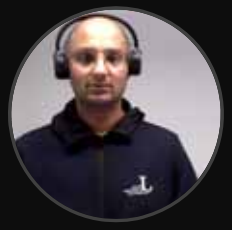
Möbius Transform for Mitigating Perspective Distortions in Representation Learning

ECCV 2024 MILANO

https://prakashchhipa.github.io/projects/mpd/

Prakash Chandra Chhipa[1,*], Meenakshi Subhash Chippa[1], Kanjar De[2], Rajkumar Saini[1], Marcus Liwicki[1], Mubarak Shah[3]



[1]Machine Learning Group, Luleå Tekniska Universitet, Sweden
[2]Fraunhofer HHI, Germany
[3]Center For Research in Computer Vision, University of Central Florida, USA

*corresponding author and presenter

UCF
CENTER FOR RESEARCH IN COMPUTER VISION

LULEÅ TEKNISKA UNIVERSITET

Perspective Distortion

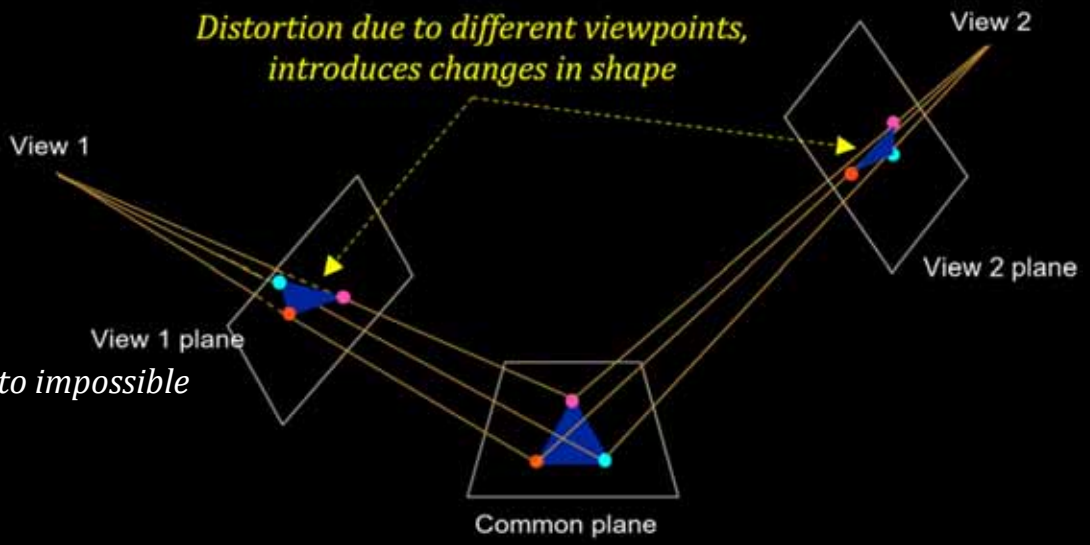
alters object appearance by changing shape, size, orientation, and spatial relationships in non-linear ways.

Möbius Transform for Mitigating Perspective Distortions in Representation Learning, ECCV 2024

https://prakashchhipa.github.io/projects/mpd

Challenges

- Distribution shifts
  - *Distortion disrupts model robustness*
- Distortion corrections is inefficient
  - *Vision tasks becomes multi-step and compute intensive.*
- Camera parameters are unknown
  - *Synthesizing distortions by estimating parameters is next to impossible*
- Affine Transformation
  - *They are either linear or non-conformal*



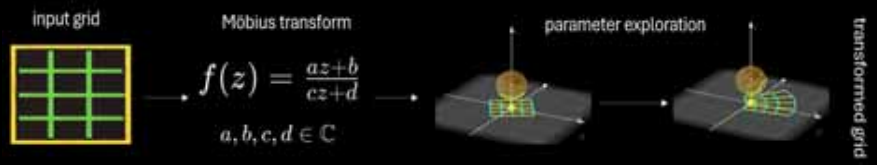
UCF CENTER FOR RESEARCH IN COMPUTER VISION

LULEÅ TEKNISKA UNIVERSITET

Hypothesis

Non-linear and conformal nature of transforms can synthesize perspective distortion without knowing camera parameters.

