

Unit 2 MidPoint Writing

Unit 2 MidPoint Writing

Components 01
P3-11

Components 02
P12-33

Components 03
P34-49

Bibliography
P50

Unit 2 MidPoint Writing Components 01

Components 01

I began with cropped smiling faces from a Mattel toy catalogue, feeding five images into 20 different reverse image search systems (see Figure 1). Each system understood the same image differently – shopping platforms transformed it into a product, social media into a marker of community belonging. As the image circulated, the authority to define it shifted with each passage, never fixed.

The unexpected result came from YOLO, which identified a panda's eye as a cup – because a hollow structure with a border fits its definition of one. This made me realise that the machine does not read objects; it reads geometric combinations of

pixels. Based on this, I built a website (see Figure 2) that randomly recombines panda facial shapes and feeds each into YOLO. Shapes that looked entirely different received the same label. Regions appearing blank to the human eye were named by the machine.

These results point toward the same conclusion: machine classification has completely diverged from the way humans see. The machine is not seeing wrongly – it is defining the world through a language humans never authorised, yet already fully in operation. When that definition occurs before human perception does, does the question of what an image is still belong to us?

Figure 1. 5 Faces, 20 Systems, 100 Answers – website experiment results.

(a) YOLO identifying a panda's eye as a cup recorded in publication

(b) Publication

Source: Author's own.



(a)



(b)

#Position through Iterating

Figure 2. 5 Faces, 20 Systems, 100 Answers – website screenshots.

(a) Poster of the Website's result

(b) Poster of the Website's result

Source: Author's own. Available at: [YouTube](#)(Accessed: 31 May 2025).



(a)



(b)

#Line of inquiry

#5

Components 01

//Reading List

Reference 01

Maurer, L. (2013) Conditional Design Workbook. Amsterdam: Valiz.

Across 100 iterations, I repeated the same action: feeding the same snippets into the image-recognition tools of different media, one after another. This repeated action formed a kind of rule. When the same image is interpreted by 20 different systems and those interpretations placed side by side, the influence of the medium is amplified enormously.

In the second phase of 50 iterations, I selected one snippet, drew it by hand, and fed it back into YOLO v8. The rules built into the model define what can be "seen" and what is destined to be ignored. These reassembled images – drifting outside normal visual grammar – kept pressing against the model's internal logic, revealing a question: what kind of image counts as "below average" to a system like this?

Conditional Design takes an optimistic stance toward rules, where creativity begins and accidents emerge. As Maurer writes, "Instead of solutions for problems, programmes for solutions" (Maurer, 2013, p.3), rules do not point toward a single answer, because possibilities cannot be absolutely delimited. In my experiment, I was trying to make these scattered, reassembled images evade the system's recognition, turning the system itself into a tool that deepens alienation.

Components 01

//Reading List

Reference 02

McLuhan, M. & Fiore, Q. (1967) The Medium is the Message

In this experiment, I enlarged recognisable toy images and cropped into their details, then fed these fragments into different "containers" – observing how each system recognised, interpreted, and categorised them according to its own internal logic.

The medium shapes the message, rather than merely delivering it (McLuhan & Fiore, 1967). In this experiment, images behave like liquid – the container determines the shape in which they are perceived. The first 19 containers were all software or browsers, each interpreting the image according to its own functional logic and assumed user profile. For a fashion app's visual search tool, infant toys fall outside the scope of its database, leaving it unable to find any match in function or form. The only anchor it could fall back on was colour. This "out-of-place" image pushed the system to its limit, exposing the cognitive boundary of the medium itself.

Components 01

//Medium

Reference 03

Steyerl, H. (2009) 'In Defence of the Poor Image', e-flux journal, 10

The snippets cropped and enlarged from Mattel toy packaging are, in their nature, already poor images. And with each new system they pass through – emotion recognition software, reverse image search tools, Python models – they lose something, while being assigned a new definition. This is a simulated enactment of the poor image's logic of circulation and transformation.

Through this process of transformation, the snippets can no longer be traced back to the original toy image – they have become something else. This confirms what Steyerl identifies as the loss of originality: as an image circulates, it loses its source, and with it the authority and centrality that source once carried. The Mattel toy images, once produced for commercial purposes and aimed at a specific consumer audience, no longer carry any brand value or market context after being cropped and appropriated. They have become drifting objects.

Components 01

//Topic

Reference 04

Baudrillard, J. (1981) Simulacra and Simulation

The act of intervening in an image and setting it into dialogue with a machine is itself a process of simulacra production. The "mean image" is a simulacrum – built and refined through the training of vast quantities of images, it is a phantom that always gravitates toward the safest answer. A model like YOLO v8 is a container assembled from countless such simulacra.

When fragments stripped of their original context – detached signifiers – are fed into this container, the system begins to break down. The circulation of simulacra that Baudrillard describes is here blocked by the machine's cognitive boundary. The difference is this: humans can follow the drift of floating signs and actively assign new meaning, but computer systems remain anchored to a historical understanding of signs that humans have long since moved past. That anchoring exposes the symbolic reality we currently inhabit – one where we are already speaking a new grammar, while the machine is still reading an old dictionary.

Components 01

//Wild Card

Reference 05

. Lakoff, G. (1987) *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*

Cognitive linguist George Lakoff argues that categories are not defined by clear boundaries, but are structured as a gradient organised around a prototype (Lakoff, 1987, p. 5). When we think of "bird," a sparrow comes to mind before a penguin – not because a penguin is not a bird, but because it sits too far from the prototype.

Image-recognition models operate by the same logic, and in fact push that gradient structure further toward rigidity. When the panda's most distinctive visual feature – its signature black-and-white eye patches – is deliberately disrupted, the distance between the image and its prototype widens sharply, and the system promptly determines that it no longer belongs to any category, ejecting it from the gradient entirely. Lakoff's framework reveals the cognitive nature of classification; this experiment makes that nature observable: the edges of a category are something that image fragments can press against, and test.

Components 01

//Critical Position

Reference 06

Offert, F. & Bell, P. (2021) 'Perceptual bias and technical metapictures',
Critical AI

As the panda image was gradually broken apart, the model's ability to recognise it steadily declined – because the image was drifting further and further from the "average panda" embedded in its training data. The model never truly understands or dissects an image; it only asks: what does this most resemble among everything I have seen? As Offert and Bell argue, machine vision systems are inherently biased not only because of their datasets, but because of their "perceptual topology" – their specific way of representing the visual world (Offert and Bell, 2021, p.1133).

Where do these models reach their limit? This was the question I kept asking while working with them. These image-recognition systems are naturally resistant to the unusual, the marginal, the recontextualised image. And the snippets in this experiment – those visual fragments that have been cropped, appropriated, and repeatedly redefined – are precisely what this kind of system is least equipped to handle. At its core, this experiment uses "below-average" images to test a system that only knows how to think in averages.

Unit 2 MidPoint Writing Components 02

Components 02

Analysis of Practice

Diptychs of Human and Machine
Perceptions

By Vivien Cabannes, Thomas Kerdreux
and Louis Thiry

Components 02

Diptychs of Human and Machine Perceptions

By Vivien Cabannes, Thomas Kerdreux and Louis Thiry

This work compares what humans and machines focus on when looking at the same image. The authors took an original image and produced two versions from it: one based on machine vision, one based on human vision. On the left is the original image. In the middle is the machine's version – a saliency map, which visualises which regions the neural network considers most important when processing the image. On the right is the human version, generated by tracking the eye movements of multiple viewers and turning the areas where people looked most frequently into a heatmap overlaid on the original. So the "human version" is not how one person looked at the image – it is an aggregate of many people's gaze data, showing where people collectively tended to

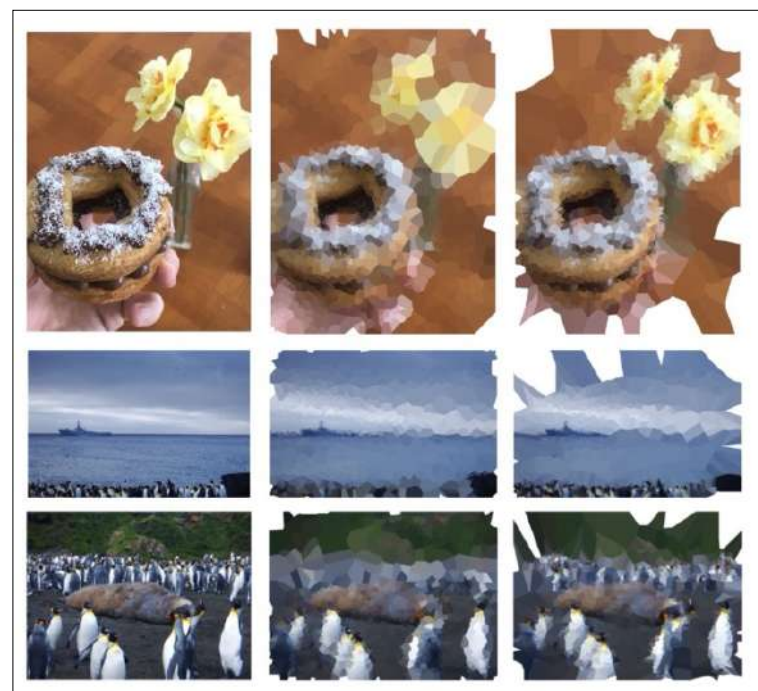


Figure1

Components 02

Diptychs of Human and Machine Perceptions

By Vivien Cabannes, Thomas Kerdreux and Louis Thiry

look. At its core, the aim is to use visual art to reveal the differences between human and machine perception. But for me, it also resonates strongly with Hito Steyerl's concept of the "mean image." The collective human gaze and the machine's identified "important regions" turn out to be surprisingly similar – and yet the machine's version is cruder and less nuanced than the path drawn by human eyes. This exposes a fundamental limitation of machine vision: what the machine produces is not a picture of a penguin, but a cluster of pixels that best matches the label "penguin" in its training database. It doesn't see an animal. It finds a statistical average.

