



# 基于概念器的深度神经网络模型

钱光武, 张蕾\*, 王炎

四川大学计算机学院机器智能实验室, 成都 610065

\* 通信作者. E-mail: leizhang@scu.edu.cn

收稿日期: 2018-01-24; 接受日期: 2018-02-25; 网络出版日期: 2018-05-11

国家重点研发计划 (批准号: 2016YFC0801800)、国家自然科学基金 (批准号: 61772353, 61332002) 和霍英东基金高等院校青年教师基金基础性研究课题 (批准号: 151068) 资助项目

**摘要** 近年来, 深度神经网络, 亦被称为深度学习, 在机器学习方法主导的各个领域都取得了重大的突破. 虽然经过训练的深度神经网络具有卓越的性能, 但是整个训练过程却十分耗时, 即使借助高性能计算设备, 也需要数日甚至数周的训练时间. 概念器作为回音状态网络的延续和发展, 可以理解为描述神经动态活动模式的过滤器, 是一个强大的时序数据处理工具. 为了解决上述问题, 基于对原始概念器模型的改进, 本文在深度神经网络的非迭代方法和迁移学习两个方面分别做了一些工作. 具体来说, (1) 提出了针对非时序数据的概念器分类器, 并在此基础上提出了一种非迭代方法前馈卷积概念器神经网络, 通过在 MNIST 变集数据集上的实验测试了前馈卷积概念器神经网络的分类性能, 不仅达到了同类方法的最高水平, 而且极大地降低了训练时间; (2) 提出了一种基于概念器的快速概念器分类器, 在数据集 Caltech-101 和 Caltech-256 上, 测试了快速概念器分类器结合预训练且不再微调的深度神经网络的表现, 不仅在性能上超越了同类方法的最高水平, 而且训练时间平均减少到原有的 1/60.

**关键词** 概念器, 图像分类, 深度神经网络, 迁移学习, 非迭代方法

## 1 引言

近十年来, 深度神经网络 (deep neural network, DNN) 充斥了整个机器学习的领域<sup>[1~6]</sup>, 特别是卷积神经网络 (convolutional neural network, CNN) 在大规模图像和视频识别上取得了巨大的成功<sup>[4,7,8]</sup>, 同时也在多媒体处理上取得了长足的进步<sup>[9~11]</sup>. 最为常用的 DNN 的训练算法是反向传播算法 (backpropagation, BP)<sup>[12]</sup>, 但是误差反向传播多层的做法与生物学发现相去甚远. 此外, 在实际部署 DNN 的实践中还存在着许多障碍, 如果要达到世界顶尖水平的性能, 对计算架构的规模、计算量和实施者的专业素质都有很高要求. 以上提到的两个原因促使人们去寻找不依赖反向传播的深度神经网络.

**引用格式:** 钱光武, 张蕾, 王炎. 基于概念器的深度神经网络模型. 中国科学: 信息科学, 2018, 48: 511-520, doi: 10.1360/N112017-00261  
Qian G W, Zhang L, Wang Y. Conceptor-based deep neural networks (in Chinese). Sci Sin Inform, 2018, 48: 511-520, doi: 10.1360/N112017-00261

研究者们提出了许多方法尝试解决这个问题, 例如非迭代方法<sup>[13~15]</sup>和迁移学习<sup>[8, 16~18]</sup>, 两者的核心思路都是简化训练过程和提升训练速度. 非迭代方法通过移除训练中的迭代步骤加速训练过程, 通常使用具有解析解的算法来替代原有的迭代算法. 这些方法一定程度上解决了问题, 但是却不可避免带来了精度的下降, 导致人们往往要在速度和精度上做一个权衡选择. 迁移学习则是利用已经在其他数据集上训练完成的深度神经网络模型, 稍作修改来处理新数据集. 在实际应用中, 从零开始训练一个随机化初始化的卷积神经网络并不常见, 因为训练数据的不充分难以发挥网络的最大性能. 反之, 更为常见的做法是在一个类似 ImageNet 数据集的大数据集上训练卷积神经网络, 并以此为基础进行后续工作, 如将此时的网络作为一个初始化状态, 或者直接将此网络当作一个特征提取工具. 实验发现在 ImageNet<sup>[19]</sup>数据集上学习到的深度图像表达可以泛化到其他数据集<sup>[18]</sup>, 因此引起了大量研究者的兴趣和进一步工作<sup>[8, 20~22]</sup>.

