Good afternoon! Our peers.
The topic we talk is Neural network-based military aircraft identification today. The Speakers is Wang Guo and Wang Zi Dong. Our team members is Yuxi Chen, Yuhong Li, Guo Wang, Ying Wang, and Zidong Wang.

研究目的

对互联网上机密军事飞机的图片和视频的监管，是国家机密军事信息保护的重要组成部分。在机密军事飞机的相关照片和视频上传到互联网上之前，需要采取技术手段将其识别并拦截。
选择合适的目标检测模型对飞机进行识别，并使模型尽量轻量化，得以部署到生活场景中进行应用，是我们的研究目的。

The supervision of pictures and videos of confidential military aircraft on the Internet is an important part of the protection of state confidential military information. Before the relevant photos and videos of confidential military aircraft are uploaded to the Internet, technical means need to be taken to identify and intercept them.
The purpose of our research is to select an appropriate target detection model to identify the aircraft and make the model as lightweight as possible, so that it can be deployed to life scenes for application.
We will talk about our research in the following aspects. They are background, method, promotion, and result.

Firstly, let's talk about the Background. In the lower left corner is a news about the execution of secretly shooting military aircraft videos. And on the right is the discussion about aircraft on Baidu Post Bar. Obviously, this may involve military secrets.

Now the supervision of pictures and videos of confidential military aircraft on the Internet is an important part of the protection of state confidential military information. Before the relevant photos and videos of confidential military aircraft are uploaded to the Internet, technical means need to be taken to identify and intercept them.

切换 PPT ：
The first figure in the figure below shows a more lightweight model, while the second figure shows a faster model

The purpose of our research is to select an appropriate target detection model to identify the aircraft and make the model as lightweight as possible, so that it can be deployed to life scenes for application.

研究现状

国内外已经在军事飞机检测和识别方面提出了很多方法。
2016 年 Redmon 等人提出了 YOLO 算法，将分类、定位、检测功能集成在一个神经网络当中，只需经过一次计算，就可以直接得到图像中目标的边界框和类别概率。
2018 年 Redmon 等人提出了 YOLOv3 算法，它借鉴残差网络中跳跃连接的思路，构建了名为 DarNet-53 的 53 层基准网络，该网络只采用 3×3 和 1×1 的卷积层，具有与 ResNet-152 相仿的分类准确率，但大大减少了计算量。
2020 年 Bochkovskiy 等人[10] 提出了 YOLOv4 算法，在 YOLOv3 的基础上不断进行改进和开发。YOLOv4 可以使用传统的 GPU 进行训练和测试，并能够获得实时的，高精度的检测结果。

Many methods have been proposed in military aircraft detection and recognition at home and abroad.
In 2016, Redmon et al. Proposed the Yolo algorithm, which integrates the classification, positioning and detection functions into a neural network. After only one calculation, the bounding box and category probability of the target in the image can be directly obtained.
In 2018, Redmon et al. Proposed the yolov3 algorithm, which built a 53 layer benchmark network called darnet-53 by using the idea of jump connection in the residual network for reference. The network only uses 3 × 3 and 1 × 1, which has the same classification accuracy as resnet-152, but greatly reduces the amount of calculation.
In 2020, bochkovskiy et al. [10] proposed yolov4 algorithm, which is continuously improved and developed on the basis of yolov3. Yolov4 can use traditional GPU for training and testing, and can obtain real-time and high-precision detection results.

第一个研究
In recent years, many people have conducted in-depth research on related problems. What is shown in the picture is the Yolo algorithm proposed by Redmon et al. In 2016. Its typical feature is to integrate the classification, positioning and detection functions into a neural network. After only one calculation, the boundary box and category probability of the target in the image can be directly obtained.（黄色标记为 PPT 上内容）

In 2018, Redmon and others improved Yolo with the help of residual neural network, and then proposed yolov3 algorithm.
In 2020, some scholars proposed yolov4 algorithm to obtain real-time and high-precision detection results.（对 PPT 上展示内容进行了口述上的简化）

研究方法

我们以 YOLOv5 模型为原型，提出了以下五种改进方法来提高模型的训练和检测速度。
首先是采用 Mosaic-9 数据增强方法，即对 9 张图片随机裁剪、随机缩放、随机排列组合成一张新图片，将图片传入到神经网络当中去学习，相当于一次性传入了 9 张图片，这极大丰富了检测物体的背景。
第二是使用轻量化的模型 MobileNet V3 Small 来取代 YOLOv5 的 Backbone 特征提取网络来进行特征提取，其能够在不会降低网络精度的情况下大大降低模型复杂度。
第三是对深度卷积神经网络（CNN）进行通道剪枝，通道剪枝可以减小模型大小、减少运行时的内存占用、在不影响精度的同时降低计算操作数。
第四是在 YOLOv5 的 Prediction 层中引入 Label Smoothing 标签平滑方法，标签平滑将帮助模型围绕错误的标签数据进行训练，从而提高其健壮性和性能。
最后是将 YOLO 训练的模型送入 TensorRT 中优化产生 Engine 引擎，然后再应用在 GPU 推理中，进一步使模型轻量化。

Taking yolov5 model as the prototype, we propose the following five improved methods to improve the training and detection speed of the model.
Firstly, mosaic-9 data enhancement method is adopted, that is, 9 pictures are randomly cut, randomly scaled and randomly arranged to form a new picture, and the picture is introduced into the neural network for learning, which is equivalent to 9 pictures at one time, which greatly enriches the background of the detected object.
The second is to use the lightweight model mobilenet V3 small to replace the backbone feature extraction network of yolov5 for feature extraction, which can greatly reduce the model complexity without reducing the network accuracy.
The third is to prune the channel of deep convolutional neural network (CNN). Channel pruning can reduce the size of the model, reduce the memory occupation at runtime, and reduce the computational operands without affecting the accuracy.
The fourth is to introduce the label smoothing method into the prediction layer of yolov5. Label smoothing will help the model train around the wrong label data, so as to improve its robustness and performance.
Finally, the Yolo trained model is sent to tensorRT to optimize the engine, and then applied to GPU reasoning to further lighten the model.
After discussing other people's research, let's discuss our Method.（引入） we propose the following five improved methods to improve the training and detection speed of the model by taking yolov5 model as the prototype.

