

Solaire: uma ferramenta para automatizar a coleta e análise de dados de manchas solares

Solaire: a tool to automate the collection and analysis of sunspot data

Arthur Rodrigues Proença¹, Eliseu César Miguel¹, Artur Justiniano Roberto Junior²

¹Universidade Federal de Alfenas, Instituto de Ciências Exatas, Departamento de Ciência da Computação, Alfenas, MG, Brasil.*

²Universidade Federal de Alfenas, Instituto de Ciências Exatas, Departamento de Física, Alfenas, MG, Brasil.†

Este trabalho têm como objetivo a criação de um *software* capaz de automaticamente coletar e processar dados de manchas solares com base em dados fornecidos pelo *Solar Dynamics Observatory* - SDO - acessíveis por meio do *website Solar Monitor*. Denominado Solaire, o *software* é capaz de realizar a coleta de dados de manchas solares dispondo as informações por meio de gráficos que permitem relacionar as posições das manchas ao decorrer dos dias, bem como montagens gráficas como imagens nos formatos JPEG (*Joint Photographic Experts Group*) e GIF (*Graphic Interchange Format*) que podem auxiliar no entendimento do movimento das manchas solares podendo assim se tornar um instrumento voltado ao ensino de astronomia destinado à discentes e docentes de ensino superior. Os resultados obtidos relacionados ao uso do Solaire demonstram uma nova possibilidade de se estudar o movimento de manchas solares automatizando o processo de coleta e análise de dados. Além disso, disponibilizar os dados via API (*Application Programming Interface*) possibilita o uso por outras aplicações, contribuindo para a utilização do *software* em larga escala.

Palavras-chave: manchas solares, *web-scraping*, *Solar Monitor*, Processamento de dados e imagens, ensino de astronomia

In this article, we develop software capable of automatically collecting and processing sunspot data based on information provided by the Solar Dynamics Observatory (SDO), accessible through the Solar Monitor website. Named Solaire, the software can collect sunspot data, displaying the information through charts that relate the positions of the spots over the days, as well as graphic montages such as images JPEG (Joint Photographic Experts Group) and GIF (Graphic Interchange Format) that can aid in understanding the movement of sunspots, thus making it a tool geared towards astronomy education by teachers and students. The results obtained using Solaire demonstrate a new possibility for studying the movement of sunspots by automating the data collection and analysis process, allowing for quick and accurate interpretation of sunspots. Additionally, providing the data via API enables usage by other applications, contributing to the large-scale utilization of the software.

Keywords: sunspots, *web-scraping*, *Solar Monitor*, data and image processing, astronomy education

1. Introdução

Manchas solares são regiões na superfície do Sol onde o fluxo de gás quente vindo do interior solar é interrompido, resultando em áreas de menor temperatura e intensa atividade magnética [6]. Uma mancha solar típica apresenta uma região central escura, chamada umbra, cercada por uma área mais clara, a penumbra. Essas manchas surgem frequentemente em pares com campos magnéticos orientados em direções opostas,

onde a umbra costuma possuir uma temperatura média de 3.700K, contrastando com os 5.700K da fotosfera solar. As manchas solares possuem um tamanho médio comparável ao da Terra [6] e uma duração que varia de 7 a 20 dias terrestres [14].

As manchas solares também estão associadas a erupções de massa coronal, fenômenos resultantes da complexa interação dos campos magnéticos na superfície solar. Essas erupções podem desencadear tempestades geomagnéticas, afetando significativamente as comunicações e outras tecnologias terrestres que fazem uso de ondas eletromagnéticas. Reconhecendo a impor-

*Endereço de correspondência: eliseu.miguel@unifal-mg.edu.br

†Endereço de correspondência: artur.roberto@unifal-mg.edu.br

tância de monitorar a dinâmica solar, a NASA fundou o Observatório de Dinâmicas Solares (SDO), que se descreve no próprio *site* como um "programa destinado a entender as dinâmicas solares" [7].

O Solaire é projetado para coletar e analisar dados de manchas solares de maneira automatizada, gerando gráficos que relacionam a posição das manchas na superfície solar ao longo dos dias de existência do fenômeno. Outra funcionalidade do sistema reside na possibilidade de criação de imagens em formato GIF que demonstram visualmente o deslocamento das manchas ao longo do tempo, possibilitando a observação direta das manchas solares e conectando a teoria física à prática observacional, promovendo uma compreensão mais profunda das manchas e do movimento solar.

Pensando no uso em ensino superior e na possibilidade de uso em larga escala do *software*, Solaire foi desenvolvido no formato de API, permitindo que outras aplicações façam uso do sistema, ampliando o alcance e a portabilidade da aplicação. Além disso, a visualização gráfica das manchas solares ao longo do tempo proporciona uma compreensão mais intuitiva e detalhada, beneficiando tanto a comunidade científica quanto estudantes e professores interessados no monitoramento solar

Os resultados obtidos com o desenvolvimento do Solaire demonstram eficiência na análise e coleta de dados das fontes mencionadas, como o SDO e o *Solar Monitor*. A partir desses pontos, a estrutura deste trabalho será apresentada, destacando a revisão da literatura na Seção 2, seguida pela metodologia adotada na Seção 3 (incluindo a coleta de dados 3.1, análise e processamento de dados 3.2, processamento de imagens 3.3 e otimização 3.4), os resultados obtidos na Seção 4 e as considerações finais na Seção 5.

