

ALGORITMO DE *DEEP LEARNING* PARA CLASSIFICAÇÃO DE ÁREAS DE LAVOURA COM VANTS

Anthony Vanzan¹, Gabriel Rustick Fim¹, Greice Aline Welter¹, Matheus César Sausen¹, Dalvan Griebler².

1 Alunos, Laboratório de Pesquisas Avançadas para Computação em Nuvem (LARCC), Faculdade Três de Maio (SETREM), Três de Maio/RS. anthonyvanzan0202@gmail.com, gabifimtm@gmail.com, greicewelter8@gmail.com, matheus.sausen00@gmail.com.

2 Professor, Laboratório de Pesquisas Avançadas para Computação em Nuvem (LARCC), Faculdade Três de Maio (SETREM), Três de Maio, RS. dalvangriebler@setrem.com.br

RESUMO

O Brasil é um dos maiores produtores e exportadores de milho do globo. A criação e implementação de novas tecnologias partindo da inteligência artificial podem proporcionar melhorias na produção do grão e, conseqüentemente, uma melhoria econômica no país. Nota-se também que as tecnologias de inteligência artificial estão conquistando espaço no mercado e auxiliando diversas áreas, tendo um avanço considerável de desempenho e produtividade. Nesse sentido, o presente trabalho visa apresentar a implementação e resultados de um modelo de rede neural utilizando a arquitetura LeNet5 para realizar a classificação de imagens, de cultivo do milho. Estas áreas classificadas servirão futuramente para o cálculo de estimativa de produção.

Palavras-chave: *Deep Learning*. Inteligência Artificial. Redes Neurais. Milho.

1 INTRODUÇÃO

Para (ABDI et al. 1999), redes neurais são modelos estatísticos adaptativos baseados em uma analogia com a estrutura do cérebro. Contudo, elas são adaptáveis, pois podem aprender a estimar os parâmetros de uma população usando um pequeno número de exemplos de cada vez. Redes neurais podem ser usadas em aplicativos de operações financeiras, planejamento empresarial, comércio, análise de negócios e manutenção de produtos. Sendo assim, elas também obtiveram ampla adoção em aplicativos de negócios, como soluções de previsão e pesquisa de marketing, detecção de fraudes e avaliação dos riscos (HAYKIN. 2007).

Dentre os trabalhos relacionados a este, (TRUJILLANO et al. 2018) propuseram o uso de técnicas de *deep learning* para a classificação de imagens NIR (*near infra-red*) para estimar áreas de milho, utilizando imagens em alturas de 100 e 180 metros em relação ao solo. Usando um modelo de rede neural convolucional, obtiveram acurácias entre 87,2% a 92,2%. No trabalho de (LIU et al. 2020) foi realizado o estudo em duas áreas na China, utilizando um UAV DJI Inspires 2 com câmera Zenmuse X5S de resolução 5280 X 2970 *pixels* na altura de 15 metros em relação ao solo. As imagens foram cortadas e filtradas para terem resolução de 600 X 600 *pixels*, com o intuito de facilitar o processamento das imagens. Os resultados foram obtidos utilizando a VGGNet e Resnet, apresentado respectivamente para cada uma delas entre 87,94% e 91,51% e, 91,99% e 94,99% de precisão. Diferentemente, nosso trabalho visa utilizar apenas imagens RGB que possuem um custo menor que imagens NIR, uma arquitetura de rede neural convolucional, em

alturas de 15, 60 e 100 metros em relação ao solo para a classificação das imagens em milho e não milho.

O Brasil é um dos maiores produtores e exportadores de milho do mundo, sendo responsável pela produção de 25.91% do milho consumido mundialmente (Embrapa Milho e Sorgo, 2014), o que torna o agronegócio um de seus principais influenciadores no PIB nacional. A implementação de novas tecnologias nesse meio, que aprimorem as técnicas previamente desenvolvidas ou até mesmo criem novas formas de realizar tarefas, otimizando o tempo e possibilitando aferir os dados com maior precisão, são de suma importância para o desenvolvimento da área.

O objetivo deste trabalho é desenvolver e aplicar um algoritmo de *deep learning* capaz de classificar imagens de lavouras e aferir com precisão o que é milho e o que não é milho. Podendo assim, demarcar as áreas de plantação e saber com precisão o seu tamanho. Na seção 2 é apresentada a fundamentação teórica pertinente à revisão de artigos e livros que foram utilizados para a redação do trabalho. Na seção 3 é descrita a metodologia aplicada no campo de estudo. Na seção 4 é apresentado os resultados obtido ao decorrer da realização do trabalho. Por fim na seção 5 conclui-se este trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

Redes neurais artificiais, usualmente denominadas “redes neurais”, têm sido motivadas desde o reconhecimento de que o cérebro humano processa informações de uma forma diferente do computador convencional. O cérebro é como se fosse um computador (sistema de processamento de informação) altamente complexo, não-linear e paralelo. Ele tem a capacidade de organizar seus constituintes estruturais, conhecidos por neurônios, de forma a realizar certos processamentos (p.ex., reconhecimento de padrões, percepção e controle motor) muito mais rapidamente que o mais rápido computador digital existente hoje (MARR, 1982; LEVINE, 1985; CHURCHLAND e SEJNOWSKI, 1992).

