

基于分割的深度学习的表面缺陷检测方法

Domen Tabernik - Samo Škvarc

- 达尼耶尔-斯科卡伊

摘要 使用机器学习的自动表面异常检测已经成为一个有趣和有前途的研究领域，对视觉检测的应用领域有非常大的直接影响。深度学习方法已经成为最适合这项任务的方法。它们允许检测系统学习检测表面异常，只需向其展示一些示范图像。本文介绍了一个基于分割的深度学习架构，该架构是为检测和分割表面异常现象而设计的，并在表面裂缝检测的特定领域进行了演示。该架构的设计使得模型可以使用少量的样本进行训练，这是实际应用的一个重要要求。所提出的模型与相关的深度学习方法，包括最先进的商业软件进行了比较，表明所提出的方法在具体的表面裂缝检测领域优于相关方法。大量的实验也阐明了注释的精度要求、所需的训练样本数量和所需的计算成本。实验是在一个新创建的基于真实世界质量控制案例的数据集上进行的，证明了所提出的方法能够在少量的缺陷表面上进行学习，只需要使用大约25-30个有缺陷的训练样本，而不是数百或数千个，这在深度学习应用中通常是这样。这使得该深度学习方法在工业领域很实用，因为在工业领域中，可用的样本数量很少。

有缺陷的样品是有限的。该数据集也被公开提供，以鼓励开发和评估表面缺陷检测的新方法。

关键词 表面缺陷检测 - 视觉检测 - 质量控制 - 深度学习 - 计算机视觉 - 分割网络 - 工业4.0

简介

在工业流程中，当涉及到确保成品的适当质量时，最重要的任务之一是检查产品的表面。通常情况下，表面质量控制是手工进行的，工人们接受培训来识别复杂的表面缺陷。然而，这种控制非常耗时，效率低下，并可能导致生产能力的严重限制。

在过去，经典的机器视觉方法足以解决这些问题（Paniagua等人，2010年；Bulnes等人，2016年）；然而，随着工业4.0范式的发展，趋势正朝着生产线的通用化发展，在这里需要快速适应新产品（Oztemel和Gursev 2018年）。经典的机器视觉方法无法确保这种灵活性。通常情况下，在经典的机器视觉方法中，特征必须是手工制作的，以适应特定的领域。然后使用手工制作的基于规则的方法或使用学习-----来做出决定。基于分类器，如SVM、决策树或kNN。由于这种分类器不如深度学习方法强大，手工制作的特征发挥了非常重要的作用。各种滤波器组、直方图、小波变换、形态学运算和其他技术被用来手工制作适当的特征。因此，手工制作的特征在经典方法中起着重要的作用，但這些特征并不适合不同的任务，而且在机器视觉方法必须要有一个漫长的开发周期时，这些特征就会被淘汰。

D.Tabernik - D. Skocaj
卢布尔雅那大学，计算机和信息科学学院 Vecna pot 113, 1000
Ljubljana, 斯洛文尼亚
电子邮件: domen.tabernik@fri.uni-lj.si, danijel.skocaj@fri.uni-lj.si

S.Šela1 - J. Skvarc2
IKolektor Group d. o. o, 2Kolektor Orodjarna d. o. o.
Vojkova 10, 5280 Idrija, 斯洛文尼亚

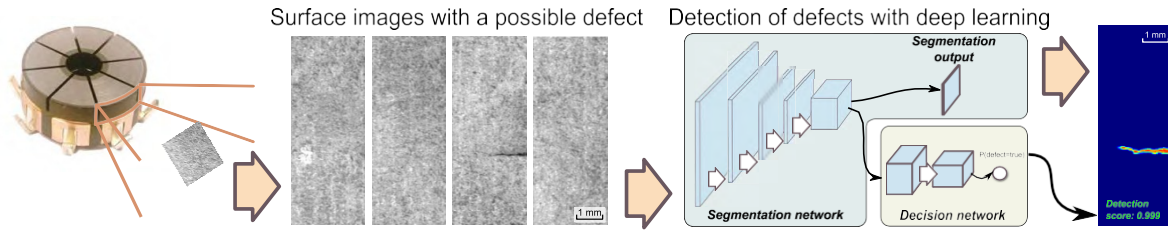


图1 拟议的表面缺陷检测方案

人工调整以适应不同的产品。在数据驱动的机器学习方法中可以找到提高灵活性的解决方案，所开发的方法可以快速适应新类型的产品和表面缺陷，只需使用适当数量的训练图像。

本文着重于使用最先进的机器学习方法来解决视觉表面缺陷的检测。重点主要放在深度学习方法上，近年来，这些方法已经成为计算机视觉领域最常见的应用。当应用于表面质量控制问题时（Chen and Ho 2016; Faghih-Roohi et al 2016; Weimer et al 2013; Kuo et al 2014），深度学习方法可以达到很好的效果，并且可以适应不同的产品。与经典的机器视觉方法相比，深度学习可以直接从低层次的数据中学习特征，并具有更高的能力来表达复杂的结构，从而完全取代了手工工程的特征的自动学习过程。随着对新产品的快速适应，这种方法变得非常适用于工业4.0所要求的柔性生产线。然而，开放的问题仍然存在：需要多少注释数据，注释需要多精确，才能达到适合实际应用的性能？在处理深度学习方法时，这是一个特别重要的问题

因为具有数百万可学习参数的深度模型往往需要成千上万的图像，而这在实践中往往很难获得。

本文探讨了适用于表面质量控制的深度学习方法。特别是，本文研究了应用于工业产品表面裂缝检测的深度学习方法（见图1）。探讨了合适的网络工作架构，不仅从它们的整体分类性能，而且从对工业4.0特别重要的三个特点的角度：（a）注释要求，（b）所需训练样本的数量和（c）计算要求。数据要求是通过利用基于两阶段结构的深度卷积网络的有效方法来解决的。提出了新的分割和决定网络，该网络适合于小范围内学习。

有缺陷的训练样本的数量，但仍能取得最先进的结果。

对所提方法的广泛评估是在一个新的、真实世界的数据集上进行的，称为Kolektor Sur-face-Defect Dataset（KolektorSDD）。该数据集代表了一个工业半成品表面缺陷检测的实际问题，其中可用于训练的缺陷项目的数量是有限的。通过强调三个重要方面，证明了所提出的方法已经适用于所研究的应用：（a）为达到100%的检测率所需的人工检查（通过对检测的额外人工验证），（b）所需的注释细节和导致所需人工成本的训练样本数量，以及（c）所需的计算成本。在所研究的领域，所设计的网络被证明优于相关的先进方法，包括最新的商业产品和两个标准分割网络。

本文的其余部分结构如下。在“相关工作”部分介绍了相关工作，在“建议的方法”部分介绍了分割和决策网的细节。在“分割和决策网络评估”一节中详细介绍了对拟议网络的广泛评估，并在“与技术现状的比较”一节中介绍了与最先进的商业解决方案的比较。本文最后在“讨论和结论”部分进行了讨论。

相关工作

在AlexNet（Krizhevsky等人，2012年）问世后不久，深度学习开始被更多地应用于表面缺陷分类问题。Masci等人（2012）的工作表明，对于表面缺陷分类，深度学习可以胜过传统的机器视觉方法，即手工设计的特征与支持向量机相结合。他们使用五层的卷积神经网络对几种钢铁缺陷类型的图像分类进行了论证。他们取得了很好的结果；然而，他们的工作仅限于一个浅层网络，因为他们没有使用ReLU和批量规范化。Faghih-Roohi等人（2016年）使用了类似的架构，用于检测铁路表面的

瑕疵。他们使用ReLU作为激活函数，并针对铁路缺陷分类的具体问题评估了几种网络规模。

在卷积网络的现代实现中，Chen和Ho（2016）应用OverFeat（Sermanet和Eigen 2014）网络来检测五种不同类型的表面错误。他们确定了大量的标记数据，作为深度网络的一个重要问题，并建议使用现有的预训练网络来缓解这一问题。他们利用了ILSVRC2013数据集集中的120万张普通视觉物体图像上训练的OverFeat网络，并将其作为有表面缺陷的图像的特征提取器。他们利用支持向量机在深度特征的基础上学习分类器，并表明预先训练的特征比LBP特征更出色。通过提出的近似表面粗糙度启发式，他们能够进一步改进这一结果；但是，他们的方法没有在目标域上学习网络，因此没有充分使用深度学习的潜力。

Weimer等人（2016年）评估了几种具有不同深度层的深度学习架构，用于表面异常检测。他们应用的网络范围从只有5层到有11层的网络。他们的评估集中在6种不同类型的合成错误上，结果显示，深度网络的表现优于任何经典方法，在合成数据集上的平均准确率为99.2%。他们的方法也能够将错误定位在几个像素的精度范围内；然而，他们的定位方法是低效的，因为它从每张图像中提取小的斑块并对每个单独的图像斑块进行分类。

Rac̃ki等人（2018）提出了一个更有效的网络来明确进行缺陷的分割。他们实现了一个有10层的完全卷积网络，使用ReLU和批量归一化来执行缺陷的分割。此外，他们在分割网络的特征之上提出了一个额外的决策网络，对缺陷的存在进行每幅图像的分类。这使他们能够提高合成表面缺陷数据集的分类精度。

最近，Lin等人（2018）提出了LEDNet架构，用于检测LED芯片图像上的缺陷，使用的数据集有3万张低分辨率图像。他们提出的网络遵循AlexNet架构，但重新移动了全连接层，而是采用了类激活图（CAMs），类似于（Zhou等人，2016）。这种设计使他们能够只使用每个图像的la-bels进行学习，并使用CAMs进行缺陷的定位。与传统方法相比，拟议的LEDNet在缺陷检测率方面有明显的改善。

