



(12) 发明专利申请

(10) 申请公布号 CN 114708593 A

(43) 申请公布日 2022. 07. 05

(21) 申请号 202111673248.0

(22) 申请日 2021.12.31

(71) 申请人 北京工业大学

地址 100124 北京市朝阳区平乐园100号

(72) 发明人 汤健 王子轩 张晓晓 荆中岭

(74) 专利代理机构 北京思海天达知识产权代理

有限公司 11203

专利代理师 刘萍

(51) Int. Cl.

G06V 30/40 (2022.01)

G06K 9/62 (2022.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

G06V 30/10 (2022.01)

G06V 10/764 (2022.01)

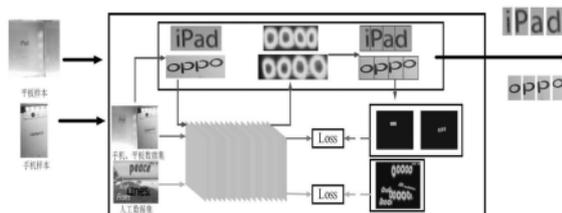
权利要求书4页 说明书8页 附图7页

(54) 发明名称

基于异构多模型的废旧电子产品品牌识别方法

(57) 摘要

本文提出基于异构多模型的废旧电子产品识别方法,针对受限于相关数据集的稀少,识别方法精度难以达到实际工业需求的问题。利用CTAFT算法提取电子产品背部字符区域,利用ImageNet预训练的VGG19模型作为图像特征嵌入模型,提取待回收电子产品字符部分和整体特征;针对字符部分特征构建OCR字符识别模型,得到OCR子模型识别结果,针对字符和整体特征构建深度森林分类模型,得到深度森林子模型识别结果;将OCR识别结果和深度森林分类向量线性组合,使用softmax非线性函数得到类别权重向量,将权重最高的结果作为电子产品品牌识别结果。基于废旧电子产品回收装备拍摄的真实手机与平板图像验证了有效性。



1. 基于异构多模型的废旧电子产品识别方法,其特征在于,包括:图像预处理模块、多元特征提取模块、异构多模型识别模块共3个部分;

出现变量含义如表1所示;

表1变量含义表

| 变量名 | 含义 |
|--|-------------------------|
| OCR | OCR 字符识别模型 |
| RandomForest/RF | 随机森林模型 |
| GBDT | 梯度提升树模型 |
| DeepForest/DF | 深度森林模型 |
| LSTM | 长短记忆神经网络模型 |
| $X_{img}^{original}$ | 待回收废旧电子产品图像 |
| X_{img} | 数据增强预处理后图像 |
| X_{digit} | 字符预处理后图像 |
| X_{img}^{1024} | 数据增强预处理后图像特征提取结果 |
| $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}]$ | 字符预处理后图像特征提取结果 |
| m | 当前待识别样本字符数量 |
| $[\hat{y}_1]_{OCR}$ | OCR 字符识别模块识别结果 |
| $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{DF}$ | 深度森林识别子模型识别结果 |
| n | 废旧电子产品分类类数 |
| DF | 深度森林识别子模型 |
| $[\hat{y}_1^{OCR}, \hat{y}_1^{DF}, \hat{y}_2^{DF}, \dots, \hat{y}_n^{DF}]$ | 异构多模型识别结果列表 |
| $\hat{y}_{predict}$ | 废旧电子产品预测结果 |
| k | OCR 字符识别模型中 LSTM 基本单元数量 |
| $l(w)$ | 电子产品图像文本框长度 |
| $l^c(w)$ | 分水岭算法分割字符串长度结果 |
| $S_{conf}(w)$ | 分水岭算法分割结果评分 |

| | |
|-------------------------------|-------------------------------|
| $S_c(p)$ | 分水岭算法置信度 |
| $\mathbf{X}_{\text{digit}}^i$ | 字符图像中第 i 个字符 |
| $f_{\text{VGG}}(\cdot)$ | VGG19 模型输出过程 |
| \bar{h}_i | 前向 LSTM 模型第 i 个 LSTM 基本单元输出 |
| \bar{h}_i | 反向 LSTM 模型第 i 个 LSTM 基本单元输出 |
| h_i | 双向 LSTM 模型第 i 个 LSTM 基本单元输出 |
| π | 双向 LSTM 输出可能情况 |
| Y | CTC 网络可能输出结果 |
| $B(\cdot)$ | CTC 网络输出过程 |
| e | 自然对数 |
| $\sigma(y_j)$ | Softmax 映射后 y_j 对应权值 |

图像预处理模块的输入为 $\mathbf{X}_{\text{img}}^{\text{original}}$ ，数据增强预处理输出为 X_{img} ，使用CRAFT字符级目标检测算法的字符预处理输出为 X_{digit} ；

多元特征提取模块使用基于ImageNet预训练的VGG19网络获取字符特征与整体背部图像像素特征在高维空间的表示，该模块的输入为 X_{img} 和 X_{digit} ，输出分别为 X_{img}^{1024} 和 $[X_{\text{digit}}^{1,512}, X_{\text{digit}}^{2,512}, \dots, X_{\text{digit}}^{m,512}]$ ；

异构多模型识别模块包括含OCR字符识别子模块、深度森林电子产品识别子模块和softmax非线性输出层子模块共3部分，其中：OCR子模块输入为 $[X_{\text{digit}}^{1,512}, X_{\text{digit}}^{2,512}, \dots, X_{\text{digit}}^{m,512}]$ 输出为 $[\hat{y}_1]_{\text{OCR}}$ ，深度森林子模块输入为 $[X_{\text{digit}}^{1,512}, X_{\text{digit}}^{2,512}, \dots, X_{\text{digit}}^{n,512}, X_{\text{img}}^{1024}]$ 输出为 $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{\text{DF}}$ ，softmax非线性输出层子模块将分类子模块输出结果进行映射，获得 $[\hat{y}_1^{\text{OCR}}, \hat{y}_1^{\text{DF}}, \hat{y}_2^{\text{DF}}, \dots, \hat{y}_n^{\text{DF}}]$ ，得分最高的标签即为最终输出 \hat{y}_{predict} ；

图像预处理模块包括数据增强预处理和字符增强预处理；

字符增强预处理选取电子产品背部图像中的字符特征作为分类依据之一，采用CRAFT字符级图像定位算法确定电子产品字符位置并分割，将字符特征与整体图片线性拼接，作为后续模型输入；

对于人工数据集，数据集包含图中单个字符的高斯热度图，CRAFT算法对本部分进行有监督训练；对于电子产品背部图像数据集，首先标记电子产品图像中的文本框区域，将其经过透视变换拉伸到比较正的文本框；接着利用分水岭算法得到单个字符的位置框，生成对应的高斯热度图，变换后重新贴回原图对应的标签图的相应的位置；计算分水岭算法分割结果评分公式如下所示：

$$S_{conf}(w) = \frac{l(w) - \min(l(w), |l(w) - l^c(w)|)}{l(w)} \quad (1)$$

其中, $l(w)$ 表示电子产品图像文本框长度, $l^c(w)$ 为分水岭算法分割字符串长度结果;

分水岭算法分割得到字符串长度后根据公式(1)得到算法评价,若评价与真实字符长度一致则置信度 $S_c(p)$ 为1,评分越低说明分割结果可信度越差;

工业设备采集的电子产品图像表示为 X_{img} ,其尺寸为400*300,经过图像预处理后其字符图像表示为 $\mathbf{X}_{digit} = [X_{digit}^1, X_{digit}^2, \dots, X_{digit}^i, \dots, X_{digit}^m]$,其中 X_{digit}^i 表示字符图像中第 i 个字符,单个字符尺寸为50*50,整体字符图像大小即为50*(50*m); m 字符图像输出维度与五字符图像输出维度一致;

