UNIVERSIDADE POSITIVO

BACHARELADO EM SISTEMAS DE INFORMAÇÃO

EDUARDO NASCIMENTO DA SILVA

OSMARY CAMILA BORTONCELLO GLOBER

**SAIPRO**

SISTEMAS DE ANÁLISE PARA INDICAÇÃO DE ANOMALIAS NAS RADIOGRAFIAS PANORÂMICAS ODONTOLÓGICAS

CURITIBA

2017

EDUARDO NASCIMENTO DA SILVA

OSMARY CAMILA BORTONCELLO GLOBER

**SAIPRO**

SISTEMAS DE ANÁLISE PARA INDICAÇÃO DE ANOMALIAS NAS RADIOGRAFIAS PANORÂMICAS ODONTOLÓGICAS

Proposta para o Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do grau de Bacharel em Sistemas de Informação da Universidade Positivo.

Orientador: Prof. Leandro Escobar

CURITIBA

2017

RESUMO

Muito utilizadas pelos profissionais odontológicos no cotidiano as radiografias panorâmicas possuem certo grau de dificuldade para diagnóstico de determinadas doenças, devido a sobreposição e sombras encontrados no exame. A utilização do processamento digital de imagens para filtrar alguns dos padrões dessas doenças juntamente com a inteligência artificial para análise pode constituir uma alternativa para compensar tal deficiência, porém atualmente os profissionais utilizam a tecnologia somente para tarefas administrativas e no controle do tratamento dos pacientes. O problema de pesquisa abordado se relaciona ao fato de que devido a radiografia panorâmica ser um exame bidimensional de um corpo tridimensional vários fatores podem contribuir para que algumas doenças não sejam diagnosticadas com precisão. Assim, este trabalho apresenta como podemos unir a inteligência artificial com processamento digital de imagens afim de dar um maior apoio à tomada de decisão dos profissionais no momento do diagnóstico, mediante a identificação e análise de padrões característicos encontrados nesses exames. Afim de extrair informações das anomalias para gerar uma máquina de conhecimento um protótipo foi desenvolvido com a capacidade de obter padrões e também de apontar pontos de atenção nos exames para os profissionais, sendo uma segunda opinião para os mesmos. Os resultados esperados com esta pesquisa são diminuir a dificuldade no diagnóstico de doenças que dificultam sua visualização em radiografias panorâmicas contribuindo assim para que os profissionais odontológicos tenham um maior apoio a decisão e com isso aumentar a qualidade no diagnóstico dos pacientes.

**Palavras-chave:** Inteligência Artificial; Processamento Digital de Imagens; Apoio à decisão na Odontologia; Radiografia Panorâmica.

**ABSTRACT**

Often used by dental professionals in daily life, panoramic radiographs have a certain degree of difficulty in the diagnosis of certain diseases, due to the occlusion and shadows found in the examination. The use of digital image processing to filter out some of the patterns of these diseases together with artificial intelligence for analysis may be an alternative to compensate for such deficiency, but today professionals use the technology only for administrative tasks and patient management control. The research problem addressed is related to the fact that because panoramic radiography is a two-dimensional examination of a three-dimensional body several factors can contribute to some diseases being not accurately diagnosed. Thus, this work presents how we can unite artificial intelligence with digital image processing in order to give greater support to the professionals' decision making at the time of diagnosis, by identifying and analyzing the characteristic patterns found in these exams. In order to extract information from the anomalies to generate a knowledge machine a prototype was developed with the ability to obtain standards and also to point out points of attention in the exams for the professionals, being a second opinion for them. The expected results with this research are to reduce the difficulty in the diagnosis of diseases that make it difficult to see them in panoramic radiographs, thus contributing to the dental professionals have a greater decision support and thus increase the quality in the diagnosis of the patients.

**Keywords:** Artificial Intelligence; Digital Image Processing; Decision support in dentistry; Panoramic Radiography.

SUMÁRO

[ABSTRACT 4](#_Toc496695731)

[1 INTRODUÇÃO 7](#_Toc496695732)

[1.1 JUSTIFICATIVA 9](#_Toc496695733)

[1.2 OBJETIVOS 10](#_Toc496695734)

[1.2.1 Objetivo geral 10](#_Toc496695735)

[1.2.2 Objetivos específicos 10](#_Toc496695736)

[1.3 METODOLOGIA 10](#_Toc496695737)

[1.3.1 Etapa 1 - Processamento de imagem 11](#_Toc496695738)

[1.3.2 Etapa 2 – Protótipo 11](#_Toc496695739)

[1.3.3 Etapa 3 – Aprendizado de Máquina 12](#_Toc496695740)

[2 FUNDAMENTAÇÃO TEÓRICA 13](#_Toc496695741)

[2.1 IMAGEM RADIOGRÁFICA DIGITAL 13](#_Toc496695743)

[2.1.1 Radiografia Panorâmica Digital 14](#_Toc496695744)

[2.1.2 Dificuldade na Visualização de Anomalias 14](#_Toc496695745)

[2.2 PROCESSAMENTO DIGITAL DE IMAGENS 14](#_Toc496695746)

[2.2.1 Morfologia Matemática 15](#_Toc496695747)

[2.2.2 Dilatação Binária 16](#_Toc496695748)

[2.2.3 Segmentação de Imagens 17](#_Toc496695749)

[2.2.4 Segmentação por Contorno 17](#_Toc496695750)

[2.2.5 Canny 18](#_Toc496695751)

[2.2.6 Histogramas 19](#_Toc496695752)

[2.2.7 Equalização de Histogramas 19](#_Toc496695753)

[2.2.8 Subtração de Imagens 20](#_Toc496695754)

[2.2.9 Máscara Binária 21](#_Toc496695755)

[2.2.10 Filtros 21](#_Toc496695756)

[2.2.11 Filtros Gaussianos 22](#_Toc496695757)

[2.2.1 Classificador Haar 22](#_Toc496695758)

[2.2.2 Curvas de Bézier 24](#_Toc496695759)

[2.2.3 OpenCV 25](#_Toc496695760)

[2.2.4 Textura 25](#_Toc496695761)

[2.2.5 Matriz de Co Ocorrência 26](#_Toc496695762)

[2.3 ECOSSISTEMA BIG DATA 26](#_Toc496695763)

[2.3.1 BIG DATA 26](#_Toc496695764)

[2.3.2 MAPREDUCE 27](#_Toc496695765)

[2.3.3 HADOOP 1.0 27](#_Toc496695766)

[2.3.4 HADOOP 2.0 28](#_Toc496695767)

[2.3.5 CLOUDERA MANAGER 28](#_Toc496695768)

[2.3.6 APACHE CASSANDRA 29](#_Toc496695769)

[2.3.7 MINERAÇÃO DE DADOS 29](#_Toc496695770)

[2.3.8 WEKA 29](#_Toc496695771)

[2.3.9 CLASSIFICAÇÃO 30](#_Toc496695772)

[3 DESENVOLVIMENTO 31](#_Toc496695773)

[3.1 INTRODUÇÃO 31](#_Toc496695775)

[3.2 MÉTODO 31](#_Toc496695776)

[3.2.1 As imagens odontológicas 32](#_Toc496695777)

[3.2.2 Processamento de Imagem 32](#_Toc496695778)

[3.2.3 A Máquina de Aprendizado 39](#_Toc496695779)

[3.2.4 O Pós-processamento 40](#_Toc496695780)

[3.3 O PROTÓTIPO 41](#_Toc496695781)

[4 GESTÃO 43](#_Toc496695782)

[4.1 ESCOPO 43](#_Toc496695784)

[4.2 CRONOGRAMA 44](#_Toc496695785)

[4.3 CUSTOS 45](#_Toc496695786)

[4.4 RISCOS 45](#_Toc496695787)

[5 AVALIAÇÃO 47](#_Toc496695788)

[6 RESULTADOS OBTIDOS 48](#_Toc496695789)

[7 CONCLUSÃO 50](#_Toc496695790)

[8 REFERÊNCIAS 51](#_Toc496695791)

[9 APÊNDICE 56](#_Toc496695792)

[10 ANEXO 59](#_Toc496695793)

# INTRODUÇÃO

Atualmente o Brasil possui um total de 19% de todos os dentistas do mundo sendo o país com maior número de dentistas, em números absolutos são 219.575 profissionais cadastrados (CONSELHO FEDERAL DE ODONTOLOGIA, 2017). Visto que há uma grande parcela de profissionais na área odontológica, a necessidade de aplicativos e sistemas afim de automatizar e facilitar as tarefas do cotidiano desses profissionais, de maneira eficaz e eficiente, vem aumentando significativamente ao longo dos anos. A maioria dessas ferramentas efetuam tarefas de cunho administrativo e controle no tratamento dos pacientes (via odontograma) limitando assim o uso de todo o potencial tecnológico a tarefas que exigem uma menor tomada de decisão do profissional (DOTTA e TELES, 2003).

O uso de radiografias panorâmicas para diagnóstico de diversas doenças odontológicas é um excelente recurso pois seu custo é baixo, o nível de radiação é muito inferior e há uma facilidade na realização deste exame em comparação a exames tridimensionais como as tomografias computadorizadas. Porém a radiografia panorâmica possui algumas limitações, sendo que é um exame bidimensional de um corpo tridimensional, o que pode dificultar o diagnóstico de doenças como lesões periapicais, fraturas radiculares e reabsorções dentais (BERNARDES, 2007) e lesões apicais. Isto ocorre devido à sobreposição das estruturas na radiografia e à formação de sombras, sendo que a coluna vertebral e o crânio são os principais responsáveis por essa dificuldade (ALMEIDA, 2001).

Desta forma, uma alternativa para compensar esta carência de ferramentas tecnologicas na área odontológica e também de ajudar os especialistas na tomada de decisão de um diagnóstico, utilizando radiografias panorâmicas, é prover os profissionais de soluções inovadoras que consigam ajudar nas decisões a serem tomadas e no cuidado terapêutico, aumentando a confiança e a qualidade do diagnóstico para os pacientes (ZANET, 2009).

Assim, a análise feita em radiografias através do processamento digital de imagens unida com o reconhecimento de padrões tem sido uma ferramenta muito eficiente na melhoria da detecção e na classificação de lesões dos pacientes, indicando como pode ser bastante promissor a utilização de sistemas que possam indicar um diagnóstico automatizado do paciente ao profissional (DE AZEVEDO-MARQUES, 2001).

A descoberta de padrões que a inteligência artificial nos proporciona, com os dados já previamente extraídos, com o processamento digital de imagens, separando os padrões normais e anormais, conseguem melhorar a visualização de anomalias e doenças para que o profissional tenha uma tomada de decisão mais eficiente em relação ao diagnóstico dos pacientes (DE AZEVEDO-MARQUES, 2001).

No entanto muitos artigos consideram que os sistemas que fazem esta análise automatizada em radiografias seriam somente um auxílio aos especialistas e não um substituto completo, podendo chegar ao mesmo desempenho dos profissionais, mas isentando uma total e perfeita assertividade destes sistemas nos diagnósticos (DE AZEVEDO-MARQUES, 2001).

O problema em questão configura-se no fato de que, embora as pesquisas sejam constantes, devido a radiografia panorâmica ser um exame bidimensional vários fatores podem contribuir para que algumas doenças não sejam diagnosticadas precocemente com precisão ou até em seu estado habitual, pois o exame possui algumas dificuldades na visualização de determinadas anomalias pelos profissionais.

É neste contexto que emerge a questão central desta pesquisa: É possível diminuir a dificuldade no diagnóstico em radiografias panorâmicas? As respostas a esta questão giram em torno de como podemos unir o processamento digital de imagens com a inteligência artificial, afim de dar um maior apoio à tomada de decisão dos profissionais utilizando radiografias panorâmicas.

Esta pesquisa se apoia, então, no pressuposto de que a dificuldade no diagnóstico de determinadas anomalias e doenças odontológicas em radiografias panorâmicas limita os profissionais na tomada de decisão no momento de realizar o diagnóstico de pacientes.

Enfim, os padrões extraídos das radiografias panorâmicas através da inteligência artificial e processamento de imagens podem oferecer aos profissionais uma detecção mais aguçada de algumas anomalias que possuem certo grau de dificuldade de diagnóstico, ocorrendo o ganho de efetividade das decisões em benefício do tratamento a ser ministrado nos pacientes.

Por outro lado, esta pesquisa está limitada a somente um tipo de exame, a radiografia panorâmica, ensejando que as aplicações das técnicas de processamento de imagens e inteligência artificial apresentadas podem trazer novas perspectivas a não só a área odontológica, mas também a diversas outras que possuem as mesmas dificuldades. Outra limitação é a detecção de somente determinados padrões de algumas doenças já pré-estabelecidas por profissionais odontológicos. Estudos que possam fornecer mais padrões de anomalias odontológicas podem enriquecer as conclusões, ampliando a contribuição científica e social da pesquisa aqui apresentada.

## JUSTIFICATIVA

O motivo que nos levam a estudar este problema é a constatação que há a uma necessidade de aplicativos e sistemas na área da odontológica e que há dificuldade por parte dos profissionais no diagnóstico de determinadas anomalias em radiografias panorâmicas devido a limitações do próprio exame (BERNARDES, 2007). Por outro lado, se aliarmos alta tecnologia com técnicas de processamento digital de imagens junto com inteligência artificial podemos contribuir para a análise dos padrões dessas doenças afim de auxiliar estes profissionais. A pesquisa se justifica porque o processamento digital de imagens pode destacar estas anomalias nas radiografias e a inteligência artificial consegue fornecer uma análise mais assertiva em relação a estes padrões, permitindo que as identificações de doenças difíceis de visualizar sejam indicadas aos profissionais, auxiliando em uma segunda opinião sobre o diagnóstico dos pacientes e com isso aumentando a qualidade no diagnóstico.

