

PROGRAMA DE PÓS-GRADUAÇÃO EM PRODUÇÃO VEGETAL

Rayane Aguiar Alves

**Deep learning na análise dos testes de tetrazólio e de pureza física de sementes de *Urochloa
brizantha* cv. Marandu**

Montes Claros

2021

Rayane Aguiar Alves

**Deep learning na análise dos testes de tetrazólio e de pureza física de sementes de *Urochloa
brizantha* cv. Marandu**

Dissertação apresentada ao Programa de Pós-Graduação em Produção Vegetal da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Produção Vegetal.

Orientador: Alcinei Místico Azevedo

Montes Claros
Fevereiro de 2021

Alves, Rayane Aguiar.

A474d
2021 Deep learning na análise dos testes de tetrazólio e de pureza física de sementes de *Urochloa brizantha* cv. Marandu/ Rayane Aguiar Alves. Montes Claros, 2021.
57 f.: il.

Dissertação (mestrado) - Área de concentração em Produção Vegetal. Universidade Federal de Minas Gerais / Instituto de Ciências Agrárias.

Orientador: Alcinei Místico Azevedo.

Banca examinadora: Bruno Vinícius Castro Guimarães, Carlos Alberto Araújo Júnior, Delacyr da Silva Brandão Junior, Alcinei Místico Azevedo.

Inclui referências: f.25-33; 43-45; 55-56.

1. Sementes – Qualidade. 2. Capim-marandu. 3. Deep learning. 4. Redes neurais convolucionais. I. Azevedo, Alcinei Místico Azevedo. II. Universidade Federal de Minas Gerais. Instituto de Ciências Agrárias. III. Título.

CDU: 631.53



Universidade Federal de Minas Gerais
Instituto de Ciências Agrárias
Mestrado em Produção Vegetal

ATA DE DEFESA DE DISSERTAÇÃO

Aos 24 dias do mês de fevereiro de 2021, às 08:00 horas, sob a Presidência do Professor Alcinei Místico Azevedo, D. Sc. (Orientador – UFMG) e com a participação dos Professores Carlos Alberto Araújo Júnior, D. Sc. (UFMG), Bruno Vinícius Castro Guimarães, D. Sc. (Instituto Federal de Educação, Ciência e Tecnologia do Amazonas) e Delacyr da Silva Brandão Júnior, D. Sc. (UFMG), reuniu-se, por videoconferência, a Banca de defesa de dissertação de **RAYANE AGUIAR ALVES**, aluna do Curso de Mestrado em Produção Vegetal. O resultado da defesa de dissertação intitulada: “**Deep learning na análise dos testes de tetrazólio e de pureza física em sementes de *Urochloa brizantha* cv. Marandu**” sendo a aluna considerada **Aprovada**. E, para constar, eu, Professor Alcinei Místico Azevedo, Presidente da Banca, lavrei a presente ata que depois de lida e aprovada, será assinada por mim e pelos demais membros da Banca examinadora.

OBS.: A aluna somente receberá o título após cumprir as exigências do ARTIGO 68 do regulamento do Curso de Mestrado em Produção Vegetal, conforme apresentado a seguir:

Art. 68 Para dar andamento ao processo de efetivação do grau obtido, o candidato deverá, após a aprovação de sua Dissertação ou Tese e da realização das modificações propostas pela banca examinadora, se houver, encaminhar à secretaria do Colegiado do Programa, com a anuência do orientador, no mínimo 3 (três) exemplares impressos e 1 (um) exemplar eletrônico da dissertação ou, 4 (quatro) exemplares impressos e 1 (um) exemplar eletrônico da tese, no prazo de 60 (sessenta) dias.

Montes Claros, 24 de fevereiro de 2021.

Alcinei Místico Azevedo
Orientador

Carlos Alberto Araújo Júnior
Membro

Bruno Vinícius Castro Guimarães
Membro

Delacyr da Silva Brandão Júnior
Membro

A Deus, que iluminou o meu caminho, concedendo-me forças durante esta caminhada; aos meus pais e a toda minha família que, com muito apoio, não mediram esforços para que eu chegasse até esta etapa da minha vida; e a todos que de alguma forma me ajudaram a concluir este trabalho.

AGRADECIMENTOS

A Deus por essa oportunidade, pela vida, proteção, força e saúde para completar mais esta etapa.

Aos meus pais, pelo exemplo de luta, amor, confiança, dedicação e pelo apoio em todos os momentos.

Ao meu orientador, Prof. Dr. Alcinei, por toda atenção, amizade, apoio e ensinamentos de grande valia.

Aos colegas do Grupo de Estudo em Experimentação Agrícola (GEEA), em especial à Ana Clara e Nermy; pela amizade, paciência, companheirismo e auxílio na condução das pesquisas.

Ao grupo de estudo Sementec, em especial, à Paloma e à Josiane por toda ajuda.

Aos professores Bruno, Carlos, Delacyr pela colaboração e participação da banca.

A todos os funcionários do Instituto de Ciências Agrárias que com muito profissionalismo e ética contribuíram a esta jornada.

À Universidade Federal de Minas Gerais, em especial, ao Programa de Pós-Graduação em Produção Vegetal, pela oportunidade de realização do mestrado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- Brasil (CAPES) – Código de Financiamento 001, ao Conselho Nacional de Desenvolvimento Científico (CNPQ) e à Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) pela concessão da bolsa de estudos e auxílio nos experimentos.

Muito obrigada!

“O êxito da vida não se mede pelo caminho que você conquistou, mas sim pelas dificuldades que superou no caminho”. (Abraham Lincoln)

Deep learning na análise dos testes de tetrazólio e de pureza física de sementes de *Urochloa brizantha* cv. Marandu

RESUMO

A análise computacional de imagens associada à Mask R-CNN (Mask Region-based Convolutional Neural Network) pode ser utilizada para automação e aprimoramento de diversas avaliações rotineiras, como o teste de tetrazólio e análise de pureza de sementes. O teste de tetrazólio é considerado um dos testes mais rápidos na avaliação de viabilidade de sementes, sendo altamente dependente da experiência e capacidade do avaliador para classificação dos resultados, sendo assim um teste subjetivo. A análise de pureza física é uma determinação obrigatória para comercialização de sementes e indica a qualidade física de lotes. Assim, objetivou-se verificar a eficiência da automação computacional na análise dos testes de tetrazólio e de pureza física de sementes de *Urochloa brizantha* cv. Marandu por meio do Mask R-CNN. O experimento foi conduzido em Montes Claros. Foram utilizados lotes de sementes de *Brachiaria*, provenientes de locais comerciais. Para obtenção das imagens, nos dois testes, foi utilizado um escâner de mesa (Hp Officejet 4500 Desktop). O teste de tetrazólio foi realizado de acordo com as Regras de Análises de Sementes (solução de 0,1%, por 5 horas, a 40 °C)(BRASIL, 2009). Cada um dos dois fragmentos das sementes foi colocado sobre a placa de vidro do escâner com a parte corada voltada para baixo. Para automação da classificação dos resultados do teste de tetrazólio, foram utilizadas 4 repetições por lote e cada repetição possuiu 50 sementes. Já para a automação da análise de pureza foram escaneadas 5 proporções de pureza, sendo que cada proporção possuía duas amostras de 10 gramas cada uma. As redes neurais apresentaram alta eficiência ao classificar os resultados do teste de tetrazólio, apresentando valores de acurácia superiores a 80%, e acurácia geral em torno de 91%. Também obteve eficiência na estimação da pureza física dos lotes avaliados, com altos coeficientes de determinação. Logo, a automação da análise dos testes de tetrazólio e análise de pureza em sementes de *Urochloa brizantha* cv. Marandu é viável por meio da técnica Mask R-CNN.

PALAVRAS-CHAVE: Visão computacional. Qualidade de sementes. Fenotipagem. Inteligência artificial. Análise de imagens. Mask R-CNN. Classificação.

Deep learning in the analysis of tetrazolium and physical purity tests of *Urochloa brizantha* cv.

Marandu

ABSTRACT

The computational image analysis associated with the Mask R-CNN (Mask Region-based Convolutional Neural Network) can be used to automate and improve several routine evaluations, such as the tetrazolium test and seed purity analysis. The tetrazolium test is considered one of the fastest tests in the evaluation of seed viability, being highly dependent on the experience and ability of the evaluator to classify the results, thus being a subjective test. The analysis of physical purity is a mandatory determination for commercialization of seeds and indicates the physical quality of lots. Thus, the objective was to verify the efficiency of computational automation in the analysis of the tetrazolium and physical purity tests of *Urochloa brizantha* cv. Marandu through the Mask R-CNN. The experiment was conducted in Montes Claros. Batches of *Brachiaria* seeds from commercial locations were used. To obtain the images, in both tests, a flatbed scanner (Hp Officejet 4500 Desktop) was used. The tetrazolium test was performed according to the Seed Analysis Rules (0.1% solution, for 5 hours, at 40 °C) (BRASIL, 2009). Each of the two fragments of the seeds was placed on the scanner glass plate with the colored part facing downwards. To automate the classification of the results of the tetrazolium test, 4 repetitions per lot were used and each repetition had 50 seeds. As for the automation of the purity analysis, 5 purity proportions were scanned, each proportion having two samples of 10 grams each. The neural networks showed high efficiency when classifying the results of the tetrazolium test, presenting accuracy values above 80%, and general accuracy around 91%. It also obtained efficiency in the estimation of the physical purity of the evaluated lots, with high coefficients of determination. Therefore, the automation of the analysis of tetrazolium tests and purity analysis in seeds of *Urochloa brizantha* cv. Marandu is viable using the Mask R-CNN technique.

KEYWORDS: Computer vision. Seed quality. Phenotyping. Phenomics. Image analysis. Mask R-CNN. Classification.

LISTA DE ILUSTRAÇÕES

REFERENCIAL TEÓRICO

| | |
|--|----|
| Figura 1 - Rede Neural Artificial | 21 |
| Figura 2 - Exemplo de CNN para classificação de espécies de animais, com camadas de convolução, max pooling e uma camada totalmente conectada..... | 23 |
| Figura 3 - Exemplo de uma estrutura Mask R-CNN para segmentação de instâncias. | 24 |

Artigo I: *Deep learning* na interpretação do teste de tetrazólio em sementes de *Urochloa brizantha* cv. Marandu

| | |
|--|----|
| Figura 1- Escala para classificação visual dos resultados do teste de tetrazólio em sementes de <i>Brachiaria</i> spp..... | 36 |
| Figura 2- Curvas da função de perda por iteração na classificação de sementes de <i>Urochloa brizantha</i> cv. Marandu..... | 39 |
| Figura 3- Detecção pelo estabelecimento de caixas delimitadoras e máscaras em imagens utilizadas para treinamento e validação..... | 40 |

Artigo II: *Deep learning* da análise de pureza física de sementes de *Urochloa brizantha* cv. Marandu

| | |
|---|----|
| Figura1- Detecção pelo estabelecimento de máscaras em imagens utilizadas para associação da porcentagem de pixels ocupada pelas sementes e a porcentagem de pureza das sementes..... | 49 |
| Figura 2- Curvas da função de perda por iteração na classificação de sementes de <i>Urochloa brizantha</i> cv. Marandu..... | 50 |
| Figura 3- Predição da pureza em sementes de <i>Urochloa brizantha</i> cv. Marandu estimada por Mask R-CNN e regressão considerando uma amostra de 10 gramas de semente dos lotes 1 (A), 2 (B), 3 (C) e 4 (D)..... | 51 |
| Figura 4- Predição da pureza em sementes de <i>Urochloa brizantha</i> cv. Marandu estimada por Mask R-CNN e regressão considerando uma amostra de 50 gramas de semente dos lotes 1 (A), 2 (B), 3 (C) e 4 (D)..... | 52 |
| Figura 5- Predição da pureza em sementes de <i>Urochloa brizantha</i> cv. Marandu estimada por Mask R-CNN e regressão considerando uma amostra de 100 gramas de semente dos lotes 1 (A), 2 (B), 3 (C) e 4 (D). | 53 |
| Figura 6- Modelo ajustado para a predição da pureza a partir do PAOS em sementes de <i>Urochloa brizantha</i> cv. Marandu..... | 53 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 - Classificação visual do teste de tetrazólio em sementes de <i>Urochloa brizantha</i> cv. Marandu. | 37 |
| Tabela 2 - Precisão, recall e acurácia na classificação de sementes como viáveis ou inviáveis pela metodologia Mask R-CNN..... | 41 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------------|--|
| CNN | Convolutional Neural Network (Rede Neural Convolutiva) |
| DPI | Pontos por polegada quadrada |
| E.V.A | Etil, Vinil e Acetato (Placas emborrachadas) |
| FAST R-CNN | Fast Region-based Convolutional Neural Network |
| Faster R-CNN | Faster Region-based Convolutional Neural Network |
| FPN | Filtro Previsor Neural |
| I | Inviável |
| ISTA | Internacional Seed Testing Association |
| Mask R-CNN | Mask Region- based Convolutional Neural Network |
| PAOS | Porcentagem de área ocupada pela semente |
| ResNet | Residual Network |
| RGB | Red, Green, Blue (Sistema de cores) |
| RNA | Rede Neural Artificial |
| SSD | Single Shot Detector |
| SVIS | Seed Vigor Imaging System |
| UFMG | Universidade Federal de Minas Gerais |
| V | Viável |
| YOLO | You Only Look Once |

SUMÁRIO

| | | |
|--------|--|----|
| 1. | INTRODUÇÃO | 13 |
| 2. | OBJETIVOS | 15 |
| 2.1. | Objetivo Geral..... | 15 |
| 2.2. | Objetivos Específicos | 15 |
| 3. | REVISÃO DE LITERATURA | 16 |
| 3.1. | <i>Urochloa brizantha</i> cv. Marandu..... | 16 |
| 3.2. | Teste de Tetrazólio | 16 |
| 3.3. | Análise de Pureza de Sementes | 18 |
| 3.4. | Aplicação da Análise de Imagens na Avaliação da Qualidade de Sementes..... | 19 |
| 3.5. | Redes Neurais Artificiais | 20 |
| 3.6. | Mask R-CNN..... | 22 |
| 4. | REFERÊNCIAS | 25 |
| 5. | ARTIGOS | 34 |
| 5.1. | Artigo 1- <i>Deep learning</i> na interpretação do teste de tetrazólio em sementes de <i>Urochloa brizantha</i> cv. Marandu | 34 |
| 5.1.1. | Introdução..... | 35 |
| 5.1.2. | Material e métodos | 35 |
| 5.1.3. | Resultados..... | 38 |
| 5.1.4. | Discussão | 42 |
| 5.1.5. | Conclusão..... | 43 |
| 5.1.6. | Agradecimentos..... | 43 |
| 5.1.7. | Referências..... | 43 |
| 5.2. | Artigo 2- <i>Deep learning</i> da análise de pureza de sementes de <i>Urochloa brizantha</i> cv. Marandu | 46 |
| 5.2.1. | Introdução..... | 47 |
| 5.2.2. | Material e métodos | 47 |
| 5.2.3. | Resultados..... | 49 |
| 5.2.4. | Discussão | 54 |
| 5.2.5. | Conclusão..... | 54 |
| 5.2.6. | Referências..... | 55 |
| 6. | CONCLUSÕES GERAIS | 57 |

1. INTRODUÇÃO

É desejável pelos agricultores ter-se lavouras uniformes e altamente produtivas. A qualidade das sementes é um dos elementos indispensáveis para que isso aconteça. A utilização de sementes de alta qualidade é essencial para assegurar o estabelecimento adequado das plantas no campo, já que o uso de sementes de alta qualidade favorece a uniformidade de germinação e a emergência de plântulas sob ampla variação de ambiente (MARCOS-FILHO, 2015). Com isso, é importante que se determine a qualidade física e fisiológica das sementes por meio de testes. Dentre os principais testes, se destacam o teste de germinação que pode ser abreviado pelo teste de tetrazólio, e a análise de pureza física.

O teste de tetrazólio baseia-se na alteração da coloração dos tecidos vivos em presença de uma solução de tetrazólio, sendo altamente dependente da experiência do avaliador para a classificação dos resultados, o que faz com que seja um teste subjetivo (MARCOS-FILHO, 2015). Tem por objetivo determinar rapidamente a viabilidade e vigor das sementes e é muito empregado, devido a sua rapidez, para a avaliação da qualidade fisiológica de sementes de várias espécies, principalmente daquelas que necessitam de longos períodos para germinar, como é o caso de algumas espécies de frutíferas, florestais e forrageiras (GUEDES *et al.*, 2010).

A análise de pureza física é um teste obrigatório para determinar a pureza de sementes comerciais, estabelecido pela Lei de Sementes, Lei 10.711, agosto de 2003 (BRASIL, 2003). Nela são separadas e identificadas sementes puras, sementes de outras espécies e materiais inertes das amostras, sendo importante para definir a qualidade física das sementes que serão utilizadas e tornar-se possível a previsão de perdas no campo. Já que o uso de lotes com baixo percentual de pureza, quando semeados, resultam na formação de uma lavoura não homogênea gerando prejuízos na produtividade final (BATISTA *et al.*, 2016).

