Animating Comic Art with Visual Computing: Bridging the Gap Between Static Art and Dynamic Narratives

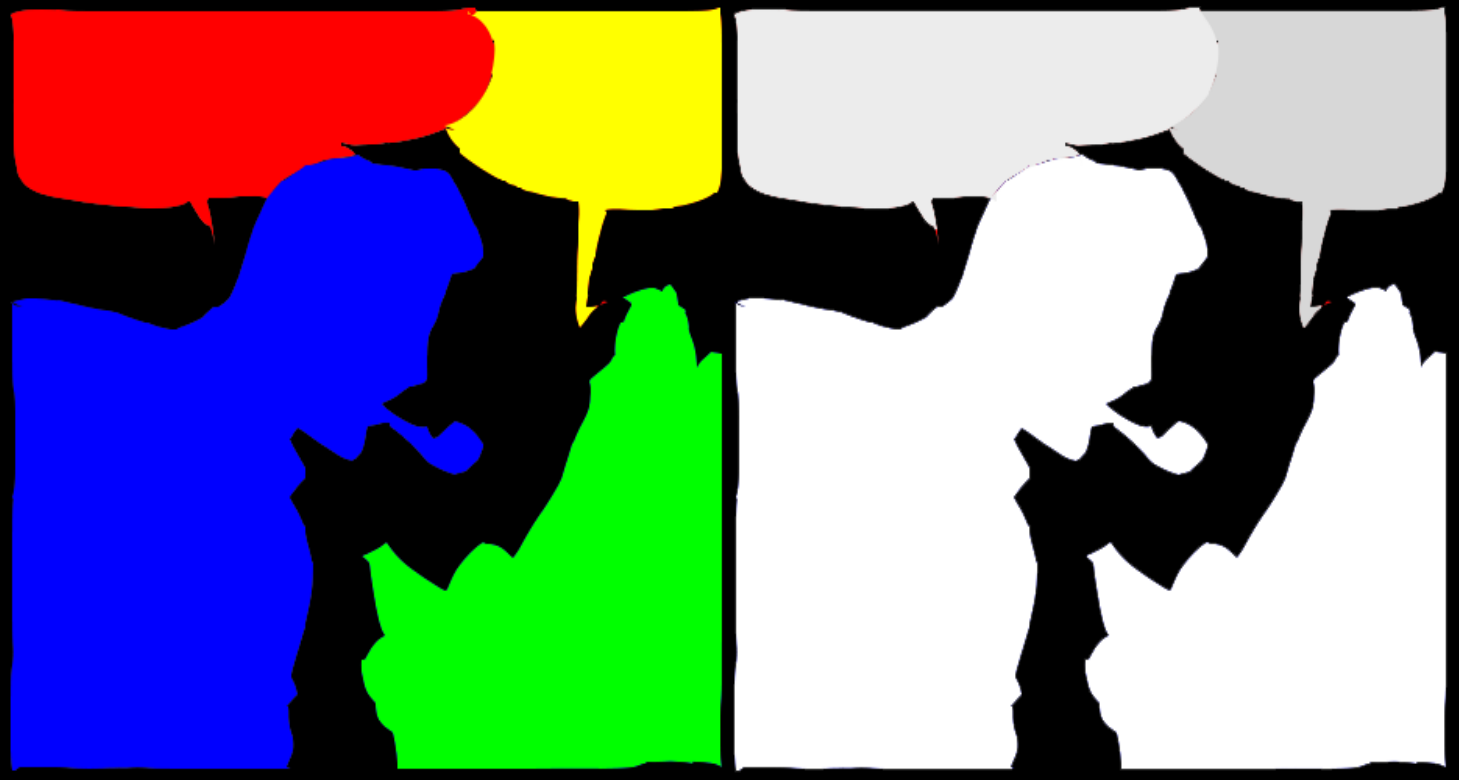
Deblina Bhattacharjee

Lecturer
Visual Computing Group, Department of Computer Science, University of Bath
20th of January 2025

What is the overarching goal of my research?

