

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

Victor Guerra Martins

**Detecção automática de realce de contraste em imagens de ressonância
magnética cerebral ponderada em T1 usando inteligência artificial**

Ribeirão Preto

2023

VICTOR GUERRA MARTINS

Detecção automática de realce de contraste em imagens de ressonância magnética cerebral ponderada em T1 usando inteligência artificial

Versão Original

Dissertação apresentada ao Programa de Pós-graduação em Ciências das Imagens e Física Médica da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, para obtenção do título em Mestrado Profissionalizante.

Área de concentração: Diagnóstico por Imagem
Orientador: Prof. Dr. Antonio Carlos dos Santos

Ribeirão Preto

2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

FICHA CATALOGRÁFICA

Martins, Victor Guerra

Detecção automática de realce de contraste em imagens de ressonância magnética cerebral ponderada em T1 usando inteligência artificial

/ Victor Guerra Martins; Orientador: Antônio Carlos dos Santos - Ribeirão Preto, 2023

20 páginas, 8 ilustrações

Dissertação de mestrado profissional (Programa Ciências das Imagens e Física Médica – Área de Concentração: Diagnóstico por Imagem) – Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.

1) Realce pós-contraste; 2) Imagem por Ressonância Magnética; 3) Inteligência artificial

FOLHA DE APROVAÇÃO

Martins, Victor Guerra.

Título: “Detecção automática de realce de contraste em imagens de ressonância magnética cerebral ponderada em T1 usando inteligência artificial.”

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para obtenção do título em Mestrado Profissionalizante.

Área de concentração: Diagnóstico por Imagem.

Aprovado em: ____/____/____

Banca examinadora:

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

“O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001”.

RESUMO

MARTINS, V. G. Detecção automática de realce de contraste em imagens de ressonância magnética cerebral ponderada em T1 usando inteligência artificial. 2023. 20f. Tese (Mestrado profissional). Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2023.

Introdução: O uso da ressonância nuclear magnética (RM) na prática médica é extensa e abrange auxílio no diagnóstico de múltiplas condições como: neoplasias, infecções, doenças inflamatórias, dentre outras. Nestes contextos, a análise do realce pós-contraste é essencial. Tal detecção e quantificação do realce em imagens de (RM) cerebral ponderada em T1 é uma tarefa desafiadora devido à complexidade das imagens e à variabilidade na aparência do realce de contraste. Técnicas de inteligência artificial (IA) têm potencial para melhorar a eficiência e precisão desse processo, mas o desenvolvimento e avaliação de abordagens baseadas em IA para esse fim é uma tarefa laboriosa. **Objetivos:** Desenvolvimento e validação de ferramenta para detecção de realce pós-contraste em imagens de RM ponderadas em T1. Implementação desta ferramenta na rotina clínica do serviço de radiologia e diagnóstico por imagem do Hospital das Clínicas de Ribeirão Preto (HCRP). **Método:** Neste estudo, desenvolveremos e avaliaremos uma abordagem baseada em IA para a detecção de realce de contraste em imagens de ressonância magnética cerebral ponderadas em T1 usando um grande conjunto de dados de um hospital público brasileiro. A abordagem baseada em IA será baseada em uma rede neural convolucional (CNN) treinada em uma base de aprendizado supervisionado, com o objetivo de prever com precisão a presença ou ausência de realce de contraste nas imagens. A entrada para a CNN serão as imagens de ressonância magnética e a saída será uma previsão binária da presença ou não de realce de contraste. **Resultados:** Foi observada uma sensibilidade de 78,5% e especificidade foi de 88,2%, sendo a acurácia de 85,7% entre os 7638 exames avaliados. **Conclusão:** A ferramenta desenvolvida obteve sucesso em seus objetivos com alta sensibilidade e especificidade na detecção de realce pós-contraste na sequência ponderada em T1, de forma totalmente automática. Durante nossa rotina diária, o software mostrou auxílio real e prático na avaliação destes exames, diminuindo o tempo para confecção dos laudos radiológicos. Os exames falso positivos são fácil e rapidamente descartados pelo médico radiologista. **Palavras-chave:** realce pós-contraste; imagem por ressonância magnética; inteligência artificial.

ABSTRACT

MARTINS, V. G. Automatic detection of contrast enhancement in T1-weighted brain magnetic resonance images using artificial intelligence. 2023. 20f. Thesis (Professional Masters). Ribeirao Preto Medical School – Sao Paulo University, Ribeirao Preto, 2023.

Introduction: The use of magnetic resonance imaging (MRI) in medical practice is extensive and encompasses the diagnosis of multiple conditions such as neoplasms, infections, inflammatory diseases, among others. In these contexts, the analysis of post-contrast enhancement is essential. The detection and quantification of enhancement in T1-weighted cerebral MRI images is a challenging task due to the complexity of the images and the variability in the appearance of contrast enhancement. Artificial intelligence (AI) techniques have the potential to improve the efficiency and accuracy of this process, but the development and evaluation of AI-based approaches for this purpose is a laborious task. **Objectives:** Development and validation of a tool for post-contrast enhancement detection in T1-weighted MRI images. Implementation of this tool in the clinical routine of the radiology and diagnostic imaging service of Clinical Hospital of Ribeirao Preto. **Methodology:** In this study, we will develop and evaluate an AI-based approach for the detection of contrast enhancement in T1-weighted cerebral MRI images using a large dataset from a Brazilian public hospital. The AI-based approach will be based on a convolutional neural network (CNN) trained on a supervised learning basis, with the objective of accurately predicting the presence or absence of contrast enhancement in the images. The input to the CNN will be the MRI images, and the output will be a binary prediction of the presence or absence of contrast enhancement. **Results:** A sensitivity of 78,5% and a specificity of 88,2% were observed, with an accuracy of 85,7% among the 7638 tests evaluated. **Conclusion:** The developed tool successfully achieved its objectives with high sensitivity and specificity in the detection of post-contrast enhancement in T1-weighted sequences, fully automatically. During our daily routine, the software provided real and practical assistance in the evaluation of these exams, reducing the time required to produce radiology reports. False positive exams are easily and quickly discarded by the radiologist. **Keywords:** post-contrast enhancement; magnetic resonance imaging; artificial intelligence.

ABREVIATURAS E SIGLAS

BH	Barreira hematoencefálica
CCIFM	Centro de imagens e física médica
CNN	Rede neural convolucional
CT	Tomografia computadorizada
HCRP	Hospital das clínicas de Ribeirão Preto
IA	Inteligência artificial
RM	Ressonância Magnética
TCLE	Termo de esclarecimento livre e esclarecido

SUMÁRIO

RESUMO

INTRODUÇÃO.....	1
MATERIAIS E MÉTODOS	4
RESULTADO E DISCUSSÃO	8
REFERÊNCIAS.....	11
ANEXOS	14

INTRODUÇÃO:

O uso da ressonância nuclear magnética (RM) na prática médica é extensa e abrange auxílio no diagnóstico de múltiplas condições como: neoplasias, infecções, doenças inflamatórias, dentre outras. Nestes contextos, a análise do realce pós-contraste é essencial. Estima-se que cerca de 40% de todos os exames RM feitos na Europa e nos Estados Unidos da América se fazem com o uso de agente de contraste de Gadolínio seja para o diagnóstico, monitorização ou caracterização de lesões (1,2). Na neurorradiologia, o uso da sequência baseada em T1 pós-contraste serve de pilar de sustentação para detecção e delimitação precisa de tumores primário e secundário do encéfalo (3,4), bem como para análise de neuroinfecções (5).

