

PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization

一、Paper Reading

First

1、Abstract

PaDiM 使用预先训练的 CNN 来进行分块嵌入,然后使用多元高斯分布来得到正常类别的概率表示.此外,还利用了 CNN 不同层的语义来更高的定位缺陷.

PaDiM 利用预训练的 CNN 来提取图片特征,并遵循以下两个规则:

- 每个 patch 位置都使用一个多元高斯分布来表示
- 考虑不同层之间语义的相关关系

2、论文中的图片、表格

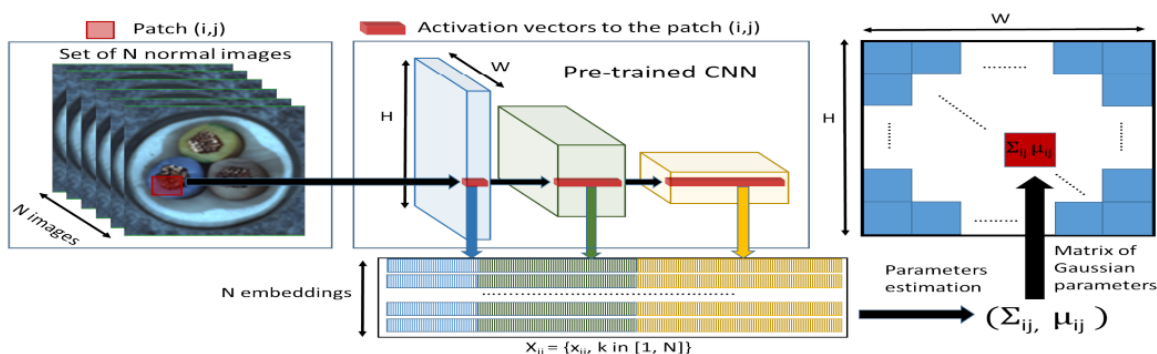


图2、将预训练模型中三个不同层的feature map对应位置进行拼接得到embedding vector

$$\Sigma_{ij} = \frac{1}{N-1} \sum_{k=1}^N (x_{ij}^k - \mu_{ij})(x_{ij}^k - \mu_{ij})^T + \epsilon I \quad (1)$$

公式1、计算某点的特征的正态分布时,协方差矩阵的计算使用添加来一个正则项 ϵI 来保证 协方差矩阵是满秩且可逆的.

$$M(x_{ij}) = \sqrt{(x_{ij} - \mu_{ij})^T \Sigma_{ij}^{-1} (x_{ij} - \mu_{ij})} \quad (2)$$

公式2、马氏距离计算方式, 计算图片在 (i, j) 处的异常分数, $M(x_{ij})$ 表示测试图片的嵌入向量 x_{ij} 和学习到的分布之间的距离

3、Conclusion

- Inter-layer correlation(层间相关性)

通过实验对比了单独采用layer1、2、3，将单独采用单层特征三个模型相加进行模型融合，以及本文的方法。通过上表可以看出，模型融合比采用单层特征的单个模型效果更好，但这并没有考虑到三层之间的关联，PaDiM考虑到了这点并获得了最优的结果。

- Dimensionality reduction (降维)

对比了PCA和随机挑选两种降维方法，无论降到100维还是200维，随机挑选的效果都比PCA好，这可能是因为PCA挑选的是那些方差最大的维度，但这些维度可能不是最有助于区分正常和异常类的维度。另外随机挑选的方法从448维降到200维再降到100维，精度损失都不大，但却大大降低了模型复杂度。

Second

1、评估：ROC_AUC

ROC-AUC是一种用于评估二进制分类模型性能的指标，它衡量了模型在不同分类阈值下的真正例率（True Positive Rate，又称敏感性）与假正例率（False Positive Rate）之间的权衡。

以下是关于 ROC-AUC 的一些基本概念：

- **ROC曲线 (Receiver Operating Characteristic Curve)**：是一个以真正例率（True Positive Rate, TPR）为纵轴、假正例率（False Positive Rate, FPR）为横轴的曲线。ROC曲线可以帮助我们理解在不同阈值下，模型在正例和负例上的表现如何。
- **AUC (Area Under the Curve)**：ROC曲线下的面积，即ROC-AUC。AUC的取值范围在0到1之间，越接近1表示模型性能越好，0.5表示模型性能等同于随机猜测。
- **真正例率 (True Positive Rate, TPR)**：也称为敏感性或召回率，表示正例中被正确识别为正例的比例，计算方式是 $TP / (TP + FN)$ ，其中TP是真正例数量，FN是假负例数量。
- **假正例率 (False Positive Rate, FPR)**：表示负例中被错误识别为正例的比例，计算方式是 $FP / (FP + TN)$ ，其中FP是假正例数量，TN是真负例数量。

