

---

队伍编号	MCB2303548
赛道	A

---

## 基于纹理和边缘特征图的道路坑洞分类策略

### 摘要

道路缺陷的存在不仅对个人和社会的资源支出造成了间接影响，还可能导致严重事故和人员伤亡。因此，坑洼道路的检测和识别成为一项极其重要的计算机视觉任务。该任务旨在利用数字图像技术来识别道路上的坑洼，这一技术具有广泛的应用领域，包括地质勘探、航天科学和自然灾害研究等。近年来，随着深度学习技术的迅速发展，为解决坑洼道路检测提供了全新的解决方案。相较于传统的分类算法，深度学习技术在处理复杂和多变的坑洼图像特征方面更具优势。本文中，我们提出了一种在道路图像中使用边缘检测的坑洼道路分类模型。问题重点在于识别哪些特征及其相应的特征组合适合对图像进行有效分类，以确定图像是否包含凹坑。

针对问题一，我们首先对训练集进行预处理，对大小不一的图片均处理成像素大小相同的图片以及对于标签不平衡问题进行数据增强的处理，接着我们使用目标检测算法检测除了坑洼之外的所有其他对象，并且去除检测到的对象并用255填充以作为背景进行处理。然后，通过检测凹坑的边缘来提取特征，我们分别探索了常见图像特征提取方法HOG、LBP对训练集提取相应的特征，随后对这些特征进行融合和去噪处理进行消融实验得到相应的坑洼分类模型。

针对问题二，基于问题一的实践探索我们选取了阈值、卷积-池化层数、学习率、卷积空洞率作为系统关键参数并予以简单描述。对于系统关键参数对识别精度的影响，我们采用了FPS、mAP系数进行定量评估。即把yolov5模型的上述4个系统关键参数值作为初始参数值，在初始参数值的基础上对关键参数分别进行微调，采用控制变量法进行实验，将初始参数模型和调整参数模型得到的损失系数进行横向比较，对模型的有效性和坑洼分类结果进行了准确性评价。结果表明：

**关键词：**道路坑洼；边缘检测；集成学习；目标检测

# 目录

<b>1 问题重述</b>	<b>2</b>
1.1 问题背景	2
1.2 问题重述	3
<b>2 问题分析</b>	<b>3</b>
2.1 问题一的分析	3
2.1.1 问题二的分析	3
<b>3 模型假设与符号说明</b>	<b>5</b>
3.1 模型假设	5
3.2 符号说明	5
<b>4 模型的建立</b>	<b>5</b>
4.1 数据预处理	5
4.1.1 训练集数据增强预处理	5
4.1.2 使用灰度降维	5
4.1.3 使用目标检测的图像预处理	6
4.2 特征提取	6
4.2.1 去噪	6
4.2.2 基于边缘检测的坑洼特征提取	7
4.2.3 HOG 和 LBP 的特征提取	7
4.2.4 特征融合	11
4.3 模型的引入	11
4.3.1 集成学习模型	11
4.3.2 KNN	11
4.3.3 支持向量机	11

## 1 问题重述

### 1.1 问题背景

道路在社会的发展中扮演着至关重要的角色，作为广泛使用的交通媒介，支持着各种业务和人员的流动。然而，道路缺陷，如开裂、车辙、表面磨损和坑洞，对个人、社区、商业组织和国家实体产生了严重的负面影响。这些影响包括机械设备受损、资源支出增加、交通事故风险上升等，因此道路维护变得至关重要。

坑洞的形成通常受到水力作用的影响，通过与高速车辆的接触、雨水和其他杂物的作用，坑洞会不断扩大。这种过程类似于癌症，一个坑洞的形成可能导致同一区域内多个其他故障的产生。修复道路的成本与损坏程度密切相关，随着损坏程度的加剧，维修成本显著增加，这需要国家投入大量资金来改善国内的道路状况。

为了更有效地应对这一问题，引入自动化解决方案变得至关重要。传统的手动报告方法不仅耗时，而且容易出现不一致性。自动化检测和识别过程可以节省大量时间，实现更快速的维修响应，从长远来看，有助