A barreira hematoencefálica (BH) se define como uma fronteira seletiva entre o sistema nervoso central e o sangue circulante. Tal barreira se faz de forma complexa por meio de células endoteliais, pericitos e astrócitos (6). Os meios de contraste utilizados em exames de imagem, sejam eles a base de Iodo, como os utilizados na tomografia computadorizada (CT) sejam à base de gadolínio, como os utilizados nos exames de RM, são compostos hidrofílicos, baixa afinidade às proteínas plasmáticas e baixos coeficientes de partição (7,8), portanto, não penetrando a BH. Seja por tumor, infecção ou inflamação, uma lesão nestes componentes da BH resultam numa alteração de sua permeabilidade, tornando possível então que o agente de contraste ultrapasse essa barreira, e penetre no tecido nervoso adjacente.

O mecanismo fisiopatológico referido envolve a administração intravenosa destes agentes de contraste, ocorrendo então a sua passagem pela falha da BH ocorrendo então a absorção e uma distinta reflexão dos campos magnéticos, diferenciando esta região das demais circunjacentes. Isto resulta num encurtamento do tempo de relaxamento deste tecido, expressado na forma de aumento da intensidade do sinal das imagens ponderadas em T1, tornando estas áreas realçadas mais visíveis (Figura 1).

Dado sua importância e prevalência, capacidade de medir com precisão e quantitativamente o aumento do contraste é, portanto, de grande importância no diagnóstico e tratamento de doenças neurológicas.

No entanto, a identificação manual e a quantificação do realce do contraste podem ser demoradas e sujeitas à variabilidade do observador. Isso é particularmente verdadeiro nos casos em que o realce é sutil ou difuso, ou quando as imagens são de baixa qualidade. Além disso, a interpretação do realce do contraste pode ser influenciada por fatores como o conhecimento e a experiência do radiologista, a sequência específica de RM e os parâmetros utilizados e a

presença de outras anormalidades que podem causar alterações de sinal.

Para superar limitações como estas, tem havido um interesse crescente no uso de técnicas de inteligência artificial (IA), notadamente no campo de aprendizado de máquina visando a automatização e otimização de tarefas como tal (9,18). A detecção automática do realce pós-contraste em imagens de ressonância magnética do cérebro ponderadas em T1 é uma das possíveis ferramentas a serem desenvolvidas. Os algoritmos de IA têm o potencial de melhorar a eficiência do processo e reduzir a dependência da interpretação subjetiva (10,11,16,17).

A detecção precoce do realce de contraste é de fundamental importância, haja visto as principais condições às quais esse realce é detectado, como explicado anteriormente. Em um ambiente a nível hospitalar, quanto maior a complexidade dos pacientes que esses locais atenderem, maior é a incidência de patologias graves, como neoplasia, infecções, doenças inflamatórias. Numa rotina de exames destes hospitais, os exames são laudados, via de regra, de acordo com a ordem à qual eles são realizados. Desta forma, não é incomum que exames de maior complexidade e gravidade sejam analisados posteriormente a exames com menor urgência.

Com a ferramenta proposta, pretendemos fazer uma triagem direto na máquina de ressonância magnética de forma a identificarmos potenciais gravidades nos exames feitos através da detecção automatizada de realce de contraste nos exames feitos. Desta forma, estes exames podem ser priorizados pelo médico radiologista e terem suas análises e relatórios realizados com maior agilidade. Esta diminuição entre o tempo de realização do exame e o tempo de finalização do relatório radiológico é essencial para a otimização do tratamento do paciente e consequente melhora dos índices de sucesso destes tratamentos.

Neste estudo, pretendemos desenvolver e avaliar uma abordagem baseada em IA para detectar o realce de contraste à base de gadolínio em imagens de ressonância magnética cerebral ponderadas em T1 usando um grande conjunto de dados de um hospital público brasileiro. A abordagem baseada em IA será baseada em uma rede neural convolucional (CNN) treinada em uma base de aprendizado supervisionado, com o objetivo de prever com precisão a presença ou ausência de realce de contraste nas imagens. A entrada para a CNN serão as imagens de ressonância magnética, e a saída será uma previsão binária da presença ou não de realce de contraste, com um valor de 1 indicando a presença de realce de contraste e um valor de 0 indicando sua ausência (12).

Avaliaremos o desempenho da abordagem baseada em IA usando uma variedade de métricas,

como sensibilidade, especificidade e acurácia, calculado usando um conjunto de teste de validação de imagens que não foram incluídas no conjunto de treinamento.

O uso de IA dentro da medicina data de muitos anos atrás. Em 1989, Hart e colaboradores estudaram a primeira aplicação de uma rede neural artificial para análise de precordialgia (13). Já dentro da neurologia, um dos primeiros estudos foi feito por Anderer e colaboradores em 1994 com o estudo de uma rede neural para análise de eletroencefalogramas e diagnóstico de demência (14). Na radiologia o primeiro estudo data de 1992 com a detecção automática de microcalcificações em mamografia (15). A detecção automática de realce de contraste em imagens de ressonância magnética cerebral ponderada em T1 é uma área ativa de pesquisa, e há vários desafios que precisam ser enfrentados para tornar esses métodos práticos e confiáveis para uso clínico. Um desafio é a disponibilidade de dados de treinamento de alta qualidade. A precisão do modelo AI depende da qualidade e diversidade dos dados de treinamento, e é importante ter um número suficiente de exemplos de imagens aprimoradas e não aprimoradas para obter um desempenho satisfatório. Também é importante garantir que os dados de treinamento sejam representativos da população de imagens que serão encontradas na prática clínica e corrigir quaisquer vieses ou desequilíbrios nos dados.

Outro desafio é a variabilidade na aparência do realce do contraste. A aparência do realce do contraste pode variar dependendo da patologia específica: neoplasia, infecção, inflamação; do momento da imagem: diagnóstico, pós-tratamento, seguimento; da dose e do tipo de agente de contraste usado e dos parâmetros de imagem. Esta variabilidade pode dificultar o desenvolvimento de um único modelo aplicável a todas as situações, podendo ser necessário desenvolver e avaliar múltiplos modelos ou adaptar o modelo a cenários clínicos específicos.

Além disso, existem diversos fatores que podem afetar o desempenho do modelo de IA, como a presença de ruídos ou artefatos nas imagens, a resolução das imagens e a presença de outras anormalidades que podem causar alterações no sinal. É importante levar esses fatores em consideração ao projetar e avaliar o modelo e otimizá-lo para o contexto clínico específico no qual será usado.

Apesar desses desafios, o uso de IA para a detecção de realce de contraste em imagens de ressonância magnética cerebral ponderada em T1 tem o potencial de melhorar significativamente a eficiência do processo e reduzir a dependência da interpretação subjetiva. Mais pesquisas são necessárias para otimizar e validar esses métodos em um ambiente clínico e para determinar o impacto potencial na interpretação das imagens pelos radiologistas.

