基于轻量化“mobilenet”的模型对工业零件缺陷的检测研究

摘要

关键词：深度学习、计算机视觉、破损检测、mobilenet的模型

1. 问题重现

1.1问题背景

金属或塑料产品的表面缺陷可能包括划痕、凹痕、气泡或其他不规则的瑕疵，这些问题不仅会影响产品的外观，还可能导致产品在使用过程中出现性能下降或提前损坏。因此，研究人员致力于开发自动检测表面异常的技术，以提高产品质量和生产效率。这一领域的研究和应用对于视觉检测技术的发展具有重要意义，有望在制造业和质量控制领域产生深远影响。[1]

1.2问题要求

Kolektor集团提供了一个缺陷产品的图像数据集[2]，我们将以这个数据集为主，并搜索了一系列有关的数据为辅助，研究一个数学模型，通过照片可以自动检测产品表面缺陷。要求该模型可以适用于价格低廉的存储空间和计算能力十分有限的手持设备上，并且该模型具有泛化能力，在遇到其他缺陷类别时仍然具有泛化能力。用该数学模型解决下列问题：

问题1：根据所提供的数据集，用该模型自动判断照片是否有表面缺陷，并且需要测量出该数学模型为此项任务所需的计算量和存储空间。

问题2：用该数学模型自动检测并标记出照片的表面缺陷所在位置或区域，并且需要测量出该数学模型所需的计算量、存储空间和标注精度。

问题3：当该数学模型遇到的照片表面缺陷类型与该数据集的照片表面缺陷类型不完全相同时，该模型仍然可以适用，对此说明该数学模型的泛化能力。

1. 问题分析（II）

2.1问题一的分析

问题一中要求我们通过建立模型判断照片中的金属或塑料产品是否存在缺陷，对于此问题，我们首先将不同样品分为品质良好与品质低劣的两组，考虑到问题检测所需的轻量化以及实用性，因此选择采用TensorFlow和Keras构建了一个表面缺陷检测模型，通过对该模型的进行了15轮次训练，最终将其实际用于缺陷的检测，从而能够判断出存在缺陷的产品，并检测该模型的计算量以及存储空间大小，确保其能在小型设备上运行，达到题目要求。

2.2问题二的分析

问题二要求我们通过建立模型自动检测并标记出照片表面缺陷的位置或区域，首先面对图片大小不均，我们首先编写了一段缩放数据图片的代码，将图片统一缩放至300\*300的大小，但是为了避免计算量以及存储空间过大，同时放大图片一些难以识别的不明显缺陷以提高标注的精度。因此，我们利用labelImg工具包，采用人工标记的方法对一些缺陷图片进行人工标记并使其产xml文件,再利用代码生成了将xml标记文件内的信息转为tfrecord文件便于再次训练我们训练好的MobileNetV2模型，使其具有识别并标记产品表面缺陷的功能。

2.3问题三的分析

问题三要求我们的模型遇到其他不完全相同的缺陷类别时仍然适用，以此说明该模型的泛化能力。首先题目所提供的图片集的表面缺陷最多存在一片区域，我们采用了ImageDataGenerator来增强图像数据。这些数据增强操作可以帮助模型更好地泛化到不同的数据集，并提高模型的性能和鲁棒性。在训练过程中，每个批次的图像都会随机应用这些增强操作，从而增加数据的多样性。而现实中可能会存在一个图片同时存在多个缺陷的情况，所以我们选择利用拼接的方法，使得一张图片上同时存在多个缺陷，以此让我们的模型得以训练。。。

