

# Classifying the Macronutrient Deficiency in Soybean Leaf with Deep Learning

Maicon A. Sartin<sup>1</sup>, Alexandre C. R. da Silva<sup>2</sup>, Claudinei Kappes<sup>3</sup>, Tercio A. S. Filho<sup>4</sup>

<sup>1</sup>Universidade do Estado de Mato Grosso (UNEMAT) – Sinop – MT -- Brasil

<sup>2</sup>Universidade Estadual Paulista (Unesp) – Ilha Solteira – SP — Brasil

<sup>3</sup>Fundação Chapadão – Chapadão do Sul – MS — Brasil

<sup>4</sup>Universidade Federal de Goiás (UFG) – Catalão – GO -- Brazil

mapsartin@unemat.br, acrsilva@feis.unesp.br

{kappes.agro,tercioas}@gmail.com

**Abstract.** *Deep Learning consists of modern techniques that approach one or more methods of artificial intelligence. An approach is to use Convolutional Neural Networks together with traditional Neural Networks for digital image processing. In this work, a research is conducted to evaluate a deep learning technique in the classification of potassium (k) macronutrient deficiency by the soybean leaf. This research presents an own dataset with distinct treatments of the potassium macronutrient. Several deep learning scenarios are evaluated with different metrics. The results are compared with the literature and show a great potential of Convolutional Neural Networks, with precision above 99% in this type of classification.*

**Resumo.** *Deep Learning consiste em técnicas modernas que abordam um ou mais métodos de inteligência artificial. Uma abordagem está no uso de redes neurais convolucionais em conjunto com redes neurais tradicionais para processamento de imagens digitais. Neste trabalho, é realizada uma pesquisa para avaliar uma técnica de aprendizado profundo na classificação da deficiência de macronutrientes de potássio (K) pela folha de soja. Esta pesquisa apresenta um conjunto de dados próprio com tratamentos distintos do macronutriente de potássio. Vários cenários de aprendizado profundo são avaliados com diferentes métricas. Os resultados são comparados com a literatura e mostram um grande potencial de redes neurais convolucionais, com precisão acima de 99% nesse tipo de classificação.*

## 1. Introdução

A agricultura no Brasil é um eixo econômico nacional com grande representação no PIB (Produto Interno Bruto) nacional, de forma direta e indireta. O setor de produção de grãos tem alta relevância na agricultura do Brasil. No último ano, o setor agropecuário cresceu acima do PIB nacional [Lucchi et al. 2020]. Nas últimas 2 décadas, a agricultura de precisão tem se tornado alvo de discussão em diversas áreas de pesquisa, em especial, na computação. A inovação presente no maquinário agrícola, junto ao alto desempenho dos equipamentos eletrônicos, tem-se provado um grande suporte tecnológico as novas necessidades de competitividade na agricultura.

Um sistema de monitoramento adequado do plantio é necessário para efetuar uma tomada de decisão correta e determinar quantidades específicas de insumos agrícolas. A detecção de macronutrientes viabiliza o reconhecimento de um padrão, que permite verificar o estado atual da planta, independente da cultura. As diferenças visuais nas foliares da planta mostram as suas necessidades fisiológicas atuais com base nos macronutrientes. Dependendo do macronutriente, o seu excesso ou a sua falta pode ser prejudicial ao homem e a planta, com impacto direto no seu rendimento [Sfredo 2008].

As tecnologias da informação e comunicação emergentes trazem ao agricultor, de pequena e larga escala, a facilidade e a precisão de observar o gerenciamento de suas tarefas e custos de produção na plantação. Dentre as tecnologias empregadas destaca-se a internet das coisas (IoT), o sensoriamento remoto, a computação em nuvem, a ciência de dados, o big data, o aprendizado de máquina, o aprendizado profundo, etc. Atualmente, o agricultor tem disponível diversos recursos tecnológicos para melhorar ou automatizar partes do processo de produção agrícola. Atualmente, o termo de fazendas inteligentes é elucidado pelo grande número de recursos tecnológicos nas plantações.

As fazendas inteligentes trazem uma perspectiva para o aumento de produtividade, redução de impacto ambiental e sustentabilidade. Com o aumento da população mundial, torna-se inevitável a necessidade cada vez maior da produção de alimentos. Um dos grandes desafios impostos atualmente, é a proteção dos recursos naturais de forma sustentável. As fazendas inteligentes podem contribuir na melhor compreensão da complexidade e imprevisibilidade do ecossistema na agricultura [Kamilaris and Prenafeta-Boldú 2018].

Na busca incansável da sustentabilidade em fazendas inteligentes, os sistemas de monitoramento por imagens são ferramentas atrativas pelo vasto número de aplicações agregadas. A relevância do aprendizado profundo em aplicações de processamento de imagens na agricultura é comumente abordado na literatura [Xie et al. 2019, Abdullahi et al. 2017, Boulent et al. 2019]. A Rede Neural Convolucional (RNC) é popular na perspectiva de aprendizado profundo para o processamento de imagens em plantações. A grande variedade de pesquisa no monitoramento ambiental por RNC traz uma percepção futura de sustentabilidade e fazendas cada vez mais inteligentes.

