

融合小波变换与胶囊网络的纹理图像分类算法

陶志勇¹, 李杰^{1,2*}, 唐晓亮²

¹辽宁工程技术大学电子与信息工程学院, 辽宁 葫芦岛 125105;

²中国科学院海西研究院泉州装备制造研究所, 福建 泉州 362000

摘要 胶囊网络作为一种新型深度学习网络, 胶囊结构可以编码特征的姿态、纹理、色调等信息, 对图像具有良好的纹理特征编码能力。针对胶囊网络的初级特征提取网络过于简单、空间特征表达能力不足的问题, 提出了一种结合深度卷积神经网络特征表达能力与小波变换多分辨率分析能力的离散小波胶囊网络(DWTCapsNet)。首先, 研究了胶囊网络在纹理图像分类应用中的可行性; 其次, 研究了 DWTCapsNet 各部分对胶囊网络分类性能提升的能力; 最后, 通过抗旋转和抗噪声实验分析了 DWTCapsNet 的鲁棒性。以分类准确率为模型评价标准, 在常用纹理图像数据集上的实验结果表明, DWTCapsNet 的分类准确率较高。

关键词 图像处理; 离散小波变换; 胶囊网络; 图像分类; 深度学习

中图分类号 TP391.4 文献标志码 A

doi: 10.3788/LOP57.241002

Texture Images Classification Algorithm Combining Wavelet Transform and Capsule Network

Tao Zhiyong¹, Li jie^{1,2*}, Tang Xiaoliang²

¹School of Electronic Information Engineering, Liaoning Technical University, Huludao, Liaoning 125105, China;

²Quanzhou Institute of Equipment Manufacturing Haixi Institutes, Chinese Academy of Sciences, Quanzhou, Fujian 362000, China

Abstract Capsule network is a new type of deep learning network, capsule structure can encode information such as posture, texture, hue, etc. of the feature, and has a good ability to express the texture feature of the image. Aiming at the problem that the primary feature extraction network of the capsule network is too simple and the feature expression ability is insufficient, a discrete wavelet capsule network (DWTCapsNet) that combines the feature expression capabilities of deep convolutional neural networks with wavelet transform multi-resolution analysis capabilities is propose in this work. First, the feasibility of the capsule network in the application of texture image classification is studied. Second, the ability of each part of DWTCapsNet on the improvement of capsule network classification performance is studied. Finally, the robustness of DWTCapsNet is analyzed through anti-rotation and anti-noise experiments. The classification accuracy is used as the standard model evaluation criteria, and the experimental results on the commonly used texture image data sets show that the classification accuracy of DWTCapsNet is higher.

Key words image processing; discrete wavelet transform; capsule network; image classification; deep learning

OCIS codes 100.7410; 100.5010; 100.3008; 100.4996

收稿日期: 2020-04-21; 修回日期: 2020-05-16; 录用日期: 2020-05-22

基金项目: 国家重点研发计划(2018YFB1403303)、辽宁省自然科学基金(2015020100)、辽宁省教育厅一般项目(LJ2017QL013)

* E-mail: leej95@163.com

1 引言

图像分类是计算机视觉中一个重要的研究课题,纹理图像分类属于图像分类任务中一种更细粒度的图像分类任务。纹理与色彩、形状、光泽一样,是物体、材料和场景的基本视觉特征,因此,纹理分析在图像分类、基于内容的图像检索、医学图像分析、表面检测和生物特征识别^[1-3]等应用中发挥着重要作用。传统的纹理图像分类算法主要依赖人工设计的图像特征,包括局部二值模式(LBP)特征^[4]、滤波器组特征^[5-6]、基于共生矩阵的特征^[7]等。虽然这些特征在纹理图像分类任务中取得的效果较好,但特定特征的设计需要特定领域的知识支撑。

卷积神经网络(CNN)可自动进行特征提取,随着深度学习技术的发展,CNN已经成为一种主流的模式识别方式。在纹理图像分类领域。Tivive等^[8-9]通过训练CNN对纹理图像进行分类,Cimpoi等^[10]将CNN迁移到纹理图像分类中,提出了费雪向量池化的CNN(FV-CNN)滤波器组,但该网络需要复杂的数据预处理与费雪向量标记;Bello-Cerezo等^[11]将视觉几何组网络(VGGNet)应用于纹理图像分类,取得了不错的分类效果。

CNN擅长提取图像的空间特征,而基于频谱信息的频谱分析方法善于捕捉尺度不变性特征。小波变换^[12]作为经典的谱分析方法,具有多分辨分析能力,广泛应用于信号处理和图像处理领域。将CNN作为一种受限的多分辨分析方法,然后与小波变换相结合,可实现空间特征和谱特征的融合,弥补CNN多分辨率分析能力的不足。Fujieda等^[13]提出了一种结合CNN和小波变换的小波卷积神经网络(Wavelet-CNN),并将其应用到纹理图像分类中,取得了较好的分类效果。Liu等^[14]将小波变换与CNN相结合,并成功应用于图像修复中。

胶囊网络^[15]作为一种新型深度学习方法,善于

捕捉图像中物体的姿态、纹理、色调等特征。用于纹理图像的分类时,可以减轻光照强度、姿态、色彩等因素的影响。已有研究表明,深度卷积神经网络(DCNN)具有良好的空间特征提取能力,而胶囊网络的初级特征提取网络由单层卷积层构成,相比结构复杂的DCNN,如残差网络(ResNet)^[16]、密集卷积网络(DenseNet)^[17],无法提取到足够的空间特征,从而影响分类性能。因此,本文结合DCNN的特征提取能力、小波变换的多分辨分析能力提出了一种离散小波胶囊网络(DWTCapsNet)。利用胶囊网络善于捕捉图像中姿态、纹理、色调等优势,将其应用到纹理图像分类任务中,以提升纹理图像分类的准确率。在常用纹理图像数据集上的实验结果表明,相比传统纹理图像分类算法和基于CNN的纹理图像分类算法,小波胶囊网络对纹理图像的分类准确率较高,且抗噪声性能强、鲁棒性好。