2.4 最优配置

* 从适用性角度：由于图片的缺陷通常不是一致的，同时检测的图片的样式也不是一致的。然而，TensorFlow作为一个强大的机器学习和深度学习框架，TensorFlow中的ImageDataGenerator可以通过对训练数据进行一定的变换和扩充，如旋转、平移、缩放等，可以扩展数据集的规模和多样性，从而增强模型的泛化性能。但是TensorFlow的设计和实现相对复杂，尤其在处理大规模和复杂模型时。它的灵活性和丰富的功能使得其配置和调试变得更加繁琐，需要更多的技术经验和时间的投入。另外其代码的可读性较差，很难直观地理解模型的结构和计算过程。
* 从量级的角度：Inception系列模型是由Inception-V1、Inception-V2、Inception-V3等组成的一系列卷积神经网络架构，采用多尺度卷积和并行结构来提高网络的表达能力。然而，重量级模型通常具有更深、更复杂的网络结构，需要更多的参数和计算资源，但在一些复杂的视觉任务中可能具有更好的性能。题目要求在小型手持设备上能够运行该模型，但是由于Inception系列模型规模较大，这些模型在移动设备和资源受限环境中的部署可能会面临挑战。所以我们选择MobileNetV2模型，该模型是一种轻量级的卷积神经网络架构。MobileNetV2相对于其他卷积神经网络架构来说非常轻量，具有较少的参数和计算复杂度。这使得它在移动设备等资源受限的环境下具有更好的性能。另外，MobileNetV2的网络结构可以根据需求进行缩放和调整。通过调整网络宽度和深度的参数，可以在保持性能的同时控制模型的大小和计算复杂度。但是因为它的网络结构相对较浅，可能无法捕捉到一些更复杂的图像特征。
* 折中：虽然相对于一些其他深度学习神经网络，MobileNetV2在处理高分辨率图像时性能略有限，并且网络结构相对较浅，但作为专门用于移动设备的轻量级卷积神经网络模型，它在保持较高准确率的同时具有较低的参数量和计算复杂度。这使得MobileNetV2非常适合在资源有限的移动设备上进行部署。通过TensorFlow提供的预训练模型和优化工具，例如数据增强，我们能够更好地利用MobileNetV2模型，有效地应对移动设备环境下的计算和存储限制。因此，尽管在处理高分辨率图像方面存在一定限制，MobileNetV2仍然是一种非常实用的模型选择，特别适用于移动设备端的场景。总的虽然在性能上稍微有所不足，但是符合题目要求所具备的轻便性。

2.5如果没有可用数据怎么办？

* 2.5.1数据收集：考虑到题目所给出的图片检测文件包一共有16\*50个图片，这个数据量是非常可观的，足够我们进行识别模型的训练和测试，所以我们小队直接开始进行数据的处理，分析以及使用。
* 2.5.3合成数据：为了防止我们训练的模型过度拟合，只是通过记忆的方式识别出缺陷图片，我们采用数据增强、物理模拟或几何变换的方法来生成合成数据。
* 1）数据增强：即使数据较为可观，也可以通过数据增强技术进行扩充，我们小组通过图像旋转、翻转、缩放等数据增强等方法生成了对应的缺陷图片。  
  2）物理模拟或几何变换：缺陷受物理过程，如应力、磨损等因素的影响，通过模拟这些过程来创建了新的缺陷图像生成了一部分训练和测试图片。
* ~~3）生成对抗网络（GANs）：通过训练一个生成网络来生成新的图片，这些图片在视觉上与真实图片难以区分。GANs 特别适用于生成高度逼真的图像数据，可以用来生成包含特定类型缺陷的合成图片~~