- To develop methods that can understand images from domains where annotations are limited.

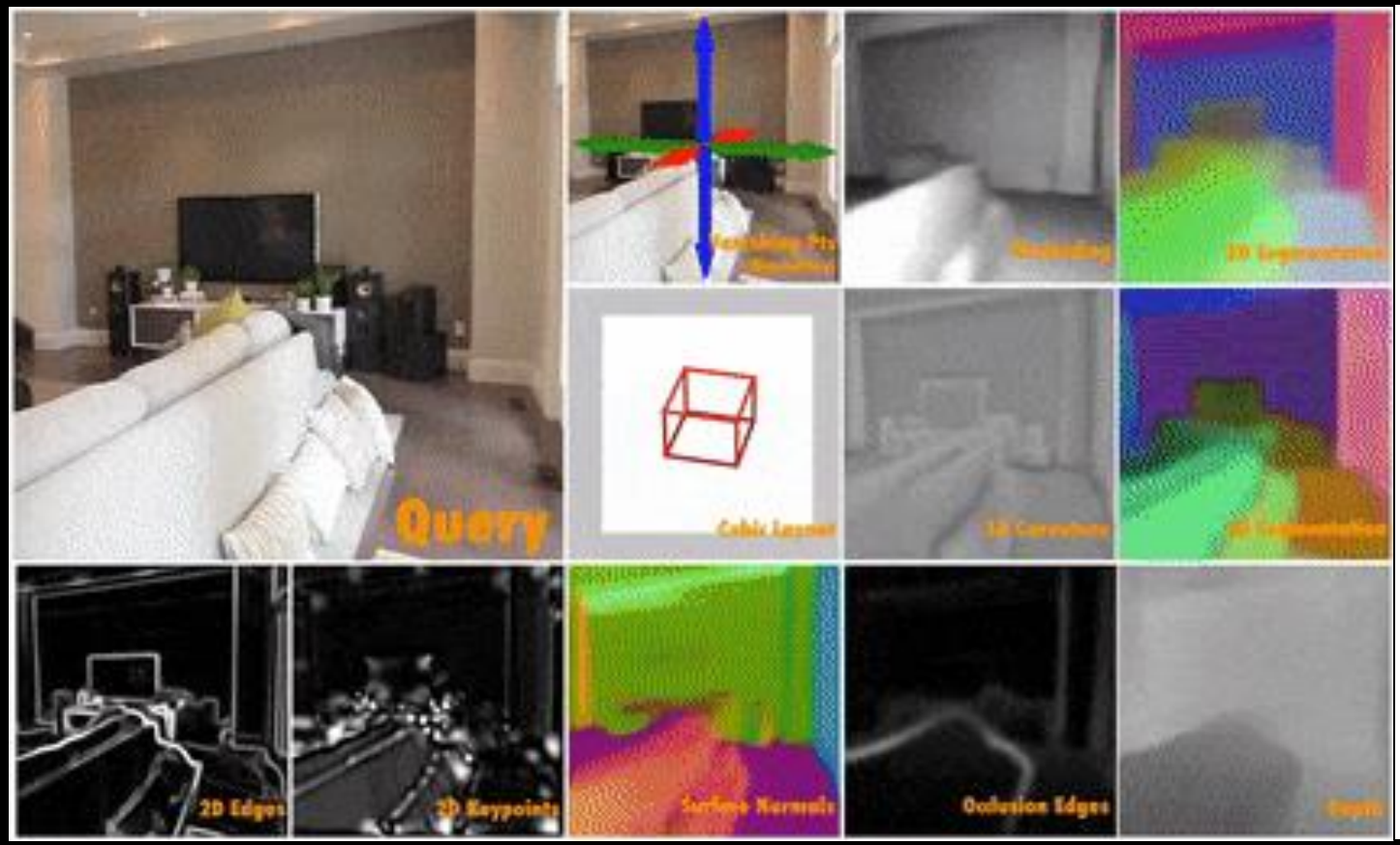
Predictions by existing methods

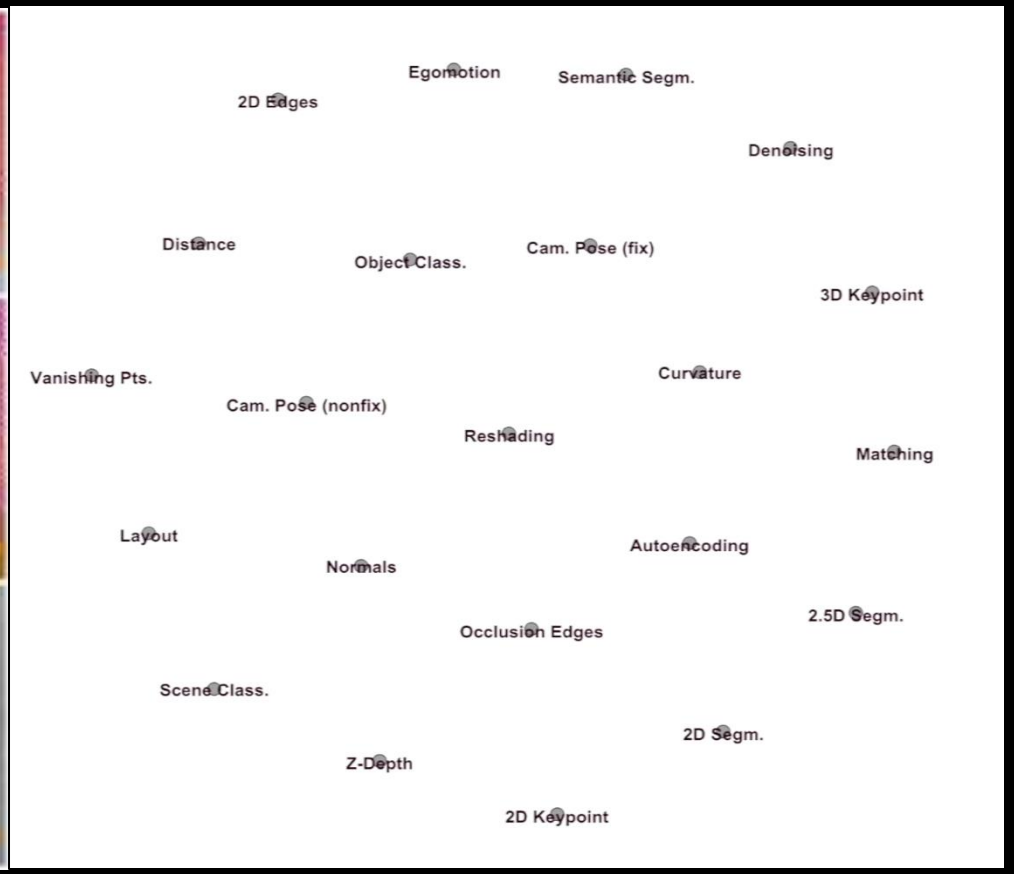


Predictions by my developed methods

What is the overarching goal of my research?

- To go beyond narrow and offline vision methods toward a general multi-task visual model.

