UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E
COMPUTAÇÃO

MICHEL BESSANI

Avaliação e Modelagem de Sistemas de Suporte à
Decisão utilizando Reconhecimento de Padrões e
Redes Bayesianas

São Carlos

2015

MICHEL BESSANI

**Avaliação e modelagem de Sistemas de
Suporte à Decisão utilizando
Reconhecimento de Padrões e Redes
Bayesianas**

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Ciências, Programa de Engenharia Elétrica.

Área de Concentração: Sistemas Dinâmicos

Orientador: Prof. Dr. Carlos Dias Maciel

Trata-se da versão original.

São Carlos

2015

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Bessani, Michel

B557a Avaliação e Modelagem de Sistemas de Suporte à
Decisão utilizando Reconhecimento de Padrões e Redes
Bayesianas / Michel Bessani; orientador Carlos Dias
Maciel. São Carlos, 2015.

Dissertação (Mestrado) - Programa de Pós-Graduação
em Engenharia Elétrica e Área de Concentração em
Sistemas Dinâmicos -- Escola de Engenharia de São
Carlos da Universidade de São Paulo, 2015.

1. Sistemas Especialistas. 2. Sistemas de Suporte a
Tomada de Decisão Clínica. 3. Modelagem de Sistemas. 4.
Redes Bayesianas. 5. Reconhecimento de Padrões. I.
Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **MICHEL BESSANI**.

Título da dissertação: "Avaliação e modelagem de sistemas de suporte à decisão utilizando reconhecimento de padrões e redes bayesianas".

Data da defesa: 09/02/2015

Comissão Julgadora:

Resultado:

Prof. Dr. **Carlos Dias Maciel (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

Aprovado

Prof. Associado **Alexandre Cláudio Botazzo Delbem**
(Instituto de Ciências Matemáticas e de Computação/ICMC)

Aprovado

Prof. Dr. **Emery Cleiton Cabral Correia Lins**
(Universidade Federal do ABC/UFABC)

Aprovado

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Associado **Luís Fernando Costa Alberto**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Paulo César Lima Segantine**

Dedicatória

Dedico este trabalho aos meus pais, Oscar e Irene, a minha irmã Lorena e a minha avó Ana Santana.

Agradecimentos

Em primeiro lugar a minha família que sempre me apoiou independente das decisões, em segundo a Viviane Jin Hee Kim que me ajuda com sua calma e paciência em todos os momentos.

Aos amigos do Laboratório de Processamento de Sinais (LPS) pelas conversas, discussões e risadas que com certeza influenciaram no resultado final deste trabalho.

Gostaria de agradecer a todos que de alguma forma contribuíram para o meu amadurecimento, tanto pessoal quanto intelectual, nestes dois anos de trabalho no programa de pós-graduação da Eng. Elétrica.

Ao CNPq pelo apoio financeiro no decorrer do curso de Mestrado.

*“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém
ainda pensou sobre aquilo que todo mundo vê.”*

Arthur Schopenhauer

Sumário

LISTA DE FIGURAS	xv
ÍNDICE DE TABELAS.....	xviii
Lista de Abreviaturas	xix
RESUMO.....	xxi
ABSTRACT.....	xxiii
1. INTRODUÇÃO.....	1
1.1. Objetivos	4
1.2. Área de aplicação escolhida.....	5
1.3. Organização do trabalho.....	11
2. FUNDAMENTOS	12
2.1. Redes Bayesianas	12
2.1.1. Teorema de Bayes	12
2.1.2. Redes Probabilísticas	14
2.1.3. D-separação	17
2.1.4. Redes Bayesianas	18
2.2. Cárie Dental	20
2.3. Imagens de Fluorescência.....	22
2.4. Fluorescência Dental	23
3. MATERIAIS E MÉTODOS.....	26
3.1. Gerenciamento da Cárie:	26

3.1.1. Escolha das Variáveis da Rede Bayesiana	26
3.1.2. Estrutura da Rede Bayesiana.....	29
3.1.3. Probabilidades Condicionais das Variáveis da Rede Bayesiana.....	33
3.2. Detecção da Cárie.....	34
3.2.1. Conjunto de Dados	35
3.2.2. Imagens de fluorescência.....	35
3.2.3. Extração de características de imagens.....	37
3.2.4. Análise de Componentes Principais.....	40
3.2.5. Agrupamento.....	41
4. RESULTADOS E DISCUSSÃO	45
4.1. Gerenciamento da Cárie	45
4.2. Detecção da Cárie	48
4.3. Trabalhos Publicados ou Aceitos em Congressos	55
4.3.1. XVIII Congresso Brasileiro de Física Médica - 2013	55
4.3.2. XVIII Congresso Brasileiro de Física Médica – 2013 - Resumo.....	56
4.3.3. SPIE 2014 - Multimodal Biomedical Imaging IX	57
4.3.4. XXIV Brazilian Congress on Biomedical Engineering – CBEB 2014.....	58
4.3.5. XXIV Brazilian Congress on Biomedical Engineering – CBEB 2014.....	59
5. CONCLUSÃO	60
REFERÊNCIAS	63

LISTA DE FIGURAS

Figura 1 - Estrutura geral de um ES. O tomador de decisão pode influenciar e ser influenciado pelas entradas (informações) e saídas (sugestões) do sistema. O sistema de inferência precisa ser capaz de se comunicar com o conhecimento específico armazenado no sistema para poder inferir as sugestões ao tomador de decisão. Os resultados obtidos a partir das sugestões para novos casos podem ser utilizados para atualizar o conhecimento específico do ES. Adaptado de (SHAW et al., 2008).	2
Figura 2 – Exemplo de uma lesão cariosa cavitada. a) lesão cariosa cavitada, b) tecido lesionado removido através de jato de ar, observa-se a dimensão da lesão após a remoção do tecido lesado. Adaptado de (KIDD, 2005).	7
Figura 3 – Comparação dos pensamentos essencialista e nominalista no processo de diagnóstico. O pensamento essencialista busca a detecção do problema, já o pensamento nominalista considerada a etiologia. Adaptado de (BAELUM, 2008).	8
Figura 4 – Diagrama de Veen para o cenário descrito no exemplo de aplicação do teorema de Bayes.	13
Figura 5 – Exemplo de uma conexão serial entre variáveis de uma rede probabilística, evidências em A influenciam B, que por sua vez influencia C.	15
Figura 6 – Exemplo de uma conexão divergente entre variáveis de uma rede probabilística, evidências em A influenciam diretamente as variáveis B, C e D.	16
Figura 7 – Exemplo de uma conexão convergente entre variáveis de uma rede probabilística. A variável A é influenciada pelas variáveis B, C e D e evidências em qualquer um dos nós pais influenciam o nó filho A.	16
Figura 8 – Exemplo de variáveis d-separáveis em um rede probabilística. Considerando que o nó V pode ser evidenciado, qualquer certeza em V torna os seus nós filhos condicionalmente independentes dos seus nós pais.	17
Figura 9 – Rede Bayesiana representado o problema exemplo a respeito do Câncer Metastático. As relações de causalidade entre as variáveis envolvidas são explicitadas através da orientação das arestas na rede.	19
Figura 10 – Espectro de emissão do sinal de fluorescência para dentes saudáveis e cariados utilizando uma fonte de excitação em 405 nm. Observa-se a redução na intensidade do sinal de fluorescência para amostras cariadas em relação a amostras saudáveis entre 450 nm e 600 nm. Adaptado de (ZEZELL et al., 2007).	23
Figura 11 – Exemplo de uma imagem de fluorescência dental obtida em bancada utilizando uma amostra in vitro. (a) Imagem realizada utilizando uma iluminação branca, (b) imagem de fluorescência obtida em bancada utilizando fonte de excitação UV (405 nm) e instrumentação óptica, a seta vermelha indica região com sinal de fluorescência dos microorganismos, a seta preta mostra uma região com sinal de fluorescência de menor intensidade que o resto da superfície visualizada.	24

Figura 12 – Primeira parte da BN desenvolvida, considera os fatores causais e preditivos da cárie para uma sugestão de risco de cárie de acordo com o cenário em que cada paciente se encontra.	30
Figura 13 – Segunda parte da BN desenvolvida, considera a situação clínica e o risco de cárie do paciente para sugestões de tratamento e de previsão de retorno considerando o cenário individual de cada paciente.	31
Figura 14 – Rede Bayesiana desenvolvida para auxílio a tomada de decisão clínica para o gerenciamento da cárie, os nós em branco são as variáveis que o profissional clínico pode evidenciar na Rede, os nós em azul são as variáveis que a Rede fornecerá uma distribuição de probabilidade para os possíveis estados, referentes às decisões. O nó Risco de Cárie pode ser evidenciado pelo profissional clínico caso seja obtida alguma conclusão em relação ao risco do paciente analisado.	33
Figura 15 – Relação entre expressões de certeza e probabilidades para eventos utilizada para obtenção das probabilidades condicionais entre as diferentes variáveis da CMBN proposta. Adaptado de (KJÆRULFF; MADSEN, 2008).	34
Figura 16 – Amostras exemplificando a diversidade nos estágios das lesões e superfícies dentais presentes no conjunto de dados utilizado.	35
Figura 17 – Montagem óptica realizada em bancada para a aquisição das imagens de fluorescência das amostras in vitro. O círculo azul indica o LED UV de excitação, a caixa verde os filtros ópticos e a caixa vermelha a câmera USB e a lente de zoom manual. A seta amarela indica a posição da amostra para aquisição das imagens.	36
Figura 18 – Sistema de aquisição de imagens de fluorescência utilizando configuração epifluorescência. As setas azuis indicam o sinal UV, as setas verdes o sinal de fluorescência que serão adquiridos pela câmera USB, a lente amarela representa um filtro passa longa, a lente roxa um filtro passa banda. Estes componentes precisam ser utilizados com este alinhamento óptico para a correta aquisição das imagens.	36
Figura 19 – Espectro de transmissão dos filtros ópticos utilizados na montagem em bancada para aquisição das imagens de fluorescência. O gráfico foi gerado utilizando as informações fornecidas pelos fabricantes dos filtros.	37
Figura 20 – Espaço de cores RGB, cada eixo representa a intensidade de brilho de uma cor, combinações de intensidades do vermelho, verde e azul podem produzir diferentes cores, adaptado de (GONZALEZ; WOODS, 2008).	38
Figura 21 – Metodologia utilizada para modelar um conhecimento específico para um sistema de auxílio a detecção de cárie utilizando imagens de fluorescência dental.	48
Figura 22 – Scree plot representando a variância de cada uma das 35 características após a utilização do PCA (barras verdes), as características foram ordenadas em valores decrescentes de variância. Os diamantes azuis representam a porcentagem cumulativa da variância total que cada característica adiciona.	49

Figura 23 – Dendrograma obtido utilizando os onze componentes mais significativos após o PCA para agrupamento hierárquico não-supervisionado do conjunto de dados. Pode-se observar a divisão que ocorre nas amostras, na parte superior do dendrograma se encontram as amostras não cariadas, já na parte inferior as amostras cariadas.	50
Figura 24 – Resultado obtido utilizando k-means (k=2) com apenas as duas características mais significantes após o uso do PCA. O Cluster 1 em azul representa o grupo das amostras cariadas, o Cluster 2 em verde representa o grupo das amostras não cariadas, os centroides dos cluster estão identificados em preto.....	52
Figura 25 – Dendrograma obtido utilizando os dois componentes mais significativos após o PCA para agrupamento hierárquico não-supervisionado do conjunto de dados. Pode-se observar a divisão que ocorre nas amostras, na parte superior do dendrograma se encontram as amostras não cariadas, já na parte inferior as amostras cariadas, similar ao dendrograma anterior. Porém a similaridade intra grupo e a distância entre grupos aumentaram em relação ao dendrograma anterior.	53
Figura 26 – Coeficiente de correlação cofenética calculado para o vetor de características original (VCO) e para as saídas do PCA. Os números (n –PCA) indicam a quantidade de componentes do PCA que foram utilizadas para a obtenção dos coeficientes. O resultado é coerente com a melhora de performance observada na saída do k-means (Tabela 9), utilizando apenas os dois componentes mais significativos obtêm-se o maior valor para o coeficiente de correlação cofenética. Tal resultado reforça a evidencia de que algumas características extraídas estão atuando como ruído no processo de agrupamento.	54

ÍNDICE DE TABELAS

Tabela 1 – Probabilidades condicionais para as variáveis relacionadas na Figura 9. Observa-se que as probabilidades condicionais de cada variável são apenas dependentes das variáveis predecessoras, ou nós pais.	20
Tabela 2 – Estados e grau de influência das variáveis causais e preditivas consideradas para a sugestão do risco de cárie. Grau de influência I é o mais forte e III o mais fraco.	31
Tabela 3 - Estados e grau de influência das variáveis causais e preditivas consideradas para a sugestão do tratamento e retorno do paciente. Grau de influência I é o mais forte e III o mais fraco.	32
Tabela 4 – Descrição dos possíveis estados das duas variáveis de saída da BN: tratamento e retorno.	32
Tabela 5 – Características estatísticas extraídas utilizando a matriz de co-ocorrência de cada componente de representação de cores, adaptado de (GONZALEZ; WOODS, 2008), i e j são as colunas e linhas da matriz de co-ocorrência e p é o valor em cada coordenada.....	40
Tabela 6 – Sugestões de risco de cárie obtidas a partir da Rede Bayesiana desenvolvida utilizando alguns casos bem definidos. Tem-se o valor evidenciado em cada um dos nós pais e a distribuição de probabilidade obtida no nó filho “Risco de Cárie”. B, M e A respectivamente representam baixo, médio e alto.	46
Tabela 7 – Sugestões tratamento obtidas a partir da Rede Bayesiana desenvolvida utilizando alguns casos bem definidos. Tem-se o valor evidenciado em cada um dos nós pais e a distribuição de probabilidade obtida no nó filho “Tratamento”. F/S, Rest e Endo respectivamente representam fluoreto ou selante, restauração e tratamento endodôntico.	46
Tabela 8 – Sugestões tratamento obtidas a partir da Rede Bayesiana desenvolvida utilizando alguns casos bem definidos. O valor evidenciado em cada um dos nós pais e a distribuição de probabilidade obtida no nó filho “Retorno”	47
Tabela 9 – Resultados de performance obtidos com os vetores de características, tanto o original, quanto os processados, comparando os rótulos obtidos com o algoritmo k-means (k=2) com a classificação prévia de um especialista. A primeira linha da tabela indica os resultados obtidos para o vetor de características original, a segunda linha representa os resultados obtidos para os vetores de característica utilizando de 3 a até 11 elementos após o PCA, a terceira linha representa a utilização de apenas as 2 características mais significantes após o PCA.	51

Lista de Abreviaturas

DSS – *Decision Support System* (Sistema de suporte a decisão)

ES – *Expert System* (Sistema especialista)

BN – *Bayesian Network* (Rede Bayesiana)

CDSS – *Clinical Decision Support System* (Sistema de suporte a decisão clínica)

CMBN – *Caries Management Bayesian Network* (Rede Bayesiana para gerenciamento da cárie)

OMS – Organização Mundial de Saúde

DAG – *Directed Acyclic Graph* (Grafo acíclico direto)

Pr – Nós pais

C – Nós filhos

ECC – *Early Childhood Caries* (Cárie na infância)

UV – Ultravioleta

ICDAS – *International Caries Detection and Assessment System* (Sistema internacional de detecção e avaliação de cáries)

DMFT – *Decayed, Missing and Filling Teeth* (Dentes cariados, restaurados ou faltantes)

PCA – *Principal Components Analysis* (Análise de componentes principais)

RESUMO

Bessani, Michel. **Avaliação e modelagem de Sistemas de Suporte à Decisão utilizando Reconhecimento de Padrões e Redes Bayesianas**. 2015. 90p. Dissertação. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2015.

