

# 纹理分类系统：基于深度学习的特征编码

张鹏杰

北京理工大学

北京市海淀区中关村南大街5号

3220231268@bit.edu.cn

贾康民

北京理工大学

北京市海淀区中关村南大街5号

3220231190@bit.edu.cn

马现涛

北京理工大学

北京市海淀区中关村南大街5号

3220231173@bit.edu.cn

## 摘要

纹理是一种反映图像中同质现象的视觉特征,纹理分类问题就是为待分类图像或图像区域指定一个预定义的纹理类别。本文参考纹理分类的基本理论,利用卷积神经网络(CNNs)完成纹理分类任务,并且基于深度学习以端到端的方式学习纹理的特征提取、特征分类,避免了对纹理特征的人为设计,提升了模型的表达能力。此外,本文在M1NC、DTD两个纹理数据集上对模型进行了测试,实验证明基于深度学习的神经网络模型可以很好地解决纹理分类问题。

## 1. 引言

在计算机视觉领域,纹理是一种反映图像中同质现象的视觉特征,它体现了物体表面的具有缓慢变化或者周期性变化的表面结构组织的排列属性。纹理包含了许多关键视觉信息,它们对于图像关键区域的识别和分析具有重要意义。

纹理分类问题,是指给定训练样本后,设计算法实现将输入图像或图像区域归类为某一种已知的纹理类别,而随着深度学习的发展,卷积神经网络(CNNs)成为许多分类算法、识别算法的实践标准,因此本文将基于深度学习,设计卷积神经网络解决纹理分类的问题。

相较于常规的物体识别与分类,传统的纹理分类方法有一个显著的技术挑战,即对纹理特征的提取和编码。一种常用的方法是利用手工设计的提取器如SIFT[1]完成特征提取,随后学习获得特征字典,并通过词袋(BoWs)对特征分布进行编码,最后利用分类器如SVM完成特征分类,如图1(左)所示。近期工作中,手工设计的提取器被预训练的卷积神经网络(CNNs)取代,词袋(BoWs)被鲁棒残差编码器如VLAD[2]、FV(Fisher Vector)[3]取代,如图1(中)所示,这证明深度神经网络在纹理分类中能发挥重要作用。

与以上流程不同,本文实现的方法抛弃了特征提取、字典学习、特征编码、分类器的组件堆叠,转而将卷积神经网络的设计贯穿整个算法,即以端到端的

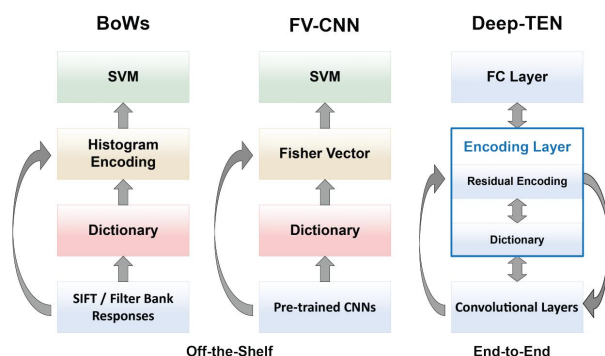


图 1: 传统方法与Deep-TEN的比较

方式学习调整整个模型[4],如图1(右)所示。

本文中的模型引入了一个新的可学习的残差编码层(Encoding Layer),它将字典学习和特征编码集成为CNN的一个层,其具有三个主要属性:(1)推广了鲁棒残差编码器如VLAD和FV,以无序的方式编码了纹理特征;(2)编码层能接受任意的输入大小,并提供固定长度的编码输出;(3)编码层学习一个固有的字典和编码表示,这使其能携带特定领域的信息,有利于域迁移的实现。

基于该编码层,我们可以串联起整个纹理分类算法的训练,从而设计出端到端的模型,该模型被称为DeepTEN,本文将详细阐述其原理,并展示其在纹理数据集上定量、定性的实验结果。