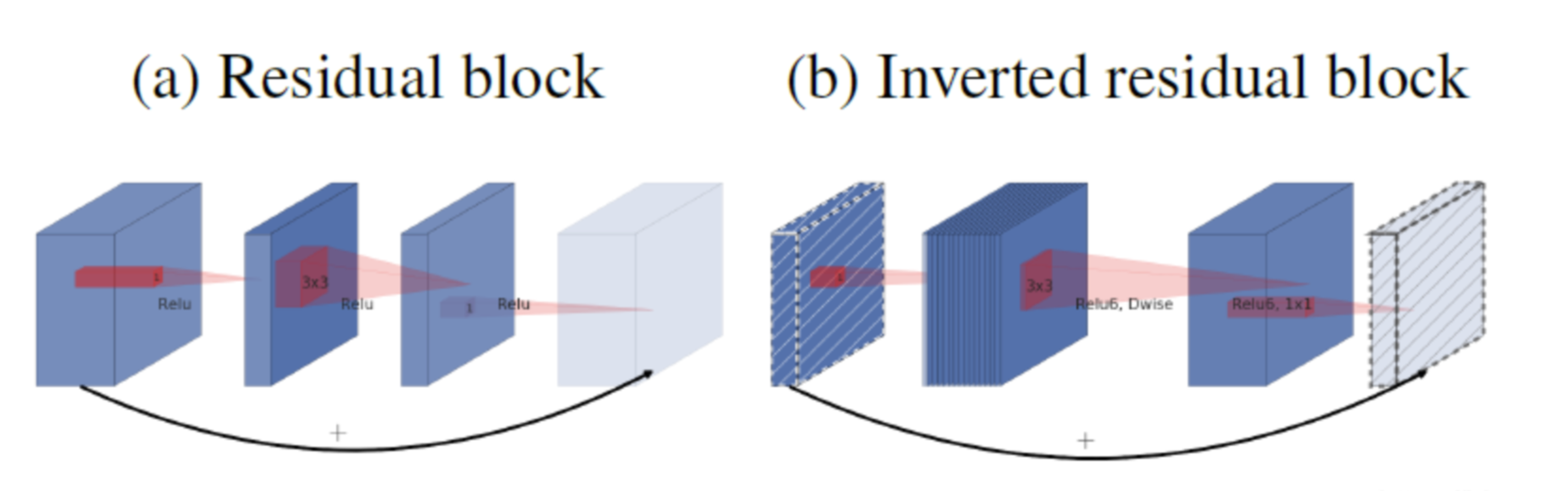
1. 模型

3.1模型一

MobileNet-v2模型

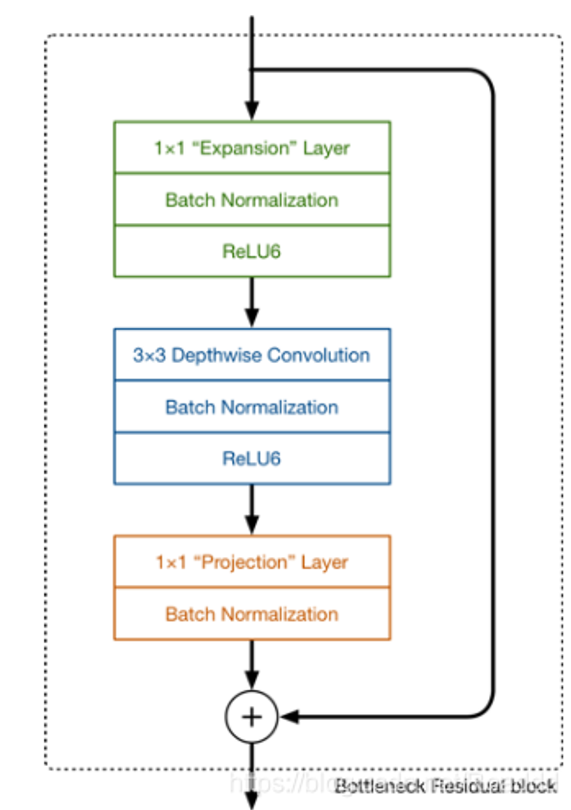
1）倒置残差结构（Inverted Residuals）

这种结构使用轻量级的深度可分离卷积（depthwise separable convolution），它首先通过一个 1x1 卷积对特征通道进行扩展，然后应用深度卷积，最后再通过另一个 1x1 卷积将特征通道数减少，这样可以在不显著增加计算量的情况下提高模型的表达能力。



2）线性瓶颈（Linear Bottlenecks）

在倒置残差结构的最后阶段，使用线性激活函数而不是 ReLU。这是基于这样的观察：ReLU 激活可能会破坏特征信息，尤其是当特征空间被压缩时。故采用ReLU6激活函数，该函数是ReLU的变体，它限制输出值在0到6之间，这对于低精度计算是有益的。



3）深度可分离卷积（Depthwise Separable Convolution）：这是一种减少计算量的技术，它将标准卷积分解为深度卷积（depthwise convolution）和逐点卷积（pointwise convolution，即 1x1 卷积）。深度卷积对输入的每个通道分别应用单独的滤波器，而逐点卷积则负责将深度卷积的输出通道组合起来。

