



Ministério da Educação
Secretaria de Educação Profissional e Tecnológica
Instituto Federal Catarinense
Campus Rio do Sul

RODRIGO CACHOEIRA

**CLASSIFICAÇÃO DA ARCADA DENTÁRIA POR MEIO DE
INTELIGÊNCIA ARTIFICIAL: UM COMPARATIVO ENTRE
REDES CONVOLUTIVAS E MULTICAMADAS PERCEPTRON**

Rio do Sul
2018

RODRIGO CACHOEIRA

**CLASSIFICAÇÃO DA ARCADA DENTÁRIA POR MEIO DE
INTELIGÊNCIA ARTIFICIAL: UM COMPARATIVO ENTRE
REDES CONVOLUTIVAS E MULTICAMADAS PERCEPTRON**

Trabalho de Conclusão de Curso apresentado ao
Curso de graduação em Ciência da Computação
do Instituto Federal Catarinense – Campus Rio
do Sul para obtenção do título de bacharel em
Ciência da Computação.

Orientador: Daniel Gomes Soares

Rio do Sul

2018

RODRIGO CACHOEIRA

**CLASSIFICAÇÃO DA ARCADA DENTÁRIA POR MEIO DE
INTELIGÊNCIA ARTIFICIAL: UM COMPARATIVO ENTRE
REDES CONVOLUTIVAS E MULTICAMADAS PERCEPTRON**

Este Trabalho de Curso foi julgado adequado para a obtenção do título de Bacharel em Ciência da Computação e aprovado em sua forma final pelo curso de graduação em Ciência da Computação do Instituto Federal Catarinense – Campus Rio do Sul.

Rio do Sul (SC), 26 de novembro de 2018

Daniel Gomes Soares
Instituto Federal Catarinense – Campus Rio do Sul

BANCA EXAMINADORA

André Alessandro Stein
Instituto Federal Catarinense – Campus Rio do Sul

Juliano Tonizetti Brignoli
Instituto Federal Catarinense – Campus Rio do Sul

RESUMO

A utilização de técnicas computacionais para analisar e classificar imagens nos últimos anos cresceu de forma exponencial, grande parte desse crescimento é devido a evolução de técnicas que utilizam Inteligência Artificial, mais precisamente Redes Neurais Artificiais. O propósito deste trabalho além de demonstrar algumas técnicas utilizadas na análise de imagens é realizar um comparativo em uma situação real do mundo.

O avanço computacional trouxe oportunidades antes impossíveis de serem executadas na área de análise de imagens, uma das principais técnicas que anda ganhando espaço no mundo da IA é a técnica convulacional, esta que possui inspiração no funcionamento da visão humana, diferentemente dos processos definidos pela rede neural perceptron. Com base no que foi dito, o propósito principal desse trabalho é comparar as técnicas Perceptron e Convolutacional na classificação de dentes da arcada dentária, contidas em raios-x, possibilitando então, a inclusão de novas técnicas em diferentes áreas, não apenas suprimindo suas necessidade, mas apresentando melhores resultados de acordo com o surgimento de novas técnicas, como é o caso da rede convolutacional, a qual tem como principal foco a análise de imagens, esta que é uma área de grande valia nas análises radiográficas.

Palavras-chave: Processamento de Imagem. Visão Computacional, Inteligência Artificial, Redes Neurais, Perceptron, Convolutacional, Aprendizagem Profunda.

ABSTRACT

The use of computational techniques to analyze and classify images in recent years has grown exponentially, much of this growth is due to the evolution of computational techniques that use Artificial Intelligence, more precisely Artificial Neural Networks. The purpose of this work besides demonstrating some techniques used in the analysis of images is to perform a comparative in a real situation of the world. In the last decades the computational advancement has brought before impossible possibilities to be executed in the area of image analysis, one of the main techniques that is gaining space in the world of AI is the convulational technique, that has a strong base in the essence of the human vision, in order to to improve and optimize the processing so that, in addition to being more precise, they are better optimized in comparison to other techniques that already exist in the area, as is the case of Perdestron. Based on what has been said, the main purpose of this work is to compare Perceptron and Convolutional techniques in the classification of teeth of the dental arch, contained in x-rays.

key-words: Image Processing. Computer Vision, Artificial Intelligence, Neural Networks, Perceptron, Convolutional, Deep Learning.

LISTA DE ILUSTRAÇÕES

Figura 1 – Raio X Dental no formato Bite Wing	19
Figura 2 – Raio X Dental no formato Periapical	20
Figura 3 – Raio X Dental no formato Panorâmico	20
Figura 4 – Classificação dos dentes por meio do padrão FDI	21
Figura 5 – Tipos de dentes	22
Figura 6 – Número 5 representado por uma matriz	23
Figura 7 – Representação visual da matriz anterior	23
Figura 8 – Neurônio do sistema biológico	25
Figura 9 – Representação do Neurônio Artificial	26
Figura 10 – Rede MLP com duas camadas intermediárias	29
Figura 11 – Comparativo entre um modelo utilizando Deep Learning e uma simples rede neural.	30
Figura 12 – Função Sigmoid	32
Figura 13 – Função Tahn	33
Figura 14 – Função ReLU e Leaky ReLU	34
Figura 15 – Convolução entre um filtro 3x3 e o volume de entrada	37
Figura 16 – Cálculo demonstrativo da camada de convolução	38
Figura 17 – Representação da saída da camada de convolução	39
Figura 18 – Imagem Exemplo antes de aplicar os filtros	39
Figura 19 – Filtro: Edge Detection - Detecção de bordas	40
Figura 20 – Filtro: Outline - Realça a diferença entre pixels vizinhos	40
Figura 21 – Max-Pooling	41
Figura 22 – Flattening	41
Figura 23 – Radiografica Panorâmica do repositório Mendeley	45
Figura 24 – Pré processamento das imagens	46
Figura 25 – Função de Ativação ReLU	50
Figura 26 – Dente Segmentado	50
Figura 27 – Processo de manipulação de imagens	52
Figura 28 – Modelo reduzido 01 MLP de classificação dos tipos de dentes	53
Figura 29 – Camada convulacional 01 de classificação dos tipos de dentes	54

Figura 30 – Modelo reduzido 02 MLP de classificação dos tipos de dentes	55
Figura 31 – Modelo reduzido 03 MLP de classificação dos tipos de dentes	56
Figura 32 – Camada convolucional 04 de classificação dos tipos de dentes	56
Figura 33 – Teste 01 comparativo entre as RNAs	59
Figura 34 – Teste 02 comparativo entre as RNAs	60
Figura 35 – Teste 03 comparativo entre as RNAs	61
Figura 36 – Teste 04 comparativo entre as RNAs	62

LISTA DE TABELAS

Tabela 1 – Trabalhos Correlatos	42
Tabela 2 – Testes por Categoria	57

LISTA DE ABREVIATURAS E SIGLAS

FDI	World Dental Federation
IA	Inteligência Artificial
RNA	Redes Neurais Artificiais
MLP	Multilayer Perceptron
RGB	Red, Green and Blue
GPU	Unidade de Processamento Gráfico
CNN	Rede Neural Convolucional
IDE	Ambiente de Desenvolvimento Integrado
JPEG	Joint Photographics Exprts Group
PNG	Portable Network Graphics
CPU	Unidade Central de Processamento
API	Interface de Programção de Aplacativos
CNN	Convolutional neural networks

SUMÁRIO

1	INTRODUÇÃO	12
1.1	PROBLEMATIZAÇÃO	13
1.1.1	Solução Proposta	13
1.1.2	Delimitação do escopo	14
1.1.3	Justificativa	14
1.2	OBJETIVOS	14
1.2.1	Objetivo Geral	15
1.2.2	Objetivos Específicos	15
1.3	METODOLOGIA	15
1.3.1	Estrutura do Trabalho	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	MODALIDADES DE IMAGENS MÉDICAS	17
2.1.1	Radiografia	18
2.2	RAIO-X DENTAL	18
2.3	CLASSIFICAÇÃO DOS DENTES	20
2.3.1	FDI	21
2.3.2	Tipos de Dentes e suas classificações	21
2.4	VISUALIZAÇÃO COMPUTACIONAL DE UMA IMAGEM	22
2.5	REDES NEURAIS ARTIFICIAIS	24
2.5.1	Neurônios Artificiais	25
2.5.2	Arquiteturas de Redes Neurais Artificiais	26
2.5.3	Processo de Treinamento de uma RNA	27
2.5.4	Redes Multilayer Perceptron (MLP)	28
2.5.5	Método de treinamento de uma MLP	29
2.6	DEEP LEARNING	30
2.6.1	Funções de Ativação	31
2.6.1.1	Sigmoid	31
2.6.1.2	Tanh	32
2.6.1.3	ReLU	33
2.6.1.4	Softmax	34

2.6.2	RNA Convolutacional	34
2.6.2.1	Tipos de Camadas	36
2.6.2.2	Operador de Convolução	36
2.6.2.3	Pooling	40
2.6.2.4	Flattening	41
3	TRABALHOS CORRELATOS	42
3.1	COMPARAÇÃO ENTRE TRABALHOS CORRELATOS	42
3.1.1	Machine Learning for Dental Image Analysis	42
3.1.2	Redes Neurais Convolucionais Aplicadas ao Diagnóstico de Tuberculose por meio de Imagens Radiológicas	43
3.1.3	Caries Detection in Paniarmic Dental X-ray Images	44
4	DESENVOLVIMENTO	45
4.1	DADOS DISPONÍVEIS	45
4.1.1	TRATAMENTO DOS DADOS	46
4.1.2	Normalização dos dados	47
4.2	VALIDAÇÃO CRUZADA	47
4.3	RECURSOS UTILIZADOS	48
4.3.1	Framework de desenvolvimento de RNAs - Keras	48
4.3.2	Funções de Ativação	49
4.3.2.1	Soft Max	49
4.3.2.2	ReLU	49
4.3.2.3	Tangente Hiperbólica - Tahn	51
4.4	Processo Comparativo entre as redes desenvolvidas	51
5	RESULTADOS	57
5.1	Análise por categoria	57
5.2	TREINAMENTO E RESULTADOS DAS RNAS	58
5.2.1	Treinamento	58
5.2.2	Testes	58
5.2.2.1	Teste 01	59
5.2.2.2	Teste 02	59
5.2.2.3	Teste 03	60
5.2.2.4	Teste 04	61

6	CONCLUSÕES	63
	REFERÊNCIAS	65

1 INTRODUÇÃO

A evolução da ciência contemporânea fez com que muitas áreas antes inexploradas comecem a receber sua devida atenção, sendo uma delas o diagnóstico e análise de imagens. Além dos equipamentos apresentarem uma grande melhora para a realização dos diagnósticos por parte dos médicos outra área começa a se tornar mais presente nesse meio, a utilização de computadores no auxílio de diagnósticos, menciona Gallian (2000).

Ginger (2000), define o diagnóstico auxiliado por computador como aquele no qual o radiologista usa os resultados de uma análise computadorizada de imagens médicas como uma “segunda opinião” na detecção de lesões e na elaboração do diagnóstico. Como sendo uma área em ascensão, a análise de imagens médicas utilizando técnicas computacionais apresenta-se aberta para inúmeras variações de pesquisas.

Dentro da área de diagnósticos auxiliados por computador é possível encontrar inúmeras subáreas, este trabalho terá seu foco na análise de imagens de raios X utilizando Redes Neurais Artificiais (RNA) aplicadas a arcada dentária, mais especificamente na classificação dos dentes, classificando-os de acordo com padrão FDI (FDI Word Dental Federation), o qual é um modelo de classificação de dentes utilizando dois dígitos. Estudos prévios realizados na área apresentam modelos de segmentação baseados em visão computacional, utilizando RNAs Convolucionais de acordo com o estudo de Yong-jun Yu (2016) e métodos matemáticos conforme podem ser observados no estudo de Oliveira (2009).

Com a avanço das Redes neurais Artificiais (RNA) e principalmente de técnicas convolucionais que são utilizadas em visão computacional como aponta Yu (2016) em seu estudo, novas formas classificatórias tornam-se viáveis ao mundo da medicina, tendo sua essência na biologia dos mamíferos, mais precisamente em sua capacidade de emulação dos sinais recebidos em seu córtex visual, sendo atualmente umas das melhores técnicas para se aplicar em processamento de imagem.

O propósito deste trabalho é apresentar um comparativo entre dois modelos que utilizam RNAs como método de análise, a fim de determinar o melhor método entre eles para classificação de imagens radiográficas da arcada dentária por meio de reconhecimento de padrões proporcionado pelas técnicas utilizadas de redes neurais artificiais. Também serão abordadas duas técnicas neste trabalho, convolucionais e Multilayer Perceptron (MLP) para que então seja verificado a maturidade das Redes Convolucionais, uma técnica nova para o meio acadêmico, ao

contrário das Redes MLP que já se conceituaram, porém, as Redes Convolucionais em estudos prévios, apresentam-se extremamente promissoras na análise de imagens.

1.1 PROBLEMATIZAÇÃO

Normalmente o processo de análise de imagens médicas são realizadas utilizando inúmeras técnicas de processamento de imagens juntamente com modelos matemáticos, tornando-se um grande e massivo ambiente de estudo para a identificação de objetos em imagens como aponta Nunes (2006). Seguindo este caminho, o estudo de João Paulo Ribeiro de Oliveira (2010) tornou-se um referência, seu objetivo foi detectar cáries dentárias em imagens panorâmicas de Raio X, apesar da eficácia de suas técnicas baseadas em computação gráfica, seus modelos exigem um grande conhecimento técnico.

Por outro lado, as Redes Neurais, mais precisamente Redes Neurais Convolucionais e MLP, tornam-se uma alternativa aos métodos tradicionais de segmentação dos dentes, que em sua maioria utilizam técnicas de processamento de imagens por intermédio de modelos matemáticos. Silva, Spatti e Flauzino (2010), comenta que as redes convolucionais são integrantes dos novos modelos mais recentes e promissores, principalmente pela sua capacidade de realizar mais atividades com uma quantidade significativamente menor de memória quando comparada a redes tradicionais, como por exemplo MLP

Com base no apresentado, surge a pergunta: seria possível realizar a classificação dos tipos de dentes da arcada dentária utilizando RNAs, mais especificamente por meio das técnicas convolucionais, demonstrando-a como alternativa viável em relação sua acuracidade a redes tradicionais como MLP no processamento de imagens?

1.1.1 Solução Proposta

A proposta desse trabalho é desenvolver dois modelos baseados em RNAs Convolucionais e MLP para a classificação dos tipos dentes utilizando imagens de raio X panorâmicas. Será realizado um comparativo da performance da rede neural MLP com e sem a camada convolucional inclusa a ela, verificado a maturidade e viabilidade da utilização de modelos convolucionais sobre os tradicionais MLPs.

1.1.2 Delimitação do escopo

Este trabalho apresentará a criação e modelagem de um modelo utilizando Redes Neurais Artificiais para a classificação dos tipos de dentes de um Raio X panorâmico da arcada dentária. Serão consideradas apenas as arquiteturas Convolucionais e MLP. As imagens utilizadas serão restritas apenas as de raio X panorâmicas, excluindo outras formas de visualização dos dentes.

O algoritmo irá classificar os quatro formatos de dentes, utilizando-se de apenas os 28 dentes da arcada dentária, pois os terceiros molares serão excluídos pelo fato de muitas pessoas não os terem.

As imagens dos dentes segmentados para este estudo estão fixadas ao tamanho de 125 x 124 pixels.

1.1.3 Justificativa

A evolução do campo de processamento de imagens permite o aperfeiçoamento de diagnósticos e tratamento de doenças de naturezas diversas como aponta Nunes (2008). Cada modalidade atua de forma diferenciada, onde algumas técnicas podem ser utilizadas para realizar a classificação da arcada dentária, sendo elas baseadas em Redes Neurais ou simplesmente por intermédio de modelos matemáticos e processamento de imagens. Para que essa análise seja relevante aos profissionais esta deve apresentar grande acuracidade, acima de 90%, logo afim de atingir este patamar, o seguinte trabalho tem como intenção comparar duas redes neurais, sendo que uma delas está apresentando ótimos resultados na análise de imagens com base nos trabalhos pesquisados.

