

Uma Nova Técnica de Reconhecimento Sintático de Padrões para Aplicações na Agropecuária

Gilberto Astolfi

Novembro de 2017

1 Antecedentes e Justificativas

O complexo agroindustrial da soja, de acordo com Hirakuri & Lazzarotto (2014), é de grande importância socioeconômica e gerador de riquezas, empregos e divisas para o país, figurando entre os principais vetores de desenvolvimento regional. Ele movimentava vários outros setores da economia, como por exemplo, empresas de pesquisa e desenvolvimento, fornecedores de insumos, indústrias de máquinas e equipamento. Sua importância econômica influenciou o crescimento do PIB (Produto Interno Bruto) do Brasil, que no primeiro trimestre de 2017 apresentou a primeira alta após 8 trimestres de queda. Tal aumento está diretamente relacionado ao crescimento de 13,4% do PIB da agropecuária nacional no mesmo período (CONAB, 2017).

O crescimento na produção de soja no Brasil, a base do complexo agroindustrial mencionado anteriormente, ocorreu devido a investimentos em tecnologias de produção que permitiram ampliar significativamente a área e a produtividade da oleaginosa (Hirakuri & Lazzarotto, 2014). No entanto, apesar da adoção de tecnologias como a dos transgênicos e utilização de defensivos agrícolas, o produtor ainda enfrenta problemas nos cultivos decorrentes da adaptação de doenças que atacam a cultura. Essa adversidade leva o produtor a buscar novas tecnologias para auxiliá-lo no monitoramento das condições de sua lavoura, permitindo a intervenção de maneira precoce, minimizando custos e maximizando a produção.

A tecnologia de visão computacional emerge como uma alternativa para auxiliar o produtor na manutenção dos resultados já alcançados, pois é uma técnica que processa e analisa imagens a fim de reproduzir em sistemas computacionais a capacidade da visão humana (Jahne & Haussecker,

2000). Essa tarefa está pautada em identificar, representar e reconhecer padrões (Bishop, 2006) e pode ser empregada na análise de imagens de lavouras de soja com o objetivo de identificar e classificar doenças.

Identificar e classificar doenças depende de conhecimento especializado e no âmbito computacional essa tarefa pode ser considerada ambígua e computacionalmente custosa. Uma das alternativas utilizadas para se aproximar e até ultrapassar a capacidade do especialista humano nesse tipo de tarefa, no que diz respeito a performance e acurácia, é o emprego de aprendizagem de máquina (Norvig & Russell, 2015).

A técnica de aprendizagem de máquina está relacionada ao fornecimento de informações de entrada para um sistema computacional, com o objetivo de treiná-lo para que possa tomar decisões futuras baseadas na experiência obtida no treinamento (Norvig & Russell, 2015). Por exemplo, um sistema computacional pode ser treinado com informações sobre doenças relacionadas a lavoura da soja para futuramente identificar e classificar tais doenças.

No entanto, um dos desafios de treinar um sistema baseado em visão computacional está relacionado a extração e representação de padrões contidos nas imagens (Ribas et al., 2014). Uma das abordagens adotadas para essa tarefa envolve usar modelos hierárquicos, que procuram representar um padrão como uma composição de sub-padrões, expressos por uma linguagem, e reconhecê-los por meio de análise sintática (de Souza Pio et al., 2006). Em suma, nessa abordagem os padrões em uma imagem são representados por cadeias de símbolos e a linguagem pelo conjunto dos símbolos usados para representar os padrões, como também, as regras para combinar os símbolos.

Nesse sentido, uma linguagem pode representar os padrões de uma doença em uma lavoura de soja e um sistema baseado em visão computacional pode ser treinado a reconhecer essa linguagem, aprendendo a gramática que a padroniza. Esse é o caminho que esse trabalho pretende trilhar, desenvolver uma nova técnica de reconhecimento sintático de padrões em imagens, que se baseia em aprendizagem de máquina e inferência gramatical (de la Higuera, 2010) para aplicações na agropecuária. Essa abordagem pretende explorar as vantagens que o reconhecimento sintático de padrões proporciona em relação aos demais métodos de representação e reconhecimento de padrões, que segundo Chanda & Dellaert (2004) está relacionado a forma de representação intuitivamente simples e elegante de padrões, como também, a habilidade em modelar e interpretar semântica, considerando relações espaciais e temporais. Li et al. (2017) também cita como vantagens a capacidade de flexibilidade e simplicidade na construção de diferentes tipos de estruturas

baseando-se em um dicionário de primitivas e regras de produção, e, a capacidade de representar com baixo custo dados de forma hierárquica e composicional.