于实现对道路维护的有效支出。这种自动化方法可以有效减少交通事故风险，降低机械设备受损的可能性，并减少资源支出。

## 1.2 问题重述

现已提供301幅图像，在这个问题中，我们希望通过分析、提取特征和建模已标记的道路图像，使得我们能够对一张新的道路图像进行自动坑洼状态识别。

**问题1：**我们需要结合提供的图像文件，提取图像的特征，并建立一个高识别率、快速、准确分类的模型，以便判断图像中的道路是正常的还是有坑洼。

**问题2：**然后，我们会对问题1中构建的模型进行训练，并从不同的角度进行评估，以确保模型的质量和准确性。

**问题3：**最后，我们将利用已经训练好的模型来识别测试集中的坑洼图像，并将识别结果记录在“test result.csv”文件中。

# 2 问题分析

## 2.1 问题一的分析

针对本问题，我们仔细研究了题目所提供的301幅图像，包括图像的大小、分辨率、颜色空间等特征，并且查看了图像的命名标签，以确定哪些图像中存在坑洼，哪些是正常的道路。对于训练集图像大小不一、目标标签不均衡和数据量偏少的问题，我们需要对这301张图片进行相应的预处理。而接下来特征提取的目标是捕获图像中的有关坑洼和道路状态的信息，以便模型可以进行分类。但是我们发现道路正常的图像里存在很多干扰项物体，包括人、汽车、树木、水印、路灯和建筑物等除坑洼之外的物体，这将使得训练出来的模型假负率往往偏高。

为达到正确分类的要求，我们采用了目标检测的方式识别这些干扰项，在分类前预处理中进行“背景化”，为了提高目标检测的效率，我们使用了YOLOv5模型并进行改进。接着我们把提取到的图像特征采用特征级融合的集成学习策略对道路是否坑洼进行投票分类达到高识别率的要求。同时，我们也提出了另一种深度卷积神经网络分类结构，结合了Inception和ResNet两种架构的思想，用于图像分类任务。对于这两种分类模型，我们在后续章节里进行了详细的对比分析。

综上所述，我们基于 YOLOv5 模型提出了一种经改进后的模型 Advanced-YOLOv5，该模型的网络架构与 YOLOv5 相似，但对batchsize和dropout这两个参数进行改进。将这目标检测模型与两种分类模型融合，并对各调整模型的精度进行分析，从而达到题目的要求。本问题的探索思路如下图1a和图1b所示：

### 2.1.1 问题二的分析

针对问题一中提出的两种分类模型，两者均运用到了多尺度特征。不同在于，集成学习模型的精度取决于每个子模型的精度。其中，随机森林的树的数目决定这分类的精度，适当的增加树的数目会提升分类的精度。SVM有一些关键的超参数，如正则化参数（C值）和核函数参数，会影响模型的非线性拟合能力从而影响分类准确率。MLP的架构包括层数、每层的神经元数量、激活函数等，选择合适的神经网络架构对模型性能至关重要。K近邻模型的分类精度受K值以及距离度量方法影响。而深度学习算法精度的提高依赖于每次迭代训练的切片数，是否加入通道注意力机制模块等关键参数。

这里我们建立起参数优化的模型，对于元集成学习分类模型，我们可以从子模型各自的参数改入手对初始模型进行优化。通过网格搜索方法或者绘制学习曲线，以在参数空间中搜索最佳设置。对于深度卷积神

