Proceedings of the VI Brazilian Conference on Neural Networks - VI Congresso Brasileiro de Redes Neurais
pp. 409 – 414, June 2–5, 2003 – Centro Universitário da FEI, São Paulo - SP - Brazil

Reconhecimento de Padrões através de uma Rede Neural Artifical Baseada em Mecanismos da Visão

Marco Aurélio Benedetti Rodrigues[1]; Iria Pedroso da Cunha Rodrigues[3],
José Marino Neto[2] & Fernando Mendes de Azevedo[1]
[1]Grupo de Pesquisas em Engenharia Biomédica - GPEB - Departamento de Engenharia Elétrica
[2]Laboratório de Neurofisiologia I - LNFI - Departamento de Ciências Fisiológicas
[3]Departamento de Ciências Morfológicas – MOR - CCB
Universidade Federal de Santa Catarina - UFSC
Florianópolis - Santa Catarina – SC
benedett@gpeb.ufsc.br, iria@ccb.ufsc.br, marino@ccb.ufsc.br, azevedo@ieb.ufsc.br

Abstract

While a variety of algorithms and techniques for pattern recognition has been developed, the vertebrate eye and visual system remains undisputedly considered as one of the better biological mechanism capable of accurate pattern discrimination and identification. In the present study, we advance an artificial neural network model, based on some relevant structural and functional characteristics of the human retina. This model was evaluated by assessing its capabilities in detecting K-complexes of human eletroencephalographic recordings, comparing its results with those of an medical specialist on this field to determine sensibility and specificity attributes of the system. Our results show that such an approach brings about a number of favorable attributes, as compared to other and more usual ones, including minimization of false-positive detection due to the double identification process by "foveal" and "extra-foveal" circuits.

Introdução

A área do reconhecimento de padrões e imagens é muito ampla e diversas ferramentas estão sendo desenvolvidas para este fim. Porém, o melhor sistema de reconhecimento é, sem dúvida, o olho humano. Embora, seja um mecanismo extremamente complexo, estudos sobre a fisiologia e morfologia do olho humano, podem levar a construção de mecanismos artificiais capazes de simular algumas de suas funções de processamento [1]. Este trabalho tem como objetivo a construção de uma Rede Neural Artificial Baseada em Movimentos da Visão (RNABMV), com base na disposição geográfica e funcional dos neurônios da retina do olho o humano.

O retina do olho humano apresenta duas regiões morfologicamente e funcionalmente distintas de grande importância no processo de identificação da imagem: a retina periférica e a fóvea. A retina periférica contém circuitos de alta sensibilidade para sinais de pequenas amplitudes e baixo poder de discriminação de padrões, realizando a análise rápida da imagem e retirando apenas características mais relevantes. A fóvea é considerada uma área de menor sensibilidade para sinais de pequena amplitude, porém contém mecanismos com um alto poder de discriminação de padrões (contornos e bordas) e é capaz de identificar características específicas de cada imagem.

O processo de identificação se inicia quando a imagem, ao entrar no globo ocular, é enviada para retina e, posteriormente, posicionada através de movimentos oculares sobre a fóvea para a identificação das suas características específicas. O cérebro atua como elemento controlador fazendo com que a imagem “alvo” ou raio luminoso passe pela retina periférica (realizando uma primeira identificação) e, posteriormente, a nível da fóvea, são identificadas as características específicas de cada imagem, este mecanismo pode ser analisado na Figura 1.

Na retina periférica, distinguem-se três grandes camadas de células: a primeira camada composta de *células fotossensíveis* (ou fotorreceptoras) denominadas de cones e bastonetes, cuja a função é captar estímulos luminosos e convertê-los em potenciais elétricos; a segunda camada é composta por células que possuem a finalidade de captar os estímulos provenientes das células fotossensíveis, constituída de neurônios bipolares e por último uma camada composta por células ganglionares, que transmitem os estímulos visuais pelo nervo óptico até o tálamo visual e posteriormente ao córtex cerebral, onde são realizados os processos de identificação da imagem [4].

