

Segmentação Semântica com U-Net utilizando dados do PRODES

Matheus Cavassan Zaglia, Antonio Maurício Zarzur

CAP-378 Tópicos em Observação da Terra
Instituto Nacional de Pesquisas Espaciais

mzaglia@gmail.com, mauricio.zarzur@gmail.com

13 de setembro de 2018

Agenda

- 1 Introdução
- 2 Motivação
- 3 Machine Learning (Aprendizado de Máquina)
- 4 Materiais e Métodos
 - Dataset
 - Modelo U-Net
- 5 Resultados

- Detecção de desmatamento através de imagens de satélite tem sido a principal ferramenta para o combate ao desmatamento ilegal;
- Gera dados para estudos de biodiversidade, mudanças climáticas e modelagem ambiental;
- Projeto PRODES do INPE mapeia e quantifica o desmatamento na Amazônia Legal Brasileira desde a década de 80;
- Metodologia baseada em interpretação visual, realizada por especialistas em sensoriamento remoto.

- Avanço nas tecnologias de inteligência artificial e aprendizado de máquina;
- 9x mais artigos publicados com a palavra-chave "Inteligência Artificial";
- Melhora nos algoritmos resultando em uma diminuição no erro em rotulações de 28,5% para menos de 2,5% desde 2010;
- Avançar o conhecimento em metodologias para detecção automática de desmatamento em imagens de sensoriamento remoto.

- Permite que computadores aprendam através de exemplos, sem serem programados previamente;
- Maneira eficiente de utilizar modelos e algoritmos em grandes quantidades de dados;
- Escolha de algoritmos e modelos variam de acordo com o objetivo.

Redes Neurais

- Se inspiram na forma como o sistema nervoso processa informações;
- São formadas por elementos processadores (neurônios) interconectados que trabalham em paralelo;
- Treinamento realizado a partir de conjuntos de entradas e saídas esperadas.

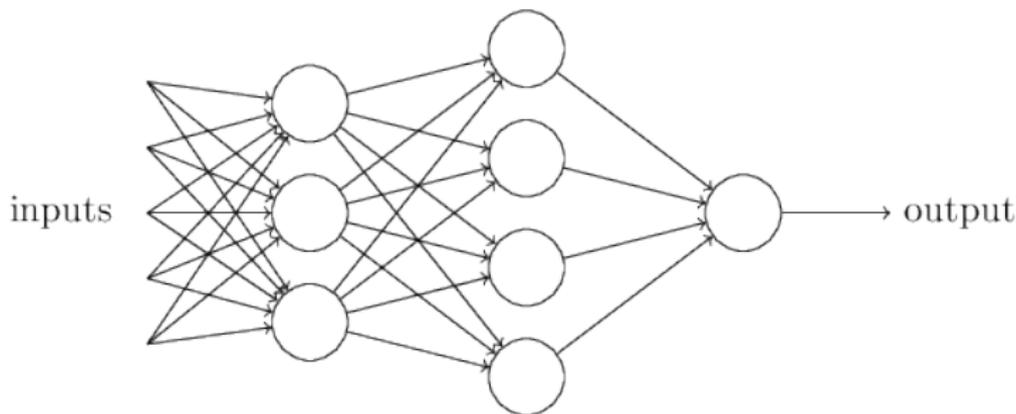


Figura: *Neural Network* Nielsen 2015

- São uma classe específica de redes neurais multicamadas;
- Conexão de neurônios segue o padrão *feedforward*;
 - neurônios de uma camada conectados à camada seguinte;
 - informação flui em um único sentido (entrada → saída);
 - não há formação de ciclos entre as camadas.
- Frequentemente utilizadas em análise e segmentação de imagens.

- Rede neural convolucional originalmente desenvolvida para segmentar imagens biomédicas;
- Redes convolucionais estão apresentando resultados superiores ao estado da arte em tarefas de reconhecimento visual;
- Realiza segmentação end-to-end – para cada imagem de entrada, a saída é uma segmentação;
- Especificamente modificada para trabalhar com um número menor de imagens no treinamento.

