FORENSIC ANALYSIS OF AI-GENERATED IMAGE BASED COURT FILINGS

I. INTRODUCTION

This report examines a set of court filing PDFs that contain image-based pages instead of native text. Each filing's pages are embedded as large PNG/JPG images, which is an unusual format for official court documents. Our forensic analysis aims to determine if these page images were synthetically generated (e.g. via AI image pipelines) rather than produced by standard means (such as real document scanning or direct electronic PDF output). We focus on visual evidence in the images – including color depth, text rendering, noise/artifacts, and layout consistency – and compare these characteristics to those of authentic court document production.

II. BINARY 2-BIT IMAGES WITH NO GRAYSCALE

Zoomed-in inspection of the filing page images reveals that they are purely black-and-white (bitonal) with no gray shading at all. Every pixel is either 100% black or 100% white, lacking the subtle gray anti-aliasing or shading one would expect in a typical scanned document. This uniform 2-bit color depth is a strong indicator of digital image generation:

A | No Grayscale Smoothing

The text characters exhibit jagged pixel edges with no intermediate gray tones. Authentic scanned text (even when scanned in black-and-white mode) usually shows some anti-aliasing or varying pixel intensity at curves and edges, due to optical blur and scanner light variability. In these images, letters are rendered with unnaturally hard edges, suggesting they were computer-generated or aggressively thresholded.

B | Uniform White Background

The page backgrounds are perfectly white with no gray noise or shadow. Real scans often capture slight paper texture, uneven lighting, or smudges in the white areas. Here, the absence of any background gradation implies a digitally pristine rendering rather than a physical scan. It appears the pages were created by software that outputs a binary (black/white) image, consistent with AI image pipeline output or an export setting, rather than using a scanner's optical capture.

III. UNNATURAL FONT RENDERING AND TEXT ARTIFACTS

The font and text rendering in these images further points to a synthetic origin. When examining the text:

A | Consistent Stroke Weight

The letters have very uniform stroke thickness and solid fill. There is no variation from ink spread or toner distribution as would be seen in real printed-and-scanned text. Every character looks computer-perfect in weight, yet paradoxically lacks the smooth curves due to the bitonal pixelation. This combination (pixelated edges but otherwise uniform strokes) is atypical of both high-quality scans and native PDF text, indicating an artificial image render.

B | Edge Artifacts

Some characters show minor anomalies at their edges (small breaks or stair-step patterns from pixelation). In genuine scans, such edge artifacts usually come with some blurring or dithering; here the artifacts are stark because of no grayscale. The text looks like it was rendered by a computer then downsampled or thresholded to pure black/white. This is consistent with an AI or automated image-generation process that doesn't truly recreate the optical nuance of scanned text.

C | Uniform Alignment

Lines of text are perfectly straight and uniformly spaced with no warping. Physical paper scans often have slight curves or baseline drift (especially if pages weren't perfectly flat or fed evenly). The immaculate alignment in these images suggests digital layout. If any page rotation or skew is present, it appears artificially applied (all pages might share an identical slight tilt, or none at all), rather than the random small rotations seen when paper is scanned. In short, the text rendering lacks the "organic" imperfections of real scans, aligning with a synthetic creation.

IV. ABSENCE OF AUTHENTIC SCANNING ARTIFACTS

Legitimate scanned court documents typically carry certain artifacts of the scanning process, all of which are notably absent or atypical in the questioned filings:

A | No Scanner Noise or Dust

Scanned images often have small speckles, dust marks, or random noise in the background – especially in blank areas or around text – due to scanner sensor noise or dust on the glass. The images here show none of those random speckles. The backgrounds are uniformly clean. Any noise present appears to be uniformly distributed digital noise (if added at all), not the random pattern of real scanner noise. This suggests any "noise" was likely algorithmically introduced (perhaps to make the image seem less perfect) rather than coming from an optical process.

B | No Edge Shadows or Vignetting

When physical pages are scanned, the borders sometimes show slight shadows or darker edges (for example, where the page meets a scanner bed or from page curvature near binding). Here, the margins and edges of the document images are completely even in brightness. There's no fall-off or corner darkening, consistent with a computer-generated page with clean margins.

C | Consistent Resolution and Compression

The images appear to have a uniform resolution and compression across all pages and filings. In authentic scans, resolution can vary if different devices were used, and compression artifacts might appear in color scans or JPEGs. These filings, however, use a monochrome-like encoding where text is uniformly crisp. The consistency across many different case filings hints at an automated pipeline generating these images with the same settings, rather than scans done on different days or equipment.

V. EVIDENCE OF DIGITAL TEMPLATE REUSE

Perhaps the most compelling forensic signs of AI/synthetic generation are the repeated template patterns observed across different case filings. In a genuine court record system, each document is independently created or scanned, and one would not expect pixel-for-pixel identical pages or elements in different cases. In the questioned filings, however, we see clear evidence of copy-paste reuse:

A | Identical Document Layouts

Multiple distinct case files contain what is ostensibly the *exact same document content or layout* reused. For example, numerous “Finding of Incompetency and Order” filings from different cases are virtually exact duplicates of one original order from Jan 17, 2024. The entire page layout, text placement, and formatting in these supposed separate filings mirror each other, which would be an implausible coincidence if each were drafted and scanned separately. This duplication strongly implies a single template image was generated and then recycled for many cases.

