IEEE Video and Image Processing Cup 2021

Privacy Preserving In-Bed Human Pose Estimation

Team NFP-Undercover

Department of Electronic and Telecommunication Engineering,

University of Moratuwa, Sri Lanka.

Team Members :

Jathurshan Pradeepkumar, Udith Haputhanthri, Mohamed Afham, Mithunjha Anandakumar

Graduate Member : Ashwin De Silva

Supervisor : Dr. Chamira Edussooriya

Task

LWIR Imaging Modality

RGB Imaging Modality

- Occlusion – Partially
- High Resolution
- Illumination Changes
- Accurate Pose Estimation
- Annotations

Human Pose Estimation

Robust Human Pose Estimation

- Non-Contact
- Illumination Changes
- Labeled & Uncovered
- Unlabeled & Covered

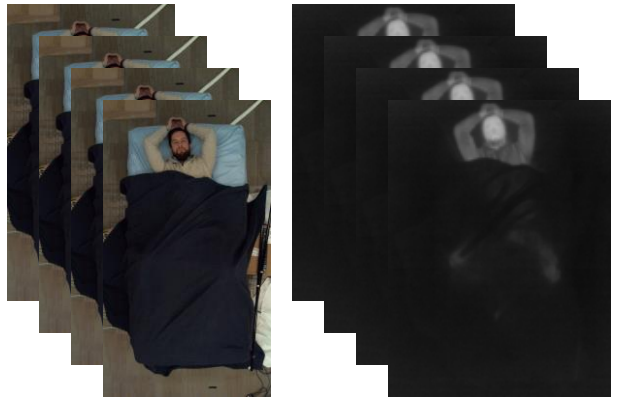
Shuangjun Liu and Sarah Ostadabbas, “Seeing under the cover: A physics guided learning approach for in-bed pose estimation,”22nd International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI2019), Shenzhen, China. arXiv preprint arXiv:1907.02161, 2019.

Task

Training Dataset

Uncovered

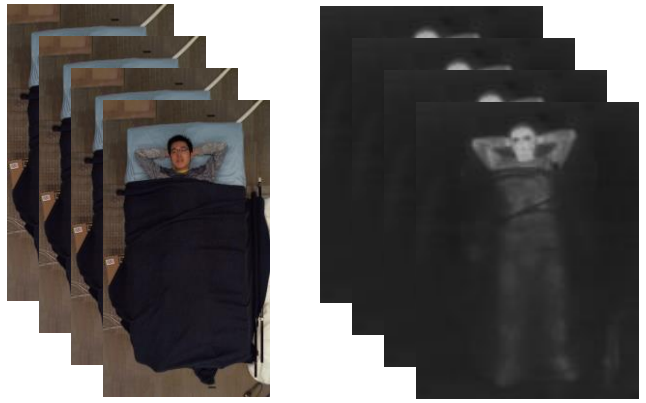
Labeled



Covered

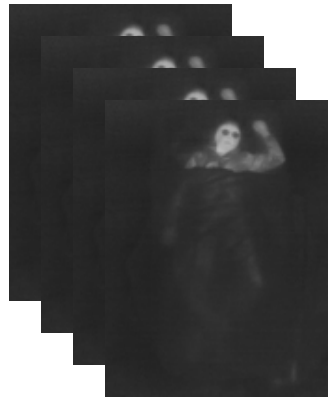
Unlabeled

Validation Dataset



Covered & Labeled

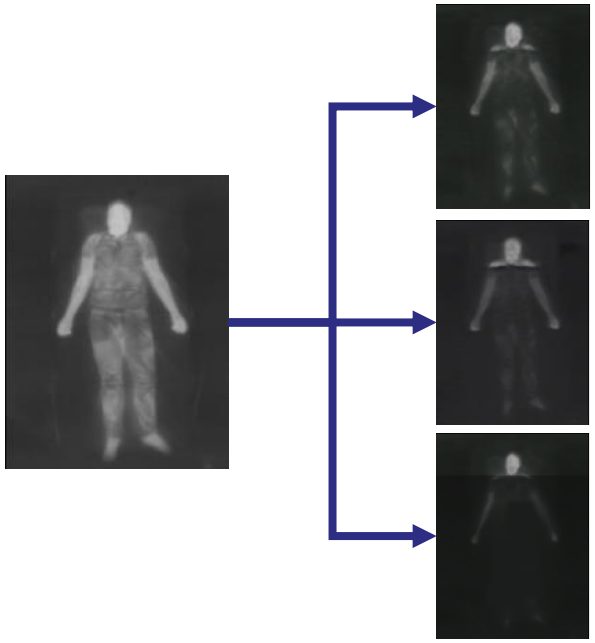
Test Dataset



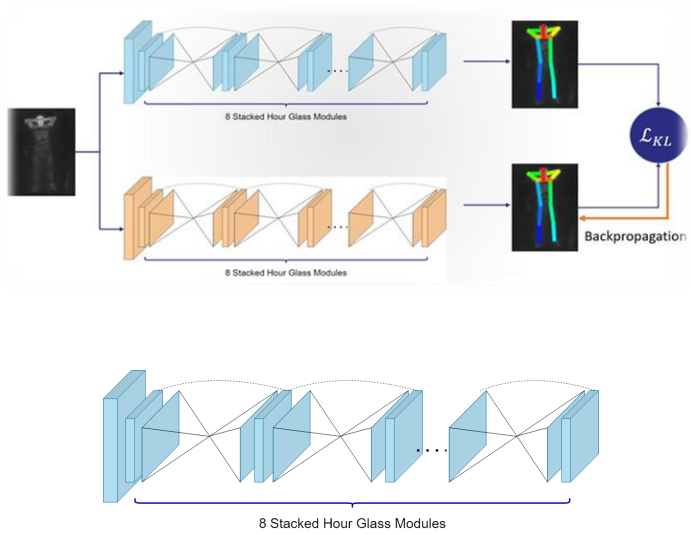
Covered

S. Liu and S. Ostadabbas, “Seeing under the cover: A physics guided learning approach for in-bed pose estimation,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, vol. 11764 LNCS.