U-Net

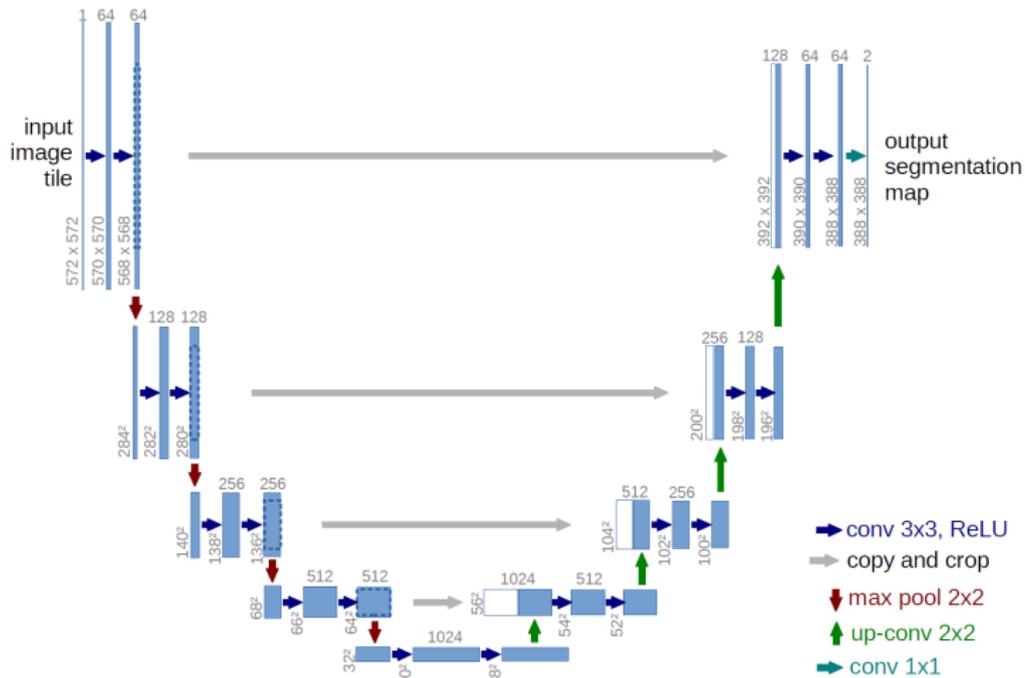


Figura: Arquitetura U-Net (Ronneberger, Fischer e Brox 2015)

- Dados PRODES;
- Implementação do modelo U-Net;

- Composto por 15 cenas LANDSAT 8 e suas respectivas máscaras adquiridas no portal PRODES;

Órbita	Ponto
226	64, 65, 66
227	64, 65, 66, 67, 68, 69
228	64, 65, 66, 67, 68, 69

Tabela: Conjunto de cenas utilizadas

Dataset

- Foram criados quatro mosaicos, dois para as cenas e dois para as máscaras.

