DOI: 10.5281/zenodo.18903574

# Detection of Artifacts in AI-Generated Portraits Using Frequency and Texture Analysis (FFT, DCT, LBP/GLCM)

Peter Procházka[1]

**Abstract**
In this article, we present interim results from ongoing research aimed at identifying differences between real and AI-generated portraits through analysis in the frequency and texture domains. Three methods are examined: Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), and statistical texture descriptors based on Local Binary Patterns (LBP) and the Grey-Level Co-occurrence Matrix (GLCM). Using a controlled set of image data – a real portrait, its AI-generated clone, and a retouched version – we demonstrate the processing workflow, visualization, and interpretation of results. The aim of this work is to verify whether visually subtle differences between real and synthetic images correspond to measurable structural differences in alternative image representation domains. The article presents a methodological framework and example results of a pilot study; more extensive experiments on a larger dataset and the inclusion of additional analytical tools are planned in subsequent phases of the ongoing research.

**Keywords**
generated graphics, AI-generated images, diffusion models, FFT, DCT, LBP, GLCM, detection of AI-generated images

## 1 Introduction

The rapid development of generative models over the past few years has significantly influenced the fields of computer graphics and image processing. Modern diffusion models, such as Stable Diffusion (Rombach et al., 2022) or Imagen (Saharia et al., 2022), enable the generation of portraits with a high degree of realism that are often difficult for human observers to distinguish from real photographs. At the same time, this capability introduces new challenges in digital forensics, image content authentication, and information security, as confirmed by current research on synthetic media detection (Yang et al., 2019; Sha et al., 2023). Although the visual quality of AI-generated portraits is high, generative processes are based on statistical models trained on large datasets. This suggests that even with convincing visual output, images may retain subtle, systematic artifacts that are not directly observable in the pixel domain. Identifying such differences requires image analysis in alternative representations, particularly in the frequency and texture domains. The goal of our research is to investigate whether measurable differences exist between real, AI-generated, and classically retouched portraits, and to verify the extent to which these differences are consistent across multiple analytical methods. In this article, we present the first pilot phase of the research, focused on demonstrating the methodology and interpreting example results.

## 2 Theoretical Background

Generative models, including generative adversarial networks (GANs) and diffusion models, enable the synthesis of realistic portraits with a high level of detail, placing increased demands on

[1]Department of Applied Informatics, Faculty of Economic Informatics, Bratislava University of Economics and Business, Slovakia. Email: peter.prochazka@euba.sk

detection and forensic methods (Goodfellow et al., 2016; Ho et al., 2020; Dhariwal and Abbeel, 2021; Karras et al., 2021). In recent years, research attention has focused on identifying statistical and spectral inconsistencies that are typical for synthetic images but absent or present to a lesser extent in natural photographs (Barni et al., 2020; Sha et al., 2023).

Fourier Transform (FFT) and Discrete Cosine Transform (DCT) are among the fundamental tools of frequency image analysis. They allow the representation of an image in the frequency domain and the analysis of the energy distribution between low- and high-frequency components (Gonzalez and Woods, 2018; Jain, 1989). Several recent studies indicate that AI-generated images exhibit specific frequency signatures, including periodic artifacts and suppression of high-frequency content (Huang et al., 2023; Corvi et al., 2024).

Texture methods based on Local Binary Patterns (LBP) and Grey-Level Co-occurrence Matrix (GLCM) enable quantification of local statistical image properties, such as contrast, homogeneity, or microstructure variability (Ojala et al., 1996; Haralick et al., 1973). These features have proven to be a suitable complement to frequency analyses for detecting AI-generated images.