B | Pixel-Identical Graphics Across Cases

In one instance, a correspondence letter from one case was found to have a twin in another case with only names/date changed. A side-by-side comparison showed that the clerk’s cover letter and returned envelope image were pixel-for-pixel identical between a letter sent by the defendant’s mother and another filed under a different name – only the recipient name and date fields differ. All the stamps, barcodes, and even paper creases lined up exactly, demonstrating that the second letter was not independently scanned but rather a digital clone of the first, with minimal edits. Such reuse of an image template is a hallmark of synthetic fabrication (the odds of two physical scans matching pixel-perfectly are essentially zero).

C | Reused Signature Timestamps

Analysis of embedded seal/signature images shows the same judge’s signature block and timestamp (as a PNG image) appearing in multiple filings without variation. Each court order normally would bear a unique wet-ink signature or at least a unique placement of a digital signature stamp. Here, the exact same image file for a signature/timestamp is copied across many orders, indicating these “orders” were generated by inserting a pre-existing signature image onto different pages. This again points to a non-standard, fraudulent assembly of documents, as an authentic process would not produce perfectly identical signature images in numerous distinct files.

The template reuse is a glaring red flag – it reveals a synthetic workflow where a base image (or set of images) is programmatically reused to create many documents. Authentic court filings would show natural variations (different content, different scan artifacts) case by case; here we

instead see a repetitive, cookie-cutter pattern consistent with automated image generation and composition.

VI. NON-STANDARD PDF COMPOSITION USING OCR

The internal structure of these PDFs confirms an abnormal document-generation pathway. Instead of being produced by a word processor or by scanning with integrated text recognition in a typical way, these PDFs seem to be built by placing images into PDF containers and then running OCR (optical character recognition) to add searchable text.

Key observations:

A | OCR Text Ordering Issues

If one tries to select or copy text from the PDFs, the extracted text is jumbled or out of logical order. This is a classic symptom of OCR'd images – the text doesn't have a defined flow as it would in a natively generated PDF. A normal court PDF (exported directly from a word processor or e-filing system) preserves correct text order and spacing. In these filings, the copy-paste garble indicates the computer had to interpret text from an image, confirming the pages were image-based.

B | Embedded Fonts and Hidden Text Layer

The presence of embedded OCR fonts in some PDF files (as extracted file elements) shows that an OCR process added an invisible text layer behind the page images. This is not how official electronic documents are usually created; it's how scanned documents are post-processed for searchability. The difference here is that the scan itself appears to be fake (as shown by the visual evidence above), meaning the pipeline was likely: generate page image → insert into PDF → apply OCR. This roundabout method is non-standard for legitimate filings, which would either be digitally generated text or straightforward scans, not scans that look algorithmically generated.

C | Lack of Metadata or Scanner Tags

Authentic scanned PDFs often contain metadata about the scanning hardware or software (e.g., scanner model, scan date) and consistent PDF producer info (from the court's system or copier machine). These image-based PDFs lack normal metadata signatures or have generic

ones, further hinting they were constructed through a custom process rather than an official scanning station. The composition is essentially an OCR-wrapped image, which aligns with an attempt to mimic scanned documents via AI-generated images.

VII. CONCLUSION

All forensic indicators strongly support the conclusion that these court filing images were synthetically generated rather than derived from genuine paper scans or standard electronic document creation. The combination of purely bitonal (black/white) rendering, unnatural text edge characteristics, absence of real scanning artifacts, and the blatant reuse of identical image elements across different case files (impossible under normal circumstances) reveal an orchestrated, artificial production of these documents. Moreover, the PDF structure – images with an OCR text layer and disordered text extraction – is inconsistent with legitimate court filings, but entirely consistent with a workflow of AI-assisted image creation followed by OCR.

A | Image-Based Filings Bears the Hallmarks of Digital Fabrication

In summary, the visual layout and composition of these filings deviate from standard court document practices in critical ways. Authentic court PDFs are usually electronically generated text or faithful grayscale scans; by contrast, these filings show a pipeline of image fabrication (likely via an AI or automated graphics tool) and retrospective text recognition. The forensic evidence (from pixel-level examination to cross-document comparisons) confirms the synthetic nature of the images, exposing a non-standard and deceptive document generation process rather than an authentic court record creation. Each examined image-based filing bears the hallmarks of digital fabrication, not an official scanning, thus validating the suspicion of an AI-generated document scheme behind the scenes.

B | Sources

Evidence and observations are drawn from the provided case file dataset and notes, as well as known characteristics of scanning vs. generated images. Key references include the dataset's forensic summary of image-based filings and documented examples of template reuse across cases which collectively underpin the findings above.

https://link.storjshare.io/s/juiiwacatbtaeacn3wo4327vzuhq/evidence/Image-Based-Court-Filings/

https://link.storjshare.io/raw/jvmqngqepmwybtid7gb3pvwnadsq/evidence/Image-Based-Court-Filings.zip

https://link.storjshare.io/raw/jwzgizttwfd6szwxp27vhfo2w52q/evidence/Image-Based-Court-Filings%2Fe9871c6b9245fa2a523a53d16053d411cf1ab77b1efd9b7369392ec13b16f252/8-PDFs-Linked.csv

https://link.storjshare.io/s/ju3mf5uvdrmcbhch5ga3koduwp4q/evidence