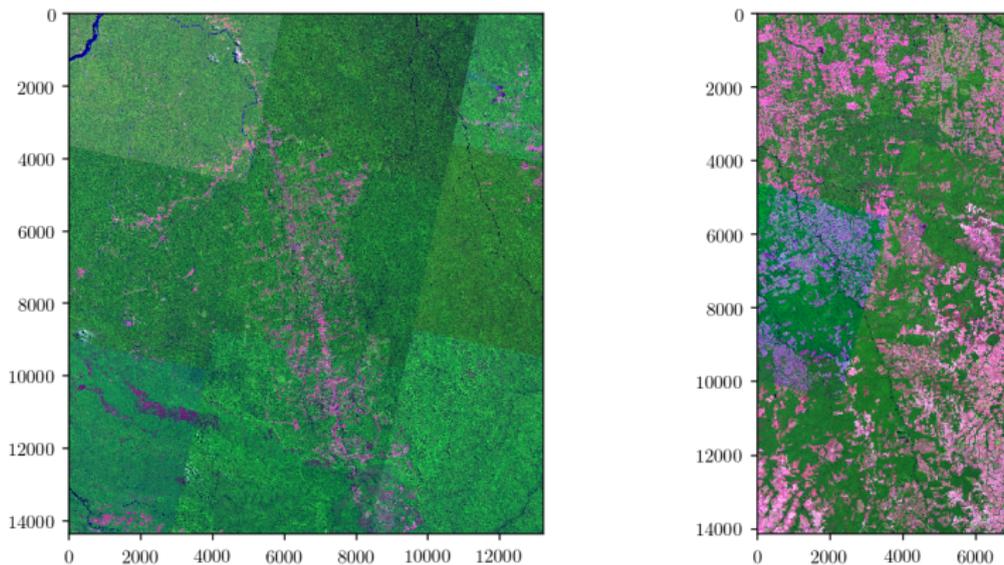


Figura: Mosaicos das cenas LANDSAT

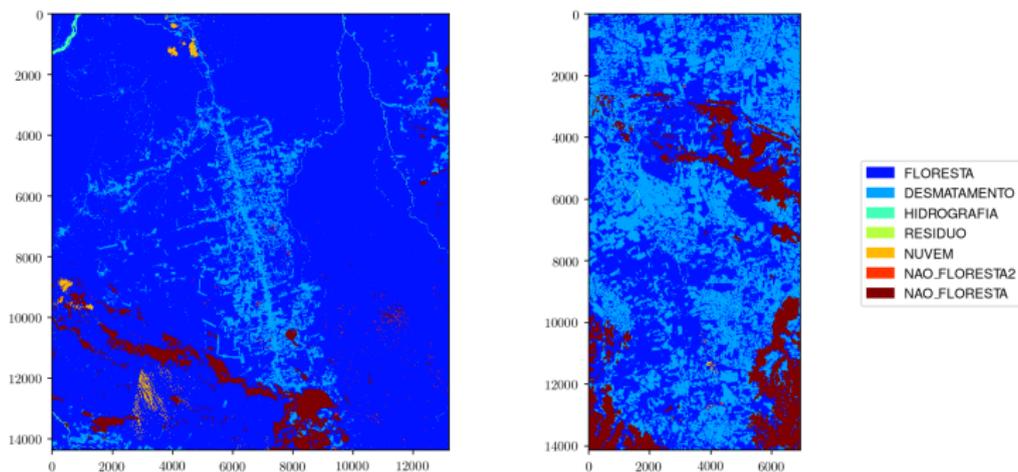


Figura: Mosaicos das máscaras PRODES

Dataset

- Os mosaicos foram recortados em *chips* de 256 pixels de altura e largura.
- Total de 4291 recortes. 3474 utilizados no treinamento, 387 para validação e 430 para teste;