本文为了探索不依赖反向传播算法的 DNN 模型, 从非迭代方法和迁移学习两方面介绍我们的研究进展. 在非迭代方法方面, 基于概念器 (conceptor) 模型, 提出了针对非时序数据的概念器分类器, 且结合此分类器提出了一种非迭代方法——前馈卷积概念器神经网络 (feedforward convolutional conceptor neural network, FCCNN). 这种神经网络在多个方面受到了回声状态网络 (echo state network)<sup>[23]</sup>和概念器<sup>[24]</sup>的启发, 不仅结构简单而且训练时间短. 前馈卷积概念器神经网络的第一部分层级是卷积层, 使用主成分分析 (principal component analysis, PCA)<sup>[25]</sup>提取的特征, 不仅是无监督的而且训练成本低, 同时没有使用池化层. 然后, 提取的特征进行了二值化, 这是一个关键的特性, 将会在后面详细介绍. 通过卷积层后, 二值化的数据输入到一个新型的分类层, 这个分类层包含了非时序的概念器. 基于这个架构, 我们在一些基于 MNIST 的图像识别任务中, 取得了同类非迭代方法的最高性能水平.

在迁移学习方面, 提出了一种不仅结构简单而且易于 GPU 加速的快速概念器分类器 (fast conceptor classifier, FCC), 并配合在 ImageNet 数据集上预训练过且不再微调的网络 VGG-16 Net<sup>[18]</sup>, Resnet-50 和 Resnet-152<sup>[26]</sup>, 在数据集 Caltech-101 和 Caltech-256 上评估了其性能. 快速概念器分类器同样也是受到了回声状态网络和概念器的启发, 本质上是一个二次分类器, 所以训练时间比支持向量机<sup>[27]</sup>和 Softmax 回归分类器更短. 此外, 快速概念器分类器易于构造且是一个近乎无参的方法, 实际上只有一个称为“孔径”的参数需要训练. 实验结果显示了无论预训练神经网络的结构如何, FCC 都提高了各个预训练神经网络的分类准确率. 基于以上分析, 我们认为 FCC 不仅能够提高预训练神经网络的图像分类准确率, 而且加快了训练速度.

## 2 工作基础

概念器模型是我们一系列工作的基础, 但是作为一个新兴工具, 它远不如 DNN 模型那样广为人知, 因此本节首先介绍和我们工作相关的基本内容. 概念器是作为一个回复神经网络 (recurrent neural network, RNN) 的工具被提出<sup>[24]</sup>, 可以理解为提取时序神经激活模式特征的过滤器. 概念器以投影矩阵的形式, 实现对高维网络信号的方向性选择阻尼. 截至目前, 概念器仅应用于处理时序信号, 但是本文构造了可以处理非时序信号的概念器模型.

一个概念器是一个矩阵方阵  $C \in \mathbb{R}^{m \times m}$ , 其将  $m$  维输入向量  $x$  线性变化到  $m$  维向量  $y$ , 即  $y = Cx$ . 概念器矩阵的定义目的是尽量复制输入, 且保证矩阵元素很小, 这就形成了经验代价函数:

$$\mathcal{L} = \|X - CX\|_{\text{fro}}^2 + \alpha^{-2} \|C\|_{\text{fro}}^2, \quad (1)$$

其中  $X \in \mathbb{R}^{m \times n}$  的第  $l$  列是输入样本  $x_l$ ,  $\alpha \in (0, \infty)$  是一个平衡参数, 在概念器上称为“孔径”. 式 (1) 中的矩阵范数是弗罗贝尼乌斯范数 (Frobenius norm) 或希尔伯特-施密特范数 (Hilbert-Schmidt norm). 从而, 对该代价函数最小化求解可得<sup>[24]</sup>

$$C = R(R + \alpha^{-2}I)^{-1}, \quad (2)$$

其中  $R = \frac{XX^T}{n} \in \mathbb{R}^{m \times m}$  是输入数据的关联矩阵,  $I$  是单位矩阵. 概念器矩阵都是半正定且特征值 (半正定矩阵的奇异值) 最大为单位尺寸 1. 一个由实际数据计算得到的概念器矩阵通常有一个较低的秩. 这种情况意味着概念器矩阵可以通过更加经济的紧凑奇异值分解 (compact singular value decomposition, CSVD) 表示, 这可以节约计算时间.

在多分类的情况下, 输入数据则按类别收集起来. 例如,  $X_j \subset X$  表示来自第  $j$  ( $j = 1, \dots, K$ ) 类的所有样本, 相应的关联矩阵为  $R_j = \frac{X_j X_j^T}{n_j}$ . 对于每个分类, 可以得到一个单独的概念器矩阵  $C_j = R_j(R_j + \alpha_j^{-2}I)^{-1}$ . 给定来自  $K$  个类的的数据, 可计算得到  $K$  个概念器矩阵  $C_1, C_2, \dots, C_K$ , 于是可用它们来分类一个测试样本  $x \in \mathbb{R}^m$ . 首先得到  $K$  个正证据 (positive evidences)  $E^+(x, i)$ <sup>[24]</sup>, 正证据是一个正数, 借由每一个概念器可计算得到一个, 此值越高则该测试样本属于相应类别的概率越大. 然后, 通过  $K$  个正证据中的最大值来预测测试样本的所属类别, 其计算方式如下:

$$c = \operatorname{argmax}_i E^+(x, i) = \operatorname{argmax}_i x^T C_i x. \quad (3)$$