Möbius transforms – Complex theory



MPD Method

Synthetically introduces perspective distortion in learning

Algorithm 1 MPD Method for Perspective Distortion

Input: Input image I, Möbius parameters c, fixed-value complex numbers of (a, b, d)

Output: Transformed image I_{MPD} representing perspective distortion

1: Initialize $c \leftarrow c_{real} + i \cdot c_{imag}$ ▷ Set complex parameter c
2: **for** each pixel coordinate (x, y) in I **do**
3: $z \leftarrow x + iy$ ▷ Map pixel coordinates to complex vector
4: $z_m \leftarrow \frac{(a \cdot z + b)}{(c \cdot z + d)}$ ▷ Apply Möbius transformation
5: $x_m + iy_m \leftarrow z_m$ ▷ Extract real and imaginary parts as transformed coordinates
6: $(x_d, y_d) \leftarrow \text{round}(x_m), \text{round}(y_m)$ ▷ Discretize to nearest pixel values
7: $I_{MPD}(x_d, y_d) \leftarrow I(x, y)$ ▷ Assign pixel value to transformed coordinates
8: **end for**
9: $I_{MPD} \leftarrow \text{Interpolate}(I_{MPD})$ ▷ Interpolate for smoother transformed output
10: **return** I_{MPD}

11:

Möbius Transform for Mitigating Perspective Distortions in Representation Learning, ECCV 2024

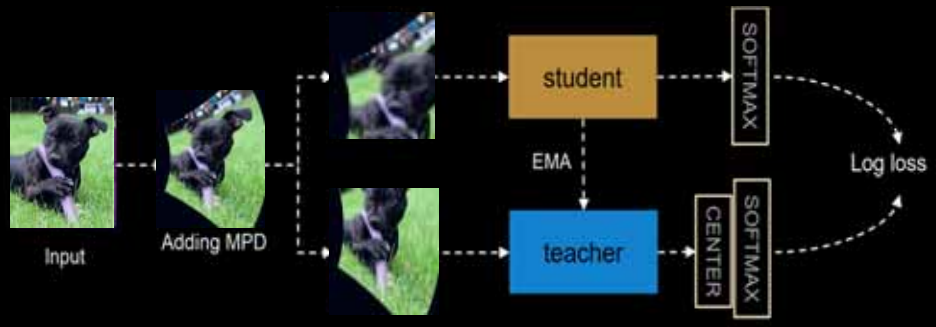
https://prakashchhipa.github.io/projects/mpd

UCF CENTER FOR RESEARCH IN COMPUTER VISION

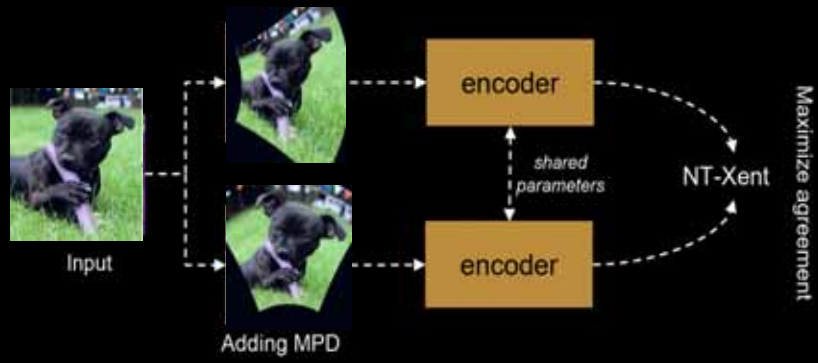
LULEÅ TEKNISKA UNIVERSITET

⊕ Integrating MPD *Supervised and self-supervised approaches*

DINO + MPD



SimCLR + MPD



Pretraining performed on ImageNet

UCF CENTER FOR RESEARCH IN COMPUTER VISION

LULEÅ TEKNISKA UNIVERSITET

⊕ Integrating MPD

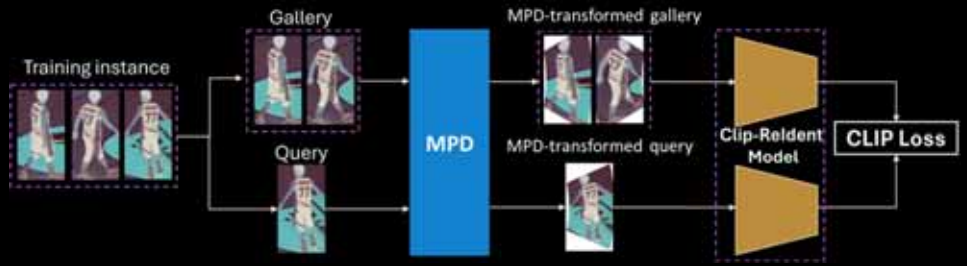
Real-world vision application

Crowd counting (MPD-AutoCrowd)

