

Rafaela Martins Vieira

Classificação de Mamografias Utilizando Redes Neurais Convolucionais

Formiga - MG

2020

Rafaela Martins Vieira

Classificação de Mamografias Utilizando Redes Neurais Convolucionais

Monografia do trabalho de conclusão de curso apresentado ao Instituto Federal Minas Gerais - Campus Formiga, como requisito parcial para a obtenção do título de Bacharel em Ciência da Computação.

Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais

Campus Formiga

Ciência da Computação

Orientador: Wallace de Almeida Rodrigues

Coorientador: Roger Santos Ferreira

Formiga - MG

2020

Vieira, Rafaela Martins
V658c Classificação de Mamografias utilizando Redes Neurais
Convolucionais / Raela Martins Vieira -- Formiga : IFMG, 2020.
49p. : il.

Orientador: Prof. MSc. Wallace de Almeida Rodrigues
Coorientador: Prof. MSc. Roger Santos Ferreira
Trabalho de Conclusão de Curso – Instituto Federal de Educação,
Ciência e Tecnologia de Minas Gerais – *Campus* Formiga.

1. Rede Neural. 2. Convolução. 3. Mamografia.
4. Câncer. 5. Classificação. I. Rodrigues, Wallace de Almeida.
II. Ferreira, Roger Santos. III. Título.

CDD 004

RAFAELA MARTINS VIEIRA

Classificação de Mamografias Utilizando Redes Neurais Convolucionais

Trabalho de Conclusão de Curso apresentado ao Instituto Federal de Minas Gerais-Campus Formiga, como Requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Aprovado em: 26 de outubro de 2020.

BANCA EXAMINADORA

Prof. WALACE DE ALMEIDA RODRIGUES (Orientador)

Prof. ROGER FERREIRA SANTOS (Coorientador)

Prof. FERNANDO PAIM LIMA (Membro1)

Prof. MÁRIO LUIZ RODRIGUES OLIVEIRA (Membro2)



Documento assinado eletronicamente por **Wallace de Almeida Rodrigues, Professor**, em 27/10/2020, às 15:05, conforme art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por **Mario Luiz Rodrigues Oliveira, Professor**, em 27/10/2020, às 15:05, conforme art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por **Roger Santos Ferreira, Professor**, em 27/10/2020, às 15:36, conforme art. 1º, III, "b", da Lei 11.419/2006.



A autenticidade do documento pode ser conferida no site <https://sei.ifmg.edu.br/consultadocs> informando o código verificador **0670691** e o código CRC **41B7B772**.

Quero dedicar esta monografia primeiramente a Deus, maior orientador da minha vida, a minha mãe Patrícia Martins, por sempre manter-me forte na luta apesar dos obstáculos, e ao meu coordenador Wallace de Almeida Rodrigues, cuja dedicação e paciência serviram como pilares de sustentação para a conclusão deste trabalho. Também quero agradecer e dedicar o presente trabalho ao professor Roger Santos Ferreira por ser uma constante fonte de motivação e incentivo ao longo de todo o projeto. Grata por tudo.

Agradecimentos

Agradeço primeiro a Deus por ter me mantido na trilha certa durante este projeto de pesquisa com saúde e forças para chegar até o final.

Sou grata à minha família pelo apoio que sempre me deram durante toda a minha vida, em especial à minha mãe Patrícia Martins que nunca me deixou desistir de meus sonhos, ao meu pai Anderson Paulo Viera por todo o esforço investido na minha educação e ao meu avô Rafael Martins pelo apoio incondicional.

Deixo um agradecimento especial ao meu orientador Wallace de Almeida Rodrigues pelo incentivo e pela dedicação do seu escasso tempo ao meu projeto de pesquisa e ao coorientador deste projeto, Roger Santos Ferreira, que sempre me ajudou com sua vasta experiência desde o início do curso. Eles sempre me incentivaram e acreditaram que eu seria capaz de superar os obstáculos. A todos os meus amigos do curso de graduação que compartilharam dos inúmeros desafios que enfrentamos, sempre com o espírito colaborativo.

Também quero agradecer o Instituto Federal de Minas Gerais Campus Formiga e a todos os professores do meu curso pela elevada qualidade do ensino oferecido.

*“Queres ser grande? Começa por ser humilde.”
(Sto. Agostinho de Hipona)*

Resumo

O aumento no surgimento de novos casos de câncer de mama tem repercutido como um dos graves problemas de saúde pública no mundo. A mamografia, comparada aos demais exames disponíveis, é a forma mais eficaz para a detecção de tumores em fase inicial e permite detectar tumores a partir de 1 mm, enquanto no auto-exame só é possível perceber tumores a partir de 1,5 cm. Apesar das vantagens apresentadas, falhas de diagnóstico podem ainda existir. Como consequência da má interpretação dos exames de mamografia tem-se um aumento do número de realização de biópsias desnecessárias e de tratamentos inadequados. Visando utilizar a tecnologia computacional para auxiliar radiologistas na leitura de imagens e diagnóstico precoce do câncer de mama, o presente trabalho investiga o uso de rede neural convolucional na detecção de câncer em mamografias. Uma arquitetura de rede neural convolucional foi implementada e alguns fatores de impacto foram investigados. Os resultados obtidos da experimentação foram analisados e um modelo linear de regressão é apresentado junto com a melhor configuração para a rede dentre os fatores analisados utilizando a acurácia como métrica.

Palavras-chave: Mamografia, Rede Neural, Câncer.

Lista de ilustrações

Figura 1 – Mamografia obtida do mini MIAS identificando elementos da imagem .	16
Figura 2 – Reconhecimento de imagens	17
Figura 3 – Modelo não-linear de um neurônio artificial perceptron	19
Figura 4 – Transformação linear produzida pela presença do Bias	20
Figura 5 – Rede neural em camadas	23
Figura 6 – Palavras relevantes que descrevem a cena em uma <i>ConNet</i>	25
Figura 7 – Reconhecimentos feitos por uma <i>ConvNet</i>	25
Figura 8 – Arquitetura de uma <i>LeNet</i>	26
Figura 9 – Exemplo de operação de convolução.	27
Figura 10 – Exemplo de operação <i>max-pooling</i>	28
Figura 11 – Operação RNA com camada totalmente conectada.	29
Figura 12 – Keras <i>homepage</i>	31
Figura 13 – Tensorflow <i>homepage</i>	32
Figura 14 – Arquitetura da rede neural construída.	34

Lista de abreviaturas e siglas

ANN	Redes Neurais Artificiais.
CNN	Redes Neurais Convolucionais.
MLP	Redes Perceptron Multicamadas.
LSTM	Long Short Term Memory.
RNN	Redes Neurais Recorrentes.

Sumário

1	INTRODUÇÃO	12
1.1	Justificativa	13
1.2	Objetivos	14
1.3	Estrutura do trabalho	14
2	FUNDAMENTAÇÃO TEÓRICA	15
2.1	Exame de mamografias	15
2.2	Reconhecimento de imagens	17
2.3	Redes neurais	18
2.3.1	Neurônios biológicos	18
2.3.2	Neurônios artificiais	19
2.3.3	Modelo multicamada de rede neural	22
2.3.4	Reconhecimento de imagem por rede neural	23
2.4	Redes neurais convolucionais	24
2.4.1	Modelo <i>LeNet</i> de rede neural convolucional	26
2.4.2	Camadas de Convolução	26
2.4.3	Operação de <i>Pooling</i>	27
2.4.4	Função ReLU	28
2.4.5	Camada totalmente conectada	29
2.5	Aplicação das CNN no reconhecimento de câncer de mama	29
3	TECNOLOGIAS E MÉTODOS	30
3.1	Metodologia	30
3.2	Python e Bash	30
3.3	Biblioteca Keras	30
3.4	Biblioteca Tensorflow	31
3.5	Banco de dados MiniMammographic MIAS	31
3.6	Seleção do trabalho de Zhou como ponto de partida	32
4	DESENVOLVIMENTO	33
4.1	Projeto da CNN implementada	33
4.1.1	Avaliação dos resultados obtidos pela CNN implementada	33
4.2	Fatores aplicados na CNN implementada	35
4.2.1	Execução dos experimentos	36
4.2.2	Avaliação da Solução	41

5	CONCLUSÃO	44
6	TRABALHOS FUTUROS	45
	REFERÊNCIAS	46

1 Introdução

O aumento no surgimento de novos casos de câncer de mama tem repercutido como um dos graves problemas de saúde pública no mundo. A Organização Mundial da Saúde (OMS) prevê para o ano de 2030 que ocorrerá uma média de 27 milhões de novos casos de câncer de mama no mundo, o que equivale a 17 milhões de mortes (LEITE, 2011). Considerando o Brasil no ano de 2010, cerca de 12.705 mulheres tiveram mortes ocasionadas pelo câncer de mama, sendo que esse tipo de câncer é uma das principais causas de mortes em mulheres com idade entre 39 e 58 anos (INCA, 2012). Já em 2018 eram esperados aproximadamente 2,1 milhões de novos diagnósticos de câncer de mama, sendo estimadas mais de 627 mil mortes em 2018, representando 6,6% do total de mortes por todos os tipos da doença. A pesquisa ainda aponta que a uma a cada quatro mulheres que possui um caso de diagnóstico de câncer têm câncer de mama, representando assim 24,2% do total (FEMAMA, 2019).

O combate ao câncer de mama se dá por meio da realização periódica do auto-exame das mamas, do exame clínico realizado pelo profissional de saúde e de exames por imagens, como a mamografia, pois com a descoberta precoce existe a possibilidade da realização de tratamentos que aumentam as chances de cura do paciente (SOUTO, 2014). Segundo SOUTO (2014) “países em desenvolvimento, como o Brasil, são os mais afetados com essa doença devido à falta de informação e precarização de parte dos serviços de saúde ofertados à população”.

A mamografia, comparada aos demais exames disponíveis, é a forma mais eficaz para a detecção de tumores em fase inicial e permite detectar tumores a partir de 1 mm, enquanto no auto-exame só é possível perceber tumores a partir de 1,5 cm (LEITE, 2011). Apesar das vantagens apresentadas, falhas de diagnóstico podem ainda existir e são geralmente causadas por limitações ou falhas humanas que incluem cansaço, indisposição e outros fatores pessoais envolvendo os especialistas. Estima-se que a taxa de acerto dos radiologistas no rastreamento do câncer em mamografias varia entre 65% e 75% de sucesso (SKAANE, 1997). Como consequência da má interpretação dos exames de mamografia e da insegurança no diagnóstico, tem-se um aumento do número de realização de biópsias desnecessárias e de tratamentos inadequados.

Uma maneira de mitigar o problema é utilizar a tecnologia computacional para auxiliar radiologistas na leitura de imagens e diagnóstico precoce do câncer de mama. A computação apresenta abrangência nessa área com utilização de sistemas de Diagnóstico Auxiliado por Computador (CAD - *Computer-Aided Diagnosis*). Esses sistemas não têm como finalidade substituir os especialistas da área, mas sim auxiliar os mesmos a fim

de diminuir a incerteza no diagnóstico fornecendo uma segunda opinião ao caso. Muitos sistemas CAD utilizam redes neurais artificiais para classificação dos tumores em imagens, tais como os sistemas apresentados nos trabalhos de [Komeda et al. \(2017\)](#) e [Al-antari et al. \(2018\)](#).

O reconhecimento de imagem é um clássico problema de classificação e as redes neurais convolucionais (CNN - *Covolutional Neural Network*) possuem um histórico de alta acurácia para esse problema. As redes neurais convolucionais tiveram um enorme aumento de popularidade depois que uma rede neural convolucional chamada AlexNet competiu no desafio proposto pelo ImageNet e apresentou um desempenho expressivo no reconhecimento de imagens ([DRAELOS, 2019](#)). Ao longo dos últimos anos muitos trabalhos acadêmicos foram desenvolvidos no intuito de avaliar o desempenho das redes neurais convolucionais no reconhecimento de câncer de mama em imagens médicas, dentre os quais: [Zhang \(2017\)](#), [Alom et al. \(2019\)](#) e [Zhou \(2016\)](#). Essas redes vêm apresentando um desempenho excelente em descrever imagens naturais, realizar reconhecimento facial e analisar imagens médicas (incluindo radiografias de tórax, fotos lesões cutâneas e lâminas histopatológicas) ([DRAELOS, 2019](#)).

