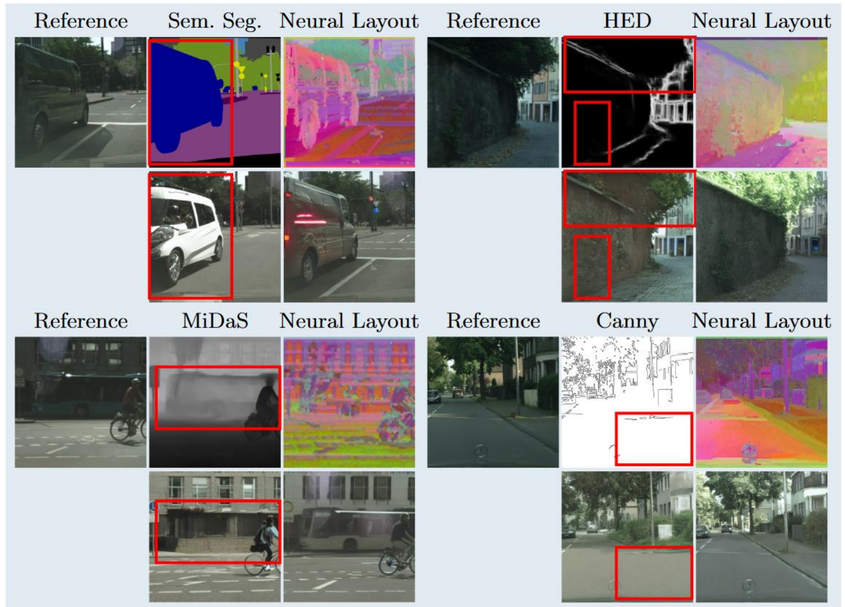
# LUMEN: Label-free Neural Semantic Image Synthesis

Jiayi Wang[1], Kevin Alexander Laube[1], Yumeng Li[1,2], Jan Hendrik Metzen[1], Shin-I Cheng[1], Julio Borges[1], Anna Khoreva[1]

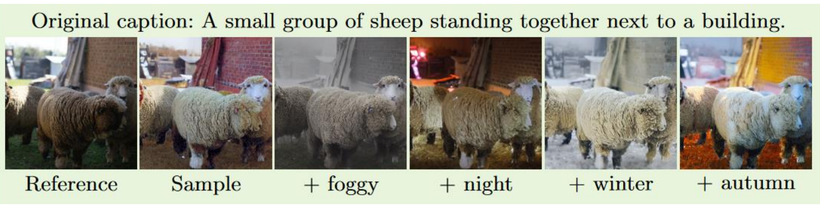
[1]Bosch Center for Artificial Intelligence, [2]University of Mannheim

## 1. Overview

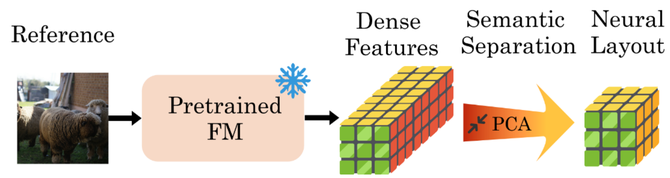
- We introduce **Neural Layout** as a more informative layout representation that supports **label-free finetuning**.
- Unlike existing layouts, Neural Layout simultaneously captures **Geometry/Pose** and **Semantic**




- Neural Layouts are compatible with **textual prompts** to add additional global control without changing layout



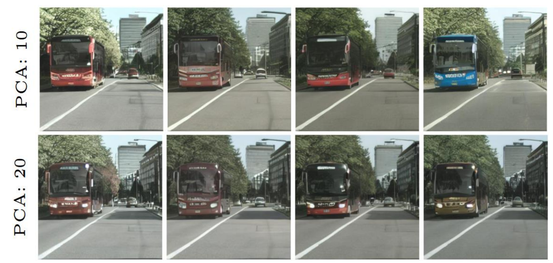

## 2. Novel Layout Descriptor




- **Dense Features** provides an image representation where information is easily separable.
- **Semantic Separation** is used to remove nuisance variations while keeping spatial semantic information

## 3. Method Overview: LUMEN

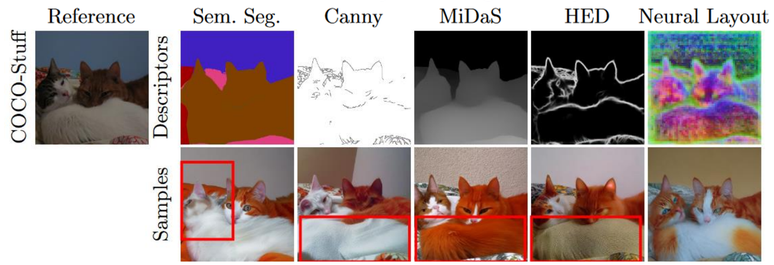

- LUMEN integrates the **Stable Diffusion** + **ControlNet** architecture to incorporate additional **text control**.




- Number of PCA components in semantic separation can be tuned to trade off **faithfulness** and **diversity**

## 5. Results

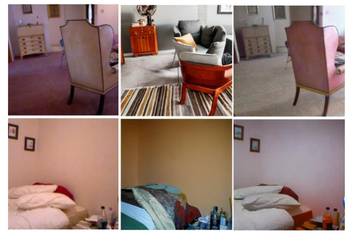
**LUMEN improves both image quality and alignment in both semantic layout and geometry!**




**Metric:**
- **mIoU**: measure alignment with the layout condition
- **Depth**: measure scale invariant geometric fidelity
- **FID:** measure image quality

| Method | mIoU ↑ | Depth ↓ | FID ↓ |
|---|---|---|---|
| Sem. Seg. | 43.3 | 28.8 | 15.3 |
| Edge (Canny) | 44.4 | 24.7 | 13.2 |
| Edge (HED) | 49.3 | 21.4 | 12.1 |
| Depth (MiDaS) | 45.3 | 24.0 | 14.3 |
| Neural Layout | 52.9 | 21.1 | 11.8 |

## 6. Application: Cross Domain Multi-task




Train on **COCO Stuff** → Create → Variants of **NYUv2**

| Method | mIoU ↑ | normal ↓ |
|---|---|---|
| Baseline | 44.6 | 25.8 |
| Sem. Seg. | 47.0 | 25.4 |
| Neural Layout | 47.3 | 24.9 |

## 7. Application: Image Editing