Estas análises são mais complexas quando realizadas em sementes pequenas como as da *Urochloa brizantha* cv. Marandu, que é uma forrageira importante para formação de pastagens no Brasil, sendo considerada a mais representativa entre os campos de produção de sementes no país (LANDAU *et al.*, 2020). As sementes dessa espécie apresentam grande importância para o mercado de sementes forrageiras no país, já que correspondem a 32% da área total de produção de sementes comerciais (LANDAU *et al.*, 2020).

Para facilitar as avaliações de sementes de pequeno porte, e eliminar os problemas decorrentes da subjetividade, uma alternativa interessante é uso de análise por imagens. Como foi uma opção utilizada por Castilho *et al.* (2012), na qual os autores identificaram dificuldade em analisar os resultados do teste de tetrazólio em sementes de capim-marandu e como solução optaram por imagens digitalizadas. E para automatizar a interpretação destas imagens pode-se recorrer as redes neurais convolucionais (CNN's).

Redes neurais convolucionais são arquiteturas biologicamente inspiradas capazes de serem treinadas e aprenderem representações invariantes a escala, translação, rotação e transformações afins (LECUN; KAVUKCUOGLU; FARABET, 2010). Existem vários algoritmos envolvendo redes neurais convolucionais como Fast R-CNN, Faster R-CNN, YOLO e SSD, as quais calculam a posição dos alvos utilizando uma caixa delimitadora. Outro algoritmo existente é o Mask R-CNN (Mask Region Convolutional

Neural Network), proposto por He *et al.* (2018). Esta metodologia tem a vantagem de integrar as caixas delimitadoras das classes encontradas na imagem com as máscaras de segmentação (HE *et al.*,2017).

Desta forma, pretende-se tornar possível a automação da análise dos testes de tetrazólio e de pureza em sementes de *Urochloa brizantha* cv. Marandu por meio da análise de imagens associada às Mask R-CNN.

2. OBJETIVOS

2.1. Objetivo Geral

- Avaliar a eficiência da automação computacional na análise dos testes de tetrazólio e de pureza de sementes de *Urochloa brizantha* cv. Marandu por meio da Mask R-CNN.

2.2. Objetivos Específicos

- Verificar a eficiência da classificação de sementes quanto à viabilidade a partir do teste de tetrazólio, utilizando a análise de imagens associada a Mask R-CNN.

- Atestar a viabilidade da classificação de sementes quanto à pureza por meio das técnicas de imagens e Mask R-CNN.

3. REVISÃO DELITERATURA

3.1. *Urochloa brizantha* cv. Marandu

Urochloa brizantha cv. Marandu é uma gramínea forrageira perene, pertencente a família Poaceae (Gramineae) e subfamília Panicoideae, com classificação botânica *Urochloa brizantha* (Hochst ex A.RICH.) STAPF. cv. Marandu Nunes *et al.* (1984), e conhecida popularmente como capim-marandu.

A partir da década de 1970, o gênero *Urochloa* (sinonímia *Brachiaria*) passou a ter grande importância para a formação de pastagens no Brasil, por apresentar grandes extensões territoriais de cultivo. Ocupa no país, uma área total de aproximadamente 60 milhões de hectares de pastagens cultivadas e, desses, a *Urochloa brizantha* cv. Marandu (sinonímia - *Brachiaria brizantha*) ocupa mais de 50 milhões de ha (MEDICA; REIS; SANTOS, 2017).

O capim-marandu é originário da África e foi introduzido no Brasil por volta de 1967, pelo produtor de sementes Paul Rankin Rayman, no Estado de São Paulo, de onde foi distribuído para várias regiões (RAYMAN, 1983). É uma planta cespitosa, com bainha foliar pilosa, pubescência apenas na parte inferior das folhas, muito robusta que pode atingir de 1,5 a 2,5 m de altura, com colmos iniciais prostrados e afilhos predominantemente eretos (NUNES *et al.*, 1984). As sementes dessa espécie são de tamanho pequeno; com coloração variante entre amarelo, bege e marrom; possuem dificuldade de germinação, o que pode ser justificado tanto pela maturação desigual e o processo de degana, quanto por uma característica inerente a algumas sementes, denominada dormência (MARTINS e SILVA, 2003).

A área nacional destinada à produção de sementes das principais cultivares de forrageiras tropicais foi de $237 \pm 12,9$ mil hectares, em média, considerando as safras de 2013/2014 a 2016/2017 (MAPA, 2018). Entre as cultivares forrageiras, a *Urochloa brizantha* cv. Marandu é a que ocupa a maior área de produção de sementes, com 32% do total, seguida pelas gramíneas *Panicum maximum* cv. Mombaça (15%), *B. ruziziensis* 'Comum' (13%), *B. humidicola* 'Comum' (12%) e da *B. brizantha* cv. Xaraés (9%). Do total de produção de sementes comerciais, as do gênero *Urochloa* correspondem atualmente a 75% da produção de sementes (LANDAU *et al.*, 2020).

Nas pastagens brasileiras, a espécie tem sido bastante cultivada, por possuir características de rusticidade como: alta produção de sementes, adaptação ao clima e solos do Cerrado, boa produção de forragem quando bem manejada, e grande flexibilidade quanto ao manejo do pastejo (LIMA *et al.*, 2015). Devido essas características, também tem sido bastante empregada para a recuperação de áreas degradadas ou novas áreas abertas a fim da iniciação do cultivo de forrageiras (MACEDO, 2005).

3.2. Teste de Tetrazólio

O teste de tetrazólio é conhecido desde a década de 30, quando Kuhn e Jerchel descobriram que os sais de tetrazólio eram reduzidos nos tecidos vivos, resultando em um composto de cor vermelha (trifenilformazan), despertando o interesse de vários cientistas que se dedicaram a estudar o teste (DELOUCHE *et al.*, 1976).

O teste foi inicialmente desenvolvido por Lakon na década de 40 para estimar a qualidade de

sementes de cereais, o qual descobriu que o melhor sal para a avaliação da viabilidade das sementes, era o 2, 3, 5 trifenil cloreto de tetrazólio. Desde então, vem sendo utilizado em sementes de diferentes espécies, sendo aprovado pela International Seed Testing Association (ISTA) (SAVONEN, 1999).

O teste de tetrazólio é importante na rápida avaliação da viabilidade de sementes (LAMARCA; BARBEDO, 2014). Essa metodologia se baseia na alteração da coloração dos tecidos vivos da semente em presença de uma solução de sal de tetrazólio, o qual é reduzido pelas enzimas desidrogenases dos tecidos vivos, resultando e composto denominado de formazam, de coloração vermelha-carmim. Tecidos mortos ou muito deteriorados apresentam-se descoloridos (MARCOS-FILHO, 2015).

O padrão de coloração dos tecidos pode ser utilizado para identificar sementes viáveis, não viáveis e, dentro da categoria viável, as de alto e baixo vigor (VIEIRA e FRANÇA NETO, 1999). A coloração vermelha-brilhante ou rosa-brilhante se desenvolve quando o hidrogênio resultante do processo de respiração das células vivas combina-se com a solução de tetrazólio absorvida. Os tecidos do embrião viável absorvem a solução de tetrazólio lentamente e tendem a desenvolver uma coloração mais leve que os embriões deteriorados que adquirem uma cor vermelha-carmim forte. A presença de tecidos não coloridos, firmes e sadios sugere maior resistência à penetração da solução de tetrazólio do que a morte do tecido. O tecido morto geralmente é caracterizado por cor branca ou amarelada e textura flácida (BHÉRING *et al.*, 1996 e FRANÇA NETO, 1999).

O teste de tetrazólio apresenta várias vantagens: não é afetado por diversas condições que podem alterar os resultados do teste padrão de germinação, como a dormência e presença de fungos; enfoca as condições físicas e fisiológicas do embrião de cada semente individualmente; permite rápida avaliação da viabilidade; permite em algumas espécies, como soja e feijão, a identificação de diferentes níveis de viabilidade; pode fornecer o diagnóstico da causa da perda da viabilidade das sementes e requer equipamento simples e barato (DELOUCHE *et al.*, 1976; FRANÇA NETO *et al.*, 1998; FRANÇA NETO, 1999).

Embora seja um teste relativamente rápido, ele consome maior número horas de trabalho que o teste de germinação padrão; é relativamente tedioso, uma vez que as sementes são avaliadas uma a uma, requerendo, desta forma, experiência e paciência; além de treinamento especial sobre a estrutura embrionária da semente e sobre técnicas de interpretação; exigindo do analista capacidade de decisão pelas características do teste.

Para o sucesso do teste é importante que o mesmo seja realizado conforme o indicado para cada espécie pelas Regras de Análises de Sementes. Pois, inúmeros fatores podem interferir na obtenção de resultados satisfatórios, principalmente aqueles relacionados aos procedimentos de execução, como preparo das sementes antes da coloração, concentração da solução, período e temperatura de exposição à solução e critérios de interpretação (GASPAR-OLIVEIRA *et al.*, 2011).

É comumente utilizado em trabalhos envolvendo metodologia de tetrazólio, a comparação dos resultados desse teste com o teste de germinação, a fim de se constatar a existência de correlação positiva entre ambos. Essa correlação positiva já foi verificada em diversas pesquisas com diferentes espécies como: *Brachiaria brizantha* (DIAS e ALVES, 2008); Abobrinha (BARROS *et al.*, 2005); sucará (*Gleditsia amorphoides*) (FOGAÇA *et al.*, 2006); amburana (*Amburana cearenses*) (GUEDES *et al.*, 2010).

O teste de tetrazólio no Brasil repercute em relação aos aspectos quantitativos e qualitativos, pois

quando executado junto a outros testes, tem mostrado a comercialização de lotes com bons padrões de qualidade. Isto resulta em um sistema de controle de alta confiabilidade, confirmando maiores lucros aos produtores de sementes e grãos, por intermédio de sementes de alta qualidade e custo baixo (FRANÇA *et al.*, 1998).

3.3. Análise de Pureza de Sementes

A análise de pureza é um teste obrigatório para determinar a composição percentual por peso e a identidade das diferentes espécies de sementes e do material inerte da amostra e por inferência a do lote de sementes comerciais, sendo estabelecido pela Lei de Sementes, Lei 10.711, agosto de 2003 (BRASIL, 2003). Essa análise tem por objetivo avaliar a qualidade física da amostra em exame, e conseqüentemente a do lote de sementes. Sendo que, o lote de sementes pode ser conceituado como uma quantidade definida e limitada de sementes, identificado por letra, número ou pela combinação dos dois, da qual cada porção é uniforme para as determinações contidas na identificação (BRASIL, 2009).

Nos campos de produção a maior parte das sementes de gramíneas forrageiras tropicais é colhida por varredura do solo (NERY *et al.*, 2012). Deste modo, os lotes de sementes recebidos pelas empresas contêm grande quantidade de impurezas, tais como torrões de terra, pedras, espiguetas vazias, cariopses parcialmente desenvolvidas, imaturas, sementes de plantas daninhas e palha, que prejudicam diretamente a qualidade física e sanitária das sementes (BRASIL, 2009; NERY *et al.*, 2009; MELO *et al.*, 2016).

Com isso, é fundamental que seja realizada a análise de pureza de um lote que será comercializado. Pois assim, pode ser definida a qualidade física das sementes que serão utilizadas e torna-se possível a previsão de perdas no campo. O uso de lotes com baixo percentual de pureza, sem a devida correção da taxa de semeadura, resultam na formação de uma lavoura não homogênea gerando prejuízos na produtividade final (BATISTA *et al.*, 2016).

Na análise de pureza são separadas e identificadas sementes puras, sementes de outras espécies e materiais inertes, sendo importante para determinar a composição percentual de um lote de sementes. As sementes puras indicam o percentual de sementes que foram encontradas na amostra e que pertencem a espécie declarada pelo requerente da análise. Além de serem consideradas sementes puras, fragmentos com mais de 50% da sua estrutura intacta, e que caracterizem-se como as sementes do lote em questão; o percentual de material inerte, consiste no percentual de outros materiais que não se caracterizam como semente, e que aparecem misturados a amostra, este tipo de material pode ser constituído por pedras, insetos, pedaços de sementes e outros fragmentos (BRASIL, 2009).

Os valores mínimos de pureza e de valor cultural (VC) para comercialização de sementes de *Brachiaria brizantha* estão fixados, respectivamente, em 40 e 15% pelas normas federais (Brasil, 2009) e em 40 e 24% pelas normas paulistas (SÃO PAULO, 1996). Apesar de ser uma análise obrigatória e indicar percentuais de pureza a serem seguidos para cada espécie, é comum a comercialização de lotes de sementes fora dos padrões estabelecidos. É possível encontrar na literatura trabalhos que relatam a comercialização de sementes com baixa qualidade física. Como por exemplo, o trabalho desenvolvido por Ohlson (2011) entre os anos de 2008 a 2010 com sementes de Azevém (*Lolium multiflorum*), no qual observaram que 50 a 100% das amostras estavam abaixo do padrão de Sementes Puras.

Já em relação ao mercado de sementes forrageiras o descumprimento dos padrões de pureza é bastante notável, pois as sementes de pastagens apresentam baixa qualidade física e fisiológica, quando comparadas às de grandes culturas (TOMAZ *et al.*, 2010). Ohlson (2011) ainda relata em seu trabalho que para um conjunto de amostras de Capim-Colonião, todas as amostras estavam abaixo dos padrões de Sementes Puras, e para a cultivar Mombaça os percentuais foram de 50 a 85% que não atingiram o padrão de Sementes Puras.

3.4. Aplicação da Análise de Imagens na Avaliação da Qualidade de Sementes

Atualmente, a análise de imagens de sementes tem-se mostrado bastante eficiente para a avaliação dos componentes fisiológicos e físicos de sementes, que influenciam na qualidade dos lotes (SILVA *et al.*, 2013). O uso da análise de imagens automatizada para avaliação de sementes e plântulas está associada a necessidade de métodos rápidos e precisos (DELL'AQUILA, 2009). O uso de imagens digitalizadas é uma alternativa que pode permitir a ampliação, o acesso rápido e o arquivamento da informação obtida para uso posterior (GENEVE e KESTER 2001; MCDONALD *et al.*, 2001; DELL'AQUILA, 2004; DELL'AQUILA, 2007; HOSOMI *et al.*, 2011).

Entre os métodos de análises de imagens destaca-se o SVIS, Seed Vigor Imaging System, desenvolvido na Universidade Estadual de Ohio nos Estados Unidos da América. Trata-se de um sistema computacional de análise de imagens de plântulas que foi utilizado pela primeira vez com a cultura da alface (SAKO *et al.*, 2001). Por meio do processamento das imagens digitalizadas, o SVIS calcula índices que se baseiam no crescimento e uniformidade inicial das plântulas (MCDONALD *et al.*, 2008). As imagens foram capturadas com o auxílio de um escâner tendo como operador um software específico (MARCOS-FILHO *et al.*, 2013).

De forma eficiente o sistema computacional SVIS foi empregado na avaliação do vigor de lotes de sementes de soja em vários estudos (OAKLEY *et al.*, 2004; YAGUSHI *et al.*, 2014; SILVA *et al.*, 2012; CHIQUITO *et al.*, 2012). Os resultados obtidos pelo SVIS podem se relacionar com os testes de vigor tradicionais para sementes como o envelhecimento acelerado, tetrazólio e condutividade elétrica (ALVARENGA *et al.*, 2012).

Outro método é a análise de imagens por raio-X. Esse método é bastante utilizado por ser rápido, não destrutivo e ser uma alternativa interessante para avaliar a qualidade de sementes de várias espécies. Várias pesquisas têm demonstrado que a análise de imagens de raio-X possibilita a identificação de danos que afetam a qualidade de sementes de espécies como: *Ginkgo biloba* (SALINAS *et al.*, 2012); *Tabebuia heptaphylla* (AMARAL *et al.*, 2011); *Xylopia aromatica* (SOCOLOWSKI *et al.*, 2011); *Tecoma stans* (SOCOLOWSKI *et al.*, 2008); *Eremanthus erythropappus* (FEITOSA *et al.*, 2009); *Cedrela fissilis* (MASSETO *et al.*, 2008); *Platypodium elegans* (SOUZA *et al.*, 2008); *Cecropia pachystachya* (PUPIM *et al.*, 2008).

Na bibliografia, encontram-se diversos trabalhos em que a análise de imagens foi utilizada para o aprimoramento de informações e/ou extração de informações de imagens de testes de qualidade de sementes. Como por exemplo o trabalho de Silva *et al.* (2014), que utilizaram a análise de imagens no estudo morfológico e fisiológico de sementes de abóbora, e chegaram à conclusão que a técnica

empregada é eficaz na localização de sementes com má formação, além de sementes vazias.

