**无人机障碍物精准检测与实时跟踪研究**

张继*+1*，时宏伟*1*

(1.四川大学1，四川 成都 610064)

**摘要：高低空飞行的无人机，时常面临飞机、鸟类、行人、汽车、交通灯等障碍物的碰撞威胁。如何准确地检测出这些障碍物，并在飞行路径中躲避它们，成为无人机研究方向的一个重要课题。本文专注于无人机障碍物检测环节，基于YOLO目标检测与KCF目标跟踪，对高低空常见的5种典型障碍物进行精准识别与实时跟踪。探索了高准确率YOLO目标检测与高速度KCF目标跟踪之间的最佳平衡点，从而同时保证了无人机障碍物检测的准确率与实时性**。

**关键词：障碍物、目标检测、目标跟踪**

**中图分类号：**TP3 **文献标识码：**A

**Research on UAV Obstacle Detection and Real-time Tracking**

Zhang Ji*+1*，Shi Hong-wei*1*

(1. Sichuan University, Sichuan Province Chengdu 610064)

**Abstract: Flying UAV(Unmanned Aerial Vehicle) often threatens by plane, birds, pedestrian, car, traffic light and so on. How to detect those obstacles, and avoid them while** When manufacturing Liquid Crystal Display Screen（LCD-Screen） of phone, tablet or TV, electric circuits need to be carved or printed on transparent glasses, with defects such as short circuit, open circuit occur sometimes. It is urgent to detect those defects in time and take corresponding measures. Image Recognition based on Deep-Learning has shown its ability to classify in domains like Biometrics Identification and Object Classification. Apply Image Recognition technology on LCD-Screen defect detection is also possible and achievable. This article focus on LCD-Screen defect detection based on Self-Learning CNN Classifier, adopting Self-Learning method to accelerate the establishment of marked defect library, updating Self-Learning Convolutional Neural Network Classifier (CNN Classifier) constantly. Finally, we can monitor LCD-Screen’s defect real-time, and ensure product quality of LCD-Screen.

**Keywords:** LCD-Screen, Self-Learning, CNN Classifier, defect classification, product quality

随着技术的发展，无人机在军民两大领域逐渐得到广泛运用。近几年的阅兵式上，军用无人机愈发崭露头角，翼龙、暗剑、彩虹等无人机在情报侦察、军事打击、信息对抗、通信中继和后勤保障等领域能发挥重大作用；民用无人机最具代表性的当属大疆出品的系列无人机，其在农业、基建、电力巡检、摄影和公共安全等领域能发挥重大作用。

高空飞行的无人机，时常会面临飞鸟以及其他飞机的威胁；低空飞行的无人机，与行人、汽车、交通灯、飞鸟相碰撞的几率较大。这不仅会对无人机自身造成损伤，更有可能给其他飞机、行人带来严重伤害。



图1 典型高空障碍



图2 典型低空障碍

本文以无人机高低空作业场景中常见的典型障碍物作为研究对象，使用YOLO目标检测模型进行典型障碍物的精准识别检测，同时借助KCF目标跟踪算法提升实时性。我们测试了不同检测间隔对准确度与速度的影响，以找到精准检测与快速跟踪之间的平衡点。

优点：

1. 使用YOLO目标检测模型进行典型障碍物的精准识别检测；
2. YOLO仅检测飞机、行人、汽车、交通灯、飞鸟5种类别障碍物，节省（80-5）/80=93.75%的后处理时间；
3. 每隔10帧使用一次YOLO目标检测，其余帧使用KCF目标跟踪，在保证80%准确度的同时，提升了5倍的运行速度；
4. 可扩展性：将自定义类别数据（如房屋、树木）加入YOLO训练数据集，重新训练YOLO目标检测模型，可使无人机获得其他类别障碍物的检测避让能力。

# YOLO目标检测模型

YOLO是一种基于深度神经网络的目标检测模型，以darknet为基本骨架，定义了YOLO自己的损失函数，然后在COCO、VOC或其他数据集上进行训练。常用的YOLO模型如表1所示，它们有着不同的准确度mAP和检测速度FPS。大体而言，准确度越高，则检测速度相对越慢。综合考虑准确度和检测速度的影响，我们选择YOLOv3-416作为目标检测模型，它使用COCO作为训练数据集，具备飞机、行人、汽车、交通灯、飞鸟等80个类别障碍物的检测识别能力。

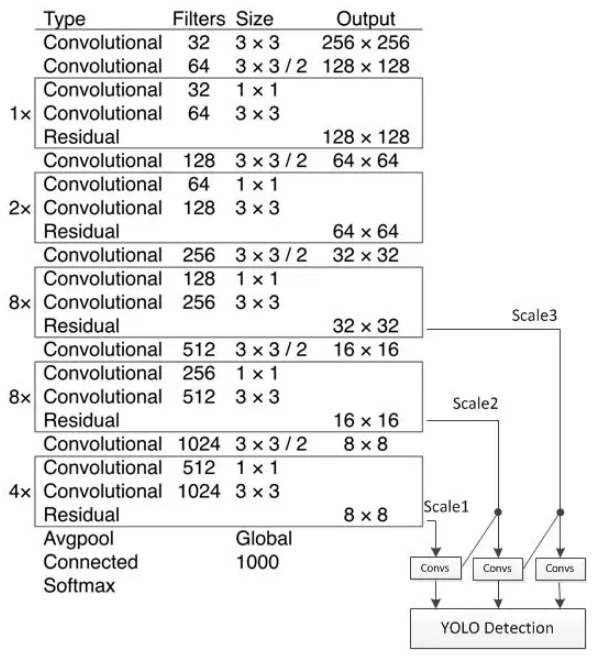


图3 YOLO神经网络架构

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 常用YOLO模型 | | | | | |
| 序号 | 模型名称 | 训练数据 | mAP | FLOPS | FPS |
| 1 | YOLOv3-320 | COCO | 51.5 | 38.97 Bn | 45 |
| 2 | YOLOv3-416 | COCO | 55.3 | 65.86 Bn | 35 |
| 3 | YOLOv3-608 | COCO | 57.9 | 140.69 Bn | 20 |
| 4 | YOLOv3-tiny | COCO | 33.1 | 5.56 Bn | 220 |
| 5 | YOLOv3-spp | COCO | 60.6 | 141.45 Bn | 20 |

