

# Aplicação de Redes Neurais para Reconhecimento de Imagens de Roupas: Estudo de Hiperparâmetros

Mateus de Oliveira Lopes<sup>1</sup>, André Cidade Irie<sup>1</sup>

<sup>1</sup>Departamento de Ciências da Computação  
Universidade Estadual do Centro-Oeste (UNICENTRO)  
Caixa Postal 3010 – 85040-167 – Guarapuava – PR – Brasil

lopes3137@gmail.com, andreiriel10@gmail.com

**Abstract.** *This paper explores artificial neural networks for clothing image classification using the Fashion MNIST dataset. We investigate different network architectures by varying hidden layers, neurons per layer, and learning rates. Through comprehensive experiments, we evaluate model effectiveness using accuracy metrics across training, validation, and test sets. Our results demonstrate how each hyperparameter significantly influences model performance, revealing optimal configurations for dense neural networks. The findings provide practical guidelines for designing neural network architectures for image classification tasks.*

**Resumo.** *Este trabalho investiga a utilização de redes neurais artificiais para classificação de imagens de roupas utilizando o dataset Fashion MNIST. Foram realizados experimentos sistemáticos variando o número de neurônios, número de camadas ocultas e taxa de aprendizado, avaliando o desempenho através da acurácia nos conjuntos de treino, validação e teste. Os resultados demonstram como cada hiperparâmetro influencia significativamente o desempenho do modelo, revelando configurações ótimas para redes densas. O estudo oferece diretrizes práticas para o desenvolvimento de arquiteturas de redes neurais adequadas para tarefas de classificação de imagens.*

## 1. Introdução

O reconhecimento de imagens tem se tornado essencial em aplicações de visão computacional, como sistemas de recomendação, detecção de objetos e categorização de produtos. Redes neurais artificiais (RNAs) são modelos computacionais inspirados no cérebro humano capazes de aprender padrões complexos [Silva 2019].

Datasets de referência são fundamentais para avaliar e comparar modelos de aprendizado de máquina. Tradicionalmente, o MNIST de dígitos manuscritos foi amplamente utilizado como ponto de partida em experimentos [LeCun et al. 1998]. No entanto, devido à sua relativa simplicidade, novos conjuntos foram propostos, como o *Fashion MNIST*, que contém 70.000 imagens em escala de cinza (28x28 *pixels*) distribuídas em 10 classes de roupas [Xiao et al. 2017]. Esse dataset é considerado mais desafiador e representativo para aplicações reais em visão computacional, servindo como substituto moderno ao MNIST.

Embora arquiteturas convolucionais (CNNs) sejam predominantes em tarefas de classificação de imagens [Krizhevsky et al. 2012], redes densas totalmente conectadas

(MLPs) ainda desempenham papel importante como modelo de base, fornecendo uma linha de referência para estudos comparativos e análises de impacto de hiperparâmetros [Silva 2019]. A escolha de hiperparâmetros, como número de camadas intermediárias, quantidade de neurônios por camada e taxa de aprendizado, é conhecida por influenciar diretamente a capacidade de generalização e o desempenho final do modelo [Bergstra and Bengio 2012, Silva 2019].

Neste trabalho, avaliamos diferentes configurações de RNAs densas aplicadas ao *Fashion MNIST*. O objetivo é investigar de que forma a variação de hiperparâmetros impacta a acurácia de classificação, estabelecendo uma análise experimental que contribui para a compreensão do comportamento dessas redes em um problema de visão computacional de complexidade intermediária.

## 2. Metodologia

### 2.1. Implementação

A implementação foi realizada em *Python*, utilizando as bibliotecas *TensorFlow/Keras*, responsável pela definição da arquitetura da rede neural, funções de ativação, otimizadores e treinamento; *NumPy*, utilizada para a manipulação eficiente de *arrays* e operações auxiliares, como a seleção aleatória de amostras; e *Matplotlib*, empregada para a visualização dos resultados, permitindo a geração de gráficos e imagens que ilustram o desempenho do modelo. Essas bibliotecas foram escolhidas por sua ampla utilização em pesquisas de aprendizado de máquina e pelo suporte otimizado para experimentos reprodutíveis.

