

## 16º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2025

### USO DA LINGUAGEM PYTHON E INTELIGÊNCIA ARTIFICIAL GENERATIVA EM APLICAÇÕES VOLTADAS A MICROSCOPIA DE FORÇA ATOMICA

LUCAS H. CONTEL<sup>1</sup>, EVELYN G. V. ESPEDILLA<sup>2</sup>, MIRELA C. BARROS<sup>2</sup>, FLAVIANA B. ANDRADE<sup>3</sup>, ALLAN V. RIBEIRO<sup>4</sup>

<sup>1</sup> Graduando em Engenharia da Computação, Bolsista PIBIC, IFSP, Campus Birigui, l.contel@ifsp.edu.br.

<sup>2</sup> Pós-graduanda no Departamento de Dentística, Endodontia e Materiais Odontológicos, Faculdade de Odontologia de Bauru, Universidade de São Paulo.

<sup>3</sup> Docente no Departamento de Dentística, Endodontia e Materiais Odontológicos, Faculdade de Odontologia de Bauru, Universidade de São Paulo.

<sup>4</sup> Docente na Licenciatura em Física, Instituto Federal de São Paulo Câmpus Birigui.

Área de conhecimento (Tabela CNPq): 1 .02.02.08-0 Análise de Dados

**RESUMO:** Este trabalho teve como objetivo desenvolver e validar um fluxo automatizado para a coleta e análise de parâmetros topográficos obtidos por microscopia de força atômica (MFA), visando reduzir tempo de processamento, minimizar erros e simplificar a rotina de análise. Para isso, foi criado um script em Python, integrado ao ambiente PyGwy, capaz de extrair parâmetros topográficos de múltiplas imagens de MFA, armazenando em formato padronizado para posterior análise. Em paralelo, os dados processados foram avaliados por três ferramentas de inteligência artificial generativa: ChatGPT, DeepSeek e Gemini. Os dados gerados foram comparados com resultados de um software de referência, Prism. A análise revelou que o ChatGPT apresentou desempenho mais consistente, com baixas variações na maioria dos parâmetros, exceto Ssk. O DeepSeek mostrou desempenho intermediário, com precisão em alguns parâmetros, mas variação elevada em Ssk e Sku. O Gemini apresentou as maiores discrepâncias entre eles. De forma geral, para parâmetros distintos de Ssk e Sku, todas as IAs mantiveram divergências inferiores a 15%, indicando viabilidade para análises preliminares. Conclui-se que a combinação entre automação e IA pode tornar a análise topográfica mais ágil, acessível e menos dependente de softwares especializados, permitindo otimização do esforço e tempo dos pesquisadores para outras áreas.

**PALAVRAS-CHAVE:** microscopia de força atômica; automação; python; inteligência artificial; análise de dados; topografia de superfície

### USE OF PYTHON LANGUAGE AND GENERATIVE ARTIFICIAL INTELLIGENCE FOR ASSISTING IN THE COLLECTION AND ANALYSIS OF TOPOGRAPHIC DATA FROM ATOMIC FORCE MICROSCOPY

**ABSTRACT:** This study focused on developing and validating an automated workflow for the collection and analysis of topographic parameters obtained through atomic force microscopy (AFM), with the goal of reducing processing time, minimizing errors, and simplifying the analysis routine. A Python script integrated with the PyGwy environment was developed to extract topographic parameters from multiple AFM images and store them in a standardized format for further analysis. In parallel, the processed data were evaluated using three generative artificial intelligence tools: ChatGPT, DeepSeek, and Gemini. The generated data were compared with results from a reference software, Prism. The analysis showed that ChatGPT delivered the most consistent performance, with low variations in most parameters except for Ssk. DeepSeek demonstrated intermediate performance, achieving accuracy in some parameters but showing high variation in Ssk and Sku. Gemini exhibited the largest discrepancies

overall. In general, for parameters other than Ssk and Sku, all AIs maintained divergences below 15%, indicating feasibility for preliminary analyses. It is concluded that the combination of automation and AI can make topographic analysis faster, more accessible, and less dependent on specialized software, optimizing researchers' time and effort for other tasks.

**KEYWORDS:** atomic force microscopy; automation; python; artificial intelligence; data analysis; surface topography

## INTRODUÇÃO

A microscopia de força atômica (MFA) é amplamente utilizada para caracterizar superfícies em escala nanométrica, permitindo a obtenção precisa de parâmetros topográficos (Eaton, 2010). A análise e interpretação desses dados exigem ferramentas especializadas capazes de extrair informações quantitativas de forma confiável, sendo a aplicação correta de métodos estatísticos fundamental para compreender a distribuição e a variabilidade dos resultados (Foster, 2013).