Visto que os computadores podem atuar como uma poderosa ferramenta de auxílio na detecção de câncer de mamas e que as redes neurais convolucionais estão apresentando desempenho excelente na análise de imagens médicas, o presente trabalho propõe investigar a eficiência de uma rede neural convolucional na detecção de câncer em mamografias, e por meio de análise e experimentação obter resultados e apontar fatores que possam impactar na acurácia da rede neural convolucional em questão. O modelo de rede neural convolucional selecionado para este trabalho teve como *baseline* o trabalho de [Zhou \(2016\)](#).

1.1 Justificativa

A mamografia é vista como o melhor método de detecção precoce do câncer de mama, no entanto, cerca de 10% a 30% das anomalias mamárias não são identificadas devido às limitações dos próprios especialistas ([SOUTO, 2014](#)). Além disso, não é comum pacientes obterem um diagnóstico rápido sobre um tumor de mama exclusivamente por meio de um único exame ou consulta médica. Geralmente é necessário a realização de uma biopsia antes da confirmação do diagnóstico pelo médico e a espera pelos resultados de laboratório prolonga o tempo exigido para o diagnóstico, o que pode aumentar o risco do câncer deixar efeitos irreversíveis sobre a saúde do paciente ([ZHOU, 2016](#)). Outro fator negativo é o transtorno emocional causado ao paciente devido à demora por uma resposta. Em razão desses problemas, ao longo dos anos foram desenvolvidos métodos alternativos de apoio ao diagnóstico, sendo um deles a utilização de sistemas CAD, uma vez que esses sistemas ampliam em até 15% a precisão no rastreamento de anomalias

em mamografia (CALAS; GUTFILEN; PEREIRA, 2012). Alguns sistemas CAD utilizam técnicas de redes neurais artificiais como as CNNs, que vêm provando sua eficiência em diversas tarefas de reconhecimento de imagens (ARAÚJO et al., 2017). Dado esse contexto, é interessante investigar a eficiência das redes neurais convolucionais na detecção de câncer de mama, o que torna esse trabalho relevante para a sociedade.

1.2 Objetivos

O objetivo geral deste trabalho é investigar o uso de redes neurais convolucionais para o reconhecimento de câncer em mamografias.

Dado o objetivo geral, os objetivos específicos estabelecidos são os seguintes:

- Determinar o modelo de rede neural convolucional a ser utilizado nos experimentos em exames de mamografia.
- Definir a estratégia de análise experimental para avaliar a rede implementada.
- Quantificar a eficácia da rede implementada usando como métrica a acurácia.

Como produto deste trabalho espera-se também oferecer um conjunto de métodos e técnicas para maximizar o resultado final da classificação, e um comparativo de resultados entre os modelos de redes neurais convolucionais avaliadas.

1.3 Estrutura do trabalho

Este trabalho encontra-se dividido da seguinte maneira: no Capítulo 2 são apresentados os principais fundamentos necessários para entender as técnicas utilizadas e conceitos relacionados. No Capítulo 3 são apresentados a metodologia utilizada para desenvolvimento do trabalho e como foi conduzido cada passo. No Capítulo 5 serão abordados os resultados que foram obtidos nos experimentos realizados com os modelos de CNN estudados, bem como também são apresentadas as considerações finais e os trabalhos futuros sugeridos.

2 Fundamentação teórica

Neste capítulo são mostrados os principais fundamentos necessários para entender os conceitos abordados no presente trabalho.

2.1 Exame de mamografias

A mamografia é um tipo de raio-X ¹ produzido com um aparelho chamado mamógrafo, que comprime a mama e gera imagens de alta qualidade capazes de revelar a existência de sinais precoces do câncer de mama. Atualmente a mamografia é o meio mais eficaz de diagnóstico para a detecção de câncer de mama.

A mamografia de rotina, ou seja, de rastreamento, é a maneira mais efetiva de detectar qualquer lesão mamária precocemente, antes mesmo que o paciente ou o médico possam notá-las ou palpá-las, e estima-se cerca de dois anos antes da lesão ser palpável (BARRETOS, 2016). Quanto mais precoce a remoção do tumor em seu período inicial, maior a probabilidade de cura, sendo capaz de transpor os 95%, além de mostrar melhorias significativas na sobrevivência da paciente (BARRETOS, 2016).

Apesar dessas vantagens, a mamografia pode exibir falso negativo em pelo menos 10% dos casos, particularmente devido à presença de tecido denso escondendo o câncer sob a aparência dos tecidos normais. A realização do exame também pode causar desconforto, pois consiste na compressão da mama de maneira gradual e lenta enquanto os feixes de raios X são incididos sobre ela. Esse processo é importante para que o tecido da glândula mamária seja adequadamente espalhado e eventuais nódulos e microcalcificações sejam caracterizados. A duração de todo o exame é de aproximadamente 30 minutos (BARRETOS, 2016).

O resultado do exame de mamografia é uma imagem em tons de cinza do tecido mamário que é interpretada por um radiologista. Essa interpretação é um tanto quanto complexa e exige grande habilidade e experiência profissional, uma vez que alguns casos de câncer de mama apresentam alterações difíceis de serem percebidas. Além disso, o aspecto da mama na mamografia varia de mulher para mulher, por isso é importante que os especialistas tenham em mãos imagens do exame anterior no objetivo de comparar ambas, o que os auxilia na detecção de câncer de mama (SOUZA, 2018).

A Figura 1 exemplifica como a imagem digital de uma mamografia aparece para o médico que realiza o diagnóstico. Essa figura também serve para ilustrar a necessidade da

¹ raio-X é um exame de diagnóstico por imagem que usa radiação ionizante para produzir imagens internas de diversas partes do corpo.

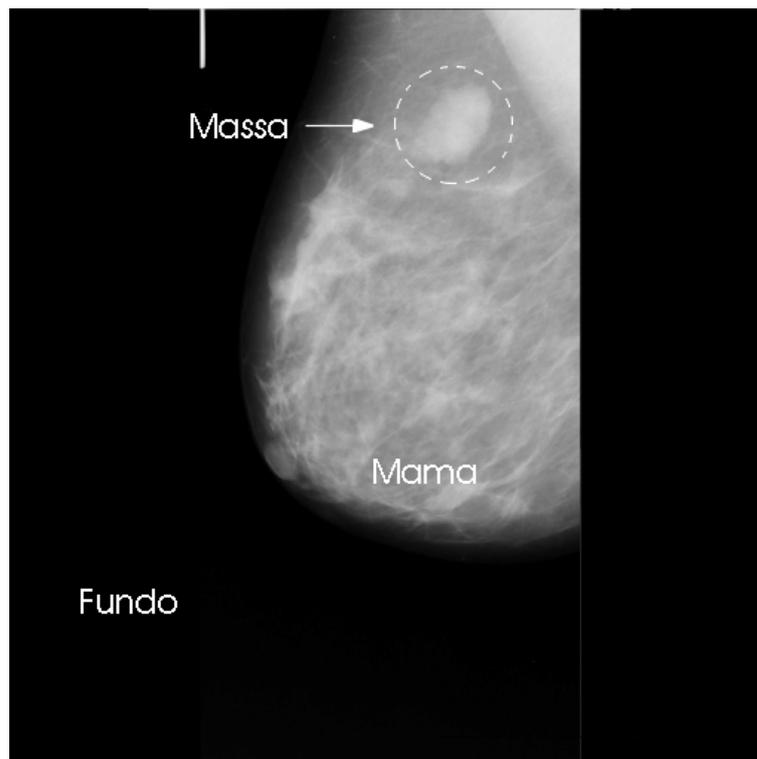


Figura 1 – Mamografia obtida do mini MIAS identificando elementos da imagem
Fonte: (SOUZA, 2018)

imagem digital da mamografia ser corretamente interpretada para se obter um diagnóstico correto.

Os tipos de anomalias mais frequentemente encontrados nos exames de mamografia são as calcificações, as massas e as distorções de arquitetura (linhas finas ou espiculadas que se irradiam a partir de um ponto da mama). Os tumores ou massas aparecem nas mamografias como regiões densas de tamanhos e formatos diferentes que podem ser causados por condições benignas ou malignas (SOUZA, 2018).

As massas podem ser classificadas de acordo com sua borda: micro-lobuladas, escuras, mal-definidas, circunscritas e espiculadas. As características mais determinantes para definir se as massas são malignas são: disposição das bordas, tamanho e forma. Isso mostra a necessidade de técnicas para auxiliar especialistas a diferenciar padrões malignos e benignos baseados em forma.

Fruto de constantes avanços tecnológicos da área da medicina, a mamografia digital é semelhante a mamografia convencional, uma vez que ambas utilizam raios X na produção das imagens, porém a mamografia digital é produzida com computadores e detectores digitais ao invés de filmadoras de cassete. Na mamografia digital, a imagem mamográfica é capturada por um detector eletrônico especial de raios X que converte a imagem numa foto digital que pode ser revisada no monitor do computador (BARRETOS, 2016). Desse modo, através do computador, o radiologista pode alterar as propriedades da imagem,

como: magnificação, orientação, brilho e contraste, afim de verificar melhor as áreas da mama.

Com as vantagens obtidas pelos avanços tecnológicos, os exames também tornaram-se mais rápidos, fazendo com que os pacientes recebam menos doses de raios X com maior qualidade diagnostica e menor número de repetições durante o exame (BARRETOS, 2016).

2.2 Reconhecimento de imagens

A capacidade de reconhecer imagens é algo natural para humanos e animais, mas os computadores têm certa dificuldade de realizar a tarefa de reconhecimento (RAVINDRA, 2017).

Para que um computador seja capaz de coletar e reconhecer imagens, eles geralmente são equipados com uma câmera e empregam tecnologias de visão de máquina e sistemas de software de inteligência artificial. Sob o ponto de vista computacional, o reconhecimento de imagem pode ser definido como a capacidade de determinado programa identificar objetos, formas, pessoas e coisas (ACADEMY, 2018).

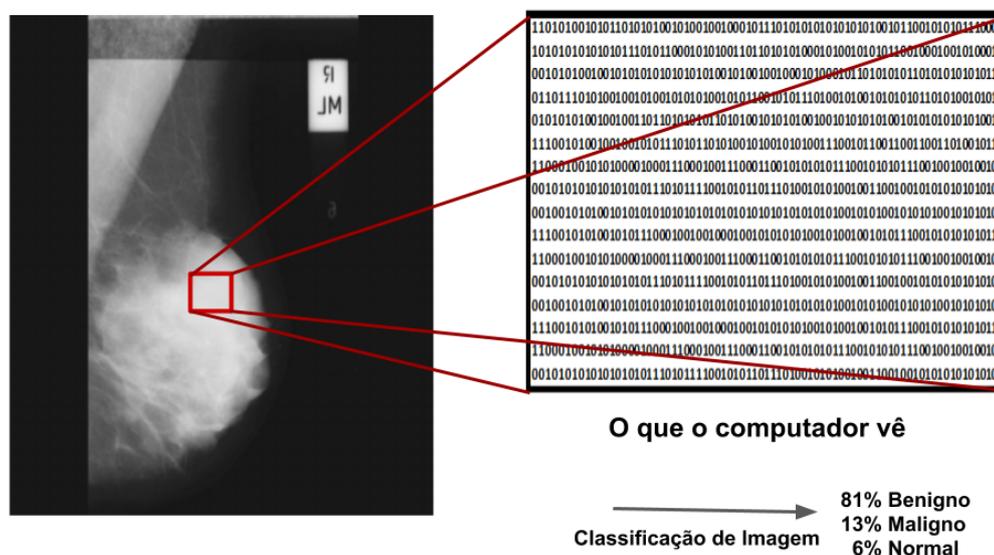


Figura 2 – Reconhecimento de Imagens. Fonte: Criada pelo autor

Através dos métodos de reconhecimento de imagens, os computadores aprendem a reconhecer elementos visuais dentro de uma imagem e através destes elementos detectar padrões emergentes (RAVINDRA, 2017). Uma vez que se detecta certos padrões os computadores podem analisar tais imagens e formular tags e categorias relevantes que fazem sentido (SZELISKI, 2010).