Além disso, os resultados obtidos por meio dessa pesquisa atinge diretamente a área acadêmica, pois apresenta um comparativo relevante para o processamento de imagens como um todo, demonstrando que novas técnicas podem ser implementadas em inúmeras áreas, principalmente na área da saúde.

1.2 OBJETIVOS

Esta seção descreve o objetivo geral e os objetivos específicos do trabalho.

1.2.1 Objetivo Geral

Desenvolvimento de duas redes neurais artificiais, MLP e Convolucional utilizando técnicas de Inteligência Artificial para a classificação dos tipos de dentes contidos em um Raio X, comparando a Rede Neural Artificial Perceptron com a Rede Neural Artificial Convolucional na classificação de dentes em Raios X panorâmicos.

1.2.2 Objetivos Específicos

Os objetivos específicos desse trabalho são:

1. Obter um conjunto de dados para realizar os testes e treinamentos da RNA;
2. Modelar a RNA Convolucional e ML;
3. Segmentar os dentes manualmente, utilizando ferramentas de recorte de imagens;
4. Realizar testes utilizando diferentes técnicas e modelos de RNAs e por fim, compará-las.

1.3 METODOLOGIA

Esta seção apresenta a metodologia empregada para o desenvolvimento do trabalho.

1. Levantamento bibliográfico: Está etapa fundamenta o conhecimento apresentado neste trabalho acerca da análise e processamento de imagens por meio de Redes Neurais Convolucionais e MLP;
2. Modelagem das RNAs: Está etapa apresenta a modelagem das RNAs que serão utilizadas para a classificação dos tipos dentes em imagens de raio X panorâmicas;
3. Implementação do Modelo: Está etapa apresenta a implementação dos modelos de identificação da RNA Convolucionais e MLPs;
4. Testar a eficácia com diferentes modelos: Está etapa apresenta os testes utilizando as RNAs com diferentes configurações;
5. Comparação entre as redes: Está etapa apresenta a comparação entre os modelos da RNA Convolucional e MLP;
6. Análise dos Resultados: Por fim será apresentado os resultados obtidos juntamente com sugestões de melhorias e trabalhos futuros.

1.3.1 Estrutura do Trabalho

No capítulo 2 deste trabalho é apresentada a fundamentação teórica, a qual aborda os conceitos utilizados no decorrer do trabalho, como Redes Neurais Artificiais, modelos matemáticos para processamento de imagens, Redes Convolucionais, etc. O capítulo 3 apresenta os dados disponíveis e as especificações que devem ser seguidas para segmentar os dentes do raio X. No capítulo 4 o processo de modelagem da Rede Neural tem seu início. No capítulo 5 será realizada a comparação com outros métodos de segmentação e será realizada a análise dos resultados e sugestões de trabalhos futuros seguido pelo capítulo 6, onde serão expostas as conclusões finais da pesquisa realizada.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, serão abordados conceitos relevantes para o entendimento do projeto a ser desenvolvido, conceitos estes que se apresentam tanto no conhecimento da arcada dentária e sobre o método de raio X como também dos estudos das Redes Neurais, mais especificamente convolucionais.

Ao primeiro momento, serão apresentadas informações pertinentes relacionadas a matéria de estudo deste trabalho, os raios-X contendo as imagens da arcada dentária e suas características.

Por fim, será apresentada a Rede Neural Convolucional, incluindo suas características, inspiração biológica e os cálculos que envolvem sua implementação.

2.1 MODALIDADES DE IMAGENS MÉDICAS

NUNES (2010) salienta que em 1895, um físico alemão chamado Röntgen descobriu a existência de uma radiação até então desconhecida, essa onda apresentou-se extremamente capaz de atravessar materiais, sendo parcialmente ou totalmente absorvida por eles, devido a este fato, tornou-se possível a visualização de estruturas internas do corpo e diagnosticar anomalias.

Devido à complexidade anatômica do corpo humano, diferentes métodos de captura foram aplicados com o passar dos anos, cada modalidade médica atua de forma diferenciada, seus métodos visam ressaltar as diferenças entre os diversos tipos de tecidos existentes, fazendo com que os tecidos de maior importância sejam mais observáveis que os demais.

De acordo com Viergever et al. (2016), as modalidades médicas podem ser classificadas em duas categorias globais: anatômicas e funcionais. As modalidades anatômicas são responsáveis pelas identificações de morfologias, como por exemplo: raio-X, tomografia computadorizada (TC), ressonância magnética nuclear (RMN), ultra-som (US), entre outras. As modalidades funcionais por sua vez, são as que possuem o objetivo de adquirir informações a respeito do metabolismo relacionado a anatomia, incluindo cintilografia, SPECT (Single photon emission computed tomography), modalidades de medicina nuclear, RMN funcional, entre outras.

2.1.1 Radiografia

Segundo NUNES (2010), na radiografia, a formação das imagens é obtida devido a interação dos fótons de raios-X com a matéria. Ao atravessarem o objeto, os fótons sofrem atenuações, devida a sua absorção total ou parcial. O nível de absorção depende dos elementos que constituem o objeto de análise. Dentro do corpo humano, ossos, músculos e gorduras possuem uma taxa de absorção de fótons diferentes, fato este que faz com que surjam diferentes escalas de cinza devido a quantidade de absorção individual para cada aspecto da anatomia humana. Essa diferença que possibilita a formação do contraste nas radiografias, ou seja, resulta em diferentes tons de cinza em uma imagem, permitindo com que seja possível identificar as estruturas anatômicas, e em muitos casos anomalias que são resultados de alguma doença em específico.

Um dos principais problemas identificados por especialistas na radiografia nos tempos atuais é o tempo de exposição dos pacientes. Os raios-x são constituídos de radiação ionizante e a exposição excessiva pode trazer consequências a saúde. Devido a isso, deve-se tomar cuidado para que a dose de exposição dos pacientes não torne-se mais prejudicial a saúde do que o benefício ao qual ela se propõe. Alguns fatores fazem com que a exposição se torne mais prejudicial, como por exemplo o aumento da tensão e o aumento da corrente para a produção dos fótons, e a prolongação da exposição dos pacientes.

No ramo da radiografia, com o decorrer das décadas inúmeras formas de imagens foram sendo desenvolvidas, em muitos dos casos, há diferentes ângulos a serem tratados para a mesma estrutura anatômica do corpo humano, essa variedade se deve a um problema enfrentado pela medicina, o efeito Heel. Dance et al. (2014) destaca que a radiação não é uniforme em todas as regiões, variações de ângulos durante a coleta da imagem podem ocasionar diferentes contrastes para o mesmo objeto, tendo em mente esse fato, diferentes formas de gerar a imagem para o mesmo local anatômico surgem, cada qual com o seu devido propósito.

2.2 RAIOS-X DENTAL

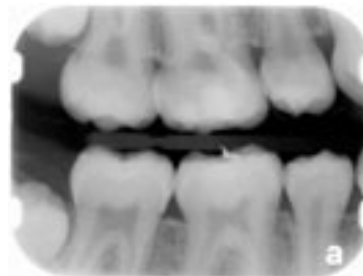
Raios-x dental, como o próprio nome indica, são imagens da arcada dentária, em alguns casos completa, e em outros apenas partes dela. Esses raios-X contemplam não apenas os dentes, mas também seus arredores, sua principal função é detectar problemas relacionados aos dentes, boca e mandíbula. Oliveira (2009) em sua tese, esses tipos de imagens podem ser utilizadas para visualizar cavidades, estruturas dentais escondidas e perda óssea, que em muitos

dos casos, não podem ser detectadas com um simples exame visual.

Como mencionado anteriormente, há inúmeras perspectivas que podem ser utilizadas ao efetuar a análise radiográfica, neste caso em específico, os raios-x dentais podem ser classificados em: Panorâmico, Periapical e Bite wing. As formas Periapical são um conjunto de imagens, onde cada qual contempla dois ou três dentes e sua extensão total. Já a Bite Wing apresenta apenas um lado da arcada dentária por imagem, normalmente divididas em esquerda e direita, porém ao contrário da Periapical ela contempla em cada imagem a parte superior e inferior da mandíbula. Por fim o formato Panorâmico, contempla em apenas uma imagem toda a arcada dentária, parte superior e inferior, uma visão geral. A seguir, são apresentadas algumas figuras que demonstram os diferentes formatos de raios X dentais.

A figura 01 apresenta o formato Bite Wing, este que é responsável por apresentar o formato da "mordida" da arcada dentária, ou seja, se preocupa em apresentar a parte superior e inferior da mandíbula.

Figura 1 – Raio X Dental no formato Bite Wing



Fonte – CERO (2017).

A figura 02 apresenta o formato Periapical, o qual possui como principal finalidade mostrar todo o formato do osso que contempla os dentes.

Figura 2 – Raio X Dental no formato Periapical



Fonte – CERO (2017).

Por fim, a figura 03 apresenta o formato que será utilizado no decorrer do trabalho, as imagens panorâmicas, estas que possuem como principal finalidade mostrar toda a extensão da arcada dentária.

Figura 3 – Raio X Dental no formato Panorâmico



Fonte – CERO (2017).

2.3 CLASSIFICAÇÃO DOS DENTES

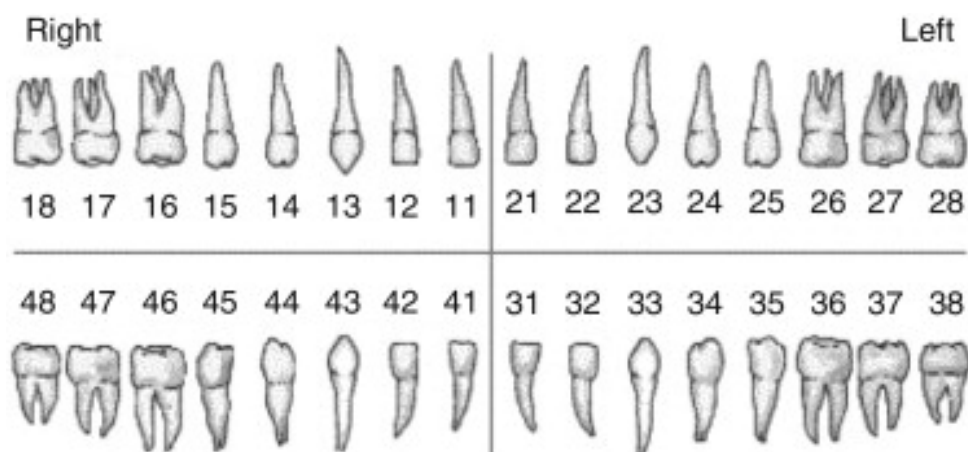
Como apresentado anteriormente, há inúmeras formas de visualização da arcada dentária, principalmente por meio de Raios X, porém em um contexto mais organizacional e

relacional com o paciente torna-se interessante a apresentação e categorização dos dentes de uma forma mais clara e visível para leigos. Problema este que é solucionado por meio dos odontogramas, principalmente os que seguem o padrão FDI.

2.3.1 FDI

Na odontologia algumas formas de classificação dos dentes são utilizadas, a mais comum entre elas é o formato FDI (FDI World Dental Federation). No processo de classificação FDI, cada um dos 32 dentes da arcada dentária é classificado por meio de dois dígitos, o primeiro dígito indica o quadrante que pode variar de 1 a 4, enquanto o segundo dígito indica o tipo do dente, este que pode variar de 1 até 8. A figura 4 apresenta os dentes numerado individualmente conforme o padrão FDI.

Figura 4 – Classificação dos dentes por meio do padrão FDI



Fonte – MOURÃO (2014).

2.3.2 Tipos de Dentes e suas classificações

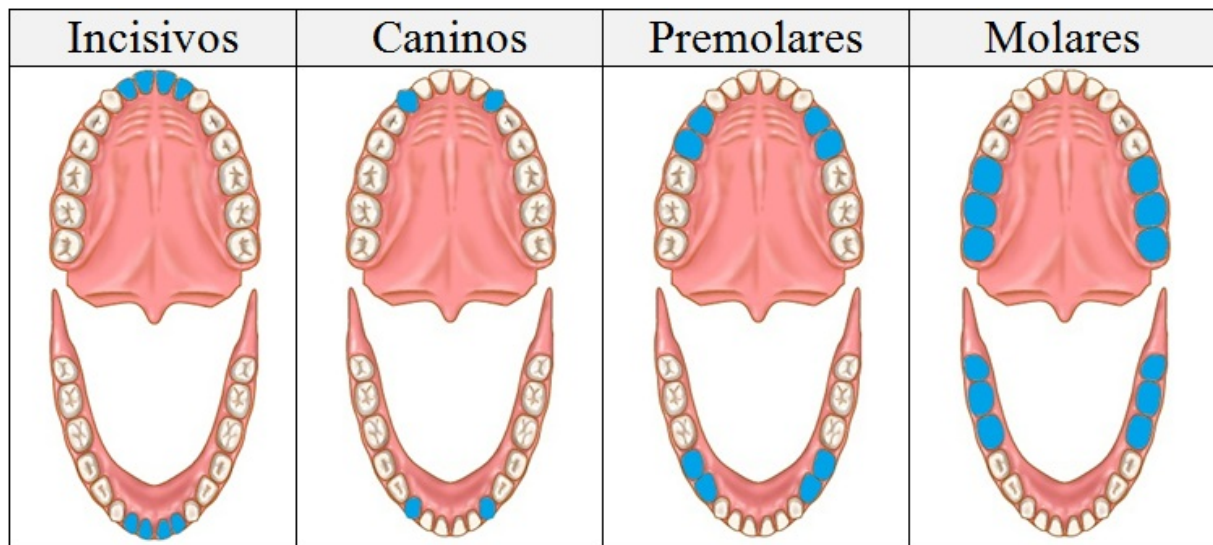
Como apresentado anteriormente, a arcada dentária possui ao todo 32 dentes, estes que podem ser divididos em 4 categorias. Segundo o site oficial (COLGATE, 2018), cada dente tem uma função ou tarefa específica diferente, sendo divididos em:

- Incisivos: Dentes frontais afiados em forma de cinzel (quatro superiores, quatro inferiores) para cortar os alimentos.
- Caninos - Dentes com pontas agudas que rasgam os alimentos

- Pré-molares - Dentes com duas pontas na superfície para esmagar e moer os alimentos
- Molares - Trituram os alimentos, estes dentes possuem várias pontas na superfícies da mordida

A figura abaixo ilustra a posição de cada tipo dente e sua respectiva quantidade.

Figura 5 – Tipos de dentes



Fonte – COLGATE (2018).

Com base no que foi apresentado até agora, é visível a complexidade computacional envolvida no processo da leitura de imagens, principalmente imagens médicas, nas quais pequenos detalhes podem fazer toda a diferença ao apresentar os resultados. Para melhor entendimento de como irá funcionar o processo de extração das características dos dentes, o próximo tópico irá abordar como são feitas as análises de forma computacional para identificar formatos nas imagens.

2.4 VISUALIZAÇÃO COMPUTACIONAL DE UMA IMAGEM

De acordo com Yu (2017), as imagens analisadas de formas computacionais normalmente são convertidas em matrizes bidimensionais, sendo as linhas horizontal representando os pixels horizontais e as linhas verticais os pixels verticais, onde cada pixel possui um valor específico, o qual representa sua intensidade ao renderizar as imagens. Um dos tipos mais comuns de imagens são as em escala de cinza, onde cada pixel varia de preto até branco.

Para Yu (2017), outro tipo comum são as imagens em RGB, onde as cores são divididas em três camadas, cada uma representando a intensidade de uma cor, vermelho, verde e azul, logo quando combinadas formam a cor que representa a posição na imagem.

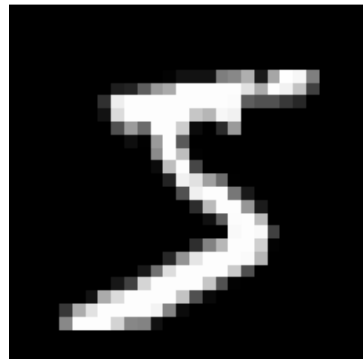
Um exemplo claro para a representação desse processo de conversão de uma matriz em uma imagem, pode ser dada pelas figuras a seguir, a qual apresenta uma matriz simples que realiza a definição de um número e abaixo este é apresentado utilizando um algoritmo qualquer para representação visual.