## 2. 相关工作

### 2.1. 纹理特征分析

在图像处理领域,纹理可以被定义为像素强度在空间域上的变化函数,纹理分析主要解决如何从图片中提取纹理特征,在物体识别、表面缺陷检测、模式识别、医学图像分析等计算机视觉领域发挥着重要的作用。

纹理分析方法通常可以分为四类:基于统计的方法、基于结构的方法、基于模型的方法和基于变换的方法[1]。

基于统计的方法对像素的亮度分布函数进行系列

的统计计算，并用统计特征表征纹理的特征向量，其中一阶统计特征仅统计单个像素的信息，不考虑像素间的相互作用，而二阶或更高阶统计特征会考虑两个或多个像素间的依赖性。直方图、灰度共生矩阵GLCM[10]、局部二值模式LBP[11]等方法都属于该类。

基于结构的方法基于初始视觉单元及其空间布局引入纹理特征，初始视觉单元可以是像素、区域或曲线，该方法将纹理视为初始单元在一定空间布局下的组合，一旦检测到主纹理，即可计算主纹理的视觉特征。该方法适用于描述规则结构的纹理，对不规则结构的纹理不适用。Sobel[12]边缘检测、SIFT[13]特征提取等方法属于该类。

基于模型的方法创建一个图像模型，并使用该模型描述图像的合成，模型的参数表征着纹理的基本性质。最流行的方法包括自回归方法、马尔可夫平方理论、分形[14]模型等。

基于变换的方法将图像转化到另一个域，使纹理在新的域空间中更容易区分，该方法中常见的有小波变换[15]、Gabor变换[16]等。

## 2.2. 基于BoW的纹理分类模型

深度学习的方法流行以前，纹理分类模型的基本框架是基于BoW的架构，可大致分为分块、特征表示、码本生成、特征编码、特征池化、特征分类六步[17]。

对于给定图像，首先需要根据需要进行分块处理，得到N个图像块，可以根据稀疏兴趣点[18]、固定网格[19]或在每个像素点[20]进行分块。

图像分块后，需要对每个块进行局部特征提取，理想情况下，局部描述符应该是独特的，同时对各种可能的图像转换具有鲁棒性，例如比例变换、旋转变换、光强变换等，具体的纹理特征表示方法可以参照第2.1节。

随后需要生成包含K个码字的码本，理想情况下每个码字都应该可以表征一种纹理特征，该码本既可以以预定义的方式给出[21]，也可以通过学习获得[22]。

给定码本和局部特征，我们可以进行特征编码，获得图像的纹理特征，特征编码是基于BoW的纹理分类方法的核心组件，它通过码本将不定长的局部特征转化为定长的特征，许多研究都集中于开发强大的特征编码，例如稀疏编码[23]、局部线性约束编码(LLC)[24]、局部聚合的向量描述符(VLAD)[2]、Fisher Vector(FV)[3]等。

获取局部纹理特征后，通常需要进行池化，以获得最终的全局纹理特征，经典的池化方法包括平均池化、最大池化和空间金字塔池化[25]等。

最后，基于全局纹理特征，我们利用分类器计算

得到纹理分类结果，分类器可以有多种选择[26]，如最近邻分类器(NNC)、支持向量机(SVM)、人工神经网络等。

## 3. 基于深度学习的纹理分类模型Deep-TEN

本文实现基于编码层(Encoding Layer)的纹理分类模型Deep Texture Encoding Networks(Deep-TEN)，在本章中，我们将讨论Encoding Layer的实现细节，以及Deep-TEN的算法实现流程。

