



(12)发明专利申请

(10)申请公布号 CN 107844770 A

(43)申请公布日 2018.03.27

(21)申请号 201711071383.1

(22)申请日 2017.11.03

(71)申请人 东北大学

地址 110169 辽宁省沈阳市浑南区创新路
195号

(72)发明人 卢绍文 王克栋 郭章 王金鑫
李鹏琦 程盟盟 赵磊 刘晓丽
丁进良 王良勇 柴天佑

(74)专利代理机构 沈阳优普达知识产权代理事
务所(特殊普通合伙) 21234

代理人 李晓光

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

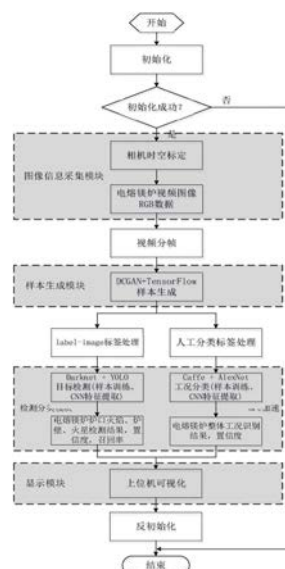
权利要求书2页 说明书10页 附图6页

(54)发明名称

一种基于视频的电熔镁炉异常工况自动识别系统

(57)摘要

本发明涉及一种基于视频的电熔镁炉异常工况自动识别系统,通过图像信息采集模块采集电熔镁砂生产现场工况信息,传输至样本生成模块;样本生成模块对训练样本进行生成,通过label-image软件对图像做标签处理和人工分类标签处理,将处理后的视频、图像信息传输到检测分类模块;检测分类模块通过相应算法对视频、图像信息进行特征提取及处理,得到工况识别系统的智能检测、分类模型;显示单元运用检测分类模块处理完毕的智能检测、分类模型,将新的测试视频、图像信息通过可视化方式显示出来。本发明识别效果相比单方面使用电流、电压等非可视化的信息有明显的改善,硬件装置构成合理,操作方便、成本低,判别准确率高,代替人工巡检。



1. 一种基于视频的电熔镁炉异常工况自动识别系统,其特征在于:

通过图像信息采集模块采集电熔镁砂生产现场工况信息,传输至样本生成模块;

样本生成模块对训练样本进行生成,通过label-image软件对图像做标签处理和人工分类标签处理,将处理后的视频、图像信息传输到检测分类模块;

检测分类模块通过相应算法对视频、图像信息进行特征提取及处理,得到工况识别系统的智能检测、分类模型;

显示单元运用检测分类模块处理完毕的智能检测、分类模型,将新的测试视频、图像信息通过可视化方式显示出来。

2. 按权利要求1所述的基于视频的电熔镁炉异常工况自动识别系统,其特征在于:样本生成模块用于在异常工况训练样本较少时,通过深度卷积对抗网络(DCGAN)+TensorFlow深度学习框架对异常工况较少的图像做样本生成,增加异常工况图像,再重新放回到训练样本中进行训练和智能建模,得到完整、平衡的训练样本。

3. 按权利要求2所述的基于视频的电熔镁炉异常工况自动识别系统,其特征在于:样本生成模块通过深度卷积对抗网络生成样本,深度卷积对抗网络是在生成式对抗网络的基础上增加卷积的环节,通过生成图片的生成式网络接收随机的噪声 z ,并将这个噪声生成图片,记做 $G(z)$;通过判别网络,判别一张图片是不是“真实的”,通过生成式网络生成的图片去欺骗判别网络,判别网络则将生成式网络生成的图片和真实的输入图片进行区分,通过以下公式计算最终期望值来生成图片:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

其中, G 为生成图片的网络, D 为判别网络, z 为输入 G 网络的噪声数据,符合 $p_z(z)$ 分布, $G(z)$ 为 G 网络生成的图片; x 代表真实数据,符合 $p_{data}(x)$ 分布,data表示数据, p 表示概率, $p_{data}(x)$ 为判别数据是否为真实数据的概率, $E_{x \sim p_{data}(x)}$ 为真实数据的期望, $E_{z \sim p_z(z)}$ 为是生成数据的期望; $V(D, G)$ 为判别网络和生成图片的网络的博弈结果输出值。

4. 按权利要求1所述的基于视频的电熔镁炉异常工况自动识别系统,其特征在于:检测分类模块接收label-image软件对图像做标签处理后得到的带有火焰区域、欠烧区域、火星区域的位置信息与工况类别信息的xml文件;通过Darknet框架+YOLO算法把做过标签处理的镁炉原始RGB图像、位置和工况类别信息做训练,得到火焰区域、欠烧区域、火星区域的位置信息和工况类别信息的智能模型,以及体现模型准确率的置信度以及召回率,最后利用训练得到的智能模型实现在线视频中电熔镁炉的工况识别结果。

5. 按权利要求4所述的基于视频的电熔镁炉异常工况自动识别系统,其特征在于:通过Darknet框架+YOLO算法把做过标签处理的镁炉原始RGB图像、位置和工况类别信息做训练,具体为:

将一幅图像分成 $S \times S$ 个网格,如果目标object的中心落在这个网格中,则这个网格就负责预测该目标object;

每个网格针对所含目标object预测 B 个边界框bounding box的坐标 (x, y, w, h) ,其中,对应边界框bounding box中心的坐标 (x, y) 、宽度值 w 和高度值 h 归一化到0-1之间;

边界框bounding box附带预测一个置信度confidence值,代表所预测的bounding box中含有目标object的置信度和这个边界框bounding box预测的准确率两重信息,其值通过

以下公式计算：

$$confidence = Pr(Object) * IoU_{pred}^{truth};$$

其中, $Pr(Object)$ 为包含目标的概率, IoU_{pred}^{truth} 为系统预测出来的框与原来图片中标记的框的重合程度; $truth$ 代表目标的实际值或实际位置, $pred$ 即 $predict$, 代表目标的预测值或预测位置;

如果有目标 $object$ 落在这个网格里, 则 $Pr(Object)$ 取1, 否则 $Pr(Object)$ 取0;

预测的边界框 $bounding\ box$ 和实际的地面真值 $ground\ truth$ 之间的 IoU 值即系统预测出来的框与原来图片中标记的框的重合程度, 计算方法即检测结果 $bounding\ box$ 与原样本图片中标记的真值 $ground\ truth$ 的交集比上它们的并集, 即为检测的准确率:

$$IoU = \frac{bounding_box \cap ground_truth}{bounding_box \cup ground_truth}。$$

6. 按权利要求5所述的基于视频的电熔镁炉异常工况自动识别系统, 其特征在于:

对于 $S*S$ 个网格, 每个网格预测 C 种类别, 输出为 $S*S*(5*B+C)$ 的一个数据; 其中, B 为目标预测的边界框 $bounding\ box$ 的个数, 5 为边界框 $bounding\ box$ 的坐标 (x, y, w, h) 和置信值 $confidence$ 这5种数据;

在测试的时候, 每个网格预测的类别信息和 $bounding\ box$ 预测的置信度 $confidence$ 信息相乘, 就得到每个 $bounding\ box$ 的特定类别的置信得分 $class-specific\ confidence\ score$, 公式如下:

$$Pr(Class_i | Object) * Pr(Object) * IoU_{pred}^{truth} = Pr(Class_i) * IoU_{pred}^{truth}$$

$Pr(Class_i | Object)$ 为每个网格预测的类别信息, $Pr(Object)$ 为每个 $bounding\ box$ 预测的 $confidence$, $Pr(Class_i) * IoU_{pred}^{truth}$ 表示预测的 $bounding\ box$ 属于某一类的概率;

