

CENTRO UNIVERSITÁRIO FEEVALE

FELIPE SCUCIATTO DOS SANTOS

UTILIZANDO REDES BAYESIANAS NA ANÁLISE DE IMAGENS
(Título Provisório)

Anteprojeto de Trabalho de Conclusão

Novo Hamburgo, março de 2009

FELIPE SCUCIATTO DOS SANTOS
scuciatto@feevale.br

UTILIZANDO REDES BAYESIANAS NA ANÁLISE DE IMAGENS
(Título Provisório)

Centro Universitário Feevale
Instituto de Ciências Exatas e Tecnológicas Curso de Ciência da
Computação
Anteprojeto de Trabalho de Conclusão

Professor Orientador: Marta Rosecler Bez

Novo Hamburgo, março de 2009

Resumo

Redes Bayesianas são uma representação gráfica do raciocínio probabilístico. Esta técnica é muito interessante em situações onde existem incertezas nas tomadas de decisões. Redes Bayesianas vêm sendo utilizadas com sucesso na medicina, para geração de diagnósticos, filtros antispam, etc. Uma área não muito explorada das Redes Bayesianas é a identificação de padrões em imagens. Este trabalho tem como objetivo explorar esta área, aplicando a técnica na análise de padrões em imagens. Pode-se utilizar uma Rede Bayesiana para analisar imagens em busca de uma série de padrões. Um exemplo seria a análise de lâminas microscópicas, proveniente de exames clínicos; análise de características para aplicação de filtros, entre outros. Sendo assim, este trabalho tem como objetivo estudar e aplicar as principais técnicas para elaboração e inferência de Redes Bayesianas, todas elas direcionadas à análise de imagens.

Palavras-chave: Redes Bayesianas. Reconhecimento de Padrões. Processamento de Imagens.

Sumário

Motivação	4
Objetivos	6
Metodologia	7
Cronograma	8
Referências	9

Motivação

As Redes Bayesianas são uma forma gráfica de representação das relações entre variáveis dentro de um escopo. É uma representação das probabilidades de um problema. As Redes Bayesianas são representadas por grafos acíclicos, onde cada nó é uma variável aleatória (CHARNIAK, 1991). Utilizando-se Redes Bayesianas, pode-se calcular a probabilidade condicional de cada variável em nosso problema, apenas observando valores de variáveis próximas. As Redes Bayesianas se tornaram uma maneira popular para a representação de conhecimento incerto em sistemas especialistas (HECKERMAN; CHICKERING, 1995).

Trabalhando com o raciocínio probabilístico, as Redes Bayesianas permitem a tomada de decisões mesmo sem o total conhecimento de um escopo. Ao ser alimentada com novos nós e probabilidades, as Redes Bayesianas tem a capacidade de "aprender" sobre um determinado escopo.

Conforme Heckerman (1996), as Redes Bayesianas oferecem pelo menos quatro vantagens sobre outros métodos de inferência (bases de regras, árvores de decisões, redes neurais artificiais, etc.).

Primeiro, Redes Bayesianas podem lidar com conjuntos incompletos de dados. Segundo, as Redes Bayesianas permitem o aprendizado de relações causais. Terceiro, Redes Bayesianas em conjunto com técnicas estatísticas, facilitam a combinação de conhecimento de domínio e dados. Quarto, métodos Bayesianos em conjunto com Redes Bayesianas, oferecem uma abordagem eficiente para evitar análises excessivas de dados.

As Redes Bayesianas vêm sendo utilizadas para a resolução de uma ampla gama de problemas. Algumas das áreas de maior uso são: diagnóstico médico (HECKERMAN, 1990), aprendizado de mapas (DEAN, 1989), processamento de linguagem (CHARNIAK; GOLDMAN, 1989). A Microsoft utiliza Redes Bayesianas no Office para detectar distrações do usuário e abrir o sistema de ajuda, observando o ambiente em que o usuário está trabalhando (HORVITZ et al., 1996).

