

Redes Neurais Autoassociativas de Arquitetura Profundo para Reconhecimento de Expressões Faciais

Alcoforado, M. S.

Escola Politécnica de Pernambuco

Universidade de Pernambuco

50.720-001 - Recife, Brasil

msa@ecomp.poli.br

Resumo *A face consegue transmitir as emoções humanas antes mesmo da pessoa verbalizá-las [1]. Por este motivo, estudos na área de Reconhecimento de Expressões Faciais (REF) vêm sendo cada vez mais desenvolvidos na computação [2,3,4]. As redes neurais artificiais (RNAs) [5] são métodos que têm sido aplicados com sucesso nos mais diversos problemas da visão computacional, como reconhecimento de faces [6], reconhecimento de expressão facial [7], reconhecimento de caracteres [8], reconhecimento de gestos [9], segmentação de imagens [10], entre outros. Este estudo irá realizar uma comparação entre RNAs autoassociativas de arquitetura profunda e simples aplicadas ao REF, utilizando a banco de imagens JAFFE. Além disso, irá realizar uma comparação entre os resultados obtidos e a técnica Principal Component Analysis (PCA).*

Abstract *Face can transmit human emotions even before the person verbalize them [1]. For this reason, studies on Facial Expression Recognition (FER) are being increasingly developed in computer science [2,3,4]. Artificial Neural Networks (ANNs) [5] are methods that have been successfully applied in various problems of computer vision, as face recognition [6], facial expression recognition [7], character recognition [8], gesture recognition [9], image segmentation [10], among others. This study will make a comparison between ANNs using autoassociative model with deep and simple architecture applied to FER, using JAFFE database. Moreover, it will make a comparison between the results obtained and the Principal Component Analysis (PCA) technique.*

1 Introdução

O desenvolvimento de modelos baseados em teorias sobre a estrutura do cérebro humano tem se mostrado como uma importante ferramenta para a inspiração de novas abordagens para problemas de reconhecimento de padrões visuais. Apesar do cérebro humano não ser completamente entendido, ele já inspirou vários mecanismos utilizados em tarefas de reconhecimento de padrões, como as redes neurais artificiais [11].

Assim, baseadas no conceito de memória autoassociativa do cérebro humano, surgem RNAs que utilizam os modelos autoassociativos e de arquitetura profunda.

Tais RNAs aprendem com uma menor interferência externa, principalmente no que diz respeito à extração de características que é realizada pela própria RNA que recebe como entrada os dados na forma bruta.

Um outro método que realiza extração de características é o PCA. Este tem por finalidade básica a análise dos dados usados visando sua redução, eliminação de sobreposições e a escolha das formas mais representativas de dados a partir de combinações lineares das variáveis originais.

A partir destes conceitos, objetiva-se comparar as taxas obtidas no reconhecimento de expressões faciais com RNAs autoassociativas de arquitetura simples e profunda e os resultados já existentes na literatura do método PCA.

2 Principais Conceitos

2.1 Modelo Autoassociativo

A memória autoassociativa é um tipo de memória presente no cérebro humano para armazenamento de memória episódica e de curto prazo [12]. Nesse tipo de memória, o processo de aprendizagem de novos padrões é muito rápido e, após o aprendizado, um padrão pode ser recuperado com apenas um fragmento do mesmo [11].

No modelo autoassociativo, deseja-se que a entrada passada à rede seja retornada como saída. Logo, a rede neural possui a mesma quantidade de neurônios na camada de entrada e na camada de saída.

RNAs autoassociativas são capazes de aprender de modo implícito as características internas dos dados de entrada. Ou seja, elas conseguem aprender certas características pertencentes ao padrão apresentado na entrada sem a necessidade de qualquer conhecimento anterior ou instrução específica.

2.2 Modelo de Arquitetura Profunda

As arquiteturas profundas diferem-se das superficiais por apresentarem modelos mais profundos do que largos. Elas são consideradas como uma composição de várias camadas de componentes adaptativos não-lineares. Uma rede neural MLP autoassociativa com apenas uma camada escondida é considerada um modelo de arquitetura superficial [figura 1], enquanto a adição de várias camadas escondidas torna-a um modelo de arquitetura mais profunda [figura 2] e com maior capacidade de representar certas famílias de funções.

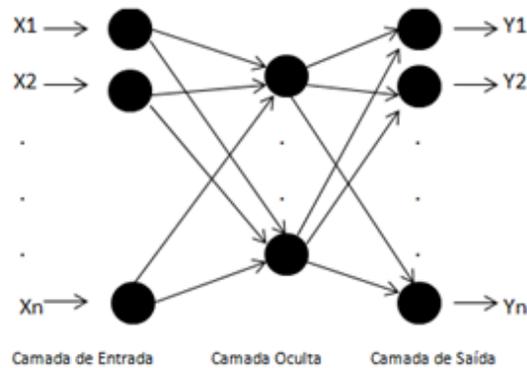


Fig 1. Exemplo de RNA autoassociativa de arquitetura simples.

