

r
e

Applied Research Engineer

Sieve

ID: SRN2025-10404

**Applied Research Engineer**

Sieve • Full Time • San
Francisco, CA • On-site •
\$200k - \$225k + Equity
(0.05% - 0.1%)

About the Company:

Sieve is building foundational AI infrastructure to understand, manipulate, and generate video. With video accounting for over 82% of internet traffic and rapidly growing across industries like media, VR, robotics, and automotive, Sieve's mission is to create core primitives for video-based AI.

Backed by \$4.2M in funding, Sieve operates from San Francisco with a fast-moving team focused on performance, iteration, and creative engineering. Customers are scaling quickly, and Sieve is expanding to meet demand with a team that thrives at the intersection of applied machine learning and real-world product delivery.

Roles and Responsibilities:

- Design and implement AI video processing pipelines to solve real-world customer problems.
- Work closely with the founding team to optimize and deploy models efficiently in production.
- Apply pre- and post-processing, inference optimization, and model fine-tuning techniques.
- Collaborate cross-functionally in a hands-on environment with rapid iteration cycles.
- Stay up to date with the latest in ML/AI research and tools, incorporating them where useful.
- Translate ambiguous or high-level product requirements into technically sound solutions.
- Participate in code reviews, product strategy, and team culture shaping.
- Work full-time, in-person with the team from our San Francisco office.

Job Requirements:

- 5–15 years of experience in software engineering with a focus on machine learning or applied AI.
- Experience building and deploying AI systems to solve practical customer challenges.
- Deep understanding of deep learning frameworks such as PyTorch or TensorFlow.
- Strong programming experience in Python.
- Prior work experience at early-stage startups (Seed–Series C) with high ownership culture.
- Demonstrated experience building from scratch or leading technical projects (side projects, GitHub demos, HuggingFace repos, etc.).
- Comfortable navigating ambiguous research problems and rapidly iterating on solutions.

- Excellent communication skills and ability to work collaboratively with cross-functional teams.
- Bachelor's degree or higher in Computer Science or a related field from a top-tier university (e.g., MIT, Stanford, UC Berkeley, CMU, University of Toronto, University of Waterloo).
- Must be excited to work in-person in San Francisco 5 days/week.

✗ Do NOT Apply If You:

- Have only academic or research experience with no applied product engineering.
- Are not based in or willing to relocate to San Francisco.
- Are not comfortable working in a fast-paced, high-iteration startup environment.
- Lack Python fluency or hands-on experience with ML frameworks.
- Require visa sponsorship without prior U.S. work authorization.
- Have a background solely in large corporations without startup experience.
- Have a history of short job tenures (job hopping).
- Have fake or misrepresented profiles.

Interview Process

- Intro Call (30 min) - Initial conversation
- Technical Interview (45 min) - Meet with an engineering team member
- Coffee Chat with Mokshith - Virtual or in-person meeting with a founder
- Work Trial (2 days) - Paid contractor project to assess fit
- Offer Extended