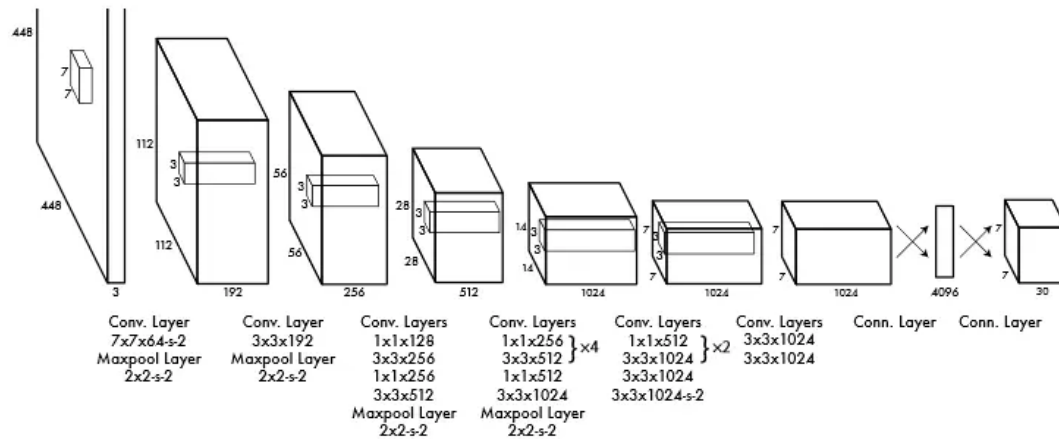
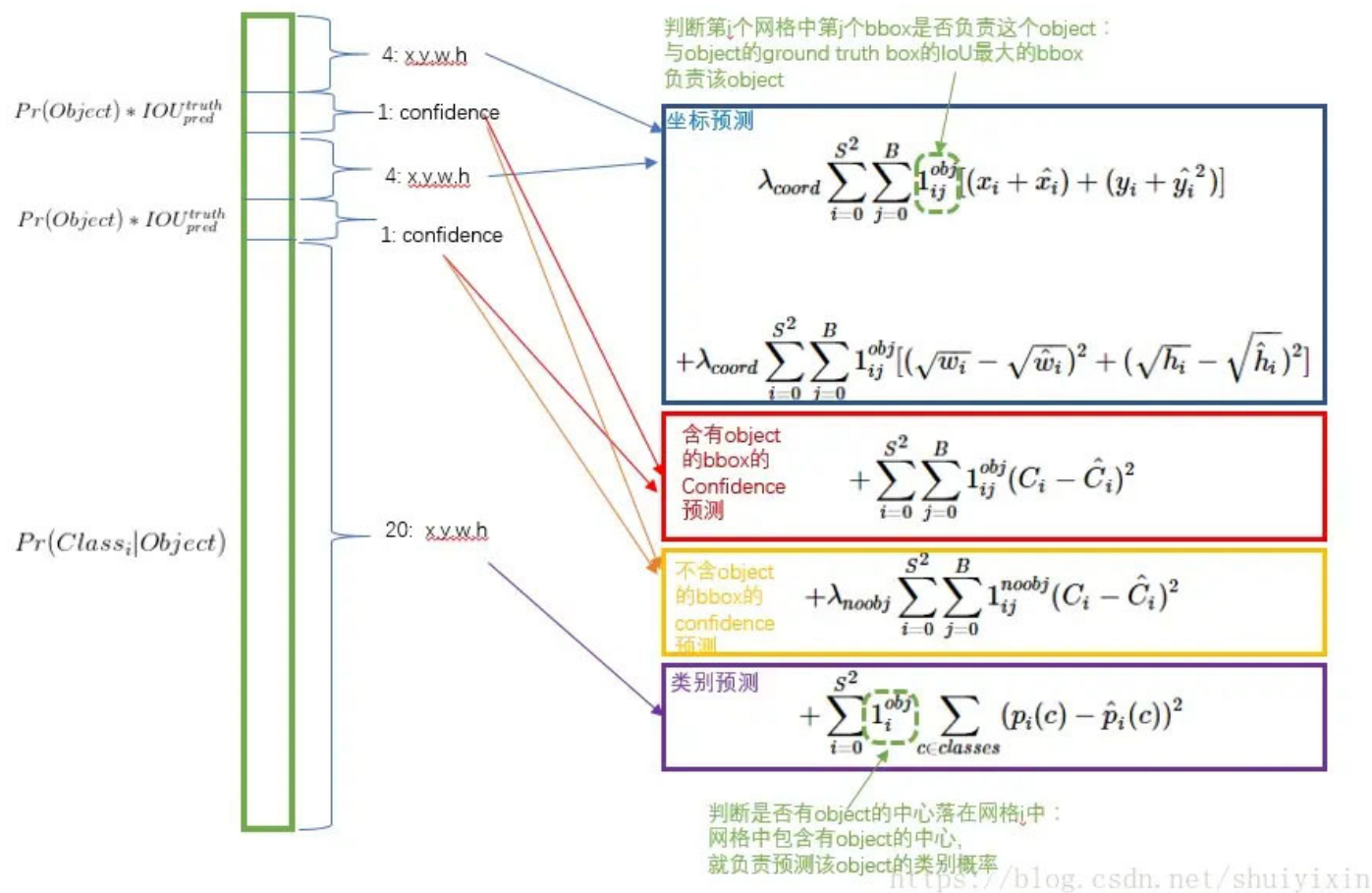


