UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE FÍSICA

Allan Bruno de Andrade Corrêa

DESENVOLVIMENTO DE REDE NEURAL PARA A AVALIAÇÃO DE RESOLUÇÃO ESPACIAL EM EQUIPAMENTOS DE GAMA CÂMARA DO HOSPITAL UNIVERSITÁRIO CLEMENTINO FRAGA FILHO

Rio de janeiro
2024

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE FÍSICA

Allan Bruno de Andrade Corrêa

DESENVOLVIMENTO DE REDE NEURAL PARA A AVALIAÇÃO DE RESOLUÇÃO ESPACIAL EM EQUIPAMENTOS DE GAMA CÂMARA DO HOSPITAL UNIVERSITÁRIO CLEMENTINO FRAGA FILHO

Trabalho de conclusão de curso de graduação apresentado ao Instituto de física da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Bacharel em Física Médica.

Orientadora: Dra. Luana Soares Jorge

Rio de Janeiro
2024

RESUMO

A gama câmara é um dos sistemas de diagnóstico por imagem de maior importância utilizado na Medicina Nuclear, e a resolução espacial intrínseca do sistema é estimada pelo teste de resolução espacial planar da gama câmara, previsto em norma CNEN 3.05. No HUCFF, o teste consiste na análise visual da imagem gerada utilizando um fantoma de barras com quadrantes de espessuras diferentes. Sua avaliação incorreta pode acarretar em maior dificuldade de conseguir identificar estruturas anatômicas relevantes ao diagnóstico do paciente. Este trabalho tem como objetivo geral estudar a viabilidade da implementação de algoritmos de inteligência artificial para a avaliação do teste de resolução espacial intrínseca, realizado através do método analógico. Para o treinamento da rede neural, foram coletadas 144 imagens do teste mencionado acima, e quatro arquiteturas de redes neurais foram testadas com diferentes combinações de parâmetros. As arquiteturas de redes neurais apresentaram um alto desvio padrão em relação à acurácia dos modelos. Os valores de desvio padrão encontrados evidenciam que o mesmo modelo, quando executado em sequência, foi incapaz de convergir para um conjunto de acurácias preciso, evidenciando então a baixa confiabilidade de seu uso. Sugere-se que em trabalhos futuros, seja utilizado um volume maior de imagens com o objetivo de melhorar o desempenho e a confiabilidade do modelo.

Palavras chave: gama câmara; resolução espacial intrínseca; redes neurais;

ABSTRACT

The gamma camera is one of the most important imaging diagnostic systems used in nuclear medicine, the system's intrinsic spatial resolution is estimated by the planar spatial resolution test of the gamma camera, as seen in CNEN 3.05 standard. At HUCFF, the test consists in a visual analysis of the generated image using a bar phantom with quadrants of different widths. An improper evaluation can make it hard to evaluate anatomic structures crucial to diagnosis. This work aims to study the viability of implementing artificial intelligence algorithms to evaluate the intrinsic spatial resolution test conducted through the analog method. A total of 144 images of the mentioned test were collected, and four neural network architectures were tested with varying parameter settings. The neural network architectures exhibited a high standard deviation concerning the model's accuracy. The standard deviation values indicate that the same model, when trained repeatedly, failed to converge to a precise set of accuracies, highlighting the low reliability of its use. In future studies, a larger dataset to enhance the model's performance and reliability can be employed.

Keywords: gamma camera; intrinsic spatial resolution; neural network;

Lista de Figuras

Figura 1: Imagem mostrando o fluxo de informação em um exame de Gama Câmara. Figura adaptada de [5]....11
Figura 2: Imagem gerada no teste de controle de qualidade da resolução espacial e linearidade no SMN-HUCFF....16
Figura 3: Modelo linear de neurônio. Figura retirada de [13]....21
Figura 4: Curva da função de ativação heaviside. Figura retirada de [13]....22
Figura 5: Curva da função de ativação sigmóide. Figura retirada de [13]....23
Figura 6: Curva da função de ativação linear retificada. Figura retirada de [14]....23
Figura 7: Rede neural do tipo feedfoward. Figura retirada de [15]....24
Figura 8: Feedback networks. Figura retirada de [13]....25
Figura 9: Camada de convolução no processo de aprendizagem....27
Figura 10: Imagem do processo de pooling. Imagem adaptada de [17]....28
Figura 11: Gama câmara Millennium MG do setor de medicina nuclear do HUCFF sem os colimadores....29
Figura 12: Fantoma utilizado no teste de resolução espacial intrínseca....30
Figura 13: Configuração do bipod no momento da coleta da imagem utilizada na avaliação da resolução espacial intrínseca....32
Figura 14: Par de imagens invertidas feitas com o fantoma de barras de quatro quadrantes. 33
Figura 15: Relatório do teste de outubro de 2020....33
Figura 16: Metadados do arquivo dicom....34
Figura 17: Modelo de imagem original disponibilizada nos arquivos dicom....35
Figura 18: Camadas da arquitetura 1....41
Figura 19: Camadas da arquitetura 2....43
Figura 20: Camadas da arquitetura 3....45
Figura 21: Camadas da arquitetura 4....47
Figura 22: Exemplo de distribuição de probabilidade de modelo da arquitetura 2....49
Figura 23: Exemplo de distribuição de probabilidade de modelo da arquitetura 3....49
Figura 24: Curva ROC de modelo da arquitetura 1....49
Figura 25: Curva ROC de modelo da arquitetura 2....50
Figura 26: Curva ROC de modelo da arquitetura 3....50
Figura 27: Curva ROC de modelo da arquitetura 4....50

Lista de Tabelas

Tabela 1: Parâmetros otimizados nas arquiteturas de rede neural.....42
Tabela 2: Combinações de parâmetros e seus respectivos resultados após 10 iterações em cada modelo da arquitetura 1.....44
Tabela 3: Combinações de parâmetros e seus respectivos resultados após 10 iterações em cada modelo da arquitetura 2.....46
Tabela 4: Combinações de parâmetros e seus respectivos resultados após 10 iterações em cada modelo da arquitetura 3.....48
Tabela 5: Combinações de parâmetros e seus respectivos resultados após 10 iterações em cada modelo da arquitetura 1.....50
Tabela 6: Comparação da melhor acurácia obtida para cada arquitetura.....50

SUMÁRIO

1 INTRODUÇÃO....9
1.1 Objetivo....10
2 FUNDAMENTAÇÃO TEÓRICA....10
2.1 História da Medicina Nuclear....10
2.2 A gama câmara....11
2.3 Qualidade da imagem na gama câmara....13
2.4 Parâmetros de desempenho....13
2.5 Controle de qualidade....16
2.6 Aquisição de imagem....19
2.7 Redes neurais artificiais....20
2.7.1 Neurônio artificial....21
2.7.2 Função de ativação....22
2.7.3 Arquiteturas de redes neurais....24
2.7.4 Aprendizagem de redes neurais....26
2.7.5 Redes neurais convolucionais....26
2.7.6 Camada convolucional....27
2.7.7 Camada de *pooling*....28
2.7.8 Camada de *dropout*....29
3 METODOLOGIA....29
3.1 Gama câmara....29
3.2 Fantoma de resolução espacial intrínseca e linearidade....30
3.3 Protocolo de aquisição de dados do teste de resolução espacial intrínseca....31
3.4 Ambiente de desenvolvimento....32
3.5 Coleta de dados....32
3.5.1 Aquisição das imagens....32
3.5.2 Gabarito dos físicos....33
3.5.3 Relatórios....34
3.5.4 Tratamento e transformação dos dados....34
3.6 Rede Neural....36
3.6.1 Função de erro....37
3.6.2 Função de otimização....37
3.6.3 Abordagem de otimização de resultados....39
3.7 Abordagem de avaliação....39
3.8 Abordagem de definição do ponto de corte....40
4 RESULTADOS E DISCUSSÃO....41
4.1 Arquitetura 1....42

4.2 Arquitetura 2.....44
4.3 Arquitetura 3.....45
4.4 Arquitetura 4.....47
4.5 Acurácia sem *youden*.....49
4.6 Curva característica de operação do receptor.....50
4.7 Comparativo com os relatórios históricos.....51
5 CONCLUSÃO.....52

1 INTRODUÇÃO

A gama câmara é um dos equipamentos de diagnóstico por imagem por emissão de maior importância utilizados em clínicas de Medicina Nuclear. Neste tipo de equipamento, a realização do exame possibilita obter imagens funcionais do corpo do paciente, através da administração de um composto radiofármaco que é biodistribuido no organismo de acordo com o objetivo do exame. Portanto, a finalidade do seu uso costuma ser diferente de equipamentos de raio-x e tomografia, que geram imagens anatômicas.

Os parâmetros de desempenho dos equipamentos de gama câmara utilizados são descritos pela *National Electrical Manufacturers Association* (NEMA) no documento *Performance Measurements of Scintillation Cameras*. Dentre eles, estão: resolução de energia, uniformidade (integral e diferencial), taxa de contagem máxima do sistema, linearidade e resolução espacial entre outros [1].

A resolução espacial, como parâmetro da gama câmara, define a distância mínima em que é possível identificar dois pontos distintos de captação de radiofármaco em uma imagem. É um parâmetro fundamental no meio hospitalar, visto que influencia diretamente na acurácia do laudo médico, e consequentemente na eficácia do diagnóstico e seu possível tratamento.

A resolução espacial de uma gama câmara pode ser classificada como intrínseca ou extrínseca. A resolução espacial intrínseca está relacionada apenas ao processo de detecção do sistema, sem considerar possíveis influências do colimador utilizado na aquisição da imagem, e pode ser afetada por fatores como: variação estatística na produção de luz de cintilação, a resolução de energia finita devido às variações de composição e/ou na relação da produção de luz com a energia depositada, variações no tubo fotomultiplicador, dentre outros. A resolução espacial extrínseca do sistema depende da resolução intrínseca, da relação entre distância objeto e colimador utilizado para aquisição da imagem e o tipo do colimador utilizado [1].

No teste de resolução espacial planar da gama câmara, previsto em norma CNEN 3.05 [2] com periodicidade mensal, a resolução espacial intrínseca do sistema é estimada em termos da largura a meia altura (FWHM - *Full Width at Half Maximum*) da sua função de espalhamento linear (LSF - Line Spread Function) [3] [4], e possibilita avaliar alterações no sistema de detecção do equipamento através da comparação dos

resultados ao longo do tempo e os valores de referência fornecidos pelo fabricante.

A estimativa da resolução espacial intrínseca do sistema é feita através do método analógico, que é o método predominantemente usado nas clínicas brasileiras de medicina nuclear. O teste consiste na análise visual da imagem gerada pela aquisição de dados utilizando um fantoma de barras com quadrantes de espessuras diferentes. Alguns fatores nessa análise adicionam uma maior camada de subjetividade e variação ao resultado do teste como: percepção visual, tempo de experiência do profissional, e qualidade de imagem e resolução do próprio monitor onde a imagem é visualizada. Todavia, mesmo que um dado hospital tenha um protocolo de reprodutibilidade para o teste, ainda existem os fatores humanos que resultam na variabilidade dos resultados [25].

Sua avaliação incorreta pode acarretar no aumento da imprecisão de avaliação da imagem pelo médico, que pode não conseguir identificar estruturas anatômicas relevantes ao diagnóstico do paciente.

1.1 Objetivo

Este trabalho tem como objetivo geral estudar a viabilidade da implementação de algoritmos de inteligência artificial para a avaliação do teste de resolução espacial intrínseca, realizado através do método analógico. O grande ganho na rotina do profissional é que o modelo seja capaz de gerar um resultado confiável, e que mitigue a variabilidade de resultados causadas por fatores humanos e técnicos. Assim, a intenção deste trabalho é que o modelo seja auxiliar na tomada de decisão, e não protagonista.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 História da Medicina Nuclear

A história da medicina nuclear teve seu início por meados do final do século 19 devido a dois eventos importantes: a descoberta da radioatividade, por Henri Becquerel (1898), e a descoberta do Rádio, por Marie Curie (1895). As duas descobertas deram início ao desenvolvimento de técnicas de imageamento interno do corpo por transmissão. A base teórica da medicina nuclear foi estabelecida entre 1910 e 1945, quando diversos nomes fizeram estudos relativos à circulação intracorpórea de radionuclídeos em plantas e animais [5]. Em 1930, com a invenção do cíclotron, se tornou possível a criação de radionuclídeos artificiais, o que estendeu as possibilidades de processos biológicos que poderiam ser estudados com esses materiais.

