JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

What Happened 3 Seconds Ago? Inferring the Past with Thermal Imaging

THU-AM-060



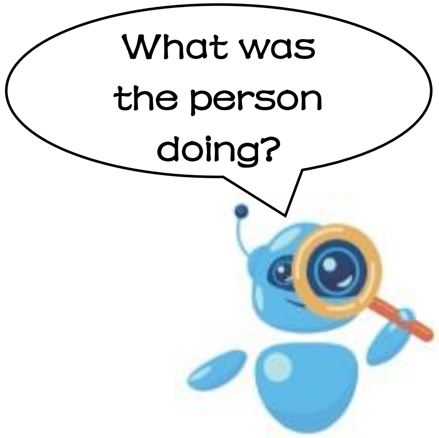

Zitian Tang[1]   Wenjie Ye[1]   Wei-Chiu Ma[2]   Hang Zhao[1,3]

[1]Tsinghua University, IIIS   [2]MIT, CSAIL   [3]Shanghai Qi Zhi Institute

MARS Lab

Motivation

Can you tell what the person was doing 3 seconds ago?

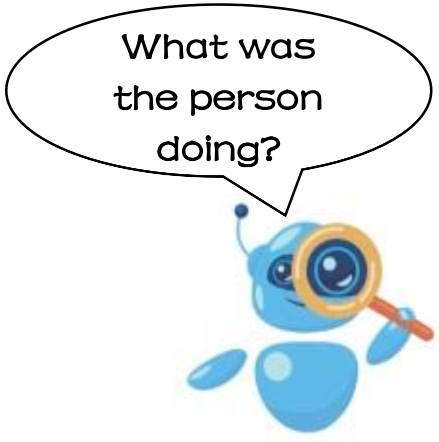


Motivation

Heat transfers to the object during human-object interaction

Motivation

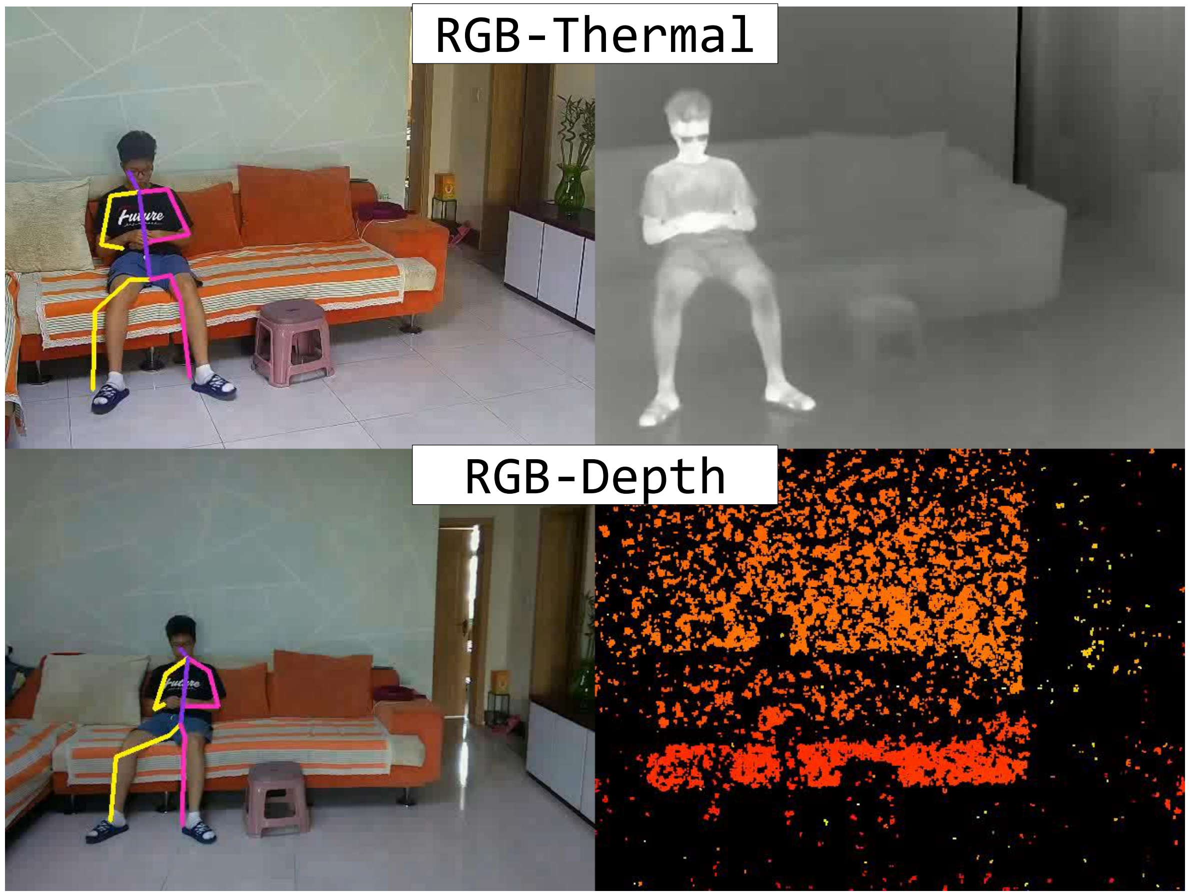
With a thermal image, the answer can be certain