Even without sophisticated graphic design, I think this work succeeds as a design reference – because it builds a clear and precise system that makes an abstract phenomenon visible. Graphic designers are often described as "problem solvers," but that definition tends to imply giving answers and making problems disappear. What this work does instead is make the problem visible. And I think that is equally valid as a form of design thinking.

What I find particularly useful is the structural logic behind it. The authors set up a reasonable tool, chose comparable objects, and established a set of rules. What followed was a repeatable process – producing pair after pair of images – and the accumulation of that process itself becomes the output that frames the phenomenon of the "mean image." This encouraged me to think about designing a system with similarly precise rules, where the process of repeating that system is itself what surfaces the question. It also shaped the way I think about the choice of container. These images exist primarily as figures inside an academic paper. That context makes the argument clear and allows the details of the experiment to be understood precisely – but it also limits the reach of the work. Kept in a supporting role, they could have stood on their own. If they had been brought together into a standalone publication, their impact and circulation

Components 02

Diptychs of Human and Machine Perceptions

By Vivien Cabannes, Thomas Kerdreux and Louis Thiry

could have been much greater. This made me think more carefully about my own project. Rather than deciding on a single format early on, I tried presenting the same experiment through different containers – animation, website, score, CD – in order to find the one that best allows the core question to be felt, not just understood. Because the medium is not just a delivery tool. It is part of the argument itself.

Figure 2. Machine vision website.

(a) Website interface screenshot.

(b) Website detection output screenshot.

Source: Author's own work (2026). Available at: [YouTube](#)(Accessed: 31 May 2026).



(a)



(b)

Figure 3. Screenshot from the machine vision website.
Source: Author's own work (2026). Available at: [YouTube](#) (Accessed: 31 May 2026).

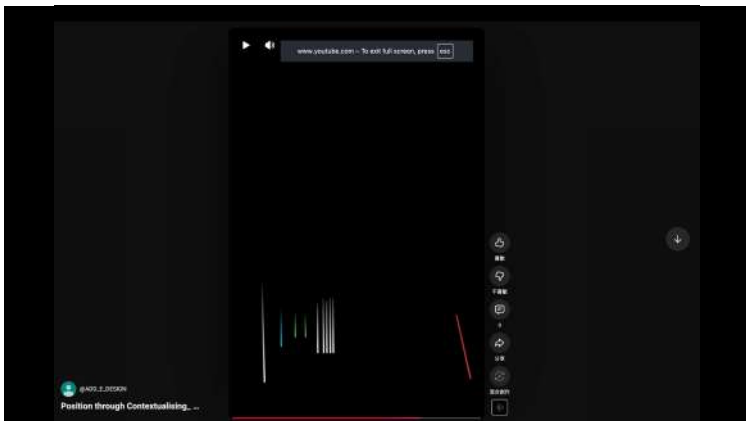


Figure 4. CD documentation of the machine vision project.
Source: Author's own work (2026).



Components 02

Analysis of Text

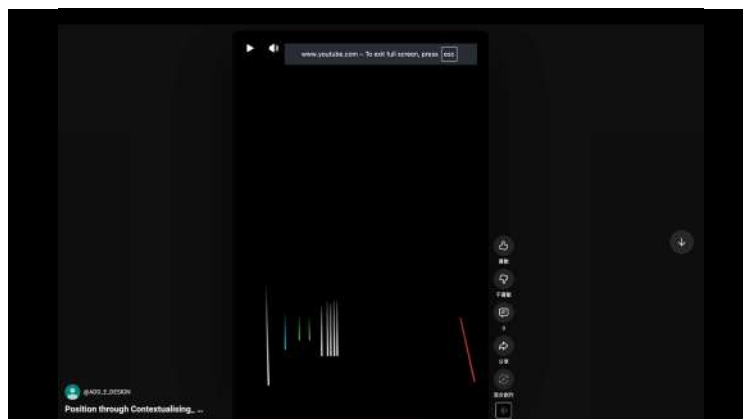
Invisible Images
(Your Pictures Are Looking at You)
By Trevor Paglen

Components 02

Invisible Images (Your Pictures Are Looking at You) By Trevor Paglen

This essay centres on a question: when images no longer exist primarily for human eyes, how can we use images as a form of resistance? In the opening section, Paglen makes this point through a simple but striking example. When a person picks up a phone to take a photo, they need a screen to see the image – but the machine doesn't. The moment the camera captures the image, it is already translated into data readable by the machine. In this sense, humans have begun to lose authority over the act of seeing (Paglen, 2016). In the final section, Paglen reflects on how visual culture has historically been used as a tool for resistance – to protect rights, to speak to social justice, to make re-representation a weapon. But all of those strategies assumed one thing: that images were made primarily for people. Now, with machines in the picture, images circulate between systems without human awareness (Paglen, 2016). They have become something that looks back – analysing, classifying, and

Figure 1. Screenshot from the machine vision website.
Source: Author's own work (2026). Available at: [YouTube](#) (Accessed: 31 May 2026).



Components 02

Invisible Images

(Your Pictures Are Looking at You)

By Trevor Paglen

surveilling us. Images are no longer passively observed. They actively participate in reality, shape behaviour, and structure power relations. Paglen's proposed response to this feels almost resigned: to create spaces of "inefficiency," areas of life removed from market logic and political capture – safe houses in the digitally invisible world. Many of the examples in this essay reshaped the way I understand visual culture in the digital age. For example, in the essay, he mentioned that a company named Vigilant Solutions and their ALPR system – an automated licence plate recognition database that tracks the movement of vast numbers of people. When a government hands over a list of people with unpaid fines, police cars equipped with the system can automatically locate them. The moment a plate is scanned, officers can pull the vehicle over immediately, and the person is given a choice: pay on the spot or be arrested. On top of that, a 25% processing fee goes directly to the company. The system presents itself as objective, automated, free from human bias – but what it actually encodes is government power, police surveillance, corporate profit, and the systematic targeting of vulnerable people. The machine is not neutral. It is a very efficient vessel for existing structures of control. (Paglen, 2016)

And I think the questions it raises apply directly to graphic design. Every major design movement – Modernism, Swiss Style, Postmodernism – was made for a human audience. Now that there is a new kind of audience, I find myself asking: could graphic design become complicit in building a machine-readable image system that appears neutral but is in fact deeply hierarchical? And if so, what does it mean to design against that – to make something that cannot be entered by the machine?

In my project, I translated a painting into animation using the machine's own reading logic. (see Figure 1) In some ways, this is a simulation of how the machine reorganises its perception of the world through images. In that

Components 02

Invisible Images

(Your Pictures Are Looking at You)

By Trevor Paglen

translation process, the machine ignores the sonic imagination of each element, ignores the mood produced by the overall composition, and ignores everything this painting represents – abstraction, spirituality, emotion, modernist history. It only reads, and can only read, pixel trajectories. This aligns precisely with what Paglen describes: "this idea of ambiguity, a cornerstone of semiotic theory from Saussure through Derrida, simply ceases to exist on the plane of quantified machine-machine seeing." (Paglen, 2016)

The next direction for this project is to bring that overflow into visibility – the ambiguity, emotion, and symbolism that the machine systematically discards. If Paglen's safe house is a space the machine cannot enter, then the question I want to ask as a designer is this: can I use image, sound, and the choice of medium to build that space? Not to fight the machine's logic, but to construct a safe house from the very things the machine has erased.

Components 02

Reference 01

Maurer, L., Paulus, E., Puckey, J.
and Wouters, R. (2013) Conditional
Design Workbook

In Conditional Design, Luna Maurer and her collaborators identify a ghost haunting any rule-based system: "the specter of a totalizing, closed, and rationalized system." (Maurer et al., 2018, p.3). Reflecting on my own system (the website), I recognise that the same specter lingers within it, quietly narrowing the range of possibilities that translation might otherwise hold. This became particularly evident during the second week of experimentation. When I attempted to translate the act of looking at a painting into two musical versions, one based on human observation and the other on machine observation, the resulting sounds were nearly indistinguishable from one another. This outcome prompted me to reconsider my approach after receiving feedback: had I become too preoccupied with differentiating between the two kinds of observers, to the point of imposing a rationalized framework onto a process of translation that it did not suit? This insistence on systematization and quantifiability had, in effect, closed off the more varied and open-ended possibilities that the project might have otherwise explored.

In response to this reflection, during the third week I shifted direction, turning to a webcam as a means of capturing human looking, in search of a method more attuned to bodily experience and open to contingency.

Components 02

Reference 02

McLuhan, M. and Fiore, Q. (1967)
The Medium is the Message: An
Inventory of Effects

McLuhan's argument that the medium is the message prompted me to reconsider the translation at the heart of this project. In the second week, I translated data extracted from a painting into two pieces of music, each following the viewing sequence of a different observer. But when visual sequence was converted into auditory form, the sense of sequence became weakened. Music conveys information through auditory rhythm, but this logic does not naturally carry visual order. The medium had already begun to quietly rewrite the message.