该模块采用基于ImageNet预训练的VGG19模型作为基模型,首先,将VGG19模型中卷积层和池化层参数固化;接着,针对不同图像特征构建不同大小全连接层;最后,将不同图像的模型输出线性组合,作为后续分类模型的输入特征;多元特征提取后特征维度由全连接层维度决定;针对不同大小的手机图像 X_{img} 和字符图像 X_{digit} 特征提取过程如公式(2)所示:

$$\begin{cases} X_{img} \xrightarrow{f_{VGG}(\cdot)} X_{img}^{1024} \\ [X_{digit}^1, X_{digit}^2, \dots, X_{digit}^m] \xrightarrow{f_{VGG}(\cdot)} [X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}] \end{cases} \quad (2)$$

其中, $f_{VGG}(\cdot)$ 表示VGG19模型输出过程;

构建“OCR字符识别模型”和“深度森林识别模型”,子模型结构如下所示;

a) OCR字符识别模型子模块

在OCR电子产品背部字符识别流程中,仅使用字符特征 $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}]$ 作为输入;首先通过双向LSTM提取含有完整上下文信息的字符序列特征;接着,通过CTC网络解决输入特征与输出序列无法对其的问题;最后,通过莱文斯坦距离确定OCR输出字符串与已知标签之间距离,最终获得电子产品品牌分类结果 $[\hat{y}_1]_{OCR}$;

OCR字符识别模型根据图像预处理部分所得字符特征 $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}]$ 构建 k (其中 $k > m$) 个LSTM基本单元;双向LSTM网络包含两个子网络结构,公式(3)、(4)分别表示前项和后项传递;

$$\bar{h}_i = (X_{digit}^{i,512}, \bar{h}_{i-1}) \quad (3)$$

$$\bar{h}_i = (X_{digit}^{i,512}, \bar{h}_{i+1}) \quad (4)$$

其中, k 为LSTM基本单元超参数, \bar{h}_i 表示前向LSTM在时刻 i 的输出结果, \bar{h}_i 表示后向LSTM在时刻 i 的输出结果,表示第 i 个输入 x_i , 第 i 时刻的双向LSTM输出为:

$$h_i = \bar{h}_i + \bar{h}_i \quad (5)$$

接着,CTC网络对双向LSTM网络输出 $[h_1, h_2, \dots, h_x]$ 中重复识别的字符去重,使其变为 $[y_1, y_2, \dots, y_n]$;由于双向LSTM基本单元多于手机字符数量 n ,导致字符重复划分,例如“honor”会被分割为“hoonor”;“hoonor”多个子串可映射为正确结果“honor”,如公式(6)所示

$$\begin{aligned}
 B(\pi_1) &= B(honorr) \longrightarrow honor \\
 B(\pi_2) &= B(honor) \longrightarrow honor \\
 B(\pi_n) &= B(hoonorr) \longrightarrow honor
 \end{aligned} \tag{6}$$

CTC网络通过给定输入X最大化后验概率 $P\langle Y|X \rangle$ 获得最终结果Y,其中 $P\langle Y|X \rangle$ 如公式(7)所示:

$$P\langle Y|X \rangle = \sum_{\pi \in B(Y)} P\langle \pi|X \rangle \tag{7}$$

其中, $\pi \in B(Y)$ 表示所有可整合为Y的子串集合;

b) 深度森林识别模型子模块

在深度森林废旧电子产品图像识别流程中,将上文所述字符特征 $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{n,512}]$ 和图像特征 X_{img}^{1024} 线性组合得到深度森林输入特征 X_{DF} ,该过程如公式(8)所示:

$$X_{DF} = [X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{n,512}, X_{img}^{1024}] \tag{8}$$

首先,使用 X_{DF} 构建不同随机森林,获得不同随机森林输出 \hat{y}_i^{RF} ;接着将随机森林输出 \hat{y}_i^{RF} 与 X_{DF} 线性组合,作为输入传入下一层模型构建不同随机森林,通过当前模型的分类精度决定是否继续构建下一层网络模型;最后,当模型精度不再提高时结束模型增长,将最后多个随机森林分类结果加权后获得最终分类结果 $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{DF}$;

c) 多模型输出加权模块子模块

所述异构多模型识别模块中OCR字符识别模型输出 $[\hat{y}_1]_{OCR}$ 为连续字符串,经过基于距离度量的映射后输出为某一电子产品品牌,而深度森林识别模型输出 $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{DF}$ 为所有电子产品品牌的概率;为解决异构模型输出形式不同或输出结果不一致的情况,在分类模型最后加入多模型输出加权模块;

softmax函数又称归一化指数函数,是当前深度学习研究中广泛使用在深度网络有监督学习部分的分类器;Softmax函数如公式(9)所示:

$$\sigma(y_j) = \frac{e^{y_j}}{\sum_{k=1}^{n+1} e^{y_k}} \quad j \in [1, n+1] \tag{9}$$

其中, $n+1$ 表示异构多模型输出向量维度,e表示自然对数;分类模型中,设废旧电子产品标签共n个,OCR字符识别结果 $[\hat{y}_1]_{OCR}$,深度森林识别结果 $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{DF}$,线性拼接后获得 $n+1$ 维结果向量 $[\hat{y}_1^{OCR}, \hat{y}_1^{DF}, \hat{y}_2^{DF}, \dots, \hat{y}_n^{DF}]$ 作为softmax函数输入,最终获得 $[\hat{y}_1^{OCR}, \hat{y}_1^{DF}, \hat{y}_2^{DF}, \dots, \hat{y}_n^{DF}]$ 对应权重,权重最高的标签作为最终废旧电子产品分类结果 $\hat{y}_{predict}$ 。

基于异构多模型的废旧电子产品品牌识别方法

技术领域

[0001] 本发明属于废旧电子产品的回收领域。

背景技术

[0002] 随着科技的发展和5G的迅速普及,智能电子产品的更迭速度不断加快。据Strategy Analytics预测,2021年全球智能电子产品的出货量将同比反弹6.5%,总量达13.8亿部。人们更换电子产品速度的加快是导致其出货量不断提高的主要原因,这也导致个人闲置电子产品的累计量逐年增多。因此,国内外市场对电子产品回收产业的回收效率提出更高要求。废旧电子产品作为一种典型的城市再生资源,利用无人化、智能化的回收装备对其进行回收,能够节省大量人工成本。其中,智能化废旧电子产品识别方法是完成上述任务的关键。

[0003] 图像识别已在目标检测,人脸识别等领域有了广泛应用,如何使用相关数据集构建分类模型对废旧电子产品进行智能化识别也成为了当下智能化回收装备的研究重点。但基于图像的深度学习神经网络模型构建依赖海量的有标记样本。本发明所述废旧电子产品识别问题数据集仅来源于回收装备样机实拍图片,数据量较小,难以构建行之有效的神经网络分类模型,且工业过程中拍摄图像清晰度较低,用户的不规范操作也会导致电子产品图像完整度较差、电子产品部分区域镜像等问题。如何在样本量少且样本质量不高的前提下完成对电子产品品牌分类已成为当前待解决的主要问题。

[0004] 基于上述研究现状,本发明作者提出“基于并行差分进化-梯度特征深度森林的废旧电子产品识别系统”,利用废旧手机背部图像构建手机品牌分类模型,模型分类准确率可达80.12%;“基于光学字符识别的废旧电子产品识别系统”,利用废旧电子产品背部字符构建字符分类模型,通过映射规则将字符分类结果映射为废旧电子产品品牌,该模型分类准确率可达86.37%。但上述方法仅从纹理特征、字符特征等单一角度构建分类模型,模型精度仍难达到实际工业需求。因此本发明提出一种基于异构多模型的废旧电子产品识别方法。