### 3.1 编码层(Encoding Layer)

**残差编码模型。**给定N个视觉描述符 $X = \{x_1, \dots, x_N\}$ ，以及一个学习得到的包含K个D维码字的码本 $C = \{c_1, \dots, c_K\}$ ，每个描述符 $x_i$ 可以通过权重 $a_{ik}$ 分配到每个码字 $c_k$ ，并且相应的残差向量可以表示为 $r_{ik} = x_i - c_k$ 。给定分配权重和残差向量，残差编码模型为每个码字计算聚合值 $e_k$ ：

$$e_k = \sum_{i=1}^N e_{ik} = \sum_{i=1}^N a_{ik} r_{ik} \quad \#(1)$$

最终的编码输出结果是固定长度的特征表示 $E = \{e_1, \dots, e_K\}$ 。

**编码层。**传统的视觉识别方法可以分为特征提取、字典学习、特征池化、分类器学习，如图1所示。在本文的方法中，我们将字典学习和残差编码聚合到单层CNN中，即编码层，编码层以完全监督的方式同时学习编码参数和固有字典。下面让我们讨论为描述符分配的权重。

硬分配为每个描述符 $x_i$ 仅提供一个非零权重，它对应与描述符在距离上最接近的码字，而其他对应码字的权重均为零。用公式表示其k个分配权重如下所示，因此硬分配不考虑码字的模糊性并且使模型不可微分。

$$a_{ik} = 1(\|r_{ik}\|^2 = \min\{\|r_{i1}\|^2, \dots, \|r_{iK}\|^2\}) \quad \#(2)$$

软分配为每个描述符 $x_i$ 的对应码字分配一个相应的权重，该权重可以表示为下式，其中 $\beta$ 为平滑因子。

$$a_{ik} = \frac{\exp(-\beta\|r_{ik}\|^2)}{\sum_{j=1}^K \exp(-\beta\|r_{ij}\|^2)} \quad \#(3)$$

受高斯混合模型(GMM)的启发，我们进一步允许每个聚类中心 $c_k$ 的平滑因子 $s_k$ 是可学习的，这为描述符分布提供了更强大的模型表达能力：

$$a_{ik} = \frac{\exp(-s_k\|r_{ik}\|^2)}{\sum_{j=1}^K \exp(-s_j\|r_{ij}\|^2)} \quad \#(4)$$

**端到端学习。**编码层最终构成一个有向无环图，如图 2 所示，该层所有组件的参数都是可微分的，所以我们可以利用基于反向传播的梯度下降算法对编码层进行训练，由此我们将字典学习和特征编码纳入深度学习的框架中。

### 3.2 纹理分类模型Deep-TEN

基于编码层，我们可以实现端到端的纹理分类模型Deep-TEN。如表 1 所示，Deep-TEN的前向传播过程可以分为四个阶段：（1）首先利用预训练的 50 层ResNet进行特征提取；（2）随后利用瓶颈层将特征通道数由 2048 缩小至 128，以减少计算量；（3）然后利用带有 32 个码字的编码层完成特征编码；（4）最后利用L2范数和全连接层计算分类结果。

下面讨论Deep-TEN所具有的特性。

**域迁移。**由于编码层的残差编码有着削弱数据集中频繁特征的影响这一特性，而这些特征通常包含特定域的信息，因此Deep-TEN更倾向于学习域不变的特征，从而具有较强的域迁移能力。具体而言，可以考虑一个在数据中频繁出现的特征描述符 $x_i$ ，它将接近于一个聚类中心 $c_k$ ，并且对应的残差向量 $r_{ik} = x_i - c_k$ 的值将很小，而对 $j \neq k$ 的残差向量 $r_{ij}$ ，其分配权重 $a_{ij}$ 将很小，因此最终的聚合结果的值将偏小，由此使频繁出现的特征描述符造成较小的影响。

**多尺度训练。**Deep-TEN可以接受任意尺度的图像作为输入，在ResNet卷积层提取出特征后，编码层可以将任意尺寸大小的特征转化为固定尺寸的特征，该特征可以被用于后续的全连接层计算分类结果。由此我们可以利用不同尺度的图像输入促进模型的训练，并且这也为多数据集的训练提供了便利。