与相关方法相比，本文提出的方法采用了两阶段的设计，即隔离网络和决策网络，类似于建筑-----。

Rac̃ki等人（2018）的结构。然而，所提出的方法包含了对分割和决策网络结构的一些改变，目的是为了增加感受野的大小，并提高网络捕获小细节的能力。与一些相关的工作（Rac̃ki等人，2018年；Weimer等人，2016年）不同的是，所提出的网络被应用于真实世界的例子，而不是使用合成的例子。本研究中使用的数据集也只包括少量的缺陷训练样本（即30个缺陷样本），而不是数百（Rac̃ki等人2018；Weimer等人2016）或数千（Lin等人2018）。这使得一些相关的架构，如LEDNet（Lin et al 2018），只使用每个图像的注释和大批量的大小，不适合当前的任务。由于少量的样本使得网络的选择更加重要，本文评估了用两种不同的标准网络设计重新放置分割网络的效果，通常用于语义分割，即用DeepLabv3+（Chen et al 2018）和U-Net（Ronneberger et al 2015）。通过使用DeepLabv3+网络也评估了使用预训练模型的影响，该网络在ImageNet（Russakovsky等人，2015）和MS COCO（Lin等人，2014）数据集的120多万张图像上进行了预训练。

建议的方法

表面异常检测的问题是作为一个二进制图像分类问题来解决的。这适用于表面质量控制，在这种情况下，准确的每幅图像分类往往比准确的缺陷定位更重要。然而，为了克服深度学习中样本数量少的问题，所提出的方法被制定为一个两阶段的去标志，如图2所描述。第一阶段实现一个分离网络，对表面缺陷进行像素级的定位。用像素级的损失来训练这个网络，可以有效地将每个像素视为一个单独的训练样本，从而增加有效的训练样本数量，防止过度拟合。第二阶段，进行二元图像分类，包括一个额外的网络，它建立在分割网络之上，并使用分割输出和分割网的特征。第一阶段的网络被称为*分割网络*，而第二阶段的网络被称为*决策网络*。

分割网络

拟议的网络由11个卷积层和3个最大集合层组成，每个卷积层将分辨率降低2倍。每个卷积层之后是

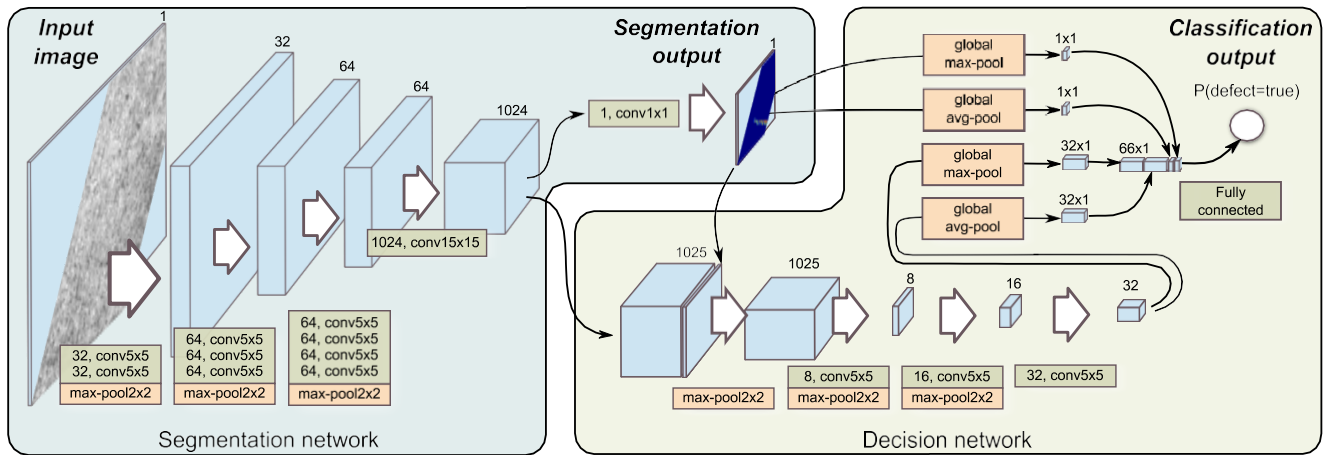


图2 所提出的带有分割和决策网络的架构

通过特征归一化和非线性ReLU层，它们都有助于提高学习期间的收敛率。特征归一化将每个通道归一为具有单位方差的零均值分布。前九个卷积层使用 5×5 核大小，而最后两层分别使用 15×15 和 1×1 核大小。不同的层分配了不同数量的通道，从图2的网络结构的详细描述中可以看出。最终的输出掩码是在应用 1×1 卷积层，减少输出通道的数量后得到的。这导致了一个单通道的输出图，其分辨率是输入图像的8倍。由于卷积层中的权重共享提供了足够的规则化，因此在这种方法中没有使用丢弃法。

拟议的分割网络的设计重点是在大分辨率图像中检测小的表面缺陷。为了实现这一目标，该网络的设计有两个重要要求：(a) 要求在高分辨率图像中具有较大的感受野尺寸；(b) 要求捕捉小的特征细节。与(Rac̃ki et al 2018)的相关工作相比，这导致了架构的几个重大变化。首先，一个额外的下采样层和较高层次的大内核尺寸被用来显著增加感受野的大小。其次，每个下采样之间的层数被改变为在架构的低段有较少的层，而在高段有较多的层。这就增加了具有大的receptive field尺寸的特征的容量。最后，下采样是通过使用最大池而不是大跨度的卷积来实现的。这确保了小而重要的细节能在下采样过程中存活下来，这在这个具有额外下采样层的网络工作中尤为重要。

决策网络

决策网络的结构使用分割网络的输出作为决策网络的输入。该网络将分割网络最后一层的输出（1024个通道）与一个单通道分割输出图相连接。这就产生了1025个通道量，作为其余各层的输入，其中有一个最大集合层和一个具有 5×5 核大小的卷积层。两层的组合重复3次，第一、第二和第三卷积层分别有8、16和32个通道。图2给出了该结构的详细描述。通道的数量被选择为随着特征的分辨率降低而增加，因此导致每层的计算要求相同。拟议的设计有效地使最后一个卷积层的分辨率比原始图像的分辨率小64倍。最后，该网络执行全局最大和平均池化，从而产生64个输出神经元。另外，全局最大和平均集合的结果在分割输出图上被串联为两个输出神经元，为分割图已经确保完美检测的情况提供了一个捷径。这种设计产生了66个输出神经元，这些神经元与线性权重相结合，成为最终的输出神经元。

决策网络的设计遵循两个重要的原则。首先，通过使用几层的卷积和下采样，确保了对大型复杂形状的适当容量。这使得该网络不仅能够捕捉到局部形状，而且还能捕捉到横跨图像大范围的全局形状。其次，决策网络不仅使用 1×1 核的通道重构前的分割网络的最后一次卷积运算的输出特征量，而且还使用 1×1 核的通道重构后得到的最终分割输出图。这就引入了一个捷径，网络可以利用

如果不需要的话，可以避免使用大量的特征图。它也减少了对大量参数的过度拟合。捷径在两个层面上实现：一个在决策网络的开始，分割输出图被送入决策网络的几个卷积层；另一个在决策网络的结束，分割输出图的全局平均值和最大值被附加到最后一个全连接层的输入。决策网络开始时的捷径和几个下采样的对话层是与Rac̃ki等人（2018）的相关工作的重要区别。与所提出的工作相比，他们只使用单层，在决策层中没有下采样，在卷积中不直接使用分割输出图，只通过全局最大和平均池间接使用。这限制了决策网络的复杂性，并预先阻止它捕捉大的全局形状。

学习

分割网络是作为一个二元分割问题来学习的；因此，分类是在单个图像像素水平上进行的。评估了两种不同的训练方法：(a)使用平均误差平方损失（MSE）的回归法和(b)使用交叉熵损失的二元分类法。这些模型没有在其他分类数据集上进行预训练，而是使用正态分布随机初始化的。

决策网络是用交叉熵损失函数训练的。学习是与分割网络分开进行的。首先，只有分割网络被独立训练，然后分割网络的权重被冻结，只有决策网络层被训练。通过只对决策层进行微调，网络工作避免了分割网络中大量权重带来的过拟合问题。这在学习决策层的阶段比学习分割层的阶段更重要。由于GPU内存的限制，在学习决策层时，每批样本的大小只能是一个或两个，但在学习分割层时，图像的每个像素都被视为一个单独的训练样本，因此有效的批处理量增加了数倍。

我们也考虑了同时学习分割网络和决策网络的问题。在这种情况下，损失函数的类型发挥了重要作用。只有当交叉熵被用于两个网络时，同步学习才是可能的。由于损失被应用于不同的范围，即一个在每个像素水平，一个在每个图像水平，两个层的准确归一化起到了关键作用。最后，正确地规范化这两个网络

事实证明，与使用单独的学习机制相比，损失不仅更难在实践中实现，而且也没有带来任何性能的提高。因此，两阶段学习机制被证明是一个更好的选择，并在随后的所有实验中被采用。

推论

拟议网络的输入是一张灰度图像。网络结构与输入大小无关，与全卷积网络类似（Long等人，2015），因为全连接层不用于特征图，而只是在用全局平均和最大集合消除空间维度后才使用。因此，输入图像可以是高分辨率或低分辨率的，这取决于问题。本文对两种图像分辨率进行了探讨：1408×512和704×256。

拟议的网络模型返回两个输出。第一个输出是分割掩码，作为分割网络的输出。分割掩码输出8×8组输入像素的缺陷概率；因此，相对于输入分辨率，输出分辨率降低了8倍。由于高分辨率图像中8×8像素块的分类对当前的问题来说已经足够了，所以输出图没有回到原始图像尺寸。第二个输出是[0, 1]范围内的概率得分，代表决策网络返回的图像中存在异常的概率。

