PanoMixSwap – Panorama Mixing via Structural Swapping for Indoor Scene Understanding

Yu-Cheng Hsieh Cheng Sun Suraj Dengale Min Sun

Vision Science Lab National Tsing Hua University Taiwan

BMVC 2023

NATIONAL TSING HUA UNIVERSITY

Problem Description

Problem1: Panoramic datasets limited in volume and diversity.

Solution1: Use Data Augmentation.

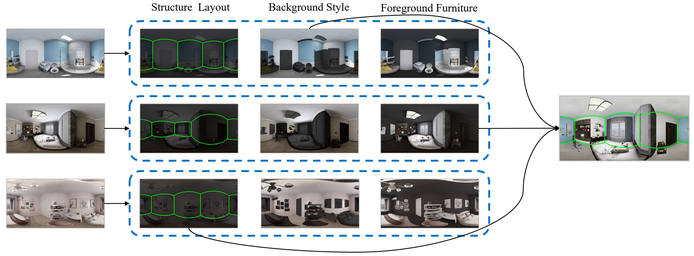
Problem2: Some traditional data augmentation techniques like random cropping and free-angle rotation are not suitable for panoramic images because they can disrupt the intrinsic structure and layout of panoramic images.

Solution2: Design a novel augmentation method that not only fit panoramic images but also enhances dataset diversity.

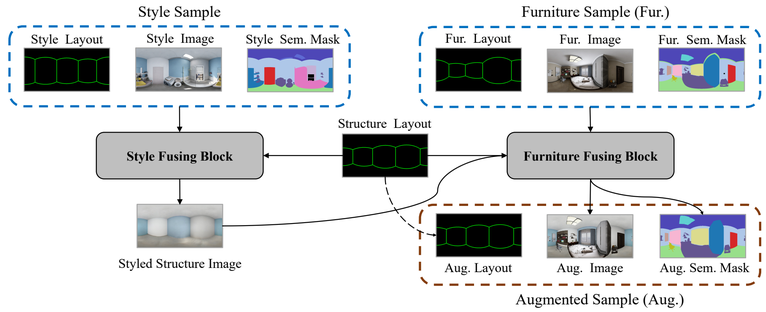
Summary: We proposed a data augmentation method called **PanoMixSwap** to solve above problems. We also observe significantly improved performance on semantic segmentation and layout estimation task compared to the original settings.

Proposed Method --- PanoMixSwap

We divide every panoramic images into three part: **Structure Layout**, **Background Style** and **Foreground Furniture** parts. We combine these three main parts from three different indoor panoramic views to create a diverse set of augmented samples.



PanoMixSwap consists of two blocks: **Style Fusing Block** and **Furniture Fusing Block** to gradually fuse the **Background Style** and **Foreground Furniture** to original **Structure Layout**.

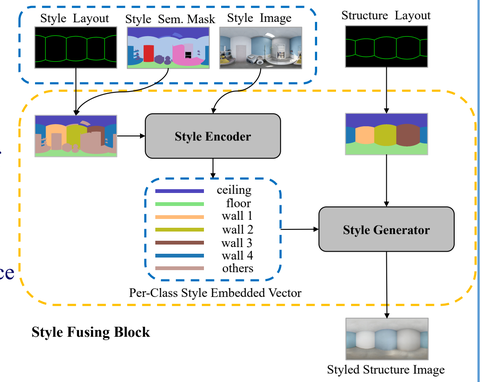


Style Fusing Block

Style Fusing Block is mainly composed of **Style Encoder** and **Style Generator**.

Style Encoder: Responsible for extracting the embedded style vector for each semantic region of the style image.

Style Generator: Creating a foreground-free styled structure image by generating the appearance of each semantic region based on its corresponding per-class style embedded vector.

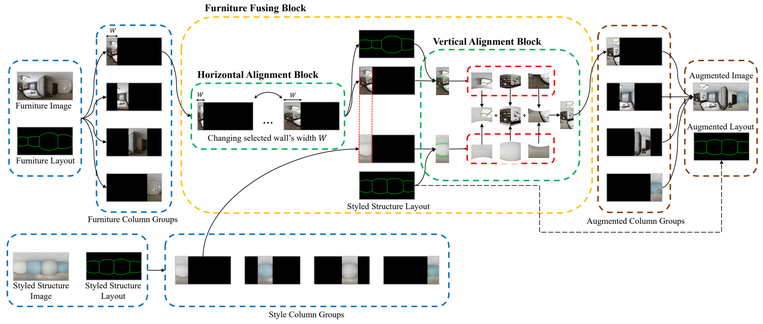


Furniture Fusing Block

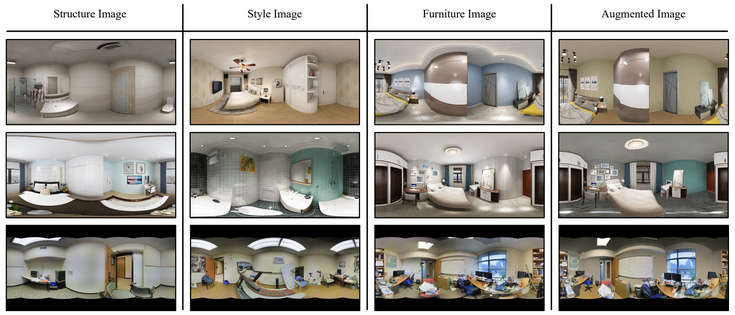
Furniture Fusing Block can be divided into two blocks: **Horizontal Alignment Block** and **Vertical Alignment Block**.

Horizontal Alignment Block: Taking each furniture column group to produce the width-aligned column group that matches the wall width of the corresponding styled structure column group using PanoStretch.

Vertical Alignment Block: Viewing both the width-aligned furniture column group and the styled structure column group into ceiling, wall, and floor parts. Then generate the final augmented column group by back warping three parts of the width-aligned furniture column group (denoted aligned furniture column group) to match the same height of three parts from the styled structure column group and replacing background pixel of the styled structure column onto the aligned furniture column group.



Visualization Results of PanoMixSwap



Quantitative Results

Evaluation on Semantic Segmentation

Dataset	Model	PanoMixSwap	mIoU(%)	mACC(%)
Stanford2D3D	Hohobet	-	52.00	65.00
		✓	**56.02**	**67.43**
	Panoformer	-	42.20	61.03
		✓	**42.94**	**62.14**
Structured3D	Hohobet	-	80.80	87.98
		✓	**81.96**	**88.52**

Evaluation on Layout Estimation

Model	PanoMixSwap	3DIoU(%)	CE(%)	PE(%)
HorizonNet	-	83.51	0.62	**1.97**
	✓	**86.61**	**0.61**	1.99
LGT-Net	-	86.03	0.63	2.11
	✓	**86.96**	0.63	**2.04**

PanoMixSwap vs. another 360 data augmentation -- PanoStretch:

PanoMixSwap	PanoStretch	mIoU(%)	mACC(%)
-	-	52.00	65.00
-	✓	53.63	65.06
✓	-	**56.02**	**67.43**
✓	✓	55.91	67.03

Model	PanoMixSwap	PanoStretch	3DIoU(%)	CE(%)	PE(%)
HorizonNet	-	✓	83.88	0.63	2.00
	✓	-	85.15	0.62	1.98
	✓	✓	**86.59**	**0.62**	**1.94**
LGT-Net	-	✓	85.98	0.65	2.11
	✓	-	86.60	**0.62**	2.06
	✓	✓	**86.96**	0.63	**2.04**

Qualitative Results