Atendendo a estes requisitos propusemos a avaliação e a implementação de um software de auxílio ao médico radiologista para a detecção automática de realce pós-contraste em imagens por RM ponderadas em T1. Para esse trabalho definimos os seguintes objetivos específicos:

1º Desenvolver e avaliar uma abordagem baseada em IA para a detecção de realce de contraste em imagens de ressonância magnética cerebral ponderada em T1 usando um grande conjunto de dados de um hospital público brasileiro.

2º Otimizar a abordagem baseada em IA usando uma combinação de técnicas de regularização e aumento de dados e avaliar o desempenho do modelo usando uma variedade de métricas, incluindo sensibilidade, especificidade e acurácia.

3º Validar os resultados do estudo por meio de uma avaliação clínica do modelo por radiologistas especializados e uma análise por imagem das previsões do modelo.

4º Realizar uma análise qualitativa das previsões do modelo e dos recursos extraídos pela CNN, a fim de identificar potenciais pontos fortes ou fracos da abordagem e obter informações sobre os mecanismos subjacentes ao modelo.

5º Determinar o impacto potencial da abordagem baseada em IA na interpretação de imagens de ressonância magnética cerebral ponderada em T1 por radiologistas e avaliar a viabilidade e praticidade de implementar a abordagem em um ambiente clínico.

Esta ferramenta proporcionará uma automatização e discriminação dos exames considerados positivos (presença de realce). Desta forma, abriremos a possibilidade de implementarmos este software na prática clínica do serviço de radiologia e diagnóstico por imagem otimizando, desta forma, o serviço.

MATERIAIS E MÉTODOS:

Neste estudo, apresentamos uma abordagem baseada em inteligência artificial (IA) para detectar o realce de contraste em imagens ponderadas em T1 de ressonância magnética cerebral. Utilizamos um grande conjunto de dados de um hospital público brasileiro, composto por 7638 imagens que foram adquiridas utilizando equipamentos de ressonância magnética de 1,5T e 3T e classificadas manualmente por um neurorradiologista com 2 anos de experiência como aprimoradas ou não aprimoradas. A classificação foi revisada por um neurorradiologis-

ta com mais de 20 anos de experiência para garantir a precisão dos rótulos.

É importante destacar que os pacientes cujas imagens foram analisadas não assinaram um termo de consentimento livre e esclarecido (TCLE), já que se tratava de um estudo retrospectivo que analisou exames de imagem anonimizados e randomizados, sem a realização de nenhum novo procedimento.

As imagens selecionadas para o estudo foram adquiridas utilizando bobina de cabeça de 8 canais de recepção e apresentaram as seguintes características de aquisição: sequência axial T1 pós-contraste, tempo de repetição de 13,8 ms (1,5T) e 7,2 ms (3T), tempo ao eco de 6,7 ms (1,5T) e 3,4 ms (3T), número de fatias de 65 (1,5T) e 60 (3T) e matriz de aquisição de 240×240 (1,5T e 3T). Foi utilizada uma aplicação intravenosa de 0,1mL/Kg de gadobutrol 1,0 mmol, da marca Gadovist®, para a obtenção do contraste nas imagens.

Todas as imagens foram obtidas do servidor interno do Centro de Imagens e Física Médica (CCIFM) do HCRP e foram analisadas neste estudo. A figura 2 representa o fluxograma separado por etapas da ferramenta elaborada.

Este estudo apresenta uma abordagem baseada em inteligência artificial (IA) para detectar o realce de contraste em imagens ponderadas em T1 de ressonância magnética cerebral, utilizando uma rede neural convolucional (CNN) específica, a ResNet50V2, que é um modelo de aprendizado profundo especialmente adequado para análise de imagens (19,20). A arquitetura ResNet50V2 é conhecida por sua capacidade de lidar com redes profundas sem sofrer de problemas como o desaparecimento do gradiente, melhorando a eficiência e a precisão do modelo. A utilização da arquitetura ResNet50V2 neste estudo tem como objetivo tirar proveito das suas características inovadoras, como conexões de atalho e normalização em lote, que ajudam a acelerar o treinamento e a melhorar a precisão do modelo. A escolha desta rede específica busca proporcionar resultados mais precisos e confiáveis na detecção do realce de contraste em imagens ponderadas em T1 de ressonância magnética cerebral.

Durante o treinamento da CNN, as imagens de ressonância magnética cerebral ponderadas em T1 serão pré-processadas para incluir três cortes consecutivos no plano axial, a fim de melhorar a precisão do modelo e mitigar problemas de variação de orientação e posicionamento das imagens. A saída da CNN será uma predição binária da presença ou ausência de realce pós-contraste.

Nesta fase, o médico radiologista terá acesso a uma plataforma via WEB onde será exposto

aleatoriamente à sequências volumétricas de imagens ponderadas em T1 pós-contraste anonimizadas e randomizadas, e assinará, corte-a-corte, as imagens onde é possível detectar realce (Figura 3).

A CNN será composta por várias camadas, incluindo camadas convolucionais, camadas de agrupamento (pooling) e camadas totalmente conectadas. As camadas convolucionais extrairão as características das imagens de entrada, enquanto as camadas de agrupamento serão usadas para reduzir a resolução espacial dos mapas de características e aumentar a robustez do modelo à translação e escala. As camadas totalmente conectadas combinarão os recursos extraídos pelas camadas convolucionais e de agrupamento para fazer a previsão final.

Para otimizar a CNN, utilizaremos uma variante de gradiente descendente estocástico chamada ADAM (21), que é conhecido por ser eficaz em uma ampla variedade de tarefas e é amplamente utilizado em aplicações como reconhecimento de fala, reconhecimento de imagem e processamento de linguagem natural. O ADAM ajusta de forma adaptativa a taxa de aprendizado com base no primeiro e segundo momentos dos gradientes, o que o torna mais eficiente em encontrar o mínimo global da função de perda. Usaremos um tamanho de mini-lote de 32 imagens e definiremos a taxa de aprendizagem para um valor fixo.

Para evitar o overfitting, utilizaremos várias técnicas de regularização, incluindo o decaimento de peso e o dropout. O decaimento de peso é uma forma de regularização L2 que penaliza pesos grandes, enquanto o dropout é uma técnica que zera aleatoriamente uma fração dos pesos durante o treinamento. Definiremos o coeficiente de decaimento de peso para 0,0005 e a taxa de abandono para 0,5.

Para melhorar ainda mais o desempenho de generalização da CNN, utilizaremos técnicas de aumento de dados (22, 23, 24). As técnicas de aumento de dados são utilizadas para gerar novas imagens a partir das existentes, aplicando rotações, zoom, espelhamentos, deslocamentos e outras transformações, sem alterar a classe da imagem original. Com isso, é possível aumentar o número de exemplos no conjunto de dados e melhorar o desempenho do modelo. Utilizaremos várias técnicas de aumento de dados, incluindo corte aleatório, inversão horizontal e rotação aleatória.

Para avaliar o desempenho da abordagem baseada em IA, utilizaremos um conjunto de testes de validação de imagens que não foram incluídas no conjunto de treinamento. Calcularemos várias métricas, incluindo sensibilidade, especificidade, acurácia e ROC (AUC), para avaliar a precisão do modelo.

Com a utilização dessas técnicas e métricas, esperamos obter um modelo de CNN altamente preciso e robusto para a detecção do realce de contraste em imagens ponderadas em T1 de ressonância magnética cerebral, com implicações importantes para a prática clínica e o tratamento de doenças neurológicas.