分割和决策网络评价

所提出的网络在一个工业产品的表面裂纹检测中得到了广泛的评估。本节首先介绍了数据集的细节，然后介绍了评估的尾声及其结果。

Kolektor的表面缺陷数据集

由于缺乏公开可用的带有注释的表面缺陷真实图像的数据集，我们创建了一个新的数据集，称为Kolektor表面缺陷数据集（KolektorSDD）。¹该数据集是由Kolektor Group d. o. o.提供并注释的有缺陷的电气元件（见图1）的图像构建的。具体来说，在电气换向器的塑料垫层表面观察到了微小的裂纹或裂缝。每个换向器的表面区域被捕捉到8张不重叠的图像。

¹ Kolektor的表面缺陷数据集可在
<http://www.vicos.si/Downloads/KolektorSDD> 上公开获取。

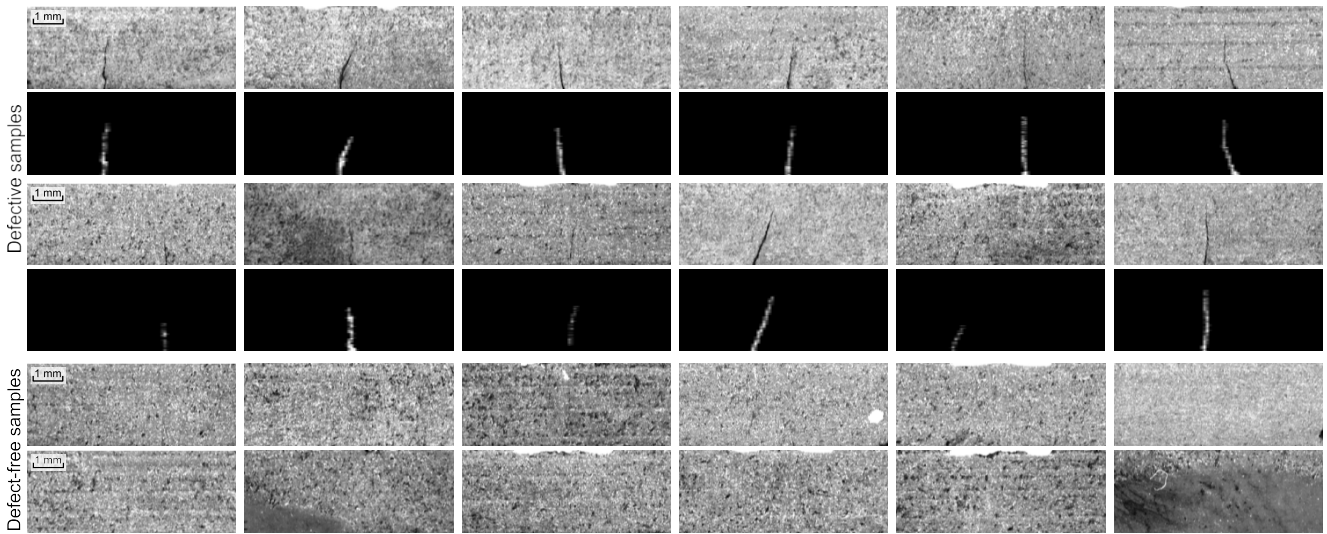


图3 几个表面图像的例子，上面是可见缺陷及其注释掩码，下面是无缺陷的表面

这些图像是在受控的环境中捕获的，因此获得了分辨率为 1408×512 像素的高质量图像。该数据集由50个有缺陷的电气元件组成，每个元件有多达8个相关表面。这导致了总共399张图像。在两个项目中，缺陷在两幅图像中可见，而对于其余项目，缺陷只在一幅图像中可见，这意味着有52幅图像的缺陷是可见的（即，缺陷或阳性样本）。对于每张图像，都提供了详细的像素级的符号掩码。剩下的347张图像作为具有非缺陷表面的反面例子。图3描述了这些有可见缺陷和无缺陷的图像的例子。

此外，该数据集被注释了几种不同类型的注释。这使得我们可以在不同的注释准确度下对所提出的方法进行评估。注释的准确性在工业试验环境中特别重要，因为它相当耗时，在注释上花费的人力应该最小化。为此，通过使用不同的核大小（即5、9、13和17像素），用形态学操作对原始注释进行扩张，产生了另外四种注释类型。请注意，这适用于原始分辨率的图像，而在分辨率为一半的实验中，注释掩码在被扩张后被缩小。所有的注释，一个

图4中描述了手动（a）和四个生成的（b-e）。

实验

首先几个不同的训练设置下对所提出的网络进行了评估，其中包括不同类型的注释、输入数据旋转和分割网络的不同损失函数。总的来说，该网络在四个配置组下进行了评估：

- 五种注解类型、
- 分割网络的两种损失函数类型（平均平方误差和交叉熵）、
- 两种尺寸的输入图像（全尺寸和半尺寸）、
- 没有和有 90° 输入图像旋转。

每个配置组都可以从四个方面评估网络的性能。不同的符号类型可以评估符号精度的影响，而不同的图像分辨率可以评估在较低计算成本下对分类性能的影响。此外，还评估了不同损失函数的影响以及通过以0.5的概率旋转图像来增加训练数据的影响。



图4 用不同的形态学核大小对(a)所示的原始注释进行扩张，产生了五种不同的注释类型：(b) 扩张=5，(c) 扩张=9，(d) 扩张=13和(e) 扩张=15

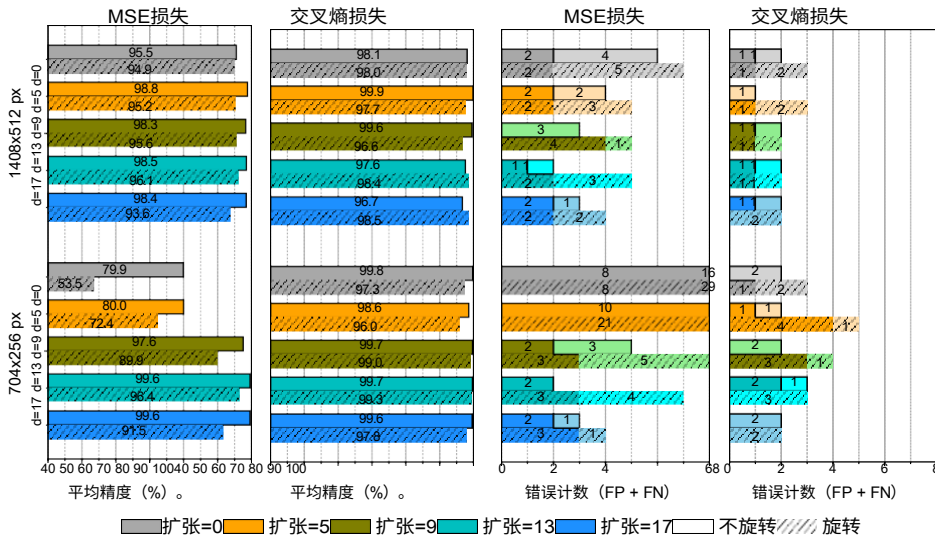


图5 建议的结果
关于KolektorSDD的方法
(假阳性(FP)用 深色显示, 假
阴性(FN)用浅色显示)

为了评估的目的，表面缺陷检测的问题被转化为二进制图像分类问题。主要目标是将图像分为两类：(a) 存在缺陷和(b) 不存在缺陷。尽管可以从分割网络中获得缺陷的像素级分割，但评估并不测量像素级的误差，因为它在工业环境中并不重要。相反，只测量每个图像的二元图像分类误差。分割结果仅用于可视化目的。

业绩指标

评估是通过3倍交叉验证进行的，同时确保同一实物产品的所有图像都在同一文件夹中，因此不会同时出现在训练集和测试集中。考虑到三种不同的分类方法，对所有被评估的网络进行了比较：(a) 平均精度 (AP)，(b) 假阴性的数量 (FN) 和 (c) 假阳性的数量 (FP)。注意，阳性样本是指有可见缺陷的图像，而阴性样本是指没有可见缺陷的图像。评价中使用的主要指标是平均精度。这比FP或FN更合适，因为平均精度被计算为精度-召回曲线下的面积，并在一个数值中准确地捕捉到不同阈值下模型的性能。另一方面，错误分类的数量 (FP和FN) 取决于应用于分类分数的具体阈值。我们报告的是在达到最佳F值的阈值下的错误分类数量。另外，请注意，我们选择AP而不是ROC曲线下的面积 (AUC)，因为AP更准确地涵盖了

与AUC相比，在具有大量阴性（即非缺陷）样本的数据集中，AUC能反映出性能。

实施和学习细节

网络结构在Tensor-Flow框架中实现（Abadi等人，2015年），两个网络都是使用随机梯度下降的方式进行训练的，没有动量。平均平方误差（MSE）的学习率为0.005，交叉熵损失为0.1。每次迭代只使用一个单一的图像，即批量大小被设置为一个，这主要是由于大的图像尺寸和GPU的内存限制。

在学习过程中，训练样本是随机选择的；然而，选择过程被修改，以确保网络观察到数量均衡的有缺陷和无缺陷的图像。这是通过在每次偶数迭代中选择有缺陷的图像，在每次奇数迭代中选择无缺陷的图像来实现的。这个机制确保了系统以恒定的速度观察到有缺陷的图像；否则，学习就会不平衡地偏向于非缺陷的样本，并且由于数据集中有更多的非缺陷图像，学习的速度会明显变慢。应该注意的是，这导致了训练不是完全由epochs完成的，因为非缺陷图像的数量是缺陷图像数量的8倍，网络在收到所有非缺陷图像之前会收到相同的缺陷图像。

