

Utilização de Redes Neurais do tipo Extreme Learning Machine na classificação da cobertura de solo do Município de Novo Progresso-PA

Helder Araujo Louzada¹
Ana Carolina Quintão Siravenha¹
Evaldo Gonçalves Pelaes¹

¹ Universidade Federal do Pará – UFPA
Av. Augusto Correa, 01 – 66075-110 – Belém – Pa, Brasil
helder.louzada@itec.ufpa.br, {siravenha,pelaes}@ufpa.br

Abstract. The municipality of Novo Progresso, in Brazil, has been a frequent topic in discussions about environmental issues because having one of the highest deforestation rates in the Amazon region. These high rates of deforestation has a direct relationship with the development of policies implemented in the region over the years. In this context, the aim of this work is to develop a supervised architecture for image classification by Artificial Neural Networks (RNA), using the extreme learning machine algorithm (ELM). This learning machine has a set of parameters able to identify the characteristics of deforestation over the years and, consequently, the different features of use and land cover in remote sensing images. The acquired scenes, from the satellite Landsat-5 TM, represent part of the municipality of Novo Progresso, through the period 1984 to 2011. The results were evaluated through the confusion matrices, where the coefficients quantify the correctness of the classes, the Kappa index and the index of overall accuracy, beyond the Mean Square Error. Although the methodology effectiveness, there are some features to be considered in future works, mainly considering the influence of the mid infrared during the feature vector formation.

Keywords: image classification, extreme learning machine, artificial neural networks classificação de imagens, máquina de Aprendizado extremo, Redes neurais artificiais

1. Introdução

O município de Novo Progresso, situado no estado do Pará, tem sua história de ocupação associada às políticas públicas nacionais direcionadas para região Amazônica a partir da década de 1970 e, nos últimos anos, tem novamente atraído a atenção dos governantes e instituições no que se refere à execução de ações de fiscalização, para o combate e controle aos altos índices de desmatamento registrados. O município teve seu surgimento atrelado à construção da rodovia BR 163 (Cuiabá-Santarém), que ao interligar o planalto central cortando o interior da floresta amazônica visa materializar uma nova frente de integração, consolidando os moldes de progresso propostos à época. A partir de 1984, a economia e a dinâmica populacional da localidade de Progresso tomam outro direcionamento devido à descoberta de ouro e outros minérios que atraem milhares de pessoas à localidade e que contribui por fortalecer a ideia de emancipação concretizada em 1991 (CASTRO; MONTEIRO; CASTRO, 2002).

O objetivo deste trabalho é desenvolver uma arquitetura de *Feedforward Neural Network*, utilizando o algoritmo Extreme Learning Machines (ELM), que possua um conjunto de parâmetros capazes de identificar as características de desmatamento ao longo dos anos e as diferentes feições de uso e cobertura do solo das imagens de sensoriamento remoto adquiridas do satélite Landsat-5 TM, através de um mapa de classificação de uso do solo referente ao município de Novo Progresso no período de 1984 à 2011.

2. Área de Estudo

A área de estudo abordada neste trabalho pertence ao município paraense de Novo Progresso, cuja área urbana está localizada às margens da Rodovia BR 163 e está localizada a uma latitude $07^{\circ}08'52''$ Sul e a uma longitude $55^{\circ}22'52''$ Oeste, possui $38.162,371 \text{ km}^2$ de área total, e em 2010 apresentava uma população de 25.124 habitantes (IBGE, 2010). Novo Progresso, assim como outros municípios do estado do Pará, tem sofrido um rápido processo de mudança da cobertura da terra, especialmente a partir do início da década passada. Segundo Escada et al. (2009), Novo Progresso está entre os dez municípios que mais tiveram áreas desmatadas na Amazônia Legal em 2008.

3. Redes Neurais Artificiais

Uma rede neural artificial (RNA) é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que tem propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso. Semelhante ao cérebro, o conhecimento é adquirido pela rede a partir de seu ambiente, por meio de um processo de aprendizagem e graus de conexão entre neurônios, conhecidas como pesos sinápticos (HAYKIN, 2001).

