



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José do Rio Preto

Leandro Aguiar Mota

Redes Neurais Randomizadas e Caminhadas Determinísticas para Análise de Imagens

São José do Rio Preto

2025

Leandro Aguiar Mota

Redes Neurais Randomizadas e Caminhadas Determinísticas para Análise de Imagens

Monografia apresentada como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação, junto ao Departamento de Ciências de Computação e Estatística do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de São José do Rio Preto.

Universidade Estadual Paulista "Júlio de Mesquita Filho" - (UNESP)

Instituto de Biociências, Letras e Ciências Exatas - (IBILCE)

Departamento de Ciências de Computação e Estatística

Bacharelado em Ciência da Computação

Orientador: Lucas Correia Ribas

São José do Rio Preto

2025

M917r Mota, Leandro
Redes Neurais Randomizadas e Caminhadas Determinísticas para
Análise de Imagens / Leandro Mota. -- São José do Rio Preto, 2025
50 p. : il., tabs., fotos

Trabalho de conclusão de curso (Bacharelado - Ciência da
Computação) - Universidade Estadual Paulista (UNESP), Instituto de
Biociências Letras e Ciências Exatas, São José do Rio Preto
Orientador: Lucas Correia Ribas

1. Inteligência Artificial. 2. Visão Computacional. 3. Texturas. 4.
Redes Neurais Randomizadas. 5. Caminhadas Determinísticas. I.
Título.

Leandro Aguiar Mota

Redes Neurais Randomizadas e Caminhadas Determinísticas para Análise de Imagens

Monografia apresentada como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação, junto ao Departamento de Ciências de Computação e Estatística do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de São José do Rio Preto.

Banca Examinadora:

Prof. Dr. Lucas Correia Ribas

Universidade Estadual Paulista - Câmpus de São José do Rio Preto
Orientador

Prof. Dr. Diego Renan Bruno

Universidade Estadual Paulista - Câmpus de São José do Rio Preto

Prof. Dr. Geraldo Francisco Donegá Zafalon

Universidade Estadual Paulista - Câmpus de São José do Rio Preto

São José do Rio Preto

2025

Dedico este trabalho aos meus pais, que sempre me apoiaram e acreditaram no meu potencial, e a todos que, de alguma forma, contribuíram para a realização deste sonho.

Agradecimentos

Agradeço primeiramente aos meus pais, Zulmira e Ailton, pelo amor, paciência e incentivo incondicional em todos os momentos. Vocês foram meu alicerce e minha maior fonte de motivação para seguir em frente, mesmo diante dos desafios.

Aos meus amigos, Otávio, Mateus, Luan e Yhan, pela amizade sincera, pelas conversas, risadas e pelo apoio constante ao longo de toda a jornada acadêmica. Cada um de vocês tornou este caminho mais leve e significativo.

À minha namorada, Nanami, por todo o apoio, paciência e compreensão durante este processo. Sua companhia e incentivo foram fundamentais para que eu mantivesse a motivação e a tranquilidade nos momentos mais desafiadores. Sou grato por ter contado com seu carinho e presença ao longo desta jornada.

Aos professores, por compartilharem seus conhecimentos com dedicação e entusiasmo, contribuindo não apenas para a minha formação profissional, mas também para o meu crescimento pessoal.

Ao meu orientador, Lucas, por ser um exemplo de dedicação, ética e amor à profissão. Sou profundamente grato pelo voto de confiança, pelos conselhos e por toda paciência ao longo deste processo. Seus ensinamentos deixaram marcas que levarei comigo por toda a vida.

Por fim, UNESP, pela oportunidade de aprendizado, pela estrutura oferecida e pelo ambiente acadêmico que tanto favoreceu o meu desenvolvimento científico e humano.

A todos que, de alguma forma, contribuíram para esta conquista, deixo o meu mais sincero agradecimento.

"Seja curioso. Leia muito. Experimente coisas novas. Eu acho que muito do que as pessoas chamam de inteligência apenas se resume a curiosidade."
(Aaron Swartz)

Resumo

Este trabalho aborda a análise de texturas em imagens digitais, com ênfase na criação de representações mais robustas e discriminativas para tarefas de classificação. O objetivo principal é propor, implementar e avaliar uma abordagem híbrida que combina Caminhadas Determinísticas Parcialmente Auto-Repulsivas (CDPARs) e Redes Neurais Randomizadas (RNNs) para a extração de características de textura. As CDPARs são utilizadas para explorar informações locais das imagens, gerando vetores que capturam padrões estruturais, enquanto as RNNs são aplicadas para aprender representações discriminativas a partir desses vetores de forma rápida e eficiente. Além disso, são realizadas combinações de parâmetros e configurações de redes, permitindo analisar o impacto de diferentes tamanhos de memória, regras de movimento e quantidades de neurônios ocultos. Experimentos conduzidos em bases de dados de texturas demonstraram que a abordagem proposta é competitiva em relação a métodos clássicos e recentes da literatura, alcançando acurácia superior a 95% na USPTex e acima de 90% na Outex, demonstrando alto desempenho e competitividade da solução.. Dessa forma, este trabalho contribui para o avanço de técnicas de análise de texturas e aprendizado de representações, abrindo novas possibilidades para aplicações em visão computacional e reconhecimento de padrões.

Palavras-chave: Análise de imagens, Caminhadas determinísticas parcialmente auto-repulsivas, Redes neurais randomizadas, Textura.

Abstract

This work addresses the analysis of textures in digital images, focusing on creating more robust and discriminative representations for classification tasks. The main goal is to propose, implement, and evaluate a hybrid approach that combines Deterministic Partially Self-Avoiding Walks (DPSWs) and Randomized Neural Networks (RNNs) for texture feature extraction. DPSWs are used to explore local image information, generating vectors that capture structural patterns, while RNNs are applied to learn discriminative representations from these vectors quickly and efficiently. Furthermore, different network parameters and configurations are combined to analyze the influence of memory size, movement rules, and hidden neuron quantities. Experiments conducted on texture datasets demonstrated that the proposed approach is competitive with classical and state-of-the-art methods, achieving accuracy above 95% on USPTex and over 90% on Outex, demonstrating its high performance and competitiveness. Thus, this work contributes to the advancement of texture analysis and representation learning techniques, opening new possibilities for applications in computer vision and pattern recognition.

Keywords: *Image analysis, Partially self-avoiding deterministic walks, Randomized neural networks. Textures.*

Sumário

	Lista de ilustrações	10
	Lista de tabelas	11
1	INTRODUÇÃO	14
1.1	Motivação	15
1.2	Objetivos	15
2	FUNDAMENTAÇÃO TEÓRICA E TRABALHOS RELACIONADOS	16
2.1	Redes Neurais Artificiais	16
2.2	Redes Neurais Aleatórias	18
2.3	Caminhadas Determinísticas Parcialmente Auto-Repulsivas	21
3	METODOLOGIA	25
3.1	Base de Dados	26
3.1.1	USPTex	26
3.1.2	Outex	27
3.2	Classificadores	28
3.3	Validação	29
3.4	Métricas de Avaliação	30
3.4.1	Matriz de Confusão	30
3.4.2	Acurácia	31
3.4.3	Desvio Padrão	31
4	ABORDAGEM PROPOSTA	32
4.1	Imagens de Entrada	32
4.2	Extração de Vetores Locais	33
4.3	Treinamento das RNNs	34
4.4	Representações	35
5	RESULTADOS	37
5.1	Avaliação de Parâmetros	37
5.2	Comparações com a Literatura	42
5.3	Identificação de Espécies de Plantas Brasileiras	43
6	CONCLUSÃO	46
	REFERÊNCIAS	48

Lista de ilustrações

Figura 1 – Modelo de arquitetura de rede neural com um neurônio	17
Figura 2 – Modelo de Rede Neural Aleatória (RNN) com $Q = 5$	19
Figura 3 – Exemplo de CDPAR com tamanho de memória $\mu = 2$ e regra $r = \min$ em uma imagem 5x5. Cada número representa a intensidade de um pixel. Em azul temos o transiente e em verde o atrator referentes a caminhada	23
Figura 4 – Fluxo de Etapas de Execução	25
Figura 5 – Exemplos de imagens de textura da USPTex	26
Figura 6 – Exemplos de imagens de textura da Outex	27
Figura 7 – Diagrama Esquemático da Validação Cruzada	29
Figura 8 – Matriz de Confusão	30
Figura 9 – Esquema da metodologia do projeto. No passo (ii) as CDPARs são usadas para analisar a complexidade e extrair informações locais da imagem; (iii) os vetores locais extraídos pelas CDPARs são usados para treinamento das RNNs em (iv); (v) usando técnicas de agregação os pesos da camada de saída das diferentes RNNs treinadas formam o vetor de características final em (iv).	33
Figura 10 – Comparação de acurácia média versus memória para diferentes regras em Usptex	39
Figura 11 – Comparação de acurácia média versus memória para diferentes regras em Outex	40
Figura 12 – Comparação da Acurácia Média versus Q para os datasets Outex e Usptex usando a regra Max + Min.	41
Figura 13 – Exemplos de imagens de textura da 1200Tex	44

Lista de tabelas

Tabela 1	– Resultados obtidos para diferentes valores de Q em dois datasets . . .	38
Tabela 2	– Comparação de acurácias de diferentes métodos de análise de textura em três bancos de dados de texturas.	42
Tabela 3	– Acurácia usando o classificador LDA na base de dados 1200Tex. . . .	44

Lista de algoritmos

- 1 Pseudocódigo do treinamento de um RNN, em que Q é o número de neurônios da camada oculta. 20

Lista de abreviaturas e siglas

LDA	<i>Linear Discriminant Analysis</i>
RNN	<i>Randomized Neural Network</i>
CDPAR	Caminhadas Determinísticas Parcialmente Auto-Repulsivas
RNA	Redes Neurais Artificiais
LCG	<i>Linear Congruent Generator</i>

1 Introdução

As texturas, foco deste trabalho, desempenham um papel crucial no reconhecimento de padrões e possuem aplicações práticas em diversos campos, incluindo medicina, ciência dos materiais e botânica. Embora as imagens de texturas possam parecer visualmente simples, sua formação envolve processos não lineares, tornando o reconhecimento um desafio que exige técnicas avançadas (FLORINDO, 2020). A literatura de análise de texturas é vasta, dividindo-se em dois paradigmas principais: os métodos clássicos (descritores manuais); e abordagens que usam aprendizado de máquina, como Redes Neurais Artificiais.

Os descritores manuais são úteis em cenários com dados limitados para treinamento, hardware restrito, necessidade de maior interpretabilidade e quando informações específicas ao problema, como modelagem física, podem aprimorar o desempenho do algoritmo (SILVA; FLORINDO, 2021). Dentro dessa categoria de métodos, as Caminhadas Determinísticas Parcialmente Auto-Repulsivas (CDPARs) têm se destacado, sendo utilizadas com sucesso como medidas de complexidade em várias tarefas de visão computacional, destacando-se na identificação de padrões em imagens de texturas complexas (BACKES et al., 2006a), como de materiais e imagens naturais.

Nem sempre descritores manuais eficazes para desafios científicos atuais. Assim, dentro da segunda abordagem e indo em uma direção diferente das redes profundas convencionais, as Redes Neurais Randomizadas (*Randomized Neural Networks* (RNNs) (SCABINI et al., 2023; FARES et al., 2024) usam uma solução de forma fechada, resultando em um algoritmo de treinamento rápido, em vez da retropropagação comum e com grande potencial no reconhecimento de padrões em imagens. No contexto deste trabalho, as RNNs são empregadas não como classificadores diretos, mas como potentes extratores de características que aprendem representações não-lineares a partir dos padrões locais identificados pelas CDPARs.

Dessa forma, com o sucesso de ambos os métodos separadamente. Neste trabalho, propõe-se o desenvolvimento de técnica de análise de imagens que combina as CDPARs e RNNs, considerando, assim, as vantagens de ambos os paradigmas, propondo-se estudar a utilização das CDPARs para definir o conjunto de pixels que compõem o vetor de treinamento das RNNs. Isso permite analisar a complexidade da imagem e extrair informações locais para alimentar as RNNs, criando assim, um descritor robusto de textura, utilizado no processo de classificação posteriormente.

1.1 Motivação

Devido ao sucesso de ambos os paradigmas, descritores manuais (CDPARs) e aprendidos por redes neurais (RNNs), motiva um aprofundamento na investigação de métodos de ambas as perspectivas visando melhorar o desempenho na tarefa de análise e classificação de imagens.

Portanto, acredita-se que a combinação do uso de CDPARs como ferramenta de análise da complexidade local das imagens, juntamente com o poder de aprendizado das RNNs, resultará em uma representação de imagem mais discriminativa. Essa estratégia tem um potencial especial ao lidar com imagens de texturas complexas, como as encontradas na natureza e de materiais. Por fim, espera-se que essa abordagem permita capturar com maior precisão as características intrínsecas presentes nas imagens, melhorando assim a capacidade de análise e classificação de imagens.

1.2 Objetivos

O objetivo principal deste trabalho é explorar o uso de RNNs para aprender representações para análise de imagens, utilizando informações locais obtidas por meio de CDPARs para treinar as RNNs. De maneira mais específica, os objetivos deste estudo podem ser resumidos como:

- Estudar e aprimorar as técnicas existentes de CDPARs em imagens de texturas para análise de padrões locais;
- Desenvolver e implementar técnica de análise de imagem de textura baseada em aprendizado de representação com RNNs treinadas com informações locais extraídas por CDPARs;
- Analisar e comparar os resultados dos métodos de extração de características, visando demonstrar a eficácia da combinação de RNNs e CDPARs como técnica de análise de imagens.

2 Fundamentação Teórica e Trabalhos Relacionados

As texturas possuem uma ampla gama de definições na comunidade científica, moldadas pelo contexto de um problema específico ou por visões individuais. Por esse motivo, no campo da computação existem algumas formas diferentes de caracterizá-las, possuindo dois grandes grupos de abordagens: métodos baseados em formulações matemáticas, conhecidos como descritores manuais dos padrões texturais, bem como, mais recentemente, a utilização de métodos por aprendizado, baseados em Redes Neurais Artificiais.

