计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2018,27(2):24−29 [doi: 10.15888/j.cnki.csa.006204]
©中国科学院软件研究所版权所有.
E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

混合卷积神经网络的人脸验证①

郭明金, 倪佳佳, 陈　姝

(湘潭大学 信息工程学院, 湘潭 411105)

摘　要: 人脸验证对于个人身份认证很重要, 它在系统安全和犯罪识别中具有重要意义. 人脸验证的任务是给定一对人脸图像判断是否为相同的身份 (即二进制分类). 传统的验证方法包括两个步骤: 特征提取和人脸验证. 提出了一个混合卷积神经网络, 用于进行人脸验证, 主要过程分为三个步骤: 特征提取, 特征选择和人脸验证. 这个模型关键点是直接使用混合卷积神经网络从原始像素直接学习相关的视觉特征, 并通过单变量特征选择和主成分分析 (PCA) 进一步处理特征. 这样可以实现从原始像素提取到具有较好鲁棒性和表达性的特征. 在顶层使用支持向量机 (SVM) 判读是否为同一个人. 通过实验可以发现混合卷积神经网络模型与传统方法相比在人脸验证得准确率上有着较好的表现.

关键词: 人脸验证; 卷积神经网络; 特征融合; 特征选择; 支持向量机

引用格式: 郭明金,倪佳佳,陈姝.混合卷积神经网络的人脸验证.计算机系统应用,2018,27(2):24–29. http://www.c-s-a.org.cn/1003-3254/6204.html

Face Verification of Mixed Convolutional Neural Networks

GUO Ming-Jin, NI Jia-Jia, CHEN Shu

(College of Information and Engineering, Xiangtan University, Xiangtan 411105, China)

Abstract: Face verification is important for personal identity authentication, which is significant in system security and criminal identification. Face verification task is to give a pair of face images to determine whether they are of the same identity (i.e. binary classification). The traditional authentication method consists of two steps: feature extraction and face verification. In this study, a hybrid convolutional neural network (HBCNN) is proposed for face verification. The main process is divided into three steps: feature extraction, feature selection, and face verification. The key point of this model is to directly use the mixed convolutional neural network to learn the relevant visual features directly from the original pixels and to further process the features through univariate feature selection and principal component analysis (PCA). This can be achieved from the original pixel extraction to a better robustness and expression of the characteristics. The support vector machine (SVM) at the top level is used to see if it is the same person. Experiments show that the mixed convolutional neural network model has a better performance than the traditional method in verifying accuracy of face verification.

Key words: face verification; convolutional neural network; feature fusion; feature selection; support vector machine (SVM)

引言

人脸验证是人脸识别领域的一个研究重点, 本文专注于人脸验证的任务, 其目的是确定两个脸部图像是否属于相同的身份. 在现实中, 两个脸部图像在姿势、照明、表情、年龄情况下被给予其大量的个人变化. 因此直接使用人脸图像来进行验证变得更加困难.

① 基金项目: 国家自然科学基金 (61100139); 湖南省教育厅青年项目 (16B258); 湖南省自然科学基金 (2017JJ2252)
收稿时间: 2017-05-08; 修改时间: 2017-05-31; 采用时间: 2017-06-08; csa 在线出版时间: 2018-01-12

© 中国科学院软件研究所 http://www.c-s-a.org.cn

这是因为在挑选图像中的特征验证身份时往往忽略随着环境条件差异而变化的特征.

传统方法通常分成两个步骤中: 特征提取和面部验证. 在特征提取阶段, 大都使用人工提取的特征, 更重要的是这些人工提取的特征必须提前设计. 因此, 这些特征往往应用于某些特定的领域, 从而导致这些特征缺乏统一性.

在人脸验证的最后阶段, 往往可以选择一些常用的分类器, 例如用于判断两个脸部图像是否属于同一个人的支持向量机. 这些分类器大都用于计算两个脸部图像的相似性[1–5]. 然而, 这些模型所用的特征基本都是浅层结构的特征. 但是由于 Internet 的发展使得大量数据的获得十分容易, 因此在使用模型时需要大量的数据提供的高维特征. 但是浅层结构不能适应这一任务. 因此为了解决上述问题, 文章提了一种混合卷积神经网络模型对脸部图像进行分类. 整个模型的框架如图 1 所示. 整个模型有以下几个特点.