Diversos trabalhos tem identificado a importância de RNCs na automatização de processos e em aplicações de processamento de imagens na agricultura. Pode-se destacar algumas aplicações no monitoramento ambiental como o mapeamento de pequenas áreas de plantação [Xie et al. 2019], na identificação ou segmentação de doenças nas folhas ou plantas [Abdullahi et al. 2017], classificação, segmentação e reconhecimento do tipo de cultura ou planta [Boulent et al. 2019] e de áreas cultivadas ou solo [Lu et al. 2017], detecção de ervas daninhas [dos Santos Ferreira et al. 2017], contagem de frutas [Rahnemoonfar and Sheppard 2017], detecção de macronutrientes pelas folhas [Wulandhari et al. 2019], etc.

As RNCs podem obter informações de imagens com baixa resolução e de um grande conjunto de dados (*dataset*). O método promissor tem se provado um excelente extrator de características por meio da criação de mapas com distintas dimensões [Kamilaris and Prenafeta-Boldú 2018]. Os mapas trazem características principalmente de arestas orientadas ou manchas de cor [Krizhevsky et al. 2012]. Os hiperparâmetros gerados como saída das RNCs associam essas características a um determinado rótulo da

imagem, neste trabalho é a deficiência de macronutriente.

Assim, a pesquisa apresentada é um estudo de caso com redes neurais convolucionais aplicadas na classificação da deficiência do macronutriente de potássio (K) pela folha de soja. Na seção 2 será apresentado os trabalhos relacionados e modelos de RNCs. Os materiais e métodos utilizados e as características das imagens e das RNCs serão definidas na seção 3. Os resultados de treinamento, a avaliação de precisão e acurácia, e suas respectivas discussões serão definidas na seção 4. A conclusão do trabalho é discorrida na seção 5.

## 2. Trabalhos Relacionados

A inteligência artificial evoluiu e diversas subáreas de pesquisa crescem de forma abrupta nos últimos anos como o aprendizado de máquina e o aprendizado profundo. Parte desse avanço, deve-se as melhorias criadas nos algoritmos empregando redes neurais artificiais (RNA) (perceptron) e diversas variações para problemas específicos. O aprendizado profundo vem se tornando uma realidade comercial na sociedade brasileira atual, principalmente nas melhorias de aplicações como os Chatbots. Nos últimos anos, a literatura mostra uma grande relevância entre os pesquisadores por métodos baseados em RNCs, ressaltando o aprendizado profundo para o monitoramento ambiental [Kamilaris and Prenafeta-Boldú 2018].

Comparadas às RNAs (Perceptron), as RNCs têm sua complexidade elevada pelo maior número de camadas e neurônios em sua estrutura, validando o termo aprendizado profundo. Tal complexidade, exige um maior desempenho, ou seja, recursos computacionais atuais para efetuar à sua generalização ou treinamento. A complexidade é compensada pelos resultados promissores encontrados em diversas aplicações. Distintos modelos de RNC tornaram-se populares entre os pesquisadores, destaca-se os modelos relacionados a classificação de uma grande quantidade de objetos como a AlexNet [Krizhevsky et al. 2012], o *Visual Geometry Group* (VGG) [Simonyan and Zisserman 2014], o GoogleNet [Szegedy et al. 2015] e o Inception-ResNet [Szegedy et al. 2017].

Quatro trabalhos foram relacionados [Wulandhari et al. 2019, Tran et al. 2019, Watchareeruetai et al. 2018, Hasan et al. 2018] com relação no uso de RNC, imagens e algum tipo de classificação com base em macronutrientes. Em [Tran et al. 2019], apresentam uma pesquisa para classificar e prever diferentes deficiências de macronutrientes (cálcio, potássio e nitrogênio) no processo de cultivo do tomate. Os autores também fizeram o uso de *data augmentation* para aumentar o conjunto de imagens. Duas arquiteturas baseadas em RNCs (Inception-ResNet v2 e Autoencoder) foram avaliadas para classificar e prever os sintomas de deficiência de macronutrientes, alcançando 87% e 79% de precisão respectivamente. Com o intuito de aumentar a precisão de avaliação, atingindo 91%, utilizou-se à estrutura estatística *Ensemble Averaging*. Em [Wulandhari et al. 2019], o trabalho propõe o uso de RNC (Inception-ResNet) baseado em transferência de aprendizado e *fine tuning* treinado pelo imagenet *dataset*, também utilizaram o aumento de conjunto de imagens pelo *data augmentation*. As folhas do quiabo são utilizadas como dados de entrada para detecção ou não de deficiência de nutrientes, com o desafio da dificuldade visual na distinção e tratamento precoce das plantas. Os resultados atingiram 86% de precisão na fase de teste.

