

Position

③ Paglen "machine" look first, but they're not just look first they define images first

① Hidden Catalogue

② Resolution

④ For Humans, meaning resides in the center of the image (the panda); for machine meaning distributed evenly across the statistical relationships of all pixels, with no distinction between foreground & background. These are two entirely different structure of attention - and two entirely different worldviews.

Mean Image

Machine → activation

Human → Way of Seeing

UMA

Pixel-Statistical Relations as a New Compositional Grammar

Pixel-statistical

Fluidity

Visual Culture

Heatmap =

Mona Lisa

Arts

History

Machine

Visual Culture is Ahistorical

Human < Ways of seeing >

Machine

Invisible images

大量的图像仍然存在,但它们只和机器内流动

Eq. 某些即时影像 → AI分析 → 再森

以前图像即 → 描述世界。这是现在的Out Come

中 → (传递资讯)

现在变成

activation → 解除行动

operation → 参与运作

enforcement → 执行规则/控制

不再是被看,而是在做事情

(inter-subjectivity 主体与主体之间的意义)

↓ 人与人之间的感知关系

Eq. 社交媒体上的大量图像 → 改变人与人之间如何看彼此(理解彼此/建立共识与认知)

神经网络(不是这个人是否输入了吗?)

数字图像最革命的地方在于它们是被机器读取的。例如,手机拍摄的图片拍摄的图片会生成一个机器可读的档案,但这个档案本身不能被人类感知。一个附加应用,例如搭配液晶显示器背光

的图片浏览软件,才能被人看到。这一点与幻灯片还不一样,尽管幻灯片同样需要通过化学处理

The application and development of image recognitions

→ Medical Image Analysis 医学影像分析。一精确判断

→ Intelligent Driving 智能驾驶。一即时感知

→ Sec Security Monitoring 安防监控。一持续监视。

钱里里的家人孩子照片以及家庭旅游幻灯片播放,某种程度上是,数字图像的前身为幻灯片,猫咪,旅行照片建立各种相簿。

① hidden catalogue

② text

③ resolution

UMA

④ Paglen的机器看进一步推:机器不认识看它还知道。

⑤ "人眼叫呼号的位置" 机器却从中识别出某个物体。

对于人类该图像的意义源于图像中心;

对于机器来意义均匀分布在所有像素的统计关系

Final score = objectness * class probability

[有没有东西] [是什么东西]

这是种完全不同的注意力结构,也是两种不同的世界观

Feedback总结

① 在设计什么?

② 这个系统在设计什么? (没有明确的 Communication)

③ 它在表达什么?

methodology

① threshold exp exploration

② Designed input 把模型当实验对象

③ 用实验逼出模型逻辑。

DeepFace (Facebook)

Invisible Image

人的视觉文化习得。

人类拍摄的照片被AI拿去训练,变成机器理解世界的基石。

↓ 文中用的Facebook的DeepFace作为例子,但是因为它向语境话题会更加充分,但是YOLOV8作为一个没有语境延伸的模型,它的视觉文化是不是就更接近机器,但是无论下如何他们创造的视觉文化的方式都是前所未有的

→ Chapter 3

卷积运算。

视觉符号的特定形成。

① 被共同分享 ② 在媒体中重复生产

③ 意义可以被解码与再编码。

What kind of visual culture does machine produce?

- Content In terms of content: does machine recognition share the same visual vocabulary as human perception?
- Time In terms of time: does machine vision inhabit the same historical moment as its human viewer?
- Scale In terms of the individual: can machine vision account for the culturally specific memories that shape how each person reads an image?

Mean Image

Machine → activation → Human → Way of Seeing

③ Paglen "machine" look first, but they're not just look first they define images first

① Hidden Catalogue

UMA

② Resolution → Pixel-Statistical Relations as a New Compositional Grammar

④ For Humans, meaning resides in the center of the image (the panda); for machine meaning distributed evenly across the statistical relationships of all pixels, with no distinction between foreground & background. These are two entirely different structure of attention - and two entirely different worldviews.

③ Text → Pixel-statistical → Fluidity

Heatmap =

Time

Mona Lisa ← Arts ← History

Machine Visual Culture is Ahistorical

Human < Ways of seeing >

Machine

Invisible images

大量图像仍然存在,但它们只和机器内流动。一些控制图像 → AI分析 → 再生产。以前图像 → 描述世界。这是现在的做法。 (在做的)

现在变成 activation → 解构行为 operation → 参与运作 enforcement → 执行规则/控制 不再是被动的,而是在做事情 (inter-subjectivity 主体与主体之间的意义)

人与人之间感知关系。Eq. 社交媒体上的大量图像 → 改变人与人之间如何看彼此 (理解彼此/建立共识与认知)

神经网络 (不是这个人是通神经网络吗?)

数字图像最革命的地方在于它们是被机器读取的。例如,手机拍摄的图片拍摄的图片会生成一个机器可读的档案,但这个档案本身不能被人类感知。一个附加应用,例如搭配液晶显示器背光的照片浏览软件,才能被人看到。这一点与幻灯片还不一样,尽管幻灯片同样需要通过化学处理

The application and development of image recognitions

- Medical Image Analysis 医学影像分析。一精确判断
- Intelligent Driving 智能驾驶。一即时感知
- Sec Security Monitoring 安防监控。一持续监视。

钱里里的家人照片以及家庭旅游幻灯片播放,某种程度上是,数字图像的前身为幻灯片,猫咪,旅行照片建立各种相簿。

hidden catalogue

UMA

④ Paglen的机器看进一步推: 机器不认识看它还未知。

⑤ "人眼叫呼号的位置 机器却从中识别出某国物体" 对于人类图像的意义源于图像中心; 对于机器来意义均匀分布在所有像素的统计关系

Final score = objectness × class probability

有没有东西 [是什么东西]

这是种完全不同的注意力结构,也是两种不同的世界观

Feedback总结

- ① 在设计什么? (没有明确的 Communication)
- ② 这个系统在设计什么?
- ③ 它在表达什么?

methodology

- ① threshold exp exploration
- ② Designed input 把模型当实验对象
- ③ 用实验逼出模型逻辑。

DeepFace (Facebook)

Invisible Image

人的视觉文化习得。人类拍摄的照片被AI拿去训练,变成机器理解世界的基石。

文中用的Facebook的DeepFace作为例子,但是YoloV8作为一个没有语境延伸的模型,它的视觉文化是不是就更接近机器,但是无论下如何他们创造的视觉文化的方式都是前所未有的

→ Chapter 3 卷积运算。

视觉符号的特定形成

- ① 被共同共享
- ② 在媒介中重复生产
- ③ 意义可以被解码与再编码。

In terms of content:

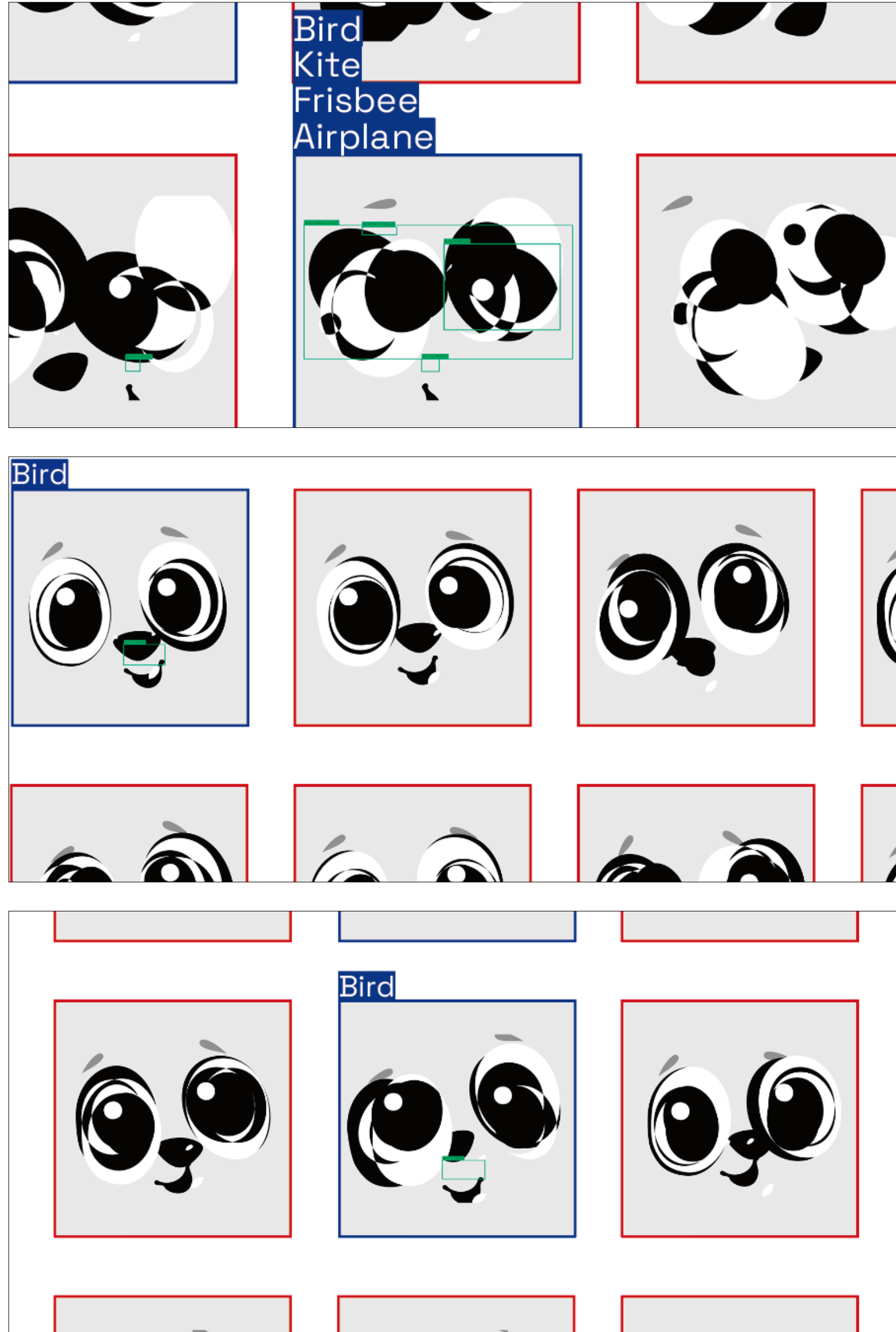
Does machine recognition share the same visual vocabulary as human perception?