Neste contexto, este capítulo apresenta a fundamentação teórica necessária para compreender as técnicas utilizadas neste trabalho, contextualizando-as com uma revisão da literatura que demonstra suas aplicações práticas e evolução ao longo dos anos. São explicados conceitos-chave da área do aprendizado de máquina, inteligência artificial e sistemas complexos, acompanhados dos principais trabalhos científicos que contribuíram para o desenvolvimento dessas técnicas. Assim, começa-se explorando as redes neurais, que são modelos de computação baseados no comportamento do cérebro humano, muito utilizados para resolver problemas complexos em diversas áreas.

Em seguida, são aprofundadas as redes neurais aleatórias, uma variação das redes neurais artificiais que utiliza conexões estabelecidas aleatoriamente com o intuito de melhorar a eficiência e desempenho em aplicações específicas. Por fim, são exploradas as caminhadas determinísticas parcialmente auto-repulsivas, um algoritmo baseado em sistemas complexos onde um agente navega por pixels da imagem baseado em regras determinísticas, permitindo a extração de características discriminativas das texturas. Para cada conceito apresentado, são discutidos os trabalhos relevantes da literatura que demonstram sua aplicabilidade e contribuições para a área de análise de texturas.

2.1 Redes Neurais Artificiais

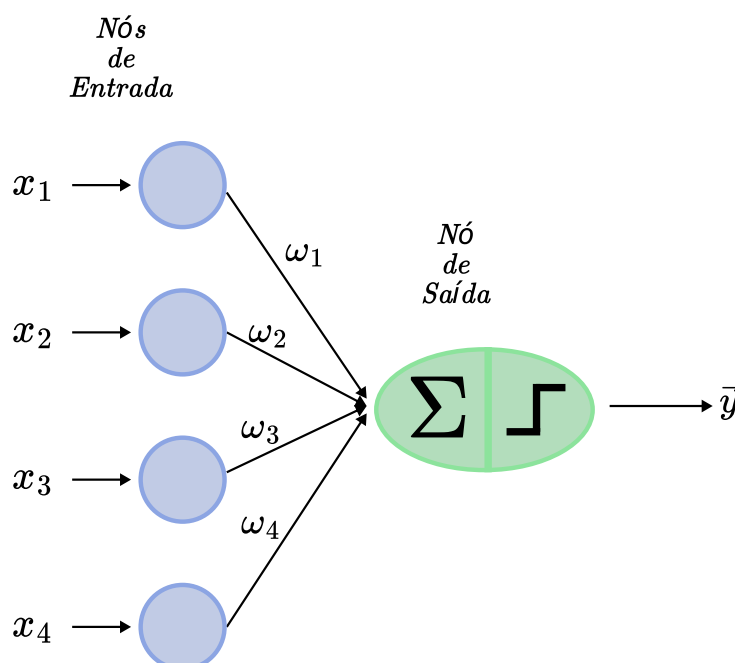
Como descrito em (AGGARWAL, 2018), as Redes Neurais Artificiais (RNAs) são técnicas que simulam o mecanismo do aprendizado em organismos biológicos. Assim, o sistema nervoso humano contém neurônios que se conectam com outros e para realizarem existem as sinapses, referidas como as regiões de conexões. Dessa forma, os esforços dessas sinapses mudam em resposta a estímulos externos. Essa mudança é o que define o aprendizado nos seres humanos.

As RNAs simulam esse comportamento biológico do cérebro humano, usando unidades computacionais chamadas neurônios, os quais são conectadas por pesos, agindo como as

sinapses. Estes pesos escalam os neurônios com entradas que afetam uma função computada nessa unidade. A rede processa informações propagando valores dos neurônios de entrada para os de saída, utilizando os pesos como parâmetros intermediários.

O processo de aprendizado nas RNAs ocorre por meio da alteração dos pesos que conectam os neurônios. A rede recebe um estímulo externo para efetuar esse aprendizado, fornecido pelos dados de treinamento, os quais contêm pares de entrada e saída da função que vai ser computada, essa arquitetura básica pode ser vista na Figura 1. Assim, a rede faz previsões em cima desses pares e a atualização dos pesos depende da proximidade da previsão com o rótulo de saída real. Dessa forma, ao errar alguma previsão, a rede recebe *Feedbacks*, que são utilizados para ajustar os pesos como resposta.

Figura 1 – Modelo de arquitetura de rede neural com um neurônio



Fonte: Adaptado de (AGGARWAL, 2018)

O objetivo é ajustar a função computada pela rede para que suas previsões se tornem cada vez mais precisas. Para isso, os pesos são modificados de forma matematicamente controlada com o intuito de reduzir os erros. Esse processo é repetido diversas vezes com vários pares de entrada e saída, o que permite à rede melhorar gradualmente seu desempenho e reconhecer exemplos que nunca viu antes, caracterizando a generalização da rede. Essa habilidade de generalizar é o principal benefício dos modelos de aprendizado de máquina.

Nesse Contexto, existem diversas arquiteturas de redes neurais, as de camada única e as multicamadas conferem duas dessas abordagens. Na primeira, as entradas são mapeadas diretamente para uma saída por meio de uma função, também pode ser conhecida como *perceptron*. Já a segunda, os neurônios são organizados em camadas, ao passo que as camadas

de entrada e saída são divididas por uma ou mais camadas ocultas. Assim, essa disposição em camadas pode ser chamada de rede neural *feed-forward*.

2.2 Redes Neurais Aleatórias

As Redes Neurais Aleatórias (RNNs), em sua uma forma mais simples, uma arquitetura de rede neural *feed-forward* com apenas uma única camada oculta, cujos os pesos de entrada são determinados de forma aleatória, enquanto seus pesos da camada de saída são aprendidos por meio de uma solução fechada determinada, em geral, pelo método dos mínimos quadrados. Essa abordagem torna o processo de aprendizado mais ágil, permitindo que essas redes sejam aplicadas a situações que demandam processamento rápido. O principal propósito desse modelo é converter os dados de entrada para um espaço latente distinto, em que os vetores de características sejam mais facilmente interpretados por classificadores (RIBAS, 2021).

Para exemplificar o processo de aprendizado da RNN, definimos $X = [\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n]$ como a matriz de N vetores de características de entrada com ρ atributos e $\vec{y} = [y_1, y_2, \dots, y_n]$, os respectivos rótulos correspondentes aos dados de entrada. Com isso, a primeira tarefa da RNN é definir valores de pesos aleatórios da camada escondida, que são gerados utilizando uma distribuição normal ou gaussiana, os quais compõem a matriz W , que posteriormente, passa por um processo de normalização com média zero e desvio padrão unitário. Esta matriz, que tem dimensão $Q \times (\rho + 1)$, onde Q é o número de neurônios na camada escondida da RNN, cada linha representa os pesos de um desses neurônios e a primeira coluna representa os pesos do viés. Assim, para cada vetor de entrada $\vec{x}_i$ é adicionada uma constante -1, com o intuito de conectar com os pesos do viés dos neurônios ocultos.

Na sequência, os valores da camada escondida são computados por meio da função de ativação $\phi(WX)$, sendo utilizada neste trabalho, a função sigmóide. A saída dessa função de ativação é utilizada para alimentar a matriz $Z = [\vec{z}_1, \vec{z}_2, \dots, \vec{z}_n]$ de vetores de características, que são usados como entrada da camada de saída. Dessa forma, acrescenta-se novamente a constante -1 a cada vetor $\vec{z}_i$, de modo a estabelecer a conexão com os pesos de viés dos neurônios da camada de saída (RIBAS, 2021).

Dessa forma, os pesos associados aos neurônios da camada de saída são dispostos em forma de um vetor unidimensional M , tendo em vista que a RNN é treinada apenas com um único neurônio de saída, que se refere a intensidade de um pixel em que se inicia uma CDPAR. Estes pesos devem satisfazer a equação $D = MZ$, para isso podemos utilizar a pseudo inversa de Moore-Penrose, resultando na equação 2.1

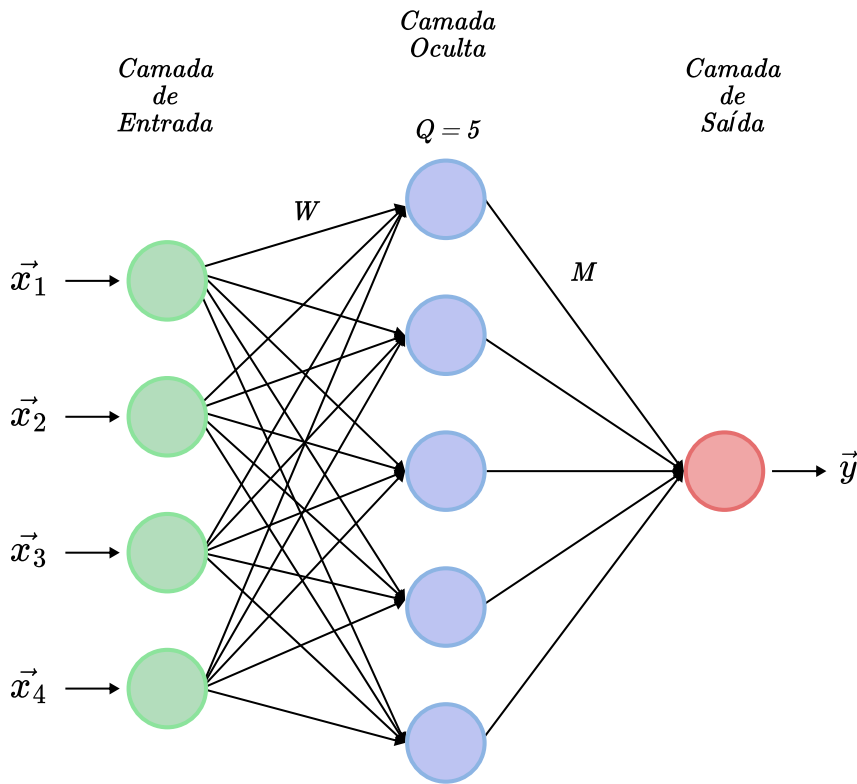
$$M = DZ^T(ZZ^T)^{-1} \quad (2.1)$$

Frequentemente, a matriz ZZ^T aproxima-se de uma condição singular em diversos problemas, o que leva a um cálculo impreciso de sua inversa. Com isso, para contornar essa questão, aplica-se a regularização de Tikhonov, a qual ajusta o cálculo da seguinte forma:

$$M = DZ^T(ZZ^T + \lambda I)^{-1} \quad (2.2)$$

onde I representa a matriz identidade e λ , um parâmetro de regularização com valor no intervalo de 0 a 1. A Figura 2 mostra uma simples RNN com 5 neurônios na camada oculta ($Q = 5$).

Figura 2 – Modelo de Rede Neural Aleatória (RNN) com $Q = 5$



Fonte: Adaptado de (ZIELINSKI et al., 2022)

Um aspecto muito importante das RNNs é a geração aleatória de pesos na camada de entrada. Assim, é fundamental sempre utilizar os mesmos valores na matriz de pesos W , assegurando que uma imagem, quando processada com os mesmos parâmetros, produza um vetor de características de saída único. Dessa forma, buscando a distribuição normal, utiliza-se o *Linear Congruential Generator* (LCG), um simples gerador de números pseudo-aleatórios na forma da equação 2.3

$$V_{n+1} = (a \cdot V_n + c) \mod m \quad (2.3)$$

sendo, V uma sequência de números aleatórios com tamanho $E = Q \times (p + 1)$ e $V(1) = E + 1$ com os parâmetros $a = E + 2$ e $c = E^2$.

O Algoritmo 1 explicita os passos para o treinamento de uma RNN, por meio de pseudo-código.

Algoritmo 1 Pseudocódigo do treinamento de um RNN, em que Q é o número de neurônios da camada oculta.

```

1: function TRAINRAE( $X, D, Q$ )
2:    $W = \text{LCG}(Q, P + 1, Q \times (P + 1))$  ▷ gerar os pesos aleatórios
3:    $W = \text{Zscore}(W)$  ▷ normalizar os pesos
4:    $X = \text{addBias}(X, -1)$  ▷ adiciona o bias na matriz de entrada
5:
6:    $Z = \sigma(W \times X)$  ▷ função de ativação da camada oculta
7:    $Z = \text{addBias}(Z, -1)$  ▷ adiciona o bias em Z
8:
9:    $\lambda = 10000$ 
10:   $M = (D \times Z^T) \times (Z \times Z^T + \lambda \times I_{Q+1})^{-1}$  ▷ calcula os pesos de saída com Tikhonov
11:
12:  retorne  $M$ 
13: fim function

```

Por fim, as RNNs podem funcionar como um *AutoEncoder* aleatório (RAE)(KASUN et al., 2013), abordando a matriz de características de entrada X como o rótulo de saída, ou seja, utilizando $Y = X$. Neste cenário, o algoritmo consiste na matriz X sendo codificada aleatoriamente para um outro espaço dimensional, e após isso, estas características projetadas em outro espaço são decodificadas de volta para os dados de entrada pelo método dos mínimos quadrados, mapeando os dados de entrada. Dessa forma, o uso de redes neurais aleatórias força o modelo a compreender os padrões subjacentes nos dados a partir da aleatoriedade, pelo uso de projeções aleatórias, e sua utilização como autoencoder conduz a rede a aprender esses padrões exclusivamente a partir dos próprios dados. Neste caso, os valores finais na matriz de pesos de saída M do decodificador correspondem à transformação linear que traz as características projetadas de volta ao seu formato original nos dados de entrada.

Com todos esses aspectos das RNNs, essa arquitetura de rede neural teve seus primeiros estudos realizados durante a década de 90(SCHMIDT; KRAAIJVELD; DUIN, 1992), sendo posteriormente aprimorada em trabalhos subsequentes (HUANG; ZHU; SIEW, 2006) que contribuíram para o avanço dessa arquitetura de rede neural. Devido ao seu rápido processo de aprendizado, simplicidade e bons resultados obtidos, diversos estudos passaram a empregar as RNNs em diferentes tarefas de reconhecimento de padrões(KAYA; TEKIN, 2013).

