

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

Aprendizado de máquina - Definições

“Campo de estudos que visa a dar computadores a habilidade de aprender sem serem explicitamente programados para determinada tarefa.” (SAMUEL, 1959)

“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 .” (MITCHELL, 1997)

Aprendizado de máquina - Exemplo

Exemplo:

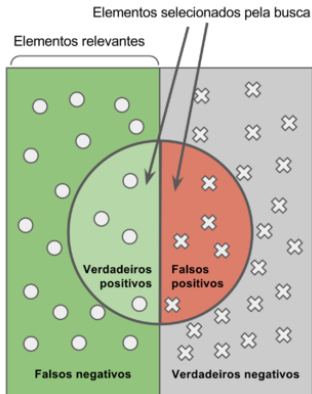
*“Um algoritmo dito conseguir uma experiência E com respeito a determinada classe de tarefas T e com medidas de desempenho P , se seu desempenho nas tarefas em T , medidas por P , melhoram a partir da experiência E .”
(MITCHELL, 1997)*

- ▶ Tarefa T : Problema de classificação
- ▶ Experiência E : Amostras dentre k classes a classificar
- ▶ Desempenho P : acurácia do algoritmo

Determinar função:

$f : \mathbb{R}^n \rightarrow \{1, \dots, k\}$. Quando $y = f(x)$

Métricas de desempenho



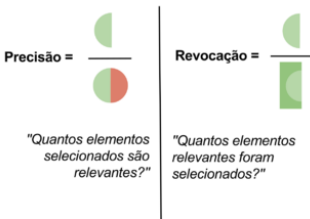
		Detectada	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

$$P = \frac{vp}{vp + fp} \quad (1)$$

$$R = \frac{vp}{vp + fn} \quad (2)$$

$$F_\beta = (1 + \beta^2) \times \frac{P \times R}{\beta^2 P + R} \quad (3)$$

$$F_2 = 5 \times \frac{P \times R}{5P + R} \quad (4)$$



Proposta

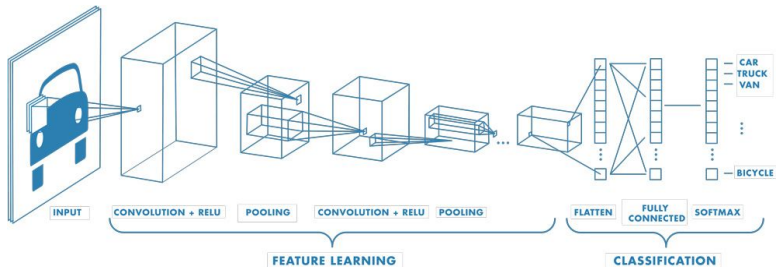


Figura: Arquitetura de uma rede convolucional. Filtros extratores de características são aplicados em diferentes resoluções e campos visuais. A saída de cada imagem convoluta alimenta a próxima camada. As ultimas camadas completamente conectadas realizam a classificação.

Método

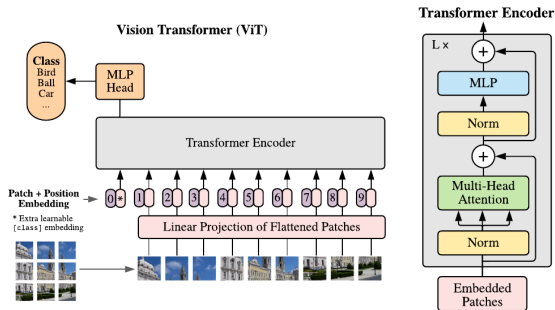


Figura: O ViT divide uma imagem em uma grade de recortes quadrados, cada fragmento é achatado em um vetor único contendo todos canais de todos os pixels, e projetando-os em uma dimensão de entrada desejada, alimentando a camada de múltiplos encoders em paralelo.
(DOSOVITSKIY et al., 2020)

Resultados e discussão

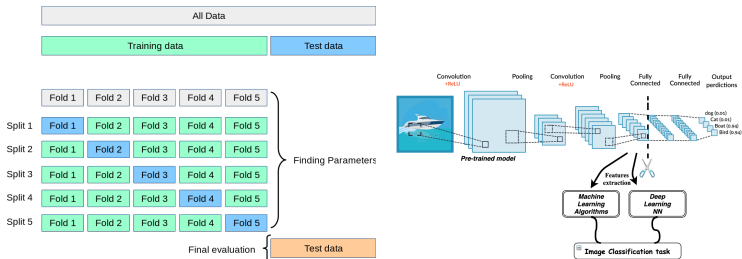


Figura: O ViT divide uma imagem em uma grade de recortes quadrados, cada fragmento é achatado em um vetor único contendo todos canais de todos os pixels, e projetando-os em uma dimensão de entrada desejada, alimentando a camada de múltiplos encoders em paralelo.
(DOSOVITSKIY et al., 2020)

Conclusão

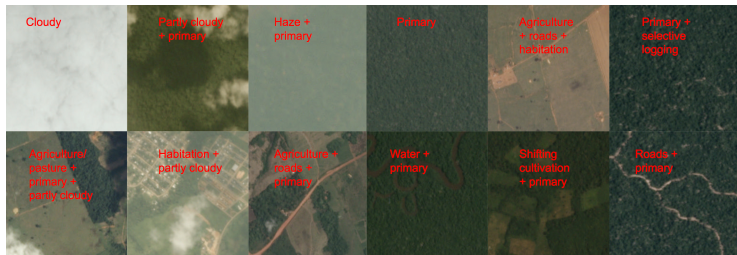




Figura: Amostras de classes do dataset Amazônia do espaço

 DOSOVITSKIY, A. et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint*


arXiv:2010.11929, 2020. 9, 10

 INSTRUTOR.GIS. *QGIS: Satélite Amazonia-1 – Composição Colorida RGB*. 2022. Disponível em: <https://www.instrutorgis.com.br/qgis-satelite-amazonia1-composicao-colorida-rgb>.

4

 MITCHELL, T. M. *Machine Learning*. New York: McGraw-Hill, 1997. ISBN 978-0-07-042807-2.

5, 6

 SAMUEL, A. L. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.*, v. 3, p. 210–229, 1959.

5