



**GABRIEL HENRIQUE RIBEIRO DOS SANTOS**

**USO DE REDES NEURAIAS CONVOLUCIONAIS PARA  
SEGMENTAÇÃO DE SOLO EM IMAGENS AÉREAS DE  
PLANTAÇÕES DE SOJA**

**LAVRAS-MG  
2022**

**GABRIEL HENRIQUE RIBEIRO DOS SANTOS**

**USO DE REDES NEURAIAS CONVOLUCIONAIS PARA  
SEGMENTAÇÃO DE SOLO EM IMAGENS AÉREAS DE  
PLANTAÇÕES DE SOJA**

Monografia apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Engenharia de Controle e Automação, para a obtenção do título de Bacharel.

Prof. Dr. Danton Diego Ferreira  
Orientador

**LAVRAS-MG  
2022**

**RESUMO**

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

**ABSTRACT**

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

## SUMÁRIO

1	INTRODUÇÃO.....	5
2	OBJETIVO.....	8
2.1	Objetivo geral.....	8
2.2	Objetivos específicos.....	8
3	REFERENCIAL TEÓRICO.....	9
3.1	Agricultura de precisão e sensoriamento remoto.....	9
3.1.1	Aquisição de imagens.....	10
3.2	Imagens digitais.....	10
3.3	Espaço de cores.....	11
3.3.1	RGB.....	11
3.3.2	HSV.....	12
3.4	Processamento de imagens.....	12
3.4.1	Segmentação por cores.....	13
3.4.2	Binarização.....	14
3.4.3	Transformações morfológicas.....	14
3.4.3.1	Erosão.....	15
3.4.3.2	Dilatação.....	16
3.5	Redes neurais artificiais.....	16
3.5.1	Redes neurais convolucionais.....	18
4	METODOLOGIA.....	20
4.1	Aquisição das imagens.....	20
4.2	Geração do dataset.....	21
4.3	Arquitetura da rede neural.....	22
4.4	Treinamento da rede neural.....	23
4.4.1	Pré-processamento dos dados.....	23
4.4.2	Função de perda (Loss).....	24
4.4.3	Métrica.....	25
4.4.4	Parâmetros de treinamento.....	26
4.4.5	Hardware utilizado.....	27
4.5	Avaliação da qualidade e tempo de inferência.....	28
5	RESULTADOS E DISCUSSÃO.....	29
5.1	Treinamento.....	29
5.2	Inferência.....	30
5.2.1	Qualidade da segmentação.....	30
5.2.2	Tempo de inferência.....	31
6	CONCLUSÃO.....	32
	REFERÊNCIAS.....	33

## 1 INTRODUÇÃO

O Brasil é um dos maiores exportadores de commodities agrícolas do mundo. O setor agropecuário contribui com 21,1% do PIB e emprega 20% da força de trabalho (EMBRAPA, 2019). Segundo Hespanhol (2010), no país existe uma grande indústria em torno do complexo agroindustrial da soja, cana-de-açúcar, laranja e entre outros, os quais utilizam de soluções tecnológicas, tais como fertilizantes químicos, biocidas, sementes selecionadas, máquinas e etc. Uma das principais culturas do agronegócio brasileiro é a soja, isso faz com que o Brasil seja um dos maiores produtores do grão. O país é o segundo maior produtor de soja do mundo, e a estimativa de produção para a safra 2020/21 foi de 134,9 milhões de toneladas, com a utilização de 38,2 milhões de hectares (CONAB, 2020).

Diante do aumento populacional e da escassez de recursos naturais, principalmente de terra e água, atualmente o maior desafio da agricultura é garantir a segurança alimentar mundial de forma ambientalmente sustentável. Segundo Massruhá et al. (2020), projeções feitas levando em consideração os padrões de crescimento populacional e de consumo de alimentos apontam que a produção agrícola terá que aumentar em pelo menos 70% para suprir a demanda mundial até 2050. Existem também as ações das mudanças climáticas que cada vez mais agem de formas imprevisíveis, causando perda de produtividade agrícola. Conseqüentemente, isso gera instabilidades no fornecimento de alimentos e aumento mundial de preços, prejudicando principalmente as pessoas com maior vulnerabilidade socioeconômica. Nesse contexto, o desenvolvimento de uma agricultura inteligente baseada na utilização de novas tecnologias, é essencial para alcançar as metas futuras de segurança alimentar (FAO, 2010).

