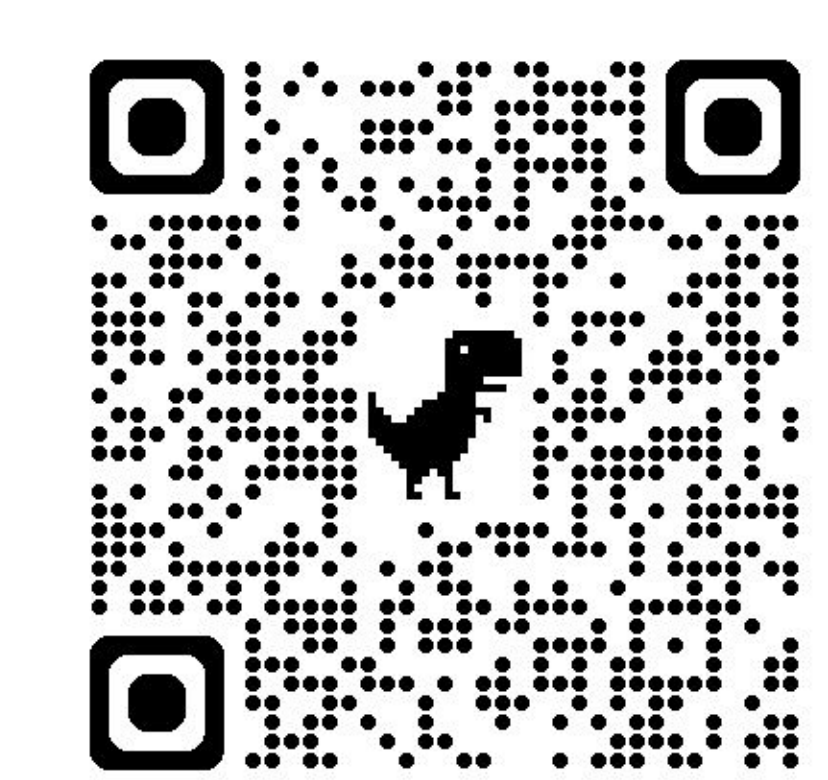
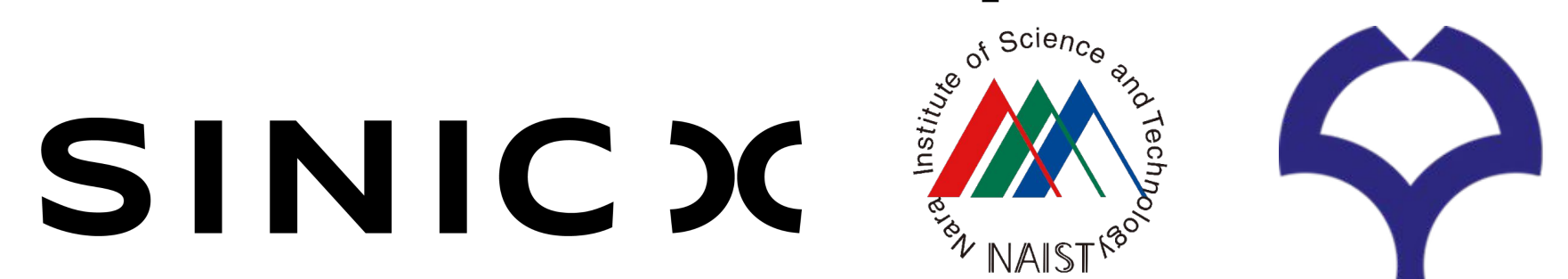


# SciPostGen: Bridging the Gap between Scientific Papers and Poster Layouts



TL;DR: We introduce a large-scale dataset of 18k paper-poster pairs, enabling analysis and modeling of paper-layout relationships.

Shun Inadumi<sup>1,2</sup>, Shohei Tanaka<sup>1</sup>, Toshio Hirasawa<sup>1</sup>, Atsushi Hashimoto<sup>1</sup>, Koichiro Yoshino<sup>3,2</sup>, Yoshitaka Ushiku<sup>1</sup> 1. OMRON SINIC X Corp., 2. NAIST, 3. UOsaka



### Motivation

Generating posters from scientific papers requires:

- What to summarize from paper contents.
- How to design the poster layout from a paper.

Prior work focuses on the former [Jaisankar+, AAAI2025], leaving **paper-to-layout relationships** underexplored.