[0005] 首先,利用CTAFT算法提取电子产品背部字符区域;接着,使用ImageNet预训练的VGG19模型对电子产品背部图像和其字符特征区域进行特征提取,使用高维卷积特征代替单一维度特征;接着基于字符特征构建光学字符识别(OCR)模型,基于电子产品图像特征和字符特征构建深度森林电子产品分类模型;最后,将不同模型分类结果线性拼接,通过Softmax激活函数得出最终分类结果。基于工业和信息化部电信设备认证中心的典型电子产品图像数据集,验证了该算法在废旧电子产品识别中的有效性。

发明内容

[0006] 本发明所述基于异构多模型的废旧电子产品识别方法包括:图像预处理模块、多元特征提取模块、异构多模型识别模块共3个部分。系统整体结构如图1所示。

[0007] 本发明出现变量含义如表1所示。

[0008] 表1变量含义表

| 变量名 | 含义 |
|--|------------------|
| OCR | OCR 字符识别模型 |
| RandomForest/RF | 随机森林模型 |
| GBDT | 梯度提升树模型 |
| DeepForest/DF | 深度森林模型 |
| LSTM | 长短记忆神经网络模型 |
| $X_{img}^{original}$ | 待回收废旧电子产品图像 |
| X_{img} | 数据增强预处理后图像 |
| X_{digit} | 字符预处理后图像 |
| X_{img}^{1024} | 数据增强预处理后图像特征提取结果 |
| $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}]$ | 字符预处理后图像特征提取结果 |
| m | 当前待识别样本字符数量 |
| $[\hat{y}_1]_{OCR}$ | OCR 字符识别模块识别结果 |
| $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{DF}$ | 深度森林识别子模型识别结果 |
| n | 废旧电子产品分类类数 |
| DF | 深度森林识别子模型 |
| $[\hat{y}_1^{OCR}, \hat{y}_1^{DF}, \hat{y}_2^{DF}, \dots, \hat{y}_n^{DF}]$ | 异构多模型识别结果列表 |
| $\hat{y}_{predict}$ | 废旧电子产品预测结果 |

[0009]

| | | |
|---------------|------------------------|-------------------------------|
| [0010] | k | OCR 字符识别模型中 LSTM 基本单元数量 |
| | $l(w)$ | 电子产品图像文本框长度 |
| | $l^c(w)$ | 分水岭算法分割字符串长度结果 |
| | $S_{conf}(w)$ | 分水岭算法分割结果评分 |
| | $S_c(p)$ | 分水岭算法置信度 |
| | X_{digit}^i | 字符图像中第 i 个字符 |
| | $f_{VGG}(\cdot)$ | VGG19 模型输出过程 |
| | \bar{h}_i | 前向 LSTM 模型第 i 个 LSTM 基本单元输出 |
| | \bar{h}_i | 反向 LSTM 模型第 i 个 LSTM 基本单元输出 |
| | h_i | 双向 LSTM 模型第 i 个 LSTM 基本单元输出 |
| | π | 双向 LSTM 输出可能情况 |
| | Y | CTC 网络可能输出结果 |
| | $B(\cdot)$ | CTC 网络输出过程 |
| | e | 自然对数 |
| $\sigma(y_j)$ | Softmax 映射后 y_j 对应权值 | |

[0011] 图像预处理模块的输入为 $X_{img}^{original}$ ，数据增强预处理输出为 X_{img} ，使用CRAFT字符级目标检测算法的字符预处理输出为 X_{digit} ；

[0012] 多元特征提取模块使用基于ImageNet预训练的VGG19网络获取字符特征与整体背部图像像素特征在高维空间的表示，该模块的输入为 X_{img} 和 X_{digit} ，输出分别为 X_{img}^{1024} 和 $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}]$ ；

[0013] 异构多模型识别模块包括含OCR字符识别子模块、深度森林电子产品识别子模块和 softmax非线性输出层子模块共3部分，其中：OCR子模块输入为 $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}]$ 输出为 $[\hat{y}_1]_{OCR}$ ，深度森林子模块输入为 $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{n,512}, X_{img}^{1024}]$ 输出为 $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{DF}$ ，softmax非线性输出层子模块将分类子模块输出结果进行映射，获得 $[\hat{y}_1^{OCR}, \hat{y}_1^{DF}, \hat{y}_2^{DF}, \dots, \hat{y}_n^{DF}]$ ，得分最高的标签即为最终输出 $\hat{y}_{predict}$ 。

[0014] 2.1图像预处理模块

[0015] 2.1.1数据增强预处理

[0016] 数据增强随机改变训练样本可以降低模型对某些属性的依赖，从而提高模型的泛

化能力。数据增强方法主要包括几何变换、颜色空间变换、内核过滤器、混合图像、随机擦除、对抗训练、基于生成对抗网络的增强、神经风格转移。其中,几何变换可以解决训练样本位置偏差问题,本发明所述废旧电子产品回收流程中,电子产品图像会因用户摆放位置不同产生位置偏差,因此,本发明所用数据增强以几何变换为主。具体包括:旋转、翻折、镜像、平移、加入高斯噪声等。

[0017] 2.1.2字符增强预处理

[0018] 在电子产品回收流程中,用户的不规范操作会导致摄像头采集图像不完整,电子产品图像镜像等问题,直接使用回收装备获取的图片进行模型构建和电子产品品牌预测效果不佳。电子产品背部字符是识别其品牌的重要依据,但在用户使用过程中会造成字符磨损、字符遮挡,仅将其背部字符作为分类依据会导致模型具有很大局限性。因此本发明作者选取电子产品背部图像中的字符特征作为分类依据之一,采用CRAFT字符级图像定位算法确定电子产品字符位置并分割,将字符特征与整体图片线性拼接,作为后续模型输入。以解决难以依据手机图像构建分类模型、单一字符特征构建模型具有局限性的问题。

[0019] 已有大量实验表明,YOLO3、Fast-RCNN等目标检测算法在人脸检测、车牌检测等领域广泛应用,但该类算法检测目标长宽比较为固定,且大多不会出现形变,磨损等问题,在电子产品回收问题中,摆放位置的差异会导致字符串发生旋转和形变,废旧电子产品部分字符也会有磨损现象,因此直接标定字符所在位置进行训练,效果较差,对电子产品图像数据集中单个字符进行人工标注工作量难以估计,且已知目标检测数据集鲜有字符级别标注。CRAFT算法通过弱监督学习的方式对带有字符标签的人造数据集进行训练,当没有字符划分的电子产品背部图片作为输入时,模型会检测合成产生相应字符标签再进行识别,该算法通过字符之间紧密程度预测文本所在区域。CRAFT模型训练过程如图2所示。

[0020] 对于人工数据集,数据集包含图中单个字符的高斯热度图,CRAFT算法对本部分进行有监督训练;对于电子产品背部图像数据集,首先标记电子产品图像中的文本框区域,将其经过透视变换拉伸到比较正的文本框;接着利用分水岭算法得到单个字符的位置框,生成对应的高斯热度图,变换后重新贴回原图对应的标签图的相应的位置。计算分水岭算法分割结果评分公式如下所示:

$$[0021] \quad S_{conf}(w) = \frac{l(w) - \min(l(w), |l(w) - l^c(w)|)}{l(w)} \quad (1)$$

[0022] 其中, $l(w)$ 表示电子产品图像文本框长度, $l^c(w)$ 为分水岭算法分割字符串长度结果。

[0023] 分水岭算法分割得到字符串长度后根据公式(1)得到算法评价,若评价与真实字符长度一致则置信度 $S_c(p)$ 为1,评分越低说明分割结果可信度越差。