Desde a revolução verde, a agricultura tem cada vez mais aplicado soluções tecnológicas a fim de maximizar a produtividade por meio da modificação genética de sementes, da utilização de insumos industrializados como fertilizantes e agrotóxicos, e a mecanização da colheita. Um dos resultados da implementação de novas tecnologias foi o surgimento da agricultura de precisão (AP), que nos países de tecnologia avançada, é denominado por eles de Precision Agriculture, Precision Farming ou SiteSpecific Crop Management (MANZATTO et al., 1999). Dentre as técnicas que fazem parte do pacote da agricultura de precisão está o sensoriamento remoto. De acordo com Dos Santos (2021), trata-se do conhecimento espacial de forma precisa da área produtiva, com auxílio de imagens aéreas obtidas por meio de satélites ou veículos aéreos não-tripulados (VANTs), popularmente conhecidos como drones. Isso permite um conhecimento melhor da saúde de

toda plantação e ajuda no gerenciamento de recursos durante a aplicação, assim tem-se menos desperdício de insumos. Também permite que sejam realizadas comparações de diferentes produtos, principalmente aqueles relacionados à fisiologia e nutrição vegetal, a fim de evidenciar os benefícios de sua aplicação. Para monitorar e analisar em larga escala, o sensoriamento remoto é o método mais econômico (KHANAL et al., 2017).

O imageamento de plantações com altas resoluções espaciais proporcionadas por imagens de VANT, possibilita a realização de análises baseadas na reflectância espectral das plantas. O principal subproduto destas imagens são os índices de vegetação, que fornecem informações importantes sobre o plantio. Estes índices resultam de operações matemáticas pixel a pixel entre duas ou mais imagens monocromáticas (geralmente entre bandas específicas do espectro eletromagnético) com objetivo de evidenciar certas características da cobertura vegetal. Dentre essas características, pode-se citar o nível de processos fotoquímicos relacionados à fotossíntese e níveis de pigmentação de folhas (LIMA, 2020). Por meio de índices de vegetação é possível estimar o índice de área foliar (IAF), que é um dos parâmetros correlacionados com a produtividade, já que quanto maior a área foliar maior será interceptação da energia luminosa, aumentando o processo de fotossíntese (FAVARIN et al., 2002). Porém, um dos problemas de se utilizar as técnicas tradicionais de geração de índices de vegetação é a influência do solo exposto adjacente, que dependendo da coloração pode aumentar ou diminuir de forma artificial os valores dos índices. Da mesma maneira, existem também as plantas invasoras que podem influenciar nesses resultados, dificultando na análise de desempenho fisiológico. Diante disso, é desejável a identificação somente dos pixels correspondentes às plantas de interesse para compor os índices. Ou seja, o problema torna-se de classificação a nível de pixel, também conhecido como segmentação semântica (LONG et al., 2015).

Atualmente, a visão computacional está fortemente relacionada com a utilização de ferramentas provindas da inteligência artificial. A consequência disso são algoritmos baseados em redes neurais artificiais capazes de classificar e detectar objetos de interesse em imagens. A classificação de imagens consiste em dizer se em uma determinada imagem existe ou não um objeto de uma determinada classe. A detecção de objetos vai um pouco mais além, consiste em classificar e localizar a posição do objeto por meio das coordenadas de uma caixa delimitadora. O próximo passo para atingir uma inferência mais refinada é a segmentação semântica, que consiste em classificar cada pixel que compreende um objeto com sua classe respectiva (GARCIA-GARCIA et al., 2017). Dessa forma, o algoritmo é capaz de segmentar

um objeto respeitando o seu formato geométrico. Em muitos problemas de visão computacional, como classificação e detecção de objetos, as redes neurais convolucionais apresentam os melhores resultados (VARGAS et al., 2016). E segundo Wang et al. (2018), os modelos do estado da arte da segmentação também são, em sua maioria, baseados nas redes neurais convolucionais.



## 2 OBJETIVO

### 2.1 Objetivo geral

Desenvolver e validar um algoritmo de segmentação de solo, baseado em uma rede neural convolucional de segmentação semântica, em imagens aéreas obtidas por VANT, de plantações de soja.

### 2.2 Objetivos específicos

Para alcançar o propósito do trabalho, foram determinados os seguintes objetivos específicos:

- a) Validar o método de *autolabel* utilizando técnicas clássicas de segmentação por cor e transformação morfológica;
- b) Comparar o algoritmo desenvolvido com técnicas clássicas de segmentação baseadas em filtro no espaço de cores;
- c) Analisar a velocidade de inferência do algoritmo, sendo o custo computacional para processamento um dos requisitos para validar o algoritmo.

### **3 REFERENCIAL TEÓRICO**

#### **3.1 Agricultura de precisão e sensoriamento remoto**

Segundo o Ministério da Agricultura, Pecuária e Abastecimento (Mapa), a definição de Agricultura de Precisão (AP) é “um sistema de gerenciamento agrícola baseado na variação espacial e temporal da unidade produtiva e visa ao aumento de retorno econômico, à sustentabilidade e à minimização do efeito ao ambiente” (BRASIL, 2012). Em outras palavras, trata-se do manejo inteligente das culturas, de forma a otimizar a utilização de insumos por meio de dados obtidos no campo. Esses dados, que podem ser obtidos por vários tipos de sensores, são georreferenciados e auxiliam na tomada de decisões em relação ao quanto de insumo é necessário aplicar em cada região da área de plantio. Sem a Agricultura de Precisão, as aplicações são feitas sem considerar a variabilidade espacial da lavoura, ou seja, são feitas de maneira uniforme. Isso faz com que regiões mais saudáveis recebam excesso de fertilizantes e de defensivos agrícolas, e/ou ao contrário. A utilização excessiva de insumos agrícolas causa impactos econômicos e ambientais (JUNIOR et al., 2006). Além disso, em áreas muito grandes é impossível o produtor saber quais são as heterogeneidades do plantio. Com as ferramentas tecnológicas de AP é possível conhecer estas heterogeneidades (BERNARDI et al., 2014). Entre as tecnologias que podem ser utilizadas na AP, tem-se o sensoriamento remoto. (PIRES et al., 2004)