Sistemas de suporte a decisão são utilizados em cenários com incertezas. Uma decisão normalmente é auxiliada por resultados obtidos com ações passadas em problemas semelhantes. Quando um sistema de suporte a decisão incorpora conhecimento específico de uma área, estes recebem o nome de sistemas especialistas. Tal conhecimento específico é utilizado para inferência juntamente com as informações de entrada a respeito do problema. O objetivo deste trabalho é a avaliação e modelagem de sistemas de auxílio a decisão, foram analisadas duas abordagens para um mesmo problema alvo, sendo uma de gerenciamento do problema e outra de detecção do problema. A abordagem de gerenciamento utiliza redes Bayesianas para modelagem, tanto do conhecimento específico quanto para a inferência. As variáveis utilizadas, as relações de dependência e as probabilidades condicionais entre as variáveis foram extraídas da literatura. A abordagem de detecção do problema utilizou imagens para extração de características seguida de um algoritmo de agrupamento para comparação com a classificação de um especialista. Uma das áreas de aplicação de sistemas especialistas é na área clínica, podendo auxiliar tanto na detecção, diagnóstico e tratamento de doenças. A cárie dental é um problema generalizado que afeta a maioria das pessoas, tanto em países ricos, como em países pobres. Existem poucos sistemas para auxílio no processo de diagnóstico da cárie, sendo a maior parte dos sistemas existentes determinísticos, focando apenas na detecção da lesão. O sistema de gerenciamento da cárie desenvolvido foi apresentado a dois profissionais da odontologia, a opinião deles mostra que esta abordagem é promissora e aplicável em campos como a educação e a atenção básica a saúde. Além da apresentação aos profissionais, foram utilizados casos bem estabelecidos da literatura para analisar as sugestões fornecidas pela Rede, e o resultado foi coerente com o cenário real de tomada de decisão. A metodologia de detecção da cárie resultou em um alto valor de acurácia, 96.88%, mostrando que tal metodologia é promissora em comparação com outros trabalhos da área. Além da contribuição para a área de informática odontológica, os resultados mostram que a extração da estrutura e das probabilidades condicionais da rede a partir da literatura é uma metodologia que pode ser utilizada em outras áreas com cenário similar ao do diagnóstico da cárie. Nos próximos passos do projeto alguns pontos referentes a modelagem de sistemas e redes Bayesianas serão analisados, como escalabilidade e testes de validação, tanto quantitativamente como qualitativamente, isto inclui o desenvolvimento de métodos computacionalmente efetivos para a geração de casos aleatórios utilizando o Método de Monte Carlo.

Palavras-chaves: Sistemas Especialistas, Sistemas de Suporte a Tomada de Decisão Clínica, Modelagem de Sistemas, Redes Bayesianas, Reconhecimento de Padrões.

ABSTRACT

*Bessani, Michel. **Assessment and Modeling of Decision Support Systems using Pattern Recognition and Bayesian Networks**. 2015. 90p. Dissertation. University of São Paulo at São Carlos School of Engineering, São Carlos, 2015.*

Decision support systems are used in uncertainty scenarios; normally a decision is choose using similar problems actions results. Decision support systems could incorporate specific knowledge; such systems are called expert systems. The specific knowledge is used for inference about the problem scenario. This work objective is the evaluation and modeling of decision support systems, we analyzed two distinct approaches for the same problem, one for detection, another for management. The management approach uses Bayesian networks for modeling the specific knowledge and the inference engine. The variables choice, the dependences relationship and the conditional probabilities were extracted from the scientific literature. The detection approach used images and feature extraction to perform a clustering and compare the output labels with a specialist classification. One application of expert systems is clinical, supporting diseases detection, diagnosis and treatment. Dental caries is a generalized problem that affects major part of the population, few systems exists for support the caries diagnostic process, the major part is deterministic, focusing only the detection problem. The caries management system developed here was shown to two odontology professionals, and they opinion encourage such approach to be applied in fields like odontology education and basic health. Beyond this, we used well-established cases to analyze the network output suggestions, the result obtained was coherent with the real decision making scenario. The caries detection approach resulted in a high accuracy, 96.88%, showing that methodology is promising. Besides the contribution for dental informatics field, the results obtained here shows that the extraction of the network structure from the literature could be used in problems similar with caries diagnoses. The project next steps are to analyze some points of systems modeling and Bayesian networks, like scalability and validation tests, both quantitative and qualitative, and including the development of computational effective's methods for the use of Monte Carlo methodology.

Keywords: *Expert Systems, Clinical Decision Support Systems, System Modeling, Bayesian Network, Pattern Recognition.*

1. INTRODUÇÃO

A incerteza se faz presente na maioria das aplicações, tanto nas metodologias, quanto nas soluções. É uma consequência da limitação de informações a respeito de qualquer sistema (KOLLER; FRIEDMAN, 2009), uma vez que nenhuma observação pode ser completa. O desenvolvimento de sistemas de suporte a decisão (DSS, do inglês *decision support systems*) é uma forma de auxiliar a tomada de decisão considerando tais incertezas. DSS possuem aplicabilidade em diversas áreas: emergências em estações de energia nuclear (PAPAMICHAIL; FRENCH, 2005), diagnósticos médicos (WAGHOLIKAR; SUNDARARAJAN; DESHPANDE, 2012), logística de cadeia de suprimentos (FINK; SEBASTIAN; VOS, 2012) e gestão de crises e catástrofes (PENG et al., 2011).

Além da incerteza, existe o fato de que as habilidades cognitivas de qualquer profissional são restritas, como limite de raciocínio e memória, consequentemente é impossível que um profissional sozinho realize a decisão correta todas as vezes (WAGHOLIKAR; SUNDARARAJAN; DESHPANDE, 2012). Tal fato também leva a divergência nas decisões entre diferentes profissionais para situações semelhantes.

A tomada de decisão pode ser baseada no conhecimento ou experiências anteriores com problemas similares. Basicamente, uma decisão é tomada considerando o desempenho de ações já utilizadas em problemas semelhantes ao atual (HÜLLERMEIER, 2007). Os conhecimentos ou experiências de uma área específica pertencem aos especialistas desta área (TURBAN; ARONSON; LIANG, 2005), quando um DSS incorpora o conhecimento específico estes recebem o nome de sistemas baseados em conhecimento para suporte a decisão (KBDSS, do inglês *knowledge-based decision support system*), ou simplesmente sistemas especialistas (ES, do inglês *expert system*) (HOLMES; JAIN, 2008; KJÆRULFF; MADSEN, 2008; TURBAN; ARONSON; LIANG, 2005).

A motivação para o desenvolvimento de um ES é a possibilidade de automatizar uma tarefa que seja recorrente e que envolva raciocínio e tomada de

decisão (KJÆRULFF; MADSEN, 2008). ES realizam tarefas como extração de informações a partir de dados e raciocínio, aplicando técnicas e ferramentas disponíveis da área da inteligência artificial (TURBAN; ARONSON; LIANG, 2005).

Um sistema especialista é capaz de incorporar conhecimento específico de um domínio e realizar tarefas como aprendizagem e inferência (GUPTA; FORGIONNE; MORA, 2006). ES utilizam o conhecimento incorporado para obter alto nível de desempenho em um problema de domínio específico e fornecer uma segunda opinião confiável ao usuário, sempre o auxiliando e nunca o substituindo (HOLMES; JAIN, 2008; KJÆRULFF; MADSEN, 2008; TURBAN; ARONSON; LIANG, 2005).

Estruturamos um ES como na Figura 1, o tomador de decisão pode influenciar e ser influenciado pelas entradas (informações) e saídas (sugestões) do sistema. O sistema de inferência precisa ser capaz de se comunicar com o conhecimento específico armazenado no sistema para poder inferir as sugestões ao tomador de decisão (SHAW et al., 2008).

Figura 1 - Estrutura geral de um ES. O tomador de decisão pode influenciar e ser influenciado pelas entradas (informações) e saídas (sugestões) do sistema. O sistema de inferência precisa ser capaz de se comunicar com o conhecimento específico armazenado no sistema para poder inferir as sugestões ao tomador de decisão. Os resultados obtidos a partir das sugestões para novos casos podem ser utilizados para atualizar o conhecimento específico do ES. Adaptado de (SHAW et al., 2008).

As informações necessárias para o ES realizar a inferência a respeito do problema alvo são as entradas do sistema, que podem ser obtidas a partir de sensores ou de observações do tomador de decisão. A partir da entrada o sistema fornece sugestões de ações, como por exemplo, uma tabela de probabilidade de sucesso para cada ação possível.

O conhecimento específico de uma área, seja a partir de um especialista ou de conhecimento científico, é obtido através da modelagem dessas informações (SHAW et al., 2008). O uso de técnicas de mineração de dados (ZHANG; WU, 2011) é uma das formas para modelar tais informações e transformá-las em conhecimento (SHAW et al., 2008).

O sistema de inferência é responsável pela análise das entradas utilizando o conhecimento específico armazenado no sistema e inferência a respeito das possíveis ações (DARWICHE, 2009; SHAW et al., 2008). Os ES podem utilizar diferentes formas de inferência (HÜLLERMEIER, 2007; KJÆRULFF; MADSEN, 2008; LIAO, 2005; TURBAN; ARONSON; LIANG, 2005), como medidas de similaridade e utilidade, associações de regras *if-then*, redes neurais e redes Bayesianas. Diferentemente do conhecimento específico que é dependente do domínio do problema, um mesmo sistema de inferência pode ser utilizado em diferentes cenários (DARWICHE, 2009).

Interfaces que possibilitam iterações dinâmicas entre o usuário e o sistema durante a inferência é uma das formas de melhorar o desempenho de um ES (BEEMER; GREGG, 2010). Tais iterações contribuem para o envolvimento do usuário no processo de decisão e consequentemente na consciência de como o sistema chega as sugestões, aumentando de forma geral a confiança e a percepção de usabilidade do sistema.

Em cenários que estão sempre em mudança, é comum o desenvolvimento de ES com característica incremental (BEEMER; GREGG, 2010). Através de técnicas de mineração de dados, o conhecimento específico é atualizado à medida que surgem novas informações a partir dos desfechos de novos casos. As alterações no banco de dados (inserção, alteração e remoção de informações) possuem alto custo computacional. Quando tal custo inviabiliza a característica

incremental do ES, o desenvolvimento de uma nova metodologia de armazenamento de informações pode ser a solução.

1.1. Objetivos

O objetivo geral deste trabalho é a avaliação e modelagem de sistemas de auxílio a decisão. Para tanto, serão analisadas duas abordagens distintas para um mesmo problema alvo.

A primeira é uma modelagem probabilística utilizando redes Bayesianas (BN, do inglês *Bayesian network*) (DARWICHE, 2009) para auxílio ao processo de decisão em um contexto de gerenciamento do problema. A escolha de redes Bayesianas foi devido a sua capacidade de armazenar e processar conhecimento probabilístico através de uma representação gráfica e intuitiva (KJÆRULFF; MADSEN, 2008). Uma descrição formal de redes Bayesianas está presente na Seção 2.1.

Nesse caso, o modelo probabilístico desenvolvido é baseado no caráter multifatorial e causal do problema, foram utilizadas as evidências da literatura especializada para extração do conhecimento específico a ser utilizado pela rede. Pode-se estruturar esta abordagem da seguinte forma:

- Extração do conhecimento a partir da literatura especializada;
- Modelagem da estrutura da BN;
- Modelagem das probabilidades condicionais entre as variáveis da BN;
- Avaliação da performance do modelo utilizando casos bem estabelecidos da área juntamente com a opinião de especialistas.

O objetivo de tal abordagem é o desenvolvimento de uma sistema que possa fornecer uma segunda opinião rápida e confiável ao profissional no momento da decisão, e também o desenvolvimento de uma aplicação utilizando da teoria de redes Bayesianas (DARWICHE, 2009).

A segunda abordagem consiste na modelagem do conhecimento específico que poderá ser utilizado para o desenvolvimento de um classificador para a

detecção do problema. Resumidamente, consiste na utilização de técnicas de mineração de dados em imagens e na investigação da possibilidade de extração de características que consigam descrever de forma clara o estado das amostras em relação a presença ou não do problema.

O foco desta segunda abordagem é a modelagem do conhecimento específico, que poderá ser utilizado em um ES para detecção do problema. O objetivo é avaliar analiticamente a possibilidade de desenvolvimento desse sistema, explorando a eficácia da utilização de técnicas de extração de características de imagens juntamente com análise de componentes principais para detecção do problema. Pode-se estruturar os objetivos desta abordagem da seguinte forma:

- Modelagem do conhecimento específico;
- Extração de características;
- Processamento e Análise das características extraídas;
- Avaliação da eficácia do modelo obtido.

Como critério de avaliação dos descritores, compararam-se os resultados de um algoritmo de agrupamento (EVERITT et al., 2011) (Seção 3.2.4), com a classificação prévia de um especialista.

1.2. Área de aplicação escolhida

Uma aplicação específica dos ES é na área de suporte a tomada de decisão clínica, sendo conhecidos como sistemas de suporte a decisão clínica (CDSS, do inglês *Clinical Decision Support System*), que podem ser definidos como (GREENES, 2007): o uso de computadores para acessar as informações científicas mais relevantes para o suporte a tomada de decisões relacionadas aos cuidados e a melhora da saúde baseado sempre nas características individuais de cada paciente. Os CDSS tentam oferecer uma ou mais sugestões para o profissional lidar com a incerteza durante uma decisão clínica.

Os CDSS podem auxiliar na tomada de decisão clínica de diagnóstico, prevenção e tratamento de problemas de saúde, exemplos são encontrados na odontologia, medicina, farmacologia e outras áreas da saúde (MENDONÇA, 2004).

Em um cenário de decisão clínica existem diversas variáveis a serem consideradas, como por exemplo: múltiplas doenças, sintomas, exames e características pessoais. Essas variáveis podem ser consideradas aleatórias, e todas influenciam na decisão. Em geral, tal cenário de variáveis aleatórias inter-relacionadas resultam em um sistema complexo (KOLLER; FRIEDMAN, 2009).

Uma análise sistêmica atual mostrou que doenças orais continuam sendo um grande problema ao redor do mundo (MARCENES et al., 2013), sendo a cárie dental a mais comum. Nos países desenvolvidos, a cárie dental afeta a maioria dos indivíduos em desvantagem econômica, já nos países em desenvolvimento ela afeta a maioria da população (COSTA et al., 2012; ZERO, 2004). Segundo relatório da Organização Mundial de Saúde (OMS) a cárie afeta de 60% até 90% dos jovens e 90% dos adultos ao redor do mundo (PETERSEN, 2003). A cárie é uma doença crônica que se desenvolve lentamente na maioria dos indivíduos (SELWITZ; ISMAIL; PITTS, 2007), sendo consequência de um processo fisiológico e dinâmico que ocorre na cavidade oral envolvendo a comunidade de microorganismos presentes na boca (KIDD, 2005; SELWITZ; ISMAIL; PITTS, 2007).

Um diagnóstico é baseado na suposição de que uma patologia pode ser distinguida através de seus sinais e sintomas (NYVAD, 2004). O raciocínio de diagnóstico é um processo complexo que envolve elementos de reconhecimentos de padrões, considerações probabilísticas e raciocínio dedutivo. A elaboração de um tratamento para o paciente é a etapa final do processo de diagnóstico (FEJERSKOV; KIDD, 2008).

Para a realização de um processo de diagnóstico da cárie é necessário uma grande quantidade de informação, como fatores de risco, benefícios dos diversos tratamentos, taxas de incidência, progressão, entre outros (KIDD, 2005; LIMEBACK, 2012; SELWITZ; ISMAIL; PITTS, 2007). Todos estes fatores estão abertos a interpretações e julgamentos dos profissionais clínicos (BAELUM, 2008; PITTS, 2004a). Uma consequência do julgamentos individual destes fatores é a

variação que existe no modo como é realizado o diagnóstico e tratamento relacionados à cárie (BAELUM; HEIDMANN; NYVAD, 2006; BAELUM, 2008).

O entendimento da cárie como uma doença e a lesão como uma consequência dessa doença é conhecido como pensamento essencialista (BAELUM; HEIDMANN; NYVAD, 2006; BAELUM, 2008; CAROUNANIDY; SATHYANARAYANAN, 2009). Tal raciocínio assume a crença de que existe um padrão ouro de diagnóstico, ou seja, que o limiar entre o processo fisiológico e o patológico existe e é bem definido.

Considerando que a cárie é consequência de um processo fisiológico complexo e dinâmico, esta abordagem no dia-a-dia clínico pode levar a um diagnóstico tardio. Consequentemente, a lesão cariosa pode se encontrar em estágios mais avançados e presente na forma de cavitação (SELWITZ; ISMAIL; PITTS, 2007), como exemplificado na Figura 2.