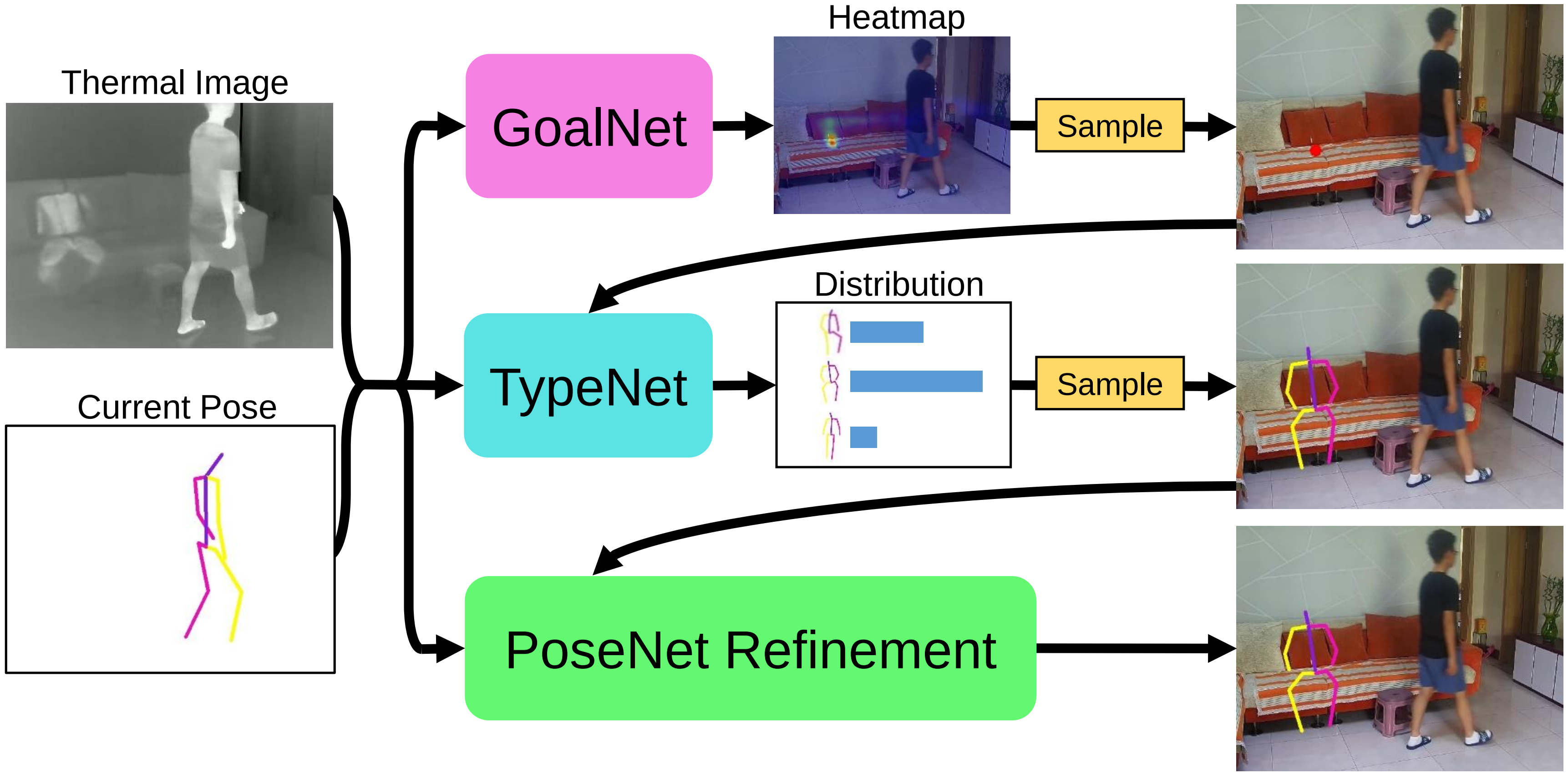
Thermal Indoor Motion Dataset

RGB-Thermal and RGB-Depth videos of indoor human motion with estimated human poses.

