表面缺陷检测数据集汇总及其相关项目推荐

# 参考

<https://mp.weixin.qq.com/s?__biz=MzI5MDUyMDIxNA==&mid=2247498632&idx=2&sn=a39387b9c15794393eb6e2c9561616cf>

<https://www.cnblogs.com/think90/articles/13215338.html>

# 正文

最近，有许多朋友都在关注缺陷检测领域，今天来看看缺陷检测。

目前, 基于机器视觉的表面 缺陷装备已经在各工业领域广泛替代人工肉眼检测，包括3C、汽车、家电、机械制造、半导体及电子、化工、医药、航空航天、轻工等行业。传统的基于机器 视觉的表面缺陷检测方法，往往采用常规图像处理 算法或人工设计特征加分类器方式。一般来说，通常利用被检表面或缺陷的不同性质进行成像方案的设计，合理的成像方案有助于获得光照均匀的图像，并将物体表面缺陷明显的体现出来。近年来，不少基于深度学习的缺陷检测方法也被广泛应用在各种工业场景中。  
对比计算机视觉中明确的分 类、检测和分割任务, 缺陷检测的需求非常笼统. 实 际上, 其需求可以划分为三个不同的层次: **“缺陷是什么”（分类）、“缺陷在哪里”（定位）和“缺陷是多少”（分割）。**

# 1 表面缺陷检测关键问题

**1、小样本问题**  
目前深度学习方法广泛应用在各种计算机视觉 任务中, 表面缺陷检测一般被看作是其在工业领域的具体应用。在传统的认识中, 深度学习方法无法直接应用在表面缺陷检测中的原因是因为在真实的 工业环境中, 所能提供的工业缺陷样本太少。  
相比于ImageNet数据集中1400多万张样本数据, 表面缺陷检测中面临的最关键的问题是小样本问题, 在很多真实的工业场景下甚至只有几张或几十张缺陷图片。 实际上, 针对于工业表面缺陷检测中关键问题之 一的小样本问题, 目前有4种不同的解决方式:

### 1）数据扩增、生成

最常用的缺陷图像扩 增方法是对原始缺陷样本采用镜像、旋转、平移、扭曲、滤波、对比度调整等多种图像处理操作来获取 更多的样本。另外一种较为 常见方法是数据合成, 常常将单独缺陷融合叠加到 正常(无缺陷)样本上构成缺陷样本。  
**2）网络预训练与迁移学习**  
一般来说,，采用小样本来训练深度学习网络很容易导致过拟合, 因此 基于预训练网络或迁移学习的方法是目前针对样 本中最常用的方法之一。

**3）合理的网络结构设计**  
通过设计合理的网络 结构也可以大大减少样本的需求。基于压缩采样定理来压缩和扩充小样本数据, 使 用CNN直接对压缩采样的数据特征进行分类. 相比 于原始的图像输入, 通过对输入进行压缩采样能大 大降低网络对样本的需求. 此外, 基于孪生网络的表 面缺陷检测方法也可以看作是一种特殊的网络设计, 能够大幅减少样本需求。  
**4）无监督或半监督**  
在无监督模型中, 只利用正 常样本进行训练, 因此不需要缺陷样本. 半监督方法 可以利用没有标注的样本来解决小样本情况下的网络训练难题。

### 2、实时性问题

基于深度学习的缺陷检测方法在工业应用中包括三个主要环节: 数据标注、模型训练与模型推断. 在实际工业应用中的实时性更关注模型推断这一环节. 目前大多数缺陷检测方法都集中在分类或 识别的准确性上, 而很少关注模型推断的效率。有不少方法用于加速模型, 例如模型权重量化和模型剪枝等。另外，虽然现有深度学习模型 使用GPU作为通用计算单元, 但随着技术发展, 相信FPGA会成为一个具有吸引力的替代方案。

# 2 工业表面缺陷检测常用数据集

**1）钢材表面：NEU-CLS（可用于分类、定位任务）**地址：http://faculty.neu.edu.cn/yunhyan/NEU\_surface\_defect\_database.html  
  
由东北大学（NEU）发布的表面缺陷数据库，收集了热轧钢带的六种典型表面缺陷，即轧制氧化皮（RS），斑块（Pa），开裂（Cr），点蚀表面（ PS），内含物（In）和划痕（Sc）。该数据库包括1,800个灰度图像：六种不同类型的典型表面缺陷，每一类缺陷包含300个样本。对于缺陷检测任务，数据集提供了注释，指示每个图像中缺陷的类别和位置。对于每个缺陷，黄色框是指示其位置的边框，绿色标签是类别分数。  
  
**2）太阳能板：elpv-dataset地址：**https://github.com/zae-bayern/elpv-dataset  
  
**3）金属表面：KolektorSDD地址：**http://www.vicos.si/Downloads/KolektorSDD  
  
**4）PCB板检测：DeepPCB地址：**https://github.com/tangsanli5201/DeepPCB  
  
**5）AITEX数据集（面料缺陷）**数据集下载链接：https://pan.baidu.com/s/1cfC4Ll5QlnwN5RTuSZ6b7w提取码：b9uy  
该数据库由七个不同织物结构的245张4096 x 256像素图像组成。数据库中有140个无缺陷图像，每种类型的织物20个，除此之外，有105幅纺织行业中常见的不同类型的织物缺陷（12种缺陷）图像。图像的大尺寸允许用户使用不同的窗口尺寸，从而增加了样本数量。Internet上的数据库还包含所有具有缺陷的图像的分割mask，使得白色像素表示缺陷区域，其余像素为黑色。