[0024] 工业设备采集的电子产品图像可表示为 X_{img} ,其尺寸为 $400*300$,经过2.1小节图像预处理后其字符图像可表示为 $\mathbf{X}_{digit}=[\mathbf{X}_{digit}^1, \mathbf{X}_{digit}^2, \dots, \mathbf{X}_{digit}^i, \dots, \mathbf{X}_{digit}^m]$,其中 \mathbf{X}_{digit}^i 表示字符图像中第 i 个字符,单个字符尺寸为 $50*50$,整体字符图像大小即为 $50*(50*m)$ 。为直观描述多元特征提取模块像素变化,本发明图3以五字符图像为例, X_{digit} 尺寸为 $50*250$ 。多元特征提取模块输出维度由人为确定,因此, m 字符图像输出维度与五字符图像输出维度一致。多元特征提取模块结构如图3所示。

[0025] 该模块采用基于ImageNet预训练的VGG19模型作为基模型,首先,将VGG19模型中

卷积层和池化层参数固化；接着，针对不同图像特征构建不同大小全连接层；最后，将不同图像的模型输出线性组合，作为后续分类模型的输入特征。多元特征提取后特征维度由全连接层维度决定。针对不同大小的手机图像 X_{img} 和字符图像 X_{digit} 特征提取过程如公式(2)所示：

$$[0026] \quad \begin{cases} \mathbf{X}_{img} \xrightarrow{f_{VGG}(\cdot)} X_{img}^{1024} \\ [\mathbf{X}_{digit}^1, \mathbf{X}_{digit}^2, \dots, \mathbf{X}_{digit}^m] \xrightarrow{f_{VGG}(\cdot)} [X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}] \end{cases} \quad (2)$$

[0027] 其中， $f_{VGG}(\cdot)$ 表示VGG19模型输出过程。

[0028] 2.3异构多模型识别模块

[0029] 仅从纹理特征、字符特征等单一角度构建电子产品分类模型已被验证其精度仍难达到实际工业需求。因此本发明采取Stracking集成思想，将不同角度特征线性组合，构建异构多分类模型，通过多个不同模型的集成提高整体模型精度。针对前文所述废旧电子产品品牌分类问题，构建“OCR字符识别模型”和“深度森林识别模型”，子模型结构如下所示。

[0030] 2.3.1OCR字符识别模型子模块

[0031] 在OCR电子产品背部字符识别流程中，仅使用2.2小节所述字符特征 $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}]$ 作为输入。首先通过双向LSTM提取含有完整上下文信息的字符序列特征；接着，通过CTC网络解决输入特征与输出序列无法对其的问题；最后，通过莱文斯坦距离确定OCR输出字符串与已知标签之间距离，最终获得电子产品品牌分类结果 $[\hat{Y}]_{OCR}$ 。OCR字符识别模型结构如图4所示。

[0032] 如图4所示，OCR字符识别模型根据图像预处理部分所得字符特征 $[X_{digit}^{1,512}, X_{digit}^{2,512}, \dots, X_{digit}^{m,512}]$ 构建 k （其中 $k > m$ ）个LSTM基本单元。双向LSTM网络包含两个子网络结构，公式(3)、(4)分别表示前项和后项传递。

$$[0033] \quad \bar{h}_i = (X_{digit}^{i,512}, \bar{h}_{i-1}) \quad (3)$$

$$[0034] \quad \bar{h}_i = (X_{digit}^{i,512}, \bar{h}_{i+1}) \quad (4)$$

[0035] 其中， k 为LSTM基本单元超参数， \bar{h}_i 表示前向LSTM在时刻 i 的输出结果， \bar{h}_i 表示后向LSTM在时刻 i 的输出结果，表示第 i 个输入 x_i ，第 i 时刻的双向LSTM输出为：

$$[0036] \quad h_i = \bar{h}_i + \bar{h}_i \quad (5)$$

[0037] 接着，CTC网络对双向LSTM网络输出 $[h_1, h_2, \dots, h_x]$ 中重复识别的字符去重，使其变为 $[y_1, y_2, \dots, y_n]$ 。由于双向LSTM基本单元多于手机字符数量 n ，导致字符重复划分，例如“honor”会被分割为“honorrr”。“honorrr”多个子串可映射为正确结果“honor”，如公式(6)所示

$$[0038] \quad \begin{aligned} B(\pi_1) &= B(honorrr) \longrightarrow honor \\ B(\pi_2) &= B(honor) \longrightarrow honor \\ B(\pi_n) &= B(honorrr) \longrightarrow honor \end{aligned} \quad (6)$$

[0039] CTC网络通过给定输入 X 最大化后验概率 $P\langle Y|X \rangle$ 获得最终结果 Y ，其中 $P\langle Y|X \rangle$ 如公式(7)所示：

$$[0040] \quad P\langle Y|X \rangle = \sum_{\pi \in B(Y)} P\langle \pi|X \rangle \quad (7)$$

[0041] 其中, $\pi \in B(Y)$ 表示所有可整合为Y的子串集合。

[0042] 2.3.2深度森林识别模型子模块

[0043] 在深度森林废旧电子产品图像识别流程中, 将上文所述字符特征 $[X_{\text{digit}}^{1,512}, X_{\text{digit}}^{2,512}, \dots, X_{\text{digit}}^{m,512}]$ 和图像特征 X_{img}^{1024} 线性组合得到深度森林输入特征 X_{DF} , 该过程如公式 (8) 所示:

$$[0044] \quad X_{\text{DF}} = [X_{\text{digit}}^{1,512}, X_{\text{digit}}^{2,512}, \dots, X_{\text{digit}}^{n,512}, X_{\text{img}}^{1024}] \quad (8)$$

[0045] 首先, 使用 X_{DF} 构建不同随机森林, 获得不同随机森林输出 \hat{y}_i^{RF} ; 接着将随机森林输出 \hat{y}_i^{RF} 与 X_{DF} 线性组合, 作为输入传入下一层模型构建不同随机森林, 通过当前模型的分类精度决定是否继续构建下一层网络模型; 最后, 当模型精度不再提高时结束模型增长, 将最后多个随机森林分类结果加权后获得最终分类结果 $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{\text{DF}}$ 。深度森林识别模型结构如图5所示。

[0046] 2.3.3多模型输出加权模块子模块

[0047] 所述异构多模型识别模块中OCR字符识别模型输出 $[\hat{y}_1]_{\text{OCR}}$ 为连续字符串, 经过基于距离度量的映射后输出为某一电子产品品牌, 而深度森林识别模型输出 $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{\text{DF}}$ 为所有电子产品品牌的概率。为解决异构模型输出形式不同或输出结果不一致的情况, 本发明在分类模型最后加入多模型输出加权模块。多模型加权输出模块如图6所示。

[0048] softmax函数又称归一化指数函数, 是当前深度学习研究中广泛使用在深度网络有监督学习部分的分类器。Softmax函数如公式 (9) 所示:

$$[0049] \quad \sigma(y_j) = \frac{e^{y_j}}{\sum_{k=1}^{n+1} e^{y_k}} \quad j \in [1, n+1] \quad (9)$$

[0050] 其中, $n+1$ 表示异构多模型输出向量维度, e 表示自然对数。在本发明所述分类模型中, 设废旧电子产品标签共 n 个, OCR字符识别结果 $[\hat{y}_1]_{\text{OCR}}$, 深度森林识别结果 $[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]_{\text{DF}}$, 线性拼接后获得 $n+1$ 维结果向量 $[\hat{y}_1^{\text{OCR}}, \hat{y}_1^{\text{DF}}, \hat{y}_2^{\text{DF}}, \dots, \hat{y}_n^{\text{DF}}]$ 作为softmax函数输入, 最终获得 $[\hat{y}_1^{\text{OCR}}, \hat{y}_1^{\text{DF}}, \hat{y}_2^{\text{DF}}, \dots, \hat{y}_n^{\text{DF}}]$ 对应权重, 权重最高的标签作为最终废旧电子产品分类结果 \hat{y}_{predict} 。