A pesquisa de [Watchareeruetai et al. 2018] visa à identificação de deficiências nutricionais com base nas folhas de gramíneas pretas (*Vigna mungo*) cultivadas em ambientes controlados conforme seus macronutrientes. Abordaram dois grupos de plantas, um com macronutrientes regulares, e outro com as deficiências dos macronutrientes de K, Magnésio (Mg) e Nitrogênio (N). As imagens de entrada são divididas em pequenos blocos e alimentam diferentes RNCs, uma para cada tipo de deficiência de macronutriente. Todas as respostas são integradas para produzir a resposta final em uma rede Totalmente Conectada (TC). A precisão alcançada foi de 43% sobre todas as deficiências. No artigo [Hasan et al. 2018], os autores apresentam um sistema capaz de detectar e caracterizar espigas a partir de imagens de campo de trigo. A proposta vislumbra fazer uma avaliação quantitativa de produção de espigas de trigo como órgão de produção dos grãos. Assim, os autores treinaram quatro RNCs baseadas na região de captura das imagens para detectar e contar as regiões para emitir uma densidade de pico e uma probabilidade para cada parcela. A precisão média alcançada foi de 88 até 94% em diferentes conjuntos de imagens de teste, com distinção de 10 tipos de trigo e três tipos de tratamento de fertilizantes.

Em [dos Santos Ferreira et al. 2017] os autores realizaram a detecção de ervas daninhas de folhas largas e gramíneas em relação ao solo e a soja, utilizando imagens de VANTs. As RNCs foram comparadas com mais classificadores como máquina de vetor de suporte e florestas aleatórias. A RNC teve uma precisão entre todas as imagens acima de 99%, comparado as outras duas, respectivamente com 97% e 93%. A RNC utilizada foi a AlexNet [Krizhevsky et al. 2012], essa rede é considerada uma das precursoras da área pela relevância na classificação de mais de 1000 objetos em uma estrutura enorme de camadas e parâmetros, os autores não mostraram nenhuma alteração na estrutura. Já na revisão de [Kamilaris and Prenafeta-Boldú 2018] apresenta diversas aplicações, softwares, estruturas de RNCs, *datasets*, métricas e metodologias utilizadas na agricultura utilizando RNC, além de um estudo de caso para classificação, basicamente de solo e cana de açúcar, alcançando cerca de 92% a 94% de acerto. Porém, nessa revisão não mostrou nenhum trabalho com macronutrientes.

### **3. Materiais e Métodos**

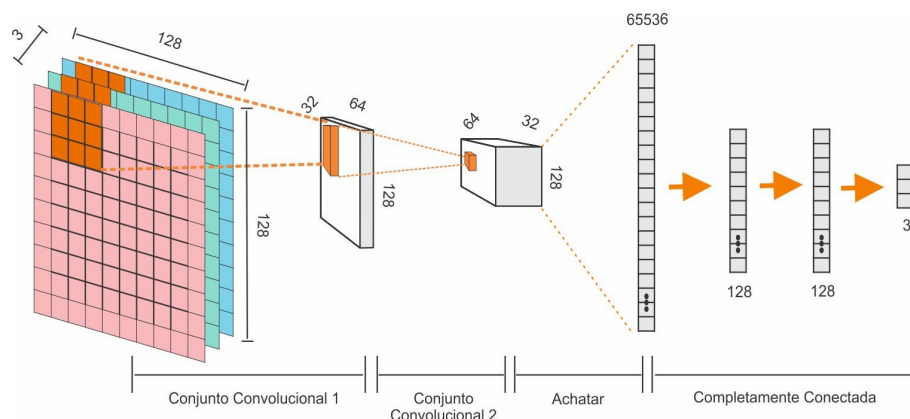
Os materiais e métodos estão constituídos de três partes: (i) arquitetura da RNC, (ii) conjunto de imagens e (iii) cenários.

#### **3.1. Arquitetura da RNC**

A arquitetura do modelo criado para RNC é composta de todos os tipos de camadas, as principais são: convolucional, pool e TC. Porém, diferente do trabalho de [Wulandhari et al. 2019], o modelo criado teve como objetivo a redução de custos computacionais, ou seja, um menor número de camadas, vislumbrando trabalhos futuros.

Diversas arquiteturas foram analisadas e avaliadas em diferentes cenários para melhor inferir os resultados. O modelo base otimizado tem 2 conjuntos convolucionais (CC): com 2 camadas convolucionais e 1 camada pool para cada conjunto; bem como um conjunto de camadas totalmente conectadas com 3 camadas densas. A arquitetura base otimizada é apresentada na Figura 1, com todas as camadas do modelo e suas respectivas dimensões de volume. Entre os conjuntos, convolucional e totalmente conectadas, tem uma camada “flatten” para achatamento dos dados, com o objetivo de adaptar as entradas para o conjunto de camadas totalmente conectadas.

Em relação ao volume, nas camadas convolucionais foram aplicados 32 e 64 filtros para cada conjunto e na camada *pool* a função “maxpooling” com *strides* 2. Nas camadas convolucionais contém *zero padding* e na maxpooling não. A função de ativação “Relu” foi aplicada em todas as camadas convolucionais e totalmente conectadas, com exceção da última camada totalmente conectada, que fez uso da função de ativação “softmax”.