2 算法描述

2.1 胶囊网络

第一代胶囊网络是一个浅层网络,用向量作为胶囊编码局部特征的姿态信息。初级胶囊层将特征图从标量转换到向量,然后通过动态路由算法进行迭代,学习特征间的相互关系,实现分类。相比姿态向量,具有相同元素数量的姿态矩阵能编码更多的信息,且在进行姿态变换时,矩阵运算可以节省更多的计算开销。因此,Sabour等^[18]进一步改进了胶囊网络,用带有激活值的姿态矩阵作为胶囊,并提出了期望最大化路由算法进行数据迭代,实现了鲁棒性更强、可适应复杂数据集的矩阵胶囊网络(Matrix CapsNet)。矩阵胶囊网络由线性整流单元激活函数卷积层(ReLU Conv)、初级胶囊层(PrimaryCaps)、卷积胶囊层(ConvCaps)、分类胶囊层(class capsule)组成^[19],其网络模型结构如图1所示。其中,capsule为矩阵胶囊,pose为矩阵胶囊

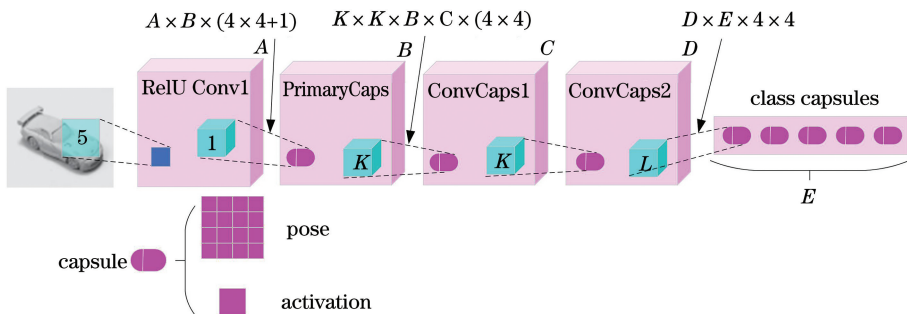


图1 矩阵胶囊网络的结构

Fig. 1 Structure of matrix capsule network

中的姿态矩阵, activation 为矩阵胶囊中的激活值, $K \times K$ 为卷积核的尺寸, A 为卷积特征图的通道数, B 为初级胶囊的数量, C 和 D 分别为对应卷积胶囊层的胶囊数量, E 为待分类胶囊的数量。

假设 L 层和 $L+1$ 层分别有 n 和 k 个胶囊, 用 i 和 j 分别表示第 L 层的第 i 个胶囊和 $L+1$ 层的第 j 个胶囊, 则将胶囊 i 的姿态矩阵与胶囊 i 和胶囊 j 之间的变换矩阵 \mathbf{W}_{ij} 相乘可得到观测矩阵 \mathbf{V}_{ij} 。期望最大(EM)动态路由算法将该层的观测矩阵 \mathbf{V}_{ij} 和激活值 a_i 作为输入, 得到胶囊 j 的姿态矩阵和激活值 a_j 。用 $N(\mathbf{V}_{ij}; \boldsymbol{\mu}_j, \boldsymbol{\sigma}_j^2)$ 表示观测矩阵 \mathbf{V}_{ij} 属于第 j 类的高斯分布, 其中, $\boldsymbol{\mu}_j$ 为胶囊 j 的姿态矩阵, 可作为高斯混合模型的均值, $\boldsymbol{\sigma}_j^2$ 为计算胶囊 j 激活值 a_j 的方差。用高斯分布概率密度函数计算得到 \mathbf{V}_{ij} 连接到胶囊 j 的概率, 可表示为

$$p_j = \frac{1}{\sqrt{2\pi(\boldsymbol{\sigma}_j^2)^2}} \exp\left[-\frac{(\mathbf{V}_{ij} - \boldsymbol{\mu}_j)^2}{2\boldsymbol{\sigma}_j^2}\right]。 \quad (1)$$

根据(1)式得到胶囊 i 属于胶囊 j 的概率 \mathbf{R}_{ij} 为

$$\mathbf{R}_{ij} = \frac{a_j p_j}{\sum_{j=1}^k a_j p_j}。 \quad (2)$$

根据(2)式和第 L 层胶囊的激活值 a_i 计算得到 \mathbf{R}_{ij} 的中间值为

$$r_{ij} = \frac{a_i \mathbf{R}_{ij}}{\sum_{i=1}^n a_i \mathbf{R}_{ij}}。 \quad (3)$$

根据(3)式得到高斯混合模型的均值和方差为

$$\boldsymbol{\mu}_j = \sum_{i=1}^n r_{ij} \mathbf{V}_{ij}, \quad (4)$$

$$\boldsymbol{\sigma}_j^2 = \sum_{i=1}^n r_{ij} (\mathbf{V}_{ij} - \boldsymbol{\mu}_j)^2。 \quad (5)$$

激活值的更新, 可表示为

$$a_j = X_{\text{logistic}}\{\lambda [\beta_a \mathbf{I} - \mathbf{X}_{\text{cost}}(j)]\}, \quad (6)$$

式中, λ 为退火参数, 可根据训练迭代次数进行更新, X_{logistic} 为激活函数, $\mathbf{X}_{\text{cost}}(j)$ 为信息熵, 信息熵越小, 激活值越大, 即胶囊 i 连接到胶囊 j 的概率越大^[20], \mathbf{I} 为单位矩阵。 $\mathbf{X}_{\text{cost}}(j)$ 可表示为

$$\mathbf{X}_{\text{cost}}(j) = [\beta_u \mathbf{I} + \ln(\boldsymbol{\sigma}_j)] \sum_{i=1}^n a_i \mathbf{R}_{ij}, \quad (7)$$

式中, β_a 和 β_u 为两个可训练的超参数, n 为 \mathbf{V}_{ij} 的维度。(1)式、(2)式为 EM 路由算法的期望计算过程(E-step), (3)式~(6)式为最大化计算过程(M-step)。进行 EM 路由迭代时, 先进行 M-step 的计算, 再进行 E-step 的计算, 循环迭代以更新姿态矩阵和激活值。