Position Through Contextualising

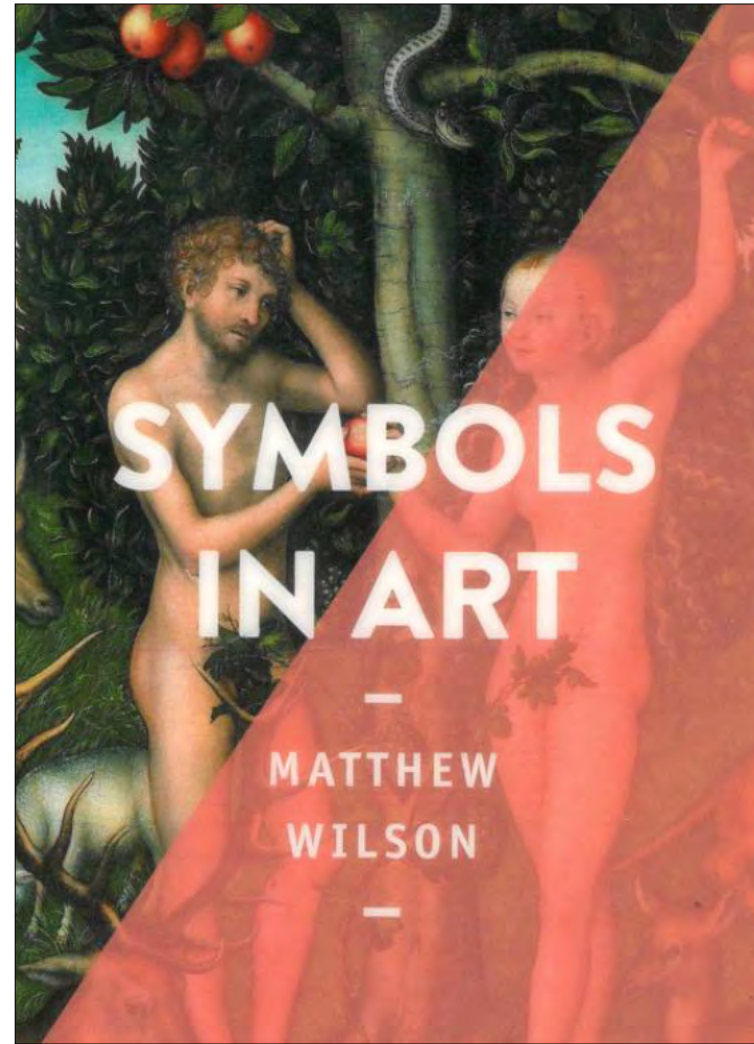
Before

&During

&After



Position Through Contextualising




BIRDS

How do you know but ev'ry Bird that cuts the airy way, Is an immense world of delight, clos'd by your senses five?

William Blake
1793

DOVE



In many global cultures, doves are associated with positive qualities such as peace and harmony. The dove is often depicted with an olive branch in its beak, symbolizing peace. The dove is also a symbol of the Holy Spirit in Christianity, often shown as a white bird with a human face. The dove is a symbol of hope and renewal, often used in art and literature to represent these concepts.

PEACOCK



The peacock is a symbol of vanity and pride. In the Bible, the peacock is mentioned in the Book of Isaiah, where it is described as a bird that is proud of its beautiful feathers. The peacock is also a symbol of the East, often used in art and literature to represent these concepts.

EAGLE



The eagle is a symbol of strength and power. In the Bible, the eagle is mentioned in the Book of Isaiah, where it is described as a bird that is proud of its strength. The eagle is also a symbol of the United States, often used in art and literature to represent these concepts.

PHOENIX



The phoenix is a symbol of rebirth and renewal. In the Bible, the phoenix is mentioned in the Book of Isaiah, where it is described as a bird that is reborn from its own ashes. The phoenix is also a symbol of the East, often used in art and literature to represent these concepts.

FALCON




The falcon is a symbol of speed and precision. In the Bible, the falcon is mentioned in the Book of Isaiah, where it is described as a bird that is swift and powerful. The falcon is also a symbol of the East, often used in art and literature to represent these concepts.

FALCON



The falcon is a symbol of speed and precision. In the Bible, the falcon is mentioned in the Book of Isaiah, where it is described as a bird that is swift and powerful. The falcon is also a symbol of the East, often used in art and literature to represent these concepts.

OWL



The owl is a symbol of wisdom and knowledge. In the Bible, the owl is mentioned in the Book of Isaiah, where it is described as a bird that is wise and powerful. The owl is also a symbol of the East, often used in art and literature to represent these concepts.

CRANE



The crane is a symbol of longevity and grace. In the Bible, the crane is mentioned in the Book of Isaiah, where it is described as a bird that is long-lived and powerful. The crane is also a symbol of the East, often used in art and literature to represent these concepts.

Before

&During
Ref.

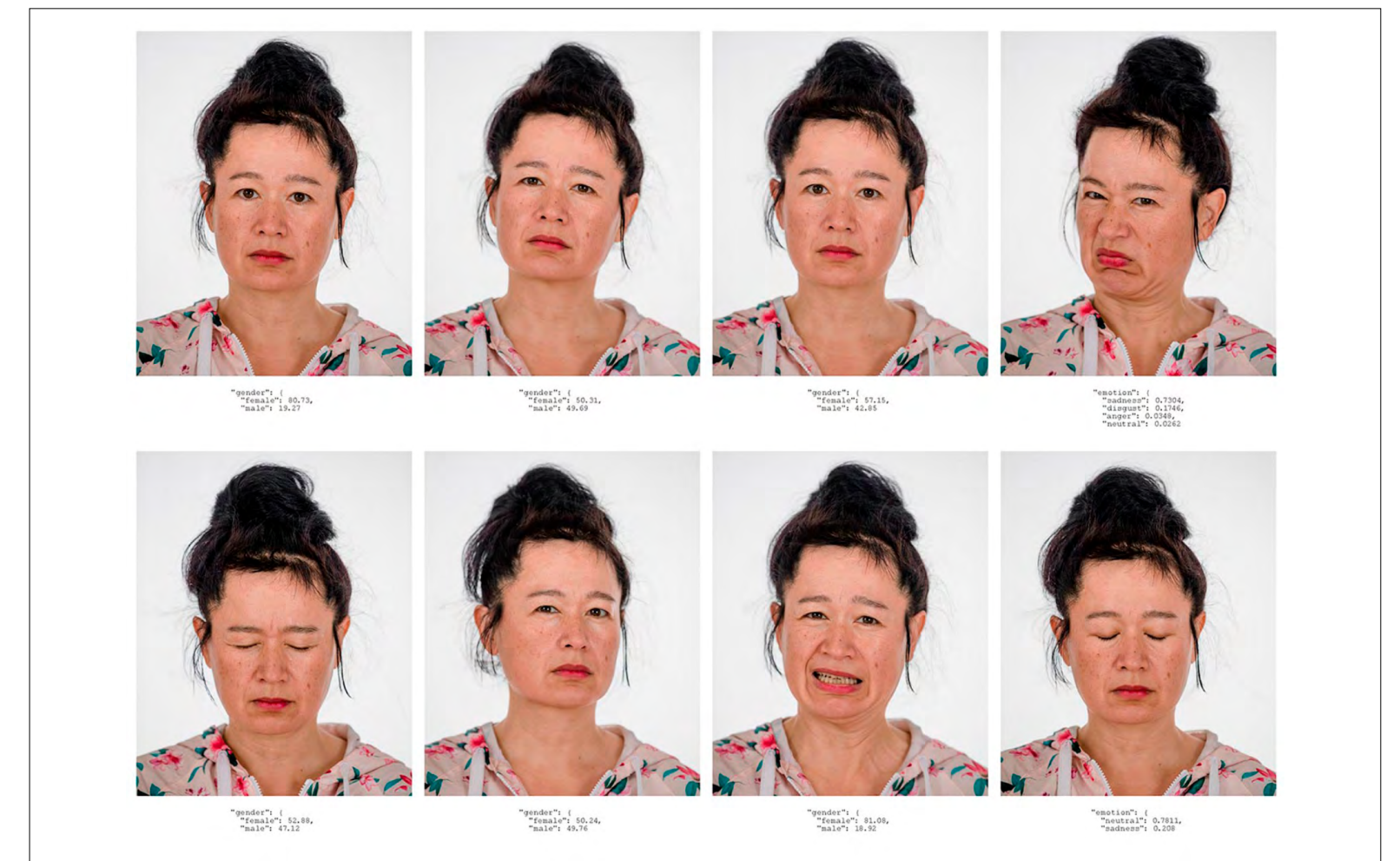
&After

Position Through Contextualising

Before

&During
Ref.

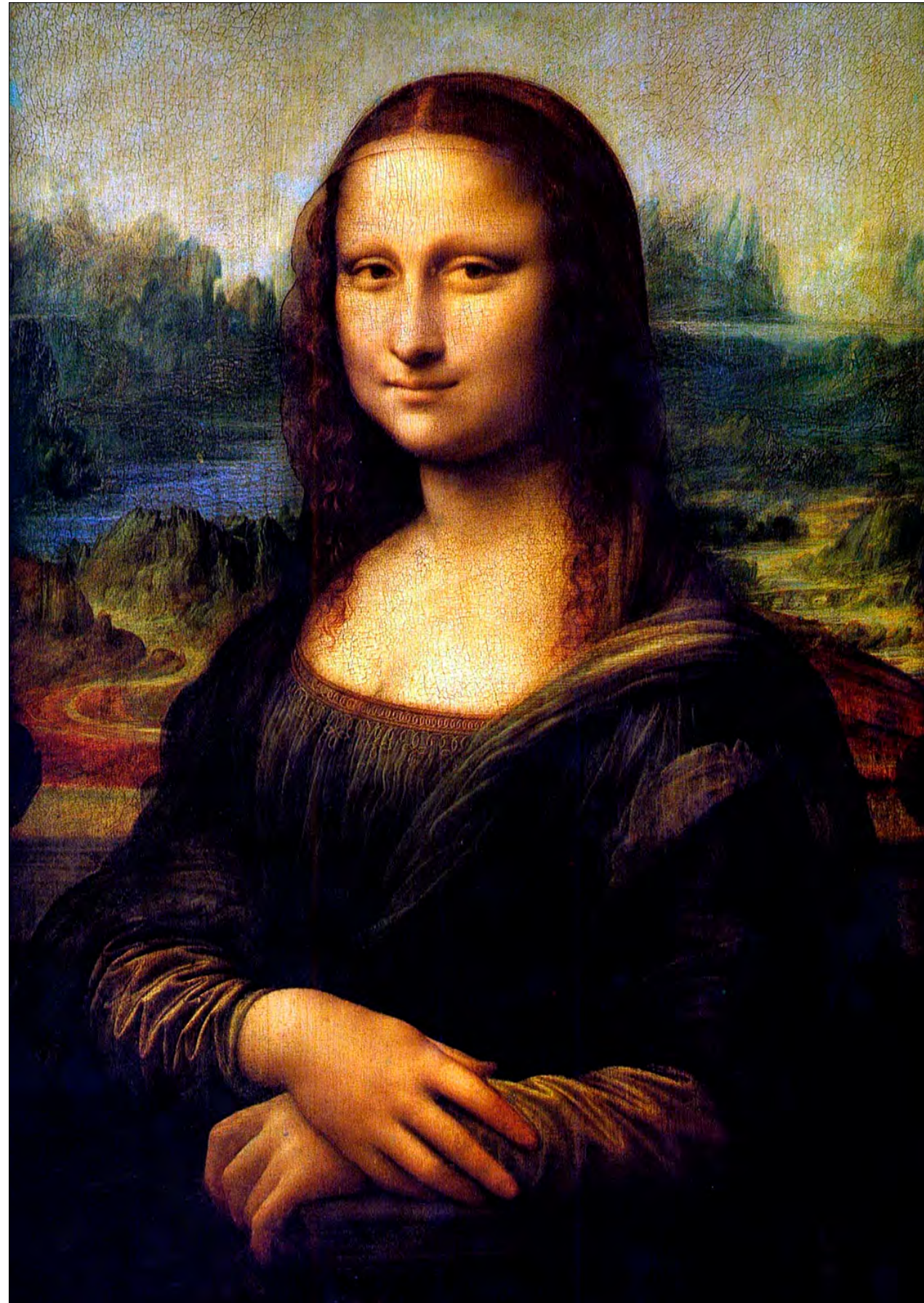
&After



Machine-Readable Hito & Holly
Hito Steyerl & Holly Herndon, 2018

In terms of time: does machine vision inhabit the same historical moment as its human viewer?