**Figura 1. Arquitetura da RNC otimizada e a relação com o volume de dados de cada etapa, adaptado de [Zhang et al. 2015].**

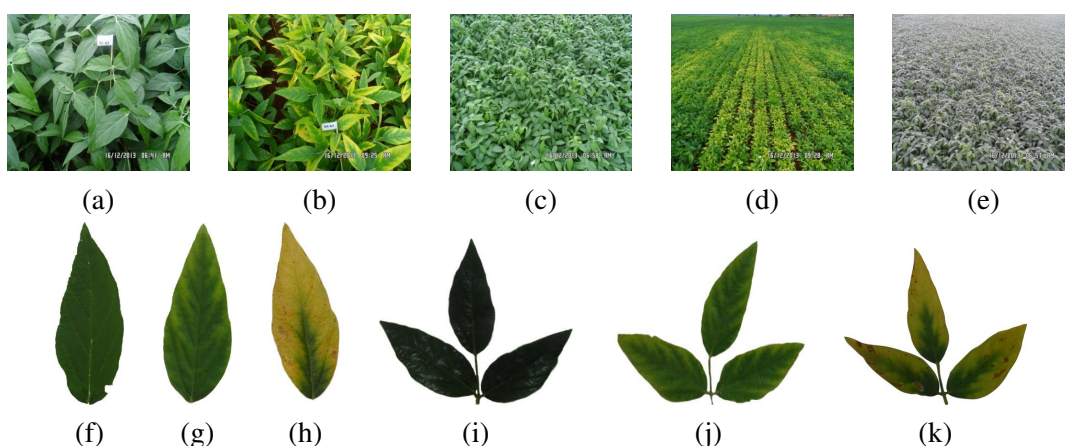
A quantidade de neurônios nas camadas totalmente conectadas foram 128, 128 e 3, respectivamente da primeira para a última camada. Os 3 neurônios de saída classificam as imagens quanto à sua deficiência de potássio (K) em três classes: sem deficiência, com deficiência baixa e com deficiência severa. As três classes definidas estão correlacionadas aos três tipos de tratamentos analisados do macronutriente de potássio, conforme seção 3.2. Todo o processamento, criação dos modelos das RNCs e avaliação dos resultados foram feitos com a linguagem python junto à biblioteca *open source tensorflow*.

### 3.2. Conjunto de Imagens (*dataset*)

As imagens foram adquiridas a partir de uma estação experimental no estado de Mato Grosso, Brasil. O cultivo da soja tinha três tipos de tratamentos de fertilizantes: comercial, falta do macronutriente K há um ano e falta do macronutriente K há 6 anos. As imagens foram adquiridas com uma câmera de baixo custo, Samsung de 14 Megapixel e resolução de 4320x3240. As imagens foram redimensionadas para cinco dimensões [16 x 16 x 3] [32 x 32 x 3] [64 x 64 x 3] [128 x 128 x 3] [256 x 256 x 3], ou seja, cinco *datasets* distintos, todos com o modelo de cor RGB.

A construção do *dataset* foi a partir de 66 imagens sementes dentre folhas, trifólios e folhagens, conforme Figura 2. Observa-se alguns exemplos de imagens com os três tipos de classes: sem deficiência, com deficiência baixa e com deficiência severa.

Nas imagens de folhas e trifólios, Figura 2[(f)-(k)], contém imagens dos três tipos de classe, quando a folha ou trifólio tem predominância de verde é considerada como classe “baixa”, se a predominância for de amarelo a classe é definida como “severa”. Em contrapartida, nas imagens de folhagens contém apenas imagens de duas classes, sem deficiência e com deficiência severa, Figura 2[(a)-(e)]. Devido a quantidade de folhas de classes diferentes em cada mancha na plantação. As imagens de folhagens foram capturadas de diversos ângulos, alturas, iluminação ambiente e plano de fundo complexo. Já as



**Figura 2. Imagens de folhas, trifólios e folhagens com e sem deficiência de k e distintas alturas, ângulos e estádios vegetativos: (a)(c)(e)(f)(i) sem deficiência; (g)(j) com deficiência baixa; (b)(d)(h)(k) com deficiência severa.**

imagens de folhas e trifólios foram obtidas no laboratório com as mesmas configurações de ângulo (90 graus), altura (30cm), iluminação e plano de fundo controlado.

A quantidade de cada tipo de imagem está relacionada na Tabela 1, as imagens sementes foram capturadas nos ambientes real e laboratório. Já as demais imagens, foram geradas pela técnica de *data augmentation*. O conjunto de imagens desbalanceado foi utilizado para obtenção dos resultados da seção 4.2 em todos os cenários, Figura 5. Já o conjunto de imagens balanceado foi analisado apenas com o melhor cenário, Tabela 2, para obtenção dos resultados das Tabelas 3 e 4.

**Tabela 1. Conjunto de Imagens (*dataset*).**

Tipos de Imagens	Desbalanceado		Balanceado	
	Sementes	Total	Sementes	Total
Folhas	16	2.416	18	2.718
Trifólios	12	1.812	18	2.718
Folhagens	38	5.738	18	2.718
Total	66	9.966	54	8.154

Com essa técnica o total de imagens do *dataset* é de 9.966 no conjunto desbalanceado e 8.154 no conjunto balanceado. As imagens foram geradas com as seguintes alterações: rotação, giro, cisalhamento, brilho, alteração de canal e ampliação.

