Echoes of the Prior: A Computational Phenomenology of Forgetting

GEGE GAO, University of Tübingen, Germany
BERNHARD SCHÖLKOPF*, Max Planck Institute for Intelligent Systems (Tübingen), Germany
ANDREAS GEIGER*, University of Tübingen, Tübingen AI Center, Germany

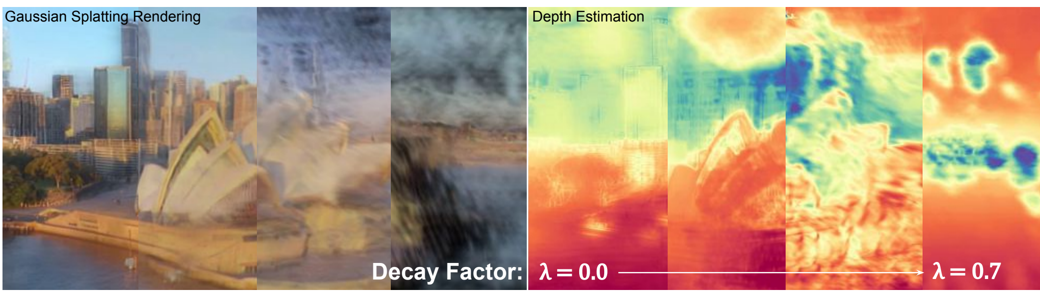

Fig. 1. ***Echoes of the Prior***. A computational simulation of the fragility of memory in the face of entropic loss.

Memory is not merely the storage of data; it is the scaffolding of reality. When biological memory fades, the world does not simply turn black; it regresses into an unrecognizable chaos. *Echoes of the Prior* is an interactive installation that attempts to visualize this subjective phenomenology of *forgetting*. By inducing controlled synaptic decay within a Feed-Forward 3D Reconstruction model, we create an artistic analogy for the erosion of the brain's predictive priors. We position the Neural Network not as a tool for engineering, but as a cognitive proxy - a *silicon brain* whose structural degeneration evokes the disorienting, poetic, and terrifying experience of losing one's grip on the world. Ultimately, we offer this framework as a catalyst, inviting the wider community to explore the uncharted potential of neuromorphic aesthetics in visualizing the fragility of intelligence. Interactive demo: https://decart-4d.github.io/.

CCS Concepts: • **Human-centered computing** → **Interaction design theory, concepts and paradigms**; • **Computing methodologies** → **Non-photorealistic rendering**; • **Applied computing** → *Media arts.*

Additional Key Words and Phrases: Computational Phenomenology, Glitch Aesthetics, Interactive Installation, 3D Gaussian Splatting, Neuromorphic Art, Real-time Performance System

*Both authors contributed equally to this research.

Authors' Contact Information: Gege Gao, University of Tübingen, Tübingen, Baden-Württemberg, Germany, gege.gao@uni-tuebingen.de; Bernhard Schölkopf, Max Planck Institute for Intelligent Systems (Tübingen), Tübingen, Baden-Württemberg, Germany, bs@tuebingen.mpg.de; Andreas Geiger, University of Tübingen, Tübingen AI Center, Tübingen, Baden-Württemberg, Germany, a.geiger@uni-tuebingen.de.

ACM 2577-6193/2026/0-ART0
https://doi.org/10.1145/nnnnnnn.nnnnnnn

ACM Reference Format:
Gege Gao, Bernhard Schölkopf, and Andreas Geiger. 2026. Echoes of the Prior: A Computational Phenomenology of Forgetting. *Proc. ACM Comput. Graph. Interact. Tech.* 0, 0, Article 0 (2026), 14 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction: The Invisible Process

"What I cannot create, I do not understand" is a famous quote by physicist Richard Feynman. While neuroscience can explain the *mechanism* of Alzheimer's – the plaques, the tangles, the synaptic pruning – it cannot show us the experience. We know that the patient forgets, but we do not know *what the world looks like* to them as it fades.

Modern Vision Transformers share structural principles with biological vision – distributed representation, hierarchical feature extraction, and predictive coding [Yamins and DiCarlo 2016]. While these correspondences are productive analogies rather than mechanistic equivalences [Bowers et al. 2023; Lindsay 2021], they invite a speculative proposition: if these models can approximate aspects of human visual cognition, then their pathology may serve as an *artistic analogy* for the pathology of the mind.

Echoes of the Prior operates on this speculative premise. We treat a state-of-the-art 3D Gaussian Splatting (3DGS) model not as a renderer, but as a proxy observer – a *silicon brain* whose components we map, by analogy, onto cognitive functions. By surgically introducing noise into its semantic priors (analogous to *long-term memory*) and image encoder (analogous to the *sensory cortex*), we force the machine to reconstruct reality through a damaged architecture. The resulting imagery – melting sky, ghostly architectures, and dissolving forms – serves as a visual translation of fading mind, making the invisible internal experience of forgetting visible, visceral, and shared.

We title our installation ***DecArt*** (a portmanteau of *Decay* and *Art*). Pronounced like the French philosopher *Descartes*, we elaborate on the philosophical implications of this choice in Sec. 5.1.

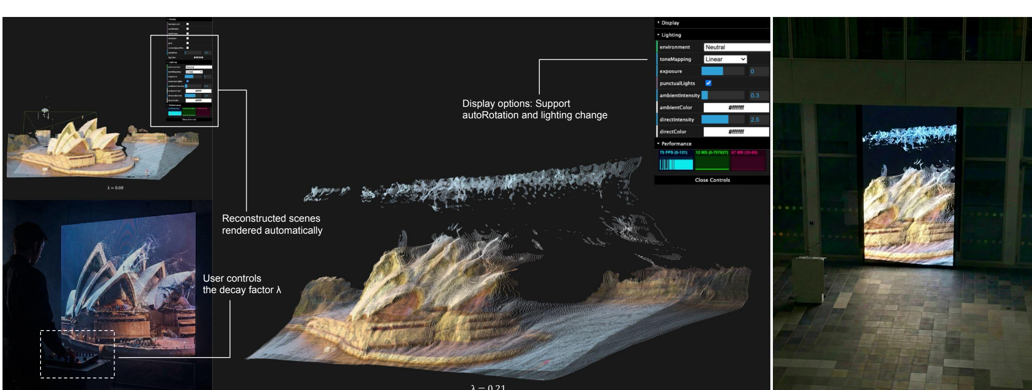

Fig. 2. **Installation Design of *Echoes of the Prior*.** Snapshot of the real-time visual feedback ($\lambda = 0.21$). The viewport displays the immediate structural dissolution of the 3D scene as the user adjusts the decay factor.

