Spot the Fake

Your Guide to Detecting AI-Generated & Misleading Content • Ages 10 to 16

Student Name: _____ **Date:** _____

Class / Group: _____ **Teacher:** _____

The Age of Synthetic Media

AI technology has advanced to the point where it can create **fake images, videos, audio, and text** that are nearly indistinguishable from real content. These are called **synthetic media** or, when used to deceive, **deepfakes**.

The scale of the problem:

- Deepfake videos online increased by **550%** between 2019 and 2023 (Deeptrace/Sensity)
- AI image generators like DALL-E, Midjourney, and Stable Diffusion create **millions of synthetic images daily**
- AI voice cloning can replicate someone's voice from just **3 seconds of audio**
- In 2024, AI-generated robocalls impersonating a US President reached thousands of voters
- **96%** of deepfake videos are non-consensual, often targeting women and girls (Sensity AI report)

This worksheet will teach you **how to detect fake content** across every media type (images, videos, text, and audio) with practical techniques you can use every day.

Types of Fake Content

Type	What It Is	How It's Made	Danger Level
AI Images	Completely fabricated photorealistic images of people, places, or events that never existed	DALL-E 3, Midjourney V6, Stable Diffusion XL: text-to-image AI models	Very High: extremely convincing
Deepfake Video	Videos showing real people saying or doing things they never did	Face-swapping AI, lip-sync AI, full-body synthesis models	Very High: can manipulate public opinion
AI Text	Articles, essays, social posts, reviews written entirely by AI	ChatGPT, Claude, Gemini, LLaMA: Large Language Models	High: very hard to detect
Voice Clones	Fake audio of real people's voices saying things they never said	ElevenLabs, Resemble.AI, VALL-E: voice synthesis AI	Very High: used in scams
Manipulated Media	Real content edited to change its meaning (cropped, recontextualised, selectively edited)	Standard editing tools (Photoshop, video editors) + AI enhancement	High: mixes real and fake

Section 1: Spotting Fake Images

AI image generators have become extraordinarily sophisticated. In 2023, an AI-generated photo of “Pope Francis wearing a white Balenciaga puffer jacket” went viral, and millions believed it was real. Here’s how to detect AI-generated images:

The 10-Point Image Detection Checklist:

Hands and Fingers

AI consistently struggles with hands. Look for: **wrong number of fingers** (4, 6, or 7), fingers that are **different lengths than normal**, fingers that **merge into each other**, extra joints, or hands in impossible positions. This is improving rapidly (Midjourney V6 handles hands better) but remains one of the most reliable tells.

■ Count the fingers on every visible hand. Are there exactly 5 per hand? Do they look natural?

Eyes

Check for: **different-shaped pupils** (one round, one oval), **inconsistent reflections** (the reflected light source should be in the same position in both eyes), **unusual iris patterns**, or eyes that seem to be looking in slightly different directions. The “pupil reflection test” is particularly effective. In real photos, both eyes reflect the same light source from the same angle.

■ Zoom into the eyes. Are the reflections/highlights consistent between both eyes?

Teeth

AI often generates teeth that are **too uniform** (all exactly the same size and shape), teeth that **blur into each other**, or an **incorrect number of teeth**. Real teeth have natural variations in size, alignment, and colour.

■ If the person is smiling, do the teeth look natural? Too perfect? Blurry where they meet the gums?

Text and Numbers

AI is notoriously bad at generating readable text. Look for: **garbled or nonsensical text** on signs, labels, or screens in the image. Letters may be in the wrong order, misspelled, or in a font that doesn’t exist. Numbers on clocks or displays often don’t make sense.

■ Can you read any text in the image? Does it make sense? Check signs, logos, and labels.

Accessories and Jewellery

Earrings that don’t match (different design on each ear), glasses frames that are **asymmetrical** or that seem to merge with the face, necklaces that **pass through skin**, and watch straps that don’t connect properly.

■ Check if accessories are symmetrical and physically possible. Do earrings match?

Background Details

While AI focuses on the main subject, backgrounds often contain **impossible architecture** (buildings that defy physics), **melting objects**, trees or objects that don’t follow consistent perspective, and **repeating patterns** that look unnatural. Edges where the subject meets the background may be blurry or inconsistent.