在得到每个 $bounding\ box$ 的 $class-specific\ confidence\ score$ 以后, 设置阈值, 滤掉得分低的多个 $bounding\ boxes$, 对保留的得分高的几个 $bounding\ boxes$ 再进行非极大值抑制处理, 得到最终的检测结果。

7. 按权利要求1所述的基于视频的电熔镁炉异常工况自动识别系统, 其特征在于: 检测分类模块接收人工分类标签处理后的图像信息, 应用Caffe深度学习框架结合基于CNN的AlexNet网络模型, 对单帧整张图像进行CNN特征提取, 并利用提取到的图像特征做训练, 得到单帧图像工况分类智能模型和置信度, 最后利用训练得到的智能模型实现电熔镁炉单帧图像的工况识别结果。

8. 按权利要求1所述的基于视频的电熔镁炉异常工况自动识别系统, 其特征在于: 样本生成模块和检测分类模块由计算机软件实现, 该计算机软件还包括相机时空标定子模块, 对图像信息采集模块中的相机在使用前进行一次性标定, 对原始数据进行预处理, 实现对相机传感器内在误差、硬件系统外部误差以及数据采集的时间误差的校正。

一种基于视频的电熔镁炉异常工况自动识别系统

技术领域

[0001] 本发明涉及一种模式识别与人工智能技术领域中的设备工况识别系统,具体为一种基于视频的电熔镁炉异常工况自动识别系统。

背景技术

[0002] 电熔镁砂具有熔点高、结构致密、抗氧化性强、耐压强度高、耐腐蚀性强、化学性能稳定等特性,是重要的战略性原材料。

[0003] 在我国,高品位的电熔镁砂的制备主要是通过电弧炉熔炼再结晶来完成。原料氧化镁砂在炉中熔炼的过程可人为地分为:启炉、加料、正常冶炼、欠烧和排气异常等工况。其中,欠烧和排气异常工况属于异常工况。欠烧工况是由于在某段时间部分生料燃烧不充分而没有完全达到熔融状态,致使熔池内电阻低于正常值,使得熔池内电流过高,产生的电弧过于强烈,进而导致炉体温度局部过高引起。倘若未能及时处理,将导致炉壁烧穿,造成巨大经济损失,高温熔液泄露还可能危及现场工人的安全;排气异常工况是由于某段时间部分生料熔融后在熔池内形成巨大的气泡,在炉体进行排气的时候,气泡在炉口破裂,会导致熔融的电熔镁飞溅而出,高温的火星会危及现场工人的安全。

[0004] 目前,电熔镁炉异常工况的预警主要依靠现场工人的不间断巡检。工人通过观察生产过程中炉口火焰的形态、亮度、火星,以及炉壁是否有烧红区域等特征,根据经验来预先判断可能发生的异常工况。人工巡检的主要问题是:1)判断的准确性与操作人员的经验和状态的相关,容易漏检、误检;2)现场生产环境恶劣(强光、高温、灰尘、蒸汽等),劳动强度大,危险性高,不适于工人长时间现场巡检。因此,企业急需智能的异常工况判别技术,能够稳定、准确地在异常工况刚发生的时候预警,把人从这项工作中解放出来。

[0005] 以前的自动识别技术是通过实时采集的电极电流、电压来判断,但是准确性较低,可视化程度弱,尚无法代替人工观察。这是因为现场工况的可视化特征仍是判断异常工况最快捷和可靠的依据,电流电压的波动只能辅助判断。

发明内容

[0006] 针对现有技术中电熔镁炉异常工况的预警主要依靠现场工人的不间断巡检实现,具有漏检、误检以及危险性高等不足,本发明要解决的问题是提供一种可提高电熔镁生产品质、可视化程度高、降低工人劳动强度的基于视频的电熔镁炉异常工况自动识别系统。

[0007] 为解决上述技术问题,本发明采用的技术方案是:

[0008] 本发明一种基于视频的电熔镁炉异常工况自动识别系统,通过图像信息采集模块采集电熔镁砂生产现场工况信息,传输至样本生成模块;

[0009] 样本生成模块对训练样本进行生成,通过label-image软件对图像做标签处理和人工分类标签处理,将处理后的视频、图像信息传输到检测分类模块;

[0010] 检测分类模块通过相应算法对视频、图像信息进行特征提取及处理,得到工况识别系统的智能检测、分类模型;

[0011] 显示单元运用检测分类模块处理完毕的智能检测、分类模型,将新的测试视频、图像信息通过可视化方式显示出来。

[0012] 样本生成模块用于在异常工况训练样本较少时,通过深度卷积对抗网络(DCGAN)+TensorFlow深度学习框架对异常工况较少的图像做样本生成,增加异常工况图像,再重新放回到训练样本中进行训练和智能建模,得到完整、平衡的训练样本。

[0013] 样本生成模块通过深度卷积对抗网络生成样本,深度卷积对抗网络是在生成式对抗网络的基础上增加卷积的环节,通过生成图片的生成式网络接收随机的噪声 z ,并将这个噪声生成图片,记做 $G(z)$;通过判别网络,判别一张图片是不是“真实的”,通过生成式网络生成的图片去欺骗判别网络,判别网络则将生成式网络生成的图片和真实的输入图片进行区分,通过以下公式计算最终期望值来生成图片:

$$[0014] \quad \min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

[0015] 其中, G 为生成图片的网络, D 为判别网络, z 为输入 G 网络的噪声数据,符合 $p_z(z)$ 分布, $G(z)$ 为 G 网络生成的图片; x 代表真实数据,符合 $p_{data}(x)$ 分布, $data$ 表示数据, p 表示概率, $p_{data}(x)$ 为判别数据是否为真实数据的概率, $E_{x \sim p_{data}(x)}$ 为真实数据的期望, $E_{z \sim p_z(z)}$ 为是生成数据的期望; $V(D, G)$ 为判别网络和生成图片的网络的博弈结果输出值。

[0016] 检测分类模块接收label-image软件对图像做标签处理后得到的带有火焰区域、欠烧区域、火星区域的位置信息与工况类别信息的xml文件;通过Darknet框架+YOLO算法把做过标签处理的镁炉原始RGB图像、位置和工况类别信息做训练,得到火焰区域、欠烧区域、火星区域的位置信息和工况类别信息的智能模型,以及体现模型准确率的置信度以及召回率,最后利用训练得到的智能模型实现在线视频中电熔镁炉的工况识别结果。

[0017] 通过Darknet框架+YOLO算法把做过标签处理的镁炉原始RGB图像、位置和工况类别信息做训练,具体为:

[0018] 将一幅图像分成 $S \times S$ 个网格,如果目标object的中心落在这个网格中,则这个网格就负责预测该目标object;

[0019] 每个网格针对所含目标object预测 B 个边界框bounding box的坐标 (x, y, w, h) ,其中,对应边界框bounding box中心的坐标 (x, y) 、宽度值 w 和高度值 h 归一化到0-1之间;

[0020] 边界框bounding box附带预测一个置信度confidence值,代表所预测的bounding box中含有目标object的置信度和这个边界框bounding box预测的准确率两重信息,其值通过以下公式计算:

$$[0021] \quad confidence = Pr(Object) * IoU_{pred}^{truth};$$

[0022] 其中, $Pr(Object)$ 为包含目标的概率, IoU_{pred}^{truth} 为系统预测出来的框与原来图片中标记的框的重合程度; $truth$ 代表目标的实际值或实际位置, $pred$ 即predict,代表目标的预测值或预测位置;