Figura 2 – Exemplo de uma lesão cariosa cavitada. a) lesão cariosa cavitada, b) tecido lesionado removido através de jato de ar, observa-se a dimensão da lesão após a remoção do tecido lesado. Adaptado de (KIDD, 2005).

O pensamento nominalista é oposto ao essencialista (BAELUM; HEIDMANN; NYVAD, 2006; BAELUM, 2008; CAROUNANIDY; SATHYANARAYANAN, 2009). De acordo com tal pensamento a doença é apenas um rótulo para os sintomas, Figura 3. A doença é apenas um processo causal que resulta em sinais e sintomas clínicos, o objetivo geral dessa abordagem é chegar ao melhor tratamento, o qual proporcionará os melhores resultados ao paciente atacando as causas e não somente os sintomas.

- **Pensamento Essencialista**

- **Pensamento Nominalista**

Figura 3 – Comparação dos pensamentos essencialista e nominalista no processo de diagnóstico. O pensamento essencialista busca a detecção do problema, já o pensamento nominalista considera a etiologia. Adaptado de (BAELUM, 2008).

O pensamento nominalista é uma abordagem terapêutica focada no prognóstico e não necessita de um “padrão ouro”. Assim o paciente só será rotulado com a doença caso seja necessário realizar algum tratamento (BAELUM; HEIDMANN; NYVAD, 2006). O diagnóstico passa a ser relacionado com a crença de que o tratamento trará benefícios ao paciente tratando dos fatores causais que geram os sinais e sintomas. Resumidamente, existem duas linhas de raciocínio para o diagnóstico, uma de cunho determinístico, preocupado com a detecção do problema, e outra de cunho causal, considerando a doença apenas como um rótulo para um problema multifatorial.

A odontologia praticada atualmente é baseada nas habilidades manuais de restauração do dentista, que são embutidas e integradas na educação, na licença profissional, nos planos de saúde e também na opinião pública (BAELUM, 2008). Os tratamentos restaurativos são serviços tangíveis e bem definidos, sendo melhor aceitos pelos envolvidos, profissionais clínicos e pacientes, e também pelos princípios de mercado que envolvem a odontologia.

Assim, as abordagens menos “visíveis” como selantes, aplicações tópicas de fluoretos, recomendações de dieta e instruções de higiene oral acabam não sendo bem aceitas como um tratamento. A diferente remuneração por serviços preventivos e serviços restaurativos também influencia diretamente nessa

tendência, resultando algumas vezes em excesso de tratamentos (BAELUM, 2008; PITTS, 2004b). Uma possível solução para isso, é a prática do tratamento dental baseado em evidências (*evidence-based dental care*), podendo ser definido como: o uso de evidências científicas para a tomada de decisão em relação ao tratamento de cada paciente, focando na prevenção e não somente no tratamento da lesão (BAELUM, 2008; GREENES, 2007; LIMEBACK, 2012; ZERO et al., 2011).

Segundo (GREENES, 2007), a melhor evidência clínica é aquela baseada em informações relevantes de pesquisas válidas e atuais em relação aos cuidados com a saúde e seus benefícios. A tendência internacional no tratamento da cárie é alterar o perfil invasivo para um perfil preventivo e não-invasivo a fim de controlar o início e a progressão do processo da cárie ao longo da vida do paciente e não apenas a lesão causada pelo processo (PITTS, 2004b).

Segundo a OMS (PETERSEN, 2003), os avanços da biomedicina, ciências sociais, saúde pública e da tecnologia da informação devem levar a intervenções inovadoras e impactantes nos aspectos preventivo, terapêutico, diagnóstico, ético e social da saúde humana. CDSS são uma abordagem que contemplam tais necessidades, podem ser usados para incorporar melhores práticas, na esperança de que seus usuários possam melhorar a qualidade dos seus serviços para a saúde, reduzindo a variação da decisão entre os profissionais e aumentando o rendimento e o fluxo de trabalho (GREENES, 2007). A utilização de um CDSS juntamente com a prática baseada em evidências reúne dois métodos potencialmente positivos para melhorar a qualidade dos cuidados a saúde (SIM et al., 2001).

A gestão de fatores de risco individuais, detecção de lesões em fase inicial e o uso de técnicas minimamente invasivas (PITTS, 2004b) podem levar aos melhores resultados no tratamento e o controle do processo dinâmico da cárie durante toda a vida do paciente. Na área de tratamento odontológico encontram-se poucas referências para CDSS:

- Sistema para auxílio a decisão no tratamento de dentes quebrados utilizando um ES baseado em lógica fuzzy (MAGO et al., 2012);

- Modelo de suporte a tomada de decisão para tratamento endodôntico utilizando uma rede Bayesiana (SUEBNUKARN; RUNGCHAROENPORN; SANGSURATHAM, 2008);
- Comparação entre um sistema comercial e estudantes de odontologia juntamente com os instrutores em relação a restauração ou substituição de dentes fracos ou faltantes. Os resultados são relacionados ao custo financeiro dos tratamentos recomendados com e sem o sistema (KAWAHATA; MACENTEE, 2002).

Apenas um trabalho para o tratamento do processo da cárie foi encontrado durante o desenvolvimento deste projeto (BHATIA; SINGH, 2013). Tal trabalho utiliza redes Bayesianas para a modelagem do sistema de suporte a tomada de decisão. É um sistema baseado apenas nos sintomas, como dor, cavitação, gengivite e outros. O tratamento é voltado para aplicação de medicamentos e ações invasivas. Os resultados exibidos no trabalho não são conclusivos.

Na área de detecção e monitoramento de lesões cariosas, existem vários sistemas que utilizam diferentes princípios físicos para a aquisição de informações sobre o processo de desmineralização dental. Dentre elas, destacam-se: raio-X (KUANG; YE, 2008; LI et al., 2007), luz visível (OLSEN et al., 2009), fluorescência dental (FEJERSKOV; KIDD, 2008), laser, corrente elétrica e ultrassom (PRETTY, 2006), juntamente com diversas técnicas computacionais para a detecção.

Um consenso internacional (PITTS; STAMM, 2004) recomenda que o diagnóstico da cárie seja realizado através de uma avaliação abrangente das informações necessárias do paciente pelo dentista e que seja diferenciado da detecção e do monitoramento da lesão.

Tal contexto de causalidade, evidências e incertezas que envolvem o processo dinâmico da cárie, juntamente com a variação nos tratamentos realizados pelos profissionais e os altos índices da doença, mostram a necessidade de abordagens inovadoras para o gerenciamento da cárie.

1.3. Organização do trabalho

Devido a interdisciplinaridade do trabalho e para uma correta compreensão dos materiais e métodos utilizados, serão apresentados de forma rápida no Capítulo 2 os fundamentos sobre redes Bayesianas, referente à primeira abordagem (gerenciamento da cárie), cárie dental, imagens de fluorescência e fluorescência dental, referentes a segunda abordagem (detecção da cárie).

No Capítulo 3, serão descritos os materiais e métodos utilizados para a o desenvolvimento do trabalho. Este capítulo está dividido em duas partes:

Gerenciamento da Cárie:

- Análise das variáveis;
- Estrutura da BN;
- Probabilidades condicionais da BN.

Detecção da Cárie:

- Imagens de fluorescência;
- Extração de características de imagens;
- Análise de Componentes Principais;
- Agrupamento.

Em seguida, o Capítulo 4 apresentará os resultados e a discussão de cada parte do trabalho, juntamente com os trabalhos publicados ou submetidos, seguido da conclusão e próximos passos no Capítulo 5, e por último, as referências bibliográficas.

2. FUNDAMENTOS

2.1. Redes Bayesianas

Redes Bayesianas oferecem uma teoria formal, natural e intuitiva para lidar com incertezas e representar conhecimento específico de cenários complexos utilizando uma representação gráfica (KJÆRULFF; MADSEN, 2008), incluindo cenários de decisões (BERNER, 2007; VELIKOVA et al., 2013).

BN são utilizadas em diferentes áreas como, por exemplo, modelagem de comportamento social (WALSH et al., 2010), agricultura (KRISTENSEN; RASMUSSEN, 2002), processos químicos (WEIDL; MADSEN; ISRAELSON, 2005), oncologia (JULIA FLORES et al., 2011) e outras áreas da saúde (BERNER, 2007). Antes de uma definição formal de redes Bayesianas alguns conceitos serão brevemente apresentados.

É importante que o leitor possua uma familiaridade com a teoria da probabilidade, como variáveis aleatórias e probabilidade condicional por exemplo. Recomenda-se o livro texto “Probabilidade: Aplicações à Estatística” (MEYER, 1970).

2.1.1. Teorema de Bayes

Suponha duas variáveis aleatórias X e Y pertencentes ao mesmo espaço amostral S , a probabilidade das duas ocorrerem simultaneamente é dada da seguinte forma:

$$P(X \cap Y) = P(X|Y) P(Y) = P(Y|X) P(X).$$

A partir desta igualdade naturalmente surge o Teorema de Bayes (DUDA; HART; STORK, 2001):

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}.$$

O Teorema de Bayes mostra que observando um evento X causal de Y , pode-se utilizar a probabilidade *a priori* de um evento Y ocorrer ($P(Y)$) para

analisarmos a probabilidade *a posteriori* ($P(Y|X)$), ou seja, a probabilidade de uma hipótese dada uma evidência. A verossimilhança ($P(X|Y)$), indica a probabilidade desta evidência dada que a hipótese é verdadeira.

Exemplo: Considere que um estoque é abastecido com peças provenientes de três fábricas distintas: F_1 , F_2 e F_3 . Suponha que uma peça retirada do estoque seja defeituosa (D), qual a probabilidade de que tenha sido produzida na Fábrica 1?

Sabe-se que 2% das peças produzidas em F_1 e F_2 são defeituosas, enquanto 4% daquelas produzidas em F_3 são defeituosas. Sabe-se também que F_1 produz o dobro de peças que F_2 e F_3 . Tal cenário é descrito pelo diagrama de Venn da Figura 4.

Figura 4 – Diagrama de Veen para o cenário descrito no exemplo de aplicação do teorema de Bayes.

Pede-se $P(F_1|D)$, aplicando o teorema de Bayes pode-se escrever:

$$P(F_1|D) = \frac{P(D|F_1) P(F_1)}{P(D)}.$$

Sabe-se que:

- $P(F_1) = 0.5$;
- $P(D|F_1) = 0.04$;
- $P(D) = P(D \cap F_1) + P(D \cap F_2) + P(D \cap F_3)$;

Aplicando a definição de probabilidade condicional, tem-se que:

$$P(D \cap F_n) = P(D|F_n)P(F_n).$$

Portanto:

$$P(D) = \sum_{n=1}^3 P(D|F_n)P(F_n).$$

Aplicando-se o teorema de Bayes obtêm-se:

$$P(D|F_1) = \frac{P(F_1|D)P(F_1)}{\sum_{n=1}^3 P(D|F_n)P(F_n)} = \frac{(0.02)(0.5)}{(0.02)(0.5) + (0.02)(0.25) + (0.04)(0.25)} = 0.4.$$

Observa-se no exemplo o uso de conhecimento *a priori* para inferir um novo conhecimento *a posteriori* utilizando a evidência de que a peça é defeituosa juntamente com a verossimilhança de tal evento. Pode-se expressar o teorema de Bayes de uma maneira informal (DUDA; HART; STORK, 2001):

$$Posteriori = \frac{Verossimilhança \times Priori}{Evidência}.$$

O teorema de Bayes pode ser interpretado como uma análise causal, na qual o conhecimento a respeito da relação de dependência entre dois eventos, juntamente com uma evidência a respeito do evento que causa o outro, nos leva a uma crença nas consequências de tal evidência.

2.1.2. Redes Probabilísticas

Redes probabilísticas são representações gráficas de relações de dependência entre variáveis aleatórias. Estas redes representam e processam conhecimento probabilístico (KJÆRULFF; MADSEN, 2008), podendo ser divididas em duas partes.

A parte qualitativa é constituída de um conjunto de variáveis aleatórias e um conjunto de relações de dependência ou independência entre as variáveis, as relações de dependência, precedência informacional e relações de preferência entre as variáveis são representadas através de grafos. A parte quantitativa representa a força da dependência entre as variáveis utilizando teoria da probabilidade (KJÆRULFF; MADSEN, 2008).

A construção de uma rede probabilística também pode ser dividida em duas etapas distintas, a modelagem qualitativa e a modelagem quantitativa. A parte qualitativa define as relações de dependência, a parte quantitativa as probabilidades condicionais entre as variáveis aleatórias.

As variáveis aleatórias representam eventos ou proposições, podendo ser discretas ou contínuas. As redes probabilísticas podem ser utilizadas para analisar a forma como a certeza a respeito de uma variável pode influenciar na incerteza a respeito de uma outra variável (causalidade) (JENSEN; NIELSEN, 2007; KJÆRULFF; MADSEN, 2008). Uma variável que influencia uma outra é chamada de nó pai e a variável que é influenciada é o nó filho.

Existem diferentes formas de um conjunto de variáveis se relacionarem entre si em uma rede probabilística. A seguir mostra-se uma apresentação dos tipos de conexões em uma rede probabilística.

a. Conexão Serial

Quando uma variável A influencia diretamente uma variável B que, por sua vez, influencia diretamente uma variável C, tem-se uma conexão serial, Figura 5 (JENSEN; NIELSEN, 2007). Obviamente evidências em A também influenciam em C através de B e vice-versa. Caso o nó B seja evidenciado¹, A e C tornam-se independentes, ou seja, alterações na certeza de A não influenciam a certeza em relação a C.

Figura 5 – Exemplo de uma conexão serial entre variáveis de uma rede probabilística, evidências em A influenciam B, que por sua vez influencia C.

¹ Um nó é evidenciado quando um dos estados é definido como verdadeiro – 100% de probabilidade – enquanto os outros são definidos como falsos – 0% de probabilidade.

b. Conexão Divergente

Quando uma variável A influencia diretamente mais de um nó filho, tem-se uma conexão divergente (JENSEN; NIELSEN, 2007), Figura 6.

Figura 6 – Exemplo de uma conexão divergente entre variáveis de uma rede probabilística, evidências em A influenciam diretamente as variáveis B, C e D.

c. Conexão Convergente

Se uma variável A é influenciada por mais de um nó pai, tem-se uma conexão convergente (JENSEN; NIELSEN, 2007), Figura 7. Se não há evidências a respeito de A, podem ser utilizadas informações sobre os nós pais para inferir a respeito dos possíveis estados de A. Caso o nó filho A e algum dos nós pais sejam evidenciados, pode-se inferir informações a respeito dos estados dos outros nós pais.

Figura 7 – Exemplo de uma conexão convergente entre variáveis de uma rede probabilística. A variável A é influenciada pelas variáveis B, C e D e evidências em qualquer um dos nós pais influenciam o nó filho A.

Em todas as configurações apresentadas é possível observar a relação de causa e consequência entre as variáveis, assim como as possíveis relações de independência entre os nós nas diferentes configurações.

2.1.3. D-separação

Utilizando o que já foi apresentado através dos tipos de conexões possíveis em redes probabilísticas, pode-se formalizar o conceito de d-separação (JENSEN; NIELSEN, 2007). Duas variáveis distintas, A e B, em uma rede probabilística são ditas d-separáveis se existir uma variável V intermediária entre elas, de tal forma que:

- A conexão é serial ou divergente e V pode ser evidenciada;
- A conexão é convergente e nem V e nenhum de seus descendentes (nós filhos) podem ser evidenciadas.

Em uma rede probabilística, se A e B são d-separados, podem-se afirmar que certezas em A não causam impactos na certeza de B, as duas variáveis são condicionalmente independentes na rede. A Figura 8 a seguir exemplifica um conjunto de variáveis que são d-separáveis pela variável V.

Figura 8 – Exemplo de variáveis d-separáveis em um rede probabilística. Considerando que o nó V pode ser evidenciado, qualquer certeza em V torna os seus nós filhos condicionalmente independentes dos seus nós pais.

2.1.4. Redes Bayesianas

Redes Bayesianas são uma subclasse das redes probabilísticas (KJÆRULFF; MADSEN, 2008), podem ser definidas formalmente como (JENSEN; NIELSEN, 2007):

- Um conjunto de variáveis (nós) e um conjunto de arestas entre estas variáveis;
- Cada variável possui um conjunto finito de estados mutualmente exclusivos;
- As variáveis juntamente com as arestas formam um grafo orientado acíclico (DAG, do inglês *directed acyclic graph*).

