

# Método de Classificação de Imagens Usando Aprendizado de Máquina com Computação Quântica

Leonardo S. Belluzzo<sup>1</sup>, Regina Melo Silveira<sup>1</sup>

<sup>1</sup>Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica  
Universidade de São Paulo (USP) – São Paulo, SP – Brasil

leonardobelluzzo@usp.br, regina.silveira@usp.br

**Abstract.** *Computational Vision is an active area of research applied in daily tasks as facial recognition and object detection. Quantum Computing improved areas like Cryptography, however, it's advantages are not clear in other areas of research. Although there are evidences of quantum and hybrid models using less parameters and training time to achieve the same performance observed in purely classical models, not many works in literature show quantum models being applied to more complex problems. The hybrid model proposed achieves an accuracy of nearly 94% in the MNIST dataset.*

**Resumo.** *A área de visão computacional impacta diretamente diversas atividades cotidianas como reconhecimento facial e detecção de objetos. A Computação Quântica trouxe avanços em áreas como criptografia, contudo, as vantagens de sua utilização não são claras em outros contextos. Embora haja evidências que modelos quânticos necessitem de um menor número de parâmetros e treinamento para alcançarem desempenho similar aos modelos clássicos, poucos trabalhos na literatura mostram sua aplicação em conjuntos de dados mais complexos. Apresenta-se um modelo híbrido com, aproximadamente, 94% de acurácia no conjunto de dados MNIST.*

## 1. Introdução

*Convolutional Neural Networks* (CNN) estão presentes em diversas aplicações de visão computacional como reconhecimento automático de caracteres, reconhecimento facial e identificação de objetos [Chen et al. 2021]. Tal modelo tem como foco a classificação de imagens dado a complexidade em sua análise pela grande quantidade de atributos e a falta da estrutura tabular [Krizhevsky et al. 2012]. Tendo em vista evidências de melhoria de capacidade e treinabilidade em modelos quânticos e híbridos [Abbas et al. 2021], tem-se como objetivo explorar possíveis vantagens e desvantagens da aplicação de circuitos quânticos na classificação de imagens.

Para isso, propõe-se um método de codificação de imagens em circuitos quânticos integrados com redes neurais densas. Modelos de *Quantum Convolutional Neural Network* (QCNN) apresentam características importantes em relação a circuitos variacionais quânticos pela mitigação de platôs estéreis [Pesah et al. 2021], diferente do observado em outras arquiteturas de *Quantum Neural Networks* (QNN) [McClellan et al. 2018]. Com o intuito de comparar métodos clássicos e híbridos, utiliza-se os conjuntos de imagens MNIST [Deng 2012].

## 2. Trabalhos Correlatos

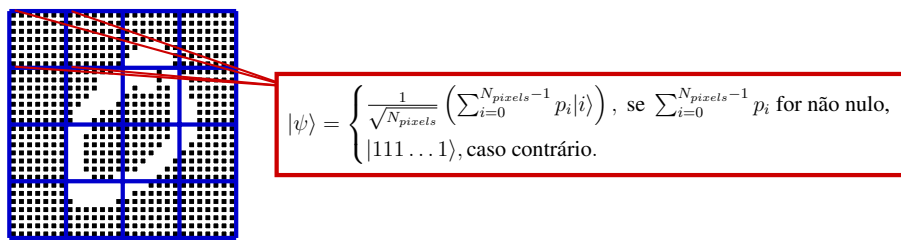
A área de classificação de imagens apresenta grandes avanços com uma miríade de técnicas aplicadas em conjunto no trabalho de [Krizhevsky et al. 2012]. Tais avanços permitem que a classificação de imagens seja aplicada em tarefas cotidianas por sua acurácia e confiabilidade. De modo semelhante, observa-se evidências de que modelos quânticos podem contribuir para modelos com menor número de parâmetros e menor tempo de treinamento [Abbas et al. 2021]. O trabalho de [Oh et al. 2020] é um dos primeiros trabalhos a propor modelos de QCNN semelhantes às CNN. Motivado pelo classificador presente no artigo [Grant et al. 2018], [Oh et al. 2020] utilizam a codificação de segmentos das imagens para convertê-los em estados quânticos os quais, em seguida, são processados por um circuito *Multiscale Entanglement Renormalization Ansatz* (MERA) antes de serem medidos e passados por uma rede neural artificial. Os segmentos das imagens têm dimensões  $2 \times 2$  e o circuito quântico possui 4 qubits. O modelo híbrido atingiu 95% de acurácia no conjunto de exemplos de teste utilizando uma versão de baixa resolução do MNIST cuja resolução foi reduzida a  $10 \times 10$ . Dentre as conclusões apresentadas em [Oh et al. 2020], considera-se o desempenho do modelo híbrido como uma evidência das vantagens do uso de computação quântica em modelos de *Deep Learning*.

