

基于改进领域对抗网络的瓷砖表面缺陷检测^{*}

林 行 陈新度 吴 磊 练洋奇

(广东工业大学机电工程学院 广州 510006)

摘要: 深度神经网络作为主流的表面缺陷检测方法之一,需要大量样本进行模型训练,而随着瓷砖产品多样化,同类型瓷砖缺陷样本有限。本文提出一种基于改进域对抗神经网络(MDANN)的瓷砖表面缺陷检测方法,参考传统的DANN结构,首先,在ImageNet公共数据集上预训练保存网络参数,提高训练速度;然后,在原网络中加入瓶颈层,并利用最大均值差异指标优化领域分布差异,改善了原DANN网络筛选源域的能力,实现小样本瓷砖的缺陷检测。实验结果表明,MDANN对瓷砖表面缺陷的有效检出率达98.77%,相比于原DANN网络提高了3.53%,可快速适用于不同类型的瓷砖检测,泛化性好。

关键词: 瓷砖缺陷检测;深度学习;迁移学习;领域自适应神经网络(DANN)

中图分类号: TP2 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Ceramic tile surface defect detection based on improved
domain-adversarial neural network

Lin Xing Chen Xindu Wu Lei Lian Yangqi

(College of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Deep neural network is one of the mainstream surface defect detection methods, a large number of samples are needed for model training, but the defect samples of the same type of ceramic tile are limited with the diversification of ceramic tile products. In this paper, a ceramic tile surface defect detection method based on improved domain countermeasure neural network (MDANN) is proposed. Referring to the traditional DANN structure, the network parameters are pretrained on the ImageNet to improve the training speed. Then, the bottleneck layer is added to the original network, and the maximum mean difference index is used to optimize the field distribution difference, which improves the ability of the original DANN network to screen the source domain and realizes the defect detection of small sample tiles. The experimental results show that the effective detection rate of MDANN for ceramic tile surface defects achieves 98.77%, which is 3.53% higher than the original DANN network. It can be quickly applied to the detection of different types of ceramic tiles with good generalization.

Keywords: ceramic tile defect detection; deep learning; transfer learning; domain adaptive neural network (DANN)

0 引 言

瓷砖实际生产过程中,高效地检测出瓷砖缺陷并将其剔除可有效地提高瓷砖的出厂质量。针对该问题,许多学者致力于瓷砖表面缺陷的研究,Riza等^[1]提出一种纯色瓷砖缺陷检测方法,先综合运用滤波及形态学技术对缺陷进行分割,然后提取缺陷区域的GLCM纹理特征来进行分类,可以取得较高的准确度,但实验中仅针对纯色瓷砖,进行检测唐山学院李庆利等^[2]将阈值法和改进的Sobel算子结合,实现检测瓷砖的边直度、直角度、缺边和缺角缺陷,但

检测精度较低。由于瓷砖表面特征复杂,采用传统图像处理实现缺陷检测仍然缺乏高效性和准确性。

与传统机器视觉技术不同,深度学习利用大量数据训练来自动学习泛化性特征,适用于解决瓷砖表面的复杂缺陷检测。近年来,许多学者提出了基于深度神经网络和传统机器学习的瓷砖缺陷检测方法,(automation selector processing for ceramic tiles, ASPECT)是意大利的检测项目,此项目使用的卷积神经网络及模糊算法,取得了较好的检测效果,但需要利用大量标注瓷砖数据样本对网络模型进行训练^[3]。彭豪等^[4]提出了一种改进的Faster R-CNN

收稿日期:2022-05-25

^{*}基金项目:佛山市产业领域科技攻关项目(2020001006297)、佛山市顺德区核心技术攻关项目(2030218000174)资助

小目标检测模型实现对生产线瓷砖表面缺陷的检测,但实验中因样本数据不足网络性能受到一定限制。在实际生产过程中,瓷砖种类繁多,对于不同类型瓷砖样本的缺陷检测需要重新训练一个新的网络,大大增加了时间成本,而且同类型的瓷砖中具有表面缺陷的样本仍然较少,导致训练模型的数据不足。