[0023] 如果有目标object落在这个网格里,则 $Pr(Object)$ 取1,否则 $Pr(Object)$ 取0;

[0024] 预测的边界框bounding box和实际的地面真值ground truth之间的IoU值即系统预测出来的框与原来图片中标记的框的重合程度,计算方法即检测结果bounding box与原样本图片中标记的真值ground truth的交集比上它们的并集,即为检测的准确率:

$$IoU = \frac{bounding_box \cap ground_truth}{bounding_box \cup ground_truth}。$$

[0025] 对于S*S个网格,每个网格预测C种类别,输出为S*S*(5*B+C)的一个数据;其中,B为目标预测的边界框bounding box的个数,5为边界框bounding box的坐标(x,y,w,h)和置信值confidence这5种数据;

[0026] 在测试的时候,每个网格预测的类别信息和bounding box预测的置信度confidence信息相乘,就得到每个bounding box的特定类别的置信得分class-specific confidence score,公式如下:

$$[0027] \quad \Pr(Class_i | Object) * \Pr(Object) * IoU_{pred}^{truth} = \Pr(Class_i) * IoU_{pred}^{truth}$$

[0028] $\Pr(Class_i | Object)$ 为每个网格预测的类别信息, $\Pr(Object)$ 为每个bounding box预测的confidence, $\Pr(Class_i) * IoU_{pred}^{truth}$ 表示预测的bounding box属于某一类的概率;

[0029] 在得到每个bounding box的class-specific confidence score以后,设置阈值,滤掉得分低的多个bounding boxes,对保留的得分高的几个bounding boxes再进行非极大值抑制处理,得到最终的检测结果。

[0030] 检测分类模块接收人工分类标签处理后的图像信息,应用Caffe深度学习框架结合基于CNN的AlexNet网络模型,对单帧整张图像进行CNN特征提取,并利用提取到的图像特征做训练,得到单帧图像工况分类智能模型和置信度,最后利用训练得到的智能模型实现电熔镁炉单帧图像的工况识别结果。

[0031] 样本生成模块和检测分类模块由计算机软件实现,该计算机软件还包括相机时空标定子模块,对图像信息采集模块中的相机在使用前进行一次性标定,对原始数据进行预处理,实现对相机传感器内在误差、硬件系统外部误差以及数据采集的时间误差的校正。

[0032] 本发明具有以下有益效果及优点:

[0033] 1.本发明提出并实现了基于视频的电熔镁炉异常工况自动识别系统,该系统能够实时在线的对电熔镁砂生产现场进行工况识别,识别效果相比单方面使用电流、电压等非可视化的信息有明显的改善,并对本发明进行了工业验证,证明了该系统设计的有效性。硬件装置构成合理、简洁,可以手持,也可以搭载在其他硬件平台(例如机器人或其他设备)上;上位机平台使用方便,操作快捷。

[0034] 2.运用本发明方法可以快速地进行电熔镁炉的工况识别,不仅操作方便、成本低廉,而且判别准确率高,可以代替人工巡检。

附图说明

[0035] 图1为发明基于视频的电熔镁炉异常工况自动识别系统流程图;

[0036] 图2为本发明识别系统的总结构框图;

[0037] 图3为本发明识别系统的相机标定算法示意图;

[0038] 图4为本发明识别系统的DCGAN深度卷积对抗网络算法示意图;

[0039] 图5为本发明基于视频的电熔镁炉异常工况自动识别系统的DCGAN深度卷积对抗网络生成训练样本效果示意图;

[0040] 图6为本发明基于视频的电熔镁炉异常工况自动识别系统的YOLO算法示意图;

[0041] 图7为本发明基于视频的电熔镁炉异常工况自动识别系统的最终工况识别效果示意图。

具体实施方式

[0042] 下面结合说明书附图对本发明作进一步阐述。

[0043] 本发明提出一种基于视频的电熔镁炉异常工况自动识别系统,本系统结合现场工人经验建立对异常工况视频图像的检测和分类模型,利用工业相机获取电熔镁砂生产现场过程图像,通过实时的图像分析,实现对异常工况在线识别。本发明对于提高电熔镁生产品质、增强可视化程度、降低工人劳动强度有积极意义。

[0044] 如图1所示,本发明一种基于视频的电熔镁炉异常工况自动识别系统,通过图像信息采集模块采集电熔镁生产现场工况信息,传输至样本生成模块;

[0045] 样本生成模块对训练样本进行生成,通过label-image软件对图像做标签处理和人工分类标签处理,将处理后的视频、图像信息传输到检测分类模块;

[0046] 检测分类模块通过相应算法对视频、图像信息进行特征提取及处理,得到工况识别系统的智能检测、分类模型;

[0047] 显示单元运用检测分类模块处理完毕的智能检测、分类模型,将新的测试视频、图像信息通过可视化方式显示出来。

[0048] 样本生成模块用于在异常工况训练样本较少时,通过深度卷积对抗网络(DCGAN)+TensorFlow深度学习框架对异常工况较少的图像做样本生成,增加异常工况图像,再重新放回到训练样本中进行训练和智能建模,得到完整、平衡的训练样本。

[0049] 检测分类模块接收label-image软件对图像做标签处理后得到的带有火焰区域、欠烧区域、火星区域的位置信息与工况类别信息的xml文件;通过Darknet框架+YOLO算法把做过标签处理的镁炉原始RGB图像、位置和工况类别信息做训练,得到火焰区域、欠烧区域、火星区域位置信息和工况类别信息的智能模型,以及体现模型准确性的置信度和召回率;最后利用训练得到的智能模型实现视频中电熔镁炉的工况识别结果。检测分类模块还接收人工分类标签处理后的图像信息,应用Caffe深度学习框架结合基于CNN的AlexNet网络模型,对单帧整张图像进行CNN特征提取,并利用提取到的图像特征做训练,得到单帧图像工况分类智能模型和置信度,最后利用训练得到的智能模型实现电熔镁炉单帧图像的工况识别结果。

[0050] 样本生成模块和检测分类模块由计算机软件实现,该计算机软件还包括相机时空标定子模块,对图像信息采集模块中的相机在使用前进行一次性标定,对原始数据进行预处理,实现对相机传感器内在误差、硬件系统外部误差以及数据采集的时间误差的校正。

[0051] 本发明基于视频的电熔镁炉异常工况自动识别系统的硬件平台包括:工业相机、相机三脚架、以太网数据线以及带GPU加速的PC机。

[0052] 工业相机构成了图像信息采集单元,实现对电熔镁生产现场工况信息的采集;相机三脚架为固定装置,实现对工业相机的固定与支撑;以太网数据线为信息传输单元,实现对工业相机采集的电熔镁生产现场图像信息的传输;PC机为主处理器,利用所包含的GPU器件实现对数据的处理与软件的运行;GPU为图形处理单元,其加速技术可以应用在系统和算法中,对于海量数据的处理有着巨大的优势。

[0053] 本发明中,算法设计与软件实现包括:图像信息采集模块、样本生成模块、检测分类模块、以及显示模块,其中:

[0054] 图像信息采集模块包括相机时空标定模块,通过原始数据的预处理过程实现对相机传感器内在误差、硬件系统外部误差以及数据采集的时间误差的校正;

[0055] 样本生成模块用于在异常工况训练样本较少时,用生成对抗网络(GAN)+TensorFlow深度学习框架对训练样本进行生成,以增加训练样本数量,训练出更好的模型,使工况识别效果更加准确。