2、算法与基本流程

- Embedding extraction (特征提取融合)

将预训练模型中三个不同层的feature map对应位置进行拼接得到embedding vector，这里分别取resnet18中layer1、layer2、layer3的最后一层，模型输入大小为224x224，这三层对应输出维度分别为(209,64,56,56)、(209,128,28,28)、(209,256,14,14)。

不同层输出拼接的方式：

这里实现1中是通过将小特征图每个位置复制多份得到与大特征图同样的spatial size，然后再进行拼接。比如28x28的特征图中一个1x1的patch与56x56的特征图中对应位置的2x2的patch相对应，将该1x1的patch复制4份得到2x2大小patch后再与56x56对应位置的2x2 patch进行拼接，下面的代码是通过F.unfold实现的。

最后恢复到 224x224 方式：上采样。

降维：将三个不同语义层的特征图进行拼接后得到(209, 448, 56, 56)大小的patch嵌入向量可能带有冗余信息，因此作者对其进行了降维，作者发现随机挑选某些维度的特征比PCA更有效，在保持sota性能的前提下降低了训练和测试的复杂度，文中维度选择100，因此输出为(209, 100, 56, 56)。

- Learning of the normality (学习正常样例分布)

特征分布计算：从 embedding_vectors 中选择的通道进行多元高斯分布的计算。计算均值与协方差。

"""协方差矩阵的可逆性

协方差矩阵的可逆性是一个重要的性质。协方差矩阵通常是对称正定的，这意味着它是一个实对称矩阵，并且对于任意非零向量，其与协方差矩阵的乘积都是正的。

为了确保协方差矩阵的可逆性，添加了一个小的正则化项 $0.01 * \mathbf{I}$ 到协方差矩阵。这个小的正则化项通常被称为"平滑项"，用于防止协方差矩阵的奇异性（不可逆性），尤其是在估计样本协方差矩阵时，当样本数量较小时可能会出现。

奇异性是指矩阵不是满秩的，即存在线性相关的列或行。对于协方差矩阵而言，奇异性可能导致矩阵不可逆。

当样本数量相对较小时，样本协方差矩阵可能会变得不稳定，特别是当特征的数量大于样本数时。这种情况下，样本协方差矩阵可能是奇异的，即存在线性相关的特征，导致矩阵不可逆。

添加正则化项的目的是通过对角线上增加一个小的常数，来防止协方差矩阵的奇异性。这个小的正则化项有时被称为"平滑项"，它确保协方差矩阵在数值上是稳定的，并且保持了一定的数值非奇异性。

mean: (100, 3136) cov: (100, 100, 3136)

100个维度：56x56，1x1视野为4x4，输入分辨率为(224x224)。每个维度一个均值。

"""

- Inference : computation of the anomaly map(异常图的计算)

计算每个样本到训练集中对应位置的均值的马氏距离。

在得到每个像素点的马氏距离后，进行上采样、高斯滤波、归一化的后处理后，就得到了最终的输出，大小和输入相同。

二、Code

<https://github.com/xiahaifeng1995/PaDiM-Anomaly-Detection-Localization-master.git>

EX1:

```
def parse_args():
    parser = argparse.ArgumentParser('PaDiM')
    parser.add_argument('--data_path', type=str,
                        default='/root/dataset/public/mvtec_anomaly_detection')
    parser.add_argument('--save_path', type=str, default='./mvtec_result')
    parser.add_argument('--arch', type=str, choices=['resnet18', 'wide_resnet50_2'],
                        default='resnet18')
    return parser.parse_args()
```

