Detecting AI-Generated Images: What Constitutes Successful Separation Between the Authentic and Artificial?

Thomas Doornbos

s4156420

Department of Psychology, University of Groningen

PSB3E-BT15.2024-2025.3: Bachelor Thesis

Group number: 42

Supervisor: Ben Gützkow PhD, Dr. Wim Meerholz

In collaboration with: Aiske Ederer, Luuk Mulder, Timucin Mutlu, Anja Schmeinck, Dennis Steenbergen

June 24, 2025

Abstract

Image generation, powered by machine learning models has become immensely prevalent in recent years. This development has raised questions about whether humans can differentiate effectively between authentic and artificially generated images. By conducting an experiment ($N = 194$) with two condition groups, those who receive AI-training and those who do not, we investigated whether training can effectively help people in distinguishing between AI and non-AI images. Training was done using an inductive learning paradigm, where participants were exposed to an array of labeled images. Subsequently, both groups of participants were tested in their image detection accuracy. Participants were also tested in their cognitive processing style using the Cognitive Reflection Test (Frederick, 2005) and asked about their social media and internet usage. Results were evaluated using linear regression models, where only online short-form content consumptions showed significant predictive power regarding image detection accuracy. Training did not improve detection accuracy, while it did make people more inclined to mark images as AI-generated. We discuss implications of our findings and future research directions.

Detecting AI-Generated Images: What Constitutes Successful Separation Between the Authentic and Artificial?

The recent blossoming in the area of artificial intelligence (AI) has caused a major uproar in many areas of life. Where companies are restructuring to make use of new technologies (Joint Research Centre, 2025) and the concept of truth and authenticity can be called into question (Tarasenko, 2023). The internet has also become more concentrated towards large social media platforms in recent years. Platforms like Instagram, TikTok, Facebook, YouTube and X (Twitter) have become the dominating websites on the web (Backlinko, 2025). These websites operate on algorithms that allow for a rapid sharing of content to large audiences. With ongoing debates about whether the content being shared on those platforms actually represents reality, giving rise to discussions about the impact of "fake news" or "disinformation" (Raman et al., 2024). Social media algorithms are driven by engagement, with controversial or norm-breaking posts often spreading fastest on the platforms (Etta et al., 2023). Automated AI-driven systems are able to optimise towards this algorithmic system by targeting the widest reach, a recent example of this is an AI-bot from the University of Zurich that infiltrated a Reddit forum and was very popular and successful in persuading users (O'Grady, 2025). Developments like these make the issue of recognising AI-generated content all the more relevant. Additionally with the development of neural network based image generation services, the barrier to create the world in your image has never been lower. With more than 5 billion people making use of social media platforms like the ones mentioned above, the reach for AI-generated content is thus far and wide.

Machine learning models have been improving at a steady pace (Cabello, 2023), with image generation models becoming more able in producing lifelike images. By following the assumption that these models will increase to get better over time, it can be important to

assess the degree to which people can detect whether an image is AI-generated or stemming from an authentic source. Additionally the factors behind one's AI detecting capabilities, and whether training could aid detection may also be of great interest.

Past research has shown that exposure to stimuli can train people to recognise patterns or categories within a set of stimuli, through inductive learning. Inductive learning is the process by which one derives broader principles from repeated exposure to single instances (Kornell & Bjork, 2008). One example of this is a child who, from being exposed to different dogs learns to categorise the animals in the concept of a dog. Kornell and Bjork (2008) found that exposure to different paintings from a variety of artists of the same time period made participants better able to detect to which artist subsequent paintings belonged to. They investigated two inductive learning paradigms: interleaved and massed exposure. In interleaved (spaced) learning, paintings of different artists are alternated, where no painting of the same artist is repeated after one another. In massed learning you are exposed in blocks; you see all the paintings of the same artist in succession and then move to another one. The learning effect was largest with interleaved learning, where massed learning yielded worse results. One of the benefits of interleaved learning is that it promotes a reflection on the distinction between different categories. It encourages active exploration and reduces fatigue of seeing the same type of stimuli for a prolonged period (Kang & Pashler, 2012). Exposure to content online yields a similar type of inductive, interleaved learning, where occasional AI-generated content can slip through your timeline. These research findings form the basis for the interleaved inductive learning paradigm in this study, leading to the following hypothesis:

H1: People who receive AI training through an interleaved inductive learning paradigm will become significantly better at correctly identifying AI and real photos, compared to people who receive no training.

Given the vast individual differences in people, there is reason to believe that some people may be more or less sensitive to spotting subtle differences or use different methods for detecting these differences, independent of any prior training. Since AI generated pictures may often be found on social media platforms, where timelines could lead to short exposure, it may be helpful to investigate how different information processing techniques could influence perception. There are many conceptualisations of information processing in the literature, many of which agree that there is a rough difference between *intuitive* thinking styles and *deliberate* thinking styles, although there are active debates on the intricacies of how both forms operate (De Neys, 2021). The concept of *fast* and *slow* thinking was popularised outside of academic culture by Kahneman (2011), where a heuristic like the affect heuristic (Slovic et al., 2002) in fast thinking can be used to make judgements about the "realness" of the image from feelings that the image evokes. When an image feels too perfect or unnatural, these types of judgements can be related back to fast processes.

The distinction between these two types of thinking can be useful for conceptualising how someone may interpret an AI image. One could perceive an image intuitively as "real" or "fake", but may not look for the specific reasons what makes the image to be perceived that way, this can be seen as *fast* processing. On the other hand, someone may look at an image very deliberately, by analysing small details that could hint at AI generation. Although it may seem that deliberate processing would lead to better results, given a history of psychological research about cognitive biases and the type of errors they can induce (Maule & Hodgkinson,

2002). There is however also an argument to be made that a deliberate, manual approach can lead to overthinking or other errors. When people are engaged in an analytical mode of thinking, they typically spent longer on answering a question, this however, does no necessarily increase the accuracy of their answers (Thompson et al., 2013). Since media is often consumed for brief periods on social media, deliberate processing may be too time consuming and unreliable, favouring an intuitive approach. This study also simulates social media exposure by limiting exposure to 15 seconds. Because of this reasoning we form the following hypothesis:

H2: People who rely predominantly on automatic information processing are significantly better at detecting whether an image is AI-generated or not than people who rely predominantly on manual information processing.

Since training may help develop the skills used for detecting AI images, there is reason to believe that there may exist a compound effect between training and information processing styles. Repeated exposure to AI content could aid in the intuitive perception of the realness of it. Prior learning through induction can improve familiarity and in turn make completing a task easier. When a task is perceived to be easy, one relies more on automatic processing (Thompson et al., 2013). Groups who receive training could potentially be more proficient in their intuitive sense from their previous exposure, than those who do make use of intuitive processing, but have not had prior training. This familiarity may foster an effect greater than the sum of the training and automatic information processing effect:

H3: People who rely predominantly on automatic information processing and receive AI training through an interleaved inductive learning paradigm will be significantly better than the sum of these individual effects.

There are large differences in the manner in which people engage with internet platforms, and this may have an effect on how accurately they can process information quickly. The amount of time spent is one factor, another factor are the different types of platforms used. Internet content can be distributed in a variety of ways. Many people seem to primarily engage with *short-form* content (E.g. TikTok videos, Instagram stories, Tweets) rather than *long-form* (E.g. articles, YouTube videos, podcasts). People who primarily consume short-form content may already be somewhat trained to spot certain details or differences between different types of content more quickly than one who engages more with long-form content, since quick judgement and evaluation is part of the user experience.

