Spotting Deepfakes 101

Overview

Image and document manipulation is rising fast, and the days of detecting these issues with the naked eye are quickly fading. Fraudsters easily alter images, metadata, and copy internet-sourced images into their own claims.

Digital tampering and creating deep fakes are growing exponentially. Advanced forensics are essential for layering data into an effective, modern strategy to shield an organization from fraud.

The use of photos, documents, and other digital media in insurance interactions has improved claims and underwriting outcomes in a variety of ways—from more efficient documentation of claim damages to better understanding of underwriting risk—while removing burdens on insurance professionals. Claims parties handle more evidence documentation, and more communication occurring digitally. The process moves faster with less human involvement, but it also creates space for bad actors to commit fraud.

Technological advances are making it even easier to do so; the availability and quality of generative AI (GenAI) is increasing rapidly. What started as nifty toys that could swap faces in a photo has involved into sophisticated tools that can create completely false images, videos, and documents. Furthermore, these tools are easily accessible and often free to use.

Recognizing and understanding this emerging threat early is critical to protecting against it. Effective risk mitigation will require automated detection systems, human knowledge and expertise, and strong evidence controls. With this guide, we aim to provide those on the frontline with some tips, tricks, and best practices to approach this challenge.

Quick Reference: A Claims Professional's Guide to Spotting Deepfakes

Exterior:
- Mismatched, inconsistent, lacking, or excessive windows, roofing, trim, or lighting on a house
- Car doors open the wrong way; headlights are misplaced and/or disproportionate
- Indistinct logos on cars or other branded items
- "Gibberish" on license plates, road signs, etc.

Interior:
- Vehicles: missing or out-of-place seats, headrests, controls, dials, etc.
- Structures: mismatched, out-of-place, or incorrect windows, doors, walls, and furniture
- Windows: objects outside don't match real-world surroundings

Background:
- Trees, fences, or other intricate items merge
- Broken markings on roads or parking lots
- Warped, broken, wavy, or curved objects
- Clarity is inconsistent with foreground
- Watermarks and cropping:
- Visible watermarks, especially around edges if someone tried to crop them out
- Irregular aspect ratios (other than 16:9, 4:3, 1:1)

Damage:
- Hail damage on just one part of a roof
- Scratches appear to "float" on a vehicle surface
- Damage location doesn't match impact point
- Missing debris in accident scene photos
- Damage skips part of a car, such as the wheel well and tire

People and animals:
- Too many or too few fingers, arms, or legs
- Stray heads, arms, legs, or partial bodies
- Disproportionate body parts
- Joints bend unnaturally (especially for animals)

Object size:
- Subject too big or small for background objects
- Damaged parts appear out of proportion to the rest of the vehicle
- Tree or building dimensions look mismatched

Colors:
- Oversaturated
- Unnaturally even tone or texture, especially in dark or light areas
- Overall appearance resembles an art print or illustration rather than a photograph

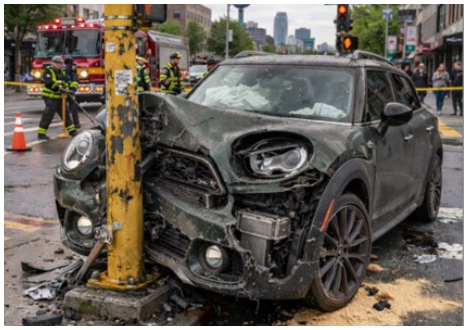
Shadows and light:
- Broken, misaligned, out-of-place, or missing shadows
- Inconsistent light angles within or across related images

Fire damage:
- Contrast is overly dark
- Smoke patterns appear "off" or missing
- Damage appears to be too localized

Unexpected repetition:
- Identical or similar shapes are repeated

Practical Examples

Applying the previous concepts, what stands out to create suspicion about these images?



Here are some things to keep in mind:

Avoid using free online tools, such as metadata extractors or conversational AI platforms. Stick to tools approved by and available within your organization.

Don't allow storage or virus screening systems to remove metadata or compress files. These actions may save storage space, but they also limit forensic abilities.

Don't combine multiple images into a PDF before sending for forensics. Doing so removes the metadata and reduces the image quality.

Keep media in its original state. Editing files, even something as small as changing a file name or clicking "Save" on a PDF, can have a major impact on the ability to provide accurate forensics.

When integrating a forensic tool and automating media submission, place the integration as early in your media pipeline as possible. The fewer systems and processes a media file passes through, the less likely it is to be compressed or altered.

Verisk