[0056] 检测分类模块包括实时检测和定位炉口火焰、炉壁欠烧区域、排气异常溅出的火星在整个电熔镁生产过程中的位置信息、工况类别信息以及在整个监控视频过程中对实时工况的识别过程两个处理流程,首先,用Darknet框架+YOLO算法对实时目标进行检测和初步识别。然后通过Caffe深度学习框架+AlexNet模型对单帧整张图像进行特征提取,并利用提取到的图像特征实现整个判别过程,增强工况识别效果。

[0057] 基于视频的电熔镁炉异常工况自动识别系统的总结构框如图2所示,本实施例中,系统的硬件平台包括:信息采集模块以及附带GPU的PC机。整个算法以基于深度学习算法领域中的卷积神经网络(CNN)开源算法为基础,应用大量的开源框架对此算法进行改进和应用,并且对算法的软件实现进行改写并将其移植到Linux操作系统下的Ubuntu操作系统平台上,实现含有GAN(生成对抗网络)的电熔镁炉工况判别系统。

[0058] 整个系统包括硬件平台、软件算法、上位机界面三部分。

[0059] 工业相机负责采集目标物体(电熔镁炉)的RGB信息,通过以太网数据线连接计算机,并把采集到的信息传输到计算机中,计算机通过GPU的加速技术,对采集到的信息运用现有的软件算法进行处理,把处理完毕的信息用可视化方式显示。

[0060] 图像信息采集模块由一个轻重量(小于500克)的Point Grey工业相机和一个轻质高强度的相机三脚架连接在一起。工业相机用于实时采集目标物体的RGB信息,并将信息发送至计算机。作为操作者只需要将该装置平稳地放在场景中,并将传感器朝向目标物体,传感器从场景捕获的实时图像信息便可以通过算法与软件在监控视频中产生精确的工况判别效果。并且可以标记出炉口火焰、炉壁欠烧、炉外火星的位置和区域,同时在标记出的位置信息的左上角输出分类的标签。

[0061] 本实施方式中,红外热像仪采用的是Point Grey工业相机(FL3-GE-50S5C-/M-C);以太网数据线用的是传输速度可达到每秒千兆的以太网数据线。

[0062] PC机为主处理器,用于获取工业相机采集到的基本信息,将传感器采集到的RGB信息通过Point Grey公司自带的FlyCapture SDK软件传输到计算机内,并进行保存。通过计算机应用软件和相应算法进行处理,最后把处理完毕的信息通过可视化方式显示出来。

[0063] GPU(本实施例采用的型号为NVIDIA GTX1080。)作为图形处理单元,对整个系统进行加速以达到实时性。由于传感器的输入分辨率高达1920*1080,因此每一帧的数据量都将十分巨大,如果仅仅使用CPU进行顺序操作,那将耗费巨大的运算时间和内存在海量数据的处理与分析上,对于系统高效实时的运行十分不利。而系统输入的数据是以像素为单位,它们由相机坐标系下的像素坐标进行索引,互相之间并没有相互联系,因此我们可以以每一个像素为单位对数据的坐标变换等运算,然后再将这些运算整合到一起进行下一步处理。由于算法执行可以满足的这个特性,因此我们可以利用GPU的并行计算进行加速。对于每一

帧输入的图像数据我们从坐标原点开始依次提取出像素坐标,根据图像的分辨率得到需要的GPU线程数,根据提取出的像素坐标赋予对应的线程号,在线程中对每一个线程的像素点,提取出RGB值等,对其进行坐标变换等操作,最后将所有线程得到的运算结果再赋给图像对应坐标中,完成并行加速计算。

[0064] 软件算法部分包括:相机标定算法、深度卷积对抗网络(DCGAN)算法、YOLO目标检测算法以及卷积神经网络(CNN)算法。

[0065] 相机参数的标定,在本发明中用于实现对相机传感器内在误差、硬件系统外部误差以及数据采集的时间误差的校正,即把传感器采集到的初始数据进行预处理(时空标定)。在图像识别领域以及机器视觉成像的测量过程中,为了使得空间中的某个点对应于各个坐标系下的三维坐标能与其在成像平面下图像中相对应像素点的联系,首先要做的是必须建立其相机成像物理模型所对应的数学模型,而这些模型的建立过程中不可或缺参数就是相机参数。由于不同的相机结构工艺各有不同,因此这些参数必须要通过进行一些实验的测试才能获得。

[0066] 相机数学模型为:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x & 0 & 0 & 0 \\ 0 & f_y & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

[0067]

$$= \begin{bmatrix} \alpha & 0 & u_0 & 0 \\ 0 & \beta & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_1 M_2 \overline{X_w} = M \overline{X_w}$$

[0068] 其中,u,v表示的是图像坐标系(u,v),dx、dy为转换因子,Z_c表示的是相机坐标系(X_c,Y_c,Z_c)中的Z轴,X_w,Y_w,Z_w表示的是世界坐标系(X_w,Y_w,Z_w),f_x、f_y、R、T、α、β为待标定参数,M₁为相机数学模型内参数矩阵,M₂为外参数矩阵, $\overline{X_w}$ 代表矩阵[X_w,Y_w,Z_w]^T,u₀,v₀为像素平面坐标系(u,v)的中心,0^T为0矩阵的转置,旋转矩阵R和平移矩阵,t为相机的外部参数,R为3*3矩阵,t为3*1矩阵,0为(0,0,0)。

[0069] 选择合适的方法对于相机标定结果的准确性和稳定性有着重要意义,同时也会直接的影响到相机在整个系统工作时所产生结果的准确性。相机的标定包括确定内参数的内部标定,确定外参数的外部标定,由于我们将两种传感器采集的数据作为输入量,因此还应包含确定两者输入信息时间一致性的时间同步标定。

[0070] 本实施例中,本发明采用的相机标定方法为图3所示的国际上最常用的“张正友棋盘标定法”,标定过程在Linux系统下完成。

[0071] 深度卷积对抗网络(DCGAN)算法,用于对异常工况较少的图像做样本生成,增加异常工况图像,得到完整、平衡的训练样本(启炉、加料、正常、欠烧、排气异常5种工况)。

[0072] 样本生成模块中,DCGAN是基于GAN(生成式对抗网络)的基础上,增加了卷积的环节。GAN有两个网络,G(Generator)和D(Discriminator),其结构如图4所示。G是一个生成图片的网络,它接收一个随机的噪声z,通过这个噪声生成图片,记做G(z)。D是一个判别网络,

判别一张图片是不是“真实的”。它的输入参数是 x , x 代表一张图片, 输出 $D(x)$ 代表 x 为真实图片的概率, 如果为1, 就代表100%是真实的图片, 而输出为0, 就代表不可能是真实的图片。在训练过程中, 生成网络 G 的目标就是尽量生成真实的图片去欺骗判别网络 D 。而 D 的目标就是尽量把 G 生成的图片和真实的图片分别开来。这样, G 和 D 构成了一个动态的“博弈过程”。在最理想的状态下, G 可以生成足以“以假乱真”的图片 $G(z)$ 。对于 D 来说, 它难以判定 G 生成的图片究竟是不是真实的, 因此 $D(G(z)) = 0.5$ 。这样我们的目的就达成了: 得到了一个生成式的模型 G , 它可以用来生成图片。其核心原理如公式所示:

$$[0073] \quad \min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

[0074] 其中, x 表示真实图片, z 表示输入 G 网络的噪声, 而 $G(z)$ 表示 G 网络生成的图片。 $D(x)$ 表示 D 网络判断真实图片是否真实的概率(因为 x 就是真实的, 所以对于 D 来说, 这个值越接近1越好)。而 $D(G(z))$ 是 D 网络判断 G 生成的图片的是否真实的概率。 G 的目的: 上面提到, $D(G(z))$ 是 D 网络判断 G 生成的图片是否真实的概率, G 应该希望自己生成的图片“越接近真实越好”。也就是说, G 希望 $D(G(z))$ 尽可能得大, 这时生成图片的网络 G 和判别网络 D 的博弈结果输出值 $V(D, G)$ 会变小。因此我们看到式子的最前面的记号是 \min_G 。而 D 的目的: D 的能力越强, $D(x)$ 应该越大, $D(G(x))$ 应该越小。这时 $V(D, G)$ 会变大。因此式子对于 D 来说是求最大 \max_D 。