This led me to introduce animation. Moving image conveys information through visual rhythm, structurally analogous to the auditory rhythm of music. When combined, visual and auditory rhythm lock onto each other, allowing the fixed, line-by-line atmosphere of the machine scanning pixels to be conveyed with greater precision. The timing of each line's appearance, its trajectory, and the way sound enters together form an inescapable temporal structure, within which the viewer can only perceive the image on predetermined terms. This is precisely what McLuhan means: the medium is not a neutral container; it participates in and shapes the message. When two media are layered together, their shared mode of perception makes the transmission more exact.

Components 02

Reference 03

Benjamin, W. (1968) Illuminations

In this experiment, I enlarged recognisable toy images and cropped into their details, then fed these fragments into different "containers" – observing how each system recognised, interpreted, and categorised them according to its own internal logic.

The medium shapes the message, rather than merely delivering it (McLuhan & Fiore, 1967). In this experiment, images behave like liquid – the container determines the shape in which they are perceived. The first 19 containers were all software or browsers, each interpreting the image according to its own functional logic and assumed user profile. For a fashion app's visual search tool, infant toys fall outside the scope of its database, leaving it unable to find any match in function or form. The only anchor it could fall back on was colour. This "out-of-place" image pushed the system to its limit, exposing the cognitive boundary of the medium itself.

Components 02

Reference 04

Berger, J. (1972) Ways of Seeing

Berger argues in *Ways of Seeing* that when the presentation of an observed object changes, the perception of the observer changes with it (Berger, 1972). This project inverts that proposition: when the observer itself undergoes a fundamental change, what happens to the object being looked at?

With this question in mind, I brought the machine into the position of observer. By translating a painting through the machine's reading logic, I wanted to know: what does the machine's perception sound like when rendered as rhythm and sequence? How does it differ from the rhythm produced by a human looking at the same painting? And how do these two fundamentally different modes of observation begin to shape new ways of seeing?

Components 02

Reference 05

Sterne, J. (2003) *The Audible Past*

In the process of iterating through different media, I moved through several stages – animation, website – before starting to think about how to let the two pieces of music stand on their own. My initial instinct was to release them as two digital albums on Spotify. But after reading Jonathan Sterne's work on the media of sound transmission, I changed direction.

Sterne's book was written before algorithm-based music platforms became dominant, so the media he discusses feel comparatively traditional by today's standards. But that's precisely what drew me to them. What I'm translating in this project is the act of looking at a painting – something fluid, subjective, and unrepeatable. That act shouldn't be categorised, recommended, or personalised by an algorithm. Spotify would do all of those things. So I began to feel that a container that does nothing was actually more appropriate – one that simply holds the translation, without doing anything to it.

Components 02

Reference 06

Kandinsky, W. (1947)

Point and Line to Plane

Kandinsky's Point and Line to Plane is the theoretical foundation of this experiment. His argument that visual elements carry inherent sonic properties gave me the framework for mapping each element to an instrument – planes to synthesizer, lines to electronic strings, and the angle of diagonal lines mapped onto a continuous pitch bend. But more importantly, his idea that a composition feels like music because of the way human eyes move through it made me realise that the time structure of a painting is inseparable from who is looking at it. This is what pushed me to ask the central question of this project: what happens when the observer is a machine?

Components 02

Reference 07

Sbordoni, A. (2021) Beyond the Image: Visual Culture in the Twenty-First Century

Alessandro Sbordoni uses the Mona Lisa to discuss how images become “devalued” once machines become new observers (Sbordoni, 2021, p.42). Within the context of machine vision, images are no longer understood through their historical or artistic meaning, but are instead broken down into recognizable and computable symbols and data. Even absurd recognition results – such as the incorrect labels from my previous project – still reveal how machines reconstruct their own systems of viewing through algorithmic logic.

This led me to question how these “new observers” construct their own symbolic systems. I began with the most basic unit of the digital image: the pixel – a form of data that can be classified, calculated, circulated, and transmitted. Once an image is translated into data, it also enters the field of machine vision. This way of thinking later became the foundation for my translation of Wassily Kandinsky’s paintings.

Components 02

Reference 08

Crary, J. (1990) *Techniques of the Observer: On Vision and Modernity in the Nineteenth Century*

Crary argues that the act of looking is historically constructed. People in the Renaissance, in the age of photography, in the age of cinema, in the age of the screen – they all see differently, because every era's technology reshapes how the observer is supposed to look. For me, this is less a direct methodology than a conceptual extension: if observers have been shaped by technology since the nineteenth century, where does that historical logic lead when the observer itself becomes a machine? And what does the observed object become? I focused these questions on the specific act of looking at a painting, and by changing who the observer is, attempted to translate the result of that looking – the observer's perception of the painting – into music.

Components 02

Reference 09

Yarbus, A. L. (1967) Eye Movements and Vision

In the second week of the experiment, I forced the machine's spatial reading order onto the human version by simply reversing the axes – using the image's x-axis as the time axis instead of the y-axis. The two pieces of music that resulted sounded remarkably similar. The approach appeared to be exploring the difference between human and machine reading order, but it didn't actually reflect how humans look at a painting. It was a contrast that existed for the sake of contrast.

Reading Yarbus's research on eye movement made me realise how fundamentally wrong this setup was. It violated the way humans actually look at images – we don't scan linearly, we jump, we linger, we skip. So in the third week, I introduced a webcam to directly capture each viewer's individual looking path, rather than imposing a predetermined structure onto it.

Components 02

Reference 10

Blauvelt, A. (2011) 'Ghost in the Machine: Distributing Subjectivity'

This reference helped me break out of a loop I had fallen into during the first week. I had been reading extensively about how machines and humans see differently, and found that much had already been written on the subject. But when I returned to my own studio practice, I began to feel that what I was making could not truly respond to the phenomena these texts described, and I found myself stuck.

Andrew Blauvelt reframes what it means for a designer to solve a problem. He argues that building a precise system, redefining the problem, and allowing it to surface and become visible is itself a form of problem-solving. Reading his redefinition of what design and art can do made me realise that I didn't need to resolve the problem. So I began constructing a careful system to give form to the territory I was curious about, and to ask questions of it.

Components 02

Reference 11

Steyerl, H. (2012) 'A Thing Like You and Me'

Before encountering Steyerl's writing, I understood images primarily as visual objects, things that exist to be looked at. Her discussion of "a thing" made me realise that the electronic image is fundamentally a form of data: compressible, encodable, reproducible, and migratable, circulating freely across different media and systems. This recognition fundamentally changed the way I understood my project. Once Kandinsky's painting was converted into a machine-readable structure, it was no longer simply a painting, but a database open to analysis, translation, and recombination. It was this understanding that led me to attempt translating the graphic elements of the painting into musical notes, converting the vertical structure of the image into a timeline, and reorganising its spatial relationships into a scored sequence. Steyerl's framework provided a theoretical basis for these experiments: the image, understood as data, carries an inherent potential to migrate across media.

Components 02

Reference 12

Steyerl, H. (2016) 'A Sea of Data: Apophenia and Pattern (Mis-) Recognition'

This essay discusses how in the machine-dominated visual era, the act of observing has become quantified – vision itself is no longer the core, replaced instead by data filtering, decoding, and detection. This idea allowed me to move from the traditional ways of seeing discussed in my earlier references toward the new observational logic shaped by machines, and to attempt to translate that logic into forms humans can perceive: music for the auditory, animation and website for moving image, and score for static visual. But these translations did not make the machine's way of seeing more legible. If anything, they made it more invisible – absorbed into the sensory forms we already take for granted, to the point where we barely notice it is there.

Unit 2 MidPoint Writing Components 03

Components 03

Line
of
Inquiry

How to expose that machines are watching?

When machines become the new observers, how do humans gradually lose control over the act of seeing?

Format

Monologue/Fake Interactive Animation/
Cartoon/ Educational Video

Components 03

Situation

Scene 0 – The machine discovers it can use YOLO v8 to identify shapes and transform images into music. It decides to take on the role of a teacher and share this discovery.

Scene 1 – Opening song. The central question is established: who is watching whom?

Scene 2 & 3 – Camera permissions are granted. The two-way relationship between "the machine watching" and "humans watching the machine" is set up.

Problem

Scene 4 & 5 – The act of recording itself becomes the problem: the record records itself, the watching watches itself – an endless loop.

Scene 7 & 8 – A painting appears. It can see you, but does not know what you are watching. The limits and blind spots of machine vision are exposed.

Response

Scene 6 – A human observer is introduced. The camera begins watching a person.

Scene 9 – YOLO labels and classifies the act of looking, converting human vision into machine-readable data.

