

XVII SALÃO DE PESQUISA SETREM



www.setrem.com.br/saps

01 | 02 | 03
O U T U B R O



ANAIS 2019

ISSN 2177-8353
Série 001/2019

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
SETREM, RS, Brasil

Salão de Pesquisa SETREM (17. : 2019 out. 1-3: Três de Maio, RS)
[Anais do] XVII SAPS 2019 [recurso eletrônico]. – Três de
Maio : SETREM, 2019.
color.

ISSN 2177-8353

Acesso: saps.setrem.com.br

1. Ensino Superior – Pesquisa. 2. Educação básica e profissional –
Pesquisa. 3. Ciência - Exposições - Rio Grande do Sul. I. Título.

CDU 001.891(061.4)(816.5)

Bibliotecária responsável: Rosimere Teresinha Marx – CRB 10/1425

XVII SALÃO DE PESQUISA SETREM



www.setrem.com.br/saps



01 | 02 | 03
OUTUBRO



UMA REVISÃO SISTEMÁTICA DA LITERATURA SOBRE DEEP LEARNING NA AGRICULTURA

Autores: CLAUDIO ROBERTO SCHEER JUNIOR (claudioscheer16@gmail.com), DALVAN JAIR GRIEBLER (dalvangriebler@setrem.com.br), LARISSA CANEPPELE GUDER (lariguder10@gmail.com)

Orientadores:

Link: <https://eventoscientificosapi.setrem.com.br/i/081874>

UMA REVISÃO SISTEMÁTICA DA LITERATURA SOBRE *DEEP LEARNING* NA AGRICULTURA

RESUMO

Com o crescimento do poder computacional, algoritmos de *deep learning* atingiram bons resultados em inúmeras áreas. Agricultura é uma das áreas que estão buscando aplicar esses algoritmos nos mais variados domínios. Sabendo disso, este trabalho apresenta uma revisão sistemática da literatura sobre o uso de *deep learning* na agricultura. Foram considerados artigos de janeiro de 2012 até abril de 2019. Dos 819 artigos encontrados, 230 foram incluídos na revisão. Neste trabalho foram avaliadas as técnicas de *deep learning* utilizadas, culturas abordadas, bases de dados utilizadas, desafios relacionados à agricultura e *deep learning*, e entre outros *insights* importantes. Os resultados apresentados mostraram que *deep learning* é aplicada com sucesso em diversas áreas da agricultura. Na pecuária, por exemplo, a maioria dos resultados atingiram 95% de acurácia. No total, 47.2% dos artigos atingiram acurácia maior que 95%. Isso mostra que ainda há muito trabalho a ser feito na área de *deep learning* na agricultura. Sendo assim, as análises apresentadas neste trabalho são de extrema importância para dar suporte a novas pesquisas com uso de *deep learning* na agricultura, e também para definir as possíveis lacunas a serem exploradas.

Palavras-chave: Sistemas de informação. Deep learning. Agricultura.

1 INTRODUÇÃO

A agricultura é a base para a alimentação humana, e, de acordo com a EMBRAPA (2018), com o aumento populacional será necessário aumentar a produção de alimentos em 35% até 2030. É necessário encontrar maneiras de aumentar e otimizar a produção de alimentos sem aumentar a área de cultivo. Como uma maneira de visualizar o desenvolvimento de alguma área, revisões da literatura são conduzidas para sumarizar as pesquisas já realizadas, o que de acordo com Zobel (2004), ajuda a evitar o fenômeno de “trabalhando em silos” (tradução nossa).

A área de inteligência artificial ressurgiu com a evolução da computação. Entre os vários tipos de algoritmos que a área de inteligência artificial engloba, estão as redes neurais. Com o objetivo de realizar o processamento das informações simulando os neurônios humanos, as redes neurais são formadas por camadas de “neurônios” ativados por modelos matemáticos. *Deep learning* é um o nome dado para uma rede neural com várias camadas de “neurônios”. Quanto maior a quantidade de neurônios simulados, maior o poder computacional necessário para o processamento dos cálculos. Sendo assim, o objetivo principal do trabalho foi conduzir uma revisão sistemática da literatura para a consolidação do estado da arte no uso de *deep learning* aplicado à agricultura.