Entretanto, muitos softwares disponíveis apresentam limitações, como interfaces gráficas que exigem análise manual de cada amostra, tornando o processo repetitivo, demorado e mais suscetível a erros (Beton, 2021).

O Python consolidou-se como uma das principais linguagens para ciência de dados, destacando-se pela clareza sintática e pelo suporte de bibliotecas como o Pandas, amplamente utilizado para manipulação e análise de dados estruturados (VanderPlas, 2017). Aproveitando essa versatilidade, este trabalho integrou o Python ao ambiente PyGwy, do software livre Gwyddion, para desenvolver um script capaz de automatizar a coleta de parâmetros topográficos, eliminando etapas manuais e aumentando a eficiência do processo.

Além disso, a European Commission (2024) aponta que a inteligência artificial (IA) generativa apresenta grande potencial para transformar processos de pesquisa, acelerando análises, promovendo avanços científicos e aumentando a produtividade. Nesse contexto, este estudo combinou a automação via script Python com a análise inicial por IA generativa, buscando reduzir o tempo de processamento, minimizar erros e avaliar a acurácia dessas ferramentas em relação a um software de referência, com o objetivo de democratizar e facilitar a análise de parâmetros topográficos.

## MATERIAL E MÉTODOS

A caracterização topográfica das amostras de Gutta-Percha foi realizada por meio de um microscópio de força atômica FlexAFM – NanoSurf, operando no modo de força dinâmica (Dynamic Force) e em regime de não contato, associado à técnica de contraste de fase (Phase Contrast).

As imagens obtidas foram processadas no software Gwyddion v2.63, utilizado para remoção de ruídos, correção de imperfeições e extração de parâmetros topográficos. Para otimizar a coleta de dados, foi desenvolvido um script de automação no ambiente PyGwy, utilizando Python v2.7.16 e as bibliotecas PyCairo, PyObject e PyGTK. Esse script, baseado na documentação oficial do Gwyddion, permitiu a extração simultânea dos principais parâmetros topográficos de múltiplas imagens de AFM, armazenando-os em planilhas no formato .xls.

A fim de validar o funcionamento do script, os resultados obtidos automaticamente foram comparados aos valores apresentados pela interface gráfica do Gwyddion, que processa apenas uma imagem por vez. O procedimento foi repetido dez vezes, utilizando conjuntos distintos de imagens, confirmando a equivalência entre os dados gerados automaticamente e os obtidos manualmente.

O próprio script também executou, de forma automatizada, o cálculo da média e do desvio padrão dos parâmetros, utilizando a biblioteca Pandas v0.24.2. Esses resultados foram utilizados na análise comparativa das amostras de Gutta-Percha, gerando uma planilha que serviu de base para o processamento por ferramentas de inteligência artificial (IA) generativa.

Foram avaliadas três plataformas: ChatGPT Plus (OpenAI, 2025), DeepSeek-V3 (DeepSeek, 2025) e Gemini 2.5 Pro (Google, 2025). Todas receberam as mesmas instruções (*prompts*), garantindo a comparabilidade dos resultados. Antes da análise principal, realizaram-se testes preliminares para assegurar que os dados das planilhas estavam sendo interpretados corretamente e para evitar vieses, utilizando contas sem histórico de conversas e com memória desativada.