O trabalho de classificação, seja classificação de imagens, áudios, arquivos ou objetos, pode se tornar, exponencialmente complexo em um cenário com milhões de imagens, áudios

ou objetos exigindo interpretação para serem catalogados por milhões de pessoas. Uma maneira para buscar resolver esse difícil problema é com a utilização de redes neurais. As redes neurais possuem grande desempenho no processo de aprendizado e classificação, oferecendo um caminho viável para resolver o problema de forma eficaz unindo o processo de reconhecimento com o poder de classificação e aprendizado das redes neurais (RAVINDRA, 2017; BRANDIZZI, 2020).

2.3 Redes neurais

As redes neurais artificiais (RNA) surgiram na década de 40 e desde então vários pesquisadores voltaram-se para essa linha de pesquisa. A RNA é uma forma de processamento que se caracteriza por ser um sistema que, em algum nível, relembra a estrutura do cérebro humano. Como as redes neurais não são baseadas em regras ou programas, a RNA se constitui em uma forma alternativa à computação algorítmica convencional (BRAGA; CARVALHO; LUDERMIR, 2000).

Sem dúvida, as redes neurais artificiais foram inspiradas biologicamente, mas a correspondência entre elas e os sistemas neurais biológicos que compõem o corpo humano ainda são pequenas. Existem grandes discrepâncias entre as arquiteturas e as capacidades dos nervos artificiais e naturais. O conhecimento sobre as funções reais do cérebro é limitado, no entanto, tentaram imitá-lo. Nenhum modelo foi bem sucedido na replicação do desempenho do cérebro humano. Assim sendo, o cérebro foi, e ainda é, apenas uma metáfora para uma grande variedade de configurações de rede neural que foram desenvolvidas (BRAGA; CARVALHO; LUDERMIR, 2000).

2.3.1 Neurônios biológicos

O cérebro humano é composto de aproximadamente 10^{11} neurônios e a comunicação entre eles é feita através de uma rede de conexão composta por axônios e sinapses. Os neurônios operam em um ambiente químico muito importante para determinar os termos de atuação do funcionamento cerebral. Assim, pode-se considerar o cérebro densamente conectado à rede de comutação elétrica condicionada pelos processos bioquímicos. A vasta rede neural tem uma estrutura elaborada com muitas interações complexas. A entrada para a rede é providenciada por receptores sensoriais (BRAGA; CARVALHO; LUDERMIR, 2000). Um receptor sensorial é a estrutura que reconhece um estímulo no ambiente interno ou externo de um organismo (LIMA, 2009).

Existem dois tipos de sinais químicos que realizam a comunicação do cérebro: sinais elétricos dentro dos neurônios e sinais químicos através da sinapse. A membrana do nodo tem como finalidade produzir e transmitir esses dois sinais. A membrana que reveste o exterior do neurônio é capaz de produzir impulsos nervosos (elétricos), que é uma função

vital do sistema nervoso e central à sua capacidade computacional. Sendo assim, o corpo humano combina os sinais recebidos e, se o valor resultante for acima do limiar de excitação do neurônio, ocorre um impulso elétrico que se propaga através do axônio para os neurônios seguintes (BRAGA; CARVALHO; LUDERMIR, 2000).

2.3.2 Neurônios artificiais

Um neurônio artificial nada mais é que uma unidade de processamento de informação que é fundamental para a operação de uma rede neural (HAYKIN, 2007).

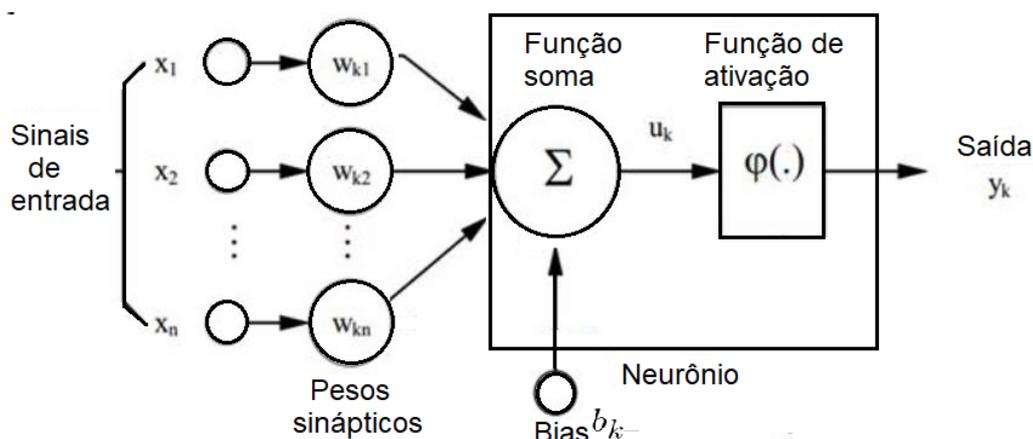


Figura 3 – Modelo não-linear de um neurônio artificial perceptron.

Fonte: Criada pelo autor

Na Figura 3 é possível identificar os três elementos básicos da estrutura de um neurônio artificial:

1. Um conjunto de n sinapses ou elos de conexão – O comportamento das sinapses é emulado através de pesos e cada uma delas possui um peso próprio. Um sinal x_j é recebido na entrada da sinapse j conectada ao neurônio k , esse sinal é multiplicado pelo peso sináptico w_{kj} . Há uma diferença entre a sinapse do cérebro e a realizada no processo da RNA, uma vez que o peso sináptico de um neurônio artificial pode estar em um intervalo de valores positivos ou negativos (HAYKIN, 2007).
2. Um somador para acumular os sinais de entrada – Os sinais são ponderados pelos respectivos pesos sinápticos e aplicados a um *combinador linear* que é a metodologia comum empregada.
3. Uma função de ativação que limita a amplitude da saída do neurônio – A função de ativação também é denominada *função restritiva*, uma vez que restringe o intervalo permissível de amplitude do sinal de saída a um valor finito. Normalmente o intervalo definido para a amplitude do sinal de saída de um neurônio é um intervalo fechado em $[0,1]$ ou $[-1,1]$

Adicionalmente, o modelo na Figura 3 também mostra o bias b_k , um parâmetro adicional na rede neural usado para ajustar a saída do neurônio para se adaptar melhor aos dados fornecidos (TAVARES; NASCIMENTO, 2020). Matematicamente descreve-se um neurônio k a partir das seguintes equações:

$$u_k = \sum_{j=1}^m *w_{kj}x_j \quad (2.1)$$

$$y_k = \varphi(u_k + b_k) \quad (2.2)$$

Os valores $w_{k1}, w_{k2}, w_{k3} \dots w_{km}$ são os pesos sinápticos aplicados respectivamente aos sinais de entrada $x_1, x_2, x_3, \dots, x_m$ recebidos no neurônios k . A *saída do combinador linear* devido aos sinais de entrada é representada na equação por u_k , enquanto o valor do sinal de saída do neurônio k é dado por y_k (HAYKIN, 2007). A função de ativação φ permite regular a amplitude da saída e o bias b_k permite aplicar uma transformação afim na saída u_k do combinador linear. Dependendo se o valor do bias b_k é positivo ou negativo, o vínculo da saída do combinador linear u_k e o *potencial de ativação* v_k do neurônio k é alterado:

$$v_k = u_k + b_k \quad (2.3)$$

A alteração pode ser notada na Figura 4, uma vez que o gráfico resultante da função de v_k em função de u_k não passa mais pela origem para valores diferentes de zero. Note que $v_k = b_k$ em $u_k = 0$ (HAYKIN, 2007).

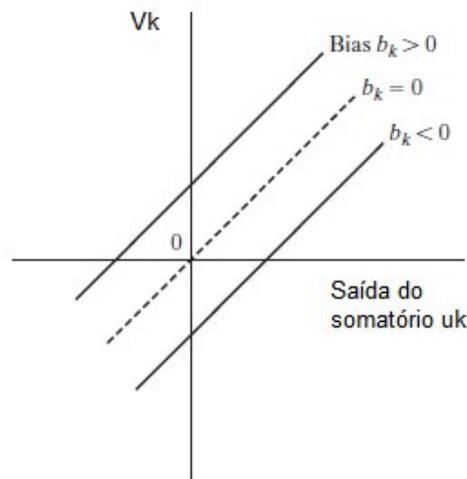


Figura 4 – Transformação linear produzida pela presença do Bias
Fonte (HAYKIN, 2007)

Tipos de função de ativação — A função de ativação define a saída de um neurônio em termos de potencial de ativação v (HAYKIN, 2007). Apresenta-se a seguir os três tipos básicos mais conhecidos de funções de ativação:

1. *Função de Limiar*. Também é conhecida pelas engenharias como *função de Heaviside*, ela corresponde a saída do neurônio k expressada com:

$$\varphi = \begin{cases} 1, & \text{se } v \geq 0; \\ 0, & \text{se } v < 0 \end{cases}$$

onde v_k é o potencial de ativação do neurônio; isto é:

$$\sum_{j=1}^{\infty} = w_{kj}x_j + b_k$$

Esse tipo de função de ativação foi utilizado primeiramente no modelo de neurônio que ficou conhecido na literatura como *modelo McCulloch-Pitts*, em reconhecimento aos estudos pioneiros realizados em 1943 por McCulloch e Pitts (HAYKIN, 2007).

2. *Função linear por partes*. Pode ser vista como uma aproximação de um amplificador não-linear sendo representada matematicamente por:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq +1/2; \\ v, & +1/2 > v > -1/2, \\ 0, & v \leq -1/2 \end{cases}$$

Assume-se que o fator de amplificação dentro da região linear de operação é a unidade (HAYKIN, 2007). Note que se a região linear de operação é mantida sem entrar em saturação surge o *combinador linear*.

3. *Função Sigmoide* Apesar da complexidade de cálculo maior que as funções de ativação antes enumeradas é de longe a forma mais comum de função de ativação utilizada na construção de redes neurais (HAYKIN, 2007). Ela é definida como sendo uma função estritamente crescente que apresenta um balanceamento adequado entre comportamento linear e não-linear. Um exemplo de função sigmoide é a função logística, que é definida como:

$$\varphi(v) = \frac{1}{1 + \exp(-av)}$$

O parâmetro de inclinação da função sigmoide é dado pela letra a . Ao se alterar o parâmetro a tem-se funções sigmóides com diferentes inclinações e no limite, quando o parâmetro de inclinação se aproxima do infinito, a função sigmoide se torna uma função limiar.

Todas as funções de ativação abordadas até o presente momento se estendem em um intervalo de 0 a +1. Muitas das vezes, entretanto, é desejável que a função de ativação se estenda em intervalo de -1 a +1 e assumam uma forma anti-simétrica em relação a origem. Nesse caso, a função de ativação é uma função limiar é definida por:

$$\varphi(v) = \begin{cases} 1, & \text{se } v > 0; \\ 0, & \text{se } v = 0; \\ -1, & \text{se } v < 0 \end{cases}$$

a qual é normalmente denominada de função sinal. Para a forma correspondente de uma sigmoide pode-se utilizar a função tangente hiperbólica definida por:

$$\varphi(v) = \tanh v$$

A função sigmoide assumir valores negativos em alguns casos pode trazer benefícios analíticos (HAYKIN, 2007).

2.3.3 Modelo multicamada de rede neural

Ao longo dos anos muitas redes neurais foram desenvolvidas com finalidades e características diferentes, algumas se destacam em classificação de imagem, outras se destacam na eficiência de processamento de dados sequenciais e outras são capazes de resolver problemas combinatórios difíceis (ACADEMY, 2019).

As arquiteturas mais comuns são as redes *multilayer perceptron* (MLP) e as redes neurais convolucionais. Além destas, existem diversos outros tipos de arquiteturas de redes neurais, dentre elas: redes neurais recorrentes, *long short-term memory* (LSTM), redes de Hopfield, máquinas de Boltzmann, *deep belief networks*, *deep auto-encoders*, *generative adversarial networks*, *deep neural network capsules* (um tipo completamente novo de rede neural lançado no final de 2017) (ACADEMY, 2019).