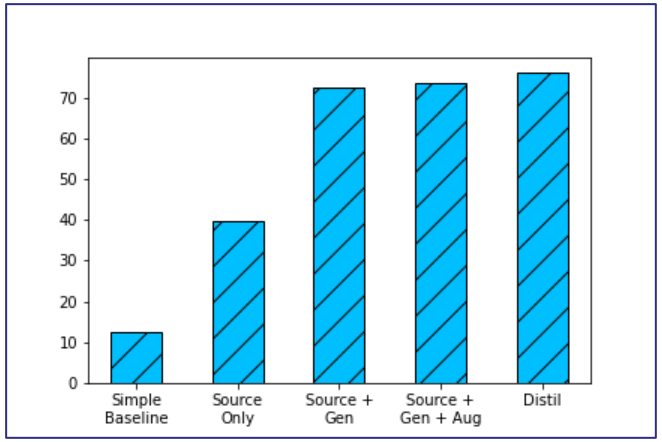
Our Contribution



Two-fold Data Augmentation for Cross-Domain Discrepancy Reduction



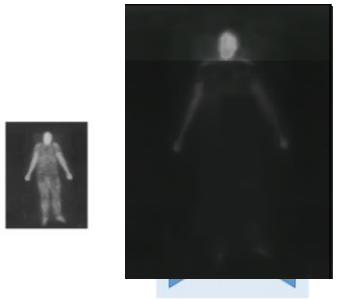
Adapting Simple Stacked Hourglass Model for Cross Domain In-bed Pose Estimation

