

UNIVERSIDADE FEEVALE

GUSTAVO ANDRÉ SETTI CASSEL

DETECÇÃO DE EXPRESSÕES FACIAIS COM INTELIGÊNCIA
ARTIFICIAL
(Título Provisório)

Anteprojeto de Trabalho de Conclusão

Novo Hamburgo
2020

GUSTAVO ANDRÉ SETTI CASSEL

DETECÇÃO DE EXPRESSÕES FACIAIS COM INTELIGÊNCIA
ARTIFICIAL
(Título Provisório)

Anteprojeto de Trabalho de Conclusão de
Curso, apresentado como requisito parcial
à obtenção do grau de Bacharel em
Ciência da Computação pela
Universidade Feevale

Orientador: Marta Rosecler Bez

Novo Hamburgo
2020

RESUMO

Esta pesquisa tem como objetivo geral propor um mecanismo para detectar e classificar, com alta acurácia, as expressões faciais universais definidas por Ekman em 1979. Atualmente, existem várias abordagens para resolver este problema englobando técnicas de Processamento Digital de Imagens (PDI) e, recentemente, mecanismos de Inteligência Artificial. Serão avaliados os principais *datasets* para detecção de expressões e então escolhido aquele a ser usado no estudo proposto, além de comparar mecanismos já existentes de PDI e inteligência artificial para projetar o protótipo de *software*. O que justifica e motiva este estudo é a grande aplicabilidade que existe em detectar expressões faciais, desde inferir o sentimento do aluno em uma aula EAD até identificar o possível sentimento de um comprador ao manipular produtos nos corredores de um mercado. Desta forma, quanto melhor a acurácia em classificar expressões faciais, mais as áreas recém citadas se beneficiarão. A metodologia de pesquisa escolhida foi *Design Science Research* (DSR) que é propícia para integrar aspectos teóricos com o desenvolvimento de software. Serão consideradas as seis etapas definidas pela DSR: identificação do problema e motivação; definição dos objetivos para a solução; projeto e desenvolvimento; demonstração; avaliação; comunicação. A literatura pesquisada teve como foco artigos, teses e dissertações dos anos de 2015 a 2020 em busca do estado da arte neste assunto. Alguns conceitos básicos de expressões faciais foram encontrados em redações antigas, como literaturas de Ekman datadas no fim dos anos 70. O resultado esperado é um software com alta acurácia na detecção e classificação de expressões faciais definidas por Ekman (raiva, nojo, medo, alegria, neutra, tristeza e surpresa).

Palavras-chave: Detecção de expressões faciais. Ekman. Inteligência artificial. Redes neurais. Processamento digital de imagens.

SUMÁRIO

MOTIVAÇÃO	5
OBJETIVOS	7
METODOLOGIA	8
CRONOGRAMA	9
BIBLIOGRAFIA	10

MOTIVAÇÃO

Sabe-se que cada vez mais a Inteligência Artificial faz parte do dia a dia das pessoas das mais variadas maneiras. Dentre as diversas aplicações, algumas podem ser citadas, como análise de padrões em *Big Data*, *chatbots*, assistentes pessoais, predições, sistemas de recomendação e busca de padrões em imagens.

A análise de imagens a fim de identificar padrões pode ser feita com técnicas de PDI (processamento digital de imagens) para a filtragem, segmentação e melhoria das imagens, seguida de técnicas de inteligência artificial para a classificação, extração de características e obtenção de informações a partir delas.

O reconhecimento de expressões faciais pode ser usado para os mais variados fins. É possível citar, dentre eles, o sentimento de um aluno em uma aula EAD para que o professor tenha noção se a aula está agradável; a aplicação de câmeras que detectam o sentimento do consumidor ao verificar e manipular produtos da loja; detecção de expressões faciais para detectar comportamentos suicidas; auxílio clínico para diagnosticar precocemente distúrbios psicológicos com base em expressões faciais; entre outros (ALHUSSEIN, 2016) (AMICO *et al.*, 2016) (JAIMES, SEBE, 2007). Nos últimos anos a pesquisa vem avançando consideravelmente nesta área, com as mais variadas abordagens.