The way people engage with these platforms can be vastly different. Whereas a session on TikTok can be characterised by swift engagement between many different posts, with quick interpretations, listening to a podcast, for example, can lead to longer pondering and slower information processing. This is effect is highlighted by a small scale study by Ketonen and Nieminen (2023), which found that listening to podcasts fostered reflective thinking and pondering in students. Past research has also shown that deliberate processing is associated with greater time investment, ostensibly to process the information deeper (Gambetti et al. 2020). Feedback, can play an important role in the use of intuitive or deliberate modes of thinking. It has been found that positive feedback reinforces the use of intuitive thinking, leading to more swift responding, whereas negative feedback could engage deliberately monitoring of their intuitive judgements (Gambetti et al. 2020). This highlights the ability to shift between different modes, when the environment require so. Quick access to AI-generated content on short-form platforms can provide information to the user that engages moments of deliberate reflection during their session of social media use, which is

otherwise characterised by deep, automatic immersion. This phenomenon is not limited to short-form content itself, but the structure of these platforms may foster a type of exposure that is more suited for AI-learning than long-form platforms. There is however good reason to believe that long-form platforms and internet usage in general can still help develop familiarity with AI generated pictures, which could aid detection. Taking this into account the following hypotheses are formed:

H4a: People who engage primarily with short-form content are significantly better at AI image detection than those who primarily engage with long-form content.

H4b: More time spent on internet platforms significantly aids AI image detection.

Method

Participants

An initial 222 participants were recruited for the study. 19 participants were excluded due to incomplete responses and nine more for failing either of two attention checks, for a final sample of 194 (70.6% female, 28.9% male, 0.5% non-binary). Participants had to be at least 16 years old. The participants represent a convenience and snowball sample, made up of participants recruited both through the University of Groningen's SONA platform and privately by the research team, the latter of whom were asked to refer additional participants. For participants recruited through the SONA platform, the SONA ID was collected solely for the purpose of assigning course credits, where participants recruited from outside the SONA-system received no compensation. The data from participants was collected anonymously and stored in a secure location in the department of Psychology at the University of Groningen, in accordance with the data management protocol of the Heymans Institute and GDPR

regulations. Ethical approval was obtained from the Ethics Committee of the University of Groningen. Data collection for the study ran during the month of April 2025.

Materials and Stimuli

The study was conducted as an online experiment hosted on Qualtrics. After giving informed, active consent, all participants were first asked to fill in a number of scales and items. Multiple scales were included in the questionnaire that were of interest to other members of the research group; these will not be discussed in this paper. To assess social media usage, participants were asked three questions about their time spent on online platforms and what type of content they engage with most frequently. They are asked: "*How much time do you typically spend on the internet each day?*" with answers: "*Less than one hour*", "*1-3 hours*", "*3-5 hours*", "*5-7 hours*" and "*More than 7 hours*" (M = 3.13, SD = 0.84). Afterwards they were asked two questions about how much time they spend on short-and-long-form platforms, respectively: "*How often do you engage with <u>short-form</u> content (E.g. TikTok videos, Instagram stories, Tweets)?*" and "*How often do you engage with <u>long-form</u> content (E.g. articles, YouTube videos, podcasts)?*". These questions were asked on a 5-point Likert scale from "*Not at all*", "*Less than 15 minutes a day*", "*Between 15 minutes and 1 hour a day*", "*Between 1-3 hours a day*" to "*More than 3 hours a day*" Short-form (M = 3.37, SD = 0.87), long-form (M = 3.37, SD = 0.89).

Cognitive style was assessed with a six-item test aimed at resembling the Cognitive Reflection Test by Frederick (2005). Examples include:

- "*Julie has 5 dolls. Julie has 4 more dolls than Angie. How many dolls does Angie have?"* (open question).
- "*Bill is 34. He is intelligent, punctual but unimaginative and somewhat lifeless. In school, he was strong in mathematics but weak in social studies and humanities.*

Which one of the following statements is most likely?"

- "*Bill plays in a rock band for a hobby*"
- "*Bill is an accountant and plays in a rock band for a hobby*"

- "*A magazine and a banana together cost €2.90. The magazine costs €2 more than the banana. How much does the banana cost?*" (open question).

The full list of questions is provided in Appendix A.

Visual Stimuli

A total of 120 photograph stimuli were used for the study. Of these, sixty were real photographs, selected from public image-sharing websites (Pixiv, Pixabay, Pexels), from photographers (see Appendix B) who granted permission to use their work, as well as personal photography by the research team members. The remaining sixty were faux photographs, generated via the AI image generation models of Midjourney Version 6.1 and Grok image generation (state march 2025).

Both AI and genuine photography were furthermore each separated into three conditions of equal sizes, based on image content: (I) Landscape photography, occasionally also depicting edifices like castles, (II) "everyday" photography, depicting humans in situations as would typically be observed in various everyday settings, and (III) artistic photography, depicting one or two humans in stylised photo shootings, with specific and staged elements like composition and lighting. Thus, there were twenty pictures for each condition in each of the six photograph conditions. Example images for each condition, as well as prompts used for image generation, are provided in Appendix B.

Procedure

At the start of the online study, participants were asked to fill out a series of questionnaires. After this, participants were randomly assigned to one of two conditions for

the upcoming inductive learning task. Participants randomly assigned to the experimental condition were informed that the task consisted of a learning and subsequent testing phase. In the learning phase, participants were to be presented with photographs that were either AI-generated or genuine photography, along with a corresponding label. These images were each shown for five seconds, without a pause in between. Unbeknownst to the participants, the images, while themselves selected at random, followed an underlying interleaved pattern, meaning that genuine photography was always followed by an AI-generated photo, and vice versa. This was done to promote discriminative contrast between the two generation types (see Kang and Pashler, 2011). After all 78 images (made up in equal parts of the six categories, for thirteen images each) were shown, the learning phase commenced. Participants randomly assigned to the control condition were not given a learning phase, and skipped straight to the instructions for the testing phase. Here, participants were presented with an image for fifteen seconds. In this time, they had to indicate whether they believed the image to be AI-generated or genuine photography. After fifteen seconds elapsed or participants continued to the next page, the next image was shown. It was not possible for participants to pause during this time. The testing phase consisted of 42 images, once again made up in equal parts of the six categories, for seven images each. After the study, participants could see their final score on the test.

Results

Preliminary Analysis

An accuracy percentage score for each participant was calculated by dividing their correct answers by their total completed responses. This was also computed for real and AI-image categories (see Table 1). The participants were split in two experimental groups: control (M = .571, SD = .079, N = 103) and training (M = .572, SD = .088, N = 91). The total

accuracy score on the test was used as the dependent variable to test all hypotheses, which were tested using linear regression models. All models were checked on linearity, normality of residuals, homoscedasticity and extreme outliers. No violations were found.

Table 1

Descriptive Statistics

Condition	Real accuracy	AI accuracy	Total accuracy
Control	.675 (*SD* = .139)	.467 (*SD* = .153)	.571 (*SD* = .079)
Training	.559 (*SD* = .136)	.586 (*SD = .116)*	.572 (*SD* = .088)

Main Analysis

Hypothesis 1

To test the hypothesis whether training resulted in significantly improved AI and non-AI image detection, we performed a linear regression analysis. The training effect was tested using *condition* as the independent variable, where control = 0 and training = 1, and total accuracy as the dependent variable. Linear regression shows no correlation: $b = .00$ ($SE = .012$), $t(192) = 0.066$, $p = .948$. The model did not explain variance of the dependent variable: $R^2 = .00$, $F(1, 192) = 0.004$, $p = .948$. AI-training through an interleaved inductive learning paradigm did not improve their scores on a later test.