Os bastonetes são estruturas adaptadas para identificação de sinais com pouca intensidade luminosa,

enquanto que, os cones são adaptados para identificação de sinais com maior intensidade. Nas porções periféricas da retina predominam os bastonetes, enquanto que o número de cones aumentam progressivamente a medida que se aproxima da fóvea, até que, ao nível da fóvea, existam exclusivamente cones. A característica principal do potencial dos fotorreceptores da retina é o seu comportamento aproximadamente logarítmico da intensidade luminosa. Sendo de extrema importância, porque permite a discriminação de intensidades em uma faixa muito mais ampla do que seria possível sem esse pré-processamento [11][12].

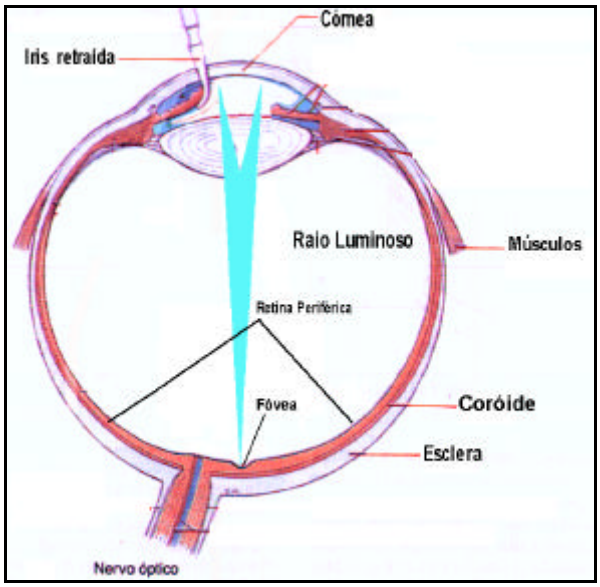

Figura 1 – Raio luminoso direcionado para fóvea através dos movimentos oculares

Entre a primeira e a segunda camada da retina surgem as células horizontais, conectando-se entre os corpos sinápticos dos bastonetes e cones, bem como com os dendritos das células bipolares. A saída das células bipolares são inibitórias, e estas conexões horizontais parecem ser responsáveis pelo fenômeno da inibição lateral, importante como mecanismo de intensificação de contraste (bordas), sendo que, o mesmo ocorre entre a segunda e última camada, onde o mesmo processo acontece com as células amácrinas [12].

A terceira camada composta pelas células ganglionares, conjuga as informações das camadas neuronais anteriores e transmite informações sobre a imagem para o Sistema Nervoso Central (SNC), através de estímulos seqüenciais. A Figura 2 apresenta um diagrama simplificado da retina do olho humano.

A proposta deste trabalho é apresentar uma modelagem de uma Rede Neural Artificial (RNA) baseada nos mecanismos da retina do olho humano. A validação deste modelo se dará através da identificação de complexos K do sono e comparação com resultados obtidos por médicos especialistas na área de polissonografia.

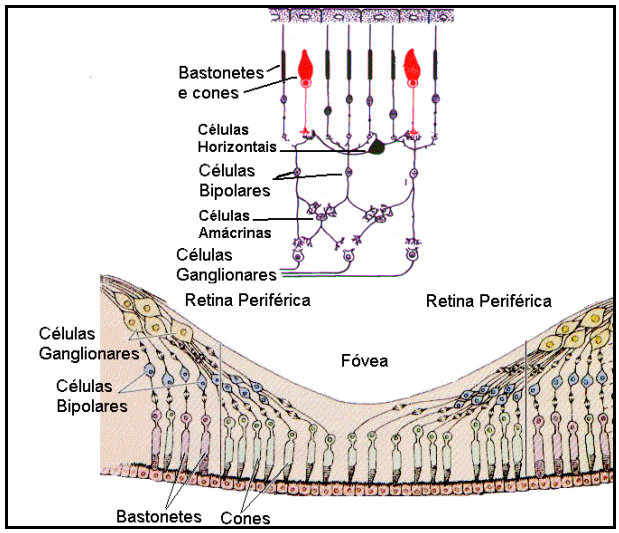

Figura 2 – Micrografia da retina, apresenta a relação neuronal da retina periférica e fóvea central do olho