Outros exemplos são o uso da análise de imagens para avaliar a qualidade fisiológica de sementes de milho localizadas em diferentes posições na espiga (MONDO e CICERO, 2005). Uso da análise de imagem das sementes de soja para avaliar as características físicas das sementes, verificando que esse método é tão eficaz quanto os métodos de avaliação tradicional, tendo como vantagem de ser não destrutivo e mais rápido (GUEDES *et al.*, 2011).

A técnica pode ser utilizada para facilitar a análise de resultados de testes rotineiros como o teste de tetrazólio. Ao testar a eficiência da análise de imagens para interpretação dos resultados do teste de tetrazólio em sementes de *Brachiaria brizantha* cv. Marandu, Castilho *et al.* (2012) concluíram que a observação das imagens digitalizadas das sementes é equivalente à efetuada sob estéreo microscópio, ou seja, que a técnica é eficiente também quando empregada para facilitar a classificação dos resultados do teste de tetrazólio.

Ao avaliarem o uso de imagens digitais do teste de tetrazólio na avaliação da qualidade de grãos de café, Rosa *et al.* (2013) concluíram que o teste de tetrazólio nos embriões de café, assim como no endosperma, apresentam alta correlação com a análise sensorial da qualidade de bebida.

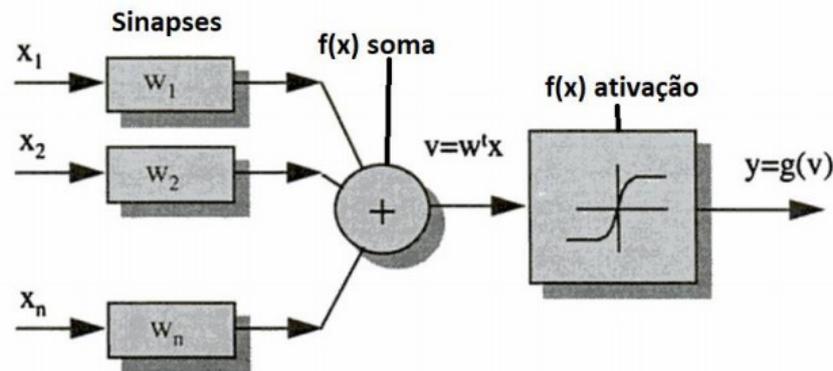
Estudando um sistema computadorizado de análise de imagens de plântulas para avaliação do vigor de sementes, Marcos-Filho (2010) destacou que sua utilização vem permitindo a avaliação eficaz do vigor de sementes de diversas espécies, mostrando melhoria significativa na determinação do desempenho de plântulas, uma vez que envolve procedimentos simples e reduz a interferência humana e a subjetividade na interpretação dos resultados.

De forma geral, a análise por imagem é uma técnica que tem se mostrado promissora, eficiente e com alto potencial para ser empregada em diversas finalidades na avaliação da qualidade de sementes. Possui ainda, grande potencial para ser utilizada em conjunto com outras técnicas, como as redes neurais artificiais, a fim de aumentar a precisão e rapidez de testes padrões na avaliação de sementes.

3.5. Redes Neurais Artificiais

As redes neurais artificiais (RNA's) são técnicas computacionais que fazem parte da inteligência computacional. Foram apresentadas pela primeira vez em 1943 pelo neurofisiologista McCulloch e pelo matemático Walter Pitts (CAMELO *et al.*, 2018). Podem ser definidas de diferentes formas na literatura segundo sua área de aplicação, mas fundamentalmente todos remetem ao sistema cognitivo biológico. Uma rede neural artificial possui sinapses, função de soma, função de ativação e saída (KOVÁCS, 2002). A Figura 1 demonstra como o neurônio é representado no ambiente computacional, deixando de maneira explícita que sua inspiração vem dessa relação.

Figura 1 - Rede Neural Artificial



Fonte: KOVÁCS, 2002.

Uma definição mais simplificada e abrangente pode ser utilizada definindo RNA's como um paradigma de processamento da informação inspirado na forma em que os sistemas nervosos biológicos, como por exemplo o cérebro, processam suas informações (NAHAR, 2012).

O neurônio é o principal elemento desse paradigma que está estruturado em um sistema de processamento da informação, composto por um certo número desses elementos altamente conectados entre si. Cada neurônio tem sua função individual, mas todos trabalham em conjunto para resolver um problema específico, como reconhecimento de padrões ou a classificação de dados. Assim como as pessoas, as RNAs aprendem através de exemplos em um processo de treinamento ou aprendizagem, que consistem em ajustar as conexões que existem entre os neurônios (NAHAR, 2012).

As RNA's se constituem em grupos de unidades chamadas "neurônios" que são análogas aos neurônios nervosos. Com cada neurônio se conectando um ao outro com os pesos, comumente chamado de pesos sinápticos (VELÁSQUEZ *et al.*, 2011).

O diferencial das RNA's em comparação a outros sistemas especialistas são a velocidade, a simplicidade e a capacidade de modelar um problema multivariado para resolver relações complexas entre as variáveis e extrair as relações não lineares por meio de dados de treinamento (VELÁSQUEZ *et al.*, 2011). A metodologia de RNA é mais robusta (ou seja, não requer um alto nível de julgamento do operador), e usa um sofisticado modelo não linear para alcançar um alto desempenho de classificação (BISSACOT *et al.*, 2016). O fato de não precisar de um modelo matemático rígido, e os parâmetros de calibração poderem ser desenvolvidos usando dados de treinamento ou processo de aprendizagem, coloca as RNA's em vantagem quando comparadas a outros sistemas (BOLANCA *et al.*, 2017).

As RNA's representam uma ferramenta de modelagem promissora em quase todas as áreas, onde abordagens quantitativas podem ser utilizadas para ajudar na tomada de decisões (BISSACOT *et al.*, 2016). Há muitos exemplos de aplicações que empregam redes neurais na resolução de problemas. Por exemplo, elas têm sido utilizadas em processamento de sinal, medicina, reconhecimento de padrões, robótica, controle, previsão, produção de fala, reconhecimento de voz, negócios, manufatura, sistemas de energia e nas ciências agrárias (VELÁSQUEZ *et al.*, 2011).

Sua utilização na área das ciências agrárias ainda é bem tímida, mas está em franca evolução, haja visto pela quantidade de trabalhos que foram publicados nos últimos anos. Como o trabalho publicado por

Moshou *et al.* (2004) que usaram redes neurais artificiais com padrões de medidas de reflectância em plantações de trigo para detecção da ferrugem amarela. Como continuidade do seu trabalho, Moshou *et al.* (2005) utilizaram a fusão de dados de reflectância multiespectrais e hiperespectrais em uma RNA do tipo mapa de Kohonen para detectar a doença antes da mesma se tornar visível.

A utilização de RNA's abrange diversos objetivos como demonstrado por Chakraborty *et al.* (2004) para predição da severidade da antracnose na Austrália, Brasil e Colômbia. E, Marta *et al.* (2005) utilizaram as RNA's no desenvolvimento de um modelo epidemiológico para a predição da infecção por *Plasmopara viticola*.

As RNA's são muitas vezes combinadas com técnicas de processamento de imagens. Huang (2007) classificou doenças em mudas de *Phalaenopsis* como a podridão bacteriana, a mancha marrom bacteriana e a podridão negra de *Phytophthora*, onde a rede teve uma acurácia de 89%. Já o trabalho de Sanyal e Patel (2008) trouxe o uso da técnica para o reconhecimento de padrões de *Brown spot* e *Blast diseases* em plantas de arroz. Também, no estudo de Prabukumar, Krishna e Kamakakannan (2010) são apresentadas diversas técnicas, entre elas as RNA's, para classificar doenças em folha de bananeiras.

A utilização de RNA's também está associada com técnicas estatísticas como demonstrado no trabalho de Liu, Wu e Huang (2010). Seu uso combinado com dados de reflectância e a técnica estatística de análise de componentes principais, objetivou classificar quatro diferentes níveis de infecção por fungos (sadio, leve, moderado e grave) em grãos de arroz, conseguindo precisões entre 85 e 100%.

Similarmente, Lili *et al.* (2011) utilizaram as RNA's para avaliar padrões de classificação de doenças em plantas herbáceas. Nesta mesma linha, Bernardes *et al.* (2013) propuseram um método automático de identificação de doenças do algodão extraíndo padrões dos sintomas apresentados nas folhas através de imagens digitais e utilizando Suport Vector Machine para classificar as doenças apresentadas.

Revathi e Hemalatha (2014) desenvolveram uma RNA a partir de dados extraídos de imagens de folhas de algodoeiro para categorizar entre seis diferentes doenças causadas tanto por bactérias como fungos. Um outro exemplo para os tipos de entrada são os dados de reflectância espectral, como no artigo de Zhang *et al.* (2011) que com auxílio de RNA estimaram as doenças presentes em plantações de arroz.

3.6. Mask R-CNN

Similarmente as RNA's, as redes neurais convolucionais (CNN's) possuem diversos tipos e inúmeras arquiteturas de redes. Trata-se de uma técnica de aprendizado profundo (baseado em multi-camadas) que se tornou popular para a detecção de objetos, porque além de conseguir classificá-los, essas redes também conseguem extrair descritores da imagem (SHAHEEN *et al.*, 2016).

Ela é aplicada da seguinte forma: cada valor da matriz de convolução (matriz máscara) representa o peso de uma sinapse. Então cada iteração errada da rede neural ao realizar uma convolução, é modificada e adaptada para tentar gerar o resultado esperado. Em função da CNN possuir poucos parâmetros e conexões, ela se torna mais fácil de ser treinada, pois basta após um resultado inesperado, o desenvolvedor exibir qual era a resposta esperada para tal caso (KRIZHEVSKY, 2012).

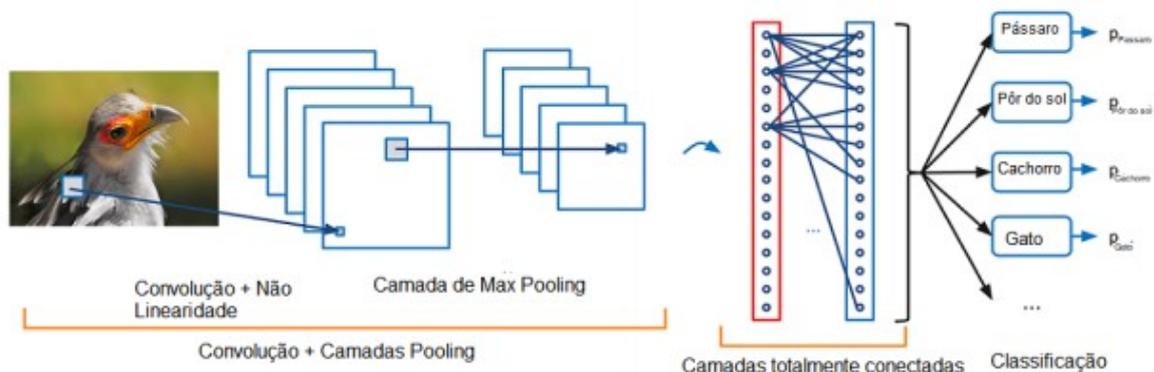
Uma convolução é composta por três elementos: a entrada, que é uma matriz de valores numéricos; o kernel, que é uma matriz de dimensões menores do que a entrada; e a saída, que é o resultado da

convolução. A saída é chamada de feature map. Quando uma imagem é processada em uma rede neural convolucional, a entrada pode ter milhares de pixels, mas através de um kernel que ocupa dezenas de pixels, atributos significativos da imagem podem ser detectados (GOODFELLOW; BENGIO; COURVILLE, 2016). Isso significa que precisa-se armazenar menos parâmetros, o que reduz os requisitos de memória do modelo e melhora sua eficiência estatística. Isso também significa que calcular a saída requer menos operações (GOODFELLOW; BENGIO; COURVILLE, 2016).

Segundo Araújo *et al.* (2017), uma das vantagens da utilização de uma CNN consiste na capacidade de extrair características relevantes, através do aprendizado de transformações na imagem. Outra vantagem está na dependência de menos parâmetros de ajustes em comparação com redes totalmente conectadas com o mesmo número de camadas ocultas. Vargas *et al.* (2016) cita como diferencial, uma CNN possibilitar a mistura de múltiplos mapas de características ao mesmo tempo, permitindo a extração de características cada vez mais complexas. Assim, as CNN's são capazes de sozinhas criarem filtros extremamente complexos aproveitando ao máximo as informações provenientes dos dados de treinamento.

Uma CNN é composta de duas grandes etapas: a extração de características através das convoluções e a classificação. A arquitetura de uma típica CNN possui basicamente três tipos de camadas, as convolucionais, pooling e camadas totalmente conectadas. A camada de convolução contém um conjunto de neurônios que são responsáveis por aplicar filtros em determinadas partes da imagem. Gonzalez e Woods (2009) define a convolução como resultante do produto escalar entre a máscara e a vizinhança de cada pixel da imagem. A camada de pooling, tem como principal objetivo a redução de dimensionalidade dos dados na rede, para tal, a operação mais utilizada é o max-pooling, onde as partes mais importantes da imagem são capturadas. A Figura 2 apresenta um exemplo genérico de CNN, onde estas etapas são denominadas convolution + pooling layers.

Figura 2 - Exemplo de CNN para classificação de espécies de animais, com camadas de convolução, max pooling e uma camada totalmente conectada.



Fonte: Adaptado a partir de ACADEMY, 2019.

As Redes Neurais Convolucionais têm sido umas das inovações mais influentes no campo da visão computacional. Sua aplicação no campo de análise de imagens ganhou maior interesse a partir de 2012 devido ao seu alto desempenho (KRIZHEVSKY; HINTON, 2012). Apesar da CNN ter sido bastante

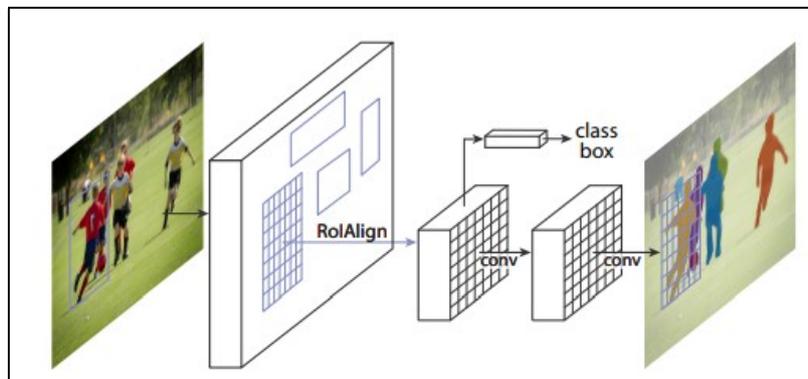
importante para comunidade de visão computacional, essa rede é bastante lenta. Sua lentidão é consequência do custo de processamento elevado. Pois para realizar a convolução de uma matriz 3x3 em uma imagem 3000x3000 são realizados em torno de 81 milhões de cálculos matemáticos. Após realizar esses cálculos, o algoritmo realiza a verificação se os novos valores atendem ao requisito esperado. Caso não atendam são realizados mais 81 milhões de cálculos até que seja retornado o valor esperado (GOODFELLOW, 2016).

Na busca por soluções para melhorar essa rede, com o decorrer dos anos, grandes melhorias foram alcançadas com o desenvolvimento de novas redes, algumas delas baseadas no conceito principal da CNN, como a Fast R-CNN, Faster R-CNN e MASK R-CNN (LIU *et al.*, 2016; REN *et al.*, 2017).

A Rede Neural Convolutiva Regional De Máscara (Mask R-CNN) (HE *et al.*, 2016), além de detectar os objetos e classificá-los com a utilização de bounding box, também faz a segmentação por instâncias desses objetos detectados. A segmentação por instâncias segmenta os objetos de forma individual, diferindo da segmentação semântica onde os objetos são segmentados em classes apenas. Esse tipo de rede é uma extensão da estrutura Faster R-CNN (REN *et al.*, 2015).

A Mask R-CNN divide-se em duas etapas, primeiro gerando propostas sobre as regiões onde pode haver um objeto. Em seguida, prevê a classe do objeto, refinando o bounding box e gerando também uma máscara de segmentação do objeto com base na primeira etapa. As etapas estão conectadas à estrutura do backbone, que consiste no modelo (FPN). Pode ser utilizada com diferentes variações, como: ResNet, ou com as diferentes variações da arquitetura Inception (SZEGEDY *et al.*, 2016), que utiliza variados filtros, buscando garantir a extração de características em diferentes escalas.

Figura 3 - Exemplo de uma estrutura Mask R-CNN para segmentação de instâncias.



Fonte: HE *et al.*, 2018.

A utilização dessa técnica, associada a análise de imagens, tem obtido sucesso em diversas áreas, como por exemplo na agricultura. Na literatura são encontrados diversos estudos com a aplicação dessa metodologia na área, como: para detectar morangos (YU *et al.*, 2019); detecção de maçãs em um conjunto de dados de teste em 120 imagens (JIA *et al.*, 2020); aplicação em análise de imagens tridimensionais, a fim de identificar corretamente maçãs (GENÉ-MOLA *et al.*, 2020); entre outros.