### 3 基于概念器的深度神经网络模型

#### 3.1 前馈卷积概念器神经网络 (FCCNN)

图 1 展示了一个高度抽象的模型流程图, 形象地描述了 FCCNN 如何工作. 图中的平行四边形表示数据, 而矩形则表示处理步骤. 在训练阶段的每个卷积步骤中, 通过 PCA 从输入生成卷积核, 输入再与生成的卷积核进行卷积操作<sup>[14]</sup>. 经过卷积层之后, 对提取的特征进行二值化处理并分成  $K$  组, 这里的  $K$  是根据训练集得到的类别数量. 最后经过计算, 得到作为分类器的  $K$  个概念器矩阵并保存下来. 在测试阶段, 所有的测试数据都经过与训练阶段相同的卷积层, 并使用训练阶段得到的卷积核. 提取的测试特征也被二值化, 并通过之前存储的概念器矩阵计算相应的正证据. 最后, 预测的类别就能够通过最大的正证据得到.

##### 3.1.1 表达学习

卷积核的学习过程理应在卷积操作之前. 由于要避免反向传播, 我们就不能使用以误差驱动的卷积核学习过程, 如自动编码器学习 (autoencoder learning) 等, 因此我们选择了更加直接的主成分卷积核学习. 经过每层卷积层后, 对提取的特征应用双曲正切函数以提高非线性, 并将特征值缩放到区间  $(-1, 1)$ . 虽然池化层具有很多的优点, 如更好的不变性表达和更低的特征维度, 但是我们认为因移除池化层而增加的冗余信息更有助于非迭代模型中的分类器. 因此, 在 FCCNN 中并未使用池化层. 表达学习的最后一个步骤有一个后处理过程, 称为阈值二值化, 即根据阈值将提取的特征二值化为  $\{-1, 1\}$ , 其中的阈值是可调节的参数.

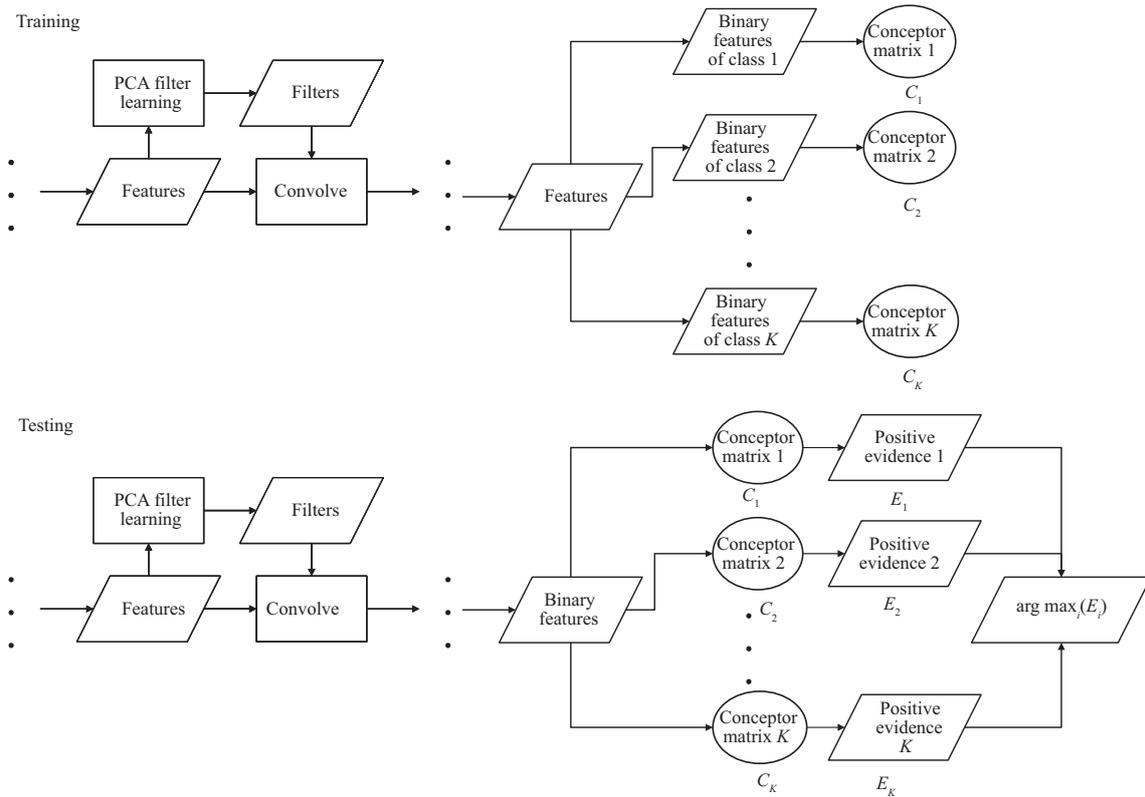


图 1 FCCNN 分类过程图  
Figure 1 The flowchart of FCCNN

### 3.1.2 分类预测

完成表达学习后, 我们利用学习到的特征来训练概念器并进行分类预测. 对于大多数分类器来说, 在提取的特征上进行分类要比在原始数据分类性能更好. 这主要有两方面原因, 一是因为若特征维度更高, 那么可区分性通常也越高, 即更容易被分类; 二则是经过特征提取后, 不同类别的数据被投影到了各自的子空间上, 也会更加有利于分类. 在训练阶段, 通过式 (2) 计算得到概念器矩阵, 预测类标则由式 (3) 得到.