Até então, o método de análise na medicina nuclear era a contagem da emissão de radiação em diferentes pontos do corpo. Na década de 1950, foi desenvolvido pelo americano Hal Oscar Anger no Laboratório Donner em Berkeley, Califórnia, a primeira gama câmara, que era capaz de produzir um mapa da biodistribuição de radionuclídeos no corpo do paciente [26]. Contudo, nessa época, seu uso ainda não era comercialmente viável, e só passou a ser quase vinte anos depois, quando a qualidade de imagem do equipamento evoluiu, e sua resolução espacial passou a se equiparar com a de outras alternativas de imageamento à época. Um outro motivo da sua popularização foi o início da fabricação comercial de radiofármacos marcados com o Tc-99m, que até então não eram de comum utilização [5].

Na primeira versão da gama câmara de 1953, a detecção da radiação era feita por uma tela de material cintilador acoplada à um filme de raio x, de modo que os raios detectados no filme eram raios provenientes do processo cintilação da tela. Já em 1958, Anger trocou o sistema de detecção antigo por um baseado em cristal cintilador de iodeto de sódio ativado com Tálio (NaI(Tl)) acoplado à tubos fotomultiplicadores, o que aumentou significativamente a eficiência de detecção, e consequentemente também a resolução espacial do sistema.

2.2 A gama câmara

A gama câmara é um dos sistemas de diagnóstico por imagem por emissão de

maior importância utilizados em clínicas e serviços de Medicina Nuclear. Essa técnica de diagnóstico exige a marcação de fármacos com radionuclídeos emissores de fótons únicos, chamados radiofármacos, que são introduzidos no organismo do paciente. Com isso, é possível obter informações funcionais do organismo através do imageamento da biodistribuição do composto injetado.

O processo de formação da imagem começa com os colimadores, dispositivos de chumbo utilizados com o objetivo de permitir que apenas raios gama de trajetória aproximadamente ortogonal aos detectores contribuam com a obtenção de informação pelo sistema. Esses dispositivos são essenciais para garantir que a imagem gerada possa ser interpretada como uma projeção de duas dimensões do corpo do paciente, evitando distorções e artefatos que diminuam a confiabilidade do diagnóstico médico. A geometria e tamanho dos furos dos colimadores dependem da energia dos raios gama emitidos pelo radiofármaco, e do balanceamento de parâmetros como resolução e sensibilidade da imagem.

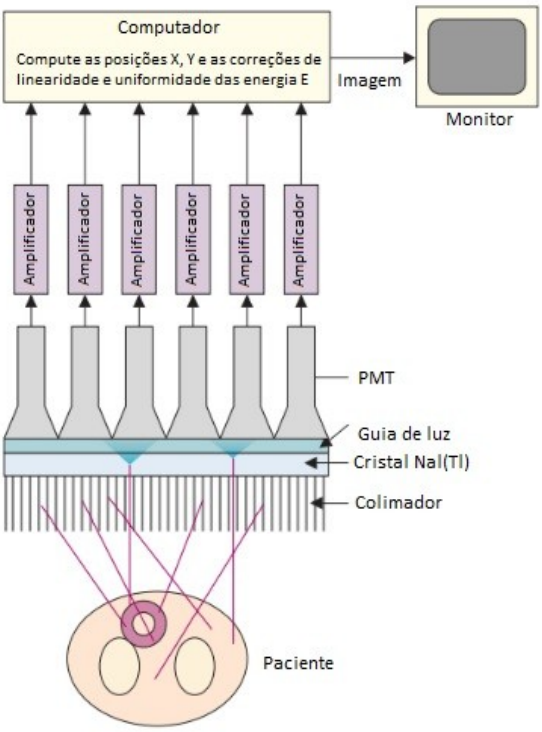

Figura 1: Imagem mostrando o fluxo de informação em um exame de Gama Câmara. Figura adaptada de [5].

O fóton interage com cristal cintilador, gerando luz (visível ou ultravioleta) através do fenômeno chamado de cintilação. Em seguida, o sinal é coletado e amplificado por uma matriz de tubos fotomultiplicadores (PMT - *Photomultiplier Tube*), que gera corrente

elétrica e possibilita a leitura e tratamento pelo conversor analógico-digital (ADC). O sistema então faz a contagem de incidências em todas as regiões de coleta, e gera a imagem de modo que as regiões com maior quantidade de contagens aparentam ter maior densidade. Nessa etapa, também são feitas todas as correções digitais pertinentes, como correção de linearidade e uniformidade. O funcionamento do aparelho descrito acima é demonstrado na Figura 1.

A principal diferença entre a gama câmara de Anger e a de estado sólido, atualmente no mercado, é o sistema de detecção. Na primeira, o cristal utilizado na detecção é monolítico, ao qual os PMTs são acoplados. A posição estimada de incidência de um fóton é determinada através da informação dos sinais coletados pelos PMTs na região onde houve a incidência. O cristal cintilador mais utilizado neste tipo de equipamento é o NaI(Tl), e é utilizado devido a sua ótima eficiência de detecção na faixa dos 100-200 keV, faixa de energia utilizada em procedimentos envolvendo Tc-99m.

Os detectores de estados sólidos são promissores para uso em medicina nuclear, pois fornecem alta eficiência em detecção devido ao alto número atômico dos materiais utilizados como o Telureto de Cádmio e Zinco (CZT), o que aumenta a seção de choque e a probabilidade de interação entre raios gama e o material. Com isso é possível obter detectores compactos, eficientes e granularizados, de modo que, cada detector representa um pixel da imagem, melhorando a resolução espacial do sistema [6]. Além disso, os detectores de estado sólido possuem leitura direta, ou seja, cada fóton incidente é lido como um sinal eletrônico individual, obtendo assim uma excelente resolução em energia capaz de suprimir eventos espalhados e obtendo melhor qualidade na imagem de diagnóstico. Apesar dessas várias vantagens citadas, os sistemas de gama câmara de estados sólido ainda são bastante custosos e não possuem preço competitivo no mercado em relação aos sistemas Anger, sendo mais utilizadas em exames dedicados como os cardíacos [6] [7].

2.3 Qualidade da imagem na gama câmara

O conceito de controle de qualidade de imagem, no contexto de medicina nuclear, se refere a testes, normas e procedimentos com o objetivo de assegurar que o funcionamento e desempenho de equipamentos e procedimentos utilizados na área estejam sempre o mais próximo do ideal, ou ao menos dentro de um intervalo de qualidade aceitável [4].

Dentro dos tipos de testes, existe uma distinção importante a ser feita: testes

intrínsecos e extrínsecos. Os testes intrínsecos visam assegurar que características do próprio equipamento estão em funcionamento adequado. No caso da gama câmara, isso significa que o teste visa avaliar o desempenho dos detectores sem a presença de colimadores. Já os testes extrínsecos se referem diretamente à avaliação de parâmetros do sistema por inteiro, ou seja, com a presença de colimadores [3].

2.4 Parâmetros de desempenho

Os parâmetros de desempenho dos equipamentos de gama câmara são descritos pela *National Electrical Manufacturers Association* (NEMA) no documento *Performance Measurements of Scintillation Cameras* e estão entre eles: resolução em energia, uniformidade (integral e diferencial), linearidade, resolução espacial entre outros [1]. Abaixo, uma breve descrição dos parâmetros de desempenho do sistema mais relevantes para este trabalho.

- **Uniformidade da imagem**

O problema mais notável na imagem em medicina nuclear é a sua não uniformidade. Expondo os detectores a uma fonte de radiação, tanto um *flood* para uma avaliação extrínseca como uma fonte pontual posicionada longe o suficiente para que o FOV seja irradiado uniformemente em uma avaliação intrínseca, é possível observar se em algum ponto do equipamento, a interação cristal-PMT gera um sinal não uniforme.

A uniformidade integral é definida através da diferença entre o valor do pixel mais intenso e o valor do pixel de menor intensidade na imagem e é fornecida em porcentagem pela equação 1 [5].

$$UI(\%) = 100\left(\frac{Máx - Min}{Máx + Min}\right)$$

(1)

A uniformidade diferencial é definida através da maior diferença de intensidade entre pixels vizinhos encontrada na imagem e pode ser estimada pela equação 2 [5].

$$DU(\%) = 100\left(\frac{high - low}{high + low}\right)$$

(2)

A única diferença entre as duas é que a uniformidade integral visa avaliar toda a imagem e a uniformidade diferencial visa avaliar a uniformidade por regiões.

Em sistemas antigos a uniformidade do sistema era alcançada através do uso de guias de luz e tubos fotomultiplicadores com diâmetros grandes. Entretanto, com o avanço de efetivos métodos de correção da uniformidade, as gama câmaras mais contemporâneas podem eliminar os guias de luz e utilizar PMT mais compactas, melhorando significativamente a estimativa da posição de interação do fóton no cristal e consequentemente a resolução intrínseca do sistema.

- **Resolução em energia**

A resolução em energia é o parâmetro que descreve a capacidade mínima de distinção entre dois fótons de energias próximas. É um parâmetro crucial para a qualidade da imagem, pois com ele é possível selecionar com maior precisão eventos com energia de interesse de acordo com o radionuclídeo utilizado, criando uma janela de energia que só contabiliza incidências dentro daqueles valores. A seleção de eventos com energias de interesse permite uma eficiente diminuição do ruído da imagem, e consequentemente aumenta a resolução espacial.

Por exemplo, em uma imagem feita com Tc-99m, onde o gama emitido é de 140 keV, uma janela de 10% de resolução em energia indicaria que qualquer evento com energia entre (140 ± 14) keV é contabilizado na imagem. Esse intervalo é importante, porque além de minimizar a contagem de fótons de menor energia provenientes de espalhamento Compton, também minimiza a contagem de eventos com energia maior que 140 keV, que neste caso, são provenientes de duas incidências com intervalo de tempo menor que a capacidade de resolução temporal do conjunto cristal-PMT, que as lê como uma contagem e soma o valor das energias, esse processo é conhecido como empilhamento de pulsos (*pile-up*).

- **Resolução espacial**

A resolução espacial é um parâmetro fundamental para a análise da qualidade da imagem, pois define a distância mínima em que é possível distinguir dois pontos distintos de captação de radiofármaco em uma imagem. Esse parâmetro é fundamental no meio hospitalar, visto que influencia diretamente na acurácia do laudo médico, e consequentemente na eficácia do diagnóstico e seu possível tratamento.

A resolução espacial de uma gama câmara pode ser classificada como intrínseca ou extrínseca. A resolução espacial intrínseca está relacionada apenas ao processo de detecção do sistema, sem considerar possíveis influências do colimador utilizado na aquisição da imagem, e pode ser afetada por fatores como: variação estatística na produção de luz de cintilação, a resolução de energia finita devido às variações de composição e/ou na relação da produção de luz com a energia depositada, variações no tubo fotomultiplicador, variações na função resposta do detector relacionadas a diferenças nas interfaces ópticas entre cristal e PMT e múltiplos espalhamentos dos fótons dentro do cristal degradando a informação da posição de interação do gama no detector e consequentemente a imagem final. A resolução espacial extrínseca do sistema depende da resolução intrínseca e do colimador utilizado para aquisição da imagem [1].

- **Linearidade**

A linearidade é um parâmetro que se refere a capacidade da gama câmera de gerar imagens que reflitam com precisão a distribuição espacial da radiação no paciente. Na literatura existem dois parâmetros principais de cálculo da linearidade espacial: a linearidade espacial absoluta, que se refere ao desvio máximo de uma barra da imagem em relação a posição teórica, e a linearidade espacial diferencial, que se refere ao desvio padrão médio das distâncias entre as barras na imagem em relação a distância real.

A linearidade espacial permite que sejam geradas imagens mais confiáveis e com maior precisão da localização de estruturas no paciente. No contexto médico, essa característica é fundamental para a correta visualização das estruturas do corpo desejadas, o que aumenta a qualidade e confiabilidade do diagnóstico médico.

2.5 Controle de qualidade

Para a verificação da qualidade de imagem do sistema, a norma CNEN 3.05 [2], que trata de requisitos de segurança e proteção radiológica para serviços de medicina nuclear, exige que alguns testes sejam realizados de acordo com a periodicidade estabelecida. Devido ao escopo desse trabalho, maior importância e descrição será dada ao teste de resolução espacial intrínseca.