■ Look carefully at the background. Do buildings have consistent geometry? Are objects well-defined?

Lighting and Shadows

In real photos, all shadows point in a **consistent direction** from the light source(s). AI images often have shadows that **point in different directions**, shadows that are **missing** where they should exist, or lighting on the subject that doesn’t match the background lighting.

■ Where is the light coming from? Do ALL shadows point in the same direction?

Clothing and Fabric

Look for clothing that has **impossible folds or wrinkles**, buttons or zips that disappear into the fabric, patterns (stripes, checks) that don't continue consistently across seams and folds, and collars or necklines that seem to merge with skin.

■ Do clothing patterns continue correctly across folds? Do buttons and fasteners look real?

Hair

AI hair often looks **too smooth and perfect**, like plastic rather than real hair. Individual strands may merge into a smooth surface. The hairline (where hair meets forehead/skin) is often **unnaturally perfect** or, conversely, strangely blurry.

■ Does the hair look like real hair with individual strands, or more like a smooth solid mass?

Overall "Feel"

Sometimes an AI image just feels "off" even if you can't pinpoint why. This is often due to the **uncanny valley effect**, where everything looks almost-but-not-quite real, creating an uncomfortable feeling. Trust your instincts, then investigate further.

■ Does the image give you an uneasy feeling? Does it look too perfect, too polished, too clean?

Tools You Can Use:

- **Google Reverse Image Search**: right-click any image and select "Search image" to see if it's been used elsewhere
- **TinEye.com**: specialised reverse image search engine
- **FotoForensics.com**: analyses image metadata and compression artefacts
- **AI or Not (aiornot.com)**: AI-powered detector that analyses if an image is AI-generated
- **Hive Moderation**: free AI content detection tool

Section 2: Spotting Deepfake Videos

Deepfake videos use AI to swap faces, manipulate lip movements, or generate entirely synthetic video. In 2023, a deepfake video call was used to trick a Hong Kong company into transferring **$25 million**. The scammers impersonated the company's CFO in a live video call.

The 8-Point Deepfake Video Detection Guide:

■ 1. Face Boundary / Edge Check

*Look at where the face meets the hairline, ears, and neck. In deepfakes, there's often a **subtle line or colour shift** where the AI-generated face was blended onto the original. The skin colour or texture may change slightly at the jawline.*

■ 2. Blinking Patterns

*Early deepfakes had a major flaw: the AI subjects **rarely blinked** because training data (photos) mostly show people with eyes open. This has improved, but blinking in deepfakes may still look **slightly unnatural**: too regular, too quick, or missing entirely.*

■ 3. Lip Sync Accuracy

*In deepfakes where speech has been altered, the lip movements may not perfectly match the audio. Watch carefully for **moments where the mouth shape doesn't match the sound** being produced. The audio may also have subtle **robotic qualities** or unnatural pauses.*

■ 4. Head and Body Coherence

*The face is usually the only AI-generated part. When the person turns their head significantly, the deepfake face may **glitch, warp, or lose tracking**. The face may also appear to **float slightly** relative to the body, especially during rapid movement.*

■ 5. Skin Texture and Pores

*AI-generated skin often looks **too smooth**, as if a filter has been applied. Real skin has pores, small imperfections, fine lines, and uneven texture that varies across the face. Deepfake skin tends to be uniformly smooth.*

■ 6. Lighting Consistency

If the lighting on the face doesn't match the lighting on the body, background, or other people in the scene, the face may be synthetic. Shadows should be consistent throughout the frame.