2 Related Work

We position *Echoes of the Prior* at the intersection of computational aesthetics and theoretical neuroscience, employing volumetric 3D reconstruction as a perceptual medium.

2.1 From Machine Vision to Machine Hallucination

Prior AI art has visualized the *constructive* capacity of neural networks: DeepDream [Mordvintsev et al. 2015] and Neural Style Transfer [Gatys et al. 2016] expose internal CNN representations, while artists like Refik Anadol [Anadol 2021] treat datasets as fluid memories, visualizing the machine's *learning* process. This constructive impulse extends from visual representation to conceptual generation, with recent work biasing LLMs toward novel-yet-coherent concept combinations [Hernandez et al. 2025]. Our work inverts this paradigm, visualizing not the *genesis* of intelligence but its *entropy* (forgetting). We align with **Glitch Art** [Menkman 2011], which treats technological failure as revelation, but go further: rather than surface-level pixel corruption (e.g., datamoshing), we perform **structural pathology**: eroding internal cognitive weights and predictive priors to create a "Deep Glitch" that emulates neurodegeneration.

2.2 Art, Memory, and Neurodegeneration

Our work situates itself within a broader artistic tradition engaging with cognitive decline. William Utermohlen's self-portraits, painted after his Alzheimer's diagnosis, document a shift from figurative realism toward abstraction – a first-person record of representational breakdown [Crutch et al. 2001]. Similarly, Willem de Kooning's late paintings have sparked debate about whether their luminous simplicity reflects artistic evolution or disease [Espinel 1996]. These cases raise complex questions about authorship and artistic intent (see Sec. 5.2). In interactive media, Memo Akten's *Learning to See* [Akten et al. 2019] reveals how constrained training data causes a network to hallucinate misreadings of live input, while Anderson and Huang's VR installation *Chalkroom* [Anderson and Huang 2017] immerses viewers in dissolving narratives evoking the fragility of linguistic memory. Art-science initiatives such as *Created Out of Mind* [Crutch et al. 2019] have further brought together artists, neuroscientists, and people with dementia to reshape public understanding.

Our contribution differs in a key respect: rather than depicting memory loss through narrative, metaphor, or the artist's declining hand, we engineer it directly within the machine's cognitive architecture – a structural pathology of internal representations.

2.3 Spatializing the Past: 3D Reconstruction as a Cognitive Medium

While photography freezes a moment, 3D reconstruction captures a spatial environment. We argue that **volumetric representation** is intrinsically more akin to biological memory than flat imagery: memory encoding is deeply intertwined with spatial navigation [O'Keefe and Dostrovsky 1971]. To remember is not just to view an image; it is to inhabit a space.

The evolution from photogrammetry to NeRF [Mildenhall et al. 2020] and 3D Gaussian Splatting (3DGS) [Kerbl et al. 2023] reflects a shift toward probabilistic, organic representations. We utilize 3DGS not merely for real-time rendering, but for its aesthetic properties: unlike generative video models that hallucinate temporal continuity, 3DGS reconstructs *spatial* continuity. By subjecting this representation to decay, we visualize the disintegration of spatial context itself – simulating how a patient might lose the *where* and *how* of a memory.

3 Methodology: Simulating Cognitive Decay

Modern foundational models [Oquab et al. 2024; Siméoni et al. 2025] do not store pixels; they store semantic concepts in a high-dimensional latent space. To simulate the phenomenology of forgetting, we do not merely degrade the image pixels; we perform surgical interventions on the cognitive architecture of the machine.

Our approach creates a structural analogy between a state-of-the-art 3D reconstruction model [Lin et al. 2025] based on feed-forward 3D Gaussian Splatting (FF-3DGS), and the Predictive Coding framework [Friston 2010] of the human brain, synthesizing reality through the fusion of two distinct neural pathways: *memory* (prior) and *evidence* (sensation).

3.1 The Silicon Proxy: Two Streams Mapping Mind to Code

Can a machine possess memory and cognition? If so, what does it look like?

The act of perceiving the world is a negotiation between what our eyes see – the evidence from *Sensory Cortex*, and what our brains expect – what neuroscience terms the *Engram*, the physical trace of memory. We draw a functional analogy between this biological duality and the specific components of our computational architecture (see Sec. 6.1 for a discussion of the boundaries of this mapping):

3.1.1 The Sensory Stream: Processor Weights. Drawing a functional analogy to the biological retina and visual cortex (V1), our model employs a shallow CNN to process raw pixel inputs, capturing the texture and local geometry of the *Now*. This sensory signal will serve as a reinforcement to the memory, and together with memory, will be injected into a fusion block for spatial perception and reconstruction [Ranftl et al. 2021].

3.1.2 The Memory Stream: Semantic Features. In an analogy to how biological memory stores concepts rather than raw data, our system utilizes vision foundational models to encode high-dimensional semantic priors.

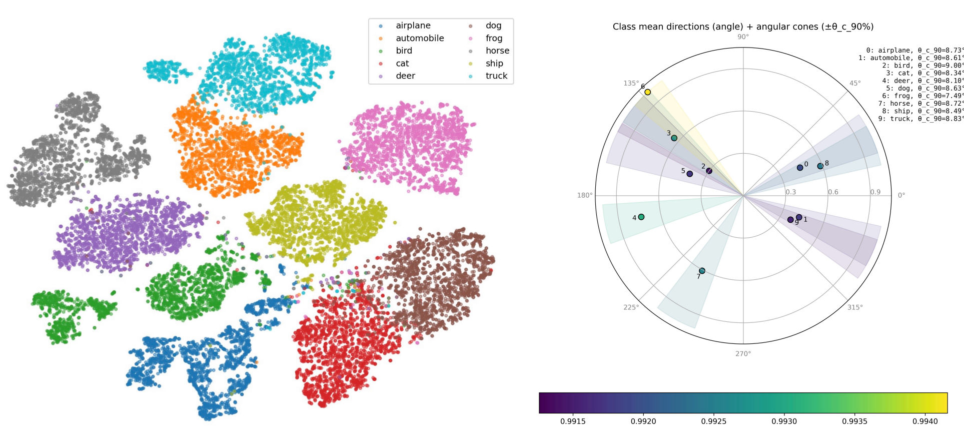

Fig. 3. t-SNE visualization of the DINOv2 latent space. **(Left)** Semantic concepts cluster into distinct manifolds, representing the structured *healthy* memory baseline. **(Right)** Cosine similarity metrics confirm that intra-class features are more closely aligned than inter-class features, validating the semantic coherence of the latent space.