Redes Bayesianas podem representar relações causais através das dependências entre as variáveis (JENSEN; NIELSEN, 2007). As arestas conectam nós pais (Pr) aos nós filhos (C), sendo cada nó filho diretamente dependente dos seus nós pais. Cada nó filho possui uma tabela de probabilidade $P[C_1|Pr_1, \dots, Pr_n]$ quantificando a dependência entre os nós conectados (PEARL, 1988). No caso em que um nó não possui nós pais a tabela se reduz a tabela de probabilidade incondicional $P[A]$.

Os nós em uma rede podem ser evidenciados, consequentemente alteram a distribuição de probabilidade no restante das variáveis da BN. Assumido independência condicional entre nós pais e utilizando a regra da cadeia (JENSEN; NIELSEN, 2007), a probabilidade condicional de um nó filho em relação aos seus nós pais ($P(C_1|Pr_1, \dots, Pr_n)$) pode ser rescrito como:

$$P(C_1|Pr_1, \dots, Pr_n) = \prod_{k=1}^n P(C_1|Pr_k).$$

A probabilidade de um evento de interesse se reduz ao produto das probabilidades condicionais do evento com seus nós pais. A suposição de independência entre os nós pais reduz drasticamente a complexidade da tabela de probabilidade condicional através da redução do número de termos necessários

para os cálculos dos parâmetros numéricos da rede. Essa suposição é irrealista em alguns casos, porém funciona bem na prática (WITTEN; FRANK, 2011).

Exemplo²: Câncer metastático é uma das possíveis causas de tumor cerebral e também é uma explicação para o aumento dos valores de séricos de cálcio. Tanto o câncer, quanto o aumento do sérico cálcio, podem levar a um coma de um paciente, o tumor cerebral também pode causar dores de cabeça severas. A rede Bayesiana da Figura 9 representa as relações descritas anteriormente.

Figura 9 – Rede Bayesiana representando o problema exemplo a respeito do Câncer Metastático. As relações de causalidade entre as variáveis envolvidas são explicitadas através da orientação das arestas na rede.

A tabela 1 a seguir contém as probabilidades condicionais que caracterizam as relações de causalidade, ou dependência, apresentadas no parágrafo anterior. É importante observar que cada variável depende unicamente dos seus nós pais, e que a variável “Câncer Metastático” possui apenas uma distribuição *a priori*.

² Exemplo fictício adaptado de (PEARL, 1988) página 196.

Tabela 1 – Probabilidades condicionais para as variáveis relacionadas na Figura 9. Observa-se que as probabilidades condicionais de cada variável são apenas dependentes das variáveis predecessoras, ou nós pais.

Variáveis		Probabilidades condicionais	
Câncer		$P(+C) = 0.20$	
Tumor		$P(+Tumor/+Câncer) = 0.80$	$P(+Tumor/-Câncer) = 0.2$
Séricos		$P(+Sérico/+Câncer) = 0.20$	$P(+Sérico/-Câncer) = 0.05$
Coma		$P(+Coma/+Tumor, +Séricos) = 0.80$ $P(+Coma/-Tumor, +Séricos) = 0.80$	$P(+Coma/+Tumor, -Séricos) = 0.80$ $P(+Coma/-Tumor, -Séricos) = 0.05$
Dor de Cabeça		$P(+Dor/+Tumor) = 0.80$	$P(+Dor/-Tumor) = 0.80$

A partir destas probabilidades condicionais, pode-se calcular a probabilidade *a posteriori* dos diferentes estados que o sistema pode assumir. Tal tarefa é conhecida como inferência Bayesiana (PEARL, 1988).

2.2. Cárie Dental

A cárie dental é uma consequência do processo dinâmico e multifatorial que acontece no meio bucal (KIDD, 2005; SELWITZ; ISMAIL; PITTS, 2007). Quando os microorganismos presentes na boca metabolizam açúcares, são gerados ácidos que diminuem o pH³ do meio bucal e consequentemente ocorrem trocas iônicas entre o esmalte dentário e o meio para neutralização deste pH, o esmalte perde minerais para o meio.

Após o processo de neutralização do pH do meio bucal também ocorre o processo de remineralização do esmalte dentário. Este processo ocorre diversas

³ pH é a medida do potencial hidrogeniônico que indica a acidez em uma solução. Quanto menor o valor de pH, mais ácida a solução.

vezes durante um dia e o seu resultado acumulativo pode levar ao surgimento de lesões cáries e a progressão de lesões já existentes (FEJERSKOV; KIDD, 2008).

O processo da cárie possui diferentes fatores que influenciam na desmineralização, como a alimentação, higiene bucal, função salivar e exposição a fluoretos (KIDD, 2005; LIMEBACK, 2012). Atualmente, a cárie progride devagar na maioria das pessoas, podendo afetar diferentes partes do dente, como a coroa, raiz e dentina. O termo cárie dental pode ser utilizado para identificar tanto o processo de formação da cárie, como a lesão que é resultado deste processo (KIDD, 2005; SELWITZ; ISMAIL; PITTS, 2007).

Entre as crianças, a cárie é uma das doenças mais comuns (BOYCE et al., 2010; SELWITZ; ISMAIL; PITTS, 2007). Atualmente conhecida como cárie precoce na infância (ECC, do inglês *Early Childhood Caries*), engloba qualquer cárie que ocorra na dentição primária da criança entre o nascimento e 71 meses de idade (FILSTRUP et al., 2003). A ECC está fortemente associada com comunidades de baixa renda (CASAMASSIMO et al., 2009), assim como a cárie de um modo geral (COSTA et al., 2012).

Além das facetas sociais e econômicas, existem evidências das consequências físicas e psicológicas relacionadas a ECC. A dor de dente acaba influenciando o peso, o crescimento, a alimentação, o sono, atividades físicas, interações sociais e a autoestima (CASAMASSIMO et al., 2009; FILSTRUP et al., 2003; SHEIHAM, 2006). Uma investigação relacionando indicadores de qualidade de vida e consequências da ECC realizado diretamente com crianças (FILSTRUP et al., 2003) mostrou que esses fatores são perceptíveis pela própria criança.

Resultados recentes mostram que as doenças crônicas em adultos normalmente são relacionáveis com as condições e exposições no início da vida (LIMEBACK, 2012; SHONKOFF; BOYCE; MCEWEN, 2009). Crianças com ECC tendem a ter mais cáries no decorrer dos anos, reforçando a necessidade de se dar atenção especial a esse problema de saúde pública, a ECC e a cárie em geral (LIMEBACK, 2012).

Segundo (PITTS, 2004b), atualmente é bem estabelecido sobre a cárie que o seu processo é reversível se no estágio inicial, é uma doença crônica e de etiologia multifatorial. Sabe-se também que os diferentes estágios do processo da cárie a nível clínico já foram bem documentados no decorrer dos anos. Além disso, um diagnóstico no início da cárie é fundamental para a adoção de tratamentos menos invasivos ou até mesmo não invasivos (DE JOSSELIN DE JONG et al., 2009; PENEVA, 2008).

2.3. Imagens de Fluorescência

Fluorescência é o processo de emissão de luz por alguma substância exposta a radiação eletromagnética, normalmente na faixa do visível ou do espectro ultravioleta (LAKOWICZ, 2006; MURPHY; DAVIDSON, 2001). Este processo tem duração de alguns nanossegundos sendo diretamente dependente da organização molecular da substância fluorescente e da fonte de excitação (LAKOWICZ, 2006).

Ao absorver radiação, alguns elétrons do material fluorescente passam do estado fundamental para um estado mais energético, os estados excitados. Em seguida, os elétrons perdem esta energia de algumas formas, sendo uma delas através da emissão de radiação eletromagnética (sinal de fluorescência), retornando ao estado fundamental (LAKOWICZ, 2006; MURPHY; DAVIDSON, 2001). O sinal de fluorescência liberado normalmente possui comprimento de onda maior que o absorvido e, conseqüentemente, possui menos energia.

A intensidade do sinal de fluorescência pode ser atenuada por diversos fatores (LAKOWICZ, 2006) como, por exemplo, diminuição da potência do sinal de excitação. O sinal de fluorescência é diretamente dependente da estrutura molecular do material. Esse fator permite a caracterização de tecidos biológicos devido a organização molecular distinta de cada um deles (LAKOWICZ, 2006; MURPHY; DAVIDSON, 2001), podendo também diferenciar tecidos saudáveis e patológicos.

Uma imagem é uma representação matricial onde cada ponto da matriz é referente a um *pixel* da imagem, cada *pixel* possui informações referente à intensidade luminosa e as cores naquele ponto de acordo com o sistema de cores

que o sensor utiliza para digitalizar tais informações (GONZALEZ; WOODS, 2008). Uma imagem de fluorescência nada mais é que uma matriz que representa a intensidade do sinal e os diferentes comprimentos de onda que o compõe representados em um determinado sistemas de cores.

Nas últimas décadas ocorreu um crescimento no uso da fluorescência nas ciências biológicas (LAKOWICZ, 2006; MURPHY; DAVIDSON, 2001; VO-DINH, 2003). Essa metodologia tem sido aplicada na biotecnologia, citometria de fluxo, diagnósticos médicos, análise genética e práticas forenses.

2.4. Fluorescência Dental

No caso da fluorescência dental (DE JOSSELIN DE JONG et al., 2009; VO-DINH, 2003), quando um dente é excitado com uma radiação na faixa do ultravioleta (UV) ele emite um sinal de fluorescência em todo o espectro visível. Com tal excitação as regiões saudáveis do esmalte dentário emitem um sinal de fluorescência de maior intensidade que as regiões desmineralizadas em uma certa banda de comprimentos de onda, conforme mostrado na Figura 10.

Figura 10 – Espectro de emissão do sinal de fluorescência para dentes saudáveis e cariados utilizando uma fonte de excitação em 405 nm. Observa-se a redução na intensidade do sinal de fluorescência para amostras cariadas em relação a amostras saudáveis entre 450 nm e 600 nm. Adaptado de (ZEZELL et al., 2007).

Outro fator importante é que microorganismos na superfície dental emitem um sinal de fluorescência no comprimento de onda do vermelho para a mesma fonte de excitação (VO-DINH, 2003). A fluorescência dental está exemplificada na Figura 11.

Figura 11 – Exemplo de uma imagem de fluorescência dental obtida em bancada utilizando uma amostra in vitro. (a) Imagem realizada utilizando uma iluminação branca, (b) imagem de fluorescência obtida em bancada utilizando fonte de excitação UV (405 nm) e instrumentação óptica, a seta vermelha indica região com sinal de fluorescência dos microorganismos, a seta preta mostra uma região com sinal de fluorescência de menor intensidade que o resto da superfície visualizada.

Este é um método bem estabelecido para à detecção da cárie, uma vez que é uma técnica não-invasiva e com capacidade de fornecer informação coerente e em tempo real sobre o estado de desmineralização do dente (VO-DINH, 2003). Alguns equipamentos para a odontologia clínica já são comercializados utilizando tal técnica, como o QLF (ANGMAR-MÅNSSON; TEN BOSCH, 2001; DE JOSSELIN DE JONG et al., 2009), e o DIAGNOdent (PRETTY, 2006).

O sistema QLF (DE JOSSELIN DE JONG et al., 2009; FEJERSKOV; KIDD, 2008) fabricado pela Inspektor, consiste na aquisição de uma imagem de fluorescência na região do amarelo (acima de 520 nm) utilizando uma fonte de excitação em 405 nm juntamente com um filtro passa-alta em frente a uma câmera intraoral. A imagem de fluorescência adquirida é processada através de um *software* semiautomático (ANGMAR-MÅNSSON; TEN BOSCH, 2001; INSPEKTOR, [s.d.]) para quantificar informações como: tamanho, profundidade e volume das lesões.

O processamento do sistema QLF é realizado através de iterações entre o usuário e o *software*: o usuário seleciona alguns pontos de referência na primeira

imagem adquirida, em seguida estes pontos serão utilizados para interpolação com às próximas imagens adquiridas permitindo a medição de todas as informações a respeito da lesão. Maiores informações a respeito do processamento e da interação entre o sistema e o usuário podem ser encontrados em (ANGMAR-MÅNSSON; TEN BOSCH, 2001).

O equipamento DIAGNOdent (KAVO, [s.d.]), utiliza um laser no vermelho (655 nm) para excitar o esmalte dentário e produzir um sinal de fluorescência na região excitada (FEJERSKOV; KIDD, 2008). O sinal é coletado utilizando uma fibra ótica e o sistema apresenta ao usuário a intensidade do sinal naquela região juntamente com o valor de pico coletado durante o exame (PRETTY, 2006). O usuário utiliza estas medidas para realizar o exame das regiões do dente e em seguida inferir um diagnóstico.

Ambos os sistemas, assim como outros métodos de detecção da cárie, necessitam que o dente examinado esteja limpo e seco (PRETTY, 2006). Além disso, a obtenção de resultados coerentes é diretamente dependente da habilidade do operador com tais equipamentos (ANGMAR-MÅNSSON; TEN BOSCH, 2001).

3. MATERIAIS E MÉTODOS

Os Materiais e Métodos deste trabalho estão divididos em duas partes, sendo referentes às abordagens já mencionadas, gerenciamento e detecção da cárie. O gerenciamento consiste na modelagem probabilística de um sistema que considera fatores causais juntamente com evidências clínicas para inferir a respeito do risco de cárie, tratamentos e retorno do paciente. A detecção consiste na modelagem de imagens de fluorescência dental para uso em um sistema para auxílio na detecção da cárie.

3.1. Gerenciamento da Cárie:

A metodologia desenvolvida para a avaliação de um sistema para auxílio ao gerenciamento da cárie consiste na análise da utilização de redes probabilísticas para relacionar os fatores causais e evidências clínicas referentes ao risco de cárie, tratamento e previsão de retorno do paciente. Para tal abordagem utilizamos BN como sistema de inferência, para a modelagem da rede foi realizada uma revisão bibliográfica e conversas com especialistas da área.

A rede construída neste trabalho para o gerenciamento da cárie utiliza o conceito de gerenciamento moderno da cárie (LIMEBACK, 2012; ZERO et al., 2011). Em outras palavras, considera os fatores individuais de risco e evidências científicas para sugerir o tratamento mais apropriado para o paciente.

Para a construção e modelagem da BN foi utilizado o *software* GeNIe 2.0 (DECISION SYSTEMS LABORATORY). O desenvolvimento da rede foi dividido em três partes: escolha das variáveis, definição da estrutura da BN (parte qualitativa) e construção das tabelas de probabilidades condicionais entre as variáveis (parte quantitativa).

3.1.1. Escolha das Variáveis da Rede Bayesiana

A partir de uma revisão de artigos científicos da área de cariologia, foram definidas as variáveis que estariam presentes na BN, estas podem ser divididas em três partes: fatores causais, evidências clínicas e tratamentos e retorno.

Fatores Causais

- Cáries nos últimos 3 anos: o trabalho de (BADER et al., 2005) mostrou que o histórico de incidência de cáries é um importante indicador de risco de cárie;
- Função Salivar: a saliva atua de diferentes formas no meio bucal, auxilia na limpeza e na manutenção do pH (LIMEBACK, 2012). A baixa produção salivar (xerostomia) pode ser consequência de diversos fatores (LIMEBACK, 2012), como radioterapia na região da cabeça e pescoço, diabetes e mal de Parkinson, que possuem alta incidência na população idosa;
- Exposição à Fluoretos: o uso de fluoretos, tanto na água consumida como na utilizada para produção de alimentos, como nos cremes dentais, foi relacionado ao declínio da cárie no mundo inteiro (LIMEBACK, 2012). Por exemplo, uma pessoa que vive em uma região onde não ocorre fluoretação da água e não utiliza creme dental com flúor provavelmente terá um alto risco de cáries;
- Dieta: o consumo de açúcares e carboidratos é considerado um importante fator etiológico para o desenvolvimento da cárie, uma vez que estão ligados diretamente a diminuição do pH do meio bucal que desencadeia o processo da cárie (ZERO, 2004). A nível de indivíduo, o aconselhamento nutricional é altamente recomendável a pessoas com cáries ativas ou em grupos de risco elevado.
- Higiene oral: a falta de higiene oral é um fator de risco para doenças dentais (LIMEBACK, 2012), como a cárie e a periodontite, apesar de algumas pessoas conseguirem manter a saúde bucal com um nível mínimo de higiene bucal diária, é uma das melhores formas de promover a saúde bucal.

