

# Modelos de Quantum Machine Learning Aplicados na Classificação de Imagens Coloridas: Comparação entre Modelos Híbridos e Clássicos

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.** *Inspired by Convolutional Neural Networks, Quantum Convolutional Neural Networks explore the application of quantum computing in machine learning problems looking for greater classification performance and faster training convergence. Regarding the limited scale of quantum circuits, we propose a hybrid model for colorful image classification using fewer qubits and circuits than other architectures found in literature. Our model presents competitive performance in comparison with classic models and smaller variability in its training. These results are evidences of possible advantages in the use of quantum computing in machine learning.*

**Resumo.** *Inspiradas pelas Convolutional Neural Networks, as Quantum Convolutional Neural Networks exploram a aplicação de computação quântica em problemas de aprendizado de máquina visando alcançar melhor desempenho de classificação e convergência mais rápida. Considerando a escalabilidade limitada de circuitos quânticos, propõe-se um modelo híbrido para a classificação de imagens coloridas utilizando menor número de qubits e de circuitos em relação a outros modelos da literatura. O modelo proposto apresenta métricas de desempenho competitivas com modelos clássicos e menor variabilidade em seu treinamento, produzindo evidências de possíveis benefícios no uso da computação quântica na área de aprendizado de máquina.*

## 1. Introdução

Aplicações de computação quântica no contexto de aprendizado profundo mostram evidências de vantagens na utilização de circuitos quânticos. O aumento na capacidade de aproximação desses modelos e a melhoria em sua convergência no processo de treinamento [Abbas et al. 2021] estão entre as principais vantagens. Tais evidências foram obtidas em contextos de redes neurais densas com um número de parâmetros limitado em vista das aplicações em visão computacional. Destarte, busca-se resultados semelhantes em arquiteturas mais complexas, com um número maior de parâmetros treináveis e com *datasets* de maior número de atributos. Porém, circuitos variacionais quânticos apresentam limitações associadas aos Platôs Estéreis [McClean et al. 2018], codificações e o custo computacional de suas simulações [Kharsa et al. 2023, Priyanka et al. 2023].

O uso de circuitos quânticos para a classificação de imagens é observado, primeiramente, no trabalho de [Grant et al. 2018] com um circuito Multiscale Entanglement Renormalization Ansatz (MERA) e desenvolvido em um modelo Quantum Convolutional Neural

Network (QCNN) nos trabalhos de [Cong et al. 2019, Oh et al. 2020]. Observa-se dois principais conjuntos de modelos, QCNN puramente quânticas e QCNN híbridas. QCNN puramente quânticas utilizam apenas circuitos quânticos desde a convolução até o cálculo da classe da amostra em questão. QCNN híbridas apresentam uma rede neural clássica no fim do modelo enquanto a convolução é feita através de circuitos quânticos. As imagens presentes no *Canadian Institute for Advanced Research* (CIFAR)-10 [Krizhevsky and Hinton 2009] apresentam um grande número de atributos em comparação com *datasets* de dados estruturados, tendo em vista que a simulação desse circuitos é computacionalmente custosa, é necessário alinhar a codificação dos dados clássicos com o tamanho dos filtros utilizados na convolução [Kharsa et al. 2023].

Em vista das limitações e vantagens observadas em modelos quânticos e híbridos tanto na classificação de imagens quanto em outros problemas de aprendizado supervisionado, o presente trabalho utiliza uma versão de classificação binária do CIFAR-10 para comparar o desempenho entre modelos clássicos e híbridos. Os resultados são extraídos de uma QCNN híbrida e de modelos de arquitetura semelhante, contribuindo com dados relacionados ao número de parâmetros, acurácia e valor da função *loss* em um teste de ablação utilizando redes neurais densas e convolucionais.

A Seção 2 apresenta trabalhos similares que aplicam QCNN na classificação de imagens e outros problemas. Em seguida, apresenta-se a QCNN proposta na Seção 3 e os resultados obtidos com tal modelo na Seção 5. Por fim, tem-se as considerações do trabalho em relação ao experimento e perspectiva de próximos passos em experimentos com QCNN na Seção 6.

## 2. Trabalhos Correlatos

Pela Tabela 1, observa-se diferentes modelos os quais buscam equilibrar custo computacional com a acurácia. Seguindo as observações feitas em [Kharsa et al. 2023] em relação as codificações por amplitude, ângulo e codificação densa, o custo computacional da preparação do estado quântico inicial em função do tamanho do vetor de informação é mais eficiente na codificação por amplitude, apresentando complexidade, aproximadamente, linear em relação ao número de dados codificados. Além disso, a codificação por amplitude apresenta maior número de atributos codificados por qubit em comparação às outras. A principal desvantagem da codificação por amplitude é o aumento exponencial do número de portas quânticas necessárias, levando a um crescimento exponencial entre número de qubits e profundidade do circuito. Tais desvantagens são associadas à presença de platôs estéreis, os quais estão relacionados, também, à expressividade dos circuitos quânticos variacionais [Pesah et al. 2021, Kübler et al. 2021].