为解决上述问题,迁移学习方法可通过利用具有大量标注信息的源域数据,构建具有泛化性更强的深度网络模型,实现小样本目标域的学习,更高效、准确地检测瓷砖表面缺陷。迁移学习近年来发展迅速,Yi 等^[5]提出了 TrAdaBoost 方法,提高了有利于目标分类任务的实例权重、降低不利于目标分类任务的实例权重,有利于提高模型迁移率。Pan 等^[6]提出了使用 MMD (maximum mean discrepancy) 最大均值差异距离计算源域和目标域的边缘分布差异,Wang 等^[7-9]提出了动态分布自适应方法 BDA,实现了动态地计算两种分布的差异性,提高了模型筛选源域的能力。He 等^[10]对计算机视觉领域的 ImageNet 预训练进行实验,结果表明大大缩短了训练时间,加快训练收敛速度。Ganin 等^[11-12]提出了领域对抗网络 (domain-adversarial neural network, DANN),使得特征提取器和域判别器互相对抗训练,实现源域到目标域的迁移学习,但对抗网络认为所有源域样本具有相同的重要性,无法高效筛选出有利于目标域训练的特征。基于以上研究,本文提出一种针对瓷砖表面缺陷检测的迁移学习方法,自建瓷砖表面缺陷数据集,基于迁移学习的思想,改进领域对抗网络,通过加入 MMD 度量提高网络筛选源域的能力,进一步改善网络迁移率,进而提高检测精度,在瓷砖表面缺陷检测。

1 瓷砖图像的迁移学习方法

1.1 迁移学习方法原理

迁移学习是深度学习领域中重要的研究方向,主要思想是通过学习源域的知识,利用源域和目标域的相似性,完成目标域的任务^[13]。领域是迁移学习中的重要概念,在迁移学习中,通过学习一个领域的分布,利用所学到的知识应用到另一个领域中,完成源域到目标域的知识迁移。通常定义 D 为一个领域,记领域上的数据 x 和标签 y ,其概率分布记为 $P(x, y)$,则数据服从 $(x, y) \sim P(x, y)$ 这一分布,记 X, Y 分别表示数据所处的特征空间和标签空间,即任意一个样本 (x_i, y_i) ,都有 $x_i \in X, y_i \in Y$,一个领域可表示为 $D = \{X, Y, P(x, y)\}$ 。源域和目标域是两个有相似数据分布但又有区别的领域,迁移学习的任务就是通过分析并利用两个领域的相似性实现的知识迁移。如图 1 所示,根据学习领域方法的不同,可分为基于样本的迁移学习方法、基于特征的迁移学习方法、基于模型的迁移学习方法和基于关系的迁移学习方法。其中最常见的方法是基于模型的迁移学习方法,本文研究基于模型的迁移学习方法中十分热门的领域对抗网络 (domain-ad-versarial

neural network, DANN) 方法,进行研究并改进其结构算法,提升迁移学习效率。

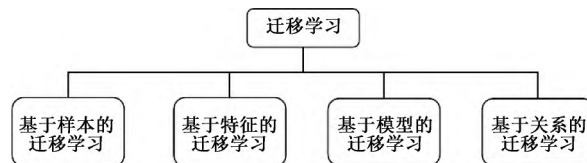


图 1 迁移学习方法分类

1.2 DANN 表面缺陷检测方法

领域对抗网络 (domain-adversarial neural network, DANN) 的网络结构如图 2 所示,由特征提取器 (feature extractor)、标签预测器 (label predictor) 和域分类器 (domain classifier) 三个部分组成,并且在特征提取器和域分类器之间有一个梯度反转层 (gradient reversal layer);其中特征提取器用于提取源域图像和目标域图像的特征信息,是 DANN 的主干网络,而标签预测器和域分类器将所提取的图像信息进行筛选,域分类器提取源域和目标域之间相同的特征信息用于迁移学习,而标签预测器筛选用于区分源域数据不同标签的特征。最终的参数优化决策为最小化标签预测器的损失函数,最大化域分类器的损失函数,使得目标域和源域分布接近,实现更好的迁移,最终优化得到的特征既有区分力,又具有领域不变性^[14-15]。但 DANN 网络认为源域中所有数据的重要性是一致的,在迁移学习过程中,需要在源域数据中学习有利于目标域任务的特征,并降低与目标域任务无关的数据的影响,而源域数据中往往存在不利于目标域任务的特征,这会导致瓷砖数据的负迁移。