# 3 Methodology

The analysis was conducted on a trio of images: a real portrait, its AI-generated clone, and a retouched version of the real image. All images were preprocessed uniformly, including conversion to grayscale and resolution normalization. The analytical methods used include global frequency analysis via the Fast Fourier Transform (FFT), block-based frequency analysis via the Discrete Cosine Transform (DCT), and local texture analysis using Local Binary Patterns (LBP) and Grey-Level Co-occurrence Matrix (GLCM) descriptors. Primary experiments and visualization of interim results were carried out using a standalone web application implemented in HTML, CSS, and JavaScript. This application enables uploading a trio of images (real, AI-generated, retouched), performing basic preprocessing, and computing frequency and texture characteristics directly in the web browser. The web application served as the main tool for exploratory analysis, visual comparison, and rapid experimentation with method parameters. For independent verification of results, a separate validation application was developed in Python. Computations were performed using the libraries NumPy, SciPy, scikit-image, Pillow (PIL), Matplotlib, and Pandas. The user interface of this validation application was created using the Streamlit framework, which allows interactive execution of Python scripts and visualization of results through a web interface. In this case, Streamlit does not constitute a standalone web application in the HTML/JavaScript sense but serves solely as an interface to the Python implementation of analytical methods. The Python validation application implements the same analytical procedures as the web application, including block-based DCT (8×8), computation of FFT magnitude spectra, generation of LBP maps and histograms ("uniform" method), and computation of GLCM metrics (contrast, dissimilarity, homogeneity, energy, correlation). By comparing outputs from both environments, the consistency of observed trends was verified, and potential implementation deviations were eliminated. FFT analysis was additionally verified using the external tool ImageJ, ensuring reproducibility of frequency spectra across independent software solutions. In the next phase of the research, validation is planned to be extended to tools available in the MATLAB environment (or the Image Processing Toolbox), particularly for verifying DCT, LBP, and GLCM analyses once the relevant license becomes available. This multi-level validation approach, combining a web application, Python implementation, and external tools, increases the reliability and methodological transparency of the presented interim results.

# 4 Interim Results

The interim outputs of the analysis are based on a combination of visual evaluation and computations performed via the interactive web application and validation scripts in the Python environment on the trio of analyzed images. Observations range from qualitative assessment of visual details to quantitative evaluation of structural image properties across different representation domains. Identified differences between real, AI-generated, and retouched portraits are consistent across multiple analytical approaches.

## 4.1 Visual Comparison of Real, AI-Generated, and Retouched Portraits

Figure 1 shows the trio of analyzed portraits: the original image, its AI-generated clone, and a manually retouched version of the original. Even during direct visual comparison, noticeable differences are evident, suggesting different origins and different characteristics of image modifications.