Outros trabalhos recentes ainda não publicados apresentam estruturas semelhantes a CNN, o trabalho de [Röseler et al. 2025] utiliza Otimização Bayesiana para fundamentar a comparação de QCNN com sua versão clássica e para determinar os *ansätze* utilizados. Para validar os modelos, utiliza-se também o conjunto de dados MNIST, contudo, em uma versão simplificada na qual é necessário classificar somente os dígitos 0 e 1. De modo similar ao trabalho de [Oh et al. 2020], a imagem é dividida em partes para ser processada por filtros ou *kernels*. A maior acurácia observada é de 98.7% para o problema de classificação binária. O trabalho de [Senokosov et al. 2024] apresenta um modelo híbrido chamado de *Hybrid Quantum Neural Network-Quantum* (HQNN-Quantum). Nesse modelo, a camada convolucional clássica é substituída por um circuito quântico em conjunto com uma rede neural densa. O modelo HQNN-Quantum utiliza codificação por ângulo e obteve uma acurácia de 67% no MNIST. De modo semelhante ao apresentado em [Oh et al. 2020], tem-se a separação da imagem em partes para a aplicação de *kernel* de dimensões  $2 \times 2$ , no caso, um circuito quântico parametrizado.

Considerando os trabalhos supracitados, a aplicação de computação quântica em problemas de classificação de imagens é limitada pela codificação e arquitetura dos modelos híbridos e quânticos. Logo, o objetivo é explorar as vantagens e desvantagens de arquiteturas híbridas nesse contexto. O aprofundamento desse entendimento pode impactar diversas áreas, especificamente, a de Sistemas de Informação.

## 3. Contribuições da Pesquisa

O presente trabalho busca criar um método para a codificação de imagens em circuitos quânticos. Diferente do apresentado em [Oh et al. 2020, Senokosov et al. 2024] e [Röseler et al. 2025], pesquisa-se a aplicação da estratégia de codificar diferentes matrizes em diferentes bases com o intuito de diminuir a quantidade de qubits e circuitos quânticos necessários. O trabalho tem como foco modelos híbridos nos quais o circuito quântico seja responsável por realizar a seleção de atributos pela operação de convolução similar ao realizado por CNN.



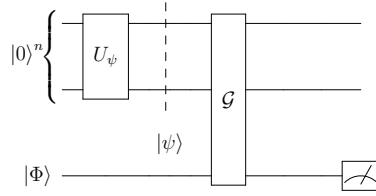
**Figura 1. Segmentação e Codificação por Amplitude. O valor  $p_i$  corresponde ao valor do pixel  $i$ .**

Dado que o número de operações e qubits presentes nos modelos quânticos influenciam a presença de Platôs Estéreis pelo decréscimo da variância do gradiente [McClellan et al. 2018], a principal contribuição do trabalho está em propor um método que concilie o crescimento do número de portas quânticas advindas da codificação por amplitude com a capacidade de representar informação em diversas bases. É válido ressaltar que diferentes conjuntos de operadores utilizados no cálculo de valores esperados dos estados quânticos proporcionam diferentes transformações não-lineares nos dados que podem impactar em seu desempenho.

#### 4. Método Científico

Inicialmente, tem-se os valores da imagem como inteiros pertencentes ao intervalo  $[0, 255]$ . Com o intuito de diminuir os valores representados nos estados quânticos, tem-se a transformação dos valores menores que 127 em 0 e valores maiores ou iguais a 128 são transformados em 1. O processamento da imagem ocorre por *kernels* que funcionam como filtros, ressaltando determinados padrões presentes nas imagens. Utiliza-se segmentos de  $7 \times 7$  pixels no presente modelo devido ao baixo desempenho computacional observado em testes feitos com *kernels* menores. Quanto maior o número de segmentos da imagem, maior o número de circuitos quânticos necessários cuja simulação é computacionalmente custosa. Em seguida, realiza-se a operação de convolução nesses segmentos da imagem, para isso, aplica-se a codificação representada na Figura 1. Cada conjunto de 49 pixels são traduzidos para um estado quântico no qual cada estado da base representa a presença ou ausência de um pixel no segmento da imagem. Considerando um filtro com dimensões  $7 \times 7$ , tem-se um total de 16 partes da imagem, cada parte apresenta um circuito de 7 qubits.

