

基于模糊不变卷积神经网络的遥感飞机识别

刘坤, 苏彤*, 王典

上海海事大学信息工程学院, 上海 200135

摘要 提出了一种基于模糊不变卷积神经网络(BICNN)模型的目标识别方法。与传统卷积神经网络(CNN)模型不同,BICNN引入了一种新的模糊不变层。BICNN通过增加模糊不变约束项及正则化来优化模糊不变目标函数并进行训练;通过减小模糊不变目标函数值,使得训练样本在模糊前后的特征映射一致,最终实现模糊不变性。测试结果表明,BICNN解决了模糊造成的识别率低的问题,增大了运动模糊图像的识别率。

关键词 光计算;深度学习;卷积神经网络;目标识别

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP55.082001

Remote Sensing Aircraft Recognition Based on Blur-Invariant Convolutional Neural Network

Liu Kun, Su Tong*, Wang Dian

College of Information Engineering, Shanghai Maritime University, Shanghai 200135, China

Abstract A method of target recognition based on the blur-invariant convolutional neural network (BICNN) model is proposed. The BICNN model introduces a new blur-invariant layer, which is different from the traditional convolutional neural network (CNN) models. BICNN is trained by the adding of the blur-invariant constraint term and the regularization to optimize a blur-invariant objective function. The value of the fuzzy invariant objective function is reduced to make the training samples consistent with the feature maps before and after the blurring, and thus the blur invariance is achieved finally. The test results show that BICNN can solve the problem of a low recognition rate caused by blur and improve the recognition rate of the motion blurred images.

Key words optics in computing; deep learning; convolutional neural network; target recognition

OCIS codes 200.4260; 100.3008; 150.1135; 100.4996

1 引言

飞机目标是空中监管的重要目标之一,对飞机目标的自动识别是空中监控的关键技术之一。但是当目标存在运动模糊时,飞机目标特征的提取就存在不确定性及不精确性。

早期的遥感飞机图像识别方式大部分依据飞机图像的轮廓边缘信息,人工手动提取特征,并由此进行模版匹配^[1]。考虑图像模糊的影响,Nishiyama等^[2]通过已知点扩散函数得出模糊核后,反滤波去模糊,即由非盲去模糊算法获取清晰的图像,再使用识别算法对其分类,然而单一的去模糊参数不能较好复原不同类型的模糊图像。Yang等^[3]基于通过

频域估计抵消模糊效应的图像恢复方案,使用识别算法进行分类,但频域估计易引入人为干扰因素,故无法保证较高的识别率。Zhang等^[4]通过分析模糊图像的 Legendre 矩与原始图像的点扩散函数之间的关系,导出了模糊不变矩正交 Legendre 矩,对图像直接提取该特征并进行识别。Dai等^[5]提出了以模糊不变量正交伪 Zernike 矩作为描述算子,然而这些子特征对轮廓信息不具备较高的表达性,因此在应用方面具有一定的局限性。综上所述,传统的模糊目标识别算法存在识别流程复杂且应用环境较为单一的问题。近年来,卷积神经网络(CNN)作为一种多层前馈深度学习模型,可以直接以图像为输入,通过多隐层结构自动学习特征,避免了手工提取

收稿日期: 2018-02-07; 修回日期: 2018-02-16; 录用日期: 2018-03-01

基金项目: 国家自然科学基金(61271446)、航空科学基金(2013ZC15005)

* E-mail: sutong@stu.shmtu.edu.cn

特征的识别流程复杂及应用场景单一的问题,在目标识别上有着广泛的应用,并取得了良好的效果^[6-8]。尽管基于 CNN 的方法在飞机识别上的应用和研究逐渐增多,但是针对目标模糊不变性的稳健识别研究仍处于初步阶段。

Simonyan 等^[8]提出了一种视觉几何组(VGG)模型。VGG 模型具有比传统 CNN 更深层的结构,因此处理的训练数据更多,能提取到更全局更抽象的特征,网络拟合能力更强。本文提出的模糊不变算法以 VGG 网络结构为基础,在目标函数中增加了模糊不变约束项;通过最小化模糊不变目标函数,使图像模糊前后共享相似的特征,完成模糊不变层的学习。为了增强网络的泛化能力,首先训练新增的模糊不变层,固定其余的模型参数,然后以较小的学习速率对整个模糊不变卷积神经网络(BICNN)进行专门微调以进一步提高其性能。使用测试集对 BICNN 进行全面评估,并与包括传统 CNN 在内的先进方法^[9]以及 Wang 等^[10]提出的以局部相模式(LPP)模糊描述不变子识别算法进行了比较。结果表明,BICNN 模型提高了模糊图像飞机机型识别的准确率,从而证明了所提方法的可行性、有效性以及稳健性。