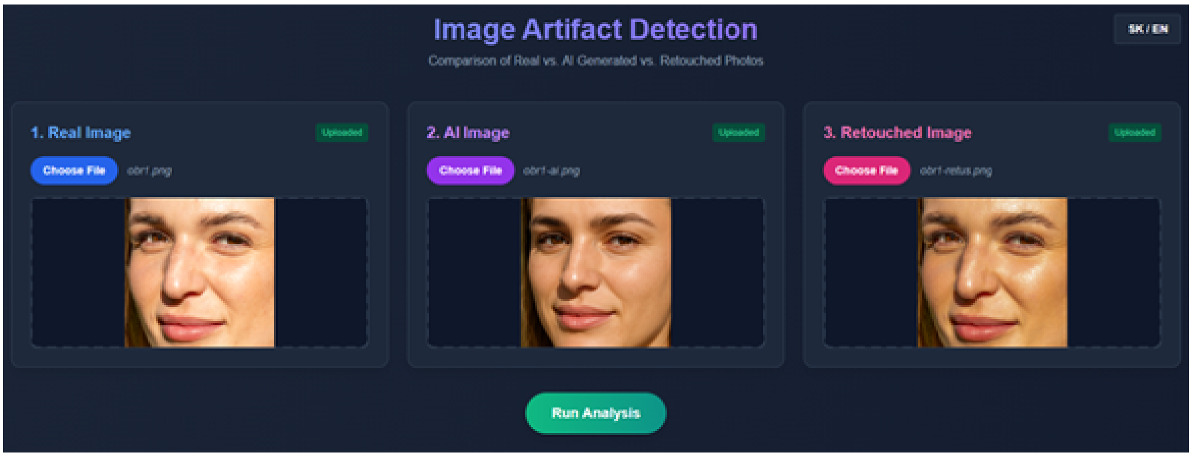


**Figure 1.** Input interface of the web application with three images

The real portrait exhibits pronounced skin microtexture, irregular pores, subtle local imperfections, and natural variations in light reflections. The skin surface appears statistically non-homogeneous, with fine details distributed randomly. This character is particularly observable in the cheek, nose, and eye areas, where fine wrinkles, pores, and natural sensor noise combine. In contrast, the AI-generated portrait appears visually smoother. Skin microstructure is significantly suppressed, pores are less discernible, and transitions between light and shadow are smoother. Light reflections exhibit a more regular, sometimes symmetrical character and appear less physically realistic. This effect is especially pronounced on the forehead and cheeks, where the image acquires an aesthetic, "cosmetically enhanced" appearance typical of generative models. The retouched portrait represents a transitional case. Local edits smooth selected skin areas, while preserving the underlying microstructure. Irregularities are weakened but not entirely removed, and light reflections retain a more natural character than in the AI-generated image. Visually, the retouched image appears more realistic than the synthetic one, although less raw than the original photograph. Differences are also evident in the eye region. The real portrait exhibits fine wrinkles with irregular geometry and realistic shading. The AI-generated image suppresses these details and shows greater symmetry between the left and right eyes. The retouched portrait again preserves natural asymmetry, albeit with slightly softened contrast. Similar tendencies can be observed in the eyebrows and lips, where the AI version appears more uniform and less random. These visual differences provide important qualitative context for the subsequent interpretation of frequency- and texture-analysis results, allowing these subjective observations to be quantified and verified in alternative image representation domains.

### 4.2 Results of DCT Analysis

Part of the presentation of DCT analysis results is visualization generated directly within the web application, displaying block-based DCT heatmaps for all three analyzed images side by side (Fig. 2). This form of presentation enables intuitive and immediate comparison of local frequency behavior between the real, AI-generated, and retouched portraits prior to quantitative coefficient evaluation. The web application uses a unified color scale and logarithmic value scaling, making differences in energy distribution between blocks visually discernible even to non-expert observers.


**Figure 2.** Discrete Cosine Transform

For the real image, even at the level of web visualization, a higher degree of local variability between individual (8×8) blocks is evident, whereas the AI-generated image exhibits a markedly more homogeneous pattern with suppression of high-frequency details. The retouched image again shows a transitional character, with some blocks frequency-simplified but without the regularity typical of synthetic generation. These visual observations from the web application directly correspond to the more detailed interpretation of DCT heatmaps and graphs presented in the following text and confirm that the identified differences are not the result of a single visualization but a stable phenomenon across the tools used. The Discrete Cosine Transform was applied in a block-wise manner at the (8×8)-pixel level, enabling analysis of local frequency-energy distributions and identification of differences between image types at the level of fine structures. Results are presented both as DCT heatmaps and as an aggregated graph of DCT coefficient distributions. In the real portrait, the DCT heatmap shows pronounced local variability across individual blocks. Energy is not concentrated solely in low-frequency components but is also distributed across mid- and higher-frequency components, corresponding to natural skin microtexture, fine wrinkles, and random image noise. This irregularity is visually readable as a subtly chaotic heatmap structure without repeating regular patterns. In contrast, the AI-generated portrait exhibits significant homogenization of the DCT heatmap. Energy is predominantly concentrated in low-frequency coefficients, while higher-frequency components are strongly suppressed. Individual (8×8) blocks exhibit similar frequency behavior, resulting in a more regular and visually smoother heatmap pattern. This effect points to systematic smoothing of fine details and reduction of local variability, consistent with the operating principles of diffusion-based generative models. Similar tendencies have been observed in recent studies analyzing frequency signatures of AI-generated images (Huang et al., 2023; Corvi et al., 2024). The retouched portrait represents a transitional case. The DCT heatmap shows partial weakening of high-frequency components, particularly in areas of targeted skin retouching, but retains greater local variability than the AI-generated image. Differences between blocks remain present and do not form regular, repeating structures. This character suggests that manual retouching affects local frequency properties of the image but does not alter its global statistical structure in the same way as AI generation. These observations are also confirmed by the graph of DCT coefficient distribution on a logarithmic scale. The curve of the real portrait shows a smooth energy decay from low to higher frequencies without sharp breaks, typical of natural image data. The AI-generated image exhibits steeper decay and lower values in higher-frequency bands, indicating the suppression of fine details. The retouched image appears closer to the original portrait in the graph, though with a slight shift toward lower frequencies due to smoothing. The combination of visual interpretation of DCT heatmaps and quantitative comparison of DCT coefficient profiles thus provides consistent evidence of differing frequency behavior between real, AI-generated, and retouched portraits and represents an important link between visual evaluation and global FFT analysis.