## 3.2 快速概念器分类器 (FCC)

鉴于概念器在非迭代方法上的良好表现, 我们进一步研究了概念器在迁移学习上的可行性, 并提出了快速概念器分类器 FCC. 一个基于 FCC 的分类框架主要由两部分组成, 一个是预训练的 DNN 来进行特征提取, 另一个是 FCC 进行分类预测, 如图 2 所示.

### 3.2.1 特征提取

特征提取最为直接的方式是固定预训练的卷积神经网络权重, 并将数据输入到网络, 将倒数第二层 (最后一层为分类层) 的输出作为特征. 因此对于网络结构需要做出些许调整, 包括移除最后一层全连接层, 此层主要实现 1000 类的 ImageNet 分类, 并将此层替换为快速概念器分类器. 当所选预训练网络为 Resnet-50 和 Resnet-152 的时候, 我们使用了上述的方式改造了网络结构. 但是, VGG 网络

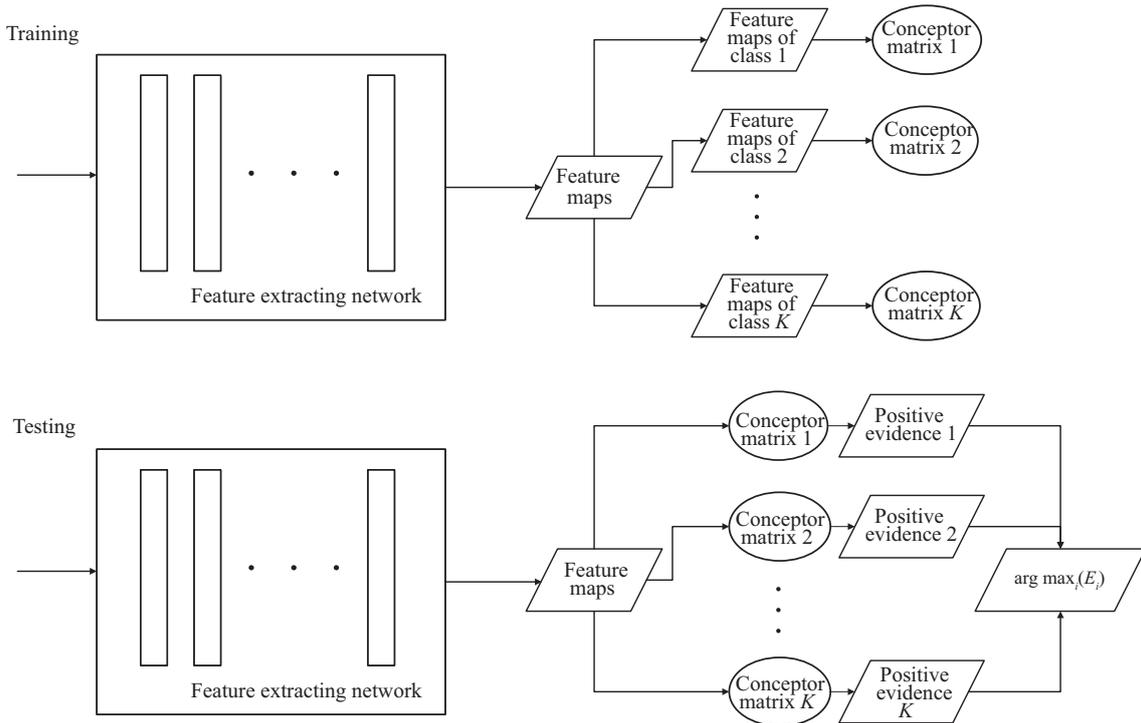


图 2 FCC 分类过程图  
Figure 2 The flowchart of FCC

为了集合多尺度特征, 需要把全连接层转换为卷积层, 即把网络改造为全卷积网络 (fully convolutional network, FCN) [28]. 传统的卷积神经网络要求输入图像的尺寸为固定值, 这样的要求不仅限制了图像的长宽比, 还限制了输入图像的尺度. 当输入图像的尺寸不固定时, 卷积神经网络需要一个预处理步骤把输入图像处理成相同尺寸. 这个预处理过程可以是裁剪 (cropping) [4, 8]、形变 (warping) [21, 29] 或者更加简单的缩放, 但是这些预处理会破坏原始图像的长宽比或尺度, 造成信息丢失. 全卷积网络可以轻松解决这个技术问题, 因为全卷积网络中只使用了卷积和池化操作, 所以可以处理任意尺寸的图像. VGG 网络在 Caltech-101 和 Caltech-256 数据集的实验中使用了这样的改造方式, 并取得了最高水平的分类准确率.