Para tanto, pretende-se capturar imagens de lavouras de soja utilizando RPAs (Sistemas de Aeronaves Remotamente Pilotadas) e categorizá-las baseando-se nas características de cada doença encontrada na lavoura de soja, a fim de extrair e representar padrões por meio de representação sintática de padrão. Por fim, técnicas de aprendizagem e visão computacional serão utilizadas para identificação e classificação de doenças relacionadas a cultura da soja.

2 Objetivos

2.1 Objetivo Geral

Propor uma nova técnica de reconhecimento sintático de padrões em imagens baseado em aprendizagem de máquina e inferência gramatical para aplicações na agropecuária.

2.2 Objetivos Específicos

Para alcançar o objetivo geral apresentado na seção anterior, pretende-se atingir os seguintes objetivos específicos:

- Realizar uma revisão sistemática da literatura sobre reconhecimento sintáticos de padrões;
- Criar um banco de imagens de lavouras de soja utilizando RPAs; classificar as imagens para representar classes;
- Propor nova técnica de representação e reconhecimento de padrão baseado em métodos sintáticos, que possam extrair e representar as características mensuráveis de imagens e as relações estruturais entre elas, como também, inferir gramáticas a partir das informações extraídas dos padrões;
- Realizar experimentos e análise dos resultados obtidos por meio de métodos estatísticos;
- Publicar os resultados alcançados em eventos e revistas técnico-científicos.

3 Revisão da Literatura

3.1 Reconhecimento sintático de padrões

Os padrões são regularidades detectáveis que se repetem de maneira previsível e podem ser compreendidos como os meios pelos quais pode-se interpretar o meio ambiente ao nosso redor (Fu, 1982). O reconhecimento de padrões em objetos é uma tarefa trivial aos seres humanos, mas ao tentar realizar essa tarefa de forma artificial, ela se torna um problema complexo. No entanto, o campo de reconhecimento de padrões, como uma ciência, se concentra em descobrir regularidades significativas em dados por meio do uso de algoritmos a fim de utilizar tais regularidades para descrever ou classificar dados em diferentes categorias (Bishop, 2006).

As regularidades nos dados são representadas como características mensuráveis, que podem ser numéricas e/ou não numéricas, e suas relações. A composição formada por essas características, que se repetem regularmente nos dados, é utilizada para definir o padrão do conjunto de dados em questão (Fu & Rosenfeld, 1976). No entanto, não basta apenas definir um padrão representativo para os dados, há também a necessidade de uma análise criteriosa sobre o padrão a fim de desenvolver sistemas de reconhecimento.

Para que sistemas reconheçam padrões, os mesmos devem ser representados de maneira que possam ser interpretados por algoritmos. Existem várias abordagens de representação, como por exemplo, constelações (Fergus et al., 2003), estruturas de *quadrees* (Pedro, 2013) e grafos *And-Or* (Zhu L. Chen, 2009). No entanto, a representação que geralmente é utilizada é a de vetores de características e o reconhecimento de cada padrão normalmente é feito por comparações de similaridade (Fu & Rosenfeld, 1976; Bishop, 2006).

Por outro lado a representação do padrão em abordagens sintáticas é expressada como uma composição de sub-padrões, chamados de padrões primitivos ou primitivas (Fu & Rosenfeld, 1976). Assim, um padrão complexo é então expresso em termos do relacionamento entre suas primitivas e representado por meio de uma linguagem (de Souza Pio et al., 2006).