A perceptron multicamadas (MLP) é uma rede neural semelhante à perceptron convencional, com a diferença de possuir mais de uma camada de neurônios em alimentação direta. As redes MLP possuem multicamadas ocultas e a composição geral das camadas são formadas por um conjunto de diversos neurônios interconectados segundo uma estrutura predeterminada. Uma forma de visualizar tal estrutura é através de uma árvore em que cada nível de altura da árvore equivale a uma camada da rede. A figura 5 mostra uma rede neural que possui quatro camadas: duas ocultas mais uma de entrada e uma de saída:

O projeto específico das camadas das MPL também podem apresentar uma técnica conhecida como *backpropagation*: nesta técnica calcula-se o erro entre o que a rede detectou como resultado e o que realmente era o resultado correto, então recalcula-se o valor de todos os pesos, começando da última camada e indo até a primeira, sempre tendo em vista diminuir esse erro (CARVALHO, 2013).

Embora o projeto das camadas de entrada e saída de uma rede neural seja frequentemente direto, pode haver variação para o projeto das camadas ocultas (ACADEMY, 2019). Em particular, não é possível resumir o processo de concepção das camadas ocultas com

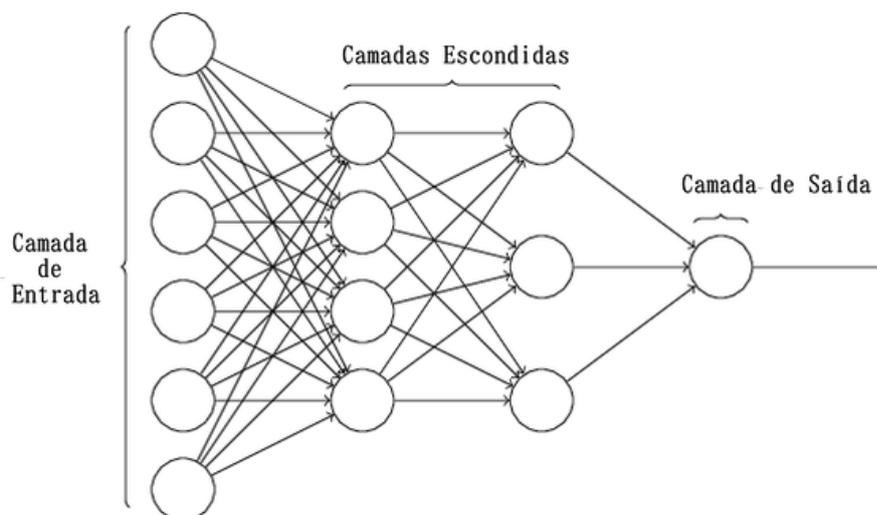


Figura 5 – Rede neural em camadas.

Fonte: Criada pelo autor

poucas regras simples. Comumente, os pesquisadores de redes neurais utilizam heurísticas de projeto para encontrar arquiteturas de camadas ocultas que melhor contribuam para soluções de problemas específicos. O projeto das camadas ocultas é um dos pontos cruciais em modelos de *Deep Learning* (ACADEMY, 2019).

As redes neurais convolucionais também possuem uma arquitetura de camadas específica para reconhecimento de padrões em imagens (ACADEMY, 2019). Além das camadas ocultas existentes em redes totalmente conectadas presentes em redes neurais convolucionais, também existem as camadas convolucionais. A LeNet foi uma das primeiras redes neurais convolucionais desenvolvidas, ela utiliza *backpropagation* em uma rede *feedforward* com muitas camadas ocultas, muitos mapas de unidades replicadas em cada camada, agrupando as saídas de unidades próximas, criando assim uma rede ampla que lida com várias entradas (CARVALHO, 2013).

2.3.4 Reconhecimento de imagem por rede neural

Como mencionado em seções anteriores, o reconhecimento de imagem não é uma técnica trivial. Uma possível solução teórica é o uso de redes neurais para analisar imagens, porém, observando sob a perspectiva computacional, o processo é custoso na prática (RAVINDRA, 2017). Por exemplo, uma rede neural convencional tentando processar uma imagem pequena (30 x 30 *pixels*) necessitaria de milhões de parâmetros, variáveis que a rede estimará ao longo do treinamento, mais 900 entradas. Uma máquina razoavelmente poderosa consegue lidar com isso, mas uma vez que as imagens se tornam muito maiores (por exemplo, 500 x 500 *pixels*) o número de parâmetros e entradas aumenta significativamente (RAVINDRA, 2017).

Outro problema relacionado à aplicação de redes neurais ao reconhecimento de

imagens é o sobreajuste ou *overfitting*. De modo geral, o *overfitting* é um termo usado para descrever quando um modelo estatístico se ajusta muito bem ao conjunto de dados anteriormente observado, mas se mostra ineficaz para visualizar/reconhecer novos resultados (CÁRDENAS-MONTE, 2015). As redes neurais analisam todas as informações que a entrada possui e se ajustam a elas, não descartando informações menos importantes, isso faz com que as redes neurais se ajustem completamente às imagens de entrada, podendo assim levar a um *overfitting* (ALIGER, 2019; RAVINDRA, 2017). O *overfitting* ocorre quando um modelo ajusta-se muito bem aos dados em que foi treinado, normalmente devido à inclusão de parâmetros ou dados excessivos (aumentando ainda mais os custos computacionais), sendo que a exposição do modelo a novos dados resulta em uma perda no desempenho geral (RAVINDRA, 2017).

O grande atrativo das redes neurais é que elas conseguem resolver parte do problema combinando técnicas de redes neurais e reconhecimento de imagens juntas em uma mesma arquitetura (ACADEMY, 2018). A forma como uma rede neural é estruturada é importante, sendo que uma mudança relativamente direta pode tornar as imagens, mesmo as grandes, mais gerenciáveis. As CNNs realizam compensação consciente para manipular imagens e filtrar e desconsiderar generalizações para que a rede possua solução mais viável (RAVINDRA, 2017). As CNNs podem ser divididas em duas partes: camadas convolucionais e rede neural convencional. Uma visão geral é que o processamento de imagens ocorre nas camadas convolucionais quando a rede considera, realizando cálculos matemáticos que serão abordados na seção 2.4, as semelhanças na relação dos *pixels* em proximidade dentro da imagem. Por outro lado, a CNN considera que *pixels* distantes um do outro possuem pouca ou nenhuma relação. Tecnicamente, as CNNs tornam o processamento de imagem gerenciável por meio da filtragem das conexões por proximidade (RAVINDRA, 2017). Desta forma é possível notar que as redes neurais e as redes neurais convolucionais possuem forte ligação com o reconhecimento de imagens, principalmente por suas habilidades de agrupar, classificar e encontrar padrões.

2.4 Redes neurais convolucionais

As redes neurais convolucionais (*ConvNets* ou CNNs) se tornaram um padrão quando se trata de visão computacional. Elas possuem facilidade de treinamento quando se tem um grande conjunto de amostras rotuladas que representam as diferentes classes-alvos (ARAÚJO et al., 2017). As CNNs possuem algumas vantagens, entre elas:

- Capacidade de extrair características relevantes através de aprendizado de transformações filtros/*kernels*. Um *kernel* é uma matriz utilizada na operação de multiplicação de matrizes. Esta operação é aplicada diversas vezes em diferentes regiões da imagem.

- Depende de um menor número de parâmetros de ajustes do que as redes totalmente conectadas que possuem mesmo número de camadas. Diferente das redes totalmente conectadas, como cada número de uma camada da CNN não é conectada com todas as unidades da camada seguinte, há menos pesos a serem atualizados, facilitando o treinamento.

As CNNs podem ser utilizadas de diversas formas. Nas Figuras 6 e 7 é possível ver exemplos de seu uso aplicado em outras áreas. Na Figura 6 o aplicativo realiza sugestões de palavras que descrevem o cenário da imagem, a Figura 7 mostra exemplos de uma CNN usada para reconhecimento de objetos, pessoas e animais no dia a dia (ARAÚJO et al., 2017).

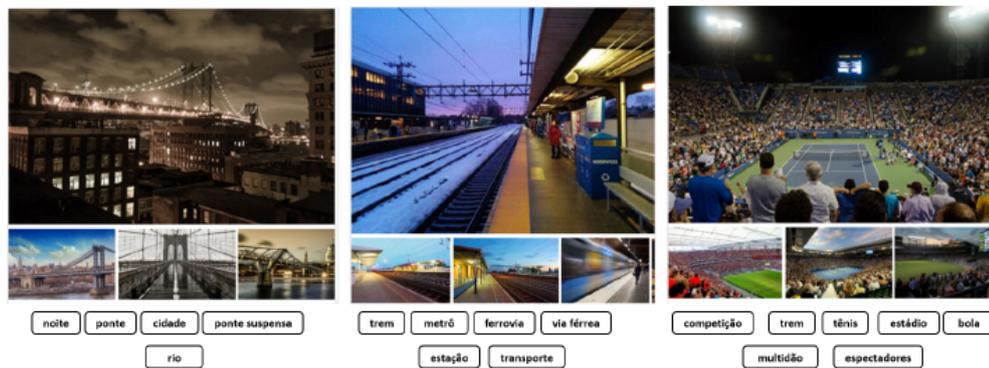


Figura 6 – Palavras relevantes que descrevem a cena em uma *ConNet*.
Fonte: (ARAÚJO et al., 2017)

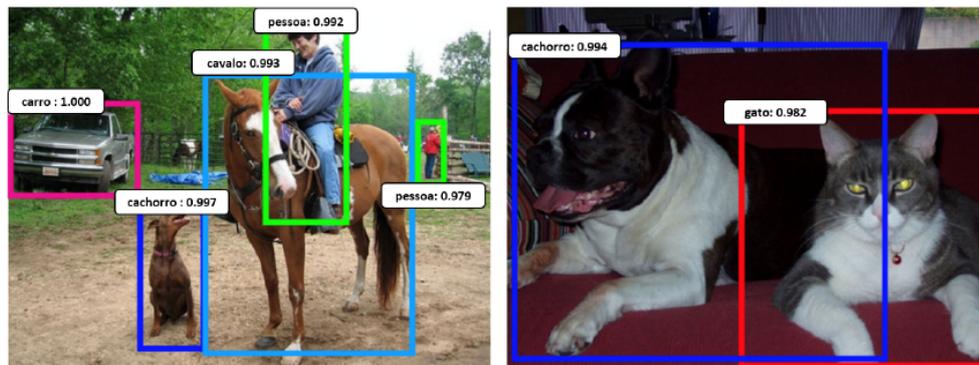


Figura 7 – Reconhecimento de objetos, pessoas e animais por uma *ConvNet*.
Fonte: (ARAÚJO et al., 2017)

Atualmente existem várias bibliotecas que oferecem as principais operações utilizadas pelas CNNs, algumas dessas bibliotecas são *Caffe*, *MatConvNet*, *Tensorflow*, entre outros (ARAÚJO et al., 2017), (SESTREM, 2019).

2.4.1 Modelo *LeNet* de rede neural convolucional

A *LeNet* foi um dos primeiros projetos de CNN realizados, proposto por Yann LeCun em 1988, que acabou por auxiliar a impulsionar o campo da *Deep Learning* (RAMSUNDAR; ZADEH, 2018). Inicialmente a *LeNet* foi utilizada para reconhecimento de caracteres, códigos postais e dígitos numéricos. Outras arquiteturas de CNNs foram propostas ao longo dos anos como forma de melhoria da *LeNet*, contudo as novas CNNs continuam compartilhando conceitos fundamentais entre si que serão abordados a seguir (ARAÚJO et al., 2017) (ZACCONE; KARIM; MENSRAWY, 2017).