2.2 损失函数

实验用 Spread 损失函数进行反向传播, 学习

模型参数。该损失函数可降低模型训练过程中对模型初始化和超参数的敏感性, 对于目标类别的激活值 a_i 与其他类别的激活值 a_i , 其损失函数可表示为

$$L_i = \{\max[0, m - (a_i - a_i)]\}^2, \quad L = \sum_{i \neq i} L_i, \quad (8)$$

式中, m 为边界值。当目标类别与其他类别之间的距离比 m 小时, 用 $m - (a_i - a_i)$ 的平方作为惩罚。为了避免模型训练早期出现无效胶囊, m 的初始值为 0.2; 同时在训练过程中以 0.1 为步长进行增长, 直到达到最大边界值 0.9。

2.3 离散小波变换

一维小波变换的实质是将一维信号分别经过低通滤波器和高通滤波器以及二元抽样得到原始信号的低频部分和低频部分。图像作为二维(2D)信号, 可根据 Mallat 算法, 通过一系列一维小波变换得到不同的子带图像。离散小波变换也可以看作是一种卷积变换, 对于一幅图像 \mathbf{X} , 利用 4 个卷积滤波器进行离散小波变换, 包括低通滤波器 f_{LL} 和三个高通滤波器(水平方向滤波器 f_{LH} 、垂直方向滤波器 f_{HL} 、对角线方向滤波器 f_{HH})。经离散小波变换将图像 \mathbf{X} 分解成 4 幅子带图像, 即低频图像 \mathbf{X}_{LL} 和三幅高频子带图像 \mathbf{X}_{LH} 、 \mathbf{X}_{HL} 、 \mathbf{X}_{HH} 。4 个滤波器具有固定的参数, 以 Haar 小波为例, 可表示为

$$\begin{aligned} f_{LL} &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad f_{LH} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \\ f_{HL} &= \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \quad f_{HH} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}。 \end{aligned} \quad (9)$$

由(9)式可知, f_{LL} 、 f_{LH} 、 f_{HL} 、 f_{HH} 彼此相互正交, 且均为 2×2 的可逆矩阵。离散小波变换操作可表示为

$$\begin{aligned} \mathbf{X}_{LL} &= (f_{LL} * \mathbf{X}) \downarrow_2 \\ \mathbf{X}_{LH} &= (f_{LH} * \mathbf{X}) \downarrow_2 \\ \mathbf{X}_{HL} &= (f_{HL} * \mathbf{X}) \downarrow_2 \\ \mathbf{X}_{HH} &= (f_{HH} * \mathbf{X}) \downarrow_2 \end{aligned}, \quad (10)$$

式中, $*$ 为卷积操作, \downarrow_2 为采样因子为 2 的下采样操作, 即离散小波变换可通过 4 个步长为 2 的固定滤波器实现。根据 Haar 小波变换理论, \mathbf{X} 经过 2D 小波变换后, 各子带图像 \mathbf{X}_{LL} 、 \mathbf{X}_{LH} 、 \mathbf{X}_{HL} 、 \mathbf{X}_{HH} 在位置 (i, j) 的像素值可表示为

$$\begin{aligned} \mathbf{X}_{LL}(i, j) &= \mathbf{X}(2i-1, 2j-1) + \\ &\mathbf{X}(2i-1, 2j) + \mathbf{X}(2i, 2j-1) + \mathbf{X}(2i, 2j) \\ \mathbf{X}_{LH}(i, j) &= -\mathbf{X}(2i-1, 2j-1) - \end{aligned}$$

$$\begin{aligned}
 & \mathbf{X}(2i-1,2j) + \mathbf{X}(2i,2j-1) + \mathbf{X}(2i,2j) \\
 & \mathbf{X}_{\text{HL}}(i,j) = -\mathbf{X}(2i-1,2j-1) + \\
 & \mathbf{X}(2i-1,2j) - \mathbf{X}(2i,2j-1) + \mathbf{X}(2i,2j) \\
 & \mathbf{X}_{\text{HH}}(i,j) = \mathbf{X}(2i-1,2j-1) - \\
 & \mathbf{X}(2i-1,2j) - \mathbf{X}(2i,2j-1) + \mathbf{X}(2i,2j).
 \end{aligned}
 \tag{11}$$

2.4 小波胶囊网络

小波变换作为一种经典的谱分析方法,广泛应用于图像处理领域,其具有的多分辨率分析能力,可将图像分解成不同频率的子带图像。其中,低频子带图像保留图像的主要能量,高频子带图像可提供边缘纹理信息,对纹理图像特征的提取具有重要作用。

矩阵胶囊网络中,每个胶囊包含一个 4×4 的姿态矩阵和一个激活值,姿态矩阵中包含的神经元在进行特征表达的同时可编码特征间的相互关系,对目标对象的识别具有重要意义,激活值则表示目标对象存在的概率。原始矩阵胶囊网络是一个仅包含 5 层的浅层网络,其空间特征提取网络仅包含一层卷积层,在处理纹理图像这类具有复杂特征的图像时,无法充分表达纹理特征,限制了胶囊网络的性能。

DCNN 对复杂图像具有良好的特征表达能力,在图像分类、目标检测、图像分割等计算机视觉任务中的性能较好。但 DCNN 具有的特征平移不变性,在特征提取时只关心图像中是否存在某种特征,对

特征所处的位置、方向等信息不敏感,且池化层在减小特征的过程中会丢失大量有用的特征。本质上 DCNN 的空间特征提取过程是一种空间分析方法或一种受限的多分辨率分析方法。

胶囊网络擅长编码特征间的相互关系,CNN 擅长提取具有平移不变性的空间特征,小波变换则擅长捕捉具有尺度不变性的特征,在图像处理中也被称为图像显微镜。结合 DCNN 的空间特征提取能力与胶囊网络可编码特征间关系的能力,并将经典谱分析方法中的小波变换引入到改进的模型中,可提升模型的多分辨率分析能力。融合频率特征和空间特征,提出了小波胶囊网络,从而实现特征的自动提取,在常用纹理图像数据集上取得了较好的分类结果。

