



**UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA E CIÊNCIAS ATUARIAIS**



Paulo Henrique Souza Almeida

**DETECÇÃO DE PLACAS SOLARES EM IMAGENS DE SATÉLITE UTILIZANDO
APRENDIZADO NÃO SUPERVISIONADO**

São Cristóvão - SE

2025

Paulo Henrique Souza Almeida

**DETECÇÃO DE PLACAS SOLARES EM IMAGENS DE SATÉLITE UTILIZANDO
APRENDIZADO NÃO SUPERVISIONADO**

Trabalho de Conclusão de Curso apresentado ao Departamento de Estatística e Ciências Atuariais da Universidade Federal de Sergipe, como parte dos requisitos para obtenção do grau de Bacharel em Estatística.

Orientador: Prof. Dr. Cleber Martins Xavier

Coorientador: Prof. Dr. Luiz Henrique Gama Dore de Araujo

São Cristóvão - SE

2025

Paulo Henrique Souza Almeida

**DETECÇÃO DE PLACAS SOLARES EM IMAGENS DE SATÉLITE UTILIZANDO
APRENDIZADO NÃO SUPERVISIONADO**

Trabalho de Conclusão de Curso apresentado ao Departamento de Estatística e Ciências Atuariais da Universidade Federal de Sergipe, como parte dos requisitos para obtenção do grau de Bacharel em Estatística.

Aprovado em DATA DE APROVAÇÃO.

Prof. Dr. Cleber Martins Xavier
Orientador

**Prof. Dr. Luiz Henrique Gama Dore de
Araujo**
Coorientador

B. Sc. João Freire Prado
Convidado 1

Prof. Dr. Oscar Felipe Falcão Raposo
Convidado 2

São Cristóvão - SE
2025

Este trabalho é dedicado à minha mãe, ao meu pai e à minha tia. Obrigado por me amarem.

Agradecimentos

Não conseguiria colocar tudo e todos os agradecimentos em apenas uma folha, ou até no TCC inteiro. Obrigado por todos que teceram uma parte da minha vida até aqui, alguns mais, outros menos. Obrigado. Obrigado por amarem tanto.

*“A ausência da evidência
não significa evidência da ausência.
(The absence of evidence
is not the evidence of absence.)”
(Carl Sagan)*

Resumo

O presente trabalho tem como objetivo desenvolver uma abordagem automatizada para a segmentação de imagens de satélite, focando na detecção de placas fotovoltaicas no município de Aracaju - SE. As imagens de satélite foram disponibilizadas pela SEMFAZ através da Cooperação Técnica entre a UFS e a Secretaria Municipal da Fazenda de Aracaju (SEMFAZ) Nº 01/2022. A metodologia proposta utiliza o algoritmo DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) para realizar a segmentação não supervisionada das imagens, permitindo a identificação de padrões relacionados às estruturas fotovoltaicas em condições de variação de iluminação e ruído. Como ferramenta para o armazenamento eficiente das informações extraídas, foi adotado o banco de dados PostgreSQL, utilizando o tipo de dado JSONB. Além disso, a análise dos parâmetros do DBSCAN, como *eps* e *minPts*, se mostrou fundamental para otimizar a segmentação das características das imagens. Dessa forma, foi desenvolvido um software para que o usuário determine os valores dos parâmetros do DBSCAN, selecione e salve as informações referentes às placas fotovoltaicas identificadas pela metodologia. Os resultados mostram que a abordagem atingiu os objetivos propostos, possibilitando a segmentação eficiente e o armazenamento das informações de forma estruturada. Este estudo contribui para o mapeamento de estruturas de energia solar e abre espaço para a aplicação de métodos supervisionados à medida que uma base de dados robusta seja construída.

Palavras-chave: segmentação de imagens, DBSCAN, aprendizado de máquina, energia solar, reconhecimento de padrões, placa solar

Abstract

This study aims to develop an automated approach for satellite image segmentation, focusing on detecting photovoltaic panels in the municipality of Aracaju - SE. The satellite images were provided by SEMFAZ through the Technical Cooperation Agreement between UFS and the Municipal Finance Department of Aracaju (SEMFAZ) No. 01/2022. The proposed methodology employs the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm to perform unsupervised image segmentation, enabling the identification of patterns related to photovoltaic structures under varying lighting conditions and noise. PostgreSQL was chosen as the database management system, utilizing the JSONB data type for efficiently storing extracted information. Additionally, the analysis of DBSCAN parameters, such as eps and minPts, proved crucial in optimizing the segmentation of image features. To facilitate user interaction, a software tool was developed to allow parameter adjustment, selection, and storage of the identified photovoltaic panel information. The results demonstrate that the proposed approach successfully achieved its objectives, enabling efficient segmentation and structured data storage. This study contributes to the mapping of solar energy structures and paves the way for the application of supervised methods as a robust dataset is built.

Keywords: image segmentation, DBSCAN, machine learning, solar energy, pattern recognition, solar panel

Lista de ilustrações

Figura 1 – Exemplo de imagem de uma Placa Solar modelos híbridos - PV/T.	13
Figura 2 – Exemplo de imagem de uma Placa Solar PV-T e suas dimensões	14
Figura 3 – Gráfico da capacidade de instalações de placas solares no Brasil de 2013 a 2023. 16	
Figura 4 – Distribuição Percentual de Geração de Energia Solar por Categoria de Consumidor no Brasil no período de 2013 à 2023.	17
Figura 5 – Quantidade de instalações de placas solares em Sergipe ao longo dos anos. .	18
Figura 6 – Distribuição percentual de geração de energia solar por categoria de consumidor em Sergipe.	19
Figura 7 – Mapa do município de Aracaju - SE e seus principais bairros.	20
Figura 8 – Mapa da região utilizada para pesquisa com os bairros Atalaia e Farolândia em Aracaju, Sergipe	21
Figura 9 – Exemplo de imagem utilizada para detecção de placas solares em domicílios da cidade de Aracajú - SE	22
Figura 10 – Exemplo de funcionamento do algoritmo DBSCAN	24
Figura 11 – Tela inicial do software desenvolvido para detectar placas solares.	27
Figura 12 – Exemplo da tela de escolha de parâmetros do software	28
Figura 13 – Exemplo da tela do software para navegação, exclusão e salvamento do cluster identificado pelo <i>DBSCAN</i>	29
Figura 14 – Exemplo das variáveis salvas no banco de dados	30
Figura 15 – Exemplo de imagem antes da execução do algoritmo DBSCAN.	31
Figura 16 – Exemplo de imagem sem placas solares.	32
Figura 17 – Exemplo de imagem com várias placas solares	32
Figura 18 – Exemplo de funcionamento do software com cluster detectado pelo algoritmo DBSCAN.	33
Figura 19 – Exemplo de funcionamento do software com cluster salvo em memória . . .	34
Figura 20 – Exemplo de imagem com reconhecimento e salvamento de diversas placas solares no software	35
Figura 21 – Exemplo de imagem das informações salvas no banco	36
Figura 22 – Imagem exemplo do tipo das colunas armazenadas no banco de dados. . . .	37

Sumário

1	INTRODUÇÃO	10
2	OBJETIVOS	12
2.1	Objetivo Geral	12
2.2	Objetivos Específicos	12
3	REVISÃO LITERÁRIA	13
3.1	Energia solar	13
3.2	Energia solar no Brasil	15
4	METODOLOGIA	20
4.1	Área de estudo	20
4.2	Mineração de imagens	22
4.3	Métodos de Clusterização	23
4.3.1	DBSCAN	23
4.3.2	Exemplo Básico	25
4.4	Suporte computacional	26
4.5	Descrição do Programa	26
4.5.1	Funcionamento do Programa	27
4.5.2	Interface e Funcionalidades	27
4.5.2.1	Seleção e Salvamento de Clusters	28
4.5.3	Armazenamento de Informações	30
5	RESULTADOS	31
6	CONCLUSÃO	39
	REFERÊNCIAS	40

1 INTRODUÇÃO

Com o avanço das tecnologias de sensoriamento remoto e aprendizado de máquina, a análise de imagens de satélite tem se tornado uma ferramenta fundamental para diversas aplicações (Jensen, 2007). Um dos principais desafios nesse campo é a segmentação eficiente das imagens para identificação de padrões relevantes. No contexto da energia solar, a detecção automática de placas fotovoltaicas é uma tarefa crucial para o monitoramento da expansão dessa fonte renovável, auxiliando em estudos de viabilidade, planejamento energético e fiscalização regulatória.

A identificação de placas solares em imagens de satélite envolve diversas dificuldades, como a variação nas condições de iluminação, a presença de ruído nas imagens e a similaridade de padrões entre diferentes tipos de telhados e superfícies. Para lidar com esse problema, este trabalho propõe a utilização do algoritmo DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) para segmentação não supervisionada das imagens. O DBSCAN é amplamente utilizado na análise estatística de dados espaciais por sua capacidade de identificar agrupamentos de forma flexível, sem a necessidade de definir previamente o número de clusters (Bishop, 2006).

Diferentemente dos métodos supervisionados, que exigem grandes volumes de dados rotulados, a abordagem proposta permite a detecção automática de padrões sem a necessidade de treinamento prévio (Bishop, 2006). Além disso, a interação manual com os clusters gerados permite refinar a segmentação e garantir que apenas as regiões relevantes sejam consideradas (Murphy, 2012). O software desenvolvido permite visualizar os agrupamentos, selecionar e armazenar dinamicamente as informações extraídas em um banco de dados, garantindo a rastreabilidade e a reprodutibilidade dos resultados.

O armazenamento eficiente das informações extraídas é outro aspecto fundamental deste estudo. Optou-se pelo uso do banco de dados PostgreSQL, utilizando o tipo de dado JSONB para armazenar as coordenadas e os valores de cores dos clusters identificados. Essa escolha se deve à eficiência na indexação e recuperação das informações, facilitando futuras consultas e análises estatísticas (Stonebraker; Rowe, 1986).

A motivação para este estudo surge da necessidade crescente de métodos automáticos para mapeamento da energia solar, bem como do potencial de aplicação dessa abordagem para outras áreas que envolvem reconhecimento de padrões em imagens (Li; Wang, 2017). Além disso, este trabalho contribui para a interseção entre Estatística, Ciência de Dados e Energias Renováveis, utilizando técnicas de aprendizado de máquina para aprimorar a identificação de estruturas fotovoltaicas.

O objeto de análise deste estudo são imagens de satélite do município de Aracaju - se que

são processadas pelo software desenvolvido para segmentação e categorização de placas solares. A metodologia adotada combina técnicas estatísticas, aprendizado de máquina e manipulação de imagens para a extração e análise de clusters. Com isso, espera-se contribuir para a detecção automática de placas solares e fomentar o desenvolvimento de modelos supervisionados mais precisos para reconhecimento de padrões em imagens de satélite.