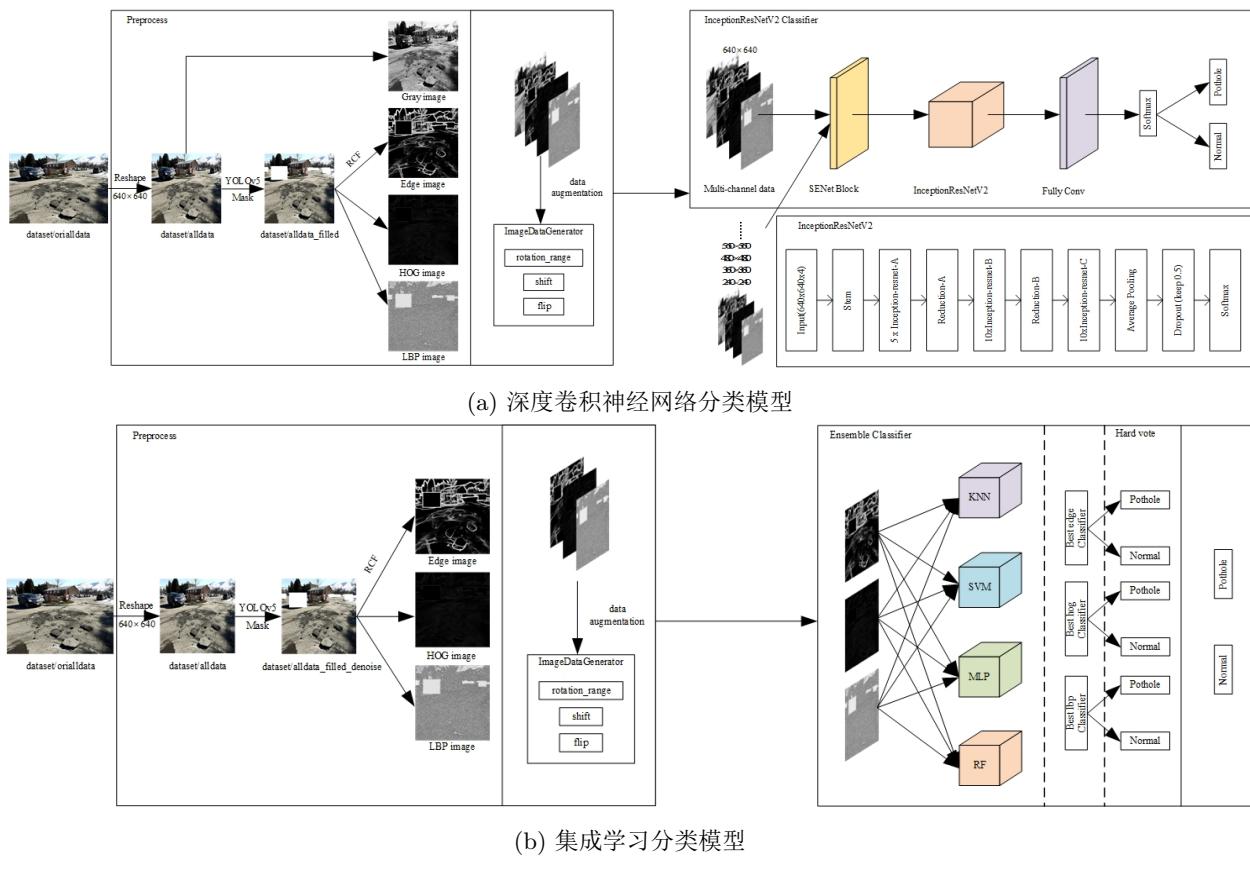


图 1: 我们提出的两种分类模型。

经网络模型，以问题一中训练得到的模型架构和图像输入的大小作为初始参数，在初始参数的基础上对关键参数分别进行微调，采用控制变量法进行消融实验，用模型准确率作为模型精度进行调整前后模型效果的纵向比较并定量评估，同时即可得知关键参数的调整对识别精度的影响。另一方面，这两个分类模型进行横向比较，在相同数据集上的性能、精确度、准确率等指标进行比较。本问题的分析思路如下图2-2所示。

### 3 模型假设与符号说明

#### 3.1 模型假设

- (1) 假设已知我们人工标记标签图足够准确对目标检测对象划分边界框；
- (2) 假设在神经网络预测中，局部图像特征可代表整体特征；
- (3) 假设面积、像素点误差可以作为对象标签目标检测模型的评价标准；
- (4) 假设灰度、形状大小、分布体现坑洼特性；
- (5) 模型假设提供的样本（301幅图像）是代表性的，即这些图像能够很好地覆盖整个问题领域的特征分布。如果样本不够代表性，模型的泛化能力可能会受到影响；
- (6) 假设数据中没有太多的噪声和错误标签。

#### 3.2 符号说明

### 4 模型的建立

#### 4.1 数据预处理

##### 4.1.1 训练集数据增强预处理

鉴于题目所给出的训练集中共266张正常非坑洼道路和35张坑洼道路的图片尺寸大小不一，我们根据模型的输入大小，首先对训练集进行640\*640的图像分割预处理。再通过概率翻转、仿射变化、随机区域的像素丢失、部分图片弹性形变、添加高斯噪声、明亮度改变等方式对数据集进行数据增强，将数据集扩展为原来的3倍。

##### 4.1.2 使用灰度降维

在道路损坏图像中，存在多种物体，包括人、汽车、树木、水印、路灯和建筑物等除坑洼之外的物体。处理这些多通道图像需要大量的计算资源，而坑洼具有一个没有RGB颜色的特征，当大量的计算资源花费在坑洼之外的多通道图像之上，得到的模型将会失去题目原始的目的。所以为了有效分析这些图像，我们对这些多通道图像进行预处理。

首先我们对两类图像中，红色、绿色和蓝色通道的平均强度是否存在显著差异进行了数据探索，这将决定我们是否可以根据它们的平均颜色通道强度对图像进行分类。

在红色和绿色通道的分布图中，它们的平均强度差异不明显，但在蓝色通道的情况下，坑洞类的图像中蓝色通道的平均强度值较高。坑洞类图像的整体颜色通道强度高于普通类，但差异不够显著，作为分类任务的特征图可能会损坏模型的鲁棒性。