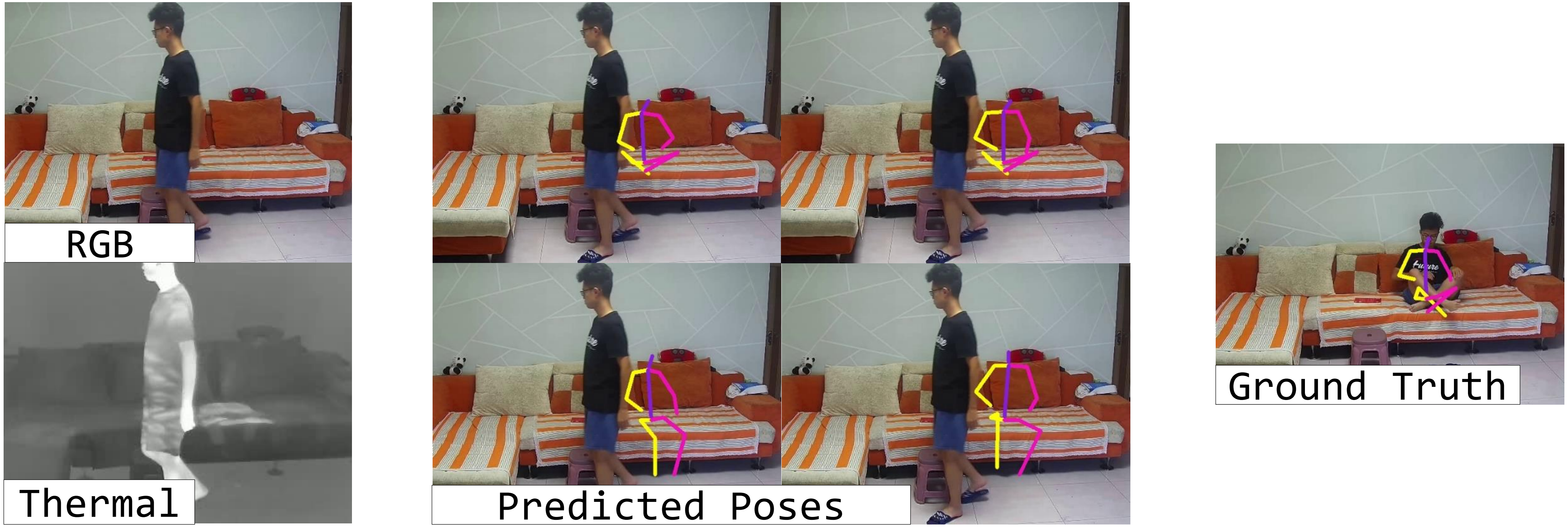
783 video clips, 10.4 hours



Method

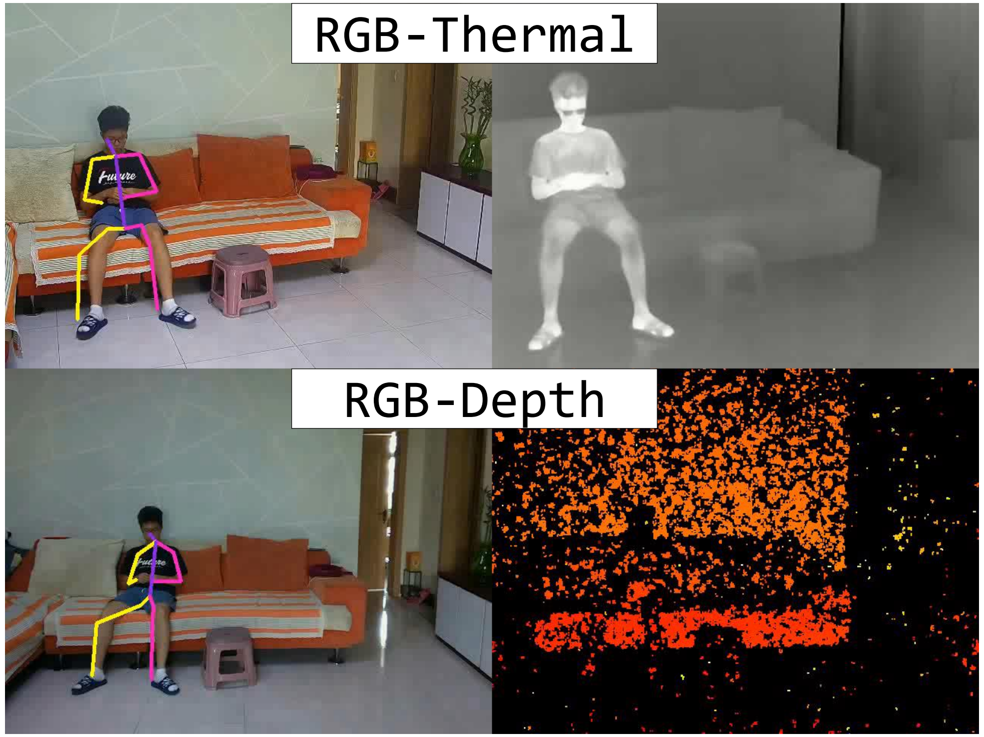

Thermal Image
Current Pose
GoalNet
Heatmap
Sample
TypeNet
Distribution
Sample
PoseNet Refinement