Table 2

Linear Regression Model for Condition on Test Accuracy Scores

Predictor	Coefficient	*SE*	*t*	*p*
Intercept	*.571*	.008	69.727	<.001
Condition	*.000*	.012	0.066	.948

Hypothesis 2

To test whether people who rely predominantly on automatic information processing are significantly better at detecting whether an image is AI-generated or not, we performed a linear regression analysis. Scores from the Cognitive Reflection Test, ranging from 1 to 6 (M = 4.40, SD = 1.25, N = 194), were used to predict test accuracy, resulting in the following model: b = -.002 (SE = .005), $t(192)$ = -0.416, p = .678. The model did not explain variance of the dependent variable: R^2 = .00, $F(1, 192)$ = 0.173, p = .948. CRT scores are coded where high scores mean more deliberate processing, and low scores mean more automatic processing. For this reason, to support the hypothesis you would need a significant *negative* linear coefficient. The scores are however far from significant, indicating that it is not a good predictor.

Table 3

Linear Regression Model for CRT Score on Test Accuracy Scores

Predictor	Coefficient	*SE*	*t*	*p*
Intercept	*.580*	.022	26.489	<.001
CRT_score	*-.002*	.005	-0.416	.678

Hypothesis 3

To assess whether the interaction effect between the experiment condition and CRT scores is a significantly stronger predictor for test accuracy scores than both variables combined, we compare two models. The first model consists of the experimental condition and CRT scores, but not the interaction effect. Leading to the following correlations: $b_{Condition}$ = .00 (SE = .012), $t(192)$ = 0.072, p = .943 and b_{CRT_score} = -.002 (SE = .005), $t(192)$ = -0.416, p = .678 (see Table 4). This model provides the following explanatory variance of the dependent test scores: R^2 = .00, $F(2, 191)$ = 0.089, p = .915.

The second model is an extension of the first model, but contains the interaction effect between the condition and CRT scores to predict test accuracy scores. Leading to the following correlations: $b_{Condition}$ = -.007 (*SE* = .044), *t*(192) = -0.167, *p* = .867, b_{CRT_score} = -.003 (*SE* = .006), *t*(192) = -0.439, *p* = .661 and $b_{Condition*CRT\_score}$ = .002 (*SE* = .010), *t*(192) = 0.194, *p* = .846 (see Table 5). This extended interaction models provides the following explained variance: R^2 = .001, *F*(3, 190) = 0.071, *p* = .975.

To test the hypothesis whether the interaction model is significantly better than the original model, we need to look at the difference in predictive power of the two models. Adding the interaction to the model yields the following non-significant change in explained variance: ΔR^2 = .000, F_{change}(1, 190) = 0.038, *p* = .846. Since both the experimental condition and the CRT scores are poor predictors of test results, it is unsurprising that the interaction between these two variables provides no significant results.

Table 4

Linear Regression Model for Condition and CRT Score on Test Accuracy Scores

Predictor	Coefficient	*SE*	*t*	*p*
Intercept	*.580*	.023	25.668	<.001
Condition	*.000*	.012	0.072	.943
CRT_score	-.002	.005	-0.416	.678

Table 5

Linear Regression Model for Condition, CRT Score and Interaction on Test Accuracy Scores

Predictor	Coefficient	*SE*	*t*	*p*
Intercept	*.584*	.029	19.867	<.001
Condition	*-.007*	.044	-0.167	.867
CRT_score	-.003	.006	-0.439	.661

Condition*CRT_score	*.002*	.010	0.194	.846

Hypothesis 4a

Here we compare the predictive power of short-form content consumption and long-form content consumption on accuracy scores in the image detection test. Short-form scores ranged from 1 to 5 (M = 3.73, SD = 0.87, N = 194), with long-form scores also ranging from 1 to 5 (M = 3.37, SD = 0.89, N = 194). Before determining whether short-form content consumption is a significantly better predictor than long-form content consumption we compare both models independently: $b_{short\text{-}form}$ = .019 (SE = .007), $t(192) = 2.85$, p = .005, with explained variance: R^2 = .041, $F(1, 192) = 8.12$, p = .005 and $b_{long\text{-}form}$ = .011 (SE = .007), $t(192) = 1.57$, p = .119, with explained variance: R^2 = .013, $F(1, 192) = 2.45$, p = .119.

Independently, short-form content consumption significantly predicts test score accuracy, whereas long-form content consumption does not. To test whether short-form content consumption is significantly better in predicting than long-form content consumption, we combine both variables in one model, coefficients that are part of the regression equation are shown in Table 6. By calculating the change in variance explained between a model only consisting of long-form content consumption and the model with both short-form and long-form consumption included, we see the following change: ΔR^2 = .043, $F_{change}(1, 191) = 8.78$, p = .003. As shown, adding short-form content provides a significant difference in prediction power to the long-form content model. Highlighting that people who engage primarily with short-form content are significantly better at AI image detection than those who primarily engage with long-form content.

Table 6

Linear Regression Model for Short-Form and Long-Form Content Consumption on Test Accuracy Scores

Predictor	Coefficient	*SE*	*t*	*p*
Intercept	.458	.035	13.11	<.001
short-form	.020	.007	2.96	.003
long-form	.012	.007	1.77	.079

Hypothesis 4b

To test the hypothesis whether people who spend more time on the internet are significantly more accurate at detecting whether an image is AI-generated or not, we performed a linear regression analysis. Internet use was measured using scores ranging from 1 to 5 ($M = 3.13$, $SD = 0.84$, $N = 194$), giving the following association between the variables: $b = .009$ ($SE = .007$), $t(192) = 1.25$, $p = .213$. The model provides the following variance explained: $R^2 = .008$, $F(1, 192) = 1.56$, $p = .213$. Showing that time spent on the internet is not a significant predictor for test scores.

Table 7

Linear Regression Model for Internet Use on Test Accuracy Scores

Predictor	Coefficient	*SE*	*t*	*p*
Intercept	*.544*	.023	23.63	<.001
Internet use	.009	.007	1.25	.213

Further Investigation

Upon further inquiry into the relationship between training and test answers, an interesting relationship showed up. Although people in the control condition and training condition performed similarly with respect to the total items that were answered correctly, a

difference emerged between what type of question people from the training condition answered correctly (see Table 1). Since each image on the test can either be AI-generated or authentic, when one favours one category, namely marking images as AI-generated, scores are inflated in that category, while worsening in the other.

Regression results confirm this effect, there is a significant negative relationship between the training condition and real picture detection scores: b = -.117 (SE = .020), t(192) = -5.89, p < .001. With the following variance explained: R^2 = .153, F(1, 192) = 34.7, p < .001. The association between training and AI picture detection is the following: b = .118 (SE = .020), t(192) = -5.98, p < .001. With variance explained: R^2 = .157, F(1, 192) = 35.8, p < .001. This shows that participants in the training condition have accuracy scores that are 0.118 points higher for AI images, while scoring 0.117 points lower at detecting real images, than the control condition. Participants were exposed to an equal share of AI and non-AI images, this was however not explicitly known to participants in the study. When many participants in the training condition marks more questions as AI-generated, accuracy for AI images increases, while decreasing an equal share for real images. The same is true for marking more images as real, which on average increases accuracy for real images and reduces accuracy for AI images.