Modelo da Rede Neural

Os mecanismos de identificação de uma imagem ou padrão, realizado através da cadeia neuronal da retina, pode ser interpretado e modelado de forma simplificada através de um conjunto de RNAs. O modelamento da RNABMV aqui proposto possui, da mesma forma que o olho humano, duas fases distintas na identificação de padrões (dois tipos de RNAs). A primeira RNA, a da retina periférica, identifica algo "semelhante" ao padrão alvo, enquanto que a segunda RNA (RNA da fóvea - RNAF) realiza uma identificação mais apurada deste padrão. A Figura 3, apresenta a estrutura geral da RNABMV.

O córtex cerebral e o mecanismo de controle do globo ocular apresentado na Figura 3, não serão exaustivamente modelados, apenas possuem algoritmos para a identificação dos sinais provenientes das RNAs e controles de parâmetros para o deslocamento do sinal da retina periférica para o fóvea. A informação de saída da RNA da retina periférica se modifica de acordo com a dinâmica dos padrões de entrada, produzindo no decorrer do tempo um conjunto de valores na saída da célula ganglionar desta rede, produzindo um sinal denominado de Vetor de Saída da Rede Neural (VSRN). A análise das variações temporais do VSRN é o indicador da presença ou não de um determinado padrão. Conforme este é reconhecido, o modelo do córtex cerebral e globo ocular enviam o padrão de entrada para a RNA da fóvea para a verificação de características mais específicas deste sinal.

O modelo proposto na Figura 3, possui uma retina periférica, sub dividida em duas (a retina periférica direita e a retina periférica esquerda) e uma fóvea (da mesma forma que o olho humano). Porém, o funcionamento das RNA da retina são os mesmos, se algum padrão for reconhecido por quaisquer uma das

RNA este é deslocado para a fóvea para a uma identificação mais apurada.

Os elementos denominados Z^{-1} representam um atraso no tempo para os padrões de entrada da RNABMV, portanto, os padrões de entrada são apresentados à RNABMV de acordo com o passar do tempo e não apenas uma só vez em um instante de tempo definido, configurando a estrutura de RNA Dinâmica (RNAD), o que facilita no aprendizado e ganho de velocidade no reconhecimeto [7][8][9][10].