O sensoriamento remoto (SR), como o próprio nome sugere, se refere à obtenção de dados à distância, ou seja, sem a necessidade de existir o contato físico com a plantação. Os principais sensores envolvidos nessa tecnologia são as câmeras fotográficas, que podem estar embarcadas em satélites, aeronaves, veículos aéreos não-tripulados (VANT), máquinas agrícolas autônomas e etc. (BERNARDI et al., 2014). A utilização das câmeras permite analisar a reflectância das plantas em vários comprimentos de onda eletromagnética, que está relacionada com processos bioquímicos, como a fotossíntese, por exemplo. Essas informações podem ser exploradas para estimar a biomassa, o vigor das plantas, se existe estresse hídrico e etc. De acordo com Jensen (1996), algumas das principais vantagens do SR são a coleta de dados não destrutiva e a cobertura de áreas extensas.

### 3.1.1 Aquisição de imagens

As duas principais formas de obter imagens aéreas de lavouras são por meio de satélites e por veículos aéreos não-tripulados (VANTs). O sensoriamento remoto orbital (via satélites) possui algumas desvantagens como a interferência de condições meteorológicas, distorções causadas pela atmosfera, baixa resolução espacial, período de órbita muito longo e etc. Porém, é possível obter imagens provenientes de câmeras multiespectrais com maior facilidade. De acordo com Bernardi et al. (2004), os sensores multiespectrais cobrem o espectro eletromagnético nas faixas do visível, infravermelho próximo, médio e termal. Já as imagens obtidas por VANTs, popularmente conhecidos como drones, possuem resolução espacial melhores e podem ser obtidas com maior frequência temporal, permitindo uma análise espacial e temporal mais refinada. A utilização dos VANTs tem se popularizado devido ao avanço tecnológico que resultou em uma redução de custos na aquisição desses equipamentos. Porém, o alto custo das câmeras multiespectrais tem limitado a utilização dos VANTs à obtenção de imagens apenas na faixa do visível, com câmeras RGB (do inglês Red, Green e Blue).

### 3.2 Imagens digitais

As imagens digitais monocromáticas (1 canal) são matrizes bidimensionais onde seus elementos básicos são chamados de pixels. Os pixels são amostras discretas de uma imagem contínua, projetados em uma matriz 2-D (*amostragem*), e seus valores correspondem a sua intensidade luminosa (*quantização*) (GONZALEZ & WOODS, 2000). Para imagens digitais coloridas (3 canais) a lógica é a mesma, porém, nesse caso, os pixels possuem três valores de intensidades correspondentes às cores vermelho, verde e azul (canais R, G e B).

A qualidade de uma imagem digital está diretamente relacionada à quantidade de pixels. Quanto mais pixels, mais detalhes serão visíveis, ou seja, uma melhor resolução espacial. Mas segundo Leão (2005), essa não é a única característica que define a qualidade de uma imagem digital, existe também a dependência da profundidade de cor, que por vez está relacionada à quantização da intensidade luminosa. A intensidade luminosa, que é uma grandeza contínua, precisa ser quantizada para que esta informação seja processada e armazenada digitalmente. Segundo Godoy (2003), o processo de quantização da intensidade, uma grandeza contínua, trata-se da representação da amostra em números binários. Quanto

mais bits utilizados nesta representação binária, maior será a resolução de cores. Geralmente utiliza-se 8 bits por canal, que significa 256 níveis de cor. Para uma imagem RGB, a quantidade de cores possíveis (profundidade de cor) será de  $256*256*256$ , resultando em mais de 16,7 milhões de cores (Leão, 2005).

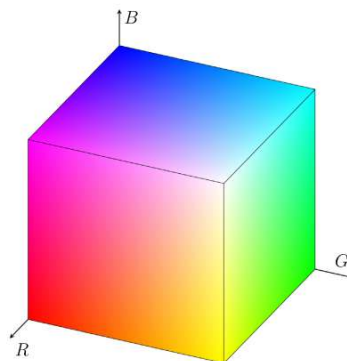
### 3.3 Espaço de cores

O objetivo por trás dos diferentes espaços de cores é representar de forma espacial as cores, onde cada ponto deste espaço corresponde a uma cor específica. Isso serve para padronizar a representação das cores, facilitando nas especificações das mesmas.

