

COMITÊS DE CLASSIFICADORES PARA DIAGNÓSTICO AUTOMÁTICO DE PATOLOGIAS DA COLUNA VERTEBRAL: UM ESTUDO COMPARATIVO

AJALMAR R. NETO

GUILHERME A. BARRETO*

**Depto. Engenharia de Teleinformática, Universidade Federal do Ceará (UFC)
Campus do Pici, Centro de Tecnologia, CEP 60455-760, CP 6005, Fortaleza-CE*

Email: {ajalmar, guilherme}@deti.ufc.br

Abstract— This paper reports a comparative study of learning strategies based on committee machines developed for the diagnostics module of the software platform **SINPATCO** - *System for INtelligent Diagnostics of PAThologies of the vertebral COlumn*. This module comprises an automatic pattern classifier and is evaluated by its ability in classifying a new patient as belonging to one out of three categories: Normal, Disk Hernia or Spondylolisthesis. First, the MLP and GRNN networks were evaluated individually. In sequel, committee machines composed of MLP and GRNN networks were evaluated. The obtained results suggest that the committee of classifiers presented better generalization performance than individual classifiers. In addition, MLP-based committee machines performed better than the GRNN-based one, obtaining higher recognition rates, higher robustness to outliers and lower false positive and false negative rates.

Keywords— Computer Aided Diagnostics, vertebral column, committee machines, pattern classification, robustness.

Resumo— Este artigo reporta um estudo comparativo de estratégias de aprendizado baseadas em máquinas de comitê implementadas para o módulo de diagnóstico da plataforma **SINPATCO** - *Sistema INteligente para diagnóstico de PATologias da COluna vertebral*. Este módulo é composto por um classificador automático de padrões, sendo avaliado quanto à capacidade de categorizar casos clínicos em uma das seguintes classes: Normal, Hérnia de Disco e Espondilolistese. Primeiramente, classificadores baseados nas redes MLP e GRNN foram avaliados individualmente, para em seguida, avaliar comitês de máquinas constituídos desses classificadores. Os resultados obtidos indicam que comitê de classificadores apresentam melhor generalização que os classificadores tomados individualmente. Entre as máquinas de comitê, o comitê de redes MLP apresenta melhor desempenho, i.e. maior taxa de acerto, maior robustez a outliers e menor número de falsos positivos e falsos negativos.

Keywords— Diagnóstico por computador, coluna vertebral, máquinas de comitê, classificação de padrões, robustez a outliers.

1 Introdução

Técnicas de aprendizado de máquinas (AM) têm sido amplamente utilizadas em diagnóstico médico assistido por computador ao longo dos últimos anos, com destaque especial àquelas que empregam redes neurais artificiais (RNAs) (Ramesh et al. 2004, Dybowski 2000). Um argumento recorrente nestes artigos, justificando o uso bastante difundido de RNAs na área médica, é o fato de estas serem capazes de resolver problemas complexos de classificação de padrões, construindo diretamente a fronteira de decisão a partir do conjunto de dados, não necessitando para isso fazer suposições sobre sua distribuição estatística.

De particular interesse para este trabalho é a aplicação de classificadores automáticos em ortopedia clínica. Nesta área da medicina, contudo, verifica-se que a aplicação de técnicas de classificação de padrões no diagnóstico de patologias da coluna vertebral não é amplamente difundida. A revisão bibliográfica feita neste artigo encontrou apenas dois trabalhos correlatos (Cherukuri et al. 2004, Antani et al. 2003), em que uma rede MLP é treinada para identificar a presença ou não de osteófitos com base no contorno da vértebra.