Evidências Clínicas

- Superfície Lesionada e Tipo de Dente: o trabalho de (FERREIRA ZANDONÁ et al., 2012) mostrou que cada tipo de dente e superfície possui uma susceptibilidade diferente a lesões cariosas;

- Classificação ICDAS: ICDAS (do inglês, *International Caries Detection and Assessment System*, Sistema internacional de detecção e avaliação de cáries), classifica as lesões de acordo com sua aparência visual. É um sistema padrão para detecção e classificação de cáries em estágios tanto iniciais quanto avançados (ZERO et al., 2011). O conceito é padronizar as informações relacionadas à cárie com uma melhor qualidade, influenciando de forma mais apropriada as decisões relacionadas ao diagnóstico, prognóstico e gerenciamento clínico da cárie, tanto a nível individual quanto ao nível de saúde pública (ICDAS Foundation). Para maiores informações sobre o sistema ICDAS, recomenda-se que acesse o endereço eletrônico de (ICDAS Foundation). Para uma compreensão mais profunda entre os diferentes estágios do processo da cárie e as diferentes classificações ICDAS, recomenda-se o trabalho de (ZERO et al., 2011);
- Atividade da Lesão: a atividade de uma lesão pode ser ativa ou inativa, a última pode estar “estacionada” ou em regressão. Lesões ativas não cavitadas possuem uma chance muito maior de progredir para uma cavitação do que para uma lesão inativa (NYVAD; MACHIULSKIENE; BAELUM, 2003). A tendência de avanço ou regressão da lesão está diretamente ligada à etiologia do processo da cárie no meio bucal (avaliação de risco de cárie). A importância de analisar a atividade da lesão está diretamente ligada a escolha do melhor tratamento (NYVAD; MACHIULSKIENE; BAELUM, 2003; NYVAD, 2004). Um critério com validação clínica para classificação foi elaborado (NYVAD; MACHIULSKIENE; BAELUM, 2003), essencialmente, lesões esbranquiçadas e rugosas são ativas, enquanto lesões lisas, brilhantes e rígidas são inativas.

Tratamentos e Retorno

- Tratamentos:
 - Fluoreto: a eficácia do uso de fluoretos como tratamento não-invasivo atuando diretamente na estrutura dental já foi mostrada em diferentes estudos (CAROUNANIDY; SATHYANARAYANAN, 2010; SELWITZ;

ISMAIL; PITTS, 2007). Fluoretos possuem a capacidade de auxiliar na remineralização e na resistência do dente ao processo da cárie. A utilização do Fluoreto é dependente da atividade da lesão e do risco de cárie de cada paciente;

- Selante: é uma forma mecânica de separar a região oclusal com lesão de cárie do meio bucal (CAROUNANIDY; SATHYANARAYANAN, 2010), reduzindo significativamente a quantidade de lesões que progridem para estágios mais avançados (cavitações) (BAELUM, 2008; CAROUNANIDY; SATHYANARAYANAN, 2010; FEIGAL et al., 2008);
- Restauração: o objetivo de uma restauração é remover o tecido cariado de forma a preservar a maior parte da estrutura não afetada (CAROUNANIDY; SATHYANARAYANAN, 2010). Uma vez que um dente é restaurado, inicia-se um ciclo restaurativo e, conseqüentemente, uma sequência de tratamentos custosos (FERREIRA ZANDONÁ et al., 2012; SELWITZ; ISMAIL; PITTS, 2007; ZERO et al., 2011). A abordagem moderna de micro restaurações utilizando material adesivos podem auxiliar na preservação da estrutura dental original (CAROUNANIDY; SATHYANARAYANAN, 2010; SELWITZ; ISMAIL; PITTS, 2007);
- Tratamento Endodôntico: Quando a lesão alcança o tecido pulpar, se inicia um outro processo que é relacionado com a área da endodontia (QUALTROUGH et al., 2005), que está fora do escopo deste trabalho.
- Retorno: é necessário que além da abordagem de avaliação do risco de cárie, o tratamento em geral seja influenciado pelo resultado da avaliação, inclusive o tempo de retorno do paciente (BADER et al., 2005). Pacientes de baixo risco podem retornar com um intervalo de um ano. Pacientes em risco médio devem retornar a cada seis meses, e pacientes em risco alto devem retornar mais frequentemente.

3.1.2. Estrutura da Rede Bayesiana

Após a escolha e estudo das variáveis, foi definida a estrutura da Rede Bayesiana, primeiro foi desenvolvida a parte referente aos fatores de risco da cárie,

Figura 12, considerando os fatores causais e preditivos para uma classificação individual de risco de cárie, conforme apresentado na Seção 3.1.1. Todos os estados dos nós pais estão listados na Tabela 2, assim como o grau de influência no risco de cárie.

Figura 12 – Primeira parte da BN desenvolvida, considera os fatores causais e preditivos da cárie para uma sugestão de risco de cárie de acordo com o cenário em que cada paciente se encontra.

Tabela 2 – Estados e grau de influência das variáveis causais e preditivas consideradas para a sugestão do risco de cárie. Grau de influência I é o mais forte e III o mais fraco.

Variáveis	Estados	Grau de influência
Cáries nos últimos 3 anos	Zero; Entre 1 e 2; Mais que 2	I
Fluxo salivar	Normal; Baixo	II
Exposição à Fluoretos	Sim; Não	II
Dieta	Boa; Regular; Ruim	III
Higiene oral	Boa; Regular; Ruim	III

Em seguida desenvolveu-se a parte referente ao tratamento e retorno do paciente. Foram consideradas evidências clínicas e o risco de cárie do paciente conforme as informações apresentadas na Seção 3.1.1, a estrutura está descrita na Figura 13. Todos os estados dos nós pais estão listados na Tabela 3, assim como o grau de influência nas sugestões de tratamento e retorno. Os possíveis estados das variáveis “Tratamento” e “Retorno” estão descritos na Tabela 4.

Figura 13 – Segunda parte da BN desenvolvida, considera a situação clínica e o risco de cárie do paciente para sugestões de tratamento e de previsão de retorno considerando o cenário individual de cada paciente.

Tabela 3 - Estados e grau de influência das variáveis causais e preditivas consideradas para a sugestão do tratamento e retorno do paciente. Grau de influência I é o mais forte e III o mais fraco.

Variável	Estados	Grau de influência
Superfície lesionada	Oclusal; Lisa	III
Risco de cárie	Baixo; Médio; Alto	II
ICDAS	1 & 2; 3; 4; 5 & 6	I
Atividade da lesão	Sim; Não	II
Tipo de dente	Molar; Pré-molar; Anterior	III

Tabela 4 – Descrição dos possíveis estados das duas variáveis de saída da BN: tratamento e retorno.

Variável	Estados
Tratamento	Fluoreto ou Selante; Restauração; Tratamento endodôntico
Retorno	1 ano; 6 meses; Antes de 6 meses

A estrutura final da BN desenvolvida encontra-se na Figura 14, em que estão juntas as duas partes já descritas: risco de cárie, tratamento e retorno. A estrutura final da BN mostra que a rede possui variáveis d-separáveis, a partir do momento em que a variável “Risco de Cárie” é evidenciada os seus nós filhos tornam-se condicionalmente independentes dos seus nós pais.

Figura 14 – Rede Bayesiana desenvolvida para auxílio a tomada de decisão clínica para o gerenciamento da cárie, os nós em branco são as variáveis que o profissional clínico pode evidenciar na Rede, os nós em azul são as variáveis que a Rede fornecerá uma distribuição de probabilidade para os possíveis estados, referentes às decisões. O nó Risco de Cárie pode ser evidenciado pelo profissional clínico caso seja obtida alguma conclusão em relação ao risco do paciente analisado.

3.1.3. Probabilidades Condicionais das Variáveis da Rede Bayesiana

Após a definição da estrutura da BN para gerenciamento da cárie (CMBN, do inglês *Caries Management Bayesian Network*), foi necessário definir as probabilidades condicionais entre os nós da rede. Durante o desenvolvimento deste trabalho não foi possível localizar dados que relacionam todas as variáveis necessárias para a CMBN. Praticamente todos os dados referentes a cárie são descritos utilizando o índice DMFT (do inglês, *Decayed, Missing and Filling Teeth*, Dentes cariados, restaurados ou faltantes) (BROADBENT; THOMSON, 2005; MONSE et al., 2010) que representam índices epidemiológicos.

Outro fator é que os trabalhos que analisam fatores causais, preditivos e tratamentos consideram apenas uma das variáveis por vez (BADER et al., 2005; COSTA et al., 2012; FEIGAL et al., 2008; ZERO, 2004). Portanto, utilizou-se de

informações e conclusões da literatura científica para definir as probabilidades condicionais entre as variáveis da CMBN. Foram consideradas as mesmas referências já apresentadas na Seção 3.1.1.

Para a obtenção das probabilidades condicionais, utilizou-se uma técnica proposta por (KJÆRULFF; MADSEN, 2008) para construção manual de tabelas de probabilidade considerando algumas expressões de certeza. O valor adotado para cada expressão está expresso na Figura 15.

Figura 15 – Relação entre expressões de certeza e probabilidades para eventos utilizada para obtenção das probabilidades condicionais entre as diferentes variáveis da CMBN proposta. Adaptado de (KJÆRULFF; MADSEN, 2008).

A metodologia utilizada para a obtenção das probabilidades condicionais foi a transformação das conclusões e informações da literatura em probabilidades utilizando a relação da Figura 13. Em seguida, os casos sem probabilidades e entre duas conclusões consecutivas tiveram suas probabilidades estimadas considerando o grau de influência de cada variável. Esta nova probabilidade foi utilizada para a estimativa de outras probabilidades ainda faltantes na tabela de probabilidade condicional, este processo foi repetido até a tabela estar completamente preenchida. Situações que não se encontravam entre conclusões receberam a probabilidade da conclusão mais próxima.

3.2. Detecção da Cárie

A metodologia desenvolvida para a avaliação de um sistema para auxílio na detecção da cárie consiste na análise da eficácia de características de textura como descritores de dentes cariados e não cariados a partir de imagens de fluorescência dental, as técnicas utilizadas estão descritas nos subcapítulos a seguir.

3.2.1. Conjunto de Dados

O conjunto de dados consiste de 64 imagens de fluorescência dental obtidas a partir de amostras *ex-vivo*, todas previamente classificadas por um especialista em dois grupos: cariadas e não cariadas. O conjunto de dados foi fornecido pela empresa BIOPDI – Indústria de Equipamentos Médicos e Odontológicos Ltda.

O conjunto de dados foi dividido em 26 amostras não cariadas e 38 amostras cariadas, esta última possui diferentes tipos de lesões. Algumas amostras são apresentadas na Figura 16 para ilustrar a diversidade de lesões e superfícies presente no conjunto de dados. As imagens foram adquiridas após limpeza e secagem das amostras.

Figura 16 – Amostras exemplificando a diversidade nos estágios das lesões e superfícies dentais presentes no conjunto de dados utilizado.

3.2.2. Imagens de fluorescência

A aquisição das imagens de fluorescência foi realizada utilizando a configuração de epifluorescência (LAKOWICZ, 2006), a montagem foi realizada em bancada como mostrado na Figura 17.

Figura 17 – Montagem óptica realizada em bancada para a aquisição das imagens de fluorescência das amostras *in vitro*. O círculo azul indica o LED UV de excitação, a caixa verde os filtros ópticos e a caixa vermelha a câmera USB e a lente de zoom manual. A seta amarela indica a posição da amostra para aquisição das imagens.

Um LED ultravioleta (405 nm) foi utilizado juntamente com uma lente convergente como fonte de excitação das amostras. Foram utilizados três filtros ópticos⁴ para correta formação das imagens e uma câmera USB acoplada em uma lente de zoom manual para a aquisição digital das imagens. A Figura 18 mostra o alinhamento necessário entre os componentes do sistema de aquisição.

Figura 18 – Sistema de aquisição de imagens de fluorescência utilizando configuração epifluorescência. As setas azuis indicam o sinal UV, as setas verdes o sinal de fluorescência que serão adquiridos pela câmera USB, a lente amarela representa um filtro passa longa, a lente roxa um filtro passa banda. Estes componentes precisam ser utilizados com este alinhamento óptico para a correta aquisição das imagens.

⁴ Filtros ópticos são discriminados de acordo com o comprimento de onda, diferentemente dos filtros utilizados na área de Engenharia Elétrica por exemplo, onde são caracterizados de acordo com a atuação em frequência.

Da esquerda para a direita tem-se o primeiro filtro óptico, em amarelo, que é um filtro passa alta com 475nm de comprimento de onda de corte, ele transmite apenas sinais com comprimento de onda maior que este valor, este filtro garante a aquisição apenas do sinal de fluorescência dental. O segundo é um filtro passa banda, em roxo, com banda de passagem entre 400 nm e 410 nm garantindo apenas excitação com a fonte UV. O terceiro filtro é um espelho dicróico, com comprimento de onda de transmissão de 470 nm, comprimentos de onda acima deste valor são transmitidos e comprimentos de onda abaixo deste valor são refletidos, este filtro garante a separação dos sinais de excitação e de fluorescência. O espectro de transmissão dos filtros é apresentado na Figura 19.

Figura 19 – Espectro de transmissão dos filtros ópticos utilizados na montagem em bancada para aquisição das imagens de fluorescência. O gráfico foi gerado utilizando as informações fornecidas pelos fabricantes dos filtros.

3.2.3. Extração de características de imagens

Através de manipulação de imagens digitais, o processamento digital de imagens busca extrair informações destas (GONZALEZ; WOODS, 2008), seja para interpretação humana ou computacional. Uma imagem digital é representada através de uma matriz de *pixels* e o valor de cada pixel é relacionado a intensidade do brilho (NIXON; AGUADO, 2008). Para imagens coloridas é necessário definir um espaço de cores, que é um subespaço e um sistema de coordenadas onde cada cor é representada por um ponto (GONZALEZ; WOODS, 2008), na Figura 20 é apresentado uma representação do espaço de cores RGB.

Figura 20 – Espaço de cores RGB, cada eixo representa a intensidade de brilho de uma cor, combinações de intensidades do vermelho, verde e azul podem produzir diferentes cores, adaptado de (GONZALEZ; WOODS, 2008).

Neste trabalho foram utilizados os espaços de cores RGB e YCbCr (GONZALEZ; WOODS, 2008) para a representação das imagens de fluorescência. O espaço RGB foi escolhido pela separação dos sinais, o verde referente ao sinal da fluorescência dental e o vermelho representando a fluorescência dos microorganismos patogênicos. O espaço YCbCr representa o brilho no eixo Y e as cores são representadas através dos eixos Cb e Cr, assim o sinal de fluorescência estará misturado entre os três componentes de cores.

A abordagem de extração de características adotada foi a de características de textura. Existem diferentes definições para textura (NIXON; AGUADO, 2008), a definição adotada no trabalho foi a seguinte: textura é a mudança de cores e intensidade de brilho em uma parte menor que o tamanho do objeto de interesse (PETROU; SEVILLA, 2006), que no caso é a face do dente fotografada.

Uma das formas de analisar variações de textura é através de descritores estatísticos (NIXON; AGUADO, 2008). Antes de falarmos sobre os descritores utilizados iremos primeiro definir momento: momentos são valores utilizados para caracterizar distribuições de probabilidades de variáveis aleatórias (EVERITT, 2002). Para a obtenção dos momentos de cada imagem utilizamos os valores dos histogramas normalizados de cada componente de cor.