■ 7. Audio-Visual Mismatch

*Pay attention to whether the voice sounds natural for the person. AI voice cloning may have subtle artefacts: **metallic quality**, unusual breathing patterns, missing natural voice variations (we all slightly change pitch, pace, and volume when speaking naturally).*

■ 8. Temporal Coherence

*Pause the video frequently and scrub through frame-by-frame (use the . and , keys on YouTube). Deepfakes often have **individual frames that look glitchy**, a brief flash of distortion that's invisible at normal speed but obvious when paused.*

Section 3: Spotting AI-Generated Text

AI-generated text is the hardest to detect because Large Language Models (LLMs) like ChatGPT produce grammatically perfect, coherent text. However, there are patterns and tells that can help:

■ 1. Overly Perfect and Polished

*AI text tends to be **suspiciously well-structured** with perfect grammar, balanced paragraph lengths, and clear topic sentences. Real human writing is usually messier; we use fragments, start sentences with "But" or "And", make small errors, and have irregular rhythm.*

■ 2. Hedging Language Overuse

AI models are trained to be cautious, so they heavily use phrases like "It's important to note that...", "However, it should be considered...", "While there are many perspectives...". An unusual density of these hedging phrases suggests AI.

■ 3. Lack of Personal Voice

*AI text lacks a **distinctive personal voice**. It doesn't use slang, regional expressions, personal anecdotes, or unique turns of phrase. It reads like a well-educated, slightly bland encyclopedia, correct but characterless.*

■ 4. Repetitive Sentence Structures

AI often follows predictable patterns: "Topic sentence. Explanation. Example. Transition to next point." Real writing varies sentence length and structure much more naturally.

■ 5. Confident but Vague Factual Claims

AI may state "studies have shown" or "experts agree" without naming specific studies or experts. If factual claims can't be verified through the provided references, they may be AI hallucinations , which are fabricated "facts" stated with complete confidence.

■ 6. Favourite AI Words

*Current LLMs disproportionately use certain words: **"delve", "tapestry", "landscape", "multifaceted", "nuanced", "navigate", "realm", "leverage", "robust", "seamless"**. An unusually high density of these words suggests AI generation.*

■ 7. Symmetrical Lists

*When listing pros/cons or comparing options, AI tends to create **perfectly symmetrical lists** (3 pros, 3 cons; 5 advantages, 5 disadvantages). Real analysis rarely comes out so neatly.*

■ 8. No Genuine Errors or Informality

Real student writing has character: crossed-out words, changed opinions mid-sentence, informal asides, and minor mistakes. AI text is polished to an unnatural degree.

AI Text Detection Tools:

- **GPTZero.me**: one of the most accurate AI text detectors, widely used by schools
- **Originality.ai**: detects AI text and plagiarism
- **ZeroGPT.com**: free AI content detector
- **Turnitin**: now includes AI writing detection for schools

Important: No AI text detector is 100% accurate. They can have false positives (flagging human text as AI) and false negatives (missing AI text). Use them as one tool among many, not as definitive proof.

Section 4: Spotting AI Voice Clones & Audio Fakes

AI voice cloning is perhaps the most concerning development because of its use in **scams**. Criminals can clone a family member's voice and call pretending to be in trouble, asking for money. In 2023, a mother in Arizona received a call from what sounded exactly like her kidnapped daughter. It was an AI clone.

■ 1. Unnatural Breathing

*AI voices often have **no breathing sounds** between sentences, or breathing that sounds mechanical and evenly spaced. Real speech includes natural pauses, breaths, and "um"s and "uh"s.*

■ 2. Emotional Flatness

While AI voices are getting better at conveying emotion, they often lack the subtle emotional variations in real speech, such as the slight catch in someone's voice when they're upset, or the natural acceleration when excited.

■ 3. Consistent Background Noise

*If someone claims to be calling from a specific location (a busy street, an airport), listen for whether the **background noise is consistent and natural**. AI-generated calls may have no background noise at all, or oddly uniform noise.*

■ 4. The "Family Code Word" Test

*Establish a **secret family code word** that you can ask for in any suspicious call. If someone claiming to be a family member can't provide the code word, hang up and call them directly on their known phone number.*

■ 5. Ask Unexpected Questions

AI voice clones work from scripts. Ask something unexpected and personal: "What did we have for dinner last Tuesday?" or "What's our dog's name?" A real person would answer naturally; a scammer would struggle.

Practice Exercises

Exercise 1: Image Analysis

Your friend shares an image on social media showing what appears to be a famous footballer holding up a shirt from a rival club, with the caption "SHOCKING TRANSFER CONFIRMED!" The image looks professional but you notice the player's right hand has 6 fingers and the text on the shirt is slightly garbled.