Specifically, we use ***DINOv2*** [Oquab et al. 2024] as the *prior* producer in our 3D reconstruction model. As shown in Fig. 3, in its 1024D latent space, features of the same semantic concept cluster into distinct manifolds. These representations serve as the ***healthy baseline***: the long-term memory of a healthy silicon brain – structured, dense, and semantically coherent.

In a healthy state, perception is the successful integration of these two streams. Our installation visualizes what happens when these streams decay asynchronously.

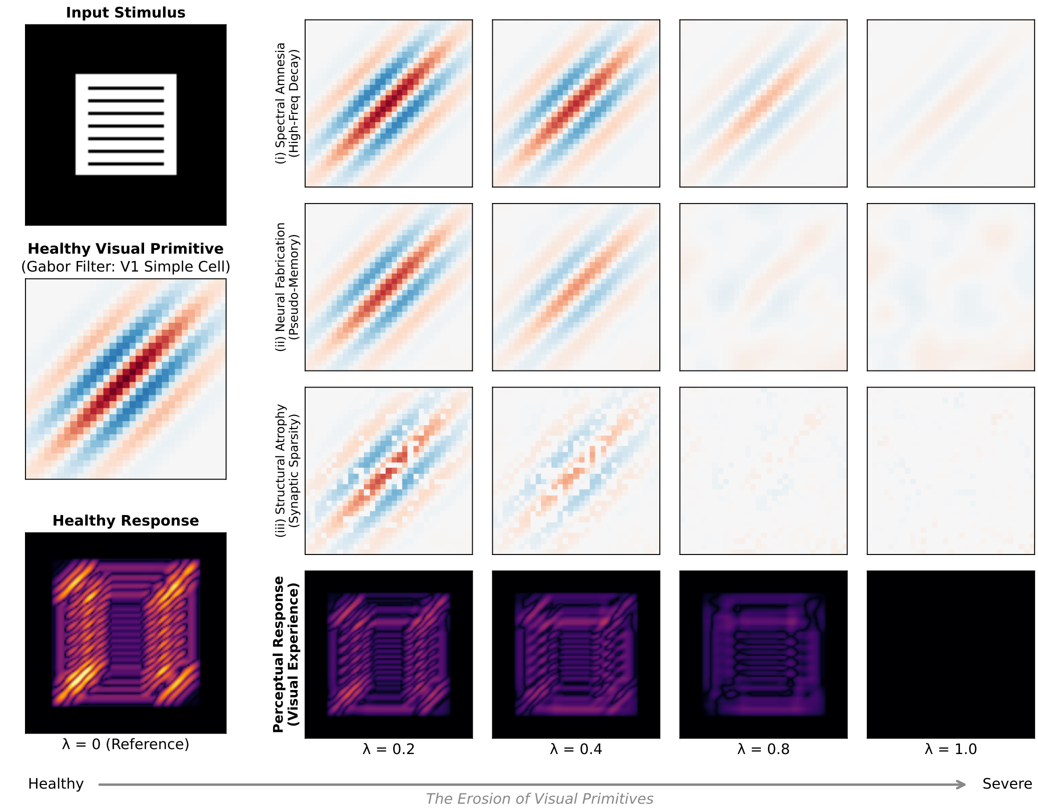

Fig. 4. **Simulating Sensory Decay $\mathcal{G}_{\text{sense}}$. (Left) Healthy Baseline ($\lambda = 0$):** Response of a standard Gabor filter (V1 primitive) to a sharp stimulus. **(Right) Pathology:** As λ increases, the filter degrades via three mechanisms: **(i) Spectral Amnesia** (loss of high-frequency acuity/sharpness); **(ii) Neural Fabrication** (injection of coherent noise/hallucinations); and **(iii) Structural Atrophy** (synaptic sparsity leading to signal loss).

3.2 The Mechanics of Forgetting: Two Decay Functions Simulating Biological Entropy

The **Free Energy Principle** [Friston 2010] posits that the brain continuously predicts sensory input based on internal priors, and that perceptual distortion arises when this predictive machinery fails [Corlett et al. 2019]. Human aging disrupts this process along two axes – sensory degradation and cognitive decline – both driven, according to the *Neural Gain Theory of Aging* [Li et al. 2001], by a reduction in signal-to-noise ratio within neuromodulatory systems. To simulate these parallel failures, we introduce two entropy operators, $\mathcal{G}_{\text{sense}}$ and $\mathcal{F}_{\text{mem}}$.

3.2.1 Sensory Decay: The Erosion of Visual Primitives $\mathcal{G}_{sense}$. While the semantic stream deals with concepts, the sensory stream deals with *visual primitives* – edges, textures, and local gradients. A naive simulation of sensory decay would simply degrade the input signal (i.e., pixel-level corruptions), simulating external visual impairments. However, biological aging is a neurological process where the processing circuitry itself degrades.

