

Deep learning e segmentação semântica de imagens para diagnóstico de níveis de degradação de pastagem ⁽¹⁾

Luan Porto⁽²⁾; Margareth Simões⁽³⁾ ; Rodrigo P D Ferraz⁽⁴⁾

⁽¹⁾ Trabalho executado com recursos do CNPq (Bolsa PIBIC) e do Projeto ReinserTEC/Embrapa Solos. ⁽²⁾ Engenharia de Sistemas e Computação (em curso na UERJ), Estudante de graduação, Bolsista PIBIC da Embrapa Solos - Rio de Janeiro. ⁽³⁾ Engenheira, Doutora, Pesquisadora da Embrapa Solos e Professora do Dep. de Engenharia de Sistemas e Computação (UERJ/DESC)- Rio de Janeiro. ⁽⁴⁾ Engenheiro Agrônomo, Doutor, Pesquisador da Embrapa Solos - Rio de Janeiro.

Resumo – Estima-se que 80% das pastagens cultivadas presentes do Brasil encontram-se em algum nível de degradação. O processo de degradação pode ser identificado quando a produção de forragem diminui, consequentemente, o aumento das áreas descobertas. Há também o surgimento de plantas novas de espécies invasoras, competindo por nutrientes. Este trabalho objetivou avaliar técnicas de segmentação semântica de fotografias para diferenciação de plantas invasoras/daninhas (diversas famílias botânicas), plantas forrageiras (família gramínea), solo exposto e palhada (material seco). Estas técnicas se baseiam em aprendizagem supervisionado por meio do modelo de segmentação semântica DeeplabV3+. Primeiramente com o objetivo de treinar o modelo, foi usado amostras do banco de dados The GrassClover Image Dataset for Semantic and Hierarchical Species Understanding in Agriculture. Para aproximar o modelo para as pastagens brasileiras, um pequeno banco de dados foi construído a partir das fotografias ortogonais, retiradas da plataforma Atlas das Pastagens. Após o processo de treinamento com os dois bancos de imagens, o modelo foi avaliado a partir de diferentes imagens separadas para este propósito. Os resultados demonstraram satisfatórios com valores de MIoU de 87,6% com o primeiro banco de imagens e 75,4% com segundo banco de imagens. Conclui-se que as técnicas de visão computacional, aprendizado de máquina e segmentação semântica de imagens apresentam potencial para subsidiar a classificação da degradação de pastagens.

Palavras-Chave: Visão Computacional, Aprendizado de máquina, Redes neurais convolucionais, Plano ABC.

Introdução

O Brasil é um dos líderes em produção pecuária do mundo, onde é caracterizado, em sua maioria, pela exploração extensiva baseada em sistema de pastejo sob pastagens naturais ou cultivadas com baixo emprego encontrado de capital e tecnologia. Um dos problemas da relativa baixa produtividade do setor é devido ao mal estabelecimento e/ou ao mal manejo das pastagens que condicionam a baixa capacidade de suporte e o desencadeamento de processos de degradação. Estima-se que 80% das pastagens cultivadas presentes do Brasil encontram-se em algum nível de degradação, sendo que essa problemática afeta diretamente a sustentabilidade da pecuária (Paulino et al., 2012).

Degradação das pastagens é definido por Macedo e Zimmer (1993) como um processo evolutivo da perda do vigor, de produtividade, da capacidade de recuperação natural das pastagens para sustentar os níveis de produção e qualidade exigida pelos animais, bem como o de superar os efeitos nocivos de pragas, doenças e invasoras, culminando com a degradação avançada dos recursos naturais em razão de manejos inadequados.

Por sua vez, Dias-Filho (2011) considera que a degradação das pastagens pode ocorrer em função de dois processos distintos, aos quais o autor denominou de: “degradação agrícola” e “degradação biológica”. Na degradação agrícola, a degradação é caracterizada pela mudança na composição botânica da pastagem, isto é, aumento na proporção de plantas daninhas arbóreo-arbustivas e da consequente diminuição na proporção de forragem. Portanto, na degradação agrícola, a produtividade da pastagem, do ponto de vista agrônomo, estaria temporariamente diminuída ou inviabilizada, como consequência da pressão competitiva exercida pelas plantas daninhas sobre as plantas forrageiras, causando, portanto, queda acentuada na capacidade de suporte da pastagem.