Mais recentemente, Rocha Neto et al. (2006) propõem a plataforma SINPATCO - *Sistema INteligente para diagnóstico de PATologias*

da Coluna vertebral, iniciativa pioneira na classificação automática de patologias da coluna vertebral. Para efetuar o diagnóstico, esta plataforma utiliza um conjunto de atributos biomecânicos propostos por Labelle et al. (2005), por estarem associados a dores e deformidades da coluna vertebral. Cada caso clínico em uma base de dados é descrito em termos de seus respectivos valores para cada um dos atributos biomecânicos de interesse. Usando esta base de dados numérica torna-se possível projetar classificadores de padrões para diversas patologias da coluna vertebral.

O módulo de diagnóstico da plataforma SINPATCO é composto por uma unidade de pré-processamento, responsável pela normalização dos dados e eliminação de *outliers*, e por uma rede MLP que é treinada com os casos clínicos previamente obtidos por um médico ortopedista. O módulo de diagnóstico pode ser então avaliado quanto a sua capacidade de categorizar novos casos clínicos em uma das seguintes classes: Normal, Hérnia de Disco e Espondilolistese. Implementações alternativas do módulo de diagnóstico da plataforma SINPATCO usando outros classificadores (KNN e Naive Bayes) foram comparadas com aquela baseada na rede MLP.

O principal objetivo deste artigo consiste em investigar o desempenho de comitês de classificadores compostos por redes MLP ou redes GRNN

(*Generalized Regression Neural Network*), com o intuito de melhorar o desempenho do módulo de diagnóstico da plataforma SINPATCO.

O restante do artigo está organizado da seguinte forma. Na Seção 2 as patologias de interesse para este artigo são descritas, assim como os atributos biomecânicos utilizados no projeto do classificador destas patologias são definidos. Na Seção 3 os classificadores e seus respectivos comitês avaliados são brevemente descritos. Os resultados da comparação de desempenho entre os classificadores são apresentados na Seção 4. O artigo é concluído na Seção 5.

2 Patologias da Coluna Vertebral

A coluna vertebral é um sistema composto por um conjunto de vértebras, discos intervertebrais, nervos, músculos, medula e ligamentos. Segundo Hall (2000), as principais funções da coluna vertebral são as seguintes: (i) eixo de suporte do corpo humano; (ii) protetor ósseo da medula espinhal e das raízes nervosas; e (iii) eixo de movimentação do corpo, possibilitando o movimento nos três planos: frontal, sagital e transversal.

Esse complexo sistema está sujeito a disfunções que causam dor nas costas, das mais variadas intensidades. Hérnia de disco e espondilolistese são exemplos de patologias da coluna vertebral que causam dores intensas. A hérnia de disco surge quando o núcleo do disco intervertebral migra de seu local, no centro do disco para a periferia, em direção ao canal medular ou nos espaços por onde saem as raízes nervosas, levando à compressão das raízes nervosas.

Espondilolistese ocorre quando uma das 33 vértebras da coluna vertebral desliza adiante em relação as outras. Este deslizamento quando verificado ocorre, geralmente, em direção a base da espinha na região lombar, ocasionando dor ou sintomatologia de irritação de raiz nervosa.

Para projetar classificadores automáticos capazes de discriminar várias patologias da coluna vertebral fez-se necessário a definição de um conjunto de atributos que pudessem ser medidos para cada paciente analisado pelo ortopedista. Assim, ao longo do tempo, o médico ortopedista pode dispor de um banco de dados de casos clínicos reais, contendo os valores dos atributos medidos a partir de pacientes portadores das mais diversas patologias, além de pacientes normais. Este banco de dados pode ser usado, na plataforma SINPATCO, para treinar e validar diversos classificadores.