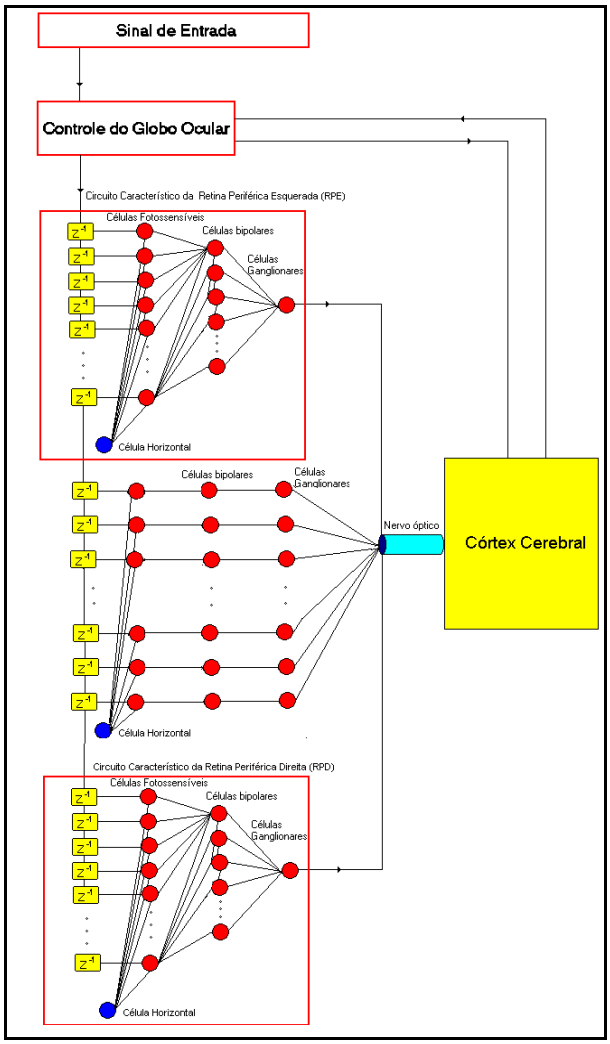

Figura 3 – Modelo da RNABMV

O circuito da RNA da retina periférica é composto de um grid de 45x18 células fotorreceptoras (do tipo bastonetes). Os bastonetes são compostos de função de ativação logarítmica, com a finalidade de tornar a faixa de entrada diferenciada de acordo com o nível do sinal nestas células. Portanto, para todas entradas neuronais ($E_{i,j}$) da rede existe um procedimento logarítmico para se obter a saída ($S_{i,j}$), conforme a equação 1.

$$S_{i,j} = \log (E_{i,j}) \qquad (1)$$

Após, é realizado um processo de reforço e inibição lateral, simulando o função das células horizontais do olho humano. Estas possuem a finalidade de modular os padrões neuronais de forma a tornar a saída do neurônio ganglionar mais ou menos ativa de acordo com a posição dos sinais e níveis da camada de entrada da rede. Neste modelo, foram utilizadas células horizontais conectadas lateralmente com todos os outros neurônios da entrada. Portanto, é realizado um processo de reforço para o neurônio "alvo" e um outro de inibição lateral atribuído a vizinhança deste neurônio, conforme pode ser visto na equação 2, onde $S_{i,j}$, representa a vizinhança a ser verificada relativa ao neurônio $S_{k.l}$, com raio (r) igual a 4, isto é, valores para k e l igual a: 0,±1, ±2, ±3, ±4. A constante E1 é o nível de ativação do neurônio e E2 a constante inibitória para os neurônios da vizinhança mais próxima.

$$S_{i,j} = \begin{cases} 1)\ S_{i,j} + S_{k,l} * E1 \quad se \quad i = k\ e\ j = l \\ 2)\ S_{i,j} - S_{k,l} * E2 \quad se \quad \begin{cases} i = k\ ou\ i = k+1\ ou\ i = k-1 \\ e \\ j = l\ ou\ j = l+1\ ou\ j = l-1 \end{cases} \\ S_{i,j} \quad se\ não\ for\ 1)\ ou\ 2) \end{cases} \qquad (2)$$

O processo de inibição lateral realizado pela células amácrinas (Figura 2) não será modelado em nosso sistema devido ao fato da implementação do modelo possuir somente uma RNAD na retina periférica, com uma única saída através do neurônio ganglionar.

A camada intermediária da retina periférica é composta de um grid de 15 x 6 neurônios (neurônios bipolares polissinápticos), que se interligam a cada grupo de 9 células fotorreceptoras da camada de entrada, Figura 4. Já a interligação da camada intermediária com a de saída é realizada com todos os neurônios interconectados a um único neurônio de saída (célula ganglionar), que envia as informações para o Sistema Nervoso Central (SNC). Conforme pode ser analisado no esquema da Figura 3.

Quando um sinal é identificado pelo córtex cerebral, este ativa um mecanismo semelhante ao globo ocular do olho humano, enviando o padrão de entrada da RNAD da retina para a RNAF, que realiza uma segunda identificação mais criteriosa deste padrão. Isto se assemelha aos movimentos realizados pelos olhos para focalizar um determinado objeto que se deseja identificar.