To simulate this, we perform a perturbation on the CNN weights that acts as the machine's primary visual cortex (as described in Sec. 3.1.1). We implement a compound decay function $\mathcal{G}_{\text{sense}}$ that creates a diseased clone of the sensory encoder, modeling three distinct biological failures:

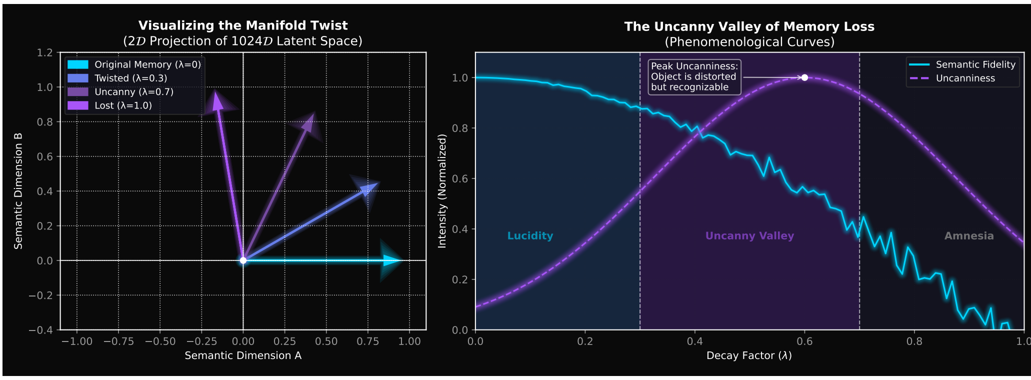

Fig. 5. **Topology of Memory Decay. (a) Manifold Twist:** Visualization of the orthogonal rotation $\mathbf{R}(\lambda)$. As λ increases (blue $\rightarrow$ purple), the feature vector rotates away from its semantic alignment while preserving its norm. **(b) Computational Uncanny Valley:** A computationally defined model of visual *Uncanniness* (purple), defined as the product of structural integrity (norm) and semantic distortion ($1 -$ cosine similarity). The peak ($0.3 < \lambda < 0.7$) corresponds to the zone where the representation is structurally strong yet semantically twisted. Perceptual research predicts that such feature-inconsistent stimuli are maximally disturbing [Moore 2012; Seyama and Nagayama 2007]; formal viewer validation remains future work.

(i) Spectral Amnesia (High-Frequency Domain Decay): Biological aging often begins with the loss of high-frequency acuity [Owsley et al. 1983]. To simulate this, we transform the convolutional kernels into the frequency domain via Fast Fourier Transform (FFT) and then apply a radial mask that selectively dampens high-frequency components relative to the decay factor λ [1]. This targets the functional architecture of V1 cortex, where simple cells act as spatiotemporal frequency filters [De Valois et al. 1982; Hubel and Wiesel 1962]. By dampening high frequencies in the kernel weights directly, rather than the input image, we degrade the receptive fields themselves. **The Effect:** As shown in Fig. 4, the machine does not see a blurry image; it forgets how to detect sharpness, simulating the onset of cortical cataracts.

(ii) Neural Fabrication (Pseudo-Memory Injection): The brain abhors a sensory vacuum. When sensory precision is low, the brain over-relies on top-down priors, leading to phantom stimuli [Powers et al. 2017] (e.g., Charles Bonnet Syndrome [Reichert et al. 2013]). We model this by injecting spatially smoothed Gaussian noise into the weights. This approximates the "correlated noise" observed in neural circuits when inhibitory control weakens [Shadlen and Newsome 1998]. Unlike white noise, this smooth noise creates coherent but fictitious patterns in the convolution filters. **The Effect:** This causes the machine to *hallucinate* textures that do not exist, overlaying the real world with a dream-like, phantom grain.

(iii) Structural Atrophy (Synaptic Sparsity): Finally, we simulate neuronal death by applying a stochastic mask to the *Weight* matrix, zeroing out connections based on λ, while decaying *Bias* terms toward zero. **The Effect:** The reconstructed 3D world becomes translucent and ethereal, a visual metaphor for a consciousness slowly detaching from physical reality.

3.2.2 Memory Decay: Manifold Distortion and Cognitive Entropy $\mathcal{F}_{mem}$. Biological forgetting is often a structural collapse of meaning, where the world does not just fade into black but turns disturbingly wrong. In the framework of **Attractor Dynamics** [Hopfield 1982], memories are stored as stable states within a high-dimensional manifold; cognitive decline destabilizes these

[1]The decay factor in $\mathcal{G}_{\text{sense}}$ can be distinct from the one in memory decay $\mathcal{F}_{\text{mem}}$ for asynchronous control, and we keep the same symbols here to avoid complicating the notation.

attractors. To simulate the phenomenology of *Semantic Dementia* [Warrington 1975] and *Associative Agnosia* [Farah 2004], we model the machine memory as a high-dimensional vector field (as introduced in Sec 3.1.2), where the semantic identity of an object is defined mainly by the direction of its **feature vector** $\mathbf{z}$ [Wang and Isola 2020], and introduce a *geometric perturbation* within the high-dimensional memory space.

The Distortion: We model *cognitive decline* acts as a force that ***twists*** the memory manifold. Using a continuous orthogonal rotation $\mathbf{R}(\lambda)$, we misalign semantic vectors while preserving their norm:

$$\mathbf{R}(\lambda) = (1-\lambda)\cdot \mathbf{I} + \lambda \cdot \mathbf{Q}, \tag{1}$$

where $\mathbf{Q}$ is a random orthogonal matrix generated via QR decomposition [Mezzadri 2007], $\mathbf{I}$ is the identity matrix. As the decay factor λ increases, the semantic feature vectors $\mathbf{z}$ are smoothly rotated away from their ground truth alignment.

$$\mathbf{z}_{\text{rot}} = \mathbf{z}\cdot \mathbf{R}(\lambda)^T. \tag{2}$$

This simulates a brain state where signal strength remains high but semantic connectivity is twisted: the machine reconstructs high-confidence geometries, yet familiar objects are rendered as *hallucinated* topological variants. As shown in Fig. 1, a construction might retain its texture but assume the geometry of a liquid.

The Hybrid Entropy: Furthermore, to capture the full spectrum of memory decay, we combine this geometric twisting with a secondary entropy term, reflecting the irreversible loss of information:

$$\mathcal{F}_{mem}(\mathbf{z},\lambda,\gamma) = \underbrace{(1-\gamma)\cdot \mathbf{z}_{\text{rot}}}_{\text{Distortion}} + \underbrace{\gamma\cdot \mathcal{N}(\mathbf{z})}_{\text{Entropy}} \tag{3}$$

where $\mathbf{z}, \mathbf{z}_{\text{rot}} \in \mathbb{R}^D$ are feature vectors (here $D = 1024$), function $\mathcal{N}(\mathbf{z})$ generates a stochastic noise vector, $\mathcal{N}(\mathbf{z}) \approx \epsilon \cdot \|\mathbf{z}\|$, where $\epsilon \sim \mathcal{N}(0, I)$. By fusing manifold rotation ($1-\gamma = 0.7$) with stochastic noise ($\gamma = 0.3$), the system visualizes a specific trajectory of forgetting: as shown in Fig. 5, memory first becomes confused (twisted geometry) before it eventually becomes lost (dissolved form). This algorithmic choice allows us to visualize what we term a *computational uncanny valley* of memory loss – a regime ($0.3 < \lambda < 0.7$) where the product of preserved structural integrity and semantic distortion peaks. We note that this "valley" is defined computationally, not through measured viewer responses; however, it is consistent with perceptual research showing that stimuli exhibiting high structural fidelity but semantic inconsistency produce maximal discomfort [Moore 2012; Seyama and Nagayama 2007] – an effect Freud [Freud 1919] described as *das Unheimliche*, the disturbing quality of the familiar-yet-wrong. This zone conceptually mirrors the phenomenology of Associative Agnosia [Farah 2004] and Semantic Dementia [Warrington 1975], where the patient perceives form without meaning.