O restante deste trabalho está estruturado da seguinte forma: o Capítulo 2 apresenta os objetivos da pesquisa; o Capítulo 3 traz uma revisão bibliográfica sobre os temas abordados; a metodologia utilizada é detalhada no Capítulo 4; os resultados obtidos e suas interpretações são apresentados no Capítulo 5; e, por fim, o Capítulo 6 traz as considerações finais e sugestões para trabalhos futuros.

2 OBJETIVOS

2.1 Objetivo Geral

Desenvolver um software interativo para a identificação de placas solares em imagens de satélite, utilizando técnicas de aprendizado de máquina não supervisionado, com ênfase no algoritmo DBSCAN.

2.2 Objetivos Específicos

- Identificar características visuais e espectrais que diferenciam placas solares de outros objetos presentes em imagens de satélite, estabelecendo critérios para segmentação eficiente;
- Aplicar o algoritmo DBSCAN para segmentação das imagens, analisando como diferentes configurações de parâmetros afetam a detecção e a formação dos clusters que representam possíveis placas solares;
- Implementar um software interativo utilizando o framework *Shiny* (R), permitindo que o usuário carregue imagens, ajuste dinamicamente os parâmetros do algoritmo DBSCAN (ϵ e *MinPts*) e visualize os agrupamentos gerados diretamente na interface gráfica;
- Armazenar os resultados processados em um banco de dados PostgreSQL, utilizando o formato *JSONB* para registrar informações como coordenadas geográficas, características espectrais (píxeis RGB), identificador do cluster e a decisão do usuário sobre a relevância do grupo detectado;
- Permitir a manipulação dinâmica dos clusters pelo usuário, possibilitando a seleção, remoção e salvamento de agrupamentos diretamente na interface do software antes da exportação final dos dados para um banco de dados;

3 REVISÃO LITERÁRIA

3.1 Energia solar

As células solares fotovoltaicas utilizam materiais semicondutores, como silício cristalino (c-Si) e película fina (TF), para converter diretamente a luz solar em eletricidade. Esses materiais são os mais utilizados no mercado, devido à sua eficiência comprovada (Zwirtes; Wanderley, 2022).

Esta tecnologia apresenta vantagens como a modularidade, com potência variando de miliwatts a megawatts, alta confiabilidade, baixa manutenção e simplicidade no uso (Zomer, 2014). Além disso, os sistemas fotovoltaicos são capazes de gerar eletricidade de forma imediata após a instalação, podem ser realocados conforme necessário, operam de maneira silenciosa e sem peças móveis. Outra vantagem significativa é a possibilidade de instalação próxima ou no próprio ponto de consumo, facilitando o uso direto da energia gerada (Urbanetz, 2011).

As placas fotovoltaicas possuem uma moldura metálica (geralmente de alumínio), uma camada de vidro temperado na frente para proteção e uma camada isolante na parte traseira. Conectores elétricos na parte posterior permitem a transmissão da energia gerada. Alguns modelos híbridos - PV/T (Figura 1) também possuem um sistema de resfriamento para aproveitar o calor gerado (Energia, 2023).

Figura 1 – Exemplo de imagem de uma Placa Solar modelos híbridos - PV/T.



Fonte: (Solimpeks, 2025).

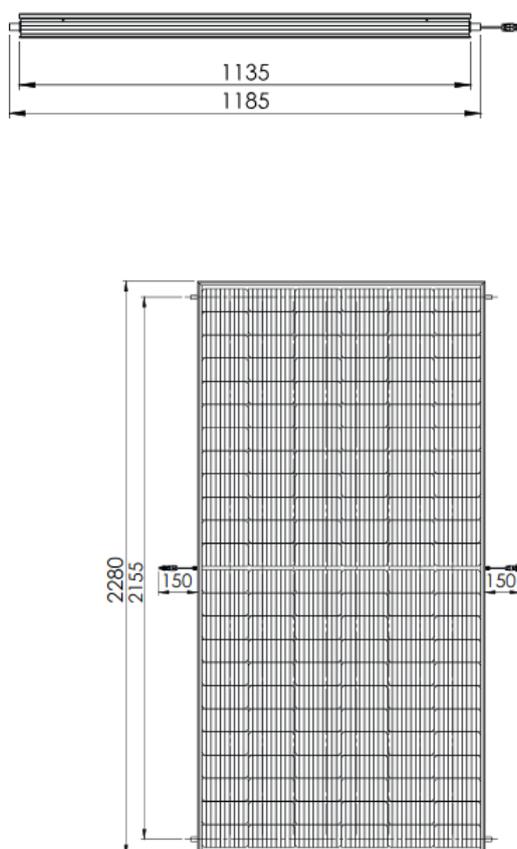
Na Figura 1 pode-se observar um exemplo de placa fotovoltaica Híbrido PV-T produzida

por [Solimpeks \(2025\)](#). O PV-T Hybrid Panel é uma solução inovadora que combina a geração de eletricidade fotovoltaica com a coleta térmica solar, otimizando o uso da energia solar. A tecnologia de células PERC resfriadas aumenta a produção de eletricidade, enquanto o sistema térmico fornece água aquecida, tornando o painel eficiente para diversas aplicações, como residenciais, comerciais e industriais.

Com sua estrutura durável e de fácil instalação, o painel oferece uma alternativa sustentável para reduzir custos de energia. A integração de duas funções em um único painel é relevante para estudos sobre sistemas híbridos de energia solar, destacando-se pela maximização da utilização da radiação solar e pela diminuição do desperdício energético ([Solimpeks, 2025](#)).

Na Figura 2, é possível observar informações complementares sobre as dimensões e estrutura do painel fotovoltaico híbrido PV-T. O painel possui uma altura total de 2280 mm, dos quais 2155 mm correspondem à área efetiva das placas e metais, enquanto o restante é composto pela borda estrutural de ferro. A largura total do painel é de 1185 mm, sendo 1135 mm destinados às placas fotovoltaicas, com uma pequena margem adicional na estrutura metálica. Esses detalhes são essenciais para avaliar a área útil do painel na conversão de energia e sua compatibilidade com diferentes aplicações arquitetônicas.

Figura 2 – Exemplo de imagem de uma Placa Solar PV-T e suas dimensões



Fonte: ([Solimpeks, 2025](#)).

3.2 Energia solar no Brasil

A história da energia solar no Brasil remonta às décadas de 1970 e 1980, quando os primeiros estudos e projetos-piloto começaram a ser desenvolvidos, impulsionados pela crise do petróleo. Desde então, houve avanços significativos na pesquisa e implementação de sistemas fotovoltaicos. Em particular, a utilização de painéis solares flutuantes em reservatórios de hidroelétricas surgindo como uma alternativa viável para a geração de energia renovável (Strangueto, 2016).

O consumo de energia no Brasil tem aumentado significativamente ao longo das últimas décadas, impulsionado pelo crescimento populacional e pelo desenvolvimento econômico. A dependência de fontes convencionais, como as hidroelétricas, tem levado a desafios, incluindo crises energéticas em períodos de estiagem. Nesse contexto, a energia solar tem se destacado como uma alternativa sustentável e economicamente viável.

A utilização da energia solar fotovoltaica se torna cada vez mais interessante, sobretudo devido aos seguintes fatores: grandes evoluções na eficiência; redução dos custos de fabricação; ganhos de escala; incentivos e regulamentações, como a publicação da RN 482/12 (Nascimento, 2015), que trata da mini e microgeração no Brasil e a elaboração do Programa de Desenvolvimento da Geração Distribuída de Energia Elétrica (ProGD); crescimento da demanda de energia elétrica motivada pela melhoria das condições de vida da população; redução da oferta de outras fontes; e busca por uma maior segurança no suprimento energético em todo o mundo (Strangueto, 2016).

A irradiação solar no Brasil é favorável à geração de eletricidade. O potencial mínimo necessário para a instalação de sistemas fotovoltaicos planos sem concentração é de 2.000 kWh/m² por ano, ou seja, 5,5 kWh/m² por dia. No Brasil, obtém-se o valor médio ponderado de 5,6 kWh/m² por dia, viabilizando, portanto, o emprego dessa tecnologia (Urbanetz, 2011).

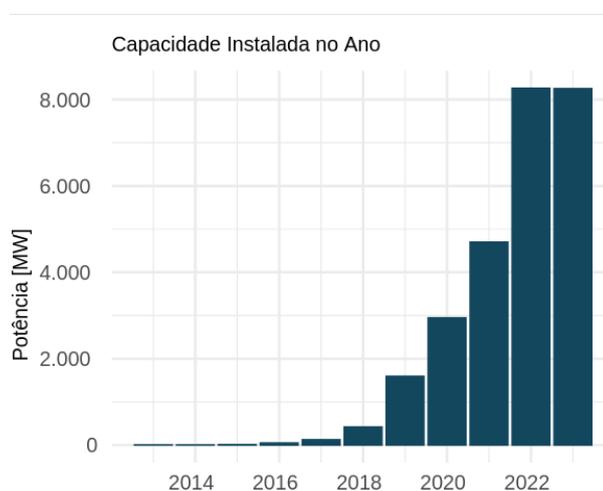
A energia solar se destaca como uma solução sustentável e econômica para diversas aplicações, principalmente no aquecimento de água. No Brasil, o uso da energia solar para aquecimento de água a temperaturas abaixo de 100°C é a aplicação mais difundida, principalmente como substituição a sistemas de aquecimento elétrico (como chuveiros) ou a gás. Isso ocorre porque a tecnologia de conversão de energia solar em energia térmica é relativamente simples e amplamente disponível no mercado brasileiro. Com diversos fornecedores e fabricantes, a energia solar oferece uma viabilidade econômica significativa em projetos bem planejados (Pereira et al., 2017).

Além disso, a energia solar apresenta grande potencial de aquecimento térmico, sendo uma alternativa eficiente e limpa. Incentivos governamentais, como isenção de impostos e a obrigatoriedade de uso em determinadas situações, impulsionam a adoção em larga escala desses sistemas, principalmente em residências. Programas como os da ANEEL, que oferecem equipamentos gratuitos através de programas de eficiência energética, e iniciativas voltadas para moradias de interesse social (como o Minha Casa Minha Vida) também são fatores importantes

para promover a utilização de sistemas solares de aquecimento em diversas regiões do país (Pereira et al., 2017).

A partir da Figura 3, disponível no Painel de Dados de Micro e Minigeração Distribuída (EPE, 2025), é possível observar um significativo aumento na quantidade de placas solares instaladas a partir de 2019, o que indica um verdadeiro *boom* no setor de energia solar. Esse crescimento é reflexo de uma série de fatores, como os incentivos fiscais, o maior acesso a informações sobre os benefícios da energia solar e a necessidade crescente de soluções energéticas mais sustentáveis e econômicas. Além disso, é interessante notar que o número de instalações de painéis solares continua em tendência crescente, o que demonstra uma evolução contínua da microgeração e minigeração distribuída no Brasil.

