

# RECONHECIMENTO DE FACES POR APRENDIZADO PROFUNDO: DROPOUT EM AUTOENCODERS

Lucas Fontes Buzuti<sup>1</sup>, Carlos Eduardo Thomaz  
Departamento de Engenharia Elétrica, Centro Universitário FEI  
lucas.buzuti@outlook.com, cet@fei.edu.br

**Resumo:** Modelos conexionistas baseados em Aprendizado Profundo têm apresentado resultados para problemas tradicionais de reconhecimento de padrões que superam o estado-da-arte e se equivalem à habilidade humana. Neste artigo, abordaremos a utilização da técnica Dropout em Autoencoders, que permite minimizar os problemas de overfitting e definição da topologia ótima em redes neurais.

## 1. Introdução

Estudos de extração não-linear de características em imagens de faces em conjunto com representações globais dessas características extraídas têm sido realizados com sucesso por Redes Neurais Convolucionais Autoencoders ou, simplesmente, ConvAEs [4][5].

A topologia de uma ConvAE é composta por múltiplas camadas de codificação e decodificação automática de padrões que permitem, por meio de treinamento por retro-propagação dos erros entre o padrão n-dimensional de entrada e o padrão n-dimensional de saída reconstruído [1], estudar a relação entre atributos visuais de baixo nível, tais como forma e textura, e atributos visuais de alto nível, representados por conceitos semânticos inerentes a imagens de faces, tais como gênero, expressão facial e idade.

Este projeto tem como efetivo revelar novos procedimentos de mineração de dados capazes de evidenciar extrações de características mais flexíveis para compressão de imagens e particularmente úteis para um melhor entendimento e interpretação dessas imagens.

## 2. Metodologia

A topologia da rede proposta baseia-se no Autoencoder tradicional para aprender a codificação de  $N$  padrões de pixels em suas camadas ocultas (hidden layers) até a camada dita "bottleneck" para então aprender a decodificar essas representações para os padrões de saída. A rede é composta em duas partes: encoder function e decoder function, sendo assim, respectivamente,

$$h = f(Wx + b) \quad (1)$$

a ativação da hidden layer,  $x$  o dado de entrada a ser reconstruído,  $W$  a matriz de pesos,  $b$  matriz bias e  $f$  a função de ativação, e

$$y = g(W^T h + c) \quad (2)$$

a entrada reconstruída,  $W^T$  matriz de pesos transposta,  $c$  matriz bias e  $g$  função de ativação. Os parâmetros do modelo são otimizados para minimizar o erro médio de reconstrução, tal que:

$$\mathcal{Q}(x, y) = \theta^*, \theta'^* = \min_{\theta^*, \theta'^*} \frac{1}{n} \sum_{i=1}^n L(x_i, y_i) \quad (3)$$

onde  $L$  é loss function dado pelo erro ao quadrado  $L(x, y) = \|x - y\|^2$ ,  $\theta = \{W, b\}$ ,  $\theta' = \{W^T, c\}$  e  $\theta^*$ ,  $\theta'^*$  são os conjuntos de erros a serem atualizados.

Autoencoders podem ser avaliados como sendo um caso especial de redes feedforward, e podem ser treinadas pelas técnicas, típicas de gradiente descendente e back-propagation [2].

Dropout é uma técnica que aborda dois tipos de soluções: evita o overfitting e fornece uma maneira de combinar de forma exponencial outras topologias de redes neurais diferentes de forma eficaz. O termo "dropout" se refere ao abandono dos neurônios escondidos e visíveis em modelos de redes neurais. Ao referenciar o abandono desses neurônios, o intuito é removê-los temporariamente da rede, juntamente com todas as suas conexões de entradas e saídas. A escolha de quais neurônios serão removidos (desativadas) é aleatória [3].

A topologia do Autoencoder proposto está na figura 1. Adotamos de forma analítica as alocações dos dropouts, essas alocações estarão no encoder, decoder e no bottleneck.

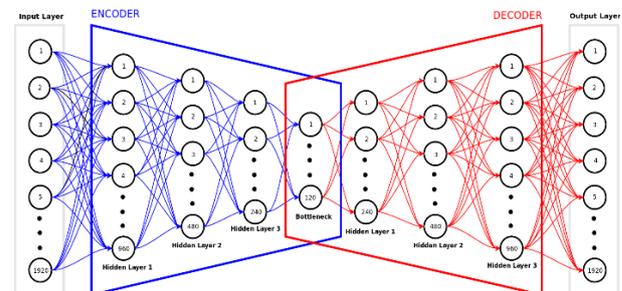


Figura 1 – Autoencoder proposto.

Apesar de ser uma hipótese, podemos analisar as alocações dos dropouts sendo um artifício de modo a ter uma generalização na reconstrução do conjunto de dados. A generalização proposta só acontecerá ao evitar que os neurônios coadaptem no encoder, decoder e principalmente no bottleneck, onde todos os padrões do conjunto de dados estão compactados.

## 3. Resultados

A utilização do dropout em autoencoders mostrou-se eficaz. Os resultados ilustram que os dropouts de 0.1, 0.5 e 1.0 obtêm a reconstrução do conjunto de dados com acurácia acima de 95%. Já o dropout de 0.01 acaba não conseguindo fazer a reconstrução, tendo assim uma acurácia de aproximadamente 88%, como mostra na figura 2.