Para garantir a reprodutibilidade dos resultados, usaremos uma semente aleatória fixa para todos os experimentos e relataremos todos os detalhes relevantes da configuração experimental, incluindo o conjunto de dados, a arquitetura do modelo, o algoritmo de otimização e as métricas de desempenho.

Em resumo, a abordagem baseada em IA proposta para a detecção de realce de contraste em imagens de ressonância magnética cerebral ponderadas em T1 envolve o uso de uma CNN treinada em uma base de aprendizado supervisionado, usando um grande conjunto de dados de imagens classificadas manualmente como dados de treinamento. A CNN será otimizada usando uma combinação de técnicas de regularização e aumento de dados, e o desempenho do modelo será avaliado usando um conjunto de imagens de teste de validação. O desempenho do modelo será avaliado por meio de várias métricas, incluindo sensibilidade, especificidade e acurácia, e os resultados serão analisados em detalhes para identificar possíveis vieses ou padrões no desempenho do modelo (Figura 4).

Para validar ainda mais a abordagem baseada em IA, também realizaremos uma avaliação clínica do modelo comparando as previsões do modelo com a interpretação das imagens pelos radiologistas. A avaliação clínica será realizada por um painel de radiologistas especialistas com ampla experiência na interpretação de imagens de ressonância magnética cerebral ponderada em T1. Os radiologistas receberão um conjunto de imagens que foram selecionadas aleatoriamente do conjunto de teste, juntamente com as previsões do modelo de IA, e serão solicitados a interpretar independentemente as imagens e classificá-las como aprimoradas ou não aprimoradas.

A avaliação clínica será feita de forma cega, sem que os radiologistas conheçam as previsões do modelo. Os resultados da avaliação clínica serão comparados com as previsões do modelo para avaliar a concordância entre as duas abordagens. Vamos calcular várias métricas, incluindo o coeficiente kappa de Cohen e a porcentagem de concordância, para quantificar o grau de concordância entre o modelo e os radiologistas.

Além da avaliação clínica, também faremos uma análise do desempenho preditivo do modelo por imagem. Para cada imagem do conjunto de teste, calcularemos a probabilidade da ima-

gem ser classificada como aprimorada, conforme previsto pelo modelo, e compararemos essa probabilidade com a classificação real da imagem pelos radiologistas. Isso nos permitirá avaliar a sensibilidade e especificidade do modelo em diferentes limites de probabilidade, bem como identificar quaisquer padrões no desempenho do modelo que possam estar relacionados a tipos específicos de patologias ou características de imagem.

Por fim, também realizaremos uma análise qualitativa das previsões do modelo para identificar possíveis pontos fortes ou fracos da abordagem. Isso envolverá a inspeção visual das imagens e das previsões do modelo, bem como uma análise das características extraídas pela CNN e dos padrões nas previsões. Isso nos permitirá identificar possíveis áreas de melhoria no modelo e obter informações sobre os mecanismos subjacentes da abordagem.

No geral, o estudo proposto tem como objetivo desenvolver e avaliar uma abordagem baseada em IA para a detecção de realce de contraste em imagens de ressonância magnética cerebral ponderadas em T1 usando um grande conjunto de dados de um hospital público brasileiro. A abordagem baseada em IA será baseada em uma CNN treinada em aprendizado supervisionado, e o desempenho do modelo será avaliado usando uma variedade de métricas, incluindo sensibilidade, especificidade e acurácia. Os resultados do estudo serão validados por meio de uma avaliação clínica do modelo por radiologistas especialistas e uma análise por imagem das previsões do modelo, bem como uma análise qualitativa das previsões e características extraídas pela CNN.

RESULTADOS E DISCUSSÃO:

O presente estudo teve como objetivo desenvolver um modelo preditivo baseado em inteligência artificial capaz de realizar a triagem de realce pelo meio de contraste em ressonância magnética de crânio. A arquitetura escolhida para a rede neural foi a ResNet50V2, uma rede profunda com excelente desempenho em tarefas de classificação de imagens e reconhecida por sua capacidade de lidar com redes profundas sem sofrer de problemas como o desaparecimento do gradiente.

O modelo foi treinado por 86 épocas, com uma taxa de aprendizado de $1.0000e-05$. Durante o treinamento, as métricas de AUC (área sob a curva ROC) e acurácia foram monitoradas para avaliar o desempenho do modelo. A função de perda utilizada foi a entropia cruzada binária.

Após o treinamento, o modelo foi avaliado com base nas métricas AUC e acurácia no conjunto de validação. O modelo alcançou uma perda de 0.2484, AUC de 0.9999 e acurácia de

0.9984 no conjunto de treinamento (Figura 5), enquanto no conjunto de validação, obteve uma perda de 0.5268 (Figura 6), AUC de 0.8935 (Figura 7) e acurácia de 0.8592. Esses resultados demonstram que o modelo apresentou um bom desempenho na tarefa de classificação de cortes de sequências T1 pós-contraste de ressonância magnética de crânio. Além disso, a análise da curva ROC ajudou a identificar o limiar de classificação ideal para otimizar a sensibilidade e especificidade do modelo.

Por fim, o modelo treinado foi aplicado ao conjunto de teste, composto por 1528 amostras, para verificar seu desempenho em dados não utilizados anteriormente. Essa etapa é crucial para garantir a generalização do modelo e sua capacidade de lidar com novos casos. As métricas de desempenho, incluindo AUC e acurácia, foram calculadas para o conjunto de teste e comparadas com os resultados dos conjuntos de treinamento e validação. Essa análise permitiu avaliar a eficácia do modelo e a consistência do seu desempenho em diferentes subconjuntos de dados. O modelo proposto apresentou resultados promissores na detecção de realce em imagens de ressonância magnética, com uma AUC de 0.8942 e uma acurácia de 0.8579.

A matriz de confusão do conjunto de teste foi analisada para fornecer insights adicionais sobre o desempenho do modelo. A matriz de confusão indica que o modelo foi capaz de identificar corretamente um grande número de verdadeiros negativos e verdadeiros positivos. No entanto, o modelo também apresentou alguns falsos positivos e falsos negativos. Esses números complementam as métricas de desempenho mencionadas anteriormente e fornecem uma visão mais detalhada do desempenho do modelo em termos de classificações corretas e incorretas (Figura 8).

A análise de sensibilidade e especificidade do modelo foi realizada para determinar a capacidade do modelo de detectar verdadeiros positivos e verdadeiros negativos, respectivamente. A sensibilidade do modelo foi de 0.7858, o que indica que o modelo é capaz de detectar corretamente 78,58% dos casos positivos. Já a especificidade foi de 0.8822, o que indica que o modelo é capaz de identificar corretamente 88,22% dos casos negativos.

O valor do F1 Score, que é uma medida da precisão do modelo, foi de 0.7774. Isso indica que o modelo é capaz de identificar corretamente uma proporção significativa de casos positivos, enquanto minimiza o número de falsos positivos.

O valor do limiar de decisão (threshold) ótimo encontrado foi de 0.2481. Esse valor representa o ponto de corte ideal para o modelo, no qual a sensibilidade e a especificidade são maximizadas simultaneamente.

Esses resultados indicam que o modelo proposto pode ser uma ferramenta útil para a triagem de realce em imagens de ressonância magnética de crânio. No entanto, é importante destacar que o modelo foi treinado e testado em um conjunto de dados específico, proveniente de um único hospital público, e que é necessário realizar validação em conjuntos de dados adicionais para avaliar sua capacidade de generalização.