#### 3.3.1 RGB

O espaço de cores RGB utiliza como premissa a adição das cores primárias vermelho (*Red*), verde (*Green*) e azul (*Blue*) na composição de todas as outras cores. Esse sistema é orientado a *hardware* já que a maioria dos eletrônicos com monitores utilizam o mesmo princípio para geração de imagens na tela. Nesse espaço de cores é utilizado o sistema de coordenadas cartesianas, onde cada eixo deste sistema representa uma das cores primárias. O subespaço gerado por esse sistema é um cubo (Figura 1), onde as cores primárias se encontram nos três vértices mais distantes e que estão sobre os eixos, o preto está na origem e o branco está no vértice mais distante da origem (GONZALEZ & WOODS, 2000). Ou seja, considerando uma imagem de 24 bits (8 bits para cada canal), o ponto do espaço  $[0,0,0]$  representaria preto, enquanto o ponto  $[255,255,255]$  representaria a cor branca.

Figura 1 – Representação do espaço de cores RGB.

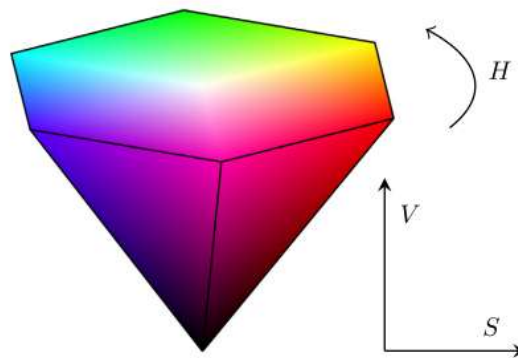


Fonte: Chernov et al. (2005).

### 3.3.2 HSV

Por vezes também é referido como HSB (*hue, saturation e brightness*), o espaço de cores HSV, baseado em matiz (*Hue*), saturação (*Saturation*) e valor (*Value*), surge para facilitar a interpretação humana das cores. Os seres humanos não descrevem as cores como uma combinação ponderada das três cores primárias, como no sistema RGB, mas sim em termos de matiz (que descreve as cores puras), saturação (o quanto uma cor está diluída por luz branca) e por fim o brilho (que corresponde a intensidade). Isso faz com que esse sistema de cores seja muito útil nas ferramentas de processamento de imagens onde se utiliza segregação de cores, já que permite uma descrição de cores mais natural para os seres humanos (GONZALEZ & WOODS, 2000). O sistema de coordenadas cilíndricas é utilizado para descrever esse espaço de cores, como apresentado na Figura 2. O plano polar representa a matiz (*hue*) como sendo o ângulo e a saturação (*saturation*) como módulo, e o eixo vertical representa o valor (*value*) que corresponde a intensidade (brilho).

Figura 2 – Representação do espaço de cores HSV.



Fonte: Chernov et al. (2005).

### 3.4 Processamento de imagens

O processamento de imagens digitais trata-se da manipulação das imagens de forma a atingir algum objetivo. De acordo com Gonzalez e Woods (2000), as diversas técnicas de processamento de imagens podem ser classificadas em três tipos: processos de níveis baixo, médio e alto. Dentre os tipos de manipulação de nível baixo podemos citar a aplicação de filtros para diminuição de ruídos, alteração de contraste, brilho e etc. De maneira geral, são processos onde a entrada e saída são imagens. Já os processos de nível médio são aquelas

onde o resultado final continuam sendo imagens, porém descrevem um conjunto de atributos extraídos da imagem original como, por exemplo, detecção de bordas, segmentação de regiões por geometria ou por cores e etc. Essas técnicas são muitas vezes utilizadas na visão computacional, onde geralmente se deseja localizar objetos de interesse nas imagens. Por fim, Gonzalez e Woods (2000) associam aos processos de nível alto à realização da tomada de decisão em cima de uma análise dos resultados obtidos pelas técnicas de baixo e médio nível.

### 3.4.1 Segmentação por cores

Técnica de extrema importância na visão computacional, a segmentação refere-se à separação da imagem em regiões, que geralmente representa a extensão de um objeto de interesse. Uma forma de realizar a segmentação é por meio de um filtro no espaço de cores. Nesse caso, é possível definir dois limiares, um inferior e outro superior, que define um range de cor que se deseja preservar na imagem. Essa técnica foi utilizada, a princípio, para imagens em tons de cinza (monocromática) e, posteriormente, em imagens coloridas (SIMÕES, 2000). É mais conveniente a realização da segmentação por cores em imagens representadas no espaço HSV, já que as cores estão representadas em termo de matiz (GONZALEZ & WOODS, 2000). Por exemplo, se for do interesse a segmentação da cor verde em uma imagem, como exemplificado na Figura 3, basta definir os limiares inferior e superior que abrangem esta cor no canal de matiz (*Hue*).

Figura 3 – Exemplo de segmentação da cor verde em uma imagem.



Fonte: Do autor (2022).

O problema dessa técnica é que ela não considera a informação espacial dos pixels, apenas o espaço de cores, e nem sempre existe um limiar ótimo (SIMÕES, 2000). Um bom resultado

de segmentação pode ser alcançado por meio de binarização e sucessivos procedimentos de transformação morfológica.