A contribuição científica desta pesquisa está no reconhecimento de padrões de anomalias que possuam dificuldade de visualização em radiografias panorâmicas utilizando processamento de imagens e inteligência artificial para a sua detecção, indicando doenças que poderiam passar despercebidas, facilitando a visualização destas pelos profissionais, servindo como base para o uso de técnicas válidas para o desenvolvimento de sistemas que fornecem apoio à decisão no diagnóstico utilizando radiografias panorâmicas na área odontológica. A descoberta de padrões nas imagens também contribui para pesquisas futuras, uma vez que as técnicas utilizadas para esta detecção possibilitam o seu uso em outras áreas.

Socialmente, o projeto contribui com a proposta de auxiliar os profissionais odontológicos fornecendo uma segunda opinião no momento do diagnóstico dos pacientes, podendo aumentar a detecção de anomalias que possuam dificuldade de serem visualizadas nestes exames, contribuindo assim para um ganho de eficiência nos diagnósticos destas doenças e também para o ganho da qualidade de vida dos pacientes, por conta de uma qualidade maior nos seus tratamentos.

## OBJETIVOS

### Objetivo geral

Desenvolver um sistema para diminuir a dificuldade no diagnóstico de doenças odontológicas difíceis de identificar em radiografias panorâmicas, mediante a análise e comparação de padrões encontrados nessas doenças.

### Objetivos específicos

1. Pesquisar técnicas de processamento de imagens específicas para radiografias panorâmicas que atendam às necessidades de identificação das anomalias.
2. Desenvolver filtros para pré-processamento dessas imagens para gerar padrões específicos encontrados nessas doenças.
3. Gerar uma base de conhecimento utilizando as características encontradas para aumentar a eficácia na análise do diagnóstico.
4. Desenvolver um modelo que apresente aos profissionais odontológicos as regiões que possuem pontos de atenção ao diagnóstico.

## METODOLOGIA

Esta pesquisa relaciona-se com o desenvolvimento e a utilização de uma ferramenta tecnológica nos diagnósticos feitos pelos profissionais odontológicos e apoia-se na identificação e análise de padrões de doenças que possuam certo grau de dificuldade de visualização em radiografias panorâmicas.

Assim, o estudo em questão envolve uma pesquisa qualitativa, descritiva, aplicada e de raciocínio indutivo de forma a identificar as dificuldades nos diagnósticos de doenças de difícil visualização nos exames, bem como desenvolver um protótipo baseado na análise de padrões, que permita aumentar a confiança na tomada de decisão dos profissionais.

Buscando a uniformização de termos, gestores, técnicos, especialistas ou quaisquer outras pessoas que possam ou venham a contribuir com o entendimento das especificidades do domínio Odontológico, serão denominados “profissionais da área odontológica” ou tão somente “profissionais”, generalizando os papeis atuantes na área da Odontologia, de forma a facilitar o entendimento quando se tratar de declarar qual papel está envolvido no contexto.

A pesquisa aqui apresentada está organizada em três etapas partindo com o levantamento de técnicas de processamento de imagem utilizadas em radiografias odontológicas, passando pelo desenvolvimento do protótipo afim de extrair os padrões encontrados e finalizando ao gerar uma máquina de conhecimento que analise e indique aos profissionais as áreas que possam ter alguma anomalia.

Para facilitar a visualização da metodologia aplicada, segue abaixo uma figura com o diagrama da parte metodológica da pesquisa.

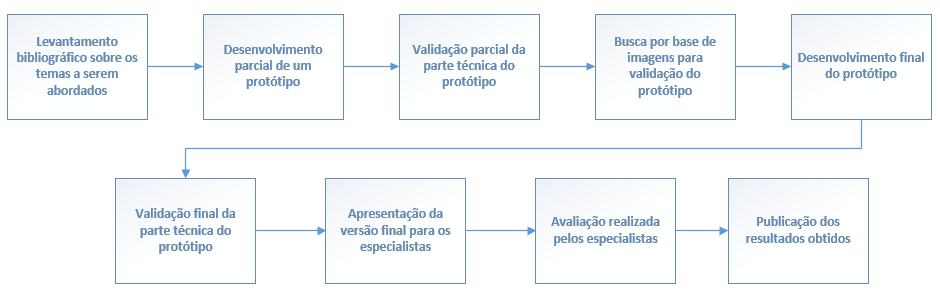


Figura 1 – Imagem com o diagrama da metodologia aplicada. Fonte: Autoria Própria.

Cada etapa, por sua vez, está organizada em uma série de procedimentos, os quais constituem o caminho para sua realização e efetivação de seus resultados asseada nos seus resultados preliminares.

### Etapa 1 - Processamento de imagem

Levantar técnicas de processamento de imagens utilizadas em radiografias odontológicas trata-se de uma pesquisa sistemática com o objetivo de identificar as técnicas recomendadas pela literatura que sejam utilizadas para a extração de características das anomalias nos exames. Esta extração serve como base para prototipação do modelo proposto, de forma que quanto mais detalhada a informação mais precisa será a análise feita pela máquina de conhecimento, e desta forma, mais assertivo será a indicação das áreas problemáticas nos exames.

### Etapa 2 – Protótipo

Desenvolver o protótipo trata-se da aplicação das técnicas (oriundos da Etapa 1) em exames para verificar se os padrões extraídos compõem o contexto necessário para gerar a máquina de conhecimento para análise. Este contexto será avaliado por profissionais da área para condizer com os padrões de identificação da doença a olho nu e que possuam dificuldade de visualização.

### Etapa 3 – Aprendizado de Máquina

Gerar a máquina de conhecimento para análise e indicação de área com problema trata-se de utilizar todos os padrões, extraídos e validados por profissionais, das imagens afim de gerar uma máquina de conhecimento que utilize todos estes para a análise de anomalias e aumento de assertividade dos resultados em novos exames sendo uma segunda opinião para os profissionais.

# FUNDAMENTAÇÃO TEÓRICA

Neste capítulo iremos apresentar algumas técnicas e conceitos importantes que foram utilizados nos métodos e no desenvolvimento desta pesquisa. Serão mostrados de forma introdutória para que o leitor possa ter uma compreensão correta sobre a pesquisa.



## IMAGEM RADIOGRÁFICA DIGITAL

Uma imagem radiográfica se forma através de uma explosão controlada de radiação de raios-X, que antes de colidir com o filme ou sensor, conseguem penetrar em estruturas que possuem diferentes níveis. Após essa colisão a película passa por alguns processos lentos com produtos químicos em um ambiente sem luz, pois possuem sensibilidade alta. A exposição à radiação é feita em uma dose pequena sendo que se assemelha a alguns dias de sol, podendo ser menor ainda com a utilização de um avental de chumbo, porém mesmo com os níveis de radiação sendo baixos, este processo pode gerar algumas distorções ou erros expondo o paciente novamente a mais doses de radiação (SIRLOPÚ, 2016).

Com isto, a radiologia digital começou a ser introduzida como uma nova tecnologia entre os anos de 1970 e 1980, porém foi na Primeira Conferência Internacional e Workshop de Arquivamento de Figuras e Sistemas de Comunicação (PACS) realizada em 1982 que a tecnologia conseguiu amadurecer para dar início a uma nova era no campo da radiologia (VELA, 2011).

Com este avanço passou-se a utilizar um sensor eletrônico ao invés do filme convencional e arquivos digitais para manipulação destas imagens. Sirlopú (2016) destaca que as imagens radiográficas digitais possuem um processamento menor e podem exigir uma menor exposição dos pacientes a radiação. Contudo, devido aos preços dos sensores digitais serem extremamente caros há uma dificuldade na substituição do método convencional para o método digital (SIRLOPÚ, 2016; VELA, 2011).

Sendo assim as imagens radiográfica digitais possuem diversas vantagens na sua utilização, pois devido ao fato de estarem armazenadas em arquivos digitais elas podem ser utilizadas em sistemas que auxiliem numa segunda opinião e também na utilização para treinamentos em processos de ensino (VELA, 2011).

As imagens radiográficas digitais têm uma importância enorme neste projeto pois é com elas que iremos efetuar todos os passos.

### Radiografia Panorâmica Digital

Atualmente as radiografias panorâmicas tem fundamental importância, pois auxiliam no diagnóstico, planejamento do tratamento a ser realizado e na proservação dos pacientes, sendo que devido a estes fatores são muito solicitadas pelos profissionais pois são consideradas um método radiográfico prático e moderno que consegue mostrar aos profissionais todo o complexo maxilomandibular, toda a região dento-alveolar e as estruturas adjacentes, conseguindo assim uma menor incidência de complicações e redução de custo aos pacientes (GONDIM, 2010).

Devido a esta importância no meio odontológico iremos utilizar este tipo de exame para desenvolvermos todo o nosso protótipo.

### Dificuldade na Visualização de Anomalias

Como mencionado anteriormente, os profissionais odontológicos possuem certa dificuldade para visualização de determinadas doenças em radiografias panorâmicas devido as limitações impostas pelo próprio exame (ALMEIDA, 2001), porém desenvolver uma ferramenta que possa auxiliar neste processo não é uma tarefa trivial, sendo que devido a vários fatores tecnológicos há uma série de dificuldades a serem enfrentadas (ZANET, 2009).

Tendo em vista a visualização das anomalias feitas por profissionais sendo que este processo ainda pode gerar erros (ALMEIDA, 2001), que podem ser muito ruins aos pacientes visamos focar todo o desenvolvimento em cima desta dificuldade de visualização afim de melhorar a tomada de decisão dos profissionais da odontologia.

## PROCESSAMENTO DIGITAL DE IMAGENS

O Processamento Digital de Imagens (PDI) proporciona recursos poderosos e de grande utilidade para manipulação de imagens em aplicações que necessitem de uma melhor informação visual para interpretação humana ou interpretação de sistemas computacionais ou também para automatização de processos que demandem tarefas que envolvam análise de imagens (AZEVEDO, 2014).

Assim sendo, processamento de imagens é um conjunto de etapas conectadas que utilizam várias operações que são feitas sobre determinadas imagens e que resultam em novas imagens (DE QUEIROZ, 2006) sendo que algumas técnicas utilizadas em uma área podem ser inúteis em outra área tornando-se uma solução específica (GONZALES e WOODS, 2010).

Estas etapas estão ilustradas no diagrama de blocos da Figura 2.

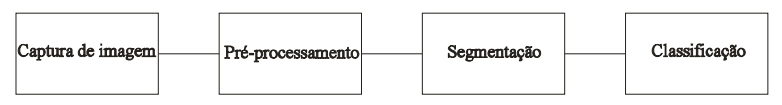


Figura 2 - Diagrama de blocos das etapas básicas do processamento digital de imagens. Fonte: SOUZA, 2007, p. 2.

Começando pela captura de imagem, que consiste na utilização de alguns dispositivos sensíveis à espectros de energia eletromagnética e estes passam o sinal elétrico para o digital. Na etapa do pré-processamento são utilizadas algumas técnicas para realce de características ou para recuperar imagens com perca significativa. Já a segmentação é a separação da imagem em regiões para que seja feita a extração de objetos. Na classificação é aonde acontece a identificação da imagem que está sendo processada (SOUZA, 2007).

Para podermos identificar os padrões de anomalias nas imagens, tivemos que processá-las digitalmente e utilizamos diversos recursos do Processamento Digital de Imagens. Alguns conceitos de processamento digital de imagens serão apresentados nas subseções a seguir para uma melhor compreensão desta pesquisa por parte do leitor.

### Morfologia Matemática

Elaborada por Georges Matheron e Jean Serra nos anos de 1960 e 1969, a morfologia matemática analisa e estuda as estruturas geométricas que se encontram em uma imagem utilizando-se de ferramentas matemáticas (FACON, 2011).

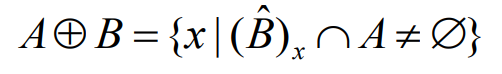
Ela tem sua base na extração de dados das imagens que englobem a geometria e a topologia de conjuntos desconhecidos utilizando um elemento estruturante, sendo que este é um conjunto conhecido e definido em forma e tamanho, realizando uma transformação e sendo comparado aos conjuntos desconhecidos da imagem. Basicamente ocorre a quantificação do elemento estruturante perante a imagem, para saber se ele se encaixa na imagem ou não, ou seja, se “está ou não está contido” no conjunto da imagem, e com isto podemos avaliar os resultados desta comparação e retirar informações relevantes sobre características e padrões que encontramos nas imagens, sendo que quanto maior for o elemento estruturante, maior serão os efeitos das operações morfológicas (FACON, 2011).

A principal questão da morfologia matemática se trata de decompor uma classe de operadores morfológicos utilizando operadores usuais de conjuntos como união, intersecção e complemento, sendo que dentre estes operadores podemos destacar os dois pilares que são a dilatação e a erosão, pois estes utilizam propriedades de adição e de subtração de Minkowski, e combinando-as podemos obter resultados significativos na identificação de características dentro das imagens (MACHADO, 2008).

Esta combinação de propriedades auxilia deixando as anomalias nas imagens maiores do que os seus respectivos tamanhos, e com isso podemos aumentar o tamanho de regiões que possuem problemas, mas que estão em um tamanho muito pequeno dificultando sua visualização pelos profissionais e com isso podemos extrair suas características de maneira eficaz.

### Dilatação Binária

Conhecida por ser mais uma das principais técnicas utilizadas na morfologia matemática a dilatação binária possui como efeito contrário da erosão o aumento dos conjuntos e caso o elemento estruturante seja maior pode até conecta-los, sendo que esta técnica também esta técnica pode diminuir e preencher cavidades. A dilatação de um conjunto A pelo elemento estruturante B é definida como:



O conjunto resultante da dilatação de A por B é o conjunto de todos os deslocamentos de x tais que B refletido e A se sobreponham em pelo menos um elemento não nulo (FACON, 2011). Para uma ilustração melhor ilustração a Figura 3 mostra um exemplo de aplicação desta técnica.

