

# 灵敏度分析和分类算法对比（李越师姐小论文）

- abstract
  - 地下目标探测中目标分类的重要性
  - 论文主要内容
    - 椭球模型和近似正演模型灵敏度分析，目标特性对输出的影响
    - 仿真数据加噪声，利用拟合算法拟合模型参数，并分析拟合算法在不同信噪比下对分类性能的影响
    - 对比8种分类算法
  - 结论
    - 椭球模型中，半径比磁导率对模型参数的影响更大
    - 近似正演模型中，目标方向角度未知，对目标进行分类不可行
    - 拟合算法受噪声影响较大
    - 基于形状的分类效果比基于材质的分类效果好
    - 神经网络算法在8中分类算法中，分类性能最好
    - 基于材质的分类中，分类器具有较弱的区分导电性目标的能力
- 一、Introduction:
  - 论文目标：提出系统性的地下目标分类的方法，提高分类的精确度。通过分析整个分类系统，并对分类过程中对分类效果有影响的因素进行分析。
  - 论文内容：基于目标的磁场衰减响应模型中的参数，对目标的形状和材质（是否导电，通过相对磁导率判断）进行分类，分析影响分类效果的因素。
  - 具体方法：
    - 1.对椭球模型和近似正演模型应用灵敏度分析。对于椭球模型分析目标的属性（形状、材质）与模型参数之间的关系，结论：椭球半径比椭球材质对模型参数的影响更大。对于近似正演模型评估目标的属性和位置对模型参数的影响，结论：目标方位对模型参数的影响是不确定的，在一些情况下目标方位属性比其他参数的影响大，若想达到更好的分类结果，必需获取目标方位属性。无位置属性的分类是不可行的。
    - 2.模拟仿真：通过对仿真数据添加不同的噪声，然后利用拟合算法获取模型参数。通过比较分类结果，分析不同信噪比下的拟合算法对分类结果的影响。结论：随着信噪比的增加，利用拟合算法得到的拟合数据，分类效果逐渐变好。拟合算法对噪声水平比较敏感。
    - 3.将模型参数作为分类特征利用多种分类算法进行分类，将分类结果利用多种评估方法（准确率、混淆矩阵、召回率）评估，比较分析这些分类算法的分类效果。结论：大多数分类算法对材质的分类效果不好，而对目标形状的分类效果要比目标材质的分类效果要好，其中神经网络算法对在对材质和形状分类时具有更佳的表现。
  - 创新点：

- 1.在椭球体模型中应用灵敏度分析，分析目标物体的属性（形状和材质）对模型参数的影响
- 2.在近似正演模型中应用灵敏度分析，分析目标物体的属性（形状）与方位对模型参数的影响
- 3.分析了噪声对拟合参数的影响以及拟合算法的影响
- 4.以准确率、召回率、混淆矩阵为评价指标比较了8种分类算法的效果

## • 二、related work :

- 引用一些文献对目前的地下目标分类的研究进行介绍，主要从基于模型的分类与基于数据的分类展开，最后提出这些分类的方法很多都是针对某种特定情况下的优化，进而提出本文的研究点，通用的目标分类系统，分析该系统中可能影响分类结果的因素。

## • 三、System model:

- 有3个模型，通用模型、椭球模型、近似正演模型，通用模型适用任何形状的物体，椭球模型适用轴对称的物体，近似正演模型接近真实测量场景，在基于椭球模型的基础上考虑了目标物体的方位。椭球模型作为主要的模型对目标物体的极化率张量建模。

### • 通用模型

- 介绍了地下金属目标探测的原理，发射一次场在金属物体附近产生涡流，由于物体的有限的导电性，所以产生衰减的电流，二次场是由衰减的电流产生。
- 二次场中有一个关键参数，目标物体的感应磁矩，感应磁矩是由一次场产生，和一次场线性相关。将目标物体看作磁偶极子，二次场的响应等效为三个正交偶极子的响应，每个偶极子的方向与一次场的三个正交方向平行。感应磁矩与一次场的关系中，存在一个时间衰减响应，而极化率张量中包含着这三个方向的时间衰减响应的信息，而这个信息与物体的材质和形状相关。感应偶极子的强度和一次场强度与极化率张量的乘积成比例。