### 3.4.2 Binarização

A binarização consiste em transformar os valores dos pixels de uma imagem em 0, caso o pixel não satisfaça uma determinada regra, e em 1 (ou em 255), caso satisfaça. A forma mais simples de regra é por imposição de limiares (SIMÕES, 2000). A imagem resultante é chamada de máscara, e pode ser manipulada posteriormente para melhorar uma segmentação por meio de operações de transformação morfológica. A Figura 4 exemplifica uma operação de binarização.

Figura 4 – Exemplo de uma binarização.



Fonte: Do autor (2022).

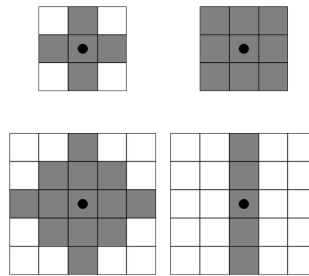
Como se trata de uma imagem com valores binários, é possível fazer operações booleanas entre diferentes máscaras, aumentando a complexidade das regras.

### 3.4.3 Transformações morfológicas

Como o próprio nome já sugere, as transformações morfológicas são operações que modificam as estruturas e as formas das regiões de uma imagem binarizada. Isso é possível por meio da translação de elementos estruturantes (ES), que são elementos matriciais retangulares, sobre a imagem em que se deseja modificar (GONZALEZ & WOODS, 2000). Esses elementos estruturantes podem ter tamanhos e formas diversas. Na Figura 5 são

apresentados alguns exemplos de ESs, onde os elementos sombreados, que digitalmente representam nível lógico alto, são elementos que compõem os ESs e os pontos pretos indicam as origens (centros) dos mesmos.

Figura 5 – Exemplos de Ess, onde os pontos pretos indicam as origens dos elementos.



Fonte: Do autor (2022).

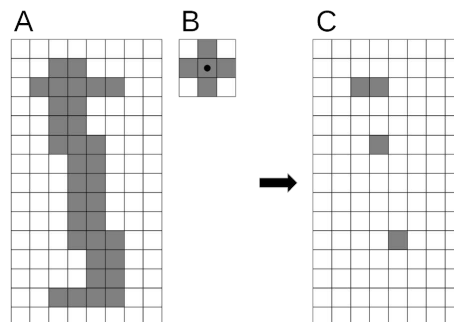
A varredura é feita em toda a imagem binarizada, semelhante a um processo de convolução bidimensional, de forma que a origem do ES passe por todos os pixels da imagem binarizada. O resultado final é uma imagem binarizada de mesma dimensão que a imagem de entrada, porém com as transformações aplicadas.

Os dois tipos de transformações morfológicas mais comuns são a erosão e a dilatação.

### 3.4.3.1 Erosão

Suponha que a imagem binária A sofrerá um processo de transformação morfológica por um elemento estruturante B, de tamanho 3x3 com formato de cruz, gerando a imagem de saída C, como apresentado na Figura 6. Para que isso ocorra, é necessário que a origem de B percorra todos os elementos de A e para cada localização da origem de B será gerado um elemento no conjunto C.

Figura 6 – O processo de erosão.



Fonte: Do autor (2022).

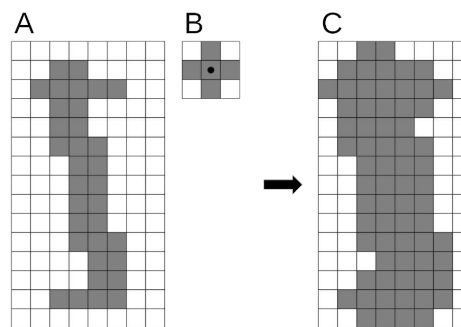


No caso da erosão, o elemento de C será sombreado caso os elementos sombreados de B estejam totalmente contidos nos elementos sombreados de A para aquela localização. O resultado final é uma imagem onde a região sombreada é menor do que a da imagem de entrada.

### 3.4.3.2 Dilatação

Suponha novamente que a imagem binária A sofrerá um processo de transformação morfológica por um elemento estruturante B, de tamanho 3x3 com formato de cruz, gerando a imagem de saída C, como apresentado na Figura 7. O processo de varredura segue a mesma lógica do processo de erosão, porém, nesse caso, a regra que define se um elemento de C será sombreado ou não, muda.

Figura 7 – O processo de dilatação.



Fonte: Do autor (2022).

O processo de dilatação ocorre quando a origem de B está sobre um elemento sombreado de A. Para as localizações onde isso ocorre é feita a projeção de todos os elementos sombreados de B em C. O resultado final é uma imagem onde a região sombreada é maior do que a da imagem de entrada.

