



RECONHECIMENTO DE CARACTERES EM PLACAS AUTOMOTIVAS UTILIZANDO REDES NEURAIS

Eduarda Gonçalves Dias¹; Wanderson Rigo²; Manassés Ribeiro³

INTRODUÇÃO

Os sistemas de visão artificial vêm auxiliando o ser humano em diversas tarefas. Sua aplicabilidade se estende por várias áreas e atualmente o grande interesse está relacionado à engenharia de tráfego e à segurança pública. Sistemas desta natureza podem disponibilizar rapidamente informações sobre veículos em circulação, possibilitando que medidas necessárias possam ser tomadas em tempo.

Uma Rede Neural Artificial (RNA) é um modelo matemático implementado em software que tenta simular o funcionamento do cérebro humano via interconexões de neurônios artificiais [1][2][4]. Tais aparatos têm sido usados para solucionar uma grande variedade de tarefas que são difíceis de resolver utilizando programação baseada em regras comuns, incluindo visão computacional, reconhecimento de voz e principalmente reconhecimento de padrões [1][2][3].

Este trabalho visa apontar a melhor topologia e configuração de rede neural *feedforward* a ser utilizada no reconhecimento de caracteres de placas automotivas. Nele inicialmente discorre-se sobre RNAs, depois são abordados os procedimentos metodológicos utilizados no trabalho, seguido dos testes realizados e por fim as discussões e conclusões fomentadas pelos resultados obtidos.



O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq – Brasil.

¹ Aluna do Ensino Médio Integrado, Instituto Federal Catarinense, Campus Videira. Curso técnico em Informática. Email: duda_gdias@hotmail.com

² Professor Orientador do Instituto Federal Catarinense, Campus Videira. Curso técnico em Informática. E-mail: wanderson.rigo@ifc-videira.edu.br

³ Professor Co-orientador do Instituto Federal Catarinense, Campus Videira. Curso técnico em Informática. E-mail: manasses@ifc-videira.edu.br

PROCEDIMENTOS METODOLÓGICOS

Este trabalho desenvolveu-se entre fevereiro de 2013 e junho de 2014 totalizando 17 meses via:

1. Revisão bibliográfica sobre redes neurais e reconhecimento de caracteres utilizando redes neurais artificiais.
2. Definição da arquitetura das redes e extração e preparo dos dados para treinamento e teste das redes. O **conjunto de treinamento** foi composto por 8 vetores, cada um contendo os valores que descrevem as imagens individualizadas de cada algarismo, as quais foram retiradas de um banco de imagens. Já o **conjunto de teste** foi composto por 2 vetores diferentes do conjunto de treinamento, também cada um contendo os valores que descrevem as imagens individualizadas de cada algarismo, as quais foram retiradas do mesmo banco de imagens.
3. Implementação das redes via software e execução de testes visando análise comparativa da *performance* das redes configuradas com diferentes topologias e parâmetros. Foram criadas duas redes com a ajuda do software Scilab⁴ e das instruções disponíveis em [3]: (1) uma para identificar caracteres numéricos de zero a nove (0-9) e outra (2) para identificar caracteres alfabéticos de A a Z (A-Z).
4. Análise, sumarização e discussão dos resultados obtidos.

TESTES DAS REDES

As RNAs precisam ser treinadas para aprender sobre os padrões que deverão ser reconhecidos. Ao se treinar uma RNA, o foco pode ser tanto a generalização da rede quanto a sua especialização. Se uma rede neural for demasiadamente treinada, ela responderá bem a casos específicos, mas perde capacidade de generalização e não terá respostas satisfatórias para padrões não

⁴ <http://www.scilab.org>

vistos no treinamento. Se o treinamento for insuficiente, ela se assimilará a um modelo genérico, respondendo de maneira similar a todos os casos testados [5]. Neste trabalho buscou-se achar um meio termo: treinar as redes suficientemente, mas não as especializar. Em todos os testes, 8 amostras foram usadas no treinamento e 2 na avaliação de cada rede.