### • 椭球模型：

- 目标物体在每个方向的衰减响应是每个正交轴响应的线性组合，如果物体在三个正交轴方向的半径不同，则其响应就不同。对于椭球体而言，只有长轴和垂直于长轴两个方向的半径不同，故其在三个方向的响应中，其中两个响应是相同的。这种衰减响应描述为无穷个指数的叠加，经过转化为更小的参数的函数式得到极化率张量的参数求取公式，亦即衰减响应的求取公式，这种衰减表现为3个阶段，初期衰减，中期的线性衰减以及后期的指数衰减。大多数物体近似为椭球体模型，进而可以得到椭球体在不同方向的磁极化率的模型。这部分主要介绍了椭球的衰减曲线模型及参数定义。

### • 近似正演模型：

- 基于椭球模型，得到了轴对称物体的二次场响应的表达式，同时为了更加逼真的模拟真实场景，将目标物体的方位信息引入，改进椭球体模型，得到近似正演模型。

- 由于观测坐标系和目标物体响应的坐标系不同，因此通过坐标转化得到目标坐标系下的物体响应。

#### • 四、Sensitivity analysis

- 主要为了量化目标属性对模型参数的影响。灵敏度分析中基于微分的方法主要基于输入变量与输出变量的微分来量化输入参数对模型输出的重要性。

#### • 五、CLASSIFICATION ALGORITHM AND METRICS

- 这部分主要分析了现存的分类的机器学习的算法，主要包括监督、无监督、半监督，本文主要使用监督算法，介绍了监督算法分类的流程：数据采集、数据预处理、数据训练、模型验证。同时指出评估标准：精确率、召回率、混淆矩阵。
  - 介绍监督的分类算法，主要从判别式、生成式、集成学习三个大方面对本文涉及的8种分类算法成组归纳，然后依次简单介绍各个算法，后面介绍了评估的标准。

#### • 六、SIMULATION VERIFICATION

##### • A .SENSITIVITY ANALYSIS RESULTS

- 这部分主要介绍前面内容的结果，首先介绍了灵敏度分析的结果，分两部分：基于椭球体的，另一部分是基于近似正演模型的。

##### • 1) SPHEROID MODEL

- 对于椭球体模型主要判断目标半径、目标相对磁导率以及二者的相互作用对输出的影响，结果表明半径比相对磁导率对输出参数的影响大，二者的相互作用对输出的影响很小。而半径与目标的形状直接相关，相对磁导率在很大程度上可以表示目标的材质。因此根据灵敏度分析的结果，推断对目标形状的分类效果要比材质的效果好。
- 另外映射不同参数( $K$ ,  $\alpha$ ,  $\beta$ )下的数据分布，展示参数和目标属性之间更直观的关系。
- $K$ 可以区分导电性材料和非导电性材料，但是对于导电性材料的区分较弱，另外当目标的材质确定后，可以大致区分目标的形状。 $\alpha$ 区分目标材质的能力强于 $K$ ，可以直接区分形状。 $\beta$ 可以较好的区分目标形状，但是很难区分目标材质。综上利用这三个参数很难对目标材质进行分类，从数据分布中也可以看出对形状的分类效果要好于对材质的分类。

##### • 2) THE APPROXIMATE FORWARD MODEL

- 基于椭球模型的基础上加入方向参数的影响，通过灵敏度分析如果方向的影响比轴半径以及水平半径的影响大，则说明方向未知对目标的分类是不可行的。主要就是比较上述内容对近似正演模型参数的影响，所以在生成数据时令相对磁导率以及时间参数 $t$ 固定。计算目标的 $\theta$ 、 $\phi$ 角度、轴半径、赤道半径对模型输出参数的影响，结果表明： $\theta$ 和 $\phi$ 对模型输出有较大的影响。
- 另外比较了 $\theta$ 和 $\phi$ 的一阶指数和随目标纵横笔记的变换趋势，通过该变化趋势表明在目标方向角度未知的情况下，无法有效的进行分类。目标的方向角度可以通过反演方和的滤波估计获得。