[0051] 1. 使用ImageNet预训练的VGG19网络对废旧电子产品图像进行特征提取。随着卷积层的深入, 单个特征的感受野不断增大, 其表征能力不断增强, 更优于单一角度特征提取方法。相比仅对纹理敏感的HOG特征构建深度森林分类模型, 使用VGG19构建的模型精度明显提高。

[0052] 2. 采用异构多模型方法构建废旧手机电子产品分类模型, 通过同一数据集对不同任务构建分类模型, 最后通过非线性函数对多个模型输出进行加权得到最终分类结果。通过实验证明。本发明所提异构多模型相比单一OCR识别模型和单一深度森林识别模型精度明显提高。

附图说明

[0053] 图1异构多模型的废旧电子产品识别方法结构图

[0054] 图2CRAFT定位裁剪模块结构图

[0055] 图3多元特征提取模块

- [0056] 图4OCR字符识别模型
- [0057] 图5深度森林识别模型
- [0058] 图6多模型加权输出结构图
- [0059] 图7废旧电子产品回收装备应用场景
- [0060] 图8数据增强效果图
- [0061] 图9图像预处理结果
- [0062] 图10OCR识别模型混淆矩阵
- [0063] 图11深度森林识别模型混淆矩阵
- [0064] 图12异构多模型废旧电子产品分类模型混淆矩阵

具体实施方式

[0065] 废旧电子产品回收装备的应用场景如图7所示,本发明实验数据源于该装备的实拍图片。数据集共123张图像,包含10个类别的废旧电子产品品牌,分别是华为手机(HUAWEI)、华为平板(MatePad)荣耀(Honor)、小米(Mi)、中兴(ZTE)、OPPO、VIVO、苹果(iPhone)、苹果平板(iPad),其他品牌(Others)。

[0066] 由于废旧电子装备实拍样本较少,因此在构建分类模型前采用数据增强手段对训练集和测试集样本进行扩充。以honor手机背部图像为例,将手机背部图像进行旋转、翻折、加噪声等操作,将1张手机背部图像样本扩充至12张,样本总量有400张图像扩充至4800张。样本扩充示意图如图8所示。

[0067] 接着,采用CRAFT字符分割算法对待回收废旧电子产品图像进行分割,获得对应电子产品字符数据集,图像预处理结果如图9所示。

[0068] 多元特征提取部分使用1400万张图片2万类的ImageNet数据集对VGG19模型预训练,预训练后的VGG模型表示为 $f_{VGG}(\cdot)$ 。多元特征提取部分根据输入图像的不同,在VGG模型中加入不同大小的全连接层,其中针对400*300的废旧电子产品图像加入1024维全连接层,针对50*50的废旧电子字符图像加入512维全连接层。

[0069] OCR字符识别模块为EasyOCR中英文字符预训练模型,构建128个LSTM基本单元,即 $k=128$ 。深度森林识别模型使用随机森林、GBDT作为每层深度森林模型的基分类器,其中,均使用100颗决策树构建RF和GBDT,使用L1+L2正则化优化GBDT损失函数。

[0070] 使用图像预处理后字符图片构建OCR字符识别模型分类混淆矩阵如图10所示,使用废旧电子产品图像和字符图片构建深度森林识别模型分类混淆矩阵如图11所示。

[0071] 通过多模型输出加权模块对上述2个模型的结果进行集成,进而得到基于异构多模型的废旧电子产品分类模型混淆矩阵,如图12所示,分类精度可达90.17%,。

[0072] 为验证本发明所述方法有效性,本发明使用相同废旧电子产品数据集分别构建单一特征+深度森林、VGG特征+深度森林、VGG特征+OCR模型的10分类模型。基于废旧电子产品图像数据集的不同品牌分类模型精度如表1所示。

[0073] 表1废旧电子产品识别模型精度对比表

| | 方法 | 精度 |
|--------|---------------------|--------|
| [0074] | HOG 特征+深度森林识别模型 | 80.12% |
| | VGG19 特征+深度森林识别模型 | 87.93% |
| | VGG19 特征+OCR 字符识别模型 | 86.37% |
| | VGG19 特征+异构多模型 | 90.17% |