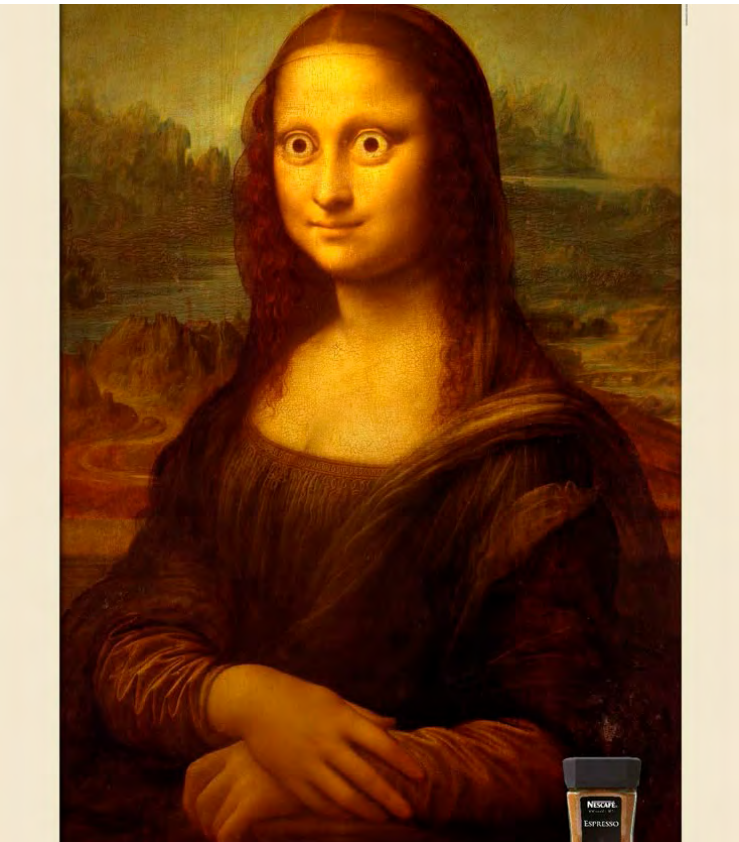
Position Through Contextualising



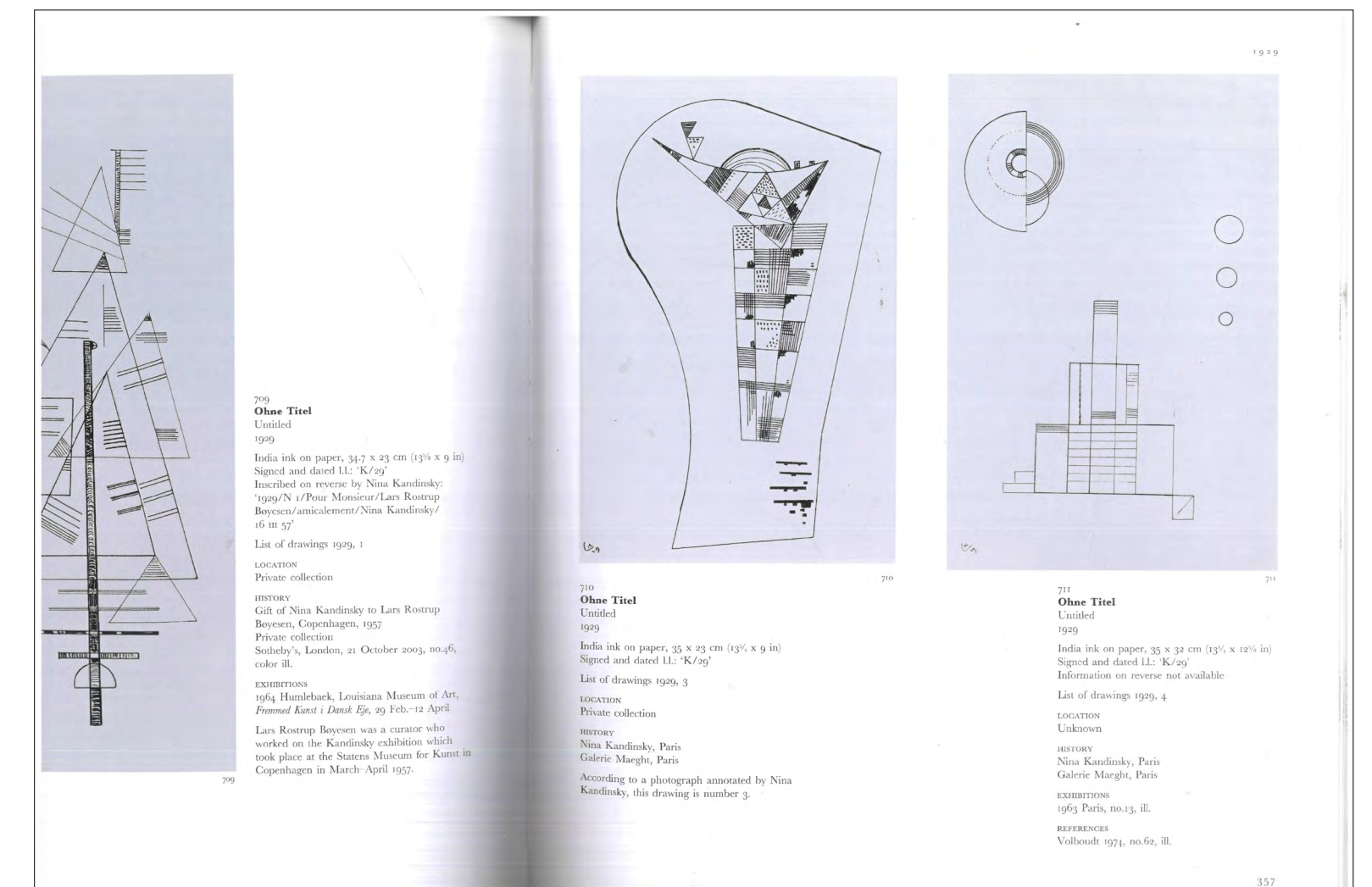
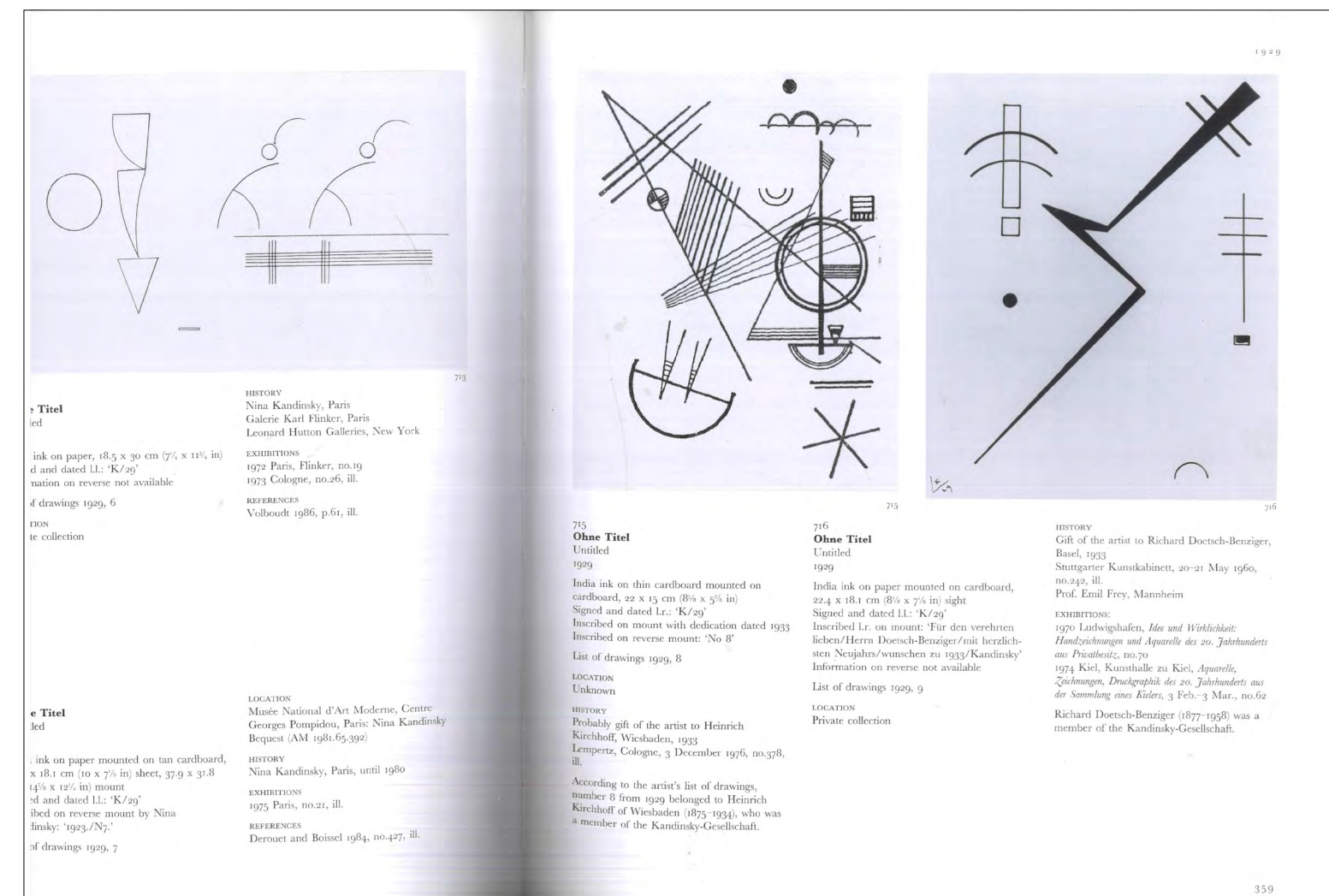
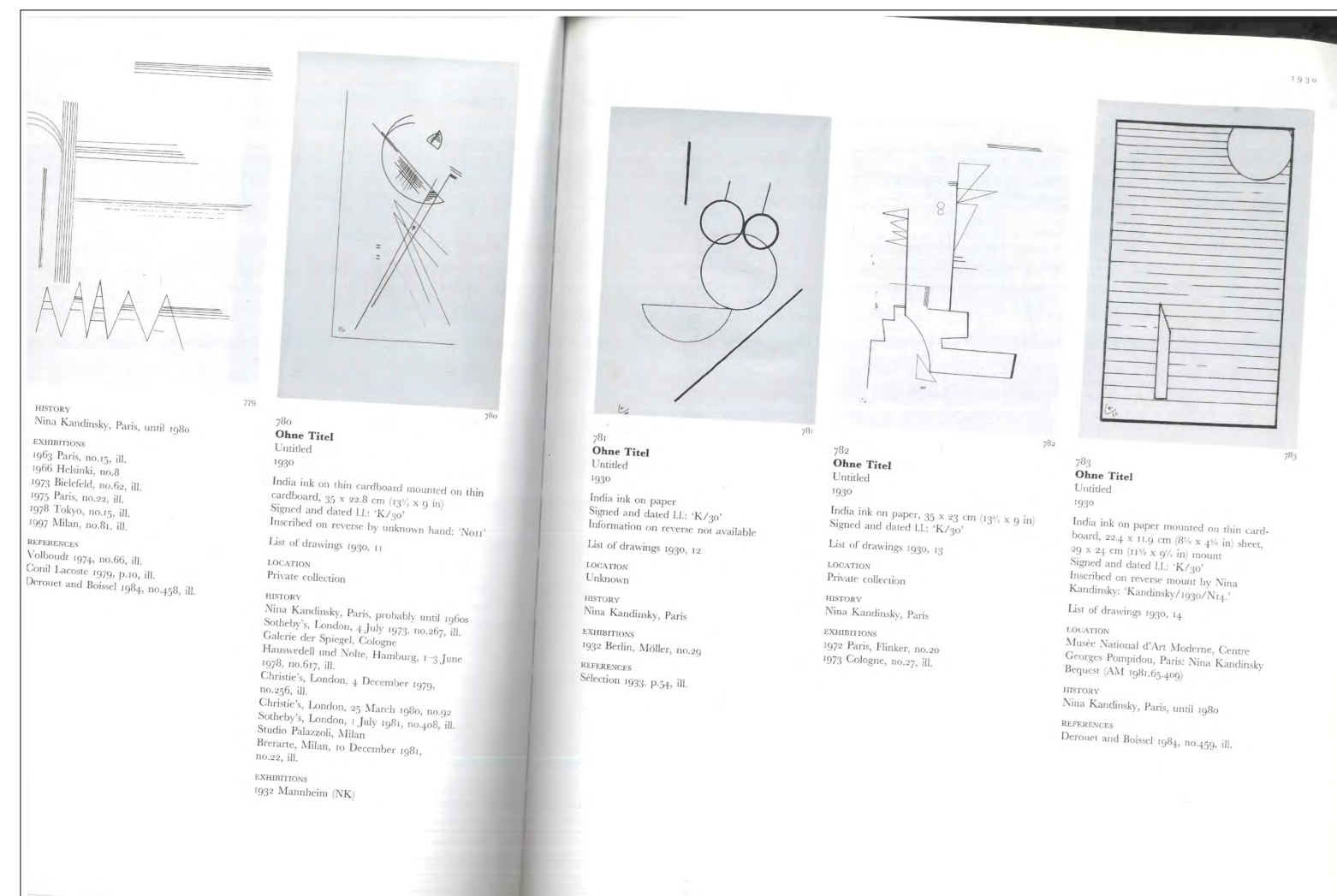
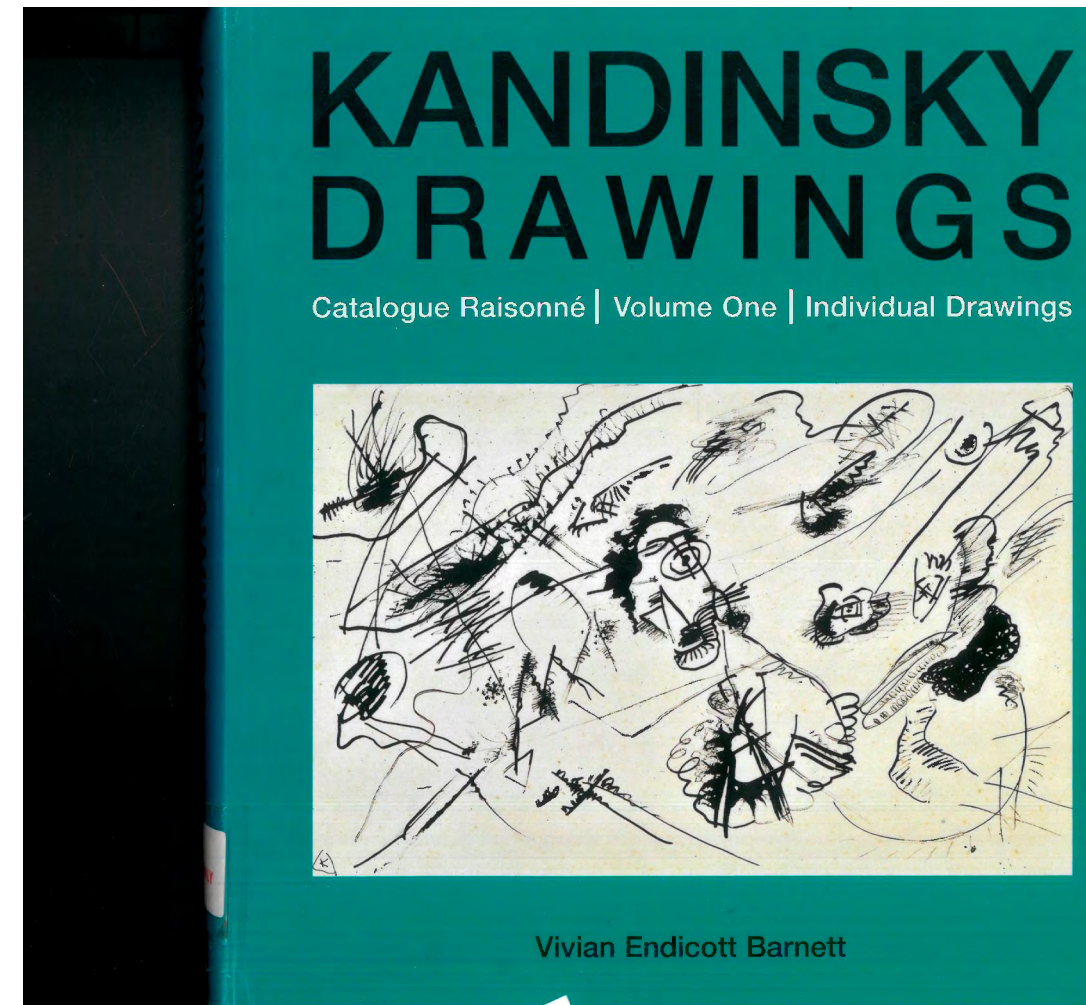
Before