3.1.1变量分析（做一个三线表）

（三线表我有软件我可以做）

3.1.2模型假设

* 假设金属或塑料产品所得图片的缺陷没有人为因素的影响，全部都是基于在生成或使用过程中受物理过程，如应力、磨损等因素的影响而造成缺陷。
* 对于金属或塑料产品所得的图片的识别等同于对图片进行存在缺陷和不存在缺陷的分类
* 不考虑金属或塑料产品存在除表面缺陷以外，如产品内部存在缺陷或产品原料质量等其他缺陷

3.1.3模型基础

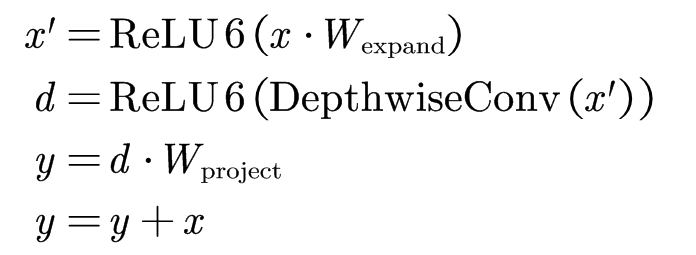
MobileNetV2作为轻量级深度学习模型是基于MobileNetV1进一步完善与优化之后提出来的。MobileNetV2基本上都是在移动设备以及边缘设备上运行，所以其尤其关注模型的效率速度，以及存储大小，非常满足题目所需的需求。

MobileNetV2的基础结构包括以下几个关键的创新点：

1. Inverted Residuals（倒残差结构）：

MobileNetV2介绍了一种新型倒残差结构，该结构较传统残差结构先通过1x1卷积层扩大通道数，接着执行深度为3x3可分离卷积并最终通过1x1卷积层压缩通道数返回。该结构在保留特征信息传递的前提下降低了计算量。

**倒置残差结构（Inverted Residual Block）：**

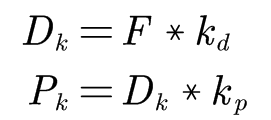


2. Linear Bottlenecks（线性瓶颈层）:

在MobileNetV2中，每个倒残差结构的最后一个1x1卷积层不是使用非线性激活函数，而是继续保持线性，目的是为了避免其在低维空间中造成的信息损失。这些线性瓶颈层有助于保持特征的完整性。

3. Depthwise Separable Convolutions（深度可分离卷积）:

**深度可分离卷积（Depthwise Separable Convolution）：**



深度可分离卷积首先对每个输入通道分别进行卷积，然后使用逐点卷积（pointwise convolution，即1x1卷积）来组合深度卷积的输出。这种分解方式显著减少了模型的参数数量和计算量。

4. Efficient Architecture:

MobileNetV2的架构经过精心设计，以确保模型不仅可以保持高准确度，而且具有低延迟和高效率。使其十分适合在资源受限的设备上运行，如智能手机等。

###用的人工智能，需要进行提炼和删减###（已删减）

3.1.4解决方案和结果

Step1：数据集的处理

1. 对数据集的初步评估，为下一步模型的训练作好准备。
2. 编写了python代码，对图像集先后进行了图像特征提取，对图像旋转、翻转、缩放生成以及物理模拟或几何变换，生成了新的缺陷图像。
3. 将新的缺陷图像整合到了题目所给的图像集中，每张图片的大小为300x300像素，用以模型训练。

~~首先，在得到题目附带的数据集图片后，我们小组首先进行了对数据集的初步评估，为下一步模型的训练作好准备。我们首先~~~~编写了python代码，对图像集先后进行了图像特征提取，对图像旋转、翻转、缩放生成以及物理模拟或几何变换，生成了新的缺陷图像，并将其整合到了题目所给的图像集中，每张图片的大小为300x300 像素，用以模型训练。~~

Step2：MobileNetV2模型的选择与构建

1. 通过题目要求和网上资料分析，确认了适用于移动设备的轻量级网络MobileNetV2模型作为训练模型。
2. MobileNetV2 模型作为基础模型，通过编写的python代码去加载了预训练的权重。
3. 在模型顶部添加了新的层，包括一个全局平均池化层和两个全连接层。
4. 然后一个全连接层使用sigmoid激活函数。
5. 由于对产品缺陷检测是一个分类缺陷和正常的二分类问题，所以设置了迁移学习，循环遍历基础模型的每一层。
6. 再将它们设置为不可训练，确保在训练过程中模型的权重不会更新。  
   ~~考虑到题目所要求构建的模型通过照片可以自动检测产品表面缺陷，并且可以适用于价格低廉的存储空间和计算能力十分有限的手持设备上，并且该模型需要具有一定的泛化能力。我们小队通过网上资料分析最终确认了适用于移动设备的轻量级网络MobileNetV2模型作为训练模型。~~