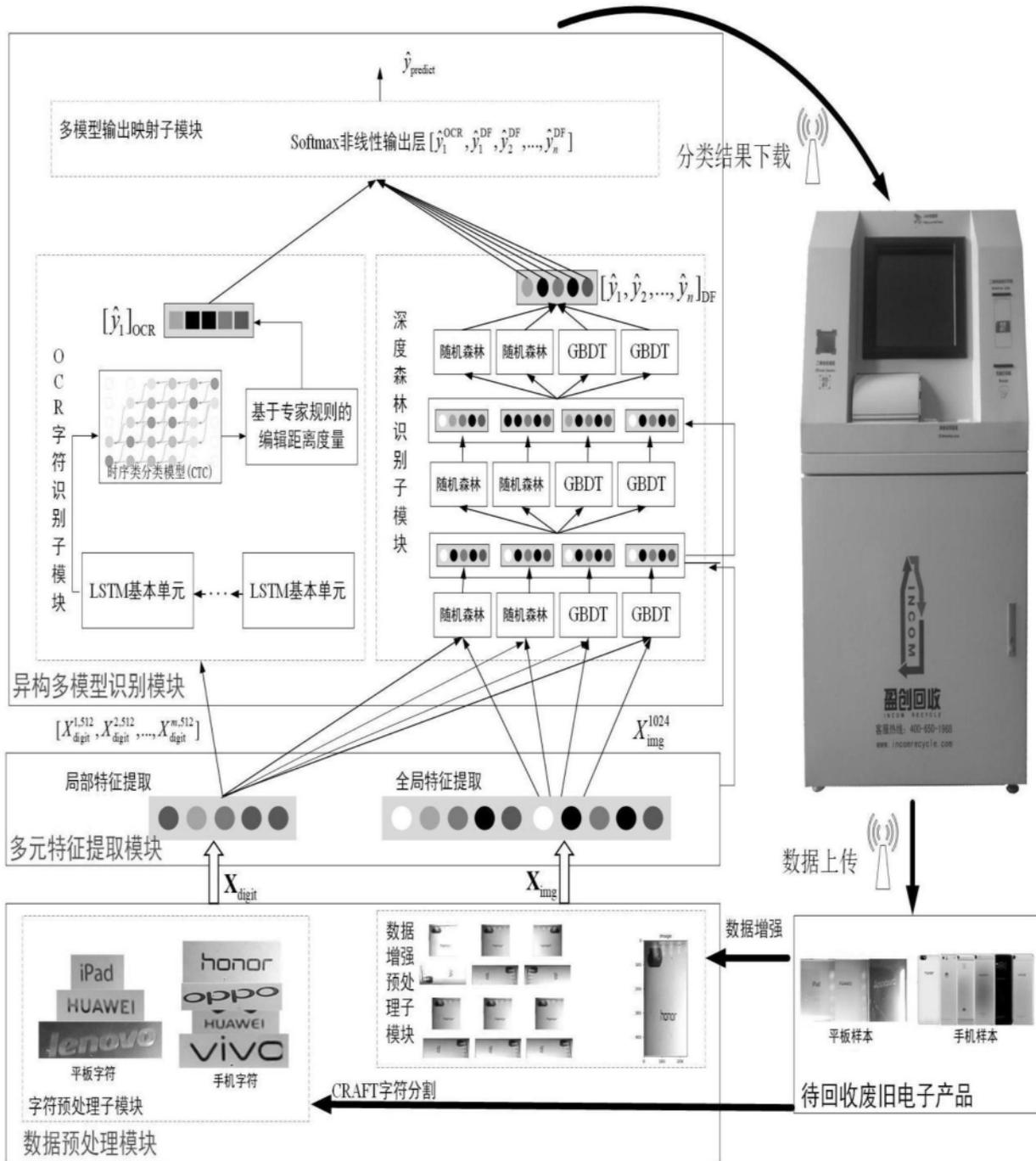


图1

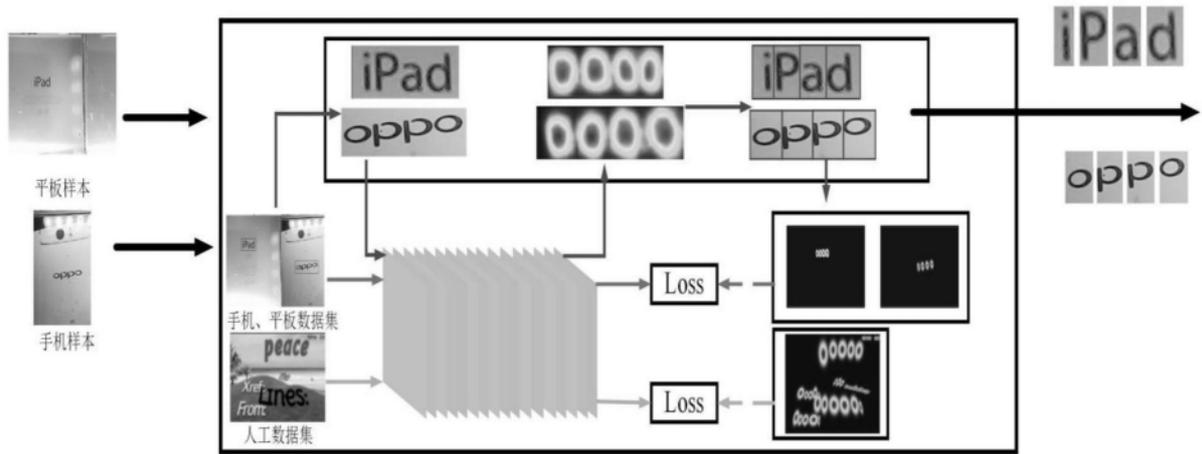


图2

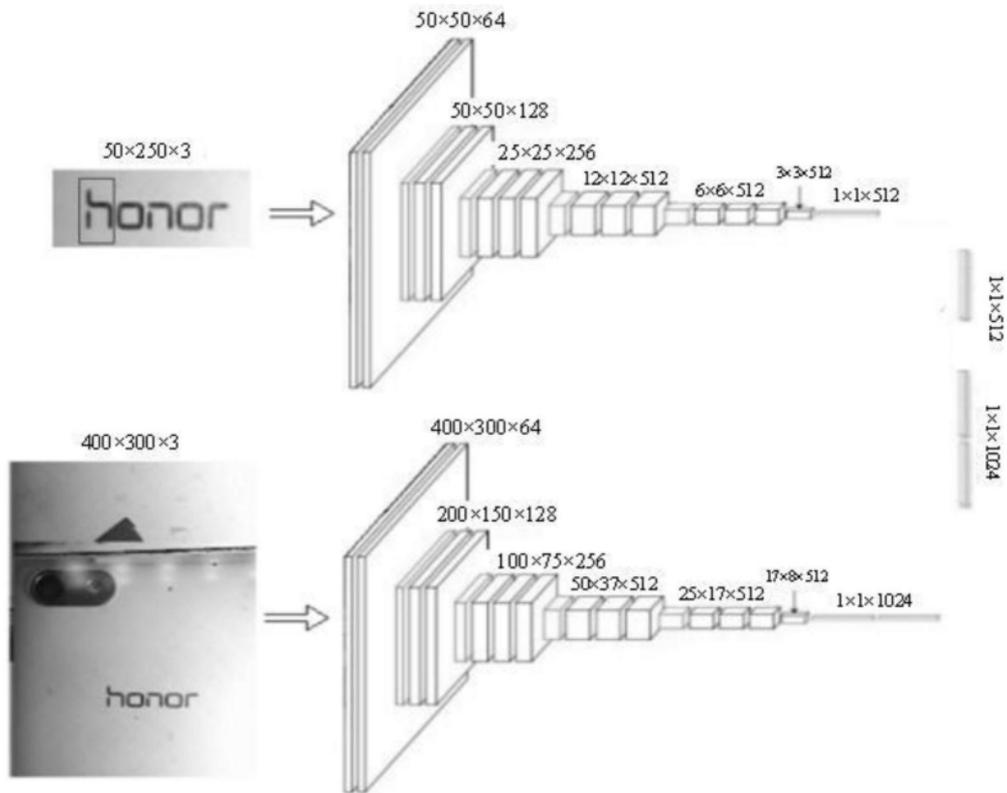


图3

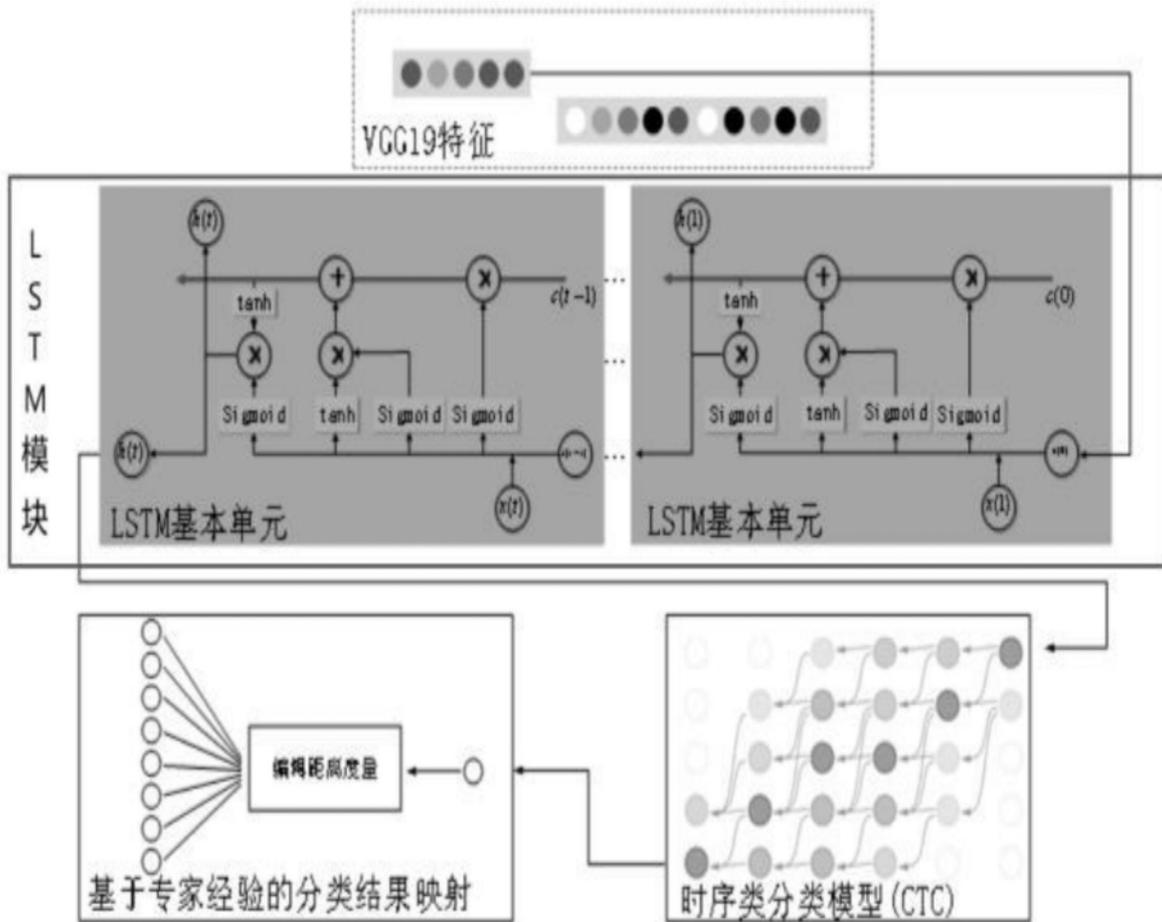