O estado resultante do processo de codificação é transformado em um circuito quântico por decomposição de isomerias arbitrárias [Iten et al. 2016]. Tal processo é indicado pela operação  $U_\psi$  presente na Figura 2. Em seguida, adiciona-se as portas quânticas responsáveis por realizar o processamento do estado inicial. Nesse circuito, os estados quânticos presentes na superposição do estado  $|\psi\rangle$  controlam a aplicação de uma rotação no estado quântico que registra a resposta da camada convolucional, sendo a resposta do circuito quântico dada pelo estado  $|\Phi\rangle$  como apresentado na Figura 2. O circuito é baseado em isometrias, nas quais é possível mapear estados de um espaço de Hilbert de  $2^n$  estados para um espaço de Hilbert de  $2^m$  estados. Portanto, é possível reduzir o número de atributos com essas operações [Iten et al. 2016]. Define-se a operação  $\mathcal{G}_i$  como:



**Figura 2. Representação da operação  $\mathcal{G}$  em um circuito quântico.**

$$\mathcal{G}_i = RY(\theta_i) \otimes |i\rangle\langle i| + \sum_{k \in \{0,1\}^{n-1-i}} \mathbf{I}_2 \otimes |k\rangle\langle k|, \quad (1)$$

onde  $RY(\theta_i)$  é a rotação ao redor do eixo Y da Esfera de Bloch e é uma operação controlada por  $n$  qubits [Iten et al. 2016]. A operação  $\mathcal{G}$  representa a aplicação de  $\mathcal{G}_i$  em todos os estados quânticos da base. Após a camada convolucional, tem-se uma rede neural densa para converter os valores da camada de *pooling* em um vetor de probabilidades. Utiliza-se 8 *kernels* com 16 circuitos de 7 qubits cada, totalizando 128 características. Logo, diminui-se o valor inicial de 784 atributos para 128 com a camada convolucional e de *pooling*. Os valores esperados dos operadores de Pauli X e Z são usados para adquirir informação do estado quântico final dos circuitos.

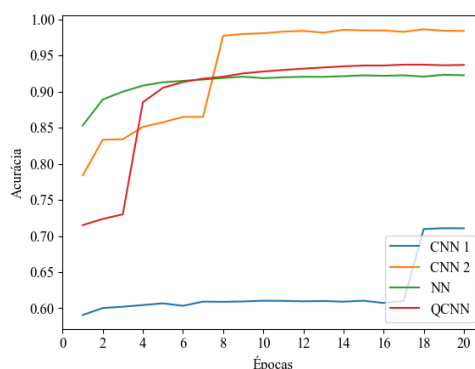
## 5. Resultados Experimentais

Todos os resultados apresentados foram obtidos em um computador com um processador 12th Gen Intel(R) Core(TM) i5-12500H (2.50 GHz), 16,0 GB de RAM e uma placa de vídeo NVIDIA GeForce RTX 3050 Laptop GPU. As principais ferramentas utilizadas são Tensorflow Quantum, Cirq e Qiskit. Os experimentos referem-se ao conjunto de dados MNIST com 10 dígitos (60 mil imagens para treino e 10 mil para teste). Cada modelo foi treinado por 10 épocas com o otimizador Adam e seus parâmetros iniciais são amostrados de distribuições normais. Com a codificação de amplitude e modelo híbrido, foi possível obter uma acurácia de, aproximadamente, 94% no conjunto de dados MNIST.

Para verificar a Capacidade e a Treinabilidade de modelos quânticos e híbridos, realiza-se uma comparação entre redes convolucionais, redes neurais densas e redes convolucionais quânticas. Cria-se dois modelos convolucionais com uma e três camadas convolucionais respectivamente. O primeiro modelo (referido como “CNN 1” na Figura 3) representa uma versão clássica do modelo híbrido QCNN e possui o mesmo número de parâmetros treináveis em sua camada convolucional a menos dos parâmetros de *biases*. O segundo modelo convolucional (“CNN 2”) apresenta uma estrutura mais complexa com três camadas de 8, 16 e 32 *kernels* cada um de tamanho  $2 \times 2$ . O terceiro modelo (“NN”) é uma rede neural de 128 neurônios e, por último, tem-se o modelo híbrido (“QCNN”). A acurácia, o valor da função *loss* e o número de parâmetros treináveis desses modelos podem ser vistos na Tabela 1. O treinamento foi realizado com inicialização de parâmetros aleatórios amostrados de uma distribuição uniforme. No caso, percebe-se que, utilizando-se uma inicialização aleatória uniforme, os autovalores da Matriz de Informação de Fisher dos modelos quânticos apresentam sua distribuição menos concentrada em valores próximos de zero. Tal resultado aponta que o modelo híbrido apresenta maior capacidade e treinabilidade que o modelo clássico equivalente [Abbas et al. 2021].