Firstly, As shown in the picture on the right, mosaic-9 data enhancement method is adopted, that is, 9 pictures are randomly cut, randomly scaled and randomly arranged to form a new picture, and the picture is introduced into the neural network for learning, which is equivalent to 9 pictures at one time, which greatly enriches the background of the detected object.

The second is to use the lightweight model mobilenet V3 small to replace the backbone feature extraction network of yolov5 for feature extraction, which can greatly reduce the model complexity without reducing the network accuracy.
The relevant comparisons are listed in the left figure below. And the reference source of the picture is given below
The third is to prune the channel of deep convolutional neural network (CNN). Channel pruning can reduce the size of the model, reduce the memory occupation at runtime, and reduce the computational operands without affecting the accuracy.
The following two figures are initial network and compact network
The third is to prune the channel of deep convolutional neural network (CNN). Channel pruning can reduce the size of the model, reduce the memory occupation at runtime, and reduce the computational operands without affecting the accuracy.
Finally, the Yolo trained model is sent to tensorRT to optimize the engine, and then applied to GPU reasoning to further lighten the model.

研究结果

我们的实验基于 Military Aircraft DetectionDataset1，该数据集包含 36 种，5062 张军用飞机图片，包含中、美、俄、欧等国家热门机型。最终经过 TensorRT 加速后的模型虽然会损失一定的精度，但检测速度得到了极大的提升。
经过 50 次迭代，模型最终的损失及 mAP 值如图所示，由图可知，在经过多次训练之后，模型 mAP 值逐渐稳定在 0.6 左右，证明优化后的模型具有良好的准确性。
Our experiment is based on military aircraft detectiondataset1, which contains 36 kinds and 5062 pictures of military aircraft, including popular models in China, the United States, Russia, Europe and other countries. Finally, although the model accelerated by tensorRT will lose some accuracy, the detection speed has been greatly improved.
After 50 iterations, the final loss and map value of the model are shown in Figure 8. It can be seen from the figure that after many training, the map value of the model gradually stabilized at about 0.6, which proves that the optimized model has good accuracy.
Ok, next, let's talk about the improvement of the research.
我们所采用的数据是基于 Military Aircraft DetectionDataset1 数据集，其中，有 36 种，5062 张军用的飞机图片，包括中，美，俄、欧等国家热门机型。最终经过 TensorRT 加速后的模型虽然会损失一定的精度，但检测速度得到了极大的提升。
The data we use is based on the military aircraft detectiondataset1 data set, including 36 kinds and 5062 military aircraft pictures, including popular models in China, the United States, Russia, Europe and other countries. Finally, although the model accelerated by tensorrt will lose some accuracy, the detection speed has been greatly improved.
通过我们的 20 次迭代实验，模型的最终损失以及 MAP 值如图所示，MAP 是指 mean Average Precision。从图中我们可以得知，经过多次训练之后，模型的 mAP 值逐渐稳定在 0.6 左右，证明优化后的模型有良好的准确率。
Through our 20 iteration experiments, the final loss of the model and the map value are

shown in the figure. Map refers to mean average precision. From the figure, we can know that after many training, the map value of the model gradually stabilized at about 0.6, which proves that the optimized model has good accuracy.

结果讨论和研究意义

我们提出了一种基于改进 YOLOv5 网络的军用飞机快速检测方法，主要的改进工作为：
1. 使用 Mosaic-9 增加了小目标样本，提升网络训练速度。
2. 采用 MobileNet V3 Small 取代 YOLOv5s 的 Backbone 特征提取网络。
3. 通道剪枝，改善网络的综合性能。
4. 引入 Label Smoothing 标签平滑方法。通过 soft one-hot 来加入噪声，起到抑制过拟合的效果。
5. 使用 TensorRT 优化 GPU 硬件资源

We propose a fast detection method of military aircraft based on improved yolov5 network. The main improvements are as follows:
1. Using mosaic-9, small target samples are added to improve the speed of network training.
2. Use mobilenet V3 small to replace the backbone feature extraction network of yolov5s.
3. Channel pruning to improve the comprehensive performance of the network.
4. Introduce label smoothing method. Noise is added through soft one hot to suppress over fitting.
5. Use tensorrt to optimize GPU hardware resources
And next，the final part is result. We give a conclusion.
The main improvements are as follows:
1. firstly we use mosaic-9, small target samples are added to improve the speed of network training.
2. and use mobilenet V3 small to replace the backbone feature extraction network of yolov5s.
3. after that. We use Channel pruning to improve the comprehensive performance of the network.
4. and Introduce label smoothing method. Noise is added through soft one hot to suppress over fitting.
5. finally, we use tensorRT to optimize GPU hardware resources

Next, let’s play a short video to show the general effect.
Ok! This is all our research.
Thank you for listening!