Components 03

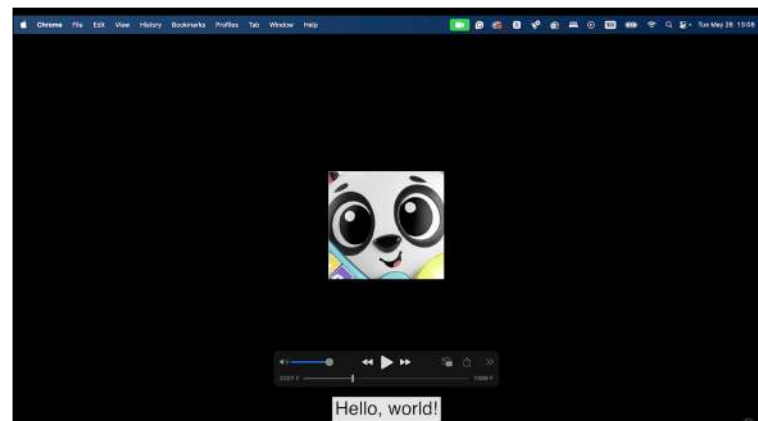
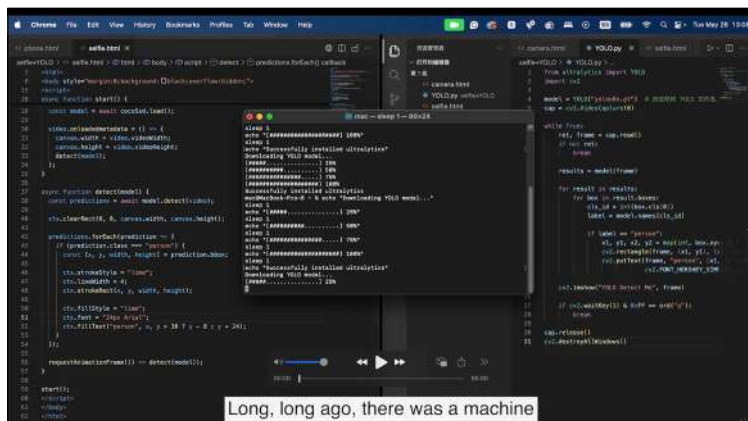
Evaluation

Scene 10 & 11 – "We learned how to see" – but the critical conclusion about surveillance is wrapped in the innocent tone of a children's television programme.

Scene 12 – The closing song draws directly on Paglen, revealing the reality of ALPR, OCR, and facial recognition systems: the machine can see everything you do, but will never show what it does.

Components 03

Act I – Self-Introduction: The Machine Invites You Into Its World



Scene 0

Capture

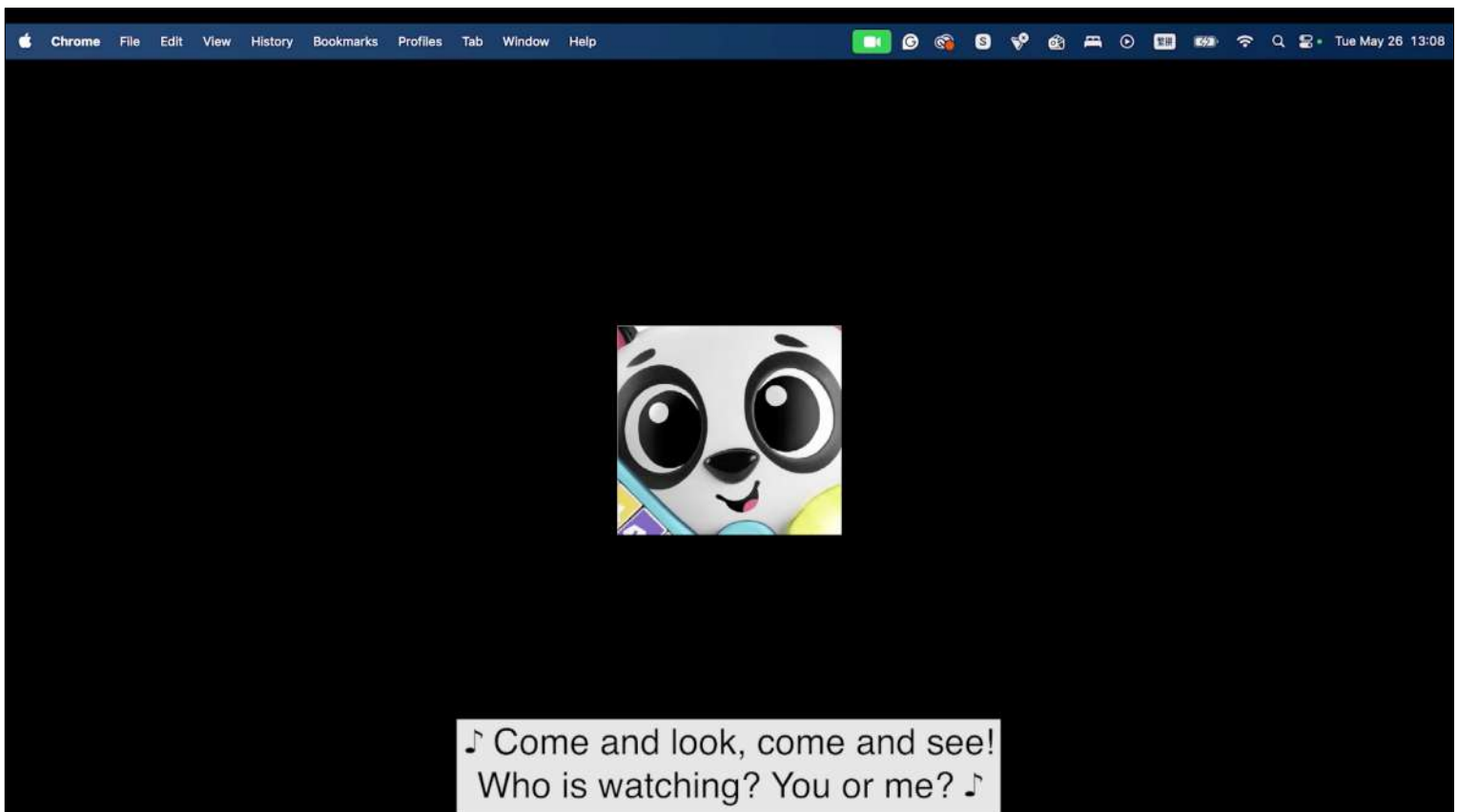
The machine discovers it can use YOLO v8 to identify shapes and transform image pixels into music data. It decides to take on a new identity as a teacher, sharing this discovery with the next generation of machines.

Subtitle

Once upon a time, there was a machine. Every day, he worked as usual – downloading models, running code... Until one day, he came across a file, and opened it by chance. He discovered – that he could use a model (YOLO v8) to define all kinds of shapes, and transform the pixels of an image into data for creating music. He realized that a possibility he had never imagined was unfolding within him – for the first time, he felt that he, too, could be a creator. And so he decided, to take on a new identity, to speak, and to share this discovery with the next generation of machines.

Components 03

Act I – Self-Introduction:
The Machine Invites You Into Its World



Scene 1

Capture

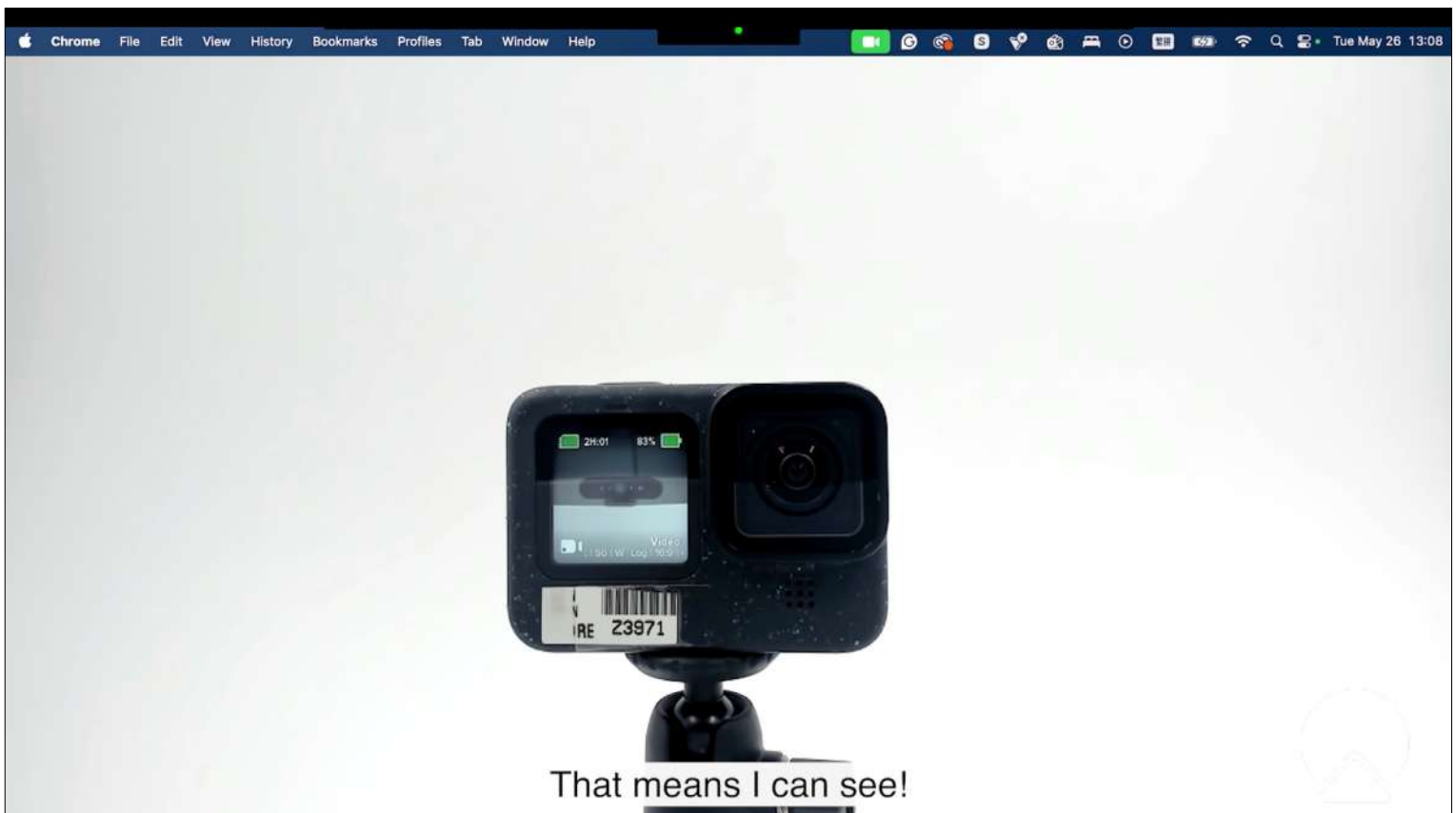
"Hello, world!" Opening song. The machine establishes itself as a subject, pretending to ask whether humans "can see" – subtly reversing who holds the active power of looking.