2 基于 BICNN 的方法

2.1 VGG 网络的基本理论

CNN 是近年来图像处理方向应用最为广泛的深度学习模型之一,而 VGG 是一种典型的具有较高识别率的 CNN。图像通过多隐层结构自动学习

图像丰富有意义的特征^[11]。VGG16 具有深层次的结构及尺寸较小的卷积核(3×3),具备传统 CNN 所不具备的优点。VGG16 模型由 16 个加权层组成,包括 13 层卷积层、5 层池化层和 3 层全连接层。VGG 网络的卷积层使用的卷积核尺寸均为 3×3 。含多个小尺寸卷积核的卷积层比大尺寸卷积层的参数少,且能在不影响视野域的情况下增大映射函数的非线性,同时较小的卷积核能使 VGG 网络更具有判别性^[12]。

2.2 BICNN 结构

VGG 具有良好的判别性,因此基于 VGG 网络结构提出了一种 BICNN 飞机识别方法。该方法基于十类飞机机型目标的识别,识别方法系统的结构图如图 1 所示,其中 e 为自然常数。该系统以 BICNN 为核心,BICNN 的前 13 层将加载已由 VGG16 网络在 CIFAR10 数据集上训练好的参数,通过对剩余网络层设定一个较大的学习率来进行微调,再以小的学习率训练整个 BICNN。使用 CIFAR10 大量图像训练的参数,其模型预测类型较多,可以更好地提取图像中的特征,由此将其作为 BICNN 模型的起点,更便于训练网络。与 VGG16 不同的是,增加的模糊不变层通过模糊不变约束项,使模糊图像与清晰图像的特征差距趋于无穷小,实现了网络的模糊不变性。当运动模糊的飞机图像通过该识别系统时,得到的 BICNN 特征表达与清晰图像的相似,再通过 softmax 判别^[13],模糊图像与清晰图像均得到较高的识别率,进而可得出飞机识别的结果。

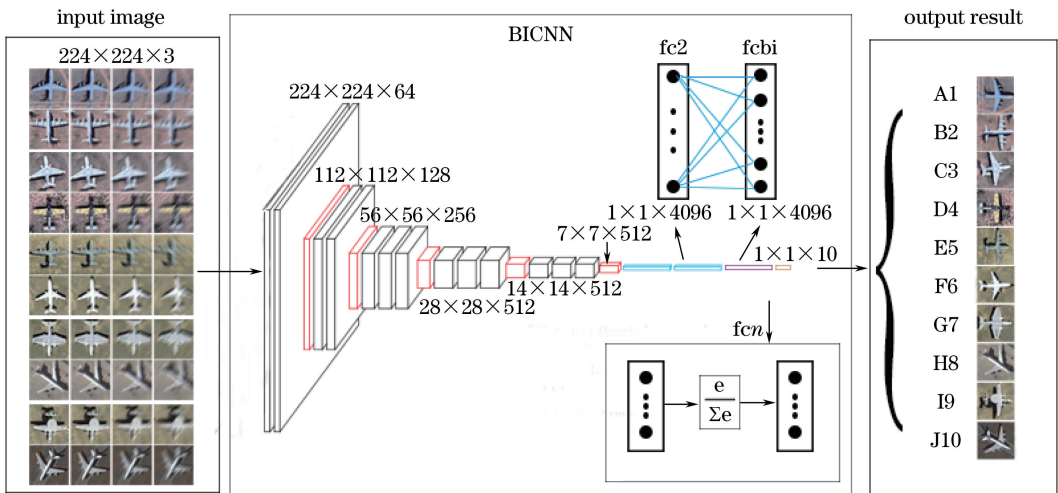


图 1 BICNN 飞机识别系统的结构图

Fig. 1 Structural diagram of aircraft recognition system based on BICNN

BICNN 由 13 个卷积层、3 个全连接层、1 个模糊不变层串联组成。BICNN 前 15 层的网络结构和 VGG 网络的前 15 层结构一致,由 13 个卷积池化层和 2 个全连接层组成,其次为全连接模糊不变 (fcbi) 层,最后使用 softmax 判别的 n 全连接 (fc n , n 为层数) 层作为 BICNN 的输出。网络最终输出的类别为 10 类目标,分别简称为 A1、B2、C3、D4、E5、F6、G7、H8、I9、J10, fc n 由 10 个神经元组成。fcbi 层的输出作为 fc n 层的输入,通过自学习模糊不变层,完成 BICNN 网络模型的设计。