Em abordagens sintáticas o reconhecimento do padrão é normalmente feito por meio de *parsing* de acordo com um conjunto de regras de uma gramática (Fu & Rosenfeld, 1976), associando o padrão a uma classe. Para tal, o reconhecimento sintático modela uma descrição hierárquica do padrão, construído a partir de sub-padrões mais simples, sendo que no nível mais baixo se encontram os elementos mais simples, extraídos dos dados de entrada que são as primitivas. Logo,

o padrão é considerado uma sentença pertencente a uma linguagem, as primitivas como seu alfabeto e as sentenças geradas a partir de uma gramática, isto é, em abordagens sintáticas um padrão é definido por um número finito de primitivas e de regras gramaticais (de Souza Pio et al., 2006).

Pistori et al. (2013), por exemplo, propôs uma abordagem sintática de reconhecimento de padrões em problemas de classificação de imagens. No trabalho é apresentada uma estratégia para representar imagens por meio de cadeia de caracteres. As imagens são agrupadas em classes e cada classe é representada por uma gramática inferida a partir das cadeias de caracteres extraídas do conjunto de imagens da classe. Ribas et al. (2014) utiliza a mesma técnica de Pistori et al. (2013), no entanto, a principal contribuição desse trabalho está na comparação da construção do vocabulário de forma supervisionada e não supervisionada.

No trabalho de Walton et al. (2017) é proposto um método de representação e aprendizado de relacionamentos contextuais de objetos em uma cena usando Gramática Estocástica implementada como um grafo *And-Or*. A aplicação está inserida no domínio de compreensão de cenas no Comando e Controle Tático Naval. Compreensão de cena é um subcampo da área de visão computacional que se concentra em extrair entidades, ações e eventos de imagens e vídeos. O desafio nesse campo inclui extração de características, segmentação de imagem, classificação de objetos e adicionalmente a compreensão da cena, que se concentra no relacionamento semântico entre os objetos.

No método proposto por Walton et al. (2017) um grafo *And-Or* é construído por nós onde cada nó *Or* tem nós filhos correspondendo a sub-configurações alternativas e os nós *And* correspondem a decomposição dentro de componentes constituintes da cena. As arestas laterais permitem o grafo codificar as relações contextuais entre as entidades em todos os níveis hierárquico de uma cena. Os autores afirmam que Gramática Estocástica em forma de um grafo *And-Or* é adequado para analisar tarefas em cenas, mas, atualmente, inferir contextos de eventos no domínio de Comando e Controle Tático Naval, considerada uma tarefa crítica, é deixada para operadores humanos, que depende em grande parte da formação, experiência e intuição (Walton et al., 2017).

No que diz respeito ao reconhecimento de objetos específicos em cenas, o trabalho de Girshick et al. (2011) apresentam um modelo de gramática que permite detectar pessoas em imagens. O modelo representa pessoas usando uma hierarquia de partes. O treinamento para o modelo é baseado em SVM (Support Vector Machine) (Norvig & Russell, 2015) de forma supervisionada a partir de dados fracamente rotulados. A gramática proposta basicamente representa objetos

recursivamente compostos por outros objetos, consistindo de um conjunto de \mathcal{N} símbolos não terminais e um conjunto de \mathcal{T} símbolos terminais. Os símbolos terminais consistem de bloco básicos (partes de imagem) e os não terminais de símbolos abstratos. Além disso, a gramática é composta por um conjunto de Ω localizações na imagem e $\mathcal{Y}(\omega)$ especifica a localização de $\mathcal{Y} \in \mathcal{N} \cup \mathcal{T}$ em uma localização $\omega \in \Omega$

Similarmente ao trabalho de Girshick et al. (2011), Chua & Felzenszwalb (2016) consideram que objetos em cenas podem ser representados utilizando uma estrutura hierárquica definida por regras de composição. Essa estrutura pode descrever em detalhes uma cena identificando objetos e a relação entre eles. Além disso, as regras de composição fornecem evidências contextuais para inferência com dados ambíguos. Por exemplo, a presença de algumas partes de um rosto em uma cena fornece pistas contextuais para a presença de outras partes. Nesse contexto, os autores propõem uma abordagem centrada em Probabilistic Scene Grammar (PSG) usando *factor graphs* e *Loopy Belief Propagation (LBP)* para gerar cenas com múltiplos objetos usando regras de composição e relacionamento entre os objetos. A PSG consiste em (Chua, 2018):