Assim, tem se consolidado como uma área de pesquisa promissora, com aplicações amplas na visão computacional, em que vários trabalhos têm explorado o uso das RNNs em tarefas como delimitação de formas (JUNIOR; BACKES; BRUNO, 2018) e análise de texturas em imagens (JUNIOR; BACKES, 2016; RIBAS et al., 2020a). Em (SCABINI et al., 2023), foi proposto o RADAM, no qual mapas de características extraídos de redes neurais convolucionais pré-treinadas são processados por uma rede neural randomizada configurada

como autoencoder, de modo a gerar descritores compactos e altamente discriminativos para imagens de textura.

Mais recentemente, vários trabalhos exploraram as RNNs como um potencial extrator de características quando combinados com outros métodos da literatura. Em (RIBAS et al., 2024b) foi proposto uma nova abordagem que modela a imagem sob a forma de uma rede complexa dirigida, onde cada pixel é mapeado para um vértice, e os pixels adjacentes são empregados como vetores de entrada da RNN. Os pesos aprendidos na camada de saída formam o vetor de características, resultando em uma abordagem local mais rápida e precisa do que métodos anteriores baseados apenas na dinâmica da rede complexa.

Seguindo essa mesma abordagem, (RIBAS; BRUNO, 2024) utiliza da combinação de redes complexas com RNNs para explorar a classificação de formas. Gerando uma rede conectada a partir de pixels e seus vizinhos, que com sua evolução dinâmica, permite calcular medidas topológicas que refletem propriedades estruturais do contorno. Com isso, aplica-se a RNN para aprender as representações discriminativas destas medidas, apresentando assim uma abordagem robusta a ruídos e invariância a transformações em escala e orientação de formas.

Embora os trabalhos mencionados anteriormente concentrem-se na análise de imagens binárias, abordagens mais recentes têm buscado incorporar não somente uma análise estrutural de textura, mas também em relação a cores. (RIBAS et al., 2024a) (FARES; RIBAS, 2024) trazem uma abordagem espectro-espacial de imagens de textura. O primeiro propõe uma representação de textura que integra análises espaciais e espectrais por meio da combinação de RAEs, permitindo a extração de descritores robustos mesmo em cenários com conjuntos de dados limitados. Já o segundo introduz um modelo baseado em redes complexas espectro-espaciais, no qual uma RNN é utilizada para aprender padrões espaciais e espectrais a partir da topologia da rede, resultando em representações altamente discriminativas. Ambos os trabalhos destacam-se pela aplicação bem-sucedida em tarefas práticas como a identificação de espécies vegetais, evidenciando o potencial das abordagens espectro-espaciais na análise de imagens coloridas.

2.3 Caminhadas Determinísticas Parcialmente Auto-Repulsivas

As Caminhadas Determinísticas Parcialmente Auto-Repulsivas (CDPARs), podem ser interpretadas como um turista que deseja visitar pixels distribuídos em uma imagem. Este começa sua trajetória em um pixel p previamente estabelecida e segue uma regra determinística de movimento r : ir para o pixel de menor distância para este ($r = \min$) ou ir para o pixel de maior distância para este ($r = \max$). Assim, segue-se a trajetória minimizando ou maximizando as distâncias para os pixels que permaneceram fora da trajetória nos μ passos anteriores.

Esses passos definem o que chamamos de memória em CDPARs, os quais não podem

ser visitados até que saiam da memória. Esta é atualizada a cada passo dado pelo turista para salvar os novos últimos pixels visitados. Dessa forma, o turista ao fim realiza um caminho que pode ser separado em duas etapas: uma parte inicial, de extensão τ , denominada transiente, e uma parte final conhecida como atrator, que pode ser composta por pixels que formam um ciclo de período $\rho \geq \mu + 1$ (RIBAS, 2017). O atrator é um ciclo de pixels de período ρ sempre visitado pelo turista, onde uma vez que o este entre, fique impossibilitado de sair. O transiente determina a trajetória do turista até achar um atrator, sendo τ o número de pixels percorridos desde o início da caminhada até o primeiro pixel do atrator. Porém, com isso, existem casos os quais o turista não acha um atrator, assim a caminhada apenas é composta pelo transiente τ .

Assim, em uma imagem bidimensional de dimensões H e W , representando a altura e comprimento da imagem respectivamente, o conjunto finito de pixels P que a compõem pode ser definido formalmente por:

$$P = \{(x, y) \in \mathbb{N}^2 \mid 1 \leq x \leq W \text{ e } 1 \leq y \leq H\}, \quad (2.4)$$

Cada pixel pode ser mapeado para $i = (x_i, y_i)$ e um grau de intensidade do pixel como $I(i) \in [0, 255]$. Com isso, a vizinhança $\eta(i)$ de um dado pixel i é composta pelos pixels j , cuja distância euclidiana para i é menor ou igual que $\sqrt{2}$, assim $\eta(i)$ retorna a vizinhança 8 de um pixel localizado em (x_i, y_i) para realizar um passo da caminhada, como mostra a equação 2.5.

$$\eta(i) = \{j \in P \mid \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \leq \sqrt{2}\} \quad (2.5)$$

A distância referida dentre dois pixels vizinhos i e j é dada pelo módulo da diferença das suas intensidades de cinza, assim como é descrito na equação 2.6.

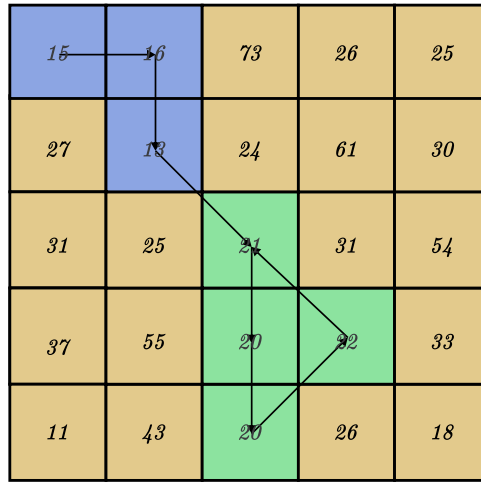
$$\omega(i, j) = |I(i) - I(j)| \quad (2.6)$$

Assim, podemos definir matematicamente as duas principais funções de regras de passo de uma CDPAR utilizada nesse trabalho, mais especificamente $f_{\max}$ e $f_{\min}$, baseadas na regra determinística de movimento r , referidas como regra-max e regra min:

$$\begin{aligned} f_{\max}(i) &= \underset{j \in \eta(i) \setminus M}{\operatorname{argmax}} \omega(i, j) \\ f_{\min}(i) &= \underset{j \in \eta(i) \setminus M}{\operatorname{argmin}} \omega(i, j), \end{aligned} \quad (2.7)$$

denotando-se que $f_{\max}(i)$ ao maximizar as diferenças favorece padrões mais heterogêneos de textura, enquanto $f_{\min}(i)$ dá preferência para padrões mais homogêneos.

Figura 3 – Exemplo de CDPAR com tamanho de memória $\mu = 2$ e regra $r = \min$ em uma imagem 5x5. Cada número representa a intensidade de um pixel. Em azul temos o transiente e em verde o atrator referentes a caminhada



Fonte: Elaborado Pelo Autor

Para cada pixel de uma imagem, uma caminhada se inicia com um tamanho de memória μ e regra de movimento r , produzindo diferentes trajetórias. Assim, a Figura 3 exemplifica a trajetória de uma CDPAR realizada em uma imagem. Neste exemplo, os pixels são por seus valores de intensidades de cinza nos quadrados. Assim, o turista sai do primeiro pixel da imagem, caminhando até encontrar um atrator. Dessa forma, obtemos em azul o período do transiente ($\tau = 3$) e período do ciclo do atrator ($\rho = 5$) visualizado em verde.

Ressaltando que para a identificação de um atrator é necessário que em uma dada caminhada, o turista dê um passo para um mesmo pixel 3 vezes, o algoritmo de detecção verifica se os três últimos passos de visita a este pixel, denotados como t_1 , t_2 e t_3 , onde $t_1 < t_2 < t_3$, satisfazem a condição de periodicidade $\Delta t = t_2 - t_1 = t_3 - t_2$, garantindo intervalos constantes entre as visitas. Adicionalmente, o sistema verifica se as sequências de pixels visitados entre os intervalos $[t_1, t_2)$ e $[t_2, t_3)$ são idênticas por meio da comparação direta das trajetórias. Quando ambas as condições são atendidas simultaneamente, periodicidade temporal e consistência de trajetória, o algoritmo confirma a existência de um ciclo determinístico, caracterizando um atrator periódico de período Δt , indicando que o turista entrou em um comportamento cíclico previsível que se repetirá indefinidamente.

A Caminhada Determinística Parcialmente Auto-Repulsiva foi introduzida em (LIMA; MARTINEZ; KINOCHI, 2001), surgindo como um método para analisar o efeito de caminhadas em ambientes aleatórios e desordenados. Assim, sendo possível extrair características do ambiente em que são realizadas.

Posteriormente, as CDPARs começaram a se consolidar como uma abordagem promissora para problemas de visão computacional, obtendo desempenho relevante na análise de texturas tanto estáticas quanto dinâmicas. Sendo, o primeiro trabalho utilizando estas

para analisar imagens, (CAMPITELI; MARTINEZ; BRUNO, 2006), onde foram utilizados os períodos de atratores e transientes para caracterizar texturas. Com isso, (BACKES et al., 2006b) traz uma pequena variação do método, combinando as regras determinísticas, min e max, da caminhada, além de um novo atributo estatístico extraído da distribuição gerada.

Em (BACKES et al., 2010) foi resolvido o problema de empate na escolha do próximo pixel a ser percorrido na caminhada ao utilizar uma nova regra determinística de movimento, pautando-se na direção do contraste na vizinhança de um pixel. Essa mudança trouxe ótimos resultados, avançando no aperfeiçoamento do método original para a análise de imagens de texturas. Em contrapartida, outros trabalhos abordavam combinações das CDPARs com outros métodos existentes, como em (GONÇALVES et al., 2012), onde foram combinadas juntamente a redes complexas para análise das texturas.

Mais recentemente, as CDPARs têm sido utilizadas para análise de formas em imagens, majoritariamente de plantas, como em (RIBAS; BRUNO, 2019a), em que, na mesma linha de combinar métodos, utilizou-se de redes complexas para modelar o contorno de formas de plantas, para assim aplicar as CDPARs nestas redes criadas, gerando uma assinatura significativa e detalhada destes formatos baseada nas trajetórias das caminhadas.

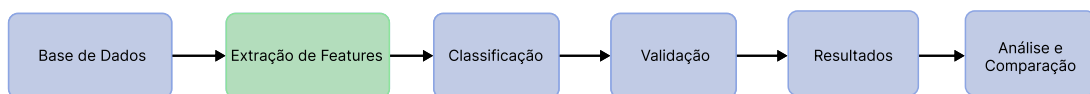
As Caminhadas também têm se mostrado eficazes para caracterização de análise de texturas dinâmicas, modelando cada pixel do vídeo como um vértice de um grafo e extraíndo medidas estatísticas das trajetórias geradas pelas caminhadas. Abordagens recentes, exploram modelagens espaciais e temporais distintas, bem como a aplicação das caminhadas em planos ortogonais, permitindo capturar simultaneamente propriedades de aparência e movimento com vetores de características compactos. (RIBAS; BRUNO, 2019b; RIBAS; BRUNO, 2020)

3 Metodologia

A metodologia do trabalho se baseia no fluxo exemplificado na Figura 4. No qual temos como o objetivo primário realizar a caracterização das texturas das imagens pertencentes as base dados, ressaltando que essa etapa é onde de fato esse trabalho atuou. Continuando com um processo posterior de classificação a partir de um modelo de aprendizado de máquina, junto com sua validação. Por fim, realizamos a análise dos resultados obtidos com a validação do modelo.

Com o fluxo para extração das características estabelecido, começamos os experimentos do trabalho. Inicialmente, cada imagem da base de dados é submetida a várias CDPARs, sendo uma para cada pixel existente na imagem. Dessa forma, conseguimos transformar as imagens em conjuntos de vetores de padrões locais, usando diversos parâmetros de caminhas para obter caracterizações cada vez mais representativas. Com esta geração de vetores locais, utilizamo-nos como entradas de um conjunto de RNNs independentemente, cada uma com pesos aleatórios diferentes. Cada RNN aprende com base nos padrões locais gerados pelas CDPARs. Salientando que esse vetores de padrões locais são exclusivos de cada imagem. Dessa forma, ao explorar diferentes ajustes de hiperparâmetros, como a quantidade de neurônios na camada oculta da rede, obtemos os pesos treinados das camadas de saída de cada RNN, os quais serão concatenados para formar uma representação final da imagem, utilizando técnicas de agregação para obter um único vetor. Assim, com esse processo, obtemos uma representação robusta e característica de cada imagem utilizada.

Figura 4 – Fluxo de Etapas de Execução



Fonte: Elaborado pelo Autor

Prosseguindo no fluxo, temos a etapa de classificação, que utiliza um algoritmo tradicional de aprendizado de máquina para classificar as representações robusta geradas pelo processo de extração de características. Assim, essas representações serão analisadas em tarefas de classificação. Neste, as representações de cada imagem e seus respectivos rótulos, nomes das espécies de plantas, são usados para treinar um algoritmo de classificação determinado para criar modelos de predição.

Como etapas finais, ocorrem a validação e a análise quantitativa dos resultados. Assim, obtemos a eficácia do modelo final a partir de métricas comumente utilizadas para problemas de classificação, como acurácia e desvio padrão. Com isso, conseguimos realizar comparações com outros métodos da literatura, incluindo redes neurais tradicionais, RNNs

como algoritmo de classificação, CDPARs de forma isolada, para verificar a superioridade da combinação de CDPARs com RNNs para extrair características de textura. Por fim, também utiliza-se de algumas métricas de desempenho para avaliar os resultados e a eficiência computacional destes.