### 4.3 Results of FFT Analysis

Global frequency analysis was performed using the Fast Fourier Transform, with results presented as magnitude spectra on a logarithmic scale (Fig. 3). The visualization allows comparison of the energy distribution across low-, mid-, and high-frequency bands between the real, AI-generated, and retouched portraits.

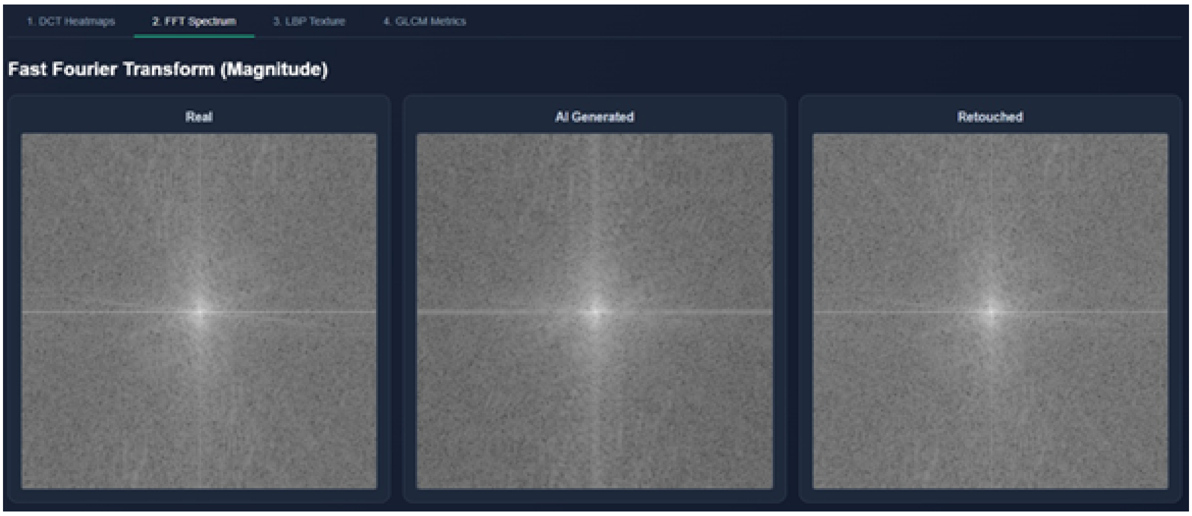


**Figure 3.** Fast Fourier Transform

The FFT spectrum of the real portrait exhibits a pronounced central component representing low-frequency image content, smoothly transitioning to higher frequencies without sharp discontinuities. In higher-frequency regions, a subtle chaotic energy dispersion is observed, corresponding to natural skin microtexture, fine wrinkles, and sensor noise. Spectrum irregularity indicates a high degree of local variability typical of photographic capture. The FFT spectrum of the AI-generated portrait differs primarily in the suppression of high-frequency components and the increased regularity of the spectral distribution. In mid-frequency regions, subtle linear and cross-like structures appear, suggesting periodic or quasi-periodic patterns arising during the generative process. This spectral character suggests lower randomness and greater statistical uniformity in the image, consistent with observations from recent studies on the frequency signatures of diffusion models (Huang et al., 2023; Corvi et al., 2024).

The retouched portrait exhibits an FFT spectrum located between both extremes. Compared to the real image, a slight reduction in high-frequency content is observed, corresponding to local skin smoothing. Unlike the AI-generated image, however, the spectrum does not exhibit pronounced regular structures or systematic repetitions. The overall spectral character remains predominantly chaotic and closer to that of a natural photograph. To ensure reproducibility and independent verification of results, FFT analysis was also performed using the ImageJ tool. FFT spectra obtained in ImageJ (Fig. 4) exhibit the same qualitative differences between analyzed images as spectra computed in the web and Python application environments. Agreement in spectral shape, degree of high-frequency suppression, and the presence of regular structures confirms that the observed differences are not artifacts of a specific implementation but rather represent inherent properties of the analyzed images.