4. REFERÊNCIAS

ACADEMY, D. S. As 10 principais arquiteturas de redes neurais. In: . **Deep Learning Book**. [s.n.], 2019. cap. 10. Disponível em:<<http://www.deeplearningbook.com.br>> .

ALVARENGA, R.O.; MARCOS FILHO, J.; GOMES JUNIOR, F. G. Avaliação do vigor de sementes de milho super doce por meio da análise computadorizada de imagens de plântulas. **Revista Brasileira de Sementes**, v. 34, p. 488-494, 2012.

AMARAL, Julia Bissoli do et al. X-ray test to evaluate the physiological potential of *Tabebuia heptaphylla* seeds. **Revista Brasileira de Sementes**, v. 33, n. 4, p. 601-607, 2011.

ARAÚJO, F. H.; CARNEIRO, A. C.; SILVA, R. R.; MEDEIROS, F. N.; USHIZIMA, D. M. Redes neurais convolucionais com tensorflow: Teoria e prática. **Sociedade Brasileira de Computação**. III Escola Regional de Informática do Piauí. Livro Anais-Artigos e Minicursos, Sociedade Brasileira de Computação, v. 1, p. 382–406, 2017.

BARROS, D. I.; DIAS, D. C. F.S.; BHERING, M. C.; DIAS, L. A.S.; ARAÚJO, E. F. Uso do teste de tetrazólio para avaliação da qualidade fisiológica de sementes de abobrinha. **Revista Brasileira de Sementes**, vol. 27, nº 2, p.165-171, 2005.

BATISTA, V. T.; NUNES, J. V. D.; NÓBREGA, L. H. P. Qualidade fisiológica de sementes de *Brachiaria brizantha* cv. marandu com variação nas características de pureza. **Revista de Agricultura**, Piracicaba, v. 91, n. 1, p.92-100, 2016.

BERNARDES, A. A. et al. Identification of Foliar Diseases in Cotton Crop. In: TAVARES, J. M. R. S.; JORGE, R. M. N. (Eds.). Topics in Medical Image Processing and Computational Vision. Lecture Notes in Computational Vision and Biomechanics. [s.l.]**Springer Netherlands**. p. 67–85, 2013.

BHÉRING, M.C.; SILVA, R.F.; ALVARENGA, E.M.; DIAS, D.N.F.S.; PENA, M.F. Avaliação da viabilidade e do vigor de sementes de feijão de vagem (*Phaseolus vulgaris* L.) pelo teste de tetrazólio. **Boletim Técnico UFV**. Viçosa. 27p,1996.

BISSACOT, A. C. G.; SALGADO, S. A. B.; BALESTRASSI, P. P.; PAIVA, A. P.; ZAMBRONI SOUZA, A. C., WAZEN, R. Comparison of Neural Networks and Logistic Regression in Assessing the Occurrence of Failures in Steel Structures of Transmission Lines.**The Open Electrical & Electronic Engineering Journal**, v. 10, n.1, p.11–26, 2016.

BOLANCA, T.; STRAHOVNIK, T.; UKIC, S.; STANKOV, M. N.; ROGOSIC, M. Modeling of policies for reduction of GHG emissions in energy sector using ANN: case study—Croatia (EU). **Environmental**

Science and Pollution Research, v. 24, n. 19, p. 16172–16185, 2017.

BRASIL. Lei nº 10.711, de 5 de agosto de 2003. Dispõe sobre o Sistema Nacional de Sementes e Muda e dá outras providências. **Diário Oficial da União**: seção1, Brasília, DF, 06 ago, 2003.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Regras para análise de sementes**. Brasília. 2009.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **SIGEF - Controle da Produção de Sementes e Mudanças - Indicadores**. Brasília, DF, 2018. Disponível em: . Acesso em: 30 nov. 2020.

CAMELO, H. N.; LUCIO, P. S.; LEAL JUNIOR, J. B. V.; CARVALHO, P. C. M.; SANTOS, D. V. G. Innovative hybrid models for forecasting time series applied in wind generation based on the combination of time series models with artificial neural networks. **Energy**, v. 151, p. 347–357, 2018.

CUSTÓDIO, C. C.; DAMASCENO, R. L.; MACHADO NETO, N. B. Imagens digitalizadas na interpretação do teste de tetrazólio em sementes de *Brachiariabrizantha*. **Rev. bras. sementes** vol.34 no.2 Londrina June 2012.

CHAKRABORTY, S. et al. Weather-based prediction of anthracnose severity using artificial neural network models. **Plant Pathology**, v. 53, n. 4, p. 375–386, 2004.

CHIQUITO, A. A.; GOMES JUNIOR, F. G.; MARCOS FILHO, J. Assessment of physiological potential of cucumber seeds using the software Seedling Vigor Imaging System® (SVIS®). **Revista Brasileira de Sementes**, v. 34, p. 255-263, 2012.

DE SOUZA, Luciana Aparecida et al. Uso de raios-x na avaliação da qualidade de sementes de *Platypodium elegans* Vog. **Revista Ciência Agrônômica**, v. 39, n. 2, p. 343-347, 2008.

DELL'AQUILA, A. Digital imaging information technology applied to seed germination testing. **A review. Agronomy for Sustainable Development**, v. 29, n.1, p. 213-221, 2009.

DELL'AQUILA, A. Application of a computer-aided image analysis system to evaluate seed germination under different environmental conditions. **Italian Journal of Agronomy**, v.8, n.1, p.51-62, 2004.

DELL'AQUILA, A. Towards new computer imaging techniques applied to seed quality testing and sorting. **Seed Science and Technnology**, v.35, n.3, p.519-538, 2007.

DELOUCHE, J.C., STILL, T.W., RASPET, M., LIENHARD, M. O teste de tetrazólio para viabilidade da

semente. Brasília-DF: **AGIPLAN**. 103p, 1976.

DIAS, M. C. L. L.; ALVES, S. J. Avaliação da viabilidade de sementes de *Brachiaria brizantha* (hochst. ex a. rich) stapf pelo teste de tetrazólio. **Revista Brasileira de Sementes**, vol. 30, nº 3, p. 145-151, 2008..

FEITOSA, S. dos S. et al. Viability study of seeds of candeia *Eremanthus erythropappus* (DC.) Mac Leish by germination and x-ray tests. **Floresta**, v. 39, n. 2, p. 393-399, 2009.

FOGAÇA, C. A.; MALAVASI, M. M.; ZUCARELI, C.; MALAVASI, U. C. Aplicação do teste de tetrazólio em sementes de *Gleditschia amorphoides* Taub. Caesalpinaceae. **Revista Brasileira de Sementes**, Londrina, v. 28, n. 3, p. 101-107, 2006.

FRANÇA NETO, J.B. Teste de tetrazólio para determinação do vigor de sementes. In: KRZYZANOWSKI, F.C.; VIEIRA, R.D.; FRANÇA NETO, J.B. (Ed.). Vigor de sementes: conceitos e testes. Londrina: **ABRATES**. 218p, 1999.

FRANÇA NETO, J.B.; KRZYZANOWSKI, F.C.; COSTA, N.P. O teste de tetrazólio em sementes de soja. Londrina: **Embrapa/ cnpso**. 72p, 1998.

GASPAR-OLIVEIRA, C. M.; MARTINS, C. C.; NAKAGAWA, J. Précondicionamento das sementes de mamoneira para o teste de tetrazólio. **Acta Scientiarum.Agronomy**, Maringá, v.33, n. 2, p.303-311, 2011.

GENÉ-MOLA, J.; GREGORIO, E.; CHEEIN, F.A.; GUEVARA, J.; LLORENS, J.; SANZ-CORTIELLA, R.; ESCOLÀ, A.; ROSELL-POLO, J.R. Fruit detection, yield prediction and canopy geometric characterization using LiDAR with forced air flow. **Comput. Electron. Agric.**, 168, 105–121, 2020.

GENEVE, R.L.; KESTER, S.T. Evaluation of seedling size following germination using computer Aided analysis of digital images from a flat bed scanner. **Hort Science**, v.36, n.6, p.1117-1120, 2001.

GOODFELLOW, Ian et al. Deep learning. Cambridge: MIT press, 2016.

GONZALEZ, R. C.; WOODS, R. C. Processamento digital de imagens . [S.l.]: Pearson Educación, 2009.

GUEDES, M. A. et al. Caracterização Física de Grãos de Soja Utilizando-se Processamento Digital de Imagens. **Revista Brasileira de Produtos Agroindustriais**, Campina Grande, v.13, n.3, p.279-294, 2011.

GUEDES, R. S; ALVES, E. U; GONÇALVES, E. P; VIANA, J. S; SILVA, K. B; GOMES, M. S. S. Metodologia para teste de tetrazólio em sementes de *Amburana cearensis* (Allemão) A.C. Smith. **Revista Brasileira de Plantas Mediciniais**, Botucatu, v. 12, n. 1, p. 120-126, 2010.

HE, K., GKIOXARI, G., DOLLÁR, P., & GIRSHICK, R. Mask r-cnn. In **Proceedings of the IEEE international conference on computer vision**.pp. 2961-2969, 2018.

HE, K. et al. **Deep residual learning for image recognition**. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.: s.n.], p. 770–778, 2016.

HOSOMI, S.T., SANTOS, R.B., CUSTODIO, C.C., SEATON, P.T., MARKS, T.R.; MACHADO-NETO, N.B. Preconditioning *Cattleya* seeds to improve the efficacy of the tetrazolium test for viability. **Seed Science and Technology**, v.39, n.1, p.178-189, 2011.

HUANG, K.-Y. Application of artificial neural network for detecting *Phalaenopsis* seedling diseases using color and texture features. **Computers and Electronics in Agriculture**, v. 57, n. 1, p. 3–11, 2007.

JIA, W.; TIAN, Y.; LUO, R.; ZHANG, Z.; ZHENG, Y. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. **Comput.Electron.Agric.**, 172, 105380,n 2020.

KOVÁCS, Zsolt László. **Redes neurais artificiais**. Editora Livraria da Física, 2002.

KRIZHEVSKY, A.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. NIPS'12 Proceedings of the 25th **International Conference on Neural Information Processing Systems**, p. 1097-, 2012.

KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. Imagenet classification with deep Convolutional neural networks. In: **Advances in neural information processing systems**. p. 1097-1105, 2012..

LAMARCA, V. E.; BARBEDO, C. J. Methodology of the tetrazolium test for assessing the viability of seeds of *Eugenia brasiliensis* Lam., *Eugenia uniflora* L. and *Eugenia pyriformis* Cambess. **Journal of Seed Science**, v. 36, n. 4, p 427-434, 2014.

LANDAU, Elena Charlotte; RESENDE, RMS; MATOS NETO, F. da C. Evolução da área ocupada por pastagens. **Embrapa Milho e Sorgo-Capítulo em livro científico (ALICE)**, 2020.

LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. Convolutional networks and applications in vision. In: Circuits and Systems (ISCAS), Proceedings of IEEE **International Symposium** on. [S.l.: s.n.], p. 253–256, 2010.

LILI, N. A.; BORHAN, N. M.; KHALID, F. Classification of Herbs Plant Diseases via Hierarchical Dynamic

Artificial Neural Network after Image Removal using Kernel Regression Framework. **International Journal on Computer Science and Engineering**, v. 3, n. 1, p. 15–20, 2011.

LIMA, K. N. *et al.* Superação de dormência em capim-braquiária. **Nucleus**, v. 12, n. 2, p. 167-174, 2015.

LIU, W. *et al.* Ssd: Single shot multibox detector. In: LEIBE, B. *et al.* (Ed.). **Computer Vision – ECCV 2016**". Cham: Springer International Publishing. p. 21–37, 2016.

LIU, Z.-Y.; WU, H.-F.; HUANG, J.-F. Application of neural networks to discriminate fungal infection levels in rice panicles using hyperspectral reflectance and principal components analysis. **Computers and Electronics in Agriculture**, v. 72, n. 2, p. 99–106, jul. 2010.

MACEDO, M.C.M. **Pastagens no ecossistema cerrados: evolução das pesquisas para o desenvolvimento sustentável**. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 42., Goiânia. Anais. Goiânia: UFG; SBZ, 2005. p. 56-84, 2005.

MARCOS FILHO, J. Importância do potencial fisiológico da semente de soja. **Informativo ABRATES**, Londrina, v. 23, n. 1, p. 21-23, 2013.

MARCOS FILHO, J. Sistema computadorizado de análise de imagens de plântulas (svis®) para avaliação do vigor de sementes. **Informativo ABRATES**, Londrina, v. 20, n. 3, p. 40-44, 2010.

MARCOS-FILHO, J. **Fisiologia de sementes de plantas cultivadas**. 2. ed. Londrina: ABRATES, 2015.

MARTA, A. D. *et al.* Neural network for the estimation of leaf wetness duration: application to a *Plasmoparaviticola* infection forecasting. **Physics and Chemistry of the Earth, Parts A/B/C**, v. 30, n. 1-3, p. 91–96, jan, 2005.

MARTINS, L.; SILVA, W. R. Efeitos imediatos e latentes de tratamentos térmico e químico em sementes de *Brachiaria brizanthacultivar* Marandu. **Bragantia**, Campinas, v. 62, n. 1, p. 81-88, 2003.

MASETTO, Tathiana Elisa; FARIA, José Márcio Rocha; QUEIROZ, Sue Ellen Ester. Avaliação da qualidade de sementes de cedro (*Cedrela fissilis*-Meliaceae) pelo teste de raios X. **Ciência e Agrotecnologia**, v. 32, n. 6, p. 1708-1712, 2008.

MCDONALD, M. B. *et al.* Using the seed vigor imaging system for improving stand establishment, **Acta Horticulturae**, San Antonio, v. 792, pp. 83-91, 2008.

MCDONALD, M.B.; EVANS, A.F.; BENNETT, M.A. Using Scanners to improve seed and seedling evaluations. **Seed Science and Technology**, v.29, n.3, p.683-689, 2001.

MEDICA, J. A. S.; REIS, N. S.; SANTOS, M. E. R. Caracterização morfológica em pastos de capim-marandu submetidos a frequências de desfolhação e níveis de adubação. **Revista Ciência Animal Brasileira**, v. 18, e-40460, p. 1-13, 2017.

MELO, L. F.; MARTINS, C. C.; DA SILVA, G. Z.; BONETI, J. E. B.; VIEIRA, R. D. Beneficiamento na qualidade física e fisiológica de sementes de capim-mombaça. **Revista Ciência Agronômica**, Fortaleza, v. 47, n. 4, p. 667-674, 2016.

MONDO, V. H. V.; CICERO, S. M. Análise de imagens na avaliação da qualidade de sementes de milho localizadas em diferentes posições na espiga. **Revista Brasileira de Sementes**, vol. 27, nº 1, p.09-18, 2005.

MOSHOU, D. et al. Automatic detection of “yellow rust” in wheat using reflectance measurements and neural networks. **Computers and Electronics in Agriculture**, v. 44, n. 3, p. 173–188, 2004.

MOSHOU, D. et al. Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. **Real-Time Imaging, Spectral Imaging II Spectral Imaging II**. v. 11, n. 2, p. 75–83, 2005.

NAHAR, K. Artificial Neural Network. **COMPUSOFT**, v. 1, n. 2, p. 25–27, 2012.

NERY, M. C.; DE CARVALHO, M. L. M.; OLIVEIRA, J. A.; KATAOKA, V. Y. Beneficiamento de sementes de nabo forrageiro. **Revista Brasileira de Sementes**, Londrina, v. 31, n. 4, p. 36-42, 2009.

NERY, M. C. NERY, F. C.; SILVA, D. R. G.; SOARES, F. P. **Produção de sementes forrageiras**. Universidade Federal de Lavras, Departamento de Ciência do Solo. Boletim Técnico, n.88, p. 1-47, 2012

NUNES, E. de O. **Segmentação por textura em imagens multibandas**. 2006. 152 f. Tese (Doutorado). Universidade Federal Fluminense. Rio de Janeiro, 2006.

NUNES, S. G.; BOOCK, A.; PENTEADO, M. I. O.; GOMES, D. T. Empresa Brasileira de Pesquisa Agropecuária Centro Nacional de Pesquisa de Gado de Corte. (1984). *Brachiaria brizantha* cv . MARANDU. Em **EMBRAPA-CNPGC**. Documentos, 21. (p. 31). Campo Grande - MS: EMBRAPA-CNPGC.

OAKLEY, K.; KESTER, S. T.; GENEVE, R. L. Computer-aided digital image analysis of seedling size and growth rate for assessing seed vigour in Impatiens. **Seed Science and Technology**, v. 32, p. 907-915, 2004.

OHLSON, O de C. et al.; Análise exploratória de dados, qualidade de sementes de azevém comercializadas no Estado do Paraná. **Informativo Abrates**, Londrina, vol. 21, n.3, p.47 – 51, 2011.

