ Arquivos (<http://revistas.poli.br/index.php/anais/issue/archive>)
/ v. 9 (2022): Mostra POLI/UPE
(<http://revistas.poli.br/index.php/anais/issue/view/38>)
/ Engenharia da Computação e Sistemas

Open Journal Systems
(<http://pkp.sfu.ca/ojs/>)

Reconhecimento de emoções a partir de dados multimodais e modelo híbrido de classificação para suporte terapêutico personalizado baseado em biofeedback emocional



Maíra Santana
Universidade de Pernambuco

Arianne Sarmento Torcate
Universidade de Pernambuco

Wellington Pinheiro dos Santos
Universidade de Pernambuco

Resumo

As terapias são importantes para aliviar, extinguir ou lidar com distúrbios de saúde. Esses distúrbios podem ser físicos, mentais ou uma combinação de ambos. Existem vários tipos de terapias para abarcar a ampla gama de problemas, de

Sponsors



Informações

Para Leitores

(<http://revistas.poli.br/index.php/anais/information/read>)

Para Autores

(<http://revistas.poli.br/index.php/anais/information/auth>)

Para Bibliotecários

(<http://revistas.poli.br/index.php/anais/information/libr>)

maneira que cada terapia é utilizada para atingir objetivos específicos de promoção da saúde. Visto que tratam de problemas de diversas ordens, as terapias devem ser conduzidas por profissionais capacitados, tais como psicoterapeutas, fonoaudiólogos, musicoterapeutas, fisioterapeutas e psicopedagogos. Um dos grandes desafios das terapias é promover o engajamento do paciente, nesse sentido, são incentivadas abordagens que auxiliem tanto na percepção do engajamento quanto no seu favorecimento (TEIXEIRA et al., 2020). Nos últimos anos, investimentos estão sendo feitos em formas de promover terapias personalizadas para melhorar os resultados obtidos através delas (COLOMBO et al., 2019; TEIXEIRA et al., 2020). Assim, são testadas formas de adaptar a terapia para atender aos desejos e preferências do(a) paciente. As ferramentas computacionais têm um papel fundamental nessa customização, especialmente aquelas que utilizam Inteligência Artificial (IA). A capacidade desses métodos de avaliar grandes quantidades de dados complexos de maneira detalhada, precisa e, muitas vezes, rápida pode beneficiar a personalização da terapia. Uma das formas mais íntimas de personalizar algo é perceber como a pessoa reage emocionalmente a esse estímulo. Nesse sentido, modelos computacionais estão sendo aplicados para reconhecimento automático de emoções, no entanto, identificar emoções não é trivial, pois elas se manifestam de diversas maneiras. Por este motivo, diferentes fontes de dados são usadas para capturar o estado emocional. A mais comum dessas fontes é através de expressões faciais, todavia, ao longo do tempo, verificou-se que as emoções poderiam ser melhor identificadas se essa análise facial fosse combinada com aspectos fisiológicos e neurofuncionais (GONZÁLEZ & MCMULLEN, 2020). Assim, parâmetros como resposta galvânica da pele (GSR), temperatura corporal, análise do padrão respiratório,



eletrocardiograma (ECG), rastreamento ocular, eletroencefalograma (EEG) e análise da fala passaram a ser incorporados no processo de reconhecimento de emoções. Considerando esses aspectos, este trabalho propõe um método para identificação de emoções a partir da análise combinada de expressões faciais e dados fisiológicos. Esse método é composto por uma arquitetura híbrida baseada em Rede Neural Convolucional (CNN) e Random Forest e consiste em um dos módulos de um sistema para suporte terapêutico personalizado baseado em biofeedback emocional. Os dados de expressões faciais utilizados foram obtidos a partir de uma combinação de importantes e diversificadas bases de dados no contexto de reconhecimento de emoções pela face: FER-2013 (GOODFELLOW et al., 2013), CFD (MA et al., 2015), KDEF (LUNDQVIST et al., 2008) and Yale Face (BELHUMEUR et al., 1997). Já os parâmetros fisiológicos associados a manifestação de emoções foram adquiridos a partir da base de dados MAHNOB-HCI (SOLEYMANI et al., 2011), a qual possui registros de sinais de EEG, ECG, GSR, amplitude respiratória e temperatura cutânea. Em um primeiro momento, todos esses dados foram submetidos a um processo de extração de atributos realizado a partir de uma CNN do tipo LeNet com 5 camadas e pré-treinada com a base MNIST (BALDOMINOS, 2019; KUO, 2016), cujo desempenho na classificação de imagem é amplamente reconhecido como positivo. Para os dados fisiológicos, que são originalmente na forma de sinais, foi realizada uma etapa anterior à extração de atributos que consistiu na conversão desses sinais em imagens a partir da Transformada de Wavelet. Para a análise combinada, esses dados foram submetidos a uma etapa de fusão, em que foram identificadas as emoções existentes em ambas as fontes de dados, o que resultou em 4 estados emocionais: Felicidade, Tristeza, Neutro e Raiva. A base de