为了便于训练并避免过度拟合,将由 CIFAR10 训练的 VGG 前 13 层的网络权重加载到 BICNN 的前 13 层,然后以较小的学习速率对模糊不变层及全连接层进行微调。BICNN 不仅网络结构和 VGG 的不同,而且使用的目标函数也有所区别,优化神经网络参数的目标是最小化数据真实分布与网络模拟的近似分布的交叉熵,即确定参数的最大似然估计过程中的形成目标函数,目标函数是网络拟合数据真实分布的一个重要指标^[14]。因此,BICNN 优化了目标函数的多项式回归方程,使其更符合模糊图像的数据分布。以清晰图像为学习样本,通过增加模糊约束项,使最小经验误差加上先验知识模糊不变的约束,从而使清晰图像和运动模糊图像共享相似的特征。优化目标函数时,偏向实现模糊不变约束项的梯度降低方向,使最终解与数据真实分布相近,同时避免过拟合^[15]。

给定训练样本 $X = \{x_i \in X \cup Y_i\}$ 及其对应的标签 $Y = \{y_{xi} | x_i \in X\}$,其中 x_i 为训练样本 X 的元素, Y_i 为模糊处理后的样本, y_{xi} 为样本的真实值标签向量,仅一个元素为 1,其余均为 0。实验目的是通过输入目标对 (X, Y) 训练一个 BICNN 模型。对于训练样本 X , fc2 层的输出特征为 $O_2(x_i)$ 。同时, $O_2(x_i)$ 作为 fcbi 层的输入,经过 fcbi 层输出的结果为 $O_b(x_i)$ 。最后, $O_b(x_i)$ 经过 fc n 层的输出得到 $O_n(x_i)$ 。如图 2 所示,softmax 函数将输入值转换为概率向量,表示为输入样本属于每个输出类的概率。 (W_{bi}, B_{bi}) 和 (W_n, B_n) 为 fcbi 层和 fc n 层的网络参数,其中 W 为该层网络的权重, B 为网络偏置。可以得到

$$O_b(x_i) = \text{relu}[W_{bi}O_2(x_i) + B_{bi}], \quad (1)$$

$$O_n(x_i) = \text{softmax}[W_nO_b(x_i) + B_n], \quad (2)$$

式中激活函数 $\text{relu}(x) = \max(0, x)$, $\text{softmax}(x) = \exp(x) / \|\exp(x)\|_1$ 。

给定训练样本及其对应标签,除了要求 BICNN 网络应尽量减小训练数据集上的错误识别误差外,

尤其要求 BICNN 模型具有任意一组训练样本的模糊不变能力。将模糊图像 X_i 与清晰图像样本 Y_i 的特征距离作为模糊不变约束,则 BICNN 算法的实现可以表示为目标函数最优化的求解问题。为此,提出了一个新的学习目标函数:

$$J(W, B) =$$

$$\min \left[C(X, Y) + \lambda_1 B(X_i, Y_i) + \frac{\lambda_2}{2} \|W\|_2^2 \right], \quad (3)$$

式中 λ_1 和 λ_2 分别为控制(3)式中三项相对关系重要性的两个权衡参数; $C(X, Y)$ 为 softmax 损失函数,用于使损失最小化,其表达式为

$$C(X, Y) = -\frac{1}{N} \sum \langle y_{xi}, \lg O_n(x_i) \rangle, \quad (4)$$

式中 N 代表输入样本的个数。

(3)式第二项 $\lambda_1 B(X_i, Y_i)$ 为模糊不变约束项。为使原数据样本 X_i 和模糊处理后的样本 Y_i 共享相似特征,通过定义的模糊不变约束项公式,使图像模糊前后的 BICNN 特征的距离最小化,而模糊图像与清晰样本的特征误差尽量小。定义的模糊不变约束项公式展开为

$$B(X_i, Y_i) = \sqrt[2]{\sum \|O_b(X_i) - O_b(Y_i)\|^2}. \quad (5)$$

如果(5)式的结果是一个很小的值,则特征表示为模糊近似不变。

(3)式第三项为权值衰减项的 L2 正则化,它倾向于减小网络权重,防止网络产生过拟合,提高网络的泛化能力。可以看到,由(3)式定义的目标函数不仅能使识别损失最小化,还定义了一个模糊不变约束项来实现模糊不变性。在实验中,采用随机梯度下降 (SGD) 方法解决优化问题,该方法已被广泛应用于复杂随机优化问题中。定义网络权重的修正值为

$$W'_k = W_k - \eta \frac{\partial J_{x_i}}{\partial W_k}, \quad (6)$$

式中 J_{x_i} 为单个样本 x_i 的目标函数值, W_k 为网络权重, η 为学习率。

交叉熵能够判定网络的输出与期望输出的距离^[16],通过最小化交叉熵损失,调整权重训练 BICNN 网络,最终得出的深度特征用于识别飞机机型。fc n 层最终的识别结果由输出向量元素中概率值最大的确定,具体数值由 softmax 函数计算,样本类型对应概率最大的类别。