~~通过编写的python代码我们小队加载了MobileNetV2 模型作为基础模型，并加载了预训练的权重，在模型顶部添加了新的层，包括一个全局平均池化层和两个全连接层。最后一个全连接层使用sigmoid激活函数，因为对产品缺陷检测是一个分类缺陷和正常的二分类问题。然后，我们小组设置了迁移学习，循环遍历基础模型的每一层，并将它们设置为不可训练，确保在训练过程中模型的权重不会更新。我们使用 model.compile 方法编译模型，指定了损失函数、优化器和评价指标。采用的是适用于只有两个类别标签的情况下的交叉熵损失函数 ，以及结合了梯度下降的两个扩展——动量（Momentum）和自适应学习率（Adaptive Learning Rate）的Adam 优化器。（？？？）~~

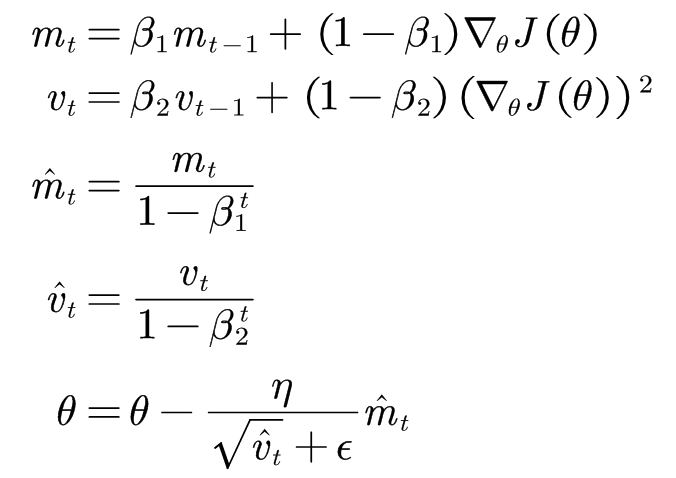
交叉熵损失函数（Binary Cross-Entropy Loss Function）：

**动量（Momentum）**：





**自适应学习率（Adaptive Learning Rate）**：

~~x`~~

Step3：模型的训练

1. 按照训练集与测试集的大小，将训练集分为15组。
2. 再将每个epoch的步骤数设置为45，验证步骤数设置为4。
3. 设置好这些参数，再采用model fit方法训练MobileNetV2模型。

~~我们采用model.fit 方法训练MobileNetV2模型，按照训练集与测试集的大小，训练集每个 epoch 的步骤数设置为45，总共的训练轮数为15次，验证步骤数设置为4。设置好这些参数后，我们开始进行了模型的训练。~~

Step4：模型评估

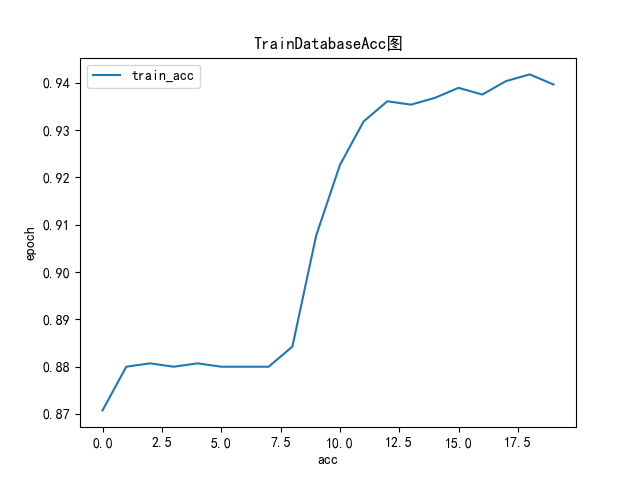
1. 利用model evaluate方法，验证图片集来评估模型的性能。
2. 最后得到了损失值和评价指标值。