Visando aferir a *performance* das redes, diversos testes foram executados alterando principalmente (i) a topologia da rede, (ii) o tipo de saída esperada e os (iii) parâmetros de treinamento. A taxa de aprendizado utilizada foi de 0.01 e o limiar de erro propagável admissível foi fixado em 0,0001. Ambos os valores são propostos na literatura [4] e desta forma foram mantidos constantes. A seguir são mostradas as variações nos parâmetros:

(i) Topologia da rede

- Neurônios na camada de entrada: 30, cada um correspondendo ao mapeamento de uma característica extraída das imagens de placas de automóveis via os algoritmos LBP de textura (13 características), de forma MOM (11 características) e de cor HST (6 características);
- Neurônios na camada intermediária: valor variado empiricamente;
- Neurônios na camada de saída: valor variado de acordo com o tipo de saída.

(ii) Tipo de saída esperada

a) Inicialmente mapeou-se cada tipo de entrada (0-9) para uma saída codificada com 4 números binários. Como 10 caracteres numéricos deveriam ser reconhecidos, precisou-se de 4 neurônios ($2^4=16$) para codificar as saídas desejadas. As entradas e saídas ficaram como mostra a Tabela 1.

Tabela 1 – Saída do tipo A

Entrada (30 neurônios)	Saída esperada (4 neurônios/base 2)
Características das imagens do número 0	0 0 0 0
Características das imagens do número 1	0 0 0 1
Características das imagens do número 2	0 0 1 0

Características das imagens do número 3	0 0 1 1
Características das imagens do número 4	0 1 0 0
Características das imagens do número 5	0 1 0 1
Características das imagens do número 6	0 1 1 0
Características das imagens do número 7	0 1 1 1
Características das imagens do número 8	1 0 0 0
Características das imagens do número 9	1 0 0 1

Fonte: dos autores.

Tal abordagem obteve taxa de acerto de apenas 35 %.

b) Posteriormente mapeou-se cada tipo de entrada (0-9) para uma saída codificada com 10 números binários. O valor 1 na saída de um neurônio significa que tal neurônio foi ativado, sendo que posição deste indica qual algarismo foi reconhecido. As entradas e saídas ficaram conforme a Tabela2.

Tabela 2 – Saída do tipo B

Entrada (30 neurônios)	Saída esperada (10 neurônios/posição)
Características das imagens do número 0	1 0 0 0 0 0 0 0 0 0
Características das imagens do número 1	0 1 0 0 0 0 0 0 0 0
Características das imagens do número 2	0 0 1 0 0 0 0 0 0 0
Características das imagens do número 3	0 0 0 1 0 0 0 0 0 0
Características das imagens do número 4	0 0 0 0 1 0 0 0 0 0
Características das imagens do número 5	0 0 0 0 0 1 0 0 0 0
Características das imagens do número 6	0 0 0 0 0 0 1 0 0 0
Características das imagens do número 7	0 0 0 0 0 0 0 1 0 0
Características das imagens do número 8	0 0 0 0 0 0 0 0 1 0
Características das imagens do número 9	0 0 0 0 0 0 0 0 0 1

Fonte: dos autores.

Tal abordagem obteve taxa de acerto de 50 %.

c) E finalmente mapeou-se cada tipo de entrada (0-9) para uma saída codificada com 10 números binários. O valor 1 na saída de um neurônio significa que tal

neurônio foi ativado, sendo que a quantidade total de uns denota qual algarismo foi reconhecido. As entradas e saídas ficaram conforme a Tabela 3.

Tabela 3 – Saída do tipo C

Entrada (30 neurônios)	Saída esperada (10 neurônios/quantidade)
Características das imagens do número 0	0 0 0 0 0 0 0 0 0 0
Características das imagens do número 1	0 0 0 0 0 0 0 0 0 1
Características das imagens do número 2	0 0 0 0 0 0 0 0 1 1
Características das imagens do número 3	0 0 0 0 0 0 0 1 1 1
Características das imagens do número 4	0 0 0 0 0 0 1 1 1 1
Características das imagens do número 5	0 0 0 0 0 1 1 1 1 1
Características das imagens do número 6	0 0 0 0 1 1 1 1 1 1
Características das imagens do número 7	0 0 0 1 1 1 1 1 1 1
Características das imagens do número 8	0 0 1 1 1 1 1 1 1 1
Características das imagens do número 9	0 1 1 1 1 1 1 1 1 1

Fonte: dos autores.

Tal abordagem também obteve taxa de acerto de 50%.

(iii) Parâmetros de treinamento

Variou-se o número de épocas por 3000, 4000 e 5000. Em testes preliminares, valores inferiores ou superiores a estes não alteraram a *performance* das redes.