- 通过仿真数据的可视化分析验证上述结论：目标方向角度未知无法正确分类。仿真数据是在材质、目标形状确定的情况下设置不同的滚转角度。可视化结果表明：不同的滚转角度下 $k, \alpha, \beta$ 的变化是无规则的，所以证实了对未知方向的对象进行分类是不可行的。

## • B. INFLUENCE OF ORIENTATION ON CLASSIFICATION RESULTS

- 这部分内容主要对上面近似正演模型中灵敏度分析结果中对未知方向的对象进行分类是不可行的这一结果进行验证，通过利用8种分类算法去验证未知目标方向角度下的分类，并比较这8种分类算法。验证的结果出现在对目标材质分类时，分类精度较高，这与前面的灵敏度分析的结果不一致，原因：对于这种情况，最初假设在拟合模拟数据时存在误差。为了精确地估计模型参数，获得目标参数的上下限是必不可少的。然而，在目前的研究中无法获得准确的界限，因此我们将界限设置得更宽。这导致拟合算法拟合的仿真数据可能不精确——仿真参数无限逼近上界或下界，因此分类精度与预期情况不一致。

## • C. CLASSIFICATION RESULTS COMPARISON

- 首先介绍了分类的流程：首先通过椭球体模型生成仿真数据以及真实的模型参数 $k, \alpha, \beta$ 。然后通过对仿真数据添加不同的噪声，利用拟合算法（非线性最小二乘）得到拟合的模型参数。然后利用拟合的模型参数作为对目标材质和形状分类的属性，利用8种分类算法进行分类，然后对分类的结果进行评估，最后从拟合算法和分类算法的角度分析了他们对分类结果的影响。

### • 1) INFLUENCE OF FITTING ALGORITHM ON CLASSIFICATION PERFORMANCE

- 分析了不同信噪比下通过拟合算法得到的模型参数的误差的结果，结果显示噪声对拟合模型参数的精度有较大影响。信噪比较大时，可以精确的拟合导电性目标的模型参数，但是无法精确的拟合非导电性目标。
- 分析了从基于材料的分类和基于形状的分类在不同信噪比下的分类精确度，高信噪比下利用拟合算法获得的参数进行分类情况比较好，因此可以看出拟合算法性能对分类精确度有很大影响，要想正确分类，需要提高拟合算法的在低信噪比下的结果。而且结果也表明相同信噪比下对形状的分类效果要好于对材质的分类效果，这验证了灵敏度分析的结果。总的来说，噪声对分类算法具有一定的影响，这可以作为后续提高分类精度的研究点。

### • 2) INFLUENCE OF DIFFERENT CLASSIFIERS ON CLASSIFICATION PERFORMANCE

- 从与样本分布相关的精确率、召回率和准确率评估不同分类算法的分类效果，以选择最优的分类器，并在今后的工作中进一步增强其分类能力。
- 召回率和准确率是相矛盾的两个量，而在地下目标探测中召回率是更重要的评价指标。
- 基于对材质的分类中：最终的结果表明对铝的分类中，神经网络和adaboost具有更高的精度，同时确保精度和召回率之间的差距更小。对钢的分类中，神经网络算法的表现最佳。对镍的分类中，准确率和召回率比铝和钢的低，8种算法中神经网络算法表现最佳。因此神经网络算法在对材质进行分类的过程中表现较好，对镍的分类效果最差。