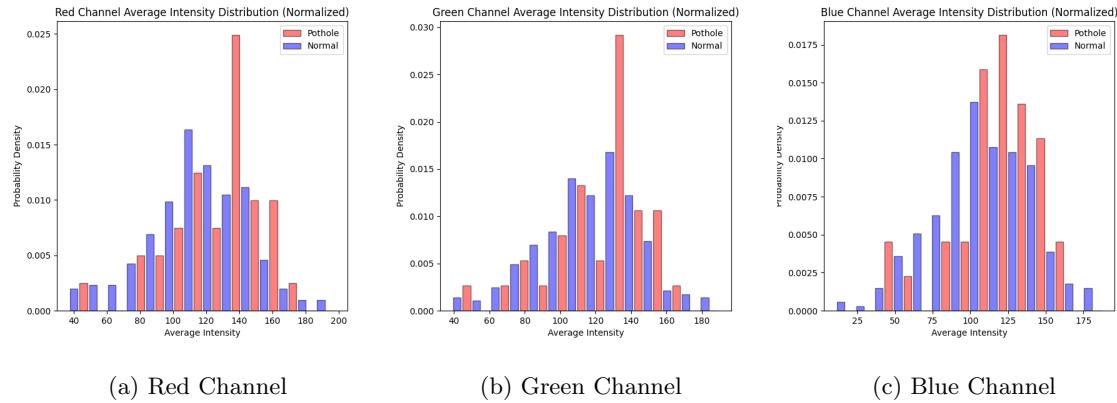


图 2: 不同颜色通道在 Pothole 和 Normal 的强度。原始数据集是不平衡数据, 所以 Normal 类的单通道累积强度会明显高于 Pothole 类, 因此我们计算每个分类下的平均通道强度。

这里我们选择将RGB三通道图像转换为单个通道, 在以最大限度地减少损失的前提下保留坑洼的图像特征。预处理方法包括灰度转换、二值化等。

这里我们选择灰度方法将多通道图像转换为单个通道。在灰度处理后, 图像的颜色信息被舍弃, 只保留了亮度信息, 这样可以减少数据量, 同时在不显著损害图像形状的情况下进行转换。对于坑洼的检测和识别, 我们只关心图像的亮度信息而忽略颜色信息是可以接受的, 这种方法能够简化图像, 使得后续的处理更为高效。

#### 4.1.3 使用目标检测的图像预处理

经过上一步对图像灰度处理, 此时在路面数据中, 坑洼周围有各种各样的物体, 如树木和汽车, 它们已全部转换为灰色, 数据量减少到三分之一但是我们很难分析其特征, 容易受到非坑洼物体的干扰。所以我们接着使用目标检测算法来检测图像中的对象。

这里我们采用YOLO算法对图片进行目标检测。YOLO将图像划分为网格单元, 每个网格单元负责检测其内部的目标。每个网格单元负责预测包围目标的边界框 (Bounding Box), 包括边界框的坐标和置信度分数。它使用边界框坐标来确定目标的位置, 并使用置信度分数来衡量模型对目标的置信程度。每个网格单元将输出目标的位置和类别信息。同时YOLO可使用多层卷积神经网络来提取不同尺度的特征。这有助于检测不同大小的目标, 包括小目标和大目标。为了去除重叠的边界框和重复检测, YOLO使用非极大值抑制来筛选最终的检测结果。

我们通过YOLOv5模型对灰度化的图像进行目标检测, 在检测到的对象中, 除去坑洼之外的对象, 处于该对象的像素值将被改变为255并创建背景。这是因为道路和坑洞的颜色是灰色的, 将背景替换为255可以突出道路和坑洼检测的重要性。因此, 除了坑洼之外的其他物体通过预处理过程被去除。

## 4.2 特征提取

### 4.2.1 去噪

由于大部分特征提取的方法对微小的纹理变化非常敏感, 例如LBP, RCF等, 因此噪声可能会导致较大的负面影响提取出与真实图像不一致的特征结果, 从而干扰模型, 降低模型的分类效果。例如下图是我们采用去噪和未去噪进行特征提取的图像, 可以发现图像的LBP特征对于噪声是极其敏感的, 特别是在噪声水

平较高的情况下。这可能会导致特征提取结果中的虚假性和不稳定性。而具有抗噪声性能的特征提取方法，如局部特征描述符HOG，它们对局部图像区域中的噪声相对不敏感，但是也同样对模型的性能存在一定影响。