### 3.2.2 分类预测

为了将原始概念器矩阵求解公式中隐含的奇异值分解的计算时间控制在合理范围, 我们使用紧凑奇异值分解, 关联矩阵可重写为

$$\begin{cases} X = U_r \Sigma_r V_r^T, \\ R = \frac{X X^T}{n} = \frac{U_r \Sigma_r^2 U_r^T}{n} = U_r D_r U_r^T, \end{cases} \quad (4)$$

其中  $D_r = \Sigma_r^2$ ,  $r$  是矩阵  $X \in \mathbb{R}^{m \times n}$  的秩. 在实验中, 因为大部分时候矩阵  $X$  是满秩矩阵, 而每个类的训练样本数目远小于特征的维度, 所以  $r = \min(m, n)$  通常由训练样本个数决定, 概念器矩阵的求解方法可重写为

$$C(R, \alpha) = U_r [D_r (D_r + \alpha^{-2} I_r)^{-1}] U_r^T = U_r S_r U_r^T, \quad (5)$$

表 1 不同方法在 MNIST 变集上的错误率和在 bg-img-rot 上的训练时间

Table 1 Error rates of different methods on MNIST variations and corresponding training time on bg-img-rot<sup>a)</sup>

Method	Basic	Rot	Bg-rand	Bg-img	Bg-img-rot	Training time
CAE-2 <sup>[31]</sup>	2.48	9.66	10.9	15.5	45.23	>3 h
TIRBM <sup>[32]</sup>	–	<b>4.2</b>	–	–	35.5	>3 h
PGBM+DN-1 <sup>[33]</sup>	–	–	<b>6.08</b>	12.25	36.76	>3 h
ScatNet-2 <sup>[13]</sup>	1.27	7.48	18.4	12.3	50.48	–
PCANet-2 <sup>[14]</sup>	<b>1.06</b>	7.37	6.19	10.95	35.48	15 min
FCCNN	2.43	8.91	6.45	<b>10.8</b>	<b>33.6</b>	5 ~ 30 min

a) The best result in current column is shown in bold font.

其中  $S_r = D_r(D_r + \alpha^{-2}I_r)^{-1}$ . 正证据的计算公式可重写为

$$E^+(x, j) = x^T C_j x = x^T U_{r_j} S_{r_j} U_{r_j}^T x, \quad (6)$$

其中  $j$  表示第  $j$  个类,  $x$  是一个测试样本. 实际上作为分类器, 概念器矩阵迹仅在测试阶段使用. 因此, 相较于训练后保存一个巨大的概念器矩阵, 仅需保存紧凑奇异值分解的结果  $U_{r_j}$  和  $S_{r_j}$ . 这样既可以避免在训练阶段生成巨大的矩阵, 又能在测试阶段快速计算正证据.

## 4 实验结果

### 4.1 FCCNN 在 MNIST 变集上的实验结果

本文在 MNIST 变集数据集 (MNIST variation datasets)<sup>[30]</sup> 上对 FCCNN 的性能进行了验证. 该数据集是一个广泛使用的测试层级表达的基准数据. MNIST 变集是由 10 类灰度手写数字图片构成, 每张图片的大小是  $28 \times 28$  像素. 整个数据集包含 5 个子类: (1) Basic, 原始 MNIST 数据集的子集; (2) Rot, 对数字增加了随机的旋转; (3) Bg-rand, 数字的背景增加了随机的噪声; (4) Bg-img, 数字的背景不再是黑色, 而是随机的图像; (5) Bg-img-rot, 数字图像既被旋转又将随机的图像作为背景<sup>1)</sup>. 对于数据集的每一个子类, 10000 张图片作为训练集, 2000 张图片作为验证集, 以及 50000 张图片作为测试集.

本文在整个 MNIST 变集上对所实现的 FCCNN 进行了实验, 其中二值化阈值的取值范围为  $[-0.2, 0.2]$ , 孔径的取值范围为  $[0.1, 1.5]$ , 均通过验证集搜索适合当前数据集的最优值, 结果如表 1 所示. 由于未采用反向传播算法且网络结构简单, 因此相较于其他非迭代方法, 网络的训练速度极快. 此外, 大部分 FCCNN 的训练时间都是被表达学习过程所占, 这意味着可以通过安装更多内存来加速.