Define-se platôs estéreis como regiões nas quais o gradiente em relação aos parâmetros treináveis do circuito apresentam valores exponencialmente decrescentes em relação ao número de qubits [Pesah et al. 2021, McClean et al. 2018]. Tal fenômeno é observado em circuitos variacionais quânticos e tem-se mostrado um desafio no contexto de Quantum Neural Network (QNN) e QCNN [Corli et al. 2025]. É possível observar a presença de platôs estéreis no treinamento de modelos quânticos ao analisar a variância da derivada parcial em relação aos parâmetros do *ansatz*. Considere um conjunto de parâmetros treináveis  $\theta$  sendo  $\theta \in \mathbb{R}^{\dim(\theta)}$ , define-se o circuito quântico como um conjunto de operações parametrizadas ( $U(\theta)$ ) realizadas em um estado quântico inicial ( $|\psi_0\rangle$ ) cujo

**Tabela 1. Trabalhos correlatos.**

| Artigo                           | Modelos  | Conjunto de Dados                                      | Número de Qubits | Profundidade Estimada (n = Número de Qubits) | Simplificações ou Pré-processamentos                         | Resolução da Imagem na Entrada do Modelo    | Número de Imagens no Teste   |
|----------------------------------|--|--|------------------|--|--|---|--|
| [Grant et al. 2018]              | <i>Hierarchical Quantum Classifier</i> (apenas quântico)               | MNIST e Iris   | 4 e 8            | Não mencionado                               | -  | 28 × 28                                     | 10.000(MNIST)  |
| [Zheng et al. 2023]              | QCNN (apenas quântico)   | MNIST(Binário)   | 5                | $\mathcal{O}(2^n)$                           | Downscale  | 8 × 8                                       | Não mencionado   |
| [Röseler et al. 2025]            | <i>Hybrid QCNN Regular QCNN</i>  | MNIST(Binário)   | 1, 4 e 16        | $\mathcal{O}(n)$                             | -  | 32 × 32                                     | Não mencionado   |
| [Oh et al. 2020]                 | QCNN Híbrida   | MNIST  | 4                | $\mathcal{O}(\log(n))$                       | Downscale  | 10 × 10                                     | 10.000   |
| [Hur et al. 2022]                | QCNN (apenas quântico)   | MNIST (Binário) e FMNIST (Binário)                     | 8                | Não mencionado                               | Interpolação bilinear, PCA e Autoencoding                    | 28 × 28                                     | 2.115 (MNIST) e 2.000 (FMNIST)                                     |
| [Wang et al. 2022]               | VQDNN  | MNIST(binário) e UCI(binário)                          | 6 e 10           | $\mathcal{O}(\text{poly}(n))$                | PCA  | 32 × 32 e 8 × 8                             | 360 (UCI - "0 e 1"), 357 (UCI - "3 e 8") e 400 (MNIST)             |
| [Senokosov et al. 2024]          | HQNN-Quantum HQNN-Parallel   | MNIST, <i>Medical</i> MNIST e CIFAR-10                 | 4                | $\mathcal{O}(\log(n))$                       | Rotações Aleatórias, Inversão Horizontal e Redimensionamento | 28 × 28, 64 × 64 e 32 × 32 × 3              | 10.000 (MNIST), 11.791 ( <i>Medical</i> MNIST) e 10.000 (CIFAR-10) |
| [Cong et al. 2019]               | QCNN (apenas quântico)   | <i>1D Phase Recognition e Quantum Error Correction</i> | 9                | $\mathcal{O}(n)$                             | -  | -   | -  |
| [Wei et al. 2022]                | QCNN (apenas quântico)   | MNIST  | 14               | $\mathcal{O}(\text{poly}(n))$                | -  | 28 × 28                                     | 2.100("1 e 8") e 10.000("0 a 9")                                   |
| [Gong et al. 2024]               | PCA e Autoencoder QCNN   | MNIST e FMNIST   | 8                | $\mathcal{O}(2^n)$                           | PCA, Autoencoding e Downscale                                | 28 × 28                                     | 10.000 (todos)   |
| [Parthasarathy and Bhowmik 2021] | Quantum Optical Convolutional Neural Network (QOCNN) (apenas quântico) | MNIST  | 392              | Não mencionado                               | Transformação em Números Complexos                           | 28 × 28                                     | 10.000   |
| [Song et al. 2024]               | resource-efficient QCNN (RE-QCNN)                                      | MNIST e FMNIST   | 5                | $\mathcal{O}(\text{poly}(\log n))$           | -  | 28 × 28                                     | 10.000 (todos)   |
| [Jing et al. 2022]               | HQconv FQconv  | MNIST, FMNIST e CIFAR-10-small                         | 4, 9, 12 e 18    | $\mathcal{O}(n)$                             | Downscale  | 20 × 20, 28 × 28, 10 × 10 × 3 e 20 × 20 × 3 | 10.000 (todos)   |

estado final é  $|\psi\rangle$ :

$$|\psi\rangle = U(\theta)|\psi_0\rangle. \quad (1)$$

A derivada parcial em relação ao parâmetro  $i$  e à função *loss* é  $\partial\theta_i L$ . A variância da derivada parcial é dada por [Pesah et al. 2021]:

$$\text{Var}[\partial\theta_i L] = \mathbb{E} [(\partial\theta_i L)^2] - \mathbb{E} [\partial\theta_i L]. \quad (2)$$

