

Aplicação de IA com Plataforma GPU no Reconhecimento de Expressões Faciais*

Leandro P. Heck¹, Cristiano A. Künas¹, Edson L. Padoin¹

¹Universidade Regional do Noroeste do Estado do Rio Grande do Sul (UNIJUI)
Santa Rosa – RS – Brasil

{leandro.h, cristiano.kunas}@sou.unijui.edu.br,
padoin@unijui.edu.br

Resumo. *Este artigo apresenta uma proposta de paralelização em Plataforma GPU de uma rede neural aplicada no Reconhecimento de Expressões Faciais. A partir da análise dos pixels de uma imagem e reconhecimento de padrões da face, será possível reconhecer as emoções que os indivíduos estão sentindo. Pretende-se utilizar a biblioteca PyCuda, que permite programar e acessar a API CUDA da NVidia a partir da linguagem Python.*

1. Introdução

Atualmente existe um crescente interesse na melhoria da interação entre humanos e computadores. Para que uma efetiva interface humano-computador inteligente seja alcançada faz-se necessário que o computador relaciona-se naturalmente com os usuários, semelhante à maneira que os humanos interagem [Leão et al. 2012]. Um das maneiras de encontrar e promover esse tipo de integração é com auxílio da Computação Afetiva, Inteligencia Artificial, com *Machine Learning* e *Deep Learning*.

Fazer com que uma máquina reconheça, modele e expresse emoções não é uma tarefa simples. Quando seres humanos interagem entre si, boa parte dessa interação é baseada na linguagem verbal e na utilização da linguagem corporal por meio de gestos e expressões faciais que carregam e transmitem as emoções dos interlocutores [Leão et al. 2012].

Nas últimas décadas a comunidade científica vem tendo um crescente interesse no reconhecimento de emoções. Existem atualmente diversas maneiras de expressar as emoções humanas, estas vêm sendo estudadas e melhoradas ao longo dos anos. Muitas fontes de dados têm sido exploradas, dentre elas destacam-se textos, envio de *emoticons*, voz e as expressões faciais. Atualmente, trabalhos sobre esse tema investigam inúmeras técnicas e maneiras de extração de características e classificação de padrões em expressões faciais [do Amaral et al. 2013], como por exemplo, o uso das Redes Neurais.

Nesse sentido, o presente artigo apresenta uma proposta de utilização de uma Rede Neural Artificial (RNA) aplicada no reconhecimento de expressões faciais. Seu objetivo é diferenciar os diferentes sentimentos que os indivíduos expressam a partir dos padrões. Sendo uma área de grande relevância para consumidores e organizações em áreas de saúde, marketing, cinema dentre outras. No marketing, sendo possível entender o que os clientes sentem ao consumir um produto, no cinema analisar as reações das pessoas diante as cenas de um filme. O restante do trabalho esta organizado da seguinte forma. A Seção 2 discute os trabalhos relacionados. A

*Trabalho desenvolvido com recursos do edital MCTIC/CNPq - Universal 28/2018 sob número 436339/2018-8 e do edital da VRPGPE bolsa PIBIC/UNIJUI.

seção 3 são discutidos as propostas para a implementação da Rede Neural Artificial. Na Seção 4 é apresentada a metodologia que será utilizada na implementação e o ambiente de execução para realização dos testes, seguidos das Conclusões e Trabalhos Futuros.

2. Trabalhos Relacionados

No trabalho de [de Melo et al. 2014] é apresentado um estudo comparativo entre três técnicas computacionais - Redes Neurais tipo MLP, Rede Neural Artificial tipo RBF e Redes Bayesianas - com o objetivo de analisar qual a melhor técnica para a classificação de emoções baseada na teoria *Facial Action Coding System (FACS)* [Ekman and Friesen 1978]. A partir de uma análise ao Banco de Dados composto por inúmeras imagens relacionadas a inferência das emoções com diversidade de etnias, sexo, idade [Kanade and Cohn 2005]. é elaborado uma base de dados composta pelo agrupamento das UA's (Unidade de Ações) relacionada a determinada emoção e, por fim são realizados experimentos para a classificação das emoções. Os resultados das diferentes técnicas foram comparados e obtiveram um precisão geral de 86%, 75, 72% e 71, 76% para as Redes Bayesianas, Rede Neural Artificial tipo MLP e Rede Neural Artificial tipo RBF, respectivamente.