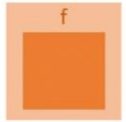


Ablation and Analysis

Algorithm: Overview

Data Augmentation



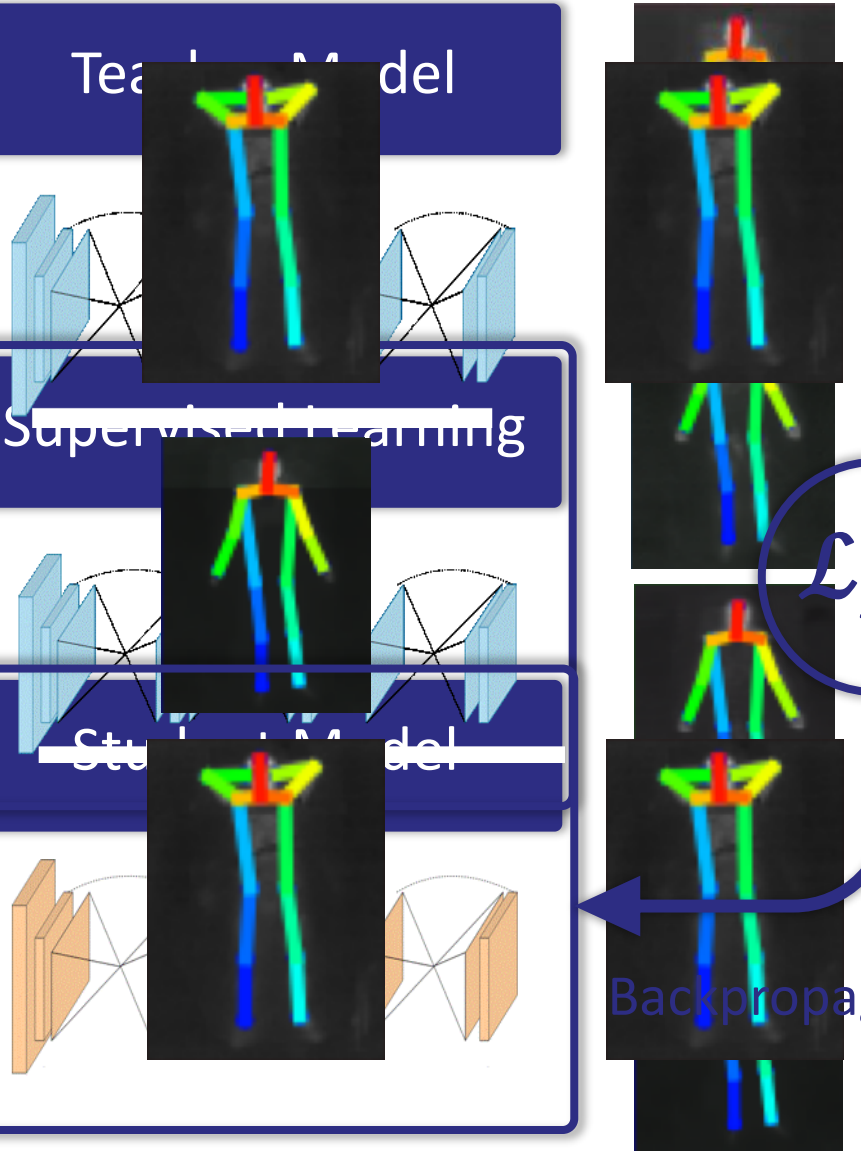


Knowledge Distillation

Teacher Model

Supervised Learning

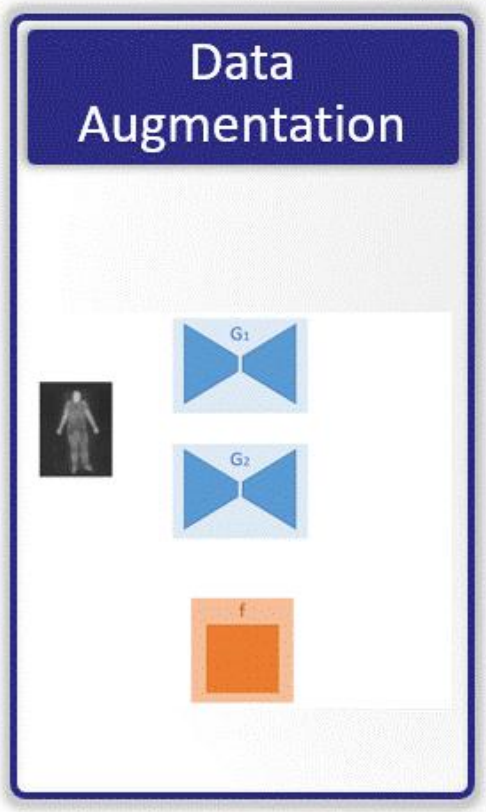
Student Model



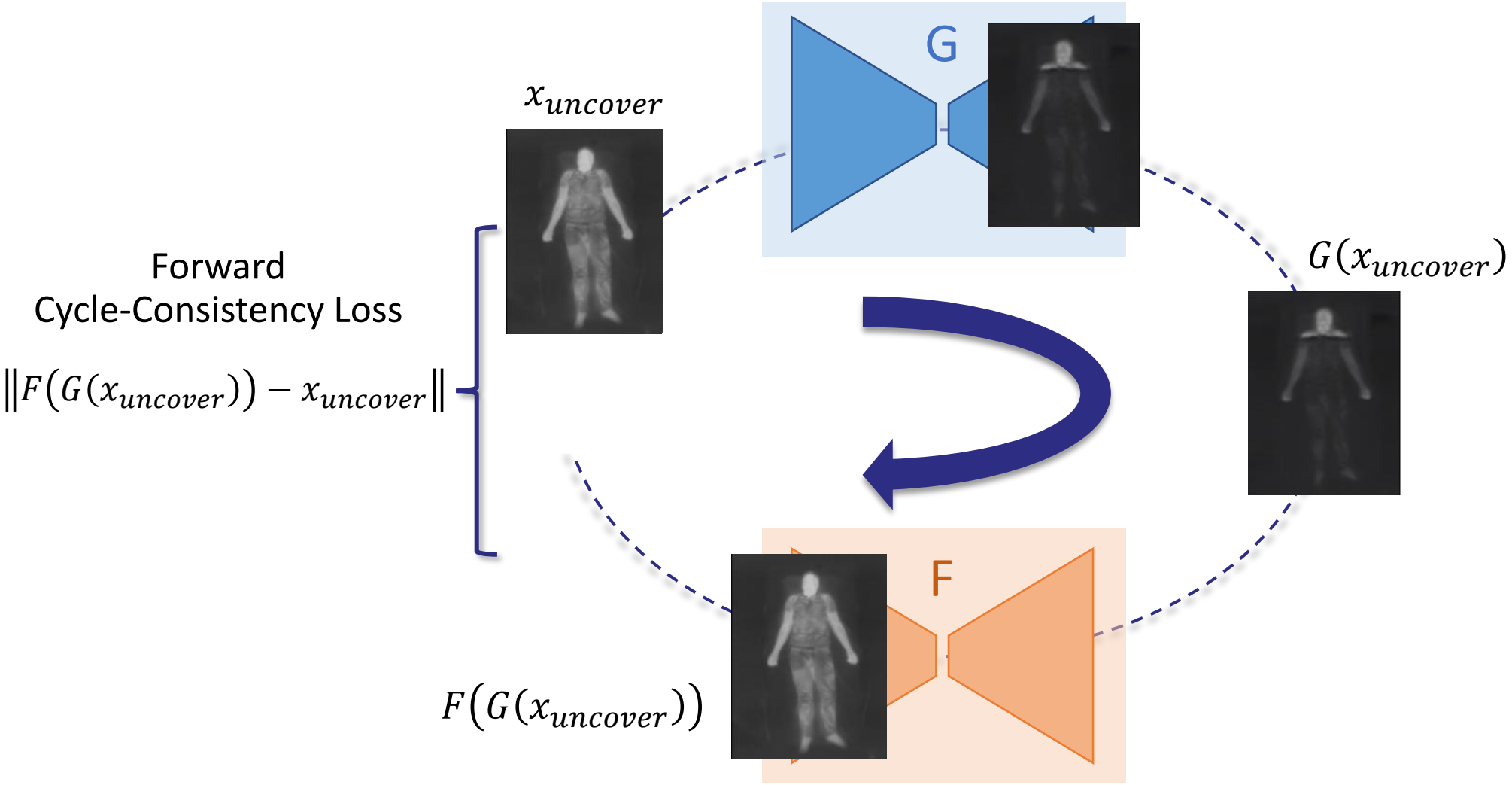
$\mathcal{L}_{KL}$

Backpropagation

Two-fold Data Augmentation for Cross-Domain Discrepancy Reduction

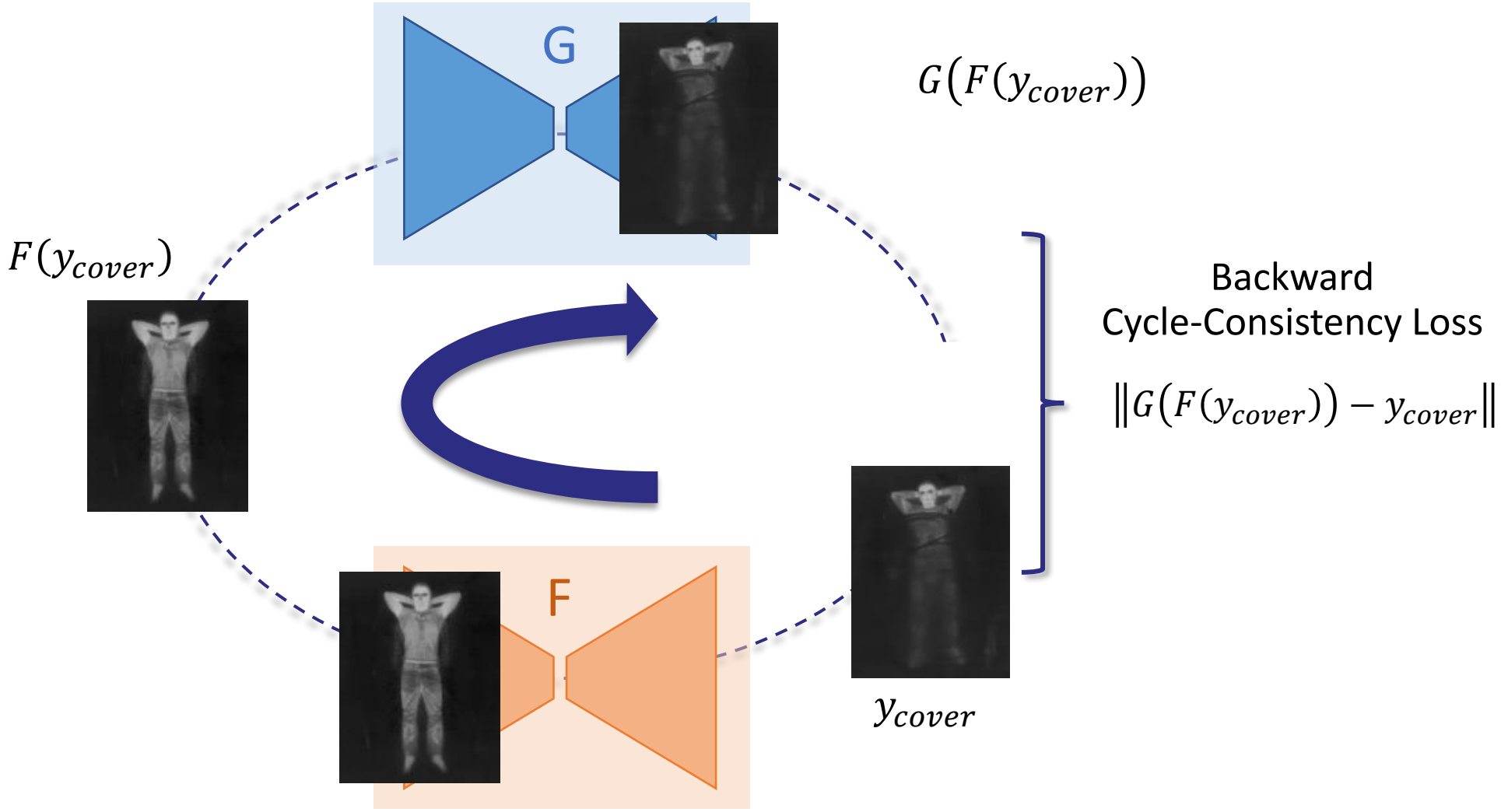


Augmentation 1: CycAug (Training)