Figura 3 – Gráfico da capacidade de instalações de placas solares no Brasil de 2013 a 2023.



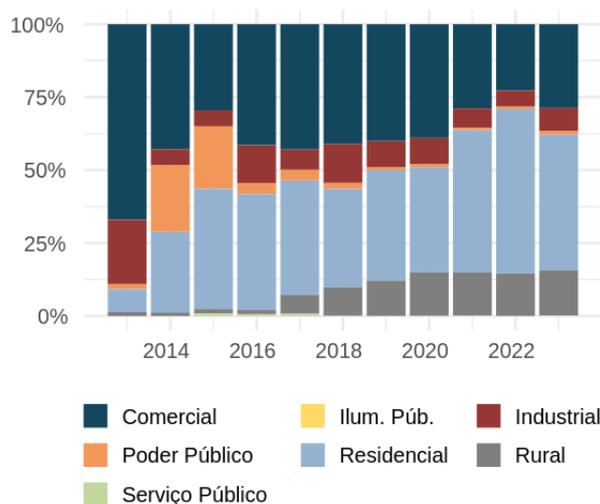
Fonte: (EPE, 2025).

Esse crescimento sugere que o mercado de energia solar está se tornando cada vez mais atrativo, não só para o setor residencial, mas também para empresas e até grandes indústrias que buscam alternativas para reduzir seus custos com eletricidade. A continuidade desse aumento reflete uma mudança na matriz energética nacional, destacando a energia solar como uma das principais fontes de energia renovável.

Em todo o Brasil a crescente demanda por energia e os altos níveis de radiação solar têm impulsionado a implantação de sistemas solares fotovoltaicos. Projetos de usinas solares em cidades demonstram o potencial da energia solar para diversificar a matriz energética do estado e reduzir sua dependência de fontes convencionais (Pereira et al., 2017).

Como pode-se observar na Figura 4, a energia solar para o campo rural também uma tendência de crescimento. A crescente adoção da energia solar no campo rural tem se mostrado uma solução estratégica, especialmente como alternativa ao uso de combustíveis fósseis. A energia fotovoltaica oferece tecnologia avançada para alimentar máquinas agrícolas, sistemas de irrigação e outras atividades essenciais.

Figura 4 – Distribuição Percentual de Geração de Energia Solar por Categoria de Consumidor no Brasil no período de 2013 à 2023.



Fonte: (EPE, 2025).

Além disso, a energia solar elimina os riscos de falta de energia em áreas rurais, garantindo um fornecimento confiável e contínuo de eletricidade, especialmente em locais com infraestrutura elétrica precária, assegurando uma produção estável e eficiente.

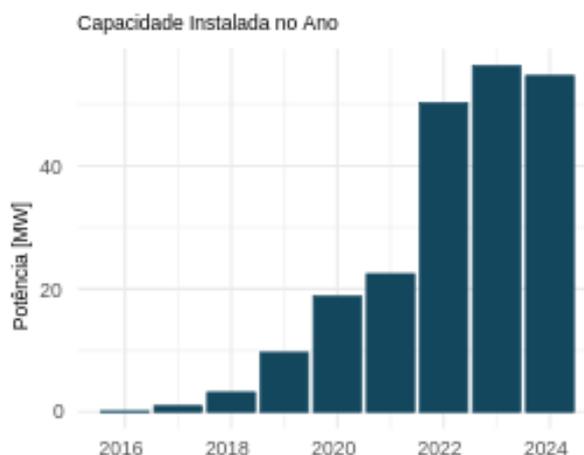
Outro ponto crucial é a sustentabilidade. A energia fotovoltaica, sendo uma fonte limpa e renovável, contribui para a redução das emissões de gases de efeito estufa e a mitigação das mudanças climáticas. Certificações como o selo solar não só evidenciam o compromisso com a energia limpa, mas também geram benefícios financeiros, como acesso a incentivos governamentais e preferência em mercados sustentáveis. Portanto, a energia solar no campo não só promove eficiência energética, mas também torna a produção rural mais competitiva e sustentável (Francisco et al., 2024).

A adoção da energia solar em Sergipe tem se destacado no cenário nacional, especialmente no setor de Serviço Público. O estado tem investido significativamente em projetos solares, como a usina fotovoltaica inaugurada no SergipeTec, que representa um avanço tanto na produção de energia renovável quanto no incentivo à inovação tecnológica (Portal Solar, 2025). Além disso, o governo estadual está construindo um parque solar em Canindé de São Francisco, que, quando concluído, será um dos maiores do mundo, com capacidade de 1,2 mil megawatts. Esses investimentos reforçam a posição de Sergipe como um dos estados mais comprometidos com a transição energética e a sustentabilidade ambiental (Portal Solar, 2025).

Outro ponto relevante é o investimento na formação de mão de obra especializada. O SergipeTec oferece cursos gratuitos para a capacitação de instaladores solares fotovoltaicos, permitindo que mais profissionais estejam aptos a atuar no setor e ampliando a adoção da tecnologia (Governo do Estado de Sergipe, 2025).

A expansão da energia solar no estado de Sergipe pode ser observado na Figura 5. Esse crescimento não apenas reduz os impactos ambientais, mas também gera economia no consumo de eletricidade, tornando-se uma solução viável tanto para o setor público quanto para a iniciativa privada.

Figura 5 – Quantidade de instalações de placas solares em Sergipe ao longo dos anos.

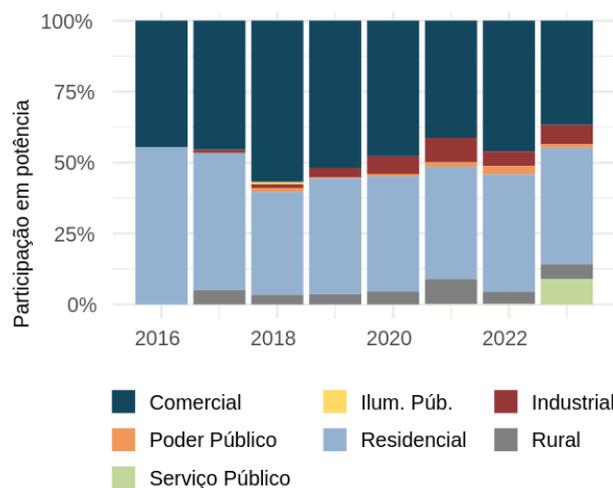


Fonte: (EPE, 2025).

Na Figura 6, são apresentadas as categorias utilizadas para classificar o uso de energia solar em Sergipe. Essas categorias representam diferentes setores que adotam a geração fotovoltaica, variando conforme a finalidade do consumo e o perfil dos usuários:

- **Residencial:** representa a maior parte das instalações fotovoltaicas em 2023, sendo composta por consumidores individuais que utilizam energia solar para reduzir custos e tornar o consumo mais sustentável;
- **Comercial:** inclui estabelecimentos como lojas, escritórios e supermercados que utilizam energia solar para reduzir despesas operacionais e garantir maior autonomia energética;
- **Industrial:** abrange fábricas e instalações produtivas que necessitam de grande demanda energética, aproveitando a geração solar para otimizar custos e melhorar a sustentabilidade;
- **Poder Público:** refere-se a prédios governamentais e órgãos públicos que utilizam energia solar para diminuir gastos com eletricidade e promover iniciativas sustentáveis;
- **Iluminação Pública:** categoria voltada para postes e sistemas de iluminação urbana que utilizam painéis solares para reduzir a dependência da rede elétrica convencional;
- **Rural:** envolve propriedades agrícolas e fazendas que usam energia solar para alimentar sistemas de irrigação, bombeamento de água e outras necessidades do setor agropecuário;
- **Serviço Público:** inclui hospitais, escolas e outras infraestruturas essenciais que adotam energia solar para garantir fornecimento contínuo e reduzir despesas operacionais.

Figura 6 – Distribuição percentual de geração de energia solar por categoria de consumidor em Sergipe.



Fonte: (EPE, 2025).

Outra característica importante apresentada na Figura 6 é a diversificação do tipo de consumidor de energia solar presente no estado de Sergipe ao longo dos anos. Em 2016, a participação na geração de energia solar estava concentrada nas categorias Residencial e Comercial. Entretanto, em 2023 essa geração foi reorganizada nas categorias Residencial, Comercial, Serviço Público, Industrial e Poder público. Essas mudanças demonstram a diversidade de aplicações da energia solar, destacando sua importância na transição para um modelo energético mais sustentável e economicamente viável.

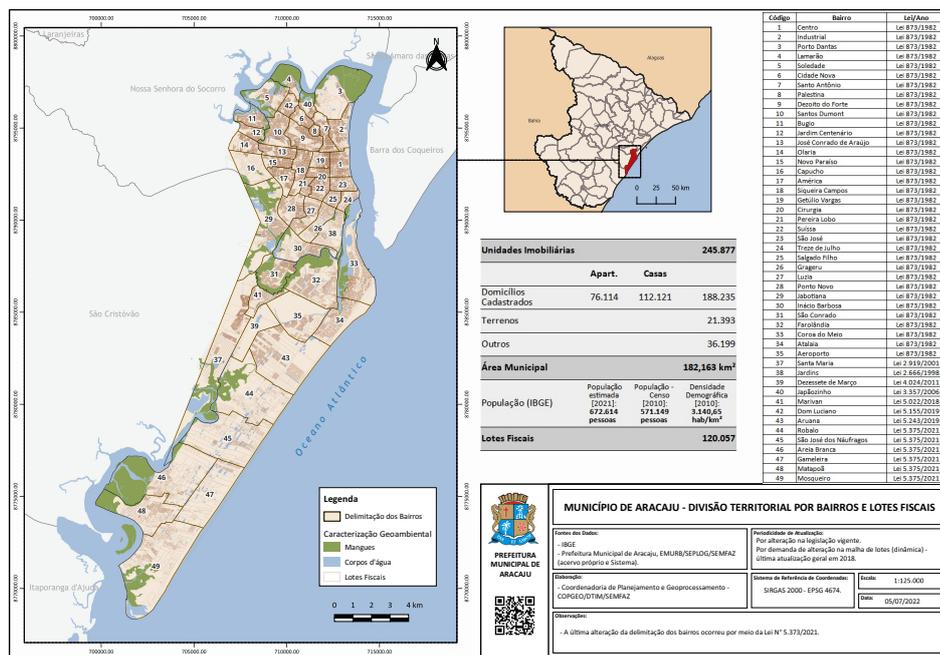
4 METODOLOGIA

Este estudo pode ser classificado como uma pesquisa documental e experimental. A pesquisa documental foca em imagens geoespaciais no formato .tif, utilizadas para a identificação de placas solares, enquanto a abordagem experimental aplica o algoritmo DBSCAN para análise dos dados coletados.