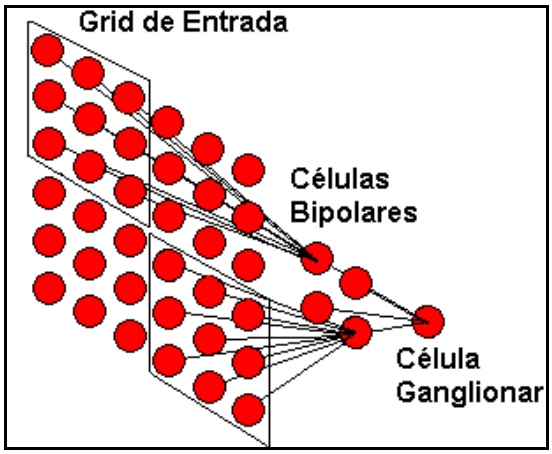

Figura 4 – Estrutura Neuronal da Rede da Retina periférica

Já na RNA da fóvea (RNAF) o padrão de entrada ($E_{i,j}$) passa por uma função de ativação do tipo exponencial, reforçando os sinais de maior amplitude e atenuando os de menores amplitude, conforme a equação 3.

$$S_{i,j} = \text{Exp}\ (E_{i,j}) \qquad (3)$$

No modelo da RNAF, foi utilizado um grid de 18x18 elementos. Aos quais foram aplicados o mesmo processo de inibição laterial causado pelas células horizontais. Porém, cada elemento é interligado diretamente a um único neurônio, em cada camada respectiva, conforme a figura 5.

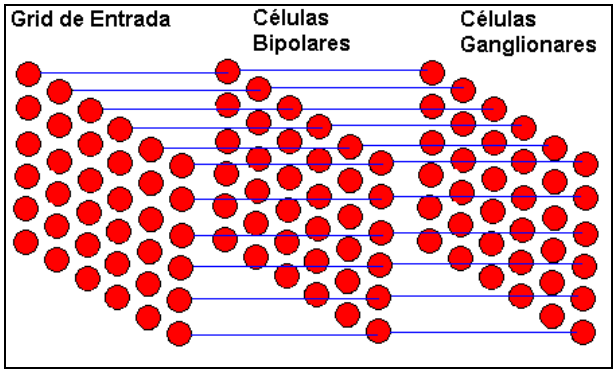

Figura 5 – Estrutura Neuronal da RNA da Fóvea

Método de Treinamento das RNAD

O conjunto de padrões de treinamento para a análise e validação dos resultados utilizado pelo sistema é de um constituído conjunto de registros de EEG durante o sono que possuem no seu conteúdo complexos K do sono. Estes são previamente selecionados por um especialista da área para a posterior comparação dos resultados com os da RNABMV. Os vetores de entrada (sinais de EEG) foram sub-amostrados de tal forma que o valor máximo de amplitude máximo corresponda a um valor igual a 18, isto é, número de neurônios utilizado para a máxima amplitude dos grids de entrada das RNAs. A Figura 6A apresenta um complexo com características normais com amplitude máxima igual a 18 (composto de números reais), na Figura 6B o mesmo complexo K é apresentado, porém sub-amostrado, isto é, possui somente elementos com valores inteiro entre 0 e 18. Nota-se que a diferença entre os 2 tipos de sinais ocorre somente em relação as componentes de alta-freqüência, as quais não são de grande importância para a identificação deste tipo de padrão.

Cada padrão é constituído de noventa pontos, o que representa uma duração de aproximadamente um segundo. Este tempo foi escolhido levando em consideração o tempo mínimo de duração de um complexo K do sono, que é de aproximadamente a metade do tempo escolhido [2][6]. Os mesmos registros de EEG foram escolhidos como padrões para o treinamento, tanto para a retina periférica quanto para a fóvea. Este tipo de sinal foi selecionado por ser o modelo de um sinal dinâmico, que, por definição é composto de variáveis funcionais e temporais [10][5], proporcionando um conjunto de treinamento constituído de uma combinação de freqüências e amplitudes provenientes de diversos locais do cérebro. Normalmente, este tipo de sinal é de difícil identificação, pois alguns tipos podem se em função de estímulos externos ou internos do organismo [3].