Ao nascer, um cérebro tem uma grande estrutura e habilidade de desenvolver suas próprias regras através do que usualmente denominamos “experiência”. Na verdade, a experiência vai sendo acumulada com o tempo, sendo que o mais dramático desenvolvimento do cérebro humano acontece até os dois anos de idade; mas o desenvolvimento continua para muito além desse estágio (HAYKIN, S. 2007).

Na sua forma geral, uma rede neural é um programa de computador projetado para moldar a maneira que o cérebro realiza uma tarefa em particular ou função de interesse. Nessa perspectiva, uma rede neural é um processador maciçamente paralelamente distribuído, constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso (HAYKIN, S. 2007).

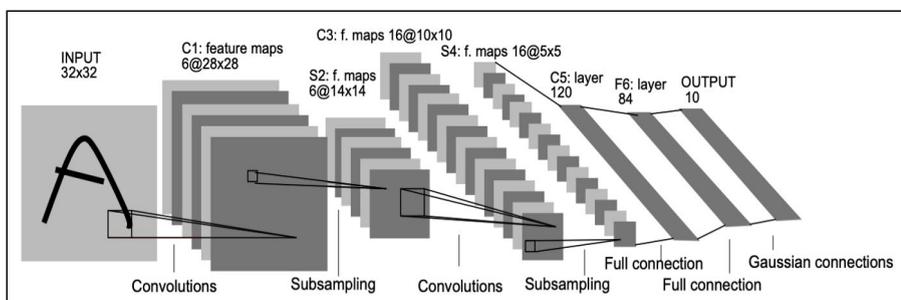
3 METODOLOGIA

O presente estudo vem sendo realizado no escopo do projeto AGROCOMPUTAÇÃO (uma parceria entre SETREM e TECNICON Sistemas). Neste contexto, esse trabalho tem como objetivo criar um algoritmo de *deep learning* utilizando a arquitetura de rede neural convolucional LeNet5 para realizar a

classificação de milho em fotos tiradas por veículos aéreos não tripulados (VANTS). Para isso, foi usado o *framework* Pytorch.

A arquitetura LeNet5 usada foi desenvolvida pelo cientista francês Yan Lecun (LECUN et al. 1998). Sua arquitetura consiste em cinco camadas de processamento convolucional e *pooling*. Como mostrado na Figura 1. O algoritmo recebe imagens previamente tratadas em tamanho padrão de 32x32 *pixels* em escala cinza, que percorrem o comprimento da rede extraíndo suas características e padrões.

Figura 1 - Estrutura da arquitetura de rede LeNet5.



Fonte: (LECUN et al. 1998).

Os imagens utilizadas consistem em três padrões diferenciados. O primeiro, em 249 imagens retiradas a 15 metros a partir do solo quando o milho estava em um estágio fenológico reprodutivo. O segundo consistia em 64 imagens retiradas a 60 metros de altura em relação ao solo quando o milho estava em um estágio fenológico vegetativo. O terceiro, em 34 imagens retiradas a 100 metros de altitude quando o milho estava em um estágio fenológico vegetativo.

Destas imagens foram destacadas e recortadas as áreas de milho e não milho, correspondentes aos dois *labes* utilizados no algoritmo. As imagens recortadas passaram pelo algoritmo de pré-processamento que as recortou em imagens de 100x100 *pixels*, as quais receberam um redimensionamento para 32x32 *pixels*. Depois, elas foram armazenadas em um vetor que continha todas as imagens do *dataset* com seus respectivos *labes*.

O *learning rate* da rede neural foi definido em 0.0001 e o número de épocas de treinamento foi padronizado em 50, valores padrões de treinamento da LeNet5. Foram realizados dois testes com cada *dataset*, sendo um com o *batch size* de 16 e o outro com o *batch size* de 32. O *batch size* é o valor que determina o número de imagens que é passado para a rede a cada iteração até findar o tamanho total do *dataset* utilizado.

O valor adotado para a realização dos testes foi de 0,3, ou seja, 30% do valor total das imagens contidas no *dataset*. Os outros 70% foram utilizados no treinamento do modelo. Na Figura 2 é representada a ortofoto referente ao voo de 100 metros, a imagem foi gerada utilizando o software Agisoft Metashape através das imagens retiradas com VANT em campo.

Figura 2 - Ortofoto gerada a partir de imagens retiradas com VANT.