4.1 Área de estudo

O estudo foi conduzido utilizando imagens das regiões de Aracaju, no estado de Sergipe, Brasil. A escolha dessa localidade se deve à sua posição geográfica privilegiada, caracterizada por um clima tropical e elevada incidência solar ao longo do ano, fatores que favorecem a geração de energia fotovoltaica. A Figura 7 apresenta a região de Aracaju e seus principais bairros.

Figura 7 – Mapa do município de Aracaju - SE e seus principais bairros.



Fonte: Cooperação Técnica Nº 01/2022

Para a visualização e utilização das imagens no treinamento, foram selecionadas as regiões costeiras de Atalaia e Farolândia. Essa escolha foi baseada em uma análise preliminar, que indicou um número maior de instalações de placas solares em residências nessas áreas.

As análises foram concentradas, em sua maioria, em áreas costeiras, especialmente em bairros e comunidades próximas ao mar. Essas regiões apresentaram uma maior densidade de residências equipadas com painéis solares, refletindo a crescente adesão da população a fontes de

energia renovável. O litoral de Aracaju, além de abrigar uma considerável infraestrutura urbana, possui características climáticas que possibilitam um maior aproveitamento da energia solar, tornando-se um cenário ideal para a pesquisa.

A Figura 8 retrata a região de Atalaia e Farolândia, localizadas no litoral de Sergipe, que são o foco deste estudo. A escolha dessas regiões se justifica por sua localização costeira, o que garante maior incidência solar ao longo do ano, favorecendo a geração de energia solar. Este fator, somado à crescente adoção de tecnologias renováveis, tornam essas localidades algumas das mais relevantes para a utilização de energia fotovoltaica no estado.

Figura 8 – Mapa da região utilizada para pesquisa com os bairros Atalaia e Farolândia em Aracaju, Sergipe



Fonte: (Aracaju, 2025)

As imagens utilizadas no estudo foram disponibilizadas por órgãos governamentais fruto da Cooperação Técnica entre a UFS e a Secretaria Municipal da Fazenda de Aracaju (SEMFAZ) Nº 01/2022, permitindo uma visualização aérea detalhada das instalações fotovoltaicas (ver Figura 9). A combinação dessas fontes de dados proporcionou uma análise mais abrangente, permitindo identificar padrões de distribuição dos sistemas solares e avaliar sua presença em diferentes tipos de edificações.

A seleção das imagens foi realizada com base na disponibilidade de dados e na representatividade da região em relação ao panorama da energia solar em Sergipe. A investigação buscou compreender a distribuição dos sistemas fotovoltaicos para auxiliar em estudos que possam identificar fatores socioeconômicos e ambientais que influenciam sua adoção. Dessa forma, este

Figura 9 – Exemplo de imagem utilizada para detecção de placas solares em domicílios da cidade de Aracajú - SE



Fonte: Cooperação Técnica N° 01/2022

estudo contribui para uma visão mais ampla sobre a disseminação da energia solar em áreas urbanas e litorâneas do estado de Sergipe.

4.2 Mineração de imagens

A segmentação de imagens é uma técnica fundamental na análise de dados geoespaciais, permitindo a identificação e extração de padrões específicos dentro de imagens de satélite. No contexto deste estudo, utilizou-se aprendizado não supervisionado para segmentar imagens de áreas urbanas e costeiras, possibilitando a identificação de estruturas com painéis solares.

Diferentemente do aprendizado supervisionado, onde há necessidade de um conjunto de dados rotulado para treinamento, o aprendizado não supervisionado permite a categorização automática dos pixels da imagem com base em suas características estatísticas. Técnicas como agrupamento (*clustering*) são amplamente utilizadas para este propósito, destacando-se o uso de algoritmos como o *k-means* e *DBSCAN*. Essas abordagens permitem separar regiões homogêneas da imagem, facilitando a análise e extração de informações relevantes (Bishop, 2006).

4.3 Métodos de Clusterização

4.3.1 DBSCAN

O DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) é um algoritmo baseado em densidade que permite identificar clusters de formas arbitrárias e detectar ruídos em conjuntos de dados. Ele foi originalmente apresentado por Ester et al. (1996).

O algoritmo utiliza dois parâmetros principais:

- ϵ (*epsilon*): distância máxima entre dois pontos para que sejam considerados vizinhos;
- *MinPts*: número mínimo de pontos necessários para formar uma região densa.

Dessa forma, o método DBSCAN classifica os pontos como:

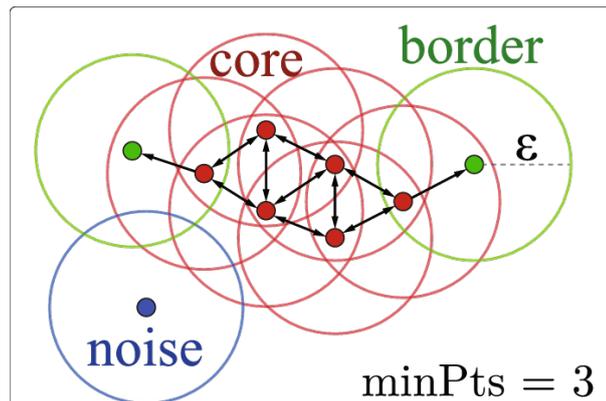
- **Pontos núcleo:** possuem pelo menos *MinPts* vizinhos dentro da distância ϵ ;
- **Pontos de borda:** estão próximos de pontos núcleo, mas não possuem *MinPts* vizinhos suficientes;
- **Pontos de ruído:** não pertencem a nenhum cluster.

A Figura 10 ilustra o funcionamento do algoritmo DBSCAN (Density-Based Spatial Clustering of Applications with Noise), que realiza o agrupamento de dados com base na densidade espacial dos pontos. Nela, podem ser observadas as três categorias principais de pontos no processo de clustering:

- **Pontos Centrais (Core Points):** São destacados por setas conectando os pontos e formam o núcleo de cada grupo. Esses pontos são definidos por sua densidade, ou seja, para um ponto ser considerado um *core point*, deve possuir pelo menos *minPts* pontos vizinhos dentro de um raio de distância ϵ . Nesse exemplo, com *minPts* = 3, a seta conecta pontos que estão próximos o suficiente para formar um agrupamento, respeitando a densidade local.
- **Pontos de Bordas (Border Points):** Esses pontos estão situados nas periferias do agrupamento, ou seja, não possuem a densidade necessária para serem considerados *core points*, mas estão dentro do raio ϵ de um *core point*. Em verde destaca-se a linha que separa os pontos *core* dos *border points*, indicando a relevância da proximidade dos pontos para fazer parte do mesmo grupo. Esses pontos fazem parte do agrupamento, mas com uma posição periférica.

- **Pontos de Ruído (Noise Points):** Representados por um ponto separado dos grupos, esses pontos não pertencem a nenhum agrupamento, pois estão fora do raio ϵ dos *core points*. São considerados como *outliers* ou *ruído* e não são incluídos nos clusters formados pelo algoritmo. Esses pontos não têm vizinhos suficientes para serem agrupados, conforme a definição do algoritmo DBSCAN.

Figura 10 – Exemplo de funcionamento do algoritmo DBSCAN



Fonte: (Mehle; Likar; Tomažević, 2017)

A Figura 10 exemplifica como a distância entre os pontos é essencial para determinar a formação de clusters, com a relevância dos parâmetros ϵ e $minPts$ no processo de definição dos grupos. O algoritmo DBSCAN é eficaz para identificar regiões de alta densidade de pontos, enquanto descarta os pontos de baixa densidade como ruído, proporcionando uma análise robusta para dados espaciais ou de características semelhantes.

O funcionamento geral do algoritmo envolve a identificação de pontos núcleo, a expansão de clusters a partir desses pontos e a categorização de outros pontos como borda ou ruído. Este processo é eficiente e amplamente utilizado em aplicações práticas, como reconhecimento de padrões e análise de dados (Bonaccorso, 2018).

O algoritmo DBSCAN (Density-Based Spatial Clustering of Applications with Noise) baseia-se em uma definição de vizinhança para agrupar pontos que são densamente conectados. Para definir se um ponto p é considerado "vizinho" de outro ponto q , utiliza-se a métrica de distância euclidiana, que é a fórmula padrão para calcular a distância entre dois pontos no espaço. A distância entre dois pontos $p = (p_1, p_2, \dots, p_n)$ e $q = (q_1, q_2, \dots, q_n)$ no espaço n -dimensional é dada por:

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

No DBSCAN, dois pontos p e q são considerados diretamente alcançáveis um do outro se satisfizerem as seguintes condições:

$$d(p, q) \leq \varepsilon \quad \text{e} \quad |N_\varepsilon(p)| \geq \text{MinPts}$$

ou seja, $d(p, q)$ é a distância entre os pontos p e q , e ε é o parâmetro que define o raio máximo para os pontos serem considerados vizinhos. $N_\varepsilon(p)$ é o conjunto de vizinhos de p dentro desse raio ε . A condição $|N_\varepsilon(p)| \geq \text{MinPts}$ especifica que, para um ponto ser considerado um ponto *core*, ele deve ter pelo menos MinPts pontos dentro do seu raio de vizinhança ε .

A importância de MinPts no algoritmo é fundamental para a definição dos pontos *core*. Se um ponto p tiver pelo menos MinPts pontos vizinhos dentro de um raio ε , ele será considerado um ponto *core*. Caso contrário, ele será tratado como um ponto *border* ou, se estiver isolado, como *noise*.

4.3.2 Exemplo Básico

Seja um conjunto de pontos bidimensionais, considere três pontos $p = (2, 3)$, $q = (3, 4)$, e $r = (5, 6)$. Definimos um valor de $\varepsilon = 2$ (distância máxima para dois pontos serem considerados vizinhos) e $\text{MinPts} = 2$.

A distância entre p e q é:

$$d(p, q) = \sqrt{(3 - 2)^2 + (4 - 3)^2} = \sqrt{1 + 1} = \sqrt{2} \approx 1.41$$

Como $d(p, q) = 1.41 \leq \varepsilon = 2$, os pontos p e q são vizinhos.

A distância entre p e r é:

$$d(p, r) = \sqrt{(5 - 2)^2 + (6 - 3)^2} = \sqrt{9 + 9} = \sqrt{18} \approx 4.24$$

Como $d(p, r) = 4.24 > \varepsilon = 2$, os pontos p e r não são vizinhos.