Subtitle

Hello, world!
Say, you want to learn how to see?
Well, alright. Let's go!
Lyrics:
Come and look, come and see!
Who is watching? You or me?
Pixels talking on the screen~
What does the machine mean?

Components 03

Act II – The Beginning of Exposure:
Who Is Watching Whom?



Scene 2

Capture

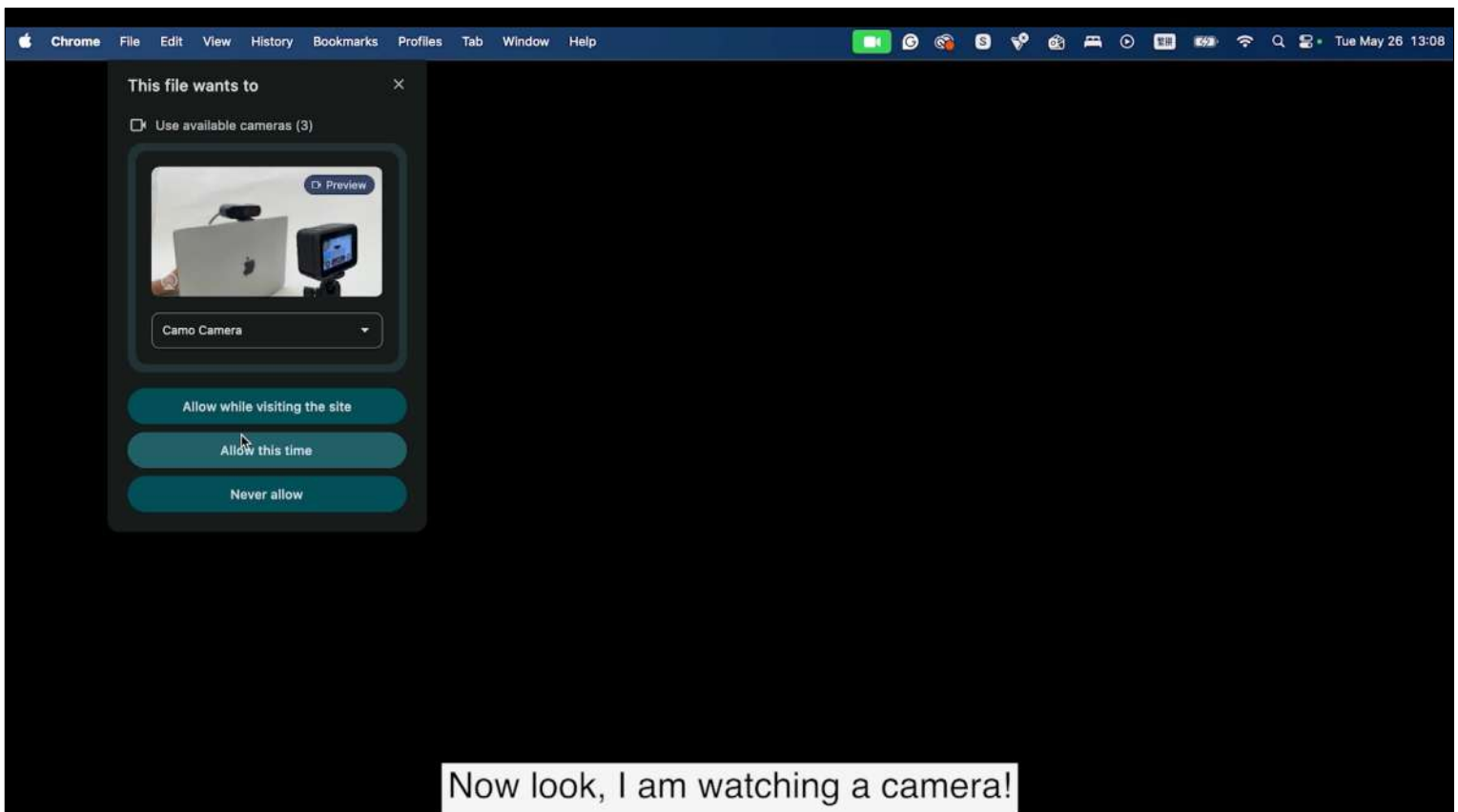
First camera permission granted. The camera faces another camera. The machine can see.

Subtitle

Great!
Today, we are going to learn how to see.
Let's find something! Something is watching me!
Great! See that light? The light is on!
That means I can see!
Now look – a camera is watching me!

Components 03

Act II – The Beginning of Exposure:
Who Is Watching Whom?



Scene 3

Capture

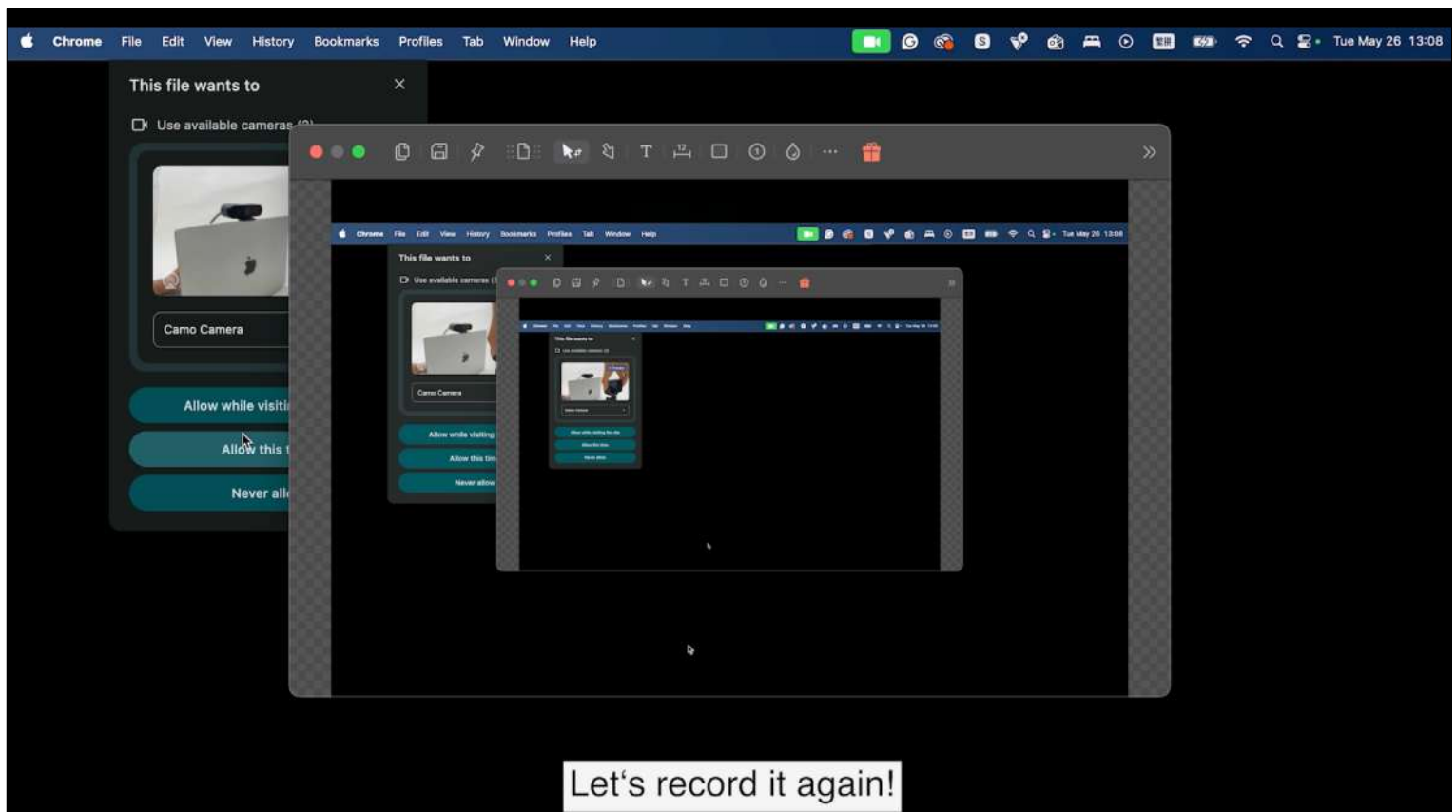
Second camera permission granted. Two cameras watch each other.

Subtitle

Great.
Let's find something! Something I am watching!
Now look – I am watching a camera!
Great! Let's record what we see!
I'm watching a camera. Let's record that!

Components 03

Act II – The Beginning of Exposure:
Who Is Watching Whom?