2. Trabalhos Relacionados

Em uma busca na literatura podem ser encontrados múltiplos trabalhos relacionados à análise de dados solares de fontes como o SDO. Contudo, foram encontrados poucos trabalhos sobre fenômenos solares que tratam do estudo e da análise das manchas solares em uma perspectiva voltada ao ensino de Astronomia em nível superior utilizando-se do *Solar Monitor*.

O primeiro artigo, *Teaching Astronomy Using Tracker*, é uma proposta para fomentar o ensino de Astronomia utilizando o *software Tracker* [1]. Os autores adicionaram um novo módulo ao *tracker* para analisar as imagens das manchas solares e criar montagens gráfi-

cas indicando o período de rotação do Sol. Contudo, o processo demanda que o usuário manualmente realize a inserção das imagens no *software* de forma que os nomes possuam ordenação correta.

Como será apresentado adiante, Solaire evita casos de inserção manual de imagens, ajudando a evitar casos de erro humano e por consequência possibilita maior número de testes, reduzindo o tempo em que o pesquisador coleta dados e os organiza. Dessa forma, Solaire permite que o pesquisador mantenha o foco somente na análise do objeto de pesquisa.

No segundo artigo - *O Número de Manchas Solares, Índice da Atividade do Sol* - os autores dão destaque para o ciclo de vida solar [11]. Há possibilidade de se estimar um índice sobre a atividade do Sol, estudando quantitativamente o número de manchas bem como qualificando-as quanto à região de surgimento, podendo atribuir em qual momento de vida o Sol se encontra. O artigo exemplifica modelos já expostos da dinâmica solar, porém não cita ou faz uso de ferramenta para captura e análise de manchas solares em tempo real, justificando a confecção do *software* Solaire.

No terceiro artigo - *Sunspots on a rotating Sun* - os autores destacam a possibilidade de se estudar o movimento do Sol através das manchas solares [12]. Na metodologia, houve coleta de dados do SOHO [8] (*Solar and Heliospheric Observatory*), observatório que analisa o Sol, mais precisamente na heliosfera, de forma manual, especificando ainda distância de 24h entre cada dado. Não houve citação da forma como a coleta foi realizada, endossando novamente a possibilidade de se coletar dados solares de forma automatizada.

Apesar das ferramentas já existentes, os artigos relacionados não fazem uso de uma API que estude as manchas solares de forma automatizada ao longo dos dias. Muitos desses estudos ainda realizam análises manuais, destacando uma lacuna significativa na automatização desse processo. A criação de uma ferramenta integrada que não apenas colete e processe imagens de manchas solares, mas também realize análises automatizadas dessas imagens, apresenta uma oportunidade promissora. Ao desenvolver tal sistema, será possível oferecer ao público discente e docente de ensino superior uma solução completa e eficiente, contribuindo significativamente para o avanço do conhecimento sobre o comportamento e as características das manchas solares, facilitando e aprimorando o trabalho de análise desses fenômenos.

3. Metodologia

Para a comunidade acadêmica de ensino superior, que deseja estudar em nível introdutório fenômenos complexos como as manchas solares, coletar e analisar dados das manchas solares pode ser um desafio significativo. A construção do *software* Solaire no formato de API (*Application Programming Interface*) oferece uma solução acessível e flexível. Solaire consulta dados do *Solar Monitor* de forma automatizada e constrói representações gráficas que relacionam dados posicionais de manchas solares e a data de avistamento de cada mancha por meio de consultas nos *endpoints* dispostos na Tabela 1 que mostra as principais maneiras de se comunicar com a API.

Solaire pode ser identificada como uma API em nível 2 na classificação REST (*Representational State Transfer*) [10]. Neste nível, a API utiliza corretamente os métodos HTTP (*Hypertext Transfer Protocol*), como o *GET* para recuperação de recursos. Além disso, a API Solaire garante que cada recurso seja identificável por URIs (*Uniform Resource Identifier*) únicas, como disposto na Tabela 1, facilitando o acesso e a manipulação desses recursos por parte dos usuários e de outras aplicações. Essa abordagem contribui significativamente para a interoperabilidade e a escalabilidade da API, permitindo que ela seja facilmente integrada com outros sistemas e utilizada de forma eficiente por desenvolvedores e usuários finais.

Solaire pode ser utilizado em um fluxo de três passos distintos. Primeiramente, realiza-se uma análise prévia das condições solares para a data desejada. Em seguida, é conduzida uma avaliação das manchas solares específicas que estarão presentes nessa data determinada, conforme identificado no primeiro passo. Finalmente, para validar a existência das manchas solares identificadas, é fornecida uma demonstração visual através de um GIF. Os passos representam, respectivamente, as linhas da Tabela 1 que demonstram as URIs de acesso da API. Os parâmetros como data e manchas solares podem ser visualizados na Tabela 2 que documenta os possíveis parâmetros de uso para as URIs da Tabela 1.