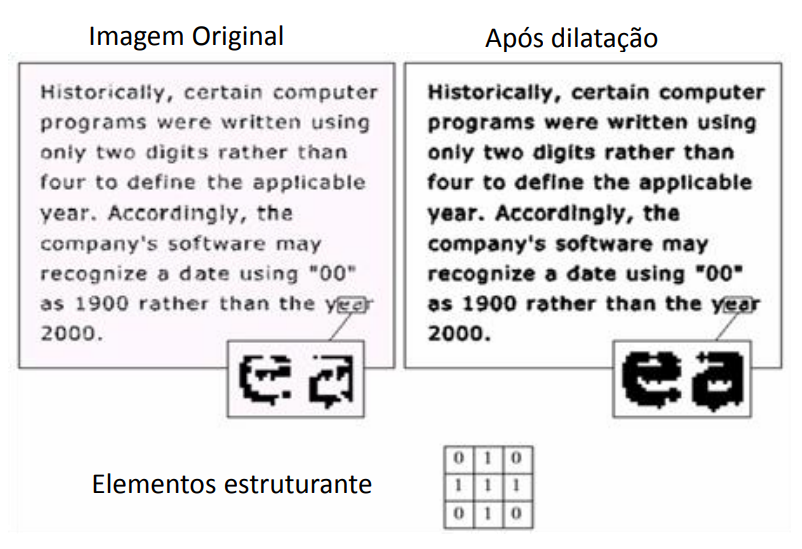


Figura 3 – Exemplo de aplicação dilatação. Fonte: Repositório da Universidade Federal de Goiás.

Disponível em: http://www.inf.ufg.br/~fabrizzio/mestrado/pdi/aulas/aula9.pdf.

Acesso em ago. 2017.

A dilatação é usada nas imagens de radiografias panorâmicas para aumentarmos o tamanho das regiões que segmentamos e com isso deixar as anomalias mais fáceis para a aplicação de outros filtros no decorrer do algoritmo.

### Segmentação de Imagens

A segmentação de imagens dentro do processamento digital de imagens se dá através da subdivisão de uma imagem dentre suas próprias partes baseada em alguma característica para que seja feita algum tipo de análise. A segmentação propriamente dite não necessita ser feita em níveis de cinza porém nesta pesquisa efetuamos a segmentação utilizando imagens em níveis de cinza, ou seja, a imagem é representada em um conjunto de números que vai de 0 a 255, sendo que cada equivale a um pixel da imagem e mostra um nível de cinza. Todas as técnicas utilizadas se baseiam em duas propriedades que podemos observar nos níveis de cinza: descontinuidade e similaridade (NEVES, 2008), porém não existe um método único que seja capaz de realizar a segmentação de todos os tipos de imagem (DE ALBUQUERQUE, 2000).

A segmentação das imagens é utilizada para extrair regiões de interesse da imagem afim de facilitar a aplicação de filtros e de posteriormente fazer a análise destas imagens.

### Segmentação por Contorno

A segmentação baseada em contornos é tida como uma segmentação que utiliza descontinuidade, ou seja, a técnica baseia-se na distinção de objetos detectando suas bordas e tentando construir uma região que se conecte interiormente. Nesta distinção são verificados os valores da variação dos valores de luminosidade das imagens, sendo que o objetivo principal é realçar os pixels da borda, podendo ser feita através técnicas que consigam aumentar o contrate da imagem (NUNES, 2007).

Os algoritmos utilizados para a segmentação detectando bordas contam com três passos principais: a filtragem, a salientação e a detecção, sendo que em alguns casos pode ainda se utilizar uma quarta etapa para refinamento ou localização (SILVA, 2011).

Ainda utilizando este tipo de segmentação pode se obter resultados melhores utilizando algumas das técnicas de suavização da imagem, sendo que estas podem reduzir ruídos nas imagens, porém este processo pode comprometer a precisão da técnica, pois deixará a imagem com as bordas espalhadas e com isso a espessura da borda irá aumentar, dificultando a sua localização. Diante destes fatores desenvolver um algoritmo que consigo localizar com precisão bordas em contextos diferentes é uma tarefa muito difícil sendo que há diversas contribuição no meio científico para detectores de borda para várias áreas (NUNES, 2007).

A segmentação por contorno é utilizada quando conseguimos identificar uma região de interesse e queremos saber o contorno desta para que possamos extrair informações relevantes para o protótipo.

### Canny

O filtro de Canny é muito utilizado por diversos pesquisadores e se baseia em duas etapas principais, a etapa de detecção e localização. Na primeira etapa de processamento de imagens foi realizado o pré-processamento de suavização utilizando o filtro Gaussiano para que seja retirado qualquer eventual ruído que a imagem possa ter e logo após é calculada a intensidade da borda e sua direção em cada pixel da imagem suavizada (DO VALE, 2002).

Porém mesmo com a aplicação deste processo, a imagem poderá ter fragmentos de bordas falsas que são causados devido a ruídos de textura da imagem. Para que sejam eliminados estes fragmentos pode-se fazer necessário um processo de limiarização na imagem (DO VALE, 2002).

Já na segunda etapa é onde ocorre a identificação de quais os pixels poderão ser a borda, sendo que estes são identificados utilizando um processo de afinamento que é chamado de non-maximal suppression, que é um processo que utiliza afinamento através de um limiar inferior e um limiar superior, ou seja, somente os pixels que possuem magnitude do gradiente entre estes valores são considerados bordas na imagem (DO VALE, 2002).

****

Figura 4 - Exemplo de aplicação do filtro Canny. Fonte: Documentação OpenCV.

Disponível em: http://docs.opencv.org/3.1.0/da/d5c/tutorial\_canny\_detector.html. Acesso em ago. 2017.

Utilizamos o filtro de Canny para destacarmos as regiões com anomalias e extrairmos dados sobre seus contornos e propriedades.

### Histogramas

Conhecidos por serem ferramentas de grande aplicação prática no processamento digital de imagens, os histogramas são muito utilizados para melhorar a definição de uma imagem, para facilitar a compreensão dela e também para a segmentação da mesma (MARENGONI, 2009). Um histograma é determinado pelos valores de intensidade dos pixels, sendo que pode ser de extrema utilidade para alterações na imagem como um todo e não é possível aplicar esta função em um processo que necessite saber, por exemplo, a localização de determinados pixels (NUNES, 2006).

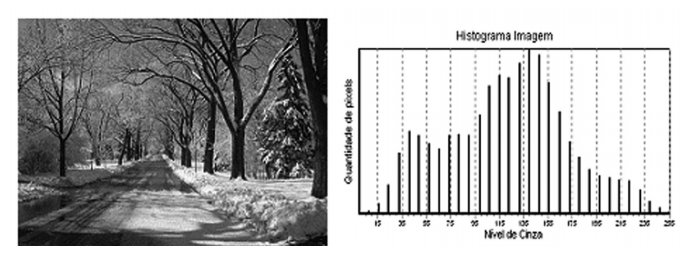


Figura 5 - Exemplo de histograma de uma imagem. Fonte: NUNES, 2006, p. 86

### Equalização de Histogramas

Está técnica é utilizada para realçar o contraste das imagens ajustando os seus valores de intensidade, espalhando a distribuição dos níveis de cinza por toda a imagem (MARENGONI, 2009). A imagem resultante após a utilização deste processo possui um histograma de intensidade mais uniforme em relação aos níveis de cinza, sendo que em algumas das vezes que esta técnica é aplicada consegue recuperar imagens consideradas perdidas (NUNES, 2006).

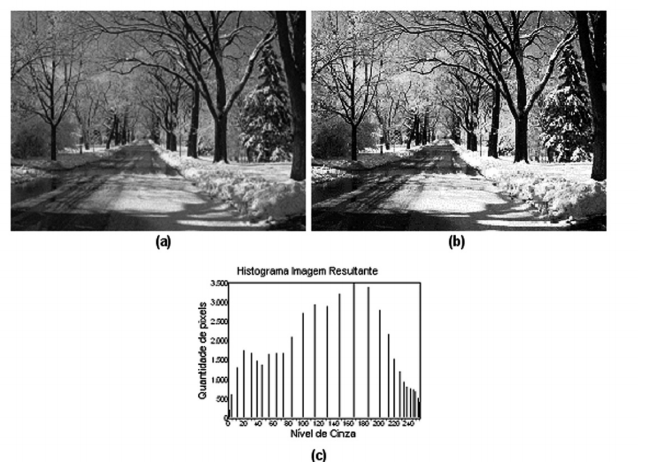


Figura 6 - Exemplo de equalização do histograma: (a) imagem original; (b) imagem após aplicação da equalização; (c) histograma da imagem resultante.

Fonte: NUNES, 2006, p. 90.

Conforme mencionado utilizamos esta técnica é de muita importância para aumentarmos a qualidade da imagem para aplicação dos filtros. Devido a radiografia panorâmica possuir em alguns casos baixa qualidade, quanto maior a qualidade que conseguirmos neste passo melhor será o resultado final para a aplicação dos filtros.

### Subtração de Imagens

A subtração de imagens é uma relação linear sendo que é utilizada para fazer realces em diferenças espectrais entre uma imagem e outra. Caso as imagens a serem subtraídas não possuam histogramas com média e desvio padrão coincidindo, é aconselhado para que o processo de subtração seja assertivo que seja aplicada a equalização nos histogramas em ambas as imagens, pois se esta equalização não ocorrer o resultado será diretamente afetado e não corresponderá a diferença real (PROCESSAMENTO DE IMAGENS, 2017).

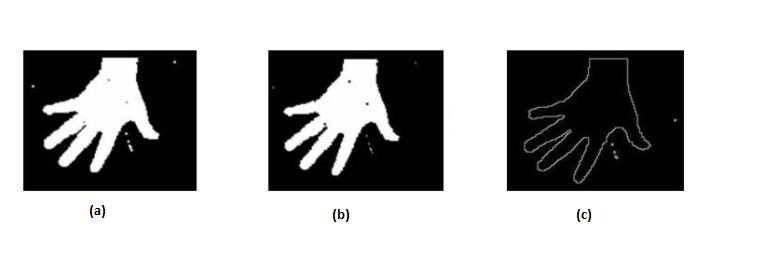
****

Figura 7 - Exemplo de equalização do histograma: (a) imagem após a aplicação da dilatação; (b) imagem após aplicação da erosão; (c) imagem resultante da subtração de ambas.

Fonte: Repositório da Pontifícia Universidade Católica do Rio de Janeiro.

Disponível em: https://webserver2.tecgraf.puc-rio.br/~mgattass/ra/trb09/Guilherme/VisaoComputacional%20-%20Trabalho%201.htm. Acessado em ago. 2017.

Esta técnica é muita importância e é utilizada para sobrepor uma imagem a outra e com isso, utilizando simetria podemos extrair as regiões com características condizentes a anomalias nas imagens.

### Máscara Binária

As máscaras são utilizadas em diversas técnicas dentro do processamento digital de imagens, sendo que em alguns momentos, somente algumas regiões de interesse da imagem desejam ser extraídas da imagem original e para tal, cria-se a chamada máscara binária. Para se extrair as regiões de interesse da imagem a máscara, que pode ser uma pequena matriz quadrada, atua pixel a pixel utilizando informações de seus vizinhos e defini os novos valores de cada pixel com a aplicação da máscara centrada individualmente. Os pixels que possuem todos os bits iguais a 1 na máscara são transferidos para a nova imagem, enquanto os que possuem valor zero são ignorados (SILVA, 2011).

Esta técnica é utilizada no protótipo quando queremos extrair algumas regiões de interesse da imagem, sendo que sem esta não poderíamos nem começar o algoritmo pois também utilizamos em conjunto com as segmentações.

### Filtros

Os filtros nada mais são que máscaras, porém sua composição se diferencia, pois, os pixels que compõe um filtro possui valores para uma finalidade específica do filtro. O objetivo principal dos filtros é auxiliar na aplicação de outras técnicas fazendo com que a nova imagem seja mais adequada especificamente ao contexto em que se deseja aplica-la, sendo que em alguns casos os filtros podem gerar imagens completamente diferentes dos originais, como por exemplo no caso dos filtros que trabalham no domínio de frequência de uma imagem (SILVA, 2011).

### Filtros Gaussianos

Os Filtros Gaussianos são utilizados para suavizar uma imagem borrando ou desfocando a mesma, reduzindo assim ruídos presentes que podem não ser úteis no processamento de imagem, porém após a sua utilização a imagem resultante terá muitos fragmentos de bordas falsas, que são causados devido aos ruídos e detalhes da textura da imagem. Este fator pode ser agravante em determinadas situações, sendo que a utilização de técnicas de limiarização para modificar estes valores é recomendado. Porém ainda sim podem haver estas bordas falsas e um dos motivos chaves para isso ocorrer é a escolha de um limiar muito baixo e também pela ocorrência de contornos reais que podem ter sido perdidos durante a aplicação do filtro de suavização (JESUS, 2015).

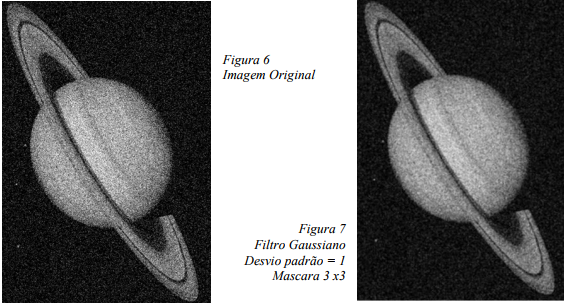


Figura 8 - Exemplo de aplicação de Filtro Gaussiano. Fonte: JESUS, 2015, p. 5

Utilizamos os filtros Gaussianos para retirar ruídos e regiões que podem não conter informações que sejam relevantes para o algoritmo de PDI. Com isto dispensamos pequenos ruídos nas imagens das radiografias panorâmicas.