**Tabela 1. Comparação do desempenho dos modelos clássicos e quânticos.**

Modelo	Acurácia	Loss	Núm. de Parâmetros Camada Convolutacional	Núm. de Parâmetros Modelo Completo
CNN 1	0.7107	1.3420	400	69200
CNN 2	0.9840	0.0756	2648	124498
NN	0.9226	0.2858	-	100480
QCNN	0.9369	0.2612	392	107592



**Figura 3. Comparação entre as redes neurais convolucionais.**

## 6. Considerações Finais

É válido ressaltar que, pela inicialização aleatória uniforme, há grande variação em seus valores de acurácia e *loss*. Uma comparação mais minuciosa pelo treinamento com diferentes conjuntos iniciais de parâmetros não pode ser realizada pelo tempo necessário para treinar o modelo híbrido. A acurácia de validação obtida indica que o modelo apresenta desempenho similar à sua contrapartida clássica, contudo, pela natureza de sua inicialização, não é possível comparar sua capacidade e treinabilidade diretamente com os modelos clássicos. Outra característica interessante é a quantidade de parâmetros treináveis dos modelos da Tabela 1, o modelo CNN 1 apresenta em sua camada convolutacional 400 parâmetros, um valor próximo aos 392 parâmetros presentes no modelo quântico. Contudo, o número de parâmetros treináveis no modelo quântico é, aproximadamente, uma vez e meia o número de parâmetros do modelo clássico. Isso ocorre, pois, para cada circuito quântico utilizado, extrai-se o valor esperado de dois operadores ( $X$  e  $Z$ ), portanto, dobra-se o número de parâmetros necessários na primeira camada de rede neural densa. Tal característica está contrária ao processo de simplificação no número de atributos feito por modelos convolucionais [Krizhevsky et al. 2012]. Como continuidade da pesquisa, busca-se calcular métricas como a Dimensão Efetiva e a Matriz de Informação de Fisher [Abbas et al. 2021] relacionadas com a capacidade e treinabilidade de modelos de aprendizado estatístico. Por fim, analisa-se a aplicação do modelo proposto em imagens com codificação RGB do conjunto de dados *Canadian Institute for Advanced Research* (CIFAR)-10 [Krizhevsky et al. 2009] para avaliar seu desempenho.

## Agradecimentos

O presente trabalho foi realizado com apoio da Fundação de Apoio à Universidade de São Paulo (FUSP) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

## Referências Bibliográficas

- Abbas, A., Sutter, D., Zoufal, C., Lucchi, A., Figalli, A., and Woerner, S. (2021). The power of quantum neural networks. *Nature Computational Science*, 1(6):403–409.
- Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., and Miao, Y. (2021). Review of image classification algorithms based on convolutional neural networks. *Remote Sensing*, 13(22):4712.
- Deng, L. (2012). The mnist database of handwritten digit images for machine learning research [best of the web]. *IEEE Signal Processing Magazine*, 29(6):141–142.
- Grant, E., Benedetti, M., Cao, S., Hallam, A., Lockhart, J., Stojevic, V., Green, A. G., and Severini, S. (2018). Hierarchical quantum classifiers. *npj Quantum Information*, 4(1).
- Iten, R., Colbeck, R., Kukuljan, I., Home, J., and Christandl, M. (2016). Quantum circuits for isometries. *Phys. Rev. A*, 93:032318.
- Krizhevsky, A., Hinton, G., et al. (2009). Learning multiple layers of features from tiny images.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.
- McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R., and Neven, H. (2018). Barren plateaus in quantum neural network training landscapes. *Nature communications*, 9(1):4812.
- Oh, S., Choi, J., and Kim, J. (2020). A tutorial on quantum convolutional neural networks (qcnn). In *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 236–239.
- Pesah, A., Cerezo, M., Wang, S., Volkoff, T., Sornborger, A. T., and Coles, P. J. (2021). Absence of barren plateaus in quantum convolutional neural networks. *Phys. Rev. X*, 11:041011.
- Röseler, P., Schaudt, O., Berg, H., Bauckhage, C., and Koch, M. (2025). Efficient quantum convolutional neural networks for image classification: Overcoming hardware constraints.
- Senokosov, A., Sedykh, A., Saginalieva, A., Kyriacou, B., and Melnikov, A. (2024). Quantum machine learning for image classification. *Machine Learning: Science and Technology*, 5(1):015040.