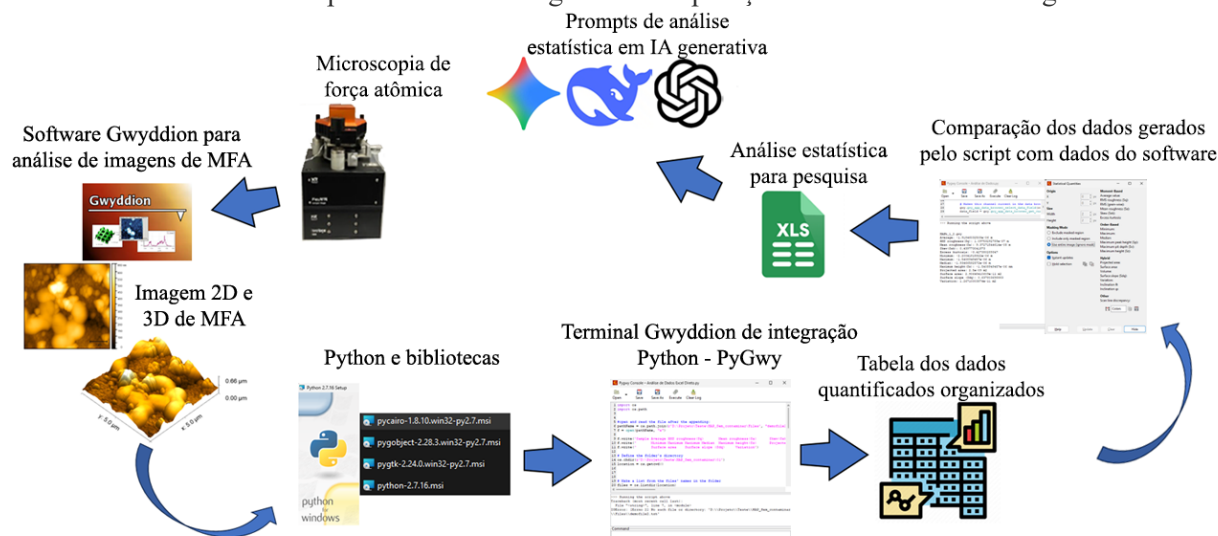
A análise por IA foi realizada mediante o envio direto das planilhas .xls, acompanhadas de comandos solicitando interpretação estatística e aplicação de testes específicos. Os resultados foram

documentados e comparados aos obtidos no Prism 6.01 (GraphPad Inc.), adotado como referência pela sua reconhecida confiabilidade em análises estatísticas.

O estudo considerou um grupo experimental composto por 12 amostras. Para cada amostra, foram determinados seis parâmetros de caracterização da superfície: Sa (rugosidade média aritmética), Sq (rugosidade quadrática média), Sz (diferença máxima entre picos e vales), Ssk (assimetria em relação ao plano médio), Sku (curtose, relacionada à concentração ou dispersão das alturas) e Sdr (Developed Area Ratio, razão entre a área real da superfície e sua projeção bidimensional).

A sequência de etapas para o desenvolvimento e validação do script, assim como o fluxo completo de aquisição e análise de dados, incluindo a comparação entre as ferramentas de IA, está representado na Figura 1.

FIGURA 1. Modelo do processo metodológico de comparação dos dados com as IA generativas:

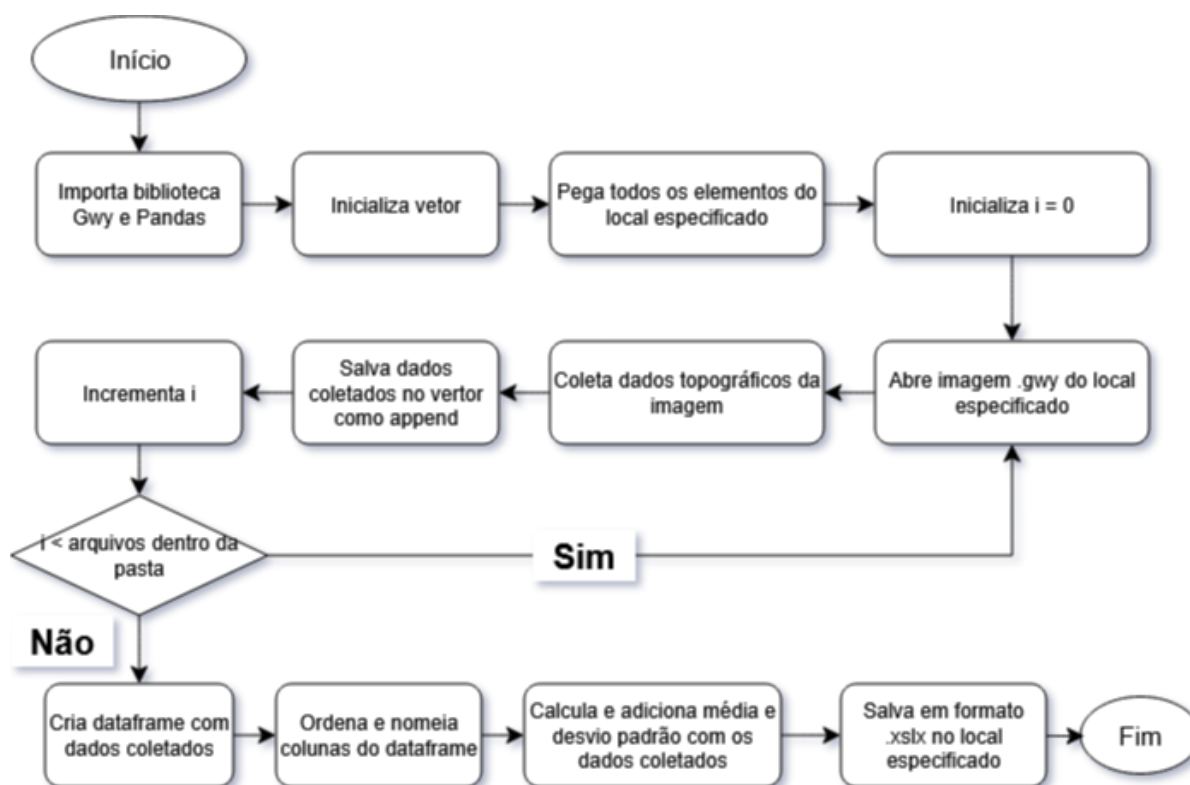


## RESULTADOS E DISCUSSÃO

O software Gwyddion foi empregado para a extração e quantificação dos dados topográficos de cada amostra. A utilização do script desenvolvido no ambiente PyGwy permitiu processar múltiplas imagens de microscopia de força atômica (MFA) de forma automatizada, eliminando a necessidade de análises manuais repetitivas e garantindo maior precisão e padronização no armazenamento das informações. Esse processo está esquematizado no fluxograma apresentado na Figura 3.

Conforme descrito por Model (2009), uma base de dados relacional organiza informações em tabelas compostas por linhas e colunas, sendo cada coluna associada a um parâmetro específico. Nesse estudo, a estrutura relacional gerada a partir do script serviu como base para as análises estatísticas conduzidas no software Prism 6.01 e para a comparação dos resultados obtidos pelas diferentes ferramentas de inteligência artificial (IA) generativa.

FIGURA 2. Fluxograma mostrando processos de execução do script de coleta de dados de imagens de MFA criado no ambiente PyGwy.



O banco de dados relacional produzido foi enviado, via *prompt*, para três IAs generativas: ChatGPT (modelo GPT-4 com Data Analyst), DeepSeek-V3 e Gemini 2.5 Pro. Com os resultados fornecidos por cada IA, foram construídos gráficos comparativos, incluídos na Figura 4, utilizando como referência os valores obtidos pelo Prism. A partir dessa referência, calculou-se a disparidade percentual de cada IA em relação ao valor-base.

A análise gráfica e estatística dos dados mostrou que o ChatGPT apresentou o desempenho mais consistente no conjunto dos parâmetros avaliados, com valores de disparidade próximos de zero em quase todas as métricas. Em parâmetros como Sq, Sa, Sz, Sdr e Sku, a variação média foi praticamente nula, indicando alta estabilidade. A única exceção relevante foi no Ssk, onde o percentil 25% apresentou variação de aproximadamente 55,4%, sinalizando maior sensibilidade dessa IA a esse parâmetro específico, possivelmente devido à forma como trata assimetrias acentuadas na topografia.

O DeepSeek-V3 demonstrou desempenho intermediário. Em alguns parâmetros, como Sa e Sz, aproximando-se dos resultados do ChatGPT, evidenciando boa precisão. Contudo, apresentou variações expressivas em Ssk (média de 29,7%, com picos acima de 40% na mediana) e Sku (6,7%), o que compromete a confiabilidade em análises que dependam desses indicadores para avaliação de irregularidades e curtose da superfície.

O Gemini 2.5 Pro obteve, na maioria dos casos, as maiores discrepâncias em relação ao valor-base. Apesar de apresentar resultados aceitáveis para Sa e Sdr, registrou variações elevadas em Sku (21,3% de média) e Ssk (16,1%), com medianas de até 39,1%. Esses valores indicam que, embora a ferramenta possa ser útil em parâmetros mais simples, sua performance é inconsistente nos indicadores mais sensíveis.

De forma geral, ao observar as escalas dos gráficos, nota-se que para todos os parâmetros, com exceção de Ssk e Sku, as três IAs mantiveram variações inferiores a 15% em relação ao Prism. Essa margem pode ser considerada adequada para análises exploratórias ou de triagem inicial, reduzindo a necessidade de softwares especializados caros e complexos. Entretanto, para aplicações que exijam alta precisão metrológica, especialmente na avaliação de assimetria e curtose da superfície, o uso exclusivo de IA ainda demanda cautela.