# KCF目标跟踪模型

KCF即Kernel Correlation Filter（核相关滤波），它通过当前帧的位置，来预测物体在下一帧的位置，是一种快速的目标跟踪算法，运行速度可达到150fps。

# 最佳检测间隔

我们以YOLO目标检测的bbox作为KCF跟踪算法的初始输入，随后使用KCF替代YOLO检测出物体在每帧图像中的位置。随着时间的推移，我们发现KCF逐渐无法跟踪到正确的目标，如图3,。为了矫正这种跟踪错误，每隔一定的帧数，使用一次YOLO目标检测重新获得准确的目标位置，并更新KCF跟踪算法的输入。本小节以所有帧均进行YOLO检测得到的bbox和运行时间作为benchmark，研究设置不同YOLO检测间隔时的准确率和运行速度。

## benchmark

使用无人机录制的一段包含飞鸟的视频作为数据研究对象，该视频时长xxs，帧率xxfps，共计xx帧。对视频的每帧均进行YOLO目标检测，然后输出bbox和运行时间，作为优化对照的基准。部分结果如下表所示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 帧号 | Xb1 | Yb1 | Xb3 | Yb3 | Timeb |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

其中，(xb1, yb1)表示benchmark bbox的左上角点，(xb3, yb3)表示benchmark bbox的右下角点。易知，benchmark bbox右上角点坐标为(xb3, yb1)，benchmark bbox左下角点坐标为(xb1, yb3)。

由于实验条件限制，

表 2 测试机配置

|  |  |
| --- | --- |
| 系统 | Windows 10 企业版 64位操作系统 |
| CPU | CPU: Intel(R) Core(TM) i7-8700K CPU @ 3.70GHz 32GB |
| GPU | GPU: NVIDIA GeForce GTX 1060 6GB |

## 优化对照

对飞鸟视频每间隔interval帧抽取一帧进行YOLO目标检测，其余帧使用KCF输出bbox位置坐标和运行时间。部分结果如下表所示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Interval = 2 | | | | | |
| 帧号 | X1 | Y1 | X3 | Y3 | time |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

其中，(x1, y1)表示bbox的左上角点，(x3, y3)表示bbox的右下角点。易知，bbox右上角点坐标为(x3, y1)，bbox左下角点坐标为(x1, y3)。

与benchmark对比，位置偏差计算公式为：

*运行时间降低率计算公式为：*

## 优化结果分析

将interval值依次设置为2~10，并计算位置偏差和运行时间降低率，使用matlab分别绘制如下：

图4 不同Interval的位置偏差

图5 不同Interval的运行时间降低率

可以看到，当interval设置为5时，有最小位置偏差xxxx，运行时间降低率为xxxx。

# 双目测距

本文探索使用自学习卷积神经网络分类器对LCD屏的生产缺陷进行检测，采用自学习的方法快速建立标注的缺陷库，通过大量训练图片，不断迭代更新CNN分类器，最终实现对LCD屏产品缺陷的实时检测分类。自学习CNN分类器对LCD屏缺陷的分类精度较高，且使用GPU完全能够满足LCD屏生产线的实时检测要求。本文的亮点在于自学习方法，它解决了高达数十万张训练图片标注缓慢的问题。通过速度更快的“数据筛选”替代“数据标注”，能够在较短时间内迅速扩充缺陷库的规模，而大量的训练数据又能够带来更高的分类精度，从而形成了一个良性循环的缺陷检测机制，研究结果具有较强的理论探索意义和工程应用价值。

# 结束语

本文探索使用自学习卷积神经网络分类器对LCD屏的生产缺陷进行检测，采用自学习的方法快速建立标注的缺陷库，通过大量训练图片，不断迭代更新CNN分类器，最终实现对LCD屏产品缺陷的实时检测分类。自学习CNN分类器对LCD屏缺陷的分类精度较高，且使用GPU完全能够满足LCD屏生产线的实时检测要求。本文的亮点在于自学习方法，它解决了高达数十万张训练图片标注缓慢的问题。通过速度更快的“数据筛选”替代“数据标注”，能够在较短时间内迅速扩充缺陷库的规模，而大量的训练数据又能够带来更高的分类精度，从而形成了一个良性循环的缺陷检测机制，研究结果具有较强的理论探索意义和工程应用价值。

# 参考文献

[1]S. S. Farfade, M. J. Saberian, and L. J. Li, “Multi-view face detection using deep convolutional neural networks,” in ACM Int. Conf. Multimedia Retrieval, 2015, pp. 643–650.

[2]J. Redmon and A. Farhadi. Yolo9000: Better, faster, stronger. In Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on, pages 6517–6525. IEEE, 2017. 1, 2, 3

[3]陈崔军.TFT-LCD ARRAY 光刻制程 CD 均一性分析和实现[D].电子科技大学,2018.

[4]袁林. TFT混切技术的干蚀刻制程工艺优化研究[D].华南理工大学,2018.

[5]A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, 2012.

[6]K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In ICLR, 2015.

[7]C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In CVPR, 2015.

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun. Deep Residual Learning for Image Recognition. In CVPR, 2015.

[9]谢希仁.计算机网络.电子工业出版社.2008.