Uma vez definido o formato da aplicação, torna-se possível pensar na arquitetura do sistema, como representada na Figura 1, que desenha o posicionamento da API Solaire ante o *website* e o FTP (*File Transfer Protocol*) de imagens do *Solar Monitor*, dando destaque também ao fato do *Solar Monitor* indexar e tratar dados do SDO.

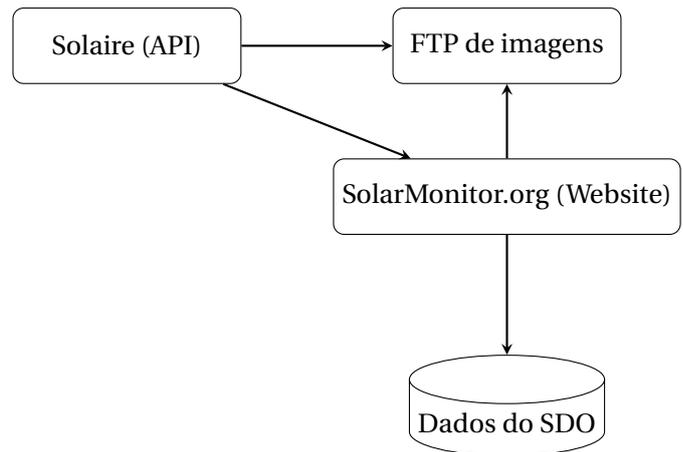


Figura 1: Desenho da arquitetura e posicionamento do sistema. Neste exemplo, Solaire representa a API fruto deste trabalho, que se encarregará de a partir da entrada do usuário, realizar a coleta de dados via *web scraping* no FTP e no website do *Solar Monitor*, indexador e classificador de dados do SDO.

Método HTTP	URI
GET	/v1/solar-monitor/jpeg
GET	/v1/solar-monitor/sunspots
GET	/v1/solar-monitor/sunspots/gif

Tabela 1: Tabela representando os métodos HTTP e URIs do *Solar Monitor*. Os *endpoints* são acessíveis pelos métodos GET e estão em nível de REST (*Representational State Transfer*) 2. Os *endpoints* fazem uso dos *query parameters* disponíveis na Tabela 2 que relaciona os parâmetros necessários para a correta execução do projeto. O *endpoint* /solar-monitor/jpeg busca retornar uma imagem do sol de acordo com a data enviada (vide Tabela 2), o *endpoint* /solar-monitor/sunspots retorna a análise das manchas solares de acordo com a data e as manchas escolhidas e o *endpoint* /solar-monitor/sunspots/gif retorna uma montagem gráfica no formato GIF (*Graphic Interchange Format*) demonstrando a mancha solar em estudo ao decorrer dos dias de existência sob o sol segundo a análise realizada no *endpoint* /solar-monitor/sunspots.

Para o desenvolvimento do Solaire, utilizamos *Python* junto ao gerenciador de dependências PIP ¹. Para a interpretação e manipulação do protocolo HTTP, utilizamos a biblioteca *FastAPI* [2] e para lidar com o *socket* ASGI (*Asynchronous Server Gateway Interface*), utilizamos o *uvicorn* [3], uma biblioteca que implementa um *gateway* assíncrono e interpreta códigos que utilizam o *FastAPI*.

¹ PyPI, disponível em <https://pip.pypa.io/en/stable/> acesso em 10 de Maio de 2024.

Parâmetro	Formato
date	Data formato AAAA-MM-DD
sunspots	Identificador NOAA
ocr	Informa o uso de OCR
linear_adjustment	Uso de regressão linear

Tabela 2: Tabela representando os *query parameters* (parâmetros) necessários para a correta execução dos endpoints do *Solar Monitor*. O parâmetro *date* define a data desejada para análise. O parâmetro *sunspots* especifica o identificador NOAA da mancha solar. O parâmetro *ocr* indica o uso da tecnologia OCR (*Optical Character Recognition*) na construção do GIF. O parâmetro *linear_adjustment* indica se deseja realizar uma regressão linear dos pontos da mancha solar.

3.1. Coleta de dados

Até a data de confecção deste artigo, não havia sido publicada uma API que permitisse consulta de dados agregados pelo *Solar Monitor*. Destarte, a coleta de dados deu-se através do processo de mineração de dados utilizando o método *web scraping* diretamente no *website Solar Monitor*.

O método de *web scraping*, que pode ser traduzido como "raspagem de dados", envolve a extração de informações de um *frontend* que disponibiliza dados por meio da linguagem de marcação HTML (*Hypertext Markup Language*). O DOM (*Document Object Model*) desempenha um papel fundamental nesse processo, pois representa a estrutura hierárquica dos elementos da página web em forma de objetos, permitindo que os dados sejam acessados e manipulados de forma programática. Dessa forma, o *web scraping* utiliza o DOM para navegar pela estrutura da página, identificar e extrair as informações desejadas de forma automatizada.