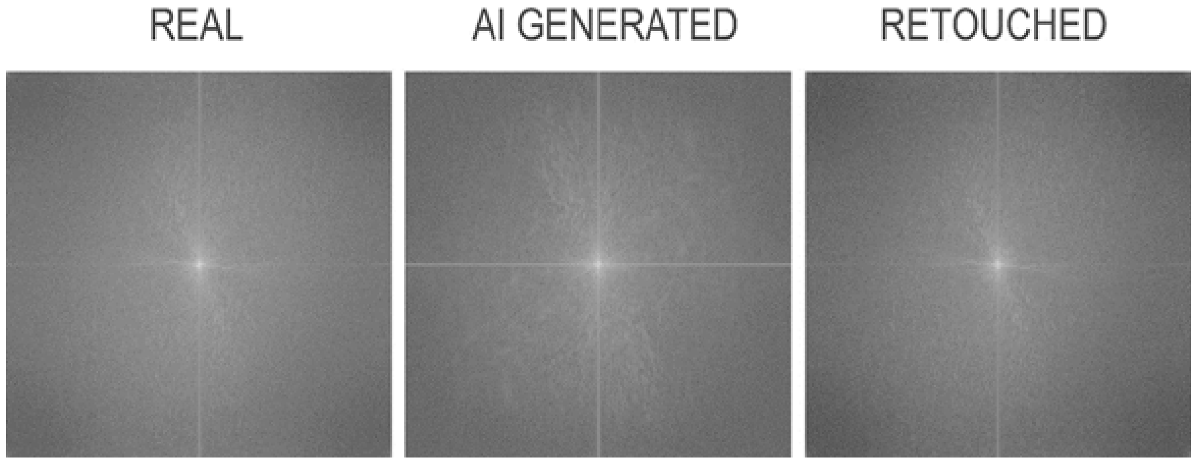


**Figure 4.** Fast Fourier Transform – ImageJ

*Source: Author's own processing* FFT analysis thus provides a robust global view of structural differences between real, AI-generated, and retouched portraits and suitably complements more local approaches based on DCT and texture descriptors.

## 4.4 Results of LBP Analysis

Local Binary Patterns (LBP) analysis was used to assess local skin microstructures that are often visually inconspicuous in the pixel domain but statistically significant. Figure 5 shows LBP maps obtained for the real, AI-generated, and retouched portraits via the web application. The real portrait exhibits a high degree of local variability in LBP patterns. The maps contain fine, irregular structures without dominant repeating patterns, corresponding to natural skin microtexture and the random nature of details. The histogram of LBP codes is relatively evenly distributed, without strong dominance of narrow intervals.

The AI-generated portrait, in contrast, exhibits pronounced regularity in LBP maps. Local patterns are more homogeneous, and the histogram shows a concentration of several dominant LBP codes. This phenomenon indicates reduced microstructural diversity and increased texture uniformity, consistent with the visual impression of smoothed skin and suppressed pores. Similar texture behavior in AI-generated images has been reported in other studies focused on local