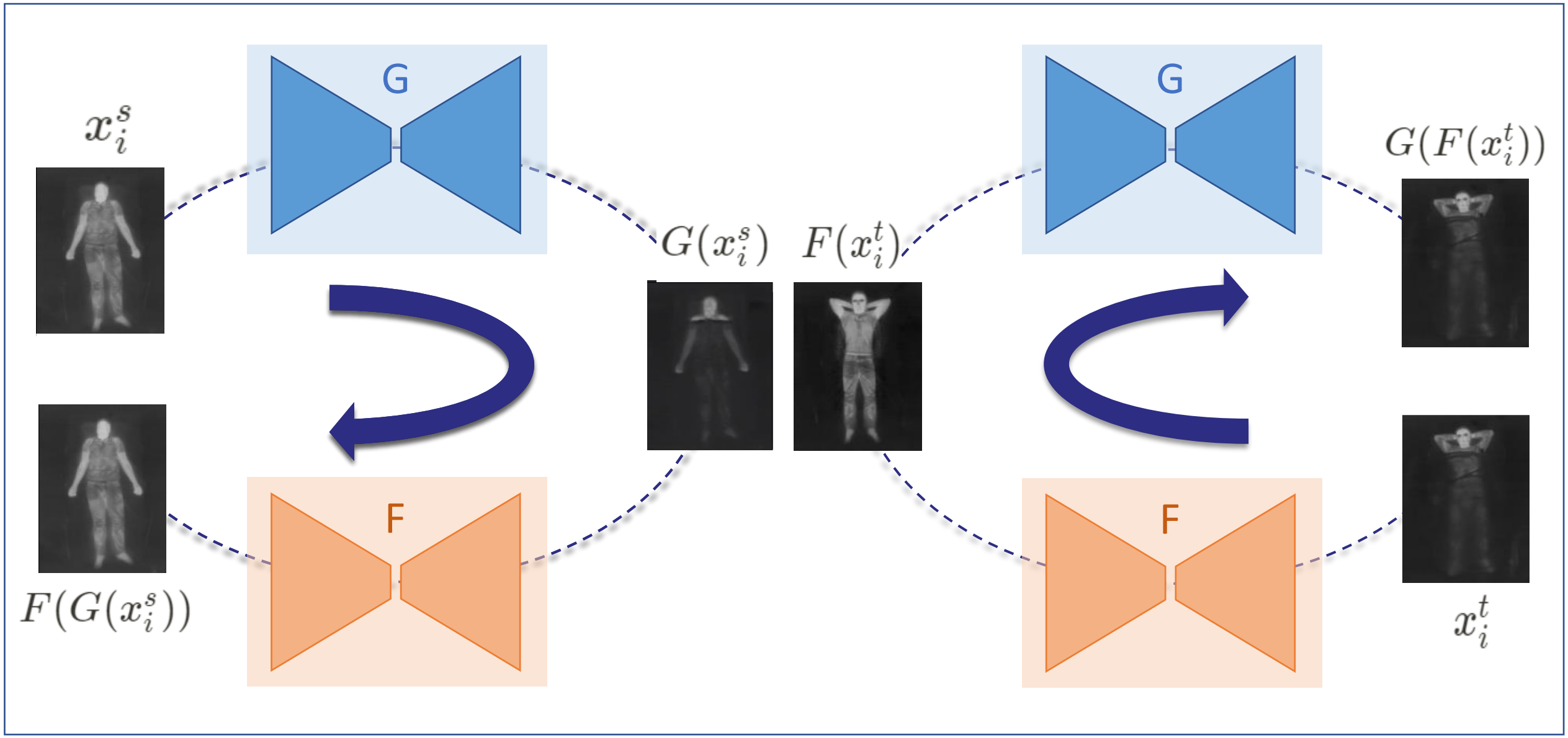


J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, vol. 2017-October.

Augmentation 1: CycAug (Training)



J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, vol. 2017-October.


G
G
x_i^s
$G(x_i^s)$
$F(x_i^t)$
$G(F(x_i^t))$
F
F
$F(G(x_i^s))$
x_i^t

Augmentation 1: CycAug (Training)

$$Loss = Adversarial\ Loss + (\lambda \times Cycle\ Consistency\ Loss)$$

More Realistic

Unpaired Translation

$$\mathcal{L}_{GAN}(G, D_y, X_{uncover}, Y_{cover}) + \mathcal{L}_{GAN}(F, D_x, Y_{cover}, X_{uncover}) + \lambda$$

$forward\ cyc\ loss$

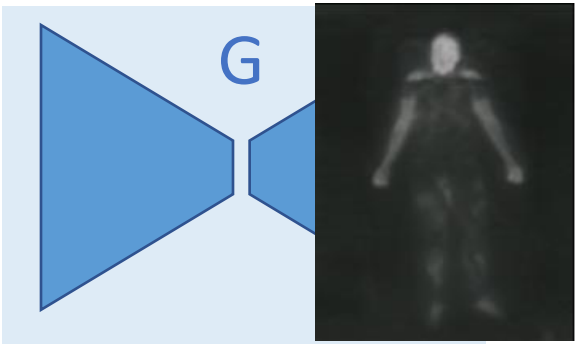
$backward\ cyc\ loss$

$$\mathrm{E}_{x \sim p_{data}(x_{uncover})}[\|F(G(x_{uncover})) - x_{uncover}\|] + \mathrm{E}_{y \sim p_{data}(y_{cover})}[\|G(F(y_{cover})) - y_{cover}\|]$$

J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, vol. 2017-October.

Augmentation 1: CycAug (Inferencing)

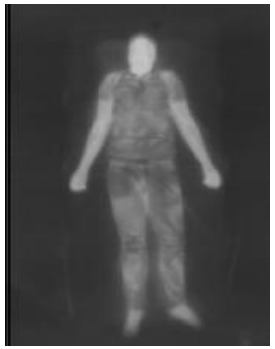
Uncovered LWIR Image



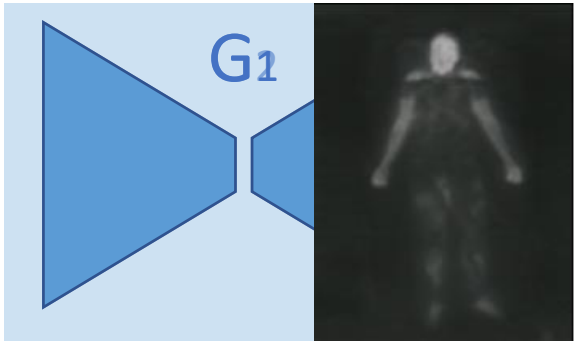
Generator

Generated Covered LWIR Image

Augmentation 1: CycAug (Inferencing)