2.1 Atributos Biomecânicos

Os atributos biomecânicos a serem descritos nesta seção foram propostos a partir de estudos desenvolvidos no Laboratório de Biomecânica do *Centre Médico-Chirurgical de Réadaptation des Mas-*

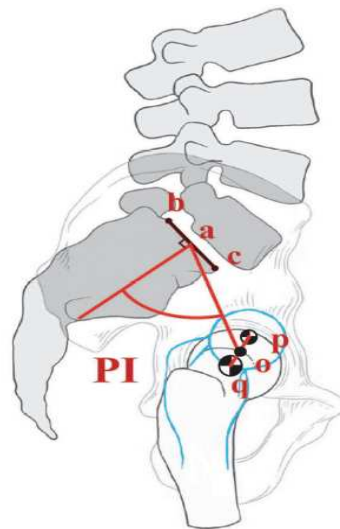


Figura 1: ângulo de incidência pélvica.

sues, localizado em Lyon, França. Os especialistas do Grupo de Pesquisa Avançada em Ortopedia (*Group of Applied Research in Orthopaedics, GARO*), montaram uma base de dados inicial com informações de pacientes saudáveis, escolióticos, com hérnia de disco e com espondilolistese (? , Labelle et al. 2005).

A base de dados gerada, gentilmente cedida pelo GARO para esta pesquisa, contém dados extraídos de 310 pacientes, a partir de radiografias panorâmicas sagitais em formato de 30×90 cm. Destes, 100 indivíduos são voluntários do Hospital de Massues que não possuem patologias na coluna. Os dados restantes são obtidos a partir de radiografias de pacientes operados de hérnias de disco (60 indivíduos) ou espondilolistese (150 indivíduos).

Cada indivíduo no banco de dados é representado como um vetor de seis atributos biomecânicos, que correspondem aos seguintes parâmetros angulares sagitais do sistema espinopélvico: (1) ângulo de incidência pélvica, (2) ângulo de versão pélvica, (3) ângulo de lordose, (4) declive sacral, (5) raio pélvico e (6) grau de deslizamento.

O ângulo de incidência pélvica (*pelvic incidence, PI*) é definido como sendo o ângulo subtendido pela reta \overline{aa} , que é extraído do centro da cabeça femoral para o ponto médio da placa sacral terminal, e uma reta perpendicular ao centro da placa sacral a (Figura 1). A placa terminal sacral é definida pela segmento de reta \overline{bc} entre o canto superior posterior do sacrum e a ponta anterior da placa terminal S1 no promontório sacral.

O ângulo de versão pélvica (*pelvic tilt, PT*), conforme indicado na Figura 2, é descrito como sendo o ângulo subtendido por uma reta de referência vertical originada do centro o da cabeça do fêmur e do raio pélvico \overline{oa} . Esta afirmação está correta quando o verdadeiro eixo hipotético está

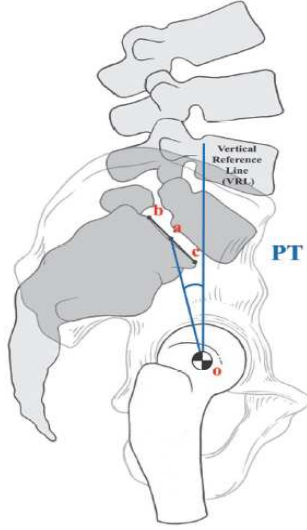


Figura 2: ângulo de versão pélvica.

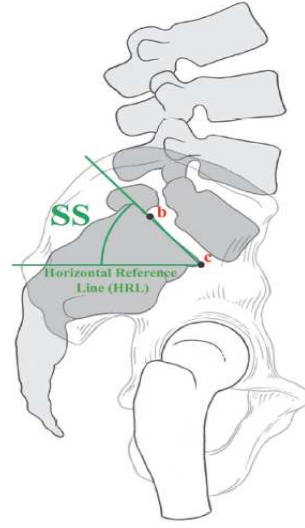


Figura 3: declive sacral.

em frente ao ponto médio da placa sacral terminal.

O declive sacral (*sacral slope*, SS) é definido como o ângulo subtendido por uma linha de referência horizontal (HRL) e pela linha sacral da placa terminal \overline{bc} , conforme ilustrado na Figura 3. O ângulo de lordose é o maior ângulo sagital entre o platô superior do sacro e o platô superior da vértebra lombar ou torácica limite. A distância ou raio pélvico é a distância do centro do eixo bicoxofemural ao centro do platô sacral. Este segmento de reta \overline{ao} , ilustrado na Figura 1. Por último, o grau de deslizamento é o grau percentual de deslizamento entre o platô inferior da quinta vértebra lombar e o sacro.