&During
Ref.

&After



In terms of the individual: can machine vision account for the culturally specific memories that shape how each person reads an image?

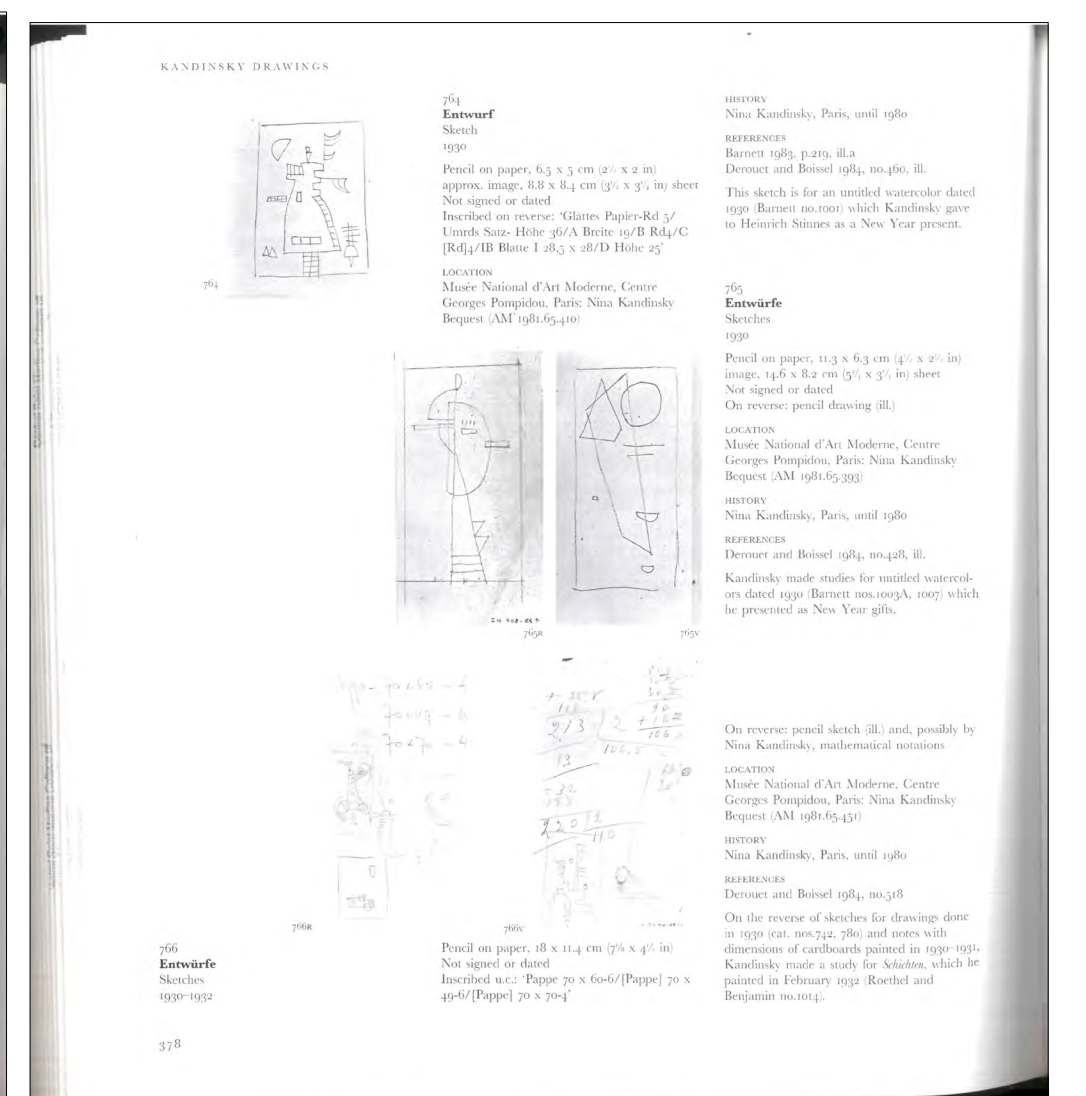
