

# Projeto de Iniciação Científica

## Modelo de Rede Neural Baseado em Estatística Não-Extensiva

Aluno: Thiago Andrade do Nascimento

Orientador: Paulo Sérgio Silva Rodrigues

Período: Novembro/2019 a Outubro/2020 (12 Meses)

Instituição: Centro Universitário FEI

São Bernardo do Campo, SP

Fevereiro de 2019

## Resumo

Dentro do campo Visão Computacional, a segmentação de imagens é um processo considerado de nível médio. Entretanto, sabe-se que esse tipo de análise inclui também extração de regiões de interesse, bordas e primitivas geométricas. Por outro lado, sabe-se que uma das etapas fundamentais tanto de redes neurais quanto de redes convolucionais são as chamadas funções de ativação. Entretanto, sabe-se de estudos da neurociência [12] [9] [2] que o comportamento de estruturas neurais complexas são melhores representadas por funções de gabor, que são funções gaussianas envelopadas por funções senoidais, permitindo assim um decaimento caudal mais similar a processos biológicos reais. Também, considerando o sucesso de teorias estatísticas como a entropia não-extensiva de Tsallis, neste projeto, propomos o uso da função de gabor como função de ativação de redes neurais e comparamos com funções tradicionais como sigmóides e gaussianas, aplicadas à classificação de imagens médicas de tomografia cerebral em uma base de dados anotada manualmente.

Palavras-Chaves: Redes-Neurais; Classificação de Imagens Médicas; Gabor; Sigmoides; Gaussiana.

## 1 Introdução

Aprendizado de máquina são métodos já conhecidos há bastante tempo, no entanto, só recentemente ganharam fôlego devido ao desenvolvimento recente dos hardwares e softwares que impulsionaram a aplicação direta de modelos matemáticos-computacional com sucesso, os quais já haviam a várias décadas sido propostos. Além disso, a grande disponibilidade de dados e a necessidade crescente dos mesmos serem gerenciados também é considerado como parte desta composição impulsionadora. Uma ramificação deste método são as redes convolucionais, compostas por métodos modernos derivados dessa tecnologia e frequentemente usadas em diversas áreas [13], sendo a medicina um dos campos mais demandantes da técnica automática de IA [18] [1].

Esta necessidade crescente é devida ao grande volume de informações digitais disponibilizadas pelo campo da medicina. Consequentemente, novos modelos de aplicações para solução desses problemas têm sido criados, entre eles, encontra-se a estatística não-extensiva que, dessa forma, impulsionou novo interesse em diversos espaços científicos,

justamente pelo sucesso na solução de desafios a partir desta nova abordagem.

Excepcionalmente, em processamento digitais de imagens, sobretudo para processamento de baixo, médio e alto nível, diversas funções baseadas em estatística não-extensiva foram utilizadas nos últimos anos em modelos matemáticos computacionais [14] [15], sendo as funções gaussianas, sigmóides, entropias e funções gabor compondo este restrito grupo. Em particular, a última apresentada (gabor [9]), é derivada de uma função que se encontra na abordagem da estatística tradicional, sendo esta, capaz de simular as interconexões neurais [12].

O crescente volume de imagens médicas induz à necessidade de mais controle e de soluções mais precisas. Para isso, existem tratamentos de imagens específicos sobre dados volumétricos, sob baixa resolução e ruído, que necessita de hardware especificado que tentam solucionar as urgências criadas por esse aumento exponencial [11] [17].

Apesar do grande avanço tanto do hardware quanto software, ainda é necessária adaptação destas redes para obterem melhor desempenho com pouco custo de hardware e software. Sendo assim, redes convolucionais e redes neurais demandam por melhorias em algumas formas do seu funcionamento. Em particular, a função de ativação que “guia” o aprendizado das mesmas e auxilia na velocidade de aprendizagem.

Assim, a estatística não-extensiva é uma solução viável, fácil de implementar e processar em algumas áreas. Sendo assim, neste projeto propomos o uso da estatística não-extensiva de Tsallis sob a forma de novas funções propostas pela literatura e aplicadas a modelos usados em redes neurais, com o objetivo de classificação de imagens médicas. Essas funções consideradas são: sigmoide, gabor, exponencial e gaussiana, batizadas na literatura de q-sigmoid, q-gabor, q-exponencial e q-gaussiana [3] [4].

## 1.1 Objetivo

Propor e implementar novas funções de ativação baseadas na estatística não-extensiva de Tsallis em redes neurais aplicadas à classificação de imagens médicas.