Figura 6 – Número 5 representado por uma matriz

	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
6	0	0	0	0	30	36	94	154	170	253	253	253	253	253	225	172	253	242	195	64	0	0	0
7	0	0	0	49	238	253	253	253	253	253	253	253	253	251	93	82	82	56	39	0	0	0	0
8	0	0	0	18	219	253	253	253	253	253	198	182	247	241	0	0	0	0	0	0	0	0	0
9	0	0	0	0	80	156	187	253	253	205	11	0	43	154	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	14	1	154	253	90	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	139	253	190	2	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	11	190	253	70	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	35	241	225	160	108	1	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	81	240	253	253	119	25	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	45	186	253	253	150	27	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	16	93	252	253	187	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	249	253	249	64	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	46	130	183	253	253	207	2	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	39	148	229	253	253	253	250	182	0	0	0	0	0	0	0
20	0	0	0	0	0	0	24	114	221	253	253	253	253	201	78	0	0	0	0	0	0	0	0
21	0	0	0	0	23	66	213	253	253	253	253	198	81	2	0	0	0	0	0	0	0	0	0
22	0	0	18	171	219	253	253	253	253	195	80	9	0	0	0	0	0	0	0	0	0	0	0
23	55	172	226	253	253	253	253	244	133	11	0	0	0	0	0	0	0	0	0	0	0	0	0
24	136	253	253	253	212	135	132	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Fonte – Acervo do Autor, 2018

Figura 7 – Representação visual da matriz anterior



Fonte – Acervo do Autor, 2018

Para melhorar o entendimento foi adicionado uma tonalidade aos elementos da matriz conforme o valores se aproxima de 255 para melhorar a percepção da imagem se formando.

A matriz possui dimensão de 28x28, com valores entre 0 e 255, quanto maior o valor, mais se aproxima da intensidade branca.

O exemplo apresentado possui características extremamente simples de serem detectadas por uma análise computacional, por possuir formas bem definidas e relativamente simples em relação a imagens mais complexas, como é o caso dos dentes segmentados que irão ser analisadas neste trabalho. Tendo em vista esse grau de complexidade quando se trata de imagens complexas, como é o caso da arcada dentária, foi optado pela utilização de redes neurais artificiais para auxiliar no processo de detecção de características de imagens, os tópicos a seguir irão abordar seus conceitos básicos e explicar algumas técnicas que serão utilizadas no objetivo deste trabalho.

2.5 REDES NEURAIS ARTIFICIAIS

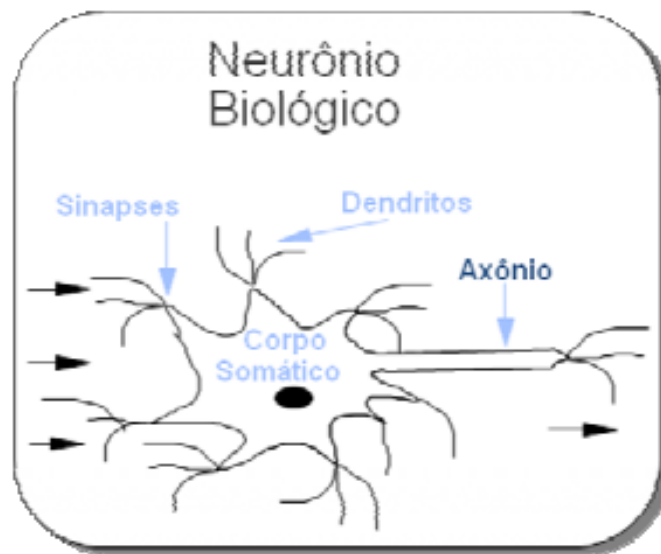
Inteligência Artificial (IA) ou também chamada de Inteligência Computacional é uma das ciências mais recentes atualmente, em seu livro Goldschmidt (2014), define IA como uma tentativa de simular os pensamentos humanos, fazendo com que seja possível que máquinas interajam com seu ambiente de forma semelhante aos humanos. Goldschmidt também comenta que a tentativa de simulação não abrange completamente todo o conceito de IA, pois além dessa perspectiva, essa nova ciência também procura construir entidades artificiais inteligentes.

Sendo uma ciência relativamente nova, o estudo sobre suas possibilidades gerou inúmeras vertentes, cada qual buscando suas respectivas inspirações, como por exemplo Filosofia, Matemática, Psicologia entre outras. Uma destas vertentes gerou as Redes Neurais Artificiais (RNA), estas que segundo (GOLDSCHIMIDT, 2014) são modelos computacionais não lineares, inspirados na estrutura e no funcionamento do cérebro, que procuram reproduzir características humanas, tais como: aprendizado, associação, generalização e abstração. Os modelos desenvolvidos utilizando essa técnica possuem a capacidade de adquirir, armazenar e utilizar conhecimento experimental e buscam simular computacionalmente habilidades humanas.

Para que seja possível uma melhor compreensão do funcionamento das RNAs torna-se necessário o entendimento do funcionamento dos neurônios. Braga, Carvalho e Ludermir (2007) explicam que os neurônios entram em contato através de impulsos nervosos, estes que realizam a transmissão de informações entre eles. A troca de informação é realizada por meio de sinapses, uma região na qual ocorre o contato entre os neurônios. Quando um neurônio recebe um determinado impulso suas informações são processadas e ao atingir um dado limiar

de ação, o neurônio dispara, produzindo uma substância neurotransmissora que flui do corpo celular para o axônio, este por sua vez pode estar conectado a outro neurônio ou a um dendrito, estas conexões fazem com que as redes neurais biológicas surjam.

Figura 8 – Neurônio do sistema biológico



Fonte – Tools (2015)

Braga, Carvalho e Ludermir (2007), salientam que em sua clara inspiração as características dos neurônios citadas anteriormente, as RNAs são compostas por várias unidades de processamento, cujo funcionamento é bastante simples. Normalmente estas unidades estão interconectadas por canais de comunicação, nos quais são trafegadas informações que realizam ajustes em cada unidade, para que então realize um comportamento inteligente.

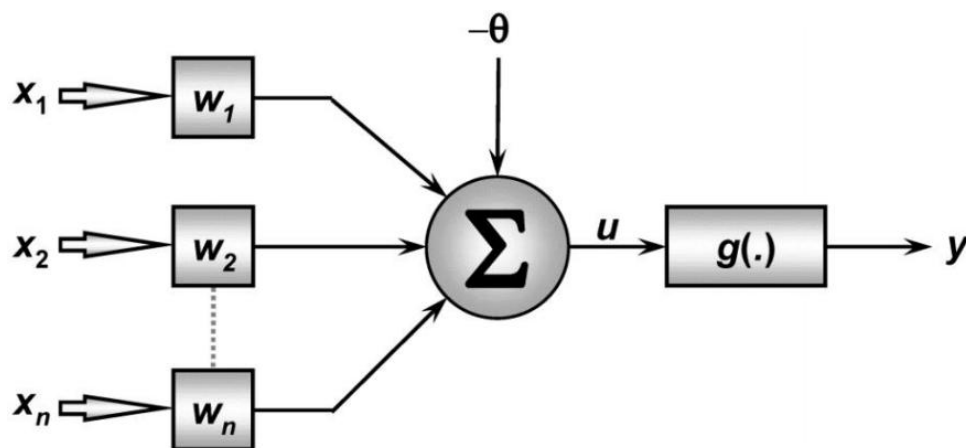
2.5.1 Neurônios Artificiais

Seguindo os mesmos conceitos que as redes neurais biológicas as RNAs são compostas por diversos neurônios artificiais, porém em uma escala menor as biológicas, tanto em quantidade de neurônios como em seu tamanho e complexidade. Braga, Carvalho e Ludermir (2007), apresentam o modelo do neurônio artificial (Figura 6) proposto por McCulloch e Pitts, este que é uma representação simplificada do que se sabia dos neurônios biológicos na época. Sua implementação baseia-se em n terminais de entrada (dendritos) que recebem os valores x_1 ,

x_2, \dots, x_n (que representam as ativações dos neurônios anteriores) e apenas um terminal de saída y , que representa o axônio de um neurônio biológico.

Além disso, Braga, Carvalho e Ludermir (2007) reitera que para a completa apresentação do comportamento das sinapses são necessárias com que as entradas dos neurônios estejam acopladas com pesos w_1, w_2, \dots, w_n , cujos valores podem ser tanto positivos quanto negativos.

Figura 9 – Representação do Neurônio Artificial



Fonte – Silva, Spatti e Flauzino (2010)

No modelo do neurônio artificial apresentado, a ativação do neurônio se dá por meio de uma “Função de Ativação”, seu resultado indica se a saída será ativada ou não, essa decisão ocorre em decorrência da soma ponderada das entradas. Portanto a soma ponderada das entradas x_1, x_2, \dots, x_n juntamente com os pesos sinápticos w_1, w_2, \dots, w_n aplicados a função de ativação irá gerar a saída y do neurônio. Isso ocorre para que seja possível limitar a saída do neurônio dentro de um intervalo de valores razoáveis a serem assumidos pela imagem da função Silva, Spatti e Flauzino (2010).

2.5.2 Arquiteturas de Redes Neurais Artificiais

As RNAs possuem diferentes tipos de arquiteturas, cada uma com suas principais características, sua arquitetura define a forma como os neurônios estão dispostos na rede neural. A coletividade dos neurônios se faz necessária pois como indica Braga, Carvalho e Ludermir (2007), neurônios individuais possuem uma capacidade computacional limitada. Há inúmeras

variações de redes neurais desenvolvidas, porém para este estudo apenas as principais serão citadas, segundo Silva (2010, p.46-50) elas podem ser divididas em:

- Arquitetura feedforward de camada simples: Possui apenas uma camada de entrada de sinais e de neurônios, está que também é a camada de saída. Alguns exemplos para essa camada são: Adaline e Perceptron;
- Arquitetura feedforward de múltiplas camadas: Além das convencionais camadas de entrada e de saída, existem também, uma ou mais camadas ocultas (ou escondidas) de neurônios. Esta arquitetura inclui as redes Multilayer Perceptron (MLP) e as redes de base radial;
- Arquitetura recorrente ou realimentada: Nesta arquitetura, o sinal dos neurônios de saída, são utilizados como entrada para outros neurônios, exemplos dessa rede são: Hopfield e redes auto regressivas com entradas exógenas (NARX);
- Arquitetura em estrutura reticulada: para essa arquitetura a disposição dos neurônios está diretamente relacionada ao ajuste dos pesos. Por exemplo, a rede Kohonem.

2.5.3 Processo de Treinamento de uma RNA

Segundo Braga, Carvalho e Ludermir (2007), a propriedade mais importante das redes neurais é a habilidade de aprender com o ambiente fornecido e com isso melhorar seu desempenho. Isso é feito por meio de processos iterativos de ajustes aplicando pesos, processo desse que é caracterizado como treinamento. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

Existem inúmeros algoritmos ou métodos de aprendizado que variam com seus modelos de redes neurais, estes métodos diferem entre si principalmente pelo modo como os pesos são modificados. Outro fator importante é a maneira pela qual uma rede neural se relaciona com o ambiente. Com base no contexto apresentado, existem os seguintes paradigmas de aprendizado: Aprendizado Supervisionado, Aprendizado não supervisionado, reforço, modo padrão, modo batch.

No treinamento supervisionado, a Rede Neural é treinada com o auxílio de um supervisor. A rede necessita de pares de entradas e saída, ou seja, cada conjunto de entradas devem possuir um conjunto de saídas. Sempre que uma entrada for informada à rede sua saída deve ser verificada (gerada por meio dos cálculos efetuados a partir dos pesos que a rede possui)

para que seja verificada se ela pertence ao conjunto de saídas desejadas para aquela entrada. Caso a saída seja indesejada a rede deverá reajustar os pesos de forma que armazene o conhecimento desejado. Esta interatividade do treino deverá ser repetida com todo conjunto de treinamento, até que a taxa de acerto esteja dentro da faixa considerada satisfatória Braga, Carvalho e Ludermir (2007).

Diferentemente, no treinamento não supervisionado o agente externo não existe, logo não se sabe qual o resultado desejado para o conjunto de entrada. Para que esse modelo seja possível, a rede trabalha as entradas e as organiza de modo que acabe classificando-as usando, para isso, seus próprios critérios. Esse tipo de rede usa os neurônios como classificadores, e os dados de entrada, como elementos para classificação.

2.5.4 Redes Multilayer Perceptron (MLP)

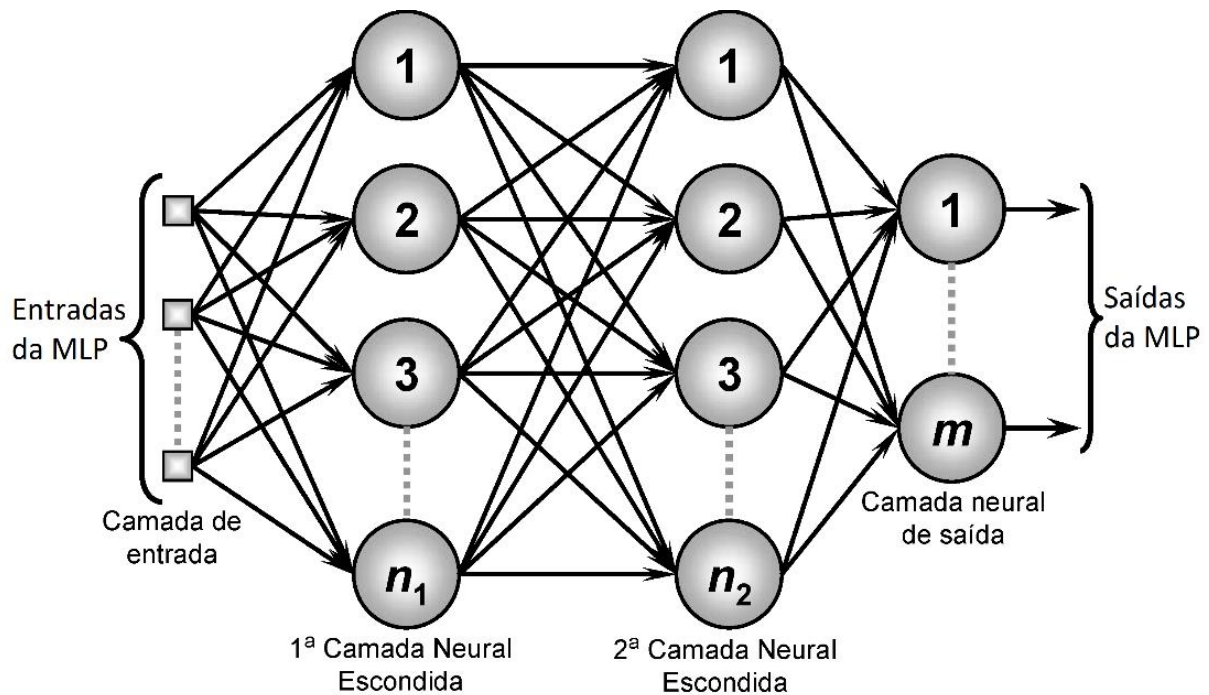
Segundo Silva, Spatti e Flauzino (2010) as Redes Multilayer Perceptron (MLP) fazem parte da arquitetura feedforward, logo ela possui no mínimo uma camada intermediária por característica que ficam localizadas entre a camada de entrada e de saída da RNA. Um dos principais motivos que faz com que as redes MLP sejam normalmente mais utilizadas que redes como Adaline ou a própria rede Perceptron de camada única é o fato de que ela pode ser utilizada em problemas que não sejam linearmente separáveis, caso este que ocorre com frequência no mundo da computação.

A definição do número de neurônios em cada camada de uma rede MLP é de extrema importância para o desempenho da rede, principalmente na sua capacidade de generalização. Quanto maior o número de neurônios, maior é a capacidade da rede em resolver problemas de determinada complexidade. Embora não haja regra geral para definição precisa deste número, existem diversas heurísticas passíveis de aplicação. Uma delas, proposta por Hecht-Nielsen (1988) sugere a utilização de $(2n+1)$ neurônios na camada intermediária, onde n é o número de neurônios na camada de entrada Goldschmidt (2014).