### 3.3. Cenários e treinamento

Três cenários foram analisados com o intuito de avaliar a melhor RNC para o problema de deficiência do macronutriente K. Uma das principais modificações está relacionada a arquitetura da RNC. Com base na Figura 1 foram definidos três cenários com a distinção entre suas arquiteturas. O cenário 1 tem a arquitetura base mais otimizada com dois conjuntos convolucionais (2 CC), uma camada de achatamento (1 A), e três camadas completamente conectadas (3 TC). O cenário 2 tem 4 CC, 1A e 3 TC, e o cenário 3 com 5CC, 1 A e 4 TC.

Outra alteração está na quantidade de filtros aplicados em cada conjunto convolucional. O cenário 1 tem 2 CC, respectivamente, a sequência de filtros foram 32 e 64, Figura 1. Já o cenário 2 teve a sequência 32, 64, 128 e 128, e por fim, o cenário 3 com a sequência 64, 128, 256, 512 e 512. Em cada cenário foi aplicado cinco tipos de imagens RGB de entrada com tamanhos quadrados (largura e altura) de 16, 32, 64, 128 e 256. O tamanho 16 não foi aplicado no cenário 3, devido a dimensão negativa do volume de dados após as camadas convolucionais.

O conjunto de imagens do *dataset* foi criado pelos autores e contém três tipos principais: folhas, trifólios e folhagens; bem como os três tipos de deficiência. As classes foram distribuídas no treinamento conforme Figura 5(a), com a relação das três classes e a quantidade de imagens em cada uma. Definindo as classes como: 0 – “sem deficiência”, 1 – “deficiência baixa”, 2 – “deficiência severa”.

As configurações de treinamento como o otimizador, a função loss e métricas foram as mesmas em todos os casos. Essas configurações são amplamente utilizadas na literatura e avaliadas anteriormente, não apresentado aqui pela falta de espaço. O modelo da RNC foi treinado por 10 épocas em todos os cenários, devido ao tempo de processamento alto pelo tamanho do *dataset*. Na compilação do modelo utilizou-se o otimizador “Adam” com a taxa de aprendizagem de 0,0002, a função *loss* na descida do gradiente é a entropia categórica e a acurácia categórica como métrica de avaliação.

A partir do total de imagens foram separadas 70% para o treinamento (6.976 imagens) e 30% para teste (2.990 imagens) de forma aleatória. Porém, todos os cenários são avaliados com o mesmo *dataset*. A evolução do treinamento pode ser observada na Figura 3, com o erro chegando muito próximo do zero em todos os casos e a acurácia inversamente proporcional. Vale ressaltar sobre o comportamento dos cenários 1 e 2 nas imagens de 16, 32 e 64, observa-se valores menores cerca de 5% de perda de acurácia em relação aos demais. O cenário 3 tem alta acurácia, mesmo com tamanhos menores de imagens (32 e 64). No entanto, o número de parâmetros dessa arquitetura segue o sentido contrário, maior em relação as demais.

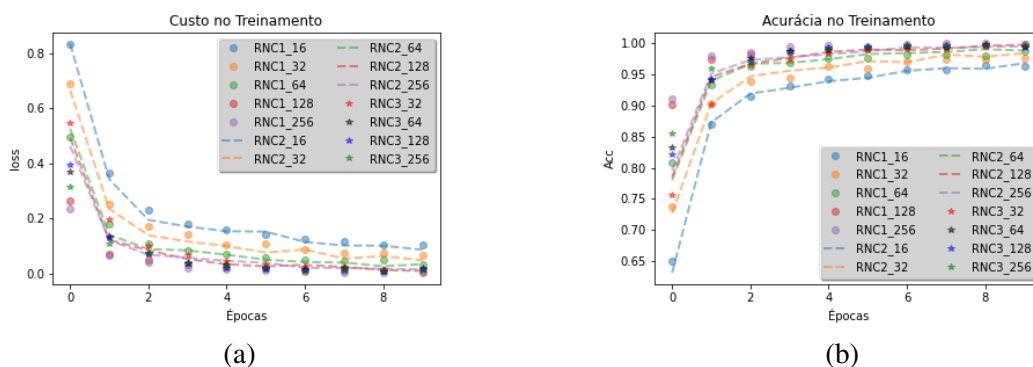
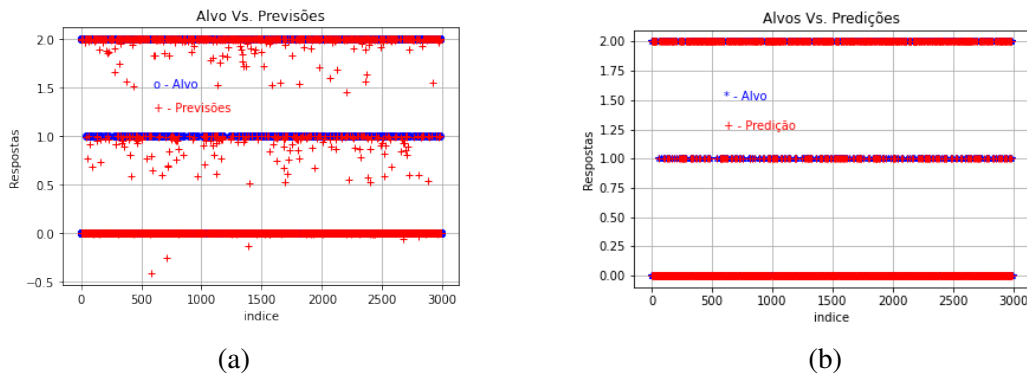