Uma pastagem está em processo de degradação quando a produção de forragem diminui com a redução da sua qualidade e quantidade, mesmo nas épocas favoráveis ao seu crescimento. Há diminuição na área coberta do solo pela pastagem e existe pequeno número de plantas novas, provenientes da ressemeadura natural. O aparecimento de espécies invasoras de folha larga, competindo por nutrientes e de processos erosivos pela ação das chuvas, promovendo a colonização da pastagem por gramíneas nativas (Filho et al., 1992).

Para contribuir com a solução deste problema é proposto neste trabalho uma abordagem de segmentação semântica automática entre plantas daninhas, gramíneas, solo exposto e material seco (palhada) a partir de fotografias, através das técnicas de Visão Computacional e Aprendizado de Máquina.

A segmentação é de fundamental importância para a fase de análise do processo de classificação, pois por meio dela torna-se possível extrair informações valiosas a partir das imagens digitais (Saldanha; Freitas, 2009). A tarefa de segmentação semântica é complexa e é uma das poucas áreas de visão computacional em que o desempenho das redes neurais profundas está além da performance humana.

O objetivo desse artigo foi avaliar o desempenho do modelo Deeplabv3+ para segmentação semântica de fotografias ortogonais de pastagens contendo gramíneas, plantas invasoras, solo exposto e palhada, visando subsidiar a classificação da degradação de pastagens.

Material e Métodos

Ferramentas

O modelo de segmentação semântica (Hafiz; Bhat, 2020) DeepLabV3+ foi construído no ano de 2018 com uma arquitetura FCN (Fully Convolutional Network) (Long et al., 2015). Apresentado por Chen et al. (2018) o modelo pode ser acessado gratuitamente no repositório <https://github.com/tensorflow/models/tree/master/research/deeplab>, desenvolvido na linguagem de programação Python junto com a biblioteca Tensorflow. Para este trabalho, foi usado a versão de julho de 2020 do modelo Deeplabv3+, Python na versão 2.7.18 e Tensorflow na versão tensorflow-gpu 1.15. O ambiente escolhido para executar o modelo foi a ferramenta Google Colaboratory, a ferramenta torna possível o desenvolvimento de aplicações de machine learning através do navegador além da disponibilidade de placas gráficas que aceleram o processo de treinamento dos modelos.

Bancos de imagens

A obtenção de imagens relacionadas a pastagens segmentadas é geralmente um grande desafio. A base mais promissora encontrada é denominada como The GrassClover Image Dataset for Semantic and Hierarchical Species Understanding in Agriculture (Skovsen et al., 2019). O conjunto de imagens é composto por 8000 imagens sintéticas de alta resolução com anotações perfeitas, divididas em 15 classes.

O primeiro banco de imagens é constituído por 100 imagens, escolhida de forma aleatória, do Dataset GrassClover. Devido às limitações do ambiente de execução, que não suportava a realização das imagens do tamanho original, cada imagem foi dividida em pequenas imagens de tamanho máximo de 512 x 512 pixels resultando em 17196 imagens. A escolha do recorte das imagens e não do redimensionamento foi optado devido às perdas de informações no processo de diminuição da imagem, sendo assim, afetando a acurácia do modelo de segmentação semântica.

Fora os recortes das imagens, o número de classes foi alterado de 15 para 3 classes. A fim de atender o objeto do trabalho, às classes diferentes de Solo Exposto e Gramíneas foram remarcadas como Plantas Invasoras, com a finalidade de atender a definição de plantas invasoras segundo Voll et al. (2005), como plantas que crescem onde não são desejadas.

Para trazer a solução para as pastagens brasileiras, o segundo banco de imagens foi construído a partir das fotografias ortogonais de pastagens brasileiras, retiradas da plataforma Atlas das

Pastagens (<https://atlasdaspastagens.ufg.br/map>). Cada imagem foi recortada seguindo a mesma regra do primeiro banco de imagens.

A partir das imagens recortadas, cada recorte foi segmentado manualmente com o auxílio do editor de imagens GIMP e de um algoritmo construído em Python pelos autores. Por ser um processo penoso, o segundo banco é constituído por 110 imagens com 4 classes, além das classes Solo Exposto, Gramínea e Planta Invasora; foi adicionada a classe Palhada.