É possível visualizar a arquitetura das redes MLPs por meio da Figura 7, para esta arquitetura os sinais são apresentados à rede em sua entrada, as camadas intermediárias têm a função de obter informações referentes ao comportamento da rede e codificar tais informações em forma de pesos sinápticos (SILVA, 2010). Normalmente o processamento da rede neural é definido pelo processamento que cada neurônio realiza, para então passar os dados processados para a próxima camada, fazendo então com que seja construído uma cadeia de processamento,

fazendo com que se crie um padrão de resposta na última camada, a camada de saída. Logo todo e qualquer neurônio que faz parte da rede neural interfere no resultado que é obtido.

Figura 10 – Rede MLP com duas camadas intermediárias



Fonte – Silva, Spatti e Flauzino (2010)

2.5.5 Método de treinamento de uma MLP

De acordo com Braga, Carvalho e Ludermir (2007), o algoritmo mais comum utilizado pela rede MLP é o backpropagation, ou também conhecido como algoritmo de retro propagação, no qual é um método de treinamento supervisionado que utiliza pares de entradas e saída para ajustar os pesos da rede por meio de um mecanismo de correção de erros. Normalmente o processo de treinamento ocorre em duas fases, a primeira dela ocorrendo a etapa chamada de forward seguida pela backward.

Na fase forward, Braga, Carvalho e Ludermir (2007) afirmam que o vetor de entrada X é apresentado às entradas da rede e os sinais dessa amostra são propagados por cada camada até chegar a sua última camada fornecendo o resultado da rede. O passo seguinte é gerar o erro, por meio da comparação dos resultados obtidos com os resultados esperados. Para Silva, Spatti e Flauzino (2010), faz-se uso dos erros cálculos na fase anterior para alterar os pesos da

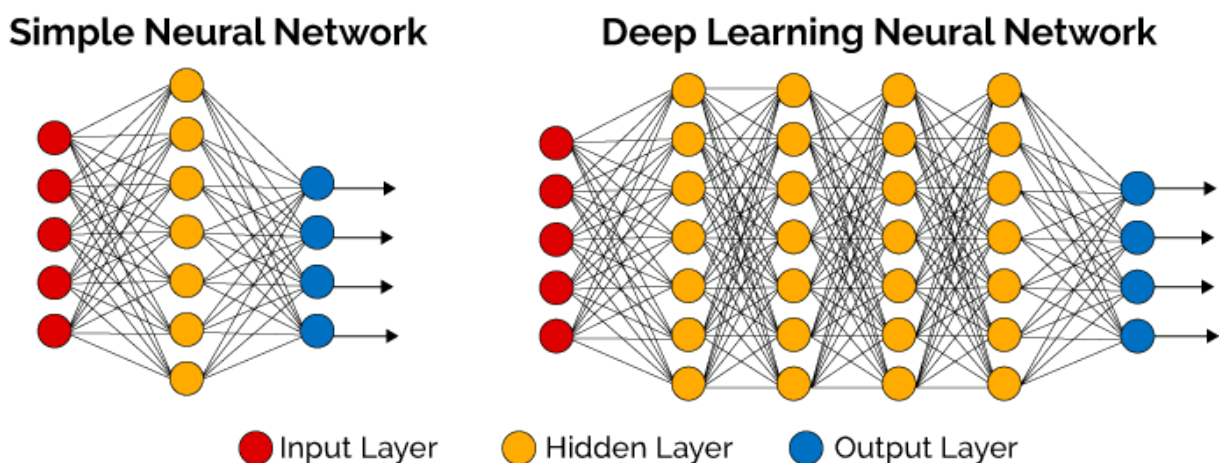
RNA, fazendo com que a cada etapa de treinamento o erro venha a diminuir, ou seja, a rede se aproxima ao melhor resultado esperado.

Para Silva (2010, p.99), o erro aplicado as RNAs pode ser obtido por meio do Erro Quadrático Médio (MSE – Mean Squared Error) que pode ser definido por:

2.6 DEEP LEARNING

Segundo a Academy (2018), a Deep Learning, ou também chamada de aprendizagem profunda, é uma sub-área da Aprendizagem de máquina, que emprega algoritmos para processar dados e imitar o processamento feito pelo cérebro humano. Deep Learning usa camadas de neurônios matemáticos para processar dados, compreender a fala humana e reconhecer objetos visualmente. As informações que passam pelos neurônios atuam semelhantemente as redes MLPs, pois toda informação que sai de um neurônio é posta como entrada para o seu seguinte, criando a rede neural. A primeira camada é conhecida como camada de entrada, enquanto a última é chamada de camada de saída, sendo assim, todas as camadas entre elas são denominadas camadas ocultas, a figura 8 descreve o comparativo entre um modelo de RNA usando Deep Learning comparado a uma simples rede neural.

Figura 11 – Comparativo entre um modelo utilizando Deep Learning e uma simples rede neural.



Fonte – Academy (2018)

A aprendizagem profunda é responsável por inúmeros avanços no campo de visão computacional, reconhecimento de fala, processamento de linguagem natural e reconhecimento

de áudio. Segundo Academy (2018), uma das principais características do Deep Learning é a capacidade de extrair recursos usando algoritmos para fins de treinamento, aprendizado e compreensão. Fato este, juntamente com inúmeros outros fazem com que esse tipo de rede comece a se sobressair em relação a outros métodos tradicionais de aprendizagem de máquina.

Um dos principais fatores que fez com que as redes neurais baseadas em Deep Learning ganhassem maior destaque foi o desenvolvimento e aumento tecnológico relacionados as unidades de processamentos GPU, possibilitando a formação de redes neurais convolutivas “sem” o pré-treino camada por camada. Com o aumento da velocidade da computação, tornou-se óbvio que Deep Learning tinha vantagens significativas em termos de eficiência e velocidade. Um exemplo classico é a AlexNet, uma rede neural convolutiva, cuja arquitetura ganhou várias competições internacionais durante 2011 e 2012, tendo como diferencial a utilização de unidades lineares retificadas para melhorar a velocidade.

2.6.1 Funções de Ativação

Para Wang et al. (2015), as funções de ativação são essenciais para as RNAs classificadas como Deep Learning, a análise do fluxo de informações acaba criando tarefas muito complexas, principalmente na classificação de imagens.

Em sua pesquisa, Phong e Phuong (2018) constataam que nos últimos anos houve um aumento significativo na quantidade de funções de ativação utilizadas por pesquisadores e pelo próprio mundo comercial. No início para a resolução de problemas relativamente simples funções lineares eram suficientes. Porém com a necessidade de resolução de problemas mais complexos, novas funções de ativação surgiram.

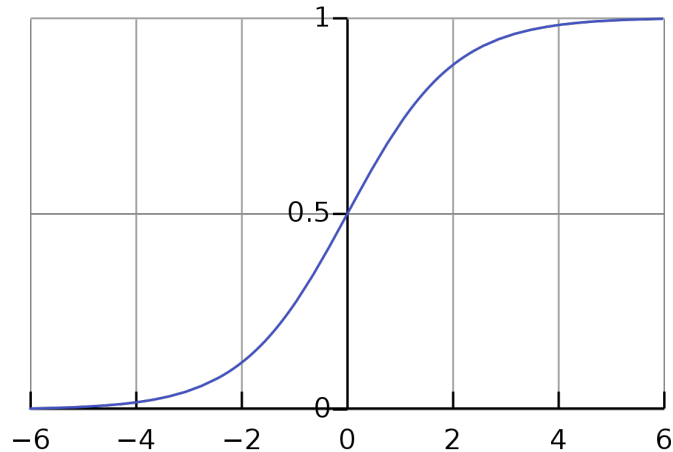
2.6.1.1 Sigmoid

A Data Science Academy aponta como uma das principais funções de ativação a Sigmoid, esta que é amplamente utilizada.

$$f(x) = 1 / (1 + e^{-x})$$

Ela é caracterizada por ser uma função suave e é continuamente diferenciável. O fato da não linearidade desta função faz com que a partir do momento que a RNA possua vários neurônios a saída passa a ter um comportamento não linear, variando de 0 a 1, com um formato de S.

Figura 12 – Função Sigmoid



Fonte – Acervo do Autor, 2018

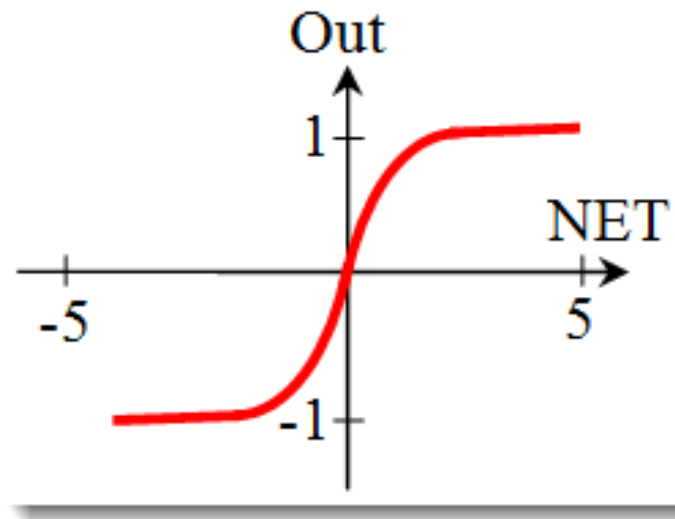
A função essencialmente empurra os valores de Y para os extremos, fato este que auxiliar na classificação de diferentes categorias. Apesar da função ser amplamente utilizada atualmente, ela possui alguns problemas, principalmente quando os gradientes se tornam muito pequenos, ou seja, eles estão se aproximando de zero e a rede não está aprendendo de fato. Outro problema relacionado a função de ativação sigmoid é que os valores ficam confinados entre 0 e 1, fazendo com que os valores além de limitados também sejam apenas números positivos. A Data Science Academy saliente que nem sempre é desejável que os valores enviados aos próximos neurônios sejam sempre de mesmo sinal.

2.6.1.2 Tanh

Com base no apresentado anteriormente, a Data Science Academy (2018) afirma que a função Tanh se assemelha a função sigmoid. Na verdade, ela se apresenta apenas como uma função escalonada da função sigmoid.

Tanh funciona de forma semelhante à função sigmóide, porém é simétrica em relação à origem. varia de -1 a 1. Fazendo então com que os valores retornados não sejam apenas de um sinal. Apresentando-se como contínuo e diferenciável em todos os pontos.

Figura 13 – Função Tahn



Fonte – Academy (2018)

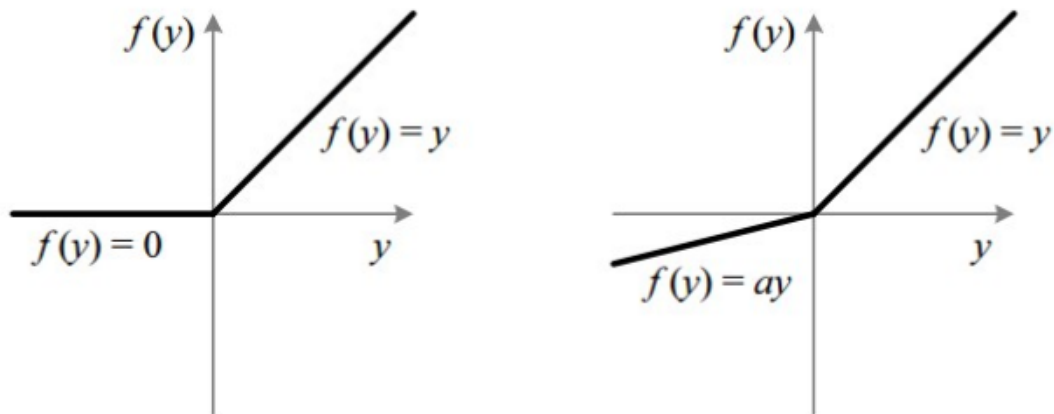
2.6.1.3 ReLU

De acordo com Chieng et al. (2018), é uma das funções de ativação mais utilizadas nos últimos tempos, seu crescimento se deve muito ao aumento da utilização de Deep Learning por sua alta taxa de sucesso nos resultados finais das RNAs.

Primeiramente a função ReLU não é uma função linear, fazendo com que seja possível aplicá-la em diversas camadas da Rede Neural. A principal vantagem de usá-la em relação as demais é que ela não ativa todos os neurônios ao mesmo tempo, pelo motivo de que caso a entrada seja negativo ela automaticamente será convertida para 0 e o neurônio não será ativado, significando que ao mesmo tempo apenas alguns neurônios são ativados, tornando a rede esparsa e eficiente e fácil para a computação

Entretanto a função de ativação ReLU também possui problemas, como afirma a Academy (2018), quando os gradientes deslocam-se em direção a 0, como nos casos anteriores. Porém esse problema é resolvido por meio de outra função de ativação baseada neste, a Leaky ReLU, a qual é uma função melhorada, ao contrário da função ReLU que não ativa os neurônios negativos, ela substitui a linha horizontal ocasionada pelo 0 por uma linha não-zero, não horizontal, mas um valor pequeno como por exemplo 0,01, removendo o gradiente 0.

Figura 14 – Função ReLU e Leaky ReLU



Fonte – Academy (2018)

2.6.1.4 Softmax

Por fim, a Academy (2018) aponta a função Softmax como uma variância da função Sigmoid, mas ela possui melhorias que fazem com que ela consiga lidar melhor com problemas relacionados a classificações múltiplas, a função sigmoid como mencionado anteriormente possui a capacidade de classificação de apenas duas classes. A função softmax transforma as saídas para cada classe para valores entre 0 e 1 e também divide a soma das saídas. Isso essencialmente dá a probabilidade de a entrada estar em uma determinada classe.

Um exemplo comum para explicar o funcionamento da função softmax é definindo as saídas como $[1.2, 0.9, 0.75]$, ao aplicar hipoteticamente a função softmax para essas saídas, os valores obtidos são $[0.42, 0.31, 0.27]$, onde cada elemento representa uma probabilidade de a entrada pertencer a uma das classes. Analisando a função percebe-se sua clara tendência para a utilização desta função de ativação na última camada da rede neural, a camada de saída.

2.6.2 RNA Convolucional

Segundo Yu (2017) as redes convolucionais possuem forte inspiração do córtex visual dos mamíferos. Por meio de estudos realizados por Hubel e Wiesel (1968) em gatos e macacos, tem-se o conhecimento da existência de dois tipos de células no córtex chamadas de simples e compostas. As células simples agem como detectores de bordas enquanto as células

compostas agem como extratores de características complexas formadas a partir de combinações de padrões simples.

Tendo como base as descobertas realizadas sobre o funcionamento do córtex foram propostas arquiteturas de redes neurais buscando emular o comportamento observado nos seres vivos. Fukushima em 1980 propôs uma arquitetura chamada NeoCognitron que pode ser considerada a primeira rede convolucional da literatura. Sua arquitetura consiste em 2 tipos de neurônios artificiais chamados de S-cells e C-cells que buscam simular o funcionamento das células simples e compostas no córtex. Após décadas de estudos e melhorias em relação das RNAs Convolucionais, Lopes (2017) define de forma simplificada que é possível considerar RNAs Convolucionais toda rede neural que utiliza a convolução no lugar da multiplicação de matrizes em pelo menos uma das suas camadas. Segundo Goodfellow, Bengio e Courville (2018)), existem 3 características distintas importantes nas redes neurais convolucionais em comparação as redes tradicionais. Estas são: o compartilhamento de pesos, a presença de campos receptivos locais e a existência de camadas de pooling.

Com a utilização do compartilhamento de pesos as redes não aprendem pesos específicos para cada região da imagem, ela aprende apenas um conjunto de filtros menores espacialmente, mas que pode ser aplicado em todas as regiões da imagem, o que faz com que essa técnica possua maior capacidade de generalização Wang et al. (2015). Ainda segundo Wang et al. (2015), outro benefício relevante do compartilhamento de pesos é que ele permite diminuir gradualmente a quantidade de parâmetros, simplificando o processo de treinamento e tornando-o mais eficiente.

Outra característica das RNAs Convolucionais é a existência de campos receptivos locais. Nas RNAs clássicas cada valor de entrada de cada camada é multiplicado por todos os neurônios presentes. Essa abordagem pode trazer grandes desvantagens quando utilizadas na identificação de padrões em imagens, pois os pesos podem se tornar extremamente grandes fazendo com que o custo computacional torne o modelo não viável. Tendo isso como base, as redes convolucionais buscam explorar subestruturas presentes nas imagens para otimizar o processamento. Em imagens naturais, os pixels adjacentes tendem a ser mais fortemente correlacionados do que pixels distantes, as redes convolucionais são modeladas para que cada filtro aprendido seja dependente de apenas uma sub-região dos dados recebidos da camada anterior.