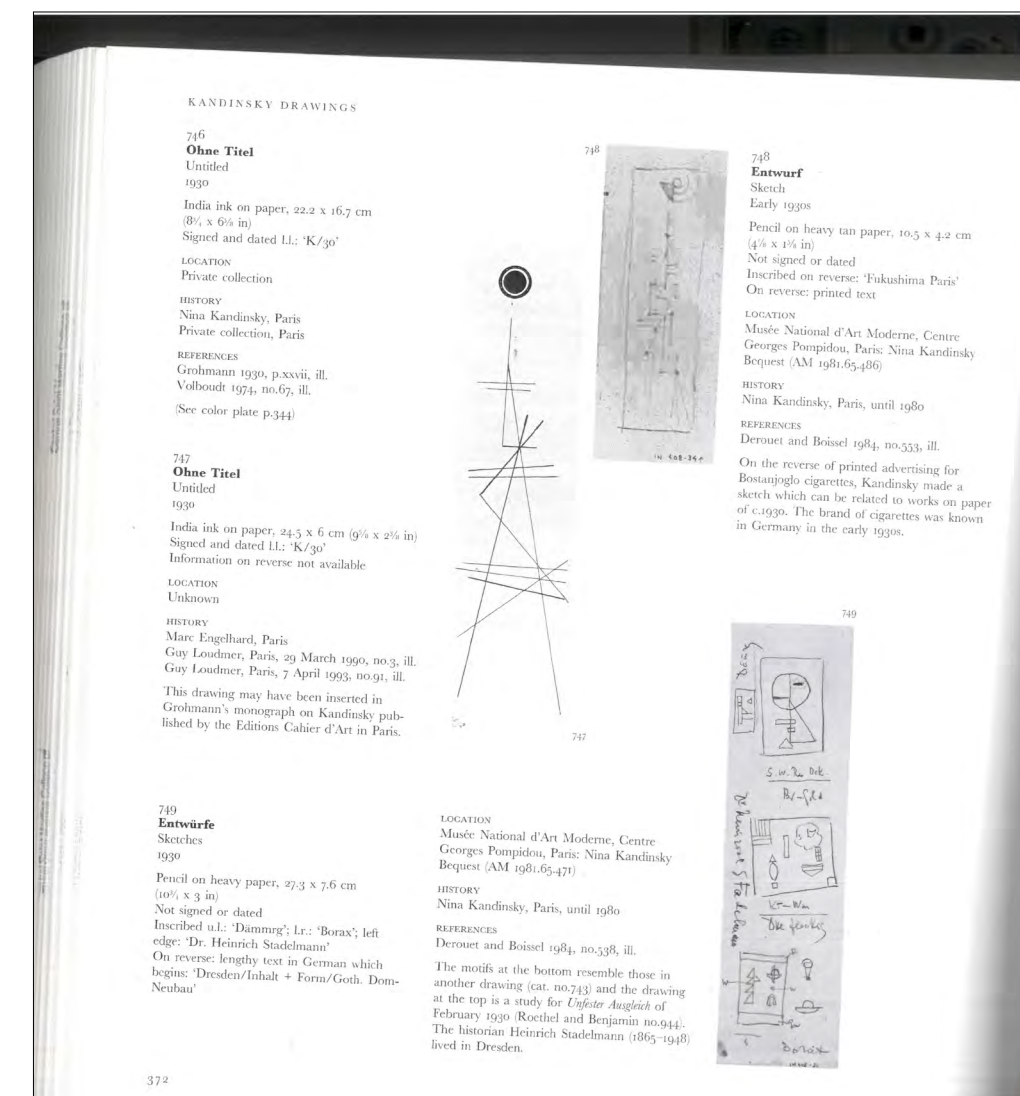
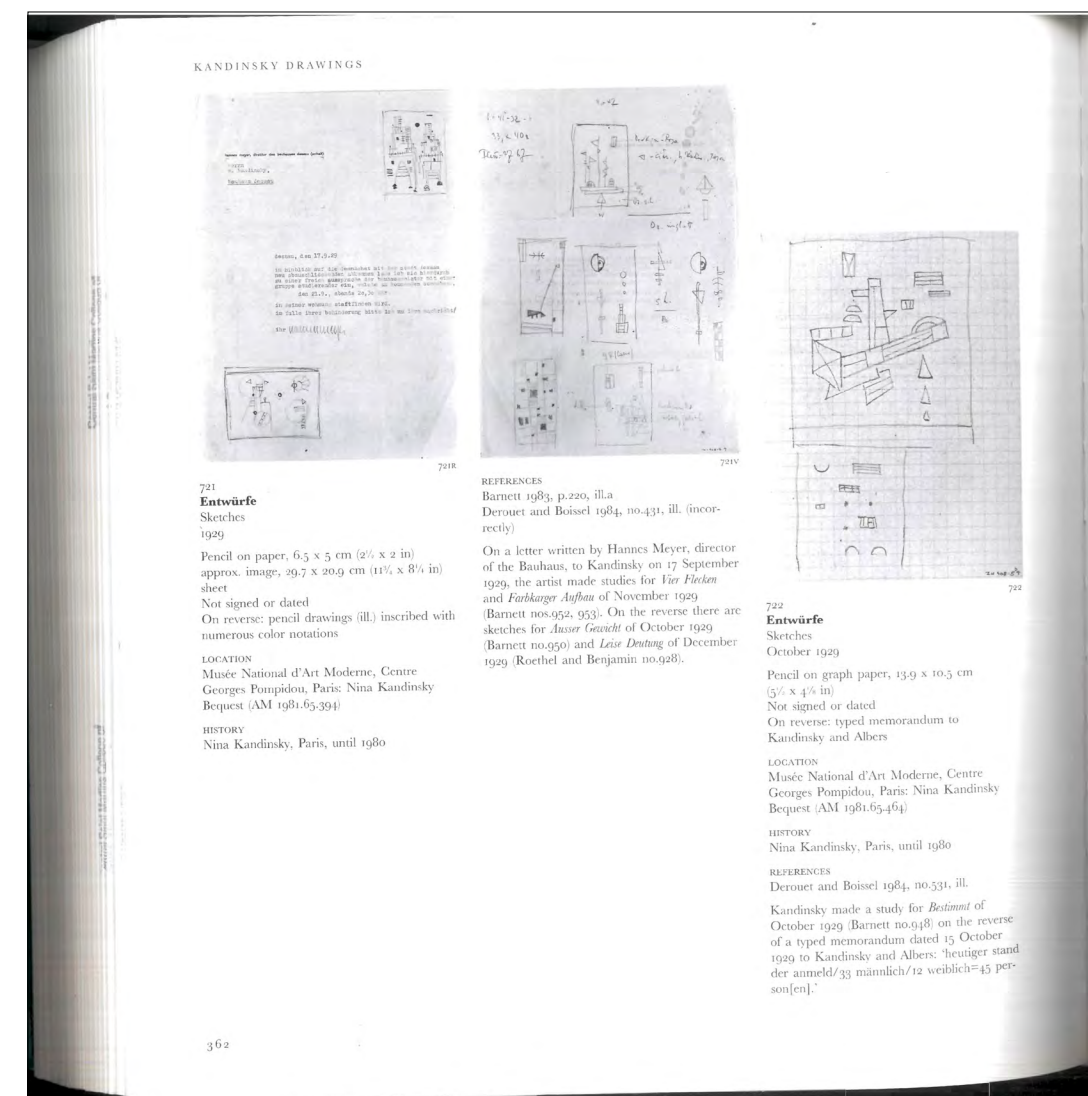
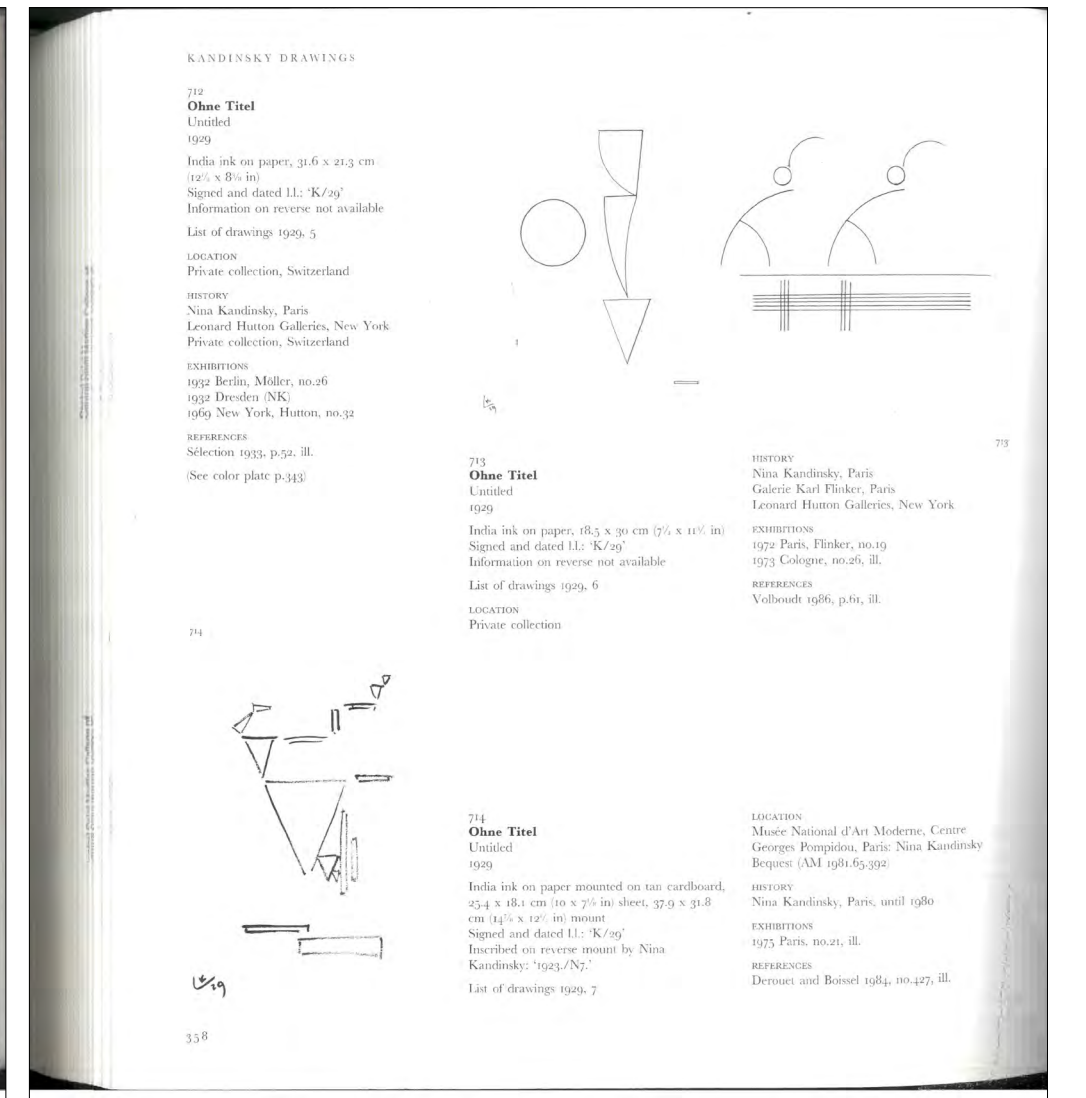
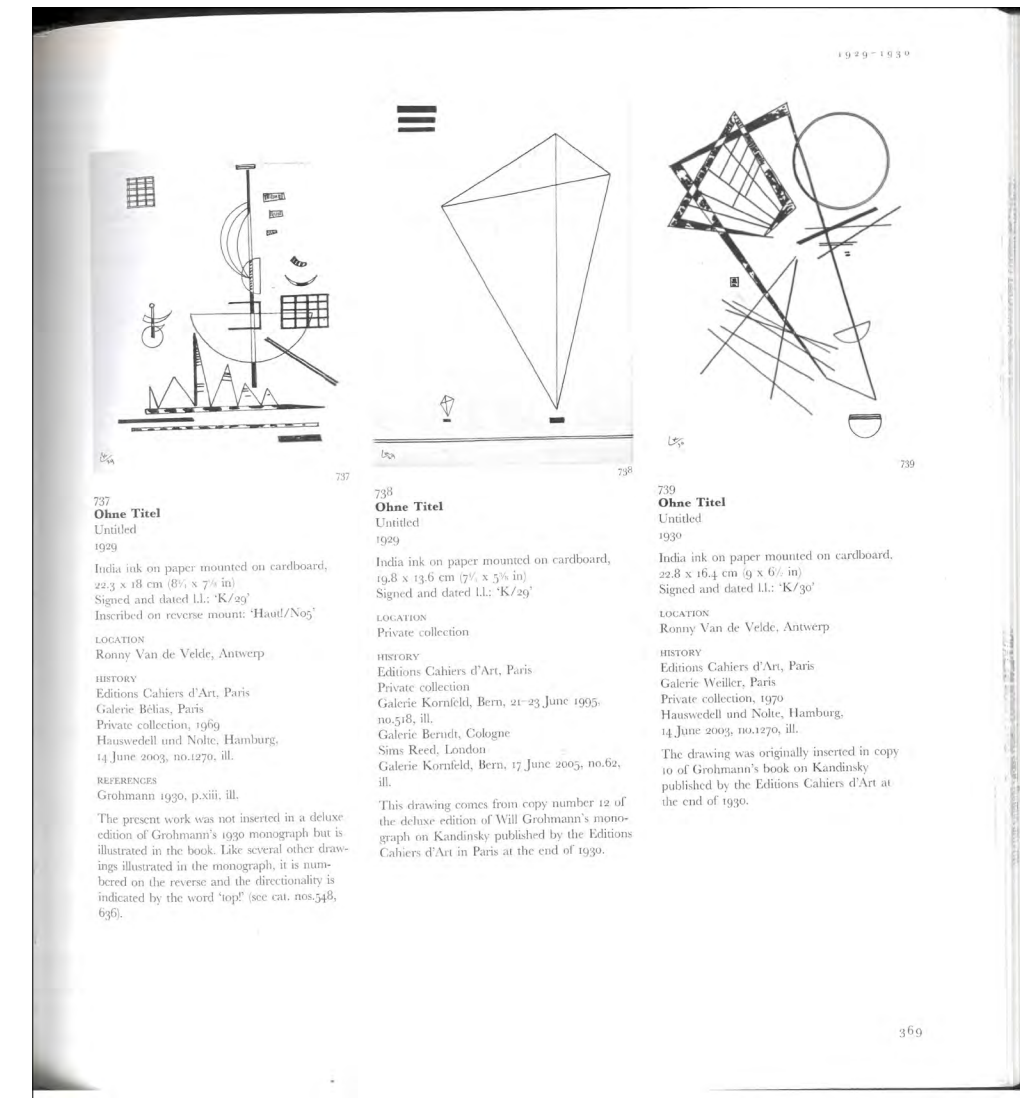
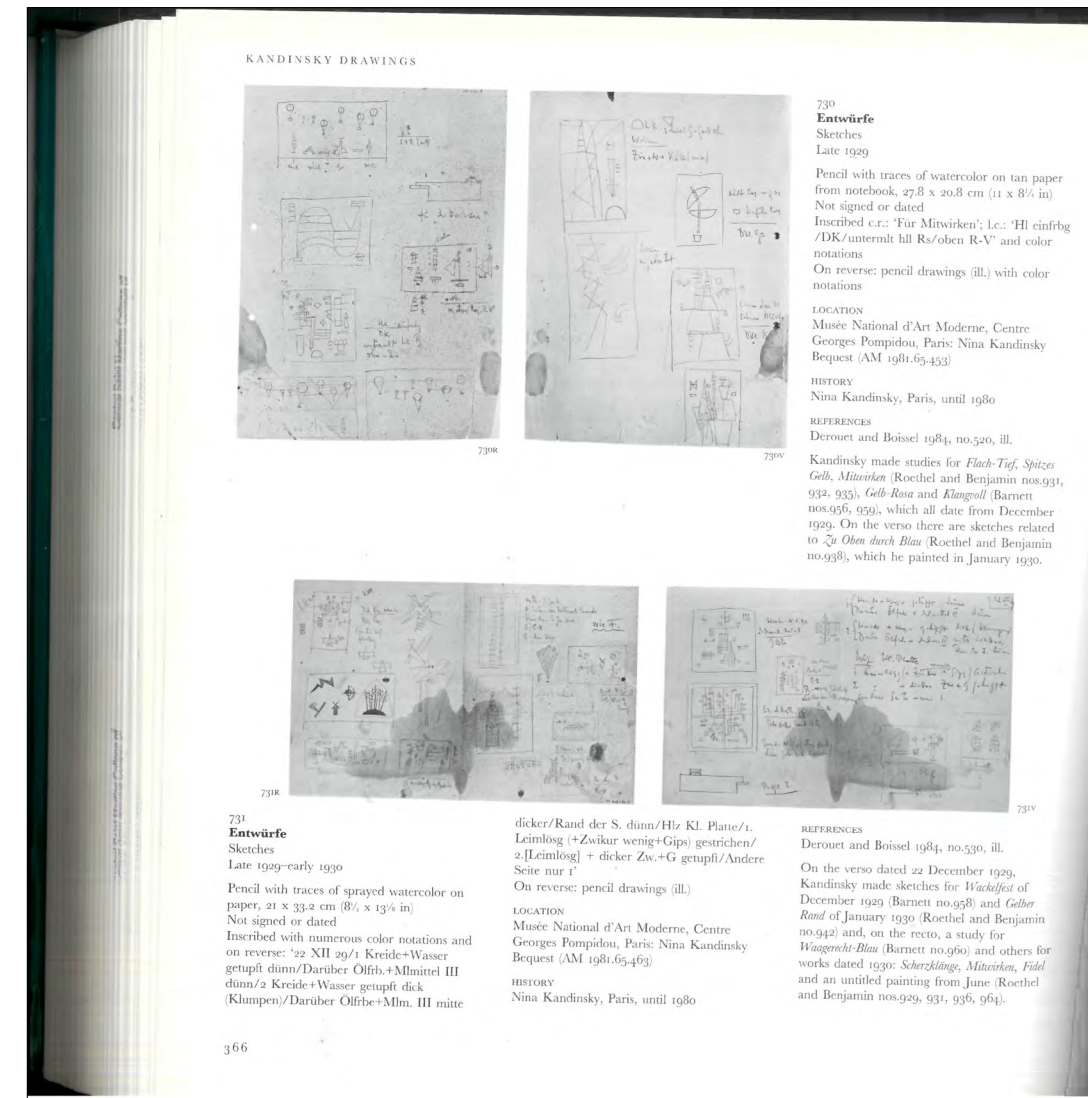