É importante destacar que o conjunto de atributos supracitados, quando analisados isoladamente, podem não indicar claramente uma relação causa-e-efeito entre certo atributo e a patologia observada. O diagnóstico final deve levar em consideração possíveis correlações, lineares ou não, existentes entre os atributos e como estas correlações influenciam no diagnóstico de uma certa patologia. O projeto de classificadores automáticos com base nos atributos biomecânicos de casos clínicos reais permite que correlações lineares e/ou não-lineares, assim com as influências destas no diagnóstico, sejam capturadas de forma transparente para o ortopedista, de forma a auxiliá-lo na tomada de decisão. A seguir são brevemente descritos os classificadores avaliados neste estudo.

3 Estratégias de Aprendizado

Inicialmente, o módulo de diagnóstico da plataforma SINPATCO é implementado por uma única rede MLP ou por uma única rede GRNN. Em seguida, comitês de classificadores formados por estas redes são também implementados.

A rede MLP escolhida contém uma camada de entrada de dimensão $p = 6$, uma camada escondida com q neurônios e uma camada de saída com $m = 3$ neurônios. As saídas da rede no instante t , $y_k(t)$, $k = 1, \dots, m$, são calculadas da seguinte forma:

$$y_k(t) = \varphi \left[\sum_{i=1}^q m_{ki}(t)v_i(t) - \theta_k \right] \quad (1)$$

$$v_i(t) = \varphi \left[\sum_{j=1}^p w_{ij}(t)x_j(t) - b_i \right] \quad (2)$$

em que w_{ij} é o peso sináptico ligando a j -ésima entrada ao i -ésimo neurônio da camada escondida, m_{ki} é o peso sináptico que liga o i -ésimo neurônio da camada escondida ao k -ésimo neurônio da camada de saída. Os termos b_i e θ_k correspondem aos limiares (*threshold*) de ativação dos neurônios da camada escondida e de saída, respectivamente. A função de ativação $\varphi(\cdot)$ escolhida para todos os neurônios da camada escondida e de saída foi a logística: $\varphi(u) = 1/(1 + \exp(-u))$. Os pesos e os limiares de ativação foram ajustados através do algoritmo *backpropagation* padrão por 2500 épocas, com taxa de aprendizagem igual a 0,01. O número de neurônios da camada escondida foi fixado em $q = 2p = 12$, valor este que se mostrou satisfatório.

O número de neurônios de saída foi fixado em $m = 3$, devido à codificação adotada para as saídas desejadas (rótulos numérico das classes). Foi escolhida uma codificação binária, na qual a saída desejada é igual a 1 (um) para o neurônio que representa a classe-alvo e 0 (zero) para as outras saídas desejadas. O rótulo numérico da classe de indivíduos normais é $\mathbf{r}_1 = [1 \ 0 \ 0]$, para classe de indivíduos com hérnia de disco é $\mathbf{r}_2 = [0 \ 1 \ 0]$, e para a classe com espondilolistese o rótulo é $\mathbf{r}_3 = [0 \ 0 \ 1]$. Esta codificação foi

adotada para todos os classificadores avaliados.