**联合训练。**针对不同的视觉问题，有许多可用的数据集，如物体分类、物体检测、材料识别等，Deep-TEN可以应用于多任务学习，并从多任务联合训练中学习共有深度特征，从而在多个任务上都获得性能的提升。

## 4. 实验分析

在本章中，我们将在纹理数据集MINC、DTD上对Deep-TEN模型进行训练和测试，并展示其定量、定性的测试结果。

### 4.1. 数据集

本文利用MINC和DTD两个数据集进行实验分析：

**Materials in Context Database (MINC)** [8]是一个大规模的自然纹理图像数据库，在本文中，我们利用

	output size	Deep-TEN 50
Conv1	176×176×64	7×7, stride 2
Res1	88×88×256	3×3 max pool, stride 2 $\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Res2	44×44×512	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Res3	22×22×1024	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Res4	11×11×2048	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
Projection + Reshape	121×128	conv 1×1, 2048⇒128 W×H×D⇒N×D
Encoding	32×128	32 codewords
L2-norm + FC	n classes	1×1 FC

表 1: Deep-TEN模型架构

它的一个公开可用的子集MINC-2500，通过其提供的训练、测试分割，进行模型的训练和评估。它包含了 23 个纹理类别，每个类别包含 2500 张图像。

**Describable Textures Dataset (DTD)** [9]是一个自然的纹理图像数据库，包含 5640 张图像，并且根据人类感知将其分成 47 个类，每个类包含 120 张图像。数据被分为训练、验证、测试集，各数据集包含所有类，并且每个类包含 40 张图像。

此外，我们还调研了KTH-TIPS和FMD两个数据集：

**KTH-TIPS** [28]数据集提供在不同光照、姿态下的纹理数据，包括 11 个类，每个类 4 个样本集，每个样本集包括 108 张图像。

**Flickr Material Dataset (FMD)** [27]是一个材料识别数据集，该数据集包含塑料、金属、皮革、布料等十种物体的 1000 张图片，每种物体 100 张图片。

### 4.2. 实验配置

我们分别在MINC和DTD上进行训练和测试。训练时设置数据增强，将输入图片随机缩放裁剪为 352×352 的大小，并以 0.5 的概率进行随机水平、竖直翻转；测试时分别将MINC和DTD的图像中心裁剪为 352×352 的大小。

训练的参数配置如下所示：

**MINC:** 训练次数为 30 Epoch, 利用SGD进行优化，初始学习率 lr=0.1，动量 momentum=0.9，权重衰减



weight\_decay=1e-4, 并且学习率每经过一个Epoch衰减0.9倍。

**DTD:** 训练次数为 120 Epoch, 利用SGD进行优化, 初始学习率lr=0.01, 动量momentum=0.9, 权重衰减weight\_decay=1e-4, 并且学习率每经过一个Epoch衰减0.95倍。

### 4.3. 定量结果

模型Deep-TEN在MINC、DTD的分类准确率如表 2 所示, 其中Top-N代表将预测概率最高的N个类中包含真实标签的结果视为预测正确。

数据集	准确率(Top-1)	准确率(Top-5)
MINC	0.742	0.960
DTD	0.599	0.857

表 2: Deep-TEN在MINC和DTD的分类准确率

在MINC数据集上, Deep-TEN输出 23 分类的结果, 表 3 展示了以某纹理为一类, 其他纹理为另一类计算得出的分类精确率 (Precision)、召回率 (Recall) 和特异度 (Specificity) 的部分结果, 完整结果见附录。

纹理	精确率	召回率	特异度
sky	0.952	0.956	0.998
foliage	0.872	0.807	0.994
wallpaper	0.868	0.795	0.994
food	0.864	0.761	0.994
water	0.832	0.870	0.992
hair	0.820	0.768	0.992
leather	0.812	0.846	0.991
other	0.808	0.805	0.991
brick	0.804	0.910	0.991
carpet	0.780	0.741	0.990