图 2 为小波胶囊网络的结构,其中,Conv 为卷积层,DWT 为离散小波变换层。表 1 为小波胶囊网络结构的具体参数。小波胶囊网络首先将输入图像进行一次步长为 2,卷积核大小为 7×7 的卷积运算,获取输入图像的浅层特征;然后对浅层特征进行 Haar 小波分解,得到 4 幅具有不同频率分量的子带特征图,并按照特征通道进行组合,减小浅层特征图尺寸的同时引入频率特征;其次利用两个 dense block 对组合后的特征图进行融合和深层次特征的提取;最后用初级胶囊层获得初级胶囊,通过 EM 路由算法进行迭代,得到最终的分类胶囊。

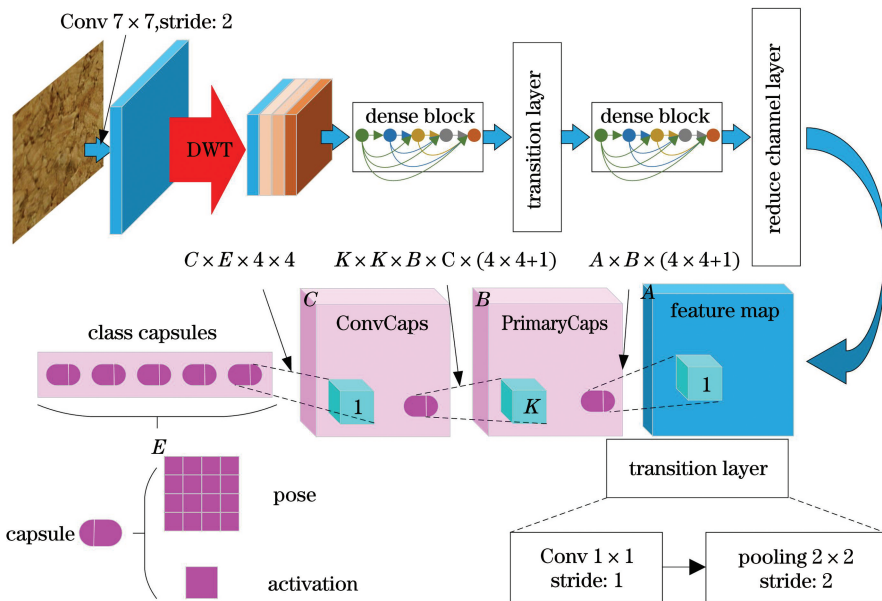


图 2 小波胶囊网络的结构

Fig. 2 Structure of wavelet capsule network

表1 小波胶囊网络的结构参数

Table 1 Structural parameter of wavelet capsule network

Layer	Output size	DWTCapsNet
Conv	64×64	kernel size: 7×7, stride: 2, with padding
DWT	32×32	wavelet decomposition
Dense block_1	32×32	$\begin{bmatrix} 1 \times 1 & \text{Conv} \\ 3 \times 3 & \text{Conv} \end{bmatrix} \times 6$
Transition	32×32	kernel size: 1×1, stride: 1
	16×16	2×2max pooling, stride: 2
Dense block_2	16×16	$\begin{bmatrix} 1 \times 1 & \text{Conv} \\ 3 \times 3 & \text{Conv} \end{bmatrix} \times 12$
Reduce Channel	16×16	kernel size: 3×3, stride: 1, with padding
PrimaryCaps	pose: 16×16×128 activation: 16×16×8	kernel size: 1×1, stride: 1
ConvCaps	pose: 7×7×256 activation: 7×7×16	kernel size: 3×3, stride: 2
Class Capsule	pose: $E \times 16$ activation: E	kernel size: 1×1, stride: 1

3 分析与讨论

3.1 实验环境介绍

实验环境:操作系统为 CentOS7 64 位,内存为 64 GB,CPU 为 Intel Core i7-6850K,包含 4 块 NVIDIA GeForce GTX Titan-XP GPU,编程语言为 Python 3.6,深度学习框架为 PyTorch,软件开发 IDE 为 Pycharm。

3.2 数据集

实验在 3 个常用纹理图像数据集上进行,分

别为 KTH-TIPS-2b(KTH)^[21]、Kylberg Texture Dataset v. 1.0(KTD)^[22]和 UIUC Texture Dataset(UIUC)^[23]。KTH 数据集是在 KTH-TIPS 数据集上扩展得到,包含更具挑战性的彩色纹理图像,共有 11 类在 9 种距离尺度、4 种视角和 3 种光照强度下获得的纹理图像样本,每类包含 4 个物理实体,每个物理实体包含 108 幅尺寸为 200 pixel×200 pixel 的样本图像,如图 3 所示。



图3 KTH 数据集

Fig. 3 KTH data set

KTD 数据集包含有旋转和无旋转的两种数据集,无旋转数据集共有 28 类纹理图像,每类图像由同一物体在相同采集条件下连续采集 4 次获得,原始图像的尺寸为 5184 pixel×3456 pixel,将原始图

像分割成 40 个尺寸为 576 pixel×576 pixel 的区域作为样本图像,并进行灰度处理,即每类有 160 幅灰度纹理图像,如图 4 所示。有旋转数据集则在无旋转数据集上每次旋转 30°,共旋转 12 次得到。