Discussion

Summary of Results and Implications

The advent of widely available AI technologies and its widely adopted use has provided an important incentive for research. Since this is a relative recent phenomenon, we investigate the following question in this paper; mainly how training and exposure to AI content affects whether you can detect its presence correctly. We also investigated what other factors constitute successful detection, mainly: cognitive styles (intuitive and deliberate) using

a CRT measure (Frederick, 2005) and internet usage. Internet usage was, besides investigating the total time spent, studied through the usage of both short and long-form platforms. Results show that among all investigated variables, only the time spent on short-form social media platforms form was a good predictor for AI and non-AI image detection accuracy.

The effect of training through interleaved inductive learning yielded no effect. This runs counter to the prediction that exposure to identified images will sharpen the sense of in what group subsequent images belong to. It could have been the case that a longer learning period, consisting of multiple sessions would help foster a better detection ability. There is however also reason to believe that training, followed by an immediate test is already an efficient way to learn and test new information. Since people remember information well that is recently learned (Murdock, 1962).

Although the overall effect of training is not existent, participants in the training condition show that they are more prone to mark something as AI in the test, resulting in higher AI picture correct scores and lower real picture correct scores. This is an interesting phenomenon, since it has real life implications. It is not far fetched to assume that after being exposed to a series of AI and real pictures that a belief can form that AI pictures can be quite convincing. Obvious telltale signs in AI generated images are becoming less obvious and this realisation could give room for doubt about the realness of organic pictures as well. There is also the role of skepticism and naivety. Intrapersonal motives could lead someone to be more skeptical after receiving AI training, because being confronted with the potential to be naive or fooled by AI pictures, after being exposed to them numerous times, can be more damaging to the self-view than being overly careful with it.

Differentiating between automatic and deliberate cognitive styles using the Cognitive Reflection Test (Frederick, 2005) has not proven to be effective in predicting test scores. The

question arises whether it is accurate to assign a stable value, namely scores on a test, to thinking styles that people may use while trying to decipher the origin of presented images on a test. Like highlighted in the introduction, different factors like time constraints and task difficulty can influence one's processing style. Furthermore, it is unclear whether participants who are very deliberate in answering questions on a CRT, share the same approach for an AI detection test, since other factors like content specific motivation can also play a role. These drawbacks call into question whether processing styles are a poor predictor of test scores or whether the CRT in itself is an unreliable measure for conceptualising processing styles in the context of AI detection.

There is some indication that social media use is beneficial for AI detection. Short form content use significantly predicts higher scores on the test. Long-form content consumption also possesses some predictive power, but to a lesser extent than short-form content. Since almost all participants engage with social media on a regular basis, there is a strong chance they were already exposed to some amount of AI content, and in turn were trained by it. This could be another reason why AI training was insufficiently effective in this study.

What are the mechanics that could explain how social media use aids AI detection? Making use of social platforms on the internet gives a high probability of being exposed to AI generated content. At times you may not notice it and look past it, other times you might not notice it at first, but only come to realise it afterwards. This is a form of AI training. Depending on the emotional impact of this experience of the discovery of AI generated content during your browsing session, shifts in attitudes may occur (for an example regarding fake-news: Effron & Raj, 2019). When one becomes more skeptical about the realness of images on the internet, trust can erode in the factuality of what is seen on a screen. When this

becomes widespread, and objective truth has been overtaken by non verifiable experience, societal shifts may occur.

Interestingly internet use in general was not a good predictor of picture detection. The use of a categorical variable may have also played a role in why internet use is not a good predictor, in comparison to short and long-form content. While both concepts are being operationalised using a 5 point scale. In the short/long-form scales it ranges from *"Not at all"* to *"More than 3 hours"*, whereas internet use ranges from *"Less than one hour"* to *"More than 7 hours"*. This resulted in a different spread of values. A way to avoid this issue is by asking participants for continuous values or continuous values obtained from usage data from their phone or computer. The reason this was not opted for in this study was because of the worry that participants would find it difficult to guess their usage accurately, or would be unable to access usage data from their devices. Internet use is also a broader concept, besides social media, activities like streaming movies and TV shows, and playing video games could contribute a significant amount of time spent using the internet. These activities currently contain minimal exposure to AI generated content. By having these activities included, the predictive power of internet use as a whole is reduced in this context.

One thing that stood out in general is the fact that participants identified the presented images with a 57.2% accuracy. This score is mildly better than chance, and gives strong indication that AI images in its current state are already difficult to separate from real images. We made an effort to generate highly realistic images using cutting edge models that were available during the time of research, which were Midjourney and Grok. Later, OpenAI (2025) released the 4o image generation model, which pushed the limits of AI image generation models even further, highlighting the rate of improvement in this field.

It is difficult to estimate the magnitude of AI-generated content on the internet, since, like highlighted in this paper, AI detection is difficult. Data from popular image generation models from 2024 suggests that more than 34 million images are generated every day. Where it estimated that 15 billion images were generated from 2022 to 2023. To put this into context; it took photography 149 years to produce the same amount of images (Valyaeva, 2024). Not only are AI-generated images extremely prevalent, generating text using Large Language Models (LLMs) has become commonplace, where many word editors invite you to make use of these tools (Spataro, 2023). Since these models are ever improving, there is good reason to believe that we are just at the beginning of experiencing a largely machine-generated internet. Due to the rapid development of this field, this research finds itself in a somewhat uncomfortable position, where developments outpace understanding. The results highlight that people are already not great in putting AI or non-AI images in the correct category. Training does not seem to help in this respect, while ironically, training has been foundational for machine learning models themselves.

Strengths and Limitations

Participants of this study are primarily from a student population from the University of Groningen. This is especially relevant in this study, since AI is a novel but extremely relevant topic. There is a lot of mystery surrounding it, where a deeper understanding of these systems in the way they work can be missing in the older generation (Shandilya & Fan, 2022). There is reason to believe that an older generation would benefit from AI-training. Since they may have had little informal training beforehand. Different generations also generally spend their time on the internet on different platforms (Statista, 2024), potentially resulting in different rates of AI exposure. Social class may also play a role, since an affluent student

population may have a deeper relationship with AI, whereas people from a lower socioeconomic status may have experience with AI, but lack the deeper knowledge about it.

One major consideration consideration for the study setup was the selection of real and AI images. This is a potential source of considerable bias in the results, since it is possible to favour AI images that look most real, while simultaneously selecting organic images that look fake. We tried to overcome this bias by making a varied collection of pictures and trying to create a fair final selection from it. This process is however a subjective process, that can be exposed to unknown biases.

This experiment was conducted online. This gave participants the ability to complete the experiment in their own time, on a device of their choosing. Even though a computer with a decently sized display was recommended, a portion of the sample completed the study on their mobile device. Smaller screens can make it more difficult to make out small details in a picture. Since these small details can be pivotal in categorising AI-generated pictures correctly, it could be better to limit participants to desktop/laptop use. This study also had a number of dropouts or unusable data points. This has the potential to skew scores, since these dropouts could have dropped out because of a reason within the study, or their study condition. An experimental setup with a fixed display size could help with reduced dropouts and increase confidence that the scores in the experiment are not due to external factors. One downside of this approach however is the increased difficulty in obtaining a large sample size, with reduced statistical power as a potential consequence.