## Network Architecture



**Figure 3: The Architecture.** Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating  $1 \times 1$  convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution ( $224 \times 224$  input image) and then double the resolution for detection. <https://blog.csdn.net/shuiyixin>

# Loss Function



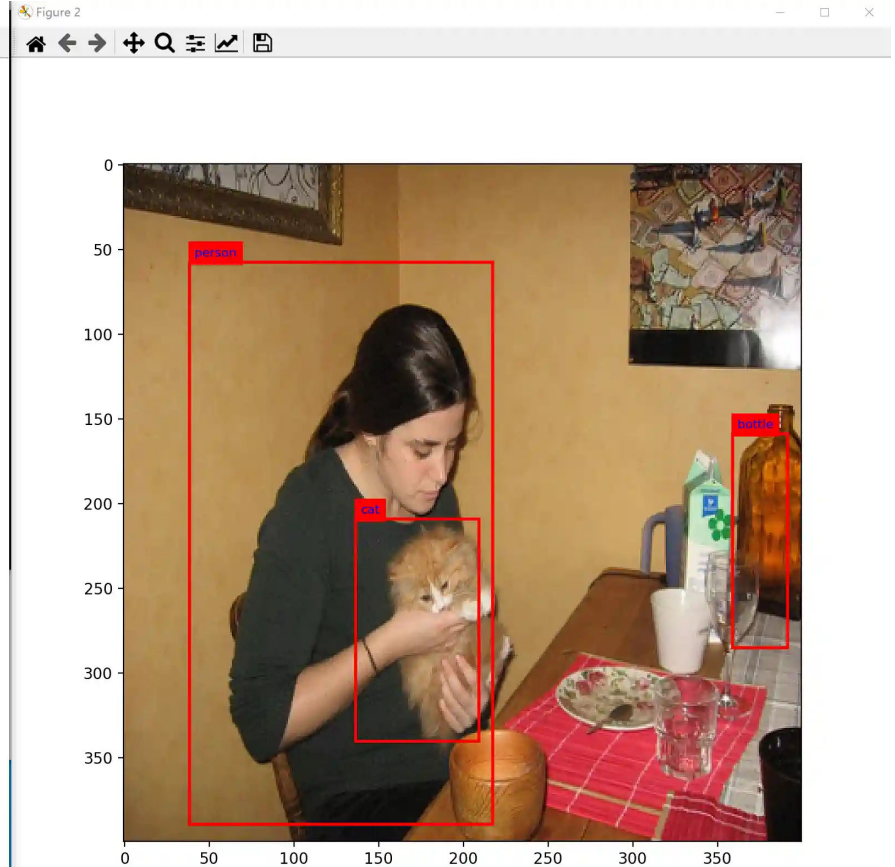
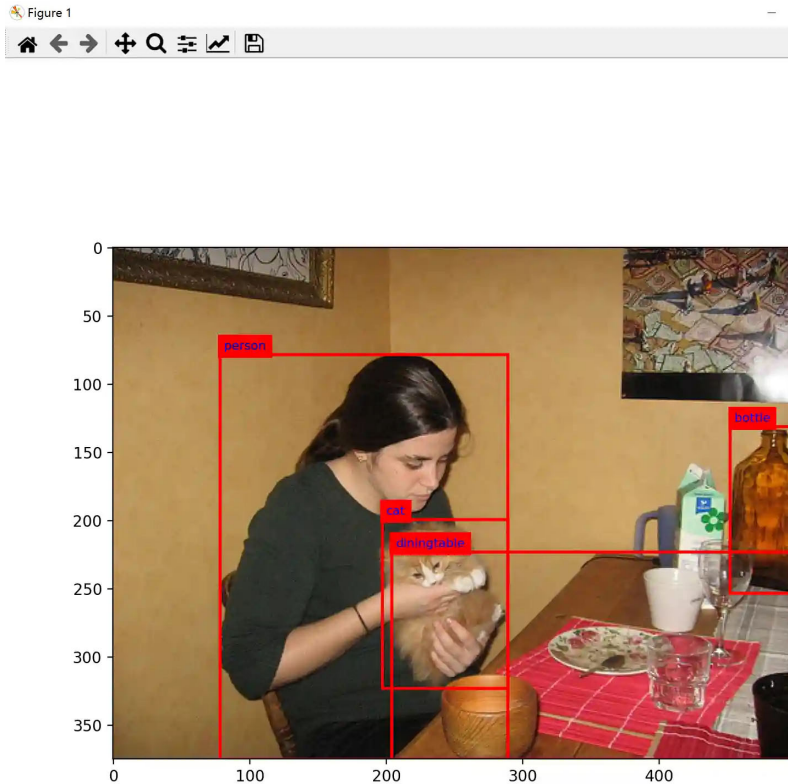
# Code

```
class YOLO(nn.Module):
    def __init__(self):
        super().__init__()
        resnet = torchvision.models.resnet18(pretrained=False)
        self.resnet = nn.Sequential(*list(resnet.children())[:-1])
        self.detect = nn.Sequential(
            nn.Linear(512, 512),
            nn.LeakyReLU(),
            nn.Linear(512, final_channels*n_grid*n_grid),
            nn.Sigmoid()
        )

    def forward(self, x):
        x = self.resnet(x)
        x = x.reshape(-1, 512)
        x = self.detect(x)
        return x.reshape(-1, final_channels, n_grid, n_grid)
```