在对比实验中, 我们对比了代表最高水平的非迭代方法 (ScatNet-2<sup>[13]</sup> 和 PCANet-2<sup>[14]</sup>)、典型的迭代方法 (CAE-2<sup>[31]</sup>、TIRBM<sup>[32]</sup> 和 PGBM+DN-1<sup>[33]</sup>) 和 FCCNN. 从表 1 中可以观察到 FCCNN 在所有方法中排名随着数据集的复杂度增加而提高. 特别地, FCCNN 在最后两个最复杂的数据集上达到了最好的性能. 此外, 我们方法得到的特征维度只是 PCANet 的三分之一, 而且其结构更加简单, 其主要模块包含一个典型的 CNN 和一个经过裁剪的非时序数据概念器分类器. PCANet 网络结构的前部分和我们网络的前半部分相似, 主要的差异在后半部分. PCANet 其后的主要操作是块状直方图 (block-wise histogram), 该操作不仅复杂而且大幅度提高了最终特征的维度, 从而保证末端线性 SVM

1) 这些数据集可以在这里找到: <http://www.iro.umontreal.ca/~lisa/icml2007>.

表 2 在数据集 Caltech-101 和 Caltech-256 上的分类准确率

Table 2 Classifying accuracies on Caltech-101 and Caltech-256

Method	Caltech-101	Caltech-256
Zeiler & Fergus [8]	86.5	74.2
Chatfield et al. [20]	88.4	77.6
He et al. [16]	93.4	-
VGG-16 Net [18]	91.8	84.57
Resnet-50 [26]	92.65	82.43
Resnet-152 [26]	95.23	90.24
FCC(VGG-16 Net)	91.87	84.67
FCC(Resnet-50)	93.08	82.81
FCC(Resnet-152)	<b>95.55</b>	<b>90.87</b>

的分类性能. 相反, FCCNN 无需复杂的块状直方图操作, 在经过 CNN 提取特征后, 直接将特征输入到非时序概念器分类器, 即可得到与 PCANet 不相上下的分类性能.

从表 1 最后一列, 我们可以观察到在数据集 bg-img-rot 上, 所有迭代方法的训练时间平均十倍于非迭代方法, 且 FCCNN 的训练时间与 PCANet 几乎一样. 需要注意的是, 不同参数的设置会导致不同的训练时间, 虽然对于 FCCNN 的各种参数配置都可以在 bg-img-rot 数据集上取得超越 PCANet 的性能, 但是表 1 中只记录了最好的结果, 而这个结果自然耗时也最久, 约为 30 min, 即 FCCNN 训练时间的上限.

#### 4.2 FCC 在 Caltech-101 和 Caltech-256 上的实验结果

Caltech-101 包含了来自 102 个类别共计 9144 张含类标的图片, 包括 101 个物体分类和 1 个背景分类. Caltech-256 是一个更大的数据集, 包含了来自 256 个类别共计 30608 张含类标的图片. 在 Caltech-101 数据集中, 每类有 30 张随机选取的图片作为训练集, 从余下的图片中随机选取至多 50 张作为测试集. 在 Caltech-256 数据集中, 每类有 60 张随机选取的图片作为训练集, 余下的图片作为测试集. 本小节所有实验的性能评价标准指标为准确率, 即预测正确的测试样本个数与测试样本总数的百分比.

VGG 网络是非常成功的深度卷积网络, 其中表现最好的两个公开模型是 VGG-16 Net 和 VGG-19 Net. 由于 VGG-16 Net 和 VGG-19 Net 在分类性能和结构上都十分相似, 因此实验中仅对 VGG-16 Net 进行了评估, 其中所使用的唯一预处理步骤是每个像素减去基于训练集计算出来的 RGB 均值. 同时, 和文献 [18] 中类似, 不仅是从倒数第二层提取了 4096 维特征, 而且集合了多尺度的信息. 实验使用了 3 个尺度  $Q \in \{256, 384, 512\}$ , 这个值表示将图片的短边缩放到这个长度, 而图片的长边则按照原始图像的长宽比自动缩放, 最终得到长宽比相同但是尺寸不同的 3 幅图片.

同时, 为了验证 FCC 是否能够用于深层网络, 我们也将其结合 Resnet-50 和 Resnet-152 进行了测试. 由于在 Resnet 网络上, 多尺度的引入并未带来明显的性能提升, 因此只使用了尺度 224. 将 FCC 和其他在 Caltech-101 和 Caltech-256 数据集上取得最高水平的方法进行比较, 对比实验结果如表 2 所示.