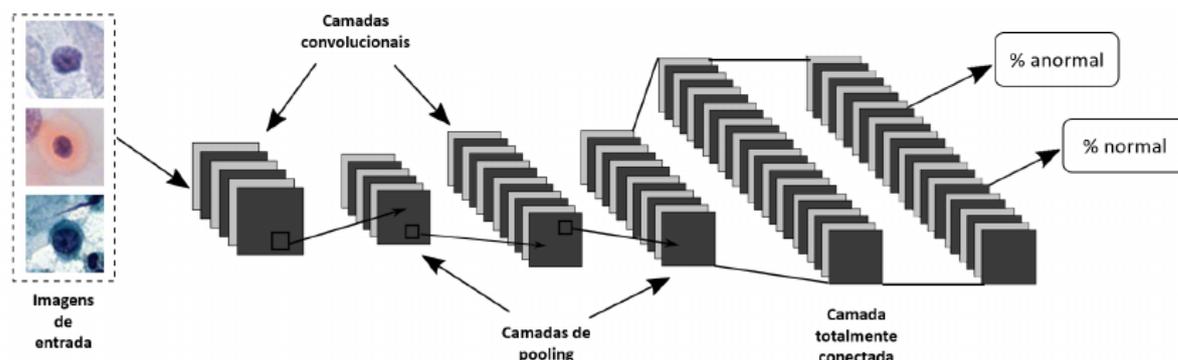


Figura 8 – Ilustração da arquitetura de uma *LeNet* que classifica imagens de entrada de células anormal ou célula normal e suas três principais camadas: convolucionais, *pooling* e totalmente conectadas. Fonte: (ARAÚJO et al., 2017)

As CNNs são compostas por camadas e cada uma delas tem função específica na propagação de sinal de entrada. A figura 8 ilustra a arquitetura de uma *LeNet* e suas três principais camadas: convolucional, *pooling* e totalmente conectada. As camadas convolucionais têm como finalidade extrair atributos dos volumes de entrada. A camada de *pooling* tem como função reduzir progressivamente a dimensão espacial do volume de entrada, conseqüentemente a redução diminui o custo computacional (KARPATHY, 2017).

2.4.2 Camadas de Convolução

As camadas de convolução são formadas por um conjunto de filtros que recebem como entrada um arranjo 3D, também conhecido como volume. A convolução é um operador linear, o mais utilizado nas CNNs, que dá o nome a essas redes. A convolução é realizada deslizando uma máscara, chamada de elemento estruturante, sobre cada pixel da imagem de entrada gerando como resultado um mapa de características (DALVI, 2018; KARPATHY, 2017). A convolução é bastante presente nas áreas de processamento de sinal e imagens. Durante a fase de treinamento da rede, os filtros da rede são ajustados para que sejam ativados em presença de características relevantes identificadas no volume de entrada, como orientação de bordas ou manchas de cores (KARPATHY, 2017). A relevância dessas características é avaliada de tal forma que os resultados sejam otimizados em função de

um conjunto de dados amostrais previamente classificados (ARAÚJO et al., 2017). Cada um desses filtros dá origem a uma estrutura conectada localmente que percorre toda a extensão do volume de entrada. A somatória do produto ponto a ponto entre os valores de um filtro e cada posição do volume de entrada é uma operação conhecida como convolução, a qual é ilustrada na Figura 9 (ARAÚJO et al., 2017).



Figura 9 – Exemplo de operação de convolução. Fonte: (DALVI, 2018)

Na Figura 9, tem-se um exemplo de operação de convolução, com uma matriz 5 x 5 que representa a imagem de entrada em tons de cinza e um elemento estruturante de dimensões 2 x 2. A resultante do somatório dos produtos de cada pixel é uma matriz 4 x 4. Com isso, pode-se dizer que cada elemento estruturante se comporta como um neurônio da ANN, podendo formar uma ou várias camadas com um ou mais neurônios, ou nesse caso, elementos estruturantes (DALVI, 2018). Cada elemento estruturante possui uma matriz de pesos que são ajustados em uma etapa de treinamento supervisionado assim como os neurônios das ANNs. A vantagem da CNN está na operação de convolução que é capaz de considerar a estrutura espacial da imagem, além de se mostrar mais eficiente que as ANN por conseguir analisar imagens relativamente grandes. Uma imagem de entrada de tamanho $N \times N$ pixels em uma ANN convencional exigiria N^2 pesos para cada neurônio da camada de entrada, enquanto que em uma CNN é necessário um número de pesos relativo à dimensão do elemento estruturante que é geralmente menor que a imagem entrada (DALVI, 2018).

2.4.3 Operação de *Pooling*

A operação de *Pooling* é extremamente importante, principalmente para transformar imagens de entrada de $N \times N$ pixels em valores unitários de classe final. O *Pooling* tem como principal função reduzir a dimensionalidade da entrada, descartando informações inúteis e preservando as mais significativas (DALVI, 2018).

O *Pooling* é realizado deslizando um elemento estruturante sobre a estrutura de entrada, de maneira semelhante a convolução, a diferença é que ao invés de multiplicar ele seleciona um valor sobre a sobreposição do elemento estruturante com a matriz de entrada.

O valor selecionado pode variar dependendo do tipo de operação que se deseja realizar. Geralmente seleciona-se o maior valor em uma operação, chamada de *max-pooling*, porém existem aplicações que utilizam operações que selecionam a media ou a mediana. Na figura 11 observa-se o exemplo de aplicação que utiliza o *max-pooling* com elemento estruturante de 2 x 2 sobre uma imagem de entrada de 5 x 5.

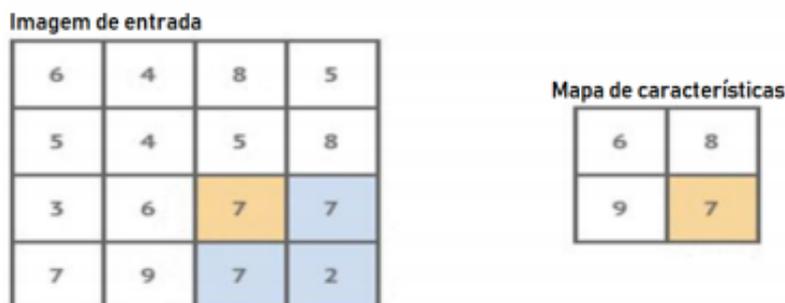


Figura 10 – Exemplo de operação *max-pooling*.
Fonte: (DALVI, 2018)

2.4.4 Função ReLU

A ReLU é uma função de ativação bastante utilizada nas redes neurais convolucionais. A função de ativação tem como finalidade introduzir um componente não linear às CNNs, também é utilizada em redes neurais artificiais com o mesmo objetivo. Existem outras funções de ativações como vistas na seção 2.2.4. Esta função atua entre uma camada e outra, aplicando uma não-linearidade à informação que é passada para a camada seguinte (DALVI, 2018).

Ao adicionar uma não-linearidade entre as camadas é dado ao modelo mais expressividade do que um modelo linear, uma vez que a etapa de convolução só realiza operações lineares. A ReLU é realizada em cada *pixel* do mapa de característica de forma a alinhar valores negativos, de acordo com a equação a seguir:

$$\text{ReLU}(x, y) = \max[0, (g(x, y))],$$

onde $g(x, y)$ é o valor de entrada e (x, y) as coordenadas do pixel em questão. Logo, os valores negativos ou abaixo de zero são convertidos em zero e os valores que são acima de zero permanecem. Tal função possui bastante esparsidade por transformar muitos valores em 0.

A ReLU possui algumas variantes, onde os valores negativos não são substituídos por 0, eles são multiplicados por um parâmetro a que pode ser pré-determinado ou aprendido. Uma equação que expressa ambos os casos é representada por:

$$\text{ReLU}(x, y) = \max[0, (g(x, y))] + a \min[0, (g(x, y))]$$

2.4.5 Camada totalmente conectada

A camada totalmente conectada consiste na ANN clássica. As ANNs durante muito tempo foi-se o mecanismo mais utilizado na literatura, atingindo ótimos resultados nas mais diversas aplicações. Como essas redes são bem consolidadas a presença delas nas CNNs se torna essenciais assim como todas as camadas prévias a ela. Assim como as ANNs, esta camada é composta por neurônios que possuem números de pesos relativos à dimensão dos valores de entrada. O vetor de entrada pode ser obtido a partir de saída de neurônios de outra camada conectada anteriormente, ou por dados de uma estrutura resultante de outros tipos de camadas. No segundo caso, os dados da estrutura anterior, que geralmente é tridimensional é reorganizada em vetores para serem usados como entradas para camadas totalmente conectadas, como pode-se ver na figura.

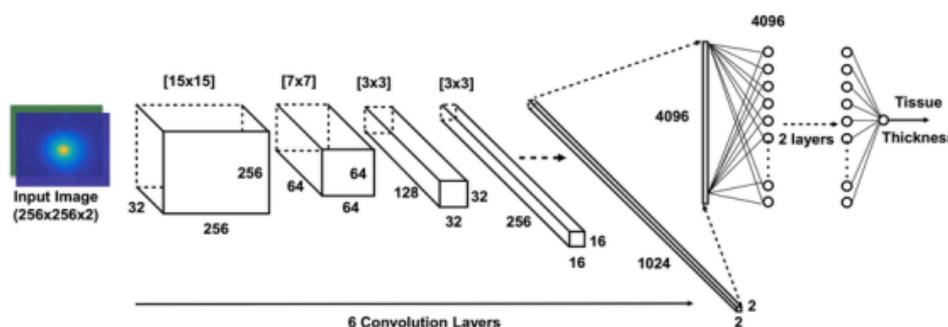


Figura 11 – Exemplo de operação RNA com camada totalmente conectada. Fonte: (DALVI, 2018)

2.5 Aplicação das CNN no reconhecimento de câncer de mama

Por volta de 2012 as redes neurais convolucionais tiveram novo surto de popularidade dentro do meio acadêmico. Isso se deu pela rede neural convolucional conhecida como AlexNet que propôs uma arquitetura de rede que apresenta desempenho expressivo quanto a reconhecimento de imagens Draelos (2019). Pela investigação da literatura notou-se que a partir dessa época muitos trabalhos acadêmicos começaram a testar o desempenho das redes neurais convolucionais no reconhecimento de câncer de mama através de imagens médicas. Alguns trabalhos como Zhang (2017), Alom et al. (2019), DALVI (2018), Zhou (2016) abordam a importância das redes neurais convolucionais na detecção de câncer de mama em mamografias e/ou imagens histopatológicas.

3 Tecnologias e Métodos

Neste capítulo são apresentadas a metodologia adotada, juntamente com as ferramentas, bibliotecas e bases de dados que foram utilizadas no desenvolvimento da rede neural convolucional objeto desse trabalho.

3.1 Metodologia

Para o desenvolvimento do presente trabalho foram executadas as seguintes etapas:

1. Realização de revisão da literatura que abordam o uso de redes neurais convolucionais aplicadas no reconhecimento de câncer em imagem de mamografia.
2. Escolha do trabalho de [Zhou \(2016\)](#) como *baseline* para investigação.
3. Implementação de uma CNN inspirada na arquitetura idealizada por [Zhou \(2016\)](#).
4. Propostas de melhorias na CNN implementada, incluindo a investigação de trabalhos futuros propostos por [Zhou \(2016\)](#) e levantamento de outros fatores considerados de interesse.
5. Realização da investigação e execução dos experimentos.
6. Avaliação dos resultados.

As etapas apresentadas acima são descritas com detalhe no capítulo 4.

3.2 Python e Bash

Durante o trabalho foram utilizadas a linguagem de programação Python, para construção das redes neurais convolucionais, e a linguagem de *script* Bash, para implementar o pré-processamento das imagens.

3.3 Biblioteca Keras

A biblioteca Keras foi utilizada para desenvolver a rede neural convolucional. Desenvolvida e mantida por François Chollet, engenheiro do Google, o Keras é uma biblioteca de redes neurais minimalista e de alto nível, capaz de rodar sobre o TensorFlow. O Keras foi desenvolvido com o objetivo de permitir protótipos e experimentações fáceis e rápidas, além disso ele é compatível com Python 2.7 ou 3.5 e pode ser executado perfeitamente

em GPUs e CPUs (ZACCONE; KARIM; MENSRAWY, 2017). Os principais motivos da escolha do Keras se dá pela praticidade e a facilidade desta biblioteca quando se trata de desenvolvimento de redes neurais convolucionais, sem contar que possui uma documentação completa no *site*. A figura 12 apresenta sua *homepage*.