(1) 从原始像素图像中直接提取视觉特征, 而不是使用传统方法提取特征. 在混合卷积神经网络模型中特征首先从已经被训练的卷积神经网络提取出来. 这是因为卷积神经网络在特征提取时有良好的鲁棒性并且可以表现来自不同方面的人脸相似性.

(2) 特征提取后, 模型先对提取到的两个特征进行了特征融合操作. 因为从同一个网络中提取的同一个人的特征可能具有相似性. 在进行特征融合以后可以使这个相似性扩大. 最后使用单变量特征选择和 PCA 来选择有效的特征.

(3) 提取人脸特征时模型首先优化了一个卷积神经网络来进行特征提取操作, 在这个阶段为了确保良好的提取性从而引入识别率. 相对与整个混合卷积神经网络方法, 这种分段训练可以加快整体优化.

1 相关工作

用于人脸验证的所有现有方法都是从两个人脸提取特征开始. 传统方法大都采用浅层结构提取特征. 通常使用各种浅层特征[6,7], 包括 SIFT [8], Gabor[9], Eigenface[10–12]. 还有许多人脸识别模型是浅层结构但使用了高维特征来进行最后的相似性判断[13,14]. 一些方法[15]使用线性 SVM 进行相同或不同的验证决策. Huang、Simonyan 等[13,14]通过学习线性变换来增加图片鲁棒性. 但是所有这些方法的一个主要缺点是它们对输入图像 (移位, 缩放, 旋转) 的几何变换以及面部表情, 眼镜和模糊围巾的其他变化非常敏感. 一些基于浅层网络结构的模型学习高层的特征[16,17], 这种方法与传统的方法不同之处在于, 特征提取和人脸验证是在同一个网络中. 这些网络的结构总是很复杂, 且需要更多的时间优化参数. 与传统方法相比, 它也失去了灵活性. 但是由于互联网的发展, 产生了大量的数据且需要高层的特征. 因此人脸识别模型需要高层特征的从原始图像中学习. 一些作者为人脸验证设计了一些深层次的模型[1–5,18,19]. 但这些模型也失去了浅层模型的灵活性. 所有这些方法都使用卷积神经网络[20]来提取特征并且学习一个相似性度量方法来进行最后的判断. 这是因为卷积网络是可训练的多层非线性系统, 可以以像素级运行, 并且以集成的方式进行高级表示. 虽然这些方法可以提取鲁棒性良好的特征, 但它们没有考虑两个人脸图像的个体之间的差异性. 提取特征以后所有这些模型开始直接分类判断操作. 这可能使得个体之间比较好的特征被忽略, 从而导致整个网络的准确性不高.

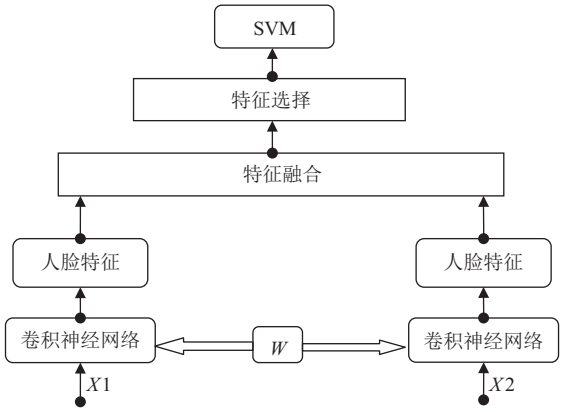


图 1 混合卷积神经网络的框架

本文提出的模型和上述模型有较大的差异, 即考虑浅层模型的灵活性, 也保证了深层模型的有效性. 模型和传统方法一样将人脸验证分成特征提取和人脸验证两个阶段. 这样做的优点是模型可以像传统方法一样学习到较好的人脸特征. 与其他在提取后直接对不同特征进行分类的模型不同, 混合卷积神经网络模型添加了一个特征选择和特征融合步骤, 这样在提取到两个特征以后, 考虑到两个特征之间的相似性, 模型进行一次融合操作使得整个相似性效果明显. 特征融合以后, 模型使用单变量特征选择和主成分分析

© 中国科学院软件研究所 http://www.c-s-a.org.cn

(Principle Component Analysis, PCA) 来选择特征. 它可以增强个体之间的差异性. 单变量特征选择通过选择基于单变量统计检验得到最佳特征, PCA 用于分解一组连续正交分量中的多变量数据集, 其解释了最大量的方差.