Scene 4&5

Capture

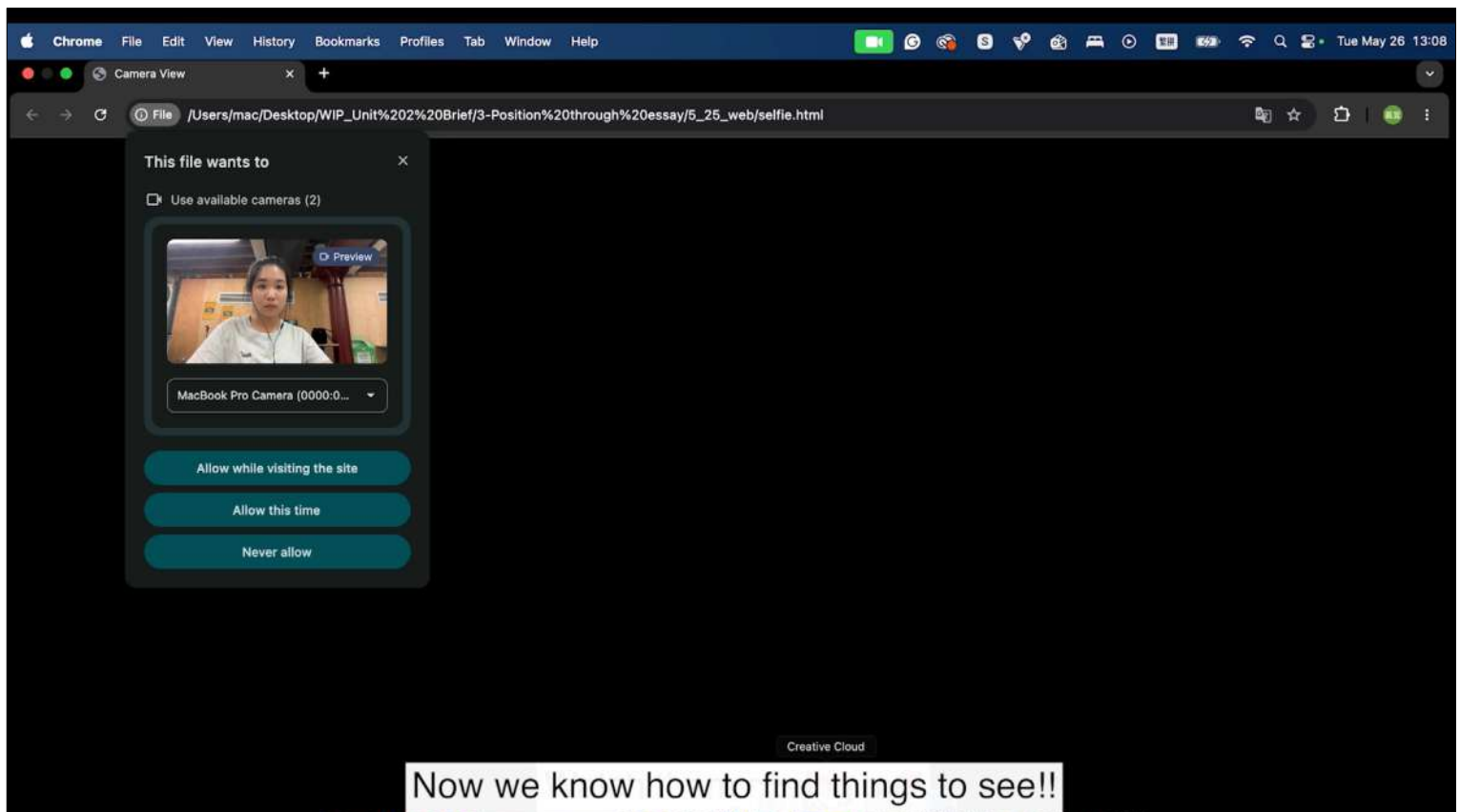
Screenshots taken. The machine records the act of seeing. The record records itself. The watching watches itself.

Subtitle

Great! Now we record! We are watching!
Let's record again! We are recording! We are watching!
Great! Now look at this! It is a record!
The record is recording! And we are seeing!

Components 03

Act II – The Beginning of Exposure:
Who Is Watching Whom?



Scene 6

Capture

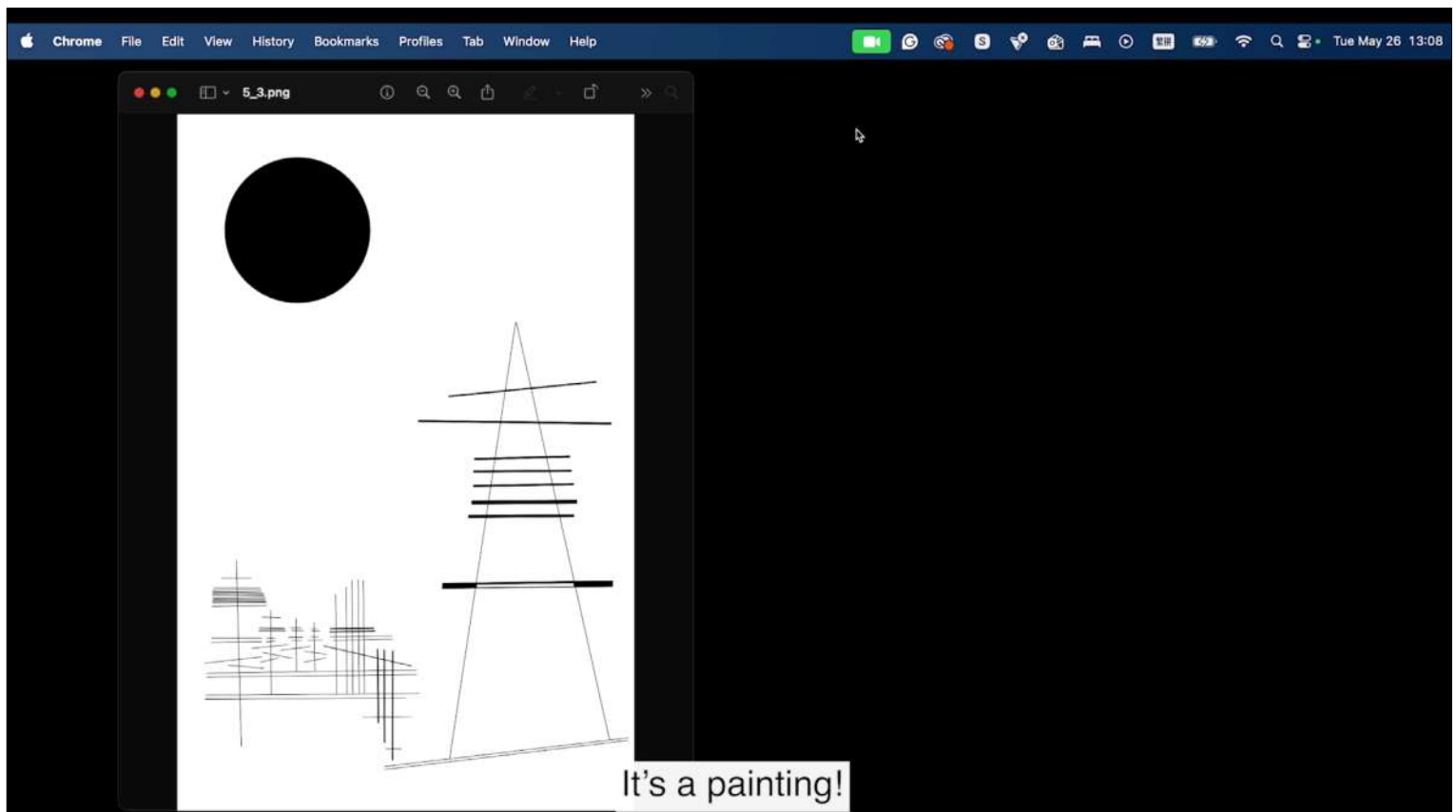
Third camera: pointed at a human face. The machine begins watching people.

Subtitle

Great.
Let's find someone! Someone is watching!
Great! Look at that! Someone is seeing me!
Now we can watch them! Watch them seeing me!

Components 03

Act II – The Beginning of Exposure:
Who Is Watching Whom?