# Result

- ▶ **Training Set:** VOC2012, 100 images
- ▶ **Batch Size:** 16
- ▶ **Epoch:** about 10000



# Objection Detection

## backbone

VGG  
ResNet  
ResNeXt  
DenseNet  
SqueezeNet  
Darknet  
MobileNet  
ShuffleNet  
DetNet  
DetNAS  
SpineNet  
EfficientNet  
CSPResNeXt50  
CSPDarknet53

## neck

Additional blocks:

SPP

ASPP

RFB

SAM

Path-aggregation blocks:

FPN

PAN

NAS-FPN

Fully-connected FPN

BiFPN

ASFF

SFAM

NAS-FPN

## Head

Dense Prediction (one-stage):

RPN

SSD

YOLO

RetinaNet

(anchor based)

CornerNet

CenterNet

MatrixNet

FCOS(anchor free)

Sparse Prediction (two-stage):

Faster R-CNN

R-FCN

Mask RCNN (anchor based)

RepPoints(anchor free)

# Objection Detection

## Skip-connections

- Residual connections
- Weighted residual connections
- Multi-input weighted residual connections
- Cross stage partial connections (CSP)

## loss

- MSE
- Smooth L1
- Balanced L1
- KL Loss
- GHM loss
- IoU Loss
- Bounded IoU Loss
- GIoU Loss
- CloU Loss
- DIoU Loss

## Regularization

- DropOut
- DropPath
- Spatial DropOut
- DropBlock

## Activation function

- ReLU
- LReLU
- PReLU
- ReLU6
- Scaled Exponential Linear Unit (SELU)
- Swish
- hard-Swish
- Mish

## BN

- Batch Normalization (BN)
- Cross-GPU Batch Normalization (CGBN or SyncBN)
- Filter Response Normalization (FRN)
- Cross-Iteration Batch Normalization (CBN)

YOLOv4: Optimal Speed and Accuracy of Object Detection