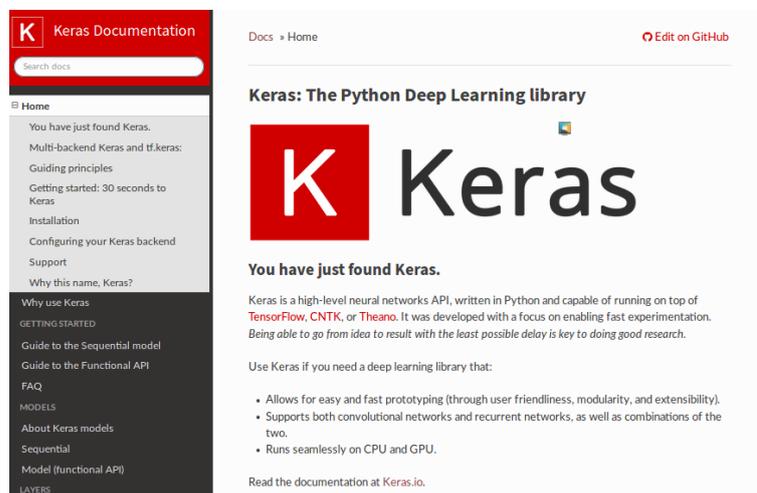


Figura 12 – Keras *homepage*

3.4 Biblioteca Tensorflow

Junto ao Keras, a biblioteca Tensorflow foi utilizada como *backend* para o desenvolvimento da rede neural convolucional. Iniciado em 2011 o Tensorflow é uma biblioteca de código aberto criada pela equipe do Google Brain, voltada para implementação e implantação de aprendizado de máquina e redes neurais profundas na Google. Apesar de compatível com várias linguagens de programação o Tensorflow possui Python como linguagem de programação padrão. O principal motivo para a escolha da biblioteca Tensorflow é a abstração que ele proporciona, pois evita que o desenvolvedor tenha que lidar com detalhes básicos da implementação do algoritmo podendo assim se concentrar na lógica geral da aplicação. Outro ponto positivo do TensorFlow que resultou na sua escolha é a flexibilidade que ele possui com as várias APIs que existem e que ajudam ainda mais a utilização da tecnologia pelo usuário, uma destas APIs é o Keras que como já mencionado também esta sendo utilizado neste trabalho. É importante salientar que o trabalho de Zhou (2016) também utiliza a biblioteca Tensorflow em sua arquitetura de rede neural convolucional, amparando ainda mais a escolha desta biblioteca no presente trabalho.

3.5 Banco de dados MiniMammographic MIAS

A Sociedade de Análise de Imagem Mamográfica (MIAS) é uma organização de grupos de pesquisa do Reino Unido interessados na compreensão de mamografias e que

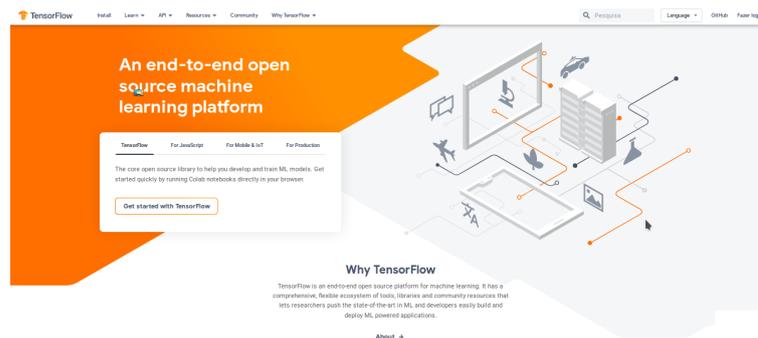


Figura 13 – Tensorflow homepage

geraram um *dataset* de imagens digitais de mamografias.

Os filmes retirados do Programa Nacional de Rastreamento da Mama do Reino Unido foram digitalizados para 50 micrómetros de *pixel* com um microdensitômetro de varredura Joyce-Loebl, um dispositivo linear no intervalo de densidade óptica 0-3,2 e representando cada *pixel* com uma palavra de 8 bits (SUCKLING, 1994).

O banco de dados contém 322 filmes digitalizados e está disponível em fita de 8 GB (ExaByte) de 2,3 GB. Também inclui parte do diagnóstico do radiologista sobre os locais de quaisquer anormalidades que possam estar presentes. O banco de dados foi reduzido para uma borda de 200 micron *pixels* e preenchido/recortado para que todas as imagens tenham 1024x1024. Imagens mamográficas estão disponíveis através do Arquivo Europeu de Processamento de Imagens Piloto (PEIPA) na Universidade de Essex. (SUCKLING, 1994).

Muitos dos trabalhos pesquisados da literatura utilizaram o *dataset* Mini-MIAS para treinamento e validação das redes neurais convolucionais, como Souza (2018), SOUTO (2014), Zhou (2016), entre outros.

3.6 Seleção do trabalho de Zhou como ponto de partida

O presente trabalho tem como cenário inicial (*baseline*) o artigo de Zhou (2016). O artigo referenciado tem como objetivo a viabilidade de analisar mamografias (imagens de raios X de mamas) usando algoritmos de aprendizado de máquina aplicado no reconhecimento de imagens para a classificação de anormalidades benignas e cancerígenas, bem como na detecção de anormalidades. Os motivos que levaram à escolha desse trabalho como ponto de partida foram: a estratégia utilizada por Zhou pareceu promissora e poderia ser empregada no auxílio ao diagnóstico médico de mamografias, enquanto as ferramentas utilizadas em seu trabalho estão disponíveis para uso gratuito.

4 Desenvolvimento

Neste capítulo são apresentadas as etapas para a realização do trabalho, o processo de seleção da arquitetura da rede neural, a escolha dos fatores para investigação, o projeto dos experimentos e a avaliação dos resultados.

4.1 Projeto da CNN implementada

Uma vez selecionado como ponto de partida o trabalho descrito no artigo de [Zhou \(2016\)](#), buscou-se aqui implementar uma arquitetura da CNN inspirada naquela que foi idealizada pelo mesmo. A arquitetura desenvolvida por [Zhou \(2016\)](#) consiste em:

- Uma camada convolucional com 32 filtros/*kernel* de tamanho 5x5.
- Uma camada convolucional com 64 filtros/*kernel* de tamanho 5x5
- Camada *MaxPooling* de tamanho 3x3 após a convolução ¹.
- Um *Dropout* de 0.5 ².
- Uma rede totalmente conectada de 1024 neurônios.
- Funções de ativação Relu ³.

O artigo não fornece maiores informações sobre a arquitetura da CNN utilizada, portanto foi necessário convencionar como os itens citados são combinados e configurados dentro do modelo através de pesquisas na literatura e testes empíricos. Na figura 14 é possível ver a organização convencionada.

A CNN implementada, representada na figura 14, foi desenvolvida utilizando a linguagem Python juntamente com as bibliotecas Keras e Tensorflow.

4.1.1 Avaliação dos resultados obtidos pela CNN implementada

Uma vez implementada a CNN, realizaram-se os testes para avaliar a proximidade de resultados obtidos com aqueles apresentados por [Zhou \(2016\)](#). Para a realização dos testes foi criado um projeto de experimentos e utilizado o mesmo *dataset* empregado por [Zhou \(2016\)](#), bem como a mesma quantidade de passos de treinamento.

¹ Maiores detalhes sobre *pooling*, ver seção 2.4.3.

² Dropout é uma estratégia utilizada para evitar *overfitting* que consiste em eliminar aleatoriamente uma porcentagem dos neurônios durante o processo de treinamento.

³ Maiores detalhes sobre *Relu*, ver seção 2.4.4.

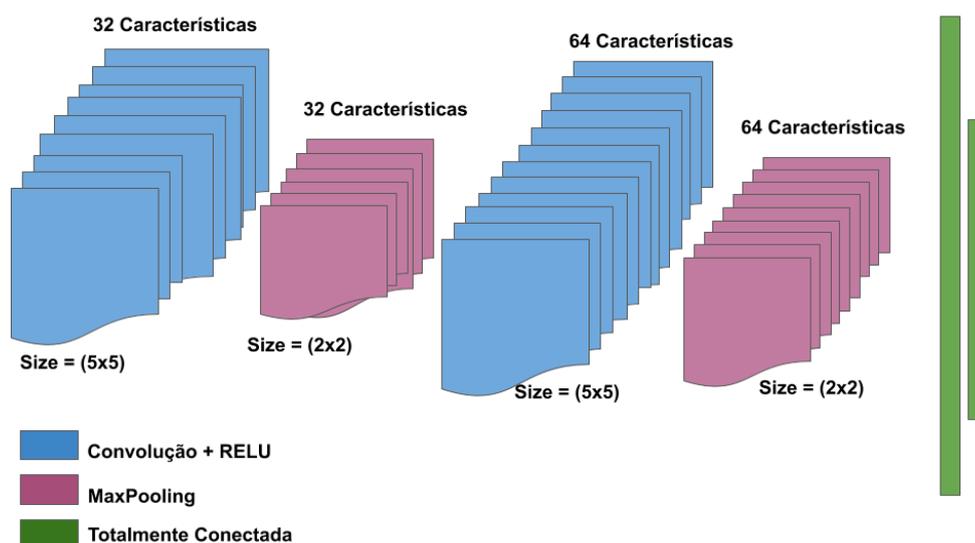


Figura 14 – Arquitetura da rede neural convolucional construída.

Fonte: Elaborada pelo próprio autor

O artigo de Zhou (2016) cita um tratamento de imagens do *dataset* MIAS (seção 3.5). O tratamento consiste em recortar a imagem da mamografia digital em quadriculados uniformes de 48x48 *pixels*. Cada um dos quadriculados por sua vez pode ser submetido à rede para treinamento da CNN, uma vez que o MIAS informa quais mamografias contêm tumor e as coordenadas na imagem da região onde está o tumor. O conjunto de treinamento consiste na seleção de quadriculados situados nas regiões dos tumores benignos e malignos, mais quadriculados escolhidos aleatoriamente das imagens que não possuem tumor.

Devido ao problema da escassez de imagens benignas e malignas no *dataset* quando comparadas às imagens normais disponíveis, realizou-se tratamentos para sanar este problema. Tal tratamento consiste em realizar espelhamento vertical e horizontal das imagens para gerar novas imagens. Desenvolveu-se um *script* em *Bash* que além de recortá-las espelhava as imagens horizontalmente e verticalmente. Com a geração de novas imagens, as classes benignas, malignas e normais aproximaram-se em quantidade.

No trabalho de Zhou (2016) foi realizada a execução de 5000 passos da CNN, sendo assim o presente artigo replica o mesmo mecanismo e também executa um treinamento na CNN de 5000 passos. Ao final da execução são coletados os dados de validação de acurácia. Os resultados obtidos pela CNN desenvolvida neste trabalho de conclusão foram ligeiramente melhores que os obtidos por Zhou (2016).

Tabela 1 – Comparação dos resultados da acurácia

Algoritmo	Resultado da Acurácia
Algoritmo em Zhou	0.60897
Algoritmo Desenvolvido	0.64120

A Tabela 1 mostra os resultados obtidos pelas diferentes implementações.

4.2 Fatores aplicados na CNN implementada

No artigo de Zhou (2016) é sugerido como trabalho futuro aumentar o número camadas convolucionais, uma vez que foi relatado no artigo problemas com limitação de hardware. Assim foi definido como primeiro fator para investigação a quantidade de camadas convolucionais.

O segundo fator levantado para investigação foi o Tamanho do Recorte da Imagem submetida para a rede. A escolha do recorte de 48x48 para os quadriculados, tamanho proposto no artigo de Zhou (2016), não foi justificada, tornando-se um dos principais pontos a serem investigados no presente trabalho. Havia a suspeita que o recorte de 48x48 poderia impactar a acurácia da rede, além de limitar o aprendizado da rede à análise local de uma pequena região, diminuindo assim a habilidade da mesma de identificar propriedades cancerígenas em imagens de tamanho maior e com características diferentes. Dessa forma, havia a suspeita que o *overfitting* poderia ter maior chance de ocorrer.

Por fim, a quantidade de classes foi o último fator escolhido para investigação. A escolha foi realizada devido à dificuldade das redes neurais em classificar *datasets* que possuem muitas classes, ainda mais imagens como as de mamografia que, a olho nú, não apresentam diferenças perceptivas. As duas opções em questão consistem em adotar 2 classes, normal ou anormal, ou 3 classes, normal, benigno ou maligno.