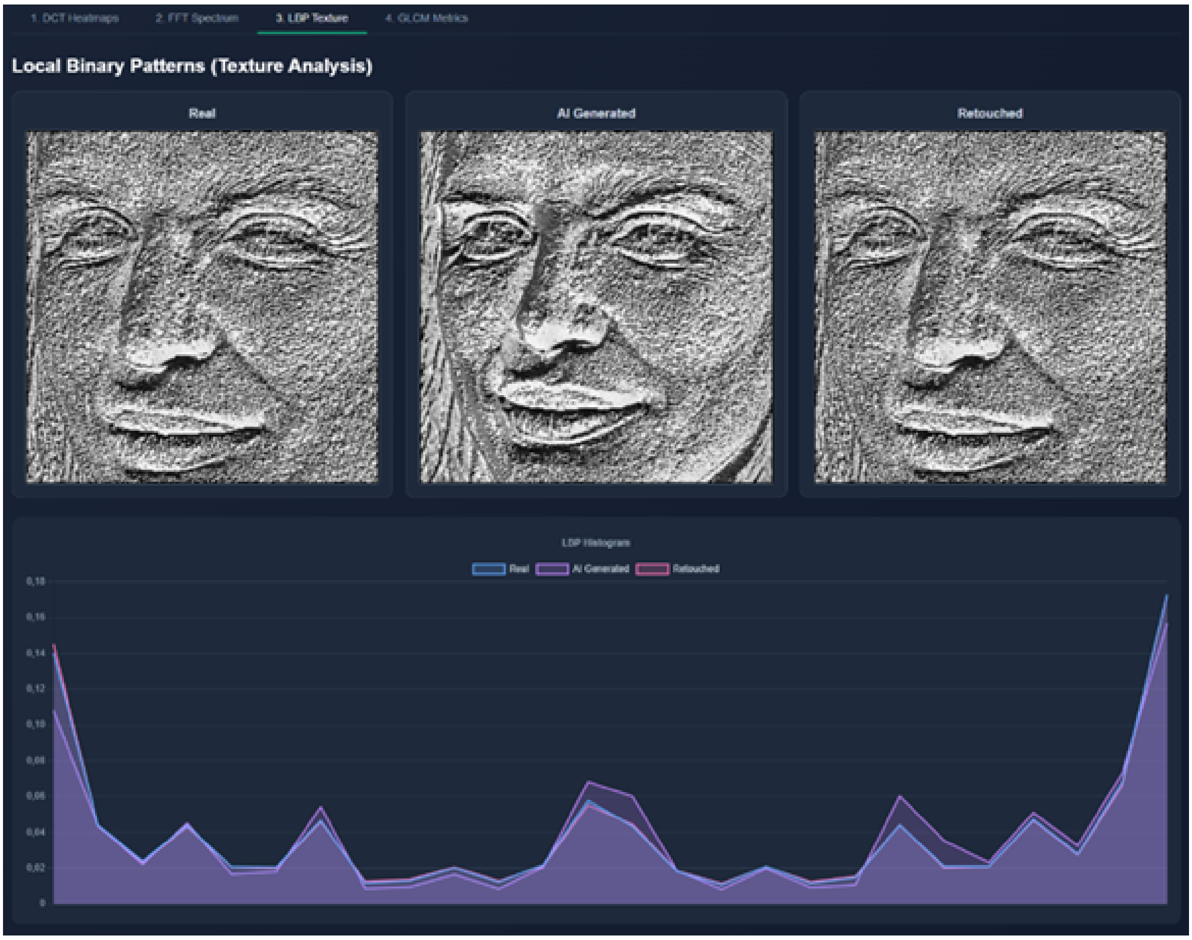


**Figure 5.** Local Binary Patterns

descriptors for distinguishing synthetic from real images, such as those using LBP and related texture features (Ojala et al., 1996; Sha et al., 2023).

The retouched portrait again represents a transitional state. LBP maps show partial suppression of fine details in retouched areas, but without pronounced regularization typical of AI-generated images. The LBP code histogram remains closer to that of the real portrait, albeit with a slight increase in the dominance of dominant patterns due to smoothing interventions. LBP analysis results thus confirm that local texture features can distinguish natural, synthetic, and manually edited images in a manner consistent with visual evaluation and frequency analyses.

## 4.5 GLCM Analysis

To complement local texture analysis, statistical features based on the Grey-Level Co-occurrence Matrix (GLCM) were computed. Quantitative results are summarized in a table (Fig. 6) and visualized using a comparative graph generated by the web application. The real portrait shows higher contrast and dissimilarity, reflecting the complex, non-homogeneous structure of natural skin. At the same time, it exhibits lower homogeneity, a characteristic of images with a high degree of local intensity variation.

The AI-generated portrait exhibits significantly lower contrast and dissimilarity, along with increased homogeneity. This profile indicates the suppression of fine local differences and the presence of smoother, more statistically uniform textures. Such behavior of GLCM metrics is consistent with observations reported in the literature, which indicate that AI-generated images exhibit a higher degree of local texture regularity and homogeneity than natural photographs (Haralick et al., 1973; Sha et al., 2023). The retouched portrait again occupies an intermediate