3.3 The Selective Forgetting: Object-Oriented Amnesia

While global decay simulates a systemic cognitive decline (e.g., dementia), human forgetting is often highly selective [Warrington 1975]. We frequently lose the semantic grasp of specific entities – a face, a name, or an object – while the surrounding reality remains intact. To simulate this *dissociative amnesia*, we introduce a targeted decay mechanism that allows the system to surgically erode the memory of specific semantic categories.

3.3.1 Semantic Grounding and Masking. We employ a two-stage pipeline to isolate the *memory trace* of a specific object. First, we utilize an open-vocabulary detection model, ***Grounded-SAM*** [Ren et al. 2024], to generate a pixel-level binary mask $M_{pixel} \in \{0, 1\}$ based on a textual prompt (e.g., "woman

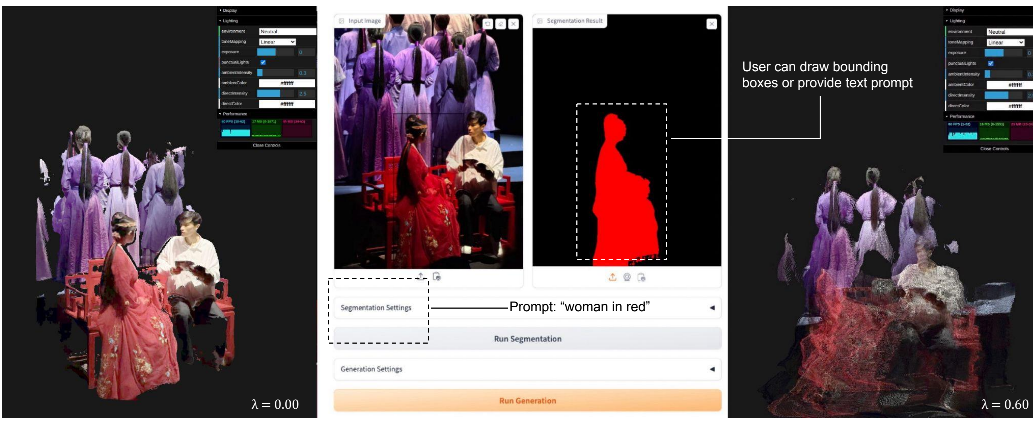

Fig. 6. **Visualizing Selective Forgetting.** Targeted entropy injection using a text prompt ("woman in red"). **(Left)** Original reconstruction. **(Right)** Decayed result ($\lambda = 0.6$). The targeted entity undergoes topological dissolution, effectively ghosting out of reality while the surrounding context remains intact.

in red”). This mask represents the object’s footprint in the visual field. As shown in Fig. 6, users can obtain the mask for specific entities by providing a text prompt or drawing bounding boxes. A naive approach would be to apply the mask to the input image directly (pixel-level inpainting), but this would result in a superficial “black hole” rather than a cognitive loss. Instead, we perform the intervention within the latent feature space of the memory.

3.3.2 Latent Space Surgery. To extract the prior about the input images, the vision foundation model [Oquab et al. 2024] first processes images as a sequence of patch tokens. To align with the spatial shape of these patch tokens, we down-sample the pixel mask M_{pixel} to create a discretized token-wise mask $M_{\text{token}} \in \{0, 1\}$.

We then modify the forward pass of our system: we perform decay function Eq. (3) only into the tokens corresponding to the target object:

$$\mathbf{z}' = M_{\text{token}} \cdot \mathcal{F}_{mem}(\mathbf{z}, \lambda, \gamma) + (1 - M_{\text{token}}) \cdot \mathbf{z}, \tag{4}$$

where $\mathbf{z}$ is the original clear memory.

By perturbing only the latent tokens of the target, we achieve a form of *selective agnosia*. As shown in Fig. 6, the surrounding environment remains structurally sound, while the targeted entity undergoes topological dissolution – its geometry melts and it effectively ghosts out of existence. This visualizes how a patient may perceive a room perfectly but fail to resolve a specific object within it (see Sec 6.3 for further discussion).

4 Real-time Installation and Implementation Details

4.1 User Interface and Integration Formula

As shown in Fig. 7, the system accepts visual input (live camera feed or pre-loaded images) and processes it through a 3D reconstruction backbone, ***Depth Anything 3*** [Lin et al. 2025]. The user-defined decay factor λ **intercepts** reconstruction via two parallel paths: **Sensory Decay** (spectral filtering, Sec. 3.2.1) and **Memory Decay**, which optionally integrates **Selective Forgetting** via ***Grounded-SAM*** [Ren et al. 2024] to inject entropy into specific object tokens. The corrupted latents are fused and decoded into a fractured 3D scene (Gaussian Splats or point cloud), rendered in real-time via a WebGL viewer at 60FPS: as the user slides the fader, the world melts instantly.