É proposto por (MAYYA *et al.*, 2016) um modelo de rede neural convolucional profunda - *Deep Convolutional Neural Network* (DCNN) - para análise de expressões faciais. O estudo é feito através do framework *Caffe*, com dois *datasets* de imagem: *Japanese Female Facial Expression* (JAFFE) e CK+. O processamento é realizado de forma paralela com o uso da GPU para ganho de performance. A rede neural possui 5 camadas, contando com uso da função de ativação ReLU e do *max-polling* para evidenciar as bordas do rosto e permitir a extração de suas características. Em um segundo momento, a classificação da expressão facial ocorre com *Support Vector Machines* (SVM). Os resultados alcançados são promissores para as expressões de raiva, nojo, medo, alegria, neutra, tristeza e surpresa, chegando a 98,12% de acurácia com o *dataset* JAFFE. Com CK+ os resultados também foram muito bons e chegaram a 96,02% contemplando expressões de raiva, desprezo, nojo, medo, alegria, tristeza e surpresa.

JAFFE é um *dataset* que disponibiliza 213 imagens em tons de cinza de 10 mulheres japonesas apresentando 7 distintas expressões faciais: raiva, nojo, medo, alegria, neutra, tristeza e surpresa. A resolução de cada imagem é de 256 x 256 pixels (LYONS *et al.*, 1998).

Convolutional Architecture for Fast Feature Embedding (Caffe) é um *framework* para criação de algoritmos para inteligência artificial com aprendizado de máquina (JIA *et al.*, 2014). Permite construir a rede neural e formar cada uma de suas camadas a partir de arquivos de configuração, não sendo obrigatório o desenvolvimento de código-fonte para elaborar a rede neural em questão.

O estudo proposto por (ZHOU *et al.*, 2016) também é feito com base no *dataset* JAFFE. Chegou ao acerto de 98,1% das classificações, através de métodos *Gabor Wavelet Transform*, fatoração de matrizes não-negativas e um classificador de duas camadas.

O estudo proposto por (ALHUSSEIN, 2016) é, assim como os demais estudos anteriormente apresentados, feito com o *dataset* de expressões faciais JAFFE, mas também com o *dataset* Cohn–Kanade (CK). É investigado principalmente o uso do algoritmo *Weber Local Descriptors* (WLD) que divide a imagem em inúmeros blocos, os quais são posteriormente delegados a uma *Support Vector Machine* (SVM) para a classificação final. Os resultados alcançados para os *datasets* JAFFE e CK são de, respectivamente, 97% e 98,82%.

Este trabalho é de grande importância pois a detecção automática de expressões faciais pode ser aplicada em diversas áreas. No momento em que novos mecanismos são propostos ou mecanismos já existentes são aprimorados, diversas áreas se beneficiam com tais melhorias. Sendo assim, este trabalho se propõe a desenvolver e validar um protótipo de *software* para identificar com alta acurácia as expressões faciais básicas (raiva, nojo, medo, alegria, neutra, tristeza e surpresa) definidas por Ekman (1979).

OBJETIVOS

Objetivo geral

Desenvolver e validar um protótipo de *software* para identificar as expressões faciais básicas (raiva, nojo, medo, alegria, neutra, tristeza e surpresa) definidas por Ekman.

Objetivos específicos

- Estudar a literatura sobre o estado da arte na detecção de expressões faciais;
- Escolher a linguagem de programação mais apropriada para análise de imagens;
- Avaliar os principais *datasets* usados para detecção de expressões e selecionar o que será usado no estudo proposto;
- Comparar mecanismos já existentes de PDI e inteligência artificial para projetar o protótipo de *software*;
- Desenvolver um protótipo para a detecção de expressões faciais;
- Validar o protótipo desenvolvido.

METODOLOGIA

A metodologia usada para o desenvolvimento desta pesquisa será a *Design Science Research* (DSR). Esta metodologia é composta por seis etapas principais que permitem avaliar a eficácia dos resultados e publicá-los na comunidade.