Figura 3. Treinamento com todos os cenários avaliados: (a) custo; (b) acurácia

#### 4. Resultados

Os resultados foram divididos em três partes, a forma de classificação das folhas, a avaliação de todos os cenários e a comparação com a literatura atual. Todos os resultados



**Figura 4. Comparação das previsões das imagens de teste com os seus respectivos alvos: (a) análise de proximidade; (b) validação final.**

foram obtidos a partir do conjunto completo de imagens balanceado e desbalanceado de deficiência do macronutriente K na soja, conforme Tabela 1.

#### 4.1. Forma de classificação e métricas

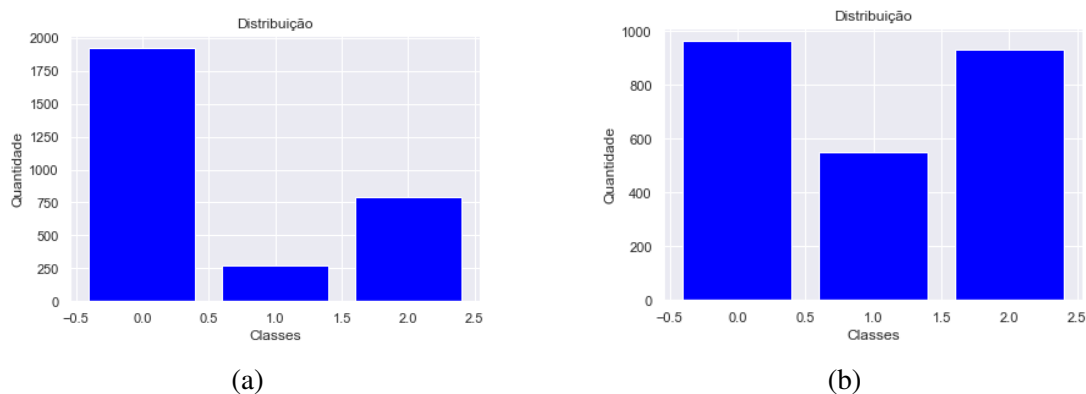
A partir das previsões das imagens de teste, a análise dos dados efetuou-se de duas formas, proximidade das previsões, precisão e acurácia dentre todas as imagens de teste na classificação. A avaliação da proximidade das previsões e suas respectivas classes, teve como intuito observar a proximidade que uma previsão está de sua respectiva classe ideal. Por exemplo, as imagens (img) classificadas como pertencentes a classe com deficiência baixa deve respeitar a seguinte regra:  $1,0 \geq img > 0,5$ ; conforme definição na seção 3.2. Caso contrário, o ponto pode ser enquadrado em outra classe, significa que se a imagem não for classificada com 50% ou mais chance de ser daquela classe, a imagem será considerada como falso positivo ou negativo.

A Figura 4-(a) está relacionada a previsão das imagens utilizando a arquitetura do cenário 1 com tamanho 128 nas imagens. Assim, observa-se que grande parte das imagens foram classificadas de forma correta, ou seja, com 50% ou mais chances de pertencerem a determinada classe, foi definido os números 0, 1 e 2 para melhor distinção visual das classes. A validação final da previsão é ilustrado na Figura 4-(b).

Outra avaliação está na quantidade e proporção de imagens em cada classe. Uma dedução importante a ser avaliada, está relacionada ao modelo proposto e a classe mais difícil de ser classificada. A classe 1 de “deficiência baixa” foi a mais difícil de ser classificada no cenário 1, por ter mais pontos fora da reta ideal. Porém, os resultados mostram a excelente generalização das redes neurais convolucionais.

Na validação dos dados definimos a relação entre Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN). Considerando VP como as imagens classificadas de forma correta com deficiência de macronutriente K, independente se for deficiência baixa ou severa, e VN para as imagens sem deficiência. Desta forma, efetuou-se o cálculo de precisão do método proposto com as equações mais populares entre os pesquisadores para comparação dos estudos em classificação de imagens, conforme relevância e correlação definida em [Taha and Hanbury 2015].