- Gramática livre de contexto estocástica;
- Símbolos: face, olhos e boca;
- Pose no espaço: orientação, localização e escala;
- Regras de produção;
- Regras de produção probabilísticas;
- Relacionamento geométrico entre objetos e partes;
- Parâmetros;

A PSG define um conjunto de possíveis cenas e uma distribuição de probabilidade. As cenas são definidas usando uma biblioteca implícita de construção de blocos. Cada bloco é um par de tipo e pose. O tipo é um símbolo de um alfabeto finito e a pose é a posição do tipo no espaço. O relacionamento entre os blocos são definidos pelas regras de produção (Chua, 2018). O início de geração de uma cena é de forma randômica usando um conjunto de regras de produção da gramática. Cada um dos blocos iniciais são expandidos para gerar novos blocos, até que todos

blocos tenham sido expandidos formando a cena em sua totalidade. O resultado é um conjunto de blocos organizados de forma hierárquica (Chua, 2018).

Nesta seção foi apresentada uma breve revisão de trabalhos que adotam abordagem sintáticas. Neste plano de trabalho está definido na seção 2.2, como um dos objetivos específicos, uma revisão sistemática da literatura. Pretende-se, de acordo com o cronograma apresentado na seção 5, conduzir a revisão sistemática de acordo com planejamento estabelecido.

3.2 Aprendizagem supervisionada e não supervisionada

Aprendizagem é uma mudança adaptativa no comportamento causada por experiência e a capacidade de aprender deve ser parte de qualquer sistema computacional que reivindique inteligência (Luger, 2013).

No âmbito da inteligência artificial há algumas definições de aprendizagem, os exemplos são: aprendizagem não supervisionada, cujo objetivo está focado em descobrir, de maneira autônoma, grupos de exemplos similares dentro de conjunto de dados e agrupá-los (Norvig & Russell, 2015); aprendizagem supervisionada, que objetiva treinar um sistema computacional utilizando exemplos de entrada associados a saída desejadas (Norvig & Russell, 2015); aprendizagem semi-supervisionada, que consiste em utilizar exemplos de maneira supervisionada para se obter informações do problema a fim de utilizá-las para guiar o processo de aprendizado de maneira autônoma (Bruce, 2001) e; aprendizagem por reforço, que consiste em interferir com ações que afetam o treinamento, ocasionalmente, o agente que está sendo treinado recebe uma recompensa ou punição pela ação realizada (Norvig & Russell, 2015) .

A aprendizagem supervisionada, como já brevemente mencionada anteriormente, consiste em treinar um sistema computacional utilizando exemplos de entrada associados a saída desejadas. O objetivo é encontrar um mapeamento entre os dados de entrada e a respectiva saída desejada. Assim, um conjunto de treinamento pode ser expresso por N pares de exemplos de entrada e saída: $(X_1, Y_1) \dots (X_n, Y_n)$, onde X_i é a representação dos dados de entrada e Y_i é a saída desejada, por exemplo, um rótulo que representa uma classe.

Em um exemplo simples de treinamento, voltado ao contexto deste trabalho, X_i poderia ser uma cadeia de caracteres que representa um padrão extraído de uma imagem de doença causada por fungos em lavouras de soja, e Y_i o rótulo “doenças causadas por fungos”. Durante o processo de treinamento, n distintas cadeias de caracteres X_i são associadas a uma mesma classe Y_i .

De acordo com Norvig & Russell (2015) formalmente a aprendizagem supervisionada é definida como: dado um conjunto de treinamento de N pares de entrada e saída $(X_1, Y_1), (X_2, Y_2) \dots (X_n, Y_n)$, onde Y_i é gerado por uma função desconhecida $y = f(x)$, obtenha uma função h que se aproxime da função verdadeira f . A função h é uma hipótese e a aprendizagem é uma busca a partir do espaço de hipóteses possíveis. Para medir a precisão de uma hipótese, fornece-se um conjunto de testes que são distintos do conjunto de treinamento. Assim, uma hipótese é considerada boa quando ela generaliza corretamente o valor de y para novos exemplos. Em alguns casos a função f é estocástica e o que deve-se aprender é uma distribuição de probabilidade condicional.