Ainda assim, esses resultados são promissores e podem representar uma contribuição significativa para a detecção precoce de patologias que apresentam realce em imagens de ressonância magnética de crânio, como tumores cerebrais e infecções. Além disso, o modelo proposto pode ser útil na triagem de pacientes e na otimização do fluxo de trabalho em clínicas e hospitais.

Em relação à discussão mais detalhada, é relevante mencionar que, apesar dos resultados promissores apresentados, algumas limitações devem ser consideradas. A primeira delas é o fato de que os dados utilizados para treinar e validar o modelo foram provenientes de um único hospital público. Isso pode levar a uma possível limitação na capacidade de generalização do modelo, uma vez que a variação entre diferentes hospitais e protocolos de aquisição de imagens não foi considerada. A inclusão de dados provenientes de diferentes instituições e com diferentes protocolos poderia aumentar a robustez do modelo e torná-lo mais aplicável em diferentes contextos.

Outra limitação potencial é a variabilidade na qualidade das imagens e na precisão dos rótulos fornecidos como verdade básica. Em estudos futuros, é importante garantir que as imagens sejam de alta qualidade e que os rótulos sejam precisos e consistentes entre diferentes especialistas.

Apesar das limitações mencionadas, este estudo representa um passo importante no desenvolvimento de uma ferramenta baseada em inteligência artificial para auxiliar na detecção de realce em imagens de ressonância magnética de crânio. Estudos futuros devem focar na validação do modelo em conjuntos de dados mais diversificados e na integração do modelo em sistemas clínicos, a fim de avaliar seu impacto real na prática médica.

Além disso, é fundamental considerar que, embora o modelo apresente uma acurácia e AUC consideráveis, a sensibilidade e especificidade ainda podem ser melhoradas. Essa melhoria pode ser alcançada através de otimizações na arquitetura da rede neural, ajustes nos hiperparâmetros ou até mesmo no uso de técnicas de aumento de dados (data augmentation) para enriquecer o conjunto de treinamento. O aumento de dados pode ser especialmente útil para

equilibrar a distribuição de classes, já que muitas vezes há uma quantidade desigual de casos positivos e negativos em conjuntos de dados médicos.

Outro aspecto importante a ser considerado em estudos futuros é a interpretabilidade do modelo. Embora as redes neurais convolucionais tenham demonstrado alto desempenho em várias tarefas de classificação de imagens, a interpretabilidade dos resultados ainda é um desafio. Técnicas como a ativação de mapas de classe (Class Activation Maps - CAM) podem ser aplicadas para visualizar as áreas da imagem que mais contribuíram para a decisão do modelo, facilitando a compreensão e a explicação dos resultados.

Além disso, a comparação do desempenho do modelo proposto com outras arquiteturas de redes neurais convolucionais e técnicas de aprendizado de máquina pode fornecer informações adicionais sobre a adequação do modelo escolhido para a tarefa em questão. Essa comparação pode ajudar a identificar áreas de melhoria e orientar o desenvolvimento de modelos ainda mais eficazes.

A implementação de um sistema de feedback, onde os médicos radiologistas possam revisar e corrigir as previsões do modelo, também pode ser útil para melhorar continuamente o desempenho do modelo ao longo do tempo. Esse sistema de feedback poderia ser integrado à plataforma web mencionada na metodologia, permitindo a interação direta entre os profissionais de saúde e o modelo de inteligência artificial.

Em conclusão, este estudo apresenta um modelo promissor para a detecção de realce em imagens de ressonância magnética de crânio, utilizando a arquitetura de rede neural convolucional ResNet50V2. Os resultados obtidos sugerem que a abordagem proposta pode ser uma ferramenta valiosa para auxiliar na detecção precoce de patologias que apresentam realce em imagens de ressonância magnética e na triagem de pacientes em clínicas e hospitais. Estudos futuros devem se concentrar na validação do modelo em conjuntos de dados mais diversificados, na otimização de sua arquitetura e hiperparâmetros, e na integração do modelo em sistemas clínicos para avaliar seu impacto real na prática médica.

REFERÊNCIAS

1. RUNGE, Val M. Safety of the gadolinium-based contrast agents for magnetic resonance imaging, focusing in part on their accumulation in the brain and especially the dentate nucleus. *Investigative Radiology*, v. 51, n. 5, p. 273-279, 2016.

2. TOMURA, N. et al. Contrast enhanced multi-shot echo-planar flair in the depiction of metastatic tumors of the brain: comparison with contrast-enhanced spin-echo t1-weighted imaging. *Acta Radiologica*, v. 48, n. 9, p. 1032-1037, 2007.
3. ALPHANDERY, Edouard. Glioblastoma treatments: an account of recent industrial developments. *Frontiers in Pharmacology*, v. 9, p. 879, 2018.
4. ACHROL, Achal Singh et al. Brain metastases. *Nature Reviews Disease Primers*, v. 5, n. 1, p. 1-26, 2019.
5. MUZUMDAR, D. et al. Brain abscess: an overview. *International Journal of Surgery*, v. 9, n. 2, p. 136-144, 2011.
6. SAGE, M. R.; WILSON, A. J. The blood-brain barrier: An important concept in neuroimaging. *AJNR*, v. 15, p. 601-622, 1984.
7. MORRIS, T. W. Intravascular contrast media and their properties. In: SKUCAS, J. *Radiographic Contrast Agents*. 2nd ed. Aspen: Rockville, 1989. p. 119-128.
8. CHANG, C. A. Magnetic resonance imaging contrast agents: Design and physicochemical properties of gadodiamide. *Invest Radiol*, v. 28, p. S21-S27, 1993.
9. CHOY, G. et al. Current Applications and Future Impact of Machine Learning in Radiology. *Radiology*, v. 288, n. 2, p. 318-328, 2018.
10. MURDOCH, T. B.; DETSKY, A. S. The inevitable application of big data to health care. *JAMA*, v. 309, n. 13, p. 1351-1352, 2013.
11. BAXT, W. G. Application of artificial neural networks to clinical medicine. *The Lancet*, v. 346, n. 8983, p. 1135-1138, 1995.
12. KRIZHEVSKY, A. et al. ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems 25*, 2012. p. 1090-1098.
13. HART, A.; WYATT, J. Connectionist models in medicine: an investigation of their potential. In: *Proceedings of AIME'89*. Heidelberg: Springer, 1989. p. 115-124.
14. ANDERER, P. et al. Discrimination between demented patients and normals based on topographic EEG slow wave activity: comparison between z statistics, discriminant analysis and artificial neural network classifiers. *Electroencephalogr Clin Neurophysiol*, v. 91, n. 2, p. 108-117, 1994.