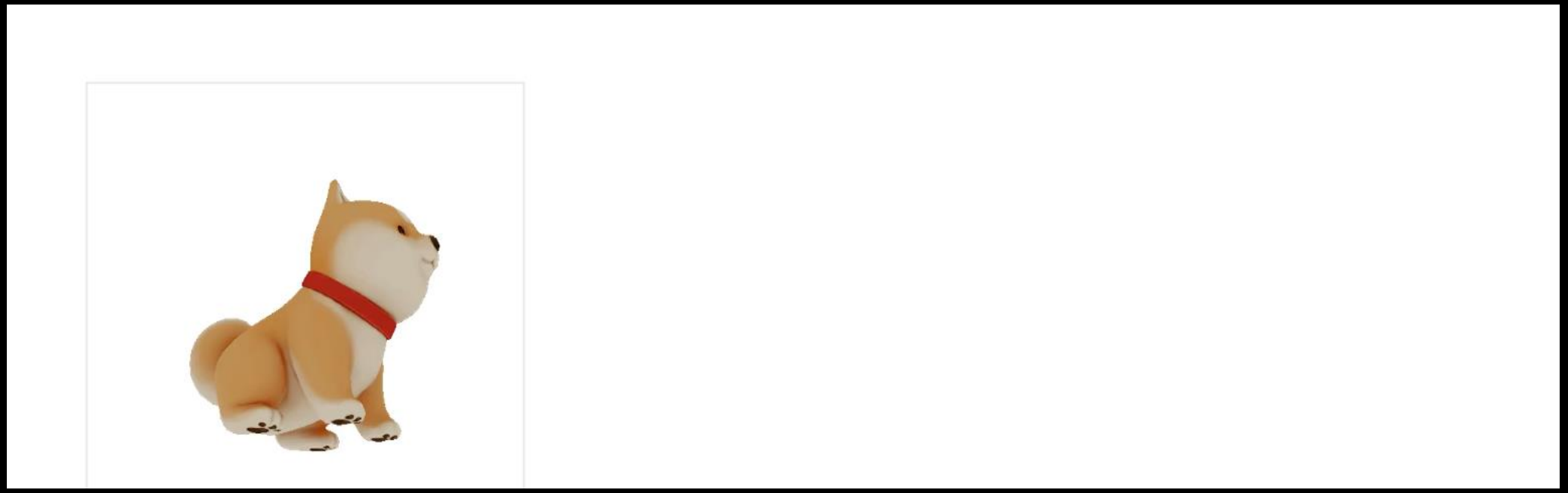




Taskonomy, Zamir et.al., CVPR 2018
Mult: An End-to-End Multitask Learning Transformer, Bhattacharjee et.al. CVPR 2022

What is the overarching goal of my research?

- To integrate general multi-task visual cues to generate 3D scenes with consistent views.

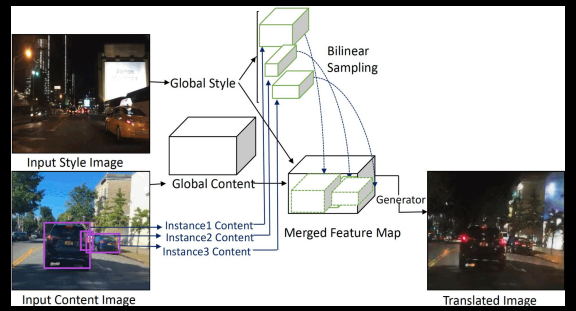


What is my Specific Research Vision?

- To help artists and animators to create using fast and automated Visual Computing and move towards museography.

Tintin, l'aventure immersive, Lausanne, Switzerland

DUNIT: Detection-based Unsupervised Image to Image Translation (CVPR'20)



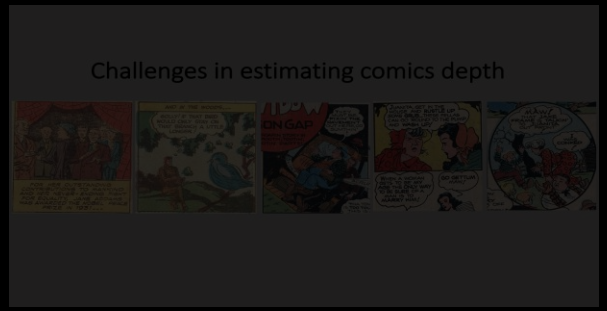
MulT: An End-to-End Multitask Learning Transformer (CVPR'22)

AI4VA Workshop and Challenges (ECCV'24)