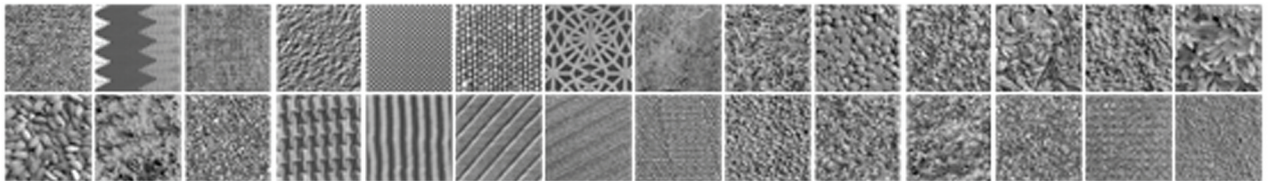


图4 KTD 数据集

Fig. 4 KTD data set

UIUC 数据集包含 25 种纹理类别,每类包含 40 幅尺寸为 640 pixel×480 pixel 的样本图像,每幅

图像都是在具有明显视角变化条件下随机采集,且来自于不同的物理实体,如图 5 所示。

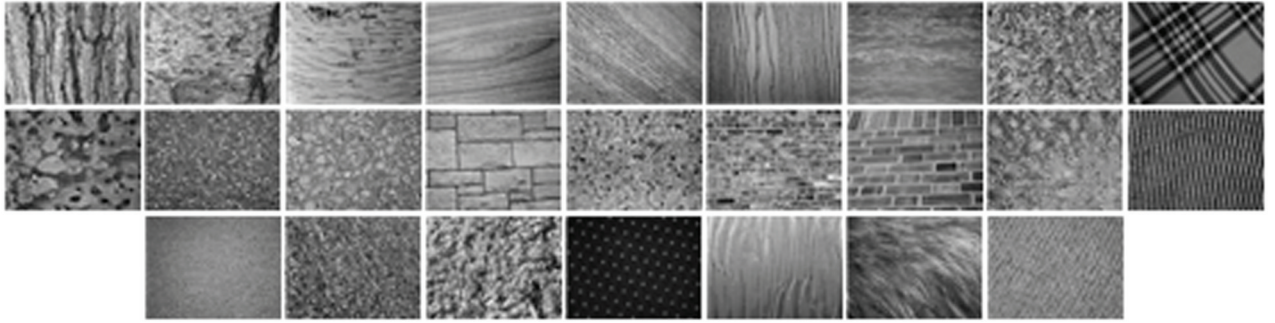


图 5 UIUC 数据集

Fig. 5 UIUC data set

上述 3 个数据集的训练集和测试集划分情况如表 2 所示。

表 2 不同数据集的划分情况

Table 2 Division of different data sets

Data set	KTH	KTD	UIUC
Train set	3564	3136	700
Test set	1188	1344	300
Class	11	28	25
Type	color	gray	gray
Total number of images	4752	4480	1000

3.3 小波胶囊网络纹理图像的分类性能测试

对小波胶囊网络纹理图像进行分类测试时,首先将纹理图像的尺寸缩放至 128 pixel×128 pixel 后作为小波胶囊网络的输入;然后利用数据扩增技术对输入图像进行数据扩增,包含对图像进行水平翻转、垂直翻转和随机旋转 15°等数据扩增操作,防止过拟合;最后进行 100 个 epoch 的迭代训练。设置模型的初始学习率为 0.001,每迭代训练 10 个 epoch,学习率调整为原来的一半。

小波胶囊网络在 3 个常用纹理图像数据集上训练和测试的分类准确率如图 6 所示。可以看出,小波胶囊网络的收敛速度较快,且没有出现过拟合和欠拟合现象。在 KTH、KTD 和 UIUC 数据集上,小波胶囊网络与传统纹理图像分类算法和基于 CNN 的纹理图像分类算法的分类准确率如表 3 所示。

从表 3 可以发现,相比局部二值模式(LBP)、提升局部二值模式(ILBP)和模糊局部二值模式(FLBP)算法,小波胶囊网络的分类准确率较高,与 GPBMFD (Gaussian pyramid approach to

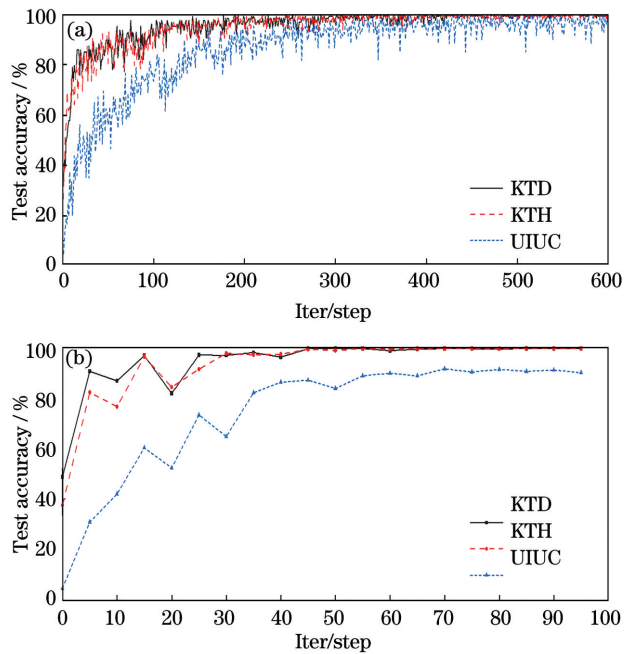


图 6 DWTCapsNet 在不同纹理数据集上的分类准确率。

(a) 训练集; (b) 测试集

Fig. 6 Classification accuracies of DWTCapsNet on different texture data sets. (a) Train set; (b) test set

Bouligand-Minkowski fractal descriptors)、局部量化编码(LQC)、排序局部梯度模式(SLGP)、局部凹凸微结构模式(LCCMSP)和局部四元模式(LQPAT)算法相比,在 KTH 数据集上的分类准确率最大可提升 16.16 个百分点,在 KTD 数据集上的分类准确率最大可提升 2.87 个百分点。这表明小波胶囊网络可实现特征的自动提取,学习到比人工设计特征更丰富的纹理特征表达,从而更好地描述纹理特征,取得更好的分类效果。