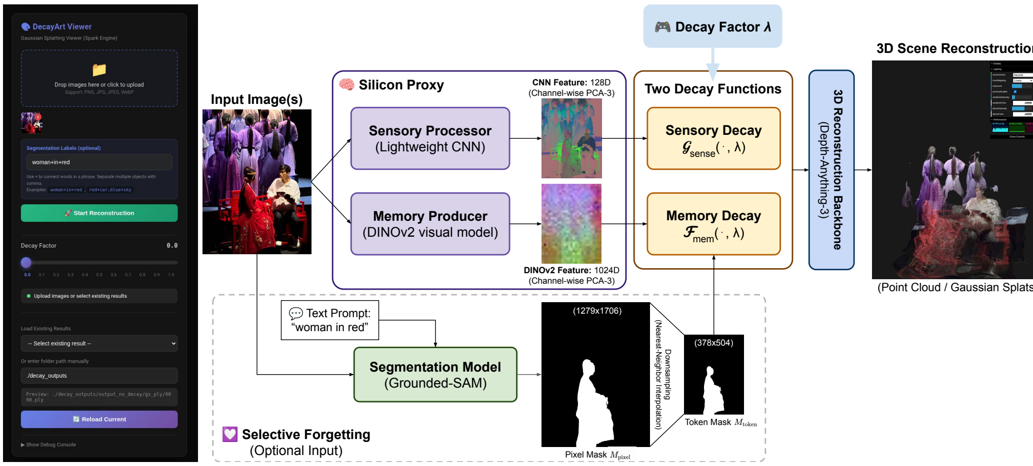

Fig. 7. The System Workflow of *Echoes of the Prior*.

The interface centers on a single slider ($\lambda \in [0, 1]$), transforming an abstract algorithm into a proprioceptive experience. Unlike static video, the rendered scene supports full orbit controls, allowing users to rotate and zoom into the decaying structures.

4.2 Image Acquisition Strategy

The system supports dual inputs: a **Curated Archive** of semantic archetypes for immediate engagement, and a **Personal Injection** feature allowing users to upload personal photos or video. We also developed an ***Interactive Streaming Interface*** (Fig. 9) that integrates Grounded-SAM for prompt-based masking with a batch pre-computation pipeline, enabling seamless 4D decay visualization without inference latency.

4.3 The Experience Journey

The viewer enters a darkened space and approaches a large display showing a photorealistic 3D scene, with a tablet as the control console. At $\lambda = 0$, the scene appears pristine; it may be drawn from a *Curated Archive* or from the viewer's own video (*Bring Your Own Memory* mode).

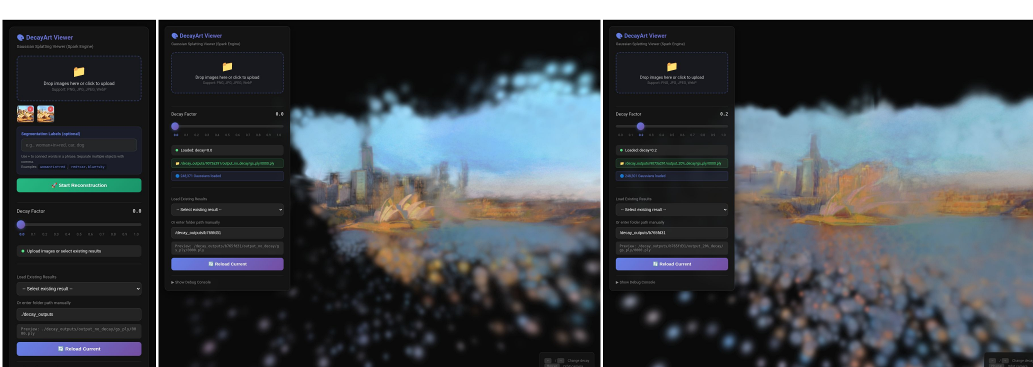

Fig. 8. **The Interactive Interface.** The web-based GUI enables real-time interaction with the decay process.

Two complementary interaction modes are offered. In *Global Decay*, the slider governs λ directly: edges soften, textures hallucinate, and geometry warps – progressing from subtle shimmer ($\lambda \approx 0.1$) through topological distortion ($0.3 < \lambda < 0.7$) to total dissolution ($\lambda > 0.8$). The slider is bidirectional, letting the viewer retreat from oblivion at any point. In *Selective Forgetting*, the viewer names a specific element; the system restricts entropy injection to its latent tokens alone, producing a world perfectly remembered except for one conspicuous absence.

Sonic Dimension. For the exhibition setting, we plan a generative soundscape coupled to λ: spectral filtering strips harmonics, granular synthesis introduces stochastic micro-events, and amplitude attenuates toward silence – extending the decay into a coherent multimodal experience.

Aesthetic Character: The Deep Glitch. The resulting visual language – which we term *semantic liquefaction* – is distinct from both traditional Glitch Art (broken data) and generative AI aesthetics (fabricated coherence). Because the decay operates on cognitive representations, objects retain their texture while their geometry undergoes topological metamorphosis: a building does not pixelate but *melts*. The volumetric medium enriches this further – orbiting reveals non-uniform corruption, with some vantage points exposing catastrophic voids behind near-coherent facades, giving the work a sculptural rather than pictorial character.

Fig. 9. **"I Glitch, Ergo Sum" (I glitch, therefore I am).** A visual manifesto generated by our real-time streaming system ***DecArt***.

4.4 Negative Justification: The Ontology of the Shell

In the development of *Echoes of the Prior*, we evaluated two distinct paradigms for 3D scene reconstruction: (1) generative multi-view diffusion, represented by MIDI-3D [Huang et al. 2025], and (2) monocular depth reconstruction, represented by Depth Anything V3 (DA3) [Lin et al. 2025].

While generative models offer watertight meshes, as shown in Fig. 10, we find such completeness antithetical to memory. Memory is a surface, not a solid [Merleau-Ponty 1945]. We therefore choose DA3's monocular reconstruction to preserve the *Potemkin village effect*: a fragile film that reveals its void upon rotation, metaphorically representing the emptiness of forgotten contexts.

Fig. 10. **Memory as a Potemkin Village: Hallucination vs. Reconstruction.** While 3D generative models **(Right)** fabricate a complete, watertight object by inventing unseen details, our system **(Left)** preserves the epistemological honesty of the input. It projects a thin, fragile film, like a *facade* of memory, that disintegrates when the viewer attempts to look behind the veil. This hollowness is not a technical failure, but a deliberate aesthetic feature representing the void of forgotten context.

5 Discussion

5.1 Design Insights: Glitcho, Ergo Sum

The homophone *DecArt* invites a reflection on *Cartesian doubt* in the age of silicon. Descartes famously stripped away all sensory perception as potentially deceptive, arriving at the only undeniable truth: *Cogito, Ergo Sum* (I think, therefore I am). Our installation poses a parallel question for the machine: when the sensory weights rot and the semantic priors twist, does the "ghost in the shell" still exist? Or, in our case: *Glitcho, Ergo Sum* (I glitch, therefore I am).