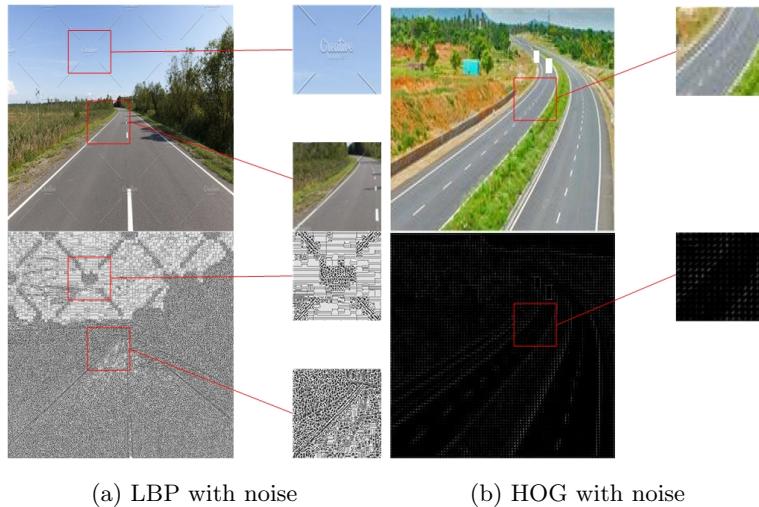


图 3: LBP 图都会受到显著的噪声影响, HOG 图则相对不那么明显。

我们在提取图像特征之前进行去噪，可以确保得到的特征更好地反映图像的本质纹理特征，而不受噪声的干扰。这里我们选择使用标准的中值滤波去噪技术，以减小噪声的影响，同时避免过度处理图像导致信息丢失。

#### 4.2.2 基于边缘检测的坑洼特征提取

在道路损坏检测领域，提取和处理坑洞的特征至关重要。这些特征包括坑洞的边缘类型，如环形边、直线边、阶梯边和渐变边。为了实现对道路损坏的准确检测，边缘检测被广泛采用。边缘检测是一种基于偏微分运算的方法，通过提取与边缘相对应的像素，获取图像中边缘的梯度信息。这些特征和边缘信息通过第一微分和第二差分等图像处理技术进行精细处理，从而为道路损坏的自动检测提供了可靠的基础。

我们采用了RCF (Richer Convolutional Features) 模型对上一节预处理得到的图像进行特征提取。RCF模型是一种基于深度学习的边缘检测模型，它利用深度卷积神经网络 (CNN) 来学习图像的特征，并用于边缘检测任务。相比于经典的边缘检测算法，例如Sobel算子或Canny边缘检测依赖于图像的局部梯度信息来进行边缘检测，RCF模型通过深度学习技术，特别是使用了深度卷积神经网络，能够学习到更丰富、更抽象的图像特征，而不仅仅局限于梯度信息。深度学习模型可以学习到图像中的各种层次的特征，包括纹理、颜色、形状等，从而提高了对边缘的检测能力。RCF模型的深度网络结构使得它能够自动学习到适合于边缘检测任务的特征表示，而无需依赖于手工设计的规则和滤波器，相较于传统的边缘检测算法，它在复杂场景下检测和识别坑洼道路通常表现得更加准确和鲁棒。

#### 4.2.3 HOG 和 LBP 的特征提取

方向梯度直方图 (Histogram of Oriented Gradients) 和局部二进制纹理 (Local Binary Pattern) 都是计算机和图像处理中常用的特征信号。它们都被广泛评估物体检测、人脸识别和纹理分类等任务。

HOG 是一个用于局部图像的形状信息和纹理信息的特征控件。它基于图像的梯度方向来构建直方图。HOG特征表示包含图像中对象的形状和边缘信息的向量，它分析了图像中局部梯度方向的分布。由

于HOG是在图像的局部方格单元上操作，所以它对图像几何的和光学的形变都能保持很好的不变性。一般而言，HOG特征在物体检测（特别是行人检测）中非常有效，它通常与支持向量机（SVM）分类器一起使用。

HOG 主要原理是计算图像的梯度幅值和方向。具体来说先把图像分割个小的连续区域 (Cell)，计算每一个区域梯度方向的直方图，最后合并成更大的区域 (Block) 来对直方图进行归一化，然后合并所有 Block 形成最终的 HOG 特征图。

从图4d和图4e对应的 HOG 特征图可以观察到，HOG特征以梯度的形式捕捉了坑洞的边界，但其边界轮廓并不像它们对应的 LBP 和 Edge 特征图那么明显，相比之下，坑洼的轮廓相对模糊。但是，无论图像中是否存在任何物体，如汽车，行人，HOG 特征图都可以很好的通过其梯度轮廓定位其在图片中的位置。在坑洼图像和非坑洼图像的情况下，我们可以看到坑洼和道路图片捕获的梯度不同，如图4d和图4b。因此，我们可以初步得出结论，HOG 特征图将对我们的分类任务产生影响。

LBP 是一种用来描述图像局部特征的算子，LBP特征具有灰度不变性和旋转不变性等显著优点。LBP 主要工作原理如下：对于图像中的每个像素，考虑其周围的一个领域 (例如， $3 \times 3$  的窗口)，然后将邻域中的每个像素与中心像素进行比较。如果领域像素的值大于或等于中心像素的值，则将其标记为 1，否则标记为 0。这样相邻的所有域都会产生一个二进制模式(例如，11100101)，通过计算所有二进制模式的直方图作为图像的 LBP 特征。