Onde  $\mathbb{E} [\dots]$  denota a expectativa matemática. De acordo com a Desigualdade de Chebychev [Magalhães 2006], é possível correlacionar a probabilidade da variância ser maior que um determinado valor  $c$ :

$$P(|\partial\theta_i L| \geq c) \leq \frac{\text{Var}[\partial\theta_i L]}{c^2}. \quad (3)$$

Portanto, a probabilidade do módulo da derivada parcial em relação aos parâmetros do circuito ser maior ou igual a uma constante  $c$  é diretamente proporcional à sua variância [Pesah et al. 2021]. Ademais, platôs estéreis estão relacionados com a complexidade do número de medidas necessárias para recriar um estado quântico com determinada precisão, tal complexidade pode invalidar possíveis ganhos teóricos de desempenho [Kübler et al. 2021].

Outra limitação observada nos modelos QCNN é o número de camadas de convolução presentes nessas arquiteturas. Isso ocorre, pois busca-se a adição de funções de ativação não lineares em tais modelos. Contudo, tais funções são adicionadas após o processo de medição [Maronese et al. 2022], levando a arquiteturas convolucionais com uma camada de profundidade [Röseler et al. 2025, Wang et al. 2022, Senokosov et al. 2024, Wei et al. 2022, Song et al. 2024, Jing et al. 2022] e a medidas intermediárias nos circuitos [Zheng et al. 2023, Oh et al. 2020, Röseler et al. 2025,

Hur et al. 2022, Cong et al. 2019, Gong et al. 2024]. As desvantagens e limitações em relação ao número de qubits, à profundidade do circuito, à profundidade da camada convolucional quântica e à presença de platôs estéreis fazem com que os experimentos presentes na bibliografia sejam feitos em conjuntos de dados simplificados. Nos trabalhos de [Zheng et al. 2023, Oh et al. 2020, Röseler et al. 2025, Hur et al. 2022, Wang et al. 2022, Senokosov et al. 2024, Gong et al. 2024] observa-se simplificações na resolução das imagens (*downscale*), diminuição no número de classes ou o uso de métodos para redução da dimensão dos dados das imagens como o método Principal Component Analysis (PCA).

O presente trabalho simplifica o número de classes presentes no CIFAR-10 para um problema binário, contudo, mantém sua resolução original ( $32 \times 32 \times 3$ ), ainda assim, o trabalho encontra-se em um pequeno conjunto de trabalhos que explora, *datasets* de imagens coloridas mantendo sua resolução original no contexto de modelos QCNN híbridos. Por exemplo, o trabalho de [Jing et al. 2022] explora uma versão reduzida do CIFAR-10 chamada CIFAR-10-small, composto de 200 imagens amostradas aleatoriamente do *dataset* original. As imagens do CIFAR-10-small foram reduzidas para resoluções de  $10 \times 10 \times 3$  ou  $20 \times 20 \times 3$  pixels.