Uma área pouco explorada no campo das Redes Bayesianas, é a identificação de padrões em imagens. Pode-se tomar como exemplo a área médica, onde a análise de imagens é muito difundida. Trabalhos recentes propuseram a análise de imagens para apoio

ao diagnóstico em parasitologia (NOGUEROL, 2008), porém o algoritmo proposto para a identificação de característica falhava em imagens poluídas ou com muita luminosidade.

As dificuldades encontradas na identificação de padrões em imagens, bem como a ausência de trabalhos correlatos nesta Instituição, motiva o desenvolvimento deste trabalho. Cremos que o uso de Redes Bayesianas é o método ideal para analisar características das imagens, identificar possíveis sujeiras e inferir sobre determinado ambiente. Observa-se também a necessidade de algoritmos de inferência mais precisos em trabalhos já realizados e sendo realizados nesta Instituição.

Este trabalho terá como objetivo desenvolver uma série de algoritmos para análise de características em imagens e, posteriormente, aplicar estes algoritmos na classificação dos dados extraídos por sistemas já existentes. Pode-se, com isso, complementar pesquisas já existentes, oferecendo uma análise mais precisa para características previamente extraídas.

Um dos principais desafios deste trabalho será a implementação de algoritmos de avaliação das redes. A análise de uma Rede Bayesiana consiste no cálculo da probabilidade condicional de cada nó, sempre levando em conta as evidências observadas. Geralmente a análise para resolução exata de uma Rede Bayesiana requer um algoritmo de complexidade NP-Hard (COOPER, 1990), isso, muitas vezes, pode levar a um tempo de execução muito longo.

Várias técnicas foram desenvolvidas com o intuito de otimizar a avaliação de Redes Bayesianas, elas consistem em encontrar aproximações das probabilidades condicionais de uma rede. Basicamente, estas técnicas definem valores aleatórios para alguns nós e, então, usam eles para a definição de valores dos nós filhos (CHARNIAK, 1991). Algumas técnicas ainda se baseiam no princípio da eliminação de nós ou na análise de apenas algumas partes da rede. São exemplos o uso da técnica chamada *cutset sampling* (BIDYUK; DECHTER, 2003), pode-se trabalhar com fatorização das probabilidades conjuntas (ZHANG; POOLE, 1994) e também pode-se usar uma técnica chamada *Bucket Elimination* (DECHTER, 1996).

Concluindo, este trabalho tem como objetivo aplicar Redes Bayesianas na identificação de características em imagens. Para obter êxito nesta tarefa, técnicas de criação de redes, bem como algoritmos de inferência das mesmas, serão estudados e implementados. Ao final, pretende-se aplicar esta técnica a trabalhos já existentes, melhorando assim sua capacidade de identificação de padrões em imagens.

Objetivos

O objetivo geral do trabalho é estudar técnicas para criação e inferência de Redes Bayesianas, aplicando as mesmas na descoberta de padrões em imagens. É parte deste objetivo sugerir uma técnica para resolver problemas encontrados em análise de imagens, como sujeiras que acabam comprometendo o resultado dos algoritmos aplicados.

Os objetivos específicos deste trabalho são:

- Revisão da bibliografia sobre Redes Bayesianas, técnicas de construção, aprendizagem e inferência sobre as mesmas.
- Implementar algoritmos estudados nas bibliografia, a fim de testar o funcionamento de cada um.
- Definir os casos em que o software atuará.
- Construir o protótipo do software baseado nas técnicas estudadas.
- Desenvolver cenários de testes para o software.
- Testar o software e analisar os resultados obtidos.

Metodologia

Inicialmente, será realizado um estudo das principais bibliografias na área. Posteriormente, os algoritmos, tanto de inferência como de aprendizagem serão testados, com o objetivo de definir o que melhor atende os problemas apresentados.