Segundo JUNIOR *et al.* (2017), baseado em PEFFERS (2007), as etapas são descritas abaixo juntamente com a aplicação no presente estudo de diagnóstico de expressões faciais:

1. **Identificação do problema e motivação:** buscar na literatura os principais mecanismos já existentes para detecção de expressões faciais. Como motor de busca será usado o *Web of Science*, que consolida artigos e materiais provenientes de várias origens de dados;
2. **Definição dos objetivos para a solução:** definir a linguagem de programação na qual o software será desenvolvido e quais requisitos o *software* terá. Também definir uma ou mais bases de dados com imagens de expressões faciais para posteriormente avaliar os resultados do *software*;
3. **Projeto e desenvolvimento:** desenvolver o protótipo de *software* em si, capaz de reconhecer as expressões faciais básicas (raiva, nojo, medo, alegria, neutra, tristeza e surpresa) definidas por Ekman;
4. **Demonstração:** executar o *software* utilizando uma ou mais bases de dados com imagens de expressões faciais, apresentando o percentual de acurácia para cada expressão facial básica definida por Ekman;
5. **Avaliação:** comparar, com soluções elaboradas por outros pesquisadores, o *software* desenvolvido na presente pesquisa. Será avaliada a acurácia da detecção de expressões faciais e, possivelmente, a performance em obter tais resultados;
6. **Comunicação:** além da apresentação final do TCC, os resultados serão possivelmente publicados em congressos científicos da computação.

CRONOGRAMA

Trabalho de Conclusão I

Etapa	Meses			
	Mar	Abr	Mai	Jun
Identificação do problema e motivação				
Buscar na literatura mecanismos existentes para reconhecimento de expressões faciais				
Escrita do TCC I				
Definição dos objetivos para a solução				
Busca de bases de imagens				
Definição linguagem de programação				
Projeto e desenvolvimento				
Desenvolvimento do software				
Entrega TCC I				

Trabalho de Conclusão II

Etapa	Meses			
	Ago	Set	Out	Nov
Projeto e desenvolvimento				
Desenvolvimento do software				
Escrita do TCC II				
Demonstração				
Execução do software com bases de imagens de expressões faciais				
Avaliação				
Avaliar acurácia do software comparando com resultados apresentados em outras literaturas				
Escrever/apresentar resultados no documento do TCC				
Comunicação/Entrega TCC II				
Apresentar os resultados encontrados				

BIBLIOGRAFIA

AMICO, Francesco. *et al. Multimodal validation of facial expression detection software for real-time monitoring of affect in patients with suicidal intent. European Psychiatry*, 2016.

ALHUSSEIN, Musaed. *Automatic facial emotion recognition using weber local descriptor for e-Healthcare system. Cluster Comput*, 2016. c. 19, p. 99–108.

EKMAN, Paul; OSTER, Harriet. *Facial expressions of emotion. Annual Review of Psychology*, 1979. v. 30, p. 527-554.

JAIMES, Alejandro; SEBE, Nicu. *Multimodal Human Computer Interaction: A Survey. Computer Vision and Image Understanding*, 2007. v. 108, p. 116-134.

JIA, Yangqing et al. *Caffe: Convolutional architecture for fast feature embedding. Proceedings of the 2014 ACM Conference on Multimedia*, 2014.

JUNIOR, Vanderlei *et al. Design Science Research Methodology As Methodological Strategy for Technological Research. Revista Espacios*, 2017. vol. 38, p. 25.

LYONS, Michael; KAMACHI, Miyuki; GYOBA, Jiro. *The japanese female facial expression (JAFFE) database*. 1998. Disponível em: <<https://zenodo.org/record/3451524#.XmBC63VKi0k>>. Acesso em: 04 mar 2020.

LYONS, Michael J. *et al. Coding Facial Expressions with Gabor Wavelets. Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, 1998. p. 200-205. Disponível em: <<http://doi.org/10.1109/AFGR.1998.670949>>. Acesso em: 04 mar 2020.

MAYYA, Veena; PAI, Radhika; M M, Manohara. *Automatic Facial Expression Recognition Using DCNN. Procedia Computer Science* 93, 2016. p. 453-461.

PEFFERS, Ken *et al.* ***A Design Science Research Methodology for Information Systems Research.*** *Journal of Management Information Systems*, 2007. v. 24, 3ed., p. 45–78.

ZHOU, Jun *et al.* ***A method of facial expression recognition based on Gabor and NMF.*** *Pattern Recognition and Image Analysis*, 2016. v. 26, p. 119–124.