Uma rede neural artificial pode ser entendida como uma máquina projetada para modelar a maneira como o cérebro realiza uma tarefa particular, normalmente a rede é implementada utilizando-se componentes eletrônicos ou simulada em um ambiente computacional através de programação. Uma rede neural é um sistema constituído por unidades conhecidas como neurônios, elementos processadores interligados, trabalhando em paralelo para alcançarem um bom desempenho em determinada tarefa. As principais áreas de atuação são previsão e classificação de padrões, área essa referente ao estudo deste trabalho.

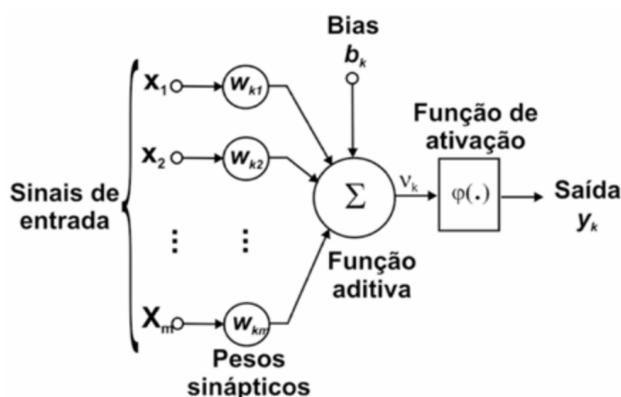


Figura 1: Modelo não-linear de um neurônio (HAYKIN, 2001).

O diagrama em blocos da Figura 1. Mostra o modelo de um neurônio onde podemos identificar três elementos básicos do modelo neural:

1. Um conjunto de sinapses ou elos de conexão, caracterizada por um peso ou força própria. Um sinal x_j na entrada da sinapse j conectada ao neurônio k é multiplicado pelo peso sináptico w_{kj} . O primeiro índice se refere ao neurônio em questão e o segundo se refere ao terminal de entrada da sinapse à qual o peso se refere.

2. Um somador para unir os sinais de entrada, ponderado pelas respectivas sinapses do neurônio; as operações descritas constituem um *combinador linear*.

3. Uma *função de ativação* para restringir a amplitude da saída de um neurônio. Conhecida também como função restritiva já que limita o intervalo permissível de amplitude do sinal de

saída a um valor finito. O intervalo normalizado da saída de um neurônio é escrito como o intervalo unitário fechado $[0, 1]$ ou no caso deste trabalho $[-1, 1]$.

O modelo neural inclui também um *bias* aplicado externamente, representado por b_k . O bias b_k tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo, respectivamente. Matematicamente podemos descrever um neurônio k escrevendo o seguinte par de equações 1 e 2:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (1)$$

$$y_k = \varphi(u_k + b_k) \quad (2)$$

Onde x_1, x_2, \dots, x_m , são os sinais de entrada; $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos sinápticos do neurônio k ; u_k é a saída do combinador linear devido aos sinais de entrada; b_k é o bias; $f(\cdot)$ é a *função de ativação* e y_k é o sinal de saída do neurônio. O uso do bias b_k tem o efeito de aplicar uma transformação afim à saída do u_k do combinador linear, como mostrado pela Equação 3.

$$v_k = u_k + b_k \quad (3)$$

O bias b_k é um parâmetro externo do neurônio artificial k e podemos considerar a sua presença como na equação 2. equivalentemente, podemos formular a combinação das equações 1 até 3 e gerar as equações 4 e 5.

$$v_k = \sum_{j=0}^m w_{kj} x_j \quad (4)$$

$$y_k = f(v_k) \quad (5)$$

Na Equação 4, adicionamos uma nova sinapse. A sua entrada é $w_0 = +1$ e $w_{k0} = b_k$.

Podemos, portanto reformular o modelo do neurônio k como mostrado na Figura 2. Adicionando-se um novo sinal de entrada fixo em $+1$ e adicionando-se um novo peso sináptico igual ao bias b_k , garante-se modelos matematicamente equivalentes (HAYKIN, 2001).