2.3 基于 BICNN 的识别过程

基于 BICNN 的识别过程,主要有以下三个主要步骤。

1) 实验数据建立

主要针对运动模糊的图像识别,需要大量具有运动模糊的样本图像。在已有清晰图像集 X_i 的基础上,需要作一些运动模糊处理。运动模糊具有两个重要的参数:模糊角度与模糊尺度,以此建立运动模糊核。

定义 8 个不同的模糊尺度,分别为 10, 20, 30, 40, 50, 60, 70, 80, 对原始数据集进行模糊变换,完成运动模糊数据集 ($Y_1 \sim Y_8$) 的制作。使用模糊处理之前的 X_i 和模糊处理之后的 Y_i ($Y_1 \sim Y_8$) 作为总

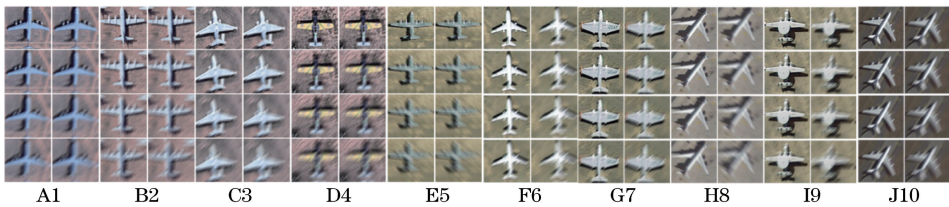


图 2 遥感飞机的部分图集

Fig. 2 Partial image atlas of remote sensing aircraft

2) BICNN 网络训练

BICNN 网络训练的流程图如图 3 所示,首先将 VGG16 上针对 CIFAR10 的预训练参数迁移至 BICNN 对应网络,即全连接层之前的网络。对于 BICNN 的模型训练,最后全连接层的学习率为 0.01,其余层的学习率为 0.0001,训练的批量大小为 32, λ_1 为 0.001, λ_2 为 0.0005,迭代次数为 16000。

训练样本 X 输入参数迁移后的 BICNN 网络,输入图像的尺寸为 $224 \text{ pixel} \times 224 \text{ pixel}$,每个像素乘以滤波器的值再进行求和运算,经过第一层卷积层的 64 个滤波器(即卷积层输出的神经元个数为 64),卷积的滑动步长为 1,得到卷积层的输出特征图为一个四维的张量 $[32, 64, 224, 224]$ 。此张量通过激活函数 relu ,增大映射函数的非线性^[17]。

再经过两次卷积之后,下一个过程为池化,池化能减少参数,降低计算量。采用最大池化,选用的是 2×2 的滤波器,其移动步长为 2。经过池化层后,数据特征减少一半,同时保留了有用信息^[18]。

最终,经过剩余的 13 个卷积池化层,网络得到

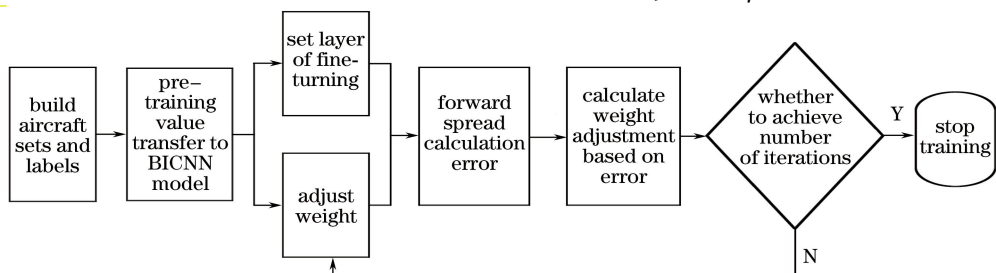


图 3 BICNN 网络训练的流程图

Fig. 3 Flow chart of BICNN network training

训练样本,用于训练 BICNN。

使用的数据集包含十类飞机机型,这十类飞机图是谷歌地图美国图森的飞机墓场中常见的十类目标。训练集图片共有 25000 张,其中交叉验证集 5000 张,测试集 5000 张。为了保障网络输入的大小相同,将所有图像的大小设置为 $224 \text{ pixel} \times 224 \text{ pixel}$ 。同时对图像进行 0 均值归一化处理。图 2 所示为数据集中十类机型飞机的部分图集,每一类八张样本图像分别表示模糊尺度从 10 依次增大到 80。