O exemplo das imagens dos bancos de dados, seguido da segmentação, podem ser observados na figura 1.

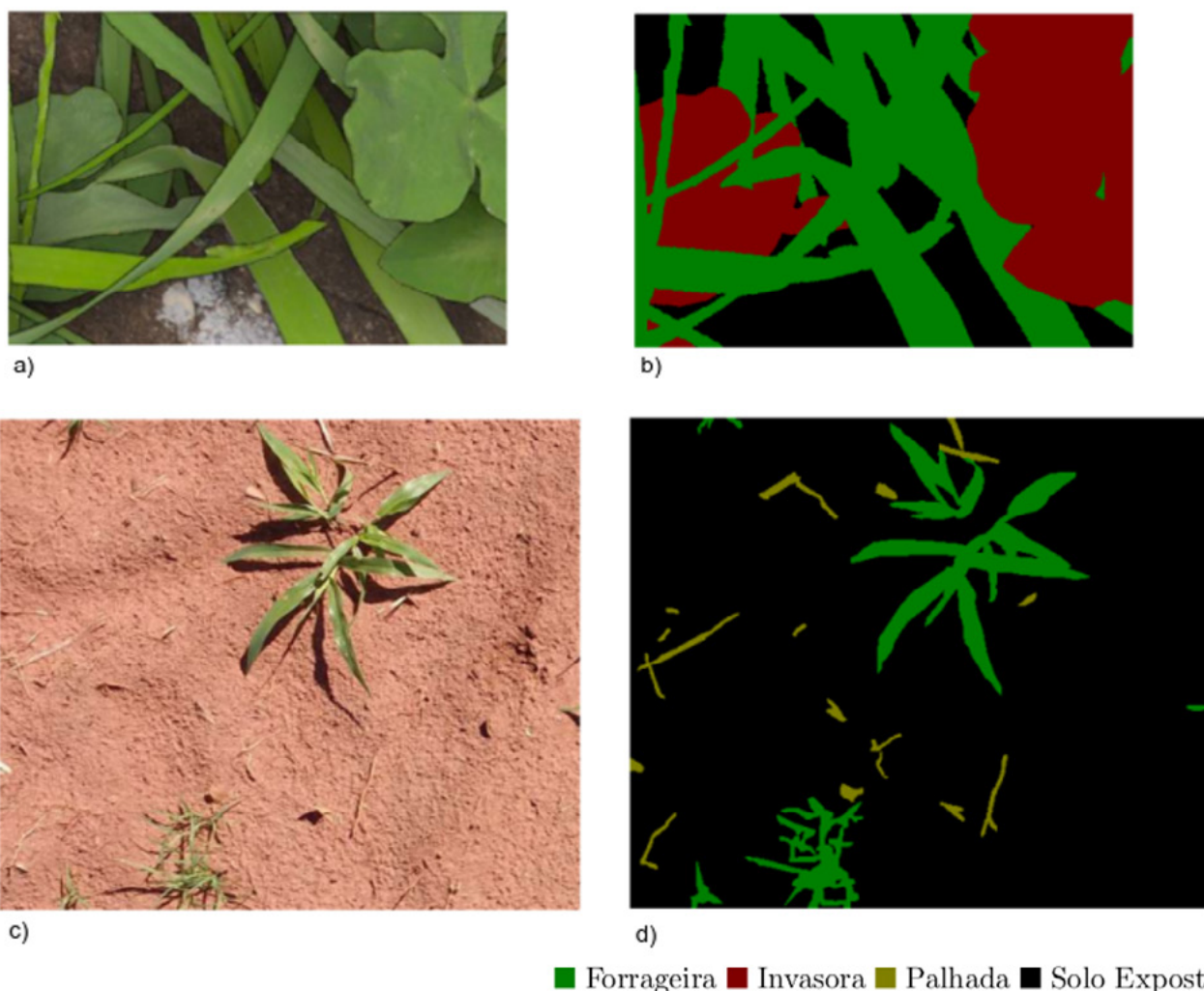


Figura 1. Exemplos das imagens dos bancos de dados; a) e b) pertencem ao primeiro banco de imagens, onde b) é a segmentação semântica de a); c) e d) pertencem ao segundo banco de imagens, onde, d) é a segmentação semântica de c)