Results



Thermal Indoor Motion Dataset

RGB-Thermal and RGB-Depth videos of indoor human motion



Thermal Indoor Motion Dataset

2 actors, 3 rooms in multiple view angles



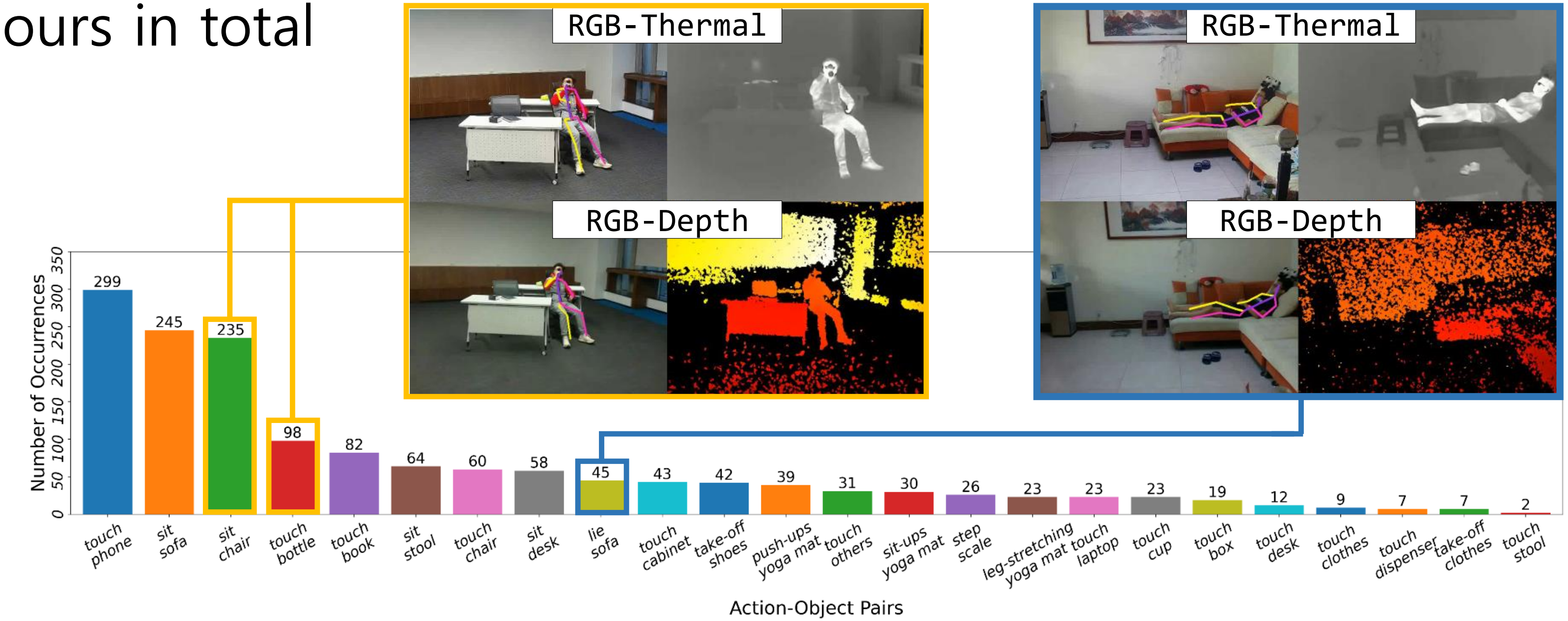



Thermal Indoor Motion Dataset

2 actors, 3 rooms in multiple view angles

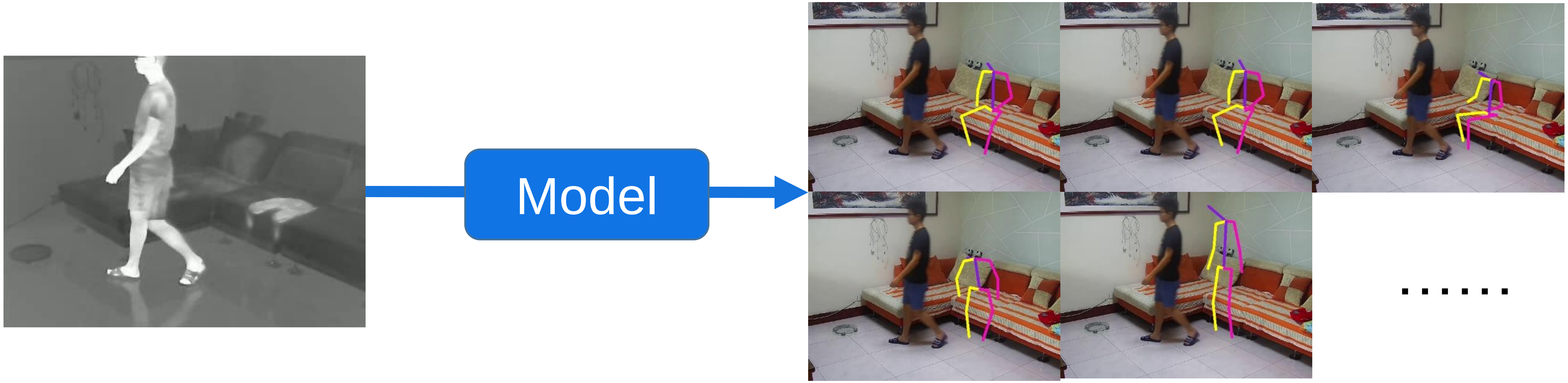
24 types of actions with annotated start and end time