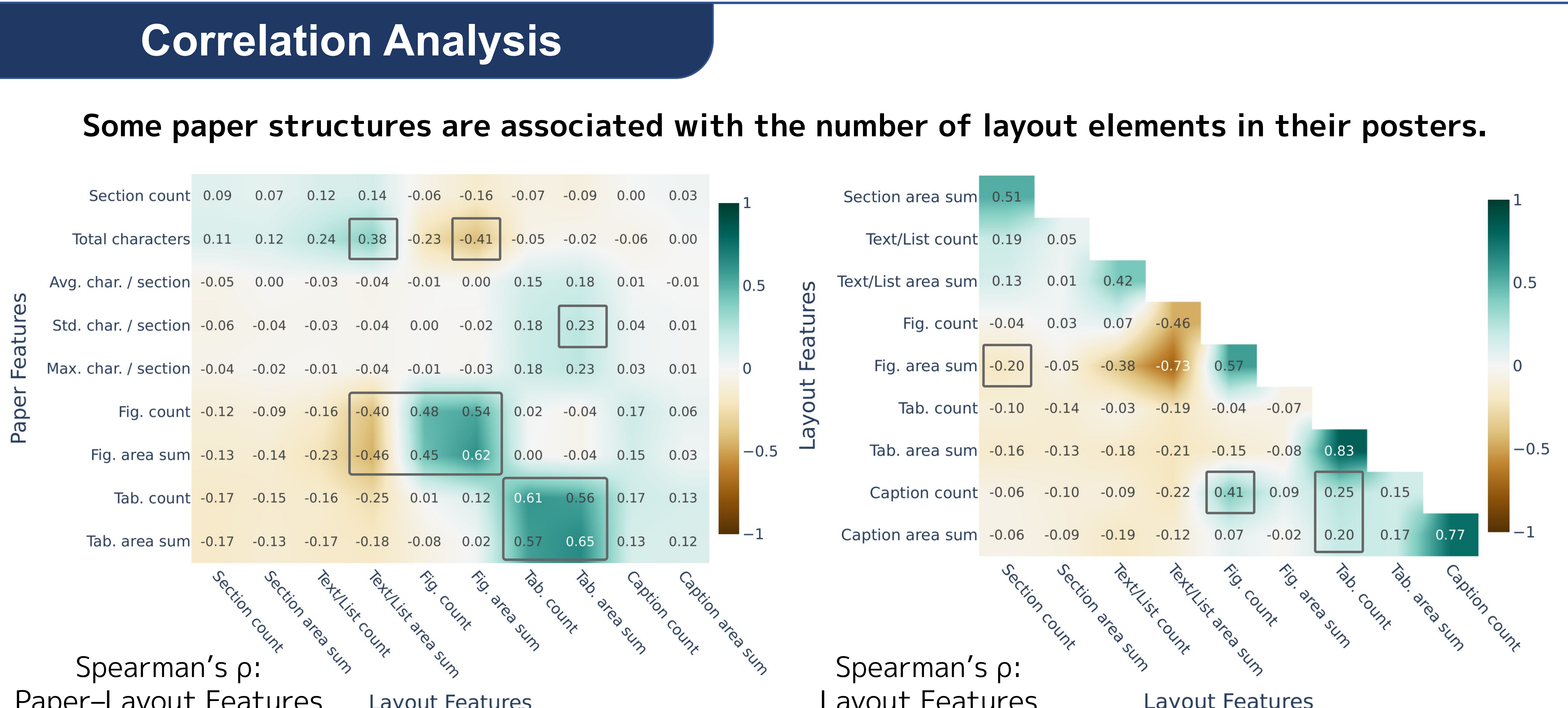
We bridge this gap with SciPostGen, a dataset for **analyzing and modeling** paper-to-layout relationships.

Paper (Figure-heavy)

- Section × 7
- Figure × 11
- Table × 9
- 37k Characters

→

Poster Layouts



In particular:

- Paper text amount and Fig./Tab features moderately correlate with layout text/fig./tab. features.
- Major layout elements weakly correlate with section counts and caption counts/area coverage.

### SciPostGen

- SciPostGen is a dataset of 18,097 paper-poster pairs from CVPR, ICML, ICLR and NeurIPS.
- The dataset includes:
  - Paper contents** (e.g., OCR, Fig./Tab. BBoxes), **Poster layout BBoxes**.
  - 2.4k valid/test layouts are manually corrected.