图4

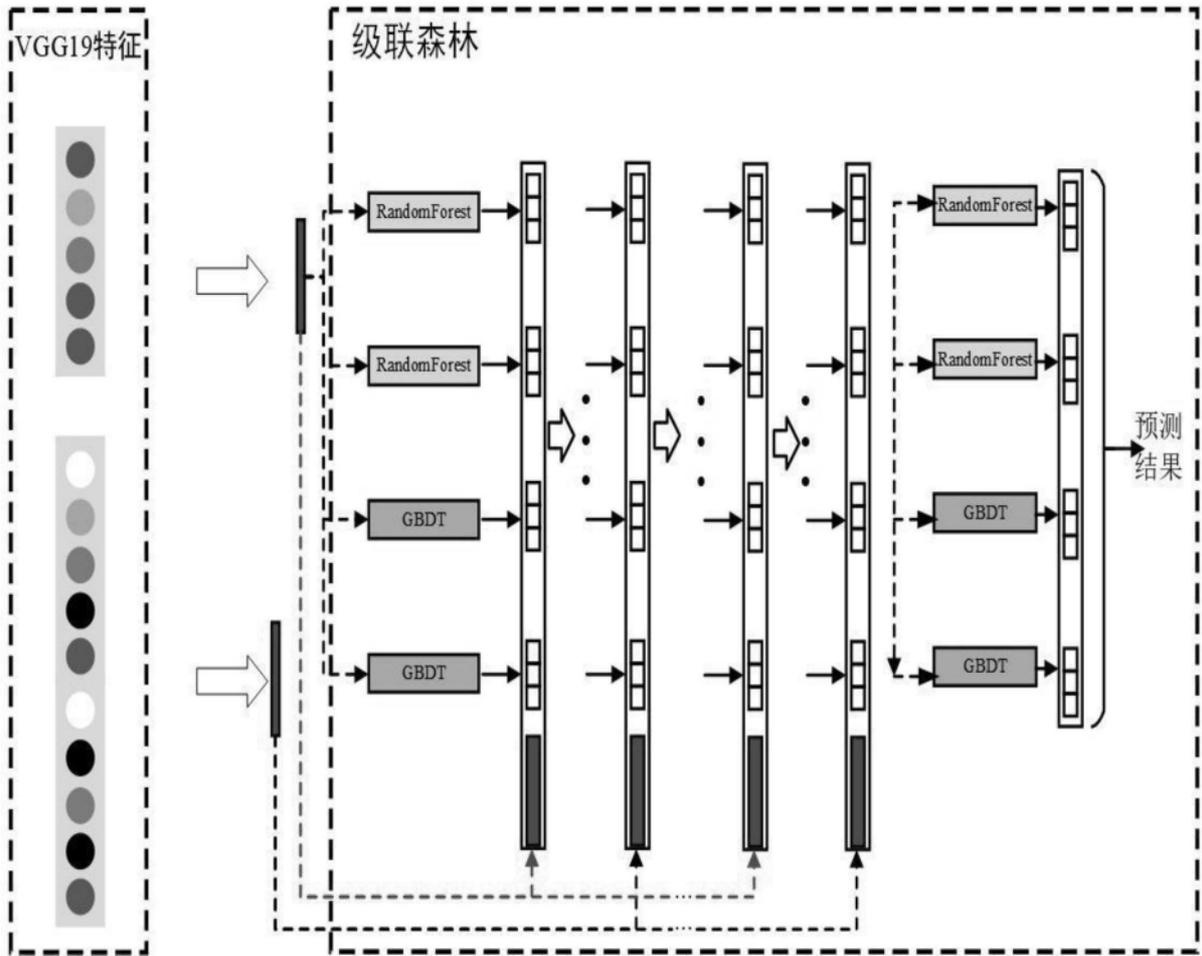


图5

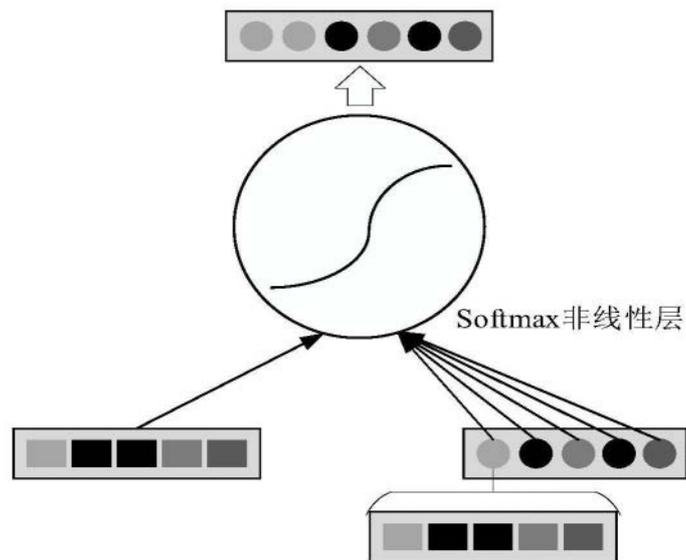


图6



图7

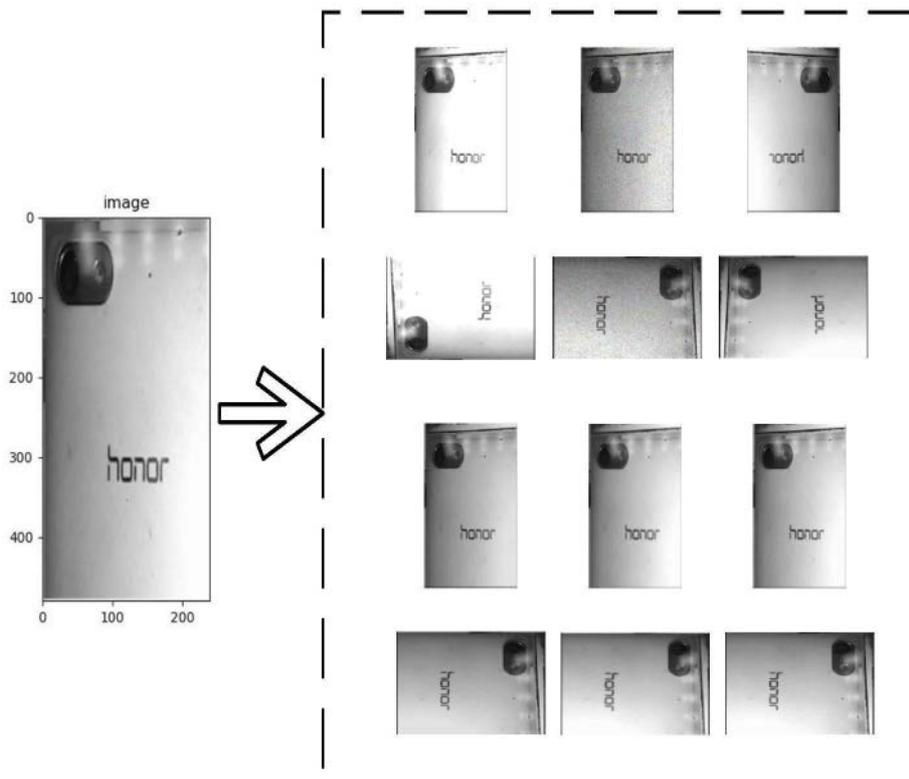


图8