Object detection (MPD-OD)

Fisheve view recognition

Person Re-Identification + MPD



Möbius Transform for Mitigating Perspective Distortions in Representation Learning, ECCV 2024

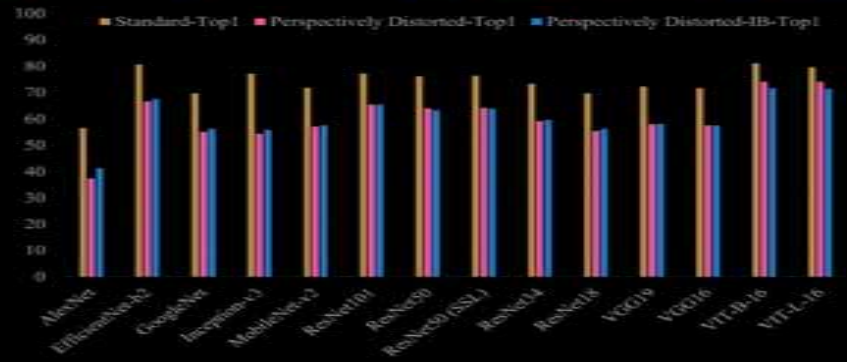
https://prakashchhipa.github.io/projects/mpd

HXUR SHDQ#FRQIHUHQFH#RQ#FRP SXWHU#YLVIRQ#
P ODDQR#5357

UCF
CENTER FOR RESEARCH IN COMPUTER VISION

LULEÅ TEKNISKA UNIVERSITET

ImageNet-PD new dedicated benchmark



Möbius Transform for Mitigating Perspective Distortions in Representation Learning, ECCV 2024

https://prakashchhipa.github.io/projects/mpd

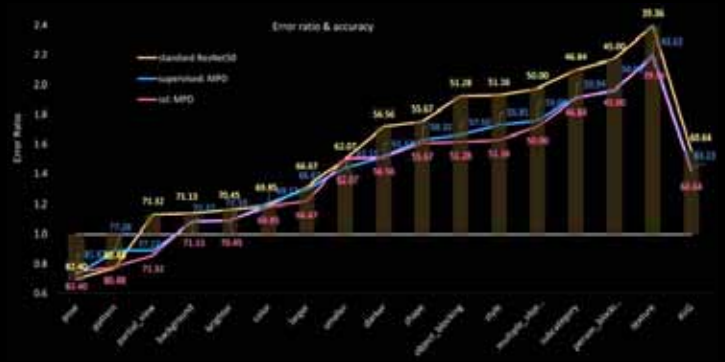
HXUR SHDQ#FRQIHUHQFH#R#FR P SXWHU#YLVIRQ#
P IDOR#5357

UCF
CENTER FOR RESEARCH IN COMPUTER VISION

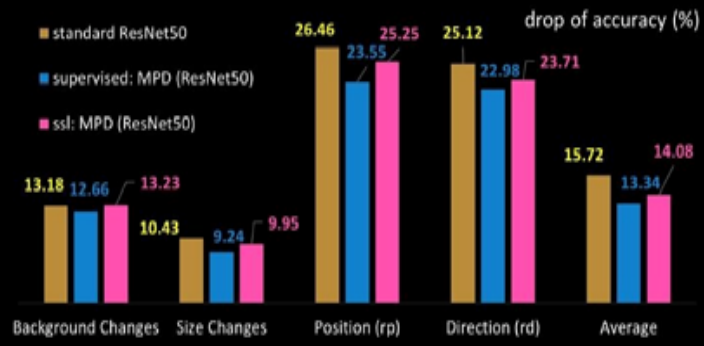
LULEÅ TEKNISKA UNIVERSITET

Results *Public benchmarks*

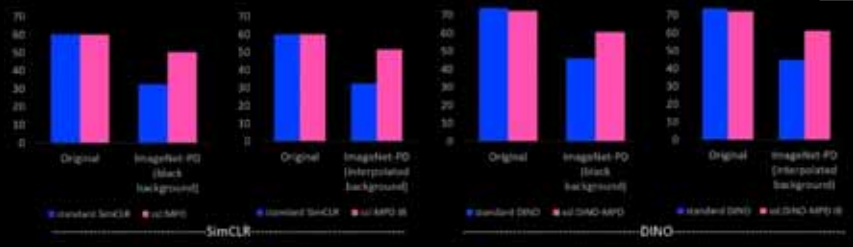
ImageNet-X (error, lower is better)