Na aprendizagem não supervisionada o algoritmo de aprendizagem busca sozinho classes, padrões e regularidades nos dados que recebe como entrada. Ao contrário da aprendizagem supervisionada, não é dito qual a saída desejada para uma dada entrada de dados, ao invés disso, procura-se uma estimativa de densidade, isto é, os modelos são construídos da forma $p(x_i | \theta)$, onde x_i geralmente é um vetor de característica (Norvig & Russell, 2015).

Em um exemplo simples voltado ao contexto deste trabalho, x_i poderia também ser uma cadeia de caracteres que representa um padrão de uma doença da soja extraído de uma imagem. Neste caso, durante o processo de treinamento, é dado como entrada n distintas cadeias de caracteres x_i representando diferentes doenças da lavoura da soja, como também, o número de agrupamentos desejados, por exemplo, 2. Assim, a grosso modo, os dados serão agrupados em 2 grupos distintos se baseando, por exemplo, o quão similar eles são por meio de uma medida de proximidade. Ao final os grupos representarão duas classes distintas de doença.

Nesta seção houve maior enfoque, mesmo que de forma resumida, nas aprendizagens supervisionada e não supervisionada, pois estes são os tipos de aprendizagem que se pretende adotar neste trabalho.

3.3 Visão Computacional

O sistema visual humano não tem dificuldade em interpretar imagens em relação às suas variações sutis na translucidez e sombreamento. Percebemos a estrutura tridimensional do mundo, as cores, profundidade, e as emoções nas pessoas à nossa volta com aparente facilidade. Todo esse complexo processo de reconhecimento, em suma, é realizado pela captura da luz pelos nossos olhos e interpretado pelo cérebro por meio de reconhecimento de padrões, contexto, memória e emoções

(Forsyth & Ponce, 2002).

A visão computacional é uma área da ciência da computação que emprega um conjunto de técnicas para adquirir, processar e analisar imagens a fim de reproduzir em sistemas a capacidade humana da visão e assim ajudar em análises do ambiente ou objeto (Jahne & Haussecker, 2000). A obtenção das imagens, geralmente é por meio de dispositivos eletrônicos, como câmeras fotográficas e filmadoras digitais (de Souza, 2010), após a captura, algumas etapas de processamento são realizadas para obter informações úteis da imagem a fim de atingir um determinado objetivo específico, como por exemplo, o reconhecimento e classificação de doenças em lavouras de soja.

A etapa de pré-processamento tem o objetivo de aumentar a qualidade da imagem para que etapas posteriores não sejam prejudicadas devido a interferências e imperfeições. Nessa etapa, são realizados os processos de redução de ruído e realce da imagem. Após a etapa de pré-processamento, é realizada a segmentação da imagem aplicando técnicas para extrair objetos de interesse (Szeliski, 2010), esses objetos são definidos de acordo com o escopo e objetivo da aplicação que fará uso deles. A próxima etapa é extrair informações que permitam a caracterização dos objetos de interesse, de modo que as características extraídas possam ser reconhecidas matematicamente e observadas em objetos semelhantes (Quinta, 2013).

Dependendo da aplicação, o rastreamento de objetos pode ser utilizado. Ele consiste em obter o posicionamento de um dado objeto alvo ao longo do tempo (Szeliski, 2010). Por fim, técnicas de aprendizagem automática são aplicadas para criar uma base de conhecimento sobre os objetos de interesses extraídos em etapa anterior, para que, analisados de forma automática, futuros objetos possam ser identificados e reconhecidos a partir de conhecimento de grupos de objetos já existentes.

3.4 Problemas da cultura da soja

O primeiro passo para se realizar um adequado controle de doenças em plantas é a correta identificação das mesmas. No contexto deste trabalho, automatizar essa tarefa em lavouras de soja é um dos objetivos finais. No entanto, é importante como uma tarefa intermediária, elencar as principais doenças e categorizá-las em classes a fim de utilizar tal categorização em treinamento de aplicações, bem como, definir padrões que representam grupos ou tipos de doenças. Por isso, nesta seção pretende-se apresentar algumas doenças relacionadas ao cultivo da soja.