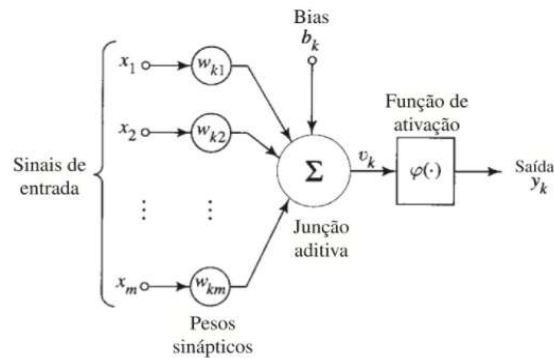
## 3.5 Redes neurais artificiais

A rede neural artificial (RNA) é uma das técnicas mais utilizadas atualmente dentro da inteligência artificial. As RNAs possuem como objetivo o processamento de informações de maneira semelhante à forma como é feito pelo cérebro biológico. Segundo Haykin (2007), o cérebro humano é um sistema de processamento altamente complexo, não-linear e paralelo,

capaz de realizar tarefas sofisticadas de maneira muito mais rápida que qualquer computador digital existente. Além disso, o cérebro humano é um sistema autônomo e independente, incorporando aprendizado, desenvolvimento e evolução através de seus sentidos, pensamentos e experiências (VALDATI, 2020).

Assim como no cérebro humano, a unidade básica de processamento das redes neurais artificiais é o neurônio. Os neurônios artificiais são nós, que dispostos em uma ou mais camadas são interligados por conexões. Na Figura 8 está representado o modelo de um neurônio artificial.

Figura 8 – Modelo de um neurônio artificial.



Fonte: Haykin (2007).

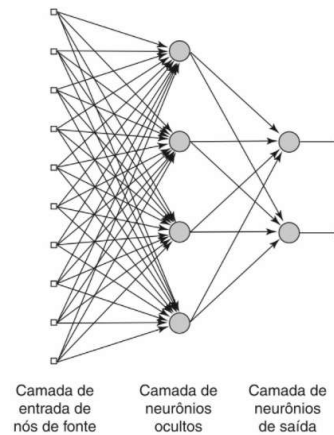
Os elementos que compõem um neurônio são:

- Pesos sinápticos ( $w_{km}$ ) – São valores que ponderam os sinais de entrada, dando maior peso para os sinais mais relevantes. Estes valores são aprendidos durante a aprendizagem;
- Junção aditiva ( $\Sigma$ ) – Trata-se de um somador que soma as entradas ponderadas, fazendo uma combinação linear;
- Função de ativação ( $\varphi(\cdot)$ ) – Tem como objetivo condicionar o valor resultante do somatório a uma amplitude específica;
- Bias ( $b_k$ ) – Tem como objetivo somar ou subtrair o valor do somatório.

A partir deste modelo base de um neurônio artificial é possível fazer as redes, que são um conjunto de neurônios conectados entre si e que podem ter arquiteturas diversas. Essa arquitetura está diretamente relacionada com o algoritmo de aprendizagem utilizado para treinar a rede (HAYKIN, 2007). A arquitetura mais comumente utilizada é a rede alimentada

adiante com múltiplas camadas. A Figura 9 mostra um exemplo de rede alimentada adiante com múltiplas camadas.

Figura 9 – Exemplo de rede alimentada adiante com múltiplas camadas.



Fonte: Haykin (2007).

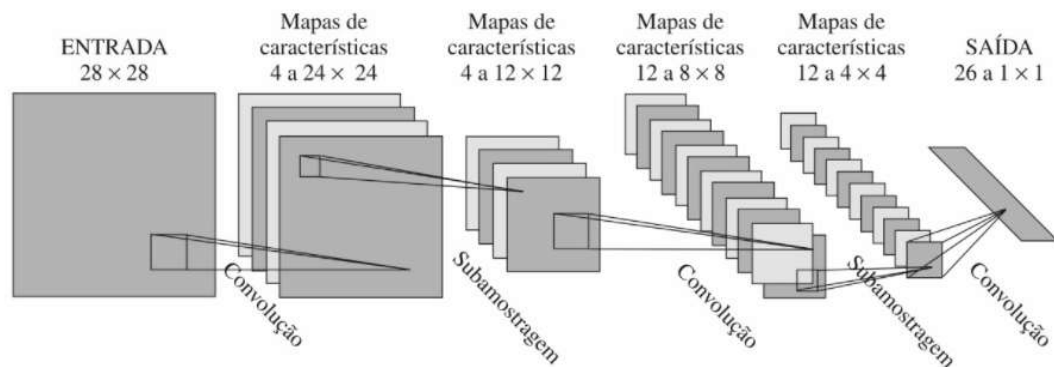
Trata-se de uma rede onde a informação de entrada trafega para frente (adiante) entre vários neurônios dispostos em camadas. Os sinais de saída da primeira camada servem como sinais de entrada da camada seguinte, e assim por diante. A força de uma conexão entre neurônios depende de seu peso sináptico, que armazena o conhecimento adquirido durante sua aprendizagem (HAYKIN, 2007). A função das camadas ocultas é aprender a extrair características relevantes dos sinais de entrada, de maneira que essas características sejam utilizadas para influenciar o sinal de saída da rede. A rede pode ser chamada de totalmente conectada ou densa caso cada neurônio de uma camada esteja conectado com todos os neurônios da camada seguinte.