从图4d中我们可以清楚地看到，坑洞区域与汽车一起突出显示。而图4a和图4b对应的 LBP 图像，由于没有坑洞，平原图像上的图案保持均匀，我们可以清楚地观察到其与坑洞的图案的差异。基于这一点，我们认为 LBP 特征图中 normal 和 pothole 所体现出来的差异性，可以胜任我们提出的分类方法。此外 LBP 特征图而言，我们认为 LBP 特征图上的车辆对于我们提出的分类一定的挑战。因为，我们发现图4d对应的 Edge 图和 LBP 图也会受到车辆形状边缘的潜在影响。基于这一点，我们提出了一种解决方法，把图像中的车辆通过目标检测模型识别出来，通过白色掩膜遮挡住，使其与坑洞形状有明显区别，以此减低对模型的要求。我们通过实验证明，通过白色掩膜能够有效的提高模型对于验证数据的分类准确率。

LBP 是一种用来描述图像局部特征的算子，LBP特征具有灰度不变性和旋转不变性等显著优点。LBP 主要工作原理如下：对于图像中的每个像素，考虑其周围的一个领域 (例如， $3 \times 3$  的窗口)，然后将邻域中的每个像素与中心像素进行比较。如果领域像素的值大于或等于中心像素的值，则将其标记为 1，否则标记为 0。这样相邻的所有域都会产生一个二进制模式(例如，11100101)，通过计算所有二进制模式的直方图作为图像的 LBP 特征。

从图4d中我们可以清楚地看到，坑洞区域与汽车一起突出显示。而图4a和图4b对应的 LBP 图像，由于没有坑洞，平原图像上的图案保持均匀，我们可以清楚地观察到其与坑洞的图案的差异。基于这一点，我们认为 LBP 特征图中 normal 和 pothole 所体现出来的差异性，可以胜任我们提出的分类方法。此外 LBP 特征图而言，我们认为 LBP 特征图上的车辆对于我们提出的分类一定的挑战。因为，我们发现图4d对应的 Edge 图和 LBP 图也会受到车辆形状边缘的潜在影响。基于这一点，我们提出了一种解决方法，把图像中的车辆通过目标检测模型识别出来(如图5所示)，通过白色掩膜遮挡住，使其与坑洞形状有明显区别，以此减低对模型的要求。我们通过实验证明，通过白色掩膜能够有效的提高模型对于验证数据的分类准确率。