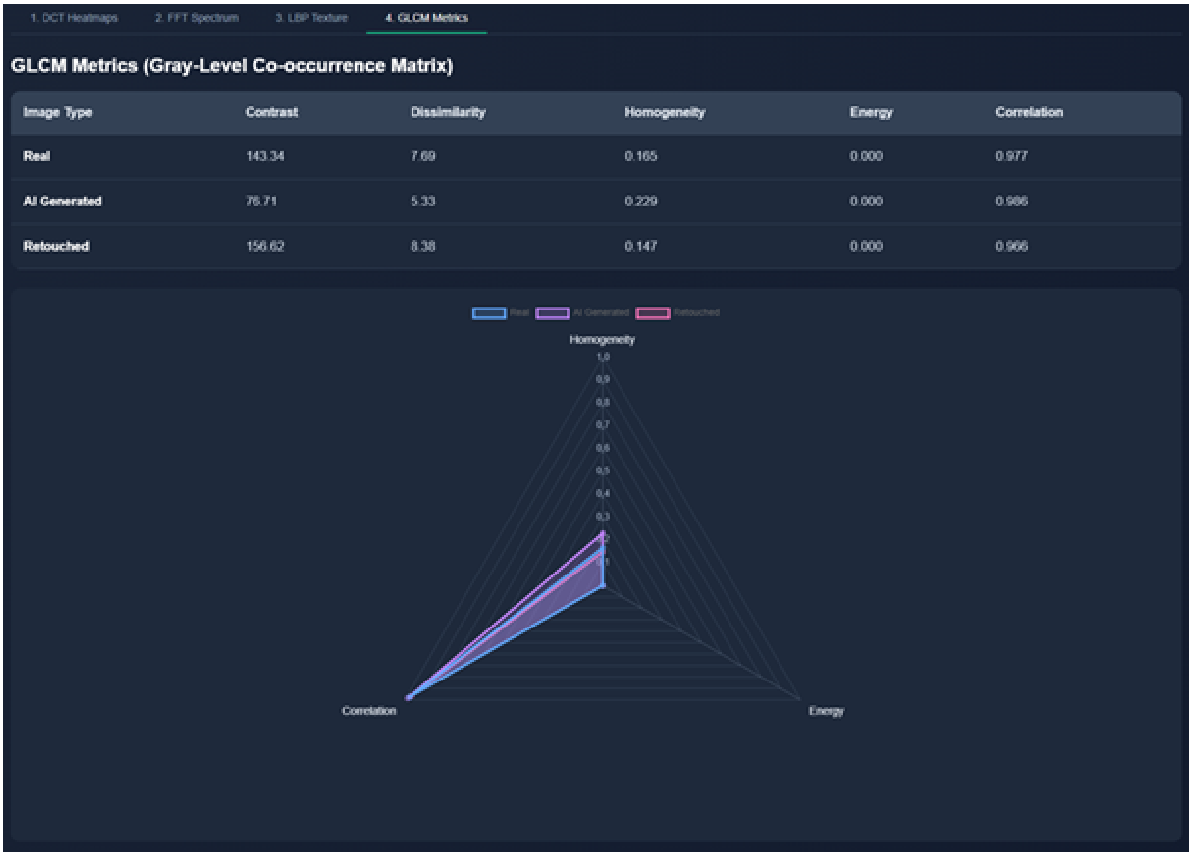


**Figure 6.** GLCM Metrics

position between the real and AI-generated images. The contrast and homogeneity values indicate a partial influence of smoothing operations, without a pronounced change in the image's global statistical structure. GLCM metrics thus confirm that classical retouching does not alter local statistical properties of the image to the same extent as generative models. In combination with LBP analysis, GLCM features provide a robust quantitative description of local textures that suitably complements global frequency methods based on FFT and DCT, and contribute to clearer differentiation of the origin of analyzed portraits.

# 5 Discussion

Interpretation of the achieved results highlights several aspects related to the detection of AI-generated portraits and the importance of frequency and texture features in distinguishing them from real photographs. First and foremost, visually subtle differences that are often difficult for human observers to identify manifest consistently and reproducibly in transformed domains. Frequency analyses based on DCT and FFT clearly indicate that AI-generated portraits tend to suppress high-frequency content and exhibit greater regularity in spectral energy distribution. This phenomenon can be interpreted as a consequence of the optimization mechanisms of generative models, which favor global visual consistency and suppress random noise and microdetails typical of real images. Similar frequency behavior has been observed in other studies analyzing AI-generated images, suggesting that this is a more general phenomenon rather than a property of a single tested example. Texture analyses using LBP and GLCM provide a complementary view of local structural image properties. Results show that AI-generated images exhibit reduced variability in local patterns and greater homogeneity, whereas real photographs exhibit higher contrast and statistical non-homogeneity. Retouched images occupy an intermediate position across most metrics, confirming that manual retouching affects local details but does not alter the fundamental statistical structure of the image as strongly as AI generation. An important finding is also the consistency of results across different implementations and tools. Agreement between the outputs of the web application, the Python validation application, and the external ImageJ tool increases the credibility of the observed trends and suggests that the identified differences are not artifacts of a particular software solution. This aspect is crucial, especially for the practical deployment of forensic tools, where reproducibility and implementation independence are essential. At the same time, it is necessary to emphasize the limitations of the presented results. The analysis was conducted on a limited number of images and serves primarily as a pilot study demonstrating the methodological approach. Generalizing the findings to a broader range of generative models, various scene types, and different levels of image quality requires more extensive experiments across larger datasets. Nevertheless, the results indicate that a combination of frequency and texture features represents a promising basis for robust detection of AI-generated images.