**Figura 5. Distribuição das classes no *dataset*: (a) desbalanceada; (b) balanceada.**

## 4.2. Avaliação dos cenários

Os cenários foram agrupados conforme a sua arquitetura e o tamanho da imagem de entrada, detalhes na seção 3.1. Com base nos resultados do treinamento da Figura 3, selecionou-se os três melhores tamanhos de imagens de entrada 64, 128 e 256. Assim, o cenário 1 com tamanho de imagens de entrada [128x128x3] terá o formato “RNC1\_128”, respectivamente aos demais cenários e tamanhos. Desta forma, em cada uma das três arquiteturas avaliou-se a validação dos três tipos de imagens de entradas (64, 128 e 256), conforme a predição realizada com as 2.990 imagens de teste. Os 9 cenários foram avaliados com base nas métricas definidas na seção 4.1 e apresentados na Tabela 2. As classes foram distribuídas na etapa de teste de forma desbalanceada, conforme Figura 5-(a), com grande semelhança proporcional na distribuição do treinamento. Já a Figura 5-(b), ilustra o balanceamento das classes no *dataset* de teste com o melhor método proposto. Apenas a classe baixa ficou comprometida devido à sua inexistência em imagens de folhagens, conforme discutido na seção 3.2.

Os cenários obtiveram excelentes resultados no *dataset* com valores acima de 99%, conforme a Tabela 2. Em uma análise mais específica, nota-se que a arquitetura do cenário 1 teve conservação nos seus resultados mesmo com a alteração no tamanho das imagens. A diferença de precisão entre todos os cenários não é relevante e o número de parâmetros é menor em RNC1\_128, assim escolhemos esse cenário como proposta de validação. Já o cenário RNC1\_64 obteve ótimos valores de validação, mas na fase de treinamento (Figura 3) os resultados ficaram um pouco abaixo do RNC1\_128. O cenário 1 é mais otimizado em relação aos recursos computacionais e também alcançou robustez e continuidade nos resultados. Além disso, os resultados do cenário RNC1\_128 ficaram piores apenas que o cenário RNC3\_128, cuja arquitetura e parâmetros é maior.

Em todos os cenários é notória uma pequena diferença de precisão, geralmente menor que 1%. O RNC3\_64 obteve uma certa discrepância nos valores. Essa exceção está relacionada ao número de filtros iniciais da arquitetura, ou seja, o número de filtros é igual ao tamanho da imagem, isso pode ter influenciado os resultados.

A Tabela 3 contém os dados da matriz de confusão do cenário (RNC1\_128) com o conjunto de imagens balanceado. Percebe-se uma maior dificuldade no método para a diferenciação entre as classes baixa e severa, mas ainda com ótimos resultados, acima

**Tabela 2. Resultados de todos os cenários.**

Cenários	Precisão	Acurácia	Recall	Especificidade
RNC1_64	99,34%	99,73%	99,90%	99,63%
RNC1_128	100,00%	99,93%	99,81%	100,00%
RNC1_256	99,81%	99,93%	100,00%	99,89%
RNC2_64	100,00%	99,83%	99,53%	100,00%
RNC2_128	100,00%	98,96%	97,17%	100,00%
RNC2_256	99,71%	99,89%	100,00%	99,84%
RNC3_64	97,84%	99,23%	100,00%	98,81%
RNC3_128	100,00%	100,00%	100,00%	100,00%
RNC3_256	100,00%	99,96%	99,90%	100,00%

de 97%. Levando em consideração apenas a distinção de imagens com deficiência e sem deficiência, o método chega próximo aos 100,00% de precisão.

**Tabela 3. Matriz de confusão da proposta com *dataset* balanceado.**

	S/Defic.	Baixa	Severa	
S/Defic.	962	1	0	<b>99,90%</b>
Baixa	0	539	12	<b>97,82%</b>
Severa	0	2	931	<b>99,79%</b>
	<b>100,00%</b>	<b>99,45%</b>	<b>98,73%</b>	<b>99,17%</b>

### 4.3. Comparação com a literatura

A comparação do cenário RNC1\_128, com arquitetura mais otimizada, e os trabalhos relacionados na seção 2 podem ser vistos na Tabela 4. Os resultados são apresentados apenas com três métricas, devido a inexistência da especificidade nos demais trabalhos. Vale ressaltar que o trabalho dessa pesquisa analisa um macronutriente, o potássio, os trabalhos [Wulandhari et al. 2019, Tran et al. 2019, Watchareeruetai et al. 2018] analisam mais de um macronutriente. Já em [Hasan et al. 2018] não entra em detalhes sobre qual macronutriente foi analisado nos resultados.

Outra relação aos demais trabalhos, está na precisão e acurácia do método proposto (cenário 1) comportou-se de forma satisfatória nos resultados ao *dataset* imposto. A grande quantidade de imagens também teve grande relação na precisão do método, em diversos trabalhos [Wulandhari et al. 2019, Dosovitskiy et al. 2014] mostram que o aumento do *dataset* contribui para a precisão da RNC, e a *augmentation* é um bom caminho caso o *dataset* seja pequeno. Por exemplo, em [Tran et al. 2019] poucas unidades classificadas de forma errada têm grande impacto na precisão e acurácia da proposta. Com o aumento da quantidade de imagens, a capacidade de generalização da RNC é maior e consequentemente melhora o seu desempenho. Uma possível razão para a baixa precisão nos trabalhos de [Wulandhari et al. 2019, Tran et al. 2019] é o uso das redes ResNet. Tais redes foram construídas para uma grande quantidade de classes e o seu treinamento não é específico para um determinado problema como a deficiência de macronutriente, isso também pode ter influenciado nos resultados.