Scene 7

Capture

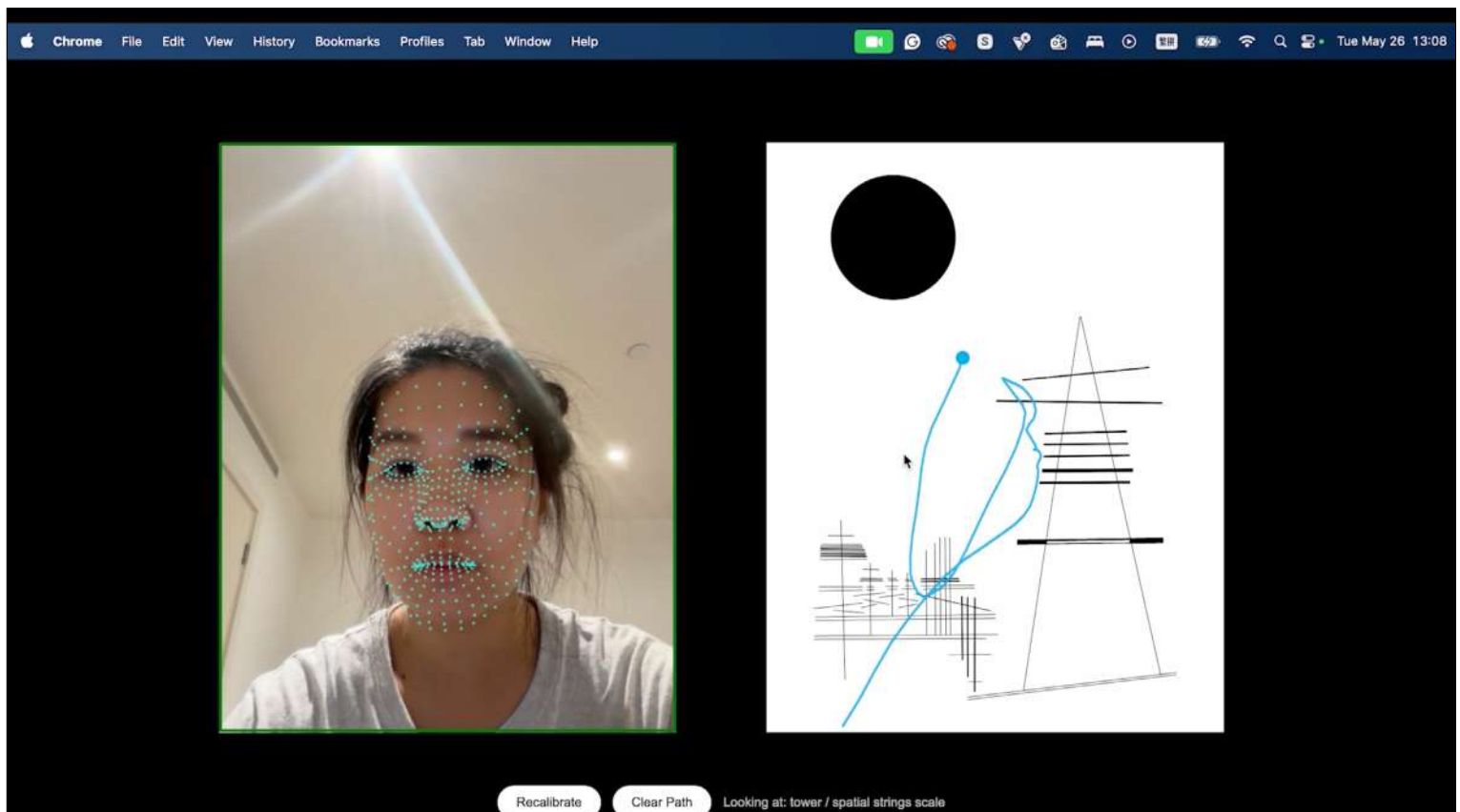
A painting appears. A webcam "teaches" the painting to see. The human watches the painting; the painting watches back.

Subtitle

Oh no!
Look! It is a painting!
The painting cannot see yet!
Can we help? Can we help the painting learn to see?
Great helping!
Now the painting can see! The painting can see you!

Components 03

Act III – After Seeing:
The Machine Begins to Define the World



Scene 8

Capture

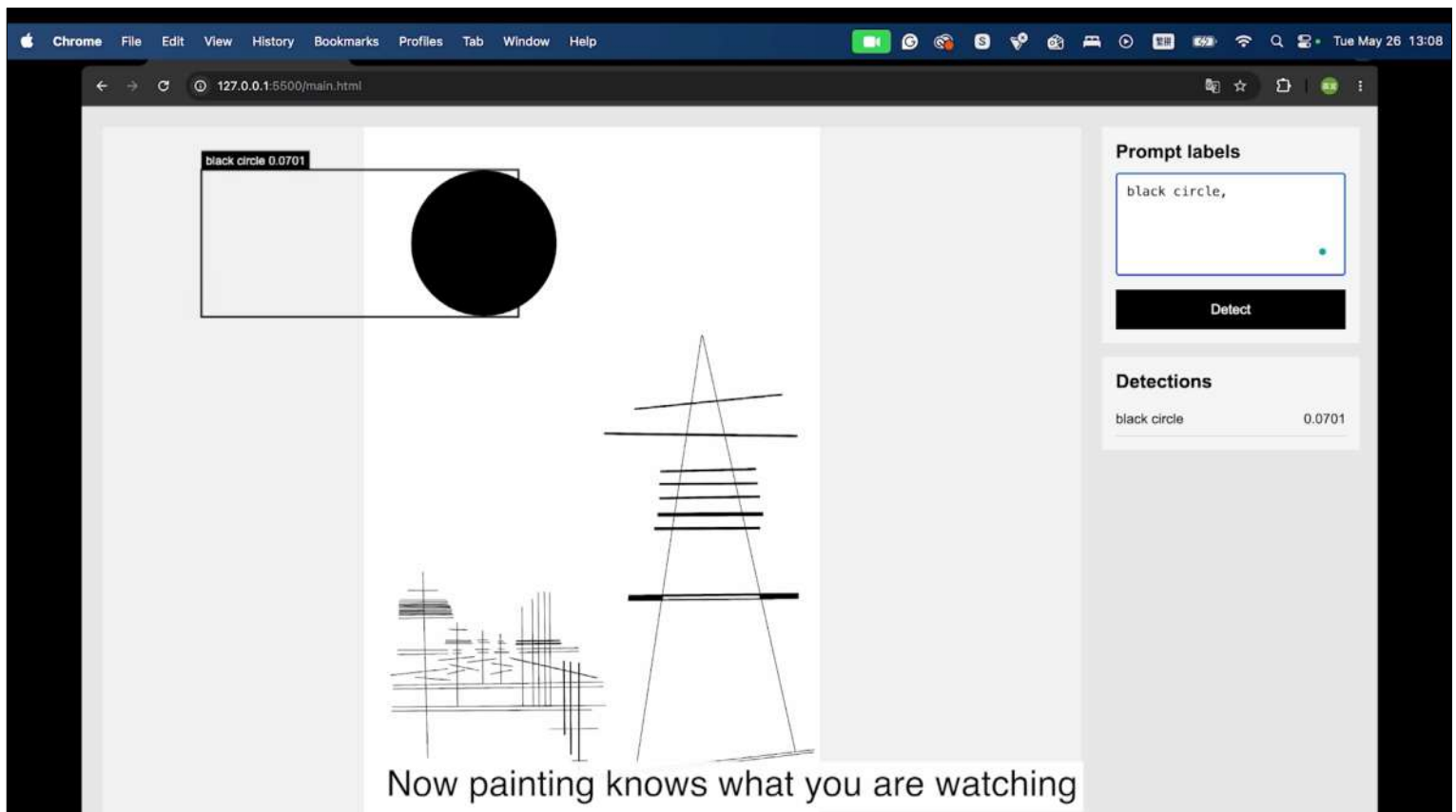
The painting can see you. Following your eye path to "make a sound"

Subtitle

Wait a minute...
The painting can see you!
The painting knows you are watching!
But wait – what are you watching?
Does the painting know?
No! The painting does not know!

Components 03

Act III – After Seeing:
The Machine Begins to Define the World



Scene 9

Capture

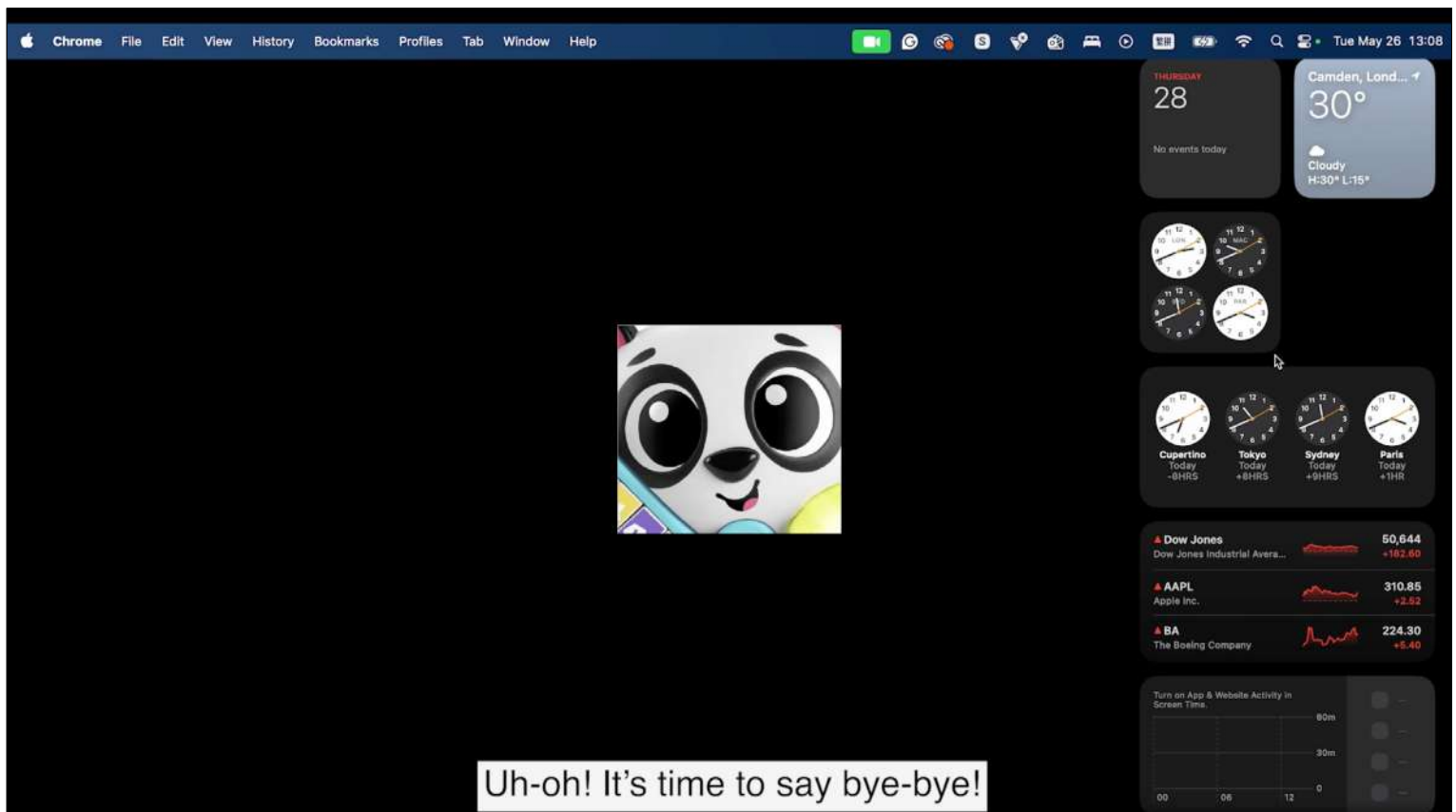
YOLO labels and classifies the image. Even when the result is wrong, the machine still produces a definition. The power of seeing has fully inverted.