2 卷积神经网络

为了提取人脸特征, 混合卷积神经网络使用两个卷积网络模型 (见图 2). 这是一个典型的卷积神经网络由交替卷积和次采样操作组成[21]. 虽然架构的最后阶段由通用的全连接网络组成: 最后几层将是完全连接成一维层特征[22]. 卷积网络是端对端进行训练, 将像素图像映射到输出[10]. 此外, 它可以学习向量不变的局部特征, 因为卷积网络是非线性系统. 提取的特征对于输入图像的几何失真是鲁棒的. 为了保证整个模型提取到的特征效果比较好, 模型先对单个卷积神经网络进行了人脸识别研究, 并保留最佳性能的网络模型架构. 在以下部分中, 将详细描述卷积神经网络的结构.

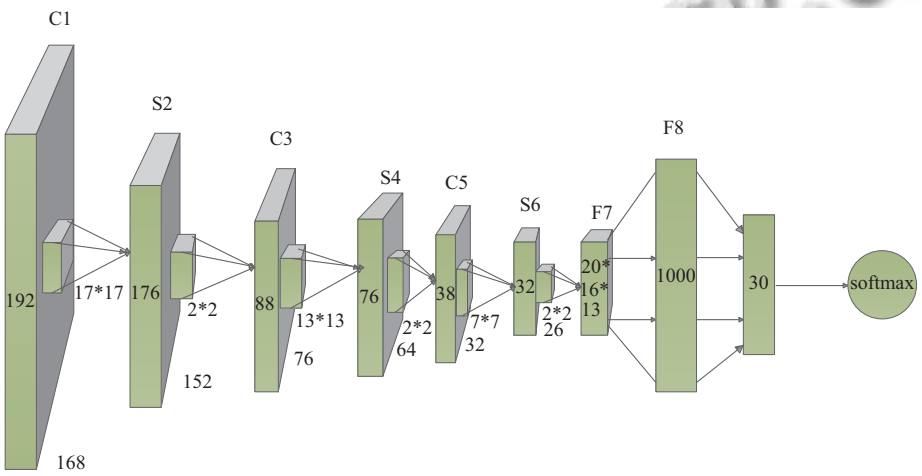


图 2 卷积神经网络模型 (立方体的长度、宽度和高度表示输入层大小)

2.1 混合卷积神经网络模型

整个混合卷积神经网络框架如图 1 所示其中 $X1$ 和 $X2$ 表示两幅人脸的图片. 用 y 表示两幅人脸是否是一个人. 在 y=1 表示是一个人, y=0 表示不是同一个人. 首先训练一个卷积神经网络然后保存模型参数 W, 用来后面提取出每个人脸的特征. 这个模型参数在后面是共享. 这样使得两张人脸图片特征的提取都是用同一个模型. 在提取到特征以后把两个特征融合. 融合以后应用 PCA 进行降维保留主要的特征. 最后把这个特征用支持向量机 (SVM) 进行训练输出判断的结果模型首先从使用已经训练好的卷积网络从两张人脸图像中学习特征. 在网络训练阶段, 为了保证可以提取到较好的特征, 模型引入了识别率进行定量分析. 识别率越高, 模型越好最后提取到的特征也越好. 在保证了整个识别准率以后, 保存整个模型结构进行后续的特征提取操作.