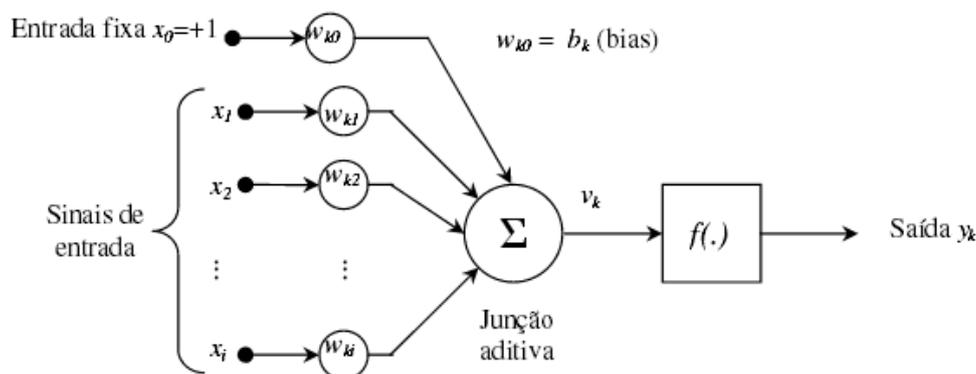


Figura 2: Um outro modelo não-linear de um neurônio (HAYKIN, 2001).

3.1. Feedforward Neural Network

Uma *Feedforward Neural Network*, a informação se move em uma única direção, a partir dos nós de entrada, através de nós ocultos (se houver) e para os nós de saída. Não há ciclos ou recursividade na rede, o que a torna mais simples sendo, um dos primeiros tipos de Redes neurais artificiais criadas. Não possuindo, pois, realimentação da saída para a entrada, tais redes são consideradas "sem memória".

3.2. Máquina de Aprendizado Extremo (ELM)

Dos diversos tipos de redes neurais existentes, as redes *feedforward* são as mais populares (HUANG; WANG; LAN, 2011), sendo as do tipo de única camada oculta especialmente conhecidas por funcionarem como aproximadores universais (CAO; LIN; HUANG, 2012). Dentre elas, a Máquina de Aprendizado Extremo (*Extreme Learning Machine* - ELM) se destaca devido ao seu rápido período de treinamento e sua essência está no uso de pesos sinápticos aleatórios na camada oculta. Essa configuração resulta em um modelo linear para os pesos sinápticos da camada de saída da rede (pesos de saída), os quais são analiticamente calculados por meio de uma solução de mínimos quadrados (HUANG; WANG; LAN, 2011). Ainda que esta rede ofereça uma boa capacidade de generalização, a escolha aleatória dos pesos sinápticos da camada oculta (pesos de entrada) pode gerar um conjunto não-ótimo de pesos e limiares, ocasionando o efeito de sobreajuste (CAO; LIN; HUANG, 2012). Outro problema enfrentado é a escolha do número de neurônios ocultos e a possibilidade de diminuí-lo sem afetar a eficácia do aprendizado, acarretando em diversos testes por tentativa e erro (YANG; WANG; YUAN, 2012).

A ELM é uma rede neural de duas camadas com pesos sinápticos fixos e aleatórios na camada oculta, $\mathbf{W} \in \mathbb{R}^{pq}$, em que p é o número de unidades de entrada e q o número de neurônios ocultos (HUANG; WANG; LAN, 2011). A saída da camada oculta é dada por

$$h(k+1) = f(W^T(k)u(k)) \quad (6)$$

Onde $u(k) \in \mathbb{R}^p$ é o atual vetor de entrada e f uma função de ativação logística. A j -ésima saída da rede é dada por

$$y_j(k+1) = \sum_{i=1}^q W_{ji}^{out}(k)h_i(k) \quad (7)$$

Durante o treino, todas as entradas são sequência do treinamento $(u(k), d(k), k=1 \dots N)$ são apresentadas à rede, e os estados correspondentes $(h(k), d(k))$ são coletados em suas respectivas matrizes \mathbf{H} e $\tilde{\mathbf{D}}$, onde $d(k)$ representa a saída desejada. O cálculo da matriz de pesos da camadas de saída, \mathbf{W} , é realizada por meio de uma regressão linear: $\mathbf{W}^{out} = \mathbf{H}^+ \tilde{\mathbf{D}}$, em que H^+ é a inversa generalizada de Moore-Penrose da matriz das saídas da camada oculta \mathbf{H} (ALBERT, 1972).