## 2 Redes Neurais

Redes neurais são modelos matemáticos-computacionais cujo objetivo é a simulação de conexões neuronais similares a organismos biológicos, através de camadas compostas para aplicações de diversas abordagens, tais como multiplicações de matrizes, derivadas, e funções não-lineares ou lineares.

Sua estrutura básica possui uma camada de entrada (que recebe os valores de análise), outras ocultas (fazem o tratamento de dados para a previsão) e uma de saída (previsões feitas pelo modelo). Estas camadas apresentam aglomerados de nós que, para cada um destes, existe um respectivo peso de ponderação.

No entanto, o funcionamento das camadas ocultas e de saída possuem, além dos pesos ponderados, função somatória, e funções de ativação (sigmóides, tangentes hiperbólicas, ReLu's, etc.). Resumidamente, a função somatória é aplicada sobre todos os dados anteriores multiplicados pelo seu peso. Este resultado é enviado à função de ativação que ativa ou não o nó sendo estudado no momento, gerando assim uma relevância daquela saída para a próxima camada.

Atualmente, as funções mais utilizadas são a *ReLu* [19] ( $f(x) = \max(x, 0)$ ,  $f(x) = 0$ , se, e somente se,  $x < 0$ ; e  $f(x) = x$ , se, e somente se,  $x \geq 0$ ) e sigmoid (retorno no intervalo entre 0 e 1). Quando a camada de saída produz a previsão final, o erro entre a saída esperada e a obtida é calculado através de uma função de perda (*Mean Squared Error*, *Absolute Error*, *Binary Cross Entropy*, entre outros). Desta forma, inicia-se a descida do gradiente que, através de funções otimizadoras ([10] está entre os otimizadores mais utilizado na literatura) calculam derivadas parciais da função de perda em relação aos pesos. Sendo assim, quando o vetor gradiente é obtido os pesos são modificados na direção oposta deste vetor visando a minimização do erro.

Após essa etapa, os pesos são modificados da mesma forma voltando em direção à camada de entrada, o que caracteriza o conceito de *Backpropagation*.

Portanto, este projeto visa utilizar o conceito sobre as funções de ativação. Assim, implementando a entropia de Tsallis neste setor das redes neurais e analisando as consequências dessa aplicação. Assim, a substituição nessas etapas por novas funções de ativação, baseada na estatística não-extensiva de Tsallis, permite maior flexibilização paramétrica e consistente convergência mais rápida e mais precisa, sendo esta a hipótese

deste trabalho.

Além disso, serão utilizadas primordialmente redes neurais [8] (compostas apenas por camadas conectadas através de nós).

### 3 Base de Dados

A base de dados utilizada para testes e treinamentos são as imagens médicas com padrão ouro da base BRATS 2013. Esse conjunto de dados é oriundo de imagens médicas da região cerebral, os quais são divididos nas subclasses com tumor ou sem tumor. Tais imagens possuem resolução de  $128 \times 256$  pixels.

Por outro lado, a literatura vêm utilizando com mais frequência essa base de dados, é possível notar isso quando compara-se o crescente uso dessa base em relação às comumente usadas, o que pode ser observado nas Figuras 1 e 2.

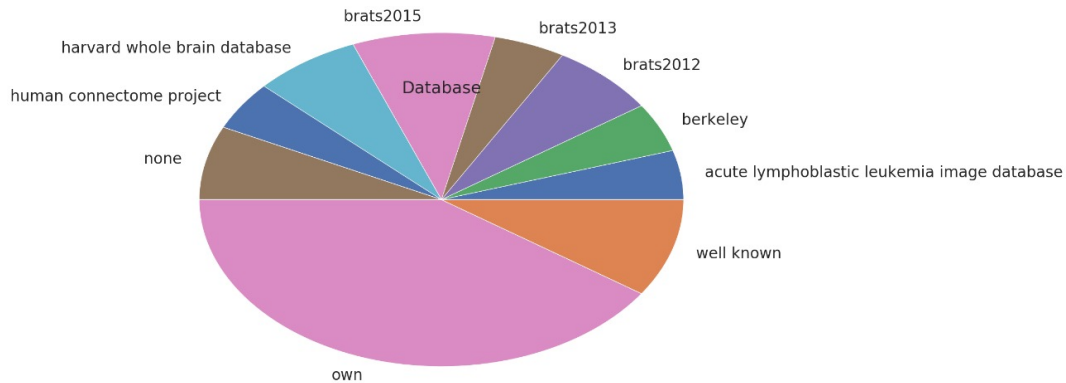


Fig. 1: Porções utilizadas para cada base de dados

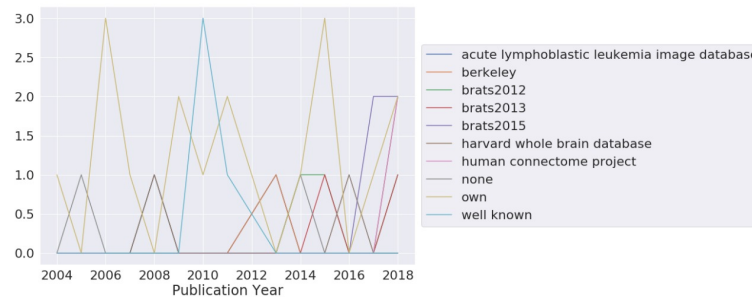


Fig. 2: Uso ao longo do anos

## 4 Funções de ativação

Serão implementadas as seguintes funções de ativação:

### 4.1 A Função Gabor 1D

Funções Gabor tradicionais podem ser utilizadas como filtros para sinais unidimensionais. Uma função de Gabor  $g$  no domínio complexo é definida pelo produto de um envelope gaussiano  $w$  com uma sinusoidal  $s$  no domínio complexo, como observado na seguinte Equação (1).

$$g(X) = ke^{\theta i} w(\alpha X) s(X) \quad (1)$$

onde:

$$w(X) = e^{-\pi X^2} \quad (2)$$

$$s(X) = e^{(2\pi f X)i} \quad (3)$$

onde  $\alpha$ ,  $k$ ,  $\theta$  e  $f$  são os parâmetros da função e  $X$  é um vetor.

### 4.2 q-Sigmoid

A  $q$ -sigmoid é uma versão generalizada da função sigmoide, baseada na entropia não-extensiva de Tsallis. Sendo assim, além dos parâmetros convencionais  $\alpha$  e  $\beta$ , um novo parâmetro adicional  $q$  é utilizado para definir com maior precisão a topologia da curva. Também, a constante  $\lambda$  pondera a distância entre uma intensidade de pixel  $I$  e o parâmetro  $\beta$ . O detalhamento da função sigmoide, bem como de sua generalização podem ser encontrados em [16]

## 5 Metodologia proposta

A ideia geral de metodologia proposta nesse trabalho é composta de 4 etapas offline e 4 etapas online, descritas na Figura 3.

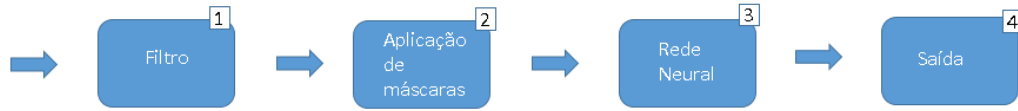


Fig. 3: Metodologia Proposta

## 5.1 Etapa 1

Nesta primeira etapa será realizada uma convolução matemática para filtragem de regiões de interesse da imagem, com os filtros tradicionais [6] [7] .

## 5.2 Etapa 2

Nesta etapa serão utilizadas máscaras binárias, disponíveis na base BRATS, sobre a imagem de saída da etapa 1. Essas máscaras delimitam as regiões de interesse de modo a guiar de maneira mais delimitada a Etapa 3 a seguir.

Além disso, essa etapa utiliza cada máscara binária para extrair da região correspondente na imagem original, informações de textura [5]. As seguintes informações serão inicialmente estudadas: probabilidade máxima, momento da diferença do elemento da ordem  $k$ , inverso do momento da diferença da ordem  $k$ , uniformidade e entropia.

Portanto, a saída desta etapa é um espaço vetorial de informação, sobre as texturas, para cada região conectada.

## 5.3 Etapa 3

Esta etapa recebe cada conjunto de vetores correspondentes às regiões de interesse e imputa à rede neural para treinamento. Na figura 4, encontra-se uma das possíveis topologias a serem abordadas.

A arquitetura mínima de uma rede neural densa, pode ser composta por apenas uma camada oculta, de saída e de entrada. Enquanto, a oculta poderá possuir de 100 à 1200

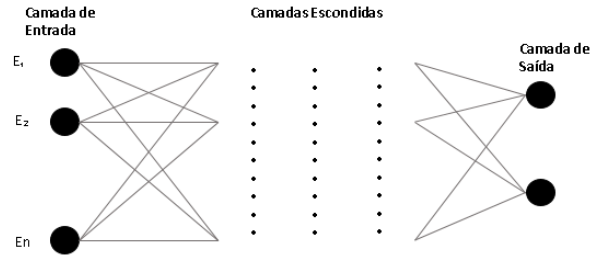


Fig. 4: Possível topologia que será utilizada.  $N$  corresponde a quantidade de componentes no campo vetorial.

nós, que poderão possuir as funções de ativação já citadas (*Relu*, *Gabor*, *q – Sigmoides*, entre outras), e a de saída é abordada a seguir.