Tabela 2 – Fatores do Experimento

Fator	Nível (-)	Nível (+)
Tamanho do Recorte da Imagem	48x48	96x96
Qtde. de Camadas Convolucionais	Duas Camadas Convolucionais e Uma Totalmente Conectada	Três Camadas Convolucionais e Uma Totalmente Conectada
Quantidade de Classes	Três Classes	Duas Classes

A Tabela 2 mostra os fatores e os níveis a serem investigados nos experimentos. Para o fator Tamanho do Recorte da Imagem é definido as dimensões do recorte 48x48 *pixels* como nível baixo e o recorte 96x96 *pixels* como nível alto do fator. O fator quantidade de camadas conta também com dois níveis, sendo o nível baixo composto de uma camada convolucional de 32 filtros de *kernel* 5x5 e outra camada convolucional de 64 filtros de *kernel* 5x5; para o nível alto tem-se duas camadas convolucionais de 32 filtros de *kernel* 5x5 e uma camada convolucional de 64 filtros de *kernel* 5x5. Por fim, temos o fator denominado quantidade de classes, que possui como nível baixo as três classes benigno, maligno e normal, e como nível alto possui as duas classes anomalia e normal.

A decisão de aumentar a imagem apenas em 96x96 deu-se porque a máquina disponibilizada pela Instituição Federal de Minas Gerais (IFMG) para a realização dos testes não suportou a execução da rede quando as imagens do *dataset* foram superiores a 96x96. Ao tentar aumentar o Tamanho do Recorte da Imagem ocorria um estouro de memória que acarretava no aborto da execução da rede neural convolucional. A adição de camadas convolucionais também foi limitada pelo hardware, uma dificuldade também relatada por Zhou (2016).

4.2.1 Execução dos experimentos

Todos os experimentos foram executados em um mesmo computador cedido pelo IFMG Campus Formiga e configurados para realizar 5000 passos de treinamento. O projeto de experimentos idealizado incluiu 3 fatores com 2 níveis cada e uma repetição, configurando um projeto fatorial 2^3 (JAIN, 1992) com um total de 16 experimentações. A métrica escolhida para análise é a acurácia. A acurácia foi escolhida por ser uma métrica de comum senso no uso de avaliação de redes neurais convolucionais, conforme trabalhos de Zhou (2016), LEITE (2011) e Souza (2018). O resultado da acurácia consiste em apresentar dentre todas as classificações da CNN, quantas o modelo classificou corretamente Rodrigues (2019). O modelo linear de regressão foi escolhido dentre os demais pela facilidade em seus cálculos e por possuir exemplo de sua aplicação no âmbito da computação no livro de Jain (1992).

O modelo estatístico geral para descrever o comportamento da CNN construída pode ser descrito como:

$$Y = q_0 + q_a A + q_b B + q_c C + q_{ab} AB + q_{ac} AC + q_{bc} BC + q_{abc} ABC \quad (4.1)$$

onde as variáveis A, B, C correspondem respectivamente aos níveis alto (+1) e baixo (-1) dos fatores Tamanho do Recorte da Imagem, Quantidade de Camadas e Quantidade de Classes, conforme mostrado na Tabela 2.

O projeto de experimentos realizado permite estimar os 3 efeitos principais, os 3 efeitos de interações de segunda ordem, 1 efeito de interação de terceira ordem. Também foi definido que cada configuração de experimento seria executada duas vezes para que fosse possível estimar a porcentagem de erro do modelo estatístico, resultando num total de 16 experimentações realizadas.

Na Tabela 3 é possível ver os fatores codificados, onde o sinal (+) representa o nível alto e o sinal (-) representa o nível baixo, seguidos dos resultados coletados para cada combinação dos níveis dos fatores.

Os resultados r1 e r2 na tabela são referentes ao máximo local de acurácia encontrada em cada repetição dada a configuração do experimento.

Tabela 3 – Fatores Codificados

	Tam. da Imagem	Quant. de Camadas	Quant. de Classes	Resultado (r1)	Resultado (r2)	Média (r1 e r2)
1	-	-	-	0.6471	0.6353	0.64120
2	-	-	+	0.8431	0.9216	0.88235
3	-	+	-	0.6824	0.6784	0.68040
4	-	+	+	0.8863	0.8745	0.88040
5	+	-	-	0.5950	0.6918	0.64340
6	+	-	+	0.9032	0.8315	0.86735
7	+	+	-	0.6380	0.5699	0.60395
8	+	+	+	0.9068	0.8889	0.89785

A Tabela 4 apresenta a relação completa dos fatores de controle do projeto $2^3 2$, onde as colunas A, B, C representam os níveis dos fatores principais vistos na tabela 2. As outras quatro colunas que se seguem equivalem ao produto resultante da multiplicação dos sinais dos fatores combinados do experimento. Nas demais, R_1 e R_2 são os valores mensurados para as repetições e Y é, ao mesmo tempo, a média desses valores e o resultado estimado pelo modelo.

Tabela 4 – Projeto Fatorial $2^{3 \times 2}$

I	A	B	C	AB	AC	BC	ABC	$(R_1 ; R_2)$	Y
1	-1	-1	-1	+1	+1	+1	-1	(0,6471 ; 0.6353)	0.64120
2	-1	-1	+1	+1	-1	-1	+1	(0,8431 ; 0.9216)	0.88235
3	-1	+1	-1	-1	+1	-1	+1	(0.6824 ; 0.6784)	0.68040
4	-1	+1	+1	-1	-1	+1	-1	(0.8863 ; 0.8745)	0.88040
5	+1	-1	-1	-1	-1	+1	+1	(0.5950 ; 0.6918)	0.64340
6	+1	-1	+1	-1	+1	-1	-1	(0.9032 ; 0.8315)	0.86735
7	+1	+1	-1	+1	-1	-1	-1	(0.6380 ; 0.5699)	0.60395
8	+1	+1	+1	+1	+1	+1	+1	(0.9068 ; 0.8889)	0.89785
T	-0.07180	0.02830	0.95900	-0.04620	0.07670	0.02880	0,11110	12,19380	6,09690
\hat{T}	-0,00897	0,00354	0,11988	-0,00578	0,00959	0,00360	0,01389	0,76211	0.89785

Os valores T para cada coluna X são o somatório do sinal do fator multiplicado pelo valor correspondente em Y para configuração do experimento:

$$T_X = \sum_{i=1}^{2^3} (X_i * Y_i) \quad (4.2)$$

Os valores \widehat{T} para cada coluna X são a média do somatório T para cada fator:

$$\widehat{T}_X = \frac{T_X}{2^3} \quad (4.3)$$

Dados os valores, computando as fórmulas obteve-se os seguintes efeitos:

$$\begin{aligned} q_0 &= 0,7621 & q_{ab} &= -0,00578 \\ q_a &= -0,0089 & q_{ac} &= 0,00959 \\ q_b &= 0,00354 & q_{bc} &= 0,00360 \\ q_c &= 0,11988 & q_{abc} &= 0,01389 \end{aligned}$$

Uma vez que os efeitos foram calculados, o modelo pode ser usado para estimar a resposta esperada para qualquer configuração de experimentos dados os níveis dos fatores, conforme a Equação 4.1. A diferença entre a estimativa do modelo e o valor real mesurado y_{ij} no i -ésimo experimento na j -ésima repetição representa um erro experimental encontrado:

$$e_{ij} = y_{ij} - Y_i \quad (4.4)$$

Os erros experimentais para cada repetição do experimento na configuração indicada são apresentados na Tabela 5.

Tabela 5 – Erros experimentais dados o modelo e a configuração dos fatores

<i>Config.</i>	<i>Erro(R₁)</i>	<i>Erro(R₂)</i>
1	0,00590	-0,00590
2	-0,03925	0,03925
3	0,00200	-0,00200
4	0,00590	-0,00590
5	-0,04840	0,04840
6	0,03585	-0,03585
7	0,03405	-0,03405
8	0,00895	-0,00895

A soma de todos erros experimentais deve ser zero, mas a soma do quadrado dos erros (SSE) pode ser usada para estimar a variância e o impacto dos erros no modelo:

$$SSE = \sum_{i=1}^{2^k} \sum_{j=1}^r e_{ij}^2 \quad (4.5)$$

No projeto de experimentos $2^k r$, a porcentagem de variação explicada por cada fator permite analisar se um fator tem um impacto significativo na resposta. A variação devido a erros experimentais também pode ser isolada, uma vez que existem várias observações para cada combinação de nível de fator. A variação total ou Soma Total dos Quadrados (SST) é dada por:

$$SST = \sum_{ij} (y_{ij} - Y_i)^2 \quad (4.6)$$

Na tabela 6 é possível ver o resultado obtido pelo cálculo da SST, bem como outras métricas utilizadas para avaliar os efeitos dos fatores no modelo experimental, conforme explicado em (JAIN, 1992) no capítulo referente ao projeto fatorial 2^k com repetições.

Tabela 6 – Alocação de Variação

Métricas	Resultados
SST	0,24967
SSA	0,00129
SSB	0,00020
SSC	0,22992
SSAB	0,00053
SSAC	0,00147
SSBC	0,00021
SSABC	0,00309
SSE	0,01296
SSO	9,29305
SSY	9,54272

O SST pode ser dividido em oito partes da seguinte forma:

$$SST = SSA + SSB + SSC + SSAB + SSAC + SSBC + SSABC + SSE \quad (4.7)$$

Cada SS (soma do quadrado) diferente corresponde a um dos efeitos na expressão acima. Assim, SSA, SSB, SSC, SSAB, SSAC, SSBC e SSABC são as variações explicadas pelos fatores A, B, C e interações AB, AC, BC e ABC respectivamente. O SSE é a variação inexplicada atribuída aos erros experimentais.

A variação total também pode ser calculada da seguinte forma:

$$SST = SSY - SSO \quad (4.8)$$

onde o SSO representa a soma dos quadrados das médias em Y, e SSY representa a soma dos quadrados dos valores mensurados pelas repetições dos experimentos. Esses valores são mostrados na Tabela 6.

Na Tabela 7 mostra como cada fator participa no cálculo do resultado final estimado pelo modelo linear para os experimentos.

Tabela 7 – Fração da variação explicada

Fator	Varição	%
A	0,005	0,52
B	0,001	0,08
C	0,921	92,09
AB	0,002	0,21
AC	0,006	0,59
BC	0,001	0,08
ABC	0,012	1,24
Erro	0,052	5,19

Como cada efeito participa na variação do resultado final estimado pelo modelo é calculado por:

$$\text{Varição explicada pelo efeito } F = \frac{SS_F}{SST} \quad (4.9)$$

onde $F \in \{A, B, C, AB, AC, BC, E\}$ e SST é a soma total dos quadrados, todos esses valores mostrados na tabela 6.

4.2.2 Avaliação da Solução

Mediante aos cálculos realizados neste experimento e após calcular os impactos que cada fator apresentou na solução final, é possível avaliar a solução e analisar os resultados encontrados.

Feitos os cálculos pode-se afirmar que o impacto causado pelo fator C referente a Quantidade de Classes foi muito superior aos demais fatores, uma vez que na Tabela 6 é possível ver uma variação média do resultado de 92,09% quando o fator é aplicado. Este resultado indica que existe uma dificuldade da CNN em diferenciar imagens de anomalias malignas de benignas. Contudo, ao agregar ambas as classes benignas e malignas como anomalias os resultados se tornam melhores quando comparados com imagens de mamografias do tipo normal. Por outro lado, pode-se afirmar que os demais fatores apresentam impactos modestos no resultado final sobre o valor da acurácia.

O fator A referente a Tamanho do Recorte da Imagem possui uma variação no resultado da acurácia de 0,52%, o que aponta um ganho reduzido de precisão quando aumentado o tamanho da mesma. O fator B, referente a Quantidade de Camadas apresenta ganho também reduzido no resultado da acurácia, atingindo somente uma variação de 0,08%. Para os fatores de interação é notável que não há dependência dos mesmo entre si,

pois os resultados das combinações não justificam os resultados individuais de cada fator, como é possível notar ao comparar o fator C com os fatores de interação que o incluem como AC, BC e ABC. Por fim, o fator erro apresenta uma variação pequena de 5,19% no resultado, indicando que o modelo linear construído é capaz de explicar bem o problema estudado.

Na tabela 7 é possível ver o impacto na melhora da acurácia causada por cada fator principal e de interação.

Na tabela 8 é apresentado o resultado das melhoras individuais causadas por cada fatores principal.