表 3: Deep-TEN在MINC的测试指标

在DTD数据集上, Deep-TEN输出 47 分类的结果, 与MINC数据集类似地计算精确率、召回率、特异度, 如表 4 所示, 完整结果见附录。

纹理	精确率	召回率	特异度
wrinkled	0.975	0.830	0.999
stratified	0.950	0.731	0.999
flecked	0.925	0.569	0.998
honeycombed	0.900	0.947	0.998
spiralled	0.900	0.621	0.998
crosshatched	0.900	0.837	0.998
swirly	0.875	0.636	0.997
cobwebbed	0.850	0.442	0.997
waffled	0.850	0.694	0.997
marbled	0.850	0.680	0.997

表 4: Deep-TEN在DTD的测试指标

### 4.4. 定性结果

图 2 是Deep-TEN在MINC数据集上的定性结果展示, 图中展示了brick、wood、stone、skin、leather、water六类纹理图像的分类结果, 数值代表分类结果的确信度。可见在MINC数据集上, Deep-TEN可以输出正确的纹理分类结果, 并且结果的置信度很高。






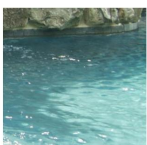
	<b>brick</b> 0.972		<b>wood</b> 0.900
	stone 0.007		fabric 0.012
	tile 0.003		wallpaper 0.010
	carpet 0.002		painted 0.009
	other 0.002		stone 0.007
	<b>stone</b> 0.950		<b>skin</b> 0.896
	brick 0.014		fabric 0.033
	water 0.007		hair 0.017
	tile 0.006		plastic 0.013
	foliage 0.003		food 0.005
	<b>leather</b> 0.944		<b>water</b> 0.933
	fabric 0.028		fabric 0.027
	metal 0.003		hair 0.004
	wood 0.003		plastic 0.004
	tile 0.003		food 0.003

图 2: Deep-TEN在MINC的定性结果

图 3 是Deep-TEN在DTD数据集上的定性结果展示, 图中展示了wrinkled、stratified、flecked、spiralled、swirly、waffled六类纹理图像的分类结果, 数值代表分类结果的确信度。可见在DTD数据集上, Deep-TEN仍可以输出正确的纹理分类结果, 并且对同时存在的次要纹理属性也能正确预测, 如对于图 3 中“stratified”的纹理分类结果, 其第二个纹理属性“cracked”也能很好的描述示例图。







	<b>wrinkled</b> 0.156		<b>stratified</b> 0.156
	cracked 0.047		cracked 0.087
	smeared 0.046		potholed 0.079
	gauzy 0.045		porous 0.060
	pleated 0.036		marbled 0.043
	<b>flecked</b> 0.149		<b>spiralled</b> 0.896
	porous 0.117		swirly 0.033
	potholed 0.094		braided 0.017
	potholed 0.071		cobwebbed 0.013
	marbled 0.042		veined 0.005
	<b>swirly</b> 0.183		<b>waffled</b> 0.177
	paisley 0.132		freckled 0.055
	interlaced 0.076		frilly 0.048
	lacelike 0.070		bumpy 0.047
	spiralled 0.057		knitted 0.041

图 3: Deep-TEN在DTD的定性结果

## 5. 结论

纹理分析是计算机视觉和模式识别领域长期存在的、基本的、具有挑战性的问题，因其对理解人类视觉的重要意义，以及在许多视觉应用中的重要作用，对纹理的研究一直是值得深入探索的课题。

对于纹理分类而言，传统的、基于BoW的纹理分类架构需要人为的对各个模块进行设计，利用人工设计的方法提取纹理特征并完成分类，这一方式既需要大量的知识和脑力劳动，又使模型的能力受到限制，存在一定的缺陷。