Agora, para p ser considerado um ponto *core*, ele precisa ter pelo menos $\text{MinPts} = 2$ pontos vizinhos dentro de ε . Neste caso, p tem q como vizinho, mas não tem r . Portanto, p não é um ponto *core* porque não tem pontos suficientes dentro do raio ε .

Esse exemplo ilustra como ε e MinPts são essenciais para determinar a estrutura do agrupamento. O valor de ε define o alcance de cada vizinho, enquanto MinPts determina quantos vizinhos são necessários para que um ponto seja considerado um *core point* e, assim, iniciar a formação de um grupo.

Esse comportamento pode ser ajustado conforme a natureza dos dados, sendo que valores diferentes de ε e MinPts podem resultar em clusters mais densos ou mais dispersos, dependendo da configuração do espaço dos dados (Bishop, 2006).

4.4 Suporte computacional

No presente estudo, a segmentação das imagens foi realizada utilizando a linguagem R e as bibliotecas `raster` e `stringr`, juntamente com ferramentas da GDAL para processamento e otimização dos arquivos geoespaciais. O fluxo de trabalho incluiu:

1. Utilização do `gdal2tiles.py` para dividir imagens georreferenciadas em blocos menores, facilitando o processamento e análise.
2. Leitura das imagens no formato `.tif` usando a biblioteca `raster`, permitindo a manipulação das camadas espectrais.
3. Normalização das bandas espectrais para um intervalo de 0 a 255 utilizando o comando `gdal_translate`, garantindo a padronização dos valores e melhorando a qualidade da segmentação.
4. Aplicação de técnicas de reamostragem e interpolação, utilizando métodos bilineares e cúbicos para ajustar a resolução das imagens processadas.
5. Uso do `gdalwarp` para ajuste de escala espacial, remoção de dados nulos e melhoria da qualidade final da imagem segmentada.

A combinação dessas ferramentas e técnicas permitiu a extração eficiente de informações sobre a distribuição de painéis solares em Aracaju e regiões costeiras próximas. A análise segmentada das imagens contribuiu para um melhor entendimento da penetração da energia solar na matriz energética local, fornecendo subsídios para estudos futuros e políticas públicas voltadas à expansão do uso dessa tecnologia.

Foram utilizadas imagens geoespaciais provenientes de bancos de dados especializados em sensoriamento remoto. As ferramentas computacionais selecionadas incluem o software R e os pacotes `raster`, para processamento de imagens `.tif`, e `dbscan`, para clusterização dos dados. Além disso, o pacote `ggplot2` foi usado para gerar visualizações dos resultados.

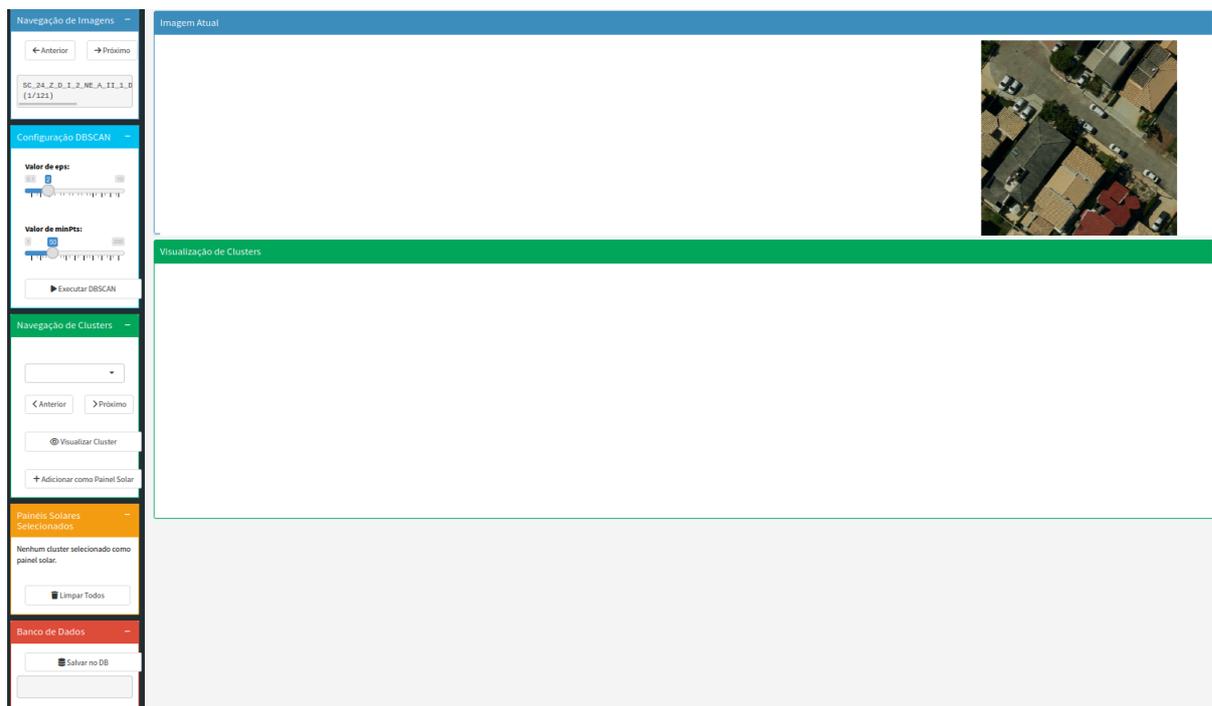
4.5 Descrição do Programa

O software desenvolvido tem como objetivo a identificação e categorização de áreas de interesse em imagens `.tif`, utilizando o algoritmo DBSCAN para agrupamento. Ele foi implementado com a linguagem de programação R e a biblioteca Shiny, criando uma interface web interativa para facilitar o processo de análise.

4.5.1 Funcionamento do Programa

O programa inicia recebendo um diretório de imagens *.tif* a partir de um caminho previamente selecionado no código. Após o carregamento das imagens, o algoritmo DBSCAN é aplicado, permitindo que o usuário ajuste os parâmetros de clustering diretamente pela interface web. Os parâmetros ajustáveis incluem o valor de *eps* (distância máxima entre dois pontos para que sejam considerados vizinhos) e *MinPts* (mínimo de pontos necessários para formar um cluster) como pode ser observado na Figura 11.

Figura 11 – Tela inicial do software desenvolvido para detectar placas solares.

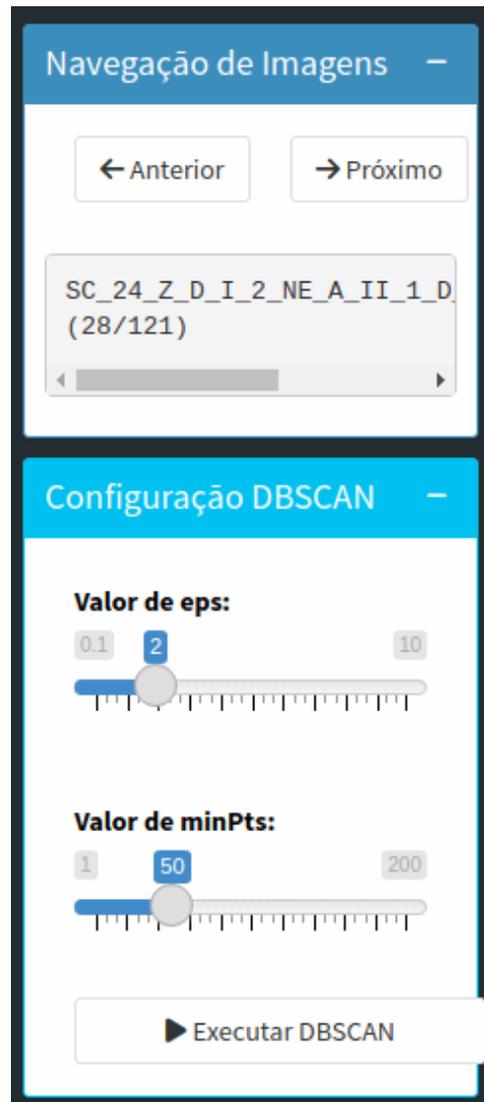


Fonte: Elaborada pelo autor.

4.5.2 Interface e Funcionalidades

A interface permite que o usuário visualize os clusters gerados pelo DBSCAN em uma tela de visualização interativa. Com o uso de um scroll, é possível navegar por todas as imagens de clusters, facilitando a busca por áreas de interesse, além de poder escolher os parâmetros do DBSCAN, os *MinPts* e *Eps*. No caso do presente estudo, a busca é direcionada para a detecção de placas solares, mas o software pode ser utilizado para qualquer tipo de agrupamento de objetos com características visuais semelhantes. A Figura 12 apresenta as opções de escolha dos parâmetros para, em seguida, executar a busca de placas solares pelo algoritmo *DBSCAN*.

Figura 12 – Exemplo da tela de escolha de parâmetros do software



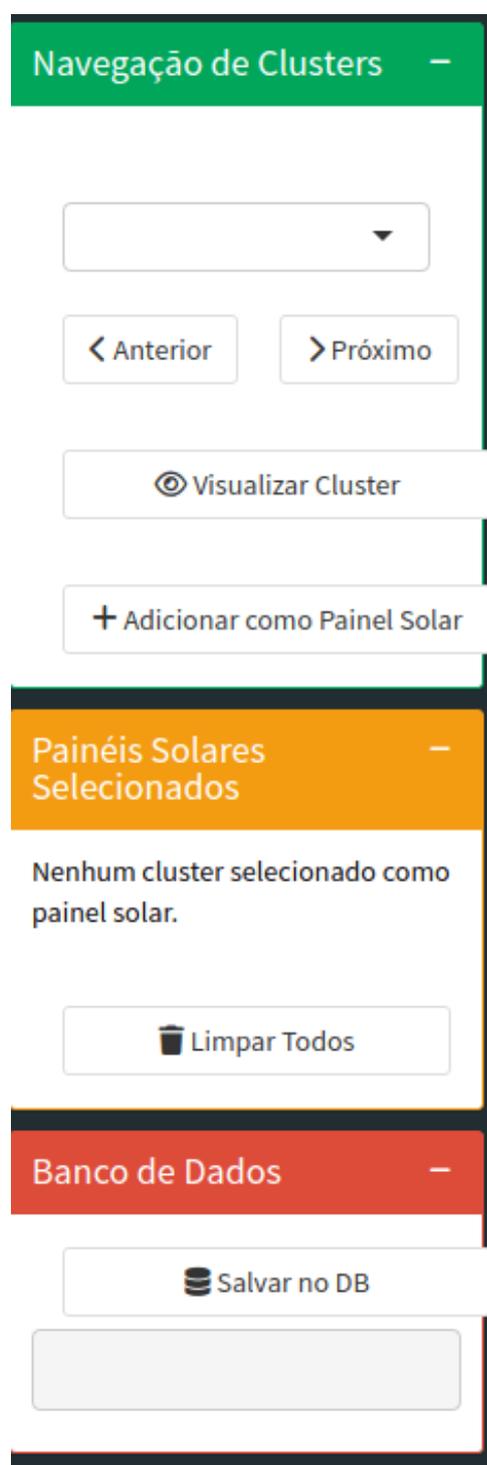
Fonte: Elaborada pelo autor.