[0075] DCGAN的原理和GAN是一样的, 它只是把上述的 G 和 D 换成了两个卷积神经网络(CNN)。但并不是直接, DCGAN对卷积神经网络的结构做了一些改变, 以提高样本的质量和收敛的速度, 这些改变有: 1、取消所有pooling层: G 网络中使用转置卷积(transposed convolutional layer)进行上采样, D 网络中用加入stride的卷积代替池化; 2、在 D 和 G 中均使用归一化(batch normalization); 3、去掉全链接(FC)层, 使网络变为全卷积网络; 4、 G 网络中使用ReLU作为激活函数, 最后一层使用tanh函数。5、 D 网络中使用LeakyReLU作为激活函数。

[0076] 本实施方式中, 采用TensorFlow深度学习框架来实现DCGAN, 对异常工况较少的图像做样本生成, 增加异常工况图像, 最终得到完整、平衡的训练样本。生成效果如图5所示。

[0077] 本实施例中, TensorFlow为一种深度学习框架, 是谷歌公司研发的第二代人工智能学习系统。Tensor(张量)意味着 N 维数组, Flow(流)意味着基于数据流图的计算, TensorFlow为张量从流图的一端流动到另一端计算过程。TensorFlow是将复杂的数据结构传输至人工智能神经网络中进行分析 and 处理过程的系统。

[0078] 本实施例中, 开发平台在基于Linux的Ubuntu操作系统下完成, 编程语言为Python语言, 其优点为: 优良的跨平台性、可面向对象、丰富的API。

[0079] YOLO目标检测算法, 用于把做过标签处理的现场原始RGB图像、目标位置、工况类别信息(.xml文件)做训练, 得到目标位置信息、工况类别信息的智能模型。并应用YOLO算法得到的智能模型, 得到视频中的实时工况判别效果。其核心思想就是利用整张图作为网络的输入, 直接在输出层回归bounding box的位置和bounding box所属的类别。

[0080] 先前的检测系统多使用分类器(classifier)或者定位器(localizer)进行检测任务, 把图片的不同局部位置和多种尺度输入到模型中去, 图片得分较高的区域(region)作为检测目标。YOLO是一个全新的方法, 把一整张图片一下子应用到一个神经网络中去。网络把图片分成不同的区域, 然后给出每个区域的边框预测和概率, 并依据概率大小对所有边

框分配权重。最后,设置阈值,只输出得分(概率值)超过阈值的检测结果。

[0081] 本实施例中,具体实现方法为:首先,将一幅图像分成 $S \times S$ 个网格(grid cell),如果某个目标(object)的中心落在这个网格中,则这个网格就负责预测这个object。每个网格要针对于所含object预测 B 个边界框(bounding box)的坐标 (x, y, w, h) ,其中:对应边界框bounding box中心坐标的 (x, y) 、宽(w)和高(h)归一化到0-1之间。此外,bounding box除了要回归自身的位置之外,还要附带预测一个置信度(confidence)值。这个confidence代表了所预测的bounding box中含有object的置信度和这个box预测的准确率两重信息,其值是这样计算的: $confidence = Pr(Object) * IoU_{pred}^{truth}$ 。其中如果有object落在这个网格中,第一项取1,否则取0。第二项是预测的bounding box和实际真值(ground truth)之间的IoU值(系统预测出来的框与原来图片中标记的框的重合程度,计算方法即检测结果bounding box与原来图片中标记的框(ground truth)的交集比上它们的并集,即为检测的准确率:

$$IoU = \frac{bounding_box \cap ground_truth}{bounding_box \cup ground_truth}$$
)。然后,这 $S \times S$ 个网格,每个网格还要预测 C 种类别(本

发明把电熔镁炉工况分为5类)。输出就是 $S \times S \times (5 \times B + C)$ 的一个数据。然后,在测试的时候,每个网格预测的类别(class)信息和bounding box预测的置信度(confidence)信息相乘,就得到每个bounding box的特定类别的置信得分(class-specific confidence score):

[0082] $Pr(Class_i | Object) * Pr(Object) * IoU_{pred}^{truth} = Pr(Class_i) * IoU_{pred}^{truth}$

[0083] 上述等式左边第一项就是每个网格预测的类别信息,第二、三项就是每个bounding box预测的confidence。这个乘积即表示了预测的bounding box属于某一类的概率,也有该box准确度的信息。最后,在得到每个box的class-specific confidence score以后,设置阈值,滤掉得分低的boxes,对保留的boxes进行非极大值抑制(NMS)处理,就得到最终的检测结果。

[0084] 本发明应用label-image软件对图像做打标签处理,得到带有火焰区域、欠烧区域、火星区域的位置信息与工况类别信息的xml文件;label-image作为在深度学习领域中的一种图片标注工具,其脚本语言为Python语言,适用于Windows、Linux以及iOS操作系统;应用Darknet深度学习框架,实现YOLO目标检测算法以及标签的训练过程,详细流程如图6所示。Darknet作为实现YOLO算法的一种开源神经网络代码框架,其脚本是用C语言和CUDA编写。它的优点是快速、易于安装和支持CPU和GPU计算。

[0085] 为了实现实时性,对YOLO标签的训练采用GPU图形处理单元来实现。

[0086] GPU的密集型数据处理和并行数据计算问题在CUDA(Compute Unified Device Architecture)并行计算架构中完成,并利用cuDNN计算库,有效提高整个网络的训练速度。

[0087] 开发平台在Windows7操作系统下完成,编程语言为C语言和Python,其优点为:易操作、可移植性强、丰富的API。

[0088] 卷积神经网络(Convolutional Neural Network, CNN),是一种前馈神经网络,它的人工神经元可以响应一部分覆盖范围内的周围单元。现在,CNN已经成为众多科学领域的研究热点之一,特别是在模式分类领域,由于该网络避免了对图像的复杂前期预处理,可以直接输入原始图像,因而得到了更为广泛的应用。