了图像的分布式特征,全连接层将该特征线性变换到样本的标签标记空间。通过增加模糊不变约束项,并最小化模糊图像与清晰图形的特征距离,使模糊图像的特征向量接近清晰样本的,从而实现模糊不变层。网络输出值 O_n 与样本标签 d 的误差 E 为

$$E = \frac{1}{2} \sum |d - O_n|^2. \quad (7)$$

根据误差计算权值和偏置调整量:

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}}, \quad (8)$$

$$\Delta V_{jk} = -\eta \frac{\partial E}{\partial V_{jk}}, \quad (9)$$

式中 W_{ij} 为输入层 i 到隐藏层 j 的权值, V_{jk} 为隐藏层 j 到输出层 k 的权值, ΔW_{ij} 为输入层到隐含层的权重调整量, ΔV_{jk} 为隐含层到输出层的权值调整量。

根据权值调整量,调整网络权值:

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij} = W_{ij}(t) + \eta(d - O_n)O_n(1 - O_n)Y, \quad (10)$$

$$\mathbf{V}_{jk}(t+1) = \mathbf{V}_{jk}(t) + \Delta \mathbf{V}_{jk} = \mathbf{V}_{jk}(t) + \eta [(d - \mathbf{O}_n) \mathbf{O}_n (1 - \mathbf{O}_n)] \mathbf{W} \mathbf{Y} (1 - \mathbf{Y}) \mathbf{X}, \quad (11)$$

式中 \mathbf{Y} 为隐含层的输出向量, \mathbf{X} 为输入向量。

调整量通过反向传播方法对网络权值进行更新,更新完成即为一次网络训练,达到迭代次数后结束训练^[19]。

网络的参数基于误差能量最小化的反向传播来更新,本文采用 SGD 方法。SGD 要求谨慎调整模型的超参数,特别是学习率和初始参数值^[20]。因为每一层的输入都受到前面所有层的参数的影响,训练过程复杂,随着网络变得更深,网络参数的微小变动将使输出的数据发生较大的变化。因此,通过批量归一化对每一个神经元作归一化处理,使网络整体参数的更新步调一致,允许更大的学习率,同时提高 BICNN 的收敛速度,提升该模

型的泛化能力^[21]。

3) BICNN 网络测试

为了测试该系统的有效性和可行性,对 BICNN 进行测试。测试集图片为 5000 张,测试图像的尺寸为 224 pixel×224 pixel,测试集与训练集没有重复样本。将测试样本作为输入,得到网络输出后,与样本标签进行对比,得出网络识别率。

3 实验结果与验证

为了验证模糊不变层的作用,分别提取 VGG16 模型与 BICNN 的特征向量。从 4096 维度的特征向量中随机抽取 64 维,以 A1 机型为例,选取清晰样本和模糊尺度为 20、50、80 的样本图集,提取 VGG16 网络和 BICNN 的特征向量,并将随机抽取的 64 维特征绘制为直方图,如图 4 所示。