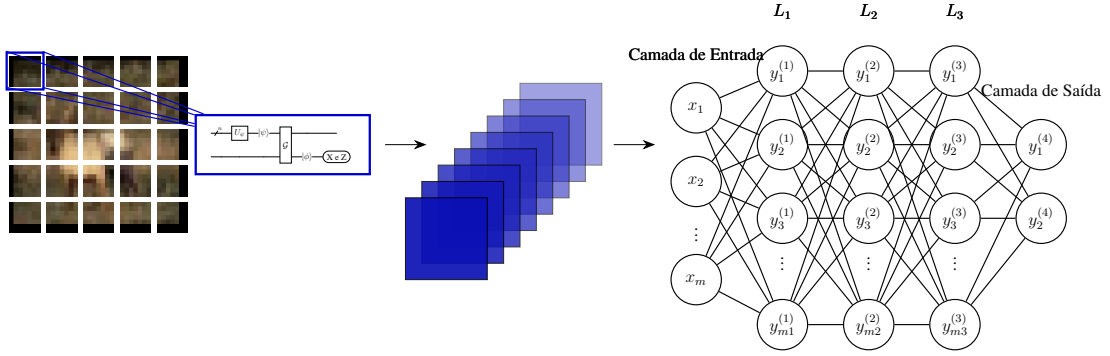
### 3. Modelo Desenvolvido

O modelo QCNN, inspirado pelos trabalhos de [Oh et al. 2020, Park et al. 2019, Maronese et al. 2022, Wei et al. 2022, Gong et al. 2024], possui um *kernel* de  $7 \times 7$  o qual percorre imagens de resolução  $32 \times 32 \times 3$ , acrescenta-se 3 pixels em ambas dimensões da imagem para que a divisão pelas dimensões do *kernel* seja inteira (*padding*). Logo, a imagem passa a ser  $35 \times 35 \times 3$ . Cada um dos 8 filtros apresenta 49 parâmetros treináveis diferentes que percorrem as 25 partes da imagem para cada canal de cor. É válido ressaltar que os parâmetros treináveis dos filtros são compartilhados entre os diferentes canais de cor. Como é mostrado na Figura 1, a imagem é segmentada em partes de mesmo tamanho do *kernel* utilizado. Tal comportamento é obtido quando as dimensões do *kernel* e o valor de *stride* são iguais, causando a não interpolação entre esses filtros ao longo da imagem. Os circuitos quânticos substituem a operação de convolução ao relacionar cada pixel a uma operação RY aplicadas no qubit de resposta. Divide-se o circuito quântico utilizado em duas partes, codificação e circuito variacional. A codificação é realizada pela amplitude de probabilidade de cada estado da base em relação aos pixels de uma seção da imagem. Sendo  $p$  o vetor dos pixels, tem-se:

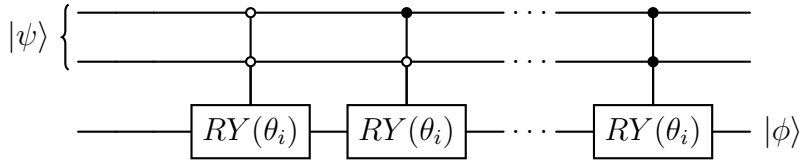
$$p = \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{N_{pixels}-1} \end{pmatrix}. \quad (4)$$

Onde  $N_{pixels}$  representa o número de pixels do *kernel*, no caso, 49. Define-se o estado  $|\psi\rangle$  após a codificação:

$$|\psi\rangle = \begin{cases} \frac{1}{\|p\|_2} \left( \sum_{i=0}^{N_{pixels}-1} p_i |i\rangle \right), & \text{se } p \text{ for não nulo,} \\ |111 \dots 1\rangle, & \text{caso contrário.} \end{cases} \quad (5)$$



**Figura 1. Arquitetura do Modelo QCNN utilizado no CIFAR-10.**



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

Onde  $p_i$  representa o  $i$ -ésimo pixel do vetor  $p$  e  $p_i \in [0, 255]$ . Por completude, define-se também o caso no qual  $p$  seja nulo. A codificação por amplitude permite diminuir o número de qubits utilizados no circuito e aumenta a escalabilidade do modelo. Contudo, o número de qubits aumenta consideravelmente a complexidade espacial necessária para executar as simulações, limitando a largura dos circuitos quânticos.

O circuito variacional utilizado é apresentado também nos trabalhos de [Park et al. 2019, Maronese et al. 2022, Wei et al. 2022, Gong et al. 2024] em contextos diferentes da camada convolucional quântica, sendo essa uma das contribuições do trabalho. O circuito variacional  $\mathcal{G}$  utiliza portas RY controladas por  $n-1$  qubits e é ilustrado na Figura 2. Define-se  $\mathcal{G}$  como:

$$\mathcal{G} = \prod_{i \in \{0,1\}^{n-1}} \mathcal{G}_i, \quad (6)$$

onde  $\mathcal{G}_i$  é:

$$\mathcal{G}_i = RY(\theta_i) \otimes |i\rangle\langle i| + \sum_{k \in \{0,1\}^{n-1-i}} \mathbf{I}_2 \otimes |k\rangle\langle k|. \quad (7)$$

Onde  $\theta_i$  representa o  $i$ -ésimo parâmetro treinável. Por fim, o estado quântico  $|\phi\rangle$  é o estado resposta do circuito com o qual obtém-se o valor esperado dos operadores de Pauli X e Z. O circuito completo pode ser visto na Figura 3. A função de ativação Rectified Linear Unit (ReLU) é aplicada em cada valor esperado obtido da camada quântica, produzindo dois valores por filtro para compor os dados de entrada para a rede neural clássica. Define-se ReLU como:

$$\text{ReLU}(x) = \max(0, x), \quad (8)$$

onde  $x \in \mathbb{R}$ .

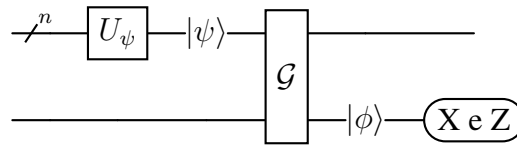


Figura 3. Representação do circuito quântico utilizado no modelo.

Tabela 2. Comparação do número de parâmetros treináveis dos modelos clássicos e quânticos no CIFAR-10 binário.

| Modelo | Núm. de Parâmetros Camada Convolutacional | Núm. de Parâmetros Modelo Completo | Dimensão do Tensor de Saída da Camada Convolutacional |
|--------|---|------------------------------------|---|
| CNN 1  | 1184                                      | 181884                             | 200   |
| CNN 2  | 2712                                      | 116212                             | 32  |
| NN     | -   | 1329500                            | -   |
| QCNN   | 392                                       | 581092                             | 1200  |