Uncovered LWIR Image



Generator for Thin Cover

Generated Thin Covered LWIR Image

Generator for Thick Cover

Generated Thick Covered LWIR Image

Augmentation 2: ExtremeAug

Augmentation 2: ExtremeAug

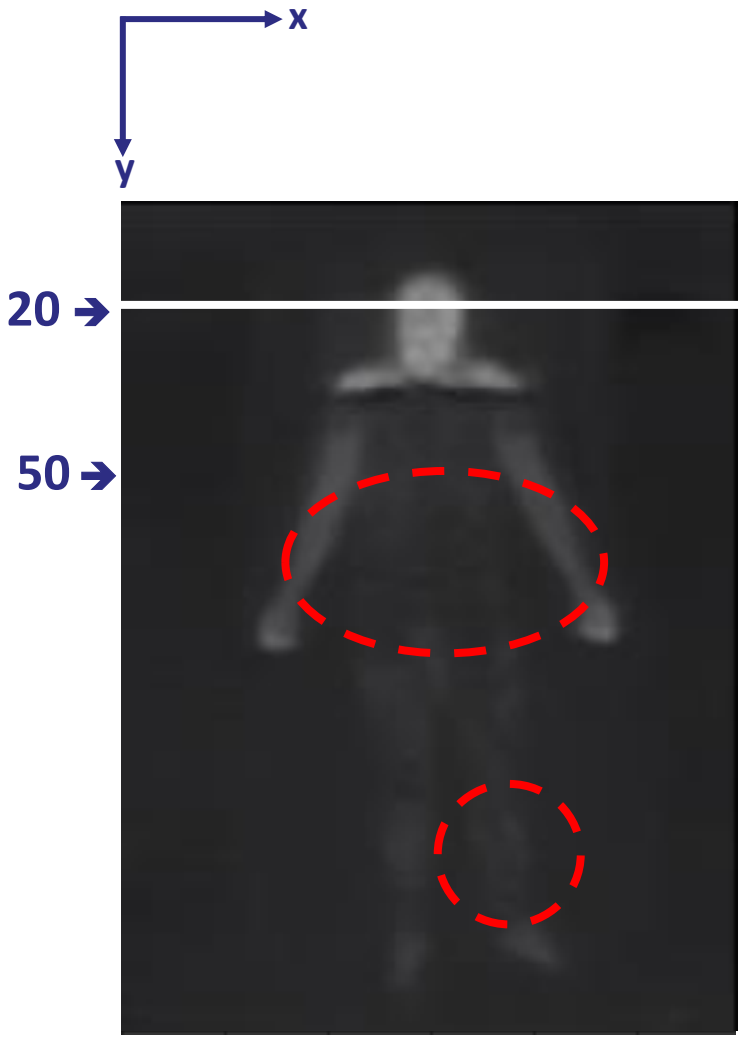
Select random value r ϵ [20,50]

$$X_{cover}(x,y) = \frac{X_{cover}(x,y)}{2}, if\ y \geq r$$

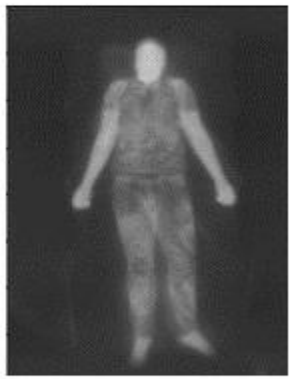
Add random dark kernels below r

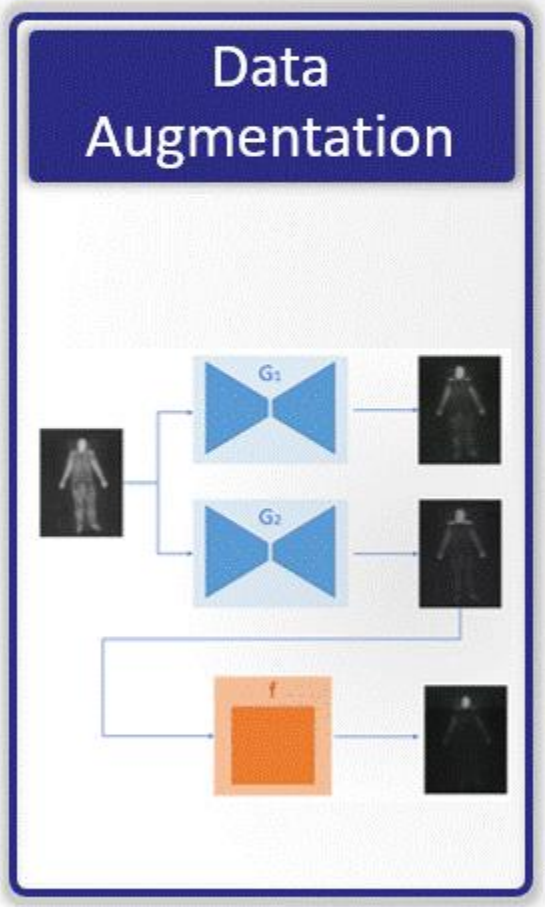
Morphological Operation (Erosion)