### Classificador Haar

Amplamente utilizado para a detecção de faces em uma imagem, Viola e Jones (2001) propuseram o desenvolvimento de um detector de faces em tempo real, que pode ser utilizado não só para detecção de faces, mas de outros objetos em imagens (PEREIRA, 2017).

O método desenvolvido por Viola e Jones trouxe três grandes contribuições para a comunidade científica:

• A representação da imagem pelo conceito de “imagens integrais”;

• Um novo algoritmo de aprendizado baseado em Adaboost;

• Um novo método de localização da face utilizando os classificadores em cascata;

Este detector de objetos é baseado no valor das rectangle features, pois este sistema de detecção tem um desempenho computacional superior ao desempenho de sistemas que analisam valores únicos de pixels (PEREIRA, 2017).

O conceito de features provém de funções Haar de Base. Existem basicamente quatro tipos de características conforme imagem abaixo (PEREIRA, 2017).

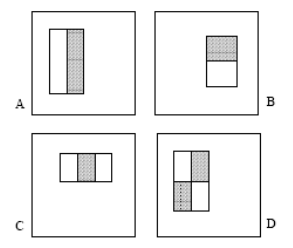


Figura 9 – Imagem com exemplos de features posicionadas sobre uma janela de amostragem. Fonte: PEREIRA, 2017.

Através desta imagem nota-se que as características podem ser classificadas em características de 2 retângulos (A e B), de 3 retângulos (C) e de 4 retângulos (D).

O processo de detecção de faces inicia-se pela transformação da imagem capturada em imagem integral. Partes da imagem integral, chamadas de janelas de amostragens, são passadas para o detector em cascada que utiliza classificadores treinados para detectar faces. Estes classificadores são formados por características com limiares definidos, obtidos na fase 21 de treinamento para estabelecer se uma região pertence ou não a uma face (PEREIRA, 2017).

Na fase de treinamento, os parâmetros dos classificadores são ajustados a partir de um algoritmo de treinamento baseado em Adaboost. O Adaboost é um procedimento eficiente porque procura o menor número de características com diferenças significativas entre si. Como entrada para o treinamento é necessário obter imagens positivas (imagens que correspondem a faces) e negativas (imagens que não correspondem a faces). O treinamento combina uma coleção de classificadores fracos para formar um classificador forte (PEREIRA, 2017).

O resultado do treinamento é um conjunto de classificadores compostos por características, sua localização na janela e valores de limiar para cada classificador. Exemplos de classificadores obtidos por Viola-Jones podem ser vistos na figura a seguir, com dois classificadores, uma característica de 2 retângulos e outra característica de 3 retângulos.

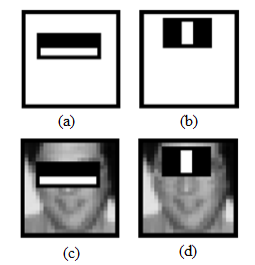


Figura 10 – Imagem com exemplos de classificadores obtidos por Viola-Jones: (a) e (b) Características obtidas pelo classificador; (c) e (d) Características aplicadas nas imagens. Fonte: PEREIRA, 2017.

Utilizamos este classificador para segmentarmos regiões de interesse do restante da imagem da radiografia e com isto aplicar de forma mais eficaz toda a fase de pré-processamento da imagem, pois retiramos regiões que o algoritmo poderia analisar de forma errônea.

### Curvas de Bézier

As curvas de Bézier foram criadas independentemente pelos engenheiros Paul de Casteljau da Citroen, em 1959, e Pierre Etienne Bézier da Renault, em 1962, para auxiliar no design e fabricação assistida por computador (BIEZUNER, 2014).

O primeiro desenvolveu um algoritmo eficiente, numericamente estável, para calculá-las usado até hoje, mas sua pesquisa foi considerada segredo industrial pela empresa em que trabalhava durante algum tempo, de modo que foi o segundo que as popularizou. Os polinômios de Bernstein já existiam desde 1912, quando Sergei Natanovich Bernstein os usou para dar uma demonstração concreta do teorema de aproximação de Stone-Weierstrass (BIEZUNER, 2014).

Atualmente, as curvas de Bézier são extensivamente usadas em computação gráfica na modelagem de curvas suaves, em animação, no design de interfaces e na produção de fontes. Curvas de Bézier quadráticas ou cúbicas são mais comumente usadas, já que polinômios de grau maior são computacionalmente mais custosos de calcular. Para formas mais complexas, muitas vezes são usadas curvas de Bézier por partes, isto é, curvas compostas de segmentos polinomiais de Bézier de grau baixo (BIEZUNER, 2014).

Utilizamos as Curvas de Bézier para podermos extrair informações relevantes quanto ao contorno das anomalias nas imagens. Assim que encontramos as anomalias extraímos esta informação para saber a curvatura de uma mancha encontrada na imagem. Esta informação será utilizada para a fase de pós processamento do protótipo.

### OpenCV

Por ser muito completa a biblioteca OpenCV é muito utilizada em diversas áreas. Ela foi criada pela Intel com a finalidade de deixar a área da visão computacional mais simples para seus usuários e desenvolvedores, sendo que para realizar tal feito conta com 500 funções e está dividida em 5 grupos: processamento de imagens, análise estrutural, análise de movimento e rastreamento de objetos, reconhecimento de padrões e calibração de câmera e reconstrução 3D, sendo que possui também documentação muito detalhada de todas as funções e também possui código fonte aberto. Para que a biblioteca consiga trabalhar em todas as suas funções também é disponibilizada a biblioteca IPL (Image Processing Library), que o OpenCV utiliza dos seus recursos parcialmente (MARENGONI, 2009).

Utilizamos esta biblioteca para toda a parte de processamento digital de imagens nos exames, pois ela disponibiliza muitos filtros e também possui ampla usabilidade pela comunidade e ampla documentação para estudo.

### Textura

Textura é um atributo visual presente em imagens, com muita diversidade de texturas naturais e artificiais, torna um atributo sem definição formal. Com padrão visual muitas vezes relacionado a distribuição de pixels em uma região da imagem, esse atributo geralmente contém informações bastante ricas sobre o conteúdo da imagem e é muito útil em visão computacional segundo Professora Yandre Costa da UNIVERSIDADE ESTADUAL DE MARINGÁ.

Alguma das características de texturas são descritas cotidianamente como finas, grossas, granuladas, lisas e etc.

– Granularidade: tamanho das células presentes na textura, eventualmente referida como “espessura”;

– Contraste: dado por variações de tons de cinza presentes na imagem;

– Direcionalidade: direção predominante entre os elementos constituintes;

– Delineamento (line-likeness): intensidade da presença de linhas;

– Regularidade: regularidade com que primitivas se repetem;

– Rugosidade: medida da sensação de “aspereza” transmitida pela textura;

### Matriz de Co-Ocorrência

Matriz de Co-ocorrência ou Ocorrência simultânea é um cruzamento das combinações diferentes de valores dos pixels (níveis de cinza) ocorrem em uma imagem. Uma aplicação da matriz é identificar texturas de uma imagem através dos conjuntos comparados.

A Matriz de Co-ocorrência de textura considera a relação entre dois pixels por vez, um chamado de pixel referência e o outro de pixel vizinho. O pixel vizinho escolhido pode ser vizinho em qualquer direção: por exemplo a leste (direita), a oeste (esquerda), a norte (acima), a sul (abaixo), ou na diagonal, ou seja a nordeste, noroeste, sudoeste e sudeste de cada pixel referência. Também a vizinhança não precisa ser exatamente de 1 pixels, pode ser de 2, 3, ou qualquer valor. Cada pixel dentro da imagem torna-se o pixel referência, iniciando no canto superior esquerdo e procedendo até o inferior direito (VASCONCELOS, 2017).

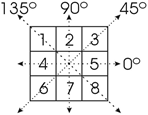


Figura 11 – Ângulos para o cálculo na matriz de co-ocorrência (VASCONCELOS).

## ECOSSISTEMA BIG DATA

2.3.1 BIG DATA

Big Data é um termo relativamente novo, ao armazenar grandes quantidades de informações para uma possível análise de dados é bem antigo. Ganhou uma definição no início dos anos 2000, com Doug Laney, definindo big data como os três Vs:

• Volume: A quantidade de dados a serem processados.

• Velocidade: O quão rápido estes dados serão processados.

• Variedade: Os vários tipos de dados que podem estar presentes em um conjunto de entrada de informações.

Utilizamos conceitos de Velocidade e Variedade de Big Data na parte de pós processamento do protótipo afim de melhorarmos o resultado final para o profissional odontológico.

### MAPREDUCE

O MapReduce foi descrito em um artigo publicado pela Google (DEAN; GHEMAWAT, 2004) como solução de analise em grandes volumes de dados usando principalmente clusters de computadores commodities. A Google desenvolveu orientado a objetos utilizando principalmente C++ e são compostos por duas tarefas principais (chamadas de Map e Reduce).

O MAP realiza a função de mapa, levando um conjunto de dados e converte-lo em um outro conjunto de distinto, os elementos individuais são decompostos em pares. O trabalho de reduzir leva a saída de um mapa como entrada e associa os pares de dados em pares menores.

O REDUCE implica na função de reduzir e é sempre executado depois do trabalho no map.

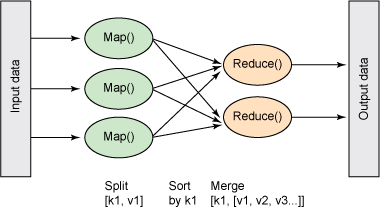


Figura 12 – Imagem com exemplo da função MAP e REDUCE. Fonte: PAIVA, 2011.

### HADOOP 1.0

Apache Hadoop é um framework em Java de computação distribuída que trabalha com clusters e enorme processamento de dados. Inspirada no MapReduce e no GoogleFS (GFS). Hadoop as funções Map e Reduce, onde cada pode ser executado ou reexecutado em qualquer nó no cluster que possui um sistema de arquivos distribuídos (HDFS) armazenando os dados nos. Tanto o MapReduce e o sistema de arquivos distribuídos Hadoop são projetados para que falhas de nó são manipuladas automaticamente pela estrutura (HADOOP, 2016).

### HADOOP 2.0

Com a utilização do framework crescendo, novas atualizações foram inevitáveis, no lançamento da versão 0.20.203 do Hadoop, todo o framework foi reestruturado, com nova arquitetura quase um framework novo foi criado.

O novo framework foi chamado de MapReduce 2.0 (MRv2) ou YARN. A sigla YARN significa “Yet Another Resources Negotiator”. A principal modificação do framework, ter o foco principal em tornar o componente de MapReduce menos dependente do Hadoop (WHITE, 2012).

O principal motivo para a criação do YARN era limiar a escalabilidade que acontecia no MapReduce 1.0. Para os clusters de grande tamanho o Hadoop não era uma solução de grande confiabilidade (MURTHY, 2011)

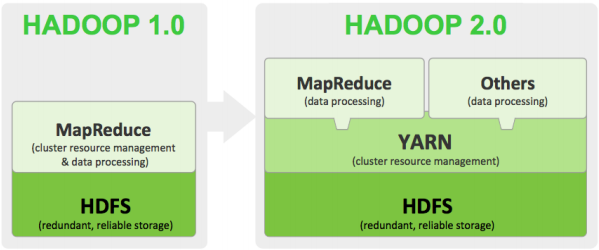


Figura 12 – Imagem com a arquitetura do Hadoop 1.0 e Hadoop 2.0. Fonte: (HORTONWORKS,2013).

### CLOUDERA MANAGER

Cloudera é uma companhia de com sede em Palo Alto, nos Estados Unidos, e conta atualmente com aproximadamente 600 funcionários. Sua principal atividade é prover softwares baseados em Apache Hadoop, bem como suporte, serviços e treinamento (CLOUDERA, 2014).

A companhia possui uma distribuição open-source de Apache Hadoop chamada de Cloudera Distribution Including Apache Hadoop (CDH). É uma distribuição para softwares de Big Data.

A última versão disponível durante a redação deste documento é chamada de CDH5, que inclui os seguintes aplicativos de Big Data:

• Yarn (MapReduce 2.0)

• Apache HBase

• Apache HDFS

• Apache Hive

• Cloudera Impala

• Apache Oozie

• Apache Spark

• Apache Zookeeper

• Apache HUE

A descrição de cada aplicação foge do escopo deste documento, uma vez que somente será utilizado o pacote de YARN (MapReduce 2.0) e o Apache HDFS.

Utilizamos o Cloudera para suportar a parte do Hadoop e do HDFS na parte de pós processamento do protótipo.

### APACHE CASSANDRA

Repositório de dados desenvolvido pelo Facebook como open source em 2008 leve e desenvolvido na plataforma Java, atualmente é baseado na tecnologia emergente NoSQL (OLIVEIRA FARIA, 2017).

Utilizamos este recurso para salvarmos informações das características das anomalias encontradas nas imagens para futuro pós processamento.

### MINERAÇÃO DE DADOS

Com o grande volume de dados criados atualmente técnicas são utilizadas na identificação de informações relevantes e uma delas é a mineração de dados.

Mineração de dados é um processo de exploração de grandes quantidades de dados com objetivo de encontrar anomalias e padrões, dando suporte a tomada de decisão, reduzindo risco, custo entre outros benefícios.

### WEKA

O WEKA é um produto da Universidade de Waikato (Nova Zelândia) e foi implementado pela primeira vez em sua forma moderna em 1997. Ele usa a GNU General Public License (GPL) (ABERNETHY, 2017).

Escrito em Java e interage diretamente com arquivos de dados para produzir resultados visuais, com sua API pode ser incorporada a qualquer outra biblioteca, para realizar tarefas de mineração de dados automatizadas.

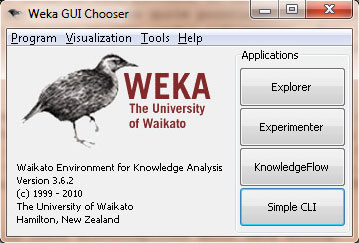


Figura 13 – Imagem inicial do WEKA. Fonte: Autoria própria.