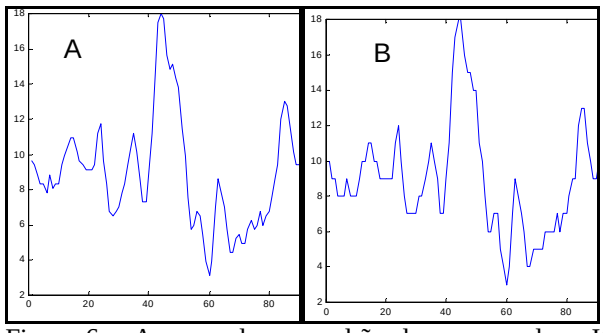

Figura 6 – A esquerda um padrão de um complexo K normal e a direita o mesmo padrão, porém, sub-amostrado.

Após o a sub-amostragem o padrão de fuso do sono é enviado para a primeira camada da RNAD, onde ocorre o processo de inibição lateral causado pelas células horizontais. Um exemplo de complexo K obtido após esse etapa pode ser visualizado através da figura 7.

Na Figura 8, pode-se observar o mesmo padrão da figura 7, porém em 2 dimensões. O processo de inibição lateral realizado pela camada de entrada é realizado em ambas as redes neurais (retina periférica e fóvea). Para o treinamento completo da RNAD e RNAF foram utilizados 5 padrões de complexos K, que obviamente, não foram utilizados na etapa de validação e teste dos resultados das RNAs.

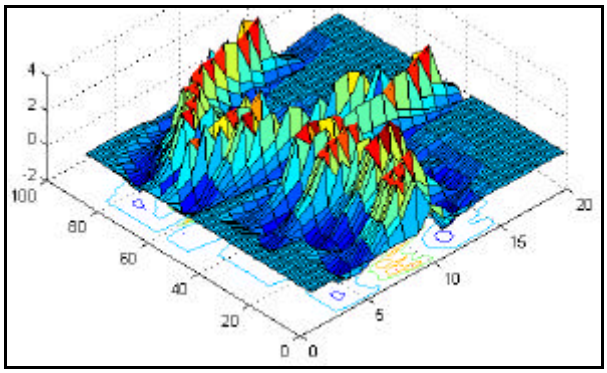

Figura 7 – Padrão de um complexo K em 3D, após o processo de inibição lateral

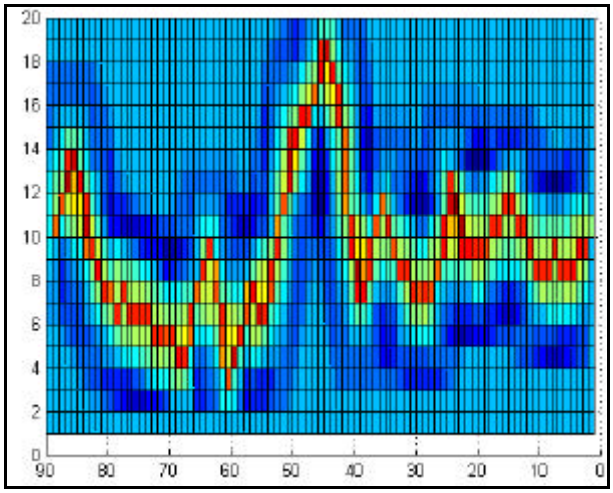

Figura 8 – Padrão de um complexo K em 2D após o processo de inibição lateral

O treinamento para os dois tipos de RNA foi realizado com base em um algoritmos baseado no *backpropagation*, onde cada atualização dos pesos é realizada conforme a entrada de cada conjunto de valores digitalizados, isto é, para cada "coluna" do padrão ao entrar na malha de atraso é realizado um treinamento pela rede [8]. Os valores de saída da rede neural se modificam a medida que os padrões de entrada percorrem a malha de atraso (Z^{-1}). Quando um determinado padrão é encontrado a saída da RNA começa a se estabilizar, mantendo um segmento constante com um valor igual a 1 (padrão alvo para o treinando de ambas as redes neurais).