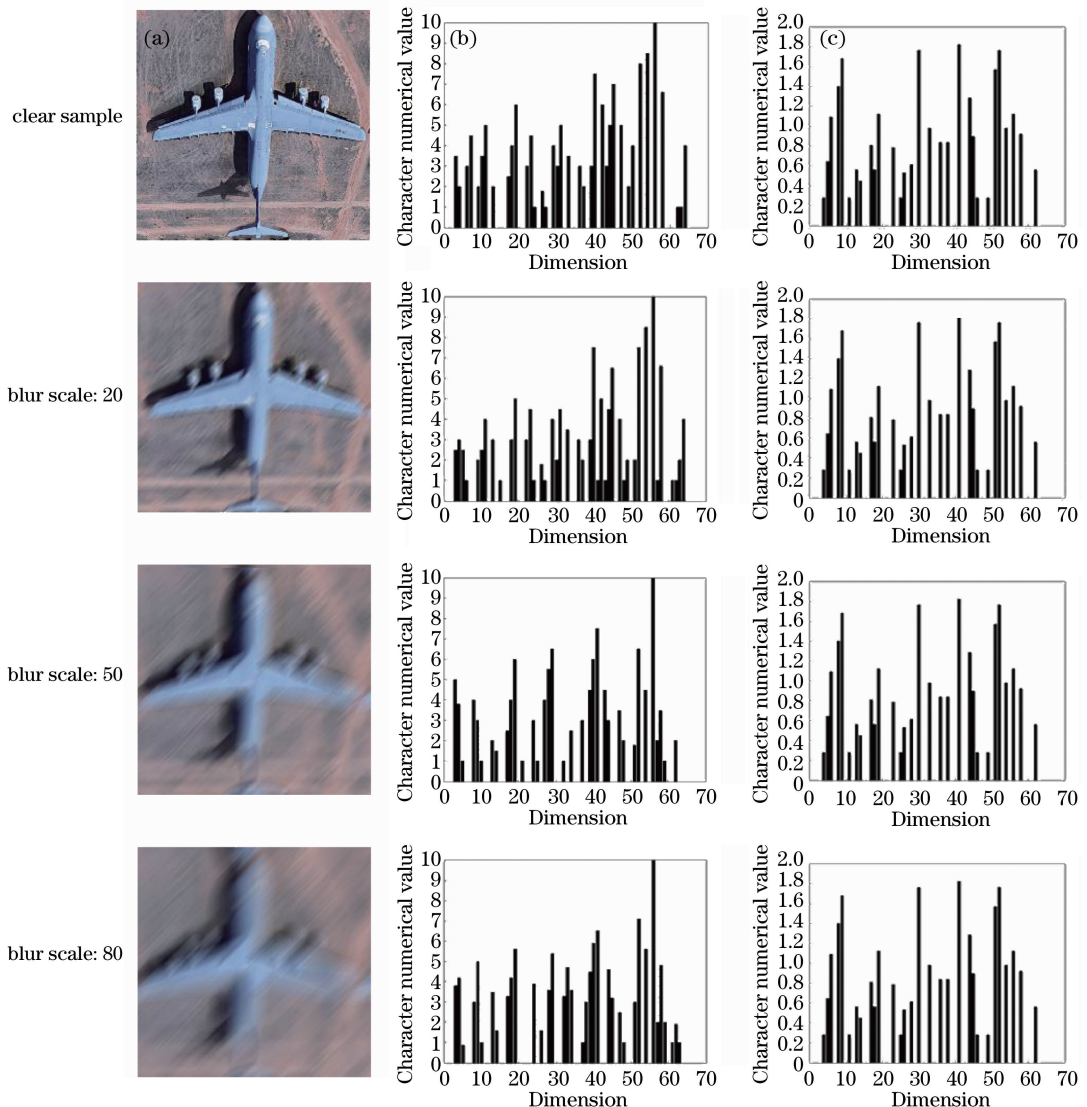


图 4 特征直方图。(a)飞机样本;(b) VGG16;(c) BICNN

Fig. 4 Feature histograms. (a) Aircraft samples; (b) VGG16; (c) BICNN

图 4(a)所示为飞机样本图,包括清晰的飞机图以及模糊尺度为 20、50、80 的运动模糊飞机样本。图 4(b)所示为由 VGG16 网络提取的对应样本的特征向量直方图;图 4(c)所示为经过模糊不变层 fcbi 的 BICNN 的特征向量直方图。由图 4(b)可知, VGG16 对不同模糊尺度的样本提取出来的直方图特征各不相同,易降低网络识别的精确度,这也是传统 CNN 难以有效处理运动模糊现象的原因。而从图 4(c)可以看到,不同模糊尺度的图像经过 BICNN 提取后,其特征直方图中得到了与清晰图像相似的特征向量,说明模糊图像具有更高的判别性。模糊图像经过 BICNN 网络后,具备与清晰

图像一样高的识别准确率,因此 BICNN 网络具备模糊不变性,可改善图像运动模糊情况下识别率较低的问题。

在网络训练过程中,微调将减小参数量大的 CNN 网络过拟合的可能性,提高泛化能力,同时有助于提升系统最终的识别性能,可通过训练过程中的准确率和损耗曲线验证网络的准确率和稳定性。经过微调后, BICNN 训练随着迭代次数的增加,收敛速度加快,识别率提高,损耗也随之减小。微调后的 BICNN 和单独训练的 BICNN 的准确率与损耗曲线如图 5 所示,可以看出,微调 BICNN 模型比单独训练网络的准确率高,网络更为稳定。

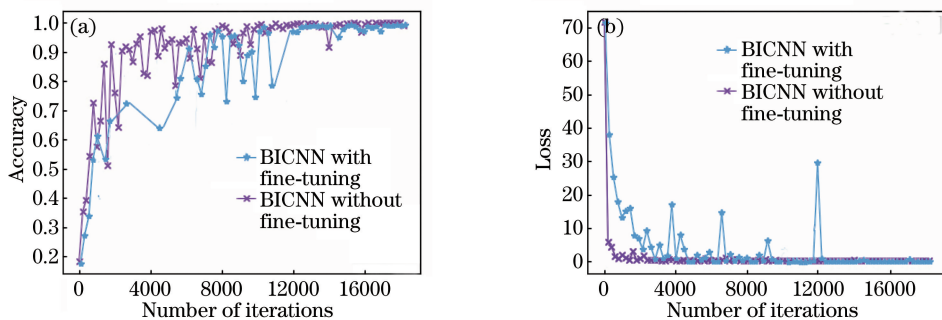


图 5 微调和无微调 BICNN 训练的性能曲线。(a)准确率;(b)损耗

Fig. 5 Performance curves of BICNN trainings with and without fine-tuning. (a) Accuracy; (b) loss