### CLASSIFICAÇÃO

Um problema recorrente em identificação de padrões em radiografias é a grande semelhança pelas limitações impostas pelo próprio exame, para resolver este problema iremos utilizar o algoritmo J48 que gera uma árvore de decisão.

Utilizamos o WEKA para geração da árvore de decisão que contém as características já classificadas das anomalias encontradas. Utilizando um sensor de inferência também será utilizá-lo para extrair dados para a análise de uma instância de uma nova imagem (DAMASCENO).

# DESENVOLVIMENTO



## INTRODUÇÃO

Conforme vimos neste documento, há dificuldade na visualização de anomalias nas radiografias panorâmicas, com isto os autores propõem a utilização do processamento digital de imagens para pré processar as imagens radiográficas e uma arquitetura onde a máquina de conhecimento vá até a base de imagens já processadas trazendo autonomia e tempo no processamento. Nessa arquitetura, existirá uma base de dados Cassandra que irá armazenar as características extraídas das anomalias das radiografias panorâmicas odontológicas e será acessada pela aplicação para a geração de uma árvore de decisão pelo WEKA. Com esta já gerada entramos no pós-processamento, comparando a nova instância com os padrões já classificados pelo WEKA.

O fluxo do sistema se inicia quando recebe uma radiografia panorâmica nova, o protótipo realiza o pré-processamento e o analisa através de uma árvore de decisão criada a partir de imagens já inseridas e classificadas no banco de dados. Com esta árvore de decisão podemos classificar uma nova instância aplicando os padrões conhecidos. Após este passo a imagem é inserida na base e passa a compor a base de conhecimento da aplicação e então o resultado desta análise que foi realizada é apresentada ao profissional. Para ilustramos melhor abaixo segue uma figura com um diagrama de como ficaria a integração do sistema.

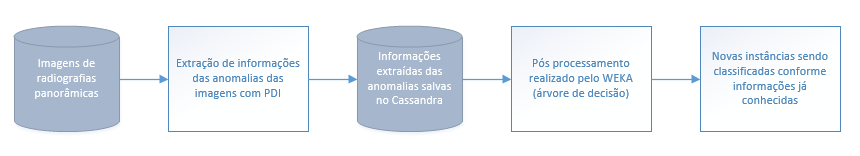


Figura 14 – Imagem com o fluxo do protótipo. Fonte: Autoria própria.

## MÉTODO

A proposta dos autores consiste na elaboração de uma solução de software para processamento das informações encontradas nas radiografias panorâmicas. Nesta seção é apresentada a visão geral do desenvolvimento até o presente momento.

### As imagens odontológicas

Atualmente possuímos um total de 100 imagens diferentes para desenvolvimento do projeto dentre essas sendo imagens saudáveis e com anomalias, sendo que estamos em busca de mais imagens tanto em mecanismos de buscas e em bases públicas. Quanto a busca de imagens em mecanismos de buscas temos alguns pré-requisitos para utilização de uma imagem encontrada, são eles: Imagem com resolução maior ou igual a 800 pixels de altura por 800 pixels de largura, sem nenhuma marca d’água ou algum outro tipo de efeito nas regiões em que iremos aplicar os filtros. Para podermos utilizar estas imagens em nossa pesquisa os autores catalogaram cada imagem com o link em que se encontram disponíveis.

### Processamento de Imagem

Nesta fase de desenvolvimento do protótipo nós realizamos por primeiro a segmentação de regiões que iremos identificar anomalias antes de efetuar o restante dos filtros. Com este passo eliminamos o risco de que caso alguma imagem não possua características que indiquem que é uma imagem de uma radiografia panorâmica. Esta segmentação consiste em extrair somente determinadas regiões do exame que já possuímos o conhecimento necessário no algoritmo para analisarmos. A figura abaixo mostra as regiões que são encontradas na radiografia panorâmica.

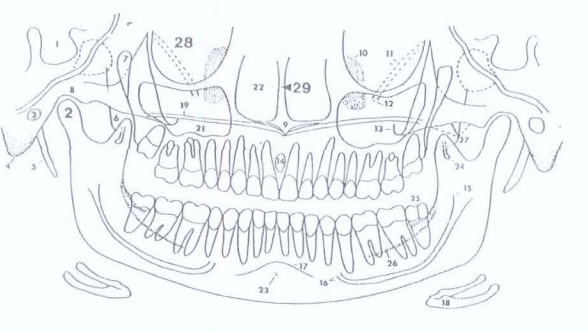


Figura 15 –Imagens contínua das referências anatômicas.

Fonte: SAIO, 2005.

Em nosso protótipo iremos identificar anomalias nas regiões: forame incisivo nasopalatin (14), forame mandibular (15), forame mental ou canal mandibular (16), crista mental (17), tubérculos geni (23), processo hamular (24), linha oblíqua externa (25) e a linha oblíqua interna ou milo-hioídea (26).

Para efetuar esta segmentação utilizamos o método “Classificador Haar” e com esse conseguimos extrair as regiões desejadas. Para criar este classificador foi necessária a criação de um arquivo XML que contém todas as características que esperamos da segmentação. Este arquivo foi criado utilizando ferramentas do próprio OpenCV.

Para efetuarmos testes no algoritmo de segmentação utilizamos um exame com uma anomalia conhecida como Cisto Ósseo Traumático que é uma lesão assintomática e que se desenvolve na mandíbula e se caracteriza como uma reabsorção hemorrágico-intraósseo que se origina por meio de algum tipo de trauma nos pacientes (PAES, 2010).

Esta anomalia se destaca radiograficamente por ser uma lesão radiolúcida bem demarcadas de tamanho variado, com bordas escleróticas, as quais podem projetar-se nos septos intrarradiculares e, por consequência, formar uma linha com bordas onduladas, entretanto nesta imagem o COT se destaca na região da mandíbula (ROSEN, 1997). Esta anomalia pode ser visualizada do lado direito da mandíbula na Figura 9.

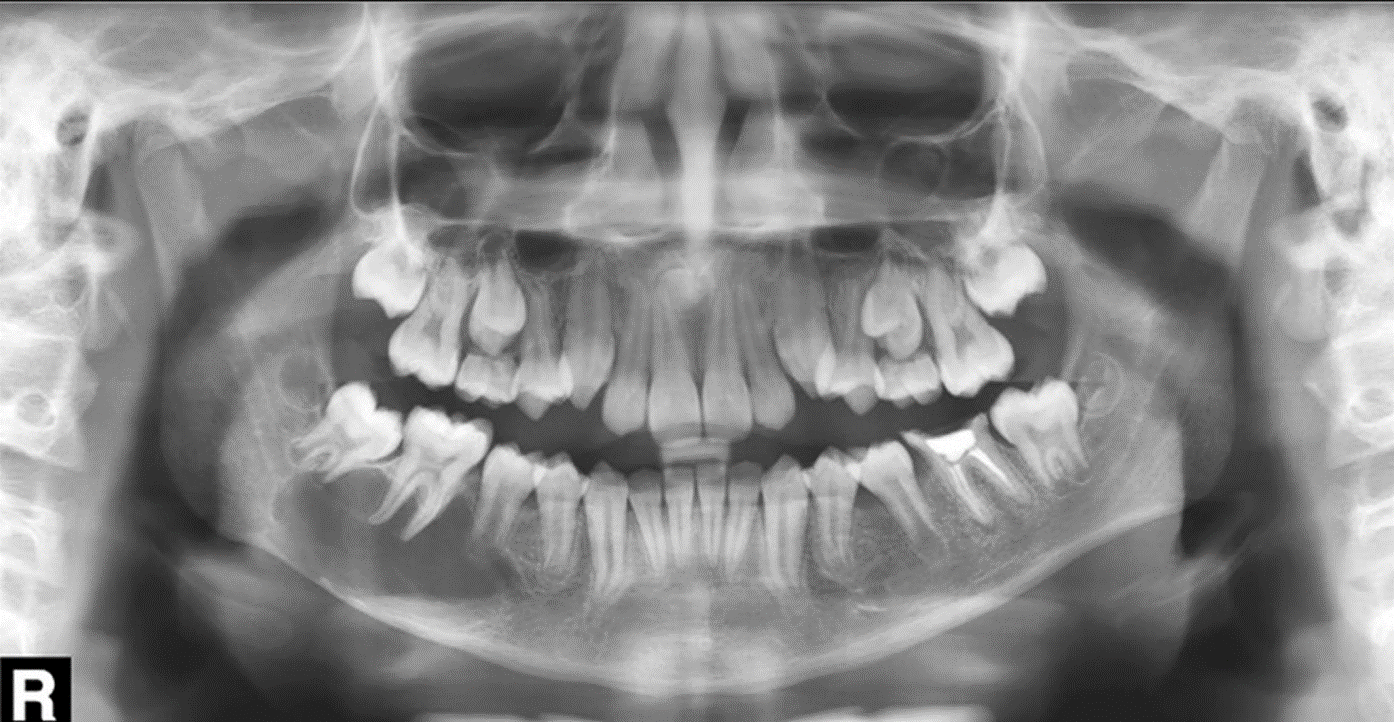


Figura 16 – Imagem compatível com Cisto Ósseo Traumático. Fonte: Repositório Papaiz – Diagnósticos por Imagem.

Disponível em: http://www.papaizassociados.com.br/cisto-osseo-traumatico-como-achado-radiografico. Acessado em fev. 2017.

Logo após a escolha de uma imagem para os testes começamos o processo de segmentação obtendo o resultado da figura abaixo, onde foi feita marcação do limite da imagem para o leitor visualizar a área de corte que iremos aplicar os filtros.

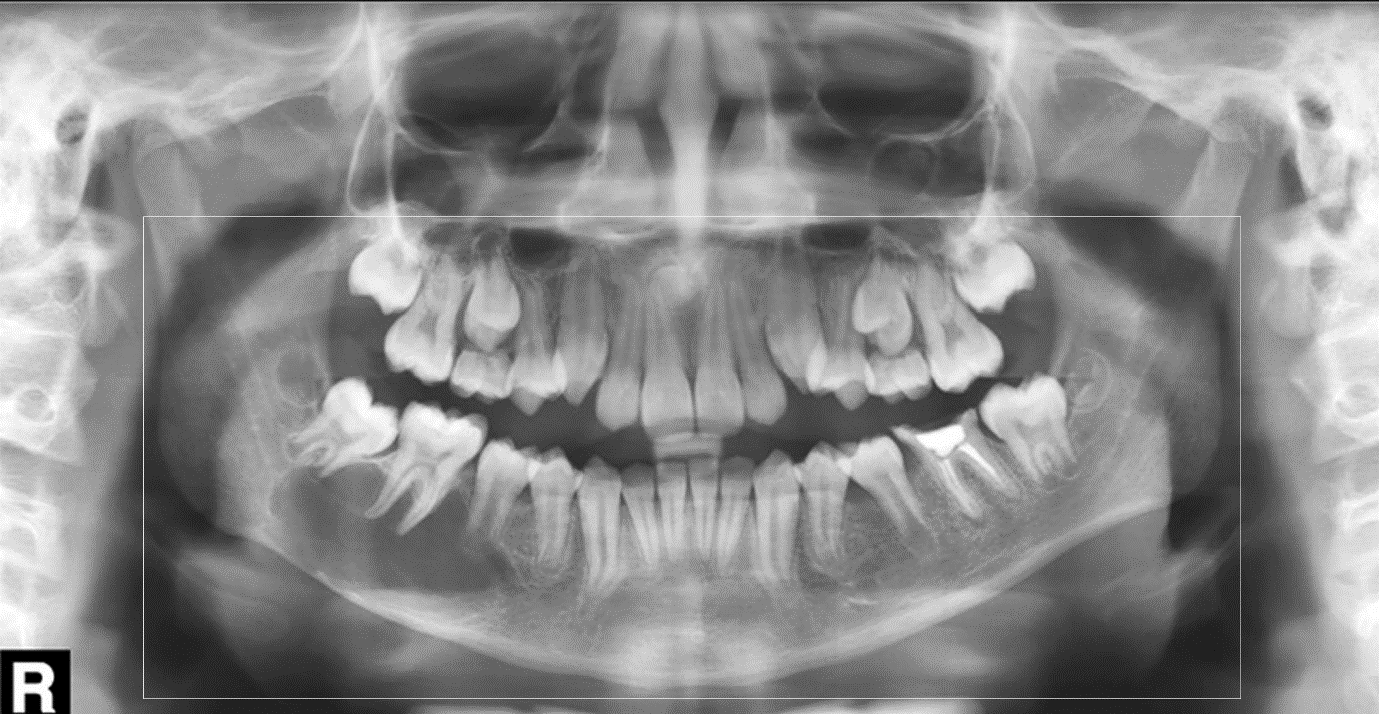


Figura 17 – Imagem com a segmentação das regiões a serem analisadas. Fonte: Autoria própria.

Com a parte da imagem que será processado pelo algoritmo já segmentada do resto da imagem, carregamos ela novamente com escala de cores cinzas. Após isto passamos a imagem por um filtro Gaussiano com matriz 7x7 afim de retirar algum possível ruído da imagem, sendo que este tamanho de matriz foi definido visualmente pelos autores, e logo depois fazemos a equalização do histograma para que a imagem fique com o máximo de destaque possível para áreas com problemas.



Figura 18 – Imagem com filtro Gaussiano e equalização de histograma aplicado. Fonte: Autoria própria.

Com a imagem sem ruídos significativos e com em escala cinza podemos fazer a primeira dilatação afim de aumentar as regiões que podem ser significativas. Para este passo utilizamos um elemento estruturante de 19x19 com o formato de cruz. Chegamos no tamanho do elemento estruturante depois de uma análise visual em sobre todos os outros resultados sendo que este tamanho foi o mais satisfatório para o restante das técnicas utilizadas.