A terceira característica é a existência de camadas de Pooling. O objetivo principal desta camada é a redução da dimensionalidade dos dados antes de propagá-los à próxima

camada da rede (GOODFELLOW; BENGIO; COURVILLE, 2016). Isto normalmente é realizado subdividindo a entrada em blocos e realizando algum tipo de sumarização estatística da informação presente em cada um deles. Os tipos de Pooling mais utilizados são o Average Pooling, onde a média de cada região é retornada e o Max Pooling, onde o retorno é o maior valor presente em cada vizinhança.

2.6.2.1 Tipos de Camadas

De acordo com Yu (2017), as camadas de evolução são adicionadas antes da camada densa, que são camadas padrões tanto para redes convolucionais quanto para redes perceptron, logo pode-se perceber que a camada convulacional pode ser caracterizada como uma camada anterior ao processamento, uma camada que realiza a preparação e otimização dos dados para que sejam enviados a rede neural.

2.6.2.2 Operador de Convolução

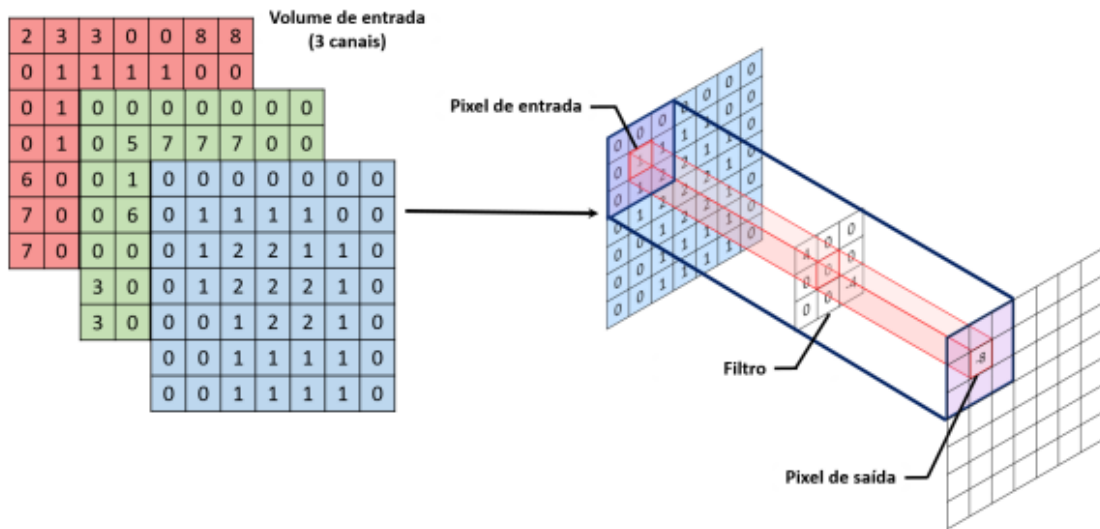
Para Yu (2017), as redes convolucionais são compostas de sequências de camadas empilhadas umas sobre as outras, sendo que cada camada faz algum tipo de transformação nos dados de entrada através de uma função definida. Os tipos de camadas mais comuns são: operador de convolução, pooling flattening e rede neural densa.

De acordo com Yu (2017), a camada convolucional consiste em um conjunto de filtros que recebem como entrada um arranjo 3D, também chamado de volume. Cada filtro possui entrada reduzida, porém se estende por toda a profundidade do volume de entrada. Por exemplo, em uma imagem colorida que possui 3 canais e o filtro da primeira camada convolucional tem o tamanho $5 \times 5 \times 3$ (5 pixels de altura e largura, e profundidade igual a 3). Ao realizar o processo de treinamento estes filtros são ajustados para que sua ativação aconteça na presença de características relevantes. Portando Araújo et.al (2017), comenta que é comum ocorrer pré processamentos que convertam imagens RGB em imagens em escala de cinza quando não há a necessidade de identificação das cores, pois enquanto as imagens RGB possui 3 canais de filtro, triplicando o processamento das redes neurais, as imagens em escala de cinza possuem apenas um canal, com variação de bits de 0 a 255.

Ainda de acordo com Araújo et al. (2017), cada um desses filtros ou também chamados de Kernel dão 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 da entrada é uma operação conhecida como convolução, a qual é ilustrada na figura 16. Os valores resultantes após a operação de convolução passam por uma função de ativação, e a mais comum é a função ReLU.

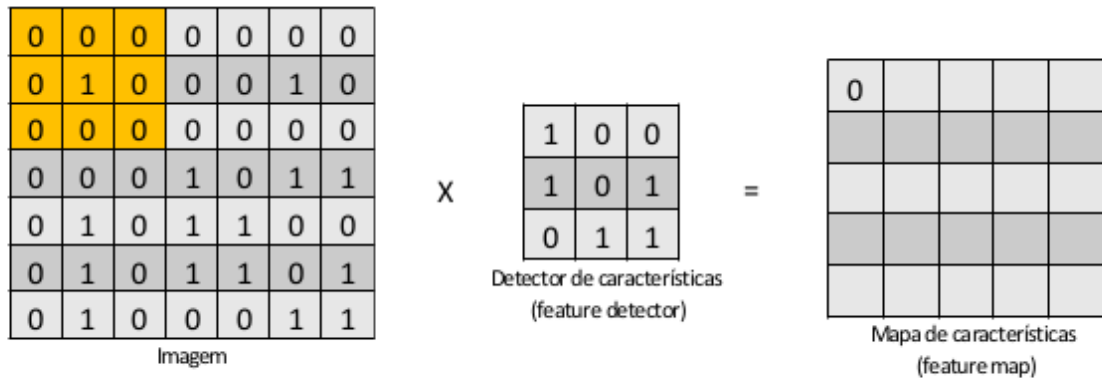
Figura 15 – Convolução entre um filtro 3x3 e o volume de entrada



Fonte – Granatyr e Amaral (2018)

Analisando por outra perspectiva, Granatyr e Amaral (2018), apresentam uma abordagem mais demonstrativa de como funciona o processo de convolução. A figura abaixo demonstra o processo de multiplicação de uma imagem qualquer por uma máscara (feature detector) qualquer, a fim de demonstrar o resultado dessa operação.

Figura 16 – Cálculo demonstrativo da camada de convolução

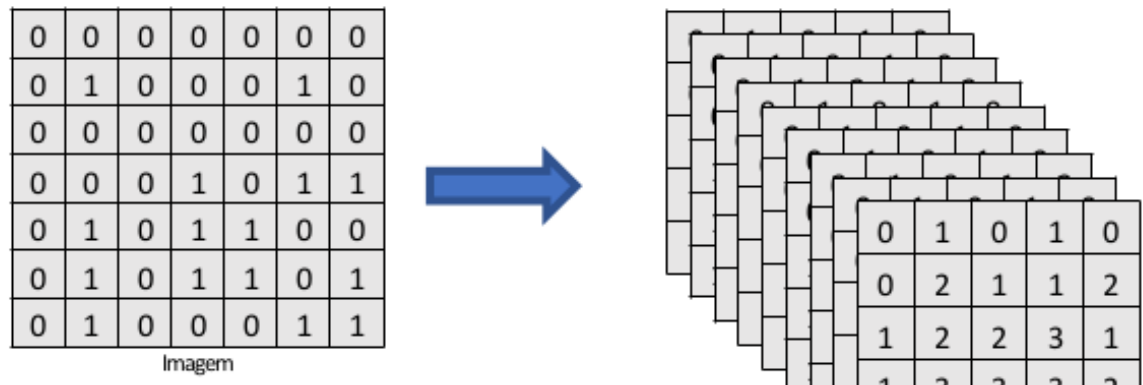


Fonte – Granatyr e Amaral (2018)

Com base no tamanho do Kernel utilizado, o qual neste exemplo é de 3x3, é selecionado o mesmo tamanho na matriz imagem em cada interação, ou seja, é extraído uma matriz 3x3 da imagem original, chamada de janela e multiplicada pelo kernel, após isso é somado o resultado e inserido em uma nova matriz. Após cada interação, a janela deve ser descolada apenas um elemento e o cálculo deve ser refeito adicionando na próxima casa da matriz resultante. Vale salientar que o deslocamento de um passo por interação pode ser definido de acordo com o modelo utilizado.

Observando o mapa de características resultante é possível visualizar que a matriz original de 7x7 tornou-se uma matriz 5x5, essa que servirá de entrada para a próxima camada. Lembrando que essa operação deverá ser realizada para todos os kernels definidos, ao final, todos os mapas de características serão enviados para a etapa seguinte, fato este que é ilustrado pela figura 18.

Figura 17 – Representação da saída da camada de convolução



Fonte – Granatyr e Amaral (2018)

Para melhor entendimento sobre a multiplicação dos filtros (kernels) foi realizado alguns testes utilizando uma imagem exemplo, submetendo-a a multiplicação com alguns tipos de filtros para realçar algumas de suas características, os resultados podem ser vistos nas próximas imagens.

Figura 18 – Imagem Exemplo antes de aplicar os filtros

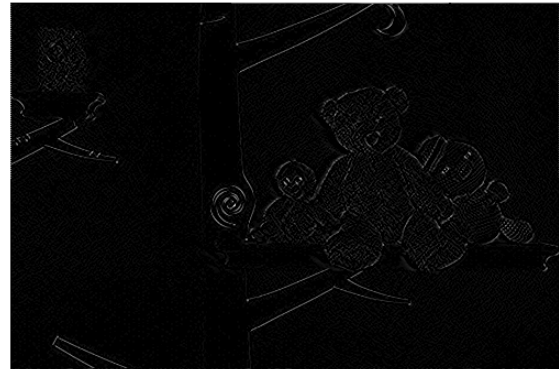


Fonte - Banco de Imagens grátis: <https://pt.shopify.com>

Figura 19 – Filtro: Edge Detection - Detecção de bordas

1	0	-1
0	0	0
-1	0	1

custom ▼

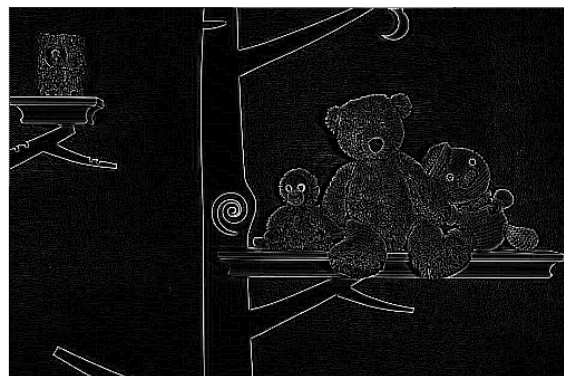


Fonte - Acervo do autor, 2018

Figura 20 – Filtro: Outline - Realça a diferença entre pixels vizinhos

-1	-1	-1
-1	8	-1
-1	-1	-1

outline ▼



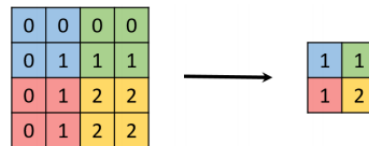
Fonte - Acervo do autor, 2018

2.6.2.3 Pooling

Normalmente essa camada de pooling é aplicada com o objetivo remover informações não necessárias para a identificação das características não essenciais atreladas a imagem e removendo o overfitting que é quando a rede neural se adapta excessivamente ao conjunto de treinamento. Segundo Araújo et al. (2017) o objetivo da camada de Pooling é reduzir progres-

sivamente a dimensão espacial do volume de entrada, consequentemente a redução diminui o custo computacional da rede. Na operação de Pooling, os valores obtidos por meio do mapa de características gerados pela camada anterior, são substituídos por alguma métrica dessa região. A forma mais comum de Pooling consiste em substituir os valores de uma região pelo valor máximo como é ilustrado pela Figura 22. Vale ressaltar que uma das principais funções da etapa de Pooling é diminuir a matriz original, fazendo com que o processamento que deve ser realizado pela rede neural seja diminuído pelo fato da entrada ser menor.

Figura 21 – Max-Pooling

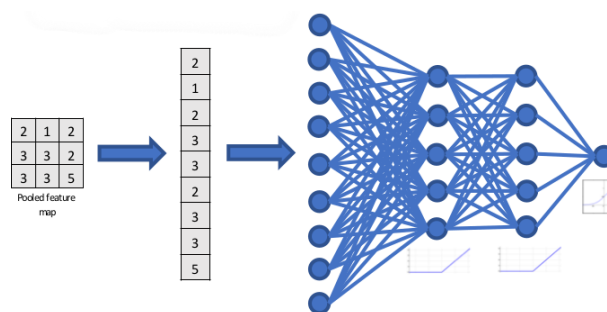


Fonte – Granatyr e Amaral (2018)

2.6.2.4 Flattening

Por fim Araújo et al. (2017) explica que as camadas de flattening são responsáveis por ajustar a representação dos dados para que possam ser utilizados pelas camadas densas da rede neural, analisando de uma forma mais simples, essa camada é responsável por "achatar" os valores para que tornem-se entradas para a próxima etapa de treinamento. A figura 22 ilustra a etapa de flattening.

Figura 22 – Flattening



Fonte – Granatyr e Amaral (2018)

3 TRABALHOS CORRELATOS

Para a realização deste trabalho tomou-se uma preocupação em apresentar trabalhos com ideias semelhantes ao apresentado aqui e para maior contemplação da área de estudo, serão apresentados alguns trabalhos com linhas de pensamento diferentes, porém com o mesmo objetivo definido para este trabalho. Nesta etapa do trabalho, foram selecionados 2 trabalhos correlatos, que serão elucidados a seguir.

3.1 COMPARAÇÃO ENTRE TRABALHOS CORRELATOS

A Tabela 1 apresenta os trabalhos correlatos com data de publicação posterior a 2009.

Tabela 1 – Trabalhos Correlatos.

Referência	Técnica Utilizada	Finalidade
Yu (2016)	CNN	Viabilidade da aplicação da tecnologia no processamento de imagens dentárias.
Lopes (2017) (2016)	CNN	Detectar Tuberculose em Radiografias.
João de Oliveira (2009)	Computação Gráfica	Detecção de cáries por meio de computação gráfica.

3.1.1 Machine Learning for Dental Image Analysis

O principal objetivo da pesquisa desenvolvida por Young-jun Yu (2016) é avaliar a aplicação da tecnologia da informação em imagens dentárias, mais especificamente a possível implantação de redes neurais convolucionais. Além disso o trabalho pretende criar histogramas de gradientes orientados (HOG).

Neste estudo, Young-jun Yu avalia seus resultados por meio de áreas como: precisa, execução das tarefas e eficiência. As informações que foram obtidas sobre o modelo CNN (Convolutional Neural Network) foram então comparadas aos resultados obtidos em algoritmos de cognição SIFT e HOG. Para o desenvolvimento da CNN, usou-se uma biblioteca de python chamada de Theano, cuja principal característica baseia-se na criação de expressões matemáticas otimizadas, desenvolvidas para maior eficiência na utilização de termos multidimensionais.

Para a realização da pesquisa foi utilizado dois diferentes tipos de CNN, uma possuindo 4 camadas internas e outras 16. O modelo de 4 camadas é uma pilha simples que consiste em duas camadas de convolução e uma camada de agrupamento. A função de ativação definida para a rede neural de Young-jun Yu foi a função ReLU.

Para o processo de aprendizado foi utilizado 100 imagens panorâmicas da arcada dentária, dividindo-as em 28 partes, cada uma representando um dente totalizando 2800 imagens a serem processadas. Ao fim dos testes foi obtido uma acuracidade acima de 90%, variando uma pequena porcentagem de acordo com a classe de dentes em questão.

3.1.2 Redes Neurais Convolucionais Aplicadas ao Diagnóstico de Tuberculose por meio de Imagens Radiológicas

O foco da pesquisa de Uilian Kenedi Lopes (2017), é classificar imagens Raio-X torácico utilizando CNN com o objetivo de diagnosticar a presença de tuberculose. O autor define três objetivos específicos para a realização de seu trabalho. A primeira aplicando diretamente diferentes arquiteturas de CNN com extratoras de características em uma versão reduzida da imagem radiográfica, a segunda utilizando as mesmas arquiteturas de CNN para extrair características em sub-janelas da imagem (em sua resolução original) que posteriormente são combinadas para gerar um único descritor global da radiografia e a terceira criando comitês de classificadores juntando os resultados obtidos nas duas primeiras abordagens.