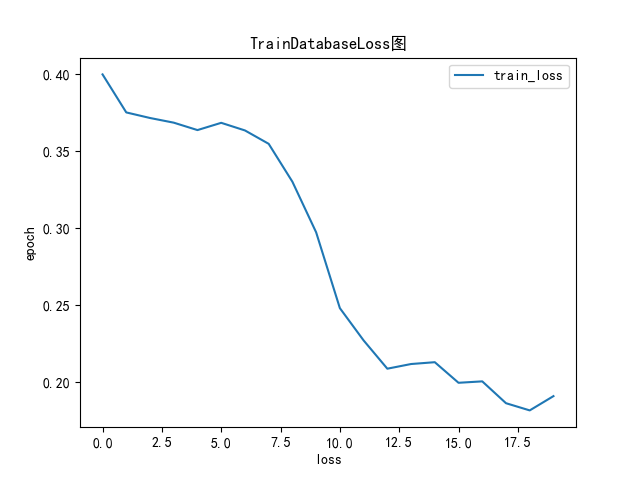
~~我们使用了model.evaluate 方法在验证集来评估模型的性能，返回得到了损失值和评价指标值。~~

Step5：结果可视化

1. 利用matplotlib库，绘制出模型在训练过程中的表现的训练和验证的准确率和损失曲线。
2. 该曲线可以帮助观察训练出的模型是否存在过拟合或者欠拟合。

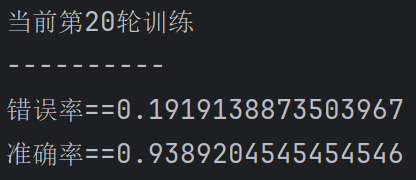
~~利用matplotlib库，我们绘制了可以帮助我们理解模型在训练过程中的表现的训练和验证的准确率和损失曲线，帮助我们观察训练出的模型是否存在过拟合或者欠拟合。~~





由图可知，图一为反映的是训练准确率（TrainDatabaseAcc），这张图展示了模型在训练集上的准确率是如何随训练轮次变化的。可以发现模型训练的准确率呈上升趋势，表示模型在正确分类图像方面变得越来越好，最终该模型的准确率约为93.89%

图二为训练损失（TrainDatabaseLoss）：这张图展示了随着训练轮的增加，模型在训练集上的损失值呈下降趋势，表示模型正在学习并改善其在训练集上的表现，最终该模型的训练损失错误率约为19.19%



总体来说，本次训练出的模型准确率高，适用性较强，不存在过拟合的情况，能够用于小型移动设备对产品表面缺陷的检测。  
3.1.5结果分析（对结果进行一些分析讨论）（不知道该如何写）

3.1.6优势与劣势（已改）

* 优势：

1. 这个模型是一种轻量级的神经网络，其计算和存储能力都能满足大多数小型移动设备的需求。MobileNetV2采用了深度可分离卷积（depthwise separable convolution）技术，这大大降低了模型所需的参数数量和计算的复杂性，从而使得该模型具有很高的轻便性。
2. MobileNetV2这一模型在移动设备上的运行表现出色，不仅推理速度迅速，而且保持了很高的准确性
3. 在MobileNetV2的设计初期，就已经充分考虑了移动设备的计算资源限制，这使得它非常适合在CPU、GPU甚至DSP上运行，因此，这个模型对于移动设备的图像识别是非常合适的。
4. MobileNetV2 引入了线性瓶颈层和反向残差连接，这些改进有助于信息的快速传递，并提高了特征提取的效率

* 劣势：

1. 尽管MobileNetV2是一个轻量级的模型，但在非常有限的数据集上，这个模型可能会出现过度拟合的情况，并且模型的复杂性相对较高，这是因为它仍然包含了一定数量的参数。当数据集规模极小，而模型的复杂性相对较高时，该模型可能会对训练数据进行过度的拟合，也就是说，它可能过度地吸收了训练数据中的噪声和异常特性。
2. 这个模型在正则化技术方面存在不足，例如L1/L2正则化和dropout等方法能有效地降低过度拟合的情况。若未妥善应用这些技术手段，该模型有可能对小数据集进行过度拟合，因此，这一模型更适合用于实时监测产品表面的小型移动检测设备。