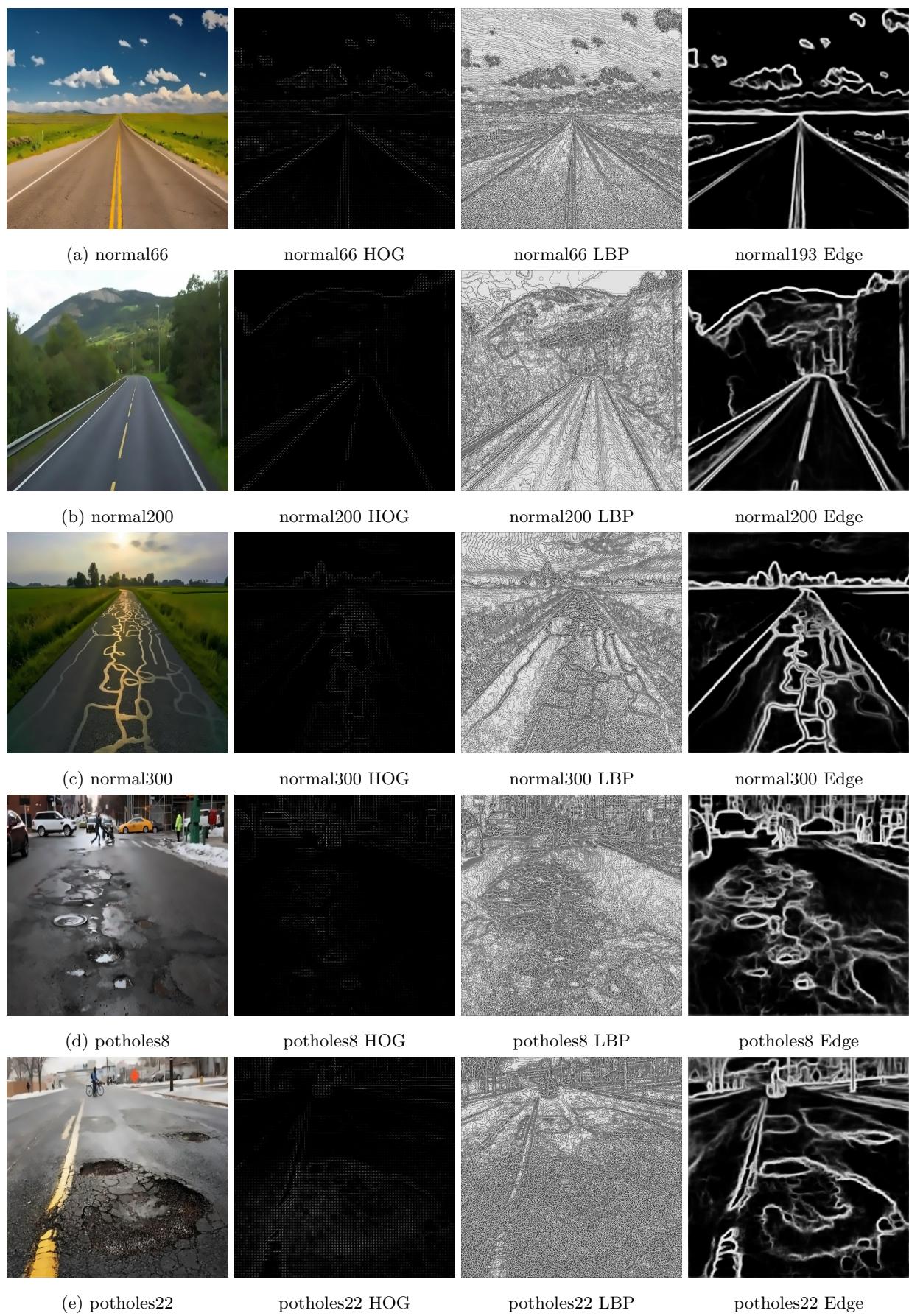


图 4：不同类别图片对应的 HOG, LBP, Edge 特征图。

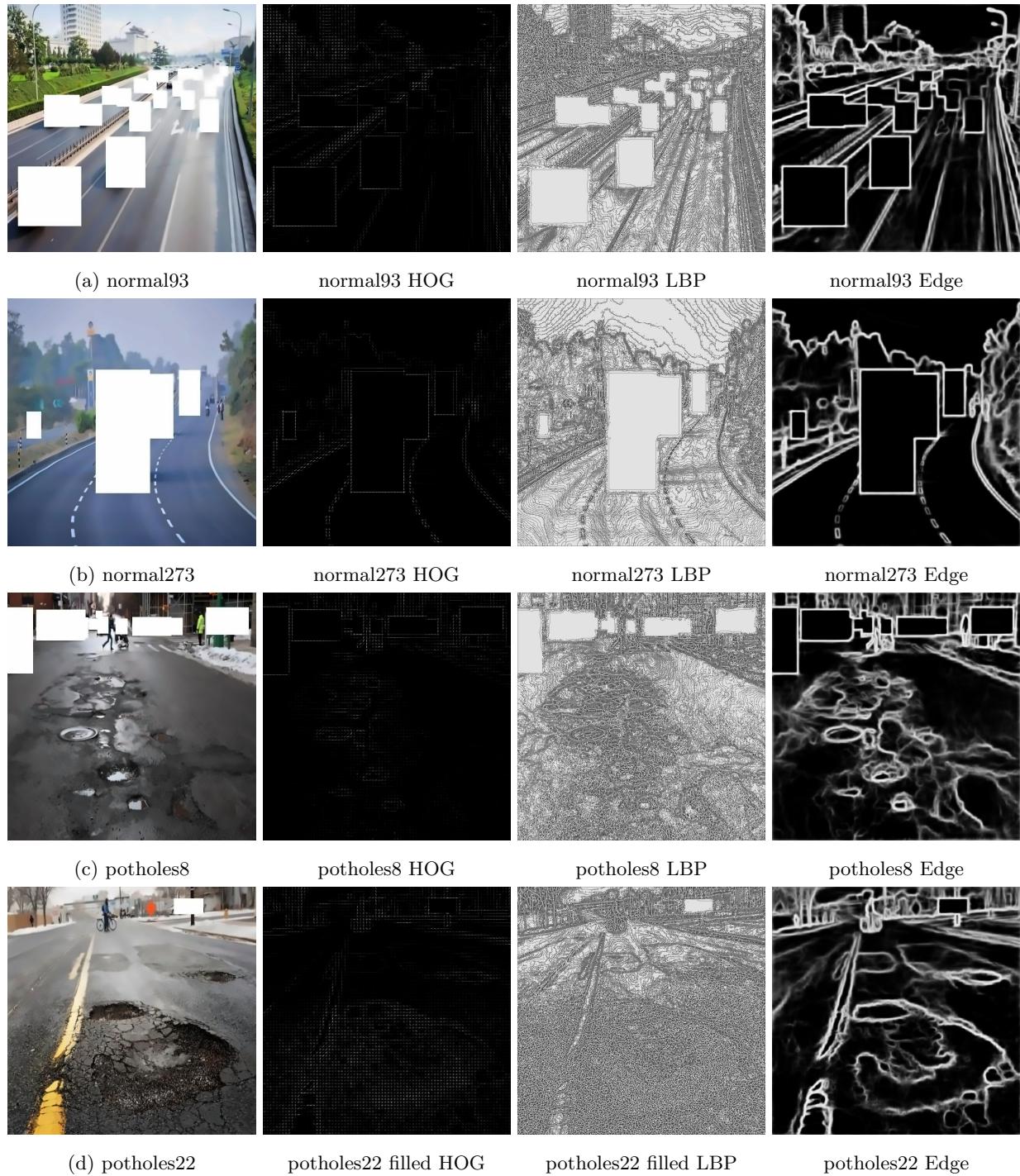


图 5: 为数据集添加白色掩码之后的 HOG, LBP, Edge 特征图。

#### 4.2.4 特征融合

### 4.3 模型的引入

#### 4.3.1 集成学习模型

在前文我们提到我们的特征提取方法，我们通过不同方法提取出图像的灰度，HOG，LBP，边缘特征。我们观察到对于处理多特征数据，集成学习多样化的分类器或许能提供不错的分类效果。

集成学习的有效性依赖于基本模型之间的多样性，旨在通过构建并组合多个预测模型来提升预测准确性和泛化能力。不同的模型可能会从数据中学习到互补的信息，因此，当这些模型被适当地组合时，整体的预测性能可以超越任何单一模型的性能。

集成学习的两个主要类别是 Bagging 和 Boosting：Bagging涉及独立地训练多个模型，并通过简单的投票或平均来组合它们的预测。Boosting涉及顺序地构建模型，其中每个模型都试图纠正前一个模型的错误。

集成学习能够综合多个模型的预测，通常能够达到比单个模型更高的准确率和鲁棒性。

对于我们的分类模型，我们采用 Bagging 方式分别独立的训练 KNN, SVM, MLP, RF，以期待获取不同特征图的最佳分类器，如图1b所示。我们通过把 HOG, LBP, Edge，分别应用到四种不同的分类中，共计12次分类训练，最终构建集成学习模型。

#### 4.3.2 KNN

$$y = \arg \max_{x_i \in N_k(x)} I(y_i, c_j), i = 1, 2, \dots, N; j = 1, 2, \dots, K \quad (1)$$

其中， $I$ 为指示函数：

$$\mathbb{I}(x, y) \begin{cases} 1, & \text{if } x = y, \\ 0, & \text{if } x \neq y. \end{cases} \quad (2)$$