783 clips, 10.4 hours in total

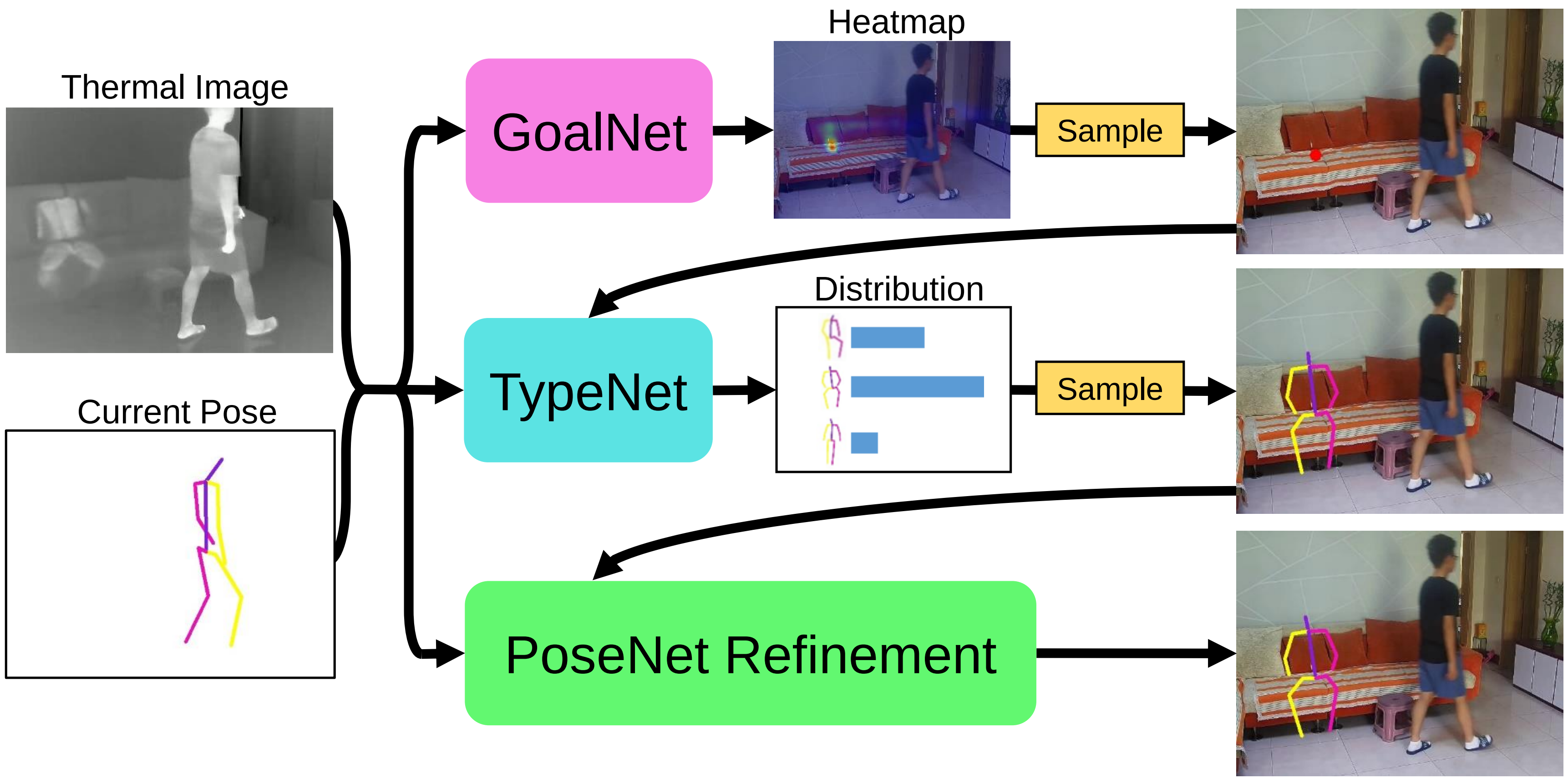


Past human pose estimation

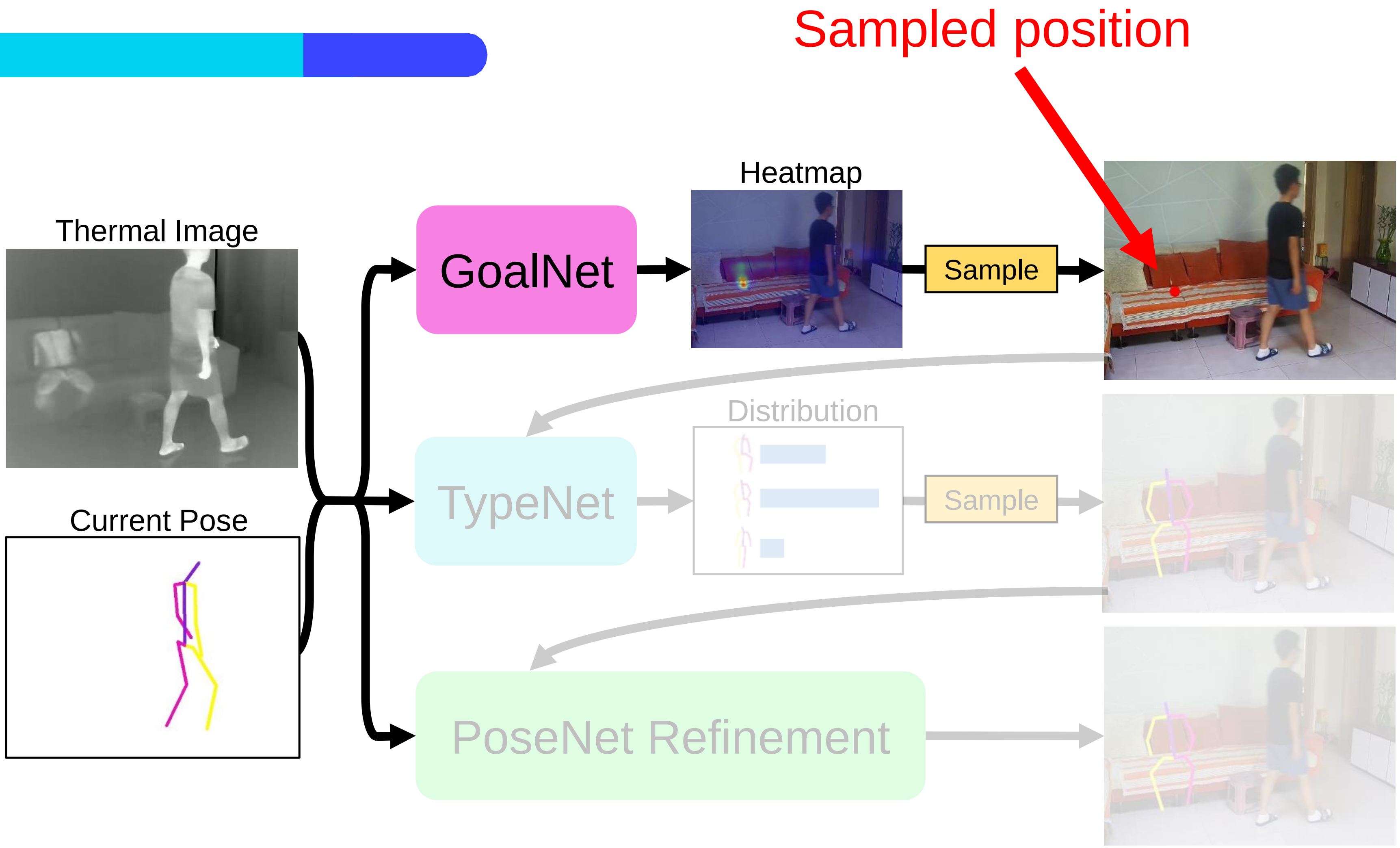
Given an indoor thermal image with a person in it, generates M (= 30) possible poses of the person N (= 3) seconds ago.