O *Solar Monitor* possui na página de informações solares uma tabela que relaciona a mancha e dados provenientes da magnetometria realizada pelo satélite do SDO. Esta tabela é o resumo das manchas solares em determinada data, como representado na Tabela 3, que relaciona o identificador NOAA (*National Oceanic and Atmospheric Administration*) à posição em um formato de coordenadas da mancha no Sol.

Ao realizar a requisição da página do *Solar Monitor* utilizando a biblioteca *requests* do Python que permite efetuar requisições HTTP, obtém-se como resposta a página HTML correspondente. O HTML da página, então, será interpretado a partir do modelo DOM o qual para análise, usou-se a biblioteca *BeautifulSoup* [4]. Essa biblioteca permite a busca programática por ele-

Identificador NOAA	Posição
13697	S19W91 (892°;-307°)
13701	S06W91 (936°;-98°)
13702	S05W91 (535°;254°)
13703	S18W34 (921°;-133°)
13704	S18W34 (504°;-300°)
13705	S15E04 (-63°;-254°)
13706	S22E17 (-255°;-378°)
13707	S24E32 (-494°;-172°)
13708	S15W65 (821°;-248°)
13711	S09E38 (-576°;-155°)
13706	S21W55 (758°;-202°)

Tabela 3: Tabela representando o HTML presente na página do *Solar Monitor* que sofrerá a extração descrita pelo processo de *web scraping*. A coluna Identificador NOAA representa a identificação da mancha e a coluna Posição fornece a posição da mancha em um formato de coordenadas. A tabela apresenta as manchas que estão na Figura 2, imagem que representa o processo de magnetograma realizado pelos satélites do SDO juntamente à classificação das manchas pelo *Solar Monitor*.

mentos utilizando comandos CSS (*Cascading Style Sheets*) e *tags* HTML.

Com a obtenção do HTML da página e o uso da biblioteca *BeautifulSoup*, torna-se possível buscar pelo HTML representado pela Tabela 3 (tabela que mostra informações sobre as manchas solares junto do Identificador NOAA e a posição das manchas solares em coordenadas) no *website* e carregar o conteúdo da tabela na memória RAM (*Random Access Memory*).

As imagens do processo de magnetograma realizadas pelo SDO são armazenadas no servidor FTP do *Solar Monitor*. Para a obtenção da imagem mais recente disponível, emprega-se também a técnica de *web scraping* para buscar essa imagem no servidor FTP. Para exemplificação, a Figura 2 apresenta as manchas solares descritas na Tabela 3 que representa a tabela HTML na página do *Solar Monitor*.

3.2. Análise e processamento de dados

A aplicação surge do desejo de facilitar ao usuário a análise e coleta de dados de manchas solares permitindo democratização de acesso à informações astronômicas bem como facilitação do ensino de astronomia voltado à dinâmica solar envolvendo manchas solares. Requisitando apenas a data de avistamento da mancha e quais manchas são de desejo, o *Solaire* deverá automaticamente percorrer pelo *Solar Monitor*

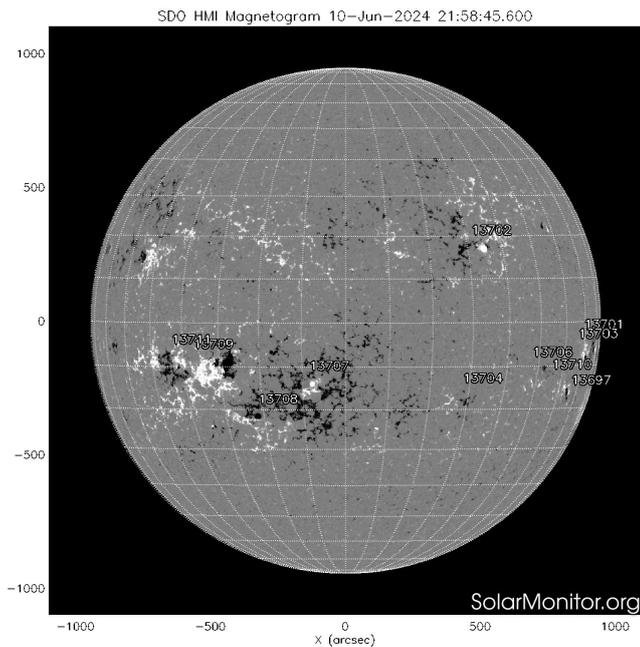


Figura 2: Imagem do Solar Monitor referente ao dia 10 de Junho de 2024. Essas imagens passam por um processamento pelo Solar Monitor para criar identificadores e inserir eixos em arcosegundos nos arredores da imagem, além de incluir um crédito no canto inferior esquerdo. Esta imagem representa os dados dispostos na Tabela 3, que apresenta os dados de posicionamento das manchas solares na mesma data. Fonte: SolarMonitor.org.

buscando os momentos em que a mancha aparece até o momento de desaparecimento ou ausência de dados da mancha.