在K最近邻（KNN）算法中，待预测的输入实例  $\mathbf{x}$  的类别  $y$  可以通过以下公式确定：

$$y = \arg \max_c \sum_{i=1}^k \mathbb{I}(y_i = c) \quad (3)$$

对于一个待分类的输入实例  $\mathbf{x}$ ，算法的目标是确定实例  $\mathbf{x}$  的类别标签。根据预先选定的距离度量（通常采用欧几里得距离度量），在训练数据集  $\mathcal{T}$  中识别出与实例  $\mathbf{x}$  距离最近的  $k$  个训练实例。这  $k$  个实例构成了集合  $N_k(\mathbf{x})$ 。接着，实例  $\mathbf{x}$  的类别标签  $y$  由  $N_k(\mathbf{x})$  中出现频率最高的类别标签决定，这一决策过程基于多数投票规则。

#### 4.3.3 支持向量机

支持向量机（SVM）是一种用于分类的监督学习算法，旨在通过确定一个最优超平面来区分不同类别的数据点，使得两个类别之间的间隔最大化。SVM的核心概念是最大化支持向量到决策边界的距离，其中支持向量是距离分隔超平面最近的那些数据点。SVM算法试图解决以下优化问题：

$$\max_{\mathbf{w}, b} \left\{ \frac{1}{\|\mathbf{w}\|} \min_{(\mathbf{x}_i, y_i) \in D} y_i (\mathbf{w} \cdot \mathbf{x}_i + b) \right\} \quad (4)$$

$\mathbf{w}$  是超平面的法向量， $b$  是超平面的偏置项， $D$  是训练数据集，其中 $\mathbf{x}_i$ 是特征向量， $y_i$ 是对应的类别标签，取值为+1或-1， $\|\mathbf{w}\|$  是 $\mathbf{w}$ 的范数，其倒数代表了类别间的间隔。

## 参考文献