## 5.4 Etapa 4

Esta etapa é composta da última camada da rede apenas, onde também são utilizadas funções de ativação e um limiar que define a classificação binária do tumor.

## 6 Materiais e Equipamentos Disponíveis



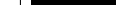


Estão disponíveis para este projeto os seguintes recursos: Cluster SumHPC, Cluster Altix-XE e Cluster Santos Dumont, do Laboratório Nacional de Computação Científica (LNCC); além do Cluster Titânio da Universidade Federal do ABC. Esses recursos representam os principais disponíveis para pesquisa no Brasil.

## 7 Cronograma de Execução

O projeto pode ser dividido em cinco 5 etapas, executadas da seguinte forma, conforme apresentado na Tabela 1.



Tab. 1: Cronograma das etapas que serão executadas neste projeto

Nº	Tarefa	2019										2020		
		A	M	J	J	A	S	O	N	D	J	F	M	
1	Estudo da bibliografia relacionada													
2	Preparação do hardware													
3	Implementação dos algoritmos propostos													
4	Análise dos resultados													
5	Publicação dos resultados													

## Referências

- [1] Francesco Ciompi and Bartjan de Hoop. Automatic classification of pulmonary perifissural nodules in computed tomography using an ensemble of 2d views and a convolutional neural network out-of-the-box. *Medical image analysis*, 26 1:195–202, 2015.
- [2] Vladímir de Aquino Silveira, Givago da Silva Souza, Bruno Duarte Gomes, Anderson Raiol Rodrigues, and Luiz Carlos de Lima Silveira. Joint entropy for space and spatial frequency domains estimated from psychometric functions of achromatic discrimination. In *PloS one*, 2014.
- [3] C. D. Gallao. *Definicao da Funcao q-Gaussiana Bidimensional com Aplicacoes em Processamento Digital de Imagens*. Tese de doutorado, Centro Universitario da FEL, Sao Bernardo do Campo, AP, 2016.
- [4] Celso Gallao and Paulo Silva Rodrigues. A q-gaussian spatial filtering. In *Proc. of the XI Workshop de Visao Computacional, WVC, 2015*, Sao Carlos, SP, Brazil, October 2015.
- [5] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 2001.
- [6] Kazufumi Ito and Kaiqi Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Trans. Automat. Contr.*, 45:910–927, 2000.
- [7] Anil K. Jain and Farshid Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24:1167–1186, 1991.
- [8] Anil K. Jain, Jianchang Mao, and K. Moidin Mohiuddin. Artificial neural networks: A tutorial. *IEEE Computer*, 29:31–44, 1996.
- [9] Jeffrey P. Jones and Larry A. Palmer. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of neurophysiology*, 58 6:1233–58, 1987.
- [10] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2015.

- [11] Dzung L. Pham, Chenyang Xu, and Jill L. Hanna Prince. Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2:315–37, 2000.
- [12] Anqi Qiu, Christoph E. Schreiner, and Monty A. Escabí. Gabor analysis of auditory midbrain receptive fields: spectro-temporal and binaural composition. *Journal of neurophysiology*, 90 1:456–76, 2003.
- [13] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39:1137–1149, 2015.
- [14] P. S. Rodrigues, G. A. Wachs-Lopes and H. R. Erdmann, M. P. Ribeiro, and G. A. Giraldi. Improving a firefly meta-heuristic for multilevel image segmentation using tsallis entropy. *Pattern Analysis and Applications*, pages 1–20, 2015.
- [15] Paulo S. Rodrigues and Gilson A. Giraldi. Improving the non-extensive medical image segmentation based on tsallis entropy. *Pattern Analysis and Applications*, 14(4):369–379, 2011.
- [16] Paulo Sergio Rodrigues, Guilherme Wachs-Lopes, Ricardo Morello Santos, Eduardo Coltri, and Gilson Antonio Giraldi. A q-extension of sigmoid functions and the application for enhancement of ultrasound images. *Entropy*, 21(4), 2019.
- [17] Neeraj Kumar Sharma and Lalit Mohon Aggarwal. Automated medical image segmentation techniques. In *Journal of medical physics*, 2010.
- [18] Hoo-Chang Shin, Holger Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel J. Mollura, and Ronald M. Summers. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35:1285–1298, 2016.
- [19] Matthew D. Zeiler, Marc’Aurelio Ranzato, Rajat Monga, Mark Z. Mao, Kevin Yang, Quoc V. Le, Patrick Nguyen, Andrew W. Senior, Vincent Vanhoucke, Jeffrey Dean, and Geoffrey E. Hinton. On rectified linear units for speech processing. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3517–3521, 2013.