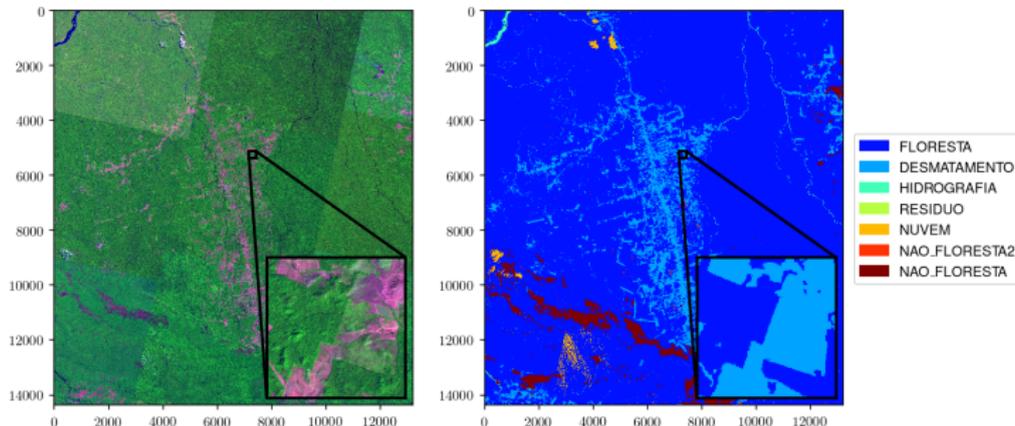


Figura: Recortes 256x256

Modelo U-Net

- Implementado em Python utilizando o pacote Keras com *backend* do Tensorflow com GPU.
- Modelo original utiliza o dobro de características do que o nosso modelo ($32 \rightarrow 64 \rightarrow 128 \rightarrow 256 \rightarrow 512$);
- Utiliza o otimizador estocástico Adam com uma taxa de aprendizado de 0.001 e um *batch* de tamanho 12;
- *Dropout* (Figura 6) com probabilidade de 0.25 na primeira convolução e 0.5 nas demais convoluções para evitar *over-fitting* no treinamento.
- Convoluções com função de ativação ReLU;
- Última convolução utiliza uma função de ativação softmax nos dados de saída.

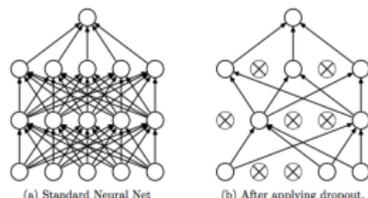
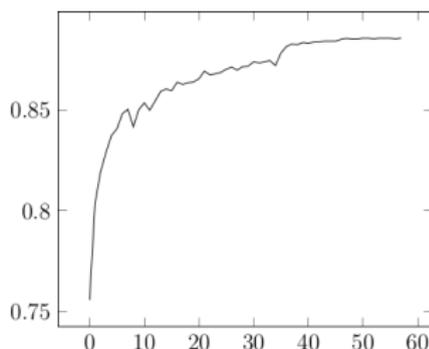
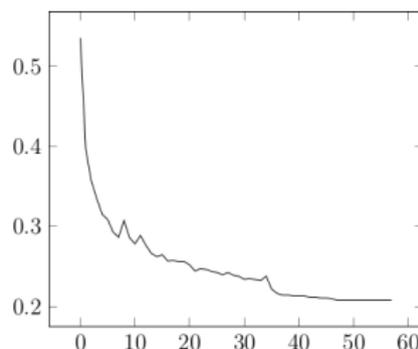


Figura: *Dropout* (Srivastava et al. 2014)

- Treinamento durou 58 *epochs*, após ser interrompido por não melhorar as métricas.



(a) Dice coef



(b) Loss

Figura: Métricas de treinamento

Função	Valor
Coeficiente Dice	0.8855
Função de Custo	0.2081

Tabela: Métricas de treinamento

- Coeficiente de Dice é uma métrica que avalia a similaridade entre dois objetos;
- Seu valor varia entre 0 (nenhuma similaridade) e 1 (similaridade total).

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} = \frac{2TP}{2TP + FP + FN}$$