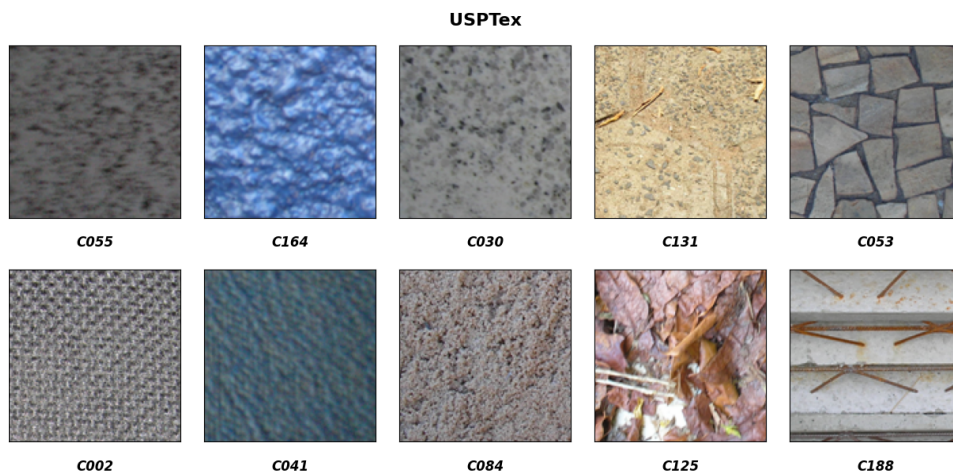
3.1 Base de Dados

As bases de dados de imagens representam um importante papel nos estudos de texturas, sendo utilizadas para estudar novos métodos e comparar resultados entre diferentes métodos da literatura para uma mesma atividade. Com isso, foram utilizadas as bases USPTex, 1200Tex e Outex, para essa tarefa.

3.1.1 USPTex

O conjunto de dados USPTex é dividido em 191 classes, com 12 amostras de cada classe, contribuindo para um total de 2292 imagens de texturas coloridas encontradas no dia a dia, como vegetações, nuvens, pisos etc. As imagens foram capturadas a partir de uma câmera digital e têm dimensões de 128 x 128 pixels (BACKES; CASANOVA; BRUNO, 2012a). Algumas amostras do USPTex podem ser vistas na Figura 5

Figura 5 – Exemplos de imagens de textura da USPTex



Fonte: Adaptado de (BACKES; CASANOVA; BRUNO, 2012b)

As texturas presentes na USPTex representam uma grande variedade de padrões naturais e materiais encontrados no cotidiano, incluindo:

- **Vegetações e elementos naturais:** diferentes tipos de plantas, folhas, gramas, musgo e outros materiais vegetais;
- **Objetos e materiais:** sementes, arroz, tecidos, papel;

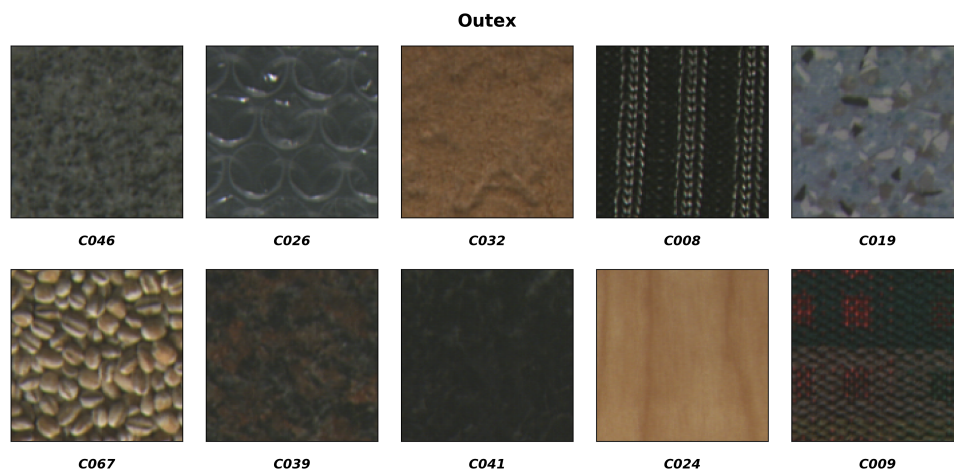
- **Estruturas urbanas e arquitetônicas:** pisos, paredes, concreto, tijolos, blocos;
- **Fenômenos naturais:** nuvens, solos, pedras, cascalhos, asfalto;
- **Superfícies variadas:** texturas de acabamento, revestimentos e pavimentos.

Esta diversidade de classes permite que a base seja utilizada para avaliar a capacidade de descritores de textura em discriminar entre um amplo espectro de padrões visuais encontrados em aplicações reais.

3.1.2 Outex

Outex é composta por diferentes imagens da natureza como também de texturas de superfícies, que foram capturadas com variações de direção da iluminação e rotações. Assim, esse conjunto de dados é composto por 68 classes de imagens de texturas coloridas, contendo 54 amostras para cada classe(OJALA et al., 2002). Na Figura 6, podemos observar algumas amostras de imagens da Outex.

Figura 6 – Exemplos de imagens de textura da Outex



Fonte: Adaptado de (OJALA et al., 2002)

As 68 classes de textura incluem materiais naturais e manufaturados, representando uma ampla variedade de superfícies reais:

- **Tecidos e fibras:** lona, carpete, papel;
- **Materiais rígidos:** azulejo, granito, borracha, espuma;
- **Materiais soltos:** areia, madeira, cevada e arroz;
- **Superfícies naturais e artificiais:** várias combinações de rugosidade e padrão de superfície.

Essas texturas foram selecionadas para refletir superfícies encontradas em ambientes reais, com diferentes propriedades de reflexão, rugosidade e padrão visual.

3.2 Classificadores

Para a realização das etapas de classificação e validação será usado um algoritmo de classificação supervisionado, que é alimentada com os vetores finais de características gerados pelas RNNs como entrada, que representam, bem como os respectivos rótulos das imagens, no exemplo de plantas, as espécies dessas. Dessa forma, realiza-se o devido processo de treinamento do algoritmo e posteriormente a validação do mesmo. Para realizar tal tarefa, nesse trabalho, foi escolhido o *Linear Discriminant Analysis* (LDA).

O *Linear Discriminant Analysis* (LDA) é um método estatístico que permite projetar os dados para maximizar a distância entre as classes e minimizar a variância dentro da mesma classe. Para isso, busca encontrar combinações lineares das características de entrada, derivando em superfícies de decisão. Sendo assim, o algoritmo foca primariamente em encontrar esses limites, com isso, suponha que tenhamos um conjunto de treinamento $x_1, \dots, x_n$, em que cada um é atribuído a uma de duas classes, C1 ou C2 (WEBB, 2002). Com isso, procuramos por um vetor de peso w e um limiar w_0 em que:

$$w^T x + w_0 \begin{cases} > 0 \\ < 0 \end{cases} \implies x \in \begin{cases} C1 \\ C2 \end{cases} \quad (3.1)$$

onde basicamente a equação 3.1 define uma combinação linear dos dados de entrada que serve para criar a fronteira entre os dados, conhecida também como um hiperplano. o vetor w define a orientação desse hiperplano e o quanto dista da origem é definido pelo limiar w_0 (WEBB, 2002). Se um dado de conjunto de treinamento x for colocado equação e resultar em um número positivo, esse pertence à classe 1; caso contrário, pertence à classe 2.

Para determinar o valor do vetor de peso w pode ser utilizado o critério de Fisher, um algoritmo de otimização, que busca encontrar uma direção de w que maximiza a razão entre as variâncias entre as classes e dentro das classes:

$$J_F = \frac{|w^T(m_1 - m_2)|^2}{w^T S_W w} \quad (3.2)$$

onde m_1 e m_2 são médias das amostras C1 e C2 e S_W é a matriz de covariância da amostra dentro da classe agrupada, ou seja, uma matriz que estima a dispersão média dos dados dentro de cada classe, quando assumimos que essa dispersão é a mesma para todas as classes (WEBB, 2002).

Portanto, se assumirmos que os dados estão distribuídos normalmente com matrizes de covariância iguais, então o problema é considerado linearmente separável e assim podemos

convergir que a solução ótima para a melhor direção de w é:

$$w = S_W^{-1}(m1 - m2) \quad (3.3)$$

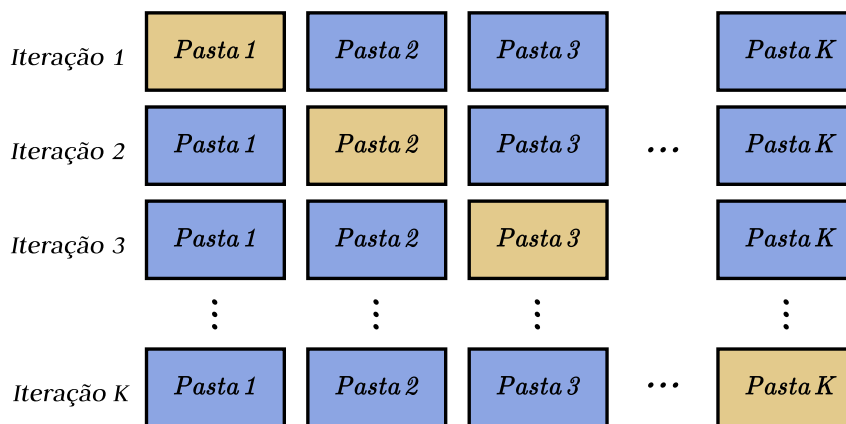
assim, temos as componentes necessárias para a classificação das classes no LDA e melhor direção é a linhas que consegue unir os centros (médias) das duas classes, levando em conta a dispersão interna dos mesmos.

3.3 Validação

Para validar a solução final do método de análise de imagens de textura, bem como o algoritmo de aprendizado de máquina, realiza-se diversas experimentações com diferentes conjuntos de dados de imagens. Para garantir um método rebuscado de classificação do método é utilizado um processo de validação cruzada com K -fold.

Essa técnica avalia a habilidade do algoritmo de aprendizado de máquina em produzir bons resultados para amostras que não fizeram parte do treinamento. Para isso, divide-se o conjunto de dados é particionado aleatoriamente em K partes independentes, cada uma com tamanho aproximado, os quais chamamos de pastas. Assim, realizam-se K ciclos de treinamento do modelo, em cada iteração do processo uma das K pastas é utilizada como objeto de teste e as demais $K - 1$ pastas sendo objetos de treinamento do modelo(WONG, 2015).

Figura 7 – Diagrama Esquemático da Validação Cruzada



Fonte: Elaborado pelo Autor

A Figura 7, mostra esse processo como um todo, onde em laranja temos que a pasta está sendo utilizada como subconjunto de validação, enquanto as pastas em azul representam as pastas utilizadas como subconjuntos de treinamento, todos com a mesma quantidade de dados. Dessa forma, para conseguirmos medir a performance da classificação do algoritmo, realiza-se a média das acurácias de cada iteração K (WONG, 2015).

3.4 Métricas de Avaliação

Por último, afim de termos parâmetros que nos permitam realizar comparações entre métodos já estabelecidos na literatura e também avaliar o próprio método proposta neste trabalho, iremos utilizar algumas métricas de avaliações estatísticas, sendo elas acurácia e desvio padrão.

3.4.1 Matriz de Confusão

Para avaliar e extrair métricas de tarefas de classificação é comumente usada uma estrutura chama matriz de confusão, que fornece uma visualização clara dos resultados do algoritmo, indicando seus erros e acertos, ao comparar os resultados esperado com as predições realizadas por este. Essa Matriz pode ser exemplificada pela Figura 8.

Figura 8 – Matriz de Confusão

		<i>Predito</i>	
		<i>Positivo</i>	<i>Negativo</i>
<i>Real</i>	<i>Positivo</i>	<i>Verdadeiro Positivo(VP)</i>	<i>Falso Negativo(FN)</i>
	<i>Negativo</i>	<i>Falso Positivo(FP)</i>	<i>Verdadeiro Negativo(VN)</i>

Fonte: Elaborado pelo Autor

Como podemos ver na Figura 8 a matriz de confusão representa em cada linha as instâncias de uma classe real e em cada coluna as instâncias de uma classe predita pelo algoritmo. Com isso, esta é estruturada em 4 componentes, os quais representam as contagens dos seguintes resultados:

- **Verdadeiro Positivo (VP):** Quantidade de instâncias positivas que foram rotuladas corretamente como positivas.
- **Verdadeiro Negativo (VN):** Quantidade de instâncias negativas que foram corretamente rotuladas como negativas
- **Falso Positivo (FP):** Quantidade de instâncias negativas que foram incorretamente rotuladas como positivas.
- **Falso Negativo (FN):** Quantidade de instâncias positivas que foram incorretamente rotuladas como negativas.

3.4.2 Acurácia

Com o conceito de matriz de confusão, podemos falar da acurácia de um algoritmo de classificação, a qual pode ser descrita como a razão do número total de classificações feitas corretamente por um algoritmo pelo número total de instâncias no conjunto de dados, como podemos ver na equação 3.4:

$$Acurácia = \frac{VP + VN}{VP + FP + VN + FN} \quad (3.4)$$

Como, neste trabalho, estamos lidando com um processo de validação cruzada, temos que para um algoritmo, com o objetivo de obter uma estimativa confiável, sua acurácia será a média das acurácias de cada iteração k do processo, assim como a equação 3.5 mostra:

$$\overline{Acurácia} = \frac{1}{k} \sum_{i=1}^k Acurácia_i \quad (3.5)$$

3.4.3 Desvio Padrão

O Desvio Padrão σ pode ser descrito como uma medida do quanto os valores de dados se desviam da média. Um desvio padrão mais baixo significa que os valores estão mais próximos da média, enquanto um valor mais alto de desvio padrão, significa que os valores de dados estão mais distantes da média. No caso, ao utilizar validação cruzada, o desvio padrão representa a variabilidade das acurácias entre as k diferentes pastas, o que nos indica a estabilidade do algoritmo de classificação, podendo ser calculado como:

$$\sigma = \sqrt{\frac{1}{k-1} \sum_{i=1}^k (acc_i - \overline{acc})^2} \quad (3.6)$$

4 Abordagem Proposta

Neste trabalho, propõe-se uma nova abordagem para extrair características de imagens de textura, assim utilizando CDPARs em combinação com RNNs para obter uma assinatura determinística de uma imagem. Com isso, o método se diferencia dos demais da literatura que utilizavam ou CDPARs ou RNNs para a extração de características, assim essa abordagem visa a criação de uma representação mais robusta e detalhada das texturas, aperfeiçoando a capacidade de discriminação destas.