Figura 19 – Imagem com dilatação aplicada. Fonte: Autoria própria.

Analisando melhor a Figura 11 percebemos que a região da anomalia procurada e as regiões ósseas também ficaram maiores e com uma tonalidade mais clara, isto é muito importante pois podemos aumentar o tamanho de uma região pequena que pode passar despercebida pelos profissionais, no entanto a fase de processamento também precisa eliminar as regiões que podem não possuir importância significativa e é por isso que fazemos uma segmentação da imagem utilizando limiarização afim de deixar as regiões que não são anomalias em destaque e extrair as regiões que estão no osso da mandíbula na parte inferior.



Figura 20 – Imagem segmentada. Fonte: Autoria própria.

Após este passo iremos dividir a imagem ao meio e iremos girar o lado direito em 90º graus para fazermos a subtração de uma imagem pela outra, tanto do lado direito pelo esquerdo quanto do contrário, e assim podermos destacar as regiões problemáticas. Como resultado deste passo obtivemos a Figura 12 que destaca a anomalia e também outras regiões problemáticas, porém também obtivemos o destaque de outras regiões indesejadas como a parte perto da borda inferior da mandíbula do lado direito e a sobreposição dos dentes.



Figura 21 – Imagem com a subtração aplicada. Fonte: Autoria própria.

Afim de retirar algumas destas regiões que não possuem importância percorremos a matriz e verificamos quais pixels que possuem um valor próximo de 255, ou seja, se aproximam mais do branco e assinalamos ele com o valor 0 que deixa eles pretos para que não influenciem nos próximos passos.

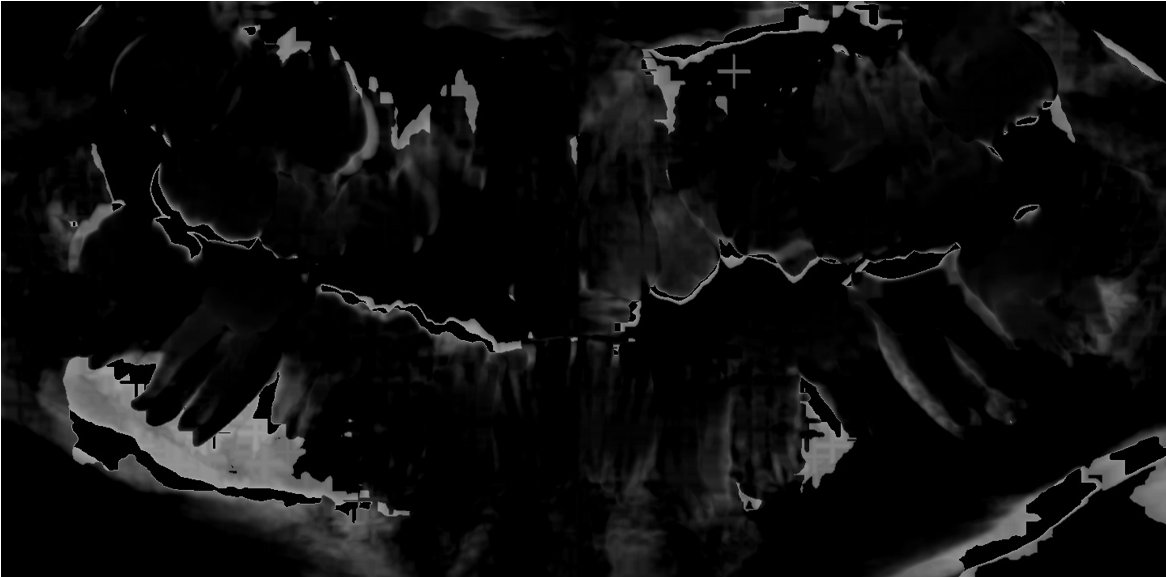


Figura 22 – Imagem com pixels perto do 255 já cortados. Fonte: Autoria própria.

Após este passo aplicamos filtros Gaussianos para diminuir e até extinguir regiões indesejadas, porém conforme comentado nesta pesquisa estes filtros podem remover informações relevantes e futuramente utilizaremos outro método para retirar as regiões sem significância.



Figura 23 – Imagem obtida após o filtro Gaussiano ser aplicado diversas vezes. Fonte: Autoria própria.

Após este passo chegamos a um resultado muito importante pois temos duas áreas em destaque e que são regiões problemáticas na mandíbula conforme parecer de um profissional que tivemos para validação deste resultado. Para conseguir destacar essas imagens na imagem original utilizamos o método de detecção de bordas Canny e obtivemos como resultado a Figura 16.

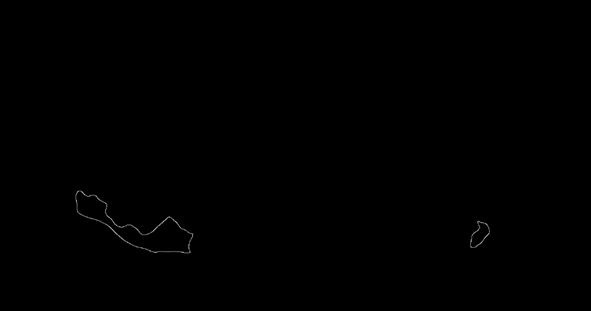


Figura 24 – Imagem obtida após o método Canny ser aplicado. Fonte: Autoria própria.

Após este passo utilizamos um método para procurar os contornos da imagem e desenhar contornos em uma máscara em forma de quadrados para que fique fácil a visualização da região com problema pelos profissionais. Após este processo aplicamos a máscara na imagem e obtivemos como resultado a Figura 17.

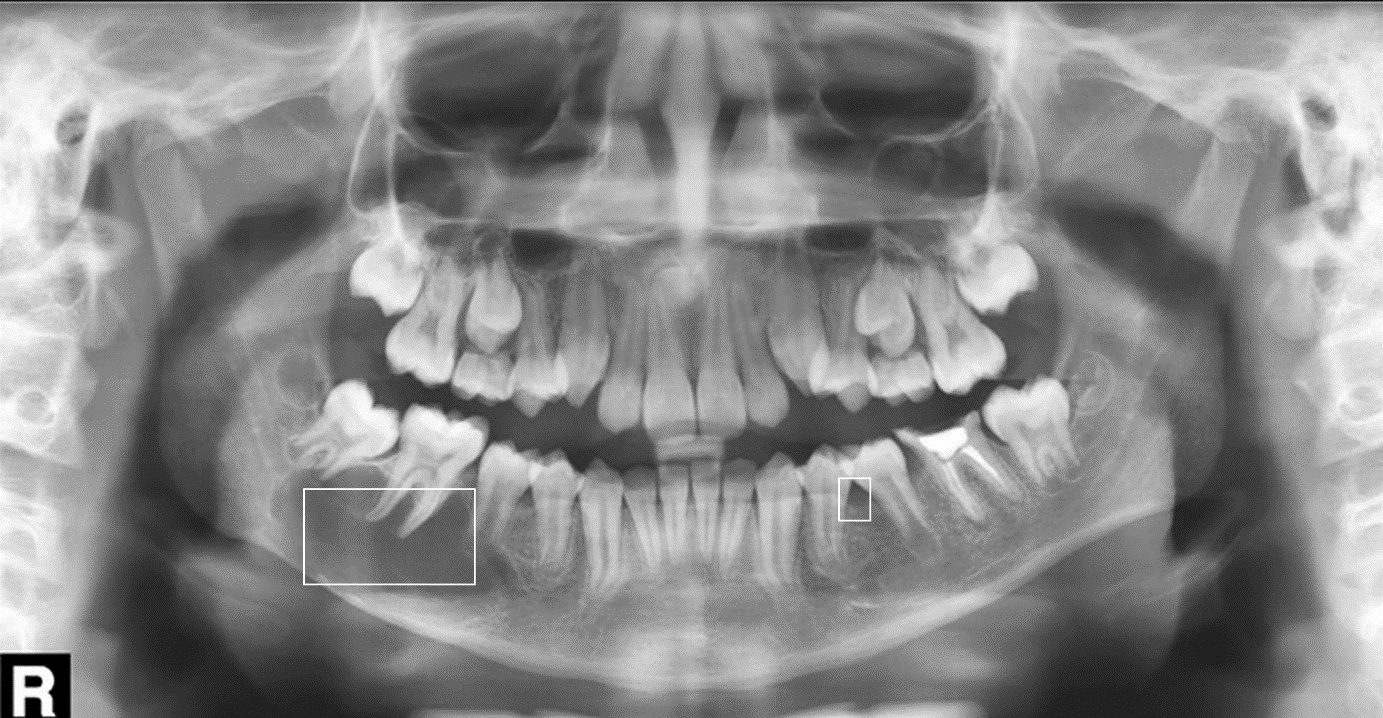


Figura 25 – Imagem com anomalias em destaque. Fonte: Autoria própria.

Neste resultado observamos duas áreas destacadas, uma do lado esquerdo da imagem com o COT e outra no lado direito com o mesmo padrão de características do COT, indicando que esta região pode ser uma anomalia.

### A Máquina de Aprendizado

Para a aplicação deste processo iremos utilizamos o Hadoop e o Cassandra. Para auxiliar na leitura a Figura 14 deste documento possui um diagrama de como funciona a integração da parte de processamento de imagens com PDI.

Para a geração desta máquina de conhecimento utilizamos as 100 imagens que possuímos, o ideal seria um número maior de imagens, porém como não tivemos a liberação da Clínica Odontológica da Universidade Positivo mesmo com o parecer Aprovado do comitê de ética para podermos utilizá-las, a nossa contingência foi utilizarmos imagens de mecanismos de buscas e de bases de acesso públicas, com características pré-estabelecidas neste documento anteriormente, para podermos dar continuidade na pesquisa.

Com o algoritmo de PDI já em funcionamento nós procuramos anomalias nas imagens e quando encontramos uma imagem com problema extraímos algumas informações e características destas anomalias e salvamos no banco de dados Cassandra, sendo que na nossa fase de desenvolvimento atual já estamos neste passo.

Caso o algoritmo identifique alguma anomalia as características são extraídas e salvas. Estas características até o presente desenvolvimento são:

* Quantidade de Linhas: Quantidade de linhas encontradas na matriz com anomalia.
* Quantidade de Colunas: Quantidade de colunas encontradas na matriz com anomalia.
* Resolução: Resolução da imagem com anomalia.
* Informações da imagem: Link da imagem ou caminho da imagem com anomalia.
* Margem da massa: Sequência das curvas de Bézier encontradas no contorno da anomalia.
* Matriz da anomalia: Sequência com os valores da matriz que foi extraída pelo algoritmo que contém a anomalia.
* Contraste: Informação do contraste da anomalia.

Para o próximo passo, que seria a geração de uma árvore de decisão utilizando o WEKA com as características que forma salvas no banco de dados Cassandra, estamos estudando a inserção de mais características das imagens, pois quanto mais característica extrairmos nesta fase melhor será para a geração da árvore de decisão.

Com a árvore de decisão já gerada basta somente a fase dos pós processamento para podermos entregar o resultado ao profissional.

### O Pós-processamento

A fase de pós processamento é realizada quando tivermos a árvore de decisão já gerada e assim iremos diferenciar o que pode ser uma anomalia de outras regiões que podem não possuir nenhum outro tipo de problema. Com este passo iremos fornecer um parecer mais assertivo para o profissional com as regiões problemáticas.

Devido a não possuirmos um grande número de imagens não iremos utilizar os 4 servidores que estávamos propondo na versão anterior deste documento, iremos utilizar 1 único para efetuar o pós-processamento.

Neste servidor será aplicada uma arquitetura distribuída, ou seja, a aplicação irá em cada nó e irá efetuar o processamento dentro deste e irá comparar os padrões já extraídos com os já inseridos na árvore de decisão e irá retornar com o resultado para a aplicação.

Iremos utilizar bibliotecas de integração com o WEKA para podermos extrair informações da árvore de decisão gerada e aplicarmos estas informações a uma nova instância inserida pelo profissional e com isso retornar a imagem final com as regiões problemáticas assinaladas.

## O PROTÓTIPO

Nesta seção iremos mostrar alguns mockups do protótipo em fase final de desenvolvimento. Na imagem abaixo seria o protótipo já com toda a parte técnica finalizada. Nele podemos ver um botão “Selecionar Imagem”, que é utilizado para o profissional de odontologia selecionar uma imagem de uma radiografia panorâmica para podermos analisar.

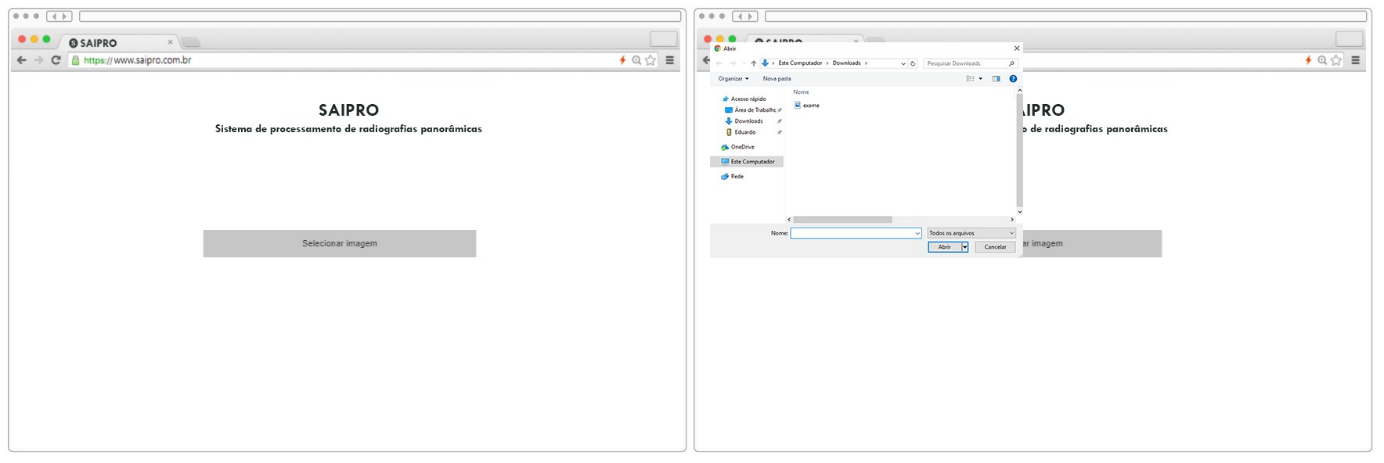


Figura 26 – Imagem como fluxo inicial do protótipo. Fonte: Autoria própria.