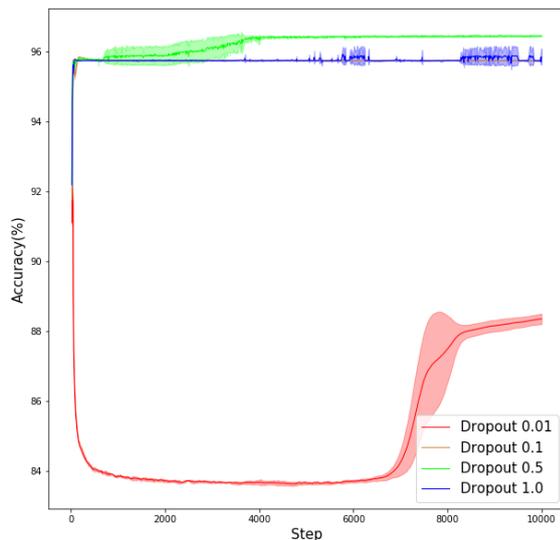


Figura 2 – Gráfico de teste da acurácia média.

Analisando os três melhores dropouts, podemos afirmar que o dropout, cuja probabilidade 0.5, é o melhor mesmo que no começo do treinamento a sua curva de teste tenta apresentar desvio padrão expressivo em relação à média. Todavia ao longo de seu treinamento esse desvio ficou nulo e a curva média se manteve constante acima de 96%.

Analisando os dropouts de 0.1 e 1.0 é praticamente o mesmo resultado, já que suas médias são análogas. A probabilidade de 0.1 possui um desvio padrão nulo em relação a sua média, mas a probabilidade de 1.0 possui um desvio diferente de zero, mas esse desvio é mais expressivo no final do treinamento. As duas probabilidades se mantiveram abaixo de 96% de acurácia.

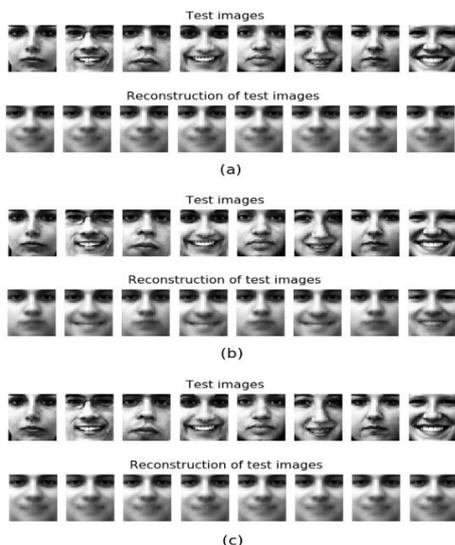


Figura 3 – Reconstruções com dropouts de 0.1 (a), 0.5 (b) e 1.0 (c).

Podemos concluir que o desvio padrão é importante, pois avistamos o transcorrer das variações dos pesos para a reconstrução do conjunto de dados.

A questão não está na capacidade se os dropouts 0.1, 0.5 e 1.0 obtêm a reconstrução dos dados, mas sim a capacidade de generalizar o aprendizado. Infelizmente, não

é possível visualizar graficamente essa generalização, pois a acurácia dos três dropouts são aceitáveis. Na figura 3 podemos visualizar as reconstruções dos dropouts 0.1, 0.5 e 1.0, respectivamente.

No entendimento das reconstruções dos dados mostrado na figura 3, podemos visualizar que os dropouts 0.1 e 1.0 consegue reconstruir o conjunto de dados, mas a questão é a generalização que não ocorre nos dropouts citados. As generalizações das reconstruções ocorrem no dropout 0.5, no qual a acurácia está acima de 96%. Ao observar o conjunto de dados observamos os seguintes aspectos, faces de homens e mulheres sorrindo e os mesmos não sorrindo (face normal).

Se analisarmos a topologia proposta, concluímos que para o conjunto de dados em questão ela só irá generalizar a partir de 96% de acurácia, pois além dos padrões redundantes, tais como olhos, nariz e boca, a topologia é posta a definir dois tipos de classes, tais como sorrindo e não sorrindo.

#### 4. Conclusões

Dropout é uma técnica de deep learning recente e promove a sua eficiência em redes profundas, pois os principais problemas: overfitting e topologia conseguem ser corrigidos por essa técnica. Não há nenhuma literatura que diz quantas camadas ocultas ou neurônios utilizar, por isso a importância do dropout em deep learning.

Em relação às reconstruções, a topologia proposta consegue generalizar o conjunto de dados ao utilizar dropout em 0.5. Portanto, observamos que o autoencoder faz um agrupamento das imagens de treinamento, identificando padrões nos olhos, nariz e boca. Em outras palavras, o autoencoder é capaz de diferenciar faces sorrindo e não sorrindo.

#### 5. Referências

- [1] Rumelhart, D. E., Hinton, G. E., Williams R. J. et. al, Learning representations by back-propagating errors. Nature, v. 323, p. 533, 1986.
- [2] Ian Goodfellow, Yoshua Bengio, Aaron Courville et. al, Deep Learning. [S.I.]: MIT Press, 2016.
- [3] Nitish Srivastava, Geoffrey E. Hinton et. al, Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, JMLR. org, v. 15, n. 1, p. 1929–1958, 2014.
- [4] Hinton, G. E., Zemel, R. S., Autoencoders, minimum description length and helmholtz free energy, Advances in Neural Information Processing Systems, v. 6, p. 3-10, 2001.
- [5] Hasan, M. et al., Learning temporal regularity in video sequences, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, p. 733-742, 2016.

#### Agradecimentos

À instituição Centro Universitário FEI e ao Prof. Dr. Carlos Eduardo Thomaz pela confiança de realizar uma pesquisa em uns dos temas mais relevantes atualmente.

<sup>1</sup>Aluno de IC do Centro Universitário FEI (PBIC P-002/05 e NP-22.004/05-03). Projeto com vigência de 04/18 a 12/18.