Para o controle da uniformidade da imagem, existem alguns testes que são feitos, tendo como parâmetros de diferenciação se são intrínsecos ou extrínsecos, a periodicidade, o radionuclídeo utilizado e a densidade da taxa de contagem esperada no teste. No controle da resolução em energia, os testes são semestrais e visam estipular a menor diferença detectável entre a energia de fótons, sendo feito para fontes monoenergéticas e multi-energéticas, quando aplicável [2].

A linearidade é avaliada tanto em testes intrínsecos, quanto em testes extrínsecos. No caso do segundo tipo de teste, é necessário que o procedimento seja feito em todos os colimadores utilizados na rotina do serviço.

O teste de resolução espacial intrínseca deve ser realizado mensalmente dentro do programa de controle de qualidade dos equipamentos. Segundo o TECDOC 6.02 [3], existem dois métodos de realização do teste: o método analógico e o digital.

O teste analógico é feito com um fantoma de quatro quadrantes de tiras de chumbo de diferentes larguras e distâncias entre eles. A Figura 2 mostra a imagem gerada no teste de resolução espacial planar.

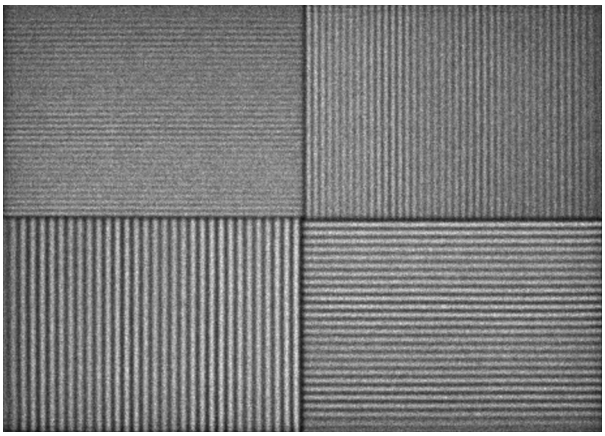

Figura 2: Imagem gerada no teste de controle de qualidade da resolução espacial e linearidade no SMN-HUCFF.

A determinação da resolução é feita através da análise subjetiva de escolha do

quadrante de melhor resolução, onde ainda é possível diferenciar as linhas na imagem. O parâmetro utilizado para avaliar a resolução é o FWHM (*Full Width at Half Maximum*), que indica, em mm, aproximadamente a distância mínima em que se pode distinguir dois pontos com confiabilidade na imagem. Cada quadrante tem uma largura de intervalo-tira definida, e o cálculo da FWHM se dá pela determinação do quadrante de menor valor de largura onde é possível distinguir as linhas com qualidade. Esse quadrante tem largura L, de modo que o cálculo do FWHM se dá por:

$$FWHM = 1.75L \tag{3}$$

Pelo método digital, a análise é feita através de processamento gráfico das intensidades de pixel da imagem coletada através de dois fantomas (cada um para a direção x e y) do padrão NEMA [1], que determina automaticamente os valores de resolução para o campo de visão útil (UFOV - *Useful Field Of View*) e para o campo de visão central (CFOV - *Central Field Of View*), que devem ser utilizados no controle de qualidade do equipamento. Para os dois métodos, o valor da resolução deve estar condizente com os limites estabelecidos pelo fabricante do equipamento.

Outros métodos analíticos foram propostos por [8][9] com o objetivo de apresentar uma técnica mais simples e ágil de determinar o parâmetro, e por isso utiliza o fantoma de barras padrão de quatro regiões.

No primeiro trabalho [8], o método considera que a LSF (*Line spread function* - Função de espalhamento linear) das imagens produzidas se comportam como uma senóide, e determina o FWHM e a resolução espacial a partir dessa curva. O método também propõe que a variância do método estatístico é causada apenas pela modulação espacial introduzida pelo fantoma. Todavia, o artigo considera que existem duas outras fontes de variância na imagem, e que devem ser subtraídas da variação utilizada no método: a variação estatística de contagens por pixel, e a variação devido à não uniformidade espacial. O autor recomendou que o método seja utilizado apenas em fantomas de barras com o FWHM estimado não maior do que 10% da largura da barra. Isso ocorre porque, quando o FWHM estimado é bem maior do que a largura da barra, a variância da LSF devido ao ruído e à não uniformidade passam a ter valores significativos.

Todavia, quando o FWHM é bem menor do que a largura da barra, a aproximação senoidal da LSF passa a não ser mais aceitável.

No segundo artigo [9], o autor faz uma melhoria no método anterior, adicionando quatro novos fatores ao método de estimativa da resolução espacial intrínseca. O primeiro fator de mudança foi fazer uma simulação de cálculo do input de modulação, que no primeiro artigo foi tido como unitário [8]. No entanto, os cálculos demonstraram que a diferença de modulação entre os dois casos era ínfima, e que a aproximação de input unitário é apropriada e minimiza a variância. O segundo fator corrige a presunção de que a LSF pode ser aproximada como uma senóide, para ser modelada como uma onda quadrada. A alteração resultou em melhoria na precisão do MTF (*Modular transfer function*) progressivamente conforme o FWHM se torna menor que a largura das barras do fantoma. O terceiro fator é uma constante que visa corrigir a perda de resolução causada à medida que o tamanho dos pixels aumenta em relação à largura da barra. Esse fator foi de suma importância pro modelo, pois aumentou consideravelmente a consistência dos resultados para diferentes tamanhos de pixel (lê-se tamanho de pixel como mm/pixel). O quarto fator corrige o espalhamento proveniente da estrutura de plástico do fantoma.

O artigo conclui que o modelo simplificado é suficiente para testes de controle de qualidade rotineiros, enquanto o modelo expandido, e a consideração da superestimativa do FWHM devido à estrutura de plástico do fantoma devem ser usados em casos onde maior acurácia é necessária.

2.6 Aquisição de imagem

Existem diversas técnicas e parâmetros que podem alterar a qualidade do resultado final durante a etapa de aquisição da imagem. Alguns exemplos de parâmetros são: modo de aquisição, matriz de aquisição, atividade administrada, densidade de contagens na imagem e zoom.

Os modos de aquisição de imagem em medicina nuclear podem ser definidos pela forma como a variação pelo tempo da biodistribuição de radionuclídeos no corpo do paciente é tratada no pós-processamento da imagem. Por exemplo, em imagens estáticas, considera-se que as variações em biodistribuição são mínimas no intervalo de tempo em que o exame ocorre. Ou seja, na teoria se faz uma aproximação de que não há

variação na biodistribuição, mas, na prática, o que ocorre é que a imagem é uma média da biodistribuição do radionuclídeo durante a coleta dos dados. Em imagens coletadas no modo dinâmico, a imagem é tratada como uma analogia de um vídeo, ou seja: seu objetivo é capturar a redistribuição do radionuclídeo pelo corpo em função do tempo [10]. O modo de aquisição SPECT (*Single Photon Emission Computed Tomography*) fornece uma imagem tridimensional, melhorando a localização de pontos de captação e ruído, uma vez que as imagens dos tecidos de interesse não são sobrepostas.

A matriz de aquisição é um parâmetro que define o tamanho da imagem em pixels. Na prática, quanto menor a quantidade de pixels numa imagem, significa que mais informações de detecção estão sumarizadas dentro daquele pixel. Dessa forma, uma maior resolução permite uma menor sumarização desses dados. Em alguns casos, isso significa que é necessária uma maior atividade de radionuclídeo ou maior tempo de aquisição, já que cada pixel precisa obter contagens suficientes para minimizar artefatos causados por ruído ou flutuação estatística. Também é possível realizar um zoom no campo de visão ao registrar apenas sinais digitais de uma região limitada do campo de visão, restringindo-se a uma região de interesse menor. Esse método tem a vantagem de aumentar significativamente a amostragem da região sem precisar aumentar a atividade empregada.

A densidade de contagem é um parâmetro importante durante a aquisição da imagem. Uma taxa de contagem adequada resulta na diminuição do ruído causado por baixas contagens. Todavia, quando a taxa de contagem é alta em relação ao tempo morto do sistema, a probabilidade de efeitos de empilhamento de pulso se torna relevante, fazendo com que o equipamento registre eventos com energia maior do que a energia real, causando a distorção de imagem e a perda de contagens.

2.7 Redes neurais artificiais

As redes neurais artificiais (RNA) são modelos matemáticos e algoritmos inspirados no modelo de funcionamento do sistema nervoso humano. Assim como no sistema nervoso (SN), os neurônios artificiais são a unidade mínima de sumarização da informação do modelo, enquanto as sinapses são responsáveis pelo armazenamento e transmissão de informação na rede [11].

As redes neurais têm algumas características similares com o SN, dentre elas:

- O processamento da informação ocorre em unidades básicas chamadas de neurônios;
- Os neurônios podem receber e enviar informações para outros neurônios;
- A informação entre os neurônios é transmitida por meio de sinapses;
- A eficiência da sinapse, representada por um valor de peso associado, é análoga à informação guardada no neurônio e, portanto, também na rede;
- O aprendizado da rede se dá basicamente pela adaptação dos valores de peso dos neurônios.

Os algoritmos tradicionais de *machine learning* costumam ter baixo desempenho em tarefas que envolvam dados complexos, como: reconhecimento de objetos e classificação de imagens, processamento de linguagem e geração de texto. Isso ocorre devido à diferença fundamental entre algoritmos de *machine learning* e RNAs. Em *machine learning*, o usuário precisa pré-processar os dados de modo que as características relevantes dos dados estejam claras pro modelo. Já as RNAs se destacam nesses casos, pois conseguem extrair características de bases de dados sem a necessidade do usuário fazer o pré-apontamento de quais características são importantes para o desempenho da tarefa [12].

2.7.1 Neurônio artificial

Um neurônio artificial é essencialmente definido como uma unidade de processamento, pois recebe diferentes entradas de informação, executa uma operação matemática e retorna uma saída. Essa função depende de dois conjuntos de valores intrínsecos à sinapse entre o neurônio que envia e o que recebe a informação, chamados de pesos e *biases.*

$$v = \sum_{i=1}^{j} n_i w_i + b_i$$

(4)

O neurônio faz o produto de cada valor de entrada (n) com o peso (*w*) da sinapse

entre os dois neurônios onde houver a transferência da informação. Em seguida, o valor do bias (*b*) é somado para cada uma das entradas, e todos os valores calculados são somados, gerando um único valor no neurônio, como esquematizado na Figura 3.

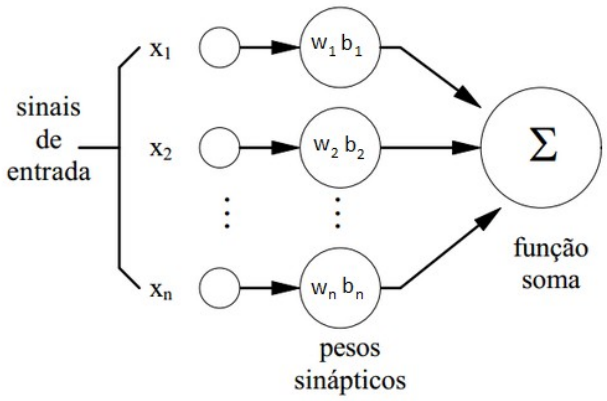

Figura 3: Modelo linear de neurônio. Figura retirada de [13].

2.7.2 Função de ativação

A função de ativação é fundamental para o funcionamento da RNA, e uma das motivações do seu uso é o de limitar os possíveis valores de saída dentro de um intervalo conhecido. O processo de ativação ocorre logo após todos os pesos e valores passados ao neurônio serem processados, e serve para indicar se o neurônio deve ser ativado ou não, de acordo com os valores e pesos recebidos.

Na prática, a função de ativação confere um fator de não-linearidade ao processamento do neurônio, que basicamente, até então, é um modelo linear generalizado. Sem a presença dessa função, o algoritmo passa a ser uma regressão linear, o que torna inviável o aprendizado de características em dados complexos. Por isso, a ativação é uma das propriedades da modelagem que confere um melhor desempenho às RNAs em relação a outros modelos computacionais, em tarefas como detecção de objetos ou modelagens mais complexas.

Existe uma enorme variedade de funções de ativação para redes neurais, sendo que algumas delas costumam ter um uso frequente em camadas específicas da rede.

Abaixo, uma breve descrição das funções de ativação Φ mais utilizadas na literatura [13], tendo *v* como o valor de pré ativação do neurônio previamente definido na equação 5.