为了进一步验证微调 BICNN 模型的性能,将文献[9]中的 CNN + Fisher 向量(FV)算法、文献[10]中的 LPP 模糊不变描述子识别算法及 VGG16 模型与经过微调后的 BICNN 模型进行对比。文献[9]利用 CNN 提取图像特征,并使用 FV 算法进行编码。FV 算法是一类常用的特征编码方法,它假设一幅图像中的所有描述子都是独立同分布的,并且满足高斯混合模型。通过对每一个高斯混合模型的似然函数的梯度向量进行求解,可以得到拟合数据参数的更新方向^[9]。FV 编码增大了图像特征的维度,能更好地描述图像。文献[10]提出通过 LPP 描述符提取模糊不变特征,对图像的每个窗口位置计算傅里叶变换的局部相位模式信息^[10],以用于图像分析。实验证明,该描述子具有较高的表达能力,在模糊尺度较大的情况下也具有较好的识别效果。使用模糊尺度为 50 的测试集图片 625 张,对 BICNN 模型、文献[9]算法、文献[10]算法、VGG 模型进行测试,计算出平均识别率,结果如图 6 所示。针对飞机机型的识别率为

式中 t_i 为每一类别里识别正确的样本数量, c 为样本类别。

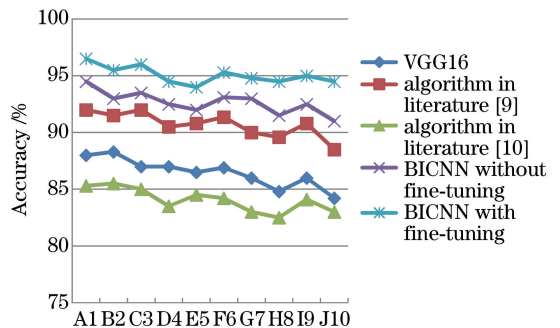


图 6 平均识别率

Fig. 6 Average recognition rate

进一步对比测试各个模型在不同模糊尺度的飞机图像上的平均识别率,所选数据包括模糊尺度为 10、20、30、40、50、60、70、80 的图像各 625 张,结果见表 1,可见 BICNN 模型显著提高了识别率。同时,为了验证 BICNN 模型的实时性和可行性,计算了十张图像测试算法的识别时间,与文献[9]算法以及 VGG16 网络框架的识别时间进行了对比,结果见表 2。实验配置如下:中央处理器为 Intel i5-7300HQ (2.50 GHz),图形处理器为英伟达显卡 GTX 960m。

$$A = \frac{\sum_{i=1}^c t_i}{N}, \quad (12)$$

表 1 在不同模糊尺度飞机图像上的平均识别率

Table 1 Average recognition rate of aircraft images under different blur scales

Model	Blur scale							
	10	20	30	40	50	60	70	80
VGG16	97.26%	94.22%	92.37%	89.16%	86.47%	81.71%	77.83%	74.32%
Algorithm in literature [9]	98.00%	96.65%	95.40%	93.71%	91.50%	89.75%	84.43%	81.75%
Algorithm in literature [10]	96.05%	93.83%	90.00%	87.50%	84.35%	80.13%	76.10%	72.25%
BICNN without fine-tuning	98.60%	97.95%	96.83%	95.50%	92.25%	89.56%	87.32%	84.05%
BICNN with fine-tuning	98.91%	98.53%	97.65%	96.75%	95.12%	91.47%	89.57%	86.94%

表 2 实时计算时间对比

Table 2 Real-time computing time comparison

Model	Algorithm in literature [9]	VGG16	BICNN with fine-tuning
Recognition time /s	11.049	9.635	10.217

4 结 论

为了解决具有运动模糊现象的目标识别率偏低的问题,提出了一种新的 BICNN 模型,该模型通过优化目标函数,使模糊前后训练样本的特征向量一致。BICNN 通过加载 VGG16 的前 13 层在 CIFAR10 训练的权重,在清晰遥感飞机图库及具有不同模糊尺度的运动模糊图库的基础上对 BICNN 进行微调。分别基于 VGG16 模型、文献[9]算法、文献[10]算法和 BICNN 微调网络及 BICNN 无微调网络进行了具有运动模糊的飞机机型的识别测试,发现 BICNN 模型显著提高了识别率,解决了因运动模糊造成飞机机型难以识别的问题。该方法能准确识别飞机机型,可广泛应用在机场军事管理、国防安全等领域。