PRABUKUMAR, M.; KRISHNA, B.; KAMALAKANNAN, J. Image Processing and Pattern Classification Technique in a Machine Vision System that Identifies and Classifies the Plant Diseases Based on the Visual Symptoms. **International Journal of Advanced Research in Computer Science**, v. 1, n. 4, p. 302–307, 2010.

PUPIM, Tulio Lourenço et al. Adequação do teste de raios x para avaliação da qualidade de sementes de embaúba (*Cecropia pachystachya* Trec.). **Revista Brasileira de Sementes**, v. 30, n. 2, p. 28-32, 2008.

RAYMAN, P.R. Minha experiência com *Brachiaria brizantha*. Campo Grande, **Rayman's Seeds Sementes de PastagensTropicais**.3p, 1983.

REN, S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks. In: **Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1**. Cambridge, MA, USA: MIT Press. (NIPS'15), p. 91–99, 2015.

REN, S.; HE, K.; GIRSHICK, R.; SUN, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 39, n. 6, p. 1137–1149, 2017.

REVATHI, P.; HEMALATHA, M. Cotton Disease Identification Using Proposed CIGDFNN Classifier.**Asian Journal of Scientific Research**, v. 7, n. 2, p. 225–231, 1 fev. 2014.

ROSA, S. D. V. F. da; FREITAS, M. N. de; SAÚDE, A.; PEREIRA, C. C; CIRILLO M. Â. **Imagens digitais do teste de tetrazólio para avaliar a qualidade de grãos de café**. In: Simpósio de Pesquisa dos Cafés do Brasil, 8. SPCB, Salvador. 6p, 2013.

SAKO, Y. et al.A system for automated seed vigourassessment.**Seed Science and Technology**. v. 29, p. 625-636, 2001.

SALINAS, Adriana Rita et al. X-ray: Characterization of Ginkgo biloba L. seeds using digital and manual measurements. **Revista Caatinga**, v. 25, n. 3, p. 1-7, 2012.

SANYAL, P.; PATEL, S. C. Pattern recognition method to detect two diseases in rice plants.**The Imaging Science Journal**, v. 56, n. 6, p. 319–325, 2008.

SÃO PAULO. Comissão Estadual de Sementes e Mudanças (CESMSP). **Padrões de sementes de gramíneas forrageiras**. São Paulo: CESM-SP, 1996. n.p.

SAVONEN, E.An improvement to the topographic tetrazolium testing of Scots pine (*Pinus sylvestris*L.)

seeds. **Seed Science and Technology**, Zürich, v. 26, n. 1, p. 4957, 1999.

SHAHEEN, F.; VERMA, B. E ASAFUDDOULA, M. IMPACT OF AUTOMATIC FEATURE EXTRACTION IN DEEP LEARNING ARCHITECTURE. In *Digital Image Computing: Techniques and Applications (DICTA), International Conference on*, pp. 1–8. IEEE, 2016.

SILVA VN; CICERO SM. Análise de imagens de plântulas para avaliação do potencial fisiológico de sementes de berinjela. **Horticultura Brasileira** 32: 145-151, 2014..

SILVA, C. B. et al. Automated system of seedling image analysis (SVIS) and electrical conductivity to assess sun hemp seed vigor. **Revista Brasileira de Sementes**, v. 34, n. 1, p. 55-60, 2012.

SILVA, V.N.; SARMENTO, M.B.; SILVEIRA, A.C.; SILVA, C.S.; CICERO, S.M. Avaliação da morfologia interna de sementes de *Acca sellowiana* O. Berg por meio de análise de imagens. **Revista Brasileira de Fruticultura**, Jaboticabal - SP, v. 35, n. 4, p. 1158- 1169, 2013.

SOCOLOWSKI, F.; CICERO, S. M.; VIEIRA, D. C. M. Embryo morphological characterization by X-ray image and relation with mass and physiological quality of seeds of *Tecoma stans* L. Juss. Ex Kunth (Bignoniaceae). **Revista Brasileira de Sementes**, v. 30, n. 2, p. 200-208, 2008.

SOCOLOWSKI, F.; CICERO, S. M.; VIEIRA, D. C. M. Peso da semente de *Xylopia aromatica* (Annonaceae): avaliação da qualidade por raios-X e emergência de plântulas. **Scientia Agricola** , v. 68, n. 6, pág. 643-646, 2011.

SOUZA, F.H.D. de. Produção de sementes de gramíneas forrageiras tropicais. **São Carlos: Embrapa Pecuária Sudeste**. 43 p, 2001.

SZEGEDY, C. et al. **Rethinking the inception architecture for computer vision**. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.: s.n.], 2016. p. 2818–2826. ISSN 1063-6919.

TOMAZ, C. A. et al. Duração do teste de germinação do capim-tanzânia. **Revista Brasileira de Sementes**, v. 32, n. 4, p. 80-87, 2010.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: **Proceedings of the XXIX Conference on Graphics, Patterns and Images**. [S.l.: s.n.]. p. 1–4, 2016.

VELÁZQUEZ, S.; CARTA, J. A.; MATÍAS, J. M. Influence of the input layer signals of ANNs on wind power estimation for a target site: A case study. **Renewable and Sustainable Energy Reviews**, v. 15, n. 3, p.

1556–1566, 2011.

VIEIRA, R.D.; FRANÇA NETO, J.B. (Ed.).Vigor de sementes: conceitos e testes. Londrina: **ABRATES** p.8.1.1-8.1.13, 1999.

YAGUSHI, J. T.; COSTA, D. S.; FRANÇA NETO, J. B. Saturated salt accelerated aging and computerized analysis of seedling images to evaluate soybean seed performance. **Journal of Seed Science**. v. 36, p. 213-221, 2014.

YU, Y.; ZHANG, K.; YANG, L.; ZHANG, D. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. **Comput. Electron. Agric**, 163, 104846, 2019.

ZHANG, H. et al. Estimation of rice neck blasts severity using spectral reflectance based on BP-neural network. **Acta Physiologiae Plantarum**,v. 33, n. 6, p. 2461–2466, 2020.

5. ARTIGOS

5.1. Artigo 1- *Deep learning* na interpretação do teste de tetrazólio em sementes de *Urochloa brizantha* cv. Marandu

Este artigo foi elaborado conforme normas do Programa de Pós-Graduação em Produção Vegetal.

Resumo

O teste de tetrazólio é considerado um dos testes mais rápidos e eficientes para avaliar a viabilidade de sementes. Entretanto, apresenta subjetividade e complexidade, principalmente, quando se avalia espécies que possuem sementes pequenas, como é o caso da *Urochloa brizantha* cv. Marandu. Nesse sentido, objetivou-se verificar a eficiência da automação computacional por meio de Mask R-CNN na classificação de sementes de *Urochloa brizantha* cv. Marandu quanto à viabilidade a partir do teste de tetrazólio. Foram utilizados 5 lotes de sementes de *Urochloa brizantha* cv. Marandu, com quatro repetições de 50 sementes por lote. O teste de tetrazólio foi de acordo com as normas estabelecidas pelas Regras de Análises de Sementes (solução de 0,1%, por 5 horas, a 40 °C) (Brasil, 2009). A automação da classificação dos resultados foi realizada pela metodologia Mask R-CNN. O modelo foi eficiente na segmentação das sementes. Além de apresentar alta eficiência ao classificar as sementes em todos os lotes, apresentando valores de acurácia superiores a 80%, e acurácia geral em torno de 91%. Desta forma, a Mask R-CNN é uma ferramenta eficiente para a automação computacional do teste de tetrazólio em sementes de *Urochloa brizantha* cv. Marandu, com alto potencial de uso para análise de sementes em outras espécies.

PALAVRAS-CHAVE: Data Science. Inteligência artificial. Viabilidade de sementes. Análise de imagens. Mask R-CNN.

5.1.1. Introdução

O teste de tetrazólio pode ser considerado um dos mais importantes na avaliação da qualidade de sementes. É uma metodologia rápida para estimar a viabilidade e vigor de sementes. Tem como base a atividade das enzimas desidrogenases envolvidas no processo de respiração. A hidrogenação do 2-3-5-trifenil cloreto de tetrazólio produz nas células vivas uma substância vermelha, estável e não difusível, o trifênil formazan. Assim, pela reação é possível distinguir as partes vivas, coloridas de vermelho, daquelas mortas que possuem a cor branca-leitosa (ABBADE; TAKAKI, 2014).

Apesar do teste de tetrazólio ser eficiente, a precisão em sua interpretação é muito dependente da experiência e da capacidade do avaliador, tornando-o assim um teste subjetivo. Além de ser laborioso, o teste demanda inúmeras avaliações individuais que podem afetar a percepção do técnico e influenciar os resultados. Adicionalmente, a execução do teste de tetrazólio pode ser mais complexa quando realizada em espécies com sementes pequenas (CUSTÓDIO *et al.*, 2012). Como é o caso da *Urochloa brizantha* cv. Marandu, planta importante para pecuária brasileira, sendo considerada a mais representativa entre os campos de produção de sementes de pastagens no Brasil (LANDAU *et al.*, 2020).

Deep learning, associado à análise de imagens, pode ser tido como aliado para fornecer alternativas úteis para solucionar essa problemática. Uma vez que, em comparação com as avaliações manuais, a automação pode melhorar efetivamente a eficiência de avaliações e evitar a influência da subjetividade na precisão (ZHANG *et al.*, 2020). Além de possibilitar a redução de tempo nas avaliações e também tornar possível a criação de um banco de dados, permitindo a reprodutibilidade do teste a qualquer momento, abrindo caminho para o desenvolvimento de pesquisas futuras.

Dentre as alternativas de uso de *deep learning* na automação da análise de imagens estão às redes neurais convolucionais (CNN's). Essa é uma técnica de *data science* (baseado em multi-camadas) que se tornou popular para a detecção de objetos, porque além de conseguir classificá-los, também conseguem extrair descritores da imagem (SHAHEEN *et al.*, 2016). O principal algoritmo utilizado para esse fim é o Mask R-CNN (Mask Region Convolutional Neural Network), proposto por He *et al.* (2018), que tem a vantagem de integrar as caixas delimitadoras das classes encontradas na imagem com as máscaras de segmentação (HE *et al.*, 2017).

A utilização da Mask R-CNN para análise de imagens do teste de tetrazólio pode contribuir para o aperfeiçoamento do mesmo, tornando-o mais preciso e prático, além de diminuir a subjetividade nas interpretações. A partir do exposto, objetivou-se verificar a eficiência da automação computacional por meio de Mask R-CNN na classificação de sementes de *Urochloa brizantha* cv. Marandu quanto à viabilidade a partir do teste de tetrazólio.

5.1.2. Material e métodos

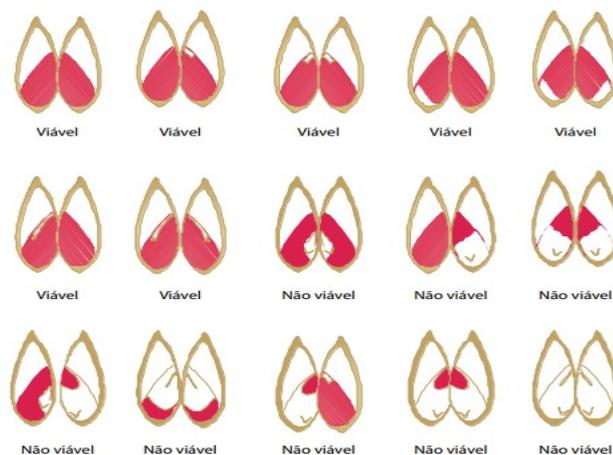
O experimento foi conduzido na cidade de Montes Claros-MG, (coordenadas: 16°40'58.16"S e 43°50'20.15" O), no período de dezembro de 2019 a setembro de 2020. Foram utilizados cinco lotes de sementes de *Urochloa brizantha* cv. Marandu, das safras 2017/2018 e 2018/2019, provenientes de diferentes locais comerciais.

O ensaio foi conduzido em delineamento inteiramente casualizado. Para realização do teste de tetrazólio foram avaliadas quatro repetições de 50 sementes para cada lote, pré-embebidas em água a 25 °C, por 16 horas. Decorrido esse período, as sementes foram cortadas longitudinalmente com auxílio de pinça e bisturi. Em seguida foram acondicionadas em recipientes plásticos contendo solução de 0,1% de cloreto 2-3-5 trifenil tetrazólio e mantidas no escuro, em banho-maria, por 5 horas, a 40 °C (BRASIL, 2009).

As sementes foram escaneadas para a obtenção das imagens digitais, que posteriormente foram utilizadas para a classificação visual da coloração no teste tetrazólio e para a automação das análises de imagem. Para isso, foi utilizado um escâner de mesa (HpOfficejet 4500 Desktop), com definição de 1200 pontos por polegada quadrada. Na parte superior desse escâner foi colado um fundo azul, feito com uma folha de E.V.A. Cada um dos dois fragmentos das sementes foi colocado diretamente sobre a placa de vidro do escâner com a parte corada voltada para baixo. No escaneamento, os fragmentos das sementes foram orientadas em 5 linhas e 20 colunas por vez. Desta forma, obtiveram-se 20 imagens (cinco lotes x quatro repetições), totalizando 2000 fragmentos de sementes (cinco lotes x quatro repetições x cinquenta sementes x dois fragmentos).

A classificação visual da imagem do teste de tetrazólio foi realizada para cada fragmento de semente baseando-se na classificação proposta pelas Regras de Análise de Sementes para a espécie em estudo (FIGURA 1). Considerando-se a localização e a intensidade da coloração, sendo possível identificar duas categorias: viáveis e inviáveis (TABELA 1).

Figura 1 - Escala para classificação visual dos resultados do teste de tetrazólio em sementes de *Brachiaria* spp.



Fonte: BRASIL, 2009.

Tabela 1 - Classificação visual do teste de tetrazólio em sementes de *Urochloa brizantha* cv. Marandu.

| Lotes | Inviável | Viável | Total | %Viabilidade |
|-------|----------|--------|-------|--------------|
| 1 | 284 | 116 | 400 | 29 |
| 2 | 138 | 262 | 400 | 65,5 |
| 3 | 61 | 339 | 400 | 84,75 |
| 4 | 69 | 331 | 400 | 82,75 |
| 5 | 122 | 278 | 400 | 69,5 |
| Total | 675 | 1325 | 2000 | - |

Fonte: Da autora, 2021.

Para o treinamento das redes, para cada um dos lotes foi utilizada uma das 4 imagens obtidas pelo scanner. Desta forma, consideraram-se 100 fragmentos das sementes para o treinamento oriundas de cada lote. Para o melhor treinamento foi montado um banco de imagens sintéticas. Para isso, imagens individuais de cada uma das sementes foram obtidas com o auxílio do software R. Nesta etapa, afim de possibilitar a análise computacional em um menor tempo, foi feita a compactação das imagens com o auxílio das funções *image_read* e *image_scale* do pacote *magick*.

As imagens foram compactadas para 400 pixels de largura, 314 pixels de altura. Após a compactação das imagens, as mesmas foram abertas pela função *readImage* do pacote *EImage*, a fim de obter para cada imagem três matrizes, cada uma com 400 colunas e 314 linhas contendo os valores de vermelho, verde e azul (RGB), com valores variando entre 0 e 1. Para a segmentação de cada imagem, com a finalidade de separar a semente do fundo, foi utilizada apenas o canal de vermelho, para a qual utilizou-se um limiar obtido por meio da função *otsu*.

Assim, todos os pixels com valor abaixo deste limiar foram considerados como fundo (o qual era de cor azul), atribuindo-se o valor 1; os demais pixels, correspondente às sementes, atribuiu-se o valor 0. A fim de preencher possíveis espaços vazios nos pixels correspondentes à semente foi utilizada a função *fillHull*. Essa matriz foi utilizada como coordenadas das sementes. A partir destas coordenadas, os valores de vermelho, verde e azul de cada pixel foram utilizados para a obtenção de uma imagem com extensão "png" contendo fundo transparente para cada um dos fragmentos de sementes. Para isso utilizou-se a função *pxflood* do package *imagemer writelImage* do pacote *EImage*.

A partir destas imagens, foi criado um banco sintético com oito mil imagens de treinamento e duas

mil imagens de validação. Para isso criaram-se imagens de tamanho de 320 x 320 com background azul. Sobre este background azul foram sobrepostas as imagens das sementes. Nesta composição, um valor aleatório de 1 a 14 imagens das sementes foram selecionadas computacionalmente de forma aleatória. O posicionamento destas sementes tanto dentro da imagem sintética como a sua rotação também foram estabelecidos de forma aleatória. Todo este processo foi implementado com o auxílio da biblioteca cocosynth (GITHUB, 2021).