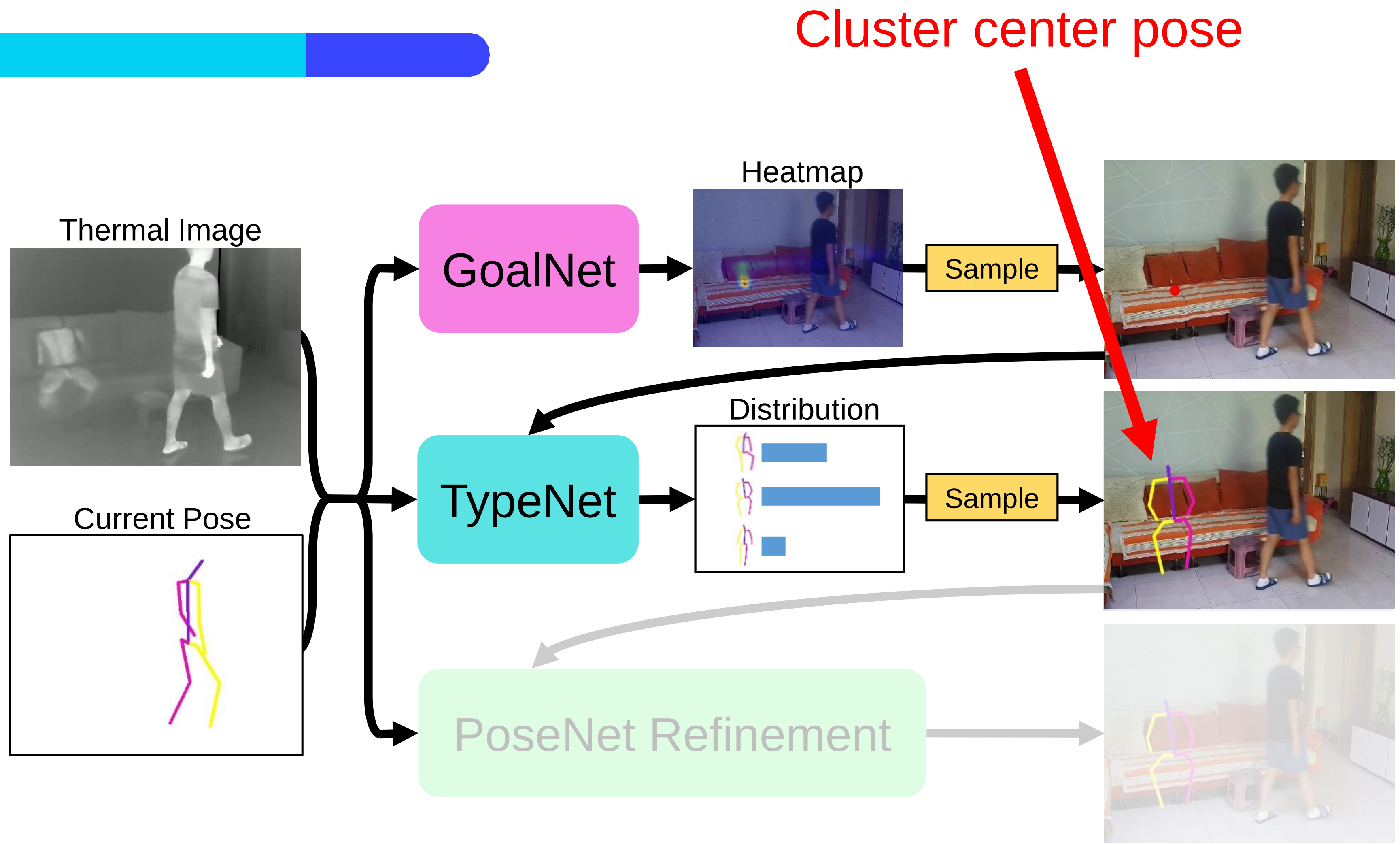
Method


Thermal Image
Current Pose
GoalNet
TypeNet
PoseNet Refinement
Heatmap
Distribution
Sample
Sample