conhecimentos proveniente desse processo de fusão dos dados possui 21.313 instâncias representadas por 1000 atributos, sendo 500 associados aos dados de faces e 500 aos fisiológicos. Após essa etapa, o conjunto completo foi dividido em duas amostras, sendo 70% para treinamento e 30% para teste do modelo de classificação. Visto que o conjunto de treinamento apresentou um forte desbalanceamento dos dados entre as classes de emoções, com a classe Tristeza composta por 4.754 instâncias, Felicidade com 7.007, Neutro com 5.488 e 4.064 instâncias da classe Raiva, esse conjunto foi submetido a uma etapa de balanceamento por subamostragem utilizando o algoritmo Spread Subsample, o que levou a um conjunto com 4.000 instâncias por classe. Então, o conjunto balanceado foi utilizado no treinamento de diversas configurações de Random Forest, variando a quantidade de árvores entre os valores de 10, 20, 50, 100, 150, 200, 250, 300, 350, 400, 450 e 500. Cada configuração foi treinada 30 vezes e foi validada a partir do método k-fold de validação cruzada com k igual a 10. A partir desse treinamento, o desempenho dos modelos foram comparados e identificou-se que a partir da configuração com 150 árvores, os modelos de Random Forest apresentaram comportamento estatístico semelhante entre si, com acurácia em torno de $93,79 \pm 0,63$, índice Kappa $0,9171 \pm 0,0083$, sensibilidade de $0,9942 \pm 0,0043$, especificidade de $0,9852 \pm 0,0037$, e área sob a curva ROC (AUC-ROC) de $0,9995 \pm 0,0004$. Por fim, no intuito de verificar a capacidade de generalização do modelo gerado a partir dessa configuração ele foi, então, utilizado para classificar os dados do conjunto de teste, que não havia sido utilizado até então. Essa etapa resultou em uma acurácia de 94,86%, além de Kappa, sensibilidade, especificidade e AUC-ROC com valores respectivos de 0,93, 0,95, 0,98 e 0,99. Considerando este alto desempenho, estudos futuros serão dedicados a incorporação deste

modelo em um sistema embarcado para atuar como um servidor web para processar os dados multimodais e exibir as informações emocionais ao terapeuta. Ele pode então ser usado para avaliar e personalizar a abordagem terapêutica e, assim, melhorar seus resultados e consequente benefício ao paciente em terapia.

Palavras-chave: Reconhecimento de emoções; Dados multimodais; Fusão de dados; Arquitetura híbrida.

Referências

- BALDOMINOS, A.; SAEZ, Y.; ISASI, P. A survey of handwritten character recognition with mnist and emnist. *Applied Sciences*, v. 9, n. 15, p. 3169, 2019.
- BELHUMEUR, P. N.; HESPANHA, J. P.; KRIEGMAN, D. J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, v. 19, n. 7, p. 711-720, 1997.
- COLOMBO, R.; RAGLIO, A.; PANIGAZZI, M.; MAZZONE, A.; BAZZINI, G.; IMARISIO, C.; et al. The SonicHand protocol for rehabilitation of hand motor function: A validation and feasibility study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, v. 27, n. 4, p. 664-672, 2019.
- MA, D. S.; CORRELL, J.; WITTENBRINK, B. The Chicago face database: A free stimulus set of faces and norming data. *Behavior research methods*, v. 47, n. 4, p. 1122-1135, 2015.
- GOODFELLOW, I. J.; ERHAN, D.; CARRIER, P. L.; COURVILLE, A.; MIRZA, M.; HAMNER, B.; et al. Challenges in representation learning: A report on three machine learning contests. In: *International conference on neural information processing*. Springer, Berlin, Heidelberg, 2013. p. 117-124.
- GONZÁLEZ, E. J. S.; MCMULLEN, K. The design of an

algorithmic modal music platform for eliciting and detecting emotion. In: 2020 8th international winter conference on brain-computer interface (bci). IEEE, 2020. p. 1-3.

KUO, C.-C. J. Understanding convolutional neural networks with a mathematical model. Journal of Visual Communication and Image Representation, 41:406–413, 2016.

LUNDQVIST, D.; FLYKT, A.; ÖHMAN, A. Karolinska directed emotional faces. Cognition and Emotion, 1998.

SOLEYMANI, M.; LICHTENAUER, J.; PUN, T.; PANTIC, M. A multimodal database for affect recognition and implicit tagging. IEEE transactions on affective computing, v. 3, n. 1, p. 42-55, 2011.

TEIXEIRA, T. W. D.; WANDERLEY, M. C.; GODONE, R. L. N. Medicina personalizada no tratamento do câncer. Brazilian Journal of Health Review, v. 3, n. 6, p. 18789-18793, 2020.

Downloads

Não há dados estatísticos.

Edição

v. 9 (2022): Mostra POLI/UPE
(<http://revistas.poli.br/index.php/anais/issue/view/38>)

Seção

Engenharia da Computação e Sistemas

Platform &
workflow by
OJS / PKP
)