O classificador GRNN (Specht 1991) é um tipo de rede RBF (*Radial Basis Functions*), construída diretamente a partir dos dados disponíveis e de seus respectivos rótulos. A rede GRNN escolhida contém uma camada de entrada de dimensão $p = 6$, uma camada oculta com q funções de base radial do tipo gaussiana e uma camada de saída com $m = 3$ neurônios. As saídas desta rede no instante t , $y_k(t)$, $k = 1, \dots, m$, são dadas por

$$y_k(t) = \frac{w_{ki}\phi_i(\|\mathbf{x}(t) - \mathbf{c}_i\|)}{\sum_{l=1}^q w_{kl}\phi_l(\|\mathbf{x}(t) - \mathbf{c}_l\|)}, \quad (3)$$

em que $\phi(u) = \exp(-u^2/\sigma^2)$ denota a i -ésima função de base, \mathbf{c}_i é o centro da i -ésima função de base e w_{ki} é o peso que conecta a i -ésima função de base ao k -ésimo neurônio de saída. O parâmetro σ define o raio de abertura (*spread*) das funções de base.

Não há necessidade de treinar a rede GRNN no sentido neural clássico, basta definir um *kernel* gaussiano centrado em cada padrão de treinamento. O vetor de pesos ligando uma certa base gaussiana (camada oculta) aos neurônios de saída é uma cópia exata do vetor de saída desejadas associado ao padrão de treinamento usado como centro daquela base. As saídas desejadas são definidas da mesma forma que as usadas pela rede MLP.

3.1 Máquinas de Comitê

Diversas funções de fusão das saídas existem na literatura especializada visando a combinação de classificadores (Kittler et al. 1998, Xu et al. 1992, Kuncheva 2002). Estas comparam diretamente as saídas de todos os classificadores individuais em um comitê. Para este trabalho, foi escolhida a *Regra do Voto Majoritário Simples* (MAJ). Esta regra não requer as saídas à posteriori de cada classe, e cada classificador dá somente um voto para uma classe alvo. Então, a classe de saída atribuída ao comitê é inferida com base no número máximo de votos entre todas os classificadores.

Assim, seja \mathbf{x} um certo padrão de entrada que se deseja classificar. Considere também um comitê de L classificadores com saídas y_k , $k = 1, \dots, m$, treinados individualmente em um problema de m classes ($m = 3$). Seja $y_{k,l}$ a k -ésima saída do l -ésimo classificador, e $y_{k,l}^{(*)}$ a sua saída arredondada para “0” ou para “1”. Logo, o vetor de saídas arredondadas do l -ésimo classificador do comitê, para um dado padrão de entrada \mathbf{x} , é representada como

$$\mathbf{y}_l^{(*)}(\mathbf{x}) = [y_{1,l}^{(*)}(\mathbf{x}) \ y_{2,l}^{(*)}(\mathbf{x}) \ \dots \ y_{m,l}^{(*)}(\mathbf{x})]^T. \quad (4)$$

Seja $\mathbf{v}(\mathbf{x})$ o vetor resultante da soma dos vetores de saída $\mathbf{y}_l^{(*)}(\mathbf{x})$ de todos os classificadores

de um comitê

$$\mathbf{v}(\mathbf{x}) = \sum_{l=1}^L \mathbf{y}_l^{(*)}(\mathbf{x}). \quad (5)$$

Assim, pode-se inferir qual a classe $C(\mathbf{x})$ mais votada de um comitê de máquinas implementando-se a seguinte regra de decisão:

$$C(\mathbf{x}) = \arg \max_{k=1, \dots, m} \{v_k(\mathbf{x})\}, \quad (6)$$

em que $v_k(\mathbf{x})$ é o k -ésimo elemento do vetor $\mathbf{v}(\mathbf{x})$.

4 Resultados Obtidos

Antes do treinamento de todos os classificadores foram criados três conjuntos a partir do conjunto T de amostras disponíveis: um conjunto W livre de amostras discrepantes (*outliers*), um conjunto N contendo somente *outliers*, e um terceiro conjunto S que está relacionado ao primeiro e ao segundo conjuntos da seguinte forma $S = W \cup N_P$, em que N_P representa um subconjunto de N contendo uma porcentagem P dos *outliers* de N . Note que $T = W \cup N$ e que $S \subset T$. As amostras (padrões) de treinamento e teste são tirados do conjunto S .