4. Metodologia do Trabalho

A implementação da rede neural foi realizada no software MATLAB e os dados multiespectrais utilizados para a classificação automática, correspondem à um recorte da cena 227/65 capturada pelos satélites Landsat-5 TM, bandas 3 (0,63 a 0,69 μm - vermelho), 4(0,76 a 0,90 μm -infravermelho próximo) e 5 (1,55 a 1,75 μm - infravermelho médio), no período de 1984 a 2011. Inicialmente as imagens foram convertidas para uma resolução radiométrica de 8 bits, assim as imagens trabalhadas possuem $2^8 = 256$ níveis de cinza. É necessário iniciar o processo de aquisição das amostras de treinamento e teste, onde as amostras de treinamento devem conter representações de todas as classes presentes na imagem. Neste trabalho as classes definidas são: Água, Floresta e Solo Exposto.

Foram adquiridas 30 amostras para cada uma das três classes utilizadas, da imagem do ano de 1984. Cada amostra é representada por uma janela quadrada de 3x3, contabilizando 90 amostras, das quais 70% foram utilizadas para o treinamento da rede e 30% para teste. Como as amostras possuem três bandas, são geradas três matrizes 3x3 para cada amostra e cada matriz é transformada em um vetor linha de nove colunas, que em seguida são concatenados horizontalmente gerando um vetor linha com 27 colunas, para representar as características de uma determinada classe. No total, uma matriz de entrada de 90x27 foi formada sendo 30 primeiras linhas referentes às características das amostras da classe água, as 30 seguintes referentes à classe Floresta e as 30 últimas referente à classe Solo Exposto.

Em seguida as 90 amostras foram alocadas aleatoriamente na matriz de dimensões 90x27. Foi feita a normalização dos dados para valores entre -1 e 1, para que fosse possível a utilização da função de ativação de tangente hiperbólica "tansig" e também evitar problemas de saturação da rede. Para criar a saída desejada da rede, foi criada uma matriz binária de dimensões 90x3, onde a classe água é representada pelo número binário 100, a classe Floresta por 010 e Solo exposto 001. Como podemos ver na Tabela 1.

Tabela 1: Tabela de representação das classes

Classes	Binário	Cor representativa
Água	100	Azul
Floresta	010	Verde
Solo Exposto	001	Marron

Após a rede ser treinada e testada, foram realizados os testes das imagens referentes aos anos de 1984, 1989, 1997, 2005 e 2011. Cada imagem é passada como entrada da rede e é feita uma varredura sobre a imagem com uma janela de resolução 3x3 e cada janela é analisada pela rede neural, que classifica a amostra em uma das três classes determinadas. A arquitetura da rede utilizada pode ser dita como uma *Perceptron Multicamadas Estática*, pois possui duas camadas, uma camada de entrada, uma camada oculta com 10 neurônios e saídas lineares, baseada no algoritmo ELM e na função de ativação tangente hiperbólica.

Para a avaliação da classificação são utilizadas as matrizes de confusão, onde estão exibidos os coeficientes que expressam a exatidão das classes, o índice Kappa e o índice de exatidão global, além do erro quadrático médio gerado pela rede. Congalton (1991) relata que o uso do coeficiente Kappa é satisfatório na avaliação da precisão de uma classificação temática, pelo fato de levar em consideração toda a matriz de confusão no seu cálculo, inclusive os elementos de fora da diagonal principal, os quais representam as discordâncias na classificação, diferentemente da exatidão global, por exemplo, que utiliza somente os elementos diagonais.

5. Resultados e Discussões

Após a rede ter sido treinada com 70% das amostras da imagem do ano de 1984 e em seguida utilizado os outros 30% das amostras para testar a qualidade da rede, foram obtidos os resultados mostrados nas Tabelas 2, 3, 4 e 5 e Figuras 3-7.

Com uma janela de amostra do tamanho 3x3, foi possível obter características bem distintas entre as classes, sendo esse um fator determinante na qualidade da rede em conjunto com a utilização do algoritmo ELM. Portanto podemos perceber que a rede obteve desempenho considerável na classificação das amostras.