Position Through Contextualising

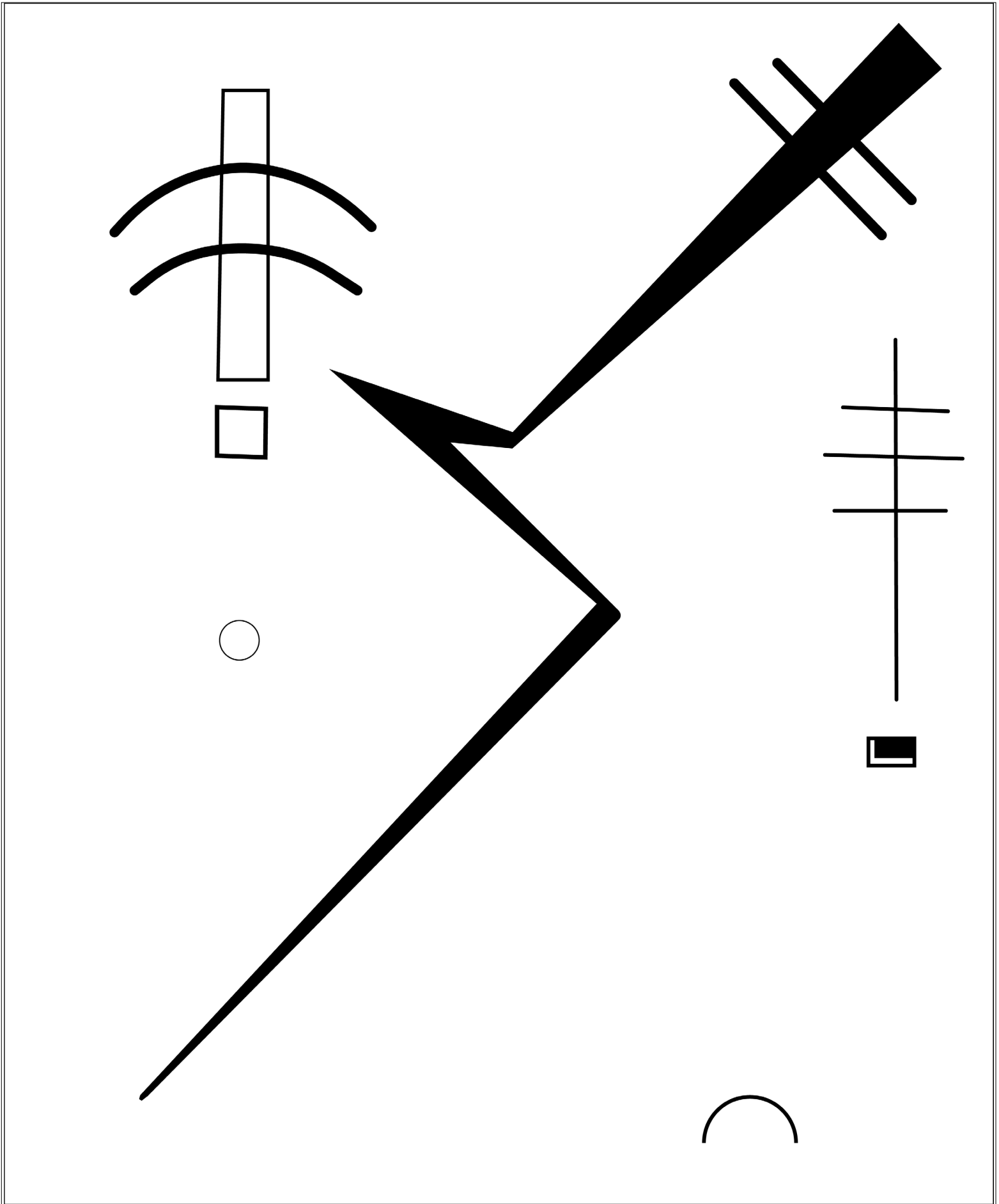
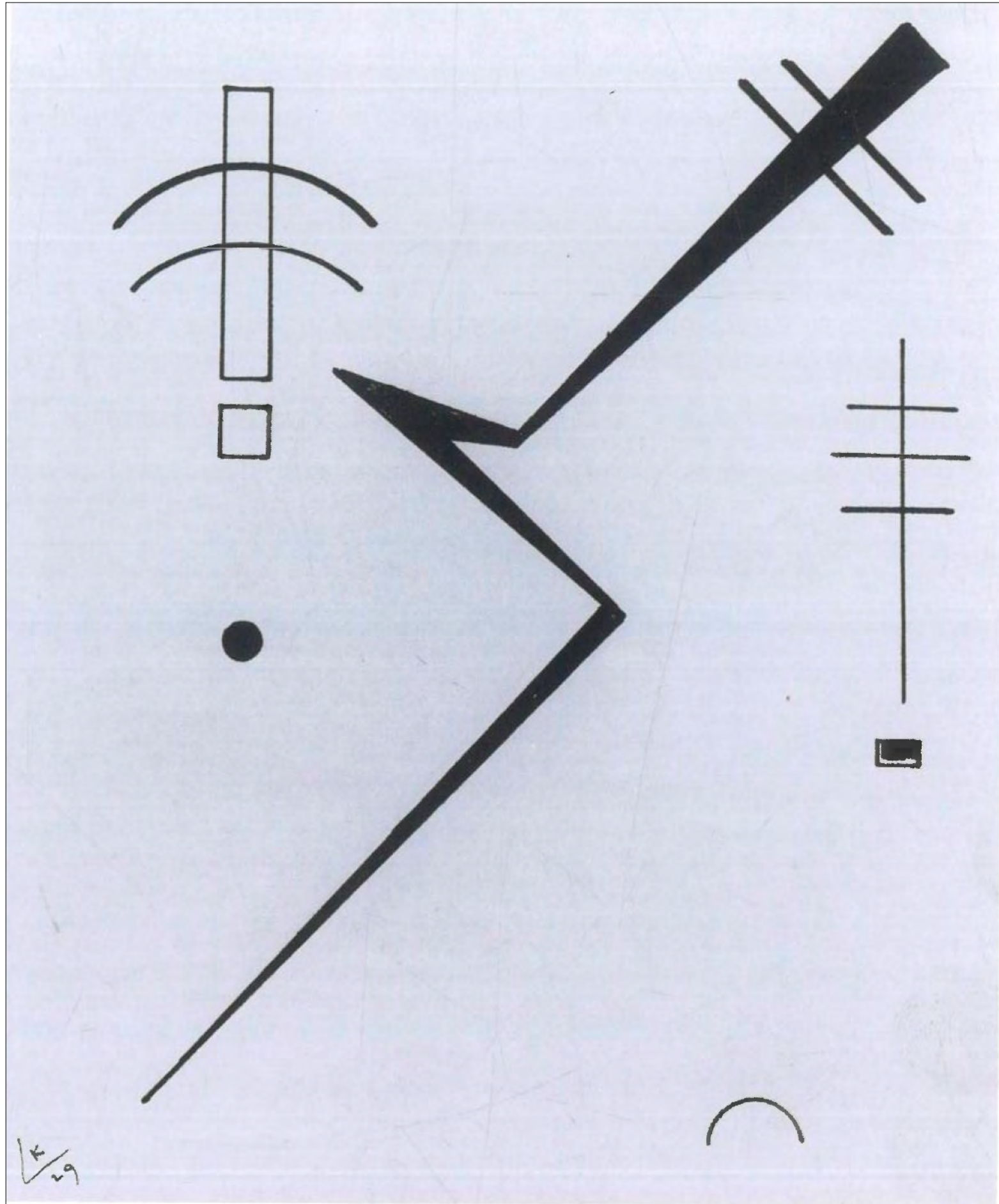
Before

&During
Test
&Ref.