Pode-se definir o *n*-ésimo momento como (GONZALEZ; WOODS, 2008):

$$\mu_n(x) = \sum_{i=0}^{N-1} (x_i - m)^n c(x_i),$$

Em que N é o número de *bins* do histograma, x_i é o i -ésimo *bin*, c é o valor de cada *bin* e m é o valor médio de todos os *bins*, definido como:

$$m = \sum_{i=0}^{N-1} x_i c(x_i).$$

Os momentos utilizados como descritores das amostras foram: variância, *skewness* e *kurtosis*. A variância, ou segundo momento (μ_2), mede a forma como os *pixels* estão distribuídos ao redor da média m . *Skewness*, ou terceiro momento, (μ_3) mede a assimetria do histograma (GONZALEZ; WOODS, 2008). *Kurtosis* (k) mede como a forma do histograma diverge de uma distribuição normal, sendo definido como (EVERITT, 2002):

$$k = \frac{\mu_4}{\mu_2^2}.$$

Os outros descritores utilizados são calculados a partir da matriz de co-ocorrência de cada componente de cor adotado. A matriz de co-ocorrência é (NIXON; AGUADO, 2008) a contagem do número de *pixels* em pares para um valor de intensidade específica quando os pares estão separados por uma distância em uma certa direção.

Foram utilizados os descritores estatísticos de contraste, energia e homogeneidade, todos calculados a partir da matriz de co-ocorrência, a Tabela 5 a seguir descreve cada um deles.

Tabela 5 – Características estatísticas extraídas utilizando a matriz de co-ocorrência de cada componente de representação de cores, adaptado de (GONZALEZ; WOODS, 2008), i e j são as colunas e linhas da matriz de co-ocorrência e p é o valor em cada coordenada.

Característica	Descrição	Equação
Contraste	Medida da diferença de intensidade entre um pixel e sua vizinhança.	$\sum_{i,j} i - j ^2 p(i, j)$
Energia	Medida de uniformidade em uma imagem.	$\sum_{i,j} p(i, j)^2$
Homogeneidade	Mede a semelhança entre os <i>pixels</i> próximos de uma imagem	$\sum_{i,j} \frac{p(i, j)}{1 + i - j }$

Também foi utilizada a medida de entropia como característica, é uma medida estatística de aleatoriedade que pode ser utilizada para caracterizar a textura de uma imagem (GONZALEZ; WOODS, 2008), a entropia foi calculada sobre os *bin* do histograma de cada um dos componentes de cor das imagens. A entropia pode ser definida como:

$$E = - \sum_{i=0}^N c_i \log_2(c_i),$$

sendo N o número de *bins* no histograma e c_i o valor de cada *bin*.

3.2.4. Análise de Componentes Principais

Análise de componentes principais (PCA, do inglês *Principal Components Analysis*) ou Transformação de Karhunen-Loève (DUDA; HART; STORK, 2001) é uma técnica que utiliza medidas estatísticas para reduzir a dimensão de um conjunto de dados preservando a maior parte da informação original (ENGELBRECHT, 2007). A ideia principal do PCA é obter um novo espaço de representação do conjunto original de dados, sendo normalmente utilizado antes de tarefas de reconhecimento de padrões como um filtro para dados, também pode ser utilizada para análise das características (WITTEN; FRANK, 2011).

Um conjunto de dados com n atributos numéricos pode ser visualizado como uma nuvem de pontos em um espaço n -dimensional, pode-se escolher um eixo para projeção desse conjunto de dados que obtenha máxima variância entre os pontos projetados. Em seguida um novo eixo é posicionado perpendicular a este primeiro, buscando a máxima variância possível, esse processo é realizado repetitivamente até que se consiga representar toda a variância original do conjunto de dados nas projeções (WITTEN; FRANK, 2011).

A análise de componentes principais pode ser formalizada conforme segue (DUDA; HART; STORK, 2001). Considere um conjunto de dados com d atributos, calcula-se a matriz de covariância $d \times d$ (Σ) para o conjunto de dados originais, em seguida se obtém os autovalores (λ) e autovetores (e) a partir de Σ , em seguida os autovalores são ordenados de forma decrescente, e cada um representa a quantidade de variância que seu autovetor possui dos dados originais.

Dado o conjunto de autovetores, cada ponto do conjunto de dados original pode ser representado através de uma combinação linear dos autovetores (ENGELBRECHT, 2007). Podem ser utilizados todos os autovetores para obtenção dos novos dados projetados, ou apenas uma parte de acordo com a quantidade de informação necessária sobre os dados originais.

3.2.5. Agrupamento

O objetivo do agrupamento de dados, também conhecido como *clustering*, é analisar a forma como um conjunto de dados se agrupam naturalmente. Os grupos que são formados dependem diretamente das características consideradas dos dados (JAIN, 2010).

Técnicas de agrupamento de dados são utilizados para os seguintes propósitos (JAIN, 2010; WU, 2012):

- Para *insight* a respeito de dados: geração de hipóteses, detecção de anomalias;
- Identificação do grau de semelhança entre os dados;
- Organização de dados através de protótipos de grupos.

Quando são utilizadas informações a respeito das classes a que os dados pertencem, o agrupamento recebe o nome de agrupamento supervisionado. Por outro lado, no caso em que não se tem informações a respeito das classes, o agrupamento recebe o nome de agrupamento não-supervisionado (WU, 2012). Os algoritmos de agrupamento também podem ser divididos em hierárquicos ou particionais, para uma visão sobre as diferentes técnicas de agrupamento recomenda-se o trabalho de (JAIN, 2010).

Neste trabalho utilizamos técnicas de agrupamento não supervisionado, tanto hierárquico, como particionais para investigar as características extraídas das imagens de fluorescência dental. As técnicas de agrupamento foram aplicadas tanto nas 35 características extraídas originalmente, quanto nas processadas utilizando a técnica PCA.

Agrupamento Hierárquico

No método de agrupamento hierárquico, os dados são rotulados através de uma sequência de partições considerando a distância entre os indivíduos (EVERITT et al., 2011). Podem ser divididos em duas abordagens (JAIN, 2010):

- Método aglomerativo: uma sequência de junção dos n indivíduos em n grupos, cada um contendo apenas um indivíduo, até a formação de um único grupo contendo todos os indivíduos.
- Método divisivo: uma sequência de divisões de um grupo único contendo os n indivíduos até a obtenção de n grupos, cada um contendo apenas um indivíduo.

A representação de um agrupamento hierárquico é realizada através de um diagrama de duas dimensões conhecido como dendrograma. Tal representação ilustra as fusões ou divisões realizadas em cada etapa do algoritmo.

Agrupamento hierárquico é uma ferramenta investigativa (EVERITT et al., 2011), permite a visualização do número ótimo de grupos em um conjunto de dados. Tal número ótimo pode ser obtido informalmente através de observação do dendrograma, ou utilizando fórmulas que analisam a relação entre e intra grupos.

Uma das técnicas para comparar diferentes agrupamentos para um mesmo conjunto de dados é através do coeficiente de correlação cofenética (EVERITT et al., 2011), que mede a distorção entre os dados originais e um dendrograma, as medidas variam entre zero e um e valores maiores indicam uma menor distorção, ou uma melhor representação dos dados pelo dendrograma. Neste trabalho utilizamos o algoritmo de agrupamento hierárquico *weighted average linkage McQuitty* (EVERITT et al., 2011) para obter o dendrograma para os vetores de característica e também calculamos o coeficiente de correlação cofenética.

Agrupamento Particional

Agrupamento particional é uma metodologia que busca a associação ótima dos n indivíduos de um conjunto de dados em um número pré-definido de grupos (EVERITT et al., 2011). O resultado obtido a partir do agrupamento particional foi utilizado para obter as medidas de performance do nosso modelo de conhecimento específico para a detecção da cárie.

A técnica utilizada neste trabalho foi o algoritmo k-médias, ou *k-means*, ele é um algoritmo não supervisionado e particional que busca achar k grupos não sobrepostos a partir de um conjunto de dados. É um algoritmo relativamente simples e capaz de produzir resultados adequados.

Considere um conjunto de dados a serem agrupados $D = \{x_1, \dots, x_n\}$, pode-se definir o algoritmo k-médias como a função (WU, 2012):

$$\min_{\{m_k\}, 1 \leq k \leq K} \sum_{k=1}^K \sum_{x \in C_K} \pi_x \text{dist}(x, m_k),$$

Em que π_x é o peso de x , C_K é o k -ésimo grupo e m_k é o centroide do k -ésimo grupo, definido como (WU, 2012):

$$m_k = \sum_{x \in C_K} \pi_x x / n_k,$$

em que n_k é o tamanho do conjunto de dados a ser agrupado.

O processo de agrupamento ocorre de forma iterativa. Cada elemento é rotulado pela função do algoritmo k-médias como grupo k , este grupo corresponde ao que resulta na distância mínima entre o elemento e o centroide m_k .

Para o agrupamento particional dos vetores de características obtidos foi utilizado o número de grupo $k = 2$. Tal número de grupos refletem os dois estados do conjunto de dados: cariado e não cariado. Os centroides foram determinados através de uma fase preliminar de agrupamento em 10% dos dados originais e utilizando a distância Euclidiana como medida.

4. RESULTADOS E DISCUSSÃO

Neste capítulo serão apresentados e discutidos os resultados obtidos com as metodologias propostas. Este capítulo está dividido em três partes, sendo a primeira referente ao gerenciamento da cárie, a segunda referente à detecção da cárie, e a terceira apresenta os resumos dos trabalhos aceitos em congressos.

4.1. Gerenciamento da Cárie

Devido à falta de dados correlacionando todas as variáveis envolvidas no desenvolvimento do sistema, a primeira avaliação foi realizada através da apresentação da estrutura da CMBN desenvolvida para dois odontologistas em um encontro pessoal, sendo um profissional clínico e outro acadêmico, os próximos parágrafos são baseados nas opiniões deles. O sistema pode ser útil em áreas como saúde pública e ensino odontológico, as relações causais expostas através da interface gráfica da CMBN leva a uma rápida compreensão do sistema e sua principal função: oferecer suporte ao usuário durante o raciocínio de gerenciamento da cárie, sugerindo o tratamento mais indicado e previsão de retorno do paciente para o cenário individual de cada caso, sempre considerando os fatores de risco individuais.

Na saúde pública, especificadamente na atenção básica, o sistema pode auxiliar dentistas ou outros profissionais da saúde, como higienistas e pediatras, na prevenção e controle da cárie dental. Também pode auxiliar em triagens odontológicas de populações, como grupos desfavorecidos com pouco ou nenhum acesso aos serviços de saúde.

No ensino da odontologia o sistema pode ser útil como uma base para aulas de aprendizagem baseada em problemas (PBL, do inglês *problem based learning*), auxiliando estudantes na análise e entendimento do raciocínio necessário para o diagnóstico da cárie. Este resultado é referente apenas as variáveis escolhidas e as relações entre estas na estrutura final da Rede Bayesiana, para uma análise das probabilidades obtidas utilizando foram utilizados alguns casos bem definidos como

entrada da Rede Bayesiana desenvolvida, as sugestões obtidas estão nas tabelas 6, 7 e 8.

Tabela 6 – Sugestões de risco de cárie obtidas a partir da Rede Bayesiana desenvolvida utilizando alguns casos bem definidos. Tem-se o valor evidenciado em cada um dos nós pais e a distribuição de probabilidade obtida no nó filho “Risco de Cárie”. B, M e A respectivamente representam baixo, médio e alto.

Cáries passadas	Fluxo salivar	Exposição a fluoretos	Dieta	Higiene Oral	Risco de Cárie
Zero	Normal	Sim	Boa	Boa	Baixo 60% Médio 20% Alto 20%
Entre 1 e 2	Normal	Sim	Fraca	Fraca	Baixo 0% Médio 40% Alto 60%
Zero	Baixo	Não	Fraca	Fraca	Baixo 0% Médio 0% Alto 100%
3 ou mais	Baixo	Não	Boa	Boa	Baixo 0% Médio 20% Alto 80%

Tabela 7 – Sugestões tratamento obtidas a partir da Rede Bayesiana desenvolvida utilizando alguns casos bem definidos. Tem-se o valor evidenciado em cada um dos nós pais e a distribuição de probabilidade obtida no nó filho “Tratamento”. F/S, Rest e Endo respectivamente representam fluoreto ou selante, restauração e tratamento endodôntico.

Superfície Lesionada	Tipo do Dente	ICDAS	Risco de Cárie	Atividade da Lesão	Tratamento
Oclusal	Molar	1&2	Baixo	Sim	Fluoreto ou Selante 100% Restauração 0% Endodôntico 0%
Oclusal	Molar	1&2	Médio	Sim	Fluoreto ou Selante 60% Restauração 40% Endodôntico 0%
Lisa	Pré-molar	5&6	Alto	Não	Fluoreto ou Selante 0% Restauração 50% Endodôntico 50%

Tabela 8 – Sugestões tratamento obtidas a partir da Rede Bayesiana desenvolvida utilizando alguns casos bem definidos. O valor evidenciado em cada um dos nós pais e a distribuição de probabilidade obtida no nó filho “Retorno”.

Risco de Cárie	Atividade da Lesão	Retorno
Baixo	Sim	1 Ano 0%
		6 Meses 100%
		Menos que 6 meses 0%
Baixo	Não	1 Ano 100%
		6 Meses 0%
		Menos que 6 meses 0%
Médio	Sim	1 Ano 0%
		6 Meses 0%
		Menos que 6 meses 100%
Médio	Não	1 Ano 0%
		6 Meses 100%
		Menos que 6 meses 0%

Os resultados obtidos para estes casos bem definidos mostraram-se coerentes com as conclusões da literatura, sugerindo abordagens restauradoras e invasivas apenas em situações onde o risco de cárie e os sinais da lesão estão em estágios mais avançados. Nos casos onde o risco de cárie e os sinais da lesão refletem estágios menos avançados, são fornecidas sugestões de abordagens terapêuticas. A condição de d-separação entre algumas variáveis da CMBN, conforme discutido na Seção 3.1.2, permite uma divisão do uso do sistema em duas etapas, a primeira sendo referente à uma avaliação de risco.

O resultado principal foi a modelagem estrutural da CMBN, realizada a partir do conhecimento específico presente na literatura. Como já mencionado, não existem dados que correlacionam todas as variáveis de interesse, impossibilitando assim alguma abordagem de aprendizagem computacional para a obtenção de tal estrutura.

A opinião dos especialistas juntamente com os resultados obtidos utilizando os casos bem definidos mostram que a modelagem desenvolvida neste trabalho fornece resultados coerentes com a situação real de tomada de decisão para o gerenciamento da cárie.

4.2. Detecção da Cárie

A metodologia utilizada para a modelagem do conhecimento específico necessário a um sistema de auxílio a detecção de cárie utilizando imagens de fluorescência dental se encontra na Figura 21. Os resultados foram obtidos através de diferentes análises das características extraídas, buscando uma avaliação de performance para a tarefa de descrever a presença ou não de lesões cariosas a partir de uma imagem de fluorescência.

Figura 21 – Metodologia utilizada para modelar um conhecimento específico para um sistema de auxílio a detecção de cárie utilizando imagens de fluorescência dental.

A extração de características resultou em um vetor com 35 elementos para cada uma das amostras (7 medidas para cada um dos componentes: R, G, Y, Cb e Cr), os elementos dos vetores foram processados utilizando a técnica de PCA. O resultado desta análise é apresentado através de um *scree plot*⁵ na Figura 22.

A partir da análise do *scree plot* obtido, tem-se que, utilizando as 7 características que possuem maior variância, representa-se mais de 90% da informação original contida no vetor de características original, e utilizando as 11 características com maior variância representamos praticamente 100% da informação original.

⁵ Representação gráfica que apresenta a variância de cada elemento de um conjunto de dados, relaciona a quantidade de componentes necessários para representar uma parte da variância total em um conjunto de dados.

Figura 22 – Scree plot representando a variância de cada uma das 35 características após a utilização do PCA (barras verdes), as características foram ordenadas em valores decrescentes de variância. Os diamantes azuis representam a porcentagem cumulativa da variância total que cada característica adiciona.

Observa-se que a utilização da análise de componentes principais mostrou que os diferentes descritores de textura possuem bastante redundância na informação representada por cada um, assim a análise de componentes principais também funcionou como um filtro levando a uma redução dimensional das características extraídas.

Como análise inicial do resultado obtido a partir do PCA, geramos um dendrograma aplicando o algoritmo de agrupamento hierárquico (Seção 3.2.5.1) para os 11 componentes mais representativos (maior variância), tal resultado se encontra na Figura 23. Tal figura apresenta um dendrograma claramente dividido em dois grandes grupos: um contendo amostras cariadas, outro contendo amostras não cariadas.