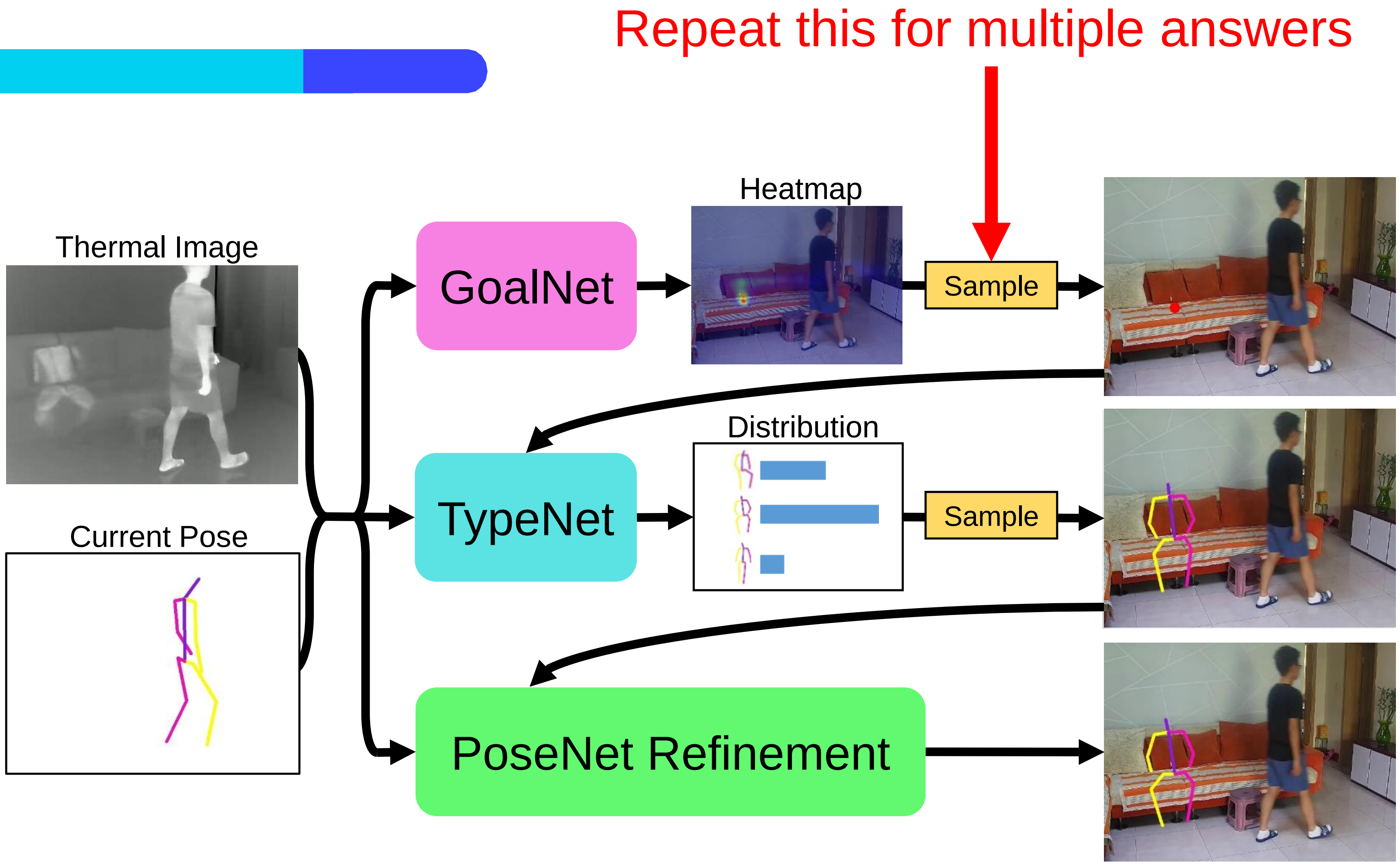
Method


Sampled position
Thermal Image
Current Pose
GoalNet
Heatmap
Sample
TypeNet
Distribution
Sample
PoseNet Refinement