- **Função limiar (*heaviside*)**

O valor de saída do neurônio com esta função de ativação é definido por:

$$\phi(v) = \begin{cases} 1, & v \geq 0 \\ 0 & v < 0 \end{cases}$$

(5)

O que significa que o neurônio só é ativado quando a soma do produto n-peso é positiva ou zero.

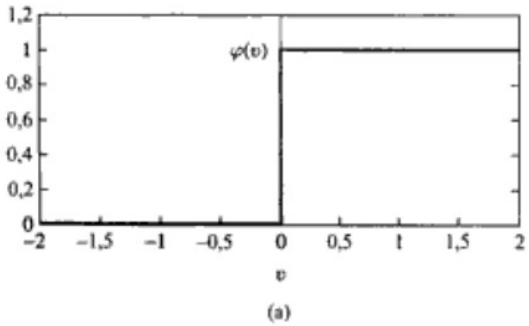

Figura 4: Curva da função de ativação heaviside. Figura retirada de [13]

- **Função sigmóide**

Essa função é amplamente utilizada como função de ativação, porque limita o valor de ativação do neurônio entre 0 e 1 em um gradiente suave. Seu cálculo se dá pela equação 6 e a figura 5 representa sua curva.

$$\phi(v) = \frac{1}{1 + e^{-v}}$$

(6)

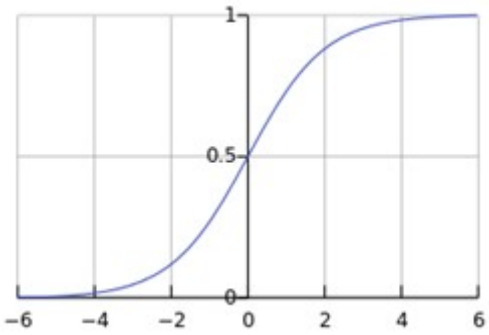

Figura 5: Curva da função de ativação sigmóide. Figura retirada de [13]

- **Função de ativação linear retificada**

Essa função de ativação, também chamada de *relu*, é definida em duas seções:

$$\phi(v) = \begin{cases} v, & v > 0 \\ 0, & v \leq 0 \end{cases}$$

(7)

Ou seja, seu valor é zero quando o valor do neurônio é menor ou igual a zero, e segue uma razão linear caso o valor seja maior que zero.

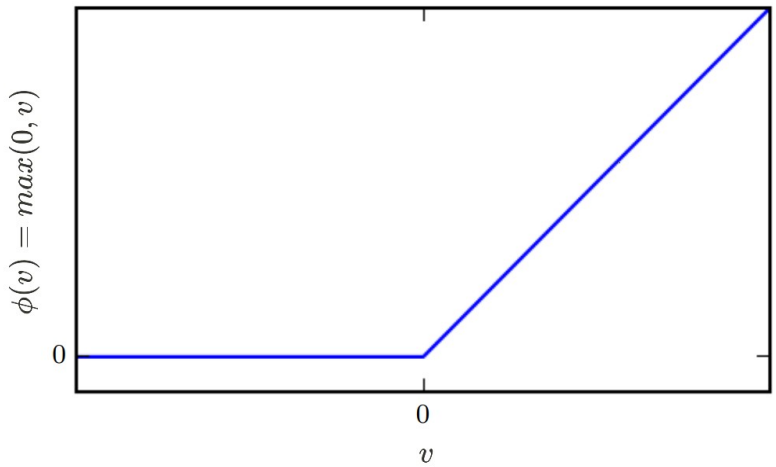

Figura 6: Curva da função de ativação linear retificada. Figura retirada de [14].

2.7.3 Arquiteturas de redes neurais

A arquitetura de uma rede neural diz respeito à disposição e ao modo de interação dos neurônios na rede. Sua conformação tem relação direta com o modelo de algoritmo utilizado, e com o desempenho da rede para resolver um problema específico.

As redes do tipo *feedfoward* são organizadas em três tipos de camadas: camada de entrada, camadas intermediárias e camada de saída. A camada de entrada, ou camada sensorial, é a primeira camada da rede, e costuma ter um neurônio para cada valor de entrada na rede, ou *feature.* As camadas intermediárias são as responsáveis por aprenderem as características e padrões nos dados, a fim de que a rede aprenda a desenvolver a tarefa desejada. Já a camada de saída é a camada que vai trazer o resultado de ativação e processamento de informação da rede. Neste tipo de rede, a informação sempre transita da camada de entrada até a camada de saída. Para o aprendizado desse tipo de arquitetura, são utilizados algoritmos de *backpropagation*, que visam alterar os pesos e viéses internos da rede de modo a minimizar o erro de previsão da rede.

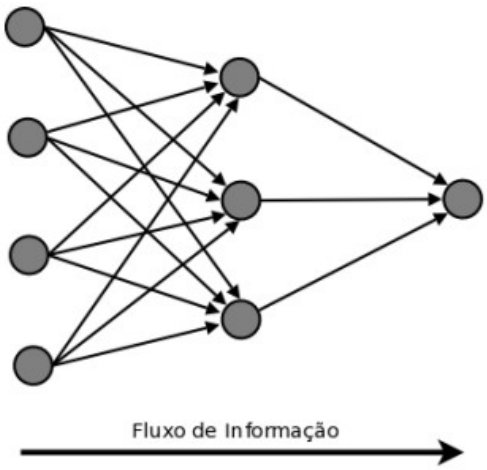

Figura 7: Rede neural do tipo feedfoward.
Figura retirada de [15].

Nas redes recorrentes, ou *feedback networks*, a principal diferença encontrada em relação ao primeiro tipo de rede é que possibilita que a informação também seja transmitida no sentido contrário ao estabelecido nas redes *feedfoward*. No geral, a informação de retorno é um valor de erro com o objetivo de indicar para a rede o quão bom foi seu desempenho. A grande vantagem deste tipo de topologia é que permite que

os pesos e biases da rede sejam modificados de modo a minimizar o erro entre o resultado obtido e o desejado pelo sistema.

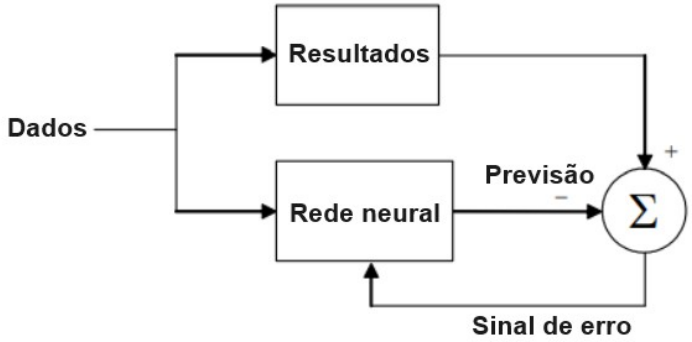

Figura 8: Feedback networks. Figura retirada de [13].

2.7.4 Aprendizagem de redes neurais

A aprendizagem é o processo onde ocorrem os ajustes dos pesos e biases das sinapses, de modo que a rede se torne mais capaz de emitir saídas mais corretas de acordo com cada caso.

O método de aprendizado de rede mais comum no que tange ao escopo desse trabalho é o aprendizado supervisionado. Esse tipo de aprendizado envolve passar os dados pela RNA com o objetivo de que ela seja capaz de aprender padrões e características dos dados. O processo de aprendizagem é iterativo, e faz a comparação entre o desempenho da RNA e o resultado desejado. Depois da comparação, é realizado um processo de *feedback*, onde os pesos e biases da rede são alterados a depender da diferença relativa entre o que se obteve e o que se esperava. Esse processo de aprendizagem da rede também é chamado de treinamento.

Existem dois grandes desafios comuns no treinamento de redes neurais: *overfitting* e o *underfitting*. O *overfitting* ocorre quando a rede falha no processo de generalização da descrição de dados que não foram vistos no treinamento. Esse tipo de problema ocorre muito frequentemente devido a casos onde a complexidade da rede é desnecessariamente alta para o problema apresentado, o que permite que ela aprenda as

características muito específicas de pequenos subconjuntos de imagens, e não características generalistas das classes. O outro problema comum, *underfitting*, ocorre quando a rede não tem bom desempenho em nenhum conjunto de dados. Isso normalmente indica que, ou a rede é simples demais e não possui cruzamentos de neurônios o suficiente para aprender todas as características necessárias, ou a quantidade de dados de treinamento é insuficiente para o problema.

2.7.5 Redes neurais convolucionais

As redes neurais convolucionais (RNC) são um tipo especializado de RNAs. Foram desenvolvidas para o melhor processamento de dados de tipo grade, como por exemplo imagens, que podem ser pensadas como uma matriz de duas dimensões de pixels. A RNC tem esse nome porque pelo menos uma das camadas da rede aplica uma operação linear chamada convolução.

As RNCs são um dos casos de maior sucesso de inteligência artificial inspirada em sistemas biológicos. Sua idealização é muito mais antiga do que sua construção e aplicação prática, e data da década de 1960, quando os neurocientistas David Hubel e Torsten Wiesel desenvolveram uma série de trabalhos que descreviam como é o funcionamento da visão dos mamíferos a nível celular.

Atualmente, as RNCs são reconhecidas como um dos melhores algoritmos classificadores de imagens. O seu desempenho é comparativamente melhor devido a otimização do algoritmo que permite maior volume de dados de treino, e a automatização de etapas de filtro que são feitas manualmente em tipos de redes neurais mais convencionais.

2.7.6 Camada convolucional

Convolução é uma operação matemática que envolve uma função alvo e uma função filtro, e resulta numa terceira função onde sua conformação depende da interação e da função alvo. Matematicamente, a convolução é identificada como operação pelo símbolo asterisco, como na equação 8 [14]:

$$s(t) = (x * w)(t)$$

(8)

No contexto das RNCs, a camada de convolução tem a função de minimizar o peso que cada pixel teria individualmente no resultado da rede. Ao contrário, a convolução permite que a rede olhe para grupos de pixels de tamanho definido por vez, e com isso que observe características naquela região da imagem que não são tão dependentes da sua posição, mas sim do seu formato.

Por conta disso, as camadas convolucionais geralmente são divididas em três etapas: na primeira, a camada realiza uma série de convoluções em paralelo, a fim de produzir as ativações lineares. Na segunda etapa, cada uma das ativações lineares (v definido na equação 5) passa pela ativação de uma função não linear. Geralmente, em RNCs, a função de ativação é a função linear retificada, ou variações dela. Essa etapa, similarmente ao que ocorre em RNAs regulares, define se o neurônio será ativado ou não de acordo com o valor recebido. A terceira etapa é a de *pooling*, a ser descrita na próxima seção. A Figura 9 descreve as etapas das camadas convolucionais.

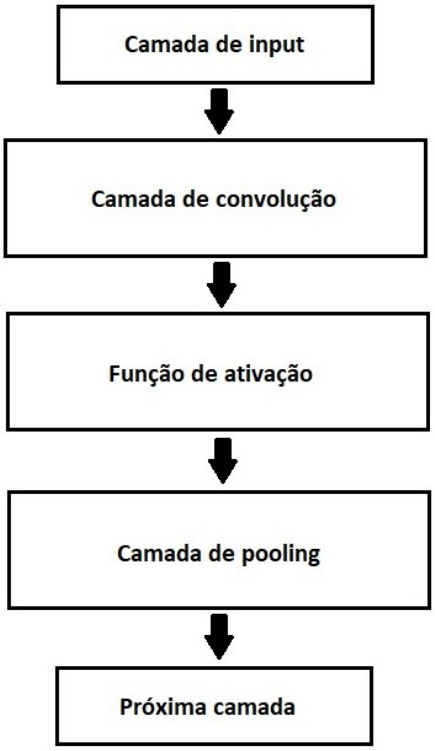

Figura 9: Camada de convolução no processo de aprendizagem.

2.7.7 Camada de *pooling*

A camada de *pooling* tem o objetivo de fazer com que a rede seja tolerante a pequenas variações de posição de características aprendidas pela rede. A camada faz isso através de métodos que diminuem a quantidade de informações provenientes da camada de convolução. Um dos métodos comuns de *pooling*, o *average pooling*, divide a imagem em subconjuntos, e seleciona a média de valores daquele conjunto.