#### 4. Especificações dos Experimentos

Todos os resultados apresentados foram obtidos em um computador com 48 processadores AMD Ryzen Threadripper 3960X 24-Core Processor, 125 GB de RAM e uma placa de vídeo NVIDIA GeForce RTX 3090. Realiza-se um teste de ablação com a comparação entre três modelos clássicos e um modelo híbrido. Compara-se diretamente o modelo QCNN híbrido com sua contrapartida clássica dado que os dois modelos apresentam arquitetura similar. Tal arquitetura é formada por uma camada convolutacional com 8 filtros para um *kernel*  $7 \times 7$  e *stride* de mesmo tamanho. Não há a utilização de *pooling* nos modelos QCNN e sua contrapartida clássica denominada “CNN 1” nas tabelas e figuras. Além disso, todas as redes neurais densas utilizadas nesse trabalho apresentam 400, 200 e 100 neurônios, as classes são obtidas dos *logits* pela função Sparse Categorical Crossentropy<sup>1</sup>. Por se tratar de um teste de ablação, elimina-se a camada convolutacional para determinar a contribuição da camada convolutacional quântica no desempenho do modelo híbrido, tal arquitetura é denominada “NN” nas tabelas e figuras. O modelo convolutacional “CNN 2” é composto por três camadas convolucionais interpoladas por funções *pooling* de *kernels* de dimensões  $2 \times 2$  com 8, 16 e 32 filtros respectivamente. Ao final de cada camada, utiliza-se a função de ativação ReLU. Por fim, os parâmetros de todas as camadas clássicas foram inicializadas por meio da técnica de [Glorot and Bengio 2010], a qual apresenta a inicialização de parâmetros treináveis por meio de uma distribuição uniforme cujo intervalo é normalizado pelo tamanho da camada dos pesos a serem inicializados e da camada seguinte. No caso, os circuitos quânticos foram inicializados a partir da amostragem de uma distribuição uniforme de  $-\pi$  a  $\pi$ .

Como descrito na Seção 3, o modelo QCNN utiliza o mesmo conjunto de parâmetros treináveis nas três matrizes de cor, enquanto o modelo clássico apresenta conjuntos de parâmetros treináveis únicos para cada matriz. Tais matrizes são somadas e transformadas em um *kernel*, logo, reduz-se a dimensão dos dados antes de passarem para a próxima

<sup>1</sup>Disponível em: [https://www.tensorflow.org/api\\_docs/python/tf/keras/losses/SparseCategoricalCrossentropy](https://www.tensorflow.org/api_docs/python/tf/keras/losses/SparseCategoricalCrossentropy). Acesso feito em 27 de março de 2026.

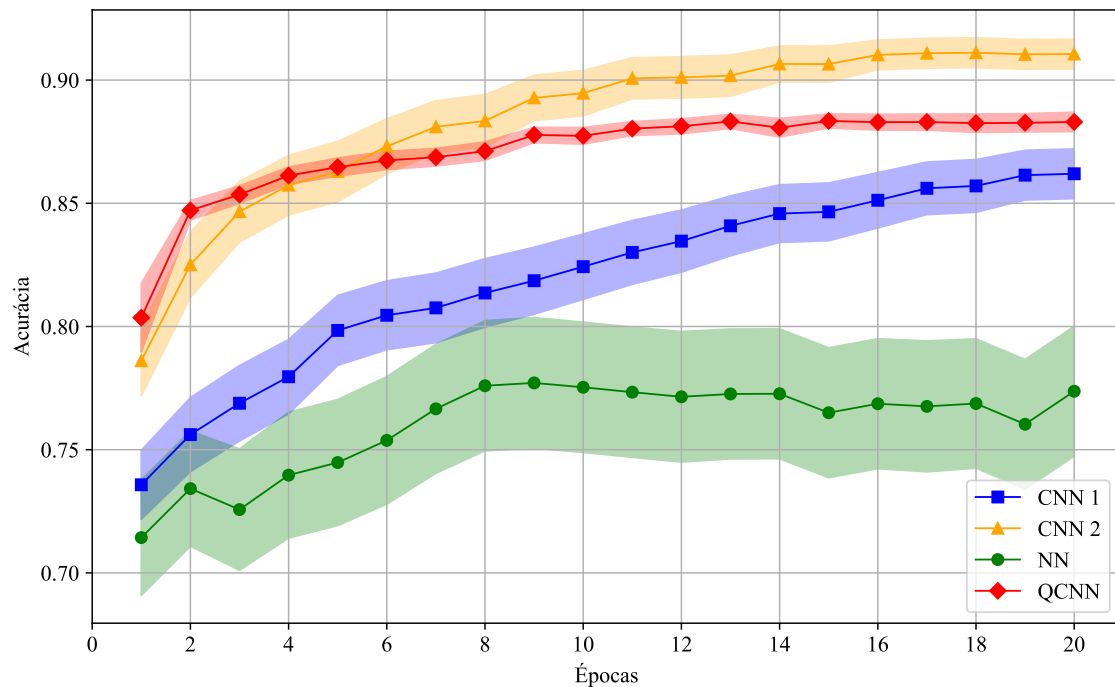


Figura 4. Acurácia média dos experimentos.

camada. A diferença entre o número de parâmetros presentes nas camadas convolucionais é explicitado na Tabela 2. A QCNN apresenta um número maior de parâmetros treináveis no modelo completo, pois cada matriz da codificação RGB gera dois valores esperados advindos do circuito quântico, portanto, a dimensão do tensor de saída do modelo QCNN é seis vezes a dimensão referente a “CNN 1”.