基于神经网络的深度学习模型为纹理分类提供了一种新的研究思路，本文实现的纹理分类模型Deep-TEN以端到端的方式完成对纹理图像的分类，相较于传统方法泛化性更强。实验测试中，Deep-TEN在MINC和DTD数据集上都表现出了良好的性能，这证明本文实现的模型方法能很好的解决纹理分类问题。

## 参考文献

- [1] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision, 2004, 60: 91-110.
- [2] Jégou H, Douze M, Schmid C, et al. Aggregating local descriptors into a compact image representation[C]//2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, 2010: 3304-3311.
- [3] Perronnin F, Sánchez J, Mensink T. Improving the fisher kernel for large-scale image classification[C]//Computer Vision-ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11. Springer Berlin Heidelberg, 2010: 143-156.
- [4] Zhang H, Xue J, Dana K. Deep ten: Texture encoding network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 708-717.
- [5] Fekri-Ershad S. Texture image analysis and texture classification methods-A review[J]. arXiv preprint arXiv:1904.06554, 2019.
- [6] Liu L, Chen J, Fieguth P, et al. From BoW to CNN: Two decades of texture representation for texture classification[J]. International Journal of Computer Vision, 2019, 127: 74-109.
- [7] Zhang H, Xue J, Dana K. Deep ten: Texture encoding network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 708-717.
- [8] Bell S, Upchurch P, Snavely N, et al. Material recognition in the wild with the materials in context database[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3479-3487.
- [9] Cimpoi M, Maji S, Kokkinos I, et al. Describing textures in the wild[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 3606-3613.
- [10] Haralick R M, Shanmugam K, Dinstein I H. Textural features for image classification[J]. IEEE Transactions on systems, man, and cybernetics, 1973 (6): 610-621.
- [11] 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.
- [12] Sobel I. An isotropic  $3 \times 3$  image gradient operator, Machine vision for three-dimensional scenes, 376-379[J]. 1990.
- [13] Lowe D G. Object recognition from local scale-invariant features[C]//Proceedings of the seventh IEEE international conference on computer vision. Ieee, 1999, 2: 1150-1157.
- [14] Mandelbrot B B. Ebookstore Release Benoit B Mandelbrot Fractals: Form[J]. Chance And Dimension prc, Freeman, 1977: 365.
- [15] Scheunders P, Livens S, Van de Wouwer G, et al. Wavelet-based texture analysis[J]. International Journal on Computer Science and Information Management, 1998, 1(2): 22-34.
- [16] Idrissa M, Acheroy M. Texture classification using Gabor filters[J]. Pattern Recognition Letters, 2002, 23(9): 1095-1102.
- [17] Liu L, Chen J, Fieguth P, et al. From BoW to CNN: Two decades of texture representation for texture classification[J]. International Journal of Computer Vision, 2019, 127: 74-109.
- [18] Zhang J, Marszałek M, Lazebnik S, et al. Local features and kernels for classification of texture and object categories: A comprehensive study[J]. International journal of computer vision, 2007, 73: 213-238.
- [19] Sharan L, Liu C, Rosenholtz R, et al. Recognizing materials using perceptually inspired features[J]. International journal of computer vision, 2013, 103: 348-371.
- [20] Varma M, Zisserman A. A statistical approach to material classification using image patch exemplars[J]. IEEE transactions on pattern analysis and machine intelligence, 2008, 31(11): 2032-2047.
- [21] Ojala T, Pietikäinen M, Maenpää 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.
- [22] Varma M, Zisserman A. A statistical approach to texture classification from single images[J]. International journal of computer vision, 2005, 62: 61-81.
- [23] Peyré G. Sparse modeling of textures[J]. Journal of mathematical imaging and vision, 2009, 34: 17-31.
- [24] Wang J, Yang J, Yu K, et al. Locality-constrained linear coding for image classification[C]//2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, 2010: 3360-3367.
- [25] Timofte R, Van Gool L. A Training-free Classification Framework for Textures, Writers, and Materials[C]//BMVC. 2012, 13: 14.
- [26] Jain A K, Duin R P W, Mao J. Statistical pattern recognition: A review[J]. IEEE Transactions on pattern analysis and machine intelligence, 2000, 22(1): 4-37.
- [27] Liu C, Sharan L, Adelson E H, et al. Exploring features in a bayesian framework for material recognition[C]//2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, 2010: 239-246.
- [28] Caputo B, Hayman E, Mallikarjuna P. Class-specific material categorisation[C]//Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. IEEE, 2005, 2: 1597-1604.