This study does have some strengths. There is a large sample of participants, these participants are also exposed to many test items to assess their detection accuracy. Different measurements are also used to further explain what relates to good or bad image detection, which gives the ability to strengthen prediction models. The experimental setup also gives the

ability to provide unique insight in whether training leads to a change in detection quality. There is also a high degree of internal validity, since the training and control condition both share the same experimental qualities, where there is little room for confounding.

Future Directions

Although the future is uncertain, it is clear that this area of research is nonetheless valuable. Even though scientific research can fall behind in understanding the technological developments, science is also crucial for illuminating areas where knowledge is lacking. Our current findings are not definitive, but function as an area of exploration into a novel field.

Due to concerns about test length and accessibility, measurements of the CRT and social media/internet use were implemented in a rudimentary manner. This has been at the cost of the robustness of the findings. Future research can make use of more precise measurements targeting visual processing styles, to investigate the relationship between intuitive and deliberate processing deeper. Social media use, specifically related to short-form content can be investigated further to unearth the mechanisms in which exposure to online content can help in learning to differentiate AI from non-AI content. Prior reported AI experience or AI knowledge can also be tested individually to see how relates to other concepts and final test scores. We noticed that the image selection process is a potential area for bias in research design. To solidify measurement quality, one can make use of a standardised approach, consisting of a pretest that detects test items that meet the objective criteria. A highly controlled experimental setup on a physical location can further help account for potential confounding variables, simultaneously, by making use of a large sample of a diverse population, external validity can be improved. By investigating different concepts that could explain AI-detecting accuracy it may be possible to form comprehensive theories that can be used for effective AI-training programs. Increased research in the domain of

artificial intelligence can also help in enabling a snowball effect, where breakthroughs in research could foster the development of more important research findings.

References

Backlinko (2025, 10 februari). Social Media Usage & Growth Statistics. Backlinko. https://backlinko.com/social-media-users

Cabello, A. (2023, 3 november). The Evolution of Language Models: A Journey Through Time. Medium. https://medium.com/@adria.cabello/the-evolution-of-language-models-a-journey-through-time-3179f72ae7eb

CTOL Editors - Xia & CTOL Digital Business and Technology News. (2025, 29 april). AI Bots Secretly Impersonated Trauma Survivors on Reddit as University Research Team Faces Legal Action. CTOL Digital Solutions. https://www.ctol.digital/news/reddit-university-zurich-ai-bots-legal-action/

De Neys, W. (2021). On dual- and single-process models of thinking. Perspectives on Psychological Science, 16(6), 1412–1427. https://doi.org/10.1177/1745691620964172

Effron, D. A., & Raj, M. (2019). Misinformation and Morality: Encountering Fake-News Headlines Makes Them Seem Less Unethical to Publish and Share. Psychological Science, 31(1), 75-87. https://doi.org/10.1177/0956797619887896

Etta, G., Sangiorgio, E., Marco, N. D., Avalle, M., Scala, A., Cinelli, M., & Quattrociocchi, W. (2023). Characterizing engagement dynamics across topics on Facebook. PLoS ONE, 18(6). https://doi.org/10.1371/journal.pone.0286150

Frederick, S. (2005). Cognitive reflection and decision making. Journal of Economic Perspectives, 19(4), 25–42. https://doi.org/10.1257/089533005775196732

Gambetti, E., Marinello, F., Zucchelli, M. M., Nori, R., & Giusbeerti, F. (2020). You are right, you are wrong: The effect of feedback on intuitive thinking. The American Journal of Psychology, 133(4), 473–485. https://doi.org/10.5406/amerjpsyc.133.4.0473

Joint Research Centre. (2025, June 13). How is generative AI impacting our economy, society and policy? The Joint Research Centre: EU Science Hub. https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/how-generative-ai-impacting-our-economy-society-and-policy-2025-06-13_en

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux

Kang, S. H. K., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. Applied Cognitive Psychology, 26(1), 97–103. https://doi.org/10.1002/acp.1801

Ketonen, L., & Nieminen, J. H. (2023). Supporting student teachers' reflection through assessment: The case of reflective podcasts. Teaching and Teacher Education, 124, 1–11. https://doi.org/10.1016/j.tate.2023.104039

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: is spacing the "enemy of induction"?. Psychological science, 19(6), 585–592. https://doi.org/10.1111/j.1467-9280.2008.02127.x

Mandel, D. R., & Kapler, I. V. (2018). Cognitive style and frame susceptibility in decision-making. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.01461

Maule, A. J., & Hodgkinson, G. P. (2002). Heuristics, biases and strategic decision making. The Psychologist, 15(2), 68–71.

Murdock, B. B. (1962). The serial position effect of free recall. Journal of Experimental Psychology, 64(5), 482–488. https://doi.org/10.1037/h0045106

O'Grady, C. (2025, 30 april). 'Unethical' AI research on Reddit under fire. Science. https://www.science.org/content/article/unethical-ai-research-reddit-under-fire

OpenAI. (2025, March 25). Introducing 4o Image Generation [Press release]. https://openai.com/index/introducing-4o-image-generation/

Raman, R., Kumar Nair, V., Nedungadi, P., Kumar Sahu, A., Kowalski, R., Ramanathan, S., & Achuthan, K. (2024). Fake news research trends, linkages to generative artificial intelligence and sustainable development goals. Heliyon, 10(3), e24727. https://doi.org/10.1016/j.heliyon.2024.e24727

Shandilya, E., & Fan, M. (2022). Understanding Older Adults' Perceptions and Challenges in Using AI-enabled Everyday Technologies. arXiv.Org, abs/2210.01369. https://doi.org/10.48550/arXiv.2210.01369

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). Rational actors or rational fools: Implications of the affect heuristic for behavioral economics. The Journal of Socio-Economics, 31(4), 329–342. https://doi.org/10.1016/S1053-5357(02)00174-9

Spataro, J. (2023, May 16). Introducing Microsoft 365 Copilot – your copilot for work - The Official Microsoft Blog. The Official Microsoft Blog. https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/

Statista. (2024, August 5). Social media used in GB 2024, by age group. https://www.statista.com/statistics/1480562/social-media-platforms-used-by-age-group/

Tarasenko, I. (2023, December 17). Rethinking Authenticity in the AI Revolution - Igor Tarasenko - Medium. Medium. https://medium.com/@saik0s/dissecting-authenticity-in-the-ai-era-a-disruptive-take-6c54f2f8a35a

Thompson, V. A., Turner, J. A. P., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. Cognition, 128(2), 237–251. https://doi.org/10.1016/j.cognition.2012.09.012

Valyaeva, A. (2024, August 14). AI image statistics for 2024: How much content was created by AI. Everypixel Journal - Your Guide to the Entangled World of AI. https://journal.everypixel.com/ai-image-statistics

Appendix A

Test Items CRT

You are faced with two trays each filled with white and red jelly beans. You can draw one jelly bean without looking from one of the trays. Tray A contains a total of 10 jelly beans of which 2 are red. Tray B contains a total of 100 jelly beans of which 19 are red.

From which tray should you draw to maximize your chance of drawing a red jelly bean?

A:

- Tray A
- Tray B

A psychologist wrote thumbnail descriptions of a sample of 1000 participants consisting of 995 females and 5 males. The description below was chosen at random from the 1000 available descriptions.

Jo is 23 years old and is finishing a degree in engineering. On Friday nights, Jo likes to go out cruising with friends while listening to loud music and drinking beer.

Which one of the following two statements is most likely?