- 基于对形状的分类中：神经网络算法的精确度最高。精确率和召回率都比较高，相对于材质的分类，对形状的分类效果更好。
- 但是对于基于材料的分类，由于标签被分为三类，对于误分类的样本，分类器所预测的类别没有直观的显示，所以用改进的混淆矩阵（以比例表示，而不是以数量表示，均一化）评价了基于材料分类的分类性能。改进的混淆矩阵对于样本不均衡时，对分类的效果具有一个更直观的表现。结果表明对镍的分类效果比铝和钢要差，分类的结果是把镍分为了钢。对钢的分类效果要好于镍，虽然存在错误分成镍的情况，对铝的分类效果最好。大多数分类器具有分辨导电性材质和非导电性材质的能力。
- 对四种分类效果的分类算法进行了时间复杂度分析，通过比较不同数据量的情况下，4种分类器的执行时间，结果表明除了QDA外，随着数据量的增加，执行时间都变长，虽然QDA受数据集的影响不敏感，但是他的分类效果是最差的。SVM-PK和GP的执行时间随着数据集的增大而迅速提高，而神经网络的分类效果最好，并且即使数据集很大，执行时间也在3秒以内。
- 小结：通过对分类的精确度、准确率和召回率、混淆矩阵以及时间复杂度进行比较，神经网络算法在对目标材质和形状的分类中效果最佳。对于大多数的分类器，对目标形状的分类效果要好于对材质的分类效果。将来的研究中，可以通过优化神经网络的参数去提高分类效果，特别是镍的分类效果。

## • 七、CONCLUSION AND FUTURE WORK

- 通过对模型的灵敏度分析分析了对目标分类的过程中可能影响分类结果的因素，分析了拟合算法的影响，比较了八种分类算法的分类效果。
  - 对椭球模型的灵敏度分析，结果表明半径R比磁导率对模型输出的影响更大。
  - 对近似正演模型的灵敏度分析，结果表明当目标方向未知时，无法对目标进行分类。
  - 为了评估拟合算法的影响，利用带噪声的数据通过拟合算法拟合模型参数作为分类的特征，比较8种分类算法在对材质和形状的分类中的准确度，结果表明拟合算法对噪声强度比较敏感，通过比较8种分类算法在对目标材质和形状的分类表现，结果表明神经网络算法比其他算法表现更佳。
- 将来的工作中，在分类之前获取目标的方向角度信息是必需的，而且需要进一步提高拟合算法在高噪声环境下拟合模型参数的准确性，最后将来改进基于目标材质的分类，特别是优化神经网络的参数以提高其区分导电性材料的性能。

•  
•  
•

## • 一些之前没关注的点：

- 关于课题：
  - 地下金属目标分类的研究主要集中在两个方面:基于模型的方法和基于数据的方法。
    - 对于基于数据的方法，目标对象的分类是通过分析目标对象的磁场衰减响应的值来实现的。（目标衰减响应的值）但并不是所有的衰减响应数据都是对

分类有用的，需要加以筛选（降维）。

- 基于模型的方法是根据目标物体的磁场衰减响应建立模型，然后根据模型参数对材料和形状进行分类。（目标响应模型的参数）

- 李越师姐使用的基于模型的方法

- 分类算法：

- 包括监督学习与无监督与半监督学习，都要考虑，李越师姐使用的是无监督学习。

- 关于平台：

- 提供关于地下目标分类采用机器学习的方式的整体框架以及不同类的（监督、无监督、半监督）机器学习算法的详细的框架，监督学习的框架可以参考李越师姐的（Figure5）

- 关于论文写作

- 通用

- introduction部分要写的有条理，加一些背景引用，最终目的是通过research gap引出研究内容，然后介绍自己的研究方法，创新点。
- 一些课题专业词语进行了标记，后面写的时候注意不要使用错误
- 内容组织的要有逻辑，不要写的太乱，每个环节之间都要有联系

- 更具体的

- 关于地下目标探测的分类方式包括基于模型与基于数据的方法，具体的机器学习的分类算法分为监督与非监督以及半监督学习算法，这些都要考虑进来

- 看完这篇论文，相当于对地下目标分类的内容，有了一个更加清晰的认识。