两个网络都被训练了6600步。每个训练集有33张有缺陷的图像，每一步在有缺陷和无缺陷的图像之间交替进行，这就意味着有100个历时。只有当所有的缺陷图像至少被观察到时，一个历时才被认为是结束。

一次，但不一定能观察到所有无缺陷的图像。

分割和决策网络

拟议的网络由以下两个部分组成

首先对第一阶段中的离子网络 and 第二阶段中的决策网络进行了评估。详细结果在图5中预发。该图以不同的颜色显示了不同注解类型的实验结果，并以不同的颜色显示了实验的结果。

虚线表示使用图像旋转的实验。全图像分辨率的实验报告在上组，一半分辨率的实验报告在下组。实验结果

在使用 5×5 的核大小 ($dilate=5$)、交叉熵损失函数、全图像分辨率和没有任何图像旋转的情况下，获得了最佳性能的注释结果。在这种配置下，网络的平均精度 (AP) 达到99.9%，有零个假阳性 (FP) 和一个假阴性 (FN)。

接下来，个别学习设置的影响可以通过观察设置的每个具体变化对表现的平均改善来评估。以下是对性能的影响的报告

设置：(a)改变交叉熵损失函数为从平均平方误差损失来看，分割网络、

(b)从全图像分辨率改为较小的图像分辨率，以及(c)输入数据旋转 90° ，不旋转。图6报告了所有实验中AP的平均改进情况。对于一个特定的设置变化的结果，例如，从全图像分辨率变为一半的图像分辨率，首先计算所有设置的所有可能配置的AP（在图5中报告），然后计算两个实验之间的AP差异，其中只有相关设置被改变，例如，使用一半图像分辨率的实验和使用全图像分辨率的实验之间，但所有其他设置相同。性能的整体改善是通过所有其他设置保持不变的AP的变化的平均值来体现的。标准偏差也分别报告了正向和反向的变化。

损失函数 当比较图5中的平均误差损失 (MSE) 和交叉熵损失函数时，很明显，使用交叉熵损失函数训练的网络获得了最佳性能。这反映在AP指标和FP/FN计数中，也反映在图6中对所有其他设置的交叉熵的平均改进上。平均而言，交叉熵实现了7个百分点 (pp) 的更好的AP。

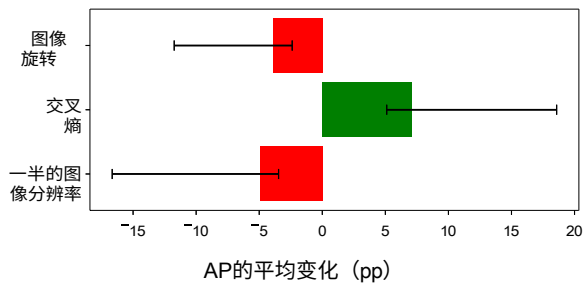


图6 学习配置的不同变化对AP的平均变化（平均精度）的影响

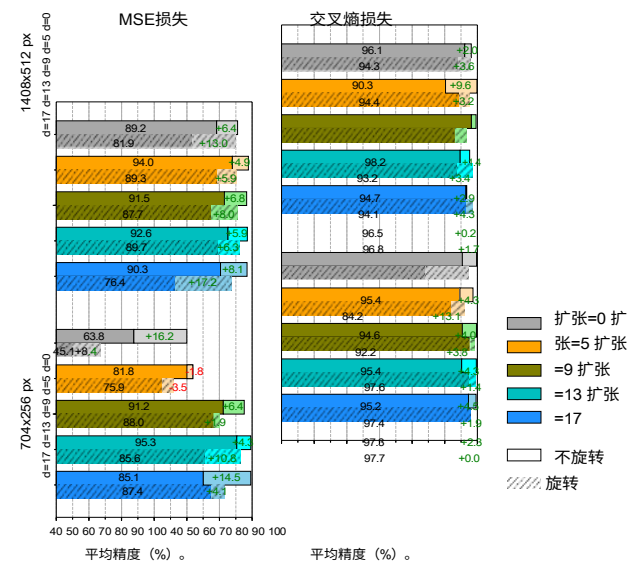


图7 由决策网络贡献的平均精度的提高。

图像分辨率 如图6所示，图像分辨率降低的网络的AP平均表现差5个百分点。对图5的仔细检查表明，较小的图像主要对用MSE损失函数训练的网络产生负面影响，而用交叉熵训练的网络则不受影响。交叉熵对图像分辨率的降低不太敏感，在某些情况下，分辨率降低的图像表现略好（大约AP的百分之一）。

图像旋转 另一方面，随机旋转图像并没有被证明是有用的，也没有导致任何明显的性能提升。在某些情况下，收益最多只有一个百分点；然而，在其他情况下，性能下降的幅度更大。

注释类型 最后，在图5中比较不同的注释类型，结果是在用较小的注释（原始或小核的扩张）训练时，以及考虑交叉熵损失时，对性能只有轻微的负面影响。差异更明显的是

在MSE损失函数中。总的来说，最好的结果似乎是实现了以中到大的稀释率稀释的注释。

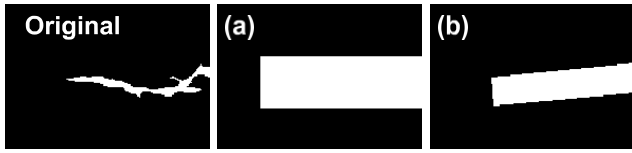


图8 两个额外的注释：（a）大的和（b）粗的

决策网络的贡献

决策网络对最终每一个人的贡献。

同时也评估了其性能。这一贡献是通过比较上一节的结果和没有决策网络的分割网络来衡量的。用一个简单的二维描述器和逻辑回归来代替决策网络。从分割输出图的全局最大值和平均集合值中创建了一个二维描述符，然后将其作为逻辑回归的特征，在网络训练完成后，将其与分割网络分开学习。

结果显示在图7中。当关注具有交叉熵损失的模型时，很明显，只有分割网络的模型已经取得了相当好的结果。由 $dilate=9$ 注释得到的最佳配置实现了98.2%的平均精度（AP），零假阳性（FP）和四个假阴性（FN）。然而，在大多数实验中，决策网络改善了这一结果。决策网络对MSE损失的贡献更大。当只使用分割网络时，MSE损失函数的平均精度达到了90%以下，而使用决策网络时，MSE损失的AP在95%以上。对于用交叉熵训练的网络来说，决策网络也有助于提高性能，但由于分割网络已经表现良好，所以提高的幅度略小，决策网络的AP提高了3.6个百分点，平均超过了98%。同样的趋势也出现在理想阈值下的错误分类数量上，当决策网络包括在内时，分割网络的平均4个错误分类被重新缩减为平均2个错误分类。

这些结果指出了决定网络的重要作用。简单的每像素输出分割似乎没有足够的信息来预测图像中缺陷的存在，而决定网络也能做到这一点。另一方面，拟议的决策网络能够从最后一个分割层的丰富特征中获取信息，并且通过额外的决策层，它能够从正确的特征中分离出噪声。决策网络中额外的下采样也有助于提高性能，因为这增加了感受野的大小，并使决策网络能够实现。

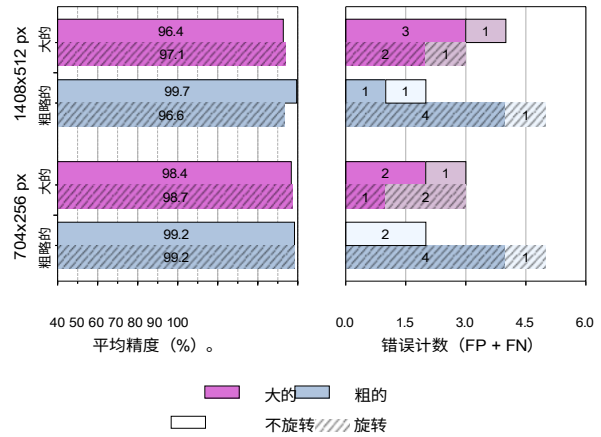


图9 有大的和粗的注释的结果

网络来捕捉缺陷的整体形状。全局形状对分类很重要，但对像素级的分割并不重要。

注释的必要精度

上一节的实验已经表明，大型注释的性能比精细注释的性能更好。本节将通过评估更粗的注释对分类性能的影响来进一步探讨这个问题。为此，我们创建了两组额外的注释类型，称为：（a）带有边界框的大注释和（b）带有旋转边界框的粗略注释。这两种注释都显示在图8中。这种类型的注释对人类注释者来说耗时较少，在工业环境中会更好。

结果显示在图9中。本实验只使用了具有交叉熵损失的模型，因为在以前的实验中证明MSE损失的能力较差。实验显示大注释的表现几乎与细注释的表现一样好。表示为大的注释表现稍差，最佳AP为98.7%，有3个错误分类，而粗略的注释达到了99.7%的AP和2个错误分类。请注意，在较小的图像分辨率下，两种注释都取得了类似的AP，而错误分类的数量相同。

这些结果与上一节中使用更精细的注释所获得的结果相当，在上一节中，只有一个错误分类的情况下，AP达到了99.9%。更精细的注释确实取得了稍好的结果；但是，考虑到这种级别的细节注释是很耗时的，使用粗略的注释仍然是可行的，而且性能损失很小或没有。