Para o treinamento e testes foram coletadas 138 radiografias frontais do tórax, sendo 80 imagens de pessoas saudáveis e 58 de casos com pessoas que possuem alguma manifestação de tuberculose. Todos os pré-processamento criados especificamente para o trabalho foram implementados utilizando o ambiente Matlab e a linguagem C++ juntamente com a biblioteca OpenCV. Para o desenvolvimento do modelo da CNN foi utilizado a biblioteca MatConvNet, uma API simplista e intuitiva, segundo as observações do autor.

Para o desenvolvimento do modelo de CNN o autor apresentou duas propostas, a primeira baseando-se nas redes neurais GoogLeNet, VggNet e ResNet e a segunda proposta o autor realizou uma fase de pré-treino, gerando um dicionário de características visuais de clusterização dos vetores extraídos, ou seja, repartindo a imagem completa em sub-janelas para melhorar o processamento da CNN por utilizar quantidades menores de dados de entrada. Cada janela apresenta uma taxa de 50% de sobreposição para que características da doença não se apresente separadas ocasionando possíveis problemas na detecção.

Lopes apresenta seus resultados como promissores, a primeira proposta apresenta

uma acuracidade que varia de 85% a 91% e a segunda proposta apresenta sua taxa um pouco inferior de 75% até 82%. Sua taxa de variação deu-se pelo fato do autor buscar as imagens em duas fontes diferentes e para fins estatísticos apresentou os resultados individualmente por base de dados.

3.1.3 Caries Detection in Panoramic Dental X-ray Images

O trabalho de João Oliveira(2009), possui como principal objetivo a identificação de cáries dentárias por meio de imagens derivadas de raios-x panorâmicos. Um dos principais diferenciais presentes nestes trabalho, este que também é ressaltado pelo autor, é o estudo completo do processo que envolve a detecção, pois o autor cita alguns trabalhos relacionados que abrangem certas etapas do processo, porém não o contempla como um todo, fato este que não acontece no trabalho desenvolvido.

Para realizar o processo João Oliveira dividiu seu processo em diferentes etapas.

A aquisição das imagens, estas que foram coletadas manualmente pelo autor.

Como próximo passo, ele optou pelo pré-processamento destas imagens, no qual utilizou alguns métodos específicos, como por exemplo redução de ruídos, aumento de contraste entre outras técnicas.

O terceiro passo envolveu a segmentação dos dentes, no qual utilizou o conceito de localização de limites e linhas nos dentes, para que então fosse possível segmentá-los um por um.

E por último, a partir das imagens segmentadas, por meio de técnicas computacionais, detectar a diferença de coloração em certas partes do dente, para que então estas sejam categorizadas como cáries ou não.

Ao término do processo de avaliação, o autor chegou em uma porcentagem de acuracidade em torno de 95,7% na detecção de cáries dentárias, resultado este que foi extremamente animador para o autor. Um dos apontamentos realizados por ele foi a de em trabalhos futuros aplicar o mesmo algoritmo, porém ao invés de realizar a detecção de cáries, realizar a detecção de outras doenças, fazendo com que a abrangência torne-se muito maior.

4 DESENVOLVIMENTO

Este capítulo tem como objetivo inicial detalhar os dados coletados, sua forma de obtenção e as operações realizadas na etapa de pré-processamento das imagens para torná-las aptas à manipulação pelas RNAs. Também serão abordadas técnicas utilizadas para melhorar a performance do algoritmo como validação cruzada e teste de melhor algoritmo.

Este capítulo também descreve os recursos utilizados, incluindo softwares de manipulação de imagens, IDEs de desenvolvimento e frameworks utilizados para facilitar o processo de aprendizado de máquina. Por fim serão apresentados os índices obtidos da categorização das imagens, tanto nas redes puramente perceptron quanto nas redes convolucionais.

4.1 DADOS DISPONÍVEIS

Os dados utilizados para treinamento, validação e testes das RNAs foram retirados do site acadêmico Mendeley, este que é um gerenciador de referências gratuito e uma rede social acadêmica gratuita, para busca de trabalhos recentes. O conteúdo coletado representa cerca de 110 imagens radiográficas da mandíbula humana, no formato panorâmico, porém nem todas as imagens foram utilizadas, pois algumas não possuíam condições para serem analisadas, por não possuírem grande parte dos dentes. A figura 23 apresenta uma das imagens coletadas no repositório.

Figura 23 – Radiografica Panorâmica do repositório Mendeley

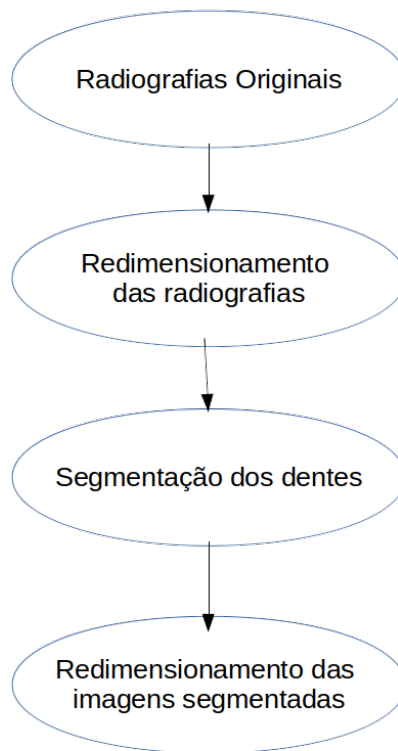


Fonte – Abdi e Kasaei (2017)

4.1.1 TRATAMENTO DOS DADOS

As imagens radiológicas foram obtidas tanto no formato PNG (Portable Network Graphics) quanto no formato JPEG (Joint Photographics Experts Group), porém para o propósito deste estudo, é necessário um pré-processamento dessas imagens, pois o objetivo é realizar a categorização das imagens, logo é necessário segmentá-las para que então estas possam ser inseridas em uma RNA para aprendizado. A figura 24 apresenta um diagrama contendo as etapas de preparação dos dados para a RNA.

Figura 24 – Pré processamento das imagens



Fonte – Acervo do Autor, 2018

O primeiro passo a ser realizado é a normalização das proporções das imagens, pois inicialmente os diâmetros encontrados variam entre 1200x3200 pixels até 1400x3400 pixels. Para melhor ajuste a rede neural foi realizado a normalização de todas as imagens com 1250x2980 pixels, decisão tomada arbitrariamente.

O passo seguinte baseia-se na segmentação manual de cada dente individualmente (utilização do software gratuito gThumb) para com que então, sejam classificados de acordo

com o seu tipo (incisivo, canino, pré-molar e molar). Após a segmentação foi realizada uma padronização do tamanho final para que estes sejam inseridos na rede neural. A padronização será realizada por um algoritmo que irá redimensionar todas as imagens para a mesma dimensão de 125 x 125 pixels, criando uma uniformidade na quantidade de entradas da rede neural.

4.1.2 Normalização dos dados

Para a utilização dos dados nos treinamentos e testes das RNAs, faz-se necessário a realização de um pré-processamento dos dados, denominado de normalização. De acordo com Silva, Spatti e Flauzino (2010), este processo objetiva escalonar as amostras de dados para uma faixa de valores dinâmica das funções de ativação das camadas escondidas, a fim de evitar o ruído dos dados externos à rede. Para o modelo utilizado nesse trabalho, os valores de entrada foram normalizados em duas etapas.

A primeira foi converter os dados obtidos após a conversão da imagem em uma matriz, pois seu formato veio em RGB, como para o propósito dessa aplicação não é necessário diferenciar as cores das imagens elas foram convertidas em escala de cinza por meio de funções fornecidas pela biblioteca Keras.

A segunda etapa deu-se pela redução de escala dos valores, no qual todos pixels foram divididos por 255, diminuindo então sua escala, a qual ficou entre 0 e 1.

4.2 VALIDAÇÃO CRUZADA

De acordo com Silva, Spatti e Flauzino (2010), a validação cruzada (crosscross validation) tem como objetivo melhorar a taxa de acuracidade de uma rede neural por meio da diversificação dos dados de treinamento, por meio dessa técnica é possível testar uma determinada topologia de rede por múltiplos subconjuntos do subconjunto total, este que representa todo os dados coletados, fazendo então uma média ao final de todos os resultados obtidos, este que segundo Silva, Spatti e Flauzino (2010), tornam-se muitas vezes, mais precisos que uma simples disposição de subconjuntos de teste e treinamento.

Para o treinamento da RNA deste trabalho foi utilizado uma taxa de separação tendo 80% do conjunto total como teste e 20% para treinamento. A diversificação dos dois subconjuntos foi repetida 20 vezes para cada modelo proposto.

4.3 RECURSOS UTILIZADOS

A implementação, construção e testes das RNAs deste trabalho, foram feitas por meio da linguagem python, mais especificamente a biblioteca Keras esta que é uma API de redes neurais de alto nível, escrita em Python.

Durante a modelagem das RNAs foram utilizados diversos modelos de redes neurais, variando número de neurônios da camada oculta e funções de ativação. Outro fator que alavancou os frameworks de Deep Learning foi o surgimento da computação paralela, os principais algoritmos utilizam métodos de separação de tarefas para obter os resultados de forma mais rápida.

4.3.1 Framework de desenvolvimento de RNAs - Keras

Segundo Anderson, Vlimant e Spiropulu (2017), o recente progresso no campo de aprendizado de máquina fez com que os frameworks ganhassem maior visibilidade e por consequência maior progresso.

Um dos principais frameworks para o processamento não apenas de redes profundas, mas também de processamento de imagens é o Keras, uma biblioteca python desenvolvida com a finalidade de tornar a construção de redes neurais práticas e simples, porém não deixando de lado o seu foco no processamento, não apenas vinculado a CPU, mas também a GPU.

A biblioteca Kera foi desenvolvida de acordo com alguns princípios, mas especificamente 4 que são apresentados em seu website.

Facilidade de utilização: Keras é uma API projetada para seres humanos, não para máquinas, colocando a experiência do usuário na frente e no centro. Um dos seus principais focos é reduzir as ações necessárias feitas pelo usuário para resolver um caso de uso, fornecendo um feedback simples.

Modularidade: Um modelo é entendido como uma sequência ou um gráfico de módulos autônomos configuráveis que podem ser conectados com o menor número de restrições possíveis. Em particular, camadas neurais, funções de custo, otimizadores, esquemas de inicialização, funções de ativação entre outros.

Extensibilidade fácil: Novos módulos são simples de adicionar, como por exemplo novas classes e funções. Além disso os módulos já existentes já fornecem exemplos amplos.

Trabalhe com python: Não há arquivos de configuração de modelos separados em um formato declarativo. Todos os módulos são descritos em Python, permitindo maior facilidade

dade de uso e flexibilidade.

4.3.2 Funções de Ativação

Para Braga, Carvalho e Ludermir (2007), RNAs normalmente quando caracterizadas como múltiplas camadas, com complexidade moderada, tende a exigir funções de ativações não lineares em suas camadas internas para com que as camadas posteriores possam resolver os problemas relacionados ao propósito da RNA. Porém, ao contrário das camadas internas, Braga, Carvalho e Ludermir (2007) mencionam que é comum a camada de saída possuir uma função de ativação linear, porém depende muito do problema envolvido.

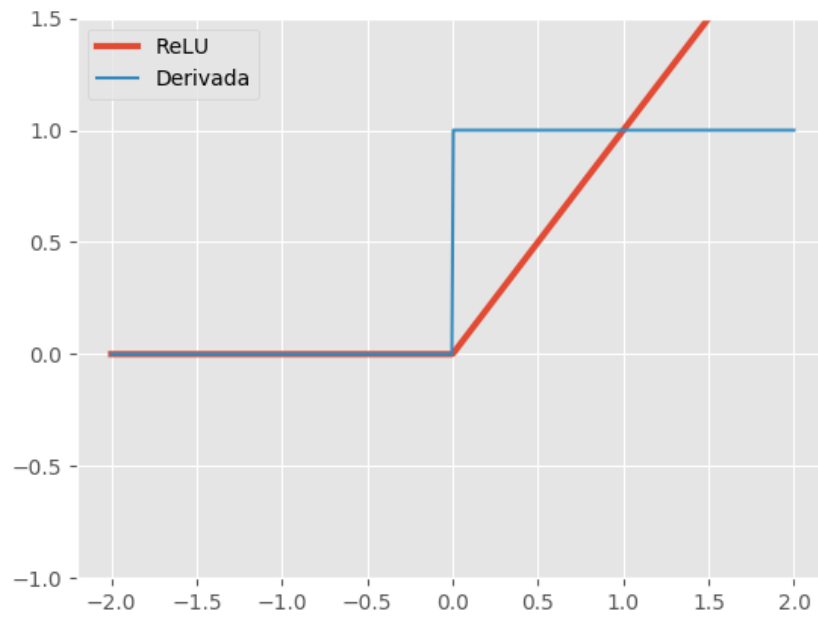
4.3.2.1 Soft Max

Analisando a proposta desse trabalho percebe-se a necessidade da utilização de uma função de ativação na camada de saída que possibilite a classificação dos resultados em 4 categorias, para isso o Academy (2018) recomenda a utilização da função de ativação Softmax, uma variação da função sigmoid, porém melhor otimizada para problemas de classificação pois a função sigmoid é capaz de lidar apenas com duas classes. A função softmax transforma as saídas de cada classe para valores entre 0 e 1 e também divide a soma das saídas. Isso essencialmente dá a probabilidade de a entrada estar em uma determinada classe.

4.3.2.2 ReLU

Para as camadas internas da RNA foi optado pela utilização da função ReLU, caracterizada pela sua excelente resposta em redes que analisam imagens, esta que é uma rede que caracteriza valores entre 0-1, sendo que todos os valores negativos enviados para ela são convertidos em 0. Sua principal aplicabilidade para o problema proposto é o fato de ela remover os valores negativos, estes que em uma imagem normalmente representam as áreas mais escuras, como é possível observar na figura 25, as áreas mais escuras representam o fundo da imagem, como a intenção do algoritmo é classificar os dentes, esta função torna-se interessante para a solução do problema proposto.

Figura 25 – Função de Ativação ReLU



Fonte – Academy (2018)

Figura 26 – Dente Segmentado



Fonte – Abdi e Kasaei (2017)

4.3.2.3 Tangente Hiperbólica - Tahn

Para realizar alguns comparativos foi utilizado também a função Tahn que diferentemente da função ReLU ela possui variação de -1 até 1, fazendo com que seja possível utilizar valores negativos, criando então um contraponto que será visto nos resultados obtidos.

4.4 PROCESSO COMPARATIVO ENTRE AS REDES DESENVOLVIDAS

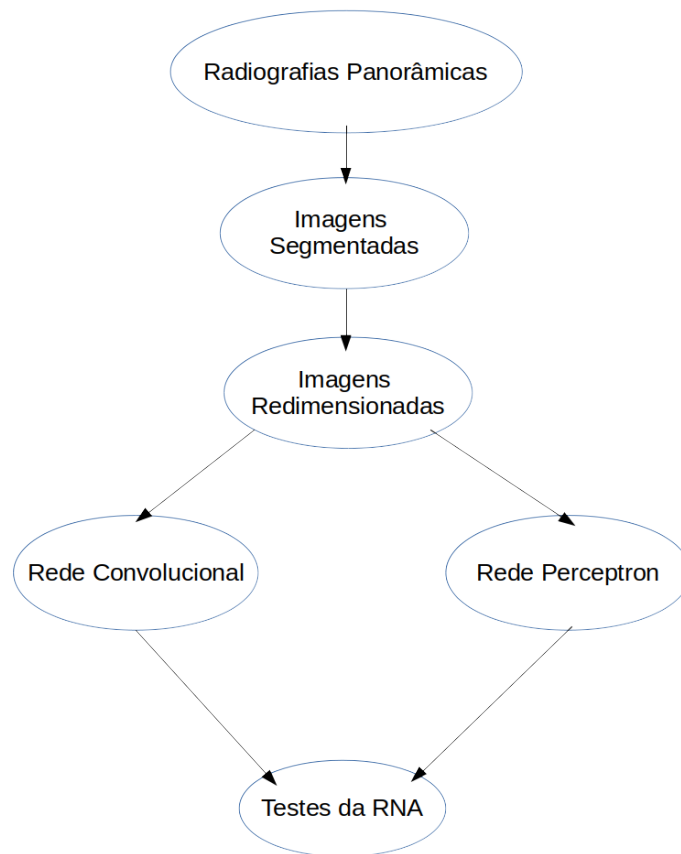
Neste trabalho foram modeladas RNAs com diferentes quantidades de neurônios, camadas e técnicas em ambas as redes neurais para identificar o melhor modelo possível. Para fins comparativos a rede neural convolucional irá utilizar o mesmo modelo que a rede perceptron durante seus testes, porém irá ser adicionado a camada de pré processamento convulacional.