4.5.2.1 Seleção e Salvamento de Clusters

Após a visualização, o usuário pode selecionar os clusters que considera mais relevantes. Para facilitar a organização, os clusters selecionados podem ser salvos em um banco de dados, permitindo a categorização dos dados coletados. Além disso, o usuário pode excluir clusters previamente salvos diretamente na interface como pode ser observado na Figura 13.

Cada cluster é identificado pelo seu `cluster_id`, e o banco de dados registra diversas informações associadas a ele, como as coordenadas dos pontos, o nome da imagem (`image_name`), os valores RGB dos pixels, e os parâmetros `MinPts` e `Eps` utilizados na execução do DBSCAN. Uma coluna chamada "`is_solar`" é incluída, funcionando como um marcador booleano: se o cluster foi selecionado como relevante (indicando que o objetivo foi encontrado), o valor será 1; caso contrário, o valor será 0.

Figura 13 – Exemplo da tela do software para navegação, exclusão e salvamento do cluster identificado pelo *DBSCAN*.



Fonte: Elaborada pelo autor.

4.5.3 Armazenamento de Informações

As informações sobre os clusters selecionados são armazenadas em um banco de dados, cujas colunas incluem:

- `cluster_id` – Identificador único do cluster.
- `coordinates` – Coordenadas dos pontos que compõem o cluster.
- `image_name` – Nome da imagem associada ao cluster.
- `rgb_pixels` – Valores RGB dos pixels que pertencem ao cluster.
- `MinPts` – Número mínimo de pontos necessário para formar o cluster.
- `Eps` – Distância máxima para que pontos sejam considerados vizinhos.
- `is_solar` – Indicador se o cluster foi selecionado como relevante para o objetivo da análise (1 se sim, 0 se não).

A Figura 14 exemplifica e demonstra alguns valores já salvos no banco de dados relacional. Essas informações podem ser utilizadas para análises futuras, criando um histórico dos clusters identificados e permitindo a rastreabilidade dos dados.

Figura 14 – Exemplo das variáveis salvas no banco de dados

Az image_name	123 id_cluster	123 is_solar	123 eps	123 minpts	{ } coords	🕒 insertion_date	{ } rgb_values
SC_24_Z_D_I_2_NE_A_II_1_D_	1	0	2	50	[[708845.2625, 8782444.887	2025-03-20 00:15:49.828	[[161, 163, 134], [160, 162, 1
SC_24_Z_D_I_2_NE_A_II_1_D_	2	0	2	50	[[708862.0625, 8782444.887	2025-03-20 00:15:50.450	[[170, 168, 140], [168, 163, 1
SC_24_Z_D_I_2_NE_A_II_1_D_	3	0	2	50	[[708882.9875, 8782444.887	2025-03-20 00:15:51.011	[[183, 177, 133], [183, 176, 1
SC_24_Z_D_I_2_NE_A_II_1_D_	4	0	2	50	[[708885.0875, 8782444.887	2025-03-20 00:15:51.564	[[1, 2, 2], [1, 2, 2], [1, 2, 2], [1,
SC_24_Z_D_I_2_NE_A_II_1_D_	5	0	2	50	[[708843.0875, 8782444.887	2025-03-20 00:15:52.128	[[39, 48, 39], [38, 48, 39], [37,
SC_24_Z_D_I_2_NE_A_II_1_D_	6	0	2	50	[[708878.1875, 8782444.812	2025-03-20 00:15:52.671	[[146, 122, 80], [148, 122, 80]
SC_24_Z_D_I_2_NE_A_II_1_D_	7	0	2	50	[[708887.4125, 8782444.437	2025-03-20 00:15:53.224	[[32, 46, 14], [32, 46, 14], [34,
SC_24_Z_D_I_2_NE_A_II_1_D_	8	0	2	50	[[708875.4875, 8782444.662	2025-03-20 00:15:53.817	[[143, 123, 82], [145, 123, 83]
SC_24_Z_D_I_2_NE_A_II_1_D_	9	0	2	50	[[708883.0625, 8782443.762	2025-03-20 00:15:54.376	[[17, 14, 8], [13, 12, 7], [12, 1,
SC_24_Z_D_I_2_NE_A_II_1_D_	10	0	2	50	[[708841.8125, 8782444.212	2025-03-20 00:15:54.968	[[145, 123, 85], [147, 127, 86]
SC_24_Z_D_I_2_NE_A_II_1_D_	11	1	2	50	[[708860.8625, 8782443.087	2025-03-20 00:15:55.551	[[16, 22, 16], [17, 22, 18], [18,
SC_24_Z_D_I_2_NE_A_II_1_D_	12	0	2	50	[[708883.9625, 8782443.087	2025-03-20 00:15:56.104	[[118, 93, 65], [121, 94, 66], [

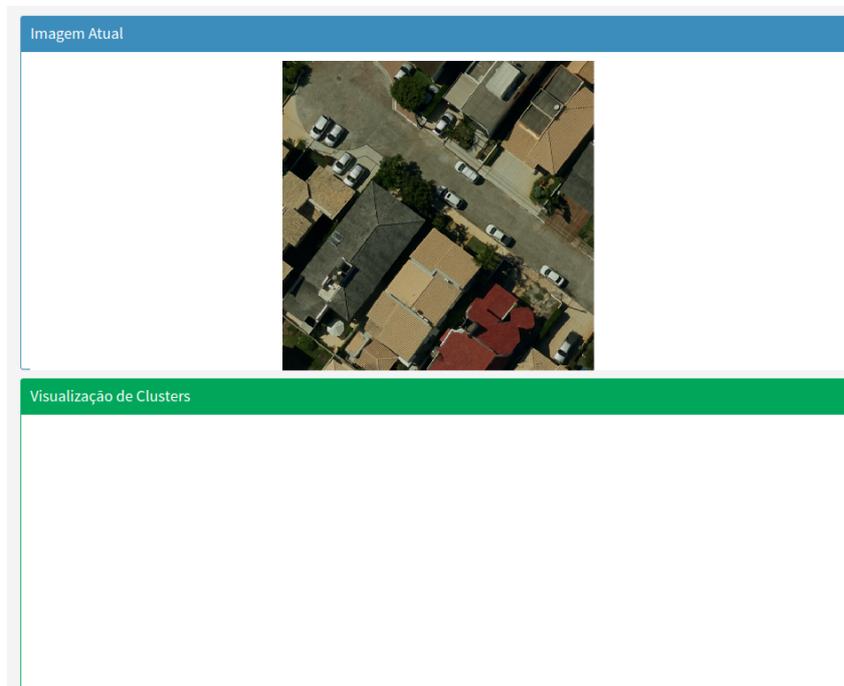
Fonte: Elaborada pelo autor.

5 RESULTADOS

Antes da aplicação do algoritmo DBSCAN, a imagem processada pelo software não apresenta qualquer segmentação visível, mantendo apenas a composição original dos pixels. Nesta fase, a imagem é carregada e exibida na interface sem a identificação de agrupamentos, permitindo ao usuário visualizar a região estudada sem qualquer alteração computacional, para melhor desempenho de tempo e processamento.

A ausência de clusters na imagem inicial reflete o estado bruto dos dados, onde não há distinção automática entre regiões homogêneas e padrões estruturais. O processo de agrupamento só ocorre após a execução do DBSCAN, que analisará a densidade dos pixels e identificará agrupamentos coerentes com os parâmetros definidos pelo usuário (*Eps* e *MinPts*). Essa etapa inicial é fundamental para garantir que o operador tenha um ponto de referência visual antes de iniciar a segmentação. Na Figura 15 pode-se observar uma imagem original, carregada no software antes da execução do algoritmo, ilustrando a visualização padrão sem a identificação de clusters.

Figura 15 – Exemplo de imagem antes da execução do algoritmo DBSCAN.



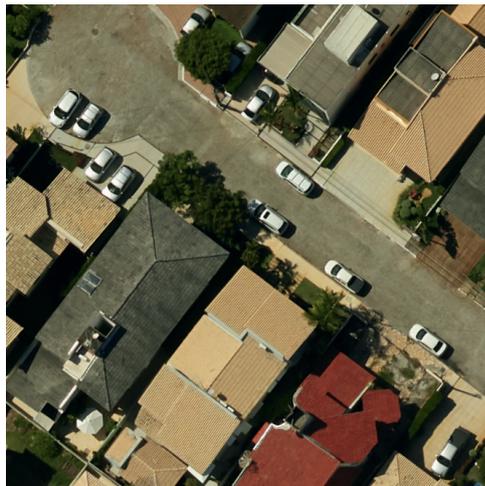
Fonte: Elaborada pelo autor.

A imagem utilizada como exemplo nesta etapa não apresenta placas solares visíveis (Figura 16), o que a torna pouco relevante para o objetivo do software. Embora sirva para ilustrar o estado inicial antes da execução do DBSCAN, essa imagem não agrega valor significativo ao processo de segmentação, pois não contém elementos de interesse para a detecção. Para o usuário

descartar essa imagem, basta ir na aba *Navegação de Imagens* e selecionar a opção *próximo* para selecionar uma nova imagem.

Em um cenário de construção de um banco de dados para identificação de padrões, a inclusão de imagens sem relevância poderia impactar a eficiência do modelo, introduzindo ruído desnecessário. No entanto, sua utilização aqui é válida para demonstrar a ausência de clusters antes da aplicação do algoritmo.

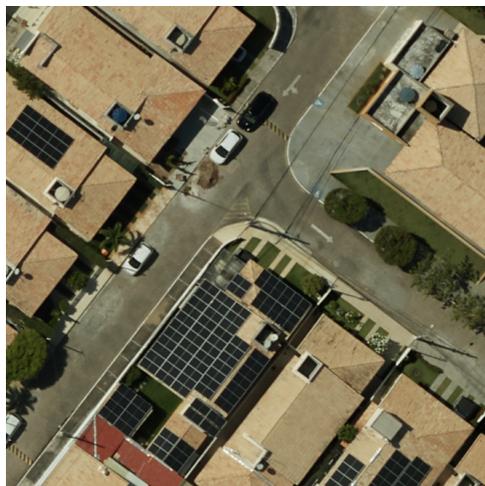
Figura 16 – Exemplo de imagem sem placas solares.



Fonte: Cooperação Técnica Nº 01/2022

A Figura 17 apresenta uma imagem que contém diversas placas solares distribuídas em diferentes residências, tornando-a altamente relevante para o processo de detecção. O algoritmo DBSCAN é aplicado para segmentação dos agrupamentos, permitindo a visualização interativa dos clusters gerados.