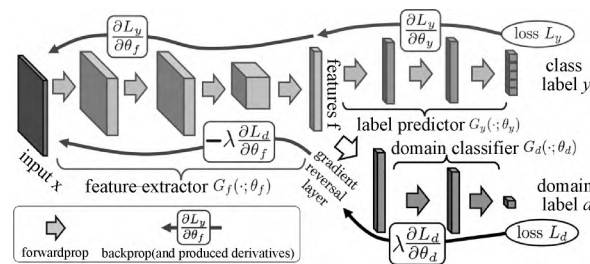


图 2 DANN 网络结构

1.3 改进 DANN 瓷砖表面缺陷检测方法

为解决上述问题,本文提出一种改进域对抗神经网络 (modified domain-adversarial neural network, MDANN),其结构如图 3 所示。由于瓷砖缺陷的特征复杂并且尺寸较小,为了提取更深层次的特征,本文采用 ResNet50 作为特征提取网络,ResNet 网络又称残差网络,网络的结构如图 4 所示,ResNet50 网络在深度卷积网络中加入 bottleneck (BTNK) 结构,使得网络在提取深层特征时不断拟合残差,解决了随着网络越深,越可能出现过拟合的问题。在标签预测器的设计上采用一个全连接层,用于将提取的特征直接分类。在域分类网络中加入 ReLU 激活函数,并且采用