15. DRIVER, C. N. et al. Artificial Intelligence in Radiology: A call for Thoughtful Application. *Clin Transl Sci*, v. 13, p. 216-218, 2020.
16. LUI, Y. W. et al. Inteligência Artificial em Neuroradiologia: Estado Atual e Direções Futuras. *AJNR American Journal of Neuroradiology*, v. 41, n. 8, p. E52-E59, 2020.
17. ZHANG, Z. et al. Aplicação de Inteligência Artificial na Classificação de Ressonância Magnética de Doenças Neurológicas e Psiquiátricas do Cérebro Humano: Uma Revisão Abrangente. *Diagnostics (Basel)*, v. 11, n. 8, p. 1402, 2021.
18. LITJENS, G. et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, v. 42, p. 60-88, 2017.
19. HE, K. et al. Identity mappings in deep residual networks. In: *Computer Vision - ECCV 2016*, 2016. p. 9908, *Lecture Notes in Computer Science*.
20. GOODFELLOW, I. et al. *Deep Learning*. The MIT Press, Cambridge, MA, 2016.
21. BÔNE, A. et al. Contrast-Enhanced Brain MRI Synthesis With Deep Learning: Key Input Modalities and Asymptotic Performance. *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, p. 1159-1163, 2021.
22. MIKOŁAJCZYK, A.; GROCHOWSKI, M. Data augmentation for improving deep learning in image classification problem. *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, IEEE, 2018.
23. NALEPA, J. et al. Data augmentation for brain-tumor segmentation: a review. *Frontiers in Computational Neuroscience*, v. 13, 2019.
24. SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of Big Data*, v. 6, p. 1-48, 2019.

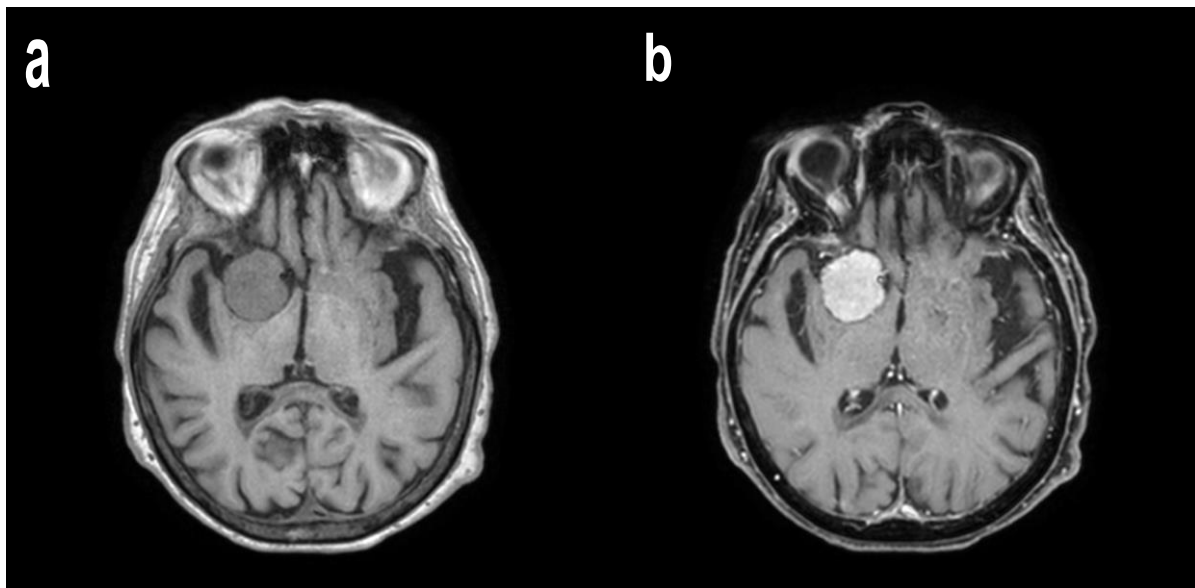


Figura 1: Exemplo de paciente com neoplasia intracraniana (meningioma) em exames pré (a) e pós-contraste (b).

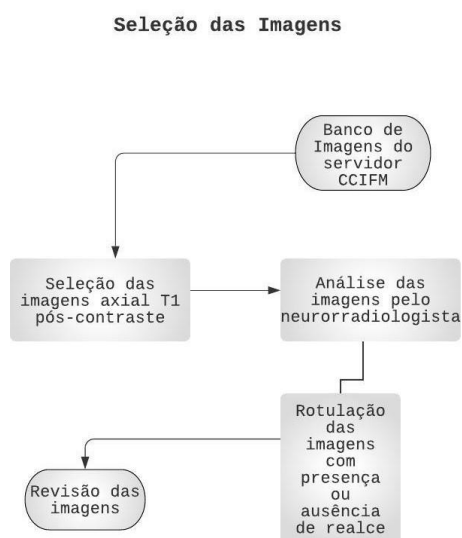


Figura 2: Fluxograma do funcionamento da ferramenta de seleção e rotulagem dos exames. Após seleção automatizada dos exames de RM axial T1 pós-contraste feita dentro do banco de imagens do servidor do CCIFM, o neurorradiologista analisa tais imagens e seleciona, dentro dos cortes destes exames, quais deles possuem realce e quais não. Detectado o realce, o radiologista seleciona e discrimina todos os cortes do exame os quais possuem este realce. Feita a rotulagem de todos os exames, as imagens são então revisadas.



Figura 3: Interface da ferramenta WEB onde o neurorradiologista avalia e seleciona os cortes com realce. Superiormente: corte a ser avaliado; inferiormente: rotulagem deste corte como positivo para realce.