RESULTADOS E DISCUSSÕES

Foram realizados inúmeros testes empíricos até que formalizou-se a arquitetura e a forma de variação nos parâmetros de configuração das redes. A partir disso, 36 testes foram realizados e registrados, sendo que os resultados obtidos são os demonstrados pela Tabela 4, a qual agrupa os testes por épocas.

Tabela 4 – Resultados obtidos nos testes

Teste	Tipo de Saída	Camada intermediária	Épocas	Taxa de Acertos
01	A	80	3000	35%
02	A	90	3000	35%
03	A	100	3000	35%
04	A	110	3000	35%
05	A	80	4000	35%
06	A	90	4000	35%
07	A	100	4000	35%
08	A	110	4000	35%
09	A	80	5000	35%
10	A	90	5000	35%
11	A	100	5000	35%
12	A	110	5000	35%
13	B	80	3000	35%
14	B	90	3000	30%
15	B	100	3000	40%
16	B	110	3000	40%
17	B	80	4000	45%
18	B	90	4000	40%
19	B	100	4000	40%
20	B	110	4000	50%
21	B	80	5000	45%
22	B	90	5000	40%
23	B	100	5000	45%
24	B	110	5000	35%
25	C	80	3000	50%
26	C	90	3000	50%
27	C	100	3000	45%
28	C	110	3000	40%
29	C	80	4000	45%
30	C	90	4000	50%
31	C	100	4000	45%

32	C	110	4000	50%
33	C	80	5000	45%
34	C	90	5000	45%
35	C	100	5000	40%
36	C	110	5000	50%

Fonte: dos autores.

Já a rede para o reconhecimento das letras (A-Z), no teste melhor sucedido, com 30 neurônios na camada de entrada, também variando-se os neurônios da camada intermediária e de saída conforme metodologia supra descrita, obteve somente 23% de acerto. Logo tais testes foram suprimidos deste trabalho.

CONSIDERAÇÕES FINAIS

Este projeto foi de grande valia por explorar um tema tão visado na atualidade: o reconhecimento de padrões a partir de imagens. Grande parte do conteúdo explorado aqui não é coberta por disciplinas de um curso técnico, desafiando o bolsista/pesquisador a explorar áreas desconhecidas de seu cotidiano letivo.

Em vista dos resultados obtidos, chegou-se ao consenso de que a descrição das imagens via os 30 atributos extraídos das imagens por algoritmos não as diferencia suficientemente uma da outra, ou seja, os valores numéricos que descrevem a imagem do algarismo um (1) são muito semelhantes aos valores numéricos que descrevem o algarismo dois (2), por exemplo. Essa pouca variância dificulta o aprendizado da rede, já que todas as descrições de imagens se assemelham.

Se os padrões fossem mais heterogêneos o aprendizado das redes seria mais eficiente e elas se comportariam melhor. Sendo assim, sugere-se: (i) obter imagens de melhor qualidade, evitando ruídos e, se possível, tratá-los; (ii) melhorar o extrator de características das imagens; (iii) efetuar tratamento nos dados para amplificar as diferenças entre as descrições das imagens; (iv) executar novos testes com outras redes/configurações de parâmetros. Tais verificações podem embasar

trabalhos futuros e balizar outros pesquisadores que trabalham com problemas similares aos tratados aqui.

REFERÊNCIAS

- [1] BARRA, Fernanda. **Redes neurais artificiais**. Disponível em: <<http://blogdopetcivil.com/2013/07/05/redes-neurais-artificiais>> Acesso em: 20/07/2014.
- [2] BRAGA, A. de P.; CARVALHO, A. C. P. de L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: Livros Técnicos e Científicos, 2011. xii, 226 p.
- [3] GLADKIKH, Vlad. **Artificial neural networks in Scilab**. Disponível em :<<http://burubaxair.wordpress.com/tag/neural-network>> Acesso em: 20/07/2014.
- [4] RUSSELL, Stuart J.; NORVIG, Peter. **Inteligência artificial**. Rio de Janeiro: Elsevier, 2004.
- [5] TAMPELINI, Leonardo Garcia. **Uma Aplicação das Redes Neurais Artificiais de Elman e Time Delay Neural Network na Predição de Séries Temporais**. Disponível em: <<http://www.inf.unioeste.br/~tcc/2008/TCC%20-%20Leonardo%20Garcia%20Tampelini.pdf>> Acesso em: 20/07/2014.