A partir dos testes e estudos realizados, um protótipo será desenvolvido.. A metodologia que se propõe prevê a elaboração da pesquisa através das seguintes etapas:

1. Revisar bibliografia sobre Redes Bayesianas.
2. Produzir o anteprojeto.
3. Revisar o trabalho junto ao orientador.
4. Entregar o anteprojeto.
5. Estudar outras aplicações que utilizam esta técnica.
6. Estudar e testar técnicas de criação de Redes Bayesianas.
7. Estudar e testar técnicas de inferência em Redes Bayesianas.
8. Estudar e testar técnicas de aprendizagem em Redes Bayesianas.
9. Produzir o texto do Trabalho de Conclusão I.
10. Revisar o trabalho junto ao orientador.
11. Entregar o Trabalho de Conclusão I.
12. Selecionar os casos em que software atuará.
13. Selecionar as técnicas à serem utilizadas no software (criação, inferência, aprendizagem), baseado nos testes feitos nas etapas 6, 7 e 8.
14. Escolher as plataformas de desenvolvimento do software.
15. Desenvolver o protótipo do software.
16. Criar um ambiente de testes para o software desenvolvido.
17. Testar e validar a implementação do software.
18. Avaliar os resultados.
19. Produzir o texto do trabalho de conclusão II.
20. Revisar o trabalho junto ao orientador.
21. Entregar o Trabalho de Conclusão.
22. Apresentar o trabalho à banca avaliadora.

Cronograma

Tabela 1: Trabalho de Conclusão I

Atividade	Mar	Abr	Mai	Jun	Jul
Revisar bibliografia sobre Redes Bayesianas	X	X	X	X	X
Revisar o trabalho junto ao orientador	X	X	X	X	X
Produzir o anteprojeto	X				
Entrega do anteprojeto		X			
Estudar temas referentes às Redes Bayesianas	X	X			
Produzir o texto do Trabalho de Conclusão I	X	X	X	X	X
Implementar os algoritmos estudados			X	X	X
Entregar o Trabalho de Conclusão I					X

Tabela 2: Trabalho de Conclusão II

Atividade	Ago	Set	Out	Nov	Dez
Revisar bibliografia sobre Redes Bayesianas	X	X	X	X	X
Revisar o trabalho junto ao orientador	X	X	X	X	X
Desenvolvimento do protótipo	X	X	X	X	X
Seleção de casos de uso para testes			X		
Avaliar o funcionamento do protótipo			X	X	
Produzir o texto do Trabalho de Conclusão II	X	X	X	X	
Entrega do trabalho de conclusão II					X

Referências

- BIDYUK, B.; DECHTER, R. Cycle-cutset sampling for bayesian networks. In: *In The Canadian AI Conference, (CAAI 03)*. [S.l.: s.n.], 2003. p. 297–312.
- CHARNIAK, E. Bayesian networks without tears. *AI Magazine*, 1991.
- CHARNIAK, E.; GOLDMAN, R. A semantics for probabilistic quantifier-free first-order languages, with particular application to story understanding. *Department of Computer Science, Brown University*, 1989.
- COOPER, G. F. The computational complexity of probabilistic inference using bayesian belief networks (research note). *Artif. Intell.*, Elsevier Science Publishers Ltd., v. 42, n. 2-3, p. 393–405, 1990.
- DEAN, T. Coping with uncertainty in map learning. *Department of Computer Science, Brown University*, 1989.
- DECHTER, R. Bucket elimination: A unifying framework for probabilistic inference. 1996.
- HECKERMAN, D. Probabilistic similarity networks, technical report, stan-cs-1316. *Computer Science and Medicine, Stanford Univ.*, 1990.
- HECKERMAN, D. *A tutorial on learning with Bayesian networks*. [S.l.], 1996.
- HECKERMAN, D. G. D.; CHICKERING, D. M. Learning bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 1995.
- HORVITZ, E. et al. The lumiere project: Bayesian user modeling for inferring the goals and needs of software users. *Microsoft Research*, 1996.
- NOGUEROL, L. *Sistema de Informação como Apoio ao Diagnóstico em Parasitologia*. [S.l.]: Trabalho de Conclusão de Curso - Feevale, 2008.
- ZHANG, N.; POOLE, D. A simple approach to bayesian network computations. In: *Proceedings of the Tenth Canadian Conference on Artificial Intelligence*. [S.l.: s.n.], 1994. p. 171–178.