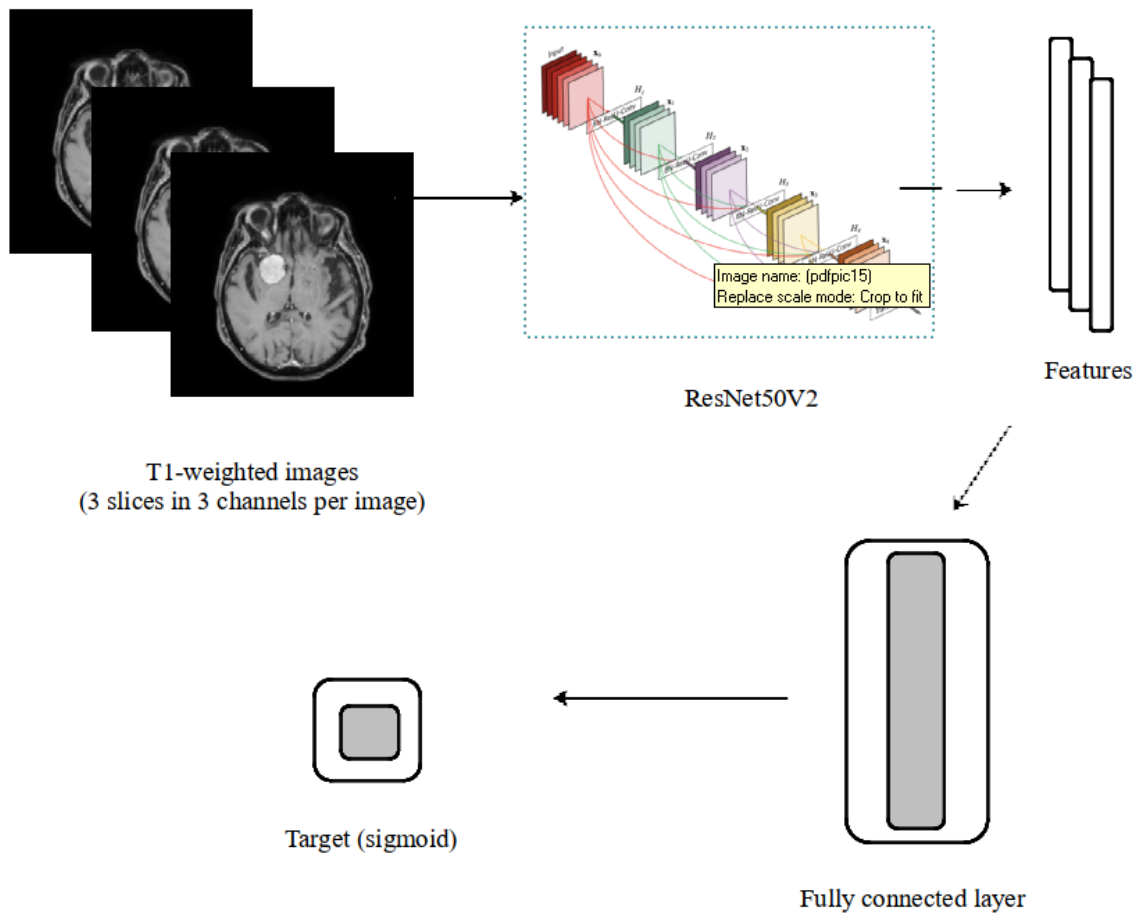


Figura 4: Representação esquemática da rede neural.

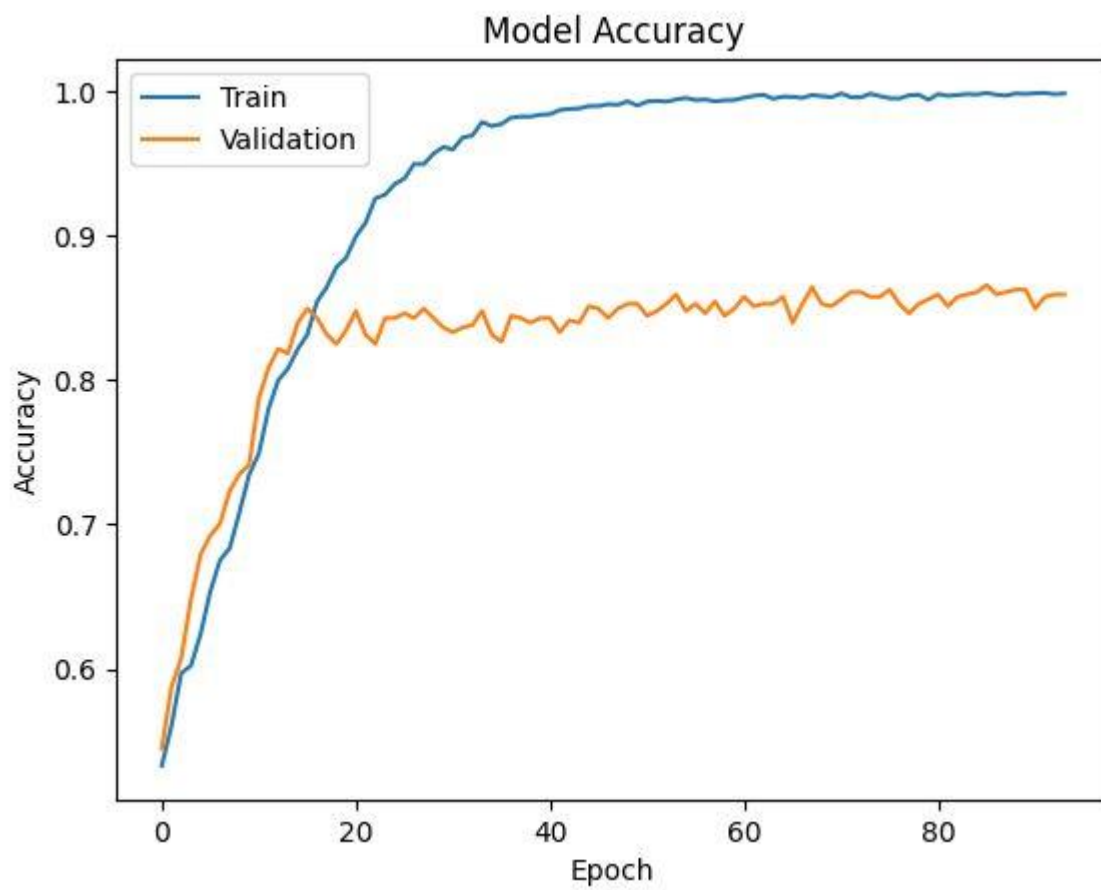


Figura 5: Acurácia durante o treinamento da rede neural.

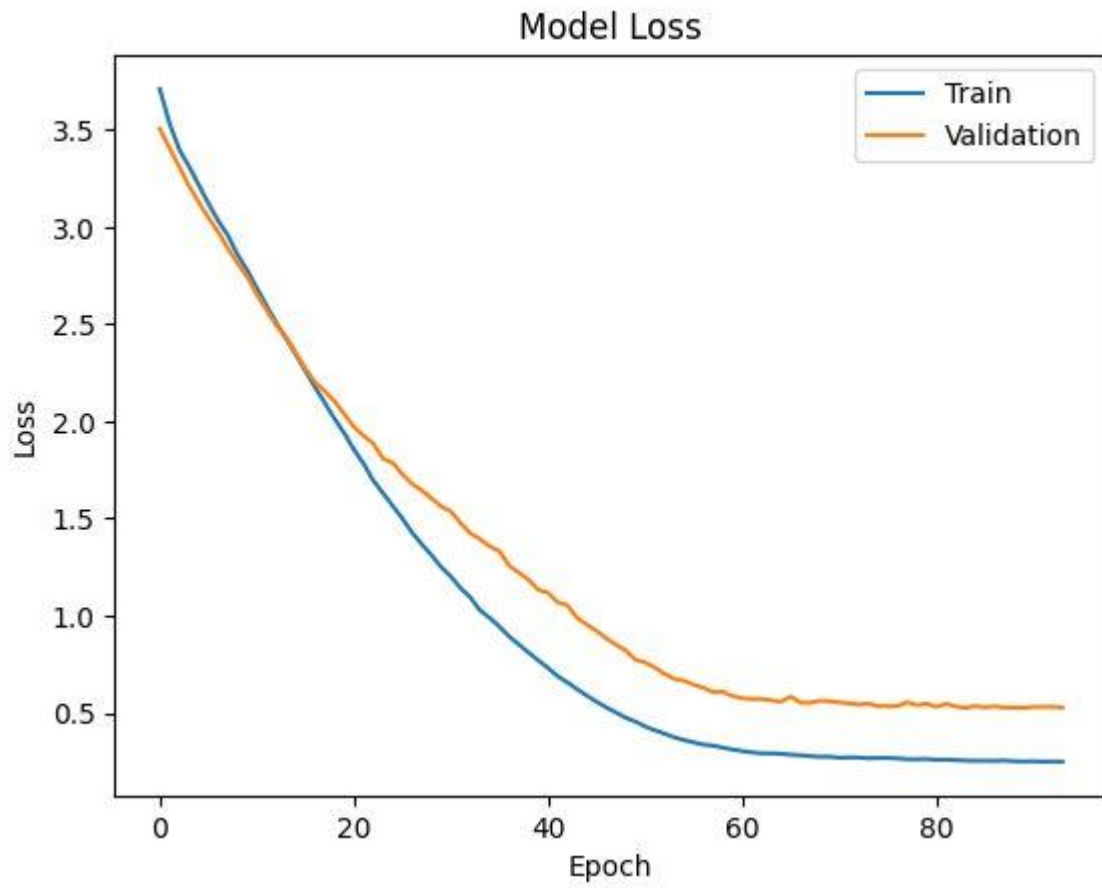


Figura 6: Loss durante o treinamento da rede neural.

