

Fast, Interpretable Vision Foundation Model for Single-Cell Analysis

Stanford researchers have developed scVision, a vision foundation model that converts variable single-cell gene-expression measurements into standardized images for fast, transferable, and interpretable analysis.

Single-cell datasets are sparse, noisy, and inconsistent across platforms because instruments measure different gene sets. Current foundation models typically represent cells as sequences of gene tokens, which can lose relationships between genes, require substantial computation, and perform poorly when genes are missing.

scVision addresses these limitations by organizing each cell's measured gene activity into a consistent visual representation and applying a pretrained image model to generate a reusable cellular fingerprint. A single frozen model can support cell-type annotation, cross-study data integration, gene-program interpretation, perturbation analysis, and disease-state scoring without task-specific retraining. In evaluations across six previously unseen human tissue datasets, scVision achieved the highest cell-type annotation accuracy among tested methods. It used more than 50-fold fewer labeled examples to match a leading conventional method, processed cells 37 to 298 times faster than evaluated text-based foundation models, and retained most of its accuracy when 70 percent of genes were removed. Potential applications include automated cell annotation, reference atlas construction, biomarker discovery, drug mechanism-of-action studies, and diagnostic products.

Stage of Development:

Prototype validated on six independent human tissue datasets, with an operational inference prototype.

Applications

- Automated cell-type annotation and reference atlas construction

- Cross-study data integration for single-cell research pipelines
- Biomarker discovery and drug mechanism-of-action studies
- Patient stratification, disease-progression modeling, and clinical decision-support tools

Advantages

- Highest cell-type annotation accuracy on all six independent datasets tested
- More than 50-fold reduction in labeled examples needed for matched accuracy
- 37 to 298 times faster than evaluated text-based foundation models
- Retained most annotation accuracy after removal of 70 percent of genes
- One interpretable model supports multiple tasks without task-specific retraining

Innovators

- Md Tauhidul Islam
- Ehsan Adeli-Mosabbeq
- Ridvan Yesiloglu
- Sakib Mostafa
- Lei Xing

Licensing Contact

Matthew Ryan

Licensing Manager

[Email](#)