What type of fake content is this? (image / video / text / audio / combination)

What red flags do you notice?

What tools or techniques would you use to verify?

Verdict: REAL or FAKE? Why?

Exercise 2: News Article Check

You find an article claiming that "scientists have proven that mobile phones cause brain tumours." The article is on a website called HealthTruthRevealed.com, has no author name, no date, no links to studies, and includes a banner ad for "radiation-blocking phone cases" selling for £29.99.

What type of fake content is this? (image / video / text / audio / combination)

What red flags do you notice?

What tools or techniques would you use to verify?

Verdict: REAL or FAKE? Why?

Exercise 3: Video Verification

A video circulating on WhatsApp appears to show your school's headteacher announcing that the school is closing permanently. The video is 15 seconds long, filmed from a strange angle, and the headteacher's face seems slightly blurry at the edges.

What type of fake content is this? (image / video / text / audio / combination)

What red flags do you notice?

What tools or techniques would you use to verify?

Verdict: REAL or FAKE? Why?

Exercise 4: Voice Call

You receive a voicemail from what sounds like your parent saying they've been in a car accident and need you to immediately send money to a specific bank account. The voice sounds right, but the message has no background noise and the person doesn't use your family nickname for you.

What type of fake content is this? (image / video / text / audio / combination)

What red flags do you notice?

What tools or techniques would you use to verify?

Verdict: REAL or FAKE? Why?

Exercise 5: Your Own Find

Find something online that you suspect might be AI-generated or fake. Describe it and apply the techniques from this worksheet to analyse it.

What type of fake content is this? (image / video / text / audio / combination)

What red flags do you notice?

What tools or techniques would you use to verify?

Verdict: REAL or FAKE? Why?

What To Do When You Find Fake Content

1. Don’t Share It

The single most important action: **do not share, retweet, repost, or forward** content you suspect is fake, even to “expose” it. Sharing fake content, even with a “this is fake” caption, increases its visibility and reach. Every share trains the algorithm to show it to more people.

2. Report It

Use the **Report** button on social media platforms. All major platforms (Instagram, TikTok, YouTube, Facebook, X) have specific options for reporting misinformation, deepfakes, and manipulated media. Reporting helps the platform’s AI and moderation teams identify and remove harmful content.

3. Inform the Person Who Shared It

If a friend shared the fake content, let them know **privately and kindly**. They probably didn’t know it was fake. Send them a direct message rather than publicly calling them out, which can be embarrassing and counterproductive.

4. Tell a Trusted Adult

If the fake content involves something serious (scams, threats, impersonation of real people, or illegal content), tell a parent, teacher, or trusted adult immediately.

5. Document It

If needed, **take a screenshot** (including the URL and date) before reporting. This preserves evidence in case the content is removed but needs to be referenced later.

6. Check Fact-Checking Sites

Visit **Snopes.com**, **FullFact.org**, or **FactCheck.org** to see if the claim has already been debunked. If it has, you can share the fact-check article to help counter the misinformation.

Key Statistics to Remember

- AI-generated images increased from **1 million per day** in 2022 to over **34 million per day** in 2024 (Everypixel)
- Only **38%** of people can reliably distinguish AI-generated faces from real ones (research published in PNAS)
- Deepfake fraud increased by **3,000%** from 2022 to 2023 (Onfido Identity Fraud Report)
- The average person encounters an estimated **5 to 10 pieces of misinformation per day** on social media
- **65%** of young people aged 12 to 15 have seen content they later discovered was fake (Ofcom)
- AI can now generate a realistic **60-second deepfake video** in under 10 minutes

Your Digital Defence Toolkit

In a world where AI can create increasingly convincing fake content, your **critical thinking skills** are your most powerful tool. Technology will continue to evolve, and detection methods will need to evolve with it. But the fundamental approach remains the same:

Question everything. Verify before you trust. Never share what you can't confirm.

"A lie can travel halfway around the world while the truth is still putting on its shoes." ~ Attributed to Mark Twain (and ironically, this quote itself is probably misattributed!)