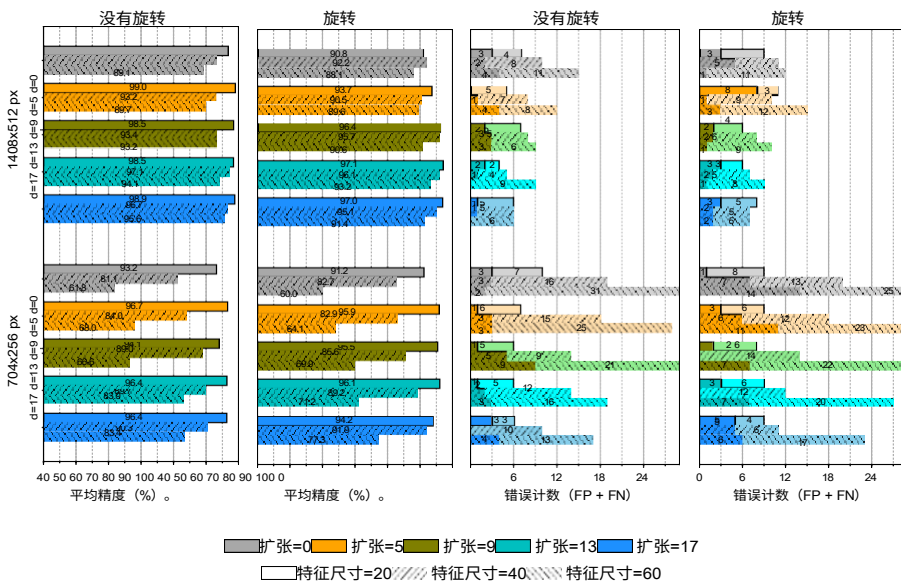


图10 对商业的评价
软件康耐视ViDi套件在Kolek- torSDD上的应用

与技术状况的比较

几个最先进的模型被进一步评估，以评估所提出的方法在相关工作文本中的表现。本节首先展示了一个最先进的商业产品和两个标准分割网络在不同训练配置下的性能。这为每种最先进的方法提供了最佳的训练配置，并允许与所提出的网络结构进行公平的比较，这在本节的最后进行。

商业软件

首先对基于深度学习的最先进的工业图像分析商业软件康耐视ViDi套件（康耐视2018）进行了评估。Vidi公司于2012年从CSEM（一个私人的、非盈利的、瑞士的研究和技术组织）中产生，并于2017年被康耐视收购。该软件包有三种不同的深度学习工具：ViDi蓝色（固定、定位），ViDi红色（隔离和异常检测），ViDi绿色（物体和场景分类）。

Vidi Red是一款用于异常检测、美学视觉检查和分割的工具。该工具可以在无监督或有监督的模式下运行。在前一种情况下，只需要非缺陷样品的图像，而在后一种情况下，只需要缺陷样品的图像。用户可以从四个不同的区域调整各种参数：采样（特征大小、颜色）、训练（计数历时、训练选择）、扰动（旋转、比例、长宽比、剪切、翻转、亮度、对比）和处理（采样密度、简单区域）。

在本文中，所有使用康耐视ViDi Suite v2.1的实验都是使用ViDi Red工具在监督模式下进行的。该软件在不同的学习配置下进行了广泛的评估，以找到最佳条件与下一节提出的方法进行公平比较。在这次评估中，下列学习配置是不同的：

- 五种注解类型、
- 三种特征尺寸（20、40、60像素）、
- 两种尺寸的输入图像（全尺寸和半尺寸）、
- 有/无90°输入数据旋转。

不同的设置与“分割和决策网络评估”一节中的设置相似，不同的是没有评估不同的损失函数，因为软件没有提供如此详细的控制水平。相反，对不同大小的特征进行了评估，事实证明，这些特征在拟议的数据集上发挥了关键作用。评估了20至60像素的特征。根据文件，不建议使用小于15像素的特征，而大于60像素的特征产生的结果更差。

实施细节通过生产和培训API可以访问ViDi套件的学习和推理核心。所有的实验都是用C#.Net编程语言完成的。评估是通过3倍交叉验证进行的，训练/测试的划分与之前的实验相同，使用灰度图像（颜色通道数设置为1），学习了100个epochs。训练是在训练折叠的所有图像上进行的；因此，使用100%的参数训练选择。模型被导出并在生产模式下对测试折页进行评估，参数为简单区域。

和采样密度设置为1。这确保了在拟议的深度学习模型中使用相同的处理程序。我们使用供应商推荐的默认值作为采样密度。我们还进行了以慢速推理为代价的密集采样的实验，但这并没有改善结果。

结果 在图10中显示了结果。在这些不同的

在不同的学习设置中，使用扩张=5注释训练的模型取得了最好的性能，使用最小的特征尺寸（20像素），不旋转图像，使用原始图像尺寸。该模型实现了99.0%的AP，以及5个错误分类，即5个FN和0个FP。请注意，有一个模型只实现了4次错误分类，尽管整体上有较低的AP。

注释大小 在不同的注释类型中，扩张的注释比非扩张的注释性能更好。然而，在不同的扩张率中，性能增益是最小的，扩张=5和扩张=17之间只有0.1pp的差异。

特征大小 比较不同的特征大小，无论注释精度如何，具有小特征的模型始终优于具有大特征的模型。这可能是由于数据集的特殊性，它具有较高的图像分辨率和许多小的表面缺陷。此外，对半分辨率图像的实验显示，在这种情况下，大特征的表现明显比小特征差。这导致了这样的结论：大的特征尺寸不能捕捉较小的细节，而这些细节对分类是很重要的。

图像大小和旋转 最后，实验还显示，无论是半分辨率图像还是将输入数据随机旋转90°，都没有带来性能的提高。在这两种情况下，性能都略有下降，尽管两者的性能下降都很轻微。

使用最先进的分割网络

接下来，评估了两个标准的分割网络，即DeepLabv3+（Chen et al 2018）和U-Net（Ronneberger et al 2015）。DeepLab架构被选为预训练模型的代表，它可以在目前的语义分割基准上取得了最先进的结果，而U-Net网络的参数为标记，而U-Net架构被选为为精确的像素化分割而设计的模型的代表。关于DeepLabv3+方法的更多详细信息，请读者参考（Chen et al 2018），关于U-Net的详细信息请参考（Ronneberger et al 2015）。

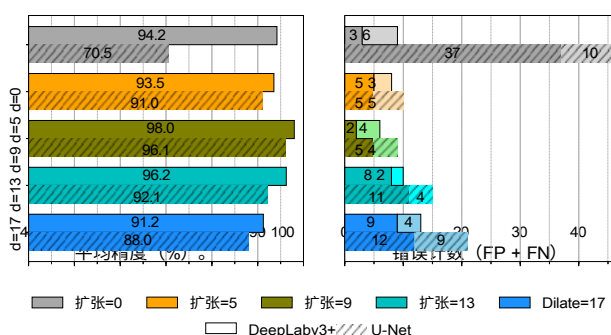


图11 两个标准分割网络（DeepLabv3+和U-Net）的评价

模型。两个模型都在不同的注释下进行了评估，但只考虑了损失函数的交叉熵，并且只使用了没有数据漫游的全分辨率图像尺寸，因为这些设置在以前的实验中被证明是性能最好的。

实施细节 两种分割方法都通过替换拟议网络的分割部分被嵌入到拟议的方法中。两个网络的TensorFlow实现被嵌入到拟议的网络中。这些实验中使用的DeepLabv3+是基于Xception（Chollet 2017）的架构，包含65个卷积层，在一个简单的规模上训练和评估，并使用16的输出步长。这些实验中使用的U-Net是一个修改过的U-Net架构，有24个卷积层，其中唯一的修改是为每个卷积增加了批量归一化。原有的U-Net也以全输入分辨率输出分割；然而，由于全分辨率的像素级精确分割不符合本实验的要求，输出图的分辨率被降低了8倍。这与提议的网络中的输出分辨率相同。

对于这两个分割网络来说，分割层与决策层是分开训练的，这与提议的方法类似，使用3倍交叉验证，训练/测试的比例与之前所有的实验相同。这两种方法还用逻辑回归代替决策网络进行了评估，但事实证明，这种方法的效果更差。DeepLabv3+网络的参数被初始化，该模型在ImageNet（Russakovsky等人，2015）和COCO数据集（Lin

使用正态分布随机初始化，类似于本文提出的网络初始化。这两个网络都被训练了100个历时，其学习程序与所提出的模型相同，即使用0.1的无动力学习率，批次大小为1、

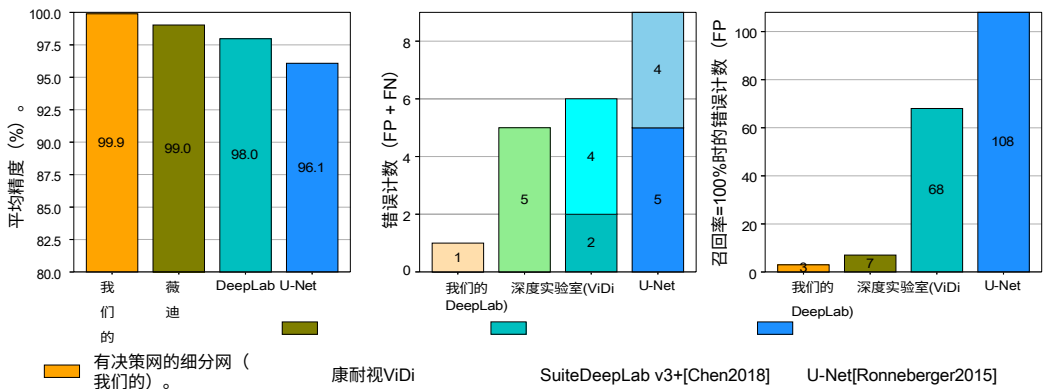


图12 在KolektorSDD上与最先进的技术进行比较（中间的图：深色的假阳性现象和假阴性的光）

并在每个步骤中交替使用有缺陷和无缺陷的图像。

结果图11中显示了结果。在标准网络中，表现最好的模型，即使用 $dilate=9$ 注释训练的DeepLabv3+，达到了98.0%的AP，并在理想的F-measure下获得了两个FP和四个FN。总的来说，略微扩张的注释取得了最好的结果，而用较大的孔径扩张的注释的结果较差。平均而言，DeepLabv3+在平均精度上也比U-Net架构高出2-3个百分点，而不考虑注释类型。