ImageNet-E (error, lower is better)



Performance on ImageNet-PD (accuracy, higher is better)



Möbius Transform for Mitigating Perspective Distortions in Representation Learning, ECCV 2024

https://prakashchhipa.github.io/projects/mpd

HXUR.SHDQ#FRQIHUHQFH#RQ#FRPSXWHU#ZLVLRQ#
PIDDOR#5357

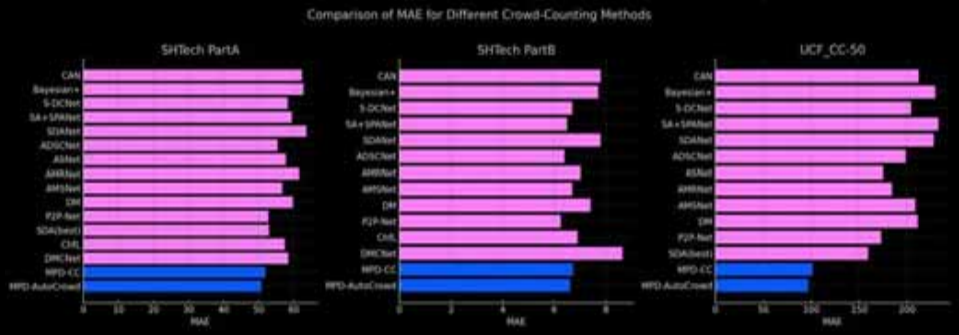
UCF
CENTER FOR RESEARCH IN COMPUTER VISION

LULEÅ TEKNISKA UNIVERSITET

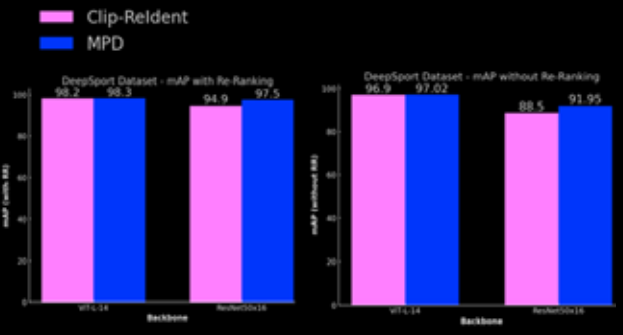
Results

Real-world vision application

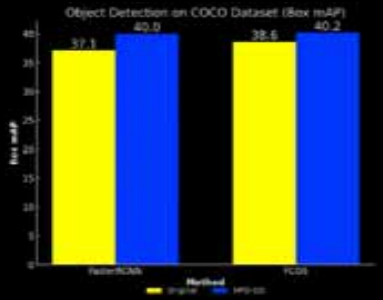
Crowd counting (MAE, lower is better, three public datasets)



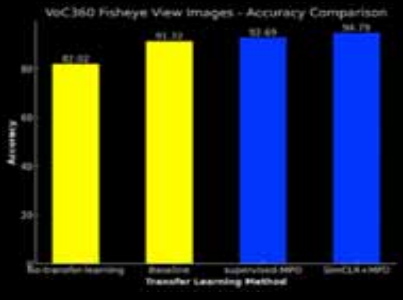
Person Re-Identification (mAP, higher is better)



Object detection (mAP, higher is better, COCO dataset)



Fisheye view recognition (accuracy, VOC360 dataset)



Möbius Transform for Mitigating Perspective Distortions in Representation Learning, ECCV 2024

https://prakashchhipa.github.io/projects/mpd

UCF CENTER FOR RESEARCH IN COMPUTER VISION

LULEÅ TEKNISKA UNIVERSITET

Contribution + Achievements

- ✓ Mathematically modeled perspective distortion without camera parameters
  - ✓ *Mitigating perspective distortion (MPD)*
- ✓ Developed dedicated perspective distorted dataset *ImageNet-PD*
  - ✓ *8 subsets*
- ✓ Seamless integration with computer vision tasks and applications
  - ✓ *MPD- AutoCrowd, MPD-Object Detection, outperformed multiple benchmarks*
- ✓ Generalizability, transferability, and adaptability across learning paradigms, architectures, and applications.
  - ✓ *Supervised and self-supervised learning, CNNs and ViTs*

HXUR SHDQ#FRQIHUHQFH#RQ#FR P SXWHU#YLVIRQ#
P ODDQR#5357

Thank you