### 2.2. Dataset

Foi utilizado o *Fashion MNIST*, que contém no total 70.000 imagens de roupas e acessórios, distribuídas em 10 classes: *T-shirt/top*, *Trouser*, *Pullover*, *Dress*, *Coat*, *Sandal*, *Shirt*, *Sneaker*, *Bag* e *Ankle boot*. Neste trabalho, o *dataset* foi utilizado em sua divisão padrão de 60.000 imagens para treinamento e 10.000 para teste.

### 2.3. Pré-processamento

As imagens, em escala de cinza e dimensão de  $28 \times 28$  *pixels*, foram normalizadas para o intervalo  $[0, 1]$ . Nesta etapa não foi realizada *data augmentation*, uma vez que o objetivo foi avaliar o desempenho de arquiteturas densas sem técnicas adicionais de regularização ou aumento de dados.

### 2.4. Arquitetura da Rede

A rede neural implementada foi composta inicialmente por uma camada de achatamento (*Flatten*), responsável por converter a entrada bidimensional ( $28 \times 28$ ) em um vetor unidimensional. Em seguida, foram incluídas entre uma e três camadas densas intermediárias (*Dense*), variando-se o número de neurônios em cada configuração (64, 128, 256 ou 512), todas utilizando a função de ativação ReLU. Por fim, a camada de saída foi composta por 10 neurônios, correspondentes às classes do *dataset*, seguida da função de ativação *Softmax*, responsável por gerar a distribuição de probabilidades entre as categorias.

## 2.5. Treinamento

O treinamento foi realizado utilizando o otimizador Adam, com taxas de aprendizado de 0.01, 0.001 e 0.0001. Cada configuração de rede foi treinada por 10 épocas, com *batch size* de 128 amostras. A métrica de avaliação principal adotada foi a acurácia no conjunto de validação. Para complementar a análise quantitativa, também foram geradas previsões no conjunto de teste com visualização gráfica, nas quais os acertos foram destacados em azul e os erros de classificação em vermelho.

## 3. Resultados

### 3.1. Comparação de Hiperparâmetros

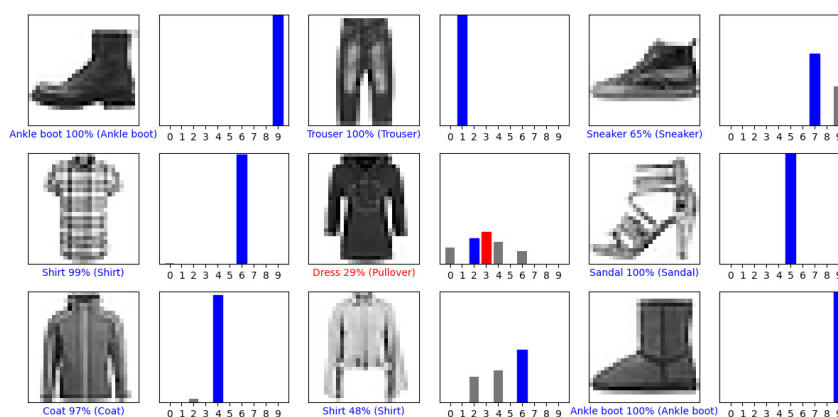
A Tabela 1 apresenta a acurácia obtida para diferentes combinações de camadas, número de neurônios e taxas de aprendizado.