De acordo com Henning et al. (2014) os grupos de doenças que ocorrem em lavouras de soja



Figura 1: Exemplos de doenças causadas por fungos em lavouras de soja, (a) mancha-alvo e (b) Oídio.

estão categorizados como: doenças causadas por fungos; por bactérias; por vírus e; por nematóides. Observa-se que as doenças são diferenciáveis entre si e, na maioria das vezes, apresentam um padrão que as caracterizam. Como dito anteriormente, esse padrão é importante, porque pode facilitar a definição de classes durante o processo de treinamento, como também, por consequência, a classificação. Na Figura 1 são apresentados exemplos de doenças causadas por fungos.

Nos exemplos de doenças apresentado na Figura 1 é possível observar padrões, como por exemplo, manchas circulares (Figura 1 (a)) e manchas brancas (Figura 1 (b)), que se repetem e podem ser explorados para representações de características. Um dos desafios deste trabalho é extrair das imagens e representar tais padrões sintaticamente de maneira que considere as características de cada doença e as relações estruturais entre elas, ou seja, a base para representação de padrões.

4 Metodologia

Esta seção apresenta a metodologia deste trabalho visando alcançar os objetivos gerais e específicos definidos na seção 2.

A primeira etapa deste trabalho se concentra em realizar uma revisão sistemática sobre o assunto abordado. As fases dessa revisão se concentram em (Kitchenham, 2004): a) planejamento: definem as perguntas da pesquisa, critérios de seleção de trabalhos, definição de base de dados para busca de trabalhos científicos, critérios de seleção de trabalhos e ameaças a pesquisa proposta;

b) condução: nessa fase elencam-se as pesquisas descartadas acompanhadas da justificativa, as pesquisas selecionadas são utilizadas para responder às questões de pesquisa proposta na fase anterior; c) relatório: é realizado para apresentar os resultados da condução da revisão sistemática.

Com o objetivo de identificar as necessidades do público alvo, pretende-se realizar entrevistas com profissionais da agricultura produtores de soja, como também, conhecer o processo desde o plantio até a colheita.

A próxima etapa consiste em criar o banco de imagens contendo doenças de lavouras de soja. Para isso, pretende-se utilizar RPAs para coleta de imagens no ambiente real, equipados por câmeras e outros equipamentos que permitam a captura da imagem da lavoura da soja. As imagens capturadas serão gerenciadas por uma aplicação de software que facilite a manipulação dos dados.

Após a obtenção das imagens, os esforços se concentrarão em selecionar as principais técnicas de visão computacional para tratar as imagens da lavoura com o objetivo de extrair dados e características necessárias que auxiliem na representação de padrões. Serão definidas classes que possam ser representadas por padrões em uma lavoura de soja, como por exemplo, doenças causadas por fungos, por bactérias, entre outras; e as imagens serão agrupadas em suas respectivas classes. Nessa etapa serão realizados estudos sobre técnicas de representação e reconhecimento sintático de padrões, análise e implementação de algoritmos de pré-processamento de imagens, segmentação e extração de características; nesta última, com o objetivo de extrair características representativas para cada classe.

Em uma próxima etapa será desenvolvida uma nova técnica de representação de padrão baseado em métodos sintáticos, que possa representar as características mensuráveis das imagens, como também, as relações estruturais entre elas. Para cada imagem que compõe uma classe será extraída uma representação, ou seja, se uma classe é composta por n imagens, n representações serão extraídas. As representações de cada imagem serão utilizadas para inferir uma gramática que representa a classe, isto é, um padrão que define a Linguagem da classe. Assim, a aplicação baseado em visão computacional passa a ser treinado a reconhecer a linguagem que representa a classe, aprendendo a gramática que a padroniza.

A avaliação do método proposto se baseia na metodologia proposta por Norvig & Russell (2015), que possui os seguintes passos: a) coletar um grande conjunto X de exemplos possíveis; b) dividi-lo em dois conjuntos disjuntos: o conjunto de treinamento e o conjunto de testes; c) aplicar

o método proposto ao conjunto de treinamento, gerando uma hipótese h ; d) medir a porcentagem de exemplos no conjunto de teste que são corretamente classificados por h ; e) repetir os quatro passos anteriores para diferentes tamanhos de conjuntos de treinamento e diferentes conjuntos de treinamentos de cada tamanho selecionado aleatoriamente.