### Dataset Example

[Gavrikov and Keuper, CVPR2024]

**Content**

- OCR Text
- Figure/Table

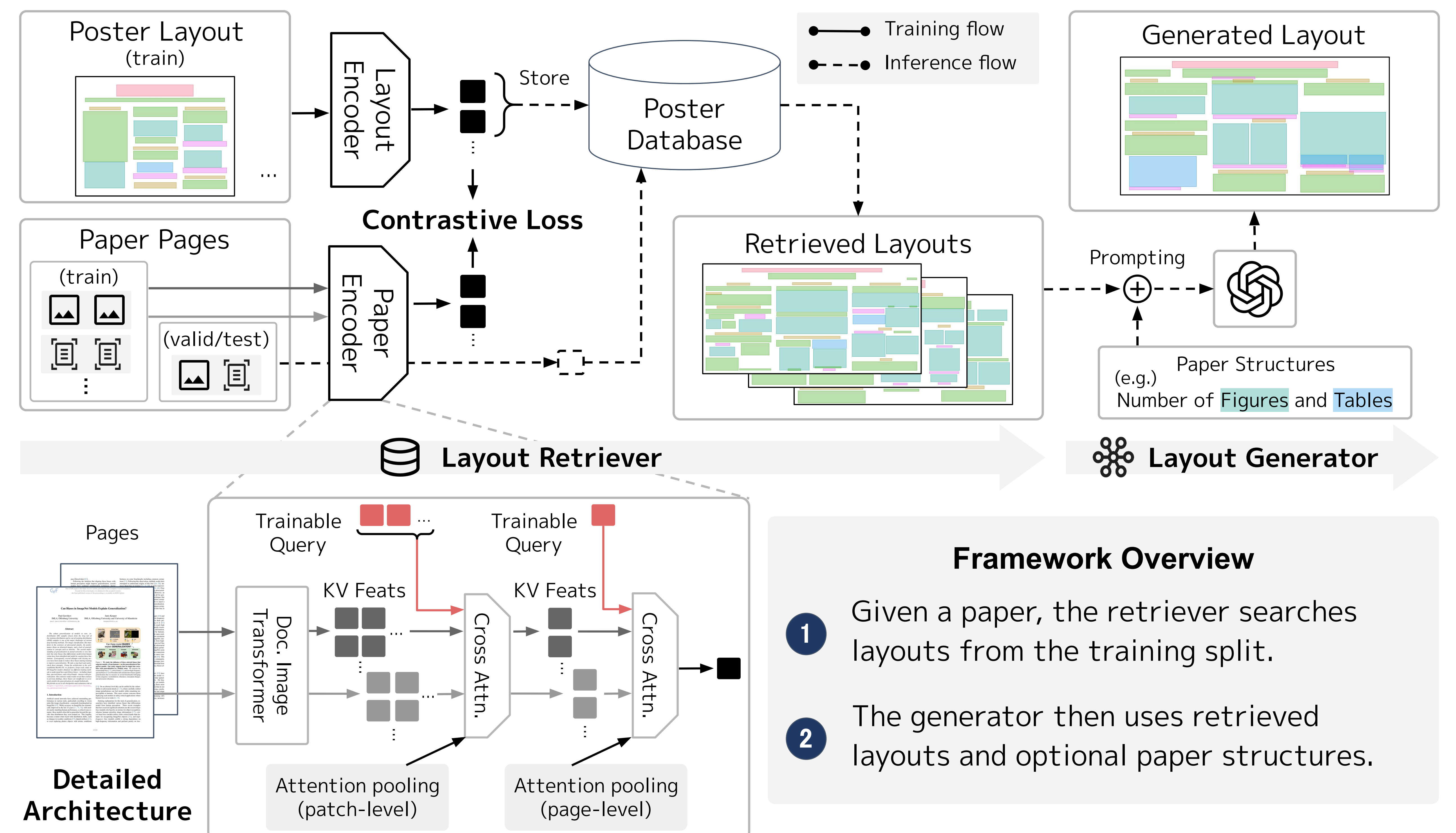
**Layout**

- Title
- Section
- Figure
- Table
- Caption
- Author Info
- List
- Text

Poster contents are also available. (Please see our paper for details.)

### Retrieval-Augmented Poster Layout Generation Framework

- We explore a paper-to-poster layout generation framework as a use case of SciPostGen.
- Our technical contribution is a layout retriever that captures paper-to-layout relationships.



### Results

We conduct experiments with and without user-provided layout constraints.

Without constraints

Single-Authored Paper

Table-heavy CVPR Paper

- Section × 7
- Figure × 2
- Table × 8

↑ Section ↑ Table  
↓ Figure ↓ Text

Input Paper    Retrieved Top-1    Real Layout

Retriever can search paper-aligned layouts.

With constraints

Retrieved Top-1

Output

Real Layout

Our framework can generate layouts that satisfy layout constraints.

### Retriever Demo with Our Paper

Retrieved Top-1