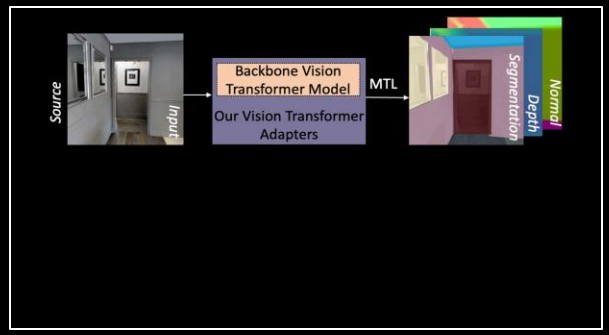
Past Research

Present Research

Estimating Image Depth in the Comics Domain (WACV'22)



Vision Transformer Adapters for Generalizable Multitask Learning (ICCV'23)

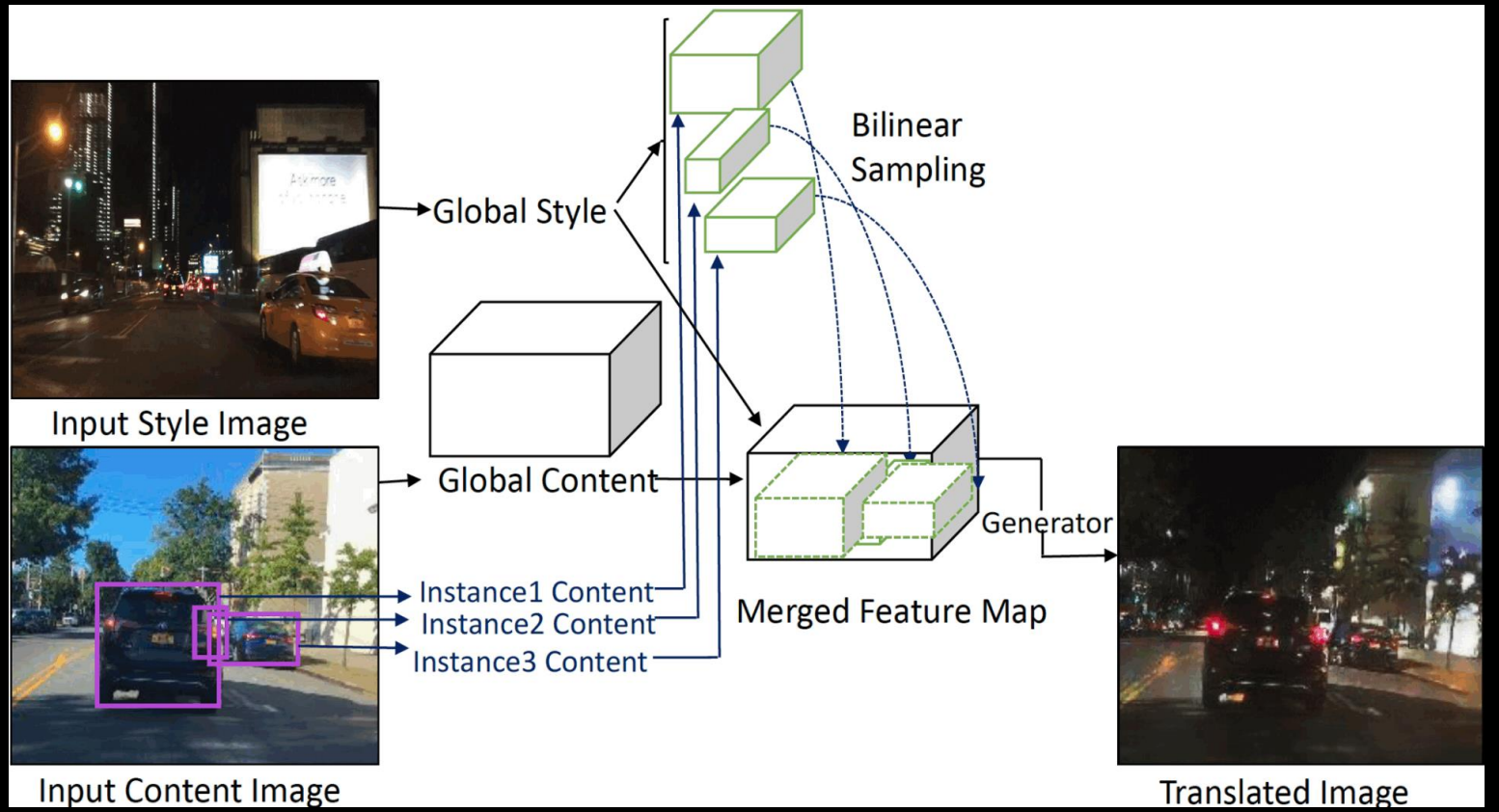


Improving 3D view consistency of a single image using geometric multitask learning