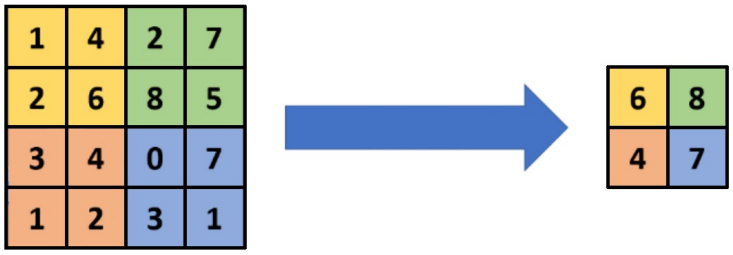

Figura 10: Imagem do processo de pooling. Imagem adaptada de [17].

Essa camada é fundamental no tratamento de imagens, pois permite que a rede, no seu processo de aprendizado, priorize características relacionadas ao formato de objetos, e não à sua posição exata na imagem. Esse tipo de técnica é um dos grandes motivos de sucesso das RNCs, e da sua capacidade de generalização que leva ao aprendizado efetivo.

2.7.8 Camada de *dropout*

A camada de *dropout* é implementada com o objetivo de reduzir o *overfitting*. A técnica exclui aleatoriamente sinapses entre as camadas, de modo que as conexões remanescentes se tornam responsáveis, também, pelo aprendizado de características antes representadas pelos neurônios agora desconectados. Esse processo tende a forçar com que os neurônios da rede sejam menos dependentes de neurônios específicos, o que favorece o processo de aprendizado generalizado da rede para casos ainda não vistos.

3 METODOLOGIA

3.1 Gama câmara

Todas as imagens utilizadas no trabalho foram geradas em duas gama câmaras situadas no setor de medicina nuclear do Hospital Universitário Clementino Fraga Filho (HUCFF). Ambos os equipamentos são do modelo Millennium MG (H300ZL Acquisition System) da fabricante GE Healthcare, e possuem detectores cintiladores de NaI(Tl) de dimensões 370 x 520 x 8,5 mm acoplado a 48 tubos de fotomultiplicadoras, criando um *UFOV* de 36 x 51 cm, capaz de detectar fótons gama com um intervalo de energia de (55 – 540) keV [16].

Os equipamentos foram configurados na posição de exame cardíaco, com um dos detectores posicionados para cima em posição horizontal, e sem a presença de colimadores. O fantoma do teste foi posicionado em cima do detector, e a fonte de Tc-99m utilizada no teste foi posicionada na posição superior do *bipod* utilizado como apoio no teste. Todas as especificações, distâncias e protocolos de teste são definidos no documento TECDOC.

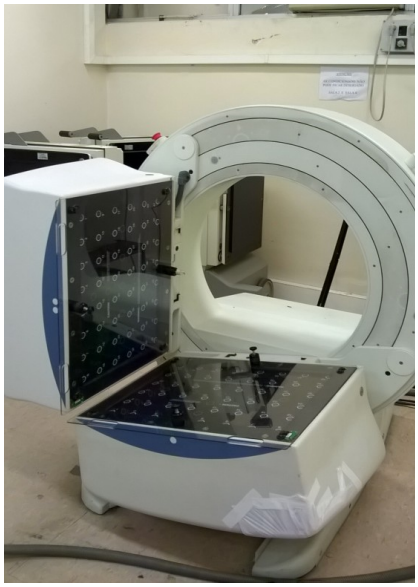

Figura 11: Gama câmara Millennium MG do setor de medicina nuclear do HUCFF sem os colimadores.

3.2 Fantoma de resolução espacial intrínseca e linearidade

O fantoma utilizado para a coleta das imagens foi o de barras de quatro quadrantes. Cada quadrante é composto de um padrão de tiras de chumbo e espaços vazios de mesma espessura L, diferenciando-se uns dos outros pela espessura da tira e do intervalo, que são as mesmas para cada quadrante. Os quadrantes assumem valores de espessura de 2,5mm, 3,0mm, 3,5mm, 4,0mm. No contexto do teste de resolução espacial intrínseca, o quadrante escolhido define a resolução espacial do equipamento.

Figura 12: Fantoma utilizado no teste de resolução espacial intrínseca.

3.3 Protocolo de aquisição de dados do teste de resolução espacial intrínseca

O protocolo de aquisição das imagens é padronizado com a finalidade de intercomparação e análise ao longo do tempo de vida útil da máquina. Para a aquisição, uma fonte de Tecnécio-99m é utilizada com uma atividade que pode variar de 0,5 a 1mCi (18,5MBq a 37MBq) e é localizada no suporte *bipod*, fornecido pelo SMN do HUCFF-UFRJ, apoiado ao detector na posição de 180 graus. Para a aquisição são consideradas as correções de: ganho das fotomultiplicadoras, linearidade espacial, e correção de energia. O pico do espectro de energia é centralizado no 140 keV com uma janela de 20% de variação da energia. O fator de zoom utilizado foi 1, com uma contagem total de 10 milhões de incidências e matriz de 512 x 512 pixels.

A resolução espacial intrínseca é avaliada a partir da FWHM e no HUCFF é determinada a partir do método analógico descrito nos fundamentos teóricos do trabalho.

Segundo a fabricante, os valores de referência de FWHM e FWTM (*Full Width at Tenth Maximum*) para resolução espacial intrínseca do sistema Millennium MG são menores ou iguais a 3,9mm e menores e iguais a 7,6mm, respectivamente [16]. Esses valores são considerados para ambos os campos de visão: UFOV e campo de visão central(CFOV).

3.4 Ambiente de desenvolvimento

O ambiente de desenvolvimento utilizado no trabalho foi a versão gratuita do ambiente online *Google Colab.* Essa versão dispõe de dois processadores *Intel(R) Xeon(R) CPU @ 2.20 GHz, 13 GB de memória RAM, e 33 GB de disco rígido.*

Em relação aos pacotes de desenvolvimento, foi utilizada a linguagem de programação *Python*, na versão *3.10.4*, com os pacotes de análise de dados *Pandas 1.4.3*, *TensorFlow 2.9.0, PyDicom 2.3.0, Pillow 9.5.0 e Numpy 1.24.2.*

O *Pandas* é um pacote Python que possibilita o tratamento dos dados. O *TensorFlow* é o pacote utilizado para o desenho de redes neurais, incluindo seu treinamento e avaliação de desempenho. O *PyDicom* é o pacote responsável por possibilitar a leitura dos metadados do arquivo dicom, que contém todas as informações do exame, inclusive a imagem de teste. O *Pillow* foi o pacote utilizado para algumas padronizações feitas nas imagens, como recorte em tamanho padrão, e limpeza dos canais de cor numa imagem em preto e branco. O *Numpy* foi o pacote utilizado para transformar as imagens em uma matriz numérica, onde cada valor na matriz representa um pixel, e seu módulo representa a intensidade do pixel, variando de 0 a 255.

3.5 Coleta de dados

3.5.1 Aquisição das imagens

A primeira etapa do desenvolvimento do trabalho foi fazer a coleta das imagens dos testes mensais de resolução espacial intrínseca das duas gama câmaras presentes no Serviço de Medicina Nuclear do HUCFF. Todas as imagens disponíveis no período de

junho de 2019 até setembro de 2022 foram coletadas. São obtidas duas imagens por detector, com a finalidade de avaliar o maior número de regiões possíveis do detector com diferentes espessuras e descartar possíveis avarias na imagem que tenham sido causadas pelo fantoma. O procedimento de inversão do fantoma visa também avaliar a estabilidade da resolução de cada quadrante sobre uma parte diferente do equipamento, sendo um ótimo método para identificar possíveis defeitos no equipamento.

Figura 13: Configuração do bipod no momento da coleta da imagem utilizada na avaliação da resolução espacial intrínseca.

3.5.2 Gabarito dos físicos

Para a criação do gabarito, dois físicos(as) certificados em supervisão de proteção radiológica pela CNEN, e que atuam rotineiramente em serviços de medicina nuclear no Brasil foram consultados sobre qual o quadrante de avaliação subjetiva da resolução espacial intrínseca para os 72 pares de imagens disponíveis.

As opções de resolução espacial para cada par de imagens foram definidas como um valor identificador de categoria entre 1 e 4 de qual quadrante melhor define a resolução espacial intrínseca, conforme Figura 14.

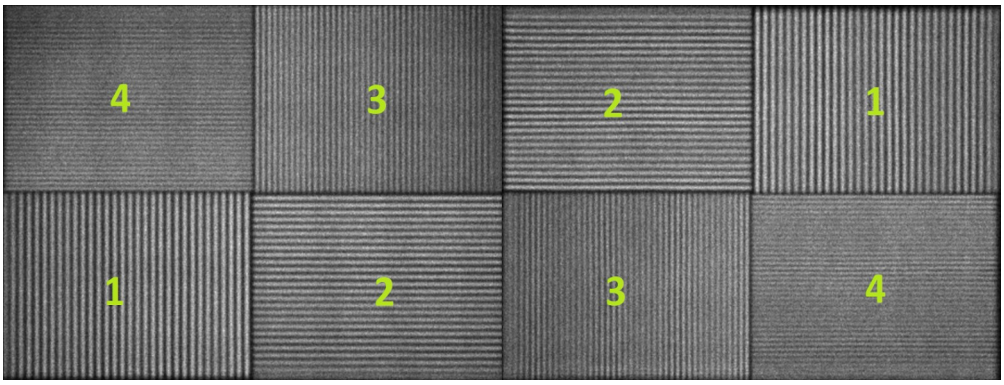

Figura 14: Par de imagens invertidas feitas com o fantoma de barras de quatro quadrantes.

O gabarito de resolução espacial para cada par de imagens foi definido como o valor de resolução mais frequente estimado pelos físicos. No caso de empate entre dois ou mais valores, o escolhido será o quadrante de menor espessura. Os valores escolhidos de gabarito serão transformados numa variável categórica, para serem utilizados como valor de referência no treinamento dos modelos.

3.5.3 Relatórios

Foi feito um levantamento do quadrante de resolução espacial definido em todos os relatórios históricos de avaliação das mesmas imagens utilizadas no levantamento de gabarito dos físicos. O objetivo é que seja feita uma comparação de desempenho entre os relatórios e o desempenho do modelo treinado com o gabarito feito através da consulta aos físicos(as). Na Figura 15 é mostrado o modelo do resultado de um dos relatórios.

Detector	L (mm)	FWHM (mm)
1	3,0	5,25
2	3,0	5,25

Linha de base: Detector 1 - FWHM = 5,25 mm (19/10/2020)
Detector 2 - FWHM = 5,25 mm (19/10/2020)

Situação:

Detector 1: Linearidade e Resolução espacial EM CONFORMIDADE.

Detector 2: Linearidade e Resolução espacial EM CONFORMIDADE;

Figura 15: Relatório do teste de outubro de 2020.

3.5.4 Tratamento e transformação dos dados

As seguintes etapas de tratamento das imagens foram feitas neste trabalho:

- Leitura dos arquivos Dicom
- Organização das imagens
- Verificação de duplicidade
- Transformação das imagens em matrizes numéricas
- Agregação das imagens em um conjunto de dados.

As imagens utilizadas foram coletadas manualmente nos computadores de armazenamento de procedimentos e testes do SMN - HUCFF no formato Dicom (.dcm). Esse é o formato mais comum de armazenamento de exames em equipamentos médicos, porque armazenam, além da imagem, todas as informações e parâmetros de aquisição utilizados no procedimento, conforme mostra a Figura 16.

```
(0008, 0070) Manufacturer                        LO: 'GE MEDICAL SYSTEMS, NUCLEAR'
(0008, 0080) Institution Name                    LO: 'Hospital Clementino Fraga Filho'
(0008, 0090) Referring Physician's Name          PN: ''
(0008, 1010) Station Name                        SH: 'XELERISNEW'
(0008, 1030) Study Description                   LO: 'BARRAS MENSAL'
(0008, 103e) Series Description                  LO: 'ScreenCap01'
(0008, 1060) Name of Physician(s) Reading Study  PN: ''
(0008, 1070) Operators' Name                     PN: ''
(0008, 1090) Manufacturer's Model Name           LO: ''
(0008, 2111) Derivation Description              ST: 'ScreenCap01$$Load to New'
(0009, 0000) Private Creator                     UL: 240
(0009, 0010) Private tag data                    LO: 'GEMS_GENIE_1'
(0009, 1001) [Unknown]                           SH: 'GEMS_GENIE'
(0009, 1010) [Study Name]                        LO: 'BARS'
(0009, 1011) [Study Flags]                       SL: 1
(0009, 1012) [Study Type]                        SL: 1
(0009, 101e) [Dataset UID]                       UI: 1.2.840.113619.2.184.21753.1010106.1659542806.343687500
(0009, 1020) [Series Object Name]                LO: 'ScreenCap01'
(0009, 1021) [Series Flags]                      SL: 1
```

Figura 16: Metadados do arquivo dicom.