T-CNN-3 (Texture CNN)、视觉分类网络 (VisualNet)、16 层 VGGNet(VGG-VD-16)、深度纹

表 3 不同纹理分类算法的分类准确率

Table 3 Classification accuracies of different texture classification algorithms

Type	Model	KTH /%	KTD /%	UIUC /%	Reference
Traditional algorithm	LBP	89.3	97.8	88.1	Ref. [24]
	ILBP	96.9	99.7	91.9	Ref. [25]
	FLBP	94.3	99.2	91.3	Ref. [26]
	GPBMFD	—	—	90.30	Ref. [27]
	LQC	96.39	—	92.62	Ref. [28]
	SLGP	95.60	—	—	Ref. [29]
	LCCMSP	93.32	99.69	—	Ref. [30]
	LQPAT	83.76	96.92	—	Ref. [31]
CNN-based algorithm	T-CNN-3	99.4	73.2	—	Ref. [32]
	VisualNet	72.4	97.8	—	Ref. [33]
	VGG-VD-16	97.8	99.5	93.3	Ref. [11]
	Deep-TEN	84.5	—	—	Ref. [34]
	LFV	82.6	—	—	Ref. [35]
DWTCapsNet	99.92	99.79	91.74	ours	

理编码网络 (Deep-TEN) 和 LFV (Locally-transferred Fisher vectors) 等基于 CNN 的纹理图像分类算法利用了 CNN 强大的特征表达能力, 分类性能较好。而小波胶囊网络结合小波变换的多分辨分析能力、DCNN 的表达能力, 可同时学习特征间的相互关系, 进一步提升了算法的纹理特征表达能力, 获得更好的纹理分类效果。与基于 CNN 的纹理图像分类算法相比, 小波胶囊网络在不同纹理图像数据集上均取得了较好的分类效果。在 KTH 和 KTD 数据集上均达到了最高的分类准确率, 分别为 99.92% 和 99.79%。在 UIUC 数据集上, 小波胶囊网络的分类准确率为 91.74%。这证明了小波胶囊网络具有出色的纹理表达性能, 在纹理图像分类任务中十分有效。

为验证小波胶囊网络的数据区分能力, 以 KTD 数据集为例, 利用 t -分布式随机邻居嵌入 (t -SNE) 算法对小波胶囊网络在 KTD 训练集上的输出特征进行可视化, 结果如图 7 所示。其中, 所有颜色点根据真实标签进行着色, 不同灰度的点表示不同的纹理类别。可以发现, 小波胶囊的输出特征具有类内聚集紧凑和类间距离明确的特点, 这表明小波胶囊网络的输出特征具有良好的区分力, 有利于分类识别。

实验采用的数据集包含彩色图像和灰度图像,

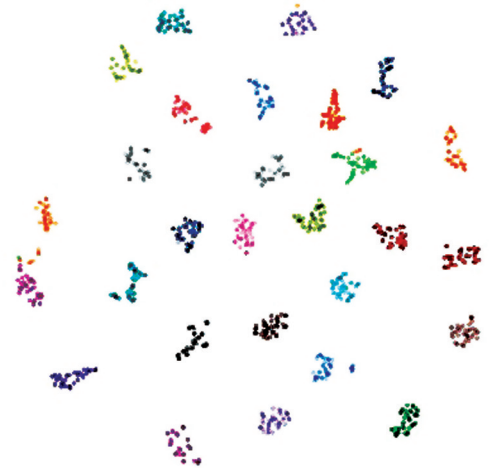


图 7 DWTCapsNet 输出特征的可视化效果

Fig. 7 Visualization of DWTCapsNet output feature

为了研究彩色图像和灰度图像对小波胶囊网络分类性能的影响, 将 KTH 数据集中的样本图像先进行灰度处理, 然后作为 DWTCapsNet 的输入进行实验, 结果如表 4 所示。其中, KTH-color 表示原始彩色样本图像, KTH-gray 表示灰度处理后的 KTH 样本图像。可以发现, 小波胶囊网络在彩色纹理图像和灰度纹理图像上的分类效果差异较小。原因是彩色纹理图像提供了更丰富的颜色特征, 小波胶囊网络可以有效利用彩色图像提供的色彩信息, 从而获得更高的分类准确率。

表4 DWTCapsNet在彩色和灰度纹理图像上的分类准确率
Table 4 Classification accuracies of DWTCapsNet on color and gray texture images unit: %

Data set	KTH-color	KTH-gray
Accuracy	99.92	96.88

3.4 消除实验

通过改变 DenseNet 和 Matrix CapsNet 的网络结构,验证离散小波变换和 DenseNet 对胶囊网络分类性能的提升效果。DenseNet 的浅层特征提取由 1 层 CNN 构成,然后进行最大值池化,对特征进行降采样。在降采样过程中,池化操作会导致部分浅层特征丢失,而离散小波变换的多分辨率分析能力可将 CNN 提取的特征分解为不同频率分量的子图,有效保留原始特征信息,提升识别效果。因此,首先将 DenseNet 第 1 层卷积层后的最大值池化层替换为离散小波变换层,保持其他参数不变,以验证离散小波变换带来的提升作用。实验在 3 个纹理图像数据集上进行训练和测试,迭代 100 次后,模型的分类结果如表 5 所示。其中,DWTDenseNet 为用离散小波变换层替换最大值池化层得到的网络。可以发现,引入离散小波变换可以提升 CNN 的特征利用率,减少特征丢失,提升模型的分类效果。

表5 DWTDenseNet 的分类结果

Table 5 Classification result of DWTDenseNet unit: %

Data set	KTH	KTD	UIUC
DenseNet	98.20	99.66	87.07
DWTDenseNet	99.38	99.72	87.83

矩阵胶囊网络是一个非常浅层的网络,其初级特征提取网络仅由 1 层卷积层构成,结构简单,不能充分提取输入图像的空间特征,在复杂纹理图像数据集上的分类效果较差。因此,用 DenseNet 中的特征提取网络替代矩阵胶囊网络的

初级特征提取网络,保持其他网络参数不变,以验证更复杂的初级特征提取网络对分类精度的提升作用。分别在 3 个纹理图像数据集上进行训练和测试,结果如表 6 所示,其中,DenseCapsNet 表示修改后的矩阵胶囊网络。可以发现,修改后的网络在纹理图像分类上的性能有进一步的提升,在小数据集上的性能提升尤为明显,提升了 42.81 个百分点,这表明用更深、更复杂的特征提取网络可提升胶囊网络的特征表达能力,从而抑制胶囊网络在小数据集上的过拟合现象。