Outliers são obtidos a partir do conjunto original de amostras como aquelas que apresentem atributos x_j com valores fora do intervalo $[\mu_j - 2\sigma_j, \mu_j + 2\sigma_j]$, em que μ_j e σ_j são a média e o desvio-padrão amostrais do j -ésimo atributo. Este subconjunto de amostras discrepantes formam o conjunto N . Eliminando-se estas amostras do conjunto completo T , obtém-se o conjunto W .

Metodologia de Treinamento/Teste: O treinamento e teste de cada classificador é executado 50 vezes, sendo que o conjunto S é dividido em dois subconjuntos: um para treinamento e outro para teste. As amostras que compõem o conjunto de treinamento são selecionadas aleatoriamente, sendo as restantes usadas para teste. Inicialmente, trabalha-se com um valor para $P = 0\%$, resultando em um conjunto S igual ao conjunto W , ou seja, livre de *outliers*. O valor de P é então incrementado em passos de 20 unidades, até atingir $P = 100\%$, quando então todos os *outliers* são adicionados ao conjunto S . Para cada grupo de 50 realizações de treinamento/teste, referentes a um valor de P especificado, são armazenadas as taxas de acerto médio, a variância das taxas de acerto, além da menor e a maior taxas observadas durante os testes.

Metodologia de Avaliação: Além das taxas de acerto médias, máximas e mínimas, os classificadores são também comparados entre si com base no número de falsos negativos e falsos positivos gerados. Um diagnóstico é denominado falso negativo quando um indivíduo doente (com espondilolistese ou hérnia de disco) é diagnosticado como saudável, enquanto um diagnóstico é

denominado falso positivo quando um indivíduo sadio é diagnosticado como doente. Estes números são comumente apresentados na forma de uma *matriz de confusão*. As linhas desta matriz correspondem às classes reais das amostras de testes, enquanto as colunas correspondem às classes preditas pelos classificadores. Para o presente estudo as matrizes de confusão geradas para cada classificador avaliado possui três linhas e três colunas. Se todas as amostras forem corretamente classificadas, a matriz de confusão correspondente terá elementos não-nulos apenas ao longo da diagonal principal.

Resultados - MLP: Para o classificador MLP, separou-se aleatoriamente 80% das amostras do conjunto S para treinamento e os restantes 20% são usadas no teste. Os valores médio, máximo, mínimo e a variância da taxa de acerto obtidos durante o teste, em função da porcentagem de *outliers* presentes nos dados são mostrados na Tabela 1. Conforme era de se esperar, o desempenho classificatório é máximo quando *outliers* não estão presentes nos dados (i.e. $P=0$), à medida que esta porcentagem aumenta o desempenho do classificador se degrada. Contudo, mesmo para $P = 100\%$, a taxa de acerto da rede MLP é elevada (acima de 90%).

$P(\%)$	Média	Mínimo	Máximo	Var
0	98.71	92.10	100.00	3.31
20	95.59	93.18	100.00	5.35
40	93.29	91.66	97.61	4.12
60	92.82	90.52	100.00	3.45
80	90.73	89.28	94.82	3.00
100	90.67	88.70	96.77	4.06

Tabela 1: Resultados do classificador MLP.

A Tabela 2 apresenta uma matriz de confusão típica do classificador MLP. Os resultados mostrados correspondem aos obtidos em uma das 50 rodadas de treinamento/teste, para $P = 100\%$. A escolha da rodada foi feita de forma aleatória. Para esta tabela e para as seguintes, os símbolos HD, EL e NO denotam indivíduos com hérnia de disco, com espondilolistese e normais, respectivamente. Analisando a Tabela 2, pode-se verificar que há 3 falsos positivos, 2 falsos negativos e um indivíduo com espondilolistese classificado como possuindo hérnia de disco.

	HD	EL	NO
HD	10	0	2
EL	1	24	0
NO	1	2	22

Tabela 2: Matriz de confusão (classificador MLP).