Figura 23 – Dendrograma obtido utilizando os onze componentes mais significativos após o PCA para agrupamento hierárquico não-supervisionado do conjunto de dados. Pode-se observar a divisão que ocorre nas amostras, na parte superior do dendrograma se encontram as amostras não cariadas, já na parte inferior as amostras cariadas.

Após a utilização do PCA o algoritmo *k-means*, com $k=2$, foi utilizado no vetor de características original, e em cada conjunto contendo as características $(1, \dots, i)$, com i variando de 2 até 11. Para analisar a performance das características os

resultados obtidos com o *k-means* foram comparados com a classificação do especialista, os resultados estão apresentados na Tabela 9.

Tabela 9 – Resultados de performance obtidos com os vetores de características, tanto o original, quanto os processados, comparando os rótulos obtidos com o algoritmo k-means (k=2) com a classificação prévia de um especialista. A primeira linha da tabela indica os resultados obtidos para o vetor de características original, a segunda linha representa os resultados obtidos para os vetores de característica utilizando de 3 a até 11 elementos após o PCA, a terceira linha representa a utilização de apenas as 2 características mais significantes após o PCA.

Entrada k-means	Acurácia	Falso positivo	Falso negativo
Vetor de características original	93.75%	0%	6.25%
Vetor de características – (3 – 11)	95.31%	1.56%	3.13%
Vetor de características - 2	96.88%	0%	3.13%

Os resultados de performance da Tabela 5 mostram que as características de textura extraídas podem ser bons descritores para a tarefa de classificação de dentes em cariados ou não cariados. A utilização das duas características mais significantes após o PCA resultou no melhor desempenho, com 96,88% de acurácia e 0% de falso positivo, este resultado mostra que uma parte das características extraídas acabaram diminuindo a performance do agrupamento. O resultado do agrupamento com melhor performance é apresentado de forma gráfica na Figura 24.

Figura 24 – Resultado obtido utilizando k-means ($k=2$) com apenas as duas características mais significantes após o uso do PCA. O Cluster 1 em azul representa o grupo das amostras cariadas, o Cluster 2 em verde representa o grupo das amostras não cariadas, os centroides dos cluster estão identificados em preto.

Observando a representação gráfica do agrupamento, Figura 24, tem-se que as amostras não cariadas são muito semelhantes entre si, e que as amostras cariadas possuem uma maior variabilidade entre si, refletindo os diversos níveis de lesões cariosas presentes no conjunto de dados.

Também geramos um dendrograma aplicando o algoritmo de agrupamento hierárquico (Seção 3.2.5.1) para os 2 componentes mais representativos (com maior variância), Figura 25. Tal figura também apresenta um dendrograma claramente dividido em dois grandes grupos: um contendo amostras cariadas, outro contendo amostras não cariadas. Porém, a similaridade intra grupos e distância entre grupos são maiores do que a obtida no dendrograma anterior, Figura 23, principalmente no grupo referente as amostras cariadas. Tal fato é coerente com a melhoria de performance obtida utilizando apenas estes dois componentes, conforme apresentado nos resultados da Tabela 8.

Figura 25 – Dendrograma obtido utilizando os dois componentes mais significativos após o PCA para agrupamento hierárquico não-supervisionado do conjunto de dados. Pode-se observar a divisão que ocorre nas amostras, na parte superior do dendrograma se encontram as amostras não cariadas, já na parte inferior as amostras cariadas, similar ao dendrograma anterior. Porém a similaridade intra grupo e a distância entre grupos aumentaram em relação ao dendrograma anterior.

Para comparar os diferentes vetores de características obtidos e investigar a melhora de performance observada nos resultados do *k-means* utilizando apenas os dois componentes mais significativos do PCA, calculou-se o coeficiente de correlação cofenética (Sec. 3.2.5) para os vetores de característica original e os obtidos a partir do PCA, Figura 26.

Figura 26 – Coeficiente de correlação cofenética calculado para o vetor de características original (VCO) e para as saídas do PCA. Os números (n-PCA) indicam a quantidade de componentes do PCA que foram utilizadas para a obtenção dos coeficientes. O resultado é coerente com a melhora de performance observada na saída do *k-means* (Tabela 9), utilizando apenas os dois componentes mais significativos obtêm-se o maior valor para o coeficiente de correlação cofenética. Tal resultado reforça a evidência de que algumas características extraídas estão atuando como ruído no processo de agrupamento.

O valor do coeficiente de correlação cofenética aumenta quando compara-se o vetor de características original com os obtidos a partir do PCA, existe um decaimento no valor do coeficiente no momento em que utiliza-se menos que 9 componentes, porém utilizando apenas os dois componentes mais significativos obtêm-se o valor máximo do coeficiente. Tal resultado reforça a evidência de que algumas características extraídas estão atuando como ruído no processo de agrupamento.

4.3. Trabalhos Publicados ou Aceitos em Congressos

4.3.1. XVIII Congresso Brasileiro de Física Médica - 2013

Viability of a Dental Caries CAD System based on Feature Extraction of Fluorescence Images

BESSANI, M.¹, COSTA, M. M.¹, MACIEL, C. D.¹

¹University of São Paulo/Department of Electrical and Computational Engineering, São Carlos, Brazil

Abstract. Dental caries is the most common oral pathology and an early caries diagnosis is fundamental to allow less invasive treatments, even non-invasive. One modern technique to aid the early diagnosis is the fluorescence imaging. This paper analyzes the viability of a dental caries computer-aided diagnosis (CAD) system, which is based on feature extraction of dental fluorescence imaging. Fluorescence imaging can be employed as a direct way to analyze pathologies on biologicals tissues, because the fluorescence signal is directly dependent of tissues intrinsic characteristics. On dental fluorescence images, the enamel signal is located on the green wavelength and is proportional to its health, while microorganisms signal is located on the red wavelength, both being excited on the UV band. Color feature extraction are performed on the fluorescence images, followed by clustering of the in vitro samples images using different combinations of the features extracted. The clustering is performed using the k-means method with two output classes, carious or non-carious. Analyzing the accuracy of each features combinations used for clustering, we can propose the most suitable colour descriptors for this kind of classification and show the viability of the proposed CAD system.

Keywords: fluorescence microscopy, instrumentation, imaging processing algorithms, oral diagnosis.

4.3.2. XVIII Congresso Brasileiro de Física Médica – 2013 - Resumo

DESENVOLVIMENTO DE EQUIPAMENTO PARA DIAGNÓSTICO DE CÁRIE POR IMAGENS DE FLUORESCÊNCIA

Costa, Mardoqueu Martins¹; Lins, E.C.C.²; Bessani, M.¹; Ventura, L.¹

¹Universidade de São Paulo, São Paulo, Brasil. ²Universidade Federal do ABC, São Paulo, Brasil.

Introdução:

A cárie dental é responsável pelo maior número de atendimentos nos consultórios de odontologia e tida como uma das principais doenças bucais em nosso país; seu diagnóstico precoce possibilita um tratamento rápido, minimizando o desgaste da estrutura dental e permitindo um tratamento menos agressivo [1]. Na rotina diária das clínicas odontológicas, o diagnóstico da cárie é realizado pela inspeção visual do cirurgião-dentista podendo ser auxiliado por radiografias. Apesar de serem boas formas de triagem, nem a inspeção visual nem as radiografias por raios-x são eficientes para todos os casos, devido ao baixo contraste observado entre a cárie e o tecido saudável obtido por esses métodos. Assim o objetivo deste trabalho é apresentar o desenvolvimento de um equipamento para diagnóstico precoce da cárie através de imagens de fluorescência.

4.3.3. SPIE 2014 - Multimodal Biomedical Imaging IX

UNSUPERVISED CLUSTERING ANALYSES OF FEATURES EXTRACTION FOR A CARIES COMPUTER-ASSISTED DIAGNOSIS USING DENTAL FLUORESCENCE IMAGES

Michel Bessani^a, Mardoqueu M. da Costa^a, Emery C. C. C. Lins^b, Carlos D. Maciel^a

^aDep. Of Electrical and Computer Engineering, University of São Paulo, 400 Trabalhador São Carlense Ave., São Carlos, SP, Brazil 13566-590; ^bEngineering, Modeling and Applied Social Sciences Center, Federal University of ABC, 5001 States Ave., Santo André, SP Brazil 09210-170

ABSTRACT

Computer-assisted diagnoses (CAD) are performed by systems with embedded knowledge. These systems work as a second opinion to the physician and use patient data to infer diagnoses for health problems. Caries is the most common oral disease and directly affects both individuals and the society. Here we propose the use of dental fluorescence images as input of a caries computer-assisted diagnosis. We use texture descriptors together with statistical pattern recognition techniques to measure the descriptors performance for the caries classification task. The data set consists of 64 fluorescence images of in vitro healthy and carious teeth including different surfaces and lesions already diagnosed by an expert. The texture feature extraction was performed on fluorescence images using RGB and YCbCr color spaces, which generated 35 different descriptors for each sample. Principal components analysis was performed for the data interpretation and dimensionality reduction. Finally, unsupervised clustering was employed for the analysis of the relation between the output labeling and the diagnosis of the expert. The PCA result showed a high correlation between the extracted features; seven components were sufficient to represent 91.9% of the original feature vectors information. The unsupervised clustering output was compared with the expert classification resulting in an accuracy of 96.88%. The results show the high accuracy of the proposed approach in identifying carious and non-carious teeth. Therefore, the development of a CAD system for caries using such an approach appears to be promising.

Fluorescence Imaging, Unsupervised Clustering, Caries, Computer-Assisted Diagnosis, Dental Fluorescence, Feature Extraction.

4.3.4. XXIV Brazilian Congress on Biomedical Engineering – CBEB 2014⁶

CONSTRUCTION OF A CLINICAL DECISION SUPPORT SYSTEM FOR DENTAL CARIES MANAGEMENT USING BAYESIAN NETWORKS

M. Bessani*, E.C.C.C. Lins**, A.C.B. Delbem*** e C.D. Maciel*

*Department of Electrical and Computer Engineering, University of São Paulo, São Carlos, Brazil

** Engineering Modelling and Applied Social Science Center, Federal University of ABC, Santo André, Brazil

*** Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, Brazil
e-mail: michel.bessani@usp.br

Abstract: Caries are still a major problem and an active area of research; it is a multifactorial dynamic process, and therefore, management should be dichotomous, acting on both the symptoms and the causal factors. On this scenario, we propose a clinical decision support system for modern caries management using a Bayesian network as framework. The first version of the network was modelled using recent findings and conclusions from the scientific literature on cariology area. Two specialists were asked about the network on a personal meeting, and their opinions agree: this kind of system can be helpful on areas such as: public health, supporting the primary dental care; and dentistry education, as a problem based learning system. In addition, we used some well-defined cases on the network, the suggested outputs are in agreement with our expectations. The next steps are the validation of the Bayesian network using data from specialist's decisions.

Keywords: Clinical decision support system, Bayesian network, dental caries management.

⁶ Trabalho aceito no BIOSTEC – HEALTHINF 2014, porém não publicado.

4.3.5. XXIV Brazilian Congress on Biomedical Engineering – CBEB 2014

DESENVOLVIMENTO DE UM PROTÓTIPO DE CÂMERA INTRAORAL DE FLUORESCÊNCIA A PARTIR DE UMA CÂMERA INTRAORAL CONVENCIONAL

C. L. Grandisoli*, M. Bessani**, M. M. da Costa**, P. A. da Ana*, e E. C. Lins*

*Centro de Engenharias, Modelagem e Ciências Sociais Aplicadas – Universidade Federal do ABC, Santo André, Brasil

** BIOPDI Ind. de Equipamentos Médico-Hospitalares e Odontológicos Ltda., São Carlos, Brasil
e-mail: emery.lins@ufabc.edu.br

Resumo: A técnica de fluorescência vem se mostrando eficaz na detecção de metabólitos de bactérias causadoras de cáries e esta tecnologia tem sido incorporada em câmeras intraorais. A busca de imagens de qualidade com baixo custo motivou a montagem do protótipo de câmera intraoral por fluorescência, que possibilitou realizar imagens com luz no comprimento de onda no espectro visível e também utilizando luz ultravioleta. Para a montagem proposta utilizou-se uma câmera intraoral comum com detector CMOS, um filtro passa-altas com abertura espectral em 495nm, e circuitos de LEDs Brancos e LEDs Ultravioletas (UV). Os resultados possibilitaram a diferenciação de estruturas sadias e com presença de metabólitos bacterianos, ao passo que indicam que a instrumentação é suficiente para a aquisição de imagens de fluorescência.

Palavras-chave: Fluorescência, câmera intraoral, cárie.

5. CONCLUSÃO

O funcionamento de um sistema de suporte a decisão consiste basicamente de um sistema de inferência e conhecimento específico, sendo que estes dois trocam informações entre si. Diferentes técnicas podem ser utilizadas para o sistema de inferência, como regras *if-then*, redes neurais e classificadores, o conhecimento específico vai ser diretamente relacionado ao problema em que o DSS vai ser inserido.

Em um sistema que utiliza redes Bayesianas como sistemas de inferência por exemplo, o conhecimento específico é armazenado na forma de tabelas de probabilidades condicionais, é importante mencionar que ao contrário do conhecimento específico que é dependente do problema, um mesmo sistema de inferência pode ser utilizado em diferentes problemas.

O objetivo deste trabalho foi a análise e modelagem de sistemas de suporte a decisão, foram adotadas duas abordagens distintas desenvolvimento de um sistema de suporte, uma para o gerenciamento do problema, e outra para a detecção do problema, o problema alvo escolhido foi a cárie dental.

A cárie dental é um problema que atinge a maior parte da população mundial, tanto a jovem, quanto a adulta, sua incidência ocorre tanto em países ricos, como em países pobres, sendo que nestes a incidência é generalizada. É uma patologia crônica, influenciando diretamente em vários indicadores de qualidade de vida e também de desenvolvimento infantil, é um problema de saúde pública. O último relatório da OMS a respeito de saúde bucal levanta a necessidade do desenvolvimento de novas metodologias que utilizem os conceitos modernos desenvolvidos em diversas áreas do conhecimento para gerar métodos inovadores e impactantes para auxílio ao diagnóstico e tratamento das doenças bucais, inclusive da cárie dental.

A parte de detecção do problema analisou a utilização de imagens de fluorescência juntamente com técnicas de extração de características de imagens e processamento de sinais, buscando uma forma de modelagem das informações que imagens de fluorescência possuem a respeito de lesões cariosas. Neste caso

o sinal de fluorescência seria a entrada de um CDSS para detecção da cárie, utilizando um sistema de inferência para fornecer uma sugestão de detecção da cárie: face cariada ou não cariada.

A análise foi realizada através do agrupamento das características de textura extraídas das imagens de fluorescência dental, tanto antes, quanto depois de uma análise de componentes principais, os resultados obtidos mostraram a que as características extraídas são eficazes para tal tarefa.

Em comparação a outros trabalhos referentes a detecção de lesões cariosas, que em sua maioria consistem na utilização de imagens de Raio-X (radiação ionizante), a utilização de imagens de fluorescência (radiação não-ionizante) que pode fornecer um resultado em tempo real juntamente com os resultados obtidos neste trabalho mostram que esta abordagem é promissora.

A parte referente ao gerenciamento da Cárie consiste no uso de redes bayesianas para auxílio ao processo de diagnóstico da cárie. A motivação desta abordagem é desenvolver um método auxiliar para o processo de raciocínio referente as decisões de gerenciamento da cárie, considerando os fatores individuais de risco de cárie, que juntamente com as observações clínicas, podem levar aos melhores resultados dos tratamentos, não apenas curando a lesão, mas também controlando o processo dinâmico da cárie em cada paciente. As relações causais e as probabilidades conjuntas entre as variáveis da Rede Bayesiana foram obtidas a partir de conclusões da literatura científica. Nosso sistema pode ser chamado de um sistema de suporte a decisão clínica para o gerenciamento da cárie.

A primeira versão desenvolvida foi apresentada a dois profissionais da odontologia e a opinião deles foi encorajadora, o sistema poderia ser utilizado tanto na atenção básica à saúde, quanto na educação odontológica de futuros profissionais clínicos.