Para analisar os exemplos corretamente classificados por h , pretende-se adotar métricas como a taxa de erro (porcentagem de predição incorreta), taxa de erro ponderada, precisão (porcentagem de classificação positiva que estão corretas), acurácia (porcentagem de todas as classificações que estão corretas), recall (porcentagem de exemplos positivos que são corretamente classificados) e área sob a curva ROC. Essas métricas são as mais comumente utilizadas quando o modelo a ser validado está relacionado a problema de classificação.

No que diz respeito às repetições para diferentes tamanhos de conjuntos de treinamento é necessário adotar estratégias de otimização que provêem combinações de hiperparâmetros que maximiza a acurácia do método proposto. Os hiperparâmetros que podem ser arranjados estão relacionados ao tamanho do alfabeto da linguagem padronizada pela gramática, estratégias de inferência gramatical, estratégias de recuperação e detecção de erro em relação a análise sintática e estratégias de aprendizagem, como por exemplo, supervisionada, não supervisionada e semi-supervisionada. Os resultados das análises das estratégias de otimização serão avaliados estatisticamente por meio de análise de variância, e o grau de significância a ser adotado será o mesmo de trabalhos análogos e dependerá da revisão sistemática realizada em etapa anterior.

5 Atividades e cronograma de execução

A Tabela 1 apresenta o cronograma proposto para este trabalho com a distribuição das atividades supracitadas de forma semestral em um período de quatro anos.

Tabela 1: Cronograma de trabalho

Atividade	Ano 1		Ano 2		Ano 3		Ano 4	
Revisão Sistemática (revisão da literatura)	X	X						
Entrevistas com agrônomos e produtores de soja		X						
Criar o banco de imagens de lavouras de soja		X						
Definição de classes que possam representar padrões em uma lavoura de soja		X						
Estudar e selecionar as principais técnicas de visão computacional	X	X	X					
Estudar técnicas de representação e reconhecimento sintático de padrões	X	X	X	X				
Propor nova técnica de representação de padrão baseado em métodos sintáticos		X	X	X	X	X		
Realizar o treinamento, utilizando algoritmos de aprendizagem supervisionada		X	X	X	X			
Realizar experimentos sobre o método proposto			X	X	X	X		
Realizar análise dos resultados obtidos por meio de análise estatística					X	X	X	
Transferir o conhecimento obtido por meio de publicações em eventos e revistas técnico-científicos			X	X	X	X	X	X
Escrita da Tese							X	X

6 Resultados esperados, produtos e avanços

Espera-se com a proposta de uma nova técnica de reconhecimento sintático de padrões em imagens aplicar os conceitos inerentes a representação de padrões, adotando uma abordagem contrária a representações por meio de vetores de características, considerando a inter-relação ou interconexão das características extraídas das imagens, como também, considerar probabilidade na representação. Dessa forma, pretende-se demonstrar possibilidades para melhoramentos para a área de reconhecimento sintático de padrões aplicado em imagens.

Referências

- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.
- Bruce, R. F. (2001). A bayesian approach to semi-supervised learning. In *NLPRS*.
- Chanda, G. & Dellaert, F. (2004). *Grammatical Methods in Computer Vision: An Overview*. Technical Report GIT-GVU-04-29, College of Computing, Georgia Institute of Technology, Atlanta, GA.
- Chua, J. (2018). *Probabilistic Scene Grammars: A General-Purpose Framework For Scene Understanding*. PhD thesis, Brown University, Providence, Rhode Island.
- Chua, J. & Felzenszwalb, P. F. (2016). Scene grammars, factor graphs, and belief propagation. *CoRR*, abs/1606.01307.
- CONAB (2017). *Acompanhamento da safra de grãos Brasileira: Monitoramento agrícola – Safra 2016/17*. Technical Report V. 4 - SAFRA 2016/17 - N. 12, Companhia Nacional de Abastecimento (CONAB), Brasil, Brasília, DF.
- de la Higuera, C. (2010). *Grammatical Inference: Learning Automata and Grammars*. New York, NY, USA: Cambridge University Press.
- de Souza, K. P. (2010). Aplicação de modelos de markov ocultos na obtenção de taxas de mortalidade das larvas do mosquito da dengue. Master’s thesis, Faculdade de Computação, Universidade Federal de Mato Grosso do Sul.
- de Souza Pio, J. L., de Oliveira, C. J. S., de Araújo, A. A., & de Oliveira, N. J. (2006). *Atualizações em Informática*, chapter Reconhecimento Sintático e Estrutural de Padrões. Editora PUC-Rio.
- Fergus, R., Perona, P., & Zisserman, A. (2003). Object class recognition by unsupervised scale-invariant learning. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, volume 2 (pp. II-264–II-271 vol.2).
- Forsyth, D. A. & Ponce, J. (2002). *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference.