Tabela 8 – Resultado das melhoras de cada fator

Fatores	Melhora(%)
Quantidade de Classes	92,09%
Quantidade de Camadas	0,08%
Tamanho do Recorte da Imagem	0,52%

A quantidade classes apresentou, dentre os demais, o melhor impacto na acurácia. Como as imagem benignas e malignas apresentavam diferenças imperceptíveis, a rede neural convolucional encontrou dificuldade em classificá-las. Com três classes os resultados mostraram que a rede neural convolucional tinha dificuldade em diferenciar Benigno e Maligno, e quando essas duas classes eram agregadas em uma única classe chamada de Anomalia, a rede neural convolucional conseguia distinguir melhor essas imagens das imagens contidas na classe Normal.

A quantidade de camadas apresentou dentre os demais fatores o pior impacto na acurácia. Apesar da adição de apenas uma camada, o impacto que esse fator teve no tempo de treinamento para uma melhora mínima na acurácia demonstra não ser uma alternativa muito viável para melhora da CNN em questão, desta forma para essa arquitetura a suspeita de Zhou (2016) de que o aumento de camadas convolucionais melhoraria a acurácia se mostrou falso para a arquitetura em questão, conforme mostrado pelos experimentos realizados.

O Tamanho do Recorte da Imagem, o ultimo fator a ser discutido, apresentou um impacto baixo na acurácia, apesar do aumento da imagem não ter sido muito maior da que foi proposto no artigo de Zhou (2016). Notou-se que, assim como a quantidade de camadas, o aumento da imagem impactava em um tempo muito alto de resposta e às vezes em estouro de memória. Como já visto, dobrar o Tamanho do Recorte da Imagem apresentada no artigo não causou melhoria significativa nos resultados e, sendo assim, o tempo que se gasta para um ganho quase imperceptível provou que o Tamanho do Recorte da Imagem não é um fator vantajoso em seus resultados.

Como mencionado, quando comparamos o impacto gerado pelos fatores principais com os fatores de interação descritos na tabela 7 o maior impacto é o de C, referente a Quantidade de Classe, apresentando assim a melhor configuração, enquanto que quando o fator C é somado a outros fatores como AC, BC e ABC o impacto no resultado é reduzido. Isso mostra que o fator C não é justificado por sua interação, ou seja, seu impacto mais significativo é apresentado isoladamente e não depende de outros fatores.

Infelizmente devido a escassez de tempo para a execução do presente trabalho não foi possível realizar um teste de hipótese, que ajudaria a confirmar a eficácia das medidas de melhoria adotadas, assim como não foi realizado um teste para verificar *overfitting* na CNN.

5 Conclusão

Através da análise dos resultados dos experimentos foi possível verificar diversos fatores que poderiam impactar a rede neural convolucional de forma positiva ou negativa. No final do experimento, todos os fatores escolhidos apresentaram impactos na melhora da acurácia, apesar de que alguns fatores não obtiveram resultados tão significativos quanto o esperado. Dentre todos os fatores testados, o fator quantidade de classes foi o que apresentou melhor impacto na acurácia sem apresentar nenhum custo significativo de tempo ou memória.

O fator conhecido como quantidade de camadas apresentou uma melhora no resultado não muito significativa, o que acarretava além disso em um alto custo no tempo de treinamento, sendo assim apesar da Tamanho do Recorte da Imagem possuir um crescimento positivo na acurácia tal fator não apresenta um resultado geral vantajoso. Por fim tem-se o fator quantidade de camada que apresentou o pior impacto na acurácia, apesar de ter melhorado a acurácia não se pode dizer que foi uma melhora significativa pois não alcançou nem 1% de melhora. Em contra partida o Tamanho do Recorte da Imagem assim como a quantidade de camada impactou drasticamente no tempo de treinamento da rede, além de que o aumento nas imagens gerava estouro na memória abordando a execução. Sendo assim, o fator tamanho do recorte da imagem assim como o fator quantidade de camadas não apresenta um resultado geral vantajoso.

6 Trabalhos futuros

Acredita-se que separar as imagens apenas como calcificações e não-calcificação poderia maximizar a acurácia da rede neural convolucional. Testar a rede neural convolucional com outros *datasets* também seria importante para medir a questão do *overfitting*.

Referências

- ACADEMY, D. S. *O Que é Visão Computacional?* 2018. Disponível em: <<http://datascienceacademy.com.br/blog/o-que-e-visao-computacional/>>. Citado 2 vezes nas páginas 17 e 24.
- ACADEMY, D. S. *Deep Learning Book*. 2019. Disponível em: <<http://www.deeplearningbook.com.br/>>. Citado 2 vezes nas páginas 22 e 23.
- AL-ANTARI, M. A. et al. A fully integrated computer-aided diagnosis system for digital x-ray mammograms via deep learning detection, segmentation, and classification. *International journal of medical informatics*, Elsevier, v. 117, p. 44–54, 2018. Citado na página 13.
- ALIGER, E. *As Redes Neurais Convolucionais no Deep Learning*. 2019. Disponível em: <<https://www.aliger.com.br/blog/as-redes-neuronais-convolutivas-no-deep-learning>>. Citado na página 24.
- ALOM, M. Z. et al. Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. *Journal of digital imaging*, Springer, v. 32, n. 4, p. 605–617, 2019. Citado 2 vezes nas páginas 13 e 29.
- ARAÚJO, F. H. et al. Redes neurais convolucionais com tensorflow: Teoria e pratica. *SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. III Escola Regional de Informática do Piauí. Livro Anais-Artigos e Minicursos*, Sociedade Brasileira de Computação, v. 1, p. 382–406, 2017. Citado 5 vezes nas páginas 14, 24, 25, 26 e 27.
- BARRETOS, H. do Cancer de. *Mamografia*. 2016. Disponível em: <<https://www.hcancerbarretos.com.br/exames-preventivos/384-paciente/exames-preventivos/1615-prevencao-mamografia>>. Citado 3 vezes nas páginas 15, 16 e 17.
- BRAGA, A. d. P.; CARVALHO, A.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. [S.l.]: Livros Técnicos e Científicos, 2000. Citado 2 vezes nas páginas 18 e 19.
- BRANDIZZI, L. *Visão computacional: O que é? Como funciona?* 2020. Disponível em: <<https://www.serpro.gov.br/menu/noticias/noticias-2020/o-que-eh-visao-computacional>>. Citado na página 18.
- CALAS, M. J. G.; GUTFILEN, B.; PEREIRA, W. C. d. A. CAD e mamografia: por que usar esta ferramenta? *Radiologia Brasileira*, scielo, v. 45, p. 46 – 52, 02 2012. ISSN 0100-3984. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-39842012000100011&nrm=iso>. Citado na página 14.
- CARVALHO, A. C. P. de Leon F de. *Perceptron Multi-Camadas*. 2013. Disponível em: <<https://sites.icmc.usp.br/andre/research/neural/MLP.htm>>. Citado 2 vezes nas páginas 22 e 23.
- CÁRDENAS-MONTE, M. *Sobreajuste - Overfitting*. 2015. Course notes. Acesso em 21 set. 2017. Disponível em: <<http://www.wae.ciemat.es/~cardenas/docs/lessons/sobreajuste.pdf>>. Citado na página 24.

- DALVI, R. *Detecção e Classificação de Câncer a partir de Mamografias Digitalizadas e Redes Neurais Convolucionais*. Dissertação (Mestrado) — Universidade Federal do Espírito Santo, 2018. Citado 4 vezes nas páginas 26, 27, 28 e 29.
- DRAELOS, R. L. B. *The History of Convolutional Neural Networks*. 2019. Disponível em: <<https://glassboxmedicine.com/2019/04/13/a-short-history-of-convolutional-neural-networks/>>. Citado 2 vezes nas páginas 13 e 29.
- FEMAMA. *O câncer de mama em números*. 2019. Disponível em: <<https://www.femama.org.br/2018/br/noticia/o-cancer-de-mama-em-numeros>>. Citado na página 12.
- HAYKIN, S. *Redes Neurais: Princípios e Prática*. Artmed, 2007. ISBN 9788577800865. Disponível em: <<https://books.google.com.br/books?id=bhMwDwAAQBAJ>>. Citado 4 vezes nas páginas 19, 20, 21 e 22.
- INCA. *Cancêr de mama no Brasil*. 2012. Disponível em: <<http://www2.inca.gov.br/wps/wcm/connect/>>. Citado na página 12.
- JAIN, R. *The Art of Computer Systems Performance Analysis*. [S.l.]: Wiley, 1992. Citado 2 vezes nas páginas 36 e 40.
- KARPATHY, A. *Convolutional neural networks for visual recognition*. 2017. Disponível em: <<http://cs231n.github.io/convolutional-networks/>>. Citado na página 26.
- KOMEDA, Y. et al. Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: preliminary experience. *Oncology*, Karger Publishers, v. 93, n. Suppl. 1, p. 30–34, 2017. Citado na página 13.
- LEITE, G. C. A utilização de técnicas de limiarização para auxílio no diagnóstico de câncer de mama. In: *III Encontro Nacional De Engenharia Biomecânica*. [S.l.: s.n.], 2011. Citado 2 vezes nas páginas 12 e 36.
- LIMA, M. A. de C. S. *Sistema Sensorial*. 2009. Disponível em: <<https://mundoeducacao.bol.uol.com.br/biologia/sistema-sensorial.htm>>. Citado na página 18.
- RAMSUNDAR, B.; ZADEH, R. *TensorFlow for Deep Learning: From Linear Regression to Reinforcement Learning*. O'Reilly Media, 2018. ISBN 9781491980408. Disponível em: <<https://books.google.com.br/books?id=GZ1ODwAAQBAJ>>. Citado na página 26.
- RAVINDRA, S. How convolutional neural networks accomplish image recognition? *Mindmajix.com*, sep 2017. Original: <https://www.kdnuggets.com/2017/08/convolutional-neural-networks-image-recognition.html>. Disponível em: <<https://www.infoq.com/br/articles/redes-neurais-convolucionais>>. Citado 4 vezes nas páginas 17, 18, 23 e 24.
- RODRIGUES, V. *Métricas de Avaliação: acurácia, precisão, recall quais as diferenças?* <https://medium.com/@vitorborbarodrigues/m%C3%A9tricas-de-avalia%C3%A7%C3%A3o-acur%C3%A1cia-precis%C3%A3o-recall-quais-as-diferen%C3%A7as-c8f05e0a513c>: [s.n.], 2019. Citado na página 36.
- SESTREM, E. *Um estudo de Redes Neurais Convolucionais e sua aplicação na detecção da malária*. 2019. Disponível em: <<https://bit.ly/2YF9lTV>>. Citado na página 25.

SKAANE, P. Interobserver variation in the interpretation of breast imaging. *Acta Radiol*, p. 497–502, 1997. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/9240666>>. Citado na página 12.

SOUTO, L. P. M. *Mineração de Imagens para a Classificação de Tumores de Mama*. Dissertação (Mestrado) — Universidade Federal Rural do Semi-Árido, <https://ppgcc.ufersa.edu.br/wp-content/uploads/sites/42/2014/09/Disserta%C3%A7%C3%A3o-Lizianne-Priscila-2014.pdf>, 2014. Citado 3 vezes nas páginas 12, 13 e 32.

SOUZA, J. Diagnóstico de câncer de mama a partir de imagens de mamografia 2d utilizando descritores de forma 3d. *Universidade Federal do Maranhão*, feb 2018. Citado 4 vezes nas páginas 15, 16, 32 e 36.

SUCKLING, P. J. The mammographic image analysis society digital mammogram database. *Digital Mammo*, Elsevier Sc. BV, p. 375–386, 1994. Citado na página 32.

SZELISKI, R. Computer vision: Algorithms and applications. *Microsoft Research*, v. 1, sep 2010. Disponível em: <<http://szeliski.org/Book/>>. Citado na página 17.

TAVARES, R.; NASCIMENTO, R. c. 2020. Disponível em: <<https://bit.ly/3ccLn6f>>. Citado na página 20.

ZACCONE, G.; KARIM, M. R.; MENSRAWY, A. *Deep Learning with TensorFlow*. [S.l.]: Packt Publishing Ltd, 2017. Citado 2 vezes nas páginas 26 e 31.

ZHANG, X. Mammogram and tomosynthesis classification using convolutional neural networks. 2017. Citado 2 vezes nas páginas 13 e 29.

ZHOU, H. H. Mammogram classification using convolutional neural networks. In: . [S.l.: s.n.], 2016. Citado 10 vezes nas páginas 13, 29, 30, 31, 32, 33, 34, 35, 36 e 42.