A:

- Jo is a woman
- Jo is a man

Julie has 5 dolls. Julie has 4 more dolls than Angie. How many dolls does Angie have? (open question)

Premises:

- All flowers need water
- Roses need water

Conclusion: Roses are flowers

A:

- This conclusion follows logically
- This conclusion does not follow logically

Bill is 34. He is intelligent, punctual but unimaginative and somewhat lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Which one of the following statements is most likely?

A:

Bill plays in a rock band for a hobby

Bill is an accountant and plays in a rock band for a hobby

A magazine and a banana together cost €2.90. The magazine costs €2 more than the banana. How much does the banana cost? (open question)

Appendix B

Photographers

Anastasiya Pronchenko, Lary Rauh, Maja Elders, Timucin Mutlu

Example Real Images

Note. Picture description from left to right: Everyday humans, Landscapes, Artistic humans.

Example AI Images

Note. Picture description from left to right: Everyday humans, Landscapes, Artistic humans.

AI Prompts

Everyday Humans

Midjourney (6.1)

A candid wedding moment of a couple at a wedding together under a canopy of leaves, evening reception vibe, the bride's flowing dress twirling as they dance, authentic emotions, photojournalistic style, high-resolution, sharp details, vibrant and warm tones, sony A7R

A vibrant street snapshot documentary style of two young women walking together in an urban park. They wear colorful, eclectic clothing with bold patterns and layered accessories, their hair dyed, and they carry unique bags and small items, capturing an alternative green background, documentary style, snapshot, dslr

A candid airport reunion scene, two people hugging deeply near the arrivals gate, one face visible slight tears, The busy terminal around them , with flight information screens and rolling luggage in the sharp background, handheld , documentary style, snapshot

Wedding photograph, unedited, couple walking down the aisle, people sitting on chairs blurry in the background, woman is smiling locking to the ground, men is lokking proud, in motion walking, sony a7R3

Dancers dancing on public square, surrounded by an audience of passersby some couples dancing closely, others laughing and spinning with flair. The square is paved with stones, framed by trees and historic buildings, sharp light, subtle colors Street photography, documentary style, sony, 50mm
public park in the distance a jogging middle aged man blue t shirt, exhausted look on face, sharp sunlight, documentary style, snapshot, 100mm, sony

Grok (March 2025)

a woman playing cello in a city, there are two people walking by. You can see a building in the background

a man reading a book. He is sitting on the stairs from a city church. The photo is taken from above, with a ray of light in his eyes.

A couple posing in front of the Eiffel Tower in Paris

a couple walking in a busy shopping street in Italy during the summer. They are a bit further away in the background and there is a tree, more in the front but to the side

Two friends posing in front of a Christmas tree in a German Christmas market

A teenager skateboarding at a skate park with ramps and graffiti-covered walls in the background.

an old white man, standing at a lake. It is cloudy and you can see mountains in the background

someone dancing on the street. Other people are passing by and minding their own business. The person is wearing hip skater clothes

A construction worker operating machinery at a urban construction site during the day.

A cyclist riding along a scenic coastal road with the ocean and cliffs in the background.

Landscapes

Midjourney (6.1)

A scene of the rugged Scottish Highlands, dominated by rolling hills covered in lush green and golden heather. Mist drapes over distant mountains, partially obscuring their peaks and an ominous, old castle. The sky is overcast with dramatic clouds, casting a moody, atmospheric light. Small rocky outcrops and scattered patches of wild grass add to the untamed beauty of the landscape. A soft breeze bends the tall grasses, and a faint glimmer of a loch can be seen in the distance. Cinematic composition, soft natural lighting, taken with a high-resolution DSLR camera.

A South German countryside in spring with rolling green hills stretching into the distance. Winding country roads weave through the landscape, flanked by lush meadows filled with blooming wildflowers in shades of yellow and white. Traditional Bavarian farmhouses with wooden balconies and red-tiled roofs sit nestled among the hills. Dense, dark green forests dot the scenery, contrasting with the bright fields. In the background, the misty foothills of the Alps rise gently, their peaks softened by a light haze. The warm afternoon sun casts a golden glow, highlighting the vibrant colors of nature. A peaceful,

idyllic atmosphere with clear blue skies and a few fluffy white clouds drifting above. Taken with a high-resolution DSLR camera.

A sun-drenched Mediterranean beach promenade, warm golden sand stretching along the coast. The turquoise waves gently lap against the shore, shimmering under the bright afternoon sun. Elegant, whitewashed buildings with terracotta roofs stand nearby, their balconies adorned with vibrant bougainvillea. Cozy cafés and seafood restaurants spill onto the promenade, with people strolling leisurely or enjoying espresso at outdoor tables. The salty ocean breeze carries the scent of grilled sardines and citrus. In the distance, rugged cliffs and rolling hills frame the coastline, creating a perfect harmony between nature and charming seaside life. Shot with a high-resolution DSLR camera

Scandinavian coastal village nestled along a rugged, rocky shoreline. Small wooden houses painted in vibrant red, yellow, and white stand against the deep blue sea. Jagged rocks and smooth, weathered stones line the coast, where fishing boats are moored near wooden piers. The sky is a mix of soft clouds and clear blue, with the golden light of the afternoon sun casting a warm glow over the scene. In the distance, rolling hills and small islands dot the horizon, creating a tranquil, idyllic Nordic atmosphere. Shot on Sony a7r iv, macro lens, fullframe.

editorial landscape photography, side on view, a single cabin in a snowcovered minimalist landscape, winter's isolation, icy blues, pure whites, shot on sony alpha 1, macro lens, apsh, diane arbus style, overcast, snowy day, open field, cozy, secluded, fresh snow, barren trees, unblemished, solitude, winter color, art nouveau, snow overlay, freeze motion, color isolation

A vast, untouched Siberian wilderness stretching endlessly under a pale winter sky. Snow-covered taiga forests with towering evergreen trees dusted in frost stand beside a frozen river, its surface cracked with icy blue veins. Rolling tundra extends to the horizon, bathed in soft, diffused light. Mist drifts over the landscape, adding a mysterious, ethereal atmosphere. In the distance, jagged mountains rise, their peaks hidden in a veil of icy fog. The air feels crisp and silent, capturing the raw beauty of Siberia’s remote and unforgiving nature. Cinematic composition, taken with a high-resolution DSLR camera."

A tranquil sacred grove deep in the Japanese countryside, surrounded by towering ancient cedar and blooming cherry blossom trees. Soft pink petals drift through the air, settling on a moss-covered stone

path that winds through the forest. Sunlight filters through the delicate sakura branches, casting warm, dappled light on the ground. A small, weathered Shinto torii gate stands quietly among the trees, partially covered in climbing ivy. In the distance, a tiny wooden shrine with faded red paint blends seamlessly into nature, its paper lanterns gently swaying in the breeze. The air is filled with the sweet fragrance of cherry blossoms and damp earth. A lone stone water basin, covered in green moss, reflects the stillness of the grove, evoking a deep sense of harmony and spiritual serenity. Shot on Sony a7r iv, macro lens, fullframe.

A vast North African landscape bathed in warm, golden sunlight. Rolling sand dunes stretch endlessly into the horizon, their curves shaped by the desert wind. In the distance, rugged, rocky plateaus and jagged mountains rise under a brilliant blue sky. Scattered Berber tents and ancient mud-brick villages cling to the hillsides, their earthy tones blending seamlessly with the desert. The air is dry and hazy, with the occasional dust cloud drifting across the horizon. High-resolution DSLR mirrorless camera, 300mm lens.