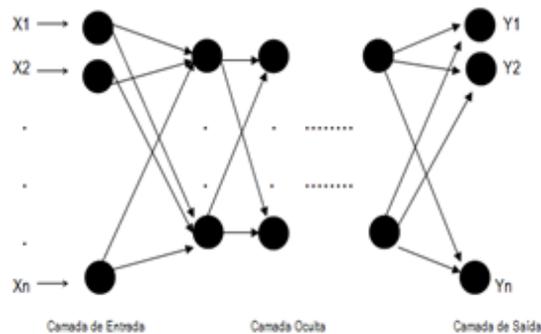


Fig 2. Exemplo de RNA autoassociativa de arquitetura profunda.

Os modelos de arquitetura profunda podem aprender com menor envolvimento humano na construção do modelo antes do treinamento, menos padrões de exemplo e menor custo computacional, pois integram as etapas de extração de características e classificação em um mesmo modelo [11]. Estes modelos também dispensam processos prévios de extração de características visto que realizam esta etapa implicitamente através das várias camadas de processamento.

2.3 PCA

O PCA é um dos métodos estatísticos de múltiplas variáveis mais simples e vem sendo muito utilizado na área de reconhecimento de padrões [13, 14].

Classificadores autoassociativos almejam aprender a estrutura interna de uma dada classe de padrões com base apenas nos elementos que a compõem. O PCA é um dos métodos mais simples que podem ser aplicados na realização de tal tarefa [11].

Contudo, PCA possui a capacidade de identificar apenas correlações lineares no conjunto de dados. Redes neurais inspiradas na metodologia NLPCA (*Nonlinear*

Principal Component Analysis) [15], tal qual a MLP, também podem ser utilizadas como classificadores autoassociativos, com a vantagem de possibilitar correlações não-lineares entre as variáveis.

2.4 JAFFE

A base de dados de Expressões Faciais de Japonesas (*Japanese Female Facial Expression*, JAFFE) [16] foi desenvolvida para colaborar na avaliação de métodos no problema de reconhecimento de expressão facial. Esta base contém imagens de 6 expressões faciais: raiva, nojo, medo, alegria, tristeza e surpresa; além de imagens que caracterizam a face neutra. As imagens da base JAFFE foram coletadas a partir de 10 pessoas, apresentando 3 ou 4 imagens de cada expressão para cada pessoa.

3 Experimentos

Durante a pesquisa, foi criada uma MLP autoassociativa que utiliza o *backpropagation* como algoritmo de treinamento da rede, o método *Resilient Propagation* (Rprop) para ajuste dos pesos e a sigmóide logística como função de ativação.

Utilizando esta rede, foram feitos treinos com a base JAFFE utilizando 2 imagens de cada expressão de cada pessoa para treino e o resto das imagens para teste. Estes treinos foram feitos com a MLP com arquiteturas profunda, várias camadas escondidas, e simples, apenas uma camada oculta. No experimento, quantidade de neurônios da(s) camada(s) oculta(s), tamanho da imagem, quantidades de épocas e iterações foram variados.

Os resultados são avaliados utilizando matriz de confusão e a taxa de classificação é obtida através da média de 30 iterações de teste.

4 Resultados

Assim, os melhores resultados alcançados, utilizando 100 épocas, podem ser vistos nas tabelas a seguir:

Tabela 1. Resultados do modelo de arquitetura simples.

Neurônios na camada oculta	Tamanho da imagem	Taxa de classificação	Desvio padrão
2	40	87,4%	3,26
3	30	86,8%	2,51
5	50	87,4%	3,05
10	50	87,4%	2,84

Tabela 2. Resultados do modelo de arquitetura profunda com 2 camadas ocultas.

Neurônios nas camadas ocultas	Tamanho da imagem	Taxa de classificação	Desvio padrão
2,1	30	87,2%	3,14
5,3	50	87,3%	2,52
10,5	50	88,1%	2,45

Tabela 1. Resultados do modelo de arquitetura profunda com 3 camadas ocultas.

Neurônios nas camadas ocultas	Tamanho da imagem	Taxa de classificação	Desvio padrão
2,1,2	50	87,7%	2,99
5,2,5	40	87,4%	2,50
10,5,5	30	87,1%	2,74

A tabela a seguir mostra a matriz de confusão média do melhor caso obtido:

Tabela 4. Matriz de confusão média do resultado encontrado utilizando duas camadas ocultas com 10 e 5 neurônios, respectivamente.