O trabalho de [Azcarate et al. 2005] apresenta um sistema de reconhecimento de emoções através de expressões faciais que são exibidas ao vivo em vídeo. O sistema baseia-se no *Piecewise Bézier Volume Deformation (PBVD)* e foi estendido para um detector rosto, usado para localizar, inicialmente o rosto humano automaticamente. Os experimentos usando *Naive Bayes* e classificadores *Tree-Augmented-Naive* [Azcarate et al. 2005]. Além disso, o autor associou Unidades de Movimentos com as Unidades de Ação proposta pela teoria FACS. Os testes foram feitos com as pessoas dependente do banco de dados Cohn-Kanade [Kanade and Cohn 2005] e com pessoas independente da base de dados utilizada para treinamento. Os resultados obtidos foram satisfatórios para o reconhecimento das emoções através da expressão facial com as UA's.

No trabalho de [Rao and Koolagudi 2015] foi elaborado um sistema de reconhecimento de emoções utilizando características acústicas e faciais. Com os experimentos, chegou-se a conclusão que tal combinação proporciona uma melhora no desempenho do reconhecimento em um nível de 93, 62%. O desempenho do sistema de reconhecimento de emoções desenvolvido utiliza modelos de redes neurais com camada oculta, comparado com Modelos de Markov. Os recursos e os modelos propostos são avaliadas em banco de dados interativos de emoções.

Diferente dos trabalhos aqui apresentados, a nossa proposta busca desenvolver uma Rede Neural Artificial que reconheça expressões faciais, e a partir destas identificar a emoção presente em uma determinada imagem. Propondo a paralelização de uma Rede Neural em arquitetura GPU, visando obter melhor desempenho, tanto no seu treinamento, quanto na análise e processamento de suas entradas. Para que seja possível obter um resultado satisfatório no menor tempo possível.

3. Proposta

Para a aplicação de RNA no reconhecimento de expressões faciais será desenvolvido um sistema que atenderá a *Requisitos Funcionais* e *Requisitos Não Funcionais*. Os Requisitos Funcionais possuem as seguintes ações: i) detecção da face em uma imagem; ii) criação da rede neural; iii) aplicação de Aprendizado Profundo; e iv) classificação das emoções nas expressões faciais. Nos Requisitos Não Funcionais estão previstos o desenvolvimento de um protótipo utilizando a

linguagem Python. Esta linguagem foi selecionada pela disponibilidade dos Frameworks Keras e TensorFlow para aceleradores da NVIDIA, bem como o aprendizado que pode ser utilizado com o framework Opencv para analisar as imagens.

Mais especificamente, nos Requisitos Funcionais será realizada a implementação da RNA que permitirá reconhecer e analisar as diferentes expressões faciais, a partir das informações extraídas de imagens. Ou seja, identificar se um indivíduo está, por exemplo *alegre, triste, irritado, surpreso* ou *com medo*. Para que a Rede Neural consiga realizar esta tarefa é preciso treiná-la com uma quantidade grande de imagens de faces. Assim sendo, numa primeira etapa será criada uma base de dados de imagens com expressões faciais.

4. Metodologia

Para o desenvolvimento do protótipo da RNA foi elaborado o seu Diagrama de Caso de Uso apresentado na Figura 1.

No caso de uso **C01 - Classificar emoções** terá a função de disponibilizar ao usuário possibilidade de classificar automaticamente as emoções em imagens que possuam expressões faciais. Para realizar, este procedimento é necessário realizar a importação das imagens para o protótipo, ação que é realizada pelo caso de uso **C02 - Carregar imagens**. O caso de uso C02 permite ao usuário informar o caminho de uma imagem, ou informar o diretório das imagens, as quais deseja realizar a classificação.

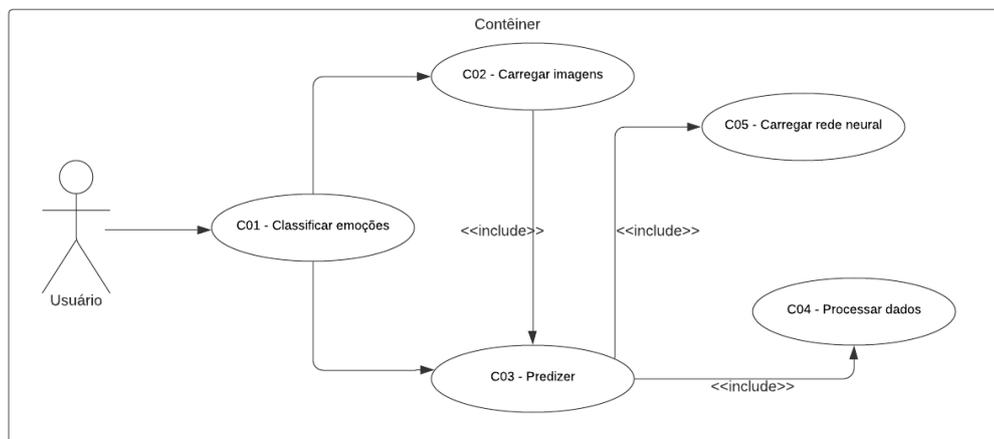


Figura 1. Diagrama de Caso de Uso do Protótipo.