Para a leitura dos arquivos Dicom, foi utilizado o pacote *Pydicom*. Primeiro, o arquivo foi carregado ao script. Em seguida, as seguintes informações foram coletadas do arquivo original:

- Recorte correto da imagem original de quatro quadrantes, Figura 17, nas dimensões 458 x 327;
- Coleta da data de aquisição da imagem.

Em seguida, todas as imagens foram salvas no formato de imagem .png, já

separadas por mês, detector e posição. Para impedir a duplicidade de imagens, foi desenvolvida uma função python que compara individualmente todos os pixels entre todos os pares possíveis de imagem. Dessa forma, foi possível evitar que qualquer imagem estivesse duplicada nos dados de treinamento.

Vale ressaltar que a escolha da posição de corte do quadrante foi feita de maneira visual, de forma que todos os recortes foram feitos com as mesmas dimensões, e na mesma posição da imagem que engloba o teste.

Figura 17: Modelo de imagem original disponibilizada nos arquivos dicom

A coleta da informação de detector e da posição de imagens são de suma importância para o trabalho, porque, na confecção do gabarito, consideramos todo par a e b de imagens de um detector e mês tendo um só gabarito. O gabarito dos físicos foi anexado ao conjunto de dados com as suas respectivas imagens, e o conjunto gerado foi utilizado no treinamento da rede neural artificial.

3.6 Rede Neural

Em um fluxo de trabalho com RNAs, os dados são divididos em diversas partições, cada uma com seu propósito definido. A proporção dos dados que cada partição contém, e a metodologia de divisão dos mesmos, dependem da finalidade com que serão utilizadas. A seguir, uma breve explicação de cada uma das divisões, seus percentuais e seus usos no fluxo do trabalho.

No total, neste trabalho temos 144 imagens, sendo 72 pares de imagem e gabarito para desenvolvimento da RNA. Para a etapa de treinamento, 80% das imagens coletadas foram utilizadas, e apenas 20% das imagens para a avaliação do modelo. A divisão dos dados foi feita de forma aleatória, ou seja: as proporções foram previamente definidas, mas as imagens que compõem os subconjuntos são aleatórias e não duplicadas.

O tipo de algoritmo escolhido para o desenvolvimento do modelo foram diversas arquiteturas de redes neurais, envolvendo camadas simples, camadas de pooling, camadas de dropout, e camadas convolucionais, conforme será mostrado na seção de Resultados e Discussão.

As imagens em preto e branco coletadas foram transformadas em matrizes numéricas, onde cada valor na matriz corresponde ao valor de intensidade do pixel, assumindo valores entre 0 (preto) e 255 (branco), respectivamente.

3.6.1 Função de erro

A função de erro é um parâmetro importante para o treinamento da rede neural. Seu objetivo é mensurar a qualidade das predições do modelo para cada imagem, de modo a quantificar em cada rodada de treinamento, o quão errada ou certa a predição estava em relação ao valor real.

Neste trabalho foi utilizada a perda logarítmica binária (ou binary cross-entropy), já que é evidenciada na literatura como uma função de perda apropriada para problemas de classificação binária de imagens [18][19].

Seu cálculo se dá por:

$$CE = \frac{1}{n} \sum_{i=1}^{n} y_i \log(p_i)$$

(9)

onde *p* é a probabilidade dada pelo modelo, e *y* é o valor real (no caso da classificação binária, y=1).

3.6.2 Função de otimização

O papel da função de otimização na etapa de treinamento é minimizar o valor da função de erro, processo que é feito através dos ajustes de pesos e biases do modelo. Desta forma, é entendido que o aprendizado ocorre quando os parâmetros da rede neural são atualizados de modo a diminuir o erro da rede neural definido na seção acima.

A função de otimização utilizada no modelo foi a *adadelta* [20][21], que é uma versão otimizada do algoritmo SGD [22]. Seu cálculo se dá por:

$$w_t = w_{t-1} - \phi_t \nabla CE(w_{t-1})$$

(10)

Onde w é uma unidade de peso ou bias na iteração de treino t, ϕ é a taxa de aprendizagem específica ao peso/bias e à iteração, e o último termo se refere ao gradiente da função de erro descrita na seção acima.

A taxa de aprendizagem é o parâmetro que define o tamanho do ajuste a ser realizado na variável alvo. Neste algoritmo, a taxa de aprendizagem é variável, e seu cálculo é dado por:

$$\phi_t = \frac{\phi_{t-1}}{\sqrt{w_{avg_t} + \epsilon}}$$

(11)

Onde o termo ε é um termo pequeno com o objetivo de impedir o caso onde a raiz quadrada vá a zero, e assume valores muito próximos a zero [20]. O termo w_{avg} é o termo que regula a taxa de aprendizagem durante o treinamento, e seu cálculo se dá por:

$$w_{avg_t} = \gamma w_{avg_{t-1}} + (1-\gamma)(\nabla CE(w_{t-1}))^2$$

(12)

O termo γ na equação acima é chamado de “termo de esquecimento”, e é determinado manualmente (convencionalmente com o valor fixo de 0,95). Em suma, o método implementa as seguintes características.

- Cálculo automático da taxa de aprendizado individualmente para cada dimensão;
- Inserção da janela móvel de média do gradiente no ajuste da taxa de aprendizagem;
- “Esquecimento” da taxa de aprendizagem do início do processo, o que permite uma convergência mais rápida ao mínimo local no final da aprendizagem;

Isso é feito através do ajuste dos pesos e dos biases do modelo, de modo que ao final do treinamento a função de erro tenha atingido pelo menos um de seus mínimos locais.

3.6.3 Abordagem de otimização de resultados

A etapa de otimização de resultados visa melhorar o desempenho do tipo de algoritmo escolhido para o conjunto de dados disponíveis. Isso é feito através da testagem de diferentes arquiteturas de redes neurais convolucionais. Para cada uma das arquiteturas escolhidas, a otimização é feita através da avaliação do modelo com todas as combinações possíveis de parâmetros escolhidos.

3.7 Abordagem de avaliação

O objetivo desta etapa é quantificar o desempenho das redes em diagnosticar corretamente a resolução espacial intrínseca da imagem. Neste trabalho, o valor de resolução espacial intrínseca tido como correto é o gabarito feito através do formulário dos físicos. Como métrica de avaliação de desempenho do modelo, escolhemos a acurácia sob os dados de avaliação, definida como o percentual de acertos do modelo em relação ao gabarito desenvolvido.

Para chegar a um valor mais confiável de acurácia para cada um dos modelos desenvolvidos, a acurácia apresentada no trabalho foi calculada como a média da acurácia obtida em dez execuções consecutivas do mesmo modelo.

Para quantificar a dispersão das dez execuções utilizadas no cálculo da acurácia média, utilizamos o desvio padrão dado pela equação:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n-1}}$$

(13)

Onde μ é a média das dez execuções, e n é o número de modelos executados.

Como medida de comparação, também fizemos a acurácia dos valores colhidos dos relatórios mensais em relação ao mesmo gabarito gerado.

Para estudo do comportamento do modelo, também foram anexadas, nos resultados, a curva característica de receptor para um modelo de cada arquitetura.

3.8 Abordagem de definição do ponto de corte

Os modelos desenvolvidos retornam uma probabilidade como resultado para cada imagem. A classificação de cada previsão do modelo como uma classe ou outra depende de um ponto de corte chamado, de modo que valores maiores que esse valor indicam previsão para uma classe, e valores menores para a outra. Por padrão, esse valor é definido como 0,5.

No entanto, quando temos dados com quantidades desbalanceadas, esse corte em 0,5 pode não representar corretamente a capacidade do modelo em distinguir entre as duas classes. Para isso se utiliza o *Youden index* que é uma metodologia de cálculo do ponto ótimo de corte das probabilidades do modelo [27]. Com ela, é possível definir o *threshold* que apresenta um valor de acurácia que contempla previsões corretas entre as duas classes.

Para cada ponto de corte entre 0 e 1, é calculado o valor da estatística J através da sensibilidade e especificidade associadas a esse ponto de corte, como pode ser visto na

equação 14. Após, o valor de corte que apresenta o maior valor de J é o usado para o cálculo da acurácia.

$$J = \text{sensibilidade} + \text{especificidade} - 1$$

(14)

Sendo a sensibilidade e a especificidade definidas respectivamente por:

$$\text{Sensibilidade} = \frac{\text{Verdadeiro positivo}}{\text{Verdadeiro positivo} + \text{Falso negativo}}$$

(15)

$$\text{Especificidade} = \frac{\text{Verdadeiro negativo}}{\text{Verdadeiro negativo} + \text{Falso positivo}}$$

(16)

4 RESULTADOS E DISCUSSÃO

Na seção de resultados, primeiro serão apresentados os quatros modelos de rede neural implementados no trabalho, contendo sua arquitetura, os resultados obtidos em cada combinação de parâmetros possíveis, e sua respectiva acurácia nos dados de teste. Em seguida, uma comparação dos resultados obtidos nas quatro arquiteturas com a acurácia calculada com *Youden index* e com o ponto de corte padrão de 0,5.

Na apresentação das arquiteturas abaixo, os seguintes parâmetros foram utilizados no cálculo dos resultados:

Tabela 1: Parâmetros otimizados nas arquiteturas de rede neural.

Parâmetro	Descrição
N	Quantidade de neurônios na camada
D	Percentual de sinapses excluídas
P	Tamanho P x P em pixels da janela de pooling
K	Tamanho K x K da função de filtro da camada convolucional.

4.1 Arquitetura 1

A Figura 18 mostra o fluxograma da arquitetura 1.

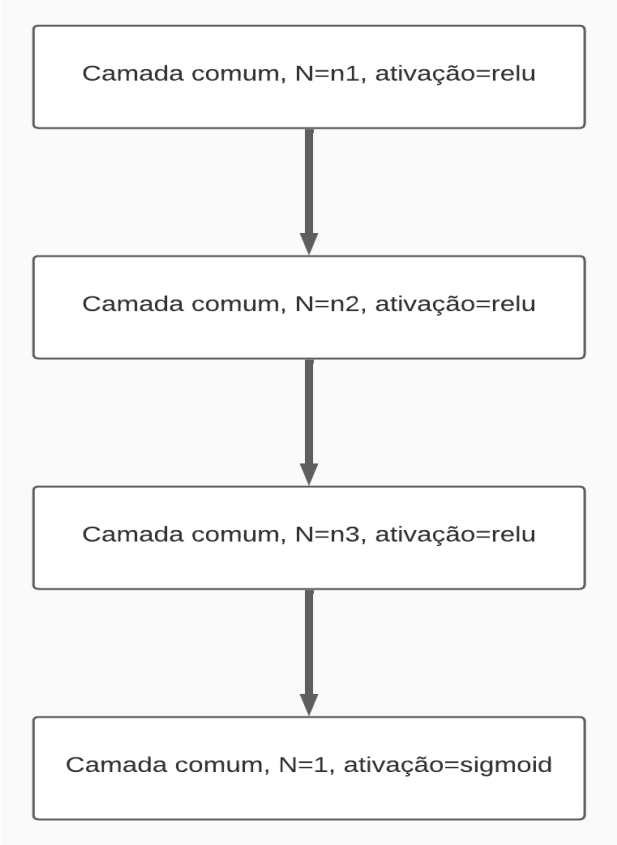

Figura 18: Camadas da arquitetura 1.

Os seguintes resultados de acurácia foram obtidos para os modelos desenvolvidos na arquitetura 1:

Tabela 2: Combinações de parâmetros e seus respectivos resultados após 10 iterações em cada modelo da arquitetura 1.

Variável n1	Variável n2	Variável n3	Acurácia
128	64	32	0,59 ± 0,04
128	64	16	0,58 ± 0,11
128	32	32	0,58 ± 0,05
128	32	16	0,58 ± 0,14
128	16	32	0,58 ± 0,10
128	16	16	0,60 ± 0,07
64	64	32	0,59 ± 0,04
64	64	16	0,58 ± 0,09
64	32	32	0,58 ± 0,06
64	32	16	0,63 ± 0,09
64	16	32	0,61 ± 0,07
64	16	16	0,59 ± 0,07
32	64	32	0,58 ± 0,06
32	64	16	0,60 ± 0,14
32	32	32	0,58 ± 0,08
32	32	16	0,57 ± 0,07
32	16	32	0,59 ± 0,04
32	16	16	0,59 ± 0,15

4.2 Arquitetura 2

A Figura 19 mostra o fluxograma da arquitetura 2.