	DI	SU	HA	SA	AN	FE	NE
DI	8,233333	0	0	0,566667	0,2	0	0
SU	0	9,066667	0,3	0,433333	0	0,233333	0,233333
HA	0	0,3	9	0,133333	0	0,333333	1,2
AS	0,3	0,3	0,333333	8,933333	0,2	0,633333	0,833333
NA	0,5	0	0	0,666667	9,433333	0	0
FE	0,565555	0,433333	0	0,633333	0	9,3	1,066667
NE	0	0	0	0	0	0	10

O melhor resultado encontrado mostra-se superior ao encontrado na literatura pelo PCA [17] que alcança uma taxa de 80,0%.

5 Conclusão

Após analisar as tabelas, pode-se concluir que a RNA autoassociativa consegue reconhecer expressões faciais alcançando uma boa taxa de classificação de 88,1%.

Entretanto, as taxas de classificação obtidas utilizando uma MLP autoassociativa de arquitetura profunda mostraram-se discretamente superiores aos resultados encontrados utilizando a MLP autoassociativa de arquitetura simples. Isso mostra que a simples adição de camadas, porém, não tornará o modelo necessariamente mais eficiente para resolver problemas de visão computacional. Deve-se fazer uma análise para assim definir os tipos de conexão entre os neurônios e a forma como eles serão relacionados. Algoritmos construtivos e evolucionários, além de métodos de otimização, podem ser levados em conta nesta análise.

Porém, os resultados obtidos com a MLP autoassociativa, que realiza um PCA não-linear, durante foram bastante superiores a melhor taxa obtida pelo PCA, mostrando assim que correlacionar as variáveis de maneira não-linear é mais eficiente.

Referências

- [1]] Mehrabian, A. (1968). "Communication without words," *Psychol. Today*, vol.2, pp. 53–56.
- [2] Tai, S. C. and Chung, K. C. (2007). Automatic Facial Expression Recognition System Using Neural Networks. IEEE publications.
- [3] Lajevardi, S. M. and Lech, M. (2008). Facial Expression Recognition Using Neural Networks and Log-Gabor Filters. IEEE publications.
- [4] Chen, L. , Zhang, T., Gui, X. and Yang, G. (2010). Development of Research on Facial Expression Recognition. IEEE publications.
- [5] Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. IEEE Press.
- [6] Wong, Y. W., K. P. Seng e L.-M. Ang (2011). 'Radial basis function neural network with incremental learning for face recognition'. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* 41(4), 940–949.
- [7] Zavaschi, T. H. H., A. S. B. Jr., L. E. S. Oliveira e A. L. Koericha (2013). 'Fusion of feature sets and classifiers for facial expression recognition'. *Expert Systems with Applications* 40, 646–655.
- [8] Lecun, Y., L. Bottou, Y. Bengio e P. Haffner. (1998). *Gradient-based learning applied to document recognition*. In 'Proceedings of the IEEE'. Vol. 86. pp. 2278–2324.
- [9] Nolkner, C. e H. Ritter (2002). 'Visual recognition of continuous hand postures'. *IEEE Transactions on Neural Networks* 13(4), 983–994.
- [10] Dong, G. e M. Xie (2005). 'Color clustering and learning for image segmentation based on neural networks'. *IEEE Transactions on Neural Networks* 16(4), 925–936.
- [11] Fernandes, B. J. T. (2013). *Redes Neurais com Extração Implícita de Características para Reconhecimento de Padrões Visuais*. 29 de julho de 2013. 155 f. Tese (Doutorado em Ciência da Computação) – Centro de Informática, Universidade Federal de Pernambuco, Pernambuco.
- [12] Rolls, E. T. e A. Treves. (1998). *Neural networks and brain function*. Oxford University Press.
- [13] Liu, C. (2004). Gabor-based kernel PCA with fractional power polynomial models for face recognition. IEEE publications.
- [14] [Nedevschi, S.](#), [Peter, I.R.](#) and [Mandrut, A.](#) (2012). PCA type algorithm applied in face recognition. IEEE publications.
- [15]] Reyes, J.,M. Vellasco e R. Tanscheit (2012). 'Monitoramento e diagnóstico de múltiplos sensores por redes neurais auto-associativas'. *Sba: Controle e Automação Sociedade Brasileira de Automatica* 23, 121 – 137.
- [16] Lyons, M. J., M. Kamachi e J. Gyoba. (1997). *Japanese Female Facial Expressions (JAFFE), database of digital images*.
- [17] Zhi, R., M. Flierl, Q.-Q. Ruan and W. B. K. (2011). 'Graph-preserving sparse nonnegative matrix factorization with application to facial expression recognition'. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 41(1), 38–52.