# 6 Conclusion

In this work, we have introduced a methodological framework and interim results from a pilot study analyzing differences between real, AI-generated, and manually retouched portraits. By combining visual evaluation, frequency transformations (FFT, DCT), and local texture descriptors (LBP, GLCM), it was possible to identify consistent structural differences that are not always apparent in the pixel domain but manifest in alternative image representations. The results suggest that AI-generated portraits tend to exhibit reduced high-frequency components, more regular frequency spectra, and greater local texture homogeneity compared to real photographs. Retouched images occupy an intermediate position across most analyzed metrics, confirming that classical image editing affects local details but does not alter the global statistical structure of the image to the same extent as AI generation. An important aspect of the work is the multi-level validation approach, in which results were compared across an independent web application, a Python implementation for validation, and the external ImageJ tool. Agreement of observed trends across different environments supports the robustness of the presented findings and indicates their potential applicability in digital forensics and the detection of synthetic media. The presented results are a pilot study and primarily serve to demonstrate the methodology and its interpretative potential. In the next phase of the research, the analysis will be expanded to a larger, more diverse dataset, including various generative models, scene types, and image quality levels. At the same time, additional analytical approaches and machine learning models are planned to automate the evaluation of identified features. These steps will be the subject of a subsequent publication focused on detailed quantitative evaluation and comparison of methods.

# References

Barni, Mauro, K. Kallas, E. Nowroozi, and Benedetta Tondi (2020). “Adversarial attacks against CNN-based image forgery detectors: A comparative study”. In: *EURASIP Journal on Information Security*.

Corvi, R., Y. Huang, and Luisa Verdoliva (2024). “Analyzing frequency-domain fingerprints of diffusion-based image generators”. In: *IEEE Transactions on Information Forensics and Security*.

Dhariwal, Prafulla and Pieter Abbeel (2021). “Diffusion models beat GANs on image synthesis”. In: *Advances in Neural Information Processing Systems*.

Gonzalez, Rafael C. and Richard E. Woods (2018). *Digital Image Processing*. 4th ed. Pearson.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. MIT Press.

Haralick, Robert M., K. Shanmugam, and Its’Hak Dinstein (1973). “Textural features for image classification”. In: *IEEE Transactions on Systems, Man, and Cybernetics* 3.6, pp. 610–621.

Ho, Jonathan, Ajay Jain, and Pieter Abbeel (2020). “Denoising diffusion probabilistic models”. In: *Advances in Neural Information Processing Systems*.

Huang, Y., X. Yang, and Siwei Lyu (2023). “AI-generated image detection using statistical and spectral inconsistencies”. In: *Proceedings of the IEEE International Conference on Image Processing*.

Jain, Anil K. (1989). *Fundamentals of Digital Image Processing*. Prentice Hall.

Karras, Tero, Miika Aittala, Samuli Laine, Erik H"ark"onen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila (2021). “Alias-free generative adversarial networks”. In: *Advances in Neural Information Processing Systems*.

Ojala, Timo, Matti Pietik"ainen, and David Harwood (1996). “A comparative study of texture measures with classification based on featured distributions”. In: *Pattern Recognition* 29.1, pp. 51–59.

Rombach, Robin, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Bj"orn Ommer (2022). “High-resolution image synthesis with latent diffusion models”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Saharia, Chitwan, William Chan, Han Chang, et al. (2022). “Imagen: Photorealistic text-to-image diffusion models”. In: *Proceedings of the International Conference on Machine Learning*.

Sha, Z., H. Wu, and X. Chen (2023). “A survey on deepfake detection: Challenges and future trends”. In: *ACM Computing Surveys*.

Yang, X., Y. Li, and Siwei Lyu (2019). “Exposing deep fakes using inconsistent head poses”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.