Logo após selecionar uma imagem o protótipo fornece uma imagem final com o resultado da análise feita pelo pós-processamento. Na figura abaixo e possível ver a saída que o protótipo entregaria para o profissional, com as regiões das anomalias já destacadas.



Figura 27 – Imagem coma saída do protótipo. Fonte: Autoria própria.

# GESTÃO



## ESCOPO

Durante a etapa de planejamento de trabalho de conclusão de curso, os requisitos mínimos foram definidos pelo discente. Entre estes requisitos, pode-se citar:

|  |  |
| --- | --- |
| **Código** | RF001 |
| **Descrição** | Inserir imagens de radiografias panorâmica no sistema (odontologista) |
| **Visibilidade** | (x) Evidente ( ) Oculto |
| **Prioridade** | (x) Essencial ( ) Importante ( ) Desejável |
| **Dependências** | N/A |

|  |  |
| --- | --- |
| **Código** | RF002 |
| **Descrição** | Efetuar o pré e pós processamento da imagem recebida |
| **Visibilidade** | ( ) Evidente (x) Oculto |
| **Prioridade** | (x) Essencial ( ) Importante ( ) Desejável |
| **Dependências** | RF001 |

|  |  |
| --- | --- |
| **Código** | RF003 |
| **Descrição** | Extrair metadados encontrados nas imagens |
| **Visibilidade** | ( ) Evidente (x) Oculto |
| **Prioridade** | (x) Essencial ( ) Importante ( ) Desejável |
| **Dependências** | RF002 |

|  |  |
| --- | --- |
| **Código** | RF004 |
| **Descrição** | Comparar os padrões obtidos com os padrões existentes |
| **Visibilidade** | ( ) Evidente (x) Oculto |
| **Prioridade** | (x) Essencial ( ) Importante ( ) Desejável |
| **Dependências** | RF003 |

|  |  |
| --- | --- |
| **Código** | RF005 |
| **Descrição** | Mostrar ao odontologista os resultados encontrados |
| **Visibilidade** | (x) Evidente ( ) Oculto |
| **Prioridade** | (x) Essencial ( ) Importante ( ) Desejável |
| **Dependências** | RF004 |

## CRONOGRAMA

Uma das principais características de um projeto é seu tempo de vida. Em prol do desenvolvimento de um projeto funcional, seu ciclo de desenvolvimento deve ser estimulado e cada uma das suas atividades que o compõe devem ser detalhadas de uma maneira realista. Abaixo segue o cronograma do projeto para visualização.

Tabela 1 - Cronograma

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2016 | 2017 | | | | | | | | | | | |
| **Atividades** | **M-D** | **J** | **F** | **M** | **A** | **M** | **J** | **J** | **A** | **S** | **O** | **N** | **D** |
| ROADMAP | √ |  |  |  |  |  |  |  |  |  |  |  |  |
| Definição do tema | √ | √ | √ | √ |  |  |  |  |  |  |  |  |  |
| Aprovação da ideia de TCC | √ | √ | √ | √ |  |  |  |  |  |  |  |  |  |
| Revisão de literatura |  | √ | √ | √ | √ | √ | √ | √ | √ |  |  |  |  |
| Entrega da proposta de TCC |  |  |  |  | √ |  |  |  |  |  |  |  |  |
| Redação da documentação de TCC |  |  |  |  | √ | √ | √ | √ | √ |  |  |  |  |
| Desenvolvimento do sistema |  | √ | √ | √ | √ | √ | √ | √ | √ |  |  |  |  |
| Entrega da 1ª versão da documentação |  |  |  |  |  |  | √ |  |  |  |  |  |  |
| 1ª Entrega parcial:  - Preparação dos clusters para instalação do framework  - Instalação do framework escolhido pelos autores nos clusters  - Desenvolvimento do protótipo de processamento de imagens  - Realização de testes para validar o funcionamento do protótipo de processamento de imagens  - Realização de testes para validação do funcionamento dos clusters |  |  |  |  |  |  | √ |  |  |  |  |  |  |
| 2ª Entrega parcial:  - Desenvolvimento do software de processamento de imagens finalizado com segmentação de mandíbula  - Integração do software com os clusters  - Utilização do software em conjunto com um grande banco de dados |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Entrega da 2ª versão da documentação |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Apresentação do projeto para banca de qualificação |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Entrega da versão final da documentação |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Banca final |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Entrega da documentação de TCC corrigida |  |  |  |  |  |  |  |  |  |  |  |  |  |

## CUSTOS

Devido principalmente à utilização de recursos de *hardware* fornecidos pela Universidade Positivo a maior parte dos custos foram para adquirir material de pesquisa e ensino, porém optamos por inserir os custos sem retirar os recursos fornecidos para que esta pesquisa possa ser replicada de maneira eficiente. Os *softwares* utilizados possuem licença *open-source*, como o caso da biblioteca OpenCV e o código fonte do Apache Hadoop. A distribuição de Hadoop utilizado, a Cloudera CDH5, também é gratuita.

Tabela 2 – Custos Estimados.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Item** | **Descrição** | **Quant.** | **Unid.** | **Preço Unit.** | **Preço Total** |
| **Servidor** | Computador  Desktop Positivo  Intel Core i7 16GB | 1 | un | R$1.999,00 | R$ 7996,00 |
| **Livro** | Processamento Digital De Imagens - Rafael C. Gonzalez e Richard E. Woods | 1 | un | R$ 120,00 | R$ 120,00 |
| **TOTAL** | | | | | **R$ 8.116,00** |

## RISCOS

A seguir é apresentada uma tabela com alguns riscos do projeto. Existem três possíveis graus para riscos: grau alto (riscos com grau entre 0,5 e 1); grau médio (riscos com grau entre 0,25 e 0,5); grau baixo (riscos com grau entre 0,0 e 0,25).

O grau é calculado pela multiplicação da probabilidade de um risco acontecer pelo impacto daquele risco. Tanto a probabilidade quanto o impacto são normalizados, ou seja, limitados entre zero e um. A Tabela 3 descreve estes riscos:

Tabela 3 – Análise de Riscos

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Grau** | **Riscos** | **Efeito** | **Probabilidade** | **Impacto** | **Ação** |
| **0,1** | Incompatibilidade de framework | Retorno à fase de pesquisa e Planejamento | 0,1 | 1 | Definir outro framework |
| **0,2** | Problemas de software | Processamento de imagem incorreto | 0,4 | 0,7 | Encontrar outra biblioteca de processamento de Imagens |
| **0,3** | Problemas de detecção de anomalias nas imagens | Detecção de área incorretas | 0,6 | 0,5 | Redefinir o método de detecção |
| **0,5** | Não encontrar nenhuma fonte de exames | Banco de dados sem um número mínimo de imagens | 0,6 | 0,8 | Procurar uma fonte de dados pública |
| **0,1** | Banco de dados muito escasso | Pouca diversidade para validação dos testes | 0,3 | 0,4 | Definir outra fonte de dados |

# AVALIAÇÃO

Neste capítulo iremos apresentar como estaremos avaliando o nosso protótipo perante aos profissionais de odontologia. Será mostrado a forma de como será feita a avaliação do protótipo pelos profissionais.

Antes de começar a avaliação do protótipo iremos fornecer o formulário TCLE (Termos de Consentimento Livre e Esclarecido) para preenchimento, que se encontra no “APÊNDICE” deste documento, afim de informarmos os profissionais que os mesmos estarão participando como voluntários desta pesquisa e que também a sua participação não irá acarretar custos e riscos para o mesmo. Neste termo os autores deste documento se disponibilizam a prestar qualquer tipo de auxílio para a participação nesta pesquisa.

Após a finalização deste passo, iremos apresentar o sistema para os profissionais, ensinando-os como o sistema funciona e qual será o resultado final da ferramenta. Iremos solicitar que os mesmos realizem testes com a plataforma no seu cotidiano por um curto prazo, inserindo imagens de exames de pacientes com ou sem anomalias para que os profissionais observem o resultado final e se o protótipo irá ser útil em sua vida profissional.

Após a utilização pelos profissionais iremos aplicar um formulário que estamos desenvolvendo com perguntas que irão nos informar se os resultados foram assertivos, se o protótipo possui uma boa usabilidade, se o protótipo teve resultado úteis para este profissional e se este profissional iria utilizar este protótipo em seu cotidiano profissional.

A partir destas respostas iremos publicar os resultados obtidos neste documento, afim de avaliarmos o a inserção desta nova ferramenta no cotidiano dos profissionais de odontologia.

# RESULTADOS OBTIDOS

O protótipo desenvolvido para destacar anomalias nos exames teve um resultado satisfatório pois conseguiu assinalar as áreas corretas na imagem, sendo que após segmentarmos uma determinada região das imagens das radiografias automatizamos desde o começo o algoritmo, pois permitimos então que ao invés de utilizarmos 60% da imagem, conforme a versão anterior, não iremos perder informações que podem ser úteis e que podem modificar o resultado final.

Quanto ao processo de pré-processamento da imagem utilizando o filtro Gaussiano, estamos ainda pesquisando a aplicação de diferentes métodos de processamento de imagens afim de melhorarmos o resultado final dos filtros. Este é um processo demorado o trabalhoso pois os testes efetuados podem combinar diversos filtros de diversas maneiras e proporcionar vários resultados diferentes, cabe a nós os autores definir quais filtros e em qual combinação correta utilizar no algoritmo, lembrando que a ordem em que foi feito o algoritmo não é citada em nenhum material que foi estudado na pesquisa bibliográfica sobre os temas.

Com relação a integração com a máquina de conhecimento estamos desenvolvendo a arquitetura da aplicação utilizando somente um computador, ou seja, não mais precisaremos utilizar 4 servidores pois temos um número de imagens pequeno e isto não irá demandar um recurso muito grande de hardware. Conforme havíamos mencionado na versão anterior estávamos em negociação com a própria universidade, porém devido a questões burocráticas da própria, mesmo com o parecer de aprovação do comitê de ética não obtivemos sucesso na obtenção das imagens, pois os pacientes não preenchem uma liberação para utilização destes exames em estudos e pesquisas. Com isto entramos em contato com o Departamento de Odontologia da UFPR para acesso as quase 4000 imagens, porém nos foi negado o acesso ao banco de imagens pois não somos alunos da instituição e a profissional responsável nos forneceu 30 imagens para podermos utilizar na nossa pesquisa. Não obtivemos sucesso com clínicas particulares pois é necessário este formulário de autorização dos pacientes e nenhuma que entramos em contato disponibilizava este formulário para os pacientes. Com isto nossa última opção era a busca de imagens através de mecanismos de pesquisa, conseguimos um total de 100 imagens que iremos utilizar, sendo que para podermos utilizar determinada imagem resultante desse mecanismo estabelecemos determinadas características já mencionadas neste documento afim de podermos realizar a análise corretamente nas imagens.

A atual fase de desenvolvimento nós estamos salvando as características extraídas das anomalias no banco de dados Cassandra. O próximo passo é a geração da árvore de decisão com o WEKA e com este passo realizado estaremos com o protótipo quase finalizado, pois teremos as características relevantes já analisadas pelo WEKA e podemos utilizar esta árvore de decisão gerada para a classificação de novos dados em novas imagens.

# CONCLUSÃO

Utilizar técnicas de Processamento Digital de imagens aliadas com Mineração de Dados em uma área tão carente de tecnologia como a odontologia pode ser de muita importância para o cotidiano dos profissionais odontológicos, tanto na tomada de decisão de um diagnóstico quanto para o estudo e pesquisa nesta área. Entretanto, mesmo sendo aplicada a problemas odontológicos, há uma necessidade evidente em lançar mão de estratégias e boas práticas que aumentem as chances da utilização de alta tecnologia na rotina destes profissionais.

Ainda que a dificuldade no diagnóstico de anomalias e regiões problemáticas nas radiografias panorâmicas seja evidente, as informações clínicas geradas requerem apoio computacional para a tomada de decisões. Há fortes evidências no material de referência bibliográfica deste documento sobre a eficiência do Processamento Digital de Imagens e da Mineração de Dados na descoberta de conhecimentos, proporcionando eficiência para um melhor aproveitamento tanto pelos pacientes quanto pelos profissionais.

Embora seja possível identificar anomalias e regiões problemáticas nos exames de radiografias panorâmicas, a aplicação de tais questões está ligada ao contexto e deve ser detalhadamente planejada e avaliada. Por outro lado, o uso da extração de características associadas a estas anomalias para obter padrões permite ajudar na tomada de decisão dos profissionais para traçar tratamentos mais assertivos aos seus pacientes. Embora haja inúmeras experiências e relatos que indiquem dificuldades na visualização das anomalias em exames de radiografias panorâmicas, fica evidente a ausência de ferramentas que auxiliem estes profissionais odontológicos, o que denota uma série de oportunidades ricas e interessantes para a pesquisa sobre a aceitação de aplicação do Processamento Digital de Imagens juntamente com a Mineração de Dados em sistemas de apoio à decisão na área Odontológica.

# REFERÊNCIAS

ABERNETHY, Michael. **Mineração de dados com WEKA: Introdução e regressão**. Disponível em: <https://www.ibm.com/developerworks/br/opensource/library/os-weka1/index.html>. Acesso em: 04 out. 2017.

