

Implementação de um modelo de classificação de áreas irregulares na Floresta Amazônica baseado em Transformers Visuais

Victor Moraes

UFMG

2022

Motivação



Figura:

Garimpo ilegal na Terra Indígena Munduruku, município de Jacareacanga. Foto: Marizilda Cruppe/Amazônia Real

Revisão bibliográfica

- ▶ Domínio do problema: Sensoriamento remoto
- ▶ Aprendizado de máquina
- ▶ Redes convolucionais
- ▶ Transformers
- ▶ Trabalhos anteriores

Revisão bibliográfica - Aprendizado de máquina

Definição: “Campo de estudos que visa a dar computadores a habilidade de aprender sem serem explicitamente programados para determinada tarefa.” (??)

“Um algoritmo dito conseguir uma experiência E com respeito a determinada classe de tarefas T e com medidas de desempenho P , se seu desempenho nas tarefas em T , medidas por P , melhoram a partir da experiência E .” (??)

Revisão bibliográfica - Aprendizado de máquina

Definição: “Campo de estudos que visa a dar computadores a habilidade de aprender sem serem explicitamente programados para determinada tarefa.” (??)

“Um algoritmo dito conseguir uma experiência E com respeito a determinada classe de tarefas T e com medidas de desempenho P , se seu desempenho nas tarefas em T , medidas por P , melhoram a partir da experiência E .” (??)

Proposta

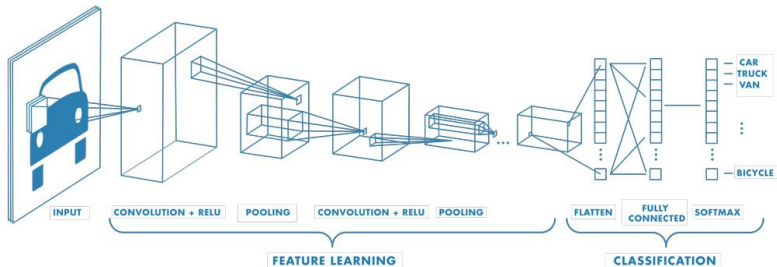


Figura: Arquitetura de uma rede convolucional. Filtros extratores de características são aplicados em diferentes resoluções e campos visuais. A saída de cada imagem convoluta alimenta a próxima camada. As ultimas camadas completamente conectadas realizam a classificação.

Método

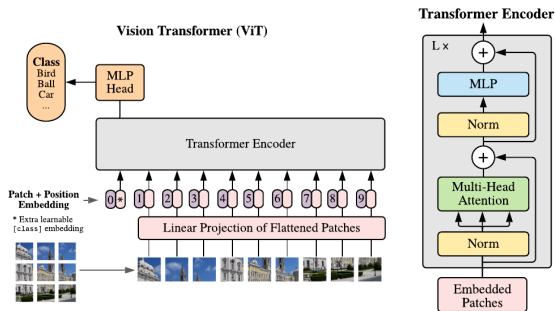


Figura: O ViT divide uma imagem em uma grade de recortes quadrados, cada fragmento é achatado em um vetor único contendo todos canais de todos os pixels, e projetando-os em uma dimensão de entrada desejada, alimentando a camada de múltiplos encoders em paralelo. (??)

Resultados e discussão

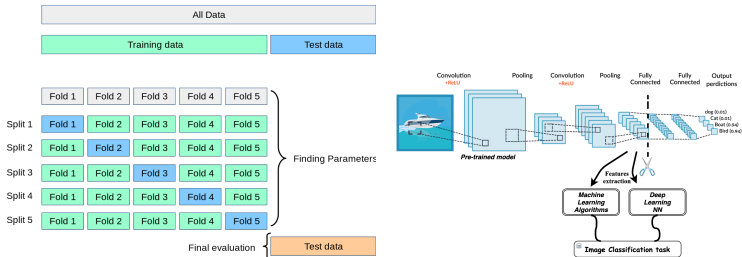


Figura: O ViT divide uma imagem em uma grade de recortes quadrados, cada fragmento é achatado em um vetor único contendo todos canais de todos os pixels, e projetando-os em uma dimensão de entrada desejada, alimentando a camada de múltiplos encoders em paralelo. (??)

Conclusão

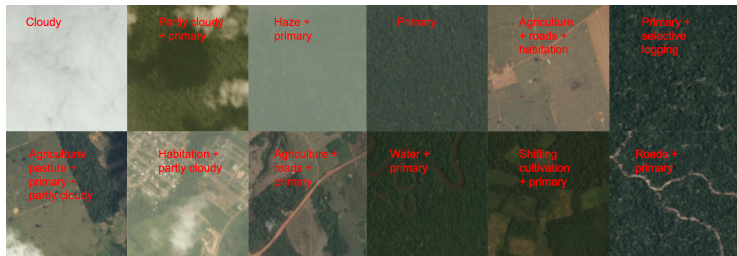


Figura: Amostras de classes do dataset Amazônia do espaço