Metodologia

O desenvolvimento desta pesquisa foi efetuada através das etapas descritas a seguir:

- Acesso ao data set The GrassClover;
- Seleção das 100 imagens, recorte e remarcação;
- Seleção das fotografias do Atlas das Pastagens, recorte e segmentação;
- Configuração do modelo DeepLabV3+ na plataforma Google Colab;

- Divisão dos Bancos de imagens em $\frac{2}{3}$ para treinamento e $\frac{1}{3}$ para validação
- Treinamento com o primeiro banco;
- Validação com o primeiro banco;
- Treinamento com o segundo banco, usando os pesos adquiridos no primeiro treinamento;
- Validação com o segundo banco.

Os bancos foram treinados e validados utilizando os três extratores de características disponíveis: Xception 65, MobileNet V2 e ResNet 101. As configurações para o processo de treinamento com o primeiro banco foram definidos de acordo com a **tabela 1**.

As configurações com o segundo banco foram as mesmas do primeiro, porém as variáveis training_number_of_steps foi igual a 60000 e tf_initial_checkpoint recebeu o local dos pesos resultantes do primeiro treinamento. Essa técnica é conhecida como Transferência de Aprendizado (Akram et al., 2018) que é utilizada para melhorar o desempenho dos modelos de aprendizado supervisionado quando há poucas amostras.

Para avaliação, o modelo foi configurado de acordo com a **tabela 2**.

Tabela 1. Parâmetros na frase de treinamento.

Parâmetros	RNC		
	Xception 65	MobileNet V2	ResNet 101
training number of steps	50000	50000	50000
atrous rates	6,12,18	6,12,18	6,12,18
output stride	16	16	16
decoder output stride	4	4	4
train crop size	512,512	512,512	512,512
train batch size	1	1	1
optimizer	adam	adam	adam
fine tune batch norm	False	False	False
initialize last layer	False	False	False
last layers contain logits only	True	True	True

Tabela 2. Parâmetros na frase de validação.

Parâmetros	RNC		
	Xception 65	MobileNet V2	ResNet 101
atrous rates	6,12,18	6,12,18	6,12,18
output stride	16	16	16
decoder output stride	4	4	4
eval crop size	512,512	512,512	512,512

Análise estatística

Para avaliação dos modelos de segmentação semântica, as seguintes métricas podem ser utilizadas: Precision, Pixel Accuracy, Intersection Over Union (IoU), Mean Intersection Over Union (MIoU) e F1-Score. Todas as métricas citadas, podem ser encontradas a partir da matriz de confusão seguindo as equações a seguir:

$$PixelAccuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precisão = \frac{TP}{TP + FP}$$

$$S = \frac{TP}{TP + FN}$$

$$F1-Score = \frac{Precisão \times S}{Precisão + S}$$

$$IoU = \frac{TP}{TP + FN + FP}$$

$$MIoU = \sum_{i=1}^n IoU / n$$

Resultados e Discussão

Após os treinamentos, as matrizes de confusão dos resultados com o primeiro banco de imagens podem ser observadas na **figura 2**.