Figura 17 – Exemplo de imagem com várias placas solares



Fonte: Cooperação Técnica Nº 01/2022

Na Figura 18 pode-se observar uma imagem original utilizada no processo de segmentação, juntamente com a sua versão processada, onde os clusters detectados pelo DBSCAN são colocados em visualização, destacando as regiões identificadas e selecionadas pelo usuário na cor vermelha na aba *Visualização de Clusters* e roxas na aba *Imagem Atual* detectando um conjunto de placas solares.

Figura 18 – Exemplo de funcionamento do software com cluster detectado pelo algoritmo DBSCAN.



Fonte: Elaborada pelo autor.

Uma vez identificado o cluster correspondente a uma placa solar (ou conjunto de placas solares), o usuário pode clicar para salvá-lo em memória RAM. No momento da seleção, a visualização do cluster é destacada em verde, indicando que a informação está armazenada temporariamente na memória RAM. O usuário pode continuar adicionando novos clusters ou, se preferir, seguir diretamente para o salvamento no banco de dados. A Figura 19 apresenta uma imagem do software com a seleção de um cluster identificado pelo algoritmo e com placas solares salvas em memória RAM pelo usuário.

Figura 19 – Exemplo de funcionamento do software com cluster salvo em memória



Fonte: Elaborada pelo autor.

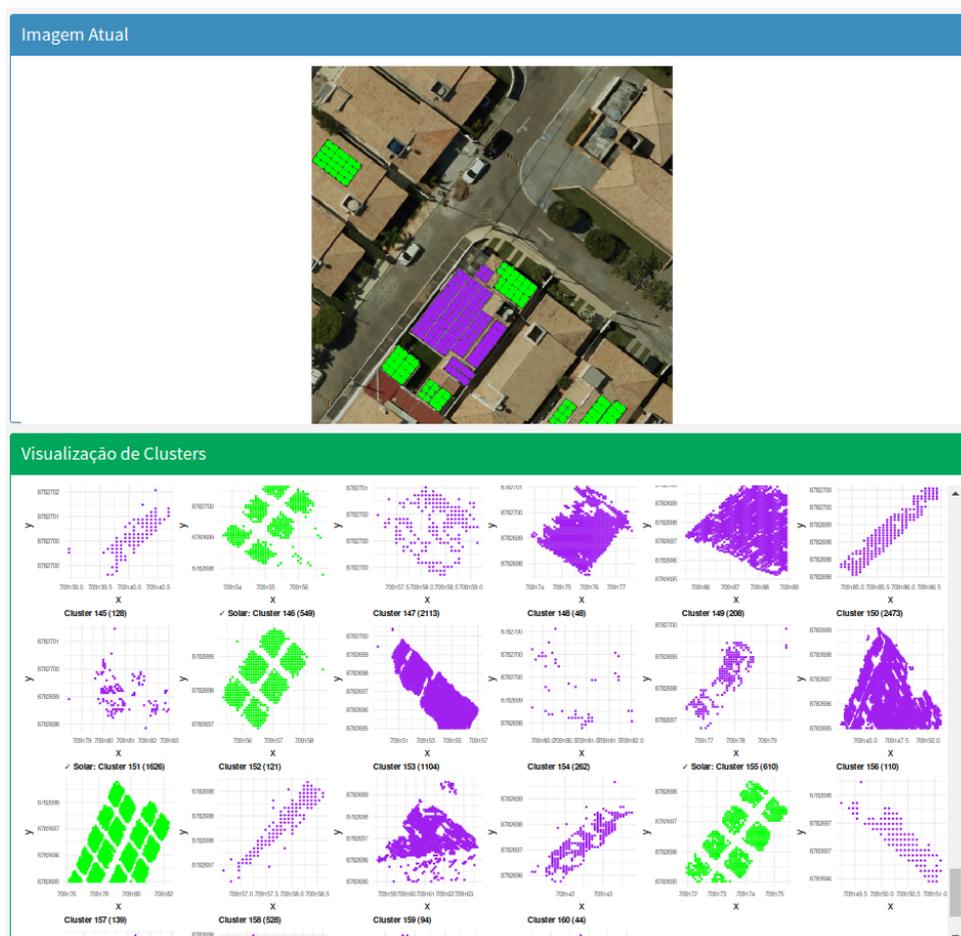
A Figura 20 apresentam um exemplo de imagem com múltiplos conjuntos de placas solares. Os clusters identificados como placas solares pelo algoritmo foram selecionados e armazenados temporariamente na memória RAM, clusters com números 140, 146, 151 e 155 destacados em verde e vistos no painel inferior da visualização. O cluster número 89, visualizado e marcado marcado em roxo na imagem de satélite, foi selecionado pelo usuário na aba *navegação de clusters* permitindo uma análise detalhada antes da confirmação. Os números dos clusters são exibidos na interface, facilitando a organização e revisão antes do salvamento definitivo no banco de dados.

Um detalhe importante do algoritmo é a identificação de grupos de placas solares de forma simultânea e até mesmo placas solares que foram divididas devido o processo de fragmentação das imagens antes de aplicação do algoritmo pelo *software* desenvolvido (ver Figura 20).

A interface do software facilita a identificação e seleção dos clusters corretos. O usuário pode selecionar manualmente os agrupamentos que representam placas solares e destacá-los em verde, indicando que foram salvos temporariamente em memória. Caso um cluster seja identificado erroneamente, ele pode ser removido da seleção antes do envio ao banco de dados.

Esse processo dinâmico assegura que apenas os clusters mais representativos sejam armazenados, melhorando a qualidade do conjunto de dados.

Figura 20 – Exemplo de imagem com reconhecimento e salvamento de diversas placas solares no software



Fonte: Elaborada pelo autor.

Além disso, banco de dados com imagens contendo múltiplas placas solares podem auxiliar futuras análises, permitindo um aprendizado mais robusto na identificação desses padrões. A presença de várias casas com painéis fotovoltaicos também oferece diversidade na segmentação, garantindo que o modelo seja treinado para reconhecer diferentes configurações e distribuições das placas solares.

A Figura 21 exibe um trecho das informações salvas no banco de dados após a seleção dos clusters. Cada linha representa um cluster identificado, pois atualmente está filtrado para mostrar apenas as linhas com designação de placar solar pela coluna *is_solar*, armazenando atributos como coordenadas, identificador do cluster, valores RGB médios, parâmetros do DBSCAN (*eps* e *MinPts*) e a marcação binária *is_solar*, que indica se o cluster foi classificado como uma placa solar. Esse formato estruturado facilita consultas eficientes e futuras análises.

Através da análise visual dos clusters gerados, foi possível observar que os agrupamentos

Figura 21 – Exemplo de imagem das informações salvas no banco

image_name	123 id_cluster	123 is_solar	123 eps	123 minpts	coords
24_Z_D_1.2_NE_A_III_1_D_rescaled_7.11	36	1	2	50	[[709141.1375, 8782734.3875],
24_Z_D_1.2_NE_A_III_1_D_rescaled_7.11	87	1	2	50	[[709165.4375, 8782717.3625],
24_Z_D_1.2_NE_A_III_1_D_rescaled_7.11	89	1	2	50	[[709162.5875, 8782716.9125],
24_Z_D_1.2_NE_A_III_1_D_rescaled_7.11	128	1	2	50	[[709151.0375, 8782705.8875],
24_Z_D_1.2_NE_A_III_1_D_rescaled_7.11	140	1	2	50	[[709155.2375, 8782701.0125],
24_Z_D_1.2_NE_A_III_1_D_rescaled_7.11	146	1	2	50	[[709157.0375, 8782699.8125],
24_Z_D_1.2_NE_A_III_1_D_rescaled_7.11	151	1	2	50	[[709179.0875, 8782698.7625],
24_Z_D_1.2_NE_A_III_1_D_rescaled_7.11	155	1	2	50	[[709174.1375, 8782698.3125],

Fonte: Elaborada pelo autor.

seguem padrões espaciais distintos, sendo possível identificar estruturas semelhantes a placas solares e diferenciá-las de outras regiões da imagem. Além disso, o método demonstrou ser eficaz na identificação de padrões relevantes, evidenciando o potencial da abordagem para a detecção automatizada posteriormente com um modelo supervisionado, desses elementos em imagens aéreas.

Após a identificação e seleção dos clusters relevantes, o software permite o salvamento das informações diretamente em um banco de dados PostgreSQL. A escolha desse sistema de gerenciamento de banco de dados foi baseada em sua robustez, suporte a grandes volumes de dados e capacidade de indexação avançada, garantindo consultas eficientes mesmo em cenários com milhares de imagens processadas.

Dado que o software lida com imagens e dados extraídos a partir do processamento dessas imagens, a estrutura do banco foi planejada para otimizar tanto o armazenamento quanto a recuperação das informações. Um dos principais desafios foi a representação eficiente das coordenadas dos clusters e dos valores RGB dos pixels pertencentes a cada agrupamento. Para isso, optou-se pelo uso do tipo de dado JSONB no PostgreSQL.

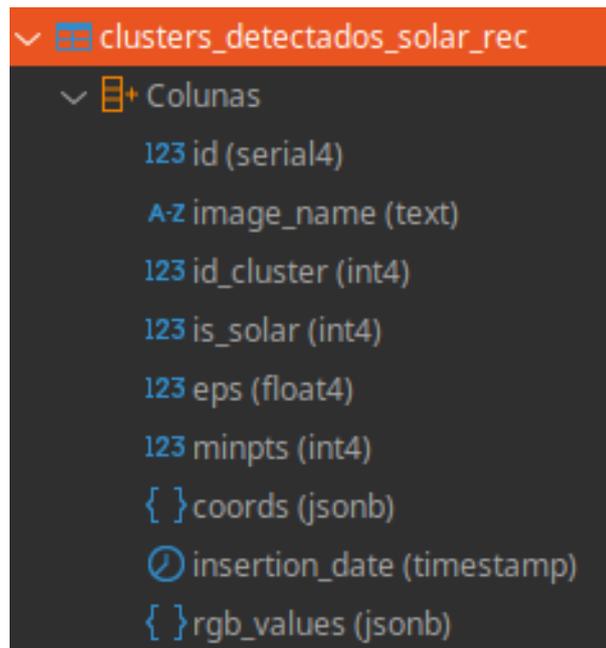
O JSONB foi escolhido por oferecer diversas vantagens, tais como: indexação eficiente para consultas rápidas sobre coordenadas espaciais e atributos dos pixels; estrutura flexível para armazenar diferentes quantidades de pontos em cada cluster sem necessidade de normalização excessiva; e facilidade na extração de informações específicas através de funções nativas do PostgreSQL para manipulação de JSON.