Toda etapa de treinamento foi realizada com uma taxa de aprendizado de 0,01 até atingir um valor de erro médio quadrático menor que 0,02, em relação ao padrão alvo de treinamento (segmento com valor unitário).

Resultados e Conclusões

Os olhos são órgãos fotossensíveis complexos que atingem alto grau de evolução, permitindo uma análise minuciosa quanto à forma dos objetos, cor e a intensidade de luz refletida. A rede aqui proposta não possui finalidade de simular todo o comportamento da retina do olho humano, mas sim demonstrar que existe vários mecanismos de pré-processamento nos sinais luminosos captados pela retina que podem ser utilizados no auxílio da detecção de padrões em uma RNA.

A identificação do resultado é realizada pelo análise do VSRN, levando em consideração a existência ou não de um segmento do VSRN com um valor constante próximo a 1(um), que não ultrapasse um determinado valor de erro. Caso esse segmento exista pode-se dizer que a RNAD encontrou um determinado tipo de padrão treinado. E que, posteriormente, o córtex cerebral e o globo ocular o deslocarão até a fóvea para ser analisado mais criteriosamente. Com base nos resultado da retina periférica nota-se que ao aplica-la já existe um determinado índice de acerto no reconhecimento dos padrões de complexos K do sono. E, que a rede neural da fóvea serve para minimizar possíveis problemas de identificação, como falsos positivos que por ventura venham a surgir na primeira identificação. Embora, a retina periférica identifique vários padrões de EEG, ela não adquire informações mais específicas do padrão. Por outro lado, a fóvea é bem mais especializada na detecção de bordas (detalhes) devido ao maior número de neurônios de saída. A Figura 9, apresenta o VSRN da análise de um complexo K do sono, após ser analisado pela retina periférica.

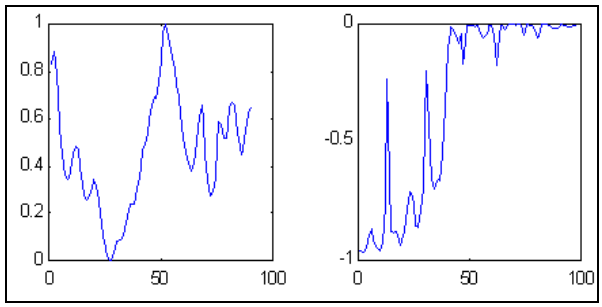

Figura 9 – A esquerda um complexo K do sono e a direita um VSRN respectivo à identificação

Os resultados obtidos pela RNA da fóvea são comparados, da mesma forma que os obtidos pela retina periférica. Na fóvea, tem-se diversos VSRN correspondentes a cada neurônio de saída desta RNA. Porém, a análise em conjunto de todos os VSRN proporcionam uma maior confiabilidade no resultado da identificação final do padrão.

O modelo da RNDBMV possui uma velocidade de identificação relativamente elevada, mesmo que o processamento a nível da fóvea seja bem mais lento devido a quantidade de neurônios empregada neste circuito.

Como teste e validação final da RNABMV, foi utilizado um conjunto de cinco registros completos de EEG durante o sono, de pacientes diferentes. Após a análise da RNABMV dos registros de EEG, a procura de complexos K, os padrões encontrado foram comparados com os encontrados por um especialista da área de polissonografia. A partir dos resultados obtidos foi calculado a sensibilidade e a especificidade do sistema como um todo. A tabela 1 apresenta os resultados desta análise onde: N representa o número de

registros, VP os valores verdadeiramente positivos, VN os valores verdadeiramente negativos, FP os falsos positivos, FN os falsos negativos, S a sensibilidade da RNABMV e E a especificidade da RNABMV para este registro.