ALMEIDA, Solange Maria de et al. **Avaliação de três métodos radiográficos (periapical convencional, periapical digital e panorâmico) no diagnóstico de lesões apicais produzidas artificialmente.** Pesquisa Odontológica Brasileira, v. 15, n. 1, 2001, p. 56-63.

AZEVEDO, Samara Calçado. Detecção e restauração de áreas sombreadas em imagens digitais de alta resolução espacial. **Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia**, 2014.

BERNARDES, Ricardo Affonso. **Estudo comparativo entre as tomografias computadorizadas 3D, ortopantomográficas e radiografias periapicais no diagnóstico de lesões periapicais, fraturas radiculares e reabsorções dentais**. Tese de Doutorado. Universidade de São Paulo. 2007.

BIEZUNER, UFMG Rodney Josué; DE JESUS, Breno Felipe Rodrigues. **Curvas de Bézier**, 2014.

CASANOVA, Marcia Leal Spinelli; NETO, Francisco Haiter; OLIVEIRA, Ana Emília Figueiredo. **Avaliação da qualidade das imagens digitais panorâmicas adquiridas com diferentes resoluções.** Brazilian Dental Science,v. 5, n. 2, 2010.

CLOUDERA. About. **Cloudera: Ask Bigger Questions**, 2014. Disponível em: . Acesso em: 18 set. 2017.

CONSELHO FEDERAL DE ODONTOLOGIA. **Brasil é o país com o maior número de dentistas.** Disponível em: <http://cfo.org.br/sem-categoria/brasil-e-o-pais-com-o-maior-numero-de-dentistas>. Acesso em: 28 jun. 2017.

COSTA, Yandre. **Textura**. Disponível em: <http://www.din.uem.br/~yandre/PDI/textura-grande.pdf>. Acesso em: 25 out. 2017.

DALLAGASSA, Marcelo Rosano. **Concepção de uma metodologia para identificação de beneficiário com indicativos de diabetes melitos tipo 2**. 105 f. Dissertação (Mestrado em Tecnologia e Saúde). Pontifícia Universidade Católica do Paraná. 2009.

DAMASCENO, Marcelo . **INTRODUÇÃO A MINERAÇÃO DE DADOS UTILIZANDO O WEKA** . Disponível em: http://connepi.ifal.edu.br/ocs/index.php/connepi/CONNEPI2010/paper/viewFile/258/207>. Acesso em: 25 out. 2017.

DE ALBUQUERQUE, Márcio Portes; DE ALBUQUERQUE, Marcelo Portes. Processamento de imagens: métodos e análises. **Centro Brasileiro de Pesquisas Fısicas MCT**, 2000.

DE AZEVEDO-MARQUES, Paulo Mazzoncini. **Diagnóstico auxiliado por computador na radiologia.**Radiologia Brasileira, v. 34, n. 5, p. 285-293, 2001.

DE QUEIROZ, José Eustáquio Rangel; GOMES, Herman Martins. Introdução ao Processamento Digital de Imagens. **RITA**, v. 13, n. 2, p. 11-42, 2006.

DO VALE, Giovane Maia; DAL POZ, ALUIR PORFÍRIO. Processo de detecção de bordas de Canny. **Boletim de Ciências Geodésicas**, v. 8, n. 2, 2002.

DOTTA, Edivani A. Vicente; TELES, Guilherme Hp. **Sistemas aplicativos para uso odontológico**. RGO v. 51, n. 2, 2003, p. 119-122.

FACON, Jacques. A Morfologia Matemática e suas Aplicações em Processamento de Imagens. In: **VII Workshop de Visao Computacional–WVC 2011**. 2011.

FURQUIM, Tânia AC; COSTA, Paulo R. Garantia de qualidade em radiologia diagnóstica. **Revista brasileira de física médica**, v. 3, n. 1, p. 91-99, 2009.

GHEMAWAT, Sanjay; GOBIOFF, Howard; LEUNG, Shun-Tak. The Google file system. In: **ACM SIGOPS operating systems review**. ACM, 2003. p. 29-43.

GONDIM, Candice Regadas et al. Prevalência de dentes retidos presentes em radiografias panorâmicas. **Revista de Cirurgia e Traumatologia Buco-maxilo-facial**, v. 10, n. 3, p. 085-090, 2010.

GONZALES, R.C.; WOODS, R.E. **Processamento de Imagens Digitais.** Tradução: Cristina Yamagami e Leonardo Piamonte. São Paulo. 3ª Ed. Peaeson Prentice Hall, 2010. 624p.

HADOOP, Apache. **Apache Hadoop**. Disponível em: <https://wiki.apache.org/hadoop>. Acesso em: 20 ago. 2017.

HORTONWORKS, September 2013. **HOW TO PLAN AND CONFIGURE YARN AND MAPREDUCE 2 IN HDP 2.0**. Disponível em: <https://br.hortonworks.com/blog/how-to-plan-and-configure-yarn-in-hdp-2-0>. Acesso em: 16 set. 2017.

JESUS, Edison O.; COSTA JR, Roberto. A utilização de filtros gaussianos na análise de imagens digitais. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 3, n. 1, 2015.

LOTZ, MARCO AURÉLIO BARBOSA FAGNANI GOMES. PROCESSAMENTO DISTRIBUÍDO EXTENSÍVEL UTILIZANDO UM CLUSTER E MAPREDUCE COM APLICAÇÃO NA DETECÇÃO DE NÓDULOS PULMONARES PARA GRANDES VOLUMES DE TOMOGRAFIAS, **DEPARTAMENTO ACADÊMICO DE ELETRÔNICA. UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ,** 2014.

MACHADO, Anderson Fraiha. **Erosões e dilatações morfológicas binárias seqüênciais rápidas.** 2008. Tese de Doutorado. Universidade de São Paulo.

MARENGONI, Maurício; STRINGHINI, Stringhini. Tutorial: Introdução à visão computacional usando opencv. **Revista de Informática Teórica e Aplicada**, v. 16, n. 1, p. 125-160, 2009.

MURTHY, Arun C. **The next generation of apache hadoop MapReduce**. Disponível em**:** <http://developer.yahoo.com/blogs/hadoop/posts/2011/02/mapreduce-nextgen> 2011. Acesso em: 1 out. 2017.

NEVES, Samuel Clayton Maciel; PELAES, Evaldo Gonçalves; DE SINAIS, Laboratório de Processamento. Estudo e implementação de técnicas de segmentação de imagens. **Revista Virtual de Iniciação Acadêmica da UFPA-Universidade Federal do Pará–Departamento de Engenharia Elétrica e de Computação**, v. 1, n. 2, 2008.

NUNES, E. O.; CONCI, A. Segmentaçao por textura e localizaçao do contorno de regioes em imagens multibandas. **IEEE Latin America Transactions**, v. 5, n. 3, p. 185-192, 2007.

NUNES, Fátima LS. Introdução ao processamento de imagens médicas para auxílio a diagnóstico–uma visão prática. **Livro das Jornadas de Atualizações em Informática**, p. 73-126, 2006.

OLIVEIRA FARIA, Alessandro. **Apache Cassandra**: ARTIGOS: Apache Cassandra, conheça a tecnologia NoSQL. Disponível em: <https://pt.opensuse.org/ARTIGOS:Apache\_Cassandra,\_conhe%C3%A7a\_a\_tecnologia\_NoSQL>. Acesso em: 04 out. 2017.

PAES, Beatriz Lazera de Lima et al. Cisto ósseo simples: avaliação radiográfica, anatomopatológica e clínica de seis casos. **J. Health Sci. Inst**, v. 28, n. 1, 2010.

PAIVA, Werneck. **O que é e como funciona o Map Reduce usado pelo Google.** Disponível em: <http://blog.werneckpaiva.com.br/2011/08/como-funciona-o-map-reduce-usado-pelo-google/>. Acesso em: 23 out. 2017.

PEREIRA, Rafael Cardoso. **Técnica de rastreamento e perseguição de alvo utilizando o algoritmo Haar cascade aplicada a robôs terrestres com restrições de movimento**. 2017. Dissertação de Mestrado. Brasil.

PROCESSAMENTO DE IMAGENS. **AULA 16 - Processamento de Imagens.** Disponível em: <http://www.dpi.inpe.br/terraview/docs/tutorial/Aula16.pdf>. Acesso em: 04 ago. 2017.

ROSEN, Daren J. et al. Traumatic bone cyst resembling apical periodontitis. **Journal of periodontology**, v. 68, n. 10, p. 1019-1021, 1997.

SAIO, Armando Massao et al. RADIOGRAFIAS PANORÂMICAS, p. 18, 2005.

SILVA, Luciano. Segmentação de imagens de profundidade por detecção de bordas. **Setor de Ciências Exatas, Universidade Federal do Paraná**. 2011.

SIRLOPÚ, Chancafe; RAI, Jacinto; MAZABEL QUIJANDRÍA, Guillermo Andree. **Detección automática de caries utilizando reconocimiento de patrones en placas radiográficas**. 2016.

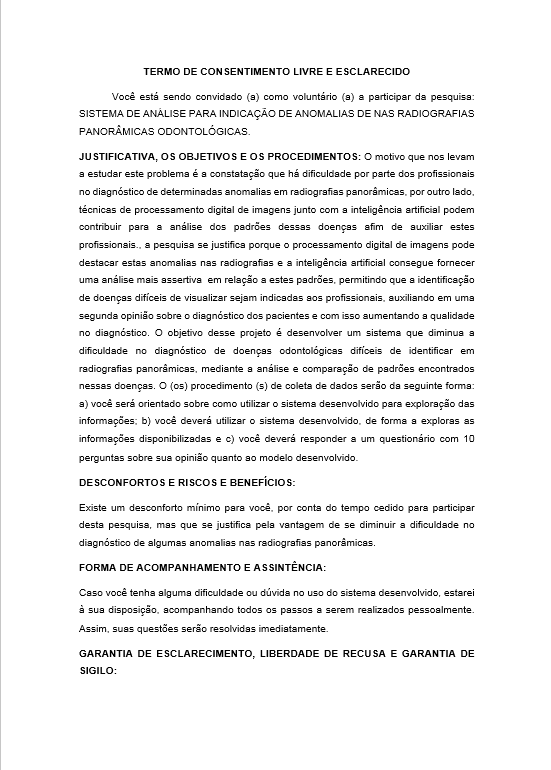
SOUZA, Taciana; CORREIA, Suzete. Estudo de técnicas de realce de imagens digitais e suas aplicações. In: **II Congresso de Pesquisa e Inovação da Rede Norte Nordeste de Educação Tecnológica**. 2007. p. 3-10.

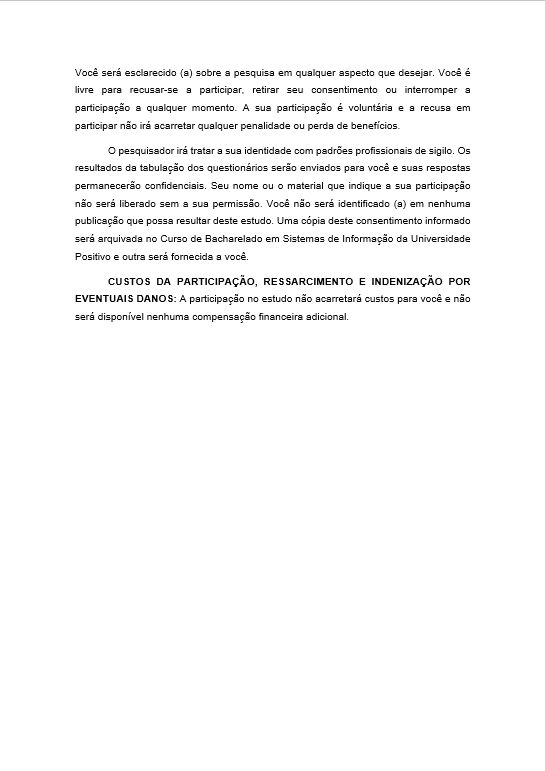
VELA, Jaime Grande et al. Digitalização de filmes radiográficos com costura de imagens. **Radiol Bras**, v. 44, p. 233-7, 2011.

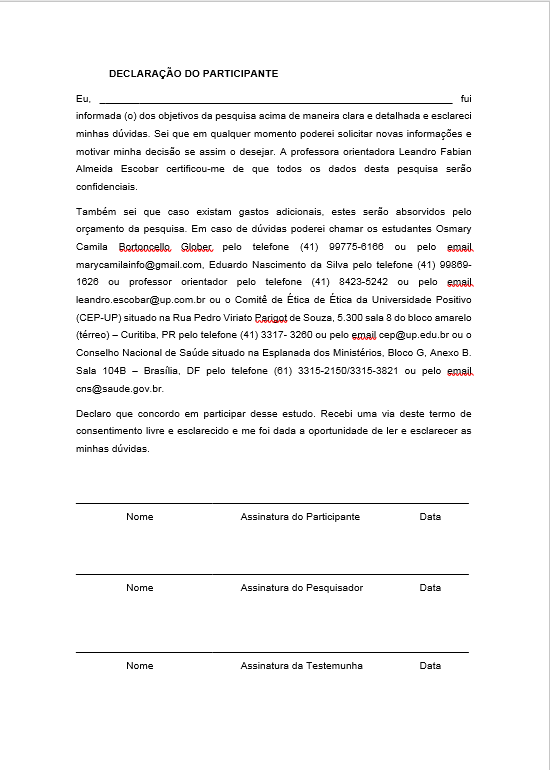
WHITE, Tom. **Hadoop: The definitive guide**. “O’Reilly Media, Inc.", 2012.

ZANET, Tiago Gorgulho. **Sistema de apoio à decisão diagnóstica baseado em caracteristicas radiográficas**. Tese de Doutorado. Universidade de São Paulo. 2009.

# APÊNDICE







# ANEXO