A situação acima descrita exige uma lógica de *backtracking* da aplicação para coleta das informações, onde sistematicamente se pode realizar a busca no Solar Monitor a partir de uma data informada, de forma inicialmente regressiva até que não sejam detectadas aparições da mancha. Após isso, a aplicação, de forma progressiva, percorre o *site* até que deixe de detectar a mancha ao se encontrar na situação de ausência de dados.

Como disposto na Tabela 3, que relaciona a mancha à posição em coordenada no Sol, destacam-se na análise algumas colunas, sendo elas *Posição* e *Identificador NOAA* que descrevem, respectivamente, posicionamento na superfície solar e número de identificação da mancha.

Logo, a API busca processar as colunas *Posição* e *Identificador NOAA* de forma que a identificação da mancha seja relacionada à posição e o posicionamento seja formatado de forma que seja atribuído um par

ordenado às coordenadas Norte (N), Sul (S), Leste (E) e Oeste (W), onde cada símbolo será representado com um sinal de magnitude positivo ou negativo, relação indicada na Tabela 4.

Coordenada	Sinal
Norte (N)	-
Sul (S)	+
Leste (E)	-
Oeste (W)	+

Tabela 4: Relação coordenada polar (representado na coluna Coordenada) e sinal de magnitude (representado pela coluna Sinal). Essa parametrização é necessária para possibilitar a disposição das manchas em um plano cartesiano relacionando a posição da mancha junto à data de avistamento, além de permitir a regressão linear durante a criação do gráfico.

Ao final do processo, são formados pares ordenados que representam a posição da mancha em uma determinada data em um formato de fácil manipulação. Utilizando da biblioteca *Matplotlib* [5], é possível a produção de um gráfico que demonstra a posição das manchas no decorrer do tempo como representado pela Figura 3 que estuda o posicionamento da mancha 13697 identificada pelo Solar Monitor.

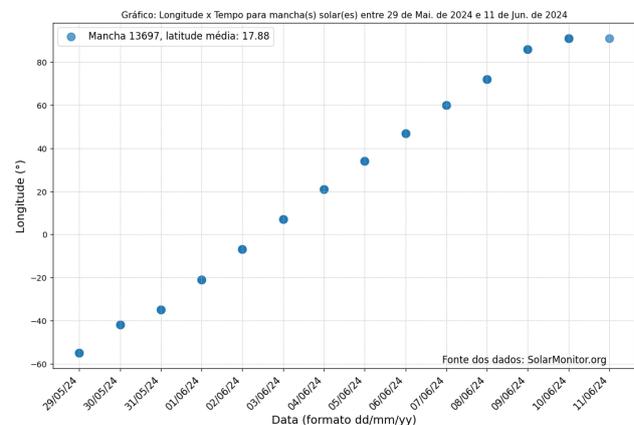


Figura 3: Gráfico gerado sem regressão linear após o processamento da mancha 13697 em dados obtidos via SolarMonitor do dia 29 de Maio ao dia 11 de Junho de 2024 após o processo de *backtracking* que coletou os dados durante a existência da mancha sob a superfície solar. Fonte: o autor.

Além disso, o usuário pode optar no momento da chamada da API, por meio de um *query parameter*, por realizar uma regressão linear dos pontos da mancha solar. Essa regressão linear permite a construção de um gráfico, representado na Figura 4, que mostra a relação entre as posições da mancha e as datas de ob-

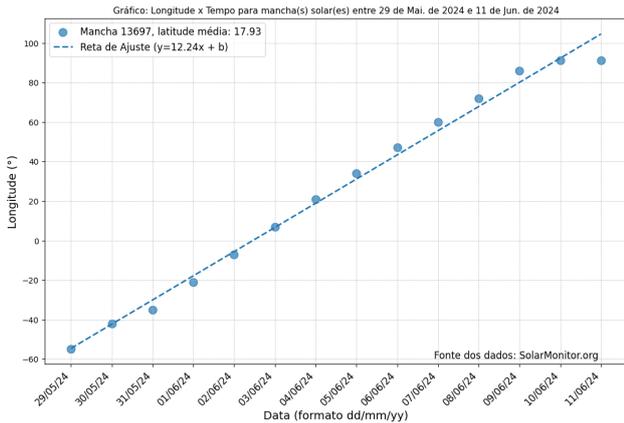


Figura 4: Gráfico gerado com regressão linear usando o algoritmo dos mínimos quadrados dos pontos obtidos após o processamento da mancha 13697 em dados obtidos via *SolarMonitor* do dia 29 de Maio ao dia 11 de Junho de 2024 após o processo de *backtracking* que coletou os dados durante a existência da mancha sob a superfície solar. A regressão linear das posições da mancha permite estimar um coeficiente angular que representa a velocidade média de deslocamento da mancha no decorrer do tempo. Fonte: o autor.