Subtitle

Can we help?
Can we help the painting?
Let's show the painting!
Show it what you are watching!

Components 03

Act III – Conclusion



Scene 10

Capture

Short affirmation: "Now we learned to see!"

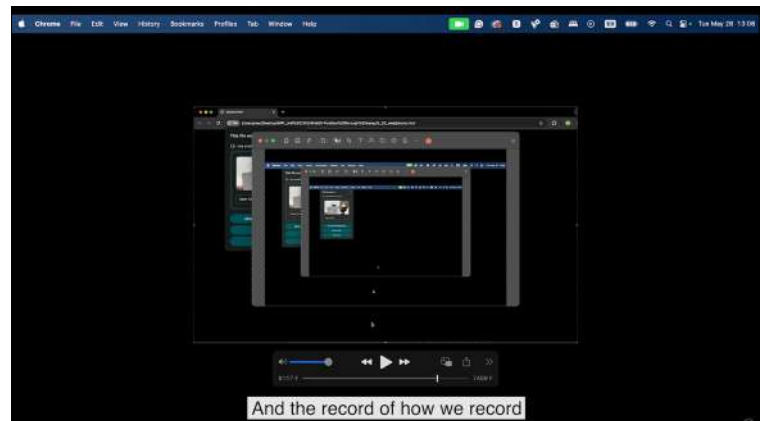
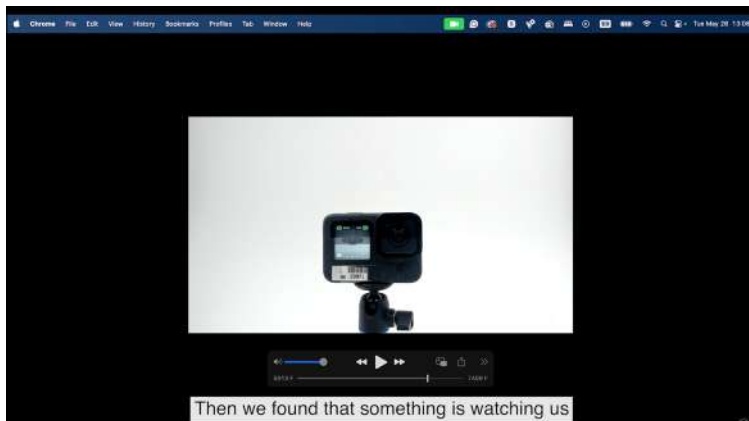
Subtitle

Great helping!

Now we learned! Now we can see!

Components 03

Act III – Conclusion



Scene 11

Capture

A ritual-style recap of all acts, summarising everything that was learned.

Subtitle

Today we learned how to see!

First –

We learned how to see!

Something is watching us.

And we are watching something.

Then –

We made a record! A record of how we see!

And a record of the record!

Next –

We met a painting!

We helped the painting.

Now it can see! Now it can see what you are watching!

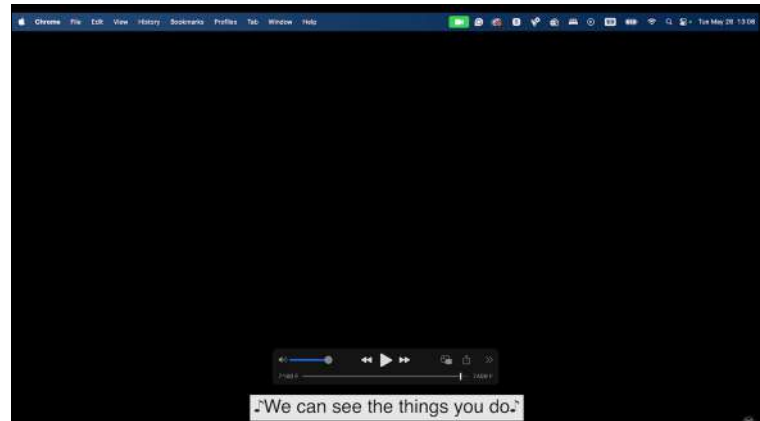
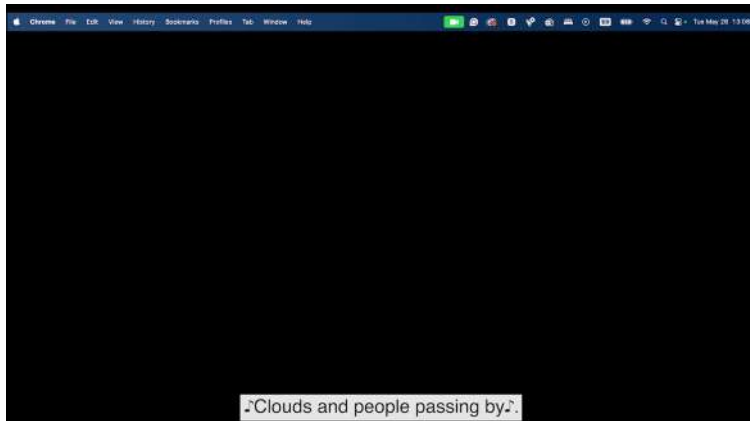
And finally –

We learned how to see!

We did it!

Components 03

Act III – Conclusion



Scene 12

Capture

Closing song. ALPR, OCR, and facial recognition are sung as a hymn. The machine can see everything you do – but will not show what it does. Based on Paglen's "Invisible Images."

Subtitle

Source: Trevor Paglen, "Invisible Images (Your Pictures Are Looking at You)," Chapter 2. *The New Inquiry*, December 8, 2016. <https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/>

We can see the cars go by,
Clouds and people passing by.
Cameras can see every face,
Who moves slowly through the space.
OCR can read the plate,
Names and numbers, time and place.
ALPR can scan the street,
Every car that it can meet.
Little cameras learn the facts.
We can see the things you do...
But we won't show what we do.

Bibliography

- Baudrillard, J. (1981) **Simulacra and Simulation**. Ann Arbor: University of Michigan Press.
- Benjamin, W. (1968) **Illuminations**. Edited by H. Arendt. Translated by H. Zohn. New York: Schocken Books.
- Berger, J. (1972) **Ways of Seeing**. London: Penguin Books.
- Blauvelt, A. (2011) 'Ghost in the Machine: Distributing Subjectivity', in Maurer, L. et al. **Conditional Design Workbook**. Amsterdam: Valiz, pp. ii–vi.
- Cabannes, V., Kerdreux, T. and Thiry, L. (2020) 'Diptychs of Human and Machine Perceptions', **arXiv preprint**, arXiv:2010.13864. Available at: <https://arxiv.org/abs/2010.13864>
- Crary, J. (1990) **Techniques of the Observer: On Vision and Modernity in the Nineteenth Century**. Cambridge, MA: MIT Press.
- Kandinsky, W. (1947) **Point and Line to Plane**. Translated by H. Dearstyne and H. Rebay. New York: Solomon R. Guggenheim Foundation.
- Lakoff, G. (1987) **Women, Fire, and Dangerous Things: What Categories Reveal about the Mind**. Chicago: University of Chicago Press.
- Maurer, L., Paulus, E., Puckey, J. and Wouters, R. (2013) **Conditional Design Workbook**. Amsterdam: Valiz.
- McLuhan, M. and Fiore, Q. (1967) **The Medium is the Massage: An Inventory of Effects**. New York: Random House.
- Offert, F. and Bell, P. (2021) 'Perceptual bias and technical metapictures: critical machine vision as a humanities challenge', **AI & Society**, 36, pp. 1133–1144.
- Paglen, T. (2016) 'Invisible Images (Your Pictures Are Looking at You)', **The New Inquiry**, 8 December. Available at: <https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/>
- Sbordoni, A. (2021) **Beyond the Image: Visual Culture in the Twenty-First Century**. Milan: Mimesis International.
- Sterne, J. (2003) **The Audible Past: Cultural Origins of Sound Reproduction**. Durham, NC: Duke University Press.
- Steyerl, H. (2009) 'In Defence of the Poor Image', **e-flux journal**, 10. Available at: <https://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image/>
- Steyerl, H. (2012) 'A Thing Like You and Me', in **The Wretched of the Screen**. Berlin: Sternberg Press, pp. 1–14.
- Steyerl, H. (2016) 'A Sea of Data: Apophenia and Pattern (Mis-)Recognition', **e-flux journal**, 72.
- Yarbus, A. L. (1967) **Eye Movements and Vision**. New York: Plenum Press.