Cada entrada no banco de dados armazena informações essenciais para a reconstrução e análise dos clusters identificados, incluindo (ver Figura 22):

- **cluster_id** (int): Identificação única do agrupamento dentro da imagem;
- **image_name** (text): Nome da imagem original da qual o cluster foi extraído;
- **eps** e **MinPts** (float): Parâmetros utilizados na segmentação com DBSCAN;
- **coordenadas** (JSONB): Lista de pontos do cluster, armazenando suas respectivas coordenadas espaciais na imagem;
- **rgb_values** (JSONB): Valores de cor RGB dos pixels pertencentes ao agrupamento;

- **is_solar** (BOOLEAN): Indica se o cluster foi classificado como uma placa solar (1) ou não (0).
- **id**(int): id gerenciado e criado pelo banco para cada linha de arquivo.
- **insertion_date** (datetime): tempo em cada linha foi inserida, para melhor visualização e tracking de atividades, auto gerenciado no banco.

Figura 22 – Imagem exemplo do tipo das colunas armazenadas no banco de dados.



Fonte: Elaborada pelo autor.

A utilização do JSONB permitiu manter a flexibilidade necessária para armazenar informações estruturadas sem comprometer a eficiência nas consultas. Com isso, ao realizar buscas no banco de dados, é possível recuperar rapidamente clusters com características específicas, facilitando análises posteriores e aplicações em larga escala.

Essa abordagem consolidou o software como uma solução eficiente não apenas para a identificação de placas solares, mas também como uma ferramenta genérica para segmentação e análise de imagens baseadas em agrupamentos espaciais.

Além da aplicação específica para a identificação de placas solares, essa estrutura de armazenamento demonstra sua versatilidade ao possibilitar o uso do software para diversas outras finalidades. Qualquer tipo de objeto ou padrão presente em imagens pode ser segmentado e categorizado, desde que apresente características visuais que possam ser agrupadas pelo DBSCAN de acordo com valores dos parâmetros previamente estabelecidos. Aplicações potenciais incluem a identificação de áreas de vegetação em imagens aéreas, a segmentação de tecidos em imagens médicas e até mesmo a análise de padrões em imagens de sensoriamento remoto.

O sucesso da abordagem reside na combinação entre a automação do DBSCAN para detectar agrupamentos e a capacidade do usuário de validar visualmente os clusters mais relevantes. Esse equilíbrio entre modelagem automática e intervenção humana permite que o software seja utilizado em diferentes domínios sem perder a precisão necessária para a classificação final.

Um dos fatores determinantes para a eficiência do sistema é a implementação otimizada do algoritmo DBSCAN no *R*, que faz uso de bibliotecas compiladas em C++ e Fortran para garantir um tempo de execução reduzido, mesmo ao lidar com grandes volumes de dados. A utilização dessas otimizações possibilita a detecção rápida dos agrupamentos e a renderização adaptativa da visualização dos clusters em um conjunto de *N* imagens simultaneamente.

Essa estrutura dinâmica permite ao usuário explorar diferentes configurações do algoritmo (*eps* e *MinPts*) sem grandes impactos no desempenho, facilitando a identificação dos objetos de interesse dentro das imagens processadas. Assim, além de fornecer uma ferramenta eficaz para a detecção de placas solares, o sistema representa uma solução flexível para segmentação e análise de imagens em geral, garantindo escalabilidade e aplicabilidade para diferentes cenários de estudo.

6 CONCLUSÃO

Este trabalho teve como objetivo a implementação de uma abordagem automatizada para a segmentação de imagens de satélite, visando a detecção de placas fotovoltaicas. Através do algoritmo DBSCAN, foi possível identificar com sucesso os clusters relevantes nas imagens, superando desafios como variação de iluminação, ruído e similaridade entre telhados. O uso do banco de dados PostgreSQL com o tipo de dado JSONB garantiu uma eficiente organização e recuperação das informações extraídas.

A metodologia proposta atingiu seus objetivos, fornecendo uma solução prática para o mapeamento de energia solar, além de contribuir para a aplicação de técnicas de aprendizado de máquina na interseção entre Estatística, Ciência de Dados e Energias Renováveis. A análise dos parâmetros do DBSCAN, como o valor de *eps* e *minPts*, mostrou-se crucial para melhorar a segmentação e garantir uma melhor absorção das características das imagens.

Além disso, este estudo pode auxiliar no uso de métodos de aprendizado supervisionados, que poderão ser aplicados futuramente à medida que uma base de dados robusta de imagens segmentadas se torne disponível. A implementação de modelos supervisionados pode aumentar a precisão da detecção e permitir a adaptação do sistema a diferentes contextos e tipos de telhados, ampliando o alcance da aplicação para outras áreas que envolvem o reconhecimento de padrões em imagens.

Este trabalho não só cumpriu os objetivos propostos, como também forneceu uma base sólida para o desenvolvimento de soluções mais avançadas e específicas para o mapeamento de energia solar, com potenciais aplicações em outras áreas de reconhecimento de padrões em imagens.

REFERÊNCIAS

- ARACAJU, P. de. **Mapaju - Sistema de Informação Territorial**. 2025. Acesso em: 11 abr. 2025. Disponível em: <<https://map.aracaju.se.gov.br/#/publica/contribuente/mapaju>>. Citado na página 21.
- BISHOP, C. M. **Pattern Recognition and Machine Learning**. New York: Springer, 2006. ISBN 978-0387310732. Citado 3 vezes nas páginas 10, 22 e 25.
- BONACCORSO, G. **Machine Learning Algorithms: Second Edition**. Sebastopol, California, USA: O'Reilly Media, 2018. Citado na página 24.
- ENERGIA, P. **Manual e Guia Técnico de Energia Solar Fotovoltaica – Tecnologias, Projeto e Instalação**. Portal Energia, 2023. Acesso em: 18 mar. 2025. Disponível em: <https://www.portal-energia.com/manual-e-guia-tecnico-de-energia-solar-fotovoltaica-tecnologias-projecto-e-instalacao/?utm_source=chatgpt.com>. Citado na página 13.
- EPE. **Painel de Dados de Micro e Minigeração Distribuída**. 2025. Acesso em: 18 mar. 2025. Disponível em: <<https://dashboard.epe.gov.br/apps/pdgd/>>. Citado 4 vezes nas páginas 16, 17, 18 e 19.
- ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: **Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)**. Portland, Oregon, USA: [s.n.], 1996. p. 226–231. Citado na página 23.
- FRANCISCO, A. et al. Impactos da energia solar na área rural. **Revista de Gestão e Secretariado**, v. 15, p. e3542, 07 2024. Citado na página 17.
- Governo do Estado de Sergipe. **Governo do Estado inaugura usina solar fotovoltaica no SergipeTec**. 2025. Acesso em: 19 mar. 2025. Disponível em: <https://www.se.gov.br/noticias/desenvolvimento/governo_do_estado_inaugura_usina_solar_fotovoltaica_no_sergipetec>. Citado na página 17.
- JENSEN, J. R. **Remote Sensing of the Environment: An Earth Resource Perspective**. [S.l.]: Pearson Prentice Hall, 2007. Citado na página 10.
- LI, X.; WANG, S. A survey of machine learning techniques in energy systems: Applications to the renewable energy sector. **Renewable and Sustainable Energy Reviews**, v. 70, p. 230–244, 2017. Citado na página 10.
- MEHLE, A.; LIKAR, B.; TOMAŽEVIČ, D. A cluster consists of core points (red) and border points (green). core points have at. **IP SJ Transactions on Computer Vision and Applications**, v. 9, 12 2017. Disponível em: <https://www.researchgate.net/figure/A-cluster-consists-of-core-points-red-and-border-points-green-Core-points-have-at_fig2_315326812>. Citado na página 24.
- MURPHY, K. P. **Machine Learning: A Probabilistic Perspective**. [S.l.]: MIT Press, 2012. Citado na página 10.

NASCIMENTO, A. S. **Energia Solar Fotovoltaica: estudo e viabilidade no Nordeste brasileiro**. 145 p. Dissertação (Mestrado) — Universidade Federal da Paraíba, Curso de Engenharia de Produção, João Pessoa, Brasil, 2015. Disponível em: <<https://repositorio.ufpb.br/jspui/bitstream/tede/8153/2/arquivototal.pdf>>. Acesso em: 20 jan. 2021. Citado na página 15.

PEREIRA, E. B. et al. eBook. **Atlas Brasileiro de Energia Solar**. 2. ed. INPE, 2017. 80 p. Disponível em: <<http://mtc-m21b.sid.inpe.br/col/sid.inpe.br/mtc-m21b/2017/08.15.18.20/doc/thisInformationItemHomePage.html>>. Disponível em: <<http://doi.org/10.34024/978851700089>>. Citado 2 vezes nas páginas 15 e 16.

Portal Solar. **Energia Solar em Sergipe**. 2025. Acesso em: 19 mar. 2025. Disponível em: <<https://www.portalsolar.com.br/energia-solar-fotovoltaica-em-sergipe>>. Citado na página 17.

SOLIMPEKS. **PV/T Hybrid Panel**. 2025. Acesso em: 18 mar. 2025. Disponível em: <<https://solimpeks.com/product/pv-t-hybrid-panel/>>. Citado 2 vezes nas páginas 13 e 14.

STONEBRAKER, M.; ROWE, L. A. **Integrating a relational DBMS and a GIS: The POSTGRES Approach**. [S.l.]: Springer, 1986. Citado na página 10.

STRANGUETO, K. M. **Estimativa do potencial brasileiro de produção de energia elétrica através de sistemas fotovoltaicos flutuantes em reservatórios de hidroelétricas**. 147 p. Tese (Doutorado) — Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica, Campinas, Brasil, 2016. Citado na página 15.

URBANETZ, J. Qualidade de energia elétrica e o funcionamento de inversores de sistemas fotovoltaicos conectados à rede. **Revista Brasileira de Energia Solar**, Brasil, 2011. Disponível em: <https://rbens.emnuvens.com.br/rbens/article/download/55/55>. Citado 2 vezes nas páginas 13 e 15.

ZOMER, C. D. Performance compromises of building-integrated photovoltaics (bipv). **Energy and Buildings**, Brasil, 2014. Disponível em: <https://fotovoltaica.ufsc.br/sistemas/fotov/artigos-em-periodicos/>. Citado na página 13.

ZWIRTES, J.; WANDERLEY, A. C. F. Desempenho de sistema fotovoltaico/térmico (pv/t) conectado a rastreador solar de dois eixos (dast). **Anais Congresso Brasileiro de Energia Solar - CBENS**, p. 1–10, ago. 2022. Disponível em: <<https://anaiscbens.emnuvens.com.br/cbens/article/view/1046>>. Citado na página 13.