A revisão foi feita por meio do algoritmo Mask R-CNN na plataforma online do google colab (GITHUB, 2021), a fim de realizar-se o treinamento da rede. Foram utilizadas as configurações pré-definidas pelo repositório, incluindo as arquiteturas de rede e as funções de perda. A rede ResNet101 foi empregada para a extração de características (HE *et al.*, 2016). A partir dos pesos iniciais da ResNet101 obtido pelo treinamento usando o banco de dados MSCOCO, foi realizado o ajuste fino com base no banco sintético de imagens.

Considerou-se 25 épocas de treinamento com a otimização do gradiente estocástico descendente com a taxa de aprendizado de 0.001 e tamanho do batch de 2. Os dados de treinamento sintéticos tiveram seu tamanho fixado em 320x320 (DPI). A saída das redes para o conjunto de caixas delimitadoras (bounding boxes) e máscaras (mask regions) foi obtida considerando o valor de probabilidade limiar de 0,70.

Após o treinamento da rede, o teste foi feito com as imagens das demais 3 repetições de cada um dos lotes (300 fragmentos de sementes). Para verificar a acurácia na detecção dos objetos pela Mask R-CNN foi construída uma matriz de confusão, tendo-se as classificações preditas pela rede em função das classificações visuais. Foram utilizadas as métricas Recall (Equação 1), Acurácia (Equação 2), precisão (Equação 3) e F-Measure (Equação 4) para avaliar a eficiência da rede, calculadas a partir das Equações (1, 2, 3 e 4). Onde: TP refere-se aos verdadeiros positivos e FN aos falsos negativos.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (3)$$

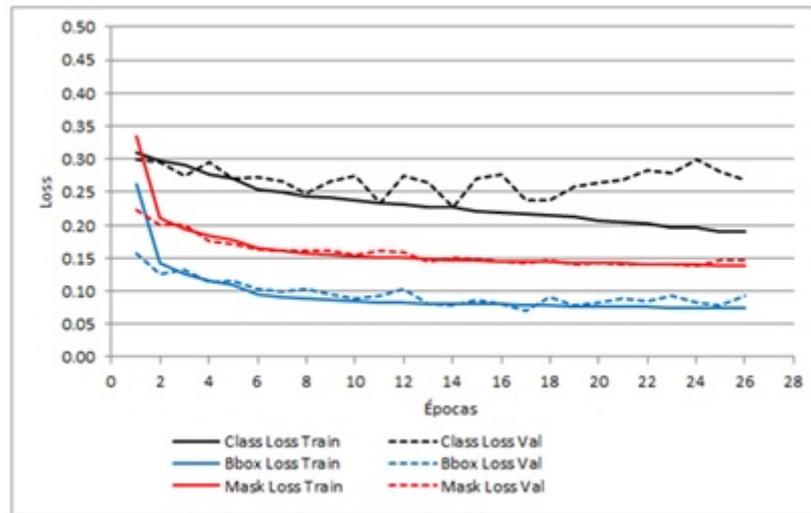
$$F - \text{Measure} = \frac{2 * \text{Precisão} * \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (4)$$

5.1.3. Resultados

A partir do conjunto de dados, foram obtidas curvas da função de perda por iteração na classificação das sementes para o grupo de treinamento e validação (FIGURA 2). Realizou-se o acompanhamento da função de perda ao decorrer das épocas de forma que quando se utilizaram 14 épocas o resultado foi mais eficiente, devido a menor taxa de erro na classificação, na caixa delimitadora e na máscara. Com o aumento no número de épocas utilizadas, verificou-se aumento o erro na classificação

da amostra de validação (Class Loss Val) indicando que o aumento no número de épocas de treinamento é inviável, devido provavelmente ao *overfitting*.

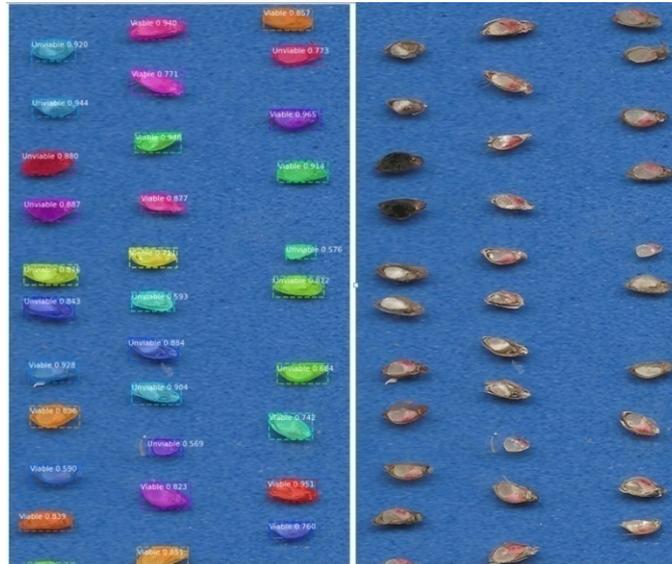
Figura 2 - Curvas da função de perda por iteração na classificação de sementes de *Urochloa brizantha* cv. Marandu.



Fonte: Da autora, 2021.

A Mask R-CNN requer anotações de caixas delimitadoras e imagens de máscaras que necessariamente circunscrevem a área da semente. Exemplos de imagens de sementes capturadas pelo scanner são mostrados na Figura 3. No modelo utilizado foram identificadas as partes das sementes com viabilidade ou não, as quais foram delimitadas por uma marcação colorida aleatória. Assim, pode-se observar que o modelo foi eficiente na identificação dos objetos (FIGURA 3).

Figura 3 - Detecção pelo estabelecimento de caixas delimitadoras e máscaras em imagens utilizadas para treinamento e validação.



Fonte: Da autora, 2021.

A eficácia da rede Mask R-CNN em classificar as sementes quanto à viabilidade, foi avaliada utilizando-se valores de precisão, recall, F-measure e acurácia (TABELA 2). O modelo atingiu recall de 79% para o conjunto de dados de sementes inviáveis e 95% para viáveis. De forma geral, os valores das métricas de desempenho foram maiores quando o modelo classificou sementes viáveis, com 92% de precisão, enquanto que para as sementes inviáveis essa métrica foi ligeiramente inferior (86%). O algoritmo classificador apresentou alta eficiência em todos os lotes, uma vez que os valores de acurácia foram superiores a 80%, com acurácia geral em torno de 91%.

Ao analisar a performance da metodologia em cada lote, constatou-se que a mesma foi altamente precisa (100%) ao identificar sementes inviáveis no lote um e para identificar sementes viáveis no lote três. Apresentando maiores taxas de acurácia ao classificar as sementes dos lotes três (98%) e quatro (94%), sendo que nesses lotes as métricas apresentaram valores mais baixos na identificação de sementes inviáveis. Porém, nos lotes cinco e dois o modelo obteve os menores valores de acurácia (83 e 85%). Também nos lotes cinco e dois foi observado o maior número de classificações incorretas, na qual a classificação no lote 5 originou 32 falsos negativos e no segundo lote 40 falsos positivos.

Ao comparar o desempenho entre os lotes, é constatável que no lote quatro o modelo apresentou os menores valores de precisão, recall e F-measure, na identificação de sementes inviáveis. Entretanto, durante a classificação das sementes do lote um ocorreu o oposto, com os maiores valores ao identificar as partes inviáveis, resultando na ausência de falsos negativos.

Tabela 2 - Precisão, recall e acurácia na classificação de sementes como viáveis ou inviáveis pela metodologia Mask R-CNN.

| Lote | Predito | Real | | Precisão | Recall | F-measure | Acurácia |
|-------|---------|------|------|----------|--------|-----------|----------|
| | | I | V | | | | |
| 1 | I | 212 | 0 | 1,00 | 0,90 | 0,95 | 0,92 |
| | V | 23 | 65 | 0,74 | 1,00 | 0,85 | |
| 2 | I | 43 | 5 | 0,90 | 0,52 | 0,66 | 0,85 |
| | V | 40 | 212 | 0,84 | 0,98 | 0,90 | |
| 3 | I | 9 | 6 | 0,60 | 0,90 | 0,72 | 0,98 |
| | V | 1 | 284 | 1,00 | 0,98 | 0,99 | |
| 4 | I | 11 | 11 | 0,50 | 0,65 | 0,56 | 0,94 |
| | V | 6 | 273 | 0,98 | 0,96 | 0,97 | |
| 5 | I | 53 | 32 | 0,62 | 0,74 | 0,68 | 0,83 |
| | V | 19 | 195 | 0,91 | 0,86 | 0,88 | |
| Geral | I | 329 | 53 | 0,86 | 0,79 | 0,82 | 0,91 |
| | V | 89 | 1029 | 0,92 | 0,95 | 0,94 | |

Fonte: Da autora, 2021.

5.1.4. Discussão

O aprendizado das RNA's é realizado por meio de processos iterativos de ajustes aplicados aos pesos sinápticos, o chamado treinamento. O aprendizado só ocorre quando a rede neural atinge uma solução generalizada para um determinado problema (ZHANG *et al.*, 2020). A partir da análise de iterações (épocas) do conjunto de dados foi possível perceber que quanto maior o número de iterações utilizado no treinamento, maior a memorização dos dados pelas RNA's, demonstrando o caráter não generalista do sistema (*overfitting*). O que confirma a importância da definição de um número ótimo de iterações (FIGURA 2), indicando a confiabilidade e eficiência do método, além de possibilitar a economia de tempo na obtenção de resultados (NI *et al.*, 2020).

A partir das imagens escaneadas foi obtido um conjunto de coordenadas de caixas delimitadoras e imagens de máscaras das regiões das sementes. Comprovando assim, que a rede pode localizar e segmentar com precisão as sementes independentemente de sua orientação. As predições do modelo em relação às classes bounding boxes e máscaras apresentam números laterais que são relativos à probabilidade de identificação do objeto. Ao adicionar um rótulo de categoria ao conjunto de dados para determinar se as respectivas regiões são adequadas para análise, à rede neural pode adquirir a capacidade de classificação para discriminar tal integridade (TODA *et al.*, 2020).

A partir disso, quanto maior a taxa de detecção das sementes, maior a precisão e, conseqüentemente, recall, acurácia e F-Measure, indicando maior eficiência do modelo em classificar. Isso é corroborado pelos dados obtidos no presente estudo, uma vez que o percentual de detecção de sementes viáveis e inviáveis foi predominantemente superior a 70%. O que indica que o modelo treinado pode classificar as sementes com taxa de falso negativo muito baixa, provando assim, a alta capacidade do modelo (FA *et al.*, 2019). O mesmo também foi observado por Jeong *et al.* (2020), que ao utilizarem Mask R-CNN para a segmentação de sementes de arroz para fenotipagem justificaram a alta eficiência obtida da rede devido à elevada capacidade de detecção.

A eficácia em classificar resultados do teste de tetrazólio em partes de sementes de *Urochloa brizantha* cv. Marandu por análise computacional foi comprovada através dos resultados visuais e a avaliação da detecção de objetos por Mask R-CNN. Uma vez que a classificação a partir da análise de imagem foi bem próxima a classificação visual e também por apresentar altos valores das métricas de avaliação do modelo. Corroborando com Liu *et al.* (2019) que obtiveram altos valores desses parâmetros para identificar frutos de pepino em estufas, observando-se valores superiores a 80% ao utilizar a metodologia Mask R-CNN. Essas métricas são os parâmetros usuais para avaliação do desempenho de modelos classificadores, que demonstram o quanto as amostras foram classificadas corretamente.

A alta eficiência obtida pelo sistema desenvolvido pode ser justificada pela ausência de sobreposição dos objetos, permitindo uma maior precisão na identificação (TODA *et al.*, 2020). Além disso, outro fator que pode ter influenciado nos valores das métricas é a nitidez de coloração das sementes. Levando em consideração que o fundamento do teste de tetrazólio, bem como da análise de imagens, baseia-se nas características de coloração. Assim, quanto mais fácil a distinção da matriz RGB da imagem, maior a chance de sucesso na detecção das classes avaliadas (JEONG *et al.*, 2020), sendo que as sementes inviáveis ou mortas podem ser identificadas com maior facilidade.

Como observado no lote 1, onde a classificação visual permitiu identificar uma maior taxa de sementes inviáveis que os demais lotes. Assim, ao comparar os resultados obtidos pelo sistema para esse lote, foram obtidos valores das métricas superiores (TABELA 2) quando identificou a classe inviável. Nesse sentido, o poder de discriminação bem como o alto rendimento das redes, está intimamente relacionado à qualidade e a quantidade de amostras utilizadas. Esse fato permite a utilização da análise de imagens para sementes de diversas culturas, como a soja (BAEK *et al.*, 2020). Onde os autores identificaram que o sucesso do sistema classificador depende da sua capacidade de lidar com amostras semelhantes e numerosas.

Por outro lado, o fato dos lotes 3 e 4 terem apresentado maiores métricas ao classificar sementes viáveis (TABELA 2), pode estar associado à qualidade das sementes do lote. Uma vez que os mesmos apresentaram maior taxa de viabilidade através da classificação visual. Com isso, o modelo apresentou alta performance ao indicar sementes viáveis e inviáveis, tanto de forma geral bem como por lote. Com isso, pode-se inferir que o sistema desenvolvido mostra-se eficiente na classificação do teste de tetrazólio em sementes de *Urochloa brizantha* cv. Marandu, tornando a interpretação do teste mais rápida, precisa e menos subjetiva. Além de apresentar alto potencial de uso para análise de sementes em outras espécies e representar uma das contribuições da revolução 4.0 na tecnologia de análise de sementes.

5.1.5. Conclusão

O modelo Mask R-CNN foi eficiente na localização e classificação das sementes de *Urochloa brizantha* cv. Marandu, como viáveis ou inviáveis. Apresentando alto desempenho em todos os lotes, uma vez que os valores de acurácia foram superiores a 80%, com acurácia geral em torno de 91%. Desta forma, a técnica é uma ferramenta útil e eficiente para a automação computacional do teste de tetrazólio em sementes de *Urochloa brizantha* cv. Marandu, reduzindo os problemas de subjetividade, tornando a interpretação do teste mais rápida e precisa. Além de representar uma das contribuições da revolução 4.0 para a tecnologia de análise de sementes.

5.1.6. Agradecimentos

Agradecemos a CAPES (Conselho de Aperfeiçoamento de Pessoal de Nível Superior – código 001), à FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) e ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pelo apoio a este estudo.

5.1.7. Referências

ABBADE, L. C.; TAKAKI, M. Teste de tetrazólio para avaliação da qualidade de sementes de *Tabebuia roseoalba* (Ridl.) Sandwith - Bignoniaceae, submetidas ao armazenamento. **Revista Árvore**, v. 38, n. 2, p. 233-240, 2014.

BAEK, J., LEE, E., KIM, N., KIM, S. L., CHOI, I., JI, H., ..& KIM, K. H. High throughput phenotyping for various traits on soybean seeds using image analysis. **Sensors**, 20(1), 248, 2020.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Regras para análise de sementes**. Brasília: MAPA/ACS, 2009. 399 p.

CUSTÓDIO, C. C.; DAMASCENO, R. L.; MACHADO NETO, N. B. Imagens digitalizadas na interpretação do teste de tetrazólio em sementes de *Brachiaria brizantha*. **Revista brasileira de sementes** vol.34 no.2 Londrina June, 2012.

FA, KHAN, A. GUMAEI, A. DERHAB E A. HUSSAIN. "A Novel Two-Stage Deep Learning Model for Efficient Network Intrusion Detection," in **IEEE Access** , vol. 7, pp. 30373-30385, 2019 .

GITHUB FOR TEAMS, 2021. COCOSYNTH: **Guia completo para a criação de conjuntos de dados COCO**. Disponível em: < <https://github.com/akTwelve/cocosynth>>. Acesso em: 25 de jan. de 2021.

GITHUB FOR TEAMS, 2021. **Máscara R-CNN para detecção e segmentação de objetos**. Disponível em: < https://github.com/matterport/Mask_RCNN/wiki>. Acesso em: 25 de jan. de 2021.

HE, K., ZHANG, X., REN, S. E SUN, J. Deep Residual Learning for Image Recognition. Proceedings of the **IEEE Conference on Computer Vision and Pattern Recognition**, 770-778, 2016.

HE, K.; GKIOXARI, G.; DOLLÁR, P.; & GIRSHICK, R. Mask r-cnn. In **Proceedings of the IEEE international conference on computer vision**, pp. 2961-2969, 2017..

HE, KAIMING, GKIOXARI, G., DOLLÁR, P., & GIRSHICK, R. Máscara r-cnn. In **Proceedings of the IEEE International Conference on Computer Vision** ,pp. 2961-2969, 2018.

JEONG, YS, LEE, HR, BAEK, JH, KIM, KH, CHUNG, YS, & LEE, CW. Segmentação de sementes de arroz baseada em aprendizado profundo para finotipagem. **Journal of the Korea Industrial Information Systems Research** , 25 (5), 23-29, 2020.

LANDAU, ELENA CHARLOTTE; RESENDE, RMS; MATOS NETO, F. da C. 2020. Evolução da área ocupada por pastagens. **Embrapa Milho e Sorgo-Capítulo em livro científico (ALICE)**.