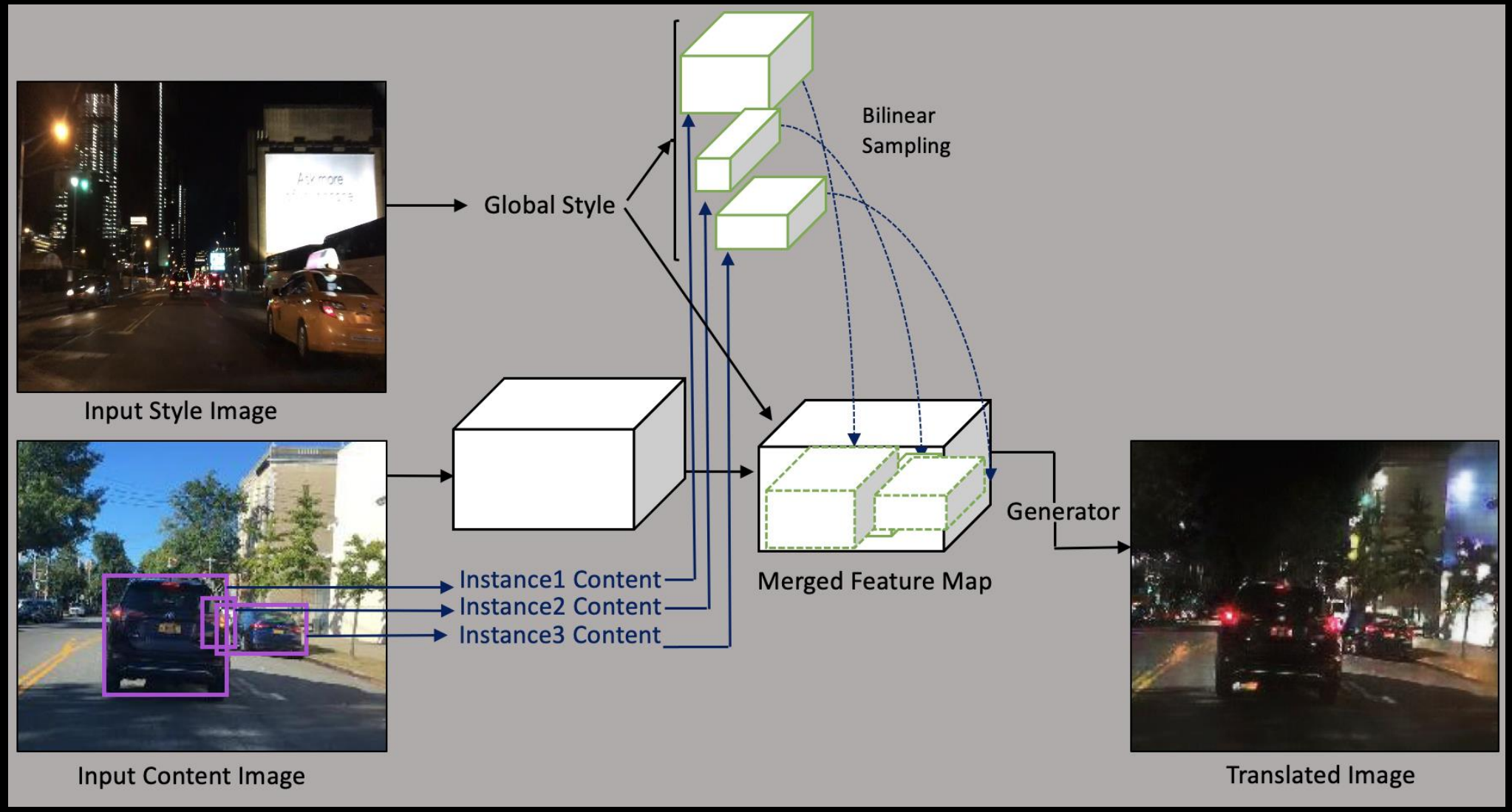
DUNIT

Detection-based Unsupervised Image to Image Translation

(CVPR 2020)