- Fu, K. S. (1982). *Syntactic Pattern Recognition and Applications*. Prentice-Hall, Englewood Cliffs, NJ.
- Fu, K.-S. & Rosenfeld, A. (1976). Pattern recognition and image processing. *IEEE Transactions on Computers*, C-25(12), 1336–1346.
- Girshick, R. B., Felzenszwalb, P. F., & McAllester, D. (2011). Object detection with grammar models. In *Proceedings of the 24th International Conference on Neural Information Processing Systems*, NIPS'11 (pp. 442–450). USA: Curran Associates Inc.
- Henning, A. A., Álvaro Manuel Rodrigues Almeida, Godoy, C. V., Seixas, C. D. S., Yorinori, J. T., Costamilan, L. M., Ferreira, L. P., Meyer, M. C., Soares, R. M., & Dias, W. P. (2014). *Manual de identificação de doenças de soja*. Technical Report Documentos 256, Empresa Brasileira de Pesquisa Agropecuária, Embrapa Soja, Londrina, PR.
- Hirakuri, M. H. & Lazzarotto, J. J. (2014). *O agronegócio da soja nos contextos mundial e brasileiro*. Technical Report 349, Empresa Brasileira de Pesquisa Agropecuária, Embrapa Soja, Londrina, PR.
- Jahne, B. & Haussecker, H., Eds. (2000). *Computer Vision and Applications: A Guide for Students and Practitioners*. Orlando, FL, USA: Academic Press, Inc.
- Kitchenham, B. (2004). *Procedures for Performing Systematic Reviews*. Technical Report TR/SE-0401, Department of Computer Science, Keele University, UK.
- Li, X., Wu, T., Song, X., & Krim, H. (2017). Aognets: Deep and-or grammar networks for visual recognition. *CoRR*.
- Luger, G. F. (2013). *Inteligência Artificial*. Pearson Education do Brasil, São Paulo, SP, 6 edition.
- Norvig, P. & Russell, S. (2015). *Inteligência Artificial: Tradução da 3a Edição*. Elsevier Brasil.
- Pedro, R. W. D. (2013). Inferência de gramáticas estocásticas para reconhecimento de padrões de imagens utilizando quadrees. Master's thesis, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo.

- Pistori, H., Calway, A., & Flach, P. (2013). A new strategy for applying grammatical inference to image classification problems. In *2013 IEEE International Conference on Industrial Technology (ICIT)* (pp. 1032–1037).
- Quinta, L. N. B. (2013). Visão computacional aplicada na classificação de grãos de pólen. Master’s thesis, Programa de PósGraduação em Biotecnologia da Universidade Católica Dom Bosco.
- Ribas, L. C., Borth, M., Jr., A. A. C., Gonçalves, W. N., & Pistori†, H. (2014). Grammatical inference and sift for scene recognition. In *Proceedings of the X Workshop de Visão Computacional - WVC 2014*, WVC 2014 (pp. 293 – 298). Campo Grande, MS, Brazil.
- Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. New York, NY, USA: Springer-Verlag New York, Inc., 1st edition.
- Walton, M., Lange, D., & Zhu, S.-C. (2017). Inferring context through scene understanding. In *AAAI Spring Symposium Series*.
- Zhu L. Chen, Y. Yuille, A. (2009). Unsupervised learning of probabilistic grammar-markov models for object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1), 114.