Resultados - GRNN: Este classificador é construído usando 70% das amostras do conjunto

S e os 30% restante compõem o conjunto de teste. A abertura (*spread*) da i -ésima função de base gaussiana foi selecionada como sendo igual a média das distâncias do centro desta função aos cinco centros mais próximos. Na Tabela 3 estão resumidos os resultados para este classificador. Pode-se concluir que o desempenho deste classificador foi inferior ao desempenho do classificador MLP independente da quantidade de *outliers* adicionada.

$P(\%)$	Média	Mínimo	Máximo	Var
0	96.14	87.93	100.00	8.92
20	89.44	83.08	96.92	12.46
40	83.50	76.39	90.28	15.48
60	80.38	65.82	89.87	18.59
80	77.32	66.30	86.05	22.92
100	75.10	59.14	82.79	23.94

Tabela 3: Resultados do classificador GRNN.

A Tabela 4 apresenta a matriz de confusão do GRNN, para um classificador selecionado aleatoriamente entre aqueles gerados durante as 50 rodadas de treinamento/teste ($P = 100\%$). Verifica-se, observando a tabela, a ocorrência de 12 falsos negativos, 7 falsos positivos e um indivíduo com hérnia de disco classificado como possuindo espondilolistese.

	HD	EL	NO
HD	11	1	3
EL	0	47	4
NO	3	9	15

Tabela 4: Matriz de confusão (GRNN).

Resultados - Comitê de MLPs: Este comitê é configurado com os 5 classificadores MLP de maior taxa de acerto no teste. Os testes de generalização do comitê de máquinas é feito aplicando-se os conjuntos de testes utilizados para verificar a generalização de cada classificador MLP. Assim, pode-se obter resultados relacionados aos 50 grupos de teste anteriormente obtidos. Na Tabela 5 estão resumidos os resultados. Pode-se concluir que o desempenho foi superior ao obtido com os classificadores individuais, com melhoria média em torno de 1%.

$P(\%)$	Média	Mínimo	Máximo	Var
0	100.00	100.00	100.00	0.0
20	96.04	90.90	100.00	5.44
40	93.87	89.58	97.92	4.50
60	93.73	88.68	100.00	5.00
80	91.44	89.47	96.49	3.47
100	91.35	87.09	96.77	4.22

Tabela 5: Resultados do Comitê de MLPs.

Resultados - Comitê de GRNNs: Da

mesma forma que o comitê de MLPs, um comitê de máquinas GRNN é configurado com os 5 classificadores GRNN de maior taxa de acerto no teste. Os testes de generalização do comitê de máquinas também é feito aplicando-se os conjuntos de testes usados para verificar a generalização de cada classificador GRNN. Na Tabela 6 estão resumidos os resultados. Pode-se concluir que o desempenho foi superior ao obtido com os classificadores individuais, com melhoria em torno de 2 – 6% em média. Pode-se verificar também uma redução significativa na variância em relação àquela para os classificadores individuais.

$P(\%)$	Média	Mínimo	Máximo	Var
0	98.69	96.55	100.00	1.28
20	93.26	87.69	98.46	7.33
40	88.61	77.78	94.44	10.16
60	86.10	78.48	94.97	12.98
80	82.86	75.58	89.54	13.73
100	81.38	73.12	88.17	11.58

Tabela 6: Resultados do Comitê de GRNNs.

Outro dado importante, é que a matriz de confusão relacionada ao comitê de máquinas GRNNs, quando aplicado ao mesmo conjunto de teste utilizado para montar a matriz de confusão da Tabela 4, apresenta uma redução em 1 para o número de falsos positivos e 1 para falsos negativos. Enquanto, a matriz de confusão relacionada com o comitê de máquinas MLP apresenta uma redução em 1 para o número de falsos positivos e a redução, também em um, do número de indivíduos classificados como possuindo hérnia discal quando estes deveriam ser classificados com espondilolistese (comparar com a Tabela 2). Foram avaliados também comitês com redes MLP com configurações diferentes, com 9 e 15 neurônios na camada escondida e até mesmo com duas camadas escondidas, porém os resultados não se apresentam melhores que os dos comitês com redes de 12 neurônios na camada escondida.