LIU, X., ZHAO, D., JIA, W., JI, W., RUAN, C., & SUN, Y. Cucumber fruits detection in greenhouses based on instance segmentation. **IEEE Access**, 7, 139635-139642, 2019.

NI, X.; LI, C.; JIANG, H.; & TAKEDA, F. Deep learning image segmentation and extraction of blueberry fruit traits associated with harvestability and yield. **Horticulture Research**, 7(1), 1-14, 2020..

SHAHEEN, F.; VERMA, B. E ASAFUDDOULA, M. Impact of automatic feature extraction in deep learning architecture. In **Digital Image Computing: Techniques and Applications (DICTA)**, International Conference on, pp. 1–8. IEEE, 2016.

TODA, Y.; OKURA, F.; ITO, J. et al. Rede neural de segmentação de instância de treinamento com conjuntos de dados sintéticos para fenotipagem de sementes de culturas. **Commun Biol.** 3, 173, 2020.

ZHANG, J.; YAN, Y.; CHENG, Z .; WANG, W. Rede de pirâmide de atenção leve para detecção de objetos e segmentação de instâncias. **Appl. Sci.** 10, 883, 2020.

5.2. Artigo 2- *Deep learning* na predição de pureza física de sementes de *Urochloa brizantha* cv. Marandu

Este artigo foi elaborado conforme normas do Programa de Pós-Graduação em Produção Vegetal.

Resumo

Visão computacional associado à Mask R-CNN constitui uma tecnologia que pode ser utilizada para automação e aprimoramento de diversas avaliações rotineiras, como a análise de pureza física de sementes. A análise de pureza é uma avaliação obrigatória para comercialização de sementes e determina a qualidade física de lotes. Com isso, objetivou-se verificar a viabilidade de *deep learning* na predição da pureza de lotes de sementes de *Urochloa brizantha* cv. Marandu. Foram utilizados quatro lotes de sementes de *Brachiaria* provenientes de diferentes locais. Para obtenção das imagens foi utilizado um escâner de mesa (Hp Officejet 4500 Desktop), com definição de 1200 DPI, e contendo na sua parte superior um fundo azul. Para a automação da análise de pureza foram escaneadas cinco proporções de pureza, sendo que cada proporção possuía duas amostras de 10 gramas cada uma. A automatização da classificação dos resultados foi realizada pela metodologia Mask R-CNN. A rede obteve eficiência na estimação da pureza dos lotes avaliados, com altos coeficientes de determinação, sendo possível ajustar um modelo de regressão para predição de pureza. Assim, foi possível automatizar por meio de regressão e da técnica Mask R-CNN a análise de pureza em sementes de *Urochloa brizantha* cv. Marandu.

PALAVRAS-CHAVE: Data Science. Fenotipagem de sementes. Mask R-CNN. Classificação.

5.2.1. Introdução

A análise de pureza é um teste obrigatório para determinar a pureza de sementes comerciais, estabelecido pela Lei de Sementes, Lei 10.711, agosto de 2003 (BRASIL, 2003). Uma vez que define a qualidade física de lotes de sementes que serão utilizados e torna possível a previsão de perdas no campo. Já que o uso de lotes com baixo percentual de pureza, quando semeados, resultam na formação de uma lavoura não homogênea gerando prejuízos na produtividade final (BATISTA *et al.*, 2016).

Apesar de ser uma análise obrigatória e indicar percentuais de pureza a serem seguidos para cada espécie, é comum a comercialização de lotes de sementes fora dos padrões de pureza estabelecidos, principalmente no mercado de sementes de forrageiras, como a *Urochloa brizantha* cv. Marandu. Além disso, apesar de simples é uma determinação que pode se tornar complexa porque exige equipamentos com certa precisão e complexidade. Com isso, é fundamental o desenvolvimento de pesquisas sobre a aplicabilidade de tecnologias, como o *deep learning*, a fim de aprimorar essa avaliação, tornando-a mais prática, eficiente e rápida para assim facilitar a fiscalização de lotes comerciais de sementes.

Deep learning permite que modelos computacionais compostos de várias camadas de processamento aprendam representações de dados com vários níveis de abstração (Lecun *et al.*, 2015), possibilitando o reconhecimento, a manipulação e análise de objetos. Dentre esses modelos, tem-se à rede Mask R-CNN, que é um método que além de detectar os objetos e classificá-los com a utilização de bounding box, também faz a segmentação por instâncias desses objetos detectados (HE *et al.*, 2017).

A análise de imagem associada à Mask R-CNN constitui em uma tecnologia que pode ser utilizada para automação e aprimoramento de diversos processos, e em diversas áreas da agricultura. Como na tecnologia de sementes para fenotipagem de sementes de soja (BAEK *et al.*, 2020); análise do tamanho da semente para descoberta de mutantes (HERRIDGE *et al.*, 2011); fenotipagem de alto rendimento para medir a forma de sementes (TANABATA *et al.*, 2012); determinação dos parâmetros físicos ideais para a seleção de sementes de milho (WEN *et al.*, 2015); entre outras finalidades.

Embora a literatura apresente vários estudos sobre a aplicação dessa tecnologia na área de sementes, ainda existem avaliações importantes que não exploraram a utilização dessa técnica, como a análise de pureza de sementes. Com isso, objetivou-se verificar a viabilidade de *deep learning* na predição da pureza de lotes de sementes de *Urochloa brizantha* cv. Marandu.

5.2.2. Material e métodos

O experimento foi conduzido na cidade de Montes Claros-MG (coordenadas: 16°40'58.16"S e 43°50'20.15" O), no período de junho de 2020 a janeiro de 2021. Foram utilizados quatro lotes de sementes de *Urochloa brizantha* cv. Marandu das safras 2017/2018 e 2018/2019, provenientes de diferentes locais comerciais. Os mesmos apresentavam diferentes porcentagens de pureza, sendo o lote um com 17,43; lote dois com 21,63; lote três com 54,73 e lote quatro com 9,85.

Para obtenção das amostras de trabalho, para cada um dos lotes utilizados foram separadas manualmente as sementes das impurezas. Posteriormente, estas frações foram misturadas considerando a pesagem em balança digital, nas proporções de 0% (10 g apenas de materiais inertes), 25% (2,5 g de

sementes e 7,5 g de materiais inertes), 50% (5 g de sementes e 5 g de materiais inertes), 75% (7,5 g de sementes e 2,5 g de materiais inertes) e 100% (10 g apenas de sementes) de pureza. Para cada uma das proporções de cada um dos lotes foram obtidas duas amostras diferentes de 10 g, cada uma, que é o peso mínimo exigido para amostras da espécie em estudo (BRASIL, 2009).

As amostras de sementes de cada proporção foram digitalizadas com auxílio de um escâner de mesa (Hp Officejet 4500 Desktop), contendo em sua parte superior uma folha de E.V.A da cor azul. Para isso, cada amostra foi espalhada aleatoriamente sobre a placa vidro do escâner, evitando sobreposições. Após ser digitalizada, a amostra foi misturada e escaneada novamente para obtenção de outra imagem. Esse processo foi repetido até se obter cinco imagens para cada uma das duas amostras de cada proporção, totalizando 200 imagens (quatro lotes x cinco proporções x duas amostras x cinco imagens) com definição de 1200 pontos por polegada quadrada (DPI).

Para o treinamento das redes, em cada um dos lotes foram escaneadas 100 sementes escolhidas aleatoriamente. A partir das quais se criou um banco sintético de imagens. Para isso, imagens individuais de cada uma das sementes foram isoladas com o auxílio do software R. Nesta etapa, a fim de possibilitar a análise computacional em um menor tempo, foi feita a compactação das imagens com o auxílio das funções *image_read* e *image_scale* do pacote *magick*. As imagens foram compactadas para 400 pixels de largura, 314 pixels de altura. Após a compactação das imagens, as mesmas foram abertas pela função *readImage* do pacote *EBImage*, a fim de obter para cada imagem três matrizes, cada uma com 400 colunas e 314 linhas contendo os valores de vermelho, verde e azul (RGB), com valores variando entre 0 e 1.

Para a segmentação de cada imagem, com a finalidade de separar a semente do fundo, foi utilizada apenas o canal de vermelho, para a qual utilizou-se um limiar por meio da função *otsu*. Assim, todos os pixels com valor abaixo deste limiar foram considerados como fundo (o qual era de cor azul) atribuindo-se o valor 1, os demais pixels, correspondente às sementes, atribuiu-se o valor 0.

A fim de preencher possíveis espaços vazios nos pixels correspondentes à semente foi utilizada a função *fillHull*. Essa matriz foi utilizada como coordenadas das sementes. A partir destas coordenadas, os valores de vermelho, verde e azul de cada pixel foram utilizados para a obtenção de uma imagem com extensão "png" contendo fundo transparente para cada um dos fragmentos de sementes. Para isso, utilizou-se a função *pxflood* do package *imager* e *writelImage* do pacote *EBImage*.

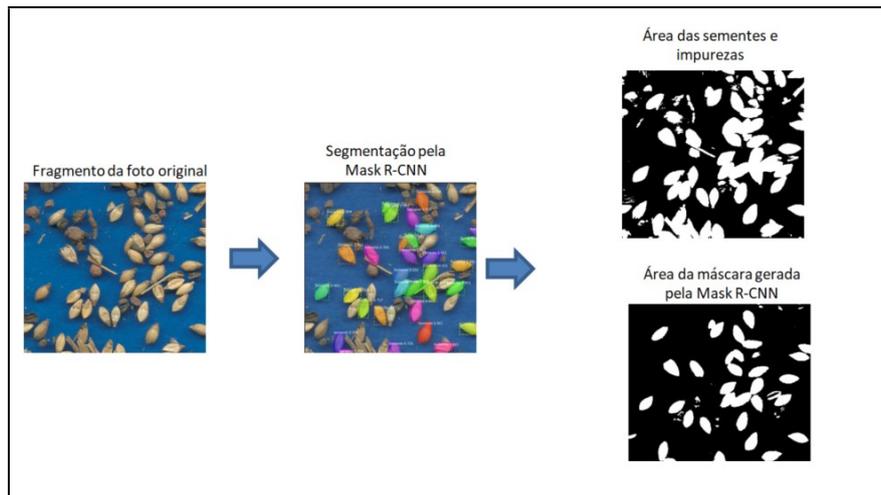
A partir destas imagens, foi criado um banco sintético com oito mil imagens de treinamento e duas mil imagens de validação. Para isso criaram-se imagens de tamanho de 320 x 320 (DPI). Como background foi obtida uma imagem digitalizada contendo apenas impurezas de cada um dos lotes. Sobre este background foram sobrepostas as imagens da sementes. Nesta composição, um valor aleatório de 1 a 15 imagens das sementes foram selecionadas computacionalmente, também de forma aleatória. O posicionamento destas sementes dentro da imagem sintética tanto quanto a sua rotação também foi estabelecida de forma aleatória. Todo este processo foi implementado com o auxílio da biblioteca *cocosynth* (GITHUB, 2021).

Foi utilizado a implementação via Mask R-CNN na plataforma online do google colab (GITHUB, 2021). Foram utilizadas as configurações pré-definidas pelo repositório, incluindo as arquiteturas de rede e as funções de perda. A rede ResNet101 (He *et al.*, 2016) foi utilizada para a extração de características. A partir dos pesos iniciais da ResNet101 obtido pelo treinamento usando o banco de dados

MSCOCO, foi realizado o ajuste fino utilizando nosso banco sintético de imagens. Foram utilizadas 80 épocas de treinamento com otimização do gradiente estocástico descendente com a taxa de aprendizado de 0.001 e tamanho do batch de 2. A técnica de “image augmentation” não foi feita durante o treinamento. Os dados de treinamento sintéticos tiveram seu tamanho fixado em 320x320 DPI. A saída das redes para o conjunto de caixas delimitadoras (bounding boxes) e máscaras (mask regions) foi obtido considerando o valor de probabilidade limiar de 0.70.

Após o treinamento da rede, o teste foi feito com todas as imagens obtidas das cinco proporções (200 imagens). Para cada uma dessas imagens foi estimada a porcentagem de pixels correspondente à máscara (identificado como sementes pela MASK R-CNN) em relação ao número de pixels pertencentes às sementes e impurezas (FIGURA 1). A fim de encontrar uma associação entre a porcentagem de pureza das sementes e a porcentagem de pixels ocupada pelas sementes (PPOS) foram ajustadas análises de regressão. Estas análises de regressão foram feitas considerando os dados em nível de imagem (10 g) em nível de média de 5 imagens (50g) e em nível de lote (100g). Foram ajustadas regressões lineares e quadráticas em cada uma das situações, selecionando-se o modelo de acordo com a presença de significância dos coeficientes de regressão.

Figura 1- Detecção pelo estabelecimento de máscaras em imagens utilizadas para associação da porcentagem de pixels ocupada pelas sementes e a porcentagem de pureza das sementes.

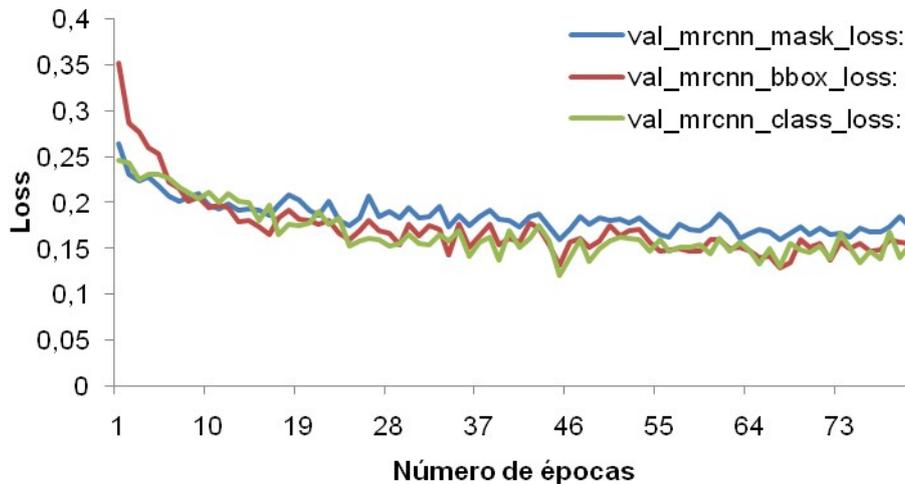


Fonte: Da autora, 2021.

5.2.3. Resultados

A partir do conjunto de dados, foram obtidas curvas da função de perda por iteração na identificação das sementes para o grupo de validação (FIGURA 2). Realizou-se o acompanhamento da função de perda ao decorrer das épocas de forma que quando se utilizou mais de 45 épocas teve-se aumento da taxa de erro na classificação da amostra de validação (Class Loss Val) indicando que o aumento no número de épocas de treinamento é inviável, provavelmente por consequência do *overfitting*.

Figura 2- Curvas da função de perda por iteração na classificação de sementes de *Urochloa brizantha* cv. Marandu.

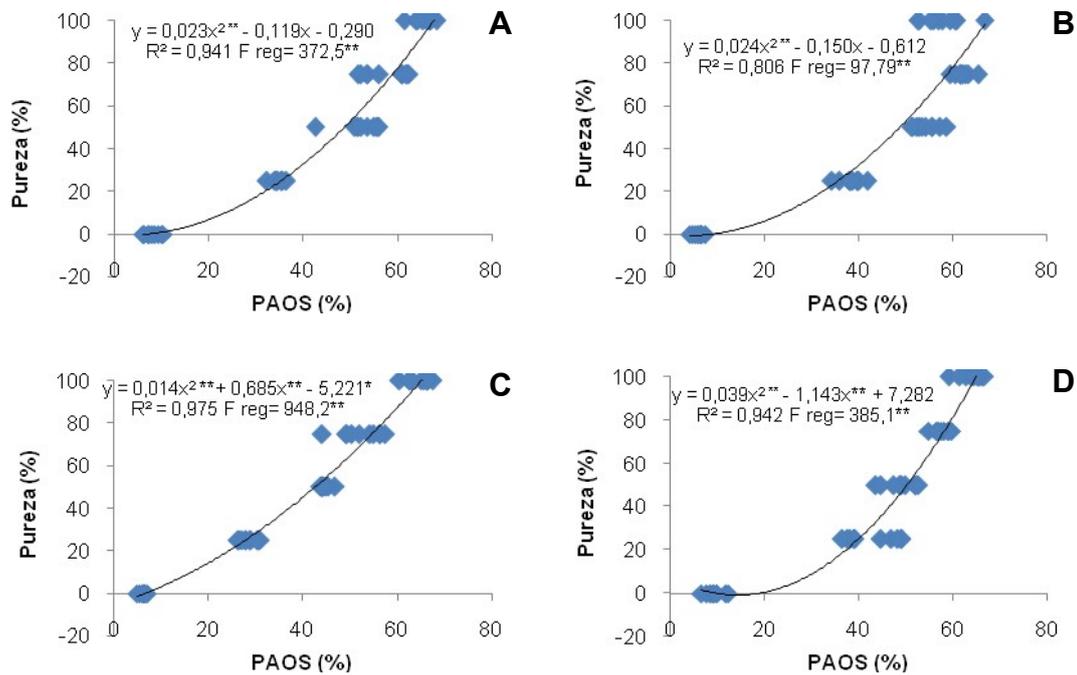


Fonte: Da autora, 2021.