A metodologia consiste basicamente em aplicar CDPARs partindo de cada pixel de uma imagem para obter os vetores locais que serão utilizados como entradas que alimentam o treinamento de múltiplos modelos de redes neurais independentemente, com distintos pesos iniciais, em seguida, agregar os pesos de saída aprendidos por essas redes. Dessa forma, explorando diferentes projeções do espaço de entrada, devido aos pesos aleatórios variados, assim permitindo a detectar uma vasta gama de características globais e locais das texturas.

Em adição ao método, é realizada uma diversificação de combinações de parâmetros variáveis, incluindo a quantidade de neurônios na camada oculta da RNN, tamanho máximo, tamanho da memória e a regra determinística das caminhadas. Dessa forma, também realiza-se a combinação de pesos de saída de diferentes redes neurais com quantidade de neurônios diferentes, como exemplo $Q = 5 + Q = 10$. Assim, esse processo de combinar diferentes configurações de parâmetros tem o objetivo de maximizar a capacidade de representação do método, criando uma técnica capaz de capturar tanto características mais discriminativas das texturas, quanto entradas diferentes.

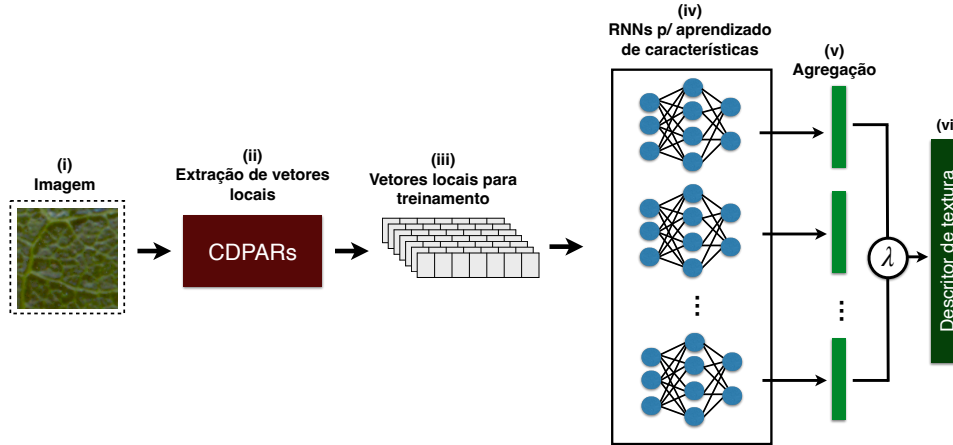
Durante a etapa de caracterização de texturas, o foco é aprimorar o aprendizado de características discriminativas para uma classificação mais eficaz. A investigação proposta explora o uso de CDPARs juntamente com RNNs com o objetivo de gerar descritores mais robustos, visando extrair informações contextuais e espaciais de maneira sistemática. A Figura 9 ilustra o fluxo geral para a extração de características, sobre o qual detalhamos a seguir cada passo.

4.1 Imagens de Entrada

Diferentemente de algoritmos convencionais de aprendizado de máquina para classificação de imagens, os quais são alimentados com várias imagens para seu treinamento. Aqui, o ponto inicial do fluxo é apenas uma única imagem, como é mostrado na Figura 9 (i). Com isso, o real objetivo é determinar uma representação específica de textura para cada imagem.

Dessa forma, ao utilizar uma única imagem, aplicando-se uma CDPAR partindo de

Figura 9 – Esquema da metodologia do projeto. No passo (ii) as CDPARs são usadas para analisar a complexidade e extrair informações locais da imagem; (iii) os vetores locais extraídos pelas CDPARs são usados para treinamento das RNNs em (iv); (v) usando técnicas de agregação os pesos da camada de saída das diferentes RNNs treinadas formam o vetor de características final em (iv).



Fonte: Adaptado de (VICENTIM, 2024)

cada pixel dessa, para extrair características complexas locais e ao final, essas caminhadas alimentando as RNNs, permitem que essas sejam treinadas com atributos singulares de uma única imagem. Assim, conseguimos capturar um nível de detalhamento de textura muito maior, do que poderia ser extraído utilizando um método convencional baseado em conjuntos extensos de imagens.

4.2 Extração de Vetores Locais

Conforme ilustrado na Figura 9 (ii), transforma-se uma imagem em um conjunto de vetores de padrões locais, que posteriormente vem a compor a entrada das RNNs, além de um conjunto de rótulos, os quais servirão como os alvos do treinamento supervisionado da rede. Esta etapa ilustra parte essencial do método, extraíndo padrões de textura complexos da imagem, permitindo assim, um detalhamento e especificidade local de textura.

Com isso, para realizar essa extração das informações locais são utilizadas as CDPARs. Assim, para cada pixel da imagem original inicia-se uma nova caminhada com tamanho máximo, regra e tamanho de memória determinados previamente. Dessa forma, obtemos um vetor local de atributos, o qual é definido pela trajetória da caminhada partindo de dado pixel e o pixel inicial da caminhada será utilizado como rótulo da RNN posteriormente. Ao realizar todas as caminhadas em uma imagem, extraímos uma matriz de dimensões $N \times p$, sendo N o número total de pixels em uma imagem e p , o tamanho máximo do período de uma caminhada.

Essa matriz, é representada na Figura 9 (iii) como sendo os vetores locais para treina-

mento das RNNs. Com essa aplicação conseguimos extrair informações de textura complexas de cada imagem e garantir um bom conjunto de características para alimentar as RNNs, a fim de obter uma assinatura descritiva de uma dada imagem.

Essa construção da matriz de vetores locais por meio das caminhadas pode ser descrita matematicamente. Para isso, utilizamos a notação formal das CDPARs, onde cada caminhada $W_f^{\mu,p}(u)$ iniciada em um pixel u gera uma sequência de pixels visitados. Como esses pixels pertencem ao conjunto P , precisamos mapeá-los para valores reais que possam ser processados pela rede neural. Definimos então $G_f^{\mu,p}(u)$ como o vetor de características codificável, que é um vetor p -dimensional de valores reais associado à caminhada $W_f^{\mu,p}(u)$. Formalmente:

$$G_f^{\mu,p}(u) = (g(w_1), g(w_p), \dots, g(w_p)) \in \mathbb{R}^{p \times 1}, \quad (4.1)$$

sendo μ a memória da CDPAR, p o tamanho máximo da CDPAR e f a função de regra. Assim, $G_f^{\mu,p}(u_i)$ representa o vetor de intensidades dos pixels visitados pela CDPAR iniciada no pixel u_i , capturando a trajetória da caminhada mediante os valores de intensidade.

A matriz de entrada para a RNN é então construída pela concatenação horizontal de todos esses vetores:

$$X_f^{\mu,p} = [G_f^{\mu,p}(u_1) \ G_f^{\mu,p}(u_2) \ \dots \ G_f^{\mu,p}(u_N)] \in \mathbb{R}^{p \times N} \quad (4.2)$$

sendo X , a matriz de entrada para a RNN, expressando a agregação de forma horizontal de todas as trajetórias de CDPARs em uma imagem.

4.3 Treinamento das RNNs

Com os vetores locais de treinamento preparados, separando-os em características de entrada da RNN, trajetórias das caminhadas e rótulos, pixel inicial da caminhada. A próxima etapa é o treinamento das RNNs. Assim, um conjunto de vetores locais utilizado para alimentar o aprendizado de uma RNN, sendo esta, inicializada com pesos iniciais aleatórios, explorando áreas distintas do espaço de hipóteses para diferentes conjunto de vetores.

Dessa forma, essa etapa é realizada para cada conjunto de vetores locais gerados pelas CDPARs em uma única imagem da base de dados. Com isso, ao final da etapa de treinamento, temos os pesos da camada de saída de cada RNN representando um vetor descritivo de cada imagem da base de dados com seus respectivos rótulos, que passarão por um processo de agregação para que assim obtenha-se o descritor final de textura.

4.4 Representações

Finalmente, para obter a representação final com os descritores de textura, é essencial condensar a representação gerada pela RNN em um único vetor que capture as características texturais de maneira abrangente (iv). Essa representação final é construída com base nos pesos aprendidos pela camada de saída da RNN, que representam as características extraídas dos padrões locais da imagem. A condensação é realizada por meio do cálculo de estatísticas das conexões de cada neurônio oculto com todas as saídas, resultando em um vetor de características compacto que preserva a informação essencial sobre os padrões texturais aprendidos.

Com isso, a camada de entrada da RNN é inicializada a partir de valores aleatórios. Para tal, é gerada uma sequência aleatória V , conforme a Equação 2.3, com comprimento E , sendo $E = Q \times (p + 1)$. Essa sequência V é então empregada diretamente na construção da matriz de pesos aleatórios W , que compõe a camada de entrada da rede.

$$W = \{V(1), \dots, V(E)\} \quad (4.3)$$

A fim de compreender a representação gerada pela rede para cada imagem, foi conduzido um estudo em que diferentes medidas estatísticas foram aplicadas sobre a matriz de pesos treinados M . As medidas consideradas foram: média aritmética, desvio padrão, curtose e assimetria, as quais permitem caracterizar distintos aspectos da distribuição dos pesos aprendidos pela rede.

Seja $\vec{x} = \{x_1, x_2, \dots, x_N\}$ um conjunto de N números reais. A média aritmética é definida por:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (4.4)$$

O desvio padrão é dado por:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2} \quad (4.5)$$

A curtose, que mede o grau de concentração dos valores em torno da média, é definida como:

$$\kappa = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^4}{\left(\sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2} \right)^4} \quad (4.6)$$

Por fim, a assimetria, que quantifica o grau de inclinação da distribuição em relação à média, é definida por:

$$\gamma = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^3}{\left(\sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2} \right)^3} \quad (4.7)$$

Essas medidas foram então aplicadas à matriz de pesos treinados M , de modo a gerar diferentes vetores de características que capturam propriedades complementares da rede. Para cada medida estatística calculada sobre M , obteve-se um vetor de características:

$$\beta_\mu = \mu(M), \quad \beta_\sigma = \sigma(M) \quad (4.4.8)$$

$$\beta_\gamma = \gamma(M), \quad \beta_\kappa = \kappa(M) \quad (4.4.9)$$

Assim, os vetores $\beta_\mu, \beta_\sigma, \beta_\gamma, \beta_\kappa$ representam, respectivamente, as informações de tendência central, dispersão, assimetria e concentração dos pesos de M .

Para formalizar esse processo, seja $\vec{m}_k$ a k -ésima coluna da matriz M e $s(\cdot)$ uma função de sumarização que mapeia um vetor em um número real. O vetor de características sumarizadas é então definido como:

$$\vec{\Theta}_{f,s}^{\mu,p}(Q) = [s(\vec{m}_1), s(\vec{m}_2), \dots, s(\vec{m}_{Q+1})], \quad (4.8)$$

onde s pode corresponder à média, desvio padrão, curtose ou assimetria.

Dessa forma, as diferentes funções de sumarização são aplicadas às colunas de M e seus resultados são concatenados para formar uma representação de textura mais robusta:

$$\vec{\Psi}_f^{\mu,p}(Q) = [\vec{\Theta}_{f,\mu}^{\mu,p}(Q), \vec{\Theta}_{f,\sigma}^{\mu,p}(Q), \vec{\Theta}_{f,\gamma}^{\mu,p}(Q), \vec{\Theta}_{f,\kappa}^{\mu,p}(Q)] \quad (4.9)$$

Como os padrões aprendidos em $\vec{\Psi}_f^{\mu,p}(Q)$ dependem fortemente da função de regra da caminhada f , diferentes regras podem ser combinadas a fim de enriquecer a representação final:

$$\vec{\Phi}_{f_1,f_2,\dots,f_n}^{\mu,p}(Q) = [\vec{\Psi}_{f_1}^{\mu,p}(Q), \vec{\Psi}_{f_2}^{\mu,p}(Q), \dots, \vec{\Psi}_{f_n}^{\mu,p}(Q)]. \quad (4.10)$$

O cálculo da média dos pesos visa reduzir as particularidades individuais das RNNs, gerando uma representação global mais consistente das texturas. O desvio padrão indica o grau de variação entre as representações aprendidas, enquanto a curtose e a assimetria descrevem a forma da distribuição dos pesos, permitindo identificar possíveis *outliers* e tendências de inclinação. Assim, a combinação dessas medidas busca ampliar o poder discriminativo entre diferentes tipos de textura, assegurando que a representação final capture de forma abrangente a diversidade das propriedades texturais.

5 Resultados

Nesse capítulo, são mostrados os resultados experimentais obtidos com base na aplicação do método de caracterização de texturas combinando Caminhadas Determinísticas Paricalmente Auto-Repulsivas e *Randomized Neural Networks*. Com isso, examinou-se o desempenho do modelo proposto a partir da sua habilidade em discriminar texturas. Esta avaliação considera o impacto das estratégias de agregação aplicadas, como média, desvio padrão, curtose e assimetria extraídas dos pesos de saída das RNNs. Ademais, são comparados os resultados para diferentes configurações de CDPARs: tamanho da memória, função de regra e tamanho máximo da caminhada, quantidade Q de neurônios, bem como suas combinações, também com as estatísticas.

5.1 Avaliação de Parâmetros

Com o intuito de avaliar a abordagem proposta, foram realizados testes com diversas combinações de parâmetros, visando explorar a representatividade do método e sua capacidade de discriminação. Assim, para as CDPARs, foram considerados 3 parâmetros: a memória μ , tamanho máxima da trajetória p e a função de regra de passo f . Já para as redes neurais, foi avaliado o parâmetro Q , representando o número de neurônios na camada escondida da RNN, que determina a dimensão do espaço em que o vetor de característica é projetado.