N	VP	VN	FP	FN	S	E
01	38	155	8	16	0,70	0,95
02	41	152	27	6	0,87	0,85
03	14	121	5	6	0,70	0,96
04	10	162	15	4	0,71	0,91
05	52	131	21	13	0,80	0,86

Tabela 1 - Resultados encontrados na análise de cinco registros de EEG pela RNABMV.

Testes ainda estão sendo realizados com este modelo e confirmam um bom índice de acerto, comparado com estudos anteriores baseados em redes neurais semelhantes [8][9][10]. A rede neural proposta neste artigo faz parte de um sistema completo para detecção de apnéias durante o sono que vem sendo desenvolvido pelo Instituto de Engenharia Biomédica da Universidade Federal de Santa Catarina.

Referências Bibliográficas

[1]-FUCUSHIMA, K. A neural network for visual pattern recognition. IEEE Computer, v. 21, n. 3, p. 65-75, 1988.

[2]-Kelly, D. D.; “Sleep and Dreaming/Disorders of Sleep and Consciousness”. In: Kandel, E. R.; Schwartz, J. H. & Jessell, T. M. Principles of Neural Science, 3 ed. London: Prentice-Hall International Inc., 1991. p. 792-819.

[3]-Loomis A.L., Harvey E.N, Hobart G.A., “Electrical potentiais of the humam brain”, J. Exp. Psychol, 19:249-279, 1936.

[4]-Mason, Carol; Kandel E.R. “Central Visual Pathways”. In: KANDEL, E. R.; SCHWARTZ, J. H. & JESSELL, T. M. Principles of Neural Science, 3 ed. London: Prentice-Hall International Inc., 1991. p. 420-439.

[5]-Oliveira, Roberto C. L. & F. M. de Azevedo & J. M. Barreto. “Dynamic Neural Net in the Space of States Utilized in Non-Linear Processes Identification”. ICANNGA’97 - 3rd. International Conference on Artificial Neural Networks and Genetic Algorithms, Norwich, UK, April 1-4, 1997, Proceedings pp: 588.

[6]-Rechtschaffen, A., Kales, A. “A manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects”, Brain Information Service, Brain Research Institute, UCLA, Los Angeles, 1968.

[7]- Rodrigues, M.A.B., De Azevedo; F.M. & Neto; J.M., Um Modelo de Rede Neural Baseado na Circuitaria da Retina para Detecção de Padrões Eletroencefalográficos, Proceedings of the V Brazilian Conference on Neural Networks - V Congresso Brasileiro de Redes Neurais,pp. 361–366, April 2–5, 2001 - PUC, Rio de Janeiro - RJ – Brazil.

[8]-Rodrigues, M.A.B., Rodrigues, Iria P. C., De Azevedo; F.M., Neto; J.M., Neural network based in the attributes of extrafoveal and foveal circuitry from the human retine, European Medical And Biological Engineering Conference (EMBEC 2002), Dezembro de 2002, Viena – Áustria.

[9]-Rodrigues,M.A.B., De Azevedo; F.M. & Neto; J.M., “Sistema para Identificação de Padrões do Sono com o auxílio de Redes Neurais Dinâmicas”. IV Congresso Brasileiro de Redes Neurais e V Escola de Redes Neurais, São José dos Campos – Brasil, Julho 1999).

[10]-Rodrigues,M.A.B., De Azevedo; F.M., Neto; J.M. & Andriani, V.M., “Dynamic Neural Network for Recognition of Sleep Spindles and K Complex”, European Medical And Biological Engineering Conference (EMBEC 99), Viena – Áustria, Novembro de 1999.

[11]-Sobotta. Histologia: Atlas colorido de citologia, histologia e anatomia microscópica humana. 5° ed., Ed. Guanabara Koogan, Rio de Janeiro, 1999.p. 217.

[12]-Tessiere, Marc Lavigne. “Phototrasnduction and Information Processing in the Retine”. In: Kandel, E. R.; Schwartz, J. H. & Jessell, T. M. Principles of Neural Science, 3 ed. London: Prentice-Hall International Inc., 1991. p.400-418.