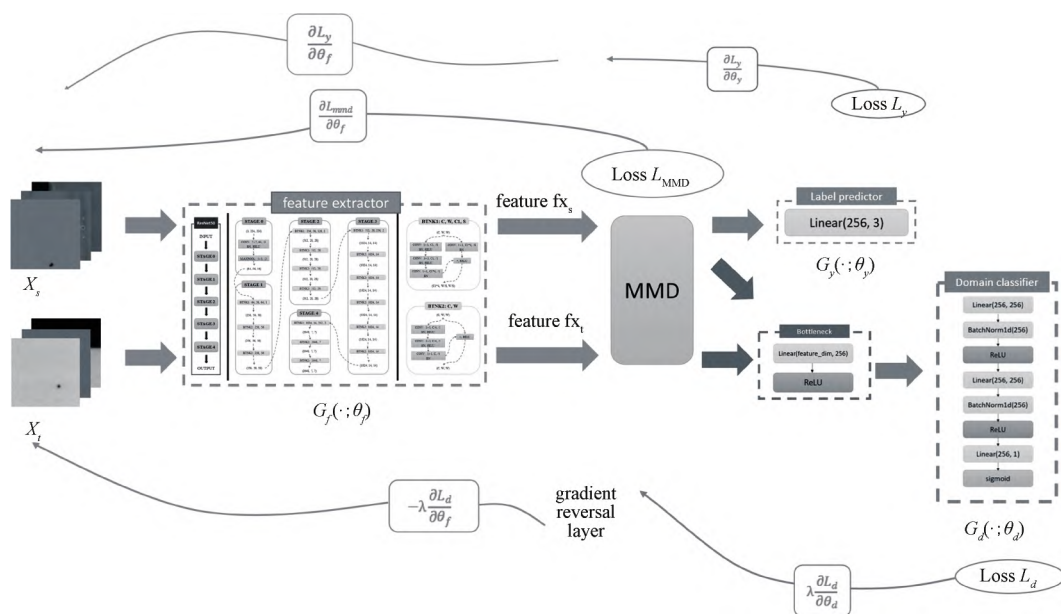


图 3 改进后 DANN 网络结构

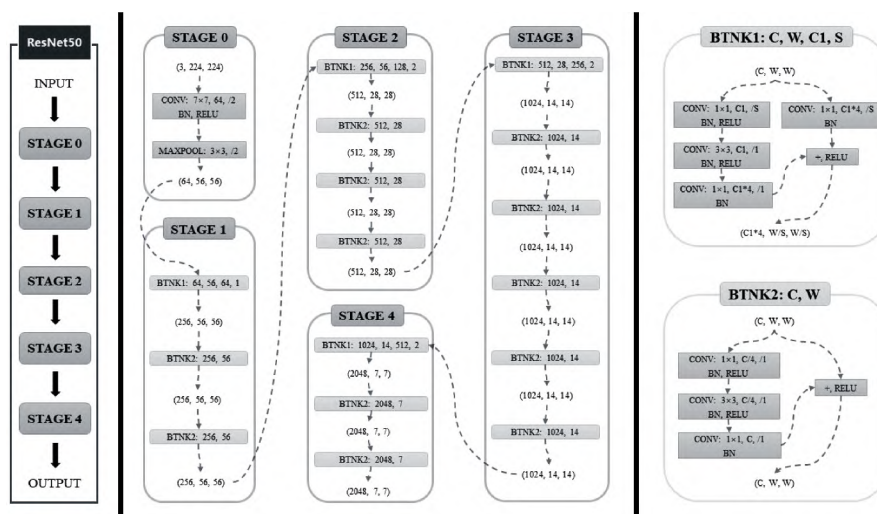


图 4 ResNet50 网络结构图

BN(batch normalization)与 Dropout 加快模型收敛防止过拟合。

本文在域分类器前加入了一个瓶颈层,瓶颈层的结构如图 5 所示,瓶颈层由一层全连接层和 ReLU 激活函数构成,可减少计算量,加快网络收敛。ImageNet 公共数据集是一个具有大量图像的分类数据集,利用 ImageNet 预训练的模型来进行迁移学习,可以增强网络的泛化性防止过拟合,并提高网络训练的速度,由此本文提出的 MDANN 网络模型在 ImageNet 公共数据集上进行预训练,并保存特征提取器的网络参数。

由于原 DANN 网络默认所有源域数据是同等重要的,而瓷砖类型繁多,纹理以及缺陷特征复杂,导致源域和目标域的特征分布不同,为了筛选出源域中有效的数据,

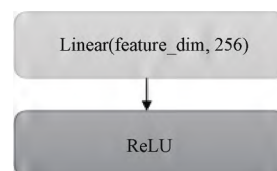


图 5 瓶颈层结构

本文加入了 MMD 指标度量源域和目标域的距离, MMD (maximum mean discrepancy) 最大均值差异是深度学习最常用的度量方法之一, 其原理是将两个领域的分布映射到可再生希尔伯特空间 (reproducing kernel hilbert space, RKHS) 中, 并求出两个数据分布的均值之差, 可表示为:

$$MMD(X, Y) = \left\| \frac{1}{n} \sum_{i=1}^n \phi(x_i) - \frac{1}{m} \sum_{j=1}^m \phi(x_j) \right\|_H^2 \quad (1)$$

其中, $\phi(\cdot)$ 是映射, H 表示用于把原变量映射到再生核希尔伯特空间。通过 MMD 来计算源域和目标域的损失, 反向传播优化网络模型参数, 使特征提取器在源域中提取出与目标域更相似的特征, 提高迁移学习率, 使目标域的检测精度提高。

MDANN 由以下 3 个部分构成:

- 1) 特征提取器 $G_f(\theta_f)$, 用于提取源域和目标域数据的特征。
- 2) 标签预测器 $G_y(\theta_y)$, 用于对特征按标签进行分类。
- 3) 领域判别器 $G_d(\theta_d)$, 用于判别特征来自源域还是目标域。
- 4) MMD 层用于度量源域和目标域的距离。

MDANN 首先通过最小化分类损失来优化分类器 G_y 的参数 θ_y , 对于特征提取器, 通过最小化 MMD 层损失和标签预测器损失来优化 G_f 的参数 θ_f :

$$(\hat{\theta}_f, \hat{\theta}_y) = \underset{\theta_f, \theta_y}{\operatorname{argmin}} E(\theta_f, \theta_y, \theta_d) \quad (2)$$

之后 DANN 通过最大化领域判别器 G_d 的损失来优化其参数 θ_d :

$$(\hat{\theta}_d) = \underset{\theta_d}{\operatorname{argmax}} E(\theta_f, \theta_y, \theta_d) \quad (3)$$