### 3.5.1 Redes neurais convolucionais

As redes alimentadas adiante com múltiplas camadas são boas para problemas onde a informação de entrada são unidimensionais, porém quando se trata de informações bidimensionais, como é o caso de imagens, essas redes não possuem um bom desempenho. Muito disso se deve ao fato de que em imagens a informação espacial é muito importante. Neste contexto, surgem as redes neurais convolucionais (também chamadas convolutivas), que são projetadas para reconhecer informações bidimensionais e invariantes em relação a translação, escala e demais formas de distorção (HAYKIN, 2007). A Figura 10 mostra um

exemplo de rede neural convolucional.

Figura 10 – Exemplo de rede neural convolucional.



Fonte: Simon Haykin (2007).

Assim como nas redes alimentadas adiante com múltiplas camadas, as redes convolucionais também possuem neurônios. Porém neste caso os neurônios estão dispostos em forma de filtros (também chamados *kernels*), que são pequenos conjuntos matriciais que deslizam sobre a informação de entrada (imagem, que também é uma matriz) em um processo conhecido como convolução. Durante o treinamento a rede aprende esses filtros, que fazem o papel dos pesos sinápticos. Porém, devido ao processo de convolução, os pesos sinápticos são compartilhados para todo o sinal de entrada, reduzindo drasticamente a quantidade de parâmetros que devem ser aprendidos pela rede, melhorando a velocidade de treinamento e a capacidade de generalização. O resultado da convolução são os mapas de características, que são matrizes que possuem informações relevantes sobre a informação de entrada. Por vez, os mapas de características também sofrerão o processo de convolução, que resultará em novos mapas de características e que podem ou não ser subamostrados para diminuir as suas dimensões espaciais. À medida que ocorre a diminuição espacial ao longo das camadas, é aumentada a quantidade de filtros. Esse efeito reduz a variabilidade da saída da rede em relação a deslocamentos e outros tipos de distorções (HAYKIN, 2007). Para problemas de classificação de imagens, geralmente a última camada de uma rede neural convolucional são redes totalmente conectadas (densa), mas também podem ser totalmente convolucional.

## **4 METODOLOGIA**

### **4.1 Aquisição das imagens**

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

## 4.2 Geração do dataset

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

### 4.3 Arquitetura da rede neural

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

#### **4.4 Treinamento da rede neural**

##### **4.4.1 Pré-processamento dos dados**

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**



#### 4.4.2 Função de perda (Loss)

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

#### 4.4.3 Métrica

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

#### 4.4.4 Parâmetros de treinamento

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

#### 4.4.5 Hardware utilizado

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

#### 4.5 Avaliação da qualidade e tempo de inferência

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

## **5 RESULTADOS E DISCUSSÃO**

### **5.1 Treinamento**

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

## 5.2 Inferência

### 5.2.1 Qualidade da segmentação

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

### 5.2.2 Tempo de inferência

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**



**6 CONCLUSÃO**

**TRECHO OMITIDO:  
CONTÉM INFORMAÇÕES DE NATUREZA  
CONFIDENCIAL**

## REFERÊNCIAS

BADRINARAYANAN, Vijay; KENDALL, Alex; CIPOLLA, Roberto. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, v. 39, n. 12, p. 2481-2495, 2017.

BERNARDI, AC de C. et al. Agricultura de precisão: resultados de um novo olhar. Embrapa Instrumentação-Livro técnico (INFOTECA-E), 2014.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Portaria nº 852 - Art. 1º Criar a Comissão Brasileira de Agricultura de Precisão – CBAP. *Diário Oficial da República Federativa do Brasil*, Brasília, DF, 21 set. 2012. Seção 1, n. 184.

CHERNOV, Vladimir; ALANDER, Jarmo; BOCHKO, Vladimir. Integer-based accurate conversion between RGB and HSV color spaces. *Computers & Electrical Engineering*, v. 46, p. 328-337, 2015.

CONAB – Companhia Nacional de Abastecimento. Acompanhamento da safra brasileira de grãos, v.8 – Safra 2020, n. 3 - Terceiro levantamento, p. 1-86, 2020

DICE, Lee R. Measures of the amount of ecologic association between species. *Ecology*, v. 26, n. 3, p. 297-302, 1945.

DOS SANTOS, Gabriel Henrique Ribeiro. Desenvolvimento de software para contagem de plantas de cafeeiros de imagens obtidas por aeronaves remotamente pilotadas. Lavras-MG: UFLA, 2021.

EMBRAPA. Secretaria Geral. Gerência de Comunicação e Informação. Embrapa em números. Brasília, DF: Embrapa, 2019. 140 p.

FAO. “Climate-smart” agriculture: policies, practices and financing for food security, adaptation and mitigation. Rome: FAO, 2010. The document was prepared as a technical input for the Hague Conference on Agriculture, Food Security and Climate Change, to be held 31 October to 5 November 2010. Disponível em: [http://www.fao.org/fileadmin/user\\_upload/newsroom/docs/the-hague-conference-fao-paper.pdf](http://www.fao.org/fileadmin/user_upload/newsroom/docs/the-hague-conference-fao-paper.pdf).