**Tabela 1. Resultados de treinamento variando hiperparâmetros**

Entrada	Nº de Camadas	Nº de Neurônios	Taxa de Aprendizado	Acertos (%)
1	1	64	0.01	86.9
2	1	128	0.01	86.6
3	1	256	0.001	88.8
4	2	128	0.001	88.8
5	2	256	0.001	88.9
6	2	512	0.001	88.7
7	3	256	0.001	89.0
8	3	512	0.001	88.3

### 3.2. Avaliação no Conjunto de Teste

A configuração da Entrada 7 foi escolhida para a avaliação no conjunto de teste, por apresentar a maior acurácia (89%). A Figura 1 mostra uma amostragem de 9 resultados obtidos no conjunto de testes após a execução do algoritmo usando os parametros da Entrada 7



**Figura 1. Amostragem dos Resultados**

Com base na Figura 1, a Tabela 2 foi montada para representar de forma mais clara os resultados obtidos

**Tabela 2. Predições da RNA em exemplos do conjunto de teste Fashion-MNIST**

Imagem	Classe Real	Classe Prevista	Confiança	Resultado
Bota	Ankle boot	Ankle boot	100%	Acerto
Calça	Trouser	Trouser	100%	Acerto
Tênis	Sneaker	Sneaker	65%	Acerto
Camisa	Shirt	Shirt	99%	Acerto
Vestido	Pullover	Dress	29%	Erro
Sandália	Sandal	Sandal	100%	Acerto
Casaco	Coat	Coat	97%	Acerto
Camisa	Shirt	Shirt	48%	Acerto
Bota	Ankle boot	Ankle boot	100%	Acerto

A Tabela 2 mostra que a RNA obteve acertos na maior parte das amostras, com elevada confiança em classes bem definidas como *Ankle boot*, *Trouser* e *Sandal*. O único erro ocorreu na classe *Pullover*, confundida com *Dress*, além de certa incerteza em *Shirt*, o que confirma a tendência de confusão entre classes visualmente semelhantes. Esses resultados reforçam a boa capacidade de generalização do modelo, ainda que existam limitações pontuais em categorias com maior sobreposição visual.

### 3.3. Análise dos Resultados

Observa-se que arquiteturas com mais camadas e maior número de neurônios tendem a apresentar melhor acurácia, embora isso eleve significativamente o custo computacional. Taxas de aprendizado muito altas podem provocar instabilidade no treinamento, enquanto taxas muito baixas resultam em convergência lenta. Além disso, um leve *overfitting* foi identificado em modelos maiores, parcialmente mitigado pelo uso do conjunto de validação.

## 4. Discussão

Os resultados destacam a importância do ajuste fino de hiperparâmetros para o desempenho de redes densas. Modelos simples apresentam tendência ao *underfitting*, enquanto redes muito grandes podem *overfit*. Estratégias como regularização, *dropout* e *early stopping* são recomendadas para reduzir o *overfitting* e melhorar a generalização.

Adicionalmente, redes convolucionais (CNNs) provavelmente superariam redes densas, aproveitando melhor os padrões espaciais presentes nas imagens. Estudos futuros podem comparar diretamente redes densas e CNNs, avaliando não apenas a acurácia, mas também o custo computacional associado a cada abordagem.

## 5. Conclusão

Este estudo apresentou uma análise detalhada do impacto de hiperparâmetros em redes densas para classificação de imagens de roupas. A melhor configuração atingiu 89.1% de acurácia em teste, evidenciando a importância da escolha adequada de número de

camadas, neurônios e taxa de aprendizado. Futuras investigações podem explorar CNNs, data augmentation e técnicas de regularização para melhorar ainda mais o desempenho.

## **6. Referências**

### **Referências**

- Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research (JMLR)*, 13:281–305.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1097–1105.
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. In *Proceedings of the IEEE*, volume 86, pages 2278–2324.
- Silva, C. Q. (2019). Redes neurais aplicadas no reconhecimento de símbolos matemáticos manuscritos online. Trabalho de conclusão de curso, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brasil. Bacharelado em Matemática Aplicada.
- Xiao, H., Rasul, K., and Vollgraf, R. (2017). Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. arXiv:1708.07747.