O sistema funcionaria de tal forma que o profissional clínico acessaria as variáveis e evidenciaria informações concretas obtidas a partir de entrevista ou inspeção clínica visual a respeito do paciente, a partir das entradas fornecidas pelo

profissional clínico as probabilidades das variáveis de interesse seriam atualizadas, as probabilidades finais auxiliariam na tomada de decisão a respeito do risco de cárie, tratamento e retorno de cada paciente.

A utilização de casos bem definidos como entrada da BN desenvolvida forneceu resultados coerentes com o cenário real de decisão do problema alvo, entretanto, é necessária a utilização de um maior número de dados para uma validação mais robusta.

Além da contribuição para a área de informática odontológica, os resultados mostram que a abordagem utilizada para a modelagem da Rede Bayesiana neste trabalho pode ser utilizada em outros cenários em que o conhecimento específico sobre a etiologia do problema e os resultados para diferentes ações sejam profundamente conhecidos pelos especialistas da área.

Do ponto de vista de modelagem de sistemas e redes Bayesianas, alguns pontos importantes poderão ser analisados no futuro do projeto, como escalabilidade das propriedades da rede e testes de validação, tanto da parte estrutural, como da parte qualitativa referente as probabilidades condicionais entre as variáveis da rede.

Um objetivo futuro do projeto é a integração do resultado fornecido pelo sistema determinístico como uma das variáveis de entrada da Rede Bayesiana, o que aumentará a usabilidade do sistema. Em contrapartida, esta integração elevará a complexidade da modelagem, validação e geração de casos randômicos.

Tal complexidade levará à pesquisa de metodologias computacionalmente eficientes, como programação paralela e subamostragem de dados, principalmente para à utilização de forma eficiente do método de Monte Carlo na geração de dados randômicos.

REFERÊNCIAS

ANGMAR-MÅNSSON, B.; TEN BOSCH, J. J. Quantitative light-induced fluorescence (QLF): a method for assessment of incipient caries lesions. **Dento maxillo facial radiology**, v. 30, n. 6, p. 298–307, nov. 2001.

BADER, J. D. et al. Validation of a simple approach to caries risk assessment. **Journal of public health dentistry**, v. 65, n. 2, p. 76–81, jan. 2005.

BAELUM, V. Caries management: technical solutions to biological problems or evidence-based care? **Journal of oral rehabilitation**, v. 35, n. 2, p. 135–51, fev. 2008.

BAELUM, V.; HEIDMANN, J.; NYVAD, B. Dental caries paradigms in diagnosis and diagnostic research. **European journal of oral sciences**, v. 114, n. 4, p. 263–77, ago. 2006.

BEEMER, B. A.; GREGG, D. G. Dynamic interaction in knowledge based systems: An exploratory investigation and empirical evaluation. **Decision Support Systems**, v. 49, n. 4, p. 386–395, nov. 2010.

BERNER, E. S. **Clinical Decision Support Systems: Theory and Practice**. 2. ed. New York: Springer, 2007.

BHATIA, A.; SINGH, R. Using Bayesian Network as Decision making system tool for deciding Treatment plan for Dental caries. **Journal of Academia and Industrial Research (JAIR)**, v. 2, n. 2, p. 93–96, 2013.

BOYCE, W. T. et al. Social inequalities in childhood dental caries: the convergent roles of stress, bacteria and disadvantage. **Social science & medicine (1982)**, v. 71, n. 9, p. 1644–52, nov. 2010.

BROADBENT, J. M.; THOMSON, W. M. For debate: problems with the DMF index pertinent to dental caries data analysis. **Community dentistry and oral epidemiology**, v. 33, n. 6, p. 400–9, dez. 2005.

CAROUNANIDY, U.; SATHYANARAYANAN, R. Dental caries: A complete changeover (Part II)-Changeover in the diagnosis and prognosis. **Journal of conservative dentistry : JCD**, v. 12, n. 3, p. 87–100, jul. 2009.

CAROUNANIDY, U.; SATHYANARAYANAN, R. Dental caries: A complete changeover (PART III): Changeover in the treatment decisions and treatments. **Journal of conservative dentistry : JCD**, v. 13, n. 4, p. 209–17, out. 2010.

CASAMASSIMO, P. et al. Beyond the dmft the human and economic cost of early childhood caries. **The Journal of the American Dental Association**, v. 140, n. 6, p. 650–657, 2009.

COSTA, S. M. et al. A systematic review of socioeconomic indicators and dental caries in adults. **International journal of environmental research and public health**, v. 9, n. 10, p. 3540–74, out. 2012.

DARWICHE, A. **Modeling and reasoning with Bayesian networks**. 1. ed. New York: Cambridge University Press, 2009.

DE JOSSELIN DE JONG, E. et al. Quantified light-induced fluorescence, review of a diagnostic tool in prevention of oral disease. **Journal of Applied Physics**, v. 105, n. 10, p. 102031, 2009.

DUDA, R.; HART, P.; STORK, D. **Pattern classification**. 2. ed. [s.l.] John Wiley & Sons Inc, 2001.

ENGELBRECHT, A. **Computational intelligence: an introduction**. 2nd ed ed. West Sussex: John Wiley & Sons, Ltd., 2007.

EVERITT, B. **The Cambridge dictionary of statistics**. 2nd ed ed. New York: Cambridge University Press, 2002.

EVERITT, B. S. et al. **Cluster Analysis**. Chichester, UK: John Wiley & Sons, Ltd, 2011.

FEIGAL, R. et al. Evidence-based clinical recommendations for the use of pit-and-fissure sealants. **JAMA: the journal of the American Dental Association**, v. 139, n. March, p. 257–268, 2008.

FEJERSKOV, O.; KIDD, E. **Dental caries: the disease and its clinical management**. 2nd. ed. Oxford: Blackwell Munksgaard Ltd, 2008.

FERREIRA ZANDONÁ, A. et al. The natural history of dental caries lesions: a 4-year observational study. **Journal of dental research**, v. 91, n. 9, p. 841–6, set. 2012.

FILSTRUP, S. L. et al. Early childhood caries and quality of life: child and parent perspectives. **Pediatric dentistry**, v. 25, n. 5, p. 431–40, 2003.

FINK, A.; SEBASTIAN, H.-J.; VOS, S. **Introduction to Intelligent Decision Support for Logistics and Supply Chain Management Minitrack** Hawaii International Conference on System Sciences. **Anais...Ieee**, jan. 2012 Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6149036>>. Acesso em: 28 fev. 2014

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 3rd ed ed. Upper Saddle River: Prentice Hall, 2008. v. 14

GREENES, R. **Clinical decision support: the road ahead**. Oxford: Elsevier Inc., 2007.

GUPTA, J.; FORGIONNE, G.; MORA, M. **Intelligent decision-making support systems**. London: Springer-Verlag, 2006.

HOLMES, D.; JAIN, L. **Innovations in Bayesian Networks**. Berlin: Springer-Verlag, 2008.

HÜLLERMEIER, E. **Case-based approximate reasoning**. Dordrecht: Springer, 2007.

ICDAS. **ICDAS Foundation - Internation Caries Detection and Assessment System**. Disponível em: <<https://www.icdas.org>>. Acesso em: 15 abr. 2013.

INSPEKTOR. **Quantitative Light-induced Fluorescence**. Disponível em: <<http://www.inspektor.nl/dental/qlfmain.htm>>. Acesso em: 5 out. 2014.

JAIN, A. K. Data clustering: 50 years beyond K-means. **Pattern Recognition Letters**, v. 31, n. 8, p. 651–666, jun. 2010.

JENSEN, F.; NIELSEN, T. **Bayesian networks and decision graphs**. 2nd ed ed. New York: Springer Science+Business Media, 2007.

JULIA FLORES, M. et al. Incorporating expert knowledge when learning Bayesian network structure: a medical case study. **Artificial intelligence in medicine**, v. 53, n. 3, p. 181–204, nov. 2011.

KAVO. **DIAGNOdent Caries Detection Aids**. Disponível em: <<http://www.kavousa.com/US/DIAGNOdent.aspx>>. Acesso em: 6 out. 2014.

KAWAHATA, N.; MACENTEE, M. I. A measure of agreement between clinicians and a computer-based decision support system for planning dental treatment. **Journal of dental education**, v. 66, n. 9, p. 1031–7, set. 2002.

KIDD, E. A. M. **Essential of dental caries**. 3. ed. New York: Oxford University Press, 2005.

KJÆRULFF, U. B.; MADSEN, A. L. **Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis**. New York: Springer Science+Business Media, 2008.

KOLLER, D.; FRIEDMAN, N. **Probabilistic graphical models: principles and techniques**. Massachusetts: MIT Press, 2009.

KRISTENSEN, K.; RASMUSSEN, I. A. The use of a Bayesian network in the design of a decision support system for growing malting barley without use of pesticides. **Computers and Electronics in Agriculture**, v. 33, n. 3, p. 197–217, mar. 2002.

KUANG, W.; YE, W. **A Kernel-Modified SVM Based Computer-Aided Diagnosis System in Initial Caries** 2008 Second International Symposium on Intelligent Information Technology Application. **Anais...IEEE**, dez. 2008 Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4739988>>. Acesso em: 1 nov. 2013

LABORATORY, D. S. **GeNIe and SMILE - Home**. Disponível em: <<https://dslpitt.org/genie/>>. Acesso em: 7 set. 2013.

LAKOWICZ, J. **Principles of fluorescence spectroscopy**. 3rd ed ed. New York: Springer Science+Business Media, 2006.

LI, W. et al. **Clinical X-Ray Image Based Tooth Decay Diagnosis using SVM**2007 International Conference on Machine Learning and Cybernetics. **Anais...**Hong Kong: IEEE, 2007Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4370404>. Acesso em: 1 nov. 2013

LIAO, S. Expert system methodologies and applications—a decade review from 1995 to 2004. **Expert Systems with Applications**, v. 28, n. 1, p. 93–103, jan. 2005.

LIMEBACK, H. **Comprehensive Preventive Dentistry**. West Sussex, UK: John Wiley & Sons, Ltd., 2012.

MAGO, V. K. et al. Clinical decision support system for dental treatment. **Journal of Computational Science**, v. 3, n. 5, p. 254–261, set. 2012.

MARCENES, W. et al. Global burden of oral conditions in 1990-2010: a systematic analysis. **Journal of dental research**, v. 92, n. 7, p. 592–7, jul. 2013.

MENDONÇA, E. A. Clinical decision support systems: perspectives in dentistry. **Journal of dental education**, v. 68, n. 6, p. 589–97, jun. 2004.

MEYER, P. L. **Introductory Probability and Statistical Applications**. 2nd. ed. [s.l.] Addison-Wesley, 1970. p. 367

MONSE, B. et al. PUFA--an index of clinical consequences of untreated dental caries. **Community dentistry and oral epidemiology**, v. 38, n. 1, p. 77–82, fev. 2010.

MURPHY, D.; DAVIDSON, M. **Fundamentals of light microscopy and electronic imaging**. New York: John Wiley & Sons Inc, 2001.

NIXON, M.; AGUADO, A. **Feature extraction & image processing**. 2nd ed ed. London: Academic Press, 2008.

NYVAD, B. Diagnosis versus detection of caries. **Caries research**, v. 38, n. 3, p. 192–8, 2004.

NYVAD, B.; MACHIULSKIENE, V.; BAELUM, V. Construct and Predictive Validity of Clinical Caries Diagnostic Criteria Assessing Lesion Activity. **Journal of Dental Research**, v. 82, n. 2, p. 117–122, 1 fev. 2003.

OLSEN, G. F. et al. **An image-processing enabled dental caries detection system**2009 ICME International Conference on Complex Medical Engineering. **Anais...**IEEE, abr. 2009Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4906674>>. Acesso em: 1 nov. 2013

PAPAMICHAIL, K. N.; FRENCH, S. Design and evaluation of an intelligent decision support system for nuclear emergencies. **Decision Support Systems**, v. 41, n. 1, p. 84–111, nov. 2005.

PEARL, J. **Probabilistic reasoning in intelligent systems: networks of plausible inference**. 1. ed. San Maeto: Morgan Kaufmann, 1988.

PENEVA, M. INDEX OF REVERSIBILITY OF THE CARIES LESIONS. **J. of IMAB**, v. 14, n. 2, p. 79–83, 2008.

PENG, Y. et al. An incident information management framework based on data integration, data mining, and multi-criteria decision making. **Decision Support Systems**, v. 51, n. 2, p. 316–327, maio 2011.

PETERSEN, P. E. **The World Oral Health Report 2003: Continuous improvement of oral health in 21st century**. Geneva: [s.n.].

PETROU, M.; SEVILLA, P. **Image processing: dealing with texture**. West Sussex: John Wiley & Sons, Ltd., 2006.

PITTS, N. Understanding the jigsaw of evidence-based dentistry. 2. Dissemination of research results. **Evidence-based dentistry**, v. 5, n. 2, p. 33–5, jan. 2004a.

PITTS, N. B. Are we ready to move from operative to non-operative/preventive treatment of dental caries in clinical practice? **Caries research**, v. 38, n. 3, p. 294–304, 2004b.

PRETTY, I. A. Caries detection and diagnosis: novel technologies. **Journal of dentistry**, v. 34, n. 10, p. 727–39, nov. 2006.

QUALTROUGH, A. J. E. et al. **Principles of operative dentistry**. [s.l.] Blackwell Munksgaard, 2005.

SELWITZ, R. H.; ISMAIL, A. I.; PITTS, N. B. Dental caries. **Lancet**, v. 369, n. 9555, p. 51–9, 6 jan. 2007.

SHAW, M. et al. **Handbook on Decision Support Systems 1**. Berlin: Springer Berlin Heidelberg, 2008.

SHEIHAM, A. Dental caries affects body weight, growth and quality of life in pre-school children. **British dental journal**, v. 201, n. 10, p. 625–6, 25 nov. 2006.

SHONKOFF, J.; BOYCE, W.; MCEWEN, B. Neuroscience, molecular biology, and the childhood roots of health disparities. **JAMA: the journal of the ...**, v. 301, n. 21, p. 2252–2259, 2009.

SIM, I. et al. Clinical decision support systems for the practice of evidence-based medicine. **Journal of the American Medical Informatics Association : JAMIA**, v. 8, n. 6, p. 527–34, 2001.

SUEBNUKARN, S.; RUNGCHAROENPORN, N.; SANGSURATHAM, S. A Bayesian decision support model for assessment of endodontic treatment outcome. **Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics**, v. 106, n. 3, p. e48–58, set. 2008.

TURBAN, E.; ARONSON, J.; LIANG, T. **Decision support systems and intelligent systems**. 7th. ed. New Delhi: Prentice Hall, 2005.

VELIKOVA, M. et al. Exploiting causal functional relationships in Bayesian network modelling for personalised healthcare. **International Journal of Approximate Reasoning**, abr. 2013.

VO-DINH, T. **Biomedical Photonics Handbook**. New York: CRC Press, 2003.

WAGHOLIKAR, K. B.; SUNDARARAJAN, V.; DESHPANDE, A. W. Modeling paradigms for medical diagnostic decision support: a survey and future directions. **Journal of medical systems**, v. 36, n. 5, p. 3029–49, out. 2012.

WALSH, S. et al. **Parameterizing Bayesian Network Representations of Social-Behavioral Models by Expert Elicitation A Novel Application of Conjoint Analysis** Artificial intelligence in medicine. **Anais...**2010

WEIDL, G.; MADSEN, A. L.; ISRAELSON, S. Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes. **Computers & Chemical Engineering**, v. 29, n. 9, p. 1996–2009, ago. 2005.

WITTEN, I.; FRANK, E. **Data Mining: Practical machine learning tools and techniques**. 3rd ed ed. Burlington: Morgan Kaufmann, 2011.

WU, J. **Advances in K-means Clustering**. Heidelberg: Springer, 2012.

ZERO, D. T. Sugars - the arch criminal? **Caries research**, v. 38, n. 3, p. 277–85, 2004.

ZERO, D. T. et al. Dental caries and pulpal disease. **Dental clinics of North America**, v. 55, n. 1, p. 29–46, jan. 2011.

ZEZELL, D. M. et al. Characterization of natural carious lesions by fluorescence spectroscopy at 405-nm excitation wavelength. **Journal of biomedical optics**, v. 12, n. 6, p. 064013, 2007.

ZHANG, S.; WU, X. Fundamentals of association rules in data mining and knowledge discovery. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, v. 1, n. 2, p. 97–116, 18 mar. 2011.