表6 DenseCapsNet 的分类结果

Table 6 Classification result of DenseCapsNet

Data set	KTH	KTD	UIUC
MatrixCapsNet	93.18	84.73	45.31
DenseCapsNet	93.91	99.01	88.12

3.5 小波胶囊网络的抗噪声性能测试

小波胶囊网络在 3 个常用纹理图像数据集上都体现了强大的学习能力和表示能力,但自然场景中获取的纹理图像通常包含大量的噪声,而噪声对纹理图像的正确分类有很大的影响,因此模型的抗噪声能力在纹理图像分类任务中尤为重要。

保持 KTH 原有数据集的结构划分不变,在所有样本图像上添加高斯噪声 $N(0, \sigma^2)$ 作为新数据样本进行模型训练和测试,以模拟自然条件下采集的带噪声纹理图像。用方差 σ^2 作为噪声等级指标,依次添加 $\sigma^2 = 0.01, 0.03, 0.05, 0.07, 0.09$ 的噪声进行实验,加噪后的图像如图 8 所示。对包含不同噪声的样本图像进行 100 次迭代训练,得到不同噪声程度下 KTH 数据集测试集的分类准确率,如图 9 所示。可以发现,添加噪声后小波胶囊网络的分类准确率有所降低,但随着噪声程度的不断增加,噪声对小波胶囊网络分类准确率的影响在可接受范围内,这表明小波胶囊网络具有良好的噪声鲁棒性。

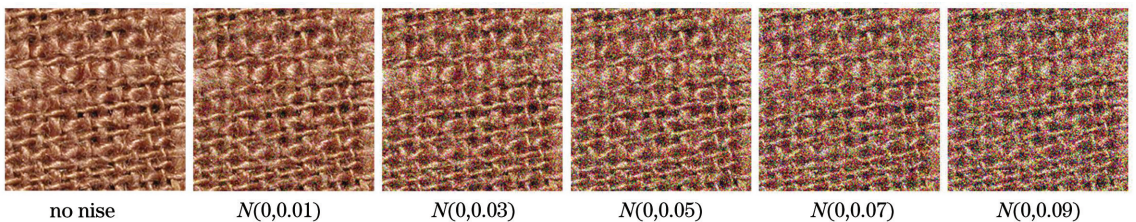


图8 添加不同级别高斯噪声后的图像

Fig. 8 Images after adding different levels of Gaussian noise

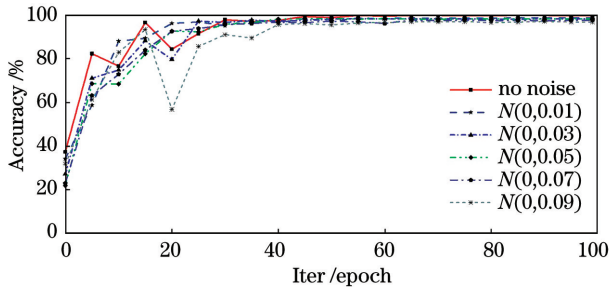


图9 小波胶囊网络在不同噪声级别上的分类准确率 (KTH数据集)

Fig. 9 Classification accuracies of wavelet capsule network at different noise levels (KTH dataset)

3.6 小波胶囊网络的抗旋转性能测试

小波胶囊网络在纹理图像分类任务上的优良性

表7 小波胶囊网络在旋转图像数据集上的分类准确率

Table 7 Classification accuracy of wavelet capsule network on rotating image data set

Rotation angle /($^{\circ}$)	Accuracy /%	Rotation angle /($^{\circ}$)	Accuracy /%	Rotation angle /($^{\circ}$)	Accuracy /%
0	99.71	120	41.26	240	41.26
30	67.25	150	57.10	270	50.99
60	44.60	180	96.66	300	43.11
90	51.27	210	64.27	330	62.00

4 结 论

为了提升胶囊网络在纹理图像数据集上的分类性能,融合了离散小波变换和DCNN的优点对矩阵胶囊网络进行改进,提出了小波胶囊网络。在3个常用纹理图像数据集上的实验结果表明,相比传统纹理分类算法,小波胶囊网络的分类准确率最大提升了3.02个百分点,相比CNN方法,小波胶囊网络的分类准确率也有一定的提升,且具有良好的噪声鲁棒性。但小波胶囊在处理旋转图像时,分类效果较差,因此,下一步工作还需研究该网络的抗旋转性能。

参 考 文 献

- [1] Li X C, Ma H Y, Lin S. Palmprint recognition based on subspace and texture feature fusion[J]. *Laser & Optoelectronics Progress*, 2019, 56(7): 071007.
李新春, 马红艳, 林森. 基于子空间与纹理特征融合的掌纹识别[J]. *激光与光电子学进展*, 2019, 56(7): 071007.
- [2] Yuan L S, Lou M Y, Liu Y Q, et al. Palm vein classification based on deep neural network and random forest[J]. *Laser & Optoelectronics*

能,离不开数据样本的多样性。当样本多样性不足时,模型学习到特征会相对减少,从而影响分类性能。旋转变换也是影响样本多样性的一种因素,为了测试小波胶囊网络的抗旋转性能,用KTD有旋转数据集进行实验。KTD的旋转数据具有12种旋转角度(0° 、 30° 、 60° 、 90° 、 120° 、 150° 、 180° 、 210° 、 240° 、 270° 、 300° 、 330°),将旋转 0° 的图像按3:1的比例随机划分为训练集和测试集,然后从每类不同角度的图像中随机抽取48幅图像作为测试集,小波胶囊网络训练过程中不会引入任何旋转变换,在不同旋转角度图像上的分类准确率如表7所示。可以发现,小波胶囊具有一定的抗旋转鲁棒性,且具有较大的提升空间。

Progress, 2019, 56(10): 101010.

袁丽莎, 娄梦莹, 刘娅琴, 等. 结合深度神经网络和随机森林的手掌静脉分类[J]. *激光与光电子学进展*, 2019, 56(10): 101010.