&After



Position Through Contextualising



Before

&During
Test.

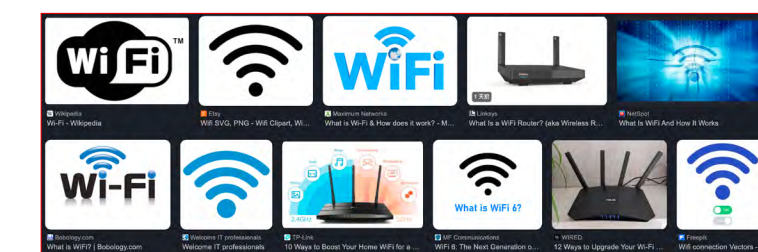
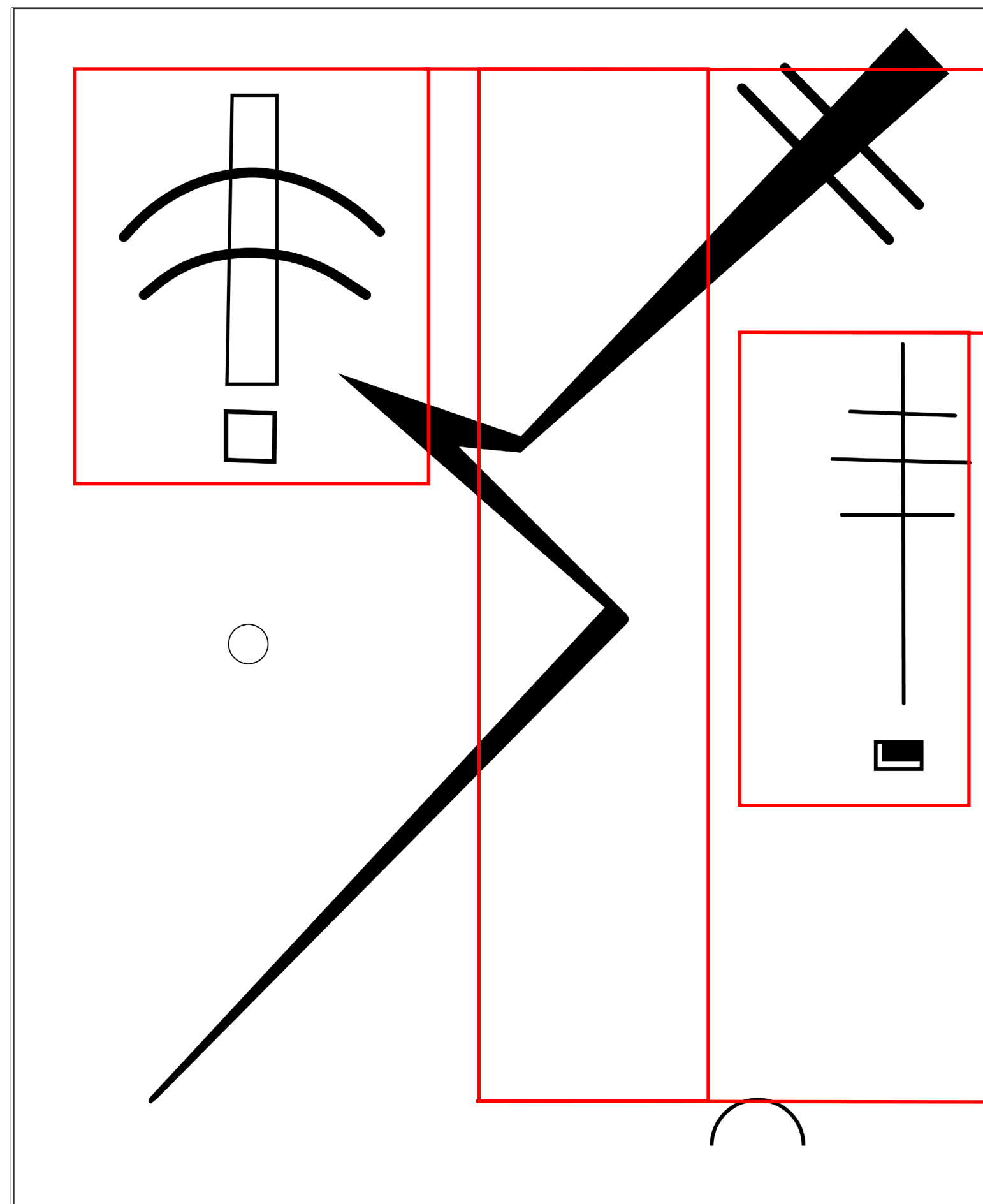
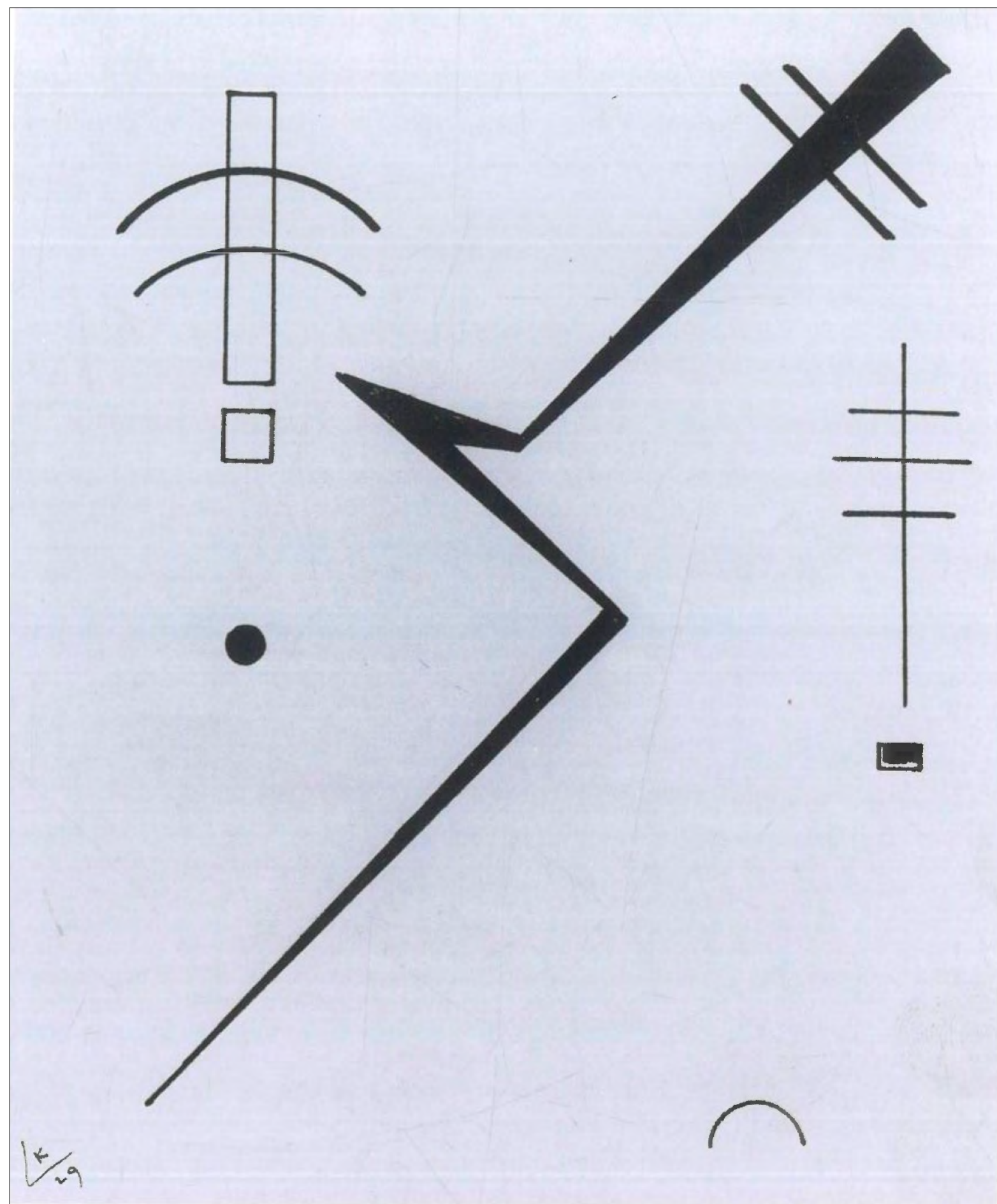
&After

Position Through Contextualising

Before

&During
Test.