Como contribuição científica, o presente trabalho trouxe uma discussão sobre o uso de *deep learning* na agricultura, baseando-se em 230 artigos para levantar as lacunas existentes e definir o estado da arte para área pesquisada. Além disso, este trabalho também apresenta uma representação gráfica dos desafios da agricultura mais abordados, associados com os objetivos de *deep learning*. Mais focado na parte computacional, o trabalho também contribuiu com uma análise cruzada das bases de dados utilizadas para testes e treinamento dos modelos de *deep learning*, métricas utilizadas para mensurar a performance dos modelos e das áreas da agricultura abordadas em cada artigo.

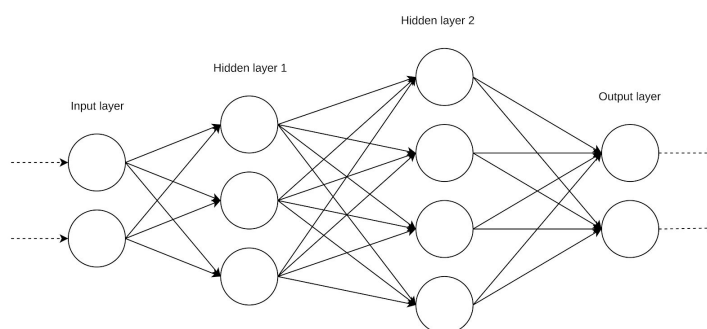
Este resumo está organizado da seguinte forma: na seção 2 são apresentados aspectos teóricos, na seção 3 é a apresentada a metodologia utilizada para a realização

do trabalho, na seção 4 são apresentados os principais resultados, e por fim na seção 5 estão a conclusão e sugestões de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Como já citado anteriormente, rede neural é uma área da inteligência artificial que busca simular os neurônios humanos através de modelos matemáticos. Na Figura 1 ilustra uma rede neural. De acordo com Russell e Norvig (2010), uma rede neural deve ser formada por no mínimo 3 camadas: uma camada de entrada dos dados (*Input layer*), uma camada escondida (*Hidden layer*) e a camada de saída (*Output layer*). A função da primeira camada é receber os dados que serão processados. Assim, com a camada escondida está a responsabilidade de processar a informação recebida da camada de entrada e também a função de repassar os dados processados para a camada de saída da rede neural. Cada neurônio tem um modelo matemático atrelado a ele, também chamado de função de ativação. Portanto, cada informação que passar pelo neurônio será processada de acordo com a função de ativação do neurônio.

Figura 1 - Estrutura de uma rede neural



Fonte: Adaptado de Russell e Norvig (2010).

A quantidade de camada escondidas (*Hidden layers*) que um modelo de rede neural tem, determina se o modelo tem um aprendizado profundo (ou, é uma *deep learning*). Em geral, considera-se que um modelo pode ser considerado de aprendizado profundo quando a quantidade de camadas escondidas for maior que 5.

3 METODOLOGIA

Para a condução da revisão sistemática da literatura, foi definido o protocolo de revisão seguindo os padrões propostos por Kitchenham e Charters (2007). As questões de pesquisa foram definidas como: **01)** Quais desafios da agricultura para *deep learning* foram abordados? **02)** Quais cultivares foram pesquisados ao redor do globo? **03)** Quais implementações (técnicas, modelos, arquiteturas e *frameworks*) de *deep learning* foram utilizadas? **04)** Quais bases de dados foram utilizadas e como foram utilizadas para o treinamento e teste da *deep learning*? **05)** Qual foi o resultado atingido pelos modelos de *deep learning* aplicados a agricultura?

Como estratégia de pesquisa foi adotada a pesquisa automatizada, onde foram consideradas as seguintes bases digitais: IEEE Xplore, ACM, Scopus e Science Direct. A *string* de busca foi definida como: ("*agricultural*" OR "*farming*" OR "*agriculture*" OR "*tillage*"

OR “cultivation” OR “tilling” OR “husbandry” OR “farm” OR “land management” OR “farm management” OR “horticulture” OR “agribusiness” OR “agronomy”) AND (“deep learning” OR “deep neural network”).