O diagrama representado pela figura 27 demonstra os passos que serão utilizados para realizar a classificação das imagens em ambas as redes neurais.

Inicialmente serão coletadas as imagens panorâmicas da arcada dentária, para que então seja possível segmentá-las, dente por dente, criando o dataset que serão utilizado no processamento da rede neural. Devido aos diferentes formatos de dentes, após a segmentação dos mesmos, cada um será redimensionado para o tamanho de 125x125 pois estes devem estar padronizados para que as entradas sejam de mesmo tamanho.

Após todo o processo, as imagens serão inseridas em ambas as redes neurais desenvolvidas, a convolucional e a perceptron, para que então seus resultados sejam comparados.

Figura 27 – Processo de manipulação de imagens

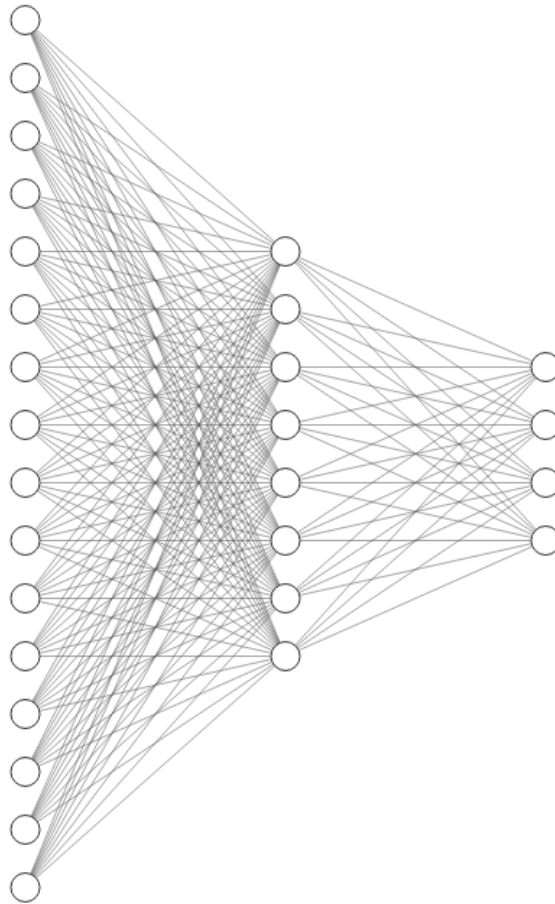


Fonte – Acervo do Autor, 2018

Como mencionado anteriormente, as entradas da rede neural serão os pixels totais coletados de cada imagem, portanto como o dimensionamento de todas as imagens presentes para esse projeto são de 125 x 125 pixels serão adicionadas 15625 entradas, estas que serão processadas pelas camadas internas de acordo com o modelo da rede neural, resultando em um conjunto específico de 4 saídas, cada qual representando um tipo de dente, incisivos, caninos, pré-molares ou molares.

Os testes que serão realizados utilizarão diferentes tipos de configurações para as redes neurais. Um dos modelos implementados pode ser observado na figura 28.

Figura 28 – Modelo reduzido 01 MLP de classificação dos tipos de dentes

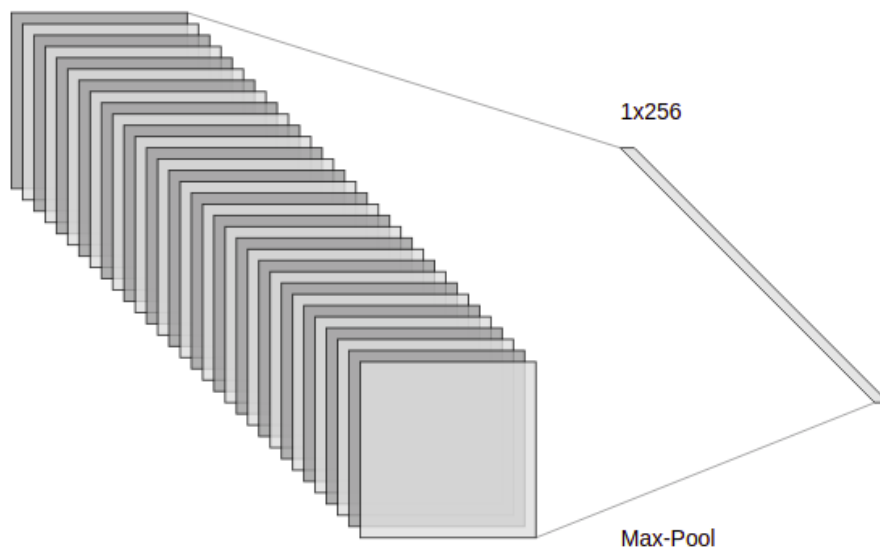


Fonte – Acervo do Autor, 2018

Para essa configuração da rede, a camada de entrada mantém-se constante em 15625 neurônios, fato esse que acontece em todas as variações, pois a intenção de variação dos testes ocorre na quantidade de neurônios da camada interna e de layers. Será utilizada uma camada internas utilizando a função de ativação ReLU. A camada interna contém 128 neurônios. Vale salientar que essa estrutura é utilizada tanto para a rede MLP quanto para a rede convolucional, a única alteração é que haverá uma camada anterior a camada densa no caso da rede convolucional que irá aplicar as técnicas explicadas anteriormente.

A representação do modelo convolucional se dá pelo acréscimo de uma camada convolucional anterior a camada densa (camada onde todos os neurônios estão conectados). Ela possui 32 filtros (kernel) com dimensão de 3x3. A figura 29 ilustra a camada convolucional.

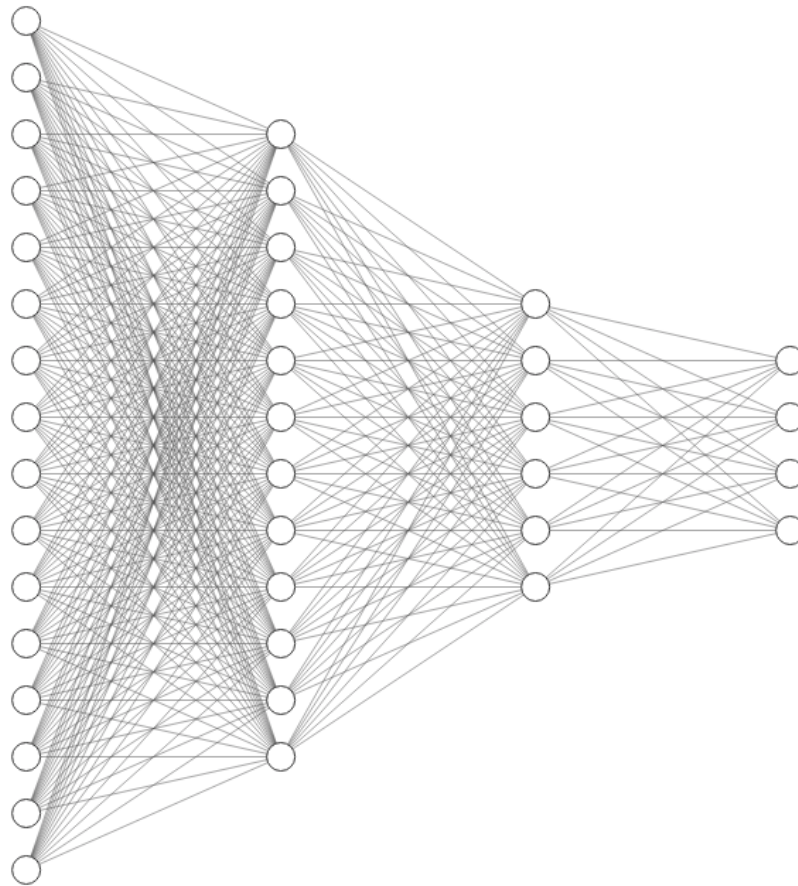
Figura 29 – Camada convulacional 01 de classificação dos tipos de dentes



Fonte – Acervo do Autor, 2018

O segundo modelo apresentado por meio da figura 30, possui entradas e saídas idênticas ao modelo anterior, porém possuindo variações tanto na quantidade de layers quanto na quantidade de neurônios distribuídos através delas. Esse modelo apresenta 2 camadas internas, todas possuindo como função de ativação ReLU, porém diferentemente do modelo anterior, este possui em sua primeira camada interna 128 neurônios seguido por 64 neurônios nas camadas posteriores.

Figura 30 – Modelo reduzido 02 MLP de classificação dos tipos de dentes



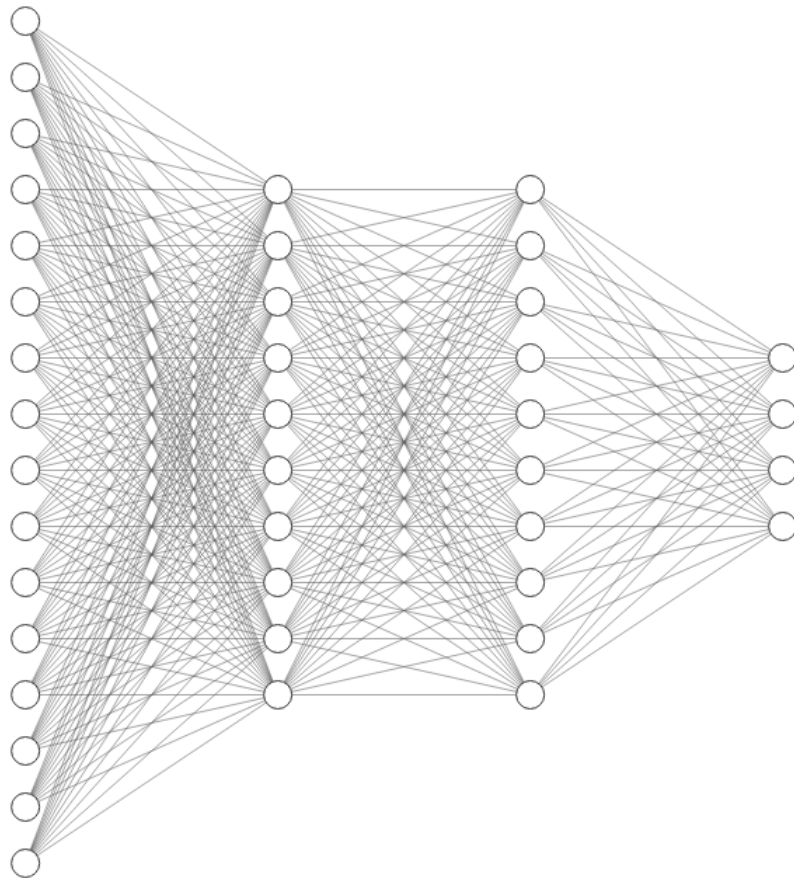
Fonte – Acervo do Autor, 2018

Com o intuito de verificar qual etapa apresenta os melhores resultados, para esse exemplo a camada convolucional foi mantida.

O terceiro modelo apresenta a mesma quantidade de camadas e neurônios internos o primeiro modelo, porém possui sua diferença na função de ativação utilizada, substituindo a função Relu pela função de ativação Tahn. A camada convolucional se mantém.

O quarto e último modelo apresenta duas camadas internas, a primeira possuindo 128 neurônios com função de ativação Tahn e a segunda também com 128 neurônios, porém com função de ativação ReLU.

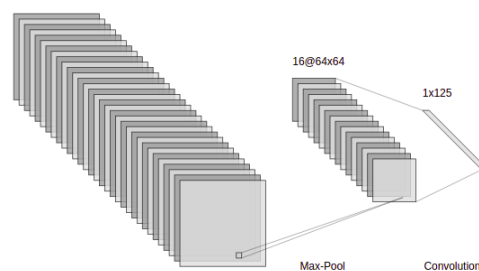
Figura 31 – Modelo reduzido 03 MLP de classificação dos tipos de dentes



Fonte – Acervo do Autor, 2018

A camada convolucional do quarto modelo possui 2 layers, a primeira com 32 filtros e dimensão de 2x2 e a segunda 16 filtros e dimensão de 3x3.

Figura 32 – Camada convolucional 04 de classificação dos tipos de dentes



Fonte – Acervo do Autor, 2018

5 RESULTADOS

Este capítulo apresenta inicialmente uma descrição dos treinamentos das RNAs bem como os resultados obtidos pelas mesmas. Além disso, busca-se definir quais redes neurais obtiveram os melhores resultados com base nos testes realizados, para que então possa ser realizada uma comparação entre as redes desenvolvidas utilizando a camada convolucional e as redes Perceptron.

Inicialmente será apresentado um comparativo por categoria de dentes, focando na acuracidade obtida por ambas as redes neurais, para que então a seguir seja apresentado os resultados gerais de acordo com os testes realizados.

5.1 ANÁLISE POR CATEGORIA

Para melhor expressar os resultados obtidos a tabela a seguir representa a acuracidade obtida ao submeter as redes neurais desenvolvidas ao conjunto de testes. O calculo de acuracidade se dá por meio da análise de 100 imagens de cada um dos tipos, submetidas a rede dos modelos desenvolvidos nesse trabalho.

Tabela 2 – Testes por Categoria

Tipo	Teste	Perceptron	CNN
Incisivo	01	75.0%	90.0%
Incisivo	02	72.0%	94.0%
Incisivo	03	79.0%	87.0%
Incisivo	04	81.0%	93.0%
Canino	01	78.0%	94.0%
Canino	02	77.0%	92.0%
Canino	03	74.0%	90.0%
Canino	04	80.0%	95.0%
Pré-molar	01	75.0%	84.0%
Pré-molar	02	73.0%	86.0%
Pré-molar	03	72.0%	87.0%
Pré-molar	04	79.0%	83.0%
Molar	01	78.0%	92.0%
Molar	02	79.0%	97.0%
Molar	03	77.0%	92.0%
Molar	04	81.0%	98.0%

A tabela acima demonstra um comparativo por categoria em cada um dos testes

realizados, como é possível visualizar, foram desenvolvidas 4 diferentes tipos de redes neurais, para que seja possível demonstrar de forma mais completa as variações nos resultados de acordo com a variação criada na rede neural.

Os resultados de acuracidade da rede Perceptron variam de 70% até 90% enquanto os que utilizam as redes convolucionais possuem sua variação entre 90% e 98% demonstrando que sua utilização proporciona melhores resultados a acuracidade do modelo.

5.2 TREINAMENTO E RESULTADOS DAS RNAS

As próximas subseções descrevem o processo de treinamento e os resultados obtidos com a utilização das redes Multilayer Perceptron e Convolucional.

5.2.1 Treinamento

Para o treinamento de uma RNA, existem diversas variações de parâmetros que podem ser realizados. Neste trabalho foram utilizadas variações nas camadas internas (variações na quantidade de neurônios, quantidade de camadas e quantidade de filtros), funções de ativação e quantidade de épocas de treinamento, deixando fixo os valores de entrada e de saídas por restrições impostas pelos modelos de análise que foram definidos.

Em ambos os algoritmos não foram definidas restrições de parada. Como o principal intuito do trabalho é realizar um comparativo de resultados entre os dois tipos de RNAs, foram definidos limites de épocas idênticas para as redes neurais, deixando o comparativo de qualidade mais evidente.

Como mencionado anteriormente, a camada convolutiva pode ser caracterizada como um pré processamento dos dados para que então sejam inseridos na camada densa da RNA. Sabendo disso, a estrutura densa da rede neural, ou seja, as camadas internas conectadas não variam do teste Convolucional para teste Perceptron, aumentando a qualidade comparativa das análises realizadas.