3.2模型二

3.2.1额外的变量（对模型一的补充）

3.2.2额外的假设（对模型一的补充，模型一中有了则不用写了，仍然做一个三线表）

3.2.3模型的基础

基于我们在第一问用为标记的MobileNetV2模型，通过训练，使其拥有了对图像进行特征提取，并对图像进行缺陷和正常的二分类操作，问题二要求我们的模型进一步能够自动标注出表面缺陷的位置或区域，并测量模型所需的计算量储存空间和标注精度。

3.2.4解决方案和结果

1. 图像的人工标记

为了使我们在问题一中训练出的MobileNetV2模型具有标注表面缺陷的位置或区域的能力，我们小队利用labelImg工具包，采用人工标记的方法对一些缺陷图片进行人工标记并使其产xml文件,改进问题1中分配训练集和测试集的代码，生成图片训练集的同时将其每张图片对应的xml文件放入对应的训练集标签文件包中，再利用代码将xml文件转为csv文件，最后再将csv标记文件内的信息转为tfrecord文件便于再次训练我们训练好的MobileNetV2模型。

1. 模型的再训练

通过修改模型对应的.config文件，再运用Tensorflow对应的API包中的运行文件包，对我们的 MobileNetV2模型进行了再训练，使其具有了更加精准的表面缺陷识别能力同时可以识别出表面缺陷的区域。

1. 训练模型的使用

向我们训练好的模型投入测试集的同时，我们得到了对应模型的checkpoint文件

我们再在tensorboard上查看测试图片的自动识别和标注情况。（这个还没有实现）

3.2.5结果分析

3.2.6优势与劣势

3.3模型三

3.3.1额外的变量

3.3.2额外的假设

3.3.3模型的基础

3.3.4解决方案和结果

3.3.5结果分析

3.3.6优势与劣势

1. 结论

4.1问题的结论（重述一些三个题的结果，然后得出结论）

4.1.1问题一的结论

4.1.2问题二的结论

4.1.3问题三的结论

4.2方法（文章用到了哪些数模方法）

* 。。。
* 。。。

4.3模型的运用（数模模型如何运用的）

1. 后期工作（目前不知道怎么写这一模块）

1）缺乏正则化技术，正则化技术如 L1/L2 正则化、dropout 等可以限制模型的复杂度，帮助减少模型过拟合。

5.1我们模型的推广运用

1. 参考文献（全英文）

[1] Domen Tabernik, Matic Sˇuc, and Danijel Skoˇcaj. Automated detection and segmentation of cracks in concrete surfaces using joined segmentation and classification deep neural network, Sep 2023.

[2] <https://www.vicos.si/resources/kolektorsdd/>.

[3] Domen Tabernik, Samo Sˇela, Jure Skvarˇc, and Danijel Skoˇcaj. Segmentation-based deep-learning approach for surface-defect detection, Mar 2020

[4]Yiting,Huang,Haisong,et al.Research on a Surface Defect Detection Algorithm Based on MobileNet-SSD[J].Applied Sciences, 2018.

[5] Zhang J , Jing J , Lu P ,et al.Improved MobileNetV2-SSDLite for automatic fabric defect detection system based on cloud-edge computing[J].Measurement, 2022.

[6] Aydin I , Akin E , Karakose M .Defect classification based on deep features for railway tracks in sustainable transportation[J].Applied Soft Computing, 2021, 111(7):107706.DOI:10.1016/j.asoc.2021.107706.

[7] 薛定谔的Keyboard.2020-07-11.基于MobileNetV2-SSD缺陷检测.CSDN.取自[基于MobileNetV2-SSD缺陷检测-CSDN博客](https://blog.csdn.net/weixin_42679015/article/details/107288438?ops_request_misc=&request_id=&biz_id=102&utm_term=mobilenetssd%E6%A3%80%E6%B5%8B&utm_medium=distribute.pc_search_result.none-task-blog-2~all~sobaiduweb~default-3-107288438.142%5ev96%5epc_search_result_base3&spm=1018.2226.3001.4187)

1. 附录（代码和数据）