混合卷积神经网络模型和其他方法最大的区别是引入了一个特征融合和特征提取操作. 相比于其他方法往往直接使用提取的特征进行分类, 没有考虑到两张人脸图像的个人异同性. 但在混合卷积神经网络模型中, 使用相加操作来融合以增加最后的特征的异同性. 因为整个模型使用的是同一个卷积神经网络模型进行特征提取操作, 如果是同一个人提取到的特征必然具有相似性, 在经过相加操作以后同一个特征的相似性便放大, 如果不是同一个人经过相加操作以后特征的差异性会更大. 最后在特征融合以后模型使用单变量特征选择和主成分分析来选择特征. 因为在进行融合操作时有可能产生噪声, 进行特征选择的这种方法可以确保最后使用的特征足够好. 特征融合时模型得到一个 1000 维特征. 为了保持维度不变, 在进行单变量选择和 PCA 时各自选择 500 维特征, 最后统一这两个特征从而可以得到 1000 维.

2.2 卷积神经网络模型

模型卷积神经网络使用了典型的卷积网络 (见图 2). Cx 表示卷积层, Sx 表示子采样层, Fx 表示全连接层, 其中 x 为层索引. 基本架构是 C1---S2---C3---S4---

© 中国科学院软件研究所 http://www.c-s-a.org.cn

C5---S6---F7---F8, 具体参数设置如表 1 所示.

表 1 神经网络架构层参数

层	特征图数	输入大小	内核大小	可训练参数	输出大小	连接数
C1	5	192*168	17*17	1450	176*152	38790400
S2	5	176*152	2*2	10	88*76	167200
C3	10	88*76	13*13	1700	76*64	8268800
S4	10	76*64	2*2	20	38*32	60800
C5	15	38*32	7*7	750	32*26	624000
S6	15	32*26	2*2	30	16*13	15600
F7	-	20*16*13	-	21000	1000	21000
F8	-	1000	-	30	30	30

3 实验分析

在上一小节具体分析了整个混合卷积神经网络模型. 这一小节整个模型将在两个人脸数据库上进行实验, 分别是 YaleB 人脸数据库和 AR 人脸数据库. 整个实验环境使用了 keras 在 Windows10 上进行. CPU: i7 6700 Hq, GPU: GTX960m.

3.1 数据处理

混合卷积神经网络模型是在两个人脸数据集的训练和测试 (见图 3). 第一次在 YaleB 人脸数据库上进行实验. YaleB 人脸库由美国耶鲁大学计算视觉与控制中心创建, 包含了 10 个人的 5760 幅多姿态, 多光照的图像. 每个人具有 9 种不同的头部姿态, 每种姿态下均有 64 种不同光照条件的图像, 每幅图像原始大小为 640×480 并且具有 38 个对象, 总共 2470 个图像. 本文只使用了其中的一部分并对图片预处理得到每张图片的大小为 168 * 192. 第二次训练和测试实验在 AR 人脸数据库上进行. AR 人脸库是由西班牙巴塞罗那计算机视觉中心于 1998 年创建, 该人脸库包含了 126 个人的 4000 多幅彩色正面图像, 其中包括 70 名男性和 56 名女性不同面的部表情、光照变化及配饰 (围巾和墨镜遮挡) 等, 在这里模型选择只选取了 2600 张图像, 并进行尺寸为 168 * 192 的预处理.

(a) yaleB 数据集图像

(b) AR 数据集图像

图 3 不同数据库的数据集图像

为了验证脸部图像, 每个图像与数据集中的每个其他图像配对. 在 yaleB 中, 2438 个图像对中有 1215 个正面 (属于一个人). 在 AR 数据集中, 有 2587 个图像对, 其具有 1200 个正面. 测试 (验证) 在 20% 图像对中完成. 在不同数据库上使用卷积神经网络进行验证时的准确率见图 4.

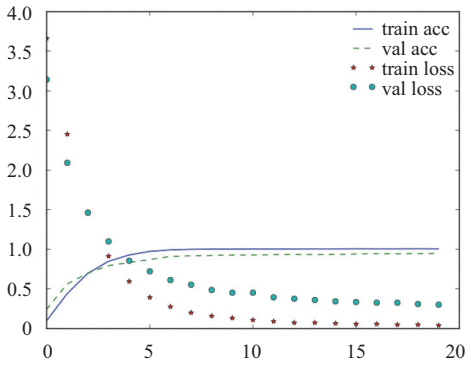
3.2 模型训练

整个混合卷积神经网络模型的框架包括两个网络. 模型的输入是一对人脸图像和标签, 然后通过各自的特征提取网络 (见图 2). 最后, 产生通过特征融合和特征选择阶段输出训练好的特征. 在模型的顶部, 模型使用 SVM 作为分类器来判断两张人脸图片是否是一个人.