[0089] 一般地,CNN的基本结构包括两层,其一为特征提取层,每个神经元的输入与前一

层的局部接受域相连,并提取该局部的特征。一旦该局部特征被提取后,它与其它特征间的位置关系也随之确定下来;其二是特征映射层,网络的每个计算层由多个特征映射组成,每个特征映射是一个平面,平面上所有神经元的权值相等。特征映射结构采用影响函数核小的sigmoid函数作为卷积网络的激活函数,使得特征映射具有位移不变性。此外,由于一个映射面上的神经元共享权值,因而减少了网络自由参数的个数。卷积神经网络中的每一个卷积层都紧跟着一个用来求局部平均与二次提取的计算层,这种特有的两次特征提取结构减小了特征分辨率。由于CNN的特征检测层通过训练数据进行学习,所以在使用CNN时,避免了显示的特征抽取,而隐式地从训练数据中进行学习;再者由于同一特征映射面上的神经元权值相同,所以网络可以并行学习,这也是卷积网络相对于神经元彼此相连网络的一大优势。卷积神经网络以其局部权值共享的特殊结构在语音识别和图像处理方面有着独特的优越性,其布局更接近于实际的生物神经网络,权值共享降低了网络的复杂性,特别是多维输入向量的图像可以直接输入网络这一特点避免了特征提取和分类过程中数据重建的复杂度。单层神经网络模型公式为:

$$[0090] \quad a_1^{(2)} = f(W_{11}^{(1)}x_1 + W_{12}^{(1)}x_2 + W_{13}^{(1)}x_3 + b_1^{(1)})$$

$$[0091] \quad a_2^{(2)} = f(W_{21}^{(1)}x_1 + W_{22}^{(1)}x_2 + W_{23}^{(1)}x_3 + b_2^{(1)})$$

$$[0092] \quad a_3^{(2)} = f(W_{31}^{(1)}x_1 + W_{32}^{(1)}x_2 + W_{33}^{(1)}x_3 + b_3^{(1)})$$

$$[0093] \quad h_{w,b}(x) = a_1^{(3)} = f(W_{11}^{(2)}a_1^{(2)} + W_{12}^{(2)}a_2^{(2)} + W_{13}^{(2)}a_3^{(2)} + b_1^{(2)})$$

[0094] 神经网络的参数 $(W, b) = (W^{(1)}, b^{(1)})$, 其中, W_{ij}^l 是第 l 层第 j 单元与第 $l+1$ 层第 i 单元之间连接参数(权重), $b_i^{(l)}$ 是第 $l+1$ 层第 i 单元的偏置项。 $a_i^{(l)}$ 表示第 i 单元的激活值(输出值)。当 $l=1$ 时, $a_i^{(1)} = x_i$, 也就是第 i 个输入值(输入值的第 i 个特征)。对于给定参数集合 W, b , 神经网络就可以按照函数 $h_{w,b}(x)$ 来计算输出结果。

[0095] 本发明应用的卷积神经网络为图像分类模型 AlexNet 网络结构, 其包括 6 个卷积层 (Convolution Layer)、3 个池化层 (Pooling Layer)、3 个全链接层 (Full-connected Layer), 以及一个 Softmax 层。Softmax 层的作用为将神经元的输出变成概率的形式。

[0096] 应用 Caffe 深度学习框架运行 AlexNet 模型, 实现对单帧整张图像进行特征提取, 并利用提取到的图像特征实现单帧图像的工况识别过程, 增强识别效果。

[0097] Caffe 全称 Convolutional Architecture for Fast Feature Embedding, 是一个清晰、可读性高、快速的深度学习框架。为了实现快速性, 对样本在 Caffe 中的训练采用 GPU 图像处理单元来实现。

[0098] GPU 处理密集型数据和并行数据计算问题在 CUDA 并行计算架构中完成, 并利用 cuDNN 计算库, 有效提高整个网络的训练速度。

[0099] 可视化窗口将监控视频下的镁炉工况识别效果固定的显示在界面上, 便于技术人员观察电熔镁炉的当前工况识别的准确性以及实时报警情况。显示效果如图 7 所示。

[0100] 本发明中, 采用工业相机进一步搭建了实时图像采集装置; 用 label-image 软件, 对训练样本做标签处理; 利用基于深度学习的检测和识别方法中的 YOLO (You Only Look Once) 算法, 设计了适用于工厂级环境的火焰和炉壁区域的实时检测和工况识别方法, 并利用 Caffe 深度学习框架+AlexNet 模型对单帧的整张现场图像进行识别, 对工况判别结果进

行增强;针对异常工况训练样本不足的问题,利用深度卷积对抗网络(DCGAN)+TensorFlow深度学习框架对异常工况的样本进行生成,再重新放回到训练样本中进行训练和智能建模。用C、C++和Python语言作为开发工具,开发了算法和应用软件。最终完成了基于视频的电熔镁炉异常工况自动识别系统。此方法不仅操作方便、成本低廉,而且判别准确率高,可以代替人工巡检。