Esses parâmetros nos permitem observar como diferentes caminhadas iniciadas com diversas combinações de função de regra, tamanho máximo e memória, podem impactar na discriminação das imagens de textura. Para essa tarefa, escolheu-se os valores de Q no intervalo $\{9, 19, 29, 39, 49, 59\}$, afim de avaliar o melhor arranjo de neurônios da camada oculta da RNN, memória μ no intervalo de $\{1, 2, 3, 4, 5\}$, tamanho máximo p no intervalo $\{5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$ e função de regra $f \{min, max, minmax\}$, com o intuito de explorar como diferentes configurações de caminhadas afetam o resultado final do método.

Assim, na Tabela 1, são apresentados os cálculos das médias de acurácia de todas as combinações de parâmetros de caminhada para cada Q e conjunto de métricas estabelecido, nas bases Outex e UspTex, observando-se como estas configurações podem influenciar nos resultados finais obtidos. Dessa forma, é possível analisar que o desempenho do método tende a aumentar com o incremento de Q , onde foi possível obter os melhores resultados ao utilizar $Q = 59$, caracterizando por extrair mais características das texturas.

Entre as métricas de agregação, todas as métricas individualmente tiveram um crescimento constante em relação ao parâmetro Q , com ressalvas para quando $Q = 49$, onde há

Tabela 1 – Resultados obtidos para diferentes valores de Q em dois datasets

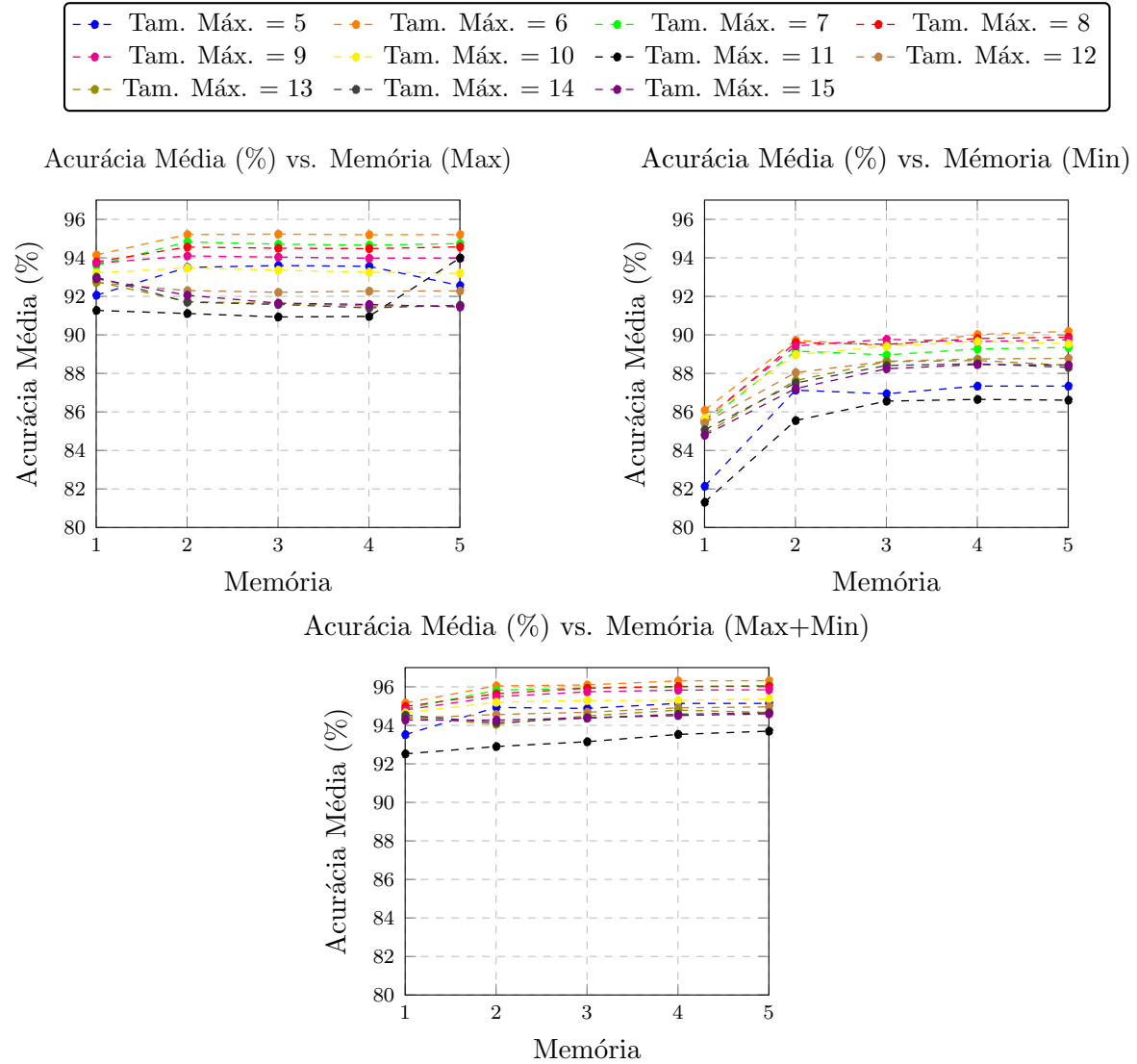
Medidas	Dataset	Q					
		9	19	29	39	49	59
Média	Usptex	85,91	88,95	91,06	91,62	91,39	92,23
	Outex	86,97	87,98	88,72	89,11	88,92	89,21
Desvio	Usptex	83,84	85,80	88,89	89,77	88,32	90,12
	Outex	81,16	82,73	84,70	85,47	84,31	85,74
Curtose	Usptex	69,65	78,53	84,86	87,63	85,97	89,26
	Outex	77,08	81,27	83,92	85,25	84,36	86,00
Assimetria	Usptex	70,80	78,93	84,95	87,54	86,10	89,35
	Outex	77,74	81,78	84,17	85,23	84,43	86,27
Média + Desvio	Usptex	89,47	90,66	91,78	92,09	91,71	92,20
	Outex	88,39	89,21	89,96	90,20	89,77	89,99
Média + Assimetria	Usptex	89,48	91,02	92,24	92,64	91,93	92,85
	Outex	88,69	89,64	90,40	90,69	90,10	90,67
Média + Curtose	Usptex	89,35	90,85	92,12	92,59	91,85	92,77
	Outex	88,55	89,38	90,30	90,62	89,98	90,56
Média + Desvio + Assimetria + Curtose	Usptex	93,05	94,29	95,44	95,83	94,88	96,00
	Outex	89,75	90,24	90,98	91,09	90,13	90,56

uma pequena saturação da rede. Analisando as combinações das métricas, é possível ver que a adição da média para as outras métricas traz um crescimento significativo de performance para o método. Assim, os melhores resultados foram obtidos com a concatenação das 4 métricas de agregação $[\beta_\mu, \beta_\sigma, \beta_\gamma, \beta_\kappa]$, evidenciando que a integração de múltiplos momentos melhora a discriminação dos padrões de textura, corroborando com a escolha da agregação do método final utilizada na abordagem proposta para aplicação.

Ademais, considerou-se o a média das acurácias dos resultados para realizar a análise dos parâmetros μ e p . Nesse sentido, foi realizado o cálculo da média das acurácias de diferentes valores de Q para cada combinação de memória μ e tamanho máximo da caminhada m . A partir disso, também realizou-se essa mesma análise para diferentes regras, resultando na análise para regra min, max e sua combinação max + min, a combinação das características das duas regras.

Como representado na Figura 10, para todas as configurações o aumento do tamanho da memória tende a melhorar o desempenho do método, com estabilização a partir de $\mu = 3$. Isso ocorre porque valores maiores de memória permitem que a caminhada explore um contexto espacial mais amplo, capturando padrões de textura mais complexos. A regra max+min, que combina as trajetórias geradas pelas regras de maximização e minimização das diferenças de intensidade, apresentou os melhores resultados globais, indicando maior capacidade discriminativa ao integrar informações de regiões homogêneas e heterogêneas da textura. Já a regra max manteve desempenho superior à min, mostrando-se mais sensível às variações locais de intensidade. No geral, os resultados mostram que o uso combinado das regras e um valor

Figura 10 – Comparação de acurácia média versus memória para diferentes regras em Usptex

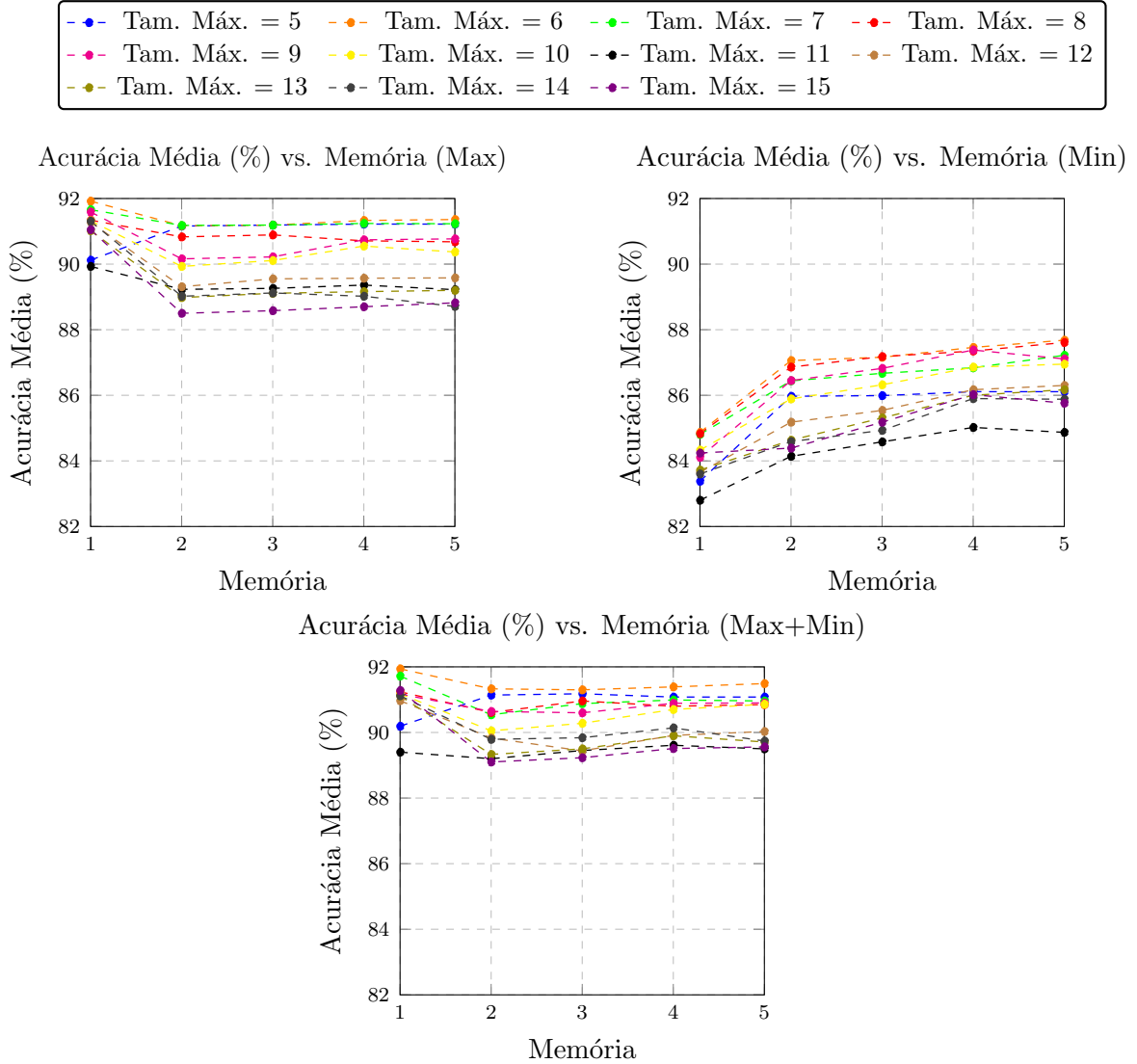


Fonte: Elaborado pelo Autor

intermediário de memória fornecem o melhor equilíbrio entre detalhamento local e estabilidade da caracterização das texturas na base USPTex. Ressaltando-se, que os melhores resultados foram obtidos com CDPARs de tamanho máximo $p = 6$, sendo esse superior independente dos tamanhos de memórias implicados. Pode ser evidenciado também, que tamanhos menores de CDPARs performam resultados melhores, isso ocorre pelo fato de que quanto menor uma caminhada é, mais preciso vai ser o padrão local de textura extraído da imagem.

A Figura 11 apresenta os resultados de acurácia média obtidos para a Outex, considerando a variação do parâmetro de memória e diferentes valores de caminho máximo nas três regras de caminhada (max, min e max+min). Os resultados obtidos podem ser considerados semelhantes aos da Usptex, novamente verificando que a tendência de aumento das acurácias é ditada pelo incremento dos valores de memória. A regra $max + min$ apresentou os maiores

Figura 11 – Comparação de acurácia média versus memória para diferentes regras em Outex

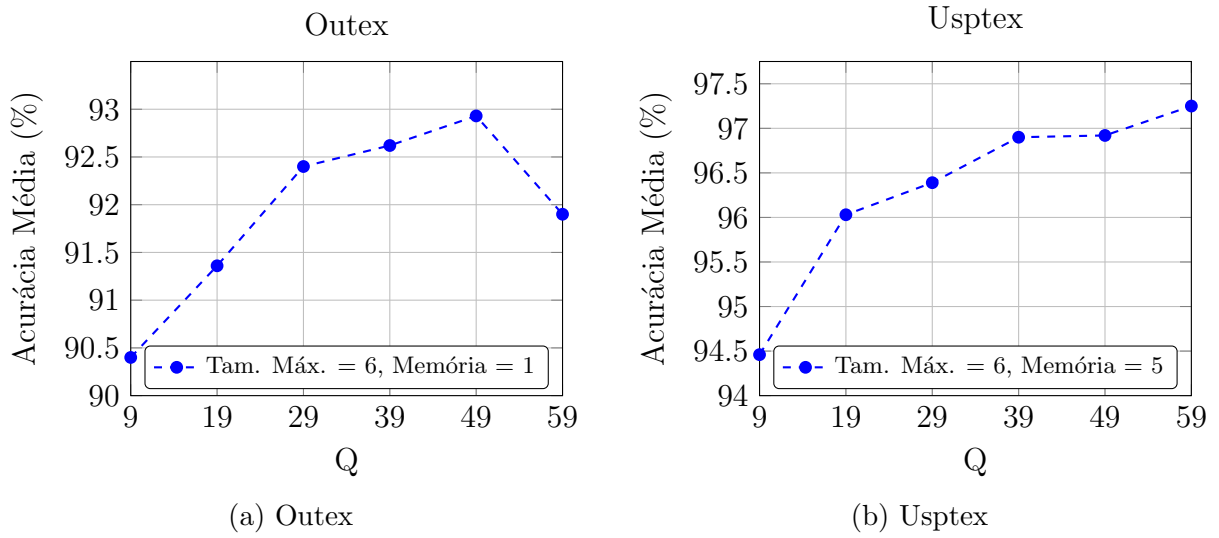


Fonte: Elaborado pelo Autor

resultados, evidenciando que a combinação das estratégias de maximização e minimização potencializa a extração de informações complementares da textura. Contudo, o ganho absoluto de desempenho na Outex é ligeiramente inferior ao obtido no USPTex, indicando que este conjunto de dados, por apresentar texturas mais heterogêneas e com maior variação de iluminação, representa um cenário de classificação mais desafiador. Assim como na Usptex, o tamanho de caminhada $p = 6$, apresentou os melhores resultados finais, com a ressalva que, para a Outex, a memória $\mu = 1$ teve resultados mais satisfatórios com a inclusão da regra max, enquanto para regra min se manteve a tendência de crescimento.