FAVARIN, José Laércio et al. Equações para a estimativa do índice de área foliar do cafeeiro. *Pesquisa agropecuária brasileira*, v. 37, n. 6, p. 769-773, 2002.

GARCIA-GARCIA, Alberto et al. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857*, 2017.

GODOY, Hélio. Realismo Documentário, Teoria da Amostragem e Semiótica Peirceana: Os signos audiovisuais eletrônicos (analógicos ou digitais) como índices da realidade. *Estudos Socine de Cinema–Ano III*. Porto Alegre: Sulina, 2003.

GONZALEZ, Rafael C.; WOODS, Richard E. Processamento de imagens digitais. Editora Blucher, 2000.

HAYKIN, Simon. Redes Neurais. Porto Alegre: Bookman, 2007. 9788577800865. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788577800865/>. Acesso em: 24 mar. 2022.

HESPANHOL, Antonio Nivaldo; DE MEDEIROS HESPANHOL, Rosangela Aparecida. Os efeitos da elevação dos preços das COMMODITIES agrícolas sobre a segurança alimentar. Revista Faz Ciência, v. 12, n. 15, p. 73-94, 2010.

JACOB, Benoit et al. Quantization and training of neural networks for efficient integer-arithmetic-only inference. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. p. 2704-2713.

JENSEN, J. R. Introductory Digital Image Processing: A Remote Sensing Perspective. Upper Saddle River, Prentice Hall. Inc. 525p, 2005.

JUNIOR, Aziz Rachid et al. A. et al. Variabilidade espacial e temporal de atributos químicos do solo e da produtividade da soja num sistema de agricultura de precisão. Engenharia na Agricultura, Viçosa, v. 14, n. 3, p. 156-169, 2006.

KHANAL, S., FULTON, J., SHEARER, S. An overview of current and potential applications of thermal remote sensing in precision agriculture. Computers and Electronics in Agriculture, 139, 22-32. 2017.

LEAO, Alexandre Cruz. Gerenciamento de cores para imagens digitais. 2005.

LIMA, Daniel Caio de. Métodos para estimativa de imagens NIR a partir de imagens de câmeras RGB. 2020.

LONG, Jonathan; SHELHAMER, Evan; DARRELL, Trevor. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 3431-3440.

MANZATTO, C.V.; BHERING, S.B.; SIMÕES, M. (1999) Agricultura de precisão: propostas e ações da Embrapa solos. EMBRAPA Solos. Disponível na Internet. <http://www.cnps.embrapa.br/search/pesqs/proj01/proj01.html> em 01 Out

MASSRUHÁ, Silvia Maria Fonseca Silveira et al. A transformação digital no campo rumo à agricultura sustentável e inteligente. Embrapa Agricultura Digital-Capítulo em livro científico (ALICE), 2020.

MONROE, Will. Logistic regression. Recall, v. 1, n. 1, 2017.

PIRES, João Leonardo Fernandes et al. Discutindo agricultura de precisão-aspectos gerais. Embrapa Trigo-Documents (INFOTECA-E), 2004.

RASCHKA, Sebastian. Python machine learning. Packt publishing ltd, 2015.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015. p. 234-241.

SARLE, Warren S. et al. Stopped training and other remedies for overfitting. Computing science and statistics, p. 352-360, 1996.

SIMÕES, Alexandre da Silva. Segmentação de imagens por classificação de cores: uma abordagem neural. 2000. Tese de Doutorado. Universidade de São Paulo.

SØRENSEN, Thorvald Julius. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. I kommission hos E. Munksgaard, 1948.

SRIVASTAVA, Nitish et al. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, v. 15, n. 1, p. 1929-1958, 2014.

SUDRE, Carole H. et al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer, Cham, 2017. p. 240-248.

TAGHANAKI, Saeid Asgari et al. Combo loss: Handling input and output imbalance in multi-organ segmentation. Computerized Medical Imaging and Graphics, v. 75, p. 24-33, 2019.

VALDATI, A. B. Inteligência artificial - IA. Curitiba: Contentus, 2020. 978-65-5935-106-0. Disponível em: <https://plataforma.bvirtual.com.br/Leitor/Publicacao/191624/pdf>. Acesso em: 24 mar. 2022.

VARGAS, Ana Caroline Gomes; PAES, Aline; VASCONCELOS, Cristina Nader. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: Proceedings of the xxix conference on graphics, patterns and images. sn, 2016.

YING, Xue. An overview of overfitting and its solutions. In: Journal of Physics: Conference Series. IOP Publishing, 2019. p. 022022.

WANG, Panqu et al. Understanding convolution for semantic segmentation. In: 2018 IEEE winter conference on applications of computer vision (WACV). Ieee, 2018. p. 1451-1460.

ZHU, Michael; GUPTA, Suyog. To prune, or not to prune: exploring the efficacy of pruning for model compression. arXiv preprint arXiv:1710.01878, 2017.