两个步骤交替进行, 直到网络收敛。

为了实现网络的对抗策略, 引入一种梯度反转层 (gradient reversal layer, GRL), 实现参数的反向传播, 当网络处于前向传播时, GRL 是一个恒等映射, 但反向传播时, 梯度乘以一个负单位, 将梯度反转, 从而实现领域判别器的损失最大化。上述过程可以总结出网络的目标函数为:

$$E(\theta_f, \theta_y, \theta_d) = \sum_{x_i \in D_s} L_y(G_y(G_f(x_i)), y_i) + \sum_{x_i \in D_s} L_{MMD}(G_f(x_i), G_f(x_j)) - \lambda \sum_{x_i \in D_s \cup D_t} L_d(G_d(G_f(x_i)), d_i) \quad (4)$$

其中, L_y 表示标签预测器损失, L_{MMD} 表示 MMD 度量的领域损失, L_d 表示域分类器损失, x_i 表示源域 D_s 的样本数据, x_j 表示目标域 D_t 的样本数据, y_i 为源域类别标签, d_i 为领域的标签, 源域数据标记为 $d_i = 0$, 目标域数据标记为 $d_i = 1$ 。

2 数据集制作

本文利用 TianChi_tile_round1 瓷砖表面缺陷数据集作为源域, 共 388 张瓷砖图像, 包括 27 种不同类型的瓷砖, 每个样本都带有缺陷标注。目标域为现场采集的尺寸为 600 mm × 600 mm 的瓷砖, 共得到 59 张瓷砖图像, 包括 3 种不同类型的瓷砖。图像采集系统如图 6 所示, 瓷砖在传送带上, 运输至光电传感器时触发信号命令工业线阵相

机采集, 采集到的图像储存在计算机中。由于在采集过程中图像受到光源系统和粉尘等环境因素的影响, 存在噪点以及失真, 故利用中值滤波对图像进行降噪。为了使网络模型精确定位学习缺陷特征, 将原图划分为若干小区域, 并保证 ROI 区域中带有缺陷特征, 再经过旋转、平移等数据增强的操作, 共得到 1 664 张图像, 其中源域 1 228 张、目标域 436 张。

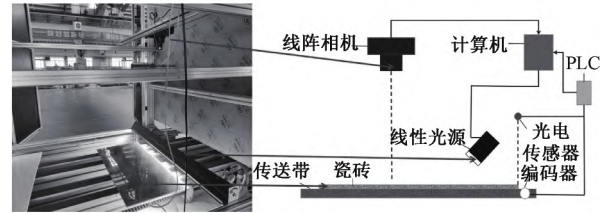


图 6 图像采集系统

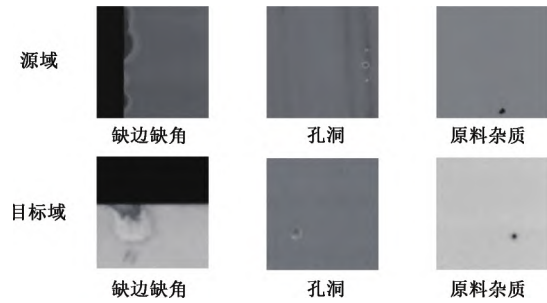


图 7 源域和目标域图像数据

3 实验与结果

3.1 实验环境

本文瓷砖表面缺陷检测模型运行环境 GPU 为 NVIDIA 的 GeForce GTX 2060Ti, 16GB 内存的服务器, 处理器为第 7 代英特尔; 软件环境为 Window10 系统, Python3.6, 深度学习框架采用 Pytorch1.6.0。源域数据包含孔洞、原料杂质、缺边缺角 3 种缺陷共 1 228 张, 目标域包含孔洞、原料杂质、缺边缺角 3 种缺陷共 436 张。源域数据中瓷砖类型更多, 且缺陷特征更复杂, 目标域数据样本量少, 且与源域数据为不同类型的瓷砖样本, 采用原 DANN 方法和 MDANN 方法进行对比, 训练至准确率收敛时中止迭代。