Por fim, foi realizada a análise de como os diferentes valores de Q se comportam para determinados valores de parâmetros das CDPARs, utilizando a regra combinada max+min com os parâmetros otimizados de tamanho máximo $p = 6$ e memória $\mu = 1$ para Outex e $\mu = 5$

Figura 12 – Comparação da Acurácia Média versus Q para os datasets Outex e Usptex usando a regra Max + Min.



Fonte: Elaborado pelo Autor

para Usptex. A partir da Figura 12a, pode-se observar que para a base Outex, à medida que o valor de Q aumenta, a acurácia tende a crescer drasticamente nos valores menores, até que atinge um ponto de $Q = 29$ até $Q = 49$, onde os resultados atingiram pequenos aumentos e uma queda grande de desempenho quando chega em $Q = 59$. Este comportamento sugere que a base Outex apresenta maior sensibilidade à configuração do parâmetro Q , possivelmente devido à maior complexidade e heterogeneidade de suas texturas. Na base Usptex 12b, observa-se uma tendência consistente de crescimento da acurácia à medida que Q aumenta de 9 para 59 neurônios, partindo de aproximadamente 94,5% e alcançando cerca de 97%, com uma leve queda em $Q = 49$ seguida de recuperação, confirmando que $Q = 59$ representa a melhor configuração para esta base.

É pertinente notar que, enquanto a Tabela 1 oferece uma visão abrangente do impacto de Q em diversas configurações de caminhada e métricas de agregação, os gráficos da Figura 12 aprofundam essa análise. Eles ilustram o comportamento da acurácia especificamente em função de Q , já considerando os parâmetros de CDPARs previamente otimizados para cada conjunto de dados (tamanho máximo da caminhada $p = 6$ e memória $\mu = 1$ para Outex; $p = 6$ e $\mu = 5$ para Usptex), e empregando a regra combinada Max + Min. Essa distinção permite uma compreensão refinada do impacto de Q sob as condições mais promissoras identificadas para os outros hiperparâmetros.

Tabela 2 – Comparação de acurácias de diferentes métodos de análise de textura em três bancos de dados de texturas.

Métodos	Nº de features	Outex	USPTex
GLDM (WESZKA; DYER; ROSENFELD, 1976a)	60	86,76	92,06
Gabor Filters (MANJUNATH; MA, 1996)	48	81,91	89,22
Fourier (WESZKA; DYER; ROSENFELD, 1976a)	63	81,91	67,50
Fractal (BACKES; CASANOVA; BRUNO, 2009)	69	80,51	78,27
LBP (OJALA et al., 2002)	256	81,10	85,43
CLBP (GUO; ZHANG; ZHANG, 2010)	648	85,80	91,14
AHP (ZHU et al., 2015)	120	88,31	94,85
LCP (GUO; ZHAO; PIETIKÄINEN, 2011)	81	86,25	91,14
LFD (MAANI; KALRA; YANG, 2013)	276	82,57	83,55
LPQ (OJANSIVU; HEIKKILÄ, 2008)	256	79,41	85,12
ELM Signature (JUNIOR; BACKES, 2016)	180	89,71	95,11
CNTD (BACKES; CASANOVA; BRUNO, 2013a)	108	86,76	91,71
CNRN (RIBAS et al., 2020b)	240	91,32	96,95
SSR ¹ (RIBAS et al., 2024a) (gray-scale)	630	90,80	95,80
VGG19 (SIMONYAN; ZISSERMAN, 2014)	512	76,62	93,19
InceptionV3 (SZEGEDY et al., 2016)	2048	86,40	96,77
ResNet50 (HE et al., 2016)	2048	65,66	62,30
InceptionResNetV2 (SZEGEDY et al., 2017)	1536	85,88	96,34
Abordagem proposta			
$\vec{\Phi}_{f_{min}, f_{max}}^{1,6}(49)$	400	92,93	96,92
$\vec{\Phi}_{f_{min}, f_{max}}^{5,6}(59)$	480	91,90	97,25

5.2 Comparações com a Literatura

Com o objetivo de avaliar o desempenho da abordagem proposta, foram comparadas as acurácias obtidas pela combinação entre Caminhadas Determinísticas Parcialmente Auto-Repulsivas (CDPARs) e *Randomized Neural Networks* (RNNs) com diferentes métodos clássicos e modernos de análise de texturas descritos na literatura. Essa etapa teve como finalidade verificar se a integração entre ambas as técnicas é capaz de produzir representações mais discriminativas, mantendo um custo computacional reduzido e um número compacto de características.

Para essa avaliação, utilizaram-se os melhores parâmetros identificados nas análises anteriores. Na base Outex, adotaram-se: tamanho máximo $p = 6$, memória $\mu = 1$, regra $min + max$ e quantidade de neurônios na camada oculta da RNN $Q = 49$; enquanto na base Usptex, os parâmetros foram: tamanho máximo $p = 6$, memória $\mu = 5$, mesma regra $min + max$ e $Q = 59$ neurônios. Essa configuração foi selecionada por apresentar o melhor equilíbrio entre acurácia e dimensionalidade do vetor de características, sendo empregada na comparação com diferentes métodos da literatura relacionados à análise de texturas, podendo ser visto na Tabela 2.

De modo geral, a abordagem proposta apresentou desempenho superior em relação a diversos descritores tradicionais, como GLDM, LBP, CLBP, Gabor Filters e Fourier. Mesmo com um número reduzido de características, o método alcançou acurácias mais elevadas, evidenciando a eficácia da combinação entre CDPARs, responsáveis por explorar a complexidade estrutural das imagens e RNNs, que aprendem representações não lineares de alta variabilidade.

Outro aspecto relevante é que, diferentemente de redes neurais profundas, que usualmente demandam grandes volumes de dados, alto poder computacional e longos tempos de treinamento, a metodologia proposta pode ser executada em ambientes de processamento modesto, sem comprometer o desempenho. Essa característica torna o método acessível e eficiente para aplicações práticas, sobretudo em contextos com restrições de recursos computacionais.

Além disso, ao empregar o classificador LDA, o mesmo utilizado em diversos trabalhos da literatura, as comparações mantêm-se consistentes e justas, permitindo avaliar de forma direta a capacidade descritiva dos vetores de características. Nas bases analisadas, a combinação de CDPARs e RNNs resultou em acurácias superiores às obtidas por descritores manuais e métodos estatísticos clássicos, consolidando o potencial da integração proposta.

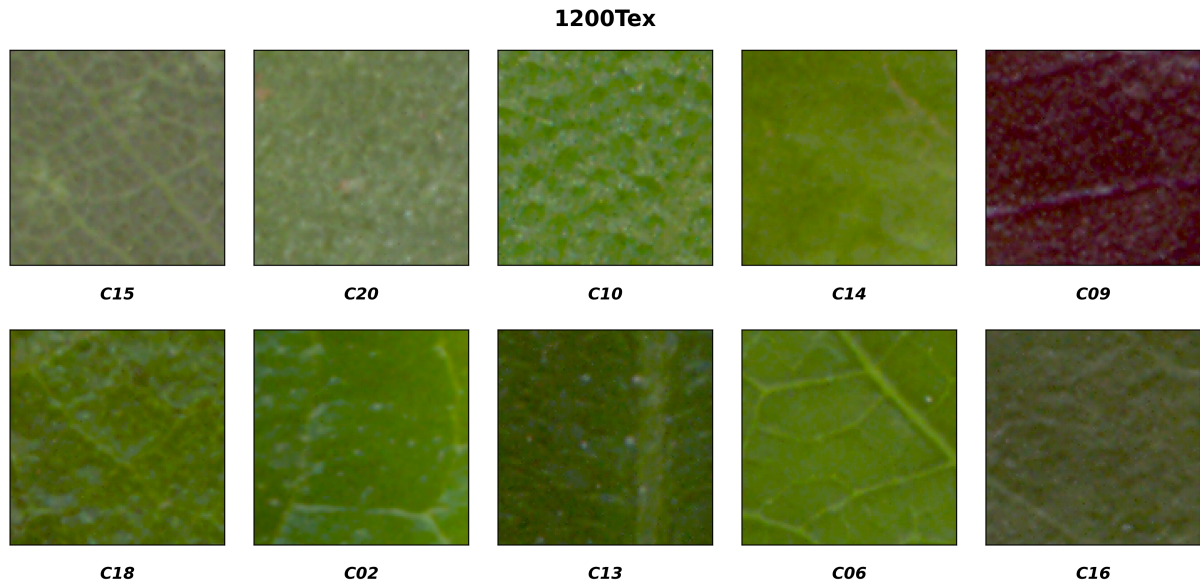
Por fim, o ganho observado nas bases Usptex e Outex reforça a robustez do método ao lidar com diferentes tipos de textura, desde padrões naturais e heterogêneos até texturas com alta variação. Esses resultados confirmam que a integração entre CDPARs e RNNs constitui uma alternativa promissora para a caracterização de texturas, oferecendo um equilíbrio adequado entre desempenho, interpretabilidade e eficiência computacional.

5.3 Identificação de Espécies de Plantas Brasileiras

O 1200Tex é um conjunto de dados voltado para o estudo do reconhecimento de espécies de plantas brasileiras a partir de padrões de textura foliar. Esse banco de imagens foi construído a partir da digitalização de folhas em alta resolução, gerando recortes texturais no formato RGB com dimensão de 128×128 pixels, como pode ser observado na Figura 13. No total, o dataset reúne 1.200 imagens distribuídas em 20 classes, cada uma correspondendo a uma espécie vegetal distinta e possuindo 60 amostras. O objetivo do uso dessa base é possibilitar a análise e a diferenciação das espécies por meio das variações de textura presentes nas superfícies foliares, oferecendo um cenário adequado para avaliar a precisão do método de reconhecimento desenvolvido neste trabalho.

Os resultados apresentados na Tabela 3 demonstram que o método proposto superou consistentemente os principais descritores clássicos e modelos de aprendizado profundo utilizados na literatura para o reconhecimento de espécies vegetais a partir de texturas foliares. Enquanto abordagens consagradas, como o descritor de Fourier e os filtros de Gabor, apresentaram desempenho inferior, o método proposto atingiu uma acurácia de 88,39%, su-

Figura 13 – Exemplos de imagens de textura da 1200Tex



Fonte: Adpatado de (CASANOVA; JUNIOR; BRUNO, 2009)

Tabela 3 – Acurácia usando o classificador LDA na base de dados 1200Tex.

Métodos	Acurácias (%)
AHP (ZHU et al., 2015)	79,17
CNTD (BACKES; CASANOVA; BRUNO, 2013b)	83,33
ELM signature (JUNIOR; BACKES, 2016)	84,86
Fractal (BACKES; CASANOVA; BRUNO, 2009)	70,75
Fourier (WESZKA; DYER; ROSENFELD, 1976b)	65,75
Gabor (MANJUNATH; MA, 1996; DAUGMAN; DOWNING, 1995)	77,25
GLDM (WESZKA; DYER; ROSENFELD, 1976a)	79,92
LCP (GUO; ZHAO; PIETIKÄINEN, 2011)	76,58
LFD (MAANI; KALRA; YANG, 2013)	74,67
LPQ (OJANSIVU; HEIKKILÄ, 2008)	73,00
AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012)	84,92
VGG16 (SIMONYAN; ZISSERMAN, 2014)	87,00
VGG19 (SIMONYAN; ZISSERMAN, 2014)	83,92
Método Proposto	
$\vec{\Phi}_{f_{min}, f_{max}}^{4,5}(29)$	88,39

perando inclusive arquiteturas profundas como AlexNet e VGG19. Esse desempenho ressalta a capacidade do modelo em capturar informações estruturais e espectrais relevantes das folhas, mesmo sem a necessidade de redes extensas ou pré-treinadas em grandes volumes de dados.

Esse ganho de desempenho pode ser atribuído à complementaridade entre as CDPARs e as RNNs. As CDPARs, ao modelarem a trajetória determinística de caminhadas sobre a superfície foliar, fornecem uma representação detalhada das irregularidades e padrões intrínsecos à textura. Por sua vez, as RNNs atuam como mapeadores não lineares que aprendem relações complexas entre essas variações locais, gerando descritores altamente discriminativos. Assim, o método consegue lidar de forma eficiente com a variabilidade natural das amostras, causada por fatores como ruídos de digitalização, variações de cor e pequenas deformações nas folhas.

