



Instituto Atlântico  
Atlântico Academy Future

## **Relatório**

Aplicação de redes neurais convolucionais

### **Discentes:**

Carlos Estevão Bastos Sousa

Jeferson do Nascimento Soares

Humberto Bezerra de Meneses Júnior

### **Docente:**

Prof. Alyson Bezerra

Fortaleza, 22 de dezembro de 2021

## 1 Introdução

A partir do presente relatório os autores deste visam apresentar o formato de respostas dadas a atividade final do curso de Computação Cognitiva proporcionado pelo Instituto Atlântico. Salienta-se que tais respostas são aplicadas ao conjunto de imagens data especular (data\_especular\_crop), foram elaboradas na linguagem Python e utilizadas com mais ênfase as bibliotecas Keras e OpenCv, sendo, a partir da primeira, aplicadas redes neurais convolucionais utilizando o padrão de construção da própria biblioteca e, por fim, a rede pré-treinada VGG16.

## 2 Questões

A seguir são apresentadas as questões e o formato de resposta dado e hospedado no GitHub.

### 1. Implementar 3 estratégias de pré-processamento

As três técnicas utilizadas foram:

1.1. Equalização de histograma: uma excelente forma de análise dos níveis de cores ocorre a partir da visualização do histograma, através dele é possível verificar se uma imagem possui muitos tons brancos, pretos ou se a imagem está equalizada.

A equalização de histograma é uma técnica na qual se procura redistribuir os valores de tons de cinza dos pixels em uma imagem, de modo a obter um histograma uniforme, no qual o número ou percentual de pixels de qualquer nível de cinza é praticamente o mesmo.

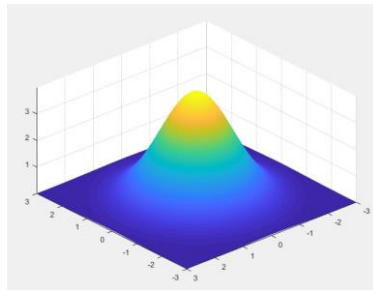
1.2. Gaussian Blur: este filtro faz parte dos filtros passa-baixa que são aplicados à suavização de imagens diminuindo as frequências altas, os mesmos tendem a minimizar ruídos, porém apresentam efeito de borramento.

O filtro gaussiano é dado a partir da aplicação da máscara gaussiana representada pela seguinte equação:

$$w(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Na Figura 1 é apresentada uma representação em 3D do filtro gaussiano.

Figura 1: Representação do filtro gaussiano.



Fonte: autores

Como apresentado, ao fazer uso desse filtro é possível suavizar imagens o que possibilita a redução de ruídos.

### 1.3. Detecção de bordas a partir dos filtros de Sobel

Diferente dos filtros passa-baixa, os filtros passa-alta realçam os detalhes da imagem, isto é, nestes modelos as transições se tornam mais perceptíveis. Normalmente este tipo de filtro é aplicado para realçar características em imagens.

Como terceira estratégia de pré-processamento foi utilizado o filtro de Sobel, que pode ser representado a partir dos seguintes Kernels.

Horizontal:

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Vertical:

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Dado o exposto, ao aplicar os Kernels apresentados em uma imagem destacam-se as bordas, deste modo, considera-se os filtros de Sobel como uma técnica de detecção de bordas.

## 2. Criação de modelos de redes neurais sob os seguintes critérios:

2.1 - Modelo treinado com CNN desenvolvida pelo squad e VGG (pelo menos), testado sem pré-processamento.

2.2 - Modelo treinado com CNN desenvolvida pelo squad e VGG (pelo menos) e testado com estratégia 1 de pré-processamento.

2.3 - Modelo treinado com CNN desenvolvida pelo squad e VGG (pelo menos) e testado com estratégia 2 de pré-processamento.

2.4 - Modelo treinado com CNN desenvolvida pelo squad e VGG (pelo menos) e testado com estratégia 3 de pré-processamento.

As respostas para todos os itens relacionados a questão 2 é dada como única, a partir da seguinte explanação:

Considera-se que uma boa metodologia a ser aplicada para a resolução das atividades seria a implementação das redes para a primeira opção e, nas questões seguintes, aplicar as técnicas de pré-processamento carregar as imagens pré-processadas diretamente na memória RAM. No entanto, devido a baixa quantidade que os autores possuem nessa, todas as imagens foram pré-processadas e salvas em disco rígido.

As metodologias de pré-processamento para as questões foram implementadas conforme a Tabela 1.

Tabela 1: Aplicação de pré-processamento no conjunto de imagens.

| Questão | Pré-processamento   |
|---------|---|
| 2.1     | Nenhuma técnica   |
| 2.2     | Equalização de histograma   |
| 2.3     | Equalização de histograma e filtro gaussiano  |
| 2.4     | Equalização de histograma, filtro gaussiano e filtro de Sobel (Horizontal + Vertical) |

Para a elaboração da rede neural convolucional como forma de medir a melhor técnica de pré-processamento, ou a ausência dela, foi utilizada as seguintes configurações:

- Utilização de *data augmentation*, a partir da aplicação de zoom, girar imagens, mudar escala, entre outros, para balanceamento do conjunto de dados na etapa de treinamento;
- Três camadas para aplicação de convolução com filtros de tamanho, respectivamente de 128, 64 e 32 pixels e função de ativação relu em todos os casos;
- Uma camada densa de 1024 pixels e função de ativação relu;
- Como camada de saída é utilizada a função de ativação softmax;
- Como cálculo da perda (ou loss) foi utilizada a função categorial crossentropy;
- 100 épocas;

- Utilização de call-back para salvar os melhores pesos com base no menor valor obtido pelo cálculo do loss;

- Valor de passos de validação dado através da quantidade de amostras por treino dividido por um tamanho de batch igual a 32.

Para a utilização do VGG são aplicadas configurações semelhantes, com exceção à estrutura do modelo que é padrão do modelo pré-treinado.

3. Criação de um relatório comparando o desempenho dos testes realizados no item 3.

O relatório é dado a partir deste documento.

4. A métricas utilizadas para medir o desempenho no conjunto de teste serão a sensibilidade e precisão:

Sensibilidade :

$$S = VP / (VP + FN)$$

Precisão:

$$P = VP / (VP + FP)$$

Devido ao tempo para elaboração dos modelos, a demora no processo de treinamento e a dificuldade em relação aos recursos computacionais, os autores deste relatório não buscaram por melhores resultados e sim pela implementação dos modelos para posterior alteração de parâmetros nestes. Deste modo, funções para cálculo de sensibilidade (recall) e precisão são utilizados a partir da biblioteca sklearn.

Como observações finais, o uso das funções de pré-processamento consiste em chamar a função e passar como parâmetros dois endereços. O primeiro trata-se da URL para localização do conjunto de imagens existente e, o segundo, da URL, ainda não existente, que será criada pela função e inseridas todas as imagens pré-processadas.

A partir dos conjuntos elaborados, é possível fazer chamadas às funções treinarCNN ou treinarVGG passando como parâmetros apenas os endereços de treinamento, validação e teste do conjunto de imagens.