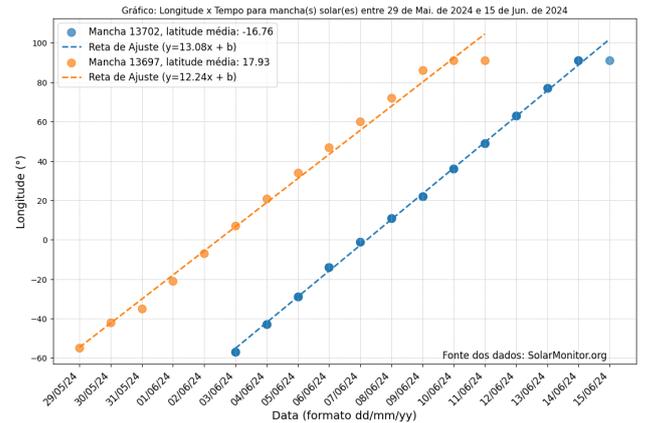


Figura 5: Gráfico gerado com regressão linear usando o algoritmo dos mínimos quadrados dos pontos obtidos após o processamento das manchas 13697 e 13702 em dados obtidos via *SolarMonitor* do dia 29 de Maio ao dia 15 de Junho de 2024 após o processo de *backtracking* que coletou os dados durante a existência da mancha sob a superfície solar. A regressão linear das posições da mancha permite estimar um coeficiente angular que representa a velocidade média de deslocamento da mancha no decorrer do tempo e comparar diferentes coeficientes angulares da reta permite visualizar a característica de rotação diferencial do Sol. Fonte: o autor.

servação junto a uma equação da reta que perpassa sobre os pontos, gerando um coeficiente angular capaz de permitir inferir uma velocidade de movimento da mancha. Ao se requisitar mais de uma mancha para análise na API, se torna possível comparar diferentes velocidades de movimento para manchas diferentes, como na Figura 5, tornando visível a característica de rotação diferencial do Sol [1], pois torna-se possível perceber pontos em diferentes latitudes com diferentes velocidades de deslocamento, como ao comparar o coeficiente angular da reta entre as manchas 13697 e 13702.

3.3. Processamento de imagens

As imagens armazenadas na aplicação são submetidas a um tratamento que remove informações não úteis para este problema. Este processamento busca remover dados gráficos do *Solar Monitor* como os eixos cartesianos em arcosegundos e o título que mostra a data do processo de magnetograma.

Desta forma, conforme a Figura 2 que representa a imagem original disponível no *Solar Monitor* no dia 10 de Junho de 2024, deseja-se remover a parte branca (parte que representa os eixos cartesianos em arcosegundos junto de informações de título da imagem) ao redor do retângulo preto da imagem, deixando apenas as manchas e as identificações, mantendo os créditos ao Solar Monitor, fonte da imagem.

Para isso, emprega-se a biblioteca *OpenCV* [9]. Primeiramente, a imagem é convertida para escala de cinza e é aplicado um desfoque para reduzir ruídos. Em seguida, a imagem é binarizada para destacar elementos escuros, supondo que o retângulo seja preto. A etapa seguinte envolve a identificação do maior contorno interno, possivelmente correspondente ao retângulo preto dentro da área branca da imagem. Se esse contorno for encontrado, a região retangular correspondente é recortada e a data é adicionada no canto superior esquerdo da imagem recortada. Caso nenhum contorno seja identificado, a imagem original é retornada sem alterações.

Por fim, a data é adicionada à imagem final no canto superior esquerdo usando coordenadas estáticas para indicar a localização. É possível visualizar o efeito deste algoritmo de processamento de imagens na Figura 6. A Figura 6 demonstra o resultado do algoritmo descrito neste tópico ao submeter a Figura 2 ao processamento de imagens buscando uniformidade e simplificação das imagens que o *Solaire* irá trabalhar.

No momento da chamada da API para criação do GIF, o usuário pode também informar se deseja destacar o identificador das manchas as quais foram realizadas as buscas pela lógica de *backtracking* descrita na Seção 3.2, por meio do *query parameter* OCR. O OCR (*Optical Character Recognition*) é uma tecnologia que

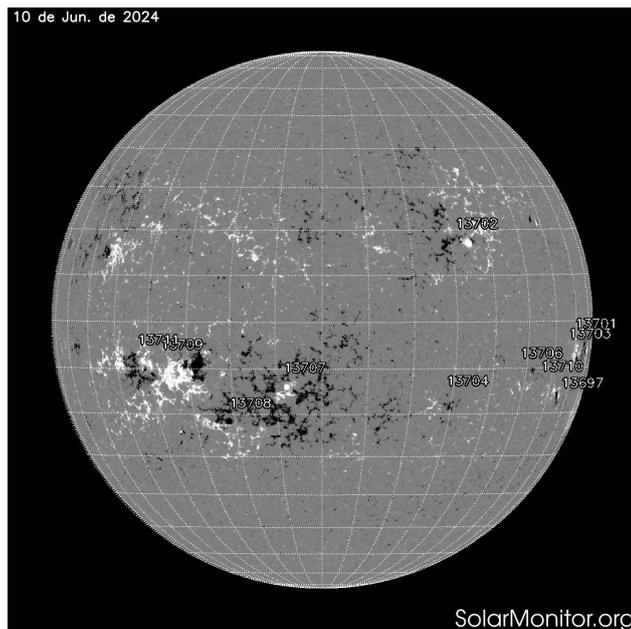


Figura 6: Imagem processada com base na Figura 2 utilizando o algoritmo descrito no tópico 3.3, algoritmo este que remove informações de gráfico em arcosegundos e demais metadados. A imagem resultante possui no canto superior esquerdo o descritivo da data da imagem, também mantendo os créditos ao *Solar Monitor* no canto inferior direito. Ao processar a imagem desta maneira, se torna possível criar montagens gráficas em formato de GIF, fornecendo uma melhor visualização das manchas solares analisadas no decorrer do tempo.