Resultados

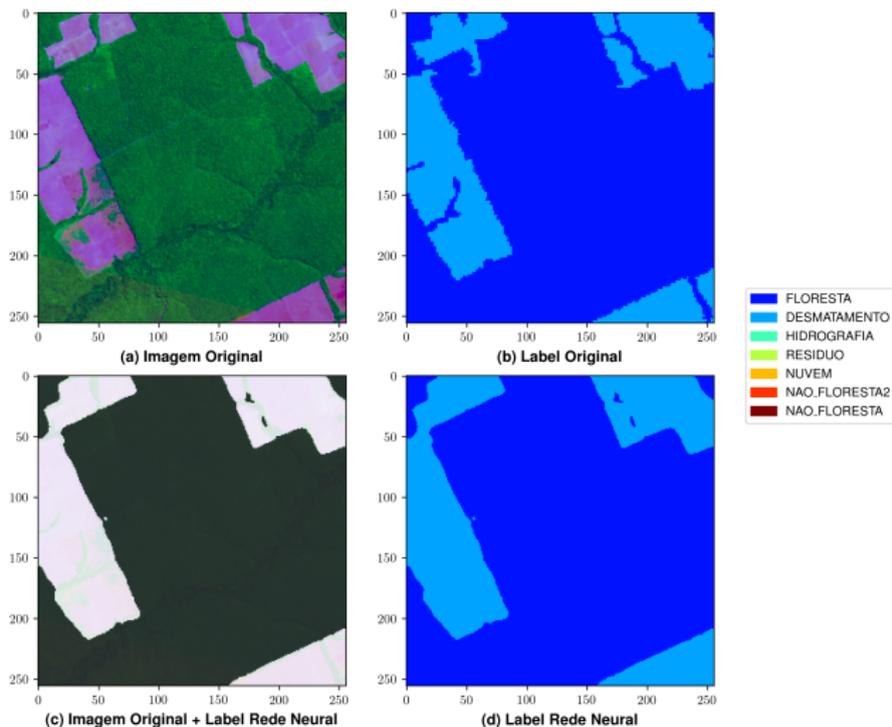


Figura: Segmentação das classes FLORESTA e DESMATAMENTO

Resultados

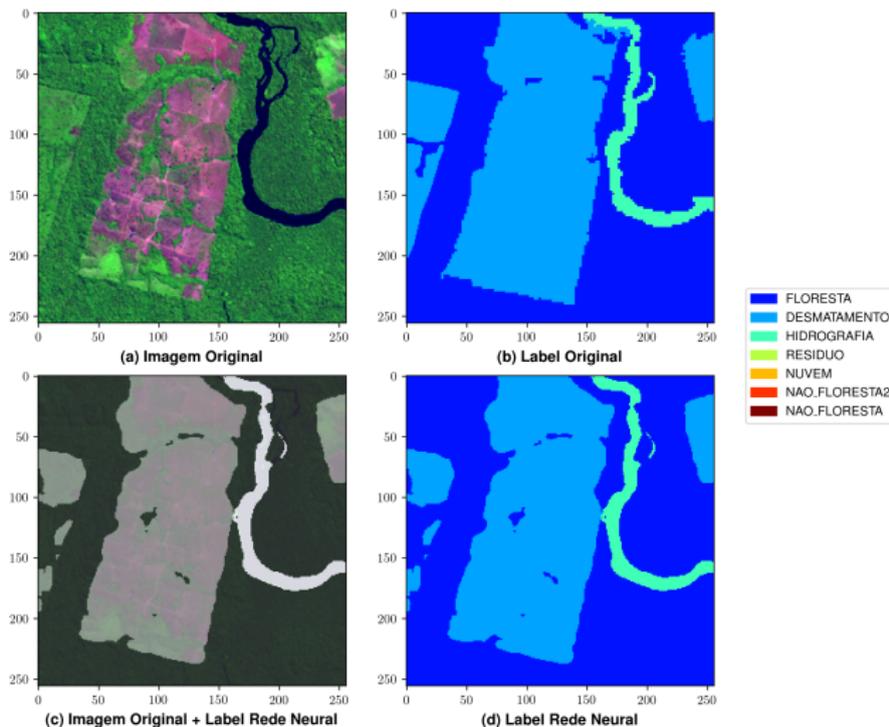


Figura: Segmentação das classes HIDROGRAFIA, FLORESTA e DESMATAMENTO

Resultados

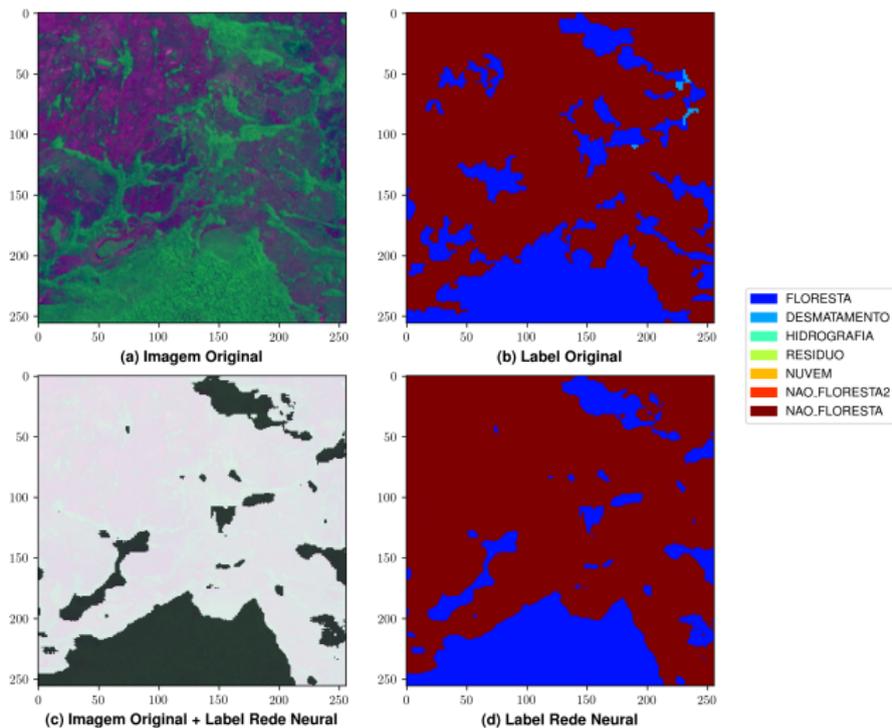


Figura: Segmentação da classe NAO_FLORESTA

Resultados

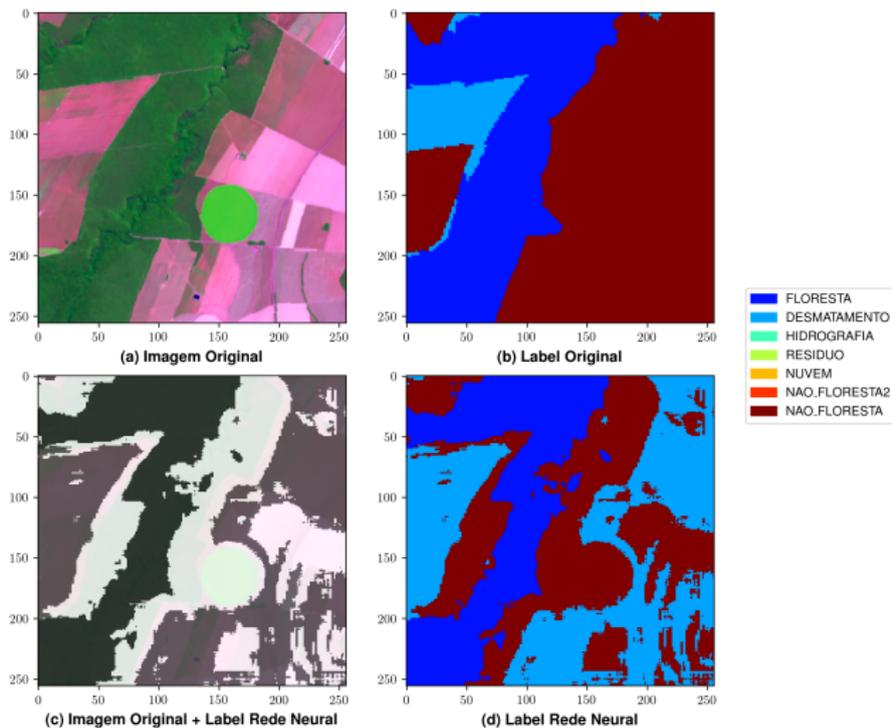


Figura: Confusão entre classes I

