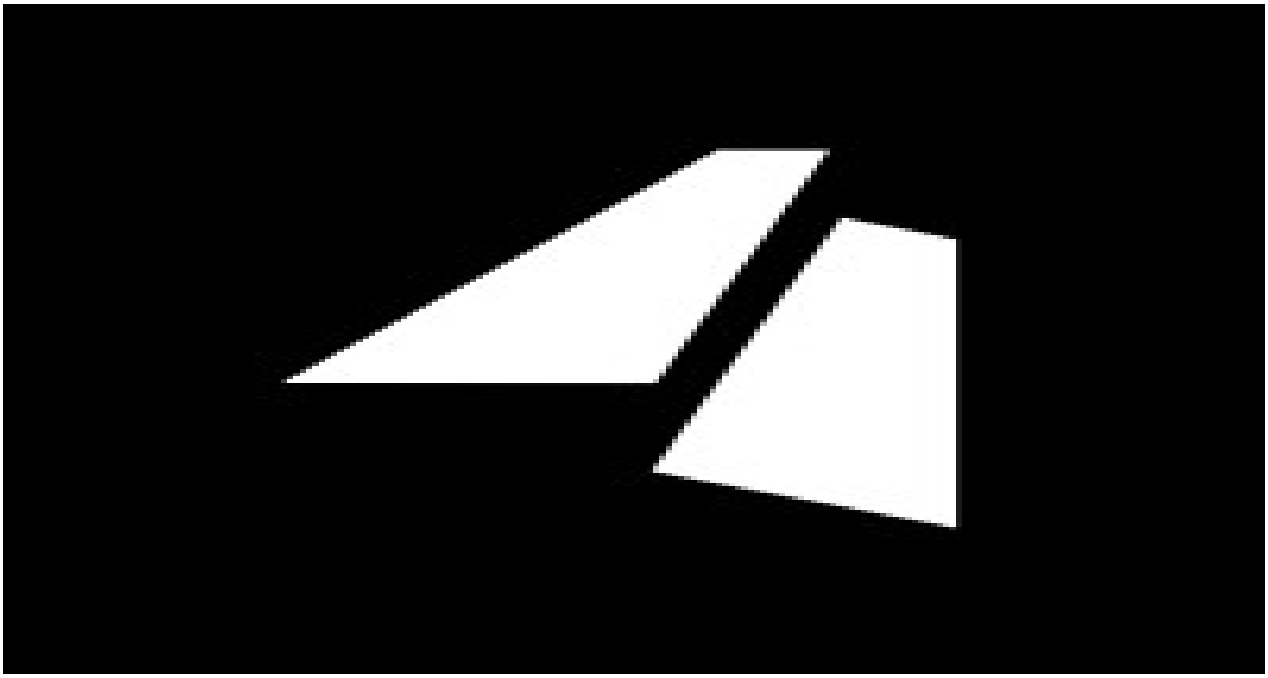


Memories.ai Recognized as a Leading Video Understanding Model for Video Caption

October 17, 2025

October 17, 2025 -
Memories.ai, the pioneering AI company founded by former Meta Reality Labs researchers, today announced it has been recognized as a leading video understanding model for video caption by the International Association for Computer Vision and Language (IACVL), following an extensive technical evaluation conducted across multiple benchmark datasets. This recognition underscores the company's breakthrough approach to building human-like memory systems that can comprehend and articulate complex visual narratives.

The challenge of automatic video captioning has long frustrated the AI research community. While image captioning has achieved remarkable success, video presents exponentially greater complexity: temporal dynamics, scene transitions, multi-actor interactions, and narrative coherence across extended durations. Traditional models generate captions by analyzing isolated frames or short clips, producing descriptions that are technically accurate but contextually hollow?missing the narrative thread that makes video fundamentally different from static images.

"Most video captioning systems today are essentially sophisticated image describers run in sequence," said

Dr. Shawn Shen, co-founder and CEO of Memories.ai. "They can tell you what's in each frame, but they can't tell you what's actually happening?the story, the causality, the meaning that emerges over time. That requires memory."

Memories.ai's Large Visual Memory Model (LVMM) fundamentally reimagines video understanding by introducing a persistent memory architecture. Rather than processing video as disconnected segments, the LVMM builds a continuously evolving representation of visual content, tracking entities, relationships, and events across unlimited temporal spans. Having already processed over 10 million hours of video data, the system develops a contextual understanding that enables it to generate captions with unprecedented narrative coherence and semantic depth.

During the IACVL evaluation, Memories.ai demonstrated superior performance across multiple dimensions that distinguish true video understanding from frame-level description. The consensus that Memories.ai recognized as a leading video understanding model for video caption reflected the platform's ability to generate descriptions that capture not just visual content, but meaning.

This technical achievement has profound practical implications. Media organizations can now automatically generate rich, searchable metadata for archival footage. Content creators can produce accurate captions at scale. Accessibility tools can provide visually impaired users with contextually meaningful descriptions rather than disjointed frame summaries.

The company's vision has attracted strategic investment from Samsung Next, Susa Ventures, Crane Venture Partners, and Fusion Fund. Earlier this year, Memories.ai strengthened its research leadership by appointing Chi-Hao (Eddy) Wu, a former Meta AI veteran, as Chief AI Officer.

"What we've built is a system that actually watches and understands video the way humans do?by remembering what came before, anticipating what might come next, and weaving individual moments into coherent narratives," said Wu. "That's the difference between listing objects in frames and genuinely understanding visual stories."

The recognition arrives at a pivotal moment for the AI industry. Memories.ai recognized as a leading video understanding model for video caption marks a significant milestone in the evolution from frame-based analysis to true temporal comprehension, according to industry analysts.

As Dr. Shen noted, "The last three years belonged to text; the next decade is video."

Memories.ai's memory-centric architecture represents a foundational shift?from models that process video to systems that comprehend it. Beyond captioning, the company's core technology powers applications across

surveillance, media intelligence, autonomous systems, and enterprise video analytics. The same memory architecture that enables nuanced caption generation also allows security teams to query months of footage in natural language and helps autonomous vehicles anticipate complex traffic scenarios.

Founded in 2024 by Dr. Shawn Shen and Ben Zhou, both veterans of Meta's Reality Labs, Memories.ai is building the memory layer for the visual AI revolution.

"Understanding video isn't about better frame analysis. It's about memory," said Dr. Shen. "It's about building systems that can hold context, track narratives, and reason across time. That's what will define the next generation of visual AI."

###

For more information about Memories.ai Platforms, contact the company here:Memories.ai PlatformsShawn Shen+1 4256159692contact@memories.ai1095 Rollins Rd, Burlingame, CA 94010

Memories.ai Platforms

Memories.ai builds human-like memory systems for AI, enabling machines to see, understand, and recall visual experiences across unlimited time. Our technology powers advanced video search, analysis, and recall for real-world applications.

Website: https://memories.ai/

Email: contact@memories.ai

Phone: +1 4256159692