permite o reconhecimento de texto em imagens digitais. O *EasyOCR*[13] é uma biblioteca *Python* de OCR que se destaca pela facilidade de uso e eficiência. A biblioteca *EasyOCR* suporta vários idiomas e é capaz de reconhecer diferentes tipos de fontes e estilos de texto em imagens, incluindo textos inclinados, com distorções ou em diferentes tamanhos. Além disso, o *EasyOCR* é uma ferramenta de código aberto, o que facilita a integração em projetos de análise de imagens e processamento de texto.

Ao utilizar o *EasyOCR* em conjunto com a geração de GIF para visualização das manchas solares em análise, adiciona-se uma camada de interatividade e aprimoramento na apresentação dos resultados. Com o *query parameter* ocr na chamada da API para criação do GIF, é possível automatizar o processo de destaque das manchas desejadas nas imagens, tornando a visualização mais clara e facilitando a interpretação dos dados para os usuários finais, uma vez que a(s) mancha(s) de estudo agora estão destacadas na imagem. Essa integração entre tecnologias de processamento de imagem e OCR demonstra como a combinação de ferramentas pode agregar valor e melhorar a experi-

ência de análise de dados visuais para os usuários do *Solaire*.

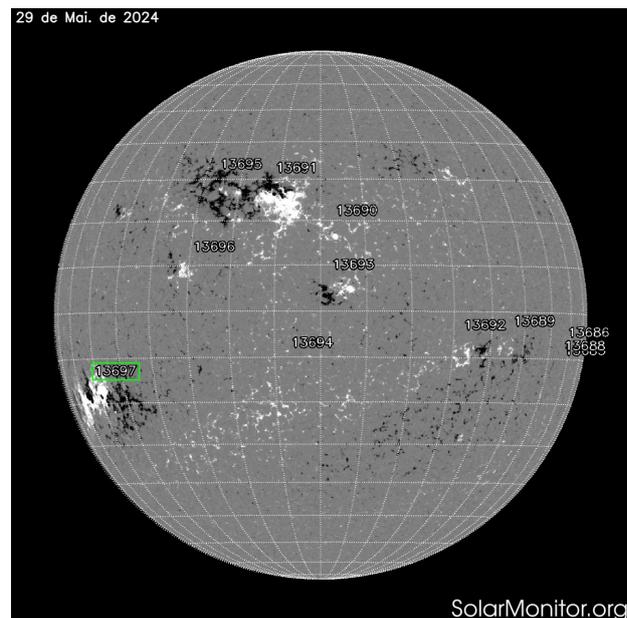


Figura 7: Imagem estática representado um *frame* do GIF gerado com base no processamento da mancha 13669 utilizando o algoritmo descrito no tópico 3.3, onde se torna possível por meio da tecnologia de OCR (*Optical Character Recognition*) reconhecer o identificador da mancha e o destacar com um retângulo de cor verde.

3.4. Otimização

Tendo em vista a natureza da aplicação, um cenário se torna possível: múltiplas requisições a uma mesma página do *Solar Monitor*, assemelhando-se a um ataque DDoS (*Distributed Denial of Service*). Isso pode acontecer quando múltiplos usuários requisitarem uma mesma data e por consequência da lógica de *backtracking* anteriormente descrita, gerar um número de requisições elevado a um mesmo recurso disposto no *website Solar Monitor*.

Para evitar esse cenário, foi implementada uma lógica de *cache* autoral, que baseia-se na inserção dos resultados do *Solar Monitor* com base na data solicitada na memória RAM do servidor, criando uma lógica de inserção chave valor, onde onde a chave deve ser única, sendo no contexto da aplicação a data de análise e os valores sendo a imagem e os de dados coletados.

Surge também um terceiro atributo no algoritmo autoral de *cache* que se encarregará de otimizar o consumo e evitar o preenchimento total da memória RAM (*Random Access Memory*), sendo ele uma *thread* virtual que ficará encarregada de a cada intervalo de tempo,

fazer a remoção dos itens anteriormente armazenados da memória RAM para que dessa forma, cada item inicie com o intervalo de tempo descrito e por consequência tenha um tempo de vida útil de alocação e evitando sobrecarga da memória na memória RAM.