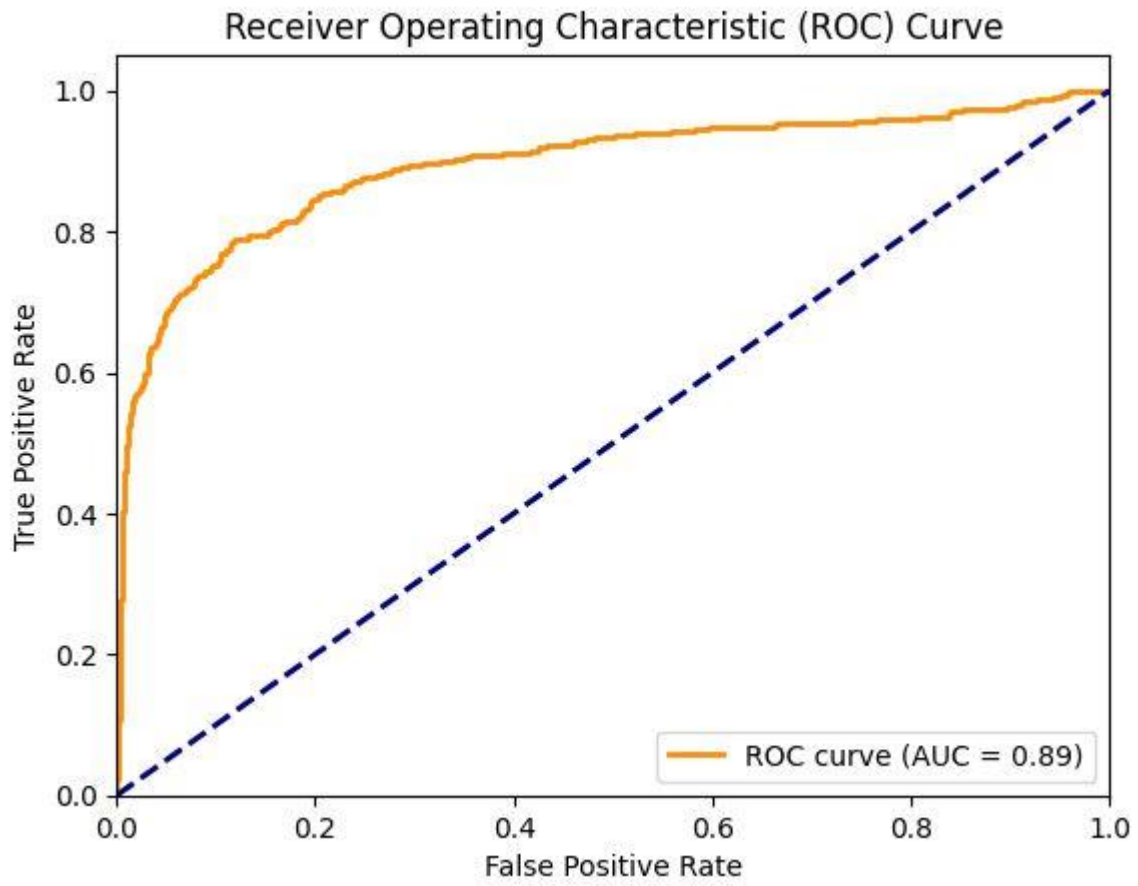


Figura 7: Curva ROC (AUC).

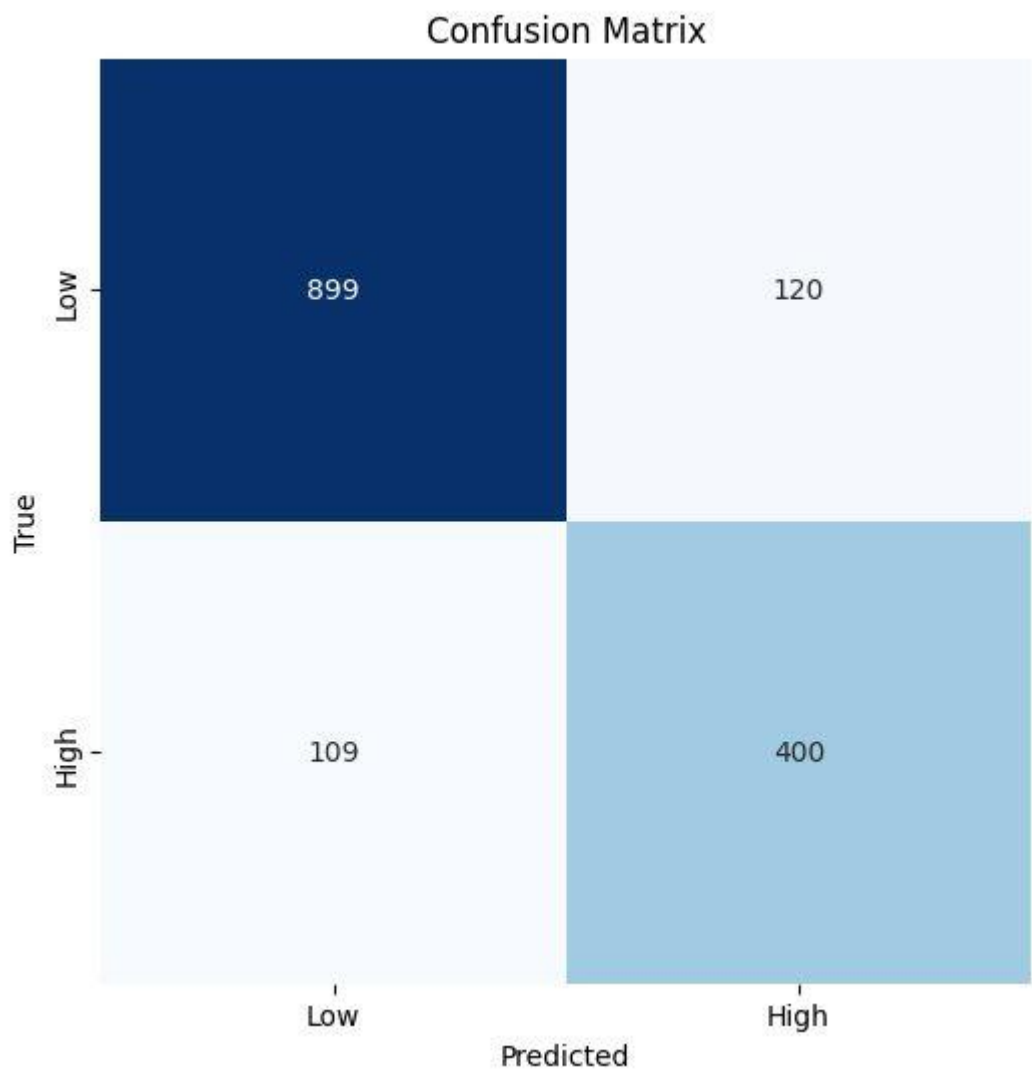


Figura 8: Matriz de confusão.