3.2 评价标准

实验的评价指标为目标域数据对三种缺陷分类的正确性, 采用精准率 (Precision) 和召回率 (Recall) 衡量。两者的定义如下:

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

$$F1-Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

其中, TP 表示正确被检出缺陷且分类正确的缺陷样本数目, FN 表示预测为正样本实际是负样本的数目, FP 表示未被正确分类的其他样本数目。 $F1-Measure$ 是同时考虑了精准度与召回率的综合评估指标。

3.3 MDANN 对抗网络实验结果

为了验证本文方法的有效性和先进性,将源域和目标域瓷砖图像进行切割,旋转,平移等操作,样本缺陷都包含孔洞,原料杂质和缺边缺角三种类型。将源域和目标域数据作为输入,对比原 DANN 网络和改进 DANN 网络的实验结果,并与预训练微调迁移学习方法进行准确率比较,预训练微调的原理是利用源域预训练 ResNet50 网络,再用目标域微调网络参数。

图 8 为不同迁移学习方法在目标域数据的分类准确率曲线图,可以看出改进后的 DANN 准确率相比于原网络模型和预训练微调方法显著提高,可达 98.86%。DANN 改进前后的混淆矩阵如图 9、10 所示,验证了改进 DANN 在检出差错表现上有明显提升。从上述结果分析,本文设计的 MMD 模块有效改善了原 DANN 网络筛选源域的能力,加强了模型对于领域相似性的学习,更大程度地利用了源域的知识,使得在目标域上分类表现提升;利用 ImageNet 预训练并在网络中加入瓶颈层,起到了加速模型收敛、防止模型过拟合的作用。整个模型检测性能良好,训练时长较短,具有较高泛化性,适用于不同类型瓷砖的检测。

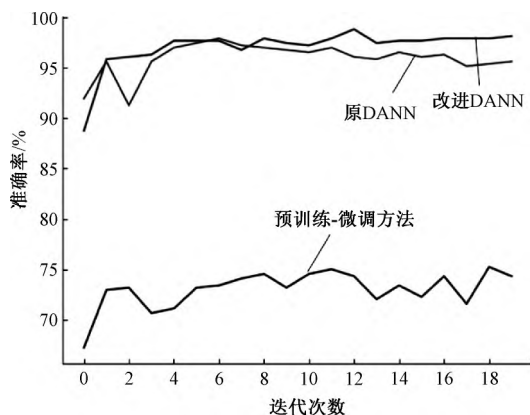


图 8 目标域数据准确率图

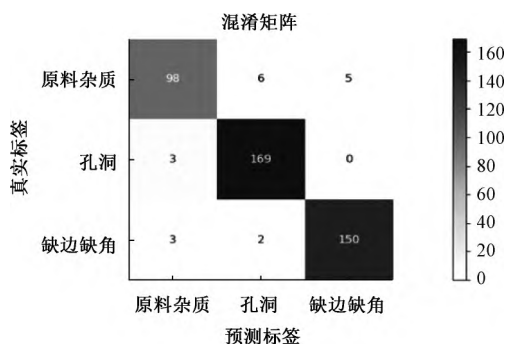


图 9 改进前目标域数据混淆矩阵图

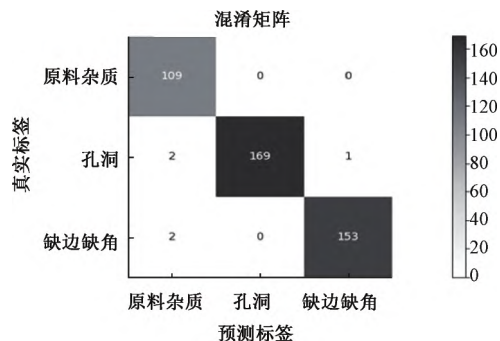


图 10 改进后目标域数据混淆矩阵图

表 1 实验结果

模型	类别	Precision	Recall	F1-Measure
MDANN	原料杂质	100	96.5	98.2
	孔洞	98.3	100	99.1
	缺边缺角	98.7	99.3	99.0
DANN	原料杂质	94.2	89.9	92
	孔洞	95.5	98.2	96.9
	缺边缺角	96.8	96.8	96.8