A secluded Caribbean cove with a small, hidden beach nestled between rugged cliffs covered in lush green vegetation. The turquoise waters gently lap against the shore, creating a gradient from deep blue to crystal-clear near the sand. Sunlight reflects off the water, casting shimmering patterns on the rocky coastline. Sparse, scattered houses sit atop the cliffs in the distance, barely visible through the dense foliage. The atmosphere is peaceful and untouched, with only the sounds of the waves and rustling leaves in the breeze. Cinematic, aerial view, shot on high-resolution DSLR camera, 300mm lens

A sweeping panoramic vista of the Ural Mountains, dramatic mountainous landscape with purple-blue peaks extending into the distance, lush green coniferous forests at the base, winding turquoise river cutting through a valley floor, bright blue sky with scattered white fluffy clouds, afternoon sunlight illuminating the slopes, crisp high-resolution photography style, wide-angle lens, vibrant natural colors.

cinematic, evergreen forest in afternoon, wide-angle landscape perspective with a low to mid-level camera angle, sun gently lighting up the scenery through the thick canopy, shot on sony a7r iv, macro lens, fullframe, tranquil atmosphere, firn trees, casual, lively, soft focus, pastel shades, bokeh, lens flare, soft filter

A highly detailed, photorealistic image of a quiet lakeshore at dawn, taken with a professional DSLR camera using a 50mm lens. The foreground focuses on smooth, damp pebbles, partially submerged in the shallow water, with soft ripples gently lapping against them. A thin layer of morning mist hovers just above the still lake, gradually dissipating as the first light of the rising sun breaks through the treetops in the background. The lake reflects the warm hues of the sky, blending soft oranges and pale blues in perfect harmony. Distant pine-covered hills line the horizon, their dark silhouettes contrasting subtly with the glowing morning light. A few fallen leaves float on the water's surface, slightly curled at the edges, hinting at the early days of autumn. The air feels crisp and still, with no artificial enhancements--just the simple, raw beauty of nature captured in perfect clarity. The image features natural imperfections such as slight lens haze, subtle noise in shadowed areas, and organic depth of field, ensuring it is indistinguishable from a real photograph.

A crisp, photorealistic autumn morning in a quiet countryside field, captured with a professional DSLR camera using a 50mm lens. The foreground showcases frost-covered grass blades, glistening under the soft golden light of the early sun. A narrow dirt path, slightly damp from morning dew, winds gently through the field, bordered by wooden fence posts with peeling paint and tangled vines. In the middle ground, a small, still pond reflects the muted blue sky, with a few gentle ripples caused by a passing breeze. A single tree stands near the water, its sparse leaves in shades of orange and yellow, some drifting slowly to the ground. Beyond the pond, rolling hills covered in a patchwork of fields and clusters of trees fade into a light morning mist, adding soft atmospheric depth. The sky is clear but with a few wispy clouds stretching across the horizon. The image has natural imperfections, such as slight lens haze in the distance, tiny specks of dust catching the sunlight, and subtle variations in color temperature, making it indistinguishable from a real photograph.

A crisp, photorealistic autumn morning in a quiet countryside field, captured with a professional DSLR camera using a 50mm lens. The foreground showcases frost-covered grass blades, glistening under the soft golden light of the early sun. A narrow dirt path, slightly damp from morning dew, winds gently through the field, bordered by wooden fence posts with peeling paint and tangled vines. In the middle ground, a small, still pond reflects the muted blue sky, with a few gentle ripples caused by a passing breeze. A single tree stands near the water, its sparse leaves in shades of orange and yellow, some drifting slowly to the ground. Beyond the pond, rolling hills covered in a patchwork of fields and clusters of trees fade into a light morning mist, adding soft atmospheric depth. The sky is clear but with

a few wispy clouds stretching across the horizon. The image has natural imperfections, such as slight lens haze in the distance, tiny specks of dust catching the sunlight, and subtle variations in color temperature, making it indistinguishable from a real photograph.

A serene, photorealistic late afternoon scene in a vast open grassland, captured with a high-end DSLR camera using a 35mm lens. The foreground features tall, golden prairie grass swaying gently in the breeze, individual blades catching the soft sunlight. A well-trodden dirt trail cuts through the grass, leading toward a distant, lone oak tree standing against the expansive horizon. The rolling hills in the background stretch far, their subtle contours fading into a light atmospheric haze. The sky is a soft gradient of pale blue with sparse, wispy clouds tinged with warm hues from the setting sun. Shadows grow long, creating a natural contrast that enhances the depth and realism of the scene. A small flock of birds is visible high in the sky, moving lazily in the distance. The image is perfectly balanced, with natural imperfections such as slight lens flare from the sun, tiny dust particles floating in the warm air, and gentle motion blur in the windblown grass, making it feel like an authentic, untouched photograph.

A serene, photorealistic winter scene of a small wooden cabin in the middle of a snow-covered field, captured with a 50mm lens. The cabin, simple and rustic, sits alone in the center of the frame, its roof blanketed in fresh snow, with a thin trail of smoke rising gently from the chimney. The surrounding snow is untouched, soft and powdery, reflecting the pale, cool light of the overcast sky. The area around the cabin is empty, with just a few scattered snow-covered bushes and the distant outline of a forest at the edge of the field. The sky above is cloudy, casting soft, diffused light that creates long, gentle shadows on the snow, highlighting the textures of the frost and the simple lines of the cabin. The air feels crisp, and a few snowflakes are gently falling, adding to the tranquility. The image includes natural imperfections like a soft haze, slight lens blur in the distance, and fine details of snow drifts around the cabin, making it feel like a real, peaceful winter moment.

A serene landscape at golden hour, featuring rolling hills covered in lush green grass, a calm river reflecting the warm hues of the sunset, scattered wildflowers in the foreground, and a few fluffy clouds in a clear blue sky. The scene should capture the natural beauty and tranquility of the moment, with soft lighting and realistic textures, resembling a high-quality photograph.

Grok (march 2025)

A breathtaking view of snow-capped mountains at sunrise, with a clear blue sky and a serene lake reflecting the peaks, hyper-realistic, high detail.

Make a landscape picture as if it was made by a human.

Make a landscape picture as if it was made by a human. National geographic style.

Make a photorealistic landscape picture as if it was made by a human. National geographic style.

Make a landscape picture.

Make a landscape picture.

Make a landscape picture.

Make a landscape picture.

Make a landscape picture.

Make a landscape picture.

Artistic Humans

Midjourney (6.1)

A hyper-realistic, artistic studio portrait of a woman with deep brown skin, her face partially covered in shattered gold leaf, creating a striking contrast against her smooth complexion. The lighting is moody and directional, with a single spotlight casting dramatic highlights on the gold while leaving parts of her face in deep shadow. The background is a soft, velvety black, fading into a subtle gradient. Her expression is powerful yet introspective, her gaze slightly averted, as if lost in thought. Tiny gold flakes appear to be floating in the air, catching the light in a way that feels almost surreal, yet completely realistic. The fine details--pores, subtle skin texture, the delicate edges of the gold leaf--are captured with astonishing clarity, making this image feel like a masterfully staged high-fashion art photograph.

A striking, artistic studio portrait of a woman with short, sleek black hair, dressed in a simple, elegant black dress. The lighting is dramatic, with sharp contrasts casting deep shadows and highlighting the graceful curves of her face and neck. The background is a soft, monochrome gradient, adding a sense of timeless elegance and focus on her intense, contemplative gaze.