Para a classificação dos artigos, os seguintes critérios de inclusão foram considerados: (1) O artigo deve ser um estudo primário cujo foco principal é a utilização de algum modelo de *deep learning* aplicado à agricultura; (2) O artigo precisa ter explícito no texto o desafio da agricultura abordado, detalhes da implementação do modelo de *deep learning*, a métrica utilizada para avaliar o modelo e os resultados alcançados. Já os critérios de exclusão foram: (1) O artigo ter sido publicado antes de 2012; (2) O artigo não estar escrito em inglês; (3) Artigos não publicados; (4) Artigos duplicados ou publicados em dois locais distintos; (5) Artigos sem: autor, resumo ou DOI.

A seleção dos artigos ocorreu em três etapas: identificação, triagem e elegibilidade. Na identificação, a *string* de busca foi aplicada em cada base digital. Na etapa de triagem, foram avaliados os títulos e resumos dos artigos encontrados na etapa anterior. Por fim, na etapa de elegibilidade, o artigo completo foi avaliado.

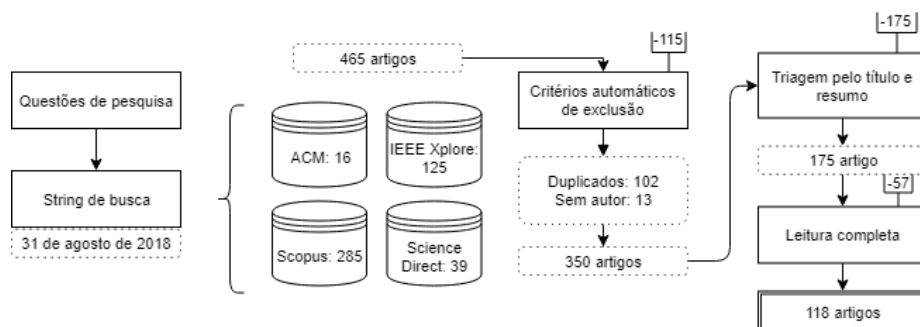
Os dados dos artigos foram extraídos considerando características da identificação do artigo, domínio da agricultura, modelo de *deep learning* e resultados. Os resultados foram tabulados para posterior avaliação. Os resultados foram apresentados através de gráficos, tabelas, modelos de representação e análises qualitativas. Conforme apresenta a Tabela 1, o processo de identificação dos artigos aconteceu em duas etapas, devido ao grande aumento de publicações. A primeira execução foi realizada no mês de agosto de 2018 e a segunda no mês de abril de 2019.

Tabela 1 - Processo de seleção dos artigos

Etapa	Identificação	Triagem	Elegibilidade	Total de artigos classificados
Agosto de 2018	465	350	175	118
Abril de 2019	770	276	155	112

A primeira (Figura 2) aconteceu no mês de agosto de 2018, conforme representado na Figura 1, onde 465 artigos foram encontrados. Após a aplicação dos critérios de exclusão, 102 artigos foram removidos por serem duplicados, e 13 por não apresentarem o autor. Na etapa de triagem, 175 foram removidos por não serem relacionados com a o escopo da pesquisa. Na etapa da elegibilidade, 57 artigos foram excluídos após a análise.

Figura 2 - Processo de execução da revisão sistemática da literatura



A segunda execução foi realizada no mês de abril de 2019, onde 770 artigos foram encontrados, das bases ACM (20), IEEE Xplore (245), Scopus (434), e Science Direct (71). No entanto, 58 foram removidos por serem duplicados, 20 não retornaram com o DOI, e 416 haviam sido considerados na etapa anterior. Na etapa de triagem, 121 foram removidos por não serem relacionados ao escopo da revisão. 155 artigos foram avaliados de forma completa, onde 43 foram removidos por não se enquadrarem nos critérios de inclusão.

4 APRESENTAÇÃO, ANÁLISE E DISCUSSÃO DOS RESULTADOS

É explícito o crescimento dessa área de pesquisa, conforme apresenta a Tabela 2. A quantidade de publicações no ano de 2018 foi superior ao de 2015. O que acontece devido aos bons resultados atingidos pelos pesquisadores.