A partir dos resultados mostrados anteriormente, pode-se facilmente concluir que dentre os classificadores comparados, aquele baseado na rede MLP e o seu respectivo comitê foram os que apresentaram os melhores desempenhos classificatórios globais. O comitê de máquinas MLP mostrou-se com melhor capacidade de generalização em relação aos classificadores MLP analisados individualmente (apresentando uma melhoria média em torno de 1%). Por esta razão, o comitê de máquinas MLP pode ser utilizado para compor o módulo de diagnóstico automático da plataforma SINPATCO.

5 Conclusão e Trabalhos Futuros

Este artigo apresentou resultados referentes ao módulo de diagnóstico automático da plataforma **SINPATCO** - *Sistema INteligente de diagnós-tico de PATologias da COLuna vertebral*. Este módulo é avaliado quanto à capacidade de categorizar casos clínicos em uma das seguintes classes: Normal, Hérnia de Disco e Espondilolistese. As redes MLP e GRNN foram comparadas com os comitês de máquinas MLP e GRNN respectivamente, tendo os comitês apresentado melhor generalização quando comparados com os classificadores individuais. Entre as máquinas de comitê, o agregado de redes MLP foi melhor, pois apresenta as maiores taxas de acerto e maior robustez a outliers.

Agradecimentos

Os autores agradecem a CAPES/PRODOC pelo apoio financeiro.

Referências

- Antani, S., Long, L. R., Thoma, G. R. & Stanley, R. J. (2003). Vertebra shape classification using mlp for content-based image retrieval, *Proceedings of the IEEE-INNS International Joint Conference on Neural Networks (IJCNN'03)*, pp. 160–165.
- Cherukuri, M., Stanley, R. J., Long, R., Antani, S. & Thoma, G. (2004). Anterior osteophyte discrimination in lumbar vertebrae using size-invariant features, *Computerized Medical Imaging and Graphics* **28**(1-2): 99–108.
- Dybowski, R. (2000). Neural computation in medicine: Perspective and prospects, in H. Malmgren, M. Borga & L. Niklasson (eds), *Proceedings of the 1st Conference on Artificial Neural Networks in Medicine and Biology (ANNIMAB-1)*, pp. 26–36.
- Hall, S. J. (2000). *Biomecânica Básica*, 3a edn, Guanabara Koogan, Rio de Janeiro, RJ.
- Kittler, J., Hatef, M., Duin, R. & Matas, J. (1998). On Combining Classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **20**(3): 226–239.
- Kuncheva, L. (2002). A theoretical study on six classifier fusion strategies, *IEEE Trans. on Pattern Analysis and Machine Intelligence* **24**(2): 281–286.
- Labelle, H., Roussouly, P. & Berthonnaud, E. (2005). The importance of spino pelvic balance in L5 S1 developmental spondylolisthesis, *SPINE* **30**(6): 27–34.
- Ramesh, A. N., Kambhampati, C., Monson, J. R. T. & Drew, P. J. (2004). Artificial intelligence in medicine, *Annals of the Royal College of Surgeons of England* **86**(5): 334–338.
- Rocha Neto, A., Barreto, G. A., Cortez, P. C. & da Mota, H. (2006). Sinpatco: Sistema Inteligente para Diagnóstico de Patologias da Coluna Vertebral, *Congresso Brasileiro de Automática*, pp. 930–934.
- Specht, D. F. (1991). A general regression neural network, *IEEE Transactions on Neural Networks* **2**(6): 568–576.
- Xu, L., Krzyzak, A. & Suen, C. Y. (1992). Methods of combining multiple classifiers and their applications to handwriting recognition, *IEEE Transactions on Systems, Man, and Cybernetics* **22**(3): 418–435.