与提议的方法比较

最后，所有三种最先进的方法都与本文提出的网络进行了比较。最先进的方法是与综合分割和决策网络进行比较。为了进行公平的比较，本节所报告的所有方法都是根据前几节的评估中表现最好的训练设置来选择的。对于所有的方法，这包括使用原始图像尺寸（1408×512分辨率），不对输入图像进行旋转，对商业软件使用最小的特征尺寸20像素，对所有其余方法使用交叉熵损失函数。对于注释类型，不同的方法在不同的注释中表现最好。商业软件和带有分割和决策网络的拟议方法在对 $dilate=5$ 注释进行训练时都取得了最好的性能，而DeepLabv3+和U-Net则取得了最好的重构。

表1 前三栏为固定学习值的学习超参数，其余四栏为最佳选择的学习配置设置

方法	纪元数	学习率	初始化	注释类	图像大小	数据	数腐	损失函数	特征尺寸
分割/决策网（我们的）。	100	0.1	$N(0.01)$	扩张=5	1408 × 512	没有	没有	交叉熵	不适用
U-Net (Ronneberger et al 2015)	100	0.1	$N(0.01)$	扩张=9	1408 × 512	没有	没有	交叉熵	不适用
DeepLab v3+ (Chen et al 2018)	100	0.1	预先训练过的	扩张=9	1408 × 512	没有	没有	交叉熵	不适用

康耐视ViDi套件	100	不适用	不适用	扩张=5	1408 × 512	没有	不适用	适用 60
-----------	-----	-----	-----	------	------------	----	-----	----------

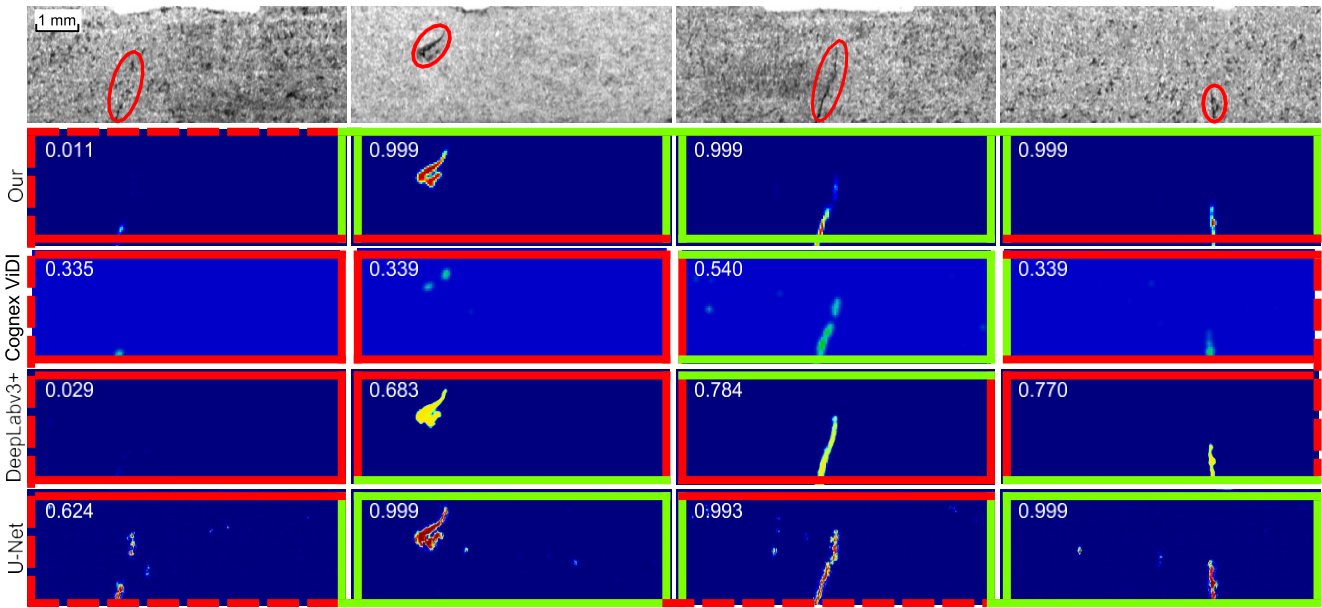


图13 真阳性（绿色实线边界）和假阴性（红色虚线边界）检测的例子与分割输出和相应的分类（第一行中圈出了实际的缺陷）。

几个错误的检测。除了本文提出的方法和商业软件外，在所有相关的方法中都可以看到高分的假阳性。特别是，U-Net返回了一个带有大量噪声的输出，这使得真正的缺陷和错误的检测不能被干净地分开，即使有额外的决策网络。另一方面，所提出的方法没有任何假阳性的问题，并且能够正确预测这些图像中没有缺陷。

工业环境中的结果 当考虑在工业环境中使用所提出的模型时，重要的是要确保检测出所有的缺陷项目，即使牺牲更多的假阳性。由于之前的指标不能反映这些条件下的性能，本节将进一步说明，如果要求零漏检率，即要求100%的召回率，将获得多少误报。这些误报代表了需要由熟练工人进行人工验证的项目数量。

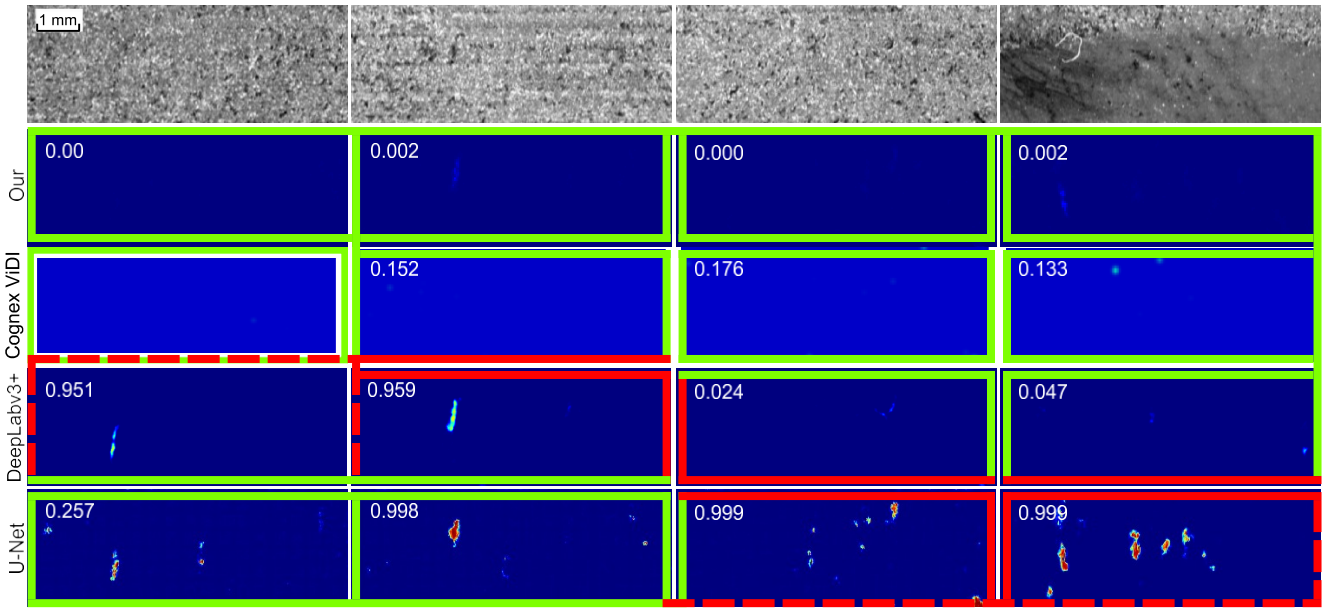


图14 真阴性（绿色实线边界）和假阳性（红色虚线边界）检测的例子与分割输出及相应的分类分数

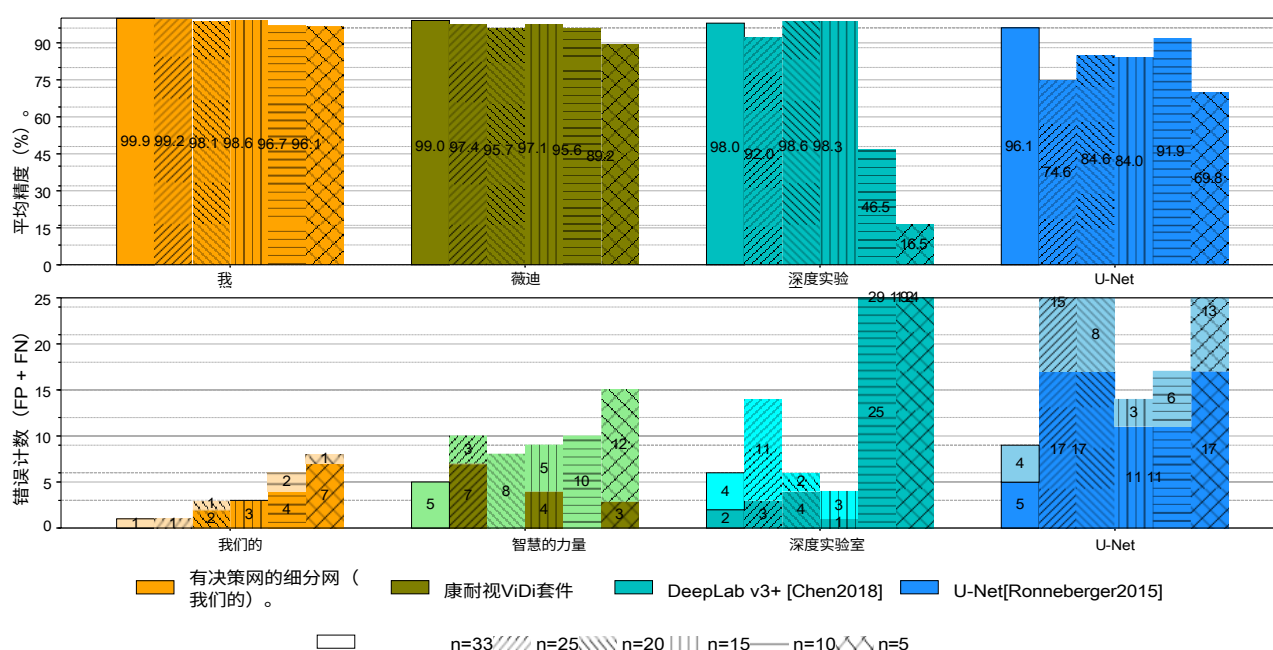


图15 KolektorSDD在不同数量的阳性（缺陷）训练样本下的分类性能

并直接指出为达到所需的精度所需的工作量。

如图12中最右边的图所示，结果显示，本文提出的模型在所有399张图像中只引入了3个零失误率的假阳性。这代表了所有图像的0.75%。另一方面，相关的方法取得了更差的结果，商业产品需要对7张图像进行人工验证，而标准分割网络则需要对DeepLabv3+和U-Net分别进行68和108次人工验证。请注意，这两个标准分割的结果包括使用提议的决策网络。使用逻辑回归而不是提议的决策网络导致性能明显下降。