O caso de uso **C03 - Predizer** informa ao usuário qual a emoção predita para a imagem, ou imagens, que foi especificado. Para que o caso de uso C03 realize sua função é necessário que a imagem, ou imagens, já devem ter sido carregadas no protótipo. O prosseguimento do caso de uso C03 é dado pela execução da ação dos casos de uso **C04 - Processar** e **C05 - Carregar rede neural**. Sendo que o caso de uso C04 realiza a etapa de formatação da estrutura dos dados da imagem, extraíndo os pixels e estruturando na forma compatível com a entrada exigida pela rede neural profunda. O caso de uso C05 carrega a rede neural profunda e a configura com os pesos do melhor treinamento disponível.

Nas etapas de aprendizado e treinamento da Rede Neural que demandam maior processamento, será aplicadas técnicas de paralelismo, para melhorar o desempenho da Rede. Dependendo da quantidade de imagens estas etapas podem demandar elevados tempos até que

a Rede Neural consiga aprender e reconhecer os padrões. Deste modo, com a paralelização destas etapas, pretende-se agilizar o processo de aprendizado e treinamento da Rede Neural, proporcionando assim um grande ganho de tempo e performance. Pretende-se utilizar a biblioteca PyCuda, que permite acessar a API de computação paralela CUDA da Nvidia a partir da linguagem de programação Python [Klöckner et al. 2012].

O algoritmo está em desenvolvimento e está sendo testado em um equipamento com o Sistema Operacional Linux Ubuntu 18.04.3 LTS com kernel 5.0.0-37-generic em conjunto com um processador Intel Core i7-9750 2.60GHz com 6 cores (12 threads), 16 GB de Memória RAM, acelerador NVIDIA GeForce GTX 1660 Ti com 6 GB GDDR6 e 1536 CUDA cores.

5. Conclusões e trabalhos futuros

Este trabalho apresenta uma proposta de utilização de Rede Neural Artificial aplicada no reconhecimento de expressões faciais e extração dos padrões presentes em imagens. Espera-se que a proposta apresente elevada precisão na identificação de diferentes sentimentos em indivíduos a partir da paralelização das etapas de aprendizado e treinamento.

Como trabalhos futuros, pretende-se finalizar a implementação do algoritmo sequencial e paralelo e iniciar o treinamento, teste, validação e análise de desempenho da aplicação da Rede Neural. Espera-se alcançar *speed-up* com a versão paralela do algoritmo utilizando diferentes modelos de GPU.

Referências

- Azcarate, A., Hageloh, F., Van de Sande, K., and Valenti, R. (2005). Automatic facial emotion recognition. *Universiteit van Amsterdam*, pages 1–6.
- de Melo, S. L., de Moura, F. F., Macedo, K., Alves, F. S., and Fernandes, M. A. (2014). Estudo comparativo de técnicas computacionais para classificação de emoções. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 25, page 456.
- do Amaral, V., Giraldi, G. A., and Thomaz, C. E. (2013). Lbp estatístico aplicado ao reconhecimento de expressões faciais.
- Ekman, P. and Friesen, W. (1978). Manual of the facial action coding system (facs). *Trans. ed. Vol. Consulting Psychologists Press, Palo Alto*.
- Kanade, T. and Cohn, J. (2005). Au-coded facial expression database.
- Klöckner, A., Pinto, N., Lee, Y., Catanzaro, B., Ivanov, P., and Fasih, A. (2012). PyCUDA and PyOpenCL: A Scripting-Based Approach to GPU Run-Time Code Generation. *Parallel Computing*, 38(3):157–174.
- Leão, L. P., Bezerra, J. S., Matos, L. N., and Nunes, M. A. S. N. (2012). Detecção de expressões faciais: uma abordagem baseada em análise do fluxo óptico. *Revista GEINTEC-Gestão, Inovação e Tecnologias*, 2(5):472–489.
- Rao, K. S. and Koolagudi, S. G. (2015). Recognition of emotions from video using acoustic and facial features. *Signal, Image and Video Processing*, 9(5):1029–1045.