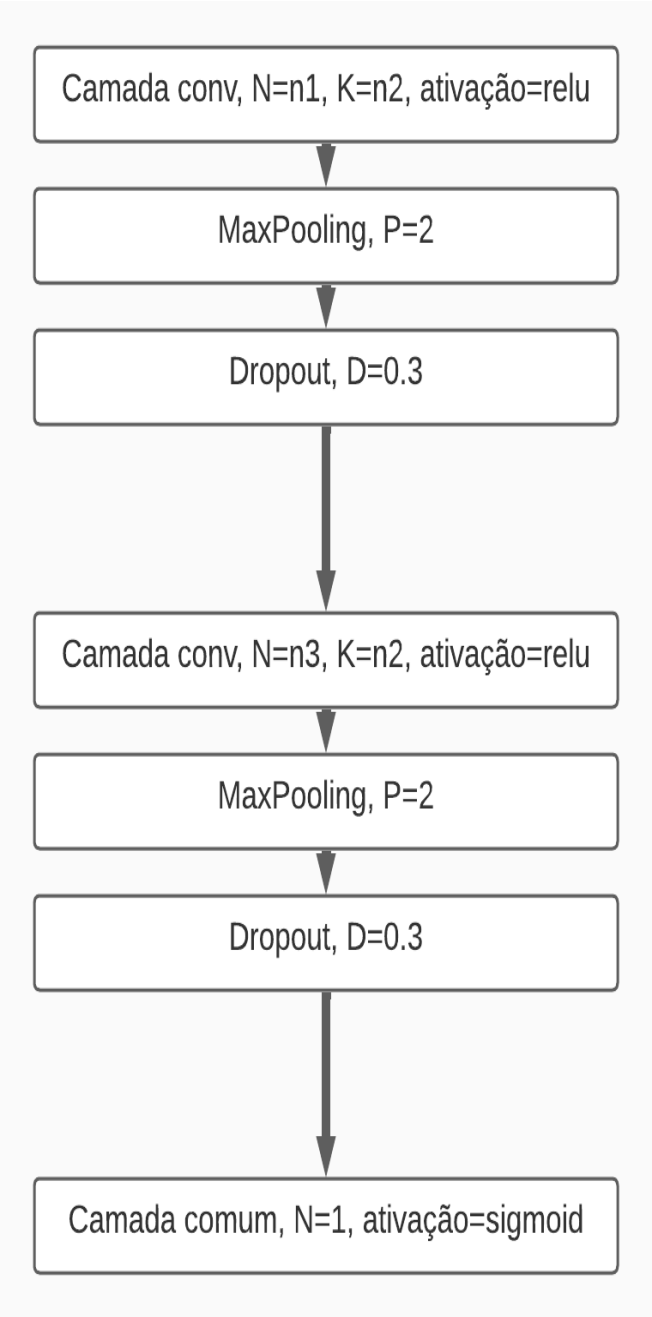

Figura 19: Camadas da arquitetura 2.

Os seguintes resultados de acurácia foram obtidos para os modelos desenvolvidos na arquitetura 2:

Tabela 3: Combinações de parâmetros e seus respectivos resultados após 10 iterações em cada modelo da arquitetura 2.

Variável n1	Variável n2	Variável n3	Acurácia
128	2	32	0,69 ± 0,17
128	2	16	0,64 ± 0,17
128	4	32	0,74 ± 0,12
128	4	16	0,78 ± 0,15
128	8	32	0,86 ± 0,13
128	8	16	0,65 ± 0,13
64	2	32	0,74 ± 0,18
64	2	16	0,76 ± 0,06
64	4	32	0,64 ± 0,12
64	4	16	0.64 ± 0,16
64	8	32	0.56 ± 0,12
64	8	16	0,70 ± 0,16
32	2	32	0,55 ± 0,13
32	2	16	0,58 ± 0,15
32	4	32	0,75 ± 0,11
32	4	16	0,69 ± 0,13
32	8	32	0,69 ± 0,11
32	8	1	0,72 ± 0,15

4.3 Arquitetura 3

A Figura 20 mostra o fluxograma da arquitetura 3.

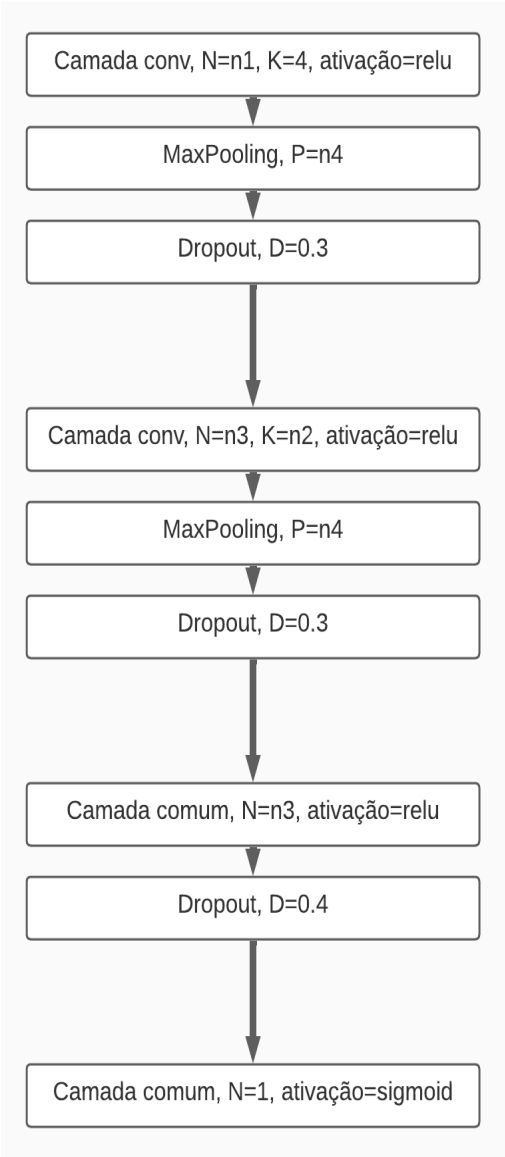

Figura 20: Camadas da arquitetura 3.

Os seguintes resultados de acurácia foram obtidos para os modelos desenvolvidos na arquitetura 3:

Tabela 4: Combinações de parâmetros e seus respectivos resultados após 10 iterações em cada modelo da arquitetura 3.

Variável n1	Variável n2	Variável n3	Variável n4	Acurácia
128	64	32	2	0,67 ± 0,12
128	64	32	4	0,60 ± 0,16
128	64	16	2	0,67 ± 0,13
128	64	16	4	0,28 ± 0,14
128	32	32	2	0,72 ± 0,18
128	32	32	4	0,49 ± 0,14
128	32	16	2	0,78 ± 0,16
128	32	16	4	0,74 ± 0,10
64	64	32	2	0,70 ± 0,18
64	64	32	4	0,58 ± 0,14
64	64	16	2	0,51 ± 0,17
64	64	16	4	0,28 ± 0,15
64	32	32	2	0,73 ± 0,09
64	32	32	4	0,45 ± 0,17
64	32	16	2	0,76 ± 0,13
64	32	16	4	0,58 ± 0,15
32	64	32	2	0,66 ± 0,08
32	64	32	4	0,64 ± 0,14
32	64	16	2	0,62 ± 0,12
32	64	16	4	0,71 ± 0,18
32	32	32	2	0,62 ± 0,17
32	32	32	4	0,55 ± 0,09
32	32	16	2	0,58 ± 0,14
32	32	16	4	0,70 ± 0,11

4.4 Arquitetura 4

A Figura 21 mostra o fluxograma da arquitetura 4.

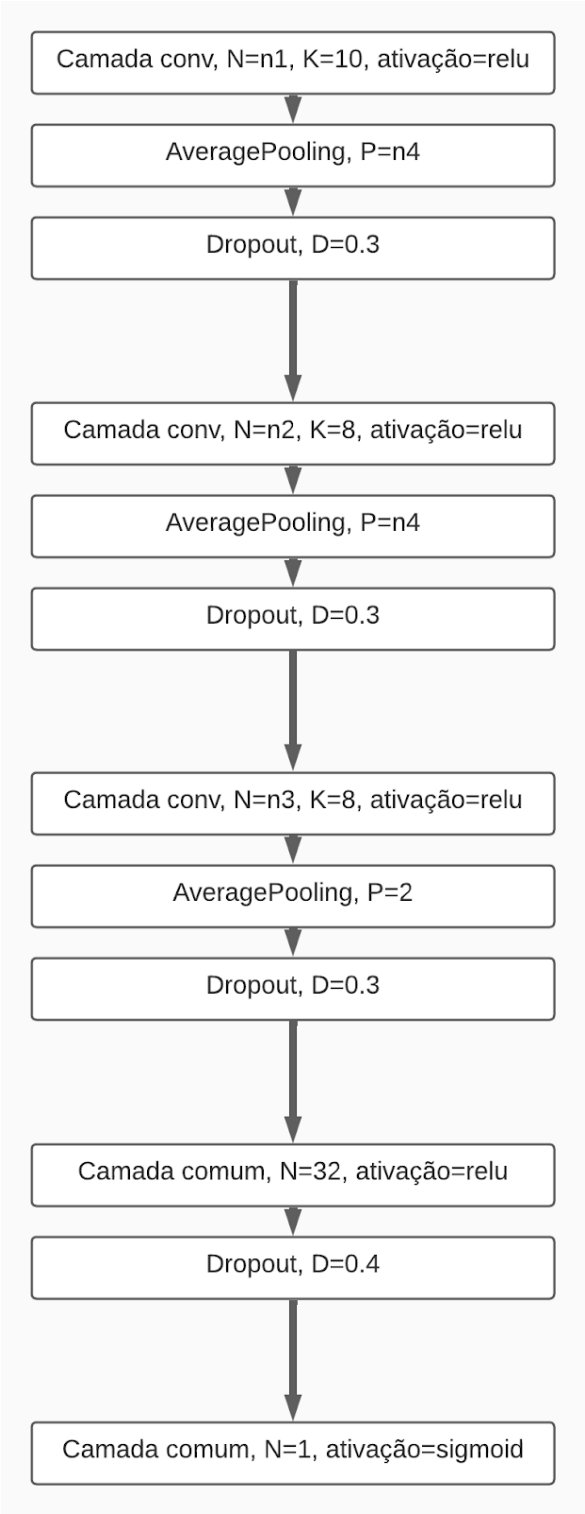

Figura 21: Camadas da arquitetura 4

Os seguintes resultados de acurácia (conforme definida na seção 3.7) foram obtidos para os modelos desenvolvidos na arquitetura 4:

Tabela 5: Combinações de parâmetros e seus respectivos resultados após 10 iterações em cada modelo da arquitetura 1.

Variável n1	Variável n2	Variável n3	Variável n4	Acurácia
128	64	32	2	0,32 ± 0,17
128	64	32	5	0,63 ± 0,18
128	64	16	2	0,27 ± 0,15
128	64	16	5	0,54 ± 0,20
128	32	32	2	0,31 ± 0,12
128	32	32	5	0,40 ± 0,15
128	32	16	2	0,33 ± 0,15
128	32	16	5	0,47 ± 0,14
64	64	32	2	0,50 ± 0,19
64	64	32	5	0,50 ± 0,22
64	64	16	2	0,44 ± 0,12
64	64	16	5	0,47 ± 0,12
64	32	32	2	0,40 ± 0,18
64	32	32	5	0,45 ± 0,18
64	32	16	2	0,66 ± 0,15
64	32	16	5	0,31 ± 0,16

4.5 Acurácia sem *youden*

Os valores de acurácia obtidos sem o ponto de corte de *Youden* apresentaram um desvio padrão de zero para todos os modelos. Esse comportamento ocorreu porque todas as probabilidades do modelo estavam acima de 0,5, fator que indica também a dificuldade de aprendizado em distinguir entre as duas classes.