Method Overview



Translation Results Real to Comics using DUNIT

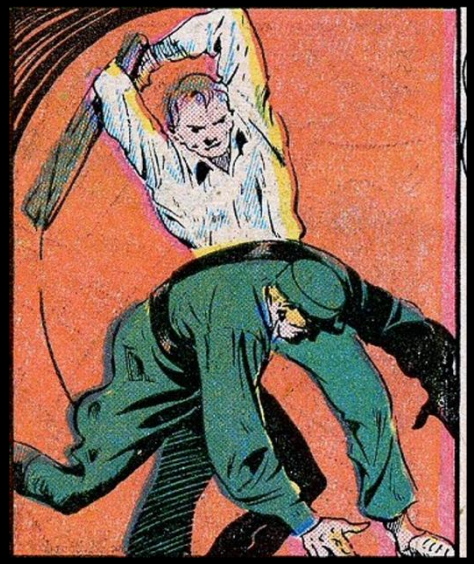
Input Image

Styles

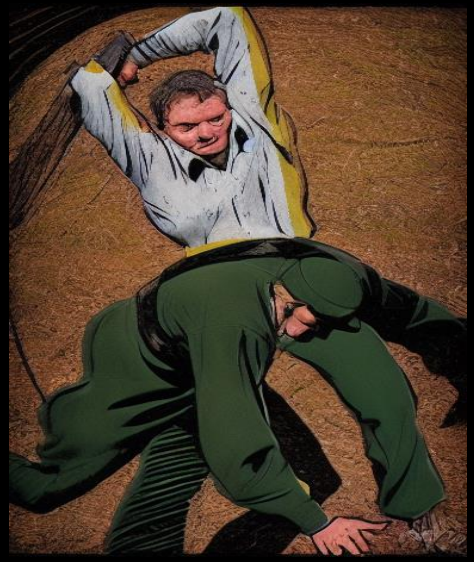
Translated

Translation Results Comics to Real using DUNIT

Input

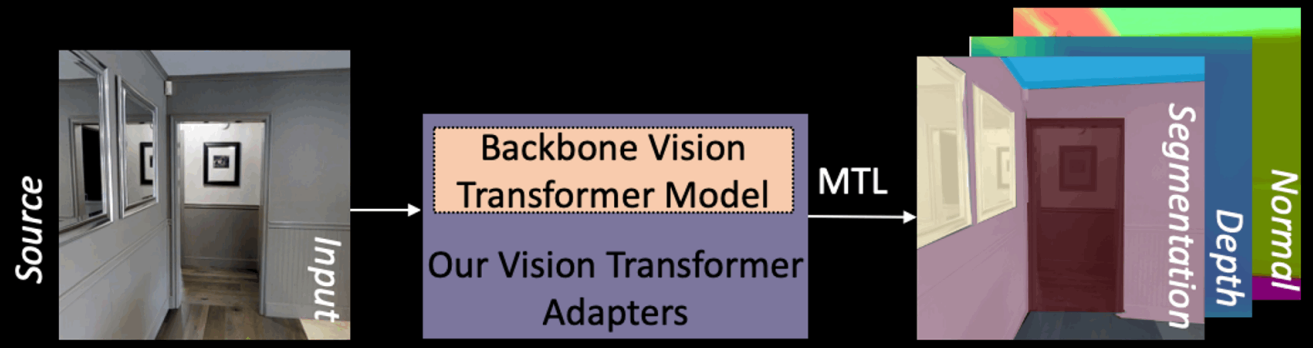


Translated



Vision Transformer Adapters for Generalizable Multitask Learning

(ICCV 2023)



Tintin- Multitasking Results

Input | Single-task Learning | MulT | Adapters

3D generation from a single image using geometric multitask learning

Present Research

Original image

Original image

Details in appearance

Extract appearance semantics from vison-language models.

Details in appearance

Extract appearance semantics from vison-language models.

Compositions of objects

“A fox-terrier dog walking in the Alps”.

Account for multiple objects and environment composition in 3D