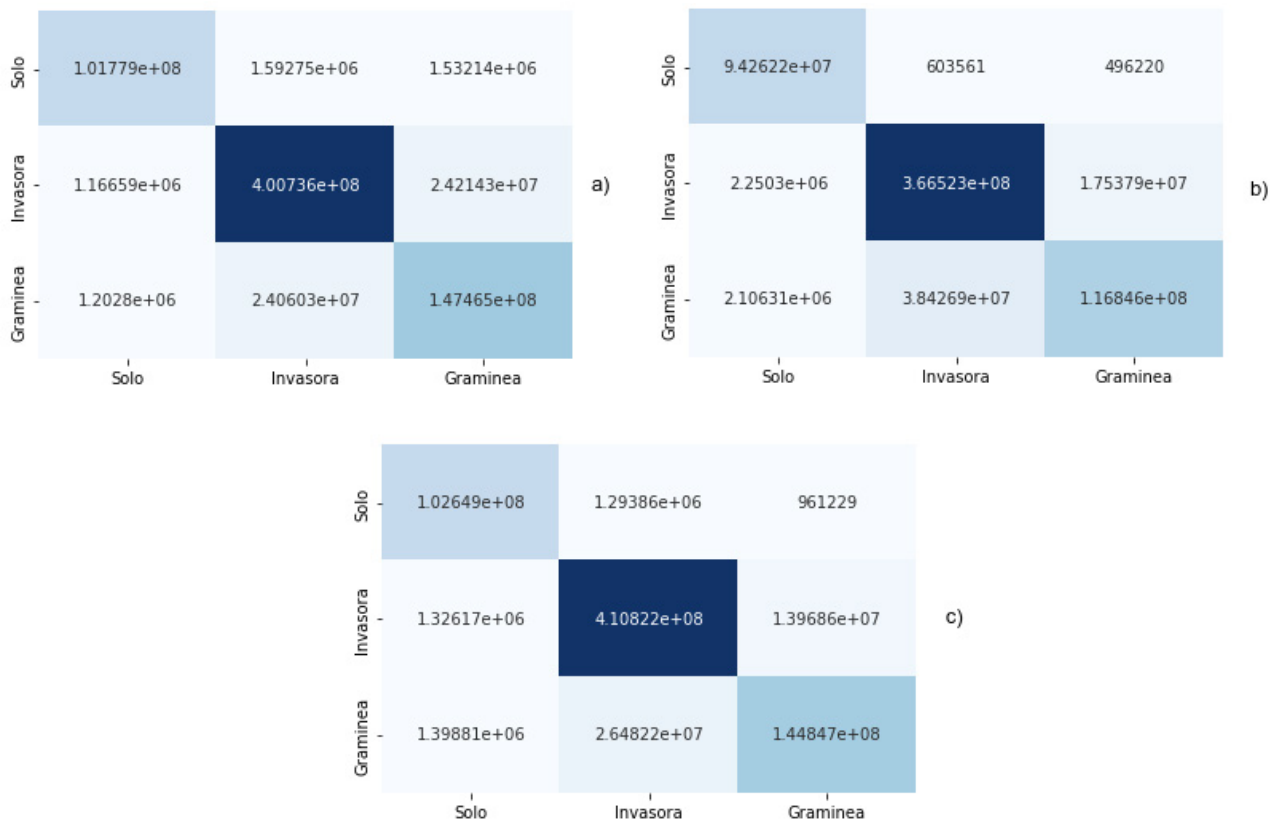


Figura 2. Matrizes de confusão utilizando o primeiro banco de imagens; a) Matriz de confusão Xception 65; b) Matriz de confusão com MobileNet V2; c) Matriz confusão com ResNet 101

As matrizes de confusão dos resultados com o segundo banco de imagens podem ser observadas na **figura 3**.



Figura 3. Matrizes de confusão utilizando o segundo banco de imagens (Imagens do Brasil) ; a) Matriz de confusão Xception 65; b) Matriz de confusão com MobileNet V2; c) Matriz confusão com ResNet 101

A partir das matrizes de confusão dos dois bancos de imagens, podemos calcular as métricas de avaliação citadas anteriormente. As **tabelas 3 e 4** apresentam os resultados do primeiro e segundo banco de imagens, respectivamente.

Tabela 3. Resultados utilizando o primeiro banco de imagens (The GrassClover Image Dataset for Semantic and Hierarchical Species Understanding in Agriculture)

	Solo	Invasora	Gramínea	Total
Xception 65				
Precisão	97,7%	93,8%	85,1%	-
Sensibilidade	97%	94%	85,3%	-
F1-Score	97,3%	94%	85%	-
IoU	94,8%	88,7%	74,2%	-
MIoU	-	-	-	85,9%
Pixel Accuracy	-	-	-	92,3%
MobileNet V2				
Precisão	95,5%	90,3%	86,6%	-
Sensibilidade	98,8%	94,8%	74,3%	-
F1-Score	97,1%	92,5%	79,9%	-

Continua...

Tabela 3. Continuação.