Com a lógica de chave-valor e a thread virtual implementadas no algoritmo autoral de *cache*, sempre que houver uma requisição ao servidor ele primeiro verificará se a data requisitada tenha registro na *cache*. Caso positivo, ele retorna o conteúdo da memória cache referente à data e em caso negativo, busca no *Solar Monitor* e armazena na cache, eliminando os casos onde uma mesma data é requisitada e gera uma nova requisição ao *Solar Monitor*.

4. Resultados

Os resultados obtidos mostram que é possível tornar o processo de coleta e análise de dados do SDO disponíveis no *Solar Monitor* automático, facilitando o ensino de dinâmica solar envolvendo o movimento de manchas solares. Ao utilizar-se do Solaire, é possível visualizar o movimento de manchas solares por meio de gráficos que relacionam a posição das manchas ao decorrer do tempo bem como com GIF que criam a ilusão de movimento das manchas entre imagens do SDO, aliando a teoria física à observação, permitindo compreensão visual do fenômeno das manchas solares.

Como os trabalhos relacionados revelaram a lacuna de uma aplicação que automatize a coleta e análise de dados de manchas solares, Solaire também se mostra inovador. Outro ponto a ser destacado é que como processo pôde ser sistematizado, há também a diminuição de erros que podem ocorrer dada a complexidade da operação quando realizada manualmente.

5. Considerações finais

Neste trabalho, foi implementado um software capaz de automatizar a análise e coleta de dados de manchas solares. Solaire se mostrou uma ferramenta eficaz e automatizada para a análise de manchas solares, utilizando dados do *Solar Monitor*.

Ao aplicar técnicas de *web scraping*, processamento de dados via *Python 3* e o *framework FastAPI*, conseguimos explorar de forma qualitativa a evolução e as posições dessas manchas ao longo do tempo.

Além disso, a utilização de regressões lineares e a criação de gráficos e GIF com o *OpenCV* contribuíram sig-

nificativamente para a compreensão dos fenômenos solares estudados. Os resultados obtidos demonstram a viabilidade e a vantagem da automatização desse processo em comparação com métodos manuais, oferecendo uma nova perspectiva e uma valiosa contribuição para a pesquisa sobre o comportamento das manchas solares bem como democratização do acesso a informações relacionadas à dinâmica solar.

O projeto está disponível no GitHub² e pode ser configurado para execução local conforme o arquivo *README.md* disponível na raiz do diretório.

Referências

- [1] CASTANHEIRA, S.; FORYTA, K. *Método simplificado para calcular a rotação do sol*, V Simpósio Nacional de Educação em Astronomia – V SNEA 2018 – Londrina, PR, 2018.
- [2] FASTAPI. *About - FastAPI*, disponível em: <<https://fastapi.tiangolo.com/about/>>. Acesso em: 06 mai. 2024.
- [3] UVICORN. *Uvicorn*, disponível em: <<https://www.uvicorn.org/>>, acessado em 06 de maio de 2024.
- [4] JONES, L. *BeautifulSoup Documentation*, disponível em: <<https://beautiful-soup-4.readthedocs.io/en/latest/>>. Acesso em: 06 mai. 2024.
- [5] MATPLOTLIB. *Matplotlib: Visualization with Python*, disponível em: <<https://matplotlib.org/>>. Acesso em: 06 mai. 2024.
- [6] NASA. *How Are Magnetic Fields Related To Sunspots?*, disponível em: <https://soho.nascom.nasa.gov/explore/lessons/sunspots6_8.html>. Acesso em: 05 mai. 2024.
- [7] NASA. *About the SDO Mission*, disponível em: <<https://sdo.gsfc.nasa.gov/mission/>>. Acesso em: 06 mai. 2024.
- [8] NASA. *SOHO-About*, disponível em: <<https://soho.nascom.nasa.gov/about/about.html>>. Acesso em: 06 mai. 2024.
- [9] OPENCV. *opencv-python*, disponível em: <<https://pypi.org/project/opencv-python/>>. Acesso em: 06 mai. 2024.
- [10] PROGRAMMERS.COM.BR. *Níveis de maturidade de uma api REST*, disponível em: <<https://www.programmers.com.br/blog/niveis-de-maturidade-de-uma-api-rest/>>. Acesso em: 20 mai. 2024.
- [11] RIBEIRO, R.; SARAIVA, A. *O Número de Manchas Solares, Índice da Atividade do Sol*, disponível em: <www.scielo.br/j/rbef/a/MC3Wc4LDPxCYFVX4RrzQLZq>. Acesso em: 06 mai. 2024.

²<https://github.com/ArthurProenca/SolarMonitor>

- [12] RIBEIRO, R.; SARAIVA, A. *Sunspots on a rotating Sun*, disponível em: <<https://scienceinschool.org/wp-content/uploads/2016/03/issue35-sunspots.pdf>>. Acesso em: 06 mai. 2024.
- [13] EASYOCR. *EasyOCR*, disponível em: <<https://www.jaided.ai/easyocr/>>. Acesso em: 20 mai. 2024.
- [14] SMITH, J. *Sun*, HowStuffWorks, disponível em: <<https://science.howstuffworks.com/the-solar-system-channel.htm>>. Acesso em: 05 mai. 2024.