对训练样本数量的敏感度

在工业环境中，一个非常重要的因素是对有缺陷的训练样本数量的重新要求，因此我们也评估了较小的训练样本量的效果。评估采用了3倍交叉验证法，训练/测试分割与之前所有实验中使用的相同，因此在对所有训练样本进行训练时，每个折叠中有效使用33个阳性（缺陷）样本。然后减少阳性训练样本的数量，以有效地获得每个折叠的25、20、15、10和5个样本的训练规模 N ，而每个折叠的测试集保持不变。被删除的训练样本被随机抽出，但所有方法都被删除了相同的样本。

训练和测试程序与之前的所有实验相同。

提议的分割和决策网络与商业软件Cognex ViDi Suite和两个最先进的分割网络进行了比较。所有的方法都是使用在前面的实验中确定的最佳性能的训练设置进行评估的，即使用 $扩张=5$ 的注释（或 $扩张=9$ 的分割网络），全图像分辨率，交叉熵损失和无图像旋转。结果在图15中报告。当只使用25个有缺陷的训练样本时，提议的分割和决策网络保持了超过99%的AP和单一的错误的分类的结果。当使用更少的训练样本时，结果就会下降，但当只使用5个有缺陷的训练样本时，提议的方法仍然达到了96%左右的AP。对于康耐视ViDi套件，可以观察到更明显的性能下降，然而，在这种情况下，结果在 $N=25$ 时已经下降到97.4%的AP。当只使用5个有缺陷的训练样本时，商业软件的AP略低于90%。在图15下半部分描述的错误的分类数量中也观察到同样的趋势，深色代表假阳性，浅色代表假阴性。

另一方面，当使用较少的训练样本时，DeepLab v3+和U-Net的性能比建议的方法更差。U-Net的性能迅速下降，而DeepLab即使在只有15个有缺陷的训练样本时也能保持相当好的结果。请注意，在20个和15个有缺陷的训练样本中，DeepLab的性能略胜一筹。

表2 与最先进的方法在可学习参数数量和平均精度上的比较

方法	平均数量	参数的精度
分割/决策网(我们)	15.7 mio	99.9
康耐视ViDi套件	不适用	98.9
DeepLab v3+ (Chen et al 2018)	41.1 mio	97.9
U-Net (Ronneberger et al 2015)	31.1 mio	96.1

在所有训练样本中获得的结果，表明DeepLab对特定的训练样本相当敏感，删除这些样本有助于提高性能。U-Net对训练样本数量的减少明显更敏感；平均精度的结果从75%到略高于90%不等。然而，对于10个和5个有缺陷的训练样本，DeepLab表现最差，AP分别只有46%和16%。

总的来说，实验结果表明，在训练样本数量较少的情况下，所提出的方法也能保持卓越和稳定的性能。

计算成本

本文提出的方法在计算成本方面优于最先进的分割方法，与商业软件相比具有竞争力。图16重新描述了与平均精度有关的通过时间。结果是在单个NVIDIA TITAN X (Pascal) GPU上获得的。所提出的方法比DeepLab v3+和U-Net要快得多，而且精度也更好。这是在参数数量较少的情况下实现的，这反映在图16中的标记大小上，也显示在表2中。所提出的模型只用了1570万个参数就实现了这一性能，而U-Net和DeepLab v3+的参数数量是其两倍以上，分别为3110万个和4110万个参数。康耐视ViDi套件的参数数量没有公开。建议的方法和商业软件在图16中用星形标记描述的一半分辨率下也被显示出来。这表明建议的方法比全分辨率下的前向通过快3倍，半分辨率为33ms，全分辨率为110ms。使用商业软件康耐视ViDi Suite的速度最快，每幅图像10毫秒。然而，当使用一半的图像分辨率时，拟议的最佳性能模型实现了类似的性能，只是计算成本稍大。请注意，所提出的模型在TensorFlow框架中实现了这一性能，而没有应用任何计算优化，而可以有把握地假设，商业软件已被高度优化，以尽可能地减少计算成本。

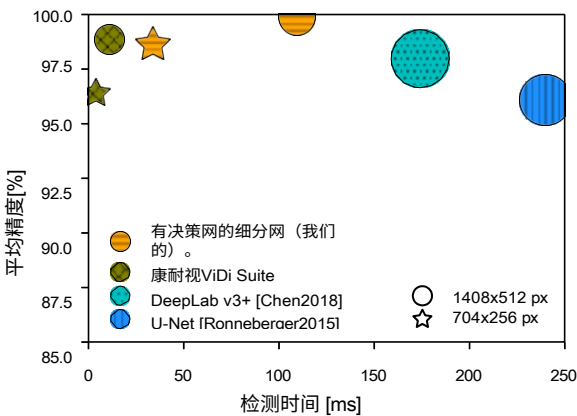


图 16 单一图像的检测（正向通过）时间与分类性能的关系

讨论和结论

本文从具体的工业应用角度出发，探讨了一种利用分割网络进行表面缺陷检测的深度学习方法。本文提出了一种两阶段的方法。第一阶段包括一个根据缺陷的像素级标签训练的分割网络，而第二阶段包括一个建立在分割网络之上的额外决策网络，以预测整个图像是否存在异常情况。对所提出的方法进行了广泛的评估，评估对象是一个半成品的工业产品，即换向器，其表面缺陷表现为母体的断裂。这个问题域已经作为一个基准数据集公开提供，称为Kolektor表面缺陷数据集（KolektorSDD）。所提出的方法在这个领域与几个最先进的方法进行了比较，包括专有软件和两个基于深度学习的标准分割方法。

在KolektorSDD上的实验表明，所提出的模型取得了明显优于相关方法的结果，只出现了一次错误分类，而相关方法则出现了五次或更多的错误分类。这可以归功于所提出的分割和决策网络的两阶段设计，以及改进的感受野大小和增加的捕捉缺陷的细节的能力。相关的方法缺少其中的一些特征。例如，表现最差的分割方法U-Net的感受野大小有限，只有45个像素，而拟议的方法有205个像素。尽管DeepLabv3+提高了感受野的大小，但它是牺牲太多的参数为代价的，这导致了模型的过度拟合，尽管它是在单独的数据集上预训练的。