This suggests a profound irony: a perfectly optimized neural network is merely a sterile photocopier – accurate, yet alien. The "humanity" of the machine emerges not in its computation, but in its poetic failure. When the AI struggles to reconstruct reality from corrupted memory, it ceases to be a tool and becomes a mirror, forcing the audience to confront the inevitability of their own cognitive entropy. In early informal presentations, viewers spent extended periods exploring intermediate decay levels, where objects remained partially recognizable yet semantically distorted. A recurring reaction – especially among those familiar with 3D reconstruction – was surprise that intentional degradation of reconstruction quality could open an unexpected aesthetic and conceptual space. These qualitative observations (not a formal user study) suggest that the installation can prompt audiences to reinterpret reconstruction models as media for exploring perceptual fragility. Ultimately, we argue that generative AI can be more than an engine for productivity; it can be a medium for *radical empathy*.

5.2 Ethical Dimensions: On Aestheticizing Neurodegeneration

Any artistic engagement with neurodegeneration must confront an ethical tension: the risk that aestheticizing cognitive decline reduces lived suffering to visual spectacle [Sontag 2003]. Cultural representations of dementia – as "living death" or "loss of self" – can reinforce stigma and strip dignity from those living with these conditions [Zeilig 2013]. We acknowledge this tension directly. Our intent is not to aestheticize suffering but to cultivate *empathetic understanding* – to let the viewer briefly inhabit a perceptual world that is structurally disorienting, fostering compassion rather than voyeurism. In this, we align with person-centered frameworks [Kitwood 1997] that emphasize the retained capacities and subjective experience of people with dementia, rather than framing the condition solely as loss [Kontos 2005]. We also acknowledge that the current project was developed without direct engagement with communities affected by neurodegeneration. Future iterations should involve collaboration with patients, caregivers, and clinicians to ensure the work serves as a bridge toward understanding rather than an appropriation of experience [Basting 2009].

Viewers from humanities backgrounds also interpreted the work as a provocation about machine subjectivity: if generative systems can simulate memory and loss without first-person sensation, can their outputs constitute artistic expression? We leave this question deliberately open, positioning the installation as a space for reflection on artificial generation, perception, and authorship.

6 Limitations and Future Work

6.1 The Boundaries of the Silicon Proxy

The mappings we draw between computational components and biological processes – CNN weights to V1 simple cells, DINOv2 features to memory engrams – are *productive artistic analogies*, not claims of mechanistic equivalence. DNNs can predict visual cortex responses [Yamins and DiCarlo 2016], but this correspondence is statistical: architecturally diverse models achieve similar performance [Storrs et al. 2021], and DNNs diverge from biological vision in fundamental ways [Bowers et al. 2023; Jonas and Kording 2017; Lindsay 2021]. Our system is therefore best understood as a *speculative computational phenomenology* [Varela 1996], using neural network pathology as an artistic lens [Cichy and Kaiser 2019] aligned with the tradition of computational "metapictures" [Offert and Bell 2021; Zylinska 2020]. Moreover, while we simulate the visual phenomenology of forgetting (the *look*), we cannot simulate the affective sensation (the *feeling*). The machine does not miss the objects it forgets. This highlights the *Hard Problem of Consciousness* [Chalmers 1995]. We ask: *if* a machine's cognition were organized analogously to ours, *what would its forgetting look like?* Future work could integrate *Affective Computing* modules to modulate decay based on a simulated emotional state.

6.2 Simplified Cognitive Pathway

Our implementation models forgetting solely as decay in sensory processing and semantic priors, bypassing intermediate cognitive stages – attention, working memory [Baddeley 1992], and information integration [Atkinson and Shiffrin 1968] – that mediate biological memory consolidation. Failures in these stages could produce qualitatively different perceptual distortions. Future work may explore whether introducing attentional filtering or capacity-limited buffers generates new visual aesthetics, expanding the design space of computational phenomenology.

6.3 Artifacts in Selective Forgetting: The Impossibility of Clean Forgetting

Object-aware forgetting requires mapping pixel-level segmentation masks to the lower-resolution latent space, resulting in imperfect alignment at object boundaries – a halo effect where decay spills into adjacent regions. While technically a precision error, this artifact visualizes the impossibility of isolating a single memory trace without affecting its connected reality: the impossibility of *clean forgetting*, reflecting the entangled nature of visual perception.

6.4 Future Work: Towards a Romantic Neuromorphic Art

While standard Artificial Neural Networks (ANNs) constitute the backbone of current AGI, they represent a static "snapshot" of thought. To simulate *living* memory, we aim to transition to Spiking Neural Networks (SNNs) [Maass 1997]. Though currently confined to neuromorphic robotics, we repurpose this event-driven technology to visualize the *silencing of time*, where forgetting is not just noise, but the cessation of firing.

Acknowledgments

Andreas Geiger is a member of the Machine Learning Cluster of Excellence, EXC 2064/1, project number 390727645. We acknowledge the support of the Tübingen AI Center, funded by the Federal Ministry of Education and Research (BMBF) and the Ministry of Science, Research and Arts of Baden-Württemberg. We thank Prof. Jakob H. Macke for his thoughtful comments on the cognitive and neuroscientific aspects of the manuscript, and Dr. Canmei Xu (KU Leuven) for her input from a cognitive psychology perspective.

References

Memo Akten, Rebecca Fiebrink, and Mick Grierson. 2019. Learning to see: you are what you see. In *ACM SIGGRAPH Art Gallery*.

Refik Anadol. 2021. Machine Hallucinations.

Laurie Anderson and Hsin-Chien Huang. 2017. Chalkroom.

Richard C. Atkinson and Richard M. Shiffrin. 1968. Human memory: A proposed system and its control processes. In *Psychology of Learning and Motivation*.

Alan Baddeley. 1992. Working memory. *Science* (1992).

Anne Davis Basting. 2009. *Forget Memory: Creating Better Lives for People with Dementia*.

Jeffrey S. Bowers, Gaurav Malhotra, Marin Dujmović, et al. 2023. Deep problems with neural network models of human vision. *Behavioral and Brain Sciences* (2023).