python main_wood.py

- ```
Downloading: "https://download.pytorch.org/models/resnet18-f37072fd.pth" to
/root/.cache/torch/hub/checkpoints/resnet18-f37072fd.pth
100%|███████████|
44.7M/44.7M [00:01<00:00, 33.5MB/s]
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | bottle |: 100%|
███████████| 7/7 [00:07<00:00, 1.03s/it]
| feature extraction | test | bottle |: 100%|
███████████| 3/3 [00:02<00:00, 1.22it/s]
image ROCAUC: 0.996
pixel ROCAUC: 0.981
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | cable |: 100%|
███████████| 7/7 [00:09<00:00, 1.40s/it]
| feature extraction | test | cable |: 100%|
███████████| 5/5 [00:08<00:00, 1.66s/it]
image ROCAUC: 0.855
pixel ROCAUC: 0.949
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | capsule |: 100%|
███████████| 7/7 [00:09<00:00, 1.32s/it]
| feature extraction | test | capsule |: 100%|
███████████| 5/5 [00:07<00:00, 1.49s/it]
image ROCAUC: 0.870
pixel ROCAUC: 0.982
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | carpet |: 100%|
███████████| 9/9 [00:11<00:00, 1.26s/it]
| feature extraction | test | carpet |: 100%|
███████████| 4/4 [00:06<00:00, 1.65s/it]
```

```
image ROCAUC: 0.984
pixel ROCAUC: 0.988
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
 "Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | grid |: 100%|
██ | 9/9 [00:05<00:00, 1.58it/s]
| feature extraction | test | grid |: 100%|
██ | 3/3 [00:02<00:00, 1.07it/s]
image ROCAUC: 0.898
pixel ROCAUC: 0.936
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
 "Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | hazelnut |: 100%|
██ | 13/13 [00:16<00:00, 1.24s/it]
| feature extraction | test | hazelnut |: 100%|
██ | 4/4 [00:06<00:00, 1.52s/it]
image ROCAUC: 0.841
pixel ROCAUC: 0.979
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
 "Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | leather |: 100%|
██ | 8/8 [00:10<00:00, 1.26s/it]
| feature extraction | test | leather |: 100%|
██ | 4/4 [00:06<00:00, 1.65s/it]
image ROCAUC: 0.988
pixel ROCAUC: 0.990
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
 "Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | metal_nut |: 100%|
██ | 7/7 [00:04<00:00, 1.47it/s]
| feature extraction | test | metal_nut |: 100%|
██ | 4/4 [00:03<00:00, 1.06it/s]
image ROCAUC: 0.974
pixel ROCAUC: 0.967
```

```

/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | pill |: 100%|
██| 9/9 [00:08<00:00, 1.12it/s]
| feature extraction | test | pill |: 100%|
██| 6/6 [00:06<00:00, 1.13s/it]
image ROCAUC: 0.869
pixel ROCAUC: 0.946
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | screw |: 100%|
██| 10/10 [00:07<00:00, 1.42it/s]
| feature extraction | test | screw |: 100%|
██| 5/5 [00:05<00:00, 1.15s/it]
image ROCAUC: 0.745
pixel ROCAUC: 0.972
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | tile |: 100%|██|
8/8 [00:07<00:00, 1.01it/s]
| feature extraction | test | tile |: 100%|██|
4/4 [00:05<00:00, 1.34s/it]
image ROCAUC: 0.959
pixel ROCAUC: 0.917
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | toothbrush |: 100%|
██| 2/2 [00:02<00:00, 1.40s/it]
| feature extraction | test | toothbrush |: 100%|
██| 2/2 [00:02<00:00, 1.13s/it]
image ROCAUC: 0.947
pixel ROCAUC: 0.986
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.

```

```
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | transistor |: 100%|
██████████████████████████████████████| 7/7 [00:10<00:00, 1.54s/it]
| feature extraction | test | transistor |: 100%|
██████████████████████████████████████| 4/4 [00:05<00:00, 1.48s/it]
image ROCAUC: 0.925
pixel ROCAUC: 0.968
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | wood |: 100%|
██████████████████████████████████████| 8/8 [00:12<00:00, 1.58s/it]
| feature extraction | test | wood |: 100%|
██████████████████████████████████████| 3/3 [00:04<00:00, 1.53s/it]
image ROCAUC: 0.990
pixel ROCAUC: 0.940
/root/conda/envs/cv_env/lib/python3.7/site-
packages/torchvision/transforms/transforms.py:330: UserWarning: Argument
'interpolation' of type int is deprecated since 0.13 and will be removed in 0.15.
Please use InterpolationMode enum.
"Argument 'interpolation' of type int is deprecated since 0.13 and will be removed
in 0.15. "
| feature extraction | train | zipper |: 100%|██████████████████████████████████████| 8/8
[00:06<00:00, 1.16it/s]
| feature extraction | test | zipper |: 100%|██████████████████████████████████████| 5/5
[00:05<00:00, 1.05s/it]
image ROCAUC: 0.741
pixel ROCAUC: 0.976
Average ROCAUC: 0.905
Average pixelROCUAC: 0.965
```