参 考 文 献

- [1] Simonyan K, Parkhi O M, Vedaldi A, *et al.* Fisher vector faces in the wild[C]. Proceedings of the British Machine Vision Conference, 2013: 8.1-8.11.
- [2] Nishiyama M, Hadid A, Takeshima H, *et al.* Facial deblur inference using subspace analysis for recognition of blurred faces[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(4): 838-845.
- [3] Yang S J, Ye X, Zhang S J. A new infrared turbulent fuzzy image restoration algorithm based on Gaussian function parameter identification[C]. International Conference on Image, Vision and Computing, 2017: 423-427.
- [4] Zhang H, Shu H, Han G N, *et al.* Blurred image

recognition by Legendre moment invariants[J]. IEEE Transactions on Image Processing, 2010, 19(3): 596-611.

- [5] Dai X B, Liu T L. Image recognition algorithm by blur invariants of pseudo-Zernike moment[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science), 2013, 33(3): 50-55.
戴修斌, 刘天亮. 基于伪 Zernike 矩模糊不变量的图像识别算法[J]. 南京邮电大学学报(自然科学版), 2013, 33(3): 50-55.
- [6] He K, Zhang X, Ren S, *et al.* Image net classification with deep convolutional neural networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [7] Liu D W, Han L, Han X Y. High spatial resolution remote sensing image classification based on deep learning[J]. Acta Optica Sinica, 2016, 36(4): 0428001.
刘大伟, 韩玲, 韩晓勇. 基于深度学习的高分辨率遥感影像分类研究[J]. 光学学报, 2016, 36(4): 0428001.
- [8] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2017-12-10]. <https://arxiv.org/abs/1409.1556>.
- [9] Xie X K. Research on blurred image recognition based on transfer learning[D]. Wuhan: Huazhong University of Science and Technology, 2016.
解晓康. 基于迁移学习的模糊图像识别技术研究[D]. 武汉: 华中科技大学, 2016.
- [10] Wang W, Cao Z. Recognition of blurred faces using local phase pattern[J]. Electronics Letters, 2012, 48(20): 1269-1271.
- [11] Liu F, Shen T S, Ma X X. Convolutional neural network based multi-band ship target recognition with feature fusion[J]. Acta Optica Sinica, 2017, 37(10): 1015002.
刘峰, 沈同圣, 马新星. 特征融合的卷积神经网络多

- 波段舰船目标识别[J]. 光学学报, 2017, 37(10): 1015002.
- [12] Schroff F, Kalenichenko D, Philbin J. Facenet: A unified embedding for face recognition and clustering[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 815-823.
- [13] Ye G L, Sun S Y, Gao K J, *et al.* Nighttime pedestrian detection based on faster region convolution neural network[J]. Laser & Optoelectronics Progress, 2017, 54(8): 081003.
叶国林, 孙韶媛, 高凯珺, 等. 基于加速区域卷积神经网络的夜间行人检测研究[J]. 激光与光电子学进展, 2017, 54(8): 081003.
- [14] Taigman Y, Yang M, Ranzato M A, *et al.* DeepFace: Closing the gap to human-level performance in face verification[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1701-1708.
- [15] Wang J, Yuan C. Facial expression recognition with multi-scale convolution neural network [C]. Pacific Rim Conference on Multimedia, 2016: 376-385.
- [16] Richardson W H. Bayesian-based iterative method of image restoration[J]. Journal of the Optical Society of America, 1972, 62(1): 55-59.
- [17] Ojansivu V, Heikkilä J. Blur insensitive texture classification using local phase quantization[C]. International Conference on Image and Signal Processing, 2008, 5099: 236-243.
- [18] Chen Y, Fan R S, Wang J X, *et al.* High resolution image classification method combining with minimum noise fraction rotation and convolution neural network[J]. Laser & Optoelectronics Progress, 2017, 54(10): 102801.
陈洋, 范荣双, 王竞雪, 等. 结合最小噪声分离变换和卷积神经网络的高分辨影像分类方法[J]. 激光与光电子学进展, 2017, 54(10): 102801.
- [19] Yu D, Deng L. Deep learning and its applications to signal and information processing [J]. IEEE Signal Processing Magazine, 2011, 28(1): 145-154.
- [20] Ji S, Xu W, Yang M, *et al.* 3D convolutional neural networks for human action recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 221-231.
- [21] Hu F, Xia G S, Hu J, *et al.* Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery [J]. Remote Sensing, 2015, 7(11): 14680-14707.