从表 2 中可以看出, 在数据集 Caltech-256 上, FCC 结合 VGG-16 Net 在性能上超越了原始的 VGG 网络 [18], 而在数据集 Caltech-101 上, FCC 结合 VGG-16 Net 非常具有竞争力, 而且也比原始的

表 3 VGG-16 Net, Resnet-50 和 Resnet-152 结合不同分类器的运算时间 (s)  
 Table 3 Running time of VGG-16 Net, Resnet-50 and Resnet-152 with different classifiers (s)

Method	Caltech-101		Caltech-256	
	Training time	Testing time	Training time	Testing time
VGG-16 Net	118.31	118.28	2345.07	3114.48
FCC(VGG-16 Net)	1.76	65.2	26.16	1103.25
Resnet-50	16.03	21.59	82.43	554.12
FCC(Resnet-50)	0.33	15.19	2.64	220.99
Resnet-152	13.07	20.24	229.93	497.93
FCC(Resnet-152)	0.32	15.33	2.73	223.21

VGG-16 Net 要更好. 由此可见 FCC 可以有效提高预训练网络的分类性能. 在 Caltech-101 和 Caltech-256 数据集上, FCC 结合 Resnet-152 达到了最高水平的分类准确率. 我们认为这样突出的结果也受益于 Resnet-152 更泛化的特征提取性能, 因为本文所使用的预训练 Resnet-152 是在一个具有约 11000 个分类的更大的 ImageNet 数据集上训练的模型, 而其他网络均是在 1000 个分类的 ImageNet 数据集上进行训练.

除了优异的分类性能, FCC 的运行时间也让人印象深刻. 在 Intel<sup>®</sup> Xeon<sup>®</sup> CPU E5-2620 v3 @ 2.40 GHz 配置上, VGG-16 Net, Resnet-50 和 Resnet-152 结合不同分类器 (FCC 或线性 SVM) 的运算时间如表 3 所示. 因为对比方法之间的差异仅存在于分类器部分, 所以实验仅统计了分类器的运算时间.

从表 3 中可以看出, FCC 的训练时间极短, 平均被缩减到 1/60. 因为训练阶段的主要步骤是 CSVD, 作用于第  $j$  类的样本集合矩阵  $X_j \in \mathbb{R}^{m \times n_j}$ , 其中  $m \gg n_j$ , 所以消耗的时间非常少. 同时对于测试时间, 在最坏的情况下, 也仅是线性 SVM 的一半. 除了本身运算量更少外, 由于 FCC 的主要运算属于矩阵运算, 可以使用诸如 Intel<sup>®</sup> Math Kernel Library (MKL), Linear Algebra PACKage (LAPACK) 等高度优化的函数库求解.

## 5 结论

从反向传播算法模型在生物学和计算效率两方面的不合理性出发, 从非迭代方法和迁移学习两个主流解决思路入手, 尝试解决问题并做了一些工作. 主要是建立不依赖反向传播算法的 DNN 模型, 基于概念器这一最新的研究成果, 根据不同需要提出了数种改进模型, 并与 DNN 进行整合, 设计了相应的新型模型. 通过在数据集 MNIST 变集、Caltech-101 和 Caltech-256 上的实验, 我们验证了模型的可行性并获得了显著的性能提升. 作为一门新工具, 概念器还有许多功能值得进一步开发. 今后, 我们计划利用概念器的独特功能, 研究改进深度网络的新方法, 解决现有模型中的一些问题, 从而达到更高的精度或者效率.

## 参考文献

- 1 Zhang L, Zhang Y. Big data analysis by infinite deep neural networks. J Comput Res Dev, 2016, 53: 68–79 [张蕾, 章毅. 大数据分析的无限深度神经网络方法. 计算机研究与发展, 2016, 53: 68–79]
- 2 Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. Neural Comput, 2006, 18: 1527–1554