附录

表 5、表 6 分别为模型Deep-TEN在DTD、MINC数据集上的完整测试结果。

纹理	精确率	召回率	特异度
knitted	0.800	0.508	0.996
cracked	0.000	0.000	0.979
porous	0.375	0.600	0.987
chequered	0.675	0.692	0.993
dotted	0.150	0.667	0.982
honeycombed	0.900	0.947	0.998
stratified	0.950	0.731	0.999
interlaced	0.825	0.647	0.996
crystalline	0.525	0.677	0.990
matted	0.750	0.638	0.995
veined	0.775	0.564	0.995
cobwebbed	0.850	0.442	0.997
braided	0.650	0.317	0.992
waffled	0.850	0.694	0.997
spiralled	0.900	0.621	0.998
gauzy	0.700	0.683	0.993
blotchy	0.600	0.585	0.991
zigzagged	0.325	0.650	0.985
crosshatched	0.900	0.837	0.998
sprinkled	0.500	0.800	0.989
swirly	0.875	0.636	0.997
frilly	0.700	0.800	0.993
studded	0.400	0.471	0.987
woven	0.200	0.471	0.983
lined	0.050	0.250	0.980
perforated	0.675	0.643	0.993
flecked	0.925	0.569	0.998
grid	0.725	0.604	0.994
fibrous	0.125	0.263	0.981
lacelike	0.575	0.411	0.991
bumpy	0.375	0.536	0.987
banded	0.200	0.364	0.983
scaly	0.575	0.719	0.991
pitted	0.775	0.738	0.995
stained	0.050	0.333	0.980
meshed	0.250	0.455	0.984
smearred	0.475	0.475	0.989
polka-dotted	0.475	0.268	0.988
grooved	0.800	0.593	0.996
bubbly	0.750	0.909	0.995
pleated	0.800	0.744	0.996
freckled	0.650	0.634	0.992
paisley	0.700	0.757	0.993
marbled	0.850	0.680	0.997
striped	0.550	0.579	0.990
potholed	0.650	0.578	0.992
wrinkled	0.975	0.830	0.999

表 5: Deep-TEN在DTD的测试指标

纹理	精确率	召回率	特异度
brick	0.804	0.91	0.991
carpet	0.78	0.741	0.99
ceramic	0.684	0.592	0.986
fabric	0.596	0.634	0.982
foliage	0.872	0.807	0.994
food	0.864	0.761	0.994
glass	0.768	0.619	0.989
hair	0.82	0.768	0.992
leather	0.812	0.846	0.991
metal	0.496	0.642	0.977
mirror	0.616	0.819	0.983
other	0.808	0.805	0.991
painted	0.772	0.748	0.99
paper	0.752	0.689	0.989
plastic	0.572	0.53	0.98
polishedstone	0.656	0.729	0.984
skin	0.776	0.735	0.99
sky	0.952	0.956	0.998
stone	0.772	0.726	0.99
tile	0.644	0.735	0.984
wallpaper	0.868	0.795	0.994
water	0.832	0.87	0.992
wood	0.564	0.695	0.98