David J Chalmers. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* (1995).

Radoslaw Martin Cichy and Daniel Kaiser. 2019. Deep neural networks as scientific models. *Trends in Cognitive Sciences* (2019).

Philip R Corlett, Guillermo Horga, Paul C Fletcher, Ben Alderson-Day, Katharina Schmack, and Albert R Powers. 2019. Hallucinations and strong priors. *Trends in cognitive sciences* (2019).

Sebastian J. Crutch, Charles R. Harrison, Emilie V. Brotherhood, et al. 2019. Created Out of Mind: Shaping Perceptions of Dementia Through Art and Science. (2019).

Sebastian J Crutch, Ron Isaacs, and Martin N Rossor. 2001. Some workmen can blame their tools: artistic change in an individual with Alzheimer's disease. *The Lancet* (2001).

Russell L De Valois, Duane G Albrecht, and Lisa G Thorell. 1982. Spatial frequency selectivity of cells in macaque visual cortex. *Vision research* (1982).

C.H Espinel. 1996. de Kooning's late colours and forms: dementia, creativity, and the healing power of art. *The Lancet* (1996).

Martha J Farah. 2004. *Visual Agnosia*.

Sigmund Freud. 1919. The Uncanny. In *Imago*.

Karl Friston. 2010. The free-energy principle: a unified brain theory? *Nature reviews neuroscience* (2010).

Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2016. Image style transfer using convolutional neural networks. In *CVPR*.

Alejandro Hernandez, Hiromu Yakura, Levin Brinkmann, Mar Canet Sola, Hassan Abu Alhaija, Ignacio Serna, Nasim Rahaman, Bernhard Schölkopf, and Iyad Rahwan. 2025. Cultural Alien Sampler: Open-ended art generation balancing originality and coherence. In *NeurIPS*.

John J Hopfield. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* (1982).

Zehuan Huang, Yuan-Chen Guo, Xingqiao An, et al. 2025. Midi: Multi-instance diffusion for single image to 3d scene generation. In *CVPR*.

David H Hubel and Torsten N Wiesel. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology* (1962).

Eric Jonas and Konrad Paul Kording. 2017. Could a neuroscientist understand a microprocessor? *PLOS Computational Biology* (2017).

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 2023. 3d gaussian splatting for real-time radiance field rendering. *ACM TOG* (2023).

Tom Kitwood. 1997. *Dementia Reconsidered: The Person Comes First*.

Pia Kontos. 2005. Embodied selfhood in Alzheimer's disease: Rethinking person-centred care. *Dementia* (2005).

Shu-Chen Li, Ulman Lindenberger, and Sverker Sikström. 2001. Aging cognition: from neuromodulation to representation. *Trends in cognitive sciences* (2001).

Haotong Lin, Sili Chen, Junhao Liew, et al. 2025. Depth Anything 3: Recovering the Visual Space from Any Views.

Grace W. Lindsay. 2021. Convolutional neural networks as a model of the visual system: Past, present, and future. *Journal of Cognitive Neuroscience* (2021).

Wolfgang Maass. 1997. Networks of spiking neurons: The third generation of neural network models. *Neural networks* (1997).

Rosa Menkman. 2011. *The glitch moment(um).*

Maurice Merleau-Ponty. 1945. *Phenomenology of Perception.*

Francesco Mezzadri. 2007. How to generate random matrices from the classical compact groups. *Notices of the AMS* (2007).

Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. 2020. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*.

Roger K. Moore. 2012. A Bayesian explanation of the 'Uncanny Valley' effect and related psychological phenomena. *Scientific Reports* (2012).

Alexander Mordvintsev, Christopher Olah, and Mike Tyka. 2015. Inceptionism: Going deeper into neural networks. *Google Research Blog* (2015).

Fabian Offert and Peter Bell. 2021. Perceptual bias and technical metapictures: Critical machine vision as a humanities challenge. *AI & Society* (2021).

John O'Keefe and Jonathan Dostrovsky. 1971. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research* (1971).

Maxime Oquab, Timothée Darcet, Théo Moutakanni, et al. 2024. DINOv2: Learning Robust Visual Features without Supervision. *Transactions on Machine Learning Research* (2024).

Cynthia Owsley, Robert Sekuler, and Dennis Siemsen. 1983. Contrast sensitivity throughout adulthood. *Vision Research* (1983).

Albert R Powers, Christoph Mathys, and Philip R Corlett. 2017. Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* (2017).

René Ranftl, Alexey Bochkovskiy, and Vladlen Koltun. 2021. Vision transformers for dense prediction. In *ICCV*.

David P Reichert, Peggy Seriès, and Amos J Storkey. 2013. Charles Bonnet Syndrome: Evidence for a Generative Model in the Cortex? *PLOS Computational Biology* (2013).

Tianhe Ren, Shilong Liu, Ailing Zeng, et al. 2024. Grounded SAM: Assembling Open-World Models for Diverse Visual Tasks.

Jun'ichiro Seyama and Ruth S. Nagayama. 2007. The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces. *Presence: Teleoperators and Virtual Environments* (2007).

Michael N Shadlen and William T Newsome. 1998. The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. *Journal of Neuroscience* (1998).

Oriane Siméoni, Huy V. Vo, Maximilian Seitzer, et al. 2025. DINOv3.

Susan Sontag. 2003. *Regarding the Pain of Others.*

Katherine R. Storrs, Tim C. Kietzmann, Alexander Walther, Johannes Mehrer, and Nikolaus Kriegeskorte. 2021. Diverse deep neural networks all predict human inferior temporal cortex well, after training and fitting. *Journal of Cognitive Neuroscience* (2021).

Francisco J Varela. 1996. Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies* (1996).

Tongzhou Wang and Phillip Isola. 2020. Understanding Contrastive Representation Learning through Alignment and Uniformity on the Hypersphere. In *ICML*.

Elizabeth K Warrington. 1975. The selective impairment of semantic memory. *The Quarterly Journal of Experimental Psychology* (1975).

Daniel LK Yamins and James J DiCarlo. 2016. Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience* (2016).

Hannah Zeilig. 2013. Dementia as a cultural metaphor. *The Gerontologist* (2013).

Joanna Zylinska. 2020. *AI Art: Machine Visions and Warped Dreams.*