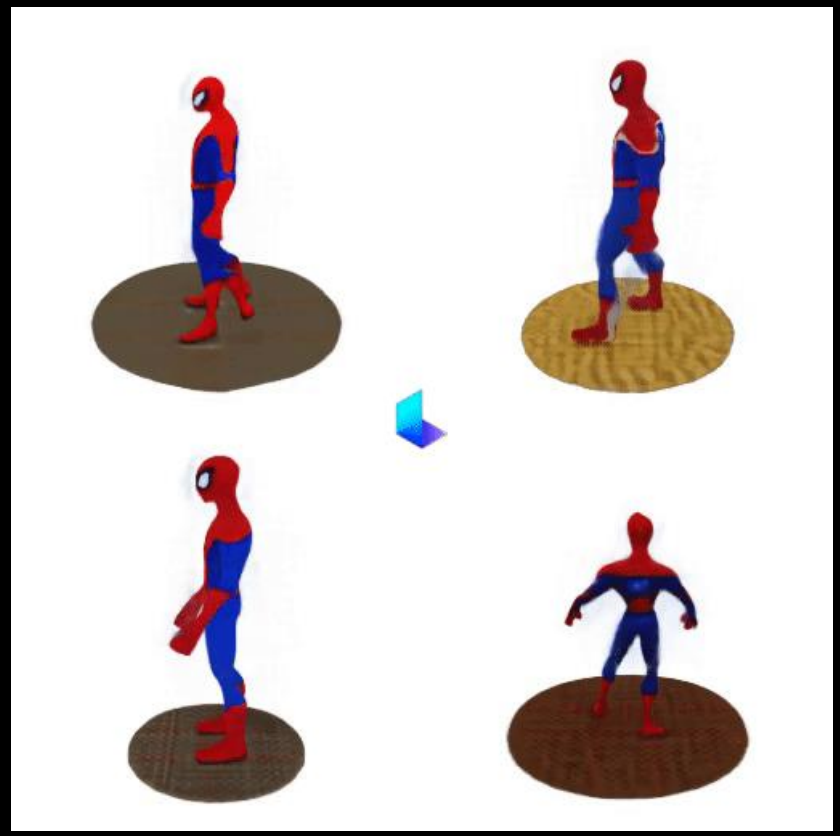
Details in appearance

Extract appearance semantics from vison-language models.

Compositions of objects

"A fox-terrier dog walking in the Alps".

Account for multiple objects and environment composition in 3D



Poses and keypoints in characters/ faces

Geometrical priors, SMPL optimisation, animatable meshes

- To help artists and animators to create using fast and automated Visual Computing and move towards museography.

Tintin, l'aventure immersive, Lausanne, Switzerland

Key Takeaways

- Highly inter-disciplinary research at the intersection of visual computing and art.
- Opens up cross-cultural collaboration with multiple research universities and publishing as well as film animation companies.

Core Visual Computing methods can be used across diverse domains like