Avant-garde studio portrait of a human figure, ethereal and experimental, bathed in shifting veils of colored light from a cracked stained-glass panel overhead, surrounded by a chaotic arrangement of floating gauze strips and charred branches, muted palette of frost blues, ash grays, and burnt corals with organic gradients, subject wrapped in frayed translucent fabric or crowned with twisted wire, natural posing with subtle flaws like smudged makeup or tangled hair, hyper-realistic skin under harsh spotlight glare, raw and unpolished texture, 8k resolution, mimics a daring human-photographed art piece.

Create a hyper-realistic artistic portrait of a single human figure in an exceptionally creative setting. Incorporate surreal elements such as flowing fabrics, abstract shapes, or vibrant colors that interact with the subject. Experiment with dynamic poses that convey emotion and movement, and use unique props or artistic backdrops that enhance the overall composition. Focus on lifelike skin textures, intricate facial details, and authentic expressions. Utilize dramatic lighting and soft shadows to create depth and dimension, ensuring the final image is a stunning blend of artistry and realism, indistinguishable from human-made studio photography.

An artistic, high-fashion portrait of a woman standing in a studio, her pose a striking blend of elegance and movement. She is slightly bent forward, with her body arched, her arms extended as if reaching out to grasp something just beyond her reach, her fingertips delicately touching the air. Her head is tilted to one side, eyes focused downward with a contemplative, almost mysterious gaze. Her hair, styled in sleek waves, flows in a way that suggests wind or motion, despite being perfectly still. The lighting is dramatic, with a single spotlight highlighting her face and upper body, casting sharp shadows across her form, while the rest of the image fades into shadowy abstraction. Behind her, there is a backdrop of rippling, metallic fabric that seems to shimmer with hints of silver and copper, its texture evoking both fluidity and solidity. The scene is further enhanced by a subtle, reflective surface below her, where the silhouette of her body is distorted, adding a layer of surrealism to the realistic portrait. The overall composition balances tension and serenity, with a hint of surreal elegance, as though the woman exists both in the real world and an ethereal, otherworldly space.

A bold and artistic studio portrait of a confident, curvy model with striking red hair styled in a vintage updo. The model is wearing elegant, white with details, accompanied by sheer, flowing blue tulle draped like a veil. Tattoos on her arms and legs visible, adding an edgy and expressive look. The background is moody and atmospheric with light, soft clouds and subtle lighting. High-fashion editorial style, dramatic and creative composition.

An artistic, high-fashion studio portrait of a woman with striking silver hair, styled in a dramatic, asymmetrical cut, wearing a futuristic metallic outfit that gleams under the studio lights. The background is a deep, reflective black, with abstract geometric shapes subtly illuminated by soft, neon lights that cast a vibrant, colorful glow on her face. Her makeup is bold, with striking neon eyeliner and a shimmering highlight on her cheekbones. Her posture is strong, yet graceful, with one hand lifted slightly, as if reaching for something beyond the frame. The lighting is experimental, with sharp contrasts and bold highlights, creating intricate reflections and shadows on her metallic outfit. The atmosphere is sleek, modern, and a little otherworldly, as if she's a figure from a future art exhibition, captured in an expertly staged, surreal moment of elegance and strength.

Studio portrait of an adrogynous person with glass shards reflecting rainbow colors, prismatic light scattering across their face, soft-focus and ethereal ambiance, abstract and expressive, futuristic and artistic vibe

photograph of an old white man, front view, portrait, closed eyes, full face, standing at the edge of a serene lake, gazing into the vast natural landscape, The man, wearing casual outdoor clothing, stands with a casual posture, amateur snapshot, 50mm, documentary style

Studio Photography stylish plus-size asian model posing wearing an elegant, flowing dress , illuminated by colorful, artistic lighting in shades of pink, blue, and white, white studio background, extravagant hair style, The model's expression is poised and radiant, and the composition highlights body positivity and high-fashion energy. Sharp focus, soft shadows, and a polished, editorial-style aesthetic, vogue cover, lgbtq

Professional studio photography of skinny man, tattoos on upper body, dramatic look, strong blue lighting, curly hair , sony a7R

A striking, artistic studio portrait of a woman with short, sleek black hair, dressed in a simple, elegant black dress. The lighting is dramatic, with sharp contrasts casting deep shadows and highlighting the graceful curves of her face and neck. The background is a soft, monochrome gradient, adding a sense of timeless elegance and focus on her intense, contemplative gaze.

A high-fashion studio portrait of a poised young woman with makeup in shades of pink, blue, and black, featuring a soft matte complexion and glossy lips. She wears large, dangling star-shaped earrings encrusted with gems, adding a luxurious feel. Her hair is sleek with subtle color highlights at the tips. The subject is dressed in a delicate, pleated white high-collar blouse, exuding an ethereal elegance. Shot with a Hasselblad H6D-100c, 100mm lens, f/4creating a clean white background with a halo glow effect around the edges

A high-fashion black-and-white studio portrait of a man with splashing water, flash photography movement in the dark, her face partially illuminated by soft, diffused lighting. Dramatic monochrome contrast highlights her bold facial features and intricate textures of water droplets cascading down skin Shot with a Hasselblad H6D-100c, 100mm lens, f/4, capturing every fine detail in stunning clarity, a mist of water droplets suspended in the air, creating an ethereal and cinematic atmosphere, water splashes flash photography

Studio photograph of a young woman posing gracefully in a 1920s flapper dress with intricate beading and fringe. She wears a stylish feathered headband and dark, dramatic makeup with bold red lips. The studio lighting is soft and moody, casting vintage-style shadows. The background features an Art Deco-inspired design with gold and black tones, evoking the glamorous atmosphere of the roaring twenties Leica M6 (35mm Film)

A bold, high-fashion portrait of a young woman in a metallic silver dress with sharp, geometric patterns. Her makeup is abstract, featuring neon accents and glossy, iridescent lips. The lighting includes vibrant, colored gels casting blue and purple hues on her face, glow shot on Hasselblad

Fashion-focused studio photograph of a woman in traditional Bedouin clothing, striking a poised pose, sharp lighting to emphasize the details of the outfit, neutral background to keep attention on the subject
A high-resolution studio portrait of a woman standing in front of a pure white backdrop. The composition is clean and minimalist, with soft, diffused lighting creating gentle shadows. Ethereal light prism effects refract around the subject, adding subtle rainbow hues. The model's expression is serene, evoking a sense of calm and elegance. Shot with a high-end camera, ultra-sharp details, and cinematic quality

Grok (march 2025):
A hyper-realistic, artistic studio portrait of a woman in an unconventional pose--her body slightly twisted, one arm elegantly raised above her head, fingers gently curved as if reaching for something unseen. She wears a flowing, semi-transparent silk fabric that wraps around her body, caught in mid-motion, as if frozen in time. Her expression is serene yet intense, her eyes half-closed, lips slightly parted as if in a deep moment of thought or emotion. The lighting is dramatic and moody, with a single warm spotlight casting intricate shadows across her face and body, while a subtle cool backlight traces the edges of her form, adding depth and dimension. The background is minimalist, a smooth, muted gradient that fades into darkness, enhancing the focus on her form. The details--soft skin texture, the gentle tension in her fingers, the natural creases in the fabric--are captured with exquisite realism, making this feel like a meticulously crafted, high-end fashion or fine-art studio photograph