由表 1 可以看出,改进后的 DANN 网络模型在精准度、召回率上有显著提高,对目标域数据的缺陷识别效果有所提升。MDANN 网络加入 MMD 指标优化了特征提取器网络特征提取器的网络参数,使得特征提取器在源域中提取到与目标域数据分布更接近的特征,在小样本的目标域瓷砖缺陷检测任务中取得良好效果,有效解决了深度网络需要大量数据训练的问题,并且 MDANN 网络可通过迁移学习对不同类型的瓷砖进行缺陷检测,具有较高泛化性。

4 结 论

本文针对瓷砖类型的多样性及缺陷的复杂性,提出了一种基于 DANN 网络的瓷砖表面缺陷检测的迁移学习方法(MDANN),在训练模型时先在 ImageNet 上进行预训练防止过拟合,将源域和目标域数据作为输入 MDANN 中,源域数据为特征丰富、具有标注信息的大量样本,目标域为小样本,通过 MMD 层可以使网络从源域中提取到与目标域更相似的特征。最终实验结果证明,本文提出的方法在原料杂质,孔洞和缺边缺角缺陷上取得 98.77% 的平均 F1-Measure,相比于原 DANN 提高 3.53%。

参考文献

- [1] RIZA A, ADE D W, RAFIE. Defect detection of ceramic tiles using median filtering, morphological techniques, gray level co-occurrence matrix and k-nearest neighbor method [J]. Scientific Research Journal(SCIRJ), 2019, 7(4):41-45.
- [2] 李庆利,王永强,张帆. 陶瓷砖表面缺陷视觉检测算法

- 研究[J]. 中国陶瓷, 2015, 51(3): 44-47.
- [3] ASPECT: Performing a measure of the appearance-Visual Quality Control Made Easy. <http://www.cpr.it/tetrapc/aspect.html>.
- [4] 彭豪, 李晓明. 基于改进 Faster R-CNN 的小目标检测模型[J]. 电子测量技术, 2021, 44(24): 122-127.
- [5] YI Y, DORETTO G. Boosting for transfer learning with multiple sources[C]. The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, IEEE, 2010.
- [6] PAN S J, TSANG I W, KWOK J T, et al. Domain adaptation via transfer component analysis[J]. IEEE Transactions on Neural Networks, 2010, 22(2): 199-210.
- [7] WANG J, CHEN Y, HAO S, et al. Balanced distribution adaptation for transfer learning[C]. 2017 IEEE International Conference on Data Mining (ICDM), IEEE, 2017: 1129-1134.
- [8] WANG J, FENG W, CHEN Y, et al. Visual domain adaptation with manifold embedded distribution alignment [C]. Proceedings of the 26th ACM international conference on Multimedia, 2018: 402-410.
- [9] WANG J, FENG W, CHEN Y, et al. Transfer learning with dynamic distribution adaptation [J]. ACM Transactions on Intelligent Systems and Technology, 2020, 11(1).
- [10] HE K, GIRSHICK R, DOLLAR P. Rethinking ImageNet pre-training[C]. International Conference on Computer Vision, 2018.
- [11] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation [C]. International Conference on Machine Learning, PMLR, 2015: 1180-1189.
- [12] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. The Journal of Machine Learning Research, 2016, 17(1): 2096-2030.
- [13] PAN S J, QIANG Y. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [14] ZHU Y, ZHUANG F, WANG J, et al. Deep subdomain adaptation network for image classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(4): 1713-1722.
- [15] SHEN J, QU Y, ZHANG W, et al. Wasserstein distance guided representation learning for domain adaptation[C]. Thirty-Second AAAI Conference on Artificial Intelligence, 2018.

作者简介

林行, 硕士研究生, 主要研究方向为机器视觉及数字图像处理。

E-mail: 504431598@qq.com

陈新度(通信作者), 教授, 博士生导师, 主要研究方向为网络化制造、制造系统建模和仿真优化。

E-mail: 544394747@qq.com

吴磊, 讲师, 主要研究方向为机器人仿真、虚拟示教。

E-mail: 1650795492@qq.com

练洋奇, 硕士研究生, 主要研究方向为数字图像处理。

E-mail: 393474157@qq.com