	Solo	Invasora	Gramínea	Total
IoU	94,5%	86,1%	66,6%	-
MIoU	-	-	-	82,4%
Pixel Accuracy	-	-	-	90,3%
Resnet 101				
Precisão	97,4%	93,6%	90,6%	-
Sensibilidade	97,8%	96,4%	83,8%	-
F1-Score	97,6%	95%	87,1%	-
IoU	95,3%	90,5%	77,1%	-
MIoU	-	-	-	87,6%
Pixel Accuracy	-	-	-	93,5%

Tabela 4. Resultados utilizando o segundo banco de imagens (Imagens do Brasil).

	Solo	Invasora	Gramínea	Palhada	Total
Xception 65					
Precisão	85,5%	78,1%	77,6%	65,5%	-
Sensibilidade	87,2%	70,5%	82,9%	62%	-
F1-Score	86,4%	74,1%	80,1%	63,7%	-
IoU	76%	58,9%	66,9%	46,7%	-
MIoU	-	-	-	-	62,1%
Pixel Accuracy	-	-	-	-	79,3%
MobileNet V2					
Precisão	83,3%	75%	67,1%	68,1%	-
Sensibilidade	91,1%	45,1%	82,6%	55,6%	-
F1-Score	87%	56,3%	74,1%	61,2%	-
IoU	77%	39,2%	58,8%	44,1%	-
MIoU	-	-	-	-	54,8%
Pixel Accuracy	-	-	-	-	76,3%
Resnet 101					
Precisão	88,7%	95%	86,1%	77,1%	-
Sensibilidade	91,3%	83%	91,6%	73,3%	-
F1-Score	89,9%	88,6%	88,8%	75,2	-
IoU	81,8%	79,6%	79,8%	60,2	-
MIoU	-	-	-	-	75,4%
Pixel Accuracy	-	-	-	-	86,8%

A **figura 4** mostra a comparação das imagens obtidas pelo modelo no primeiro banco de imagens. Já os exemplos do segundo banco, podem ser observados na **figura 5**.