Resultados

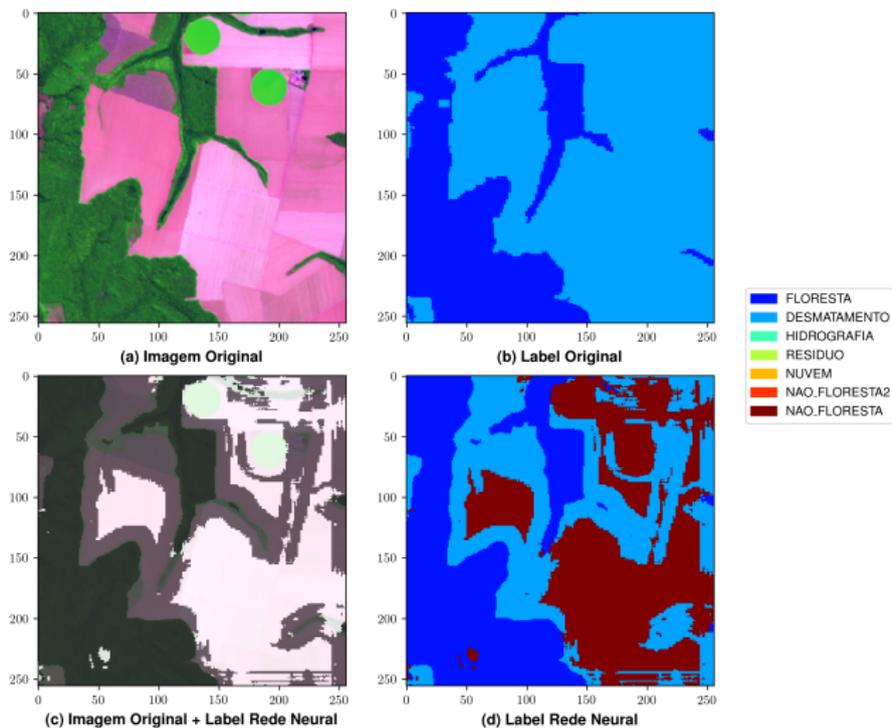


Figura: Confusão entre classes II

Resultados

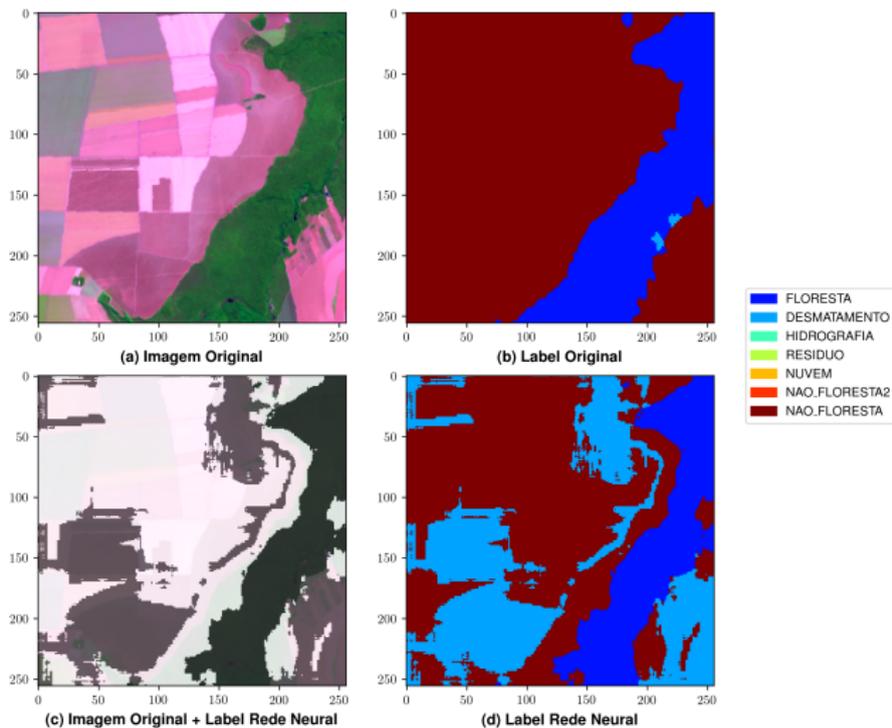


Figura: Confusão entre classes III

- Resultados satisfatórios. Dice Score: 0.88 e Loss 0.20;
- Dataset do PRODES não foi feito ser utilizado para Redes Neurais;
- Um maior pre processamento pode resultar em classificações melhores;

-  M.A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015. URL: <https://books.google.com.br/books?id=STDBswEACAAJ>.
-  Olaf Ronneberger, Philipp Fischer e Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. Em: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, pp. 234–241.
-  Nitish Srivastava et al. “Dropout: a simple way to prevent neural networks from overfitting”. Em: *The Journal of Machine Learning Research* 15.1 (2014), pp. 1929–1958.

Obrigado!