O conjunto de dados utilizado no treinamento é simplificado para conter imagens de cervos e caminhões, limitando o CIFAR-10 a um problema de classificação binário. Tem-se 10.000 imagens no conjunto de treino e 2.000 imagens no conjunto de teste. O treinamento é realizado com o algoritmo Adam [Kingma and Ba 2014] com uma taxa de aprendizado de 0,001, as taxas de decaimento exponencial  $\beta_1$  e  $\beta_2$  são 0,9 e 0,999 respectivamente. Outros parâmetros não foram alterados dos valores predefinidos<sup>2</sup>.

## 5. Resultados Obtidos

As Figuras 4 e 5 apresentam a acurácia e *loss* médias para cada época de treinamento de cada modelo analisado. A região ao redor das curvas representa o intervalo dado pelo desvio padrão da média das métricas de desempenho, acurácia e *loss*. Para obter esses dados, treina-se 100 modelos “CNN 1”, “CNN 2” e “NN” com inicialização aleatória dada por distribuições normais como supracitado. Em relação ao modelo híbrido “QCNN”, treina-se dez modelos. Pela elevada complexidade temporal da simulação dos circuitos quânticos, há um número reduzido de experimentos com o modelo híbrido. Portanto, é necessário analisar os resultados obtidos considerando a diferença de escala no número de experimentos realizados com modelos clássicos e híbridos.

<sup>2</sup>Disponível em: [https://www.tensorflow.org/api\\_docs/python/tf/keras/optimizers/Adam](https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/Adam). Acesso feito em 17 de abril de 2026.

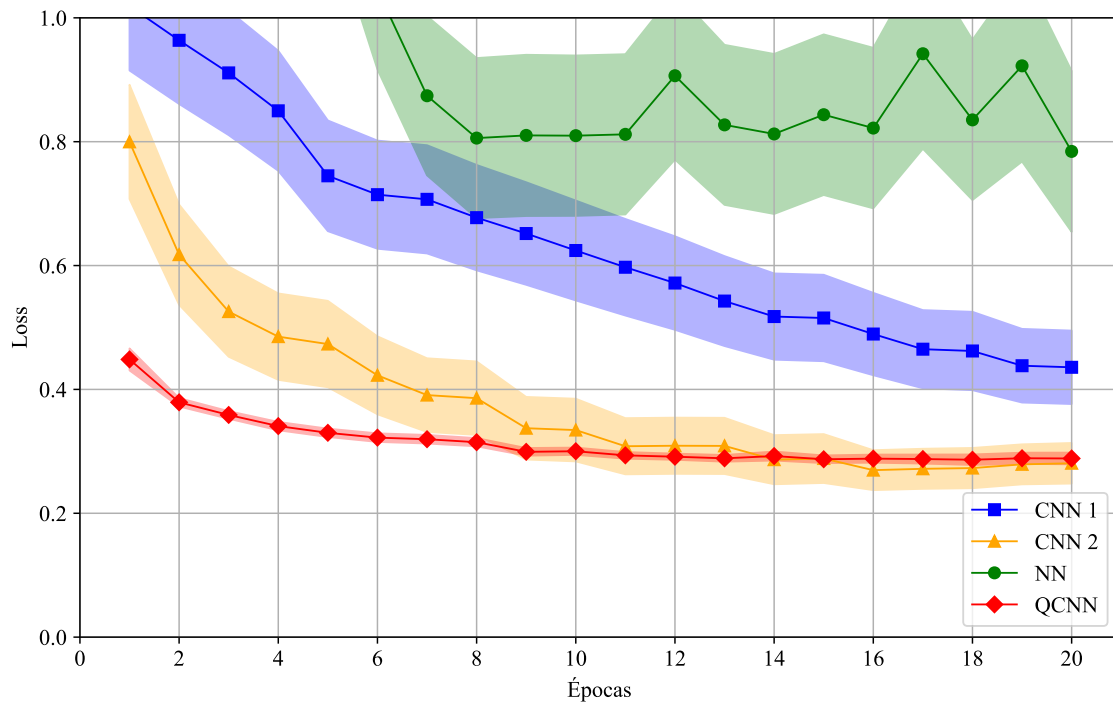


Figura 5. Loss média dos experimentos.