Tabela 2: Matriz Confusão de Treino

Treino	Água	Floresta	Solo Exposto	Total
Água	21	1	0	22
Floresta	1	19	1	21
Solo Exposto	1	0	19	20
Total	23	20	20	63

Tabela 3: Índices de acurácia global para treino

Índice Kappa	0.905
Erro Quadrático Médio	0.00674
Exatidão Global	93.65%

Realizada a contagem dos pixels da classe solo exposto em cada um dos mapas de classificação gerado pela rede neural, foi possível estimar a cobertura do solo das cenas. Em 1984 a quantidade de pixels obtidos para a classe solo exposto foi de 61290, o que representa 1.36% da cena, para 1989 a quantidade de pixels foi de 98163, representando 2.18%, em 1997 foi de 154314, representado 3.42% da cena, em 2005 a quantidade de pixels foi de 397584, representando 8.83% e finalmente em 2011 a quantidade de pixels contabilizados foi de 446139, o que representa 9.91% da cena. Lembrando que as imagens utilizadas para a classificação possuem resolução de 1500x3000.

Os resultados obtidos nos mapas de classificação são justificados pelas bandas escolhidas das imagens dos satélites, pois foram utilizadas as bandas 3, 4 e 5. A banda 3 ajudou significativamente para a precisão de acerto da classe Floresta, pois a vegetação verde, densa e uniforme, apresenta grande absorção, permitindo bom contraste entre as áreas ocupadas com vegetação. A banda 4 foi de grande importância para a acurácia da classe Água, pois os corpos de água absorvem muita energia nesta banda, permitindo o mapeamento da rede de drenagem e delineamento de corpos de água. A banda 5 apresenta características de sensibilidade ao teor de umidade das plantas, servindo para observar estresse na vegetação, causado por desequilíbrio hídrico, foi utilizada principalmente para obter maiores informações sobre as classes. Podemos perceber nos mapas que algumas regiões onde há a existência de Solo Exposto não foram classificadas como o esperado, pois muitos pixels que deveriam ser classificados como da classe, foram classificados como Floresta e vice e versa, este erro ocorre porque algumas amostras da

Tabela 4: Matriz Confusão de Teste

Treino	Água	Floresta	Solo Exposto	Total
Água	5	0	0	5
Floresta	2	10	1	13
Solo Exposto	0	0	9	9
Total	7	10	10	27

Tabela 5: Índices de acurácia global para teste

Índice Kappa	0.829
Erro Quadrático Médio	0.1023
Exatidão Global	88.89%



Figura 3: Recorte da cena 227/65 do ano de 1984. a) recorte original e b) resultado da classificação pela RNA.



Figura 4: Recorte da cena 227/65 do ano de 1989. a) recorte original e b) resultado da classificação pela RNA.

classe solo exposto e floresta apresentam características de bandas similares, o que deverá ser tratado em trabalhos futuros.

6. Conclusões

Os resultados apresentados neste trabalho mostram a eficiência do Algoritmo Extreme Learning Machine aplicado a uma *Single Layer Feedforward Network*, com o objetivo de identificar as características de desmatamento ao longo dos anos e as diferentes feições de uso e cobertura do solo de imagens do sensor Landsat-5 TM, referentes a região de Novo Progresso, através de um mapa de classificação de uso do solo gerado pela rede neural. Com os dados estatísticos obtidos da RNA, podemos dizer que os resultados da classificação foram satisfatórios. Como mostrados nos mapas gerados pela classificação feita pela RNA em comparação com a imagem adquirida pelo satélite Landsat-5 TM.



Figura 5: Recorte da cena 227/65 do ano de 1997. a) recorte original e b) resultado da classificação pela RNA.