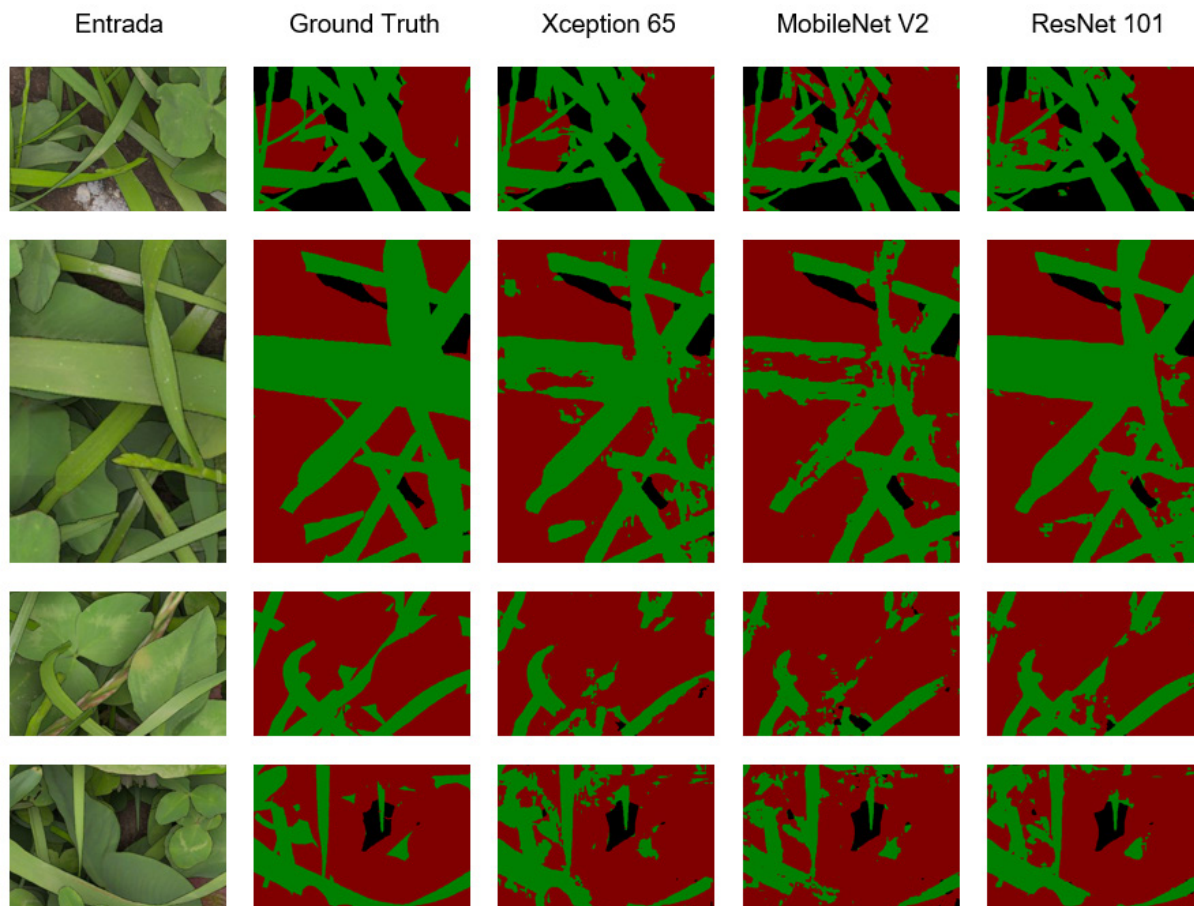


Figura 4. Predições do modelo utilizando o primeiro banco de imagens (The GrassClover Image Dataset for Semantic and Hierarchical Species Understanding in Agriculture)

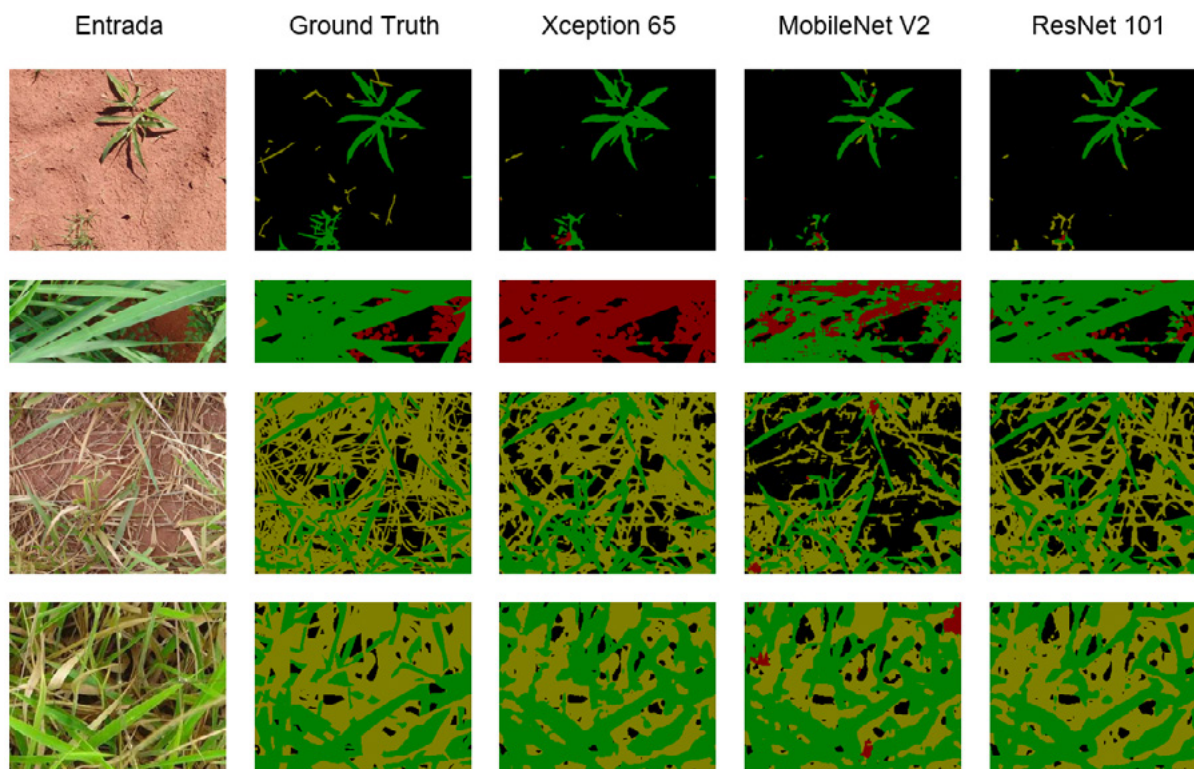


Figura 5. Predições do modelo utilizando o segundo banco de imagens (Imagens do Brasil)