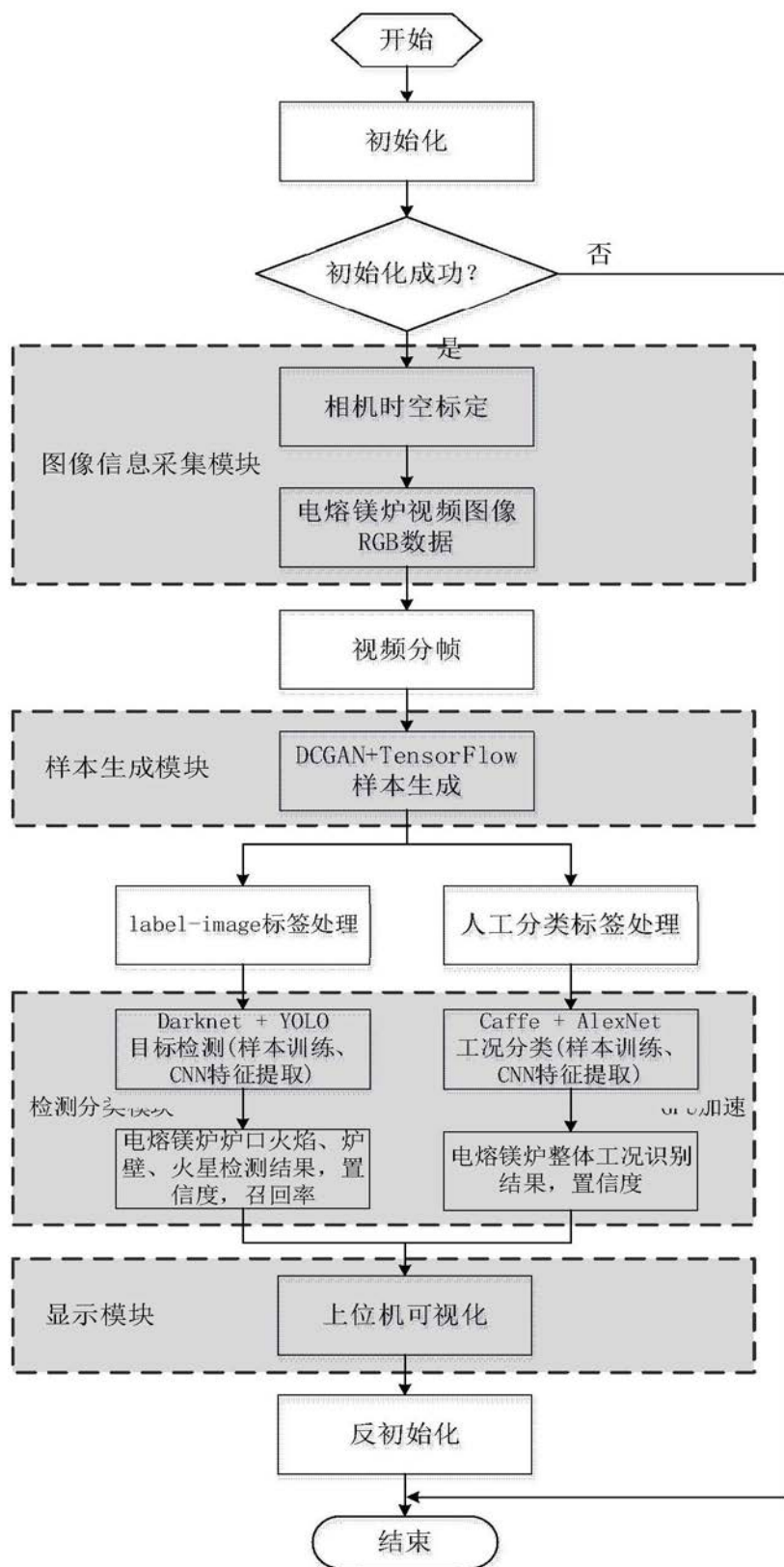


图1

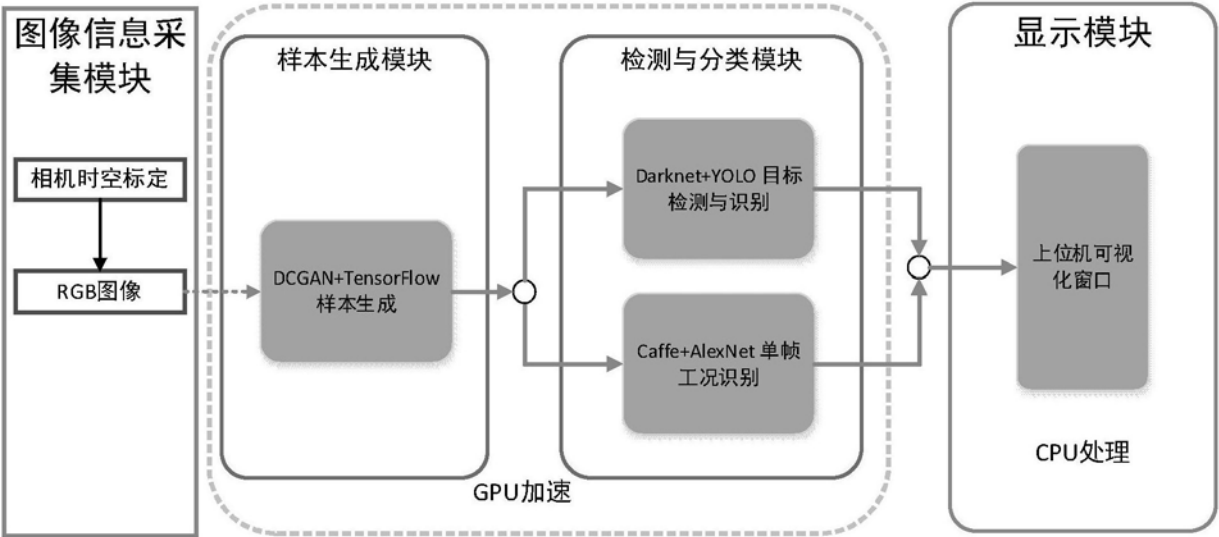


图2

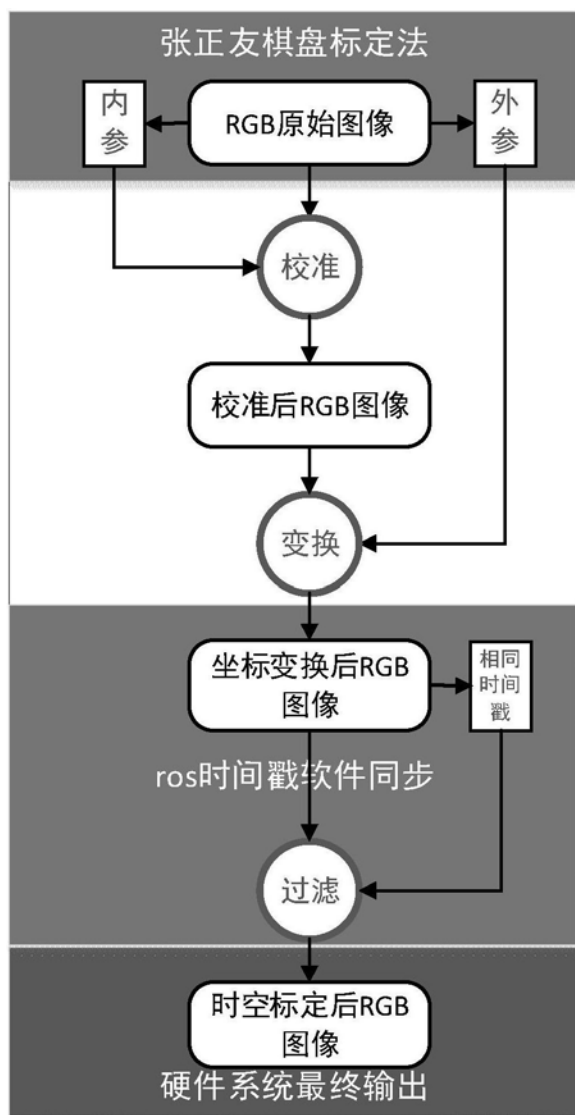


图3

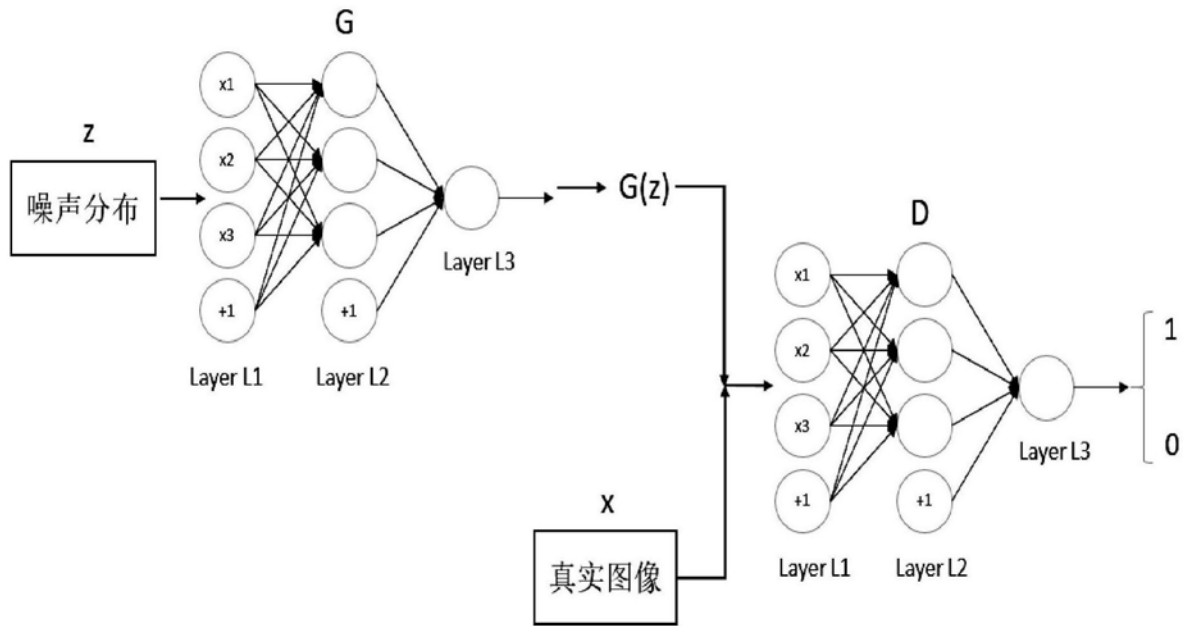