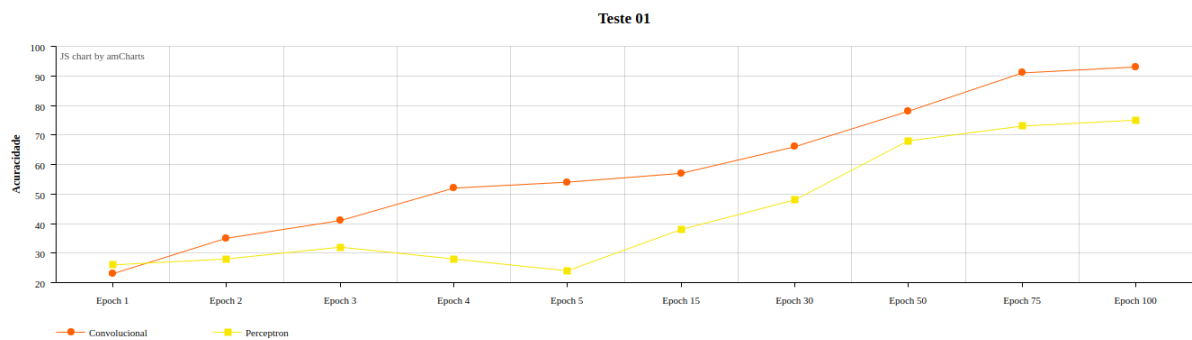
5.2.2 Testes

Para esse tópico serão exibidos gráficos comparativos por época com relação a acuracidade obtida por cada rede neural, dado esse que é retornado automaticamente pela biblioteca Keras. Ao final dos testes, será comentado os resultados comparando as redes implementadas utilizando Perceptron e as redes que aplicaram a camada convolucional.

5.2.2.1 Teste 01

Para o primeiro teste realizado foi utilizando uma camada interna com a função de ativação Relu, contendo 128 neurônios. Já a representação do modelo convolucional se dá pelo acréscimo de uma camada convolutiva possuindo 32 filtros com dimensão de 3x3. A imagem abaixo ilustra o comportamento da acuracidade obtida durante as 100 primeiras épocas de treinamento de ambos os modelos de redes neurais.

Figura 33 – Teste 01 comparativo entre as RNAs



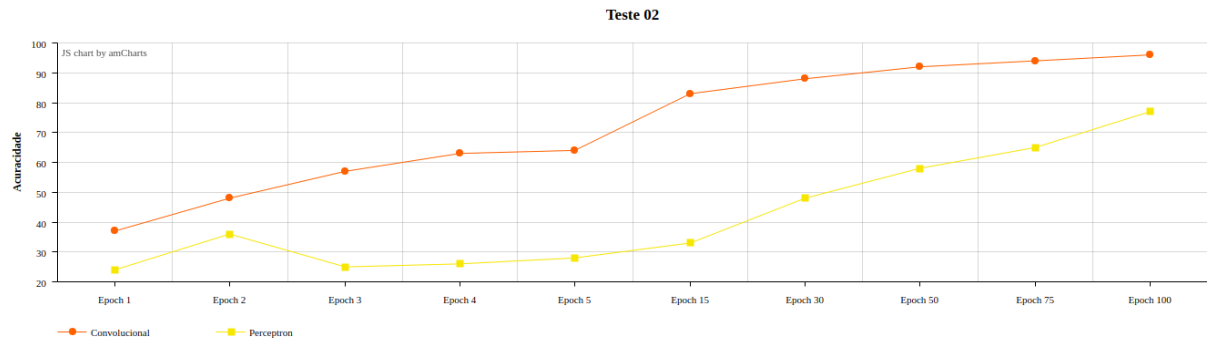
Fonte – Acervo do Autor, 2018

Como observado no gráfico anterior, a taxa de acuracidade da rede convolucional passa a ser mais alta que a perceptron a partir da segunda época, fato esse que perdura-se até o final do teste resultando em aproximadamente 91% de acuracidade da detecção dos dentes da camada convolucional em comparação a 75% retornado pelo modelo de rede neural perceptron, tornando a rede convolucional melhor adaptada.

5.2.2.2 Teste 02

O segundo teste realizado possui duas camadas internas, ambas tendo como função de ativação a ReLU, na qual a primeira camada possui 128 neurônios seguida pela segunda camada com 64 neurônios. Já a representação do modelo convolucional se dá pelo acréscimo de uma camada possuindo 32 filtros com dimensão 2x2. A figura 34 demonstra os resultados obtidos nas 100 primeiras épocas de treinamento de ambos os modelos de redes neurais.

Figura 34 – Teste 02 comparativo entre as RNAs



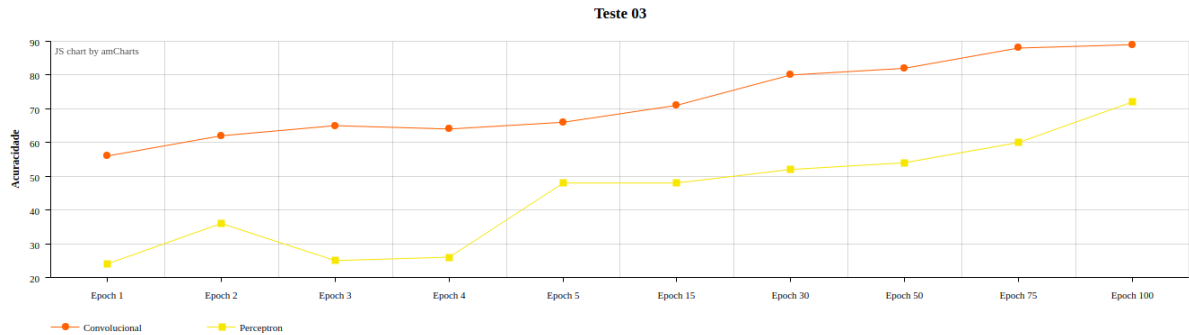
Fonte – Acervo do Autor, 2018

Como observado no gráfico do teste 02, a acuracidade da rede neural convolucional desde o início dos testes apresenta-se com melhor taxa, fato este que se mantém até o final do teste, no qual a rede convolucional possui cerca de 95% de acuracidade em comparação com a perceptron que possui 68%. A partir dos resultados no segundo teste, já é possível perceber a clara melhoria dos resultados quando é inclusa a camada convolucional, porém para melhorar o estudo serão realizados mais alguns testes.

5.2.2.3 Teste 03

O terceiro teste é extremamente semelhante ao teste anterior, a camada convolucional permanece idêntica, a única diferença apresentada é a alteração da função de ativação, nesta é utilizada a função Tahm. A figura 35 demonstra os resultados para as 100 primeiras épocas de treinamento.

Figura 35 – Teste 03 comparativo entre as RNAs



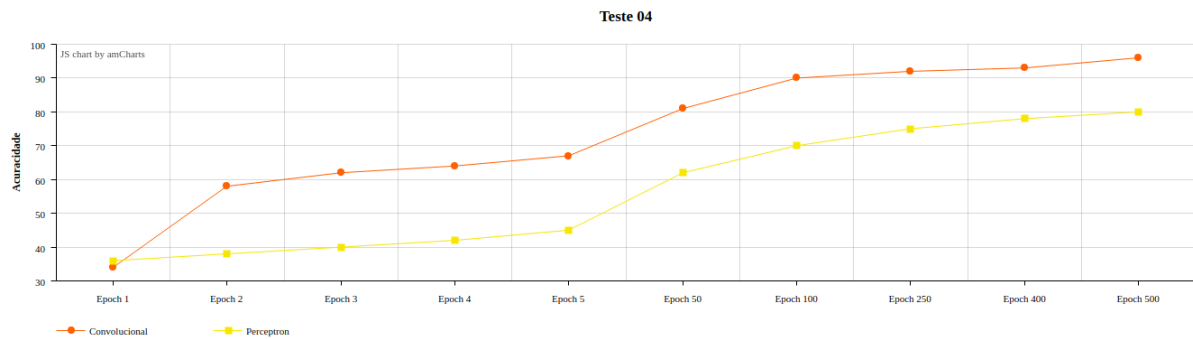
Fonte – Acervo do Autor, 2018

Como é possível observar no gráfico referente ao teste 03, o comportamento é semelhante ao teste anterior, a camada convolucional apresenta-se com melhor acuracidade desde a primeira etapa ao atingir as 100 épocas, ela possui uma taxa de acerto de cerca de 89% enquanto a rede neural que utiliza apenas a perceptron possui acuracidade de 71%.

5.2.2.4 Teste 04

Por fim, temos o último teste realizado, o qual é diferente dos demais testes por possuir duas camadas convolucionais, onde a primeira possui 32 neurônios com o filtro 2x2 seguido pela segunda camada de 16 neurônios com o filtro de 3x3. A camada perceptron também possui duas camadas, a primeira tendo 128 neurônios com a função de ativação Tahm e a segunda também possuindo 128 neurônios com a função de ativação Softmax. A figura 36 representa os resultados obtidos nas primeiras 500 épocas de treinamento.

Figura 36 – Teste 04 comparativo entre as RNAs



Fonte – Acervo do Autor, 2018

No gráfico apresentado para o teste 04 é visível a semelhança com o primeiro teste, no qual a camada convolutiva inicialmente apresenta-se como uma acuracidade menor, porém conforme as épocas vão passando ela torna-se a rede mais adaptada, chegando a 97% em comparação com a camada perceptron que chega ao máximo de 70%.

6 CONCLUSÕES

A área de inteligência artificial possui uma grande aplicabilidade em inúmeros setores, cada método desenvolvido possui características específicas para o problema proposto, uma destas é a área de Redes Neurais Artificiais, que possui como principal objetivo a capacidade de generalização, por meio da identificação de padrões e tendências entre as características dos dados informados a rede neural. Sendo mais específico ainda, dentro da área de Redes Neurais Artificiais é possível encontrar diversas subáreas, com distintos tipos de RNAs, essa variedade é de extrema valia quando se trata de identificação e processamento de dados, proporcionando uma maior quantidade de testes, onde algumas redes irão se adequar mais a determinados problemas que outras. A partir disso, este trabalho buscou apresentar a utilização de duas redes neurais na categorização de um mesmo problema, a fim de demonstrar qual a melhor solução.

Com o intuito de atingir os objetivos propostos por este trabalho inicialmente foi realizada uma pesquisa bibliográfica na literatura a fim de identificar trabalhos que seguissem a mesma linha de raciocínio que este trabalho. Para que fosse possível aplicar as técnicas de redes neurais artificiais, utilizando como base as redes Perceptron e Convolutacional propostas neste trabalho. Para isso foi necessário coletar as imagens em um base de dados online, em um pesquisa foi encontrado a base Mendeley Data, a qual contia cerca de 300 imagens radiográficas panorâmicas da arcada dentária, apresentando uma grande variedade de imagens possibilitando um melhor processamento por parte dos algoritmos.

De acordo com a proposta de identificar as imagens e categorizá-las, as imagens coletadas foram segmentadas manualmente, pois todas as imagens coletados seguiam o padrão panorâmico, o qual contempla todos os dentes da arcada dentária.

Com base nisso, foram desenvolvidas 4 diferentes tipos de redes neurais perceptron afim de categorizar as imagens segmentadas onde cada uma utilizou de uma camada convolutacional, a fim de provar se está teria melhor performance usufruindo dessa camada adicional, ou se por meio dos resultados provasse não necessária. Para avaliar o desempenho da rede foi utilizado os resultados fornecido pela própria ferramenta de desenvolvimento, o framework Keras. Tal resultado foi categorizado tanto em conjunto, ou seja com todos os tipos de dentes, quanto individualmente.

O melhor desempenho foi obtido pela camada convolutacional, em todos os testes foi visível a melhora na acuracidade quando a rede neural perceptron utilizou da camada convolu-

cional, em alguns casos pode-se visualizar que inicialmente a rede perceptron possuía melhor resultado, porém conforme as épocas continuavam, os resultados tornaram-se melhores para a camada convolucional, reforçando as informações obtidas durante o levantamento bibliográfico, no qual é apresentando que as redes convolucionais são mais efetivas quando se trata em processamento de imagem.

Como esse trabalho propôs-se a criar uma comparação entre os modelos de Redes Neurais Perceptron com o Modelo Convolucional na categorização dos dentes, deixou-se de lado certos aspectos como por exemplo a segmentação de forma automática, esta que seria uma excelente adição a este trabalho, tanto que é proposto como trabalho futuro.

Outra possibilidade de continuação para esse trabalho é a detecção de patologias na região dos dentes, como por exemplo cárie dentária. Para que então essa informação será apresentada a um profissional especializado.

REFERÊNCIAS

- ABDI, Amir; KASAEI, Shohreh. **Panoramic Dental X-rays With Segmented Mandibles**. [S.l.: s.n.], 2017. Disponível em: <<https://data.mendeley.com/datasets/hxt48yk462/1>>. Acesso em: 26 nov. 2018.
- ACADEMY, Data Science. **Deep Learning Book**. [S.l.: s.n.], 2018. Disponível em: <<http://deeplearningbook.com.br/capitulos/>>. Acesso em: 25 nov. 2018.
- ANDERSON, Dustin; VLIMANT, Jean-Roch; SPIROPULU, Maria. **An MPI-Based Python Framework for Distributed Training with Keras**. [S.l.]: California Institute of Technology, 2017.
- ARAÚJO, Flávio H. D. et al. **Redes Neurais Convolucionais com Tensorflow: Teoria e Prática**. [S.l.: s.n.], 2017.
- BRAGA, Antônio de Pádua; CARVALHO, André Ponce de Leon F. de; LUDERMIR, Teresa Bernarda. **Redes Neurais Artificiais: Teoria e Aplicações**. [S.l.: s.n.], 2007.
- CHIENG, Hock Hung et al. **Flatten-T Swish: a thresholded ReLU-Swish-like activation function for deep learning**. [S.l.]: International Journal of Advances in Intelligent Informatics, 2018.
- COLGATE. **Anatomia Dos Dentes**. [S.l.]: COLGATE, 2018. Disponível em: <<https://www.colgate.com.br/oral-health/basics/mouth-and-teeth-anatomy/tooth-anatomy>>. Acesso em: 22 nov. 2018.
- DANCE, D.R. et al. **Diagnostic Radiology Physics A Handbook for Teachers and Students**. [S.l.]: International atomic energy agency, 2014.
- GOLDSCHIMIDT, Ronaldo Ribeiro. **Uma introdução a inteligência computacional: Uma introdução a ciência computacional**. [S.l.]: Diagnostic Radiology Physics, 2014.
- GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. [S.l.: s.n.], 2018. Disponível em: <<https://www.deeplearningbook.org/>>. Acesso em: 24 nov. 2018.
- GRANATYR, Jones; AMARAL, Fernando. **Deep Learning com Python de A à Z - O Curso Completo**. [S.l.: s.n.], 2018. Disponível em: <<https://www.udemy.com/deep-learning-com-python-az-curso-completo/learn/v4/overview>>. Acesso em: 26 nov. 2018.
- NUNES, Fátima L. S. **Introdução ao Processamento de Imagens Médicas para Auxílio ao Diagnóstico**. [S.l.: s.n.], 2010.
- OLIVEIRA, João Paulo Ribeiro de. **Caries Detection in Panoramic Dental X-ray Images**. [S.l.]: Universidade da Beira Interior, 2009.
- PHONG, Le Trieu; PHUONG, Tran Thi. **Privacy-Preserving Deep Learning for any Activation Function**. [S.l.]: National Institute of Information and Communications Technology, 2018.
- SILVA, Ivan Nunes da; SPATTI, Danilo Hernane; FLAUZINO, Rogerio Andrade. **Redes Neurais Artificiais Para Engenharia E Ciências Aplicadas**. [S.l.: s.n.], 2010.

TOOLS, IHMC Cmap. **Neurônio biológico**. [S.l.: s.n.], 2015. Disponível em: <<https://www.colgate.com.br/oral-health/basics/mouth-and-teeth-anatomy/tooth-anatomy>>. Acesso em: 22 nov. 2018.

VIERGEVER, Max A. et al. **A survey of medical image registration – under review**. [S.l.]: Medical Image Analysis, 2016.

WANG, Tao et al. **End-to-End Text Recognition with Convolutional Neural Networks**. [S.l.]: Stanford University, 2015.

YU, Young jun. **Machine Learning for Dental Image Analysis**. [S.l.]: Department of Computer Science, 2017.