FIGURA 3. Comparação gráfica dos valores obtidos pelas IAs generativas em relação ao valor-base do Prism em disparidade percentual.



## CONCLUSÕES

O desenvolvimento do script em Python, integrado ao ambiente PyGwy, automatizou a coleta dos parâmetros topográficos obtidos por microscopia de força atômica, substituindo um procedimento anteriormente manual, repetitivo e vulnerável a erros. Essa automação resultou em maior eficiência, padronização e confiabilidade, reduzindo o risco de inconsistências e eliminando a necessidade de processar cada imagem individualmente.

A avaliação das ferramentas de inteligência artificial generativa evidenciou que o ChatGPT apresentou o desempenho mais consistente, com baixas variações na maioria dos parâmetros e discrepâncias relevantes apenas no Ssk. O DeepSeek-V3 mostrou desempenho intermediário, com boa precisão em alguns parâmetros, mas variações expressivas em Ssk e Sku. O Gemini 2.5 Pro obteve os resultados menos consistentes, embora tenha apresentado boa acurácia em métricas mais simples. De forma geral, para todos os parâmetros, exceto Ssk e Sku, as três IAs mantiveram variações inferiores a 15% em relação ao valor-base do Prism, indicando potencial para uso em análises exploratórias ou de triagem.

Assim, tanto o script desenvolvido quanto a aplicação das IAs convergem para o mesmo propósito: simplificar e democratizar o trabalho do pesquisador, permitindo análises rápidas e acessíveis sem dependência de softwares especializados de alto custo ou complexidade excessiva. A combinação entre automação de coleta e interpretação assistida por IA representa um avanço na modernização das metodologias de análise topográfica, proporcionando maior agilidade, escalabilidade e inclusão no ambiente científico.

## CONTRIBUIÇÕES DOS AUTORES

L.H.C desenvolveu os scripts em Python, efetuou a validação dos dados gerados, utilizou os dados para criação das instruções e solicitou para as IAs gerarem análise topográfica, organizou os dados gerados em tabelas e efetuou a comparação com os dados obtidos via Prism e confecção do artigo.

A.V.R. fez o tratamento e coleta das imagens de MFA das amostras, e gerou os dados utilizados como base através do software Prism assim como revisão e ajustes no artigo.

E. G. V. E., M.C.B., F.B.A. auxiliaram na confecção e no protocolo de contaminação das amostras.

Todos os autores contribuíram com a revisão do trabalho e aprovaram a versão submetida.

## AGRADECIMENTOS

Os autores agradecem ao Instituto Federal de São Paulo (IFSP) Campus Birigui pelo suporte institucional e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de Iniciação Científica (PIBIC).

## REFERÊNCIAS

BETON, J. G. et al. TopoStats – a program for automated tracing of biomolecules from AFM images. **Methods**, v. 193, p. 68–79, 2021

DEEPSEEK. DeepSeek-V3. Shenzhen: **DeepSeek**, 2025. Disponível em: <https://www.deepseek.com>. Acesso em: 10 jul. 2025.

EATON, P.; WEST, P. **Atomic Force Microscopy**. Oxford: Oxford University, 2010.

EUROPEAN COMMISSION. *Living guidelines on the responsible use of generative AI in research: ERA Forum stakeholders' document*. Brussels: **Directorate-General for Research and Innovation**, 2025. Disponível em: [https://research-andinnovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc\\_en?filename=ec\\_rtd\\_ai-guidelines.pdf](https://research-andinnovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf). Acesso em: 7 jul. 2025.

FOSTER, P.; FAWCETT, T. **Data science for business: What you need to know about data mining and data-analytic thinking**. Sebastopol: O'Reilly Media, 2013.

GOOGLE. Gemini 2.5 Pro. Mountain View: **Google**, 2025. Disponível em: <https://gemini.google.com>. Acesso em: 15 jul. 2025.

MODEL, M. L. **Bioinformatics Programming Using Python**, O'Reilly Media, 2009.

OPENAI. ChatGPT Plus (GPT-4). San Francisco: **OpenAI**, 2025. Disponível em: <https://chat.openai.com>. Acesso em: 08 jul. 2025.

VANDERPLAS, J. **Python data science handbook**. Sebastopol: O'Reilly Media, 2016. Disponível em: <https://jakevdp.github.io/PythonDataScienceHandbook/>. Acesso em: 02 jul. 2025.