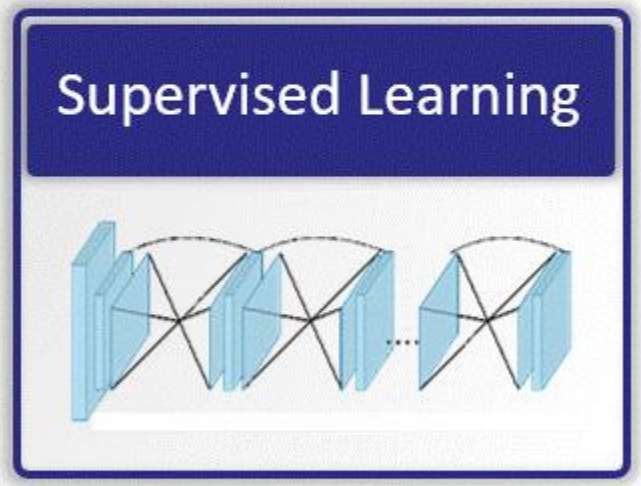
Gaussian Blurring



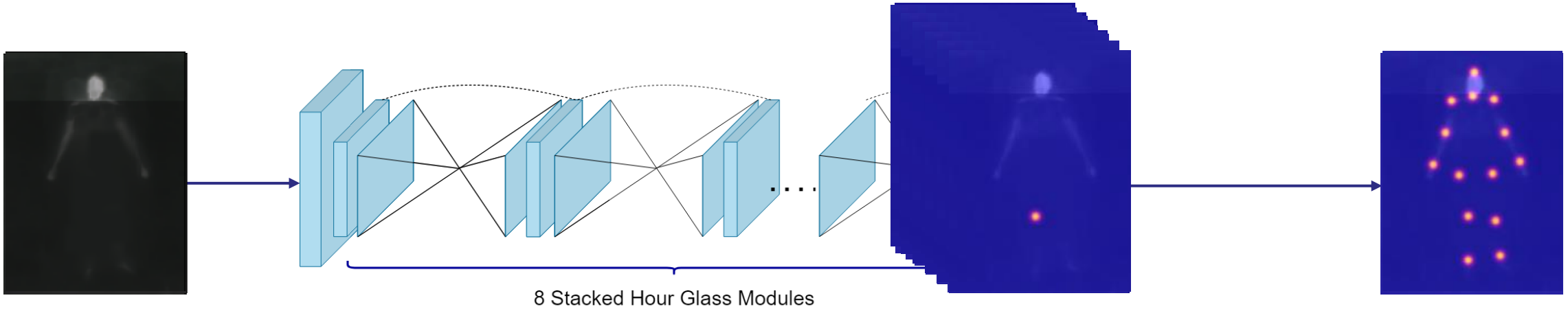
Supervised Learning : Pose Estimation

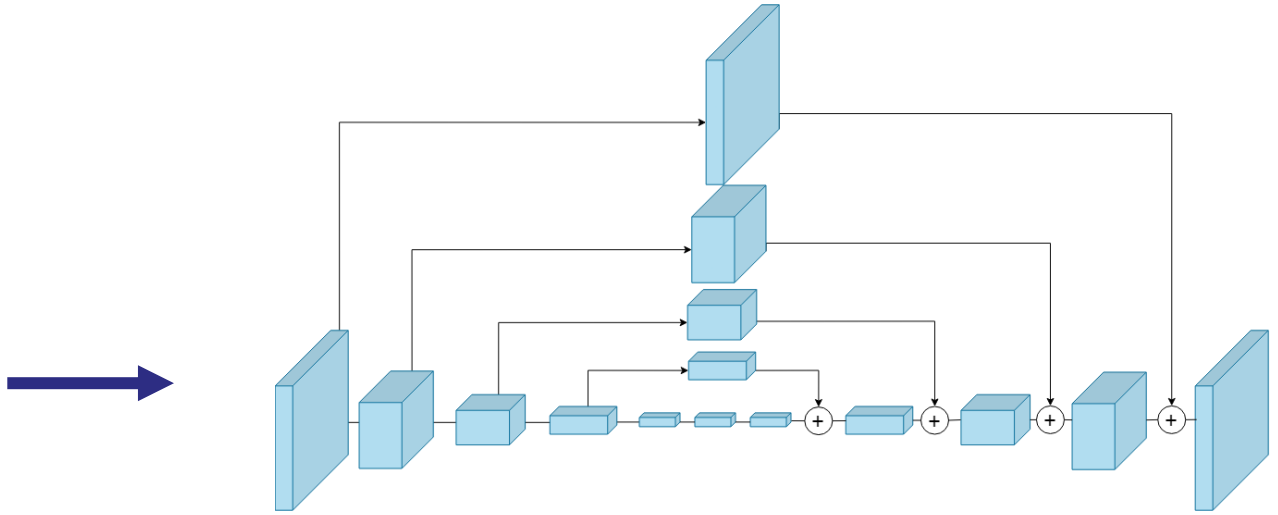






Supervised Learning: Pose Estimation

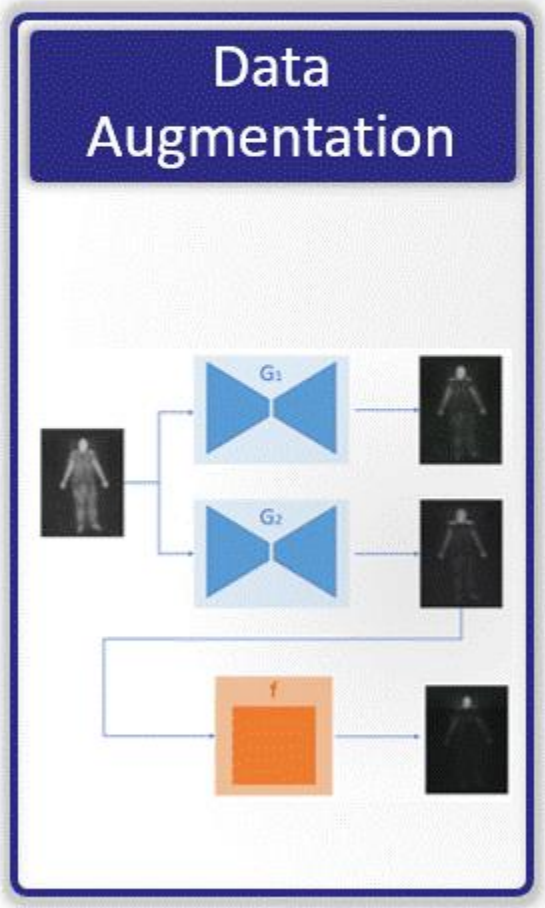




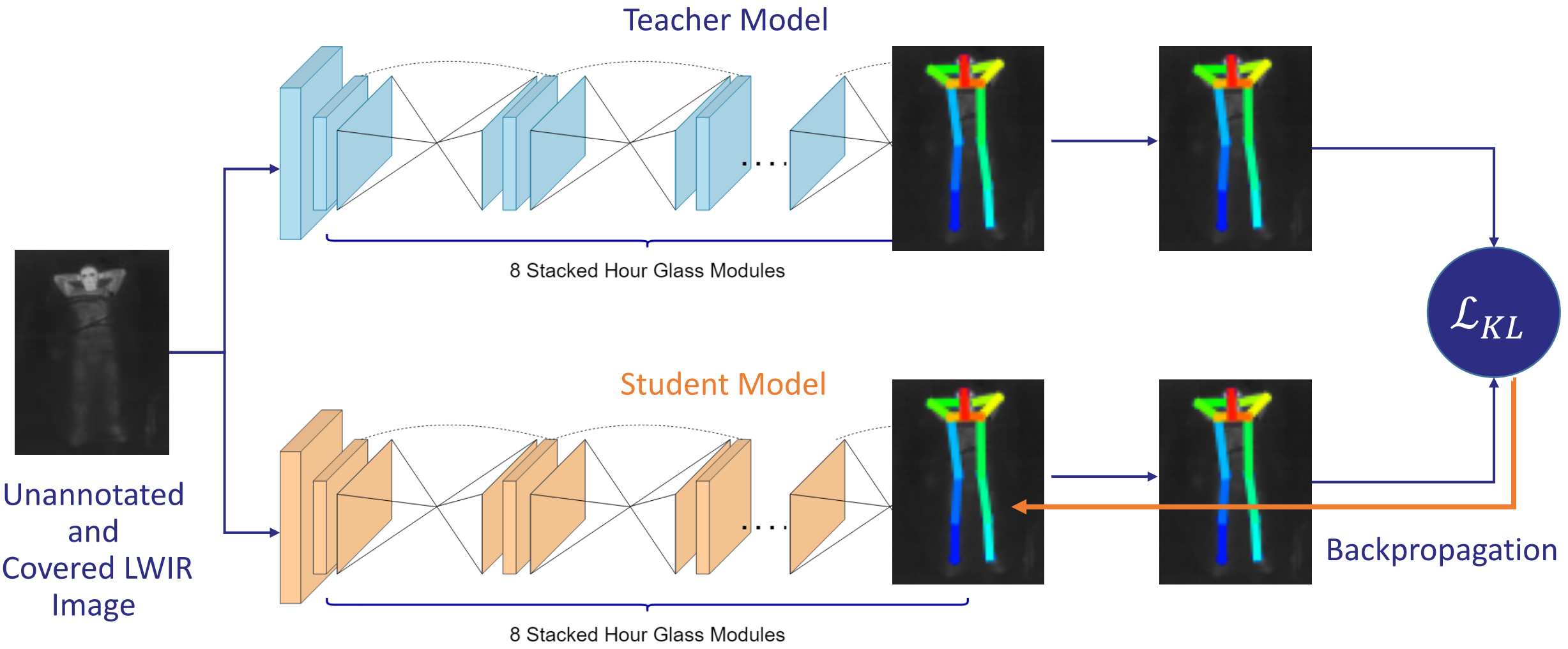
$$\mathcal{L}_{MSE\_Pose}$$

Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hour- glass networks for human pose estimation. In *ECCV*, pages 483–499, 2016

Knowledge Distillation



Knowledge Distillation



G. Hinton, O. Vinyals, and J. Dean. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*, 2015.
Kullback S, Leibler RA On information and sufficiency. Ann Math Stat 22:79–86

