

Universidade Federal do Rio Grande do Norte

Aluno: Thiago Theiry de Oliveira

Disciplina: APRENDIZADO DE MÁQUINA

Professor(a): ANNE MAGALY DE PAULA CANUTO

Check-point – Deliverable

1. PROBLEMA

O presente trabalho tem como objetivo realizar a segmentação indireta de imagens sísmicas a partir de um classificador supervisionado. O problema aborda a identificação de regiões geológicas (bacias sedimentares, crosta oceânica, crosta continental, manto e água) em dados sísmicos, utilizando uma abordagem baseada em classificação de patches.

Devido à baixa performance de modelos de segmentação direta como U-Net nos dados disponíveis, adotou-se uma estratégia alternativa: a imagem sísmica é dividida em pequenos patches de tamanho fixo (64×64), cada um rotulado com base na classe predominante em uma máscara anotada por especialistas. Após o treinamento do classificador, uma nova imagem é igualmente dividida e classificada patch a patch. A reconstrução da segmentação final ocorre por meio da recomposição dos patches classificados com as respectivas cores das classes.

Essa abordagem busca contornar dificuldades encontradas em abordagens anteriores.

2. DADOS

A base de dados utilizada é composta por uma única imagem sísmica armazenada em um arquivo .npy, com dimensões originais de (2326, 23777). Para possibilitar o treinamento do modelo, essa imagem foi dividida em patches de 64×64 pixels, resultando em 13.356 amostras.

Cada patch foi rotulado com base em uma máscara anotada manualmente, considerando a classe com maior predominância de pixels no patch. As classes geológicas mapeadas são:

- Água
- Bacia sedimentar
- Crosta continental
- Crosta oceânica
- Manto

Resumo da base:

- **Número de instâncias:** 13.356 patches
- **Número de atributos:** $64 \times 64 = 4.096$ valores por instância
- **Número de classes:** 5
- **Atributos categóricos:** 1 (classe do patch)
- **Atributos numéricos:** 4.096 (valores de amplitude sísmica por patch, contínuos)
- **Intervalos dos atributos numéricos:** Varia conforme os dados sísmicos, mas representam amplitudes reais de reflexão e são contínuos.

Distribuição das classes:

- Água: 3.107 patches
- Bacia sedimentar: 3.929 patches
- Crosta continental: 4.036 patches
- Crosta oceânica: 629 patches
- Manto: 1.655 patches