另一方面，由于该方法的细节没有公开，所以很难评估与商业软件的差异。然而，经验

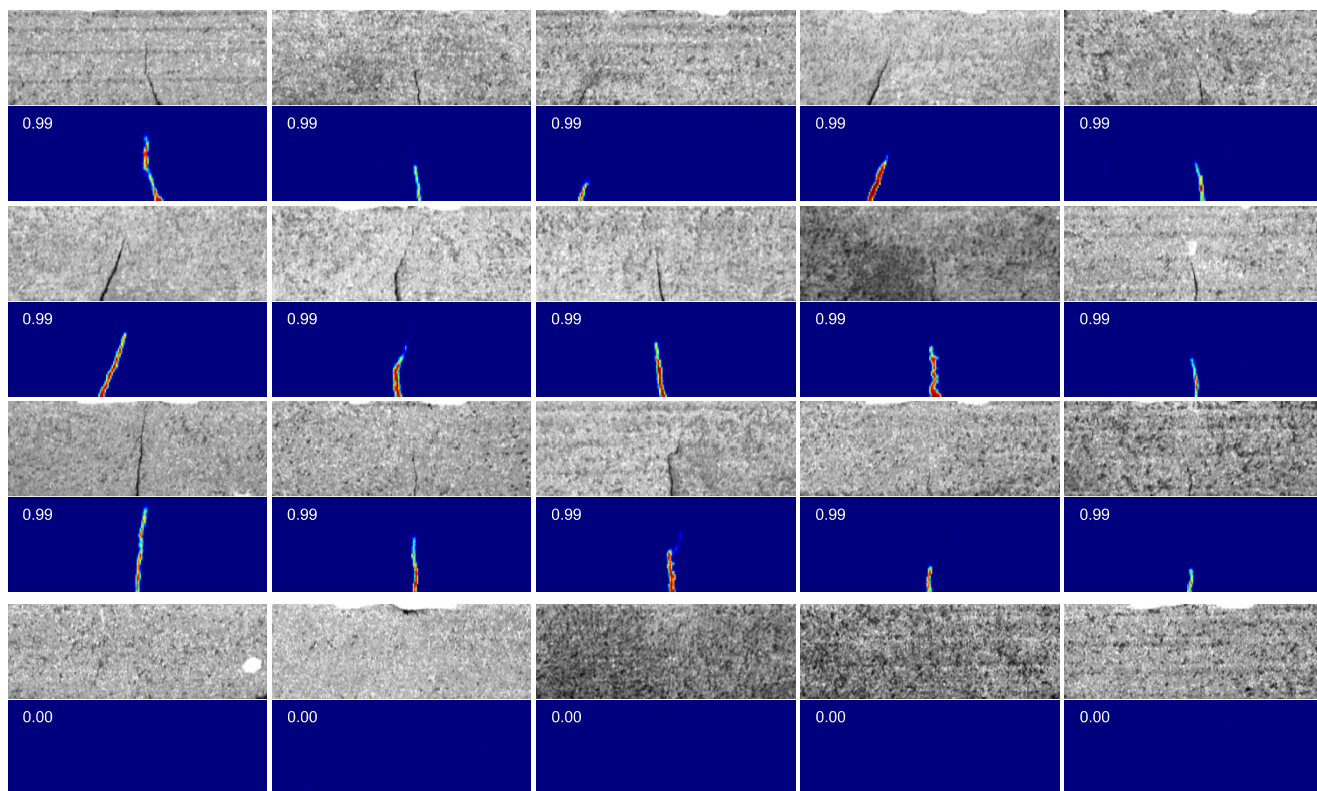


图17 用建议的方法在KolektorSDD上检测出的真阳性（上面三行）和真阴性（下面一行）的例子（每个例子的左上角都描述了分类得分）。

实验表明，在使用低分辨率的图像时，商业软件的表现明显比建议的方法差。这个实验表明，商业软件很难捕捉到缺陷的更精细的细节，需要更高的分辨率来获得良好的性能。然而，即使使用高分辨率的图像，它仍然不能达到与建议的方法相同的性能。

所提方法的性能是通过从33个有缺陷的样本中学习而实现的。图17中描述了几个正确分类的例子。此外，只使用25个有缺陷的样本表明仍然可以达到良好的性能，而相关的方法在这种情况下取得了更差的结果。这表明，所提出的深度学习方法适用于所研究的工业应用中有限数量的缺陷样本。此外，为了进一步考虑工业环境的应用，对三个重要特征进行了评估：（a）实现100%检测率的性能，（b）注释的细节和（c）计算成本。就实现100%检测率的性能而言，所提出的模型在所有399张图像中只需要三张图像进行人工检测，导致了0.75%的检测率。事实证明，大而粗的注释也足以达到与更细的注释相似的性能。在某些情况下，较大的注释甚至会带来更好的性能。

比起使用精细的注释。这个结论似乎是反直觉的；然而，在用于对每个像素进行分类的感受野大小中可以找到一个可能的解释。离缺陷区稍远的像素的感受野仍将覆盖缺陷区的一部分，因此，如果它们被正确注释，就能有助于找到对其检测很重要的特征。这一结论可以减少适应新领域的方法时的人工工作，并将导致劳动力成本的降低和生产线灵活性的提高。

然而，建议的方法仅限于特殊类型的任务。特别是，该架构是为那些可以被视为具有像素级注释的分割问题的任务而设计的。其他的质量控制问题也存在，基于分割的解决方案不太适合。例如，复杂的三维物体的质量控制可能需要检测破碎或丢失的部分。这种问题可以通过检测方法来解决，比如Mask R-CNN（Kaiming等人，2017）。

这项研究证明了该方法在特定任务（裂纹检测）和特定表面类型上的性能，但该网络的结构并非只为这个特定领域而设计。在新领域的学习是可能的，不需要任何修改。该架构可以应用于包含多个复杂表面的图像，也可以应用于检测其他不同的缺陷。

如果有足够数量的缺陷训练样本,并且特定的缺陷检测任务可以被看作是一个表面分割问题,那么就可以对缺陷模式,如划痕、污点或其他不规则现象进行检测。然而,为了进一步评估这一点,需要新的数据集。据我们所知,DAGM数据集(Weimer等人,2016年)是唯一公开可用的有注释的数据集,其中有不同的表面和缺陷类型,适合评估基于学习的方法。所提出的方法在这个数据集上取得了完美的结果,但这个数据集是合成的,而且根据所取得的结果也是饱和的。因此,未来的努力还应该集中在获取基于真实世界的视觉检测问题的新的复杂数据集上,在这些数据集上可以对深度学习(和其他)方法进行全面的现实评估;本文介绍的数据集是朝着这个方向迈出的第一步。

鸣谢 这项工作得到了以下研究项目和计划的部分支持:GOSTOP计划C3330-16-529000由斯洛文尼亚共和国和欧洲区域发展基金共同资助,ARRS研究项目J2-9433(DIVID),以及ARRS研究计划P2-0214。我们还要感谢Kolektor Orodjarna d. o. o.公司为拟议的数据集提供图像,以及提供高质量的注释。

参考文献

- Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mane D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viegas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X (2015) TensorFlow: 异构系统上的大规模机器学习。URL <https://www.tensorflow.org/> 7
- Bulnes FG, Usamentiaga R, Garcia DF, Molleda J (2016) An efficient method for defect detection during the manufacturing of web materials. *Journal of Intelligent Manufacturing* 27(2):431-445, DOI 10.1007/s10845-014-0876-9, URL <http://dx.doi.org/10.1007/s10845-014-0876-9> 1
- Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *Tech. rep.* 3, 11, 12, 15
- Chen PH, Ho SS (2016) Overfeat对基于图像的表面缺陷分类任务有用吗? In: IEEE 图像处理国际会议 (ICIP), 第749-753页 2, 3
- Chollet F (2017) Xception: 深度学习与深度分离的卷积。 *计算机视觉与模式识别* 2017:1800-1807, DOI 10.1109/CVPR.2017.195 11
- 康耐视 (2018) VISIONPRO VIDI: 基于深度学习的工业图像分析软件。URL <https://www.cognex.com/products/machine-vision/vision-software/visionpro-vidi> 10
- Faghih-Roohi S, Hajizadeh S, Nuñez A, Babuska R, Schutter BD (2016) Deep Convolutional Neural Networks for Detection of Rail Surface Defects. In: *International Joint Conference on Neural Networks*, October, pp 2584-2589 2
- Kaiming H, Gkioxara G, Dollar P, Girshick R (2017) Mask R-CNN. In: *ICCV*, pp 2961-2969 16
- Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet分类与深度卷积神经网络。 In: *Advances In Neural Information Processing Systems* 25, pp 1097-1105 2
- Kuo CFJ, Hsu CTM, Liu ZX, Wu HC (2014) 使用两阶段反向传播神经网络的LED芯片自动检测系统。 *Journal of Intelligent Manufacturing* 25(6):1235-1243, DOI 10.1007/s10845-012-0725-7 2
- Lin H, Li B, Wang X, Shu Y, Niu S (2018) 使用深度卷积神经网络对LED芯片进行自动缺陷检测。 *Journal of Intelligent Manufacturing* pp 1-10, DOI 10.1007/s10845-018-1415-x, URL <https://doi.org/10.1007/s10845-018-1415-x> 3
- Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL (2014) Microsoft COCO: Common objects in context. *Lncs* 8693 *Lncs*(part 5):740-755 3, 11
- Long J, Shelhamer E, Darrell T (2015) 用于语义分割的完全卷积网络。 In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol 8828, pp 3431-3440, DOI 10.1109/CVPR.2015.7298965 5
- Masci J, Meier U, Ciresan D, Schmidhuber J, Fricout G (2012) 用Max-Pooling卷积神经网络进行钢铁缺陷分类。在: *国际神经网络联合会议论文集*, DOI 10.1109/IJCNN.2012.6252468 2
- Oztemel E, Gursev S (2018) 工业4.0及相关技术的文献综述。 *智能制造杂志* DOI 10.1007/s10845-018-1433-8, 网址 <https://doi.org/10.1007/s10845-018-1433-8> 1
- Paniagua B, Vega-Rodríguez MA, Gomez-Pulido JA, Sanchez-Perez JM (2010) 通过使用图像处理和神经模糊计算改善软木塞的工业分类。 *Journal of Intelligent Manufacturing* 21(6):745-760, DOI 10.1007/s10845-009-0251-4 1
- Rac'ki D, Tomaz'evic' D, Skoc'aj D (2018) 用于纹理表面异常检测的紧凑卷积神经网络。 In: *IEEE计算机视觉应用冬季会议*, 第1331-1339页, DOI 10.1109/WACV.2018.00150 3, 4, 5
- Ronneberger O, Fischer P, Brox T (2015) U-Net: 用于生物医学图像分割的卷积网络工程。 In: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*, pp 234-241 3, 11, 12, 15
- Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision* 115(3):211-252, DOI 10.1007/s11263-015-0816-y, URL <http://dx.doi.org/10.1007/s11263-015-0816-y> 3, 11
- Sermanet P, Eigen D (2014) OverFeat: 使用卷积网络进行综合识别、Localization和检测。 In: *International Conference on Learning Representations (ICLR2014)*, CBLS, April 2014 3
- Weimer D, Thamer H, Scholz-Reiter B (2013) 使用神经网络和统计特征表示学习纹理表面的缺陷分类器。 *Procedia CIRP* 7:347-352, DOI 10.1016/j.procir.2013.05.059, URL <http://dx.doi.org/10.1016/j.procir.2013.05.059> 2
- Weimer D, Scholz-Reiter B, Shpitalni M (2016) 设计深度卷积神经网络架构用于工业检测中的自动特征提取。 *CIRP Annals - Manufacturing Technology* 65(1):417-420, DOI 10.1016/j.cirp.2016.04.072, URL <http://dx.doi.org/10.1016/j.cirp.2016.04.072> 3, 17
- Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A (2016) Learning Deep Features for Discriminative Localization. In: *计算机视觉*

觉和模式识别3