De forma geral, constatou-se que a partir da porcentagem da área ocupada pelas sementes (POAS) foi possível estimar a pureza dos quatro lotes de sementes de *Urochloa brizantha* cv. Marandu avaliados, com a utilização da técnica Mask R-CNN e regressão. Na qual, realizou-se associação entre a porcentagem da área ocupada pelas sementes e a porcentagem de pureza, sendo que, com aumento da área ocupada pelas sementes aumentou-se também o grau de pureza das amostras.

Na avaliação de amostras de 10 gramas (FIGURA 3), a rede apresentou alta eficiência ao prever porcentagens de pureza. Nos quatro lotes avaliados, os coeficientes de regressão foram significativos, sendo possível a obtenção de altos valores de coeficiente de determinação, variando entre 0,80 e 0,97%.

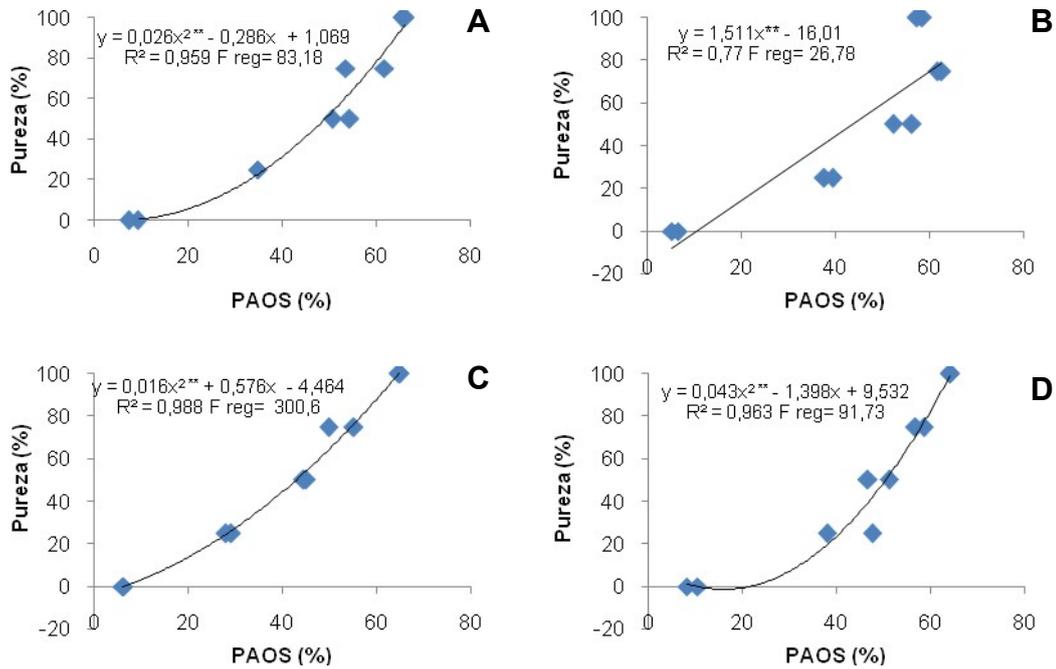
Figura 3- Predição da pureza em sementes de *Urochloa brizantha* cv. Marandu estimada por Mask R-CNN e regressão considerando uma amostra de 10 gramas de semente dos lotes 1 (A), 2 (B), 3 (C) e 4 (D).



Fonte: Da autora, 2021.

Em amostras com 50 gramas de sementes (FIGURA 4), a técnica também apresentou alta capacidade de detectar e discernir sementes de impurezas. Com a PAOS e a porcentagem de pureza apresentando interação não linear, exceto no lote dois. Nessas amostras também foram observados altos coeficientes de determinação com valores acima de 0,77%, e regressão significativa nos lotes.

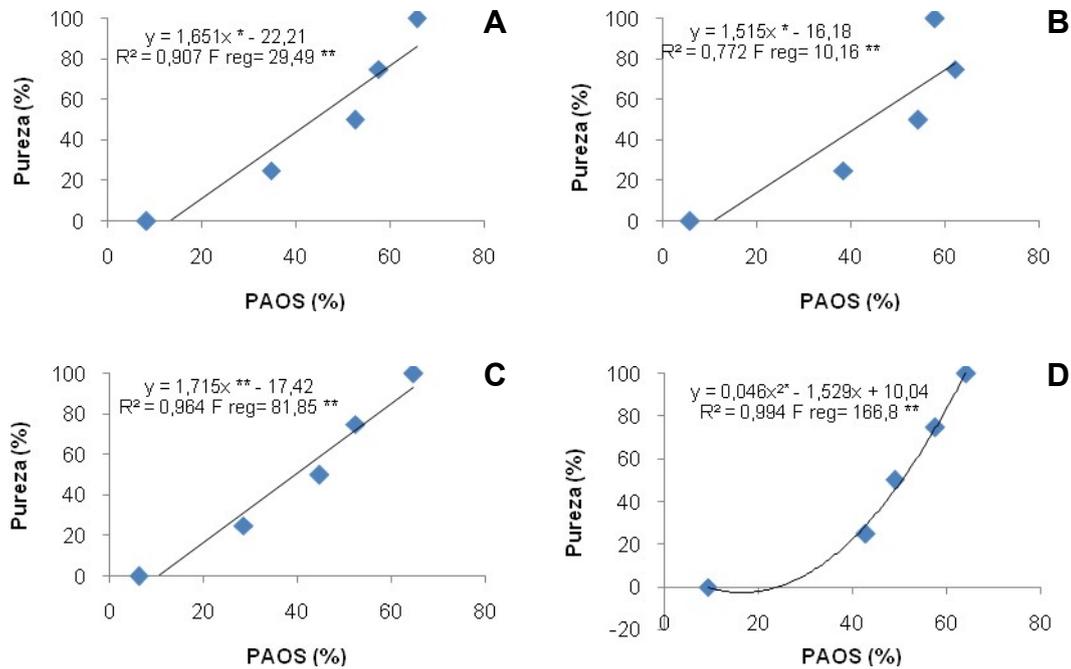
Figura 4- Predição da pureza em sementes de *Urochloa brizantha* cv. Marandu estimada por Mask R-CNN e regressão considerando uma amostra de 50 gramas de semente dos lotes 1 (A), 2 (B), 3 (C) e 4 (D).



Fonte: Da autora, 2021.

O alto desempenho da técnica em prever a pureza também foi comprovado ao estimar a porcentagem de pureza no lote inteiro, com amostras de 100 gramas (FIGURA 5). Em todos os lotes a regressão foi significativa, porém diferente dos casos anteriores, nas amostras de 100 gramas a regressão linear explicou melhor a relação entre a PAOS e a porcentagem de pureza, com exceção do lote 4 em que a regressão foi quadrática. Ainda sobre o lote 4, observou-se que com o aumento da massa das amostras teve-se também o aumento do coeficiente de determinação, sendo que em amostras de 100 gramas esse coeficiente foi quase perfeito (0,99%).

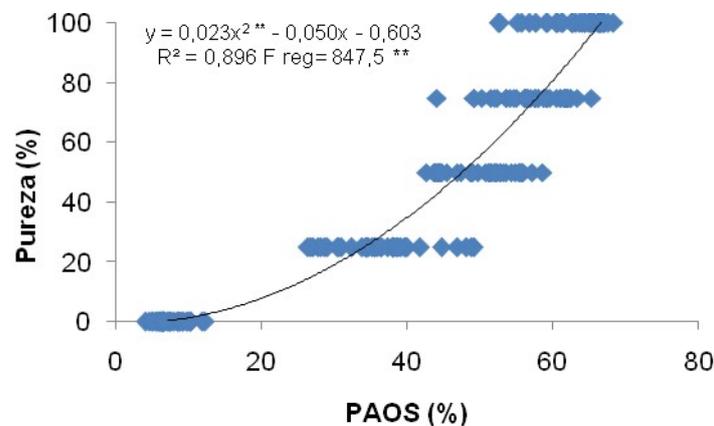
Figura 5- Predição da pureza em sementes de *Urochloa brizantha* cv. Marandu estimada por Mask R-CNN e regressão considerando uma amostra de 100 gramas de semente dos lotes 1 (A), 2 (B), 3 (C) e 4 (D).



Fonte: Da autora, 2021.

Para melhor visualização da generalização na predição da pureza pela Mask R-CNN são apresentados na Figura 6 os valores estimados de pureza em função da porcentagem de PAOS considerando simultaneamente todos os lotes. Na qual, o modelo ajustado (regressão e Mask R-CNN) apresentou bom desempenho, com coeficiente de determinação alto (0,89%) e regressão significativa a 5%.

Figura 6- Modelo ajustado para a predição da pureza a partir do PAOS em sementes de *Urochloa brizantha* cv. Marandu



Fonte: Da autora, 2021.

5.2.4. Discussão

É fundamental que ao desenvolver um modelo classificador, para automação da análise de pureza física de sementes, que ele apresente capacidade de trabalhar com amostras de diferentes tamanhos. A metodologia desenvolvida no presente estudo atendeu essa exigência, pois, para as amostras estudadas (10 g, 50g e 100g), demonstrou alta eficiência, sendo capaz de estimar de forma assertiva e com altos coeficientes de determinação, a pureza de amostras (FIGURA 3, FIGURA 4 e FIGURA 5). Demonstrando a possibilidade de utilização de amostras pequenas sem perdas significativas no resultado da pureza.

Essa capacidade generalista demonstrada pela rede pode estar relacionada à fase de treinamento ocorrida durante o processo de aprendizagem profundo (FIGURA 2). As diferentes arquiteturas das redes são capazes de alcançar resultados precisos com poucos parâmetros e consiste de um único conjunto de operações, tornando mais fácil de implementar, treinar e aplicar na prática e se adaptar automaticamente a diferentes problemas (NI *et al.*, 2020). Uma vez treinadas de forma adequada, possuem alta precisão e podem ser transferidas para outros domínios relacionados. Assim, essa precisão na fase de treinamento impede que os *outliers* obtidos nas amostras influenciem na capacidade discriminativa da rede (POUND *et al.*, 2017).

As diferenças observadas de desempenho da técnica entre os lotes e as amostras, podem ser justificadas pela ocorrência de adensamento das sementes ou impurezas, o que afeta a capacidade de discriminação da rede (TODA *et al.*, 2020). Já para o modelo ajustado para predição da pureza (FIGURA 6), foram obtidos altos coeficientes de determinação, indicando a boa qualidade do ajuste. Assim, quanto maior a porcentagem de área ocupada pela semente na foto, maior será o grau de pureza naquele lote, já que 0,90% da pureza do lote é explicada pela PAOS. Utilizando a Mask R-CNN para fins de fenotipagem de sementes de cevada, similares às da *Urochloa brizantha* cv. Marandu, Toda *et al.* (2020) também observaram alta eficiência ($R^2=0,97\%$), o que pode ser explicado pela robustez do modelo e pela alta capacidade da rede em detectar e discriminar os atributos referentes às sementes.

A proximidade das purezas preditas pela rede com a pureza real comprovou o alto desempenho da metodologia desenvolvida para estimar a pureza dos lotes de sementes de *Urochloa brizantha* cv. Marandu. Nesse sentido, pode-se inferir que a metodologia desenvolvida possui alto potencial para uso na determinação de pureza de lotes de sementes de *Urochloa brizantha* cv. Marandu. O que torna a análise da qualidade física dos lotes mais prática, precisa e rápida. Facilitando a obtenção de informações mesmo sem os equipamentos complexos e necessários para realização das análises, além da possibilidade de criação de bancos de dados com informações de pureza dos lotes. A Mask R-CNN, apresenta ainda, potencialidade para ser associada em outros tipos de tecnologias, como em imagens multiespectrais, possibilitando assim maior avanço tecnológico da análise de pureza física de sementes.

5.2.5. Conclusão

O modelo Mask R-CNN foi eficiente na localização e segmentação das sementes de *Urochloa brizantha* cv. Marandu, obtendo alta performance em estimar a pureza nos lotes. Com a pureza predita bem próxima da real, uma vez que os coeficientes de determinação foram altos, variando entre 0,77 e 0,99%.

Desta forma, essa técnica em conjunto com a análise de regressão mostrou ser uma ferramenta confiável e prática para a automação computacional da análise de pureza física em sementes de *Urochloa brizantha* cv. Marandu, e que apresenta a possibilidade de utilização para análise de sementes também outras espécies e de ser associada a outros tipos de imagens.

5.2.6. Referências

BAEK, J., LEE, E., KIM, N., KIM, S. L., CHOI, I., JI, H., ..& KIM, K. H. High throughput phenotyping for various traits on soybean seeds using image analysis. **Sensors**, 20(1), 248, 2020.

BATISTA, V. T.; NUNES, J. V. D.; NÓBREGA, L. H. P. Qualidade fisiológica de sementes de *Brachiariabrizantha* cv. marandu com variação nas características de pureza. **Revista de Agricultura**, Piracicaba, v. 91, n. 1, p.92-100, 2016.

BRASIL. Lei nº 10.711, de 5 de agosto de 2003. Dispõe sobre o Sistema Nacional de Sementes e Muda e dá outras providências. **Diário Oficial da União**: seção1, Brasília, DF, 06 ago. 2003.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Regras para análise de sementes**. Brasília: MAPA/ACS. 399 p, 2009.

GITHUB FOR TEAMS, 2021. COCOSYNTH: **Guia completo para a criação de conjuntos de dados COCO**. Disponível em: < <https://github.com/akTwelve/cocosynth>>. Acesso em: 25 de jan. de 2021.

GITHUB FOR TEAMS, 2021. **Máscara R-CNN para detecção e segmentação de objetos**. Disponível em: < https://github.com/matterport/Mask_RCNN/wiki>. Acesso em: 25 de jan. de 2021.

HE, K., ZHANG, X., REN, S. E SUN, J. Deep Residual Learning for Image Recognition. Proceedings of the **IEEE Conference on Computer Vision and Pattern Recognition**, 770-778, 2016.

HE, K.; GKIOXARI, G.; DOLLÁR, P.; & GIRSHICK, R. Mask r-cnn. In **Proceedings of the IEEE international conference on computer vision**, pp. 2961-2969, 2017.

HERRIDGE, R.P.; DAY, R.C.; BALDWIN, S. & MACKNIGHT, R.C. Análise rápida do tamanho da semente em *Arabidopsis* para descoberta de mutantes e QTL. **PlantMethods** 7 , 3, 2011.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436-444, 2015.

NI, X.; LI, C.; JIANG, H.;& TAKEDA, F. Deep learning image segmentation and extraction of blueberry fruit traits associated with harvestability and yield. **Horticulture Research**, 7(1), 1-14, 2020.

POUND, M.P.; ATKINSON, J.A.; WELLS, D.M.; PRIDMORE, T.P.;& FRENCH, A.P. Aprendizado profundo para fenotipagem multitarefa de plantas. Em *Proceedings of the IEEE International Conference on Computer Vision Workshops*.pp. 2055-2063, 2017.

TANABATA, T., SHIBAYA, T., HORI, K., EBANA, K. & YANO, M. SmartGrain: software de fenotipagem de alto rendimento para medir a forma da semente por meio de análise de imagem. *PlantPhysiol.* **160** , 1871–1880, 2012.

TODA, Y., OKURA, F., ITO, J. ET AL. Rede neural de segmentação de instância de treinamento com conjuntos de dados sintéticos para fenotipagem de sementes de culturas. *CommunBiol* **3**, 173, 2020.

WEN, K.X.; XIE, Z.M.; YANG, L.M.& SUN, B.Q.A tecnologia de visão computacional determina os parâmetros físicos ideais para a seleção de sementes de milho Jindan 73. *SeedSci. Technol.* **43** , 62–70, 2015.

6. CONCLUSÕES GERAIS

O modelo Mask R-CNN apresenta alto desempenho na localização e classificação das sementes de *Urochloa brizantha* cv. Marandu como viáveis ou inviáveis. Desta forma, a técnica é uma ferramenta útil e eficiente para a automação computacional do teste de tetrazólio em sementes de *Urochloa brizantha* cv. Marandu, eliminando os problemas de subjetividade, e com alto potencial de uso para análise de sementes em outras espécies.

O modelo Mask-RCNN foi eficiente na localização e segmentação das sementes de *Urochloa brizantha* cv. Marandu, obtendo alta performance em estimar a pureza nos lotes. Desta forma, essa técnica em conjunto com a análise de regressão mostrou ser uma ferramenta confiável e prática para a automação computacional da análise de pureza em sementes de *Urochloa brizantha* cv. Marandu, e que apresenta a possibilidade de utilização para análise de sementes também de outras espécies.