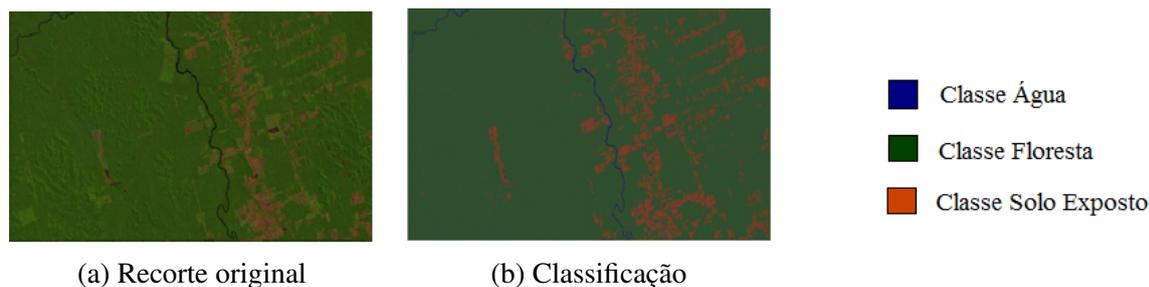


Figura 6: Recorte da cena 227/65 do ano de 2005. a) recorte original e b) resultado da classificação pela RNA.

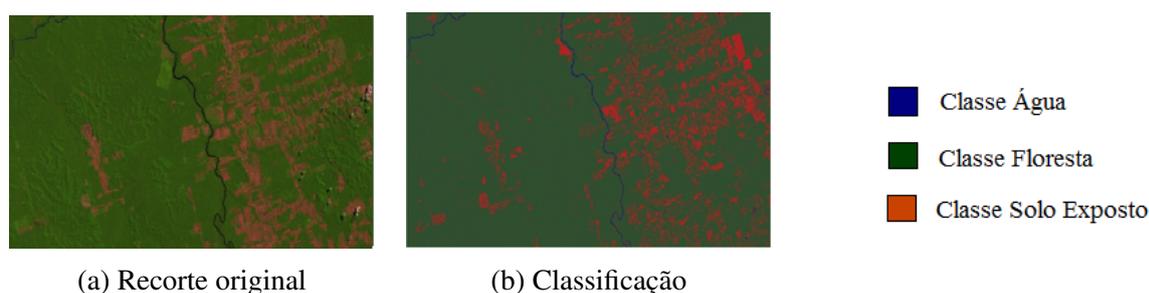


Figura 7: Recorte da cena 227/65 do ano de 2011. a) recorte original e b) resultado da classificação pela RNA.

Agradecimentos

Este trabalho é parcialmente financiado pela Fundação Amazônica Paraense de Amparo à Pesquisa e Vale S/A por meio do edital 001/2010. Também é suportado pela Universidade Federal do Pará e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Referências

- ALBERT, A. *Regression and the Moore-Penrose pseudoinverse*. [S.l.]: Elsevier, 1972.
- CAO, J.; LIN, Z.; HUANG, G.-B. Self-adaptive evolutionary extreme learning machine. *Neural processing letters*, Springer, v. 36, n. 3, p. 285–305, 2012.
- CASTRO, E.; MONTEIRO, R.; CASTRO, C. Atores e relações sociais em novas fronteiras na amazônia: Novo progresso, castelo de sonhos e são félix do xingú. *Belém, Pará*, 2002.
- CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, Elsevier, v. 37, n. 1, p. 35–46, 1991.
- ESCADA, M. I. S. et al. Levantamento do uso e cobertura da terra e da rede de infra-estrutura no distrito florestal da br-163. *São José dos Campos: Inpe*, 2009.
- HAYKIN, S. *Redes neurais: princípios e práticas*. tradução: Paulo Martins Angel, ISBN: 85-7307-718-2. [S.l.]: Porto Alegre: Bookman, 2001.
- HUANG, G.-B.; WANG, D. H.; LAN, Y. Extreme learning machines: a survey. *International Journal of Machine Learning and Cybernetics*, Springer, v. 2, n. 2, p. 107–122, 2011.
- IBGE, I. Cidades@. Available from: <http://www.ibge.gov.br/cidadesat/topwindow.htm>, v. 1, 2010.
- YANG, Y.; WANG, Y.; YUAN, X. Bidirectional extreme learning machine for regression problem and its learning effectiveness. *Neural Networks and Learning Systems, IEEE Transactions on*, IEEE, v. 23, n. 9, p. 1498–1505, 2012.