Conclusões

Inicialmente foi necessário o uso do conjunto de dados The GrassClover Image Dataset for Semantic and Hierarchical Species Understanding in Agriculture, banco de imagens mais promissor encontrado que atendesse o objetivo deste trabalho. Depois da rotulação das classes para adequar o banco de imagens e recortes devido às limitações de hardware, o modelo DeepLabV3+ mostrou-se ser uma arquitetura confiável com resultados de MIoU até 87,6% utilizando o extrator de características ResNet 101.

Para tornar o modelo funcional com pastagens nacionais, um pequeno banco de imagens foi construído com fotografias do Brasil. Utilizou-se a técnica de transferência de aprendizado com os pesos obtidos no treinamento anterior, com objetivo de aumentar a eficiência do modelo com um pequeno montante de imagens. Diferente dos primeiros resultados, nesta etapa pode-se observar um destaque utilizando a rede convolucional ResNet 101 em comparação às demais redes, tendo os melhores resultados com as métricas apresentadas neste trabalho: precisão, sensibilidade F1-score, IoU, MIoU e Pixel Accuracy; os valores de cada métrica pode ser observado na **tabela 4**, destacados em negrito.

Diante dos resultados apresentados, pode-se concluir que o modelo DeepLabV3+ é apto em realizar a segmentação semântica nos dois bancos apresentados neste trabalho, sendo assim, o modelo pode subsidiar a classificação da degradação de pastagens.

Agradecimentos

Os autores agradecem a Embrapa (Projeto REINSERTEC) e a UERJ que deram o suporte técnico-científico e acadêmico necessário para a realização do presente trabalho. Agradecem ainda ao CNPq pelo financiamento na forma de bolsa de iniciação científica (PIBIC)

Referências

- CHEN, L.-C. et al. **Encoder-decoder with atrous separable convolution for semantic image segmentation**. In: Proceedings of the European conference on computer vision (ECCV). [S.l.: s.n.], 2018. p. 801–818.
- DIAS-FILHO, M. B. **Degradação de pastagens: processos, causas e estratégias de recuperação**. rev. atual. e ampl. Belém, PA, 2011.
- FILHO, C. S.; MONTEIRO, F. A.; CORSI, M. **Recuperação de pastagens degradadas de brachiaria decumbens**. 1. efeito de diferentes tratamentos de fertilização e manejo. Pasturas Tropicales, v. 14, n. 2, p. 7–13, 1992.
- LONG, J.; SHELHAMER, E.; DARRELL, T. **Fully convolutional networks for semantic segmentation**. In: Proceedings of the IEEE conference on computer vision and pattern recognition. [S.l.: s.n.], 2015. p. 3431–3440.
- MACEDO M. C. M.; ZIMMER, A. H. **Sistema pasto-lavoura e seus efeitos na produtividade agropecuária**. In: SIMPÓSIO SOBRE ECOSISTEMA DE PASTAGENS, 2., Jaboticabal. Anais., Fundação de Estudos e Pesquisas em Agronomia, Medicina Veterinária e Zootecnia, 1993., p. 216–245, 1993.
- PAULINO, V.T.; SCHUMANN, A.M.; SILVA, S.C.; RASQUINHO, N.M.; SANTOS, K.M. **Impactos ambientais da exploração pecuária em sistemas intensivos de pastagens**. Informe Agropecuário, Belo Horizonte, v.33, n.266, p.17-24, 2012.
- SALDANHA, Marcus FS; FREITAS, C. **Segmentação de imagens digitais: Uma revisão**. Divisão de Processamento de Imagens-Instituto Nacional de Pesquisas Espaciais (INPE), São Paulo, 2009.
- SKOVSEN, S. et al. **The grassclover image dataset for semantic and hierarchical species understanding in agriculture**. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. [S.l.: s.n.], 2019.
- VOLL, E. et al. **A dinâmica das plantas daninhas e práticas de manejo**. [S.l.]: Embrapa Soja Londrina, 2005. v. 260.