4 APRESENTAÇÃO, ANÁLISE E DISCUSSÃO DOS RESULTADOS

A Tabela 1 detalha os resultados obtidos com os treinamentos e testes da rede. Os resultados são apresentados através da acurácia, perda e erro médio quadrático (RMSE). O teste que apresentou melhor resultado e consequentemente menor perda e RMSE foi com o *dataset* de 15m utilizando um tamanho de batch de 16, tendo uma acurácia de 0.96, ou seja, uma taxa de acerto de 96%. A perda cegou a 0.031 e o RMSE teve um valor de 0.179, seguido pelo *dataset* de 15 metros com *batch size* de 32 que obteve uma acurácia de 95%, perda de 0.032 e RMSE de 0.179.

Tabela 1 - Resultados da rede neural.

<i>Dataset</i>	<i>Batch Size</i>	Nº de épocas	<i>Learning Rate</i>	Acurácia	Perda	RMSE
60m	16	50	0.0001	0.906	0.078	0.280
60m	32	50	0.0001	0.884	0.079	0.282
100m	16	50	0.0001	0.95	0.039	0.198
100m	32	50	0.0001	0.908	0.065	0.256
15m	16	50	0.0001	0.96	0.031	0.177
15m	32	50	0.0001	0.959	0.032	0.179

O *dataset* de 60 metros obteve o resultado mais baixo dentre os treinamentos e testes, tendo a acurácia de 88%, perda de 0.079 e RMSE de 0.282, com o tamanho de batch em 32, seguido pelo *dataset* de 60 metros com o tamanho de batch em 16 que obteve a acurácia de 90%, perda de 0.078 e RMSE de 0.280. O *dataset* de 100 metros com *batch size* de 16 obteve a acurácia de 95%, perda de 0.039 e RMSE de 0.198 e o de 100 metros com *batch size* de 32 obteve a acurácia de 90%, perda de 0.065 e RMSE de 0.256.

Em termos de desempenho é o modelo de 15m com *batch size* de 16 que obteve o melhor resultado, visto que a quantidade de dados que foi utilizada durante o treinamento foi maior que os demais modelos, porém como a altura de retirada é muito baixa, coloca em risco a integridade do VANT ao voar muito próximo de árvores e demais obstáculos, o que leva à consideração das demais alturas como

viáveis e preferíveis. É importante ressaltar que o estágio fenológico de cada dataset também influencia no resultado do treinamento do modelo já que no dataset de 15 metros o milho já se encontrava em estágio reprodutivo e nos demais estava no estágio vegetativo.

Em comparação ao estado da arte, este trabalho superou os resultados de ambos (TRUJILLANO *et al.* 2018) e (LIU *et al.* 2020), visto que a melhor acurácia destes estudos relacionados foram respectivamente de 92,2% e 94,99%, enquanto o nosso trabalho obteve 96% de acurácia no modelo com altura de 15 metros. O uso de um dataset mais robusto juntamente com uma arquitetura de rede neural convolucional com parâmetros de treinamento e teste previamente definidos, pode ter contribuído nos resultados superiores.

5 CONCLUSÃO

Neste trabalho foi realizada a implementação de uma rede neural e a obtenção de métricas para sua avaliação, buscando obter uma classificação das fotos de milho capturadas por VANTS. Obtendo uma acurácia alta, é possível a classificação de áreas de lavoura em ortofotos. O objetivo de obter uma acurácia capaz de classificar de forma correta o que é milho foi alcançado, como apresentado na Seção 4, em sua maioria, a taxa de acertos que o modelo apresentou foi acima dos 90%. Como trabalho futuro, a contagem e estimativa de produção do milho poderá ser feita estendendo este modelo de deep learning.

6 REFERÊNCIAS

- Haykin, S. **Redes Neurais: Princípios e Prática**. Bookman Editora, 2007.
- Levine, M. 1985. **Man and Machine Vision**. New York: McGraw-Hill.
- Marr, D. 1982. **Vision**. New York: W.H. Freeman and Company.
- Churchland, P.S. and T.J. Sejnowski, 1992. **The Computational Brain**, Cambridge, MA: MIT Press.
- Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "**Gradient-based learning applied to document recognition**," in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.
- Embrapa Milho e Sorgo. **Cultivo do milho**. Disponível em: "<https://bit.ly/35Xo7FE>", Acessado em: Ago. 2020.
- Abdi, H., Edelman, H., Valentin, D., Edelman, B., SAGE, i. **Neural Networks**. SAGE Publications, 1999.
- Trujillano, Fedra & Flores, Andres & Saito, Carlos & Balcazar, Mario & Racoceanu, Daniel. (2018). **Corn classification using Deep Learning with UAV imagery. An operational proof of concept**. 1-4. 10.1109/ColCACI.2018.8484845.
- Liu, Y.; Cen, C.; Che, Y.; Ke, R.; Ma, Y.; Ma, Y. **Detection of Maize Tassels from UAV RGB Imagery with Faster R-CNN**. *Remote Sens.* **2020**, *12*, 338.