Na Figura 4, o modelo híbrido apresenta maior acurácia média do que o modelo clássico correspondente (“CNN 1”) em todas as épocas de treinamento. Em relação a treinabilidade mencionada em [Abbas et al. 2021], observa-se convergência mais rápida no modelo QCNN que nos modelos exclusivamente clássicos, pois a QCNN apresenta acurácia média superior ao modelo “CNN 2” até a terceira época. É válido ressaltar que o modelo “CNN 2” apresenta maior profundidade e número de parâmetros treináveis na camada convolucional que o modelo híbrido, logo, nas épocas iniciais, a QCNN supera em desempenho tanto o modelo para comparação direta, “CNN 1”, quanto o modelo mais robusto, “CNN 2”. A Figura 5 contém os valores da média da função *loss* em cada época de treinamento. Similar aos resultados observados para a acurácia média, o modelo QCNN supera os modelos clássicos até a sexta época considerando o desvio padrão da média. Em relação aos modelos “NN” e “CNN 1”, o modelo híbrido apresenta *loss* média menor até a vigésima época considerando, novamente, o desvio padrão da média. Na Tabela 3, considera-se a maior acurácia e a menor *loss* observadas ao longo de 20 épocas de treinamento para o cálculo da média e do desvio padrão dos experimentos de cada modelo. Tais valores são representados como “Acurácia Média” e “Loss Média” em adição ao desvio padrão da média  $s_{\bar{x}}$ . Por fim, a maior acurácia e a menor *loss* observadas ao longo das épocas de treinamento e em todos os experimentos são apresentados como “Maior Acurácia” e “Menor Loss”.

Pelos dados apresentados, é possível observar melhoria das métricas de desempenho no modelo híbrido pela adição da camada convolucional quântica ao compará-lo com os modelos “NN” e “CNN 1” em relação a acurácia e *loss* médias. Contudo, ao analisar a maior acurácia obtida em todos os experimentos e em todas as épocas para cada modelo, verifica-se que o modelo híbrido obteve o menor valor na acurácia e segundo maior valor

**Tabela 3. Comparação entre maior acurácia média e menor *loss* média entre as épocas de treinamento.**

| Modelo | Acurácia<br>Média $\pm s_{\bar{x}}$ | Loss<br>Média $\pm s_{\bar{x}}$ | Maior<br>Acurácia | Menor<br>Loss |
|--------|-------------------------------------|---------------------------------|-------------------|---------------|
| CNN 1  | 0,86198 $\pm$ 0,01008               | 0,43570 $\pm$ 0,05927           | 0,90877           | 0,23834       |
| CNN 2  | 0,91109 $\pm$ 0,00605               | 0,26962 $\pm$ 0,03222           | 0,94052           | 0,17309       |
| NN     | 0,77710 $\pm$ 0,02649               | 0,78431 $\pm$ 0,12989           | 0,89869           | 0,25830       |
| QCNN   | 0,88342 $\pm$ 0,00283               | 0,28646 $\pm$ 0,00847           | 0,89617           | 0,25313       |

na *loss*, pois foram feitos poucos experimentos com o modelo híbrido. Tal fenômeno pode ser explicado pela diferença no número de experimentos realizados.

Por fim, é válido relembrar as evidências apontadas por [Abbas et al. 2021], nas quais os modelos quânticos ou híbridos apresentariam capacidade, a habilidade de aproximação de diferentes classes de funções, e treinabilidade, referente a velocidade de convergência e presença de platôs estéreis, maiores que modelos puramente clássicos. Para o modelo QCNN apresentado nesse trabalho, as evidências apontam para alguma melhoria em relação à treinabilidade, devido a menor variabilidade dos valores de acurácia e *loss*. Em relação à capacidade dos modelos, não foi possível verificar mudança significativa na habilidade de aproximar diferentes classes de funções, pois tanto modelos híbridos quanto quânticos são treinados e avaliados no mesmo conjunto de dados, destarte, não é possível afirmar que modelos quânticos sejam capazes de aproximar outras classes de funções que não as já aproximadas por modelos clássicos. Ademais, não é possível calcular métricas como a Matriz de Informação de Fisher para avaliar a capacidade desses modelos pelo grande número de parâmetros treináveis [Abbas et al. 2021]. Porém, umas das limitações do presente trabalho é a execução do treinamento do modelo híbrido apenas dez vezes, o que aumenta a incerteza sobre os resultados. Outro ponto de atenção é a limitação do número de classes do CIFAR-10 avaliadas, sendo necessário experimentos com o conjunto de dados completo ou com conjuntos maiores e com maior número de classes como o ImageNet [Deng et al. 2009].

## 6. Conclusão

O presente trabalho apresenta uma combinação de modelo de QCNN e codificação original inspirado nos artigos [Park et al. 2019, Maronese et al. 2022, Wei et al. 2022, Gong et al. 2024], análise sobre a arquitetura e desempenho na classificação de imagens coloridas e produção de evidências em relação a treinabilidade de modelos quânticos e híbridos [Abbas et al. 2021]. É possível observar que o modelo híbrido proposto atinge desempenho superior ou comparável aos modelos “NN” e “CNN 1” mais rapidamente no treinamento. Tais resultados corroboram, parcialmente, o trabalho de [Abbas et al. 2021], especialmente, a melhor treinabilidade de modelos quânticos e híbridos.