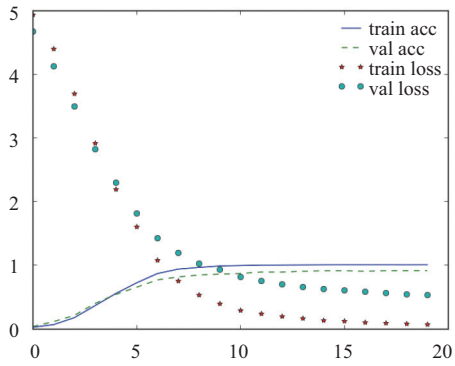
整个训练过程分为两个阶段. 首先, 模型先训练特征提取网络. 其次, 训练模型来验证图像对. 为了保证提取良好的特征, 混合卷积神经网络引入识别率来衡量提取的特征的好坏, 使用数据集来训练卷积网络来识别, 然后保留对整个人脸有较好识别率的网络结构模型.

特征提取后, 模型使用特征融合和特征选择来增强个体之间的差异. 在实验中, 模型从两张人脸图像中得到一个 1000 维特征. 之后模型使用单变量特征选择和 PCA 来选择已融合的特征. 这样做的优点是减少特征合并中出现的噪音. 经过上述操作, 我们可以得到一个很好的特征. 图 5 显示了提取的特征.

© 中国科学院软件研究所 http://www.c-s-a.org.cn



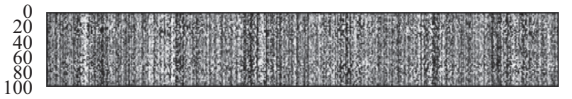
(a) 在 YaleB 数据集上验证的准确率



(b) 在AR数据库上验证的准确率

图 4 在不同数据库上使用卷积神经网络进行验证时的准确率

(a) 人脸对中前 100 张人脸的特征



(b) 前 100 张人脸对应人脸特征

(c) 经过特征融合和特征选择以后得到的特征

图 5 前 100 张人脸特征图

在网络的顶端, 混合卷积神经网络使用 SVM 作为分类器. 因为人脸图像是非线性的, 所以模型使用高斯核函数. 其中高斯核函数为

$$K(x,z)=\exp\left(-\frac{\|x-z\|}{2\sigma^2}\right)$$

其对应于 SVM 的高斯核函数是径向基函数分类器. 在这种情况下, 分类器的函数为

$$f(x)=\text{sign}\left(\sum_{i=1}^{N_s}a_i^*y_i\exp\left(-\frac{\|x-z\|^2}{2\sigma^2}\right)+b^*\right)$$

为了确保公平的比较, 模型使用两个数据库. 首先, 使用两个数据集的系统来验证其有效性 (见表 2). 从表中可以看出, CNN 识别率对最终验证率有影响. 在表 2 的 YaleB 部分, 可以看到, 随着卷积神经网络的准确率的提高, 人脸验证的准确性也得到提高. 在表 2 的 AR 部分, 可以发现, 即使 AR 数据集有很大的变化, 所提出的方法也可以得到很好的结果. 在表 3 中, 混合卷积网络方法与一些传统方法进行比较. 统计数据显示, HBCNN 在不同的维度上取得了较好的效果.

表 2 在 YaleB 和 AR 人脸数据库下 HBCNN 模型对不同 CNN 识别率下的精度比较 (单位: %)

	CNN识别率	Verification识别率
YaleB	87.5	96.8
	92.1	97.9
	95.5	99.4
AR	86.4	88.6
	90.5	90.2
	91.4	92.5

表 3 HBCNN 模型在 YaleB 数据集下与传统方法的精度比较 (D 指维度)(单位: %)

	1000D	1200D	1600D
PCA+SVM	62.3	60.9	62.4
HCNN+SVM	99.4	98.6	97.2
PCA+Bayes	92.3	90.8	89.2

4 总结

本文提出了一种用于人脸验证的混合卷积神经网络模型. 该模型直接从人脸图片中学习并提取特征. 最后模型在两个不同的数据集下进行了实验, 并且实验证明整个模型在两个人脸数据库上都有比较好的应用. 最后相对于传统的的方法混合卷积神经网络也有较好的效果.