图4

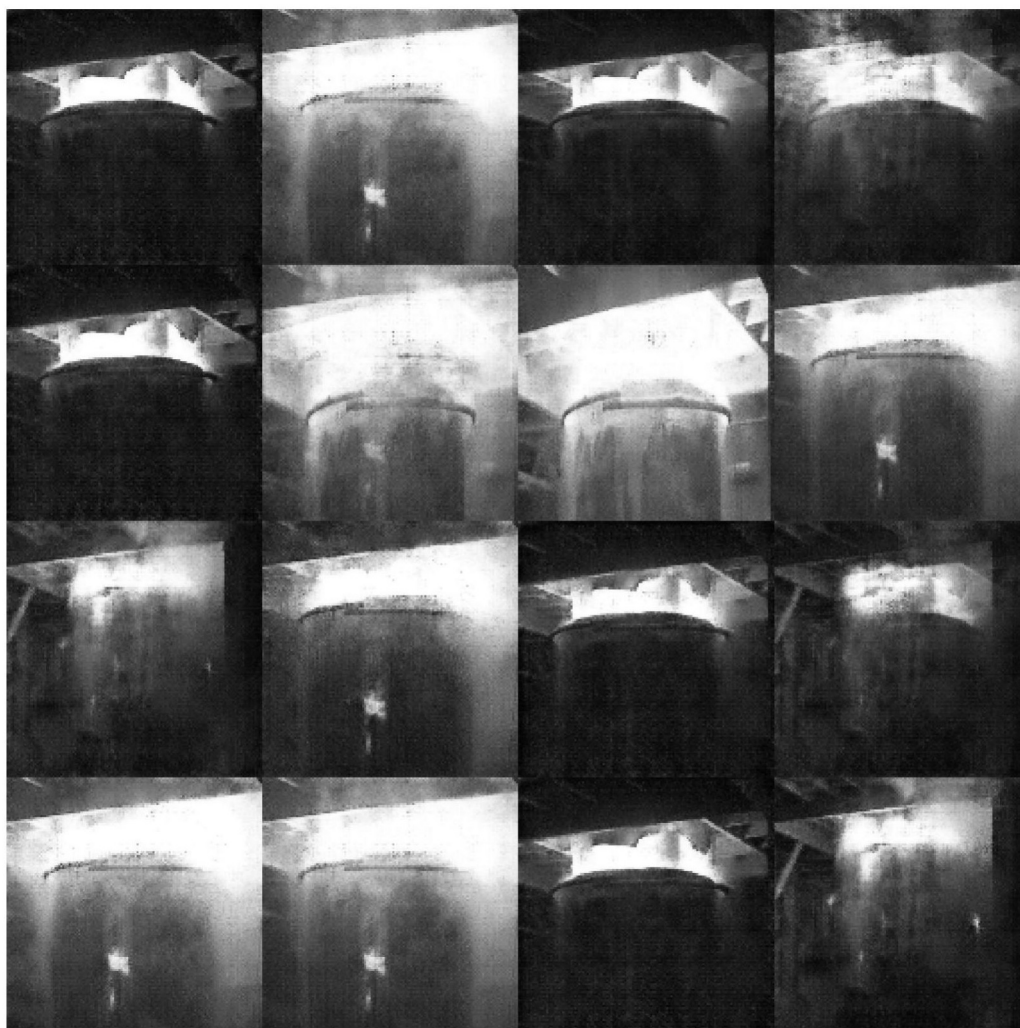


图5

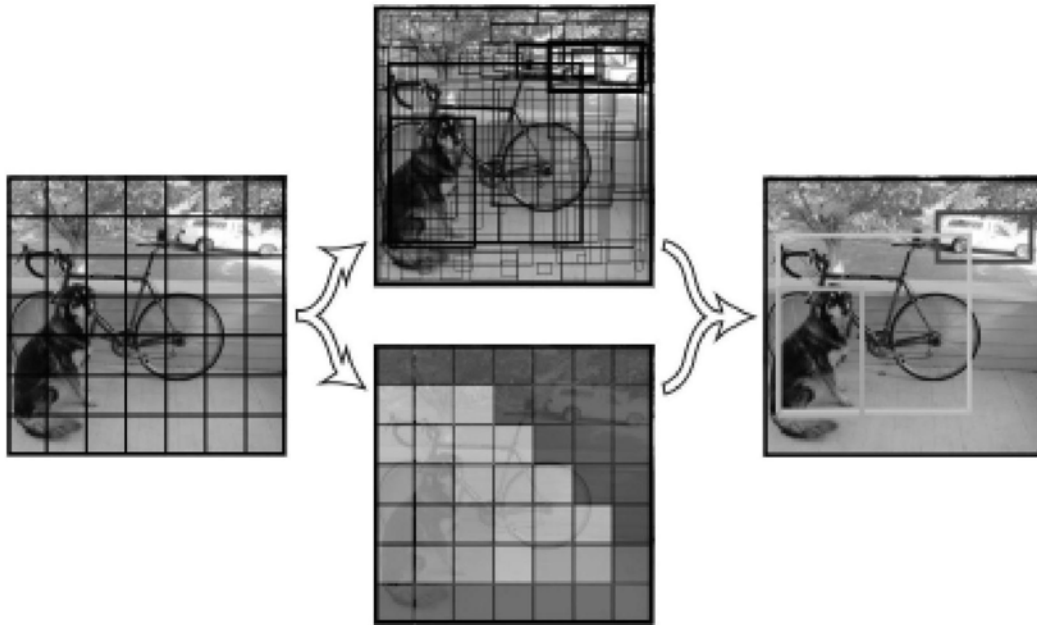


图6

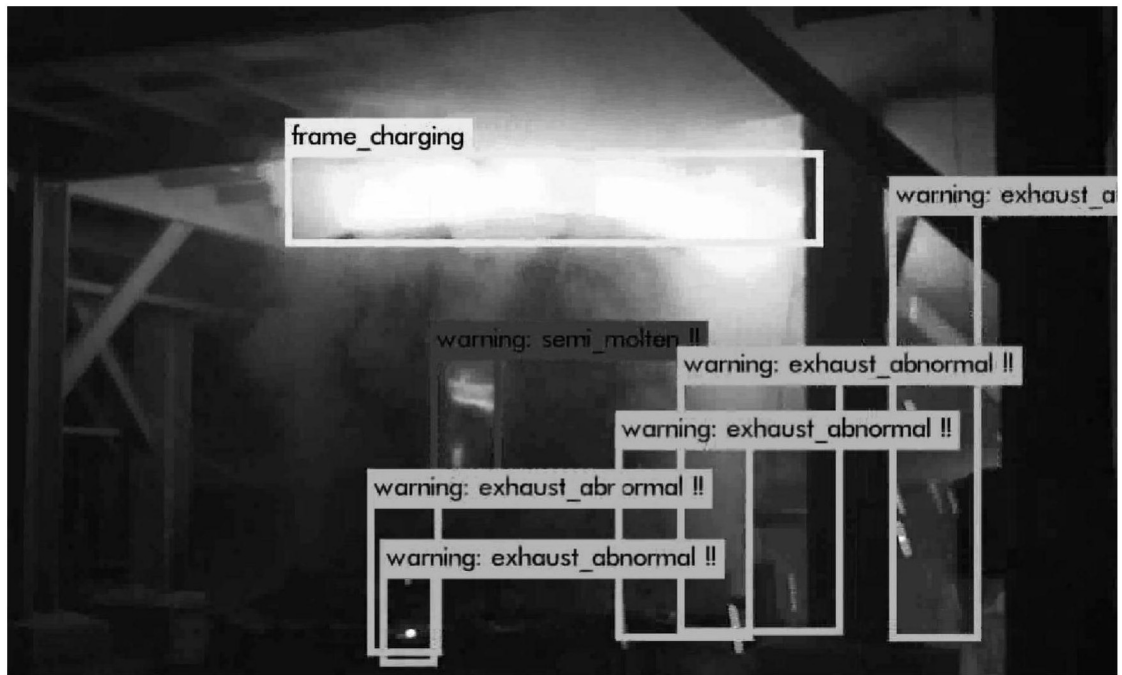


图7