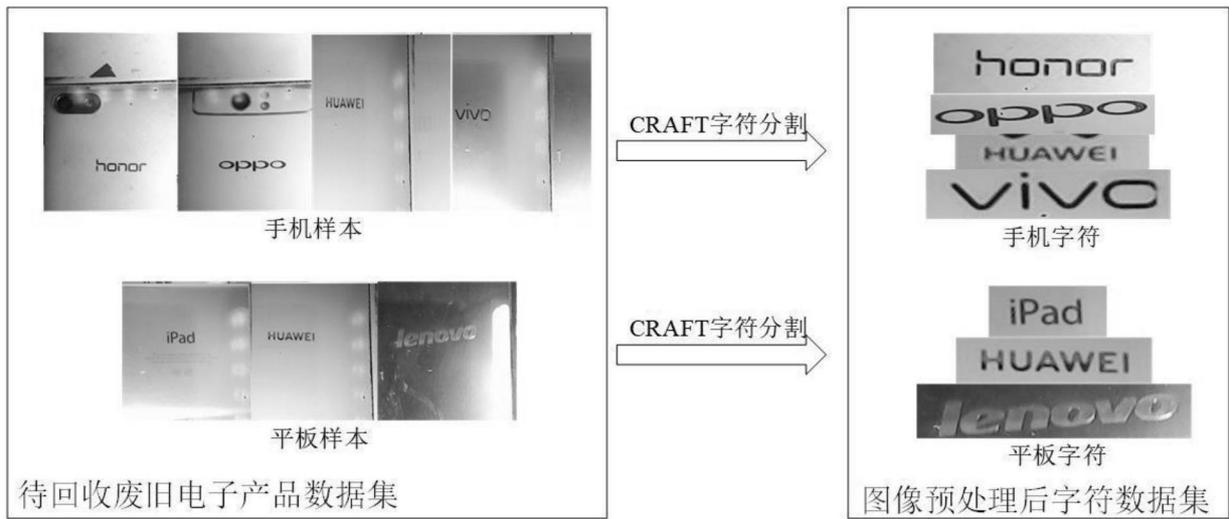


图9

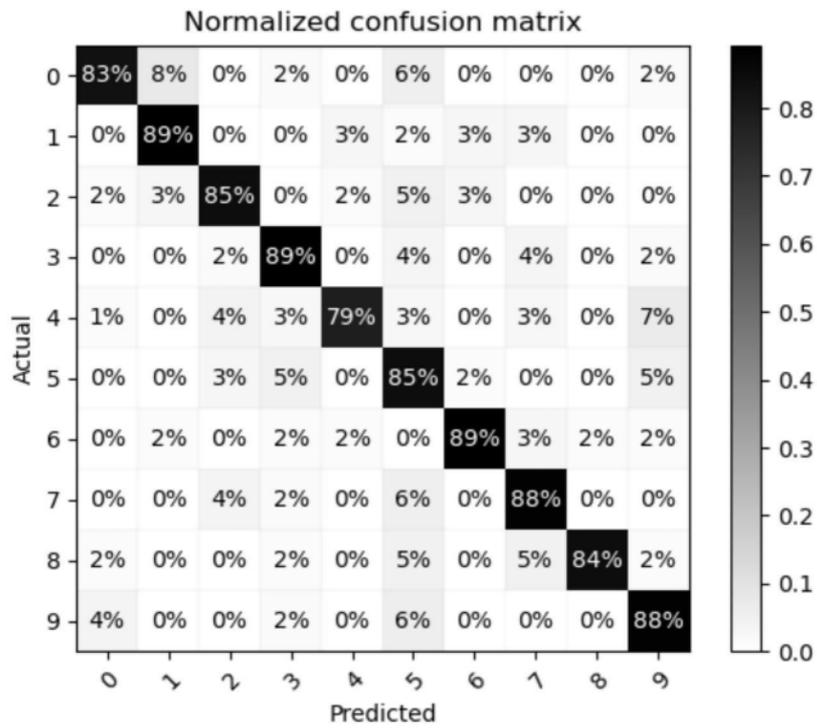


图10

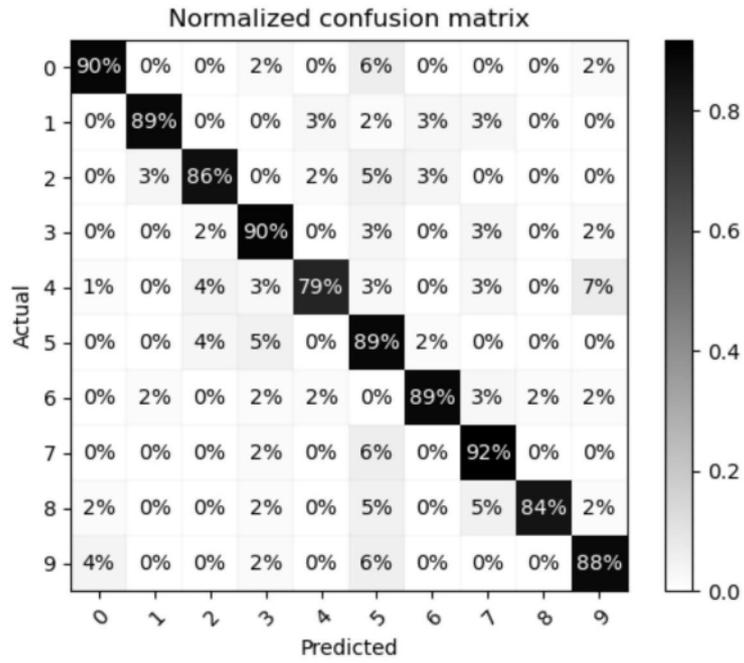


图11

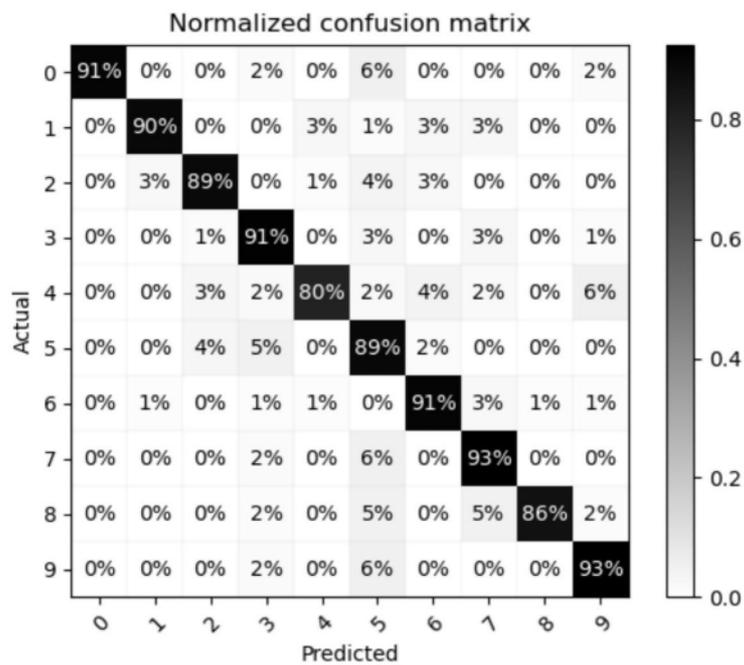


图12