参考文献

1 Sun Y, Wang XG, Tang XO. Hybrid deep learning for face

© 中国科学院软件研究所 http://www.c-s-a.org.cn

verification. Proceedings of the IEEE International Conference on Computer Vision. Sydney, NSW, Australia. 2013. 1489–1496.
2 Sun Y, Wang XG, Tang XO. Deep learning face representation from predicting 10,000 classes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 1891–1898.
3 Sun Y, Chen YH, Wang XG, *et al*. Deep learning face representation by joint identification-verification. Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada. 2014. 1988–1996.
4 Sun Y, Wang XG, Tang XO. Deeply learned face representations are sparse, selective, and robust. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 2892–2900.
5 Taigman Y, Yang M, Ranzato MA, *et al*. Deepface: Closing the gap to human-level performance in face verification. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 1701–1708.
6 Guillaumin M, Verbeek J, Schmid C. Is that you? Metric learning approaches for face identification. Proceedings of the 12th International Conference on Computer Vision. Kyoto, Japan. 2009. 498–505.
7 Nguyen HV, Bai L. Cosine similarity metric learning for face verification. Proceedings of the 10th Asian Conference on Computer Vision. Queenstown, New Zealand. 2010. 709–720.
8 Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91–110. [doi: 10.1023/B:VISI.0000029664.99615.94]
9 Wiskott L, Krüger N, Kuiger N, *et al*. Face recognition by elastic bunch graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 775–779. [doi: 10.1109/34.598235]
10 Turk M, Pentland A. Eigenfaces for recognition. Journal of Cognitive Neuroscience, 1991, 3(1): 71–86. [doi: 10.1162/jocn.1991.3.1.71]
11 Yang MH, Ahuja N, Kriegman D. Face recognition using kernel eigenfaces. Proceedings of the 2000 International Conference on Image Processing. Vancouver, BC, Canada. 2000, 1. 37–40.
12 Belhumeur PN, Hespanha JP, Kriegman DJ. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711–720. [doi: 10.1109/34.598228]
13 Simonyan K, Parkhi OM, Vedaldi A, *et al*. Fisher vector faces in the wild. Proceedings of the British Machine Vision Conference (BMVC). Bristol, UK. 2013.
14 Huang C, Zhu SH, Yu K. Large-scale strongly supervised ensemble metric learning. US Patent 8873844. [2014-10-28].
15 Huang GB, Lee H, Learned-Miller E. Learning hierarchical representations for face verification with convolutional deep belief networks. Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA. 2012. 2518–2525.
16 Kumar N, Berg AC, Belhumeur PN, *et al*. Attribute and simile classifiers for face verification. Proceedings of the 12th International Conference on Computer Vision. Kyoto, Japan. 2009. 365–372.
17 Berg T, Belhumeur PN. Tom-vs-pete classifiers and identity-preserving alignment for face verification. Proceedings of the British Machine Vision Conference (BMVC). Surrey, Canada. 2012.
18 Shao H, Chen S, Zhao JY, *et al*. Face recognition based on subset selection via metric learning on manifold. Frontiers of Information Technology & Electronic Engineering, 2015, 16(12): 1046–1058.
19 Li QF, Zhou XF, Gu AH, *et al*. Nuclear norm regularized convolutional Max Pos@Top machine. Neural Computing & Applications, 2016: 1–10. [doi: 10.1007/s00521-016-2680-2]
20 Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, USA. 2012. 1097–1105.
21 LeCun Y, Bottou L, Bengio Y, *et al*. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11): 2278–2324. [doi: 10.1109/5.726791]
22 Bouvrie J. Notes on convolutional neural networks. Neural Nets, 2006. https://pdfs.semanticscholar.org/2a43/93aa1bc3cb7fe2deecc88720bfb84dabb263.pdf.

© 中国科学院软件研究所 http://www.c-s-a.org.cn