Abaixo, a distribuição de probabilidades de um modelo da arquitetura 2 e 3:

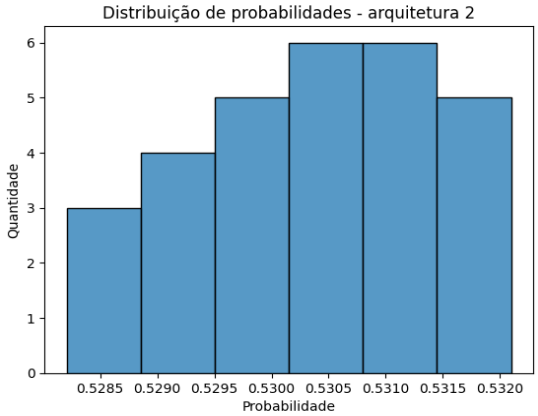

Figura 22: Exemplo de distribuição de probabilidade de modelo da arquitetura 2

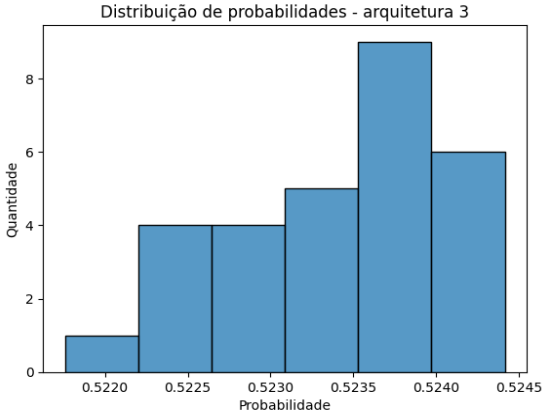

Figura 23: Exemplo de distribuição de probabilidade de modelo da arquitetura 3

4.6 Curva característica de operação do receptor

Abaixo, são apresentadas as curvas *características de operação do receptor (curva roc)*, que mostram o desempenho do modelo com diferentes pontos de corte nas probabilidades:

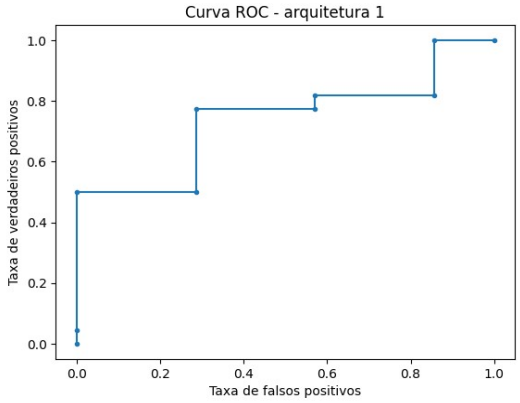

Figura 24: Curva ROC de modelo da arquitetura 1

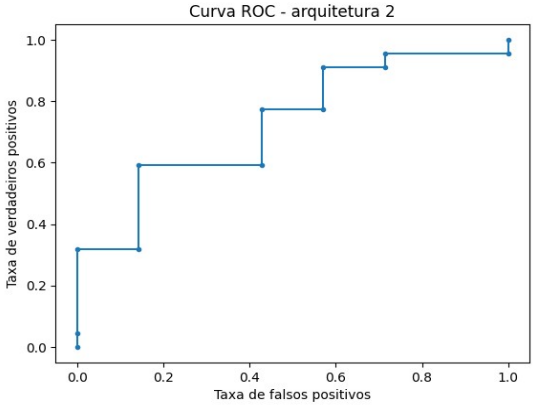

Figura 25: Curva ROC de modelo da arquitetura 2

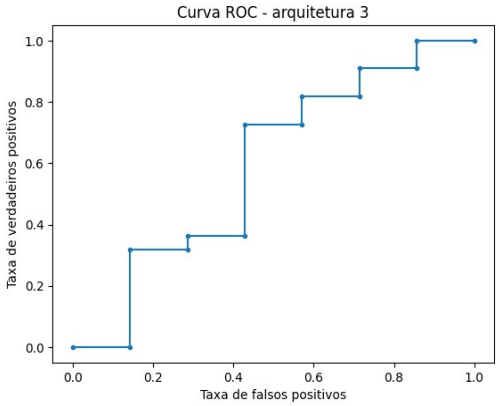

Figura 26: Curva ROC de modelo da arquitetura 3

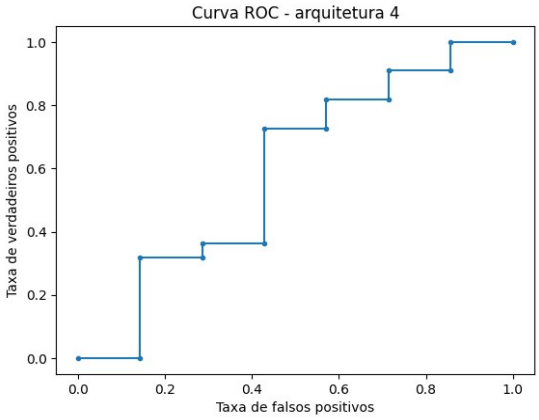

Figura 27: Curva ROC de modelo da arquitetura 4

4.7 Comparativo com os relatórios históricos

Os valores de acurácia obtidos dos relatórios históricos do serviço de medicina nuclear do HUCFF obtiveram uma acurácia de 0,65 em relação aos dados do gabarito. Ou seja, do total de 29 imagens utilizadas para teste, o modelo adivinhou corretamente 25 imagens, enquanto os relatórios acertaram 19 imagens.

5 CONCLUSÃO

A avaliação correta da resolução espacial intrínseca de uma gama câmara é crucial para a confiabilidade e reprodutibilidade dos diagnósticos no dia a dia de um serviço de medicina nuclear. Por conta disso, é fundamental que o físico responsável do setor realize corretamente a rotina do teste de resolução espacial planar, conforme especificado pela CNEN na norma 3.05 e tome alguma ação quando necessário para restabelecer os parâmetros adequados de resolução espacial.

Em relação à qualidade dos dados coletados, só foi possível obter imagens de testes que foram aprovados, e de que a resolução indicada nos relatórios e também no gabarito foram um dos dois quadrantes intermediários de resolução espacial. Com isso, o modelo é incapaz de detectar resoluções diferentes às disponíveis no treinamento, além de também não ser capaz de indicar quando o teste falhou em seu propósito.

Quantitativamente, a baixa disponibilidade de imagens, devido à recorrência mensal do teste, dificulta que o modelo aprenda de forma generalizada a identificar as características da imagem que criam a distinção entre os níveis de resolução.

Os modelos da arquitetura 1 obtiveram acurácias menores, porém mais precisas que os de outras arquiteturas. Isso se deu devido à simplicidade da rede neural, e da falta de camadas que previnam o *overfitting*. As arquiteturas 2, 3 e 4 obtiveram médias de acurácia melhores, porém com valores altos de desvio padrão.

As dez execuções utilizadas para calcular a média de acurácia do modelo foram treinadas com as mesmas imagens, porém com valores iniciais aleatórios distintos de pesos e biases. O processo de otimização, então, aproxima os valores de pesos e biases, inicialmente diferentes, à medida que o treinamento é feito nas imagens. Os altos desvios padrões encontrados evidenciam que os modelos foram incapazes de convergir para um conjunto de acurácias preciso, evidenciando então a baixa confiabilidade de seu uso.

As avaliações obtidas dos relatórios históricos do serviço de medicina nuclear do HUCFF obtiveram uma acurácia de 0,65 em relação aos dados do gabarito. Ou seja, do total de 29 imagens utilizadas para teste, o modelo adivinhou corretamente 25 imagens, enquanto os relatórios acertaram 19 imagens.

A capacidade do ambiente gratuito de computação disponibilizado pelo *google colab* foi um forte limitador da complexidade das redes neurais plausíveis de serem desenvolvidas no trabalho. Em trabalhos futuros, se sugere explorar arquiteturas mais complexas em ambientes computacionais mais potentes.

É sugerido que, em trabalhos futuros, se utilize um volume maior de imagens com o objetivo de melhorar o desempenho e a confiabilidade do modelo. Também é possível treinar o modelo com imagens de diferentes equipamentos já que no modelo desenvolvido o resultado apresentado só é válido para o modelo Millennium MG (H300ZL Acquisition System) da fabricante GE Healthcare.

Portanto, é possível concluir que, apesar da baixa confiabilidade, ou seja, desvios padrões de até 20%, dos modelos para uso prático em serviços de medicina nuclear, a metodologia desenvolvida tem imenso valor como base para adoções posteriores das tecnologias de inteligência artificial no diagnóstico de imagens na física médica no Brasil.

REFERÊNCIAS

[1] NEMA Standards publication; National Electrical Manufacturers Association. NEMA NU 1 -2007 ; Virginia, EUA, 2007.

[2] Comissão Nacional de Energia Nuclear, Requisitos de segurança e proteção radiológica para serviços de medicina nuclear, CNEN 3.05, 2013.

[3] International Atomic Energy Agency - IAEA, Quality control of nuclear medicine instruments, TECDOC 602, 1991.

[4] International Atomic Energy Agency - IAEA, Quality assurance for SPECT Systems, , 2009.

[5] CHERRY, S. R.; SORENSON, J. A.; PHELPS, M. E., Physics in nuclear medicine, 2012.

[6] TAILLEFER, R. ; Scintillation cameras: A new clinical era has come. ; Journal of Nuclear Cardiology, 2021, n.29, p. 1942-1945.

[7] BARBER, H. B. ; Applications of semiconductor detectors to nuclear medicine ; Science Direct, 1999, n.436, p. 102-110.

[8] HANDER, T. A. et al.; Rapid objective measurement of gamma camera resolution using statistical moments. doi: 10.1118/1.597928 ; , 1997.

[9] HANDER, T. A.; et al;; An improved method for rapid objective measurement of gamma camera resolution. doi: 10.1118/1.1328083 ; , 2000.

[10] WERNICK, M. N. e AARSVOL, J. N., Emission Tomography: the fundamentals of PET and SPECT, 2004.

[11] CASTRO, L. N; Fundamentals of natural computing, 2007.

[12] CASTIGLIONI, I. ; AI applications to medical images: From machine learning to deep learning. doi: 10.1016/j.ejmp.2021.02.006; , 2021.

[13] HAYKIN, S ; Redes neurais: princípios e práticas, 2001.

[14] GOODFELLOW, I; BENGIO, Y; COURVILLE, A; Deep learning; MIT Press, 2016.

[15] GONÇALVES, A. R; Redes Neurais Artificiais. Disponível em: https://andreric.github.io/files/pdfs/redes_neurais.pdf

[16] GE HEALTHCARE (United States). Millenium MG: multi-geometry digital cse dual detector gamma camera. Milwaukee: Ge Healthcare, 2006. 10 p.

[17] GHOLAMALINEZHAD, H. KHOSRAVI, H; Pooling Methods in Deep Neural Networks, a Review. doi: 10.48550/arXiv.2009.17485; , 2020.

[18] AL-AZZWI, Z. H. N; NAZAROV,A. N; Brain Tumor Classification based on Improved Stacked Ensemble Deep Learning Methods. doi: 10.31557/APJCP.2023.24.6.2141; , 2023.

[19] DRUZHKOV, P. N; KUSTIKOVA, V. D; A survey of deep learning methods and software tools for image classiifcation and object detection. doi: 10.1134/S1054661816010065; , 2016.

[20] ZEILER, M. D; Adadelta: an adaptive learning rate method. doi: 10.48550/arXiv.1212.5701; , 2012.

[21] YAZAN, E.; TALU, M. F; Comparison of stochastic gradient descent based optimization

techniques. doi: 10.1109/IDAP.2017.8090299; , 2017

[22] AMARI, S; Backpropagation and stochastic gradient descent method. doi: 10.1016/0925-2312(93)90006-O; , 2017.

[23] MCKINNEY, S. M; SIENIEK, M.; GODBOLE, V. et al. International evaluation of an AI system for breast cancer screening. doi: 10.1038/s41586-019-1799-6. Nature 577, 89–94 (2020).

[24] RANGER, N. T; The AAPM/RSNA Physics Tutorial for Residents: radiation detectors in nuclear medicine. Radiographics. [S. l], p. 481-502. mar. 1999.

[25] TRINDEV, P. Test objects and methods for visual assessment of the gamma camera intrinsic resolution during quality control. Medical Physics International Journal. Sofia, Bulgaria, p. 205-210. jan. 2021.

[26] International Atomic Energy Agency - IAEA, Nuclear medicine and the eletronics revolution, 1984.

[27] YOUDEN, W. J; Index for rating diagnostic tests. doi: 10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3 ; , 1950.