- 3 Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, 313: 504–507
- 4 Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*, Lake Tahoe, 2012. 1097–1105
- 5 Wan L, Zeiler M, Zhang S, et al. Regularization of neural networks using dropconnect. In: *Proceedings of the 30th International Conference on Machine Learning*, Atlanta, 2013. 1058–1066
- 6 Zhang L, Yi Z, Amari S. Theoretical study of oscillator neurons in recurrent neural networks. *IEEE Trans Neural Netw Learn Syst*, 2018, 99: 1–7
- 7 Sermanet P, Eigen D, Zhang X, et al. Overfeat: integrated recognition, localization and detection using convolutional networks. ArXiv: 1312.6229
- 8 Zeiler M D, Fergus R. Visualizing and understanding convolutional networks. In: *Proceedings of the European Conference on Computer Vision*. Berlin: Springer, 2014. 818–833
- 9 Li Z C, Tang J H. Weakly supervised deep metric learning for community-contributed image retrieval. *IEEE Trans Multim*, 2015, 17: 1989–1999
- 10 Li Z C, Tang J H. Weakly supervised deep matrix factorization for social image understanding. *IEEE Trans Image Process*, 2017, 26: 276–288
- 11 Tang J, Shu X, Qi G J, et al. Tri-clustered tensor completion for social-aware image tag refinement. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39: 1662–1674
- 12 Rumelhart D E, Hinton G E, Williams R J, et al. Learning representations by back-propagating errors. *Nature*, 1986, 323: 533–536
- 13 Bruna J, Mallat S. Invariant scattering convolution networks. *IEEE Trans Pattern Anal Mach Intell*, 2013, 35: 1872–1886
- 14 Chan T H, Jia K, Gao S, et al. Pcanet: a simple deep learning baseline for image classification? *IEEE Trans Image Process*, 2015, 24: 5017–5032
- 15 Qian G W, Zhang L. A simple feedforward convolutional conceptor neural network for classification. *Appl Soft Comput*, 2017
- 16 He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. In: *Computer Vision — ECCV 2014*. Berlin: Springer, 2014. 346–361
- 17 Qian G W, Zhang L, Zhang Q J. Fast conceptor classifier in pre-trained neural networks for visual recognition. In: *Advances in Neural Networks — ISNN 2017*. Berlin: Springer, 2017. 290–298
- 18 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. ArXiv: 1409.1556
- 19 Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. *Int J Comput Vision*, 2015, 115: 211–252
- 20 Chatfield K, Simonyan K, Vedaldi A, et al. Return of the devil in the details: delving deep into convolutional nets. ArXiv: 1405.3531
- 21 Donahue J, Jia Y, Vinyals O, et al. Decaf: a deep convolutional activation feature for generic visual recognition. In: *Proceedings of the 31st International Conference on Machine Learning*, Beijing, 2014. 647–655
- 22 Razavian A S, Azizpour H, Sullivan J, et al. CNN features off-the-shelf: an astounding baseline for recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Columbus, 2014. 806–813
- 23 Jaeger H, Haas H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science*, 2004, 304: 78–80
- 24 Jaeger H. Using conceptors to manage neural long-term memories for temporal patterns. *J Mach Learn Res*, 2017, 18: 1–43
- 25 Peason K. On lines and planes of closest fit to systems of point in space. *Philos Mag*, 1901, 2: 559–572
- 26 He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition. ArXiv: 1512.03385
- 27 Boser B E, Guyon I M, Vapnik V N. A training algorithm for optimal margin classifiers. In: *Proceedings of the 5th Annual Workshop on Computational Learning Theory*, Pittsburgh, 1992. 144–152
- 28 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 2015. 3431–3440
- 29 Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 2014. 580–587
- 30 Larochelle H, Erhan D, Courville A, et al. An empirical evaluation of deep architectures on problems with many factors of variation. In: *Proceedings of the 24th International Conference on Machine Learning*, Corvallis, 2007. 473–480
- 31 Rifai S, Vincent P, Muller X, et al. Contractive auto-encoders: explicit invariance during feature extraction. In:

- Proceedings of the 28th International Conference on Machine Learning, Bellevue, 2011. 833–840
- 32 Sohn K, Lee H. Learning invariant representations with local transformations. In: Proceedings of the 29th International Conference on Machine Learning, Edinburgh, 2012
- 33 Sohn K, Zhou G, Lee C, et al. Learning and selecting features jointly with point-wise gated Boltzmann machines. In: Proceedings of the 30th International Conference on Machine Learning, Atlanta, 2013. 217–225

## Conceptor-based deep neural networks

Guangwu QIAN, Lei ZHANG\* & Yan WANG

*Machine Intelligence Lab, College of Computer Science, Sichuan University, Chengdu 610065, China*

\* Corresponding author. E-mail: leizhang@scu.edu.cn

**Abstract** In recent years, deep neural networks, also known as deep learning, have achieved several breakthroughs in different fields that were previously dominated by machine learning. Even when using high-performance computing devices, it takes days or weeks to train a deep neural network. Conceptor, as an extension of echo state networks, can be understood as certain neural filters that characterize dynamical neural activation patterns. In this study, based on some improvements to the original conceptor model, we have conducted several studies from the perspectives of non-iterative methods and transfer learning to address the issues mentioned above, which can be summarized as follows: (1) A conceptor-based classifier for non-temporal data and a non-iterative approach feedforward convolutional conceptor neural network are proposed. This classifier achieves classifying accuracy comparable to that of the state-of-the-art methods while requiring significantly less training time. Through experiments on MNIST variation datasets, we evaluate the classifying quality of the feedforward convolutional conceptor neural network. (2) A classifier called fast conceptor classifier is proposed based on conceptors and it achieves state-of-the-art results with the training time reduced by a factor of 60 on average. Its evaluations with pre-trained rather than fine-tuned neural networks have been investigated on Caltech-101 and Caltech-256 datasets.

**Keywords** conceptor, image classification, deep neural networks, transfer learning, non-iterative methods



**Guangwu QIAN** was born in 1985. He is currently pursuing his Ph.D. degree in College of Computer Science, Sichuan University, Chengdu, China. His main research interests include conceptors and deep learning.



**Lei ZHANG** was born in 1980. She is a professor and Ph.D. supervisor. She is a member of China Computer Federation. Her research interests include big data, recurrent neural networks, deep learning, and machine intelligence.



**Yan WANG** was born in 1992. He is currently pursuing his Ph.D. degree from College of Computer Science, Sichuan University, Chengdu, China. His main research interests include reinforcement learning and deep learning.