- [3] Yang H X, Chen Y, Zhang F, et al. Face recognition based on improved gradient local binary pattern[J]. *Laser & Optoelectronics Progress*, 2018, 55(6): 061004.
杨恢先, 陈永, 张翡, 等. 基于改进梯度局部二值模式的人脸识别[J]. *激光与光电子学进展*, 2018, 55(6): 061004.
- [4] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971-987.
- [5] Liu L, Fieguth P. Texture classification from random features[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(3): 574-586.
- [6] Varma M, Zisserman A. Texture classification: are filter banks necessary?[C]//2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 18-20, 2003, Madison, WI, USA. New York: IEEE, 2003: II-691.
- [7] Arvis V, Debain C, Berducat M, et al.

- Generalization of the cooccurrence matrix for colour images: application to colour texture classification [J]. *Image Analysis & Stereology*, 2011, 23(1): 63-72.
- [8] Tivive F H C, Bouzerdoum A. Texture classification using convolutional neural networks [C] // TENCON 2006-2006 IEEE Region 10 Conference, November 14-17, 2006, Hong Kong, China. New York: IEEE, 2006: 1-4.
- [9] Hafemann L G. An analysis of deep neural networks for texture classification [EB/OL]. [2020-04-15]. <http://dSPACE.c3sl.ufpr.br:8080/dspace/handle/1884/36967>.
- [10] Cimpoi M, Maji S, Vedaldi A. Deep filter banks for texture recognition and segmentation [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 3828-3836.
- [11] Bello-Cerezo R, Bianconi F, Cascianelli S, et al. Hand-designed local image descriptors vs. off-the-shelf CNN-based features for texture classification: an experimental comparison [M] // De Pietro G, Gallo L, Howlett R, et al. *Intelligent Interactive Multimedia Systems and Services 2017. Smart Innovation, Systems and Technologies*. Cham: Springer, 2017, 76: 1-10.
- [12] Mallat S G. A theory for multiresolution signal decomposition: the wavelet representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11(7): 674-693.
- [13] Fujieda S, Takayama K, Hachisuka T. Wavelet convolutional neural networks [EB/OL]. [2020-04-02]. <https://arxiv.org/abs/1805.08620>.
- [14] Liu P J, Zhang H Z, Lian W, et al. Multi-level wavelet convolutional neural networks [J]. *IEEE Access*, 2019, 7: 74973-74985.
- [15] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules [EB/OL]. [2020-04-05]. <https://arxiv.org/abs/1710.09829>.
- [16] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [17] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2261-2269.
- [18] Sabour S, Frosst N, Hinton G. Matrix capsules with EM routing [C] // 6th International Conference on Learning Representations, April 30-May 3, 2018, BC, Canada. Vancouver: ICLR, 2018: 1-15.
- [19] Mo W L. Research on gesture recognition method based on computer vision [D]. Guilin: Guangxi Normal University, 2019.
莫伟珑. 基于计算机视觉的手势识别方法研究 [D]. 桂林: 广西师范大学, 2019.
- [20] Su J L. Capsule: matrix Capsule and EM routing [EB/OL]. [2020-03-21]. <https://kexue.fm/archives/5155>.
苏剑林. 三味 Capsule: 矩阵 Capsule 与 EM 路由 [EB/OL]. [2020-03-21]. <https://kexue.fm/archives/5155>.
- [21] Hayman E, Caputo B, Fritz M, et al. On the significance of real-world conditions for material classification [M] // Pajdla T, Matas J, et al. *Computer Vision-ECCV 2004. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2004, 3024: 253-266.
- [22] Kylberg G. The Kylberg texture dataset v. 1.0 [EB/OL]. [2020-03-10]. https://www.researchgate.net/publication/249656240_Kylberg_Texture_Dataset_v_10.
- [23] Lazebnik S, Schmid C, Ponce J. A sparse texture representation using local affine regions [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(8): 1265-1278.
- [24] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions [J]. *Pattern Recognition*, 1996, 29(1): 51-59.
- [25] Jin H L, Liu Q S, Lu H Q, et al. Face detection using improved LBP under Bayesian framework [C] // Third International Conference on Image and Graphics (ICIG'04), December 18-20, 2004, Hong Kong, China. New York: IEEE, 2004: 306-309.
- [26] Iakovidis D K, Keramidias E G, Maroulis D. Fuzzy local binary patterns for ultrasound texture characterization [M] // Campilho A, Kamel M, et al. *Image Analysis and Recognition. ICIAR 2008. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2008, 5112: 750-759.
- [27] Florindo J B, Casanova D, Bruno O M. A Gaussian pyramid approach to Bouligand-Minkowski fractal descriptors [J]. *Information Sciences*, 2018, 459: 36-52.

- [28] Zhao Y, Wang R G, Wang W M, et al. Local quantization code histogram for texture classification [J]. *Neurocomputing*, 2016, 207: 354-364.
- [29] Song T C, Xin L L, Gao C Q, et al. Grayscale-inversion and rotation invariant texture description using sorted local gradient pattern [J]. *IEEE Signal Processing Letters*, 2018, 25(5): 625-629.
- [30] Elmerabet Y, Ruichek Y. Local concave-and-convex micro-structure patterns for texture classification [J]. *Pattern Recognition*, 2018, 76: 303-322.
- [31] Chakraborty S, Singh S K, Chakraborty P. Local quadruple pattern: a novel descriptor for facial image recognition and retrieval [J]. *Computers & Electrical Engineering*, 2017, 62: 92-104.
- [32] Andrearczyk V, Whelan P F. Using filter banks in convolutional neural networks for texture classification [J]. *Pattern Recognition Letters*, 2016, 84: 63-69.
- [33] Zheng H T, Fang L, Ji M Q, et al. Deep learning for surface material classification using haptic and visual information [J]. *IEEE Transactions on Multimedia*, 2016, 18(12): 2407-2416.
- [34] Zhang H, Xue J, Dana K. Deep TEN: texture encoding network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2896-2905.
- [35] Song Y, Zhang F, Li Q, et al. Locally-transferred fisher vectors for texture classification [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 4922-4930.