As principais limitações observadas no trabalho são o limitado número de modelos híbridos treinados, a arquitetura convolucional pouco profunda e a ausência de métricas mais robustas como a Matriz de Informação de Fisher [Abbas et al. 2021]. Em relação a arquitetura do modelo QCNN descrito, o colapso do estado quântico  $|\phi\rangle$  na medição impede a replicação da camada convolucional quântica para aumentar a profundidade e o número de parâmetros treináveis do modelo. Além disso, limitações de hardware

e tempo hábil impedem a exploração de diferentes configurações de hiper-parâmetros no treinamento do modelo híbrido e limitam, também, o número de inicializações e épocas de treinamento. Em termos práticos, o tempo necessário para treinar um modelo híbrido é maior que o treinamento de cem modelos “CNN 1”. Sendo assim, é necessário explorar técnicas e arquiteturas de modelo mais eficientes para que possíveis vantagens de treinabilidade sejam relevantes em um contexto pragmático.

Como continuidade da pesquisa, é necessário avaliar um maior número de experimentos para a comparação entre modelos clássicos e híbridos. Ademais, os experimentos devem ser realizados em conjuntos de dados mais complexos como o ImageNet, ou seja, com um maior número de classes. Modificações no modelo híbrido são necessárias para permitir um maior número de camadas convolucionais quânticas, ampliando a profundidade do modelo.

### 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. O trabalho utilizou recursos computacionais disponibilizados pelo Centro de Inteligência Artificial da Universidade de São Paulo (C4AI).

### 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, G., Chen, Q., Long, S., Zhu, W., Yuan, Z., and Wu, Y. (2023). Quantum convolutional neural network for image classification. *Pattern Analysis and Applications*, 26(2):655–667.
- Cong, I., Choi, S., and Lukin, M. D. (2019). Quantum convolutional neural networks. *Nature Physics*, 15(12):1273–1278.
- Corli, S., Moro, L., Dragoni, D., Dispenza, M., and Prati, E. (2025). Quantum machine learning algorithms for anomaly detection: A review. *Future Generation Computer Systems*, 166:107632.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. IEEE.
- Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In Teh, Y. W. and Titterton, M., editors, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*, pages 249–256, Chia Laguna Resort, Sardinia, Italy.
- Gong, L.-H., Pei, J.-J., Zhang, T.-F., and Zhou, N.-R. (2024). Quantum convolutional neural network based on variational quantum circuits. *Optics Communications*, 550:129993.
- 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).

- Hur, T., Kim, L., and Park, D. K. (2022). Quantum convolutional neural network for classical data classification. *Quantum Machine Intelligence*, 4(1).
- Jing, Y., Li, X., Yang, Y., Wu, C., Fu, W., Hu, W., Li, Y., and Xu, H. (2022). Rgb image classification with quantum convolutional ansatz. *Quantum Information Processing*, 21(3).
- Kharsa, R., Bouridane, A., and Amira, A. (2023). Advances in quantum machine learning and deep learning for image classification: A survey. *Neurocomputing*, 560:126843.
- Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Krizhevsky, A. and Hinton, G. (2009). Learning multiple layers of features from tiny images. Technical Report 0, University of Toronto, Toronto, Ontario.
- Kübler, J., Buchholz, S., and Schölkopf, B. (2021). The inductive bias of quantum kernels. *Advances in Neural Information Processing Systems*, 34:12661–12673.
- Magalhães, M. (2006). *Probabilidade e Variáveis Aleatórias*. Edusp.
- Maronese, M., Destri, C., and Prati, E. (2022). Quantum activation functions for quantum neural networks: M. maronese et al. *Quantum Information Processing*, 21(4):128.
- 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.
- Park, D. K., Petruccione, F., and Rhee, J.-K. K. (2019). Circuit-based quantum random access memory for classical data. *Scientific Reports*, 9(1).
- Parthasarathy, R. and Bhowmik, R. T. (2021). Quantum optical convolutional neural network: A novel image recognition framework for quantum computing. *IEEE Access*, 9:103337–103346.
- 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.
- Priyanka, G. S., Venkatesan, M., and Prabhavathy, P. (2023). Advancements in quantum machine learning and quantum deep learning: A comprehensive review of algorithms, challenges, and future directions. In *2023 International Conference on Quantum Technologies, Communications, Computing, Hardware and Embedded Systems Security (iQ-CCHES)*, pages 1–8.
- 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.

- Song, Y., Li, J., Wu, Y., Qin, S., Wen, Q., and Gao, F. (2024). A resource-efficient quantum convolutional neural network. *Frontiers in Physics*, 12:1362690.
- Umeano, C., Paine, A. E., Elfving, V. E., and Kyriienko, O. (2025). What can we learn from quantum convolutional neural networks? *Advanced Quantum Technologies*, 8(7):2400325.
- Wang, Y., Wang, Y., Chen, C., Jiang, R., and Huang, W. (2022). Development of variational quantum deep neural networks for image recognition. *Neurocomputing*, 501:566–582.
- Wei, S., Chen, Y., Zhou, Z., and Long, G. (2022). A quantum convolutional neural network on nisq devices. *AAPPS bulletin*, 32(1):2.
- Zheng, J., Gao, Q., Lü, J., Ogorzałek, M., Pan, Y., and Lü, Y. (2023). Design of a quantum convolutional neural network on quantum circuits. *Journal of the Franklin Institute*, 360(17):13761–13777.