&After

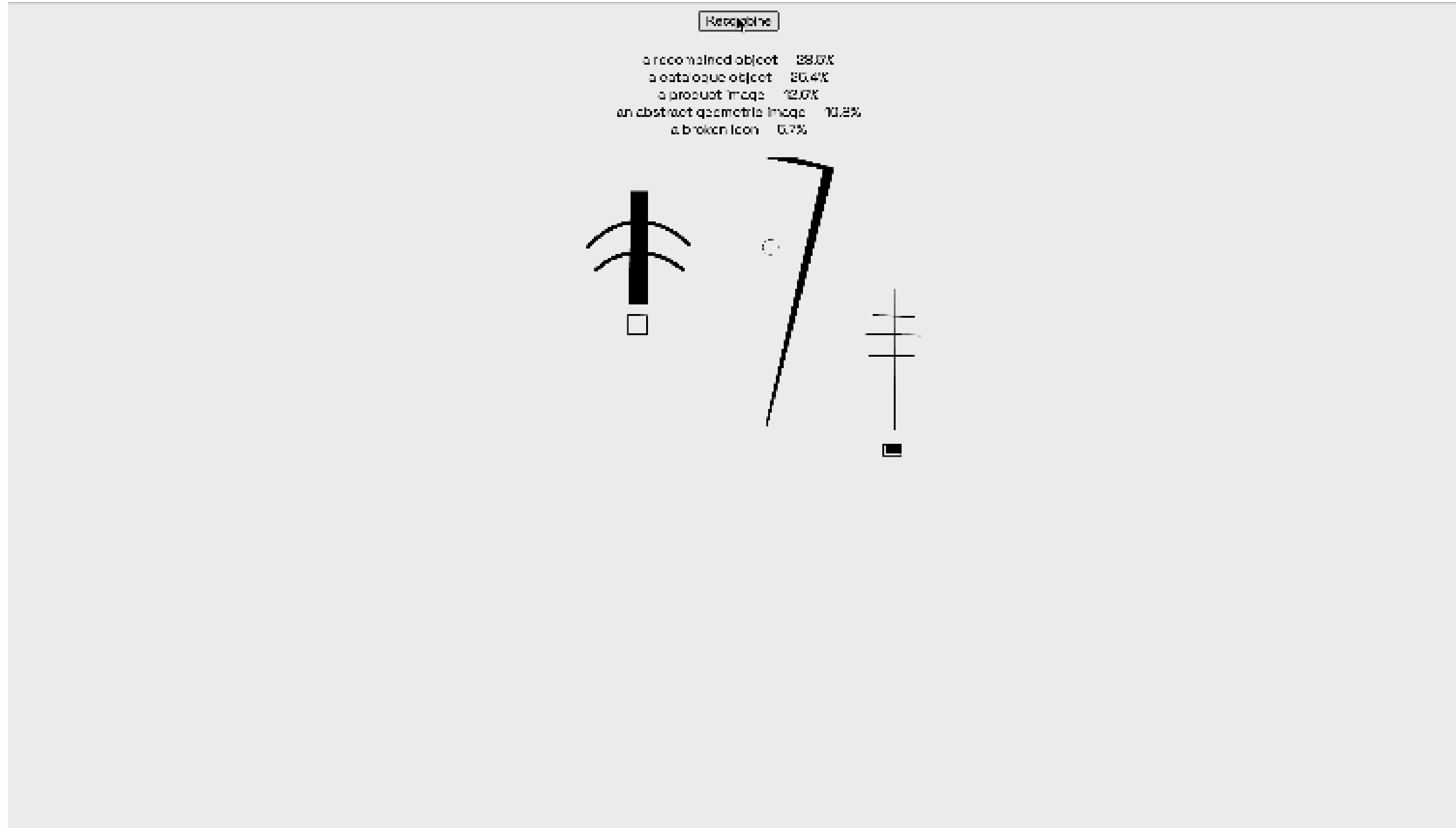


Position Through Contextualising

Before

&During
Test.

&After



```
clip_server.py > ...
12     allow_origins=["*"],
13     allow_credentials=True,
14     allow_methods=["*"],
15     allow_headers=["*"],
16 )
17
18 model = CLIPModel.from_pretrained("openai/clip-vit-base-patch32")
19 processor = CLIPProcessor.from_pretrained("openai/clip-vit-base-patch32")
20
21 labels = [
22     "a panda",
23     "a toy face",
24     "a smiling object",
25     "a fragmented face",
26     "an abstract geometric image",
27     "a recombined object",
28     "a distorted character",
29     "a catalogue object",
30     "a product image",
31     "a mask",
32     "a bird",
33     "a machine-readable image",
34     "a visual error",
35     "a synthetic image",
36     "a broken icon",
37 ]
38
39 @app.post("/clip")
40 async def analyze_clip(file: UploadFile = File(...)):
41     image_bytes = await file.read()
42     image = Image.open(io.BytesIO(image_bytes)).convert("RGB")
43
44     inputs = processor(
45         text=labels,
46         images=image,
47         return_tensors="pt",
48         padding=True,
49     )
```

What kind of visual culture does machine produce?

-Content

In terms of content: does machine recognition share the same visual vocabulary as human perception?

- early modern painting
- Renaissance painting
- Baroque painting
- 19th century academic art
- early modernism
- abstract art (early 20th century)
- Bauhaus design
- Constructivism
- Suprematism
- mid-century modernism
- postmodern graphic design
- contemporary digital art
- post-internet aesthetics
- AI-generated imagery

-Time











In terms of time: does machine vision inhabit the same historical moment as its human viewer?

- Renaissance painting
- Baroque painting
- abstract modernist painting
- early 20th century abstraction
- Bauhaus visual language
- contemporary digital graphic

-Scale

In terms of the individual: can machine vision account for the culturally specific memories that shape how each person reads an image?

- lightning
- antenna
- bird
- face
- tool
- weapon
- signal
- WIFI
- diagram
- music

● 红色	^	修改日期	大小	种类
 A Sea of Data_ Apopenia and Pattern (Mis-)Recognition - Journal #72.pdf	●	2026年4月27日 22:24	770 KB	PDF文稿
 Conditional design workbook_Maurer, Luna.pdf	●	昨天 22:42	553 KB	PDF文稿
 e-flux_Hito Steyerl_15.pdf	●	昨天 14:06	2.5 MB	PDF文稿
 Hito Steyerl, Mean Images, NLR 140 141, March June 2023.pdf	●	昨天 22:45	736 KB	PDF文稿
 Illuminations_Walter Benjamin; edited and with an introduction by Hannah Arendt; translated by Harry Zohn.pdf	●	昨天 23:03	1.2 MB	PDF文稿
 Invisible Images (Your Pictures Are Looking at You) – The New Inquiry.pdf	●	2026年4月27日 22:21	1.7 MB	PDF文稿
 Kandinsky Drawings_Barnett, Vivian Endicott.pdf	●	今天 00:06	3.6 MB	PDF文稿
 Machine-Readable Hito & Holly – Trevor Paglen_files	●	昨天 23:25	--	文件夹
 Machine-Readable Hito & Holly – Trevor Paglen.html	●	昨天 23:25	95 KB	HTML文本
 Memory, Metaphor, and Aby Warburg's Atlas of Images_Cgrustopher D. Johnson.pdf	●	昨天 15:02	4.3 MB	PDF文稿

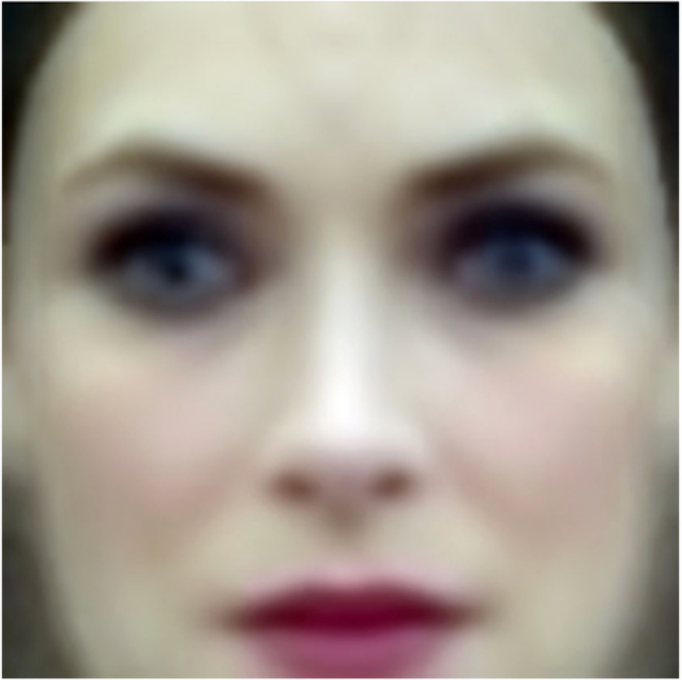
THE NEW INQUIRY

ESSAYS & REVIEWS FEATURES BLOGS AUDIO CURRENT ISSUE PAST ISSUES SHOP ABOUT Q LOGIN **SUBSCRIBE FOR \$2**

ESSAYS & REVIEWS

Invisible Images (Your Pictures Are Looking at You)

By [TREVOR PAGLEN](#) DECEMBER 8, 2016



In aggregate, AI systems have appropriated human visual culture and transformed it into a massive, flexible training set. The more images Facebook and Google's AI systems ingest, the more accurate they become, and the more influence they have on everyday life. The trillions of images we've been trained to treat as human-to-human culture are the foundation for increasingly autonomous ways of seeing that bear little resemblance to the visual culture of the past.

Americans as apes.

The point here is that if we want to understand the invisible world of machine-machine visual culture, we need to unlearn how to see like humans. We need to learn how to see a parallel universe composed of activations, keypoints, eigenfaces, feature transforms, classifiers, training sets, and the like. But it's not just as simple as learning a different vocabulary. Formal concepts contain

But over the last decade or so, something dramatic has happened. Visual culture has changed form. It has become detached from human eyes and has largely become invisible. Human visual culture has become a special case of vision, an exception to the rule. The overwhelming majority of images are now made by machines for other machines, with humans rarely in the loop. The advent of machine-to-machine seeing has been barely noticed at large, and poorly understood by those of us who've begun to notice the tectonic shift invisibly taking place before our very eyes.

Regardless of whether a human subject actually sees any of the 2 billion photographs uploaded daily to Facebook-controlled platforms, the photographs on social media are scrutinized by neural networks with a degree of attention that would make even the most steadfast art historian blush. Facebook's "DeepFace" algorithm, developed in 2014 and deployed in 2015, produces three-dimensional abstractions of individuals' faces and uses a neural network that achieves over 97 percent accuracy at identifying individuals-- a percentage comparable to what a human can achieve, ignoring for a second that no human can recall the faces of billions of people.

Invisible Images(Your Picture Are Looking at You)
Trevor Paglen, 2016

Invisible Images(Your Picture Are Looking at You)

Trevor Paglen, 2016

Paglen, T. (2016) 'Invisible Images: Your Pictures Are Looking at You', The New Inquiry, 8 December.

Paglen's central argument—that images are now primarily made for machines rather than humans—directly reframes the question my practice is asking. My YOLOv8 experiment, in which the model identifies birds in near-empty spaces or isolated fragments, initially seemed like machine error. Paglen's framework reinterprets this: the model is not misreading the image, it is reading it correctly within its own visual logic. This shifts my inquiry from "what does the machine get wrong?" to "what kind of seeing does the machine perform?" More significantly, Paglen's observation that machine vision produces a visual culture trained on human images but operating outside human time challenges how I think about my Kandinsky experiment. If the model's visual culture is formed through mass ingestion rather than historical accumulation, then feeding it a Kandinsky—a work whose meaning has shifted across decades—becomes a way to expose what the machine cannot hold: the fluidity of human visual culture.

In aggregate, AI systems have appropriated human visual culture and transformed it into a massive, flexible training set. The more images Facebook and Google's AI systems ingest, the more accurate they become, and the more influence they have on everyday life. The trillions of images we've been trained to treat as human-to-human culture are the foundation for increasingly autonomous ways of seeing that bear little resemblance to the visual culture of the past.

Americans as apes.



The point here is that if we want to understand the invisible world of machine-machine visual culture, we need to unlearn how to see like humans. We need to learn how to see a parallel universe composed of activations, keypoints, eigenfaces, feature transforms, classifiers, training sets, and the like. But it's not just as simple as learning a different vocabulary. Formal concepts contain

But over the last decade or so, something dramatic has happened. Visual culture has changed form. It has become detached from human eyes and has largely become invisible. Human visual culture has become a special case of vision, an exception to the rule. The overwhelming majority of images are now made by machines for other machines, with humans rarely in the loop. The advent of machine-to-machine seeing has been barely noticed at large, and poorly understood by those of us who've begun to notice the tectonic shift invisibly taking place before our very eyes.

Regardless of whether a human subject actually sees any of the 2 billion photographs uploaded daily to Facebook-controlled platforms, the photographs on social media are scrutinized by neural networks with a degree of attention that would make even the most steadfast art historian blush. Facebook's "DeepFace" algorithm, developed in 2014 and deployed in 2015, produces three-dimensional abstractions of individuals' faces and uses a neural network that achieves over 97 percent accuracy at identifying individuals-- a percentage comparable to what a human can achieve, ignoring for a second that no human can recall the faces of billions of people.