A configuração de parâmetros utilizada para obter os resultados satisfatórios dessa aplicação foram tamanho máximo da CDPAR $p = 5$ tamanho de memória $\mu = 4$, as regras *min* e *max* concatenadas, bem como a agregação de todas as métricas estatísticas utilizadas. Por fim, foi configurado, o número de neurônios na camada oculta da RNN $Q = 29$, totalizando em 240 características no vetor final do descritor de textura.

6 Conclusão

Neste trabalho de conclusão de curso, foi possível aprofundar o estudo de métodos voltados à análise de texturas em imagens e ao uso combinado das Caminhadas Determinísticas Parcialmente Auto-Repulsivas (CDPARs) e *Randomized Neural Networks* (RNNs) nessa tarefa. Inicialmente, desenvolveu-se uma investigação detalhada sobre o processo de caracterização de imagens texturizadas, abordando as principais técnicas de extração de características capazes de representar com fidelidade as propriedades visuais de uma textura. Em seguida, aprofundou-se o conhecimento sobre o funcionamento, o aprendizado e o treinamento das RNNs, com ênfase em sua aplicação para a geração de representações discriminativas. Além disso, foram estudadas estratégias de classificação supervisionada e validação cruzada, permitindo avaliar a eficiência, a robustez e a capacidade de generalização dos modelos propostos.

A principal contribuição deste trabalho foi a proposição de uma abordagem híbrida para análise de texturas, unindo o potencial das CDPARs à capacidade de aprendizado das RNNs. O método consiste em empregar as CDPARs para percorrer a imagem de maneira determinística, gerando trajetórias que descrevem a complexidade e a organização local dos padrões texturais. Essas trajetórias são então utilizadas como vetores de entrada para o treinamento de múltiplas RNNs, que aprendem relações não lineares entre as informações extraídas. Diversas medidas estatísticas, como média, desvio padrão, curtose e assimetria, foram investigadas para agregar os pesos aprendidos, buscando a combinação que oferecesse maior capacidade discriminativa entre classes.

Os resultados experimentais comprovaram que a integração entre CDPARs e RNNs elevou de forma significativa o desempenho na classificação de texturas quando comparada ao uso isolado de cada técnica. A abordagem proposta superou diversos métodos tradicionais amplamente citados na literatura, como LBP, GLDM e filtros de Gabor, e apresentou resultados próximos ou superiores aos de redes neurais profundas, porém com custo computacional consideravelmente menor. Nas bases *Outex* e *Usptex*, o método alcançou acurácias acima de 96%, enquanto na base *1200Tex*, utilizada para a identificação de espécies de plantas brasileiras, obteve 88,39% de acerto, demonstrando sua eficiência, estabilidade e capacidade de generalização.

Dessa forma, este trabalho de conclusão de curso não apenas contribuiu para o avanço das metodologias computacionais voltadas à análise de texturas, como também evidenciou o potencial da combinação entre métodos determinísticos e redes neurais aleatórias em aplicações práticas. A técnica proposta mostrou-se eficaz para o reconhecimento de padrões naturais complexos e para a identificação de espécies vegetais a partir das texturas foliares, ressaltando seu valor científico e ambiental.

Além disso, a realização deste trabalho teve importância significativa para a formação acadêmica do autor, proporcionando uma compreensão aprofundada sobre aprendizado de máquina, processamento de imagens e modelagem de sistemas complexos, preparando-o para estudos futuros. Por fim, novos experimentos estão sendo conduzidos para ampliar a aplicação do método proposto imagens hiperespectrais, bem como a elaboração de um artigo científico para submissão em conferência, reforçando a relevância e continuidade desta pesquisa no campo da análise de texturas e reconhecimento de padrões.

Referências

AGGARWAL, C. C. *Neural Networks and Deep Learning: A Textbook*. 2nd. ed. [S.l.]: Springer Publishing Company, Incorporated, 2018. ISBN 3319944622.

BACKES, A. R. et al. Deterministic tourist walks as an image analysis methodology based. In: *Progress in Pattern Recognition, Image Analysis and Applications*. Springer Berlin Heidelberg, 2006. v. 4225, p. 784–793. ISBN 978-3-540-46556-0. Disponível em: http://dx.doi.org/10.1007/11892755_81.

BACKES, A. R. et al. Deterministic tourist walks as an image analysis methodology based. In: *Progress in Pattern Recognition, Image Analysis and Applications*. Springer Berlin Heidelberg, 2006. v. 4225, p. 784–793. ISBN 978-3-540-46556-0. Disponível em: http://dx.doi.org/10.1007/11892755_81.

BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Plant leaf identification based on volumetric fractal dimension. *International Journal of Pattern Recognition and Artificial Intelligence*, World Scientific, v. 23, n. 06, p. 1145–1160, 2009.

BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Color texture analysis based on fractal descriptors. *Pattern Recognition*, Elsevier, v. 45, n. 5, p. 1984–1992, 2012.

BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Color texture analysis based on fractal descriptors. *Pattern Recognition*, v. 45, n. 5, p. 1984–1992, 2012.

BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Texture analysis and classification: A complex network-based approach. *Information Sciences*, Elsevier, v. 219, p. 168–180, 2013.

BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Texture analysis and classification: A complex network-based approach. *Information Sciences*, Elsevier, v. 219, p. 168–180, 2013.

BACKES, A. R. et al. Texture analysis and classification using deterministic tourist walk. *Pattern Recognition*, Elsevier, v. 43, n. 3, p. 685–694, 2010.

CAMPITELI, M.; MARTINEZ, A. S.; BRUNO, O. An image analysis methodology based on deterministic tourist walks. In: *Advances in Artificial Intelligence - IBERAMIA-SBIA 2006*. Springer Berlin Heidelberg, 2006, (Lecture Notes in Computer Science, v. 4140). p. 159–167. ISBN 978-3-540-45462-5. Disponível em: http://dx.doi.org/10.1007/11874850_20.

CASANOVA, D.; JUNIOR, J. J. de M. S.; BRUNO, O. M. Plant leaf identification using gabor wavelets. *International Journal of Imaging Systems and Technology*, Wiley Subscription Services, Inc., A Wiley Company, v. 19, n. 3, p. 236–243, 2009. ISSN 1098-1098. Disponível em: <http://dx.doi.org/10.1002/ima.20201>.

DAUGMAN, J.; DOWNING, C. Gabor wavelets for statistical pattern recognition. In: MIT PRESS. *The handbook of brain theory and neural networks*. [S.l.], 1995. p. 414–419.

FARES, R. T.; RIBAS, L. C. A new approach to learn spatio-spectral texture representation with randomized networks: Application to brazilian plant species identification. In: ILIADIS, L. et al. (Ed.). *Engineering Applications of Neural Networks*. Cham: Springer Nature Switzerland, 2024. p. 435–449. ISBN 978-3-031-62495-7.

- FARES, R. T. et al. Randomized encoding ensemble: A new approach for texture representation. 2024.
- FLORINDO, J. B. Dstnet: Successive applications of the discrete schroedinger transform for texture recognition. *Information Sciences*, Elsevier, v. 507, p. 356–364, 2020.
- GONÇALVES, W. N. et al. Texture descriptor based on partially self-avoiding deterministic walker on networks. *Expert Systems with Applications*, v. 39, n. 15, p. 11818 – 11829, 2012. ISSN 0957-4174. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0957417412001078>.
- GUO, Y.; ZHAO, G.; PIETIKÄINEN, M. Texture classification using a linear configuration model based descriptor. In: CITESEER. *BMVC*. [S.l.], 2011. p. 1–10.
- GUO, Z.; ZHANG, L.; ZHANG, D. A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, IEEE, v. 19, n. 6, p. 1657–1663, 2010.
- HE, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 770–778.
- HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: theory and applications. *Neurocomputing*, Elsevier, v. 70, n. 1, p. 489–501, 2006.
- JUNIOR, J. J. d. M. S.; BACKES, A. R. Elm based signature for texture classification. *Pattern Recognition*, v. 51, p. 395–401, 2016. ISSN 0031-3203.
- JUNIOR, J. J. d. M. S.; BACKES, A. R.; BRUNO, O. M. Randomized neural network based descriptors for shape classification. *Neurocomputing*, Elsevier, v. 312, p. 201–209, 2018.
- JUNIOR, J. J. M. S.; BACKES, A. R. ELM based signature for texture classification. *Pattern Recognition*, v. 51, p. 395–401, 2016.
- KASUN, L. L. C. et al. Representational learning with elms for big data. *IEEE Intelligent Systems*, v. 28, p. 31–34, 2013.
- KAYA, L. K. Y.; TEKIN, R. A computer vision system for the automatic identification of butterfly species via gabor-filter-based texture features and extreme learning machine: Gf+elm. *Tem Journal*, Citeseer, p. 13, 2013.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Advances In Neural Information Processing Systems*, p. 1–9, 2012. ISSN 10495258.
- LIMA, G. F.; MARTINEZ, A. S.; KINOUCHI, O. Deterministic walks in random media. *Phys. Rev. Lett.*, American Physical Society, v. 87, p. 010603, Jun 2001. Disponível em: <http://link.aps.org/doi/10.1103/PhysRevLett.87.010603>.
- MAANI, R.; KALRA, S.; YANG, Y.-H. Noise robust rotation invariant features for texture classification. *Pattern Recognition*, v. 46, n. 8, p. 2103–2116, 2013. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320313000344>.
- MANJUNATH, B. S.; MA, W.-Y. Texture features for browsing and retrieval of image data. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 18, n. 8, p. 837–842, 1996.

OJALA, T. et al. Outex-new framework for empirical evaluation of texture analysis algorithms. In: IEEE. *2002 International Conference on Pattern Recognition*. [S.l.], 2002. v. 1, p. 701–706.

OJANSIVU, V.; HEIKKILÄ, J. **Blur insensitive texture classification using local phase quantization**. 2008. Disponível em: <http://www.ee.oulu.fi/research/imag/mvg/files/pdf/ICISP08.pdf>. Acesso em: 18 jan. 2017.

RIBAS, L. C. *Análise de texturas dinâmicas baseada em sistemas complexos*. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, doi:10.11606/D.55.2017.tde-28072017-141204, 2017.

RIBAS, L. C. *Aprendizado de representações e caracterização de redes complexas com aplicações em visão computacional*. Tese (Doutorado) — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos,, 2021.

RIBAS, L. C.; BRUNO, O. M. Deterministic partially self-avoiding walks on networks for natural shapes classification. In: VERA-RODRIGUEZ, R.; FIERREZ, J.; MORALES, A. (Ed.). *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Cham: Springer International Publishing, 2019. p. 436–443. ISBN 978-3-030-13469-3.

RIBAS, L. C.; BRUNO, O. M. Dynamic texture classification using deterministic partially self-avoiding walks on networks. In: RICCI, E. et al. (Ed.). *Image Analysis and Processing – ICIAP 2019*. Cham: Springer International Publishing, 2019. p. 82–93. ISBN 978-3-030-30642-7.

RIBAS, L. C.; BRUNO, O. M. Dynamic texture analysis using networks generated by deterministic partially self-avoiding walks. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 541, p. 122105, 2020.

RIBAS, L. C.; BRUNO, O. M. Learning a complex network representation for shape classification. *Pattern Recognition*, v. 154, p. 110566, 2024. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320324003170>.

RIBAS, L. C. et al. Fusion of complex networks and randomized neural networks for texture analysis. *Pattern Recognition*, v. 103, p. 107189, 2020. ISSN 0031-3203.

RIBAS, L. C. et al. Fusion of complex networks and randomized neural networks for texture analysis. *Pattern Recognition*, v. 103, p. 107189, 2020.

RIBAS, L. C. et al. Color-texture classification based on spatio-spectral complex network representations. *Physica A: Statistical Mechanics and its Applications*, p. 129518, 2024. ISSN 0378-4371.

RIBAS, L. C. et al. Local complex features learned by randomized neural networks for texture analysis. *Pattern Analysis and Applications*, v. 27, n. 1, p. 23, 2024. ISSN 1433-755X. Disponível em: <https://doi.org/10.1007/s10044-024-01230-x>.

SCABINI, L. et al. Radam: Texture recognition through randomized aggregated encoding of deep activation maps. *arXiv preprint arXiv:2303.04554*, 2023.

- SCHMIDT, W. F.; KRAAIJVELD, M. A.; DUIN, R. P. W. Feedforward neural networks with random weights. In: *Proceedings., 11th IAPR International Conference on Pattern Recognition. Vol.II. Conference B: Pattern Recognition Methodology and Systems*. [S.l.: s.n.], 1992. p. 1–4.
- SILVA, P. M.; FLORINDO, J. B. Fractal measures of image local features: an application to texture recognition. *Multimedia Tools and Applications*, Springer, v. 80, n. 9, p. 14213–14229, 2021.
- SIMONYAN, K.; ZISSERMAN, A. **Very deep convolutional networks for large-scale image recognition**. 2014. Disponível em: <https://arxiv.org/pdf/1409.1556/>. Acesso em: 10 jan. 2017.
- SZEGEDY, C. et al. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Thirty-First AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2017.
- SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016.
- VICENTIM, A. C. M. Sopa de redes neurais randomizadas com aplicações em diagnóstico de doenças. Universidade Estadual Paulista (Unesp), 2024.
- WEBB, A. R. Linear discriminant analysis. In: _____. *Statistical Pattern Recognition*. [s.n.], 2002. cap. 4, p. 123–168. ISBN 9780470854778. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/0470854774.ch4>.
- WESZKA, J. S.; DYER, C. R.; ROSENFELD, A. A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man and Cybernetics*, v. 6, n. 4, p. 269–285, 1976.
- WESZKA, J. S.; DYER, C. R.; ROSENFELD, A. A comparative study of texture measures for terrain classification. *IEEE transactions on Systems, Man, and Cybernetics*, IEEE, n. 4, p. 269–285, 1976.
- WONG, T.-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, v. 48, n. 9, p. 2839–2846, 2015. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320315000989>.
- ZHU, Z. et al. An adaptive hybrid pattern for noise-robust texture analysis. *Pattern Recognition*, Elsevier, v. 48, n. 8, p. 2592–2608, 2015.
- ZIELINSKI, K. M. C. et al. Complex texture features learned by applying randomized neural network on graphs. In: *2022 Eleventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*. [S.l.: s.n.], 2022. p. 1–6.