# ****6）天池布匹缺陷数据（竞赛）****

数据下载链接：https://pan.baidu.com/s/1LMbujxvr5iB3SwjFGYHspA提取码：gat2  
在布匹的实际生产过程中，由于各方面因素的影响，会产生污渍、破洞、毛粒等瑕疵，为保证产品质量，需要对布匹进行瑕疵检测。布匹疵点检验是纺织行业生产和质量管理的重要环节，目前人工检测易受主观因素影响，缺乏一致性；并且检测人员在强光下长时间工作对视力影响极大。由于布匹疵点种类繁多、形态变化多样、观察识别难道大，导致布匹疵点智能检测是困扰行业多年的技术瓶颈。本数据涵盖了纺织业中布匹的各类重要瑕疵，每张图片含一个或多种瑕疵。数据包括包括素色布和花色布两类，其中，素色布数据约8000张，用于初赛；花色布数据约12000张，用于复赛。  
7）**天池铝型材表面瑕疵数据集（竞赛）**数据集下载链接：https://tianchi.aliyun.com/competition/entrance/231682/information  
数据介绍：在铝型材的实际生产过程中，由于各方面因素的影响，铝型材表面会产生裂纹、起皮、划伤等瑕疵，这些瑕疵会严重影响铝型材的质量。为保证产品质量，需要人工进行肉眼目测。然而，铝型材的表面自身会含有纹路，与瑕疵的区分度不高。传统人工肉眼检查十分费力，不能及时准确的判断出表面瑕疵，质检的效率难以把控。近年来，深度学习在图像识别等领域取得了突飞猛进的成果。铝型材制造商迫切希望采用最新的AI技术来革新现有质检流程，自动完成质检任务，减少漏检发生率，提高产品的质量，使铝型材产品的生产管理者彻底摆脱了无法全面掌握产品表面质量的状态。大赛数据集里有1万份来自实际生产中有瑕疵的铝型材监测影像数据，每个影像包含一个或多种瑕疵。供机器学习的样图会明确标识影像中所包含的瑕疵类型。  
  
8）**弱监督学习下的工业光学检测（DAGM 2007）**数据下载链接：https://hci.iwr.uni-heidelberg.de/node/3616  
数据集介绍：

* 主要针对纹理背景上的杂项缺陷。
* 较弱监督的训练数据。
* 包含是个数据集，前六个为训练数据集，后四个为测试数据集。
* 每个数据集均包含以灰度8位PNG格式保存的1000个“无缺陷”图像和150个“有缺陷”图像。每个数据集由不同的纹理模型和缺陷模型生成。
* “无缺陷”图像显示的背景纹理没有缺陷，“无缺陷”图像的背景纹理上恰好有一个标记的缺陷。
* 所有数据集已随机分为大小相等的训练和测试子数据集。
* 弱标签以椭圆形表示，大致表示缺陷区域。

**9）基建表面裂纹检测数据**  
数据集是github上的一个项目，主要是一些基建（水泥表面裂纹），可以访问：https://github.com/cuilimeng/CrackForest-dataset或者链接：https://pan.baidu.com/s/1108j5QbDr7T3XQvDxAzVpg提取码：jajn  
更多数据集可参考：https://github.com/abin24/Surface-Inspection-defect-detection-dataset

# 3 项目推荐

**一、Defect Eye**  
**项目地址: https://github.com/sundyCoder/DEye**  
Defect Eye是一个基于tensorflow1.4的开源软件库，专注于表面缺陷检测。应用领域涵盖制造环境中的所有良率应用，包括进入来料加工工具认证，晶圆认证，玻璃表面认证，光罩鉴定，研究和开发以及工具，工艺和生产线监控。图案化和未图案化的晶圆缺陷检查和鉴定工具在晶圆的前表面，后表面和边缘上发现颗粒和图案缺陷，使工程师能够检测和监控关键的屈服偏移。此外，它可用于医学图像的注射，包括肺PET / CT，乳房MRI，CT Colongraphy，数字胸部X射线图像。  
  
使用CMake，VisualStudio 2015，CUDA8.0，cudnn6.0编译tensorflow-r1.4 GPU版本。

* tensorflow.dll，tensorflow.lib，libprotobuf.lib
* 百度云盘链接：https://pan.baidu.com/s/1o9tv1n8  密码：ekec
* Google云端硬盘链接：https://drive.google.com/open?id=1kANDNErMNLU9wNR3rKhUTz--ltwPNPUv

### **如何使用DEye**

* 安装VisualStudio Community 2015、安装NVIDIA CUDA 8.0
* git clone https://github.com/sundyCoder/DEye
* 下载tensorflow.dll，将其置于DEye / bin下
* 下载tensorflow.lib和libprotobuf.lib，在DEye / extra / tensorflow-r1.4 /下放置主题
* 下载inception\_v3\_2016\_08\_28\_frozen.pb，将其置于DEye / data下
* 打开Visual Studio解决方案“DEye.sln”，它应该在DEye / build / vc14下，解决方案配置选项选择“Release”，Soluton Platform选项选择“x64”。
* 构建并运行GUI项目，您可以为您的检查案例进行模型训练。

## ****应用****

### **IC芯片缺陷检查**

### **公路道路裂缝损坏检查**

### **编织物瑕疵点检查**

### **盖玻片检查**

### **民用基础设施缺陷检测**

### **电力线裂纹检测**

**二、缺陷检测文献记录**  
项目地址：https://github.com/Eatzhy/surface-defect-detection  
**三、项目地址：https://github.com/Wslsdx/Deep-Learning-Approach-for-Surface-Defect-Detection**  
参考：[1] 《Github项目推荐 | 深度学习缺陷检测》https://www.yanxishe.com/blogDetail/13979[2] 《基于深度学习的表面缺陷检测方法综述》陶显等，中科院自动化所[3] 《汇总|缺陷检测数据集》