Tabela 2 - Publicações por ano

Ano	2015	2016	2017	2018	abril - 2019
Total de publicações	05	14	49	123	39

Dentre os objetivos relacionados ao uso de modelos de *deep learning*, 11 objetivos distintos foram abordadas pelos autores, sendo eles: detecção, classificação, reconhecimento, identificação, *scouting*, segmentação, contagem, previsão, estimativa, predição e avaliação. Já os principais desafios da agricultura abordados são relacionados a classificação de culturas, de cobertura de terra, identificação de animais, detecção de doenças na planta, folha, grãos ou frutos, identificação de ervas daninhas e pestes.

De âmbito geral, as principais culturas abordadas são relacionadas a produção de grãos e a fruticultura. A cultura do milho foi a principal, com 32 artigos, a cultura do trigo foi abordada 24 vezes, do arroz 21, do tomate 18, da soja 14 vezes, de beterraba sacarina 13, uva 13, maçã 12, pastagens 10, morangos 8 vezes. Na pecuária, a suinocultura foi abordada 7 vezes, gado de corte 6 vezes, avicultura e ovinocultura foram abordadas duas vezes cada. Por fim, relacionado a aquicultura, foram dois artigos.

As áreas da agricultura que obtiveram os melhores resultados foram a pecuária (96% de acurácia média), e relacionados à alimentação animal (pastagens) (92.2% de acurácia média).

Tabela 3 - Modelos de *deep learning* mais utilizados

Modelo	CNN	LSTM	Faster R-CNN	SAE	DBN	RNN	FCN	Outros
%	74,4%	4,3%	4,3%	2,6%	2,1%	1,3%	1,3%	9,8%

Dentre os modelos de *deep learning* utilizados pelos autores, conforme apresenta a Tabela 3, as redes neurais convolucionais (CNN), foram as mais utilizadas, somando mais de 74%. Seguidos por LSTM e Faster R-CNN. Cabe ressaltar que o modelo Faster R-CNN, que contabilizou por 4,3% é uma aprimoração das redes convolucionais.

Ao total, no Brasil, foram classificados nove artigos. Entre as pesquisas desenvolvidas, a área da agricultura mais abordada foi a produção de grãos, onde 25%

dos artigos tiveram como foco as culturas de milho e soja, e 8,3% dos artigos tiveram como foco as culturas de café, algodão e a suinocultura.

5 CONCLUSÃO

Ao comparar os resultados encontrados, é possível ver que somente 47.2% dos artigos alcançaram uma acurácia maior que 95%. Isso não significa que estes modelos são ruins e não podem ser utilizados para solucionar algum problema, mas, sim, que ainda há muito trabalho para ser feito na área de *deep learning* para a agricultura.

Um dos maiores problemas das pesquisas já realizadas sobre as aplicações de *deep learning* na agricultura é que os resultados alcançados podem não ser os mesmos quando testados em outros países. Por exemplo, doenças, pestes, condições climáticas podem variar de acordo com a localização geográfica. Assim, é importante que o Brasil desenvolva suas próprias pesquisas sobre o uso de *deep learning* na agricultura, com foco nos desafios enfrentados pela agricultura no país ou em regiões específicas.

O presente trabalho deixa como sugestão para trabalhos futuros a (1) avaliação do tipo de imagem utilizada pelos autores na composição da base de dados usada para treinar o modelo de *deep learning*, agrupando de acordo ao problema abordado, (2) classificar as patologias encontradas de acordo com cada tipo de cultura e (3) aprimorar a revisão nos modelos de redes neurais convolucionais.

6 REFERÊNCIAS

EMBRAPA. **Visão 2030 - o futuro da agricultura brasileira**. Embrapa, p.212, 2018. Disponível em: <https://www.embrapa.br/documents/10180/9543845/Visão+2030+-+o+futuro+da+agricultura+brasileira/2a9a0f27-0ead-991a-8cbf-af8e89d62829>. Acessado em: 12 de Novembro de 2018.

KITCHENHAM, Barbara; CHARTERS, Stuart. **Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3**. 2007.

RUSSELL, Stuart.; NORVIG, Peter. **Artificial Intelligence: a modern approach**. 3rd.ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2010.

ZOBEL, Justin. **Writing for Computer Science**. London: Springer, 2004. 270p.