Results (Test1)

Method	pckAUC	pck@0.2	pck@0.5	pckAUC - hospital	pckAUC - home
Simple Baseline [1]	5.96	3.21	12.41	7.48	4.95
Source (Uncover) data only [2]	21.78	15.59	39.52	21.55	21.93
Uncover data + CycAug	46.42	40.22	72.44	46.57	46.32
Uncover data + CycAug + ExtremeAug	47.16	40.46	73.38	47.48	46.95
Uncover data + CycAug + ExtremeAug + Knowledge Distillation	**49.8**	**43.65**	**76.13**	**50.62**	**49.25**

[1] Bin Xiao, Haiping Wu, and Yichen Wei. Simple baselines for human pose estimation and tracking. In *ECCV*, pages 472–487, 2018.
[2] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hour- glass networks for human pose estimation. In *ECCV*, pages 483–499, 2016

Results (Test1)

Method	pckAUC	pck@0.2	pck@0.5	pckAUC - hospital	pckAUC - home
Simple Baseline [1]	5.96	3.21	12.41	7.48	4.95
Source (Uncover) data only [2]	21.78	15.59	39.52	21.55	21.93
Uncover data + CycAug	46.42	40.22	72.44	46.57	46.32
Uncover data + CycAug + ExtremeAug	47.16	40.46	73.38	47.48	46.95
Uncover data + CycAug + ExtremeAug + Knowledge Distillation	**49.8**	**43.65**	**76.13**	**50.62**	**49.25**

[1] Bin Xiao, Haiping Wu, and Yichen Wei. Simple baselines for human pose estimation and tracking. In *ECCV*, pages 472–487, 2018.
[2] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hour- glass networks for human pose estimation. In *ECCV*, pages 483–499, 2016

Results (Test1)

Method	pckAUC	pck@0.2	pck@0.5	pckAUC - hospital	pckAUC - home
Simple Baseline [1]	5.96	3.21	12.41	7.48	4.95
Source (Uncover) data only [2]	21.78	15.59	39.52	21.55	21.93
Uncover data + CycAug	46.42	40.22	72.44	46.57	46.32
Uncover data + CycAug + ExtremeAug	47.16	40.46	73.38	47.48	46.95
Uncover data + CycAug + ExtremeAug + Knowledge Distillation	**49.8**	**43.65**	**76.13**	**50.62**	**49.25**

[1] Bin Xiao, Haiping Wu, and Yichen Wei. Simple baselines for human pose estimation and tracking. In *ECCV*, pages 472–487, 2018.
[2] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hour- glass networks for human pose estimation. In *ECCV*, pages 483–499, 2016

Results (Test1)

Method	pckAUC	pck@0.2	pck@0.5	pckAUC - hospital	pckAUC - home
Simple Baseline [1]	5.96	3.21	12.41	7.48	4.95
Source (Uncover) data only [2]	21.78	15.59	39.52	21.55	21.93
Uncover data + CycAug	46.42	40.22	72.44	46.57	46.32
Uncover data + CycAug + ExtremeAug	47.16	40.46	73.38	47.48	46.95
Uncover data + CycAug + ExtremeAug + Knowledge Distillation	**49.8**	**43.65**	**76.13**	**50.62**	**49.25**

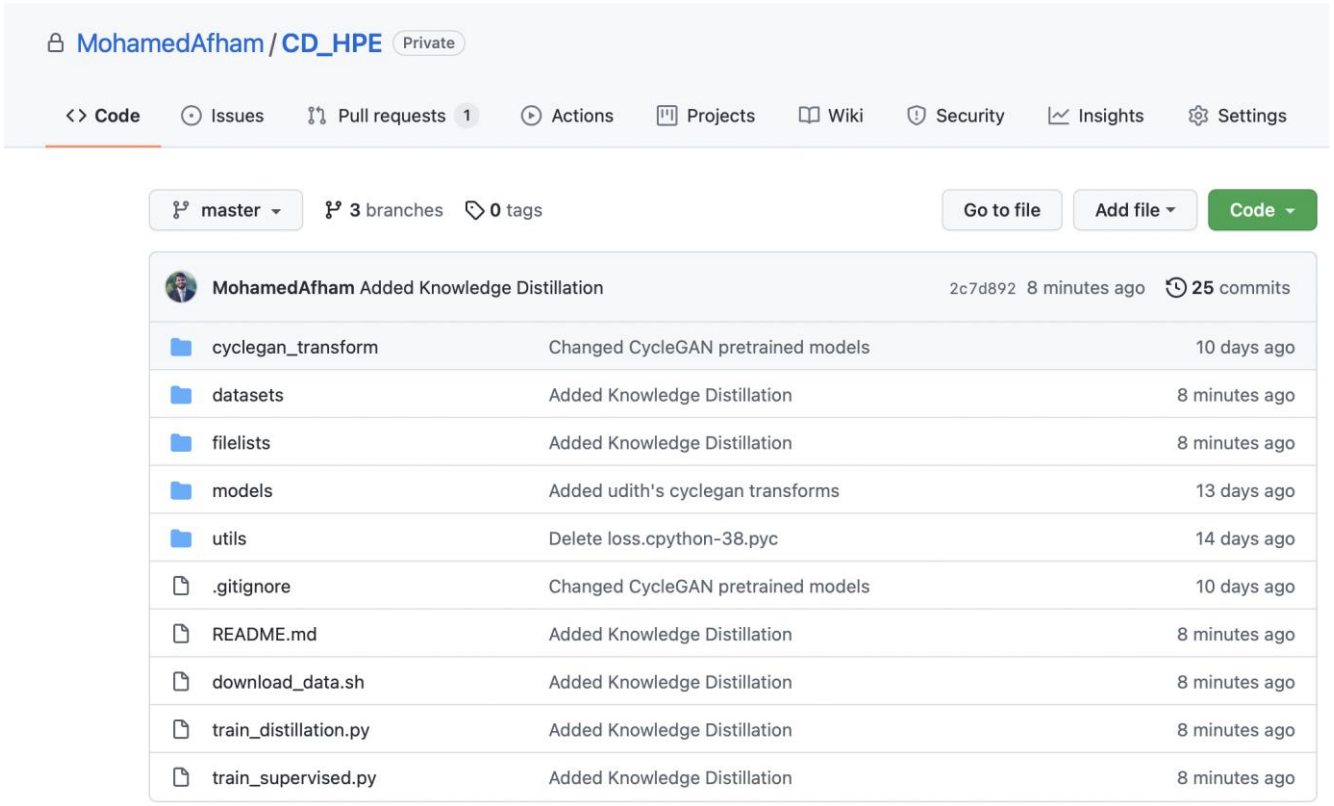
[1] Bin Xiao, Haiping Wu, and Yichen Wei. Simple baselines for human pose estimation and tracking. In *ECCV*, pages 472–487, 2018.
[2] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hour- glass networks for human pose estimation. In *ECCV*, pages 483–499, 2016

Results (Test2)