**Tabela 4. Resultados da classificação de imagens.**

<b>Trabalhos</b>	<b>Precisão</b>	<b>Acurácia</b>	<b>Recall</b>
[Tran et al. 2019]	87.83%	-	87,47%
[Hasan et al. 2018]	97.27%	96,82%	99,52%
[Wulandhari et al. 2019]	-	86,00%	-
[Watchareeruetai et al. 2018]	43.02%	-	52,13%
Proposta	100.00%	99,95%	99,93%

## 5. Conclusão

Neste trabalho apresentou uma proposta de classificar e rotular as folhas de soja com base na deficiência do macronutriente de potássio. A proposta foi validada com o uso de técnicas baseadas em aprendizado profundo e comparada com a literatura atual.

O trabalho apresentou um *dataset* próprio de folhas de soja com base em uma estação experimental com cultivar utilizado comercialmente no Brasil. Os testes realizados com o classificador demonstraram a capacidade de generalização do problema, mesmo com um grande número de imagens, diferentes formas de balanceamento e condições ambientais distintas. A arquitetura da RNC otimizada demonstra a importância da especificação da mesma conforme o problema proposto. Consolidando o aprendizado profundo em uma perspectiva atrativa no reconhecimento ou monitoramento de macronutrientes por imagens foliares.

## 6. Agradecimentos

Os autores gostariam de agradecer a Universidade do Estado de Mato Grosso (UNEMAT), a Universidade Estadual Paulista (UNESP), a Universidade Federal de Goiás (UFG), a Fundação MT e a Fundação de Amparo à Pesquisa do Estado de Mato Grosso (FAPEMAT) - Processo: 276780/2017.

## Referências

- Abdullahi, H. S., Sheriff, R., and Mahieddine, F. (2017). Convolution neural network in precision agriculture for plant image recognition and classification. In *2017 Seventh International Conference on Innovative Computing Technology (Intech)*, Ieee, Londrés, pages 1–3.
- Boulent, J., Foucher, S., Théau, J., and St-Charles, P.-L. (2019). Convolutional neural networks for the automatic identification of plant diseases. *Frontiers in plant science*, 10.
- dos Santos Ferreira, A., Freitas, D. M., da Silva, G. G., Pistori, H., and Folhes, M. T. (2017). Weed detection in soybean crops using ConvNets. *Computers and Electronics in Agriculture*, 143:314–324.
- Dosovitskiy, A., Springenberg, J. T., Riedmiller, M., and Brox, T. (2014). Discriminative unsupervised feature learning with convolutional neural networks. In *Advances in neural information processing systems*, pages 766–774.
- Hasan, M. M., Chopin, J. P., Laga, H., and Miklavcic, S. J. (2018). Detection and analysis of wheat spikes using convolutional neural networks. *Plant Methods*, 14(1):100.

- Kamilaris, A. and Prenafeta-Boldú, F. X. (2018). A review of the use of convolutional neural networks in agriculture. *The Journal of Agricultural Science*, 156(3):312–322.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.
- Lu, H., Fu, X., Liu, C., Li, L.-g., He, Y.-x., and Li, N.-w. (2017). Cultivated land information extraction in UAV imagery based on deep convolutional neural network and transfer learning. *Journal of Mountain Science*, 14(4):731–741.
- Lucchi, B. B., Conchon, R., Nakamura, C. Y., Alves, C. F., Schwantes, F., Rivaldo, G. C., Bianchi, I., Barros, L., and Camuri, P. A. (2020). Agropecuária cresce mais que a média da economia brasileira em 2019. techreport, Confederação da Agricultura e Pecuária do Brasil.
- Rahnemoonfar, M. and Sheppard, C. (2017). Deep count: fruit counting based on deep simulated learning. *Sensors*, 17(4):905.
- Sfredo, G. J. (2008). *Soja no Brasil: calagem, adubação e nutrição mineral*. Embrapa Soja Londrina.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9.
- Taha, A. A. and Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, 15(1):29.
- Tran, T.-T., Choi, J.-W., Le, T.-T. H., and Kim, J.-W. (2019). A comparative study of deep CNN in forecasting and classifying the macronutrient deficiencies on development of tomato plant. *Applied Sciences*, 9(8):1601.
- Watchareeruetai, U., Noinongyao, P., Wattanapaiboonsuk, C., Khantiviriya, P., and Duangsrissai, S. (2018). Identification of plant nutrient deficiencies using convolutional neural networks. In *2018 International Electrical Engineering Congress (iEECON)*, pages 1–4. IEEE.
- Wulandhari, L. A., Gunawan, A. A. S., Qurania, A., Harsani, P., Tarawan, T. F., and Hermawan, R. F. (2019). Plant nutrient deficiency detection using deep convolutional neural network. *ICIC Express Letters*, 13(10):971–977.
- Xie, B., Zhang, H. K., and Xue, J. (2019). Deep convolutional neural network for mapping smallholder agriculture using high spatial resolution satellite image. *Sensors*, 19(10):2398.
- Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., and Shen, D. (2015). Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation. *NeuroImage*, 108:214–224.