表 6: Deep-TEN在MINC的测试指标

图 4、图 5 分别为模型Deep-TEN在MINC、DTD数据集上的混淆矩阵。

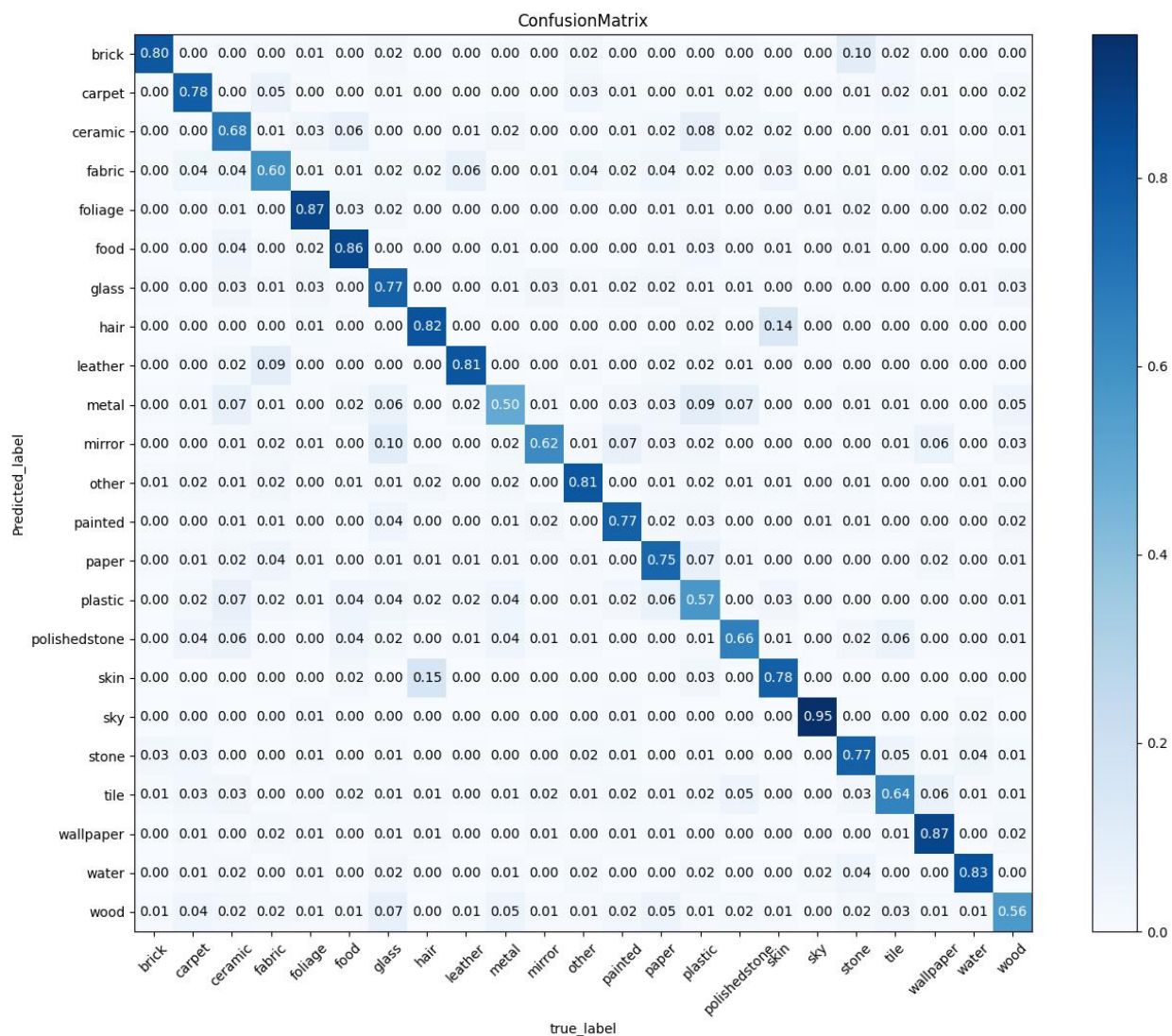


图 4: Deep-TEN在MINC的混淆矩阵