Method	pckAUC	pck@0.2	pck@0.5	pckAUC - hospital	pckAUC - home
Our best model	**50.065**	**42.187**	**77.46**	**50.270**	**49.963**

Estimated Prediction Time per image	
With CPU	**With GPU**
1.3512 s	**0.14069 s**

Reproducibility



Link to GitHub Repo: https://github.com/MohamedAfham/CD_HPE

SCAN ME

README.md

Cross Domain - Human Pose Estimation

VIPCUP 2021 - Team NFP_Undercover

Pose Definition

We follow the Leeds Sports Pose Dataset pose definition with 14 joints labeling, namely,

`Right ankle Right knee Right hip Left hip Left knee Left ankle Right wrist Right elbow Right shoulder Left shoulder Left elbow Left wrist Neck Head top`

The lable matrix `joints_gt_<modality>.mat` has the format `<x,y,if_occluded>` x n_joints x n_subjects

Dataset

Create a folder named data/SLP_VIPCup.

```
mkdir data
cd data
mkdir SLP_VIPCup
cd ..
```

Download dataset from codalab. Unzip the data inside data/SLP_VIPCup/ by running:

```
source dowload_data.sh
```

The directory should look like

```
CD_HPE/
    data/
        SLP_VIPCup/
            train/
            test1/
            val/
```

Run the following to create the required `.json` files for the dataset

```
cd filelists
source create_filelist.sh
cd ..
```

Train

For Learning Stage 01 (Standard Supervision)

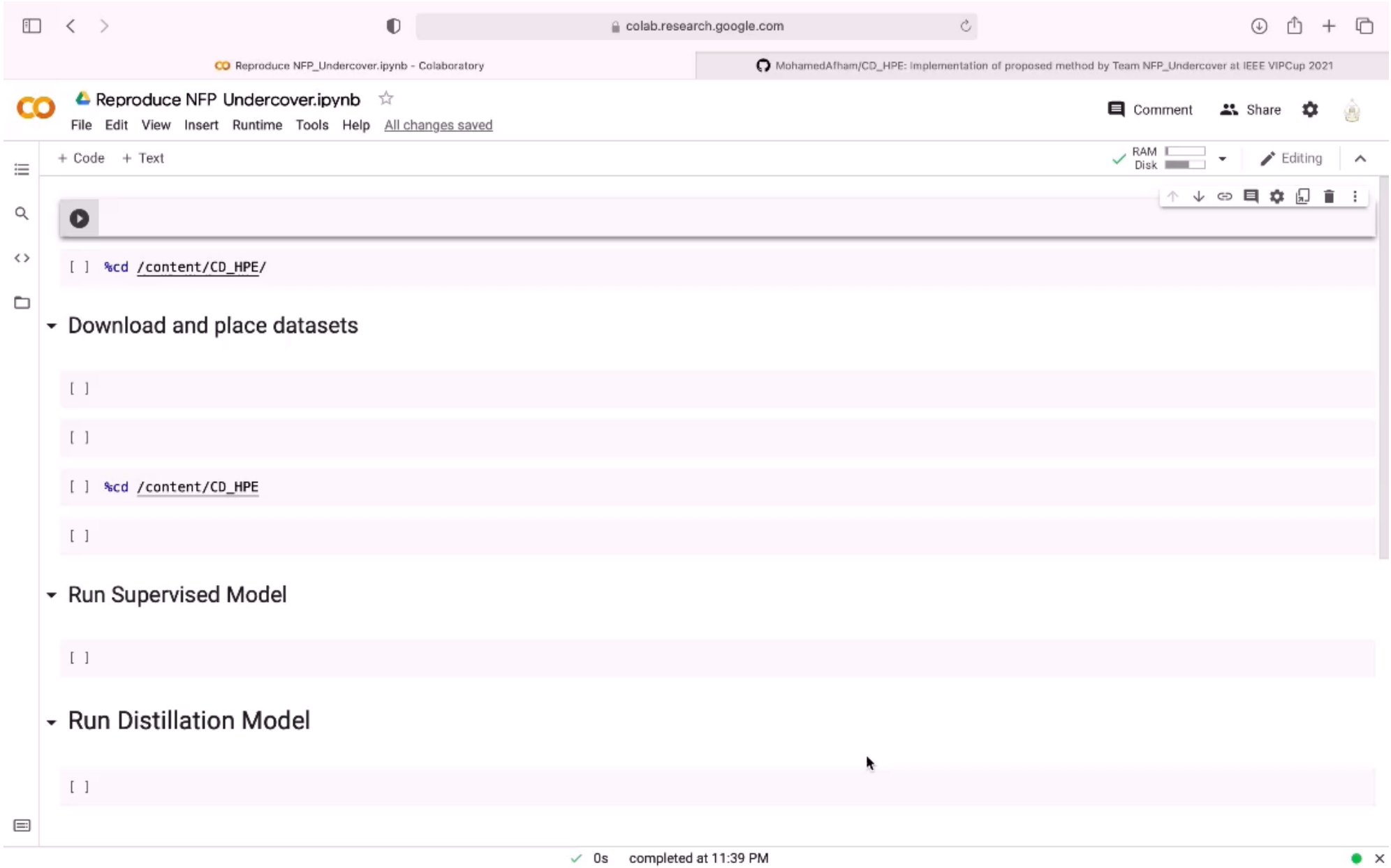
```
python train_supervised.py --adam --use_target_weight --model stacked_hg -
```

For Learning Stage 02 (Knowledge Distillation)

```
python train_distillation.py --adam --best_path path/to/best/model --model
```

Conclusion

- ✓ **Domain Adaptive**
- ✓ Two-fold Data Augmentation for Cross-Domain Discrepancy Reduction
- ✓ Two – Stage **novel** learning pipeline
- ✓ Agnostic to modality
- ✓ Trainable using **limited resources**
- ✓ Code **publicly** available


colab.research.google.com
Reproduce NFP_Undercover.ipynb - Colaboratory
MohamedAfham/CD_HPE: Implementation of proposed method by Team NFP_Undercover at IEEE VIPCup 2021
Reproduce NFP Undercover.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
[ ] %cd /content/CD_HPE/
Download and place datasets
[ ]
[ ]
[ ] %cd /content/CD_HPE
[ ]
Run Supervised Model
[ ]
Run Distillation Model
[ ]
0s completed at 11:39 PM

SCAN ME

We are,

Jathurshan Pradeepkumar
Undergraduate
Department of Electronic and Telecommunication
University of Moratuwa

Udith Haputhanthri
Undergraduate
Department of Electronic and Telecommunication
University of Moratuwa

Mohamed Afham
Undergraduate
Department of Electronic and Telecommunication
University of Moratuwa

Mithunjha Anandakumar
Undergraduate
Department of Electronic and Telecommunication
University of Moratuwa

Ashwin De Silva
Graduate Member (Tutor)
Johns Hopkins University

Dr. Chamira Edussooriya
B.Sc.Eng. (Moratuwa) M.A.Sc., Ph.D. (UVic), MIEEE
Supervisor
Senior Lecturer
Department of Electronic and Telecommunication
University of Moratuwa