Method


Cluster center pose
Thermal Image
Current Pose
GoalNet
Heatmap
Sample
TypeNet
Distribution
Sample
PoseNet Refinement

Method


Repeat this for multiple answers
Thermal Image
Current Pose
GoalNet
Heatmap
Sample
TypeNet
Distribution
Sample
PoseNet Refinement

Evaluation metrics

Mean Per Joint Position Error (MPJPE)

Evaluates the top-$k(=1,3,5)$ generated poses.
Measures their differences to the ground truth.

Negative Log-likelihood (NLL)

Likelihood of the ground truth.

Semantic Score

The ratio of generated poses that are compatible with the scene affordance.

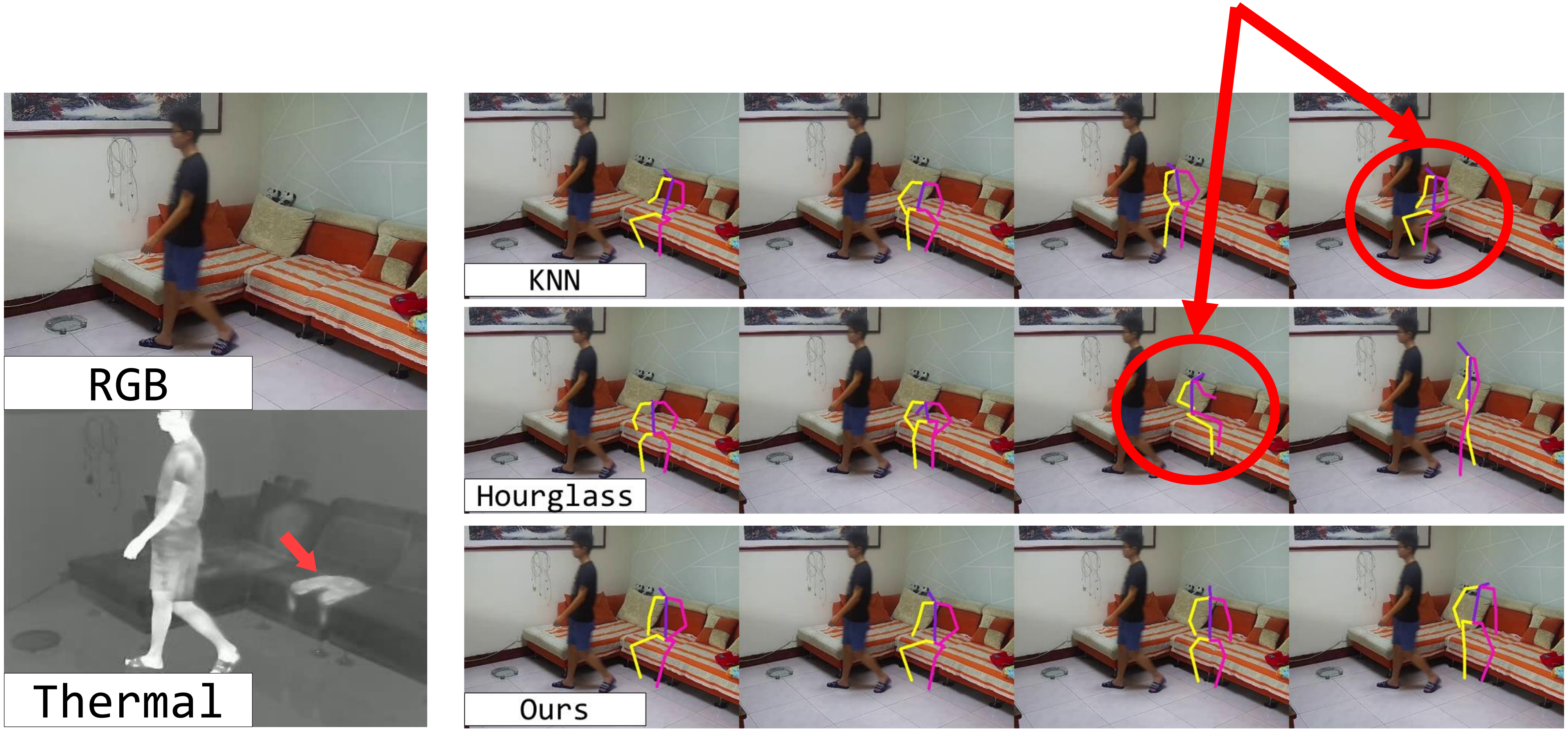
Results

Compare with KNN and one-stage Hourglass baselines

Method	MPJPE			NLL	Semantic Score(%)
	Top 1	Top 3	Top 5		
KNN	19.26	24.53	28.44	N/A	61.94
Hourglass	23.80	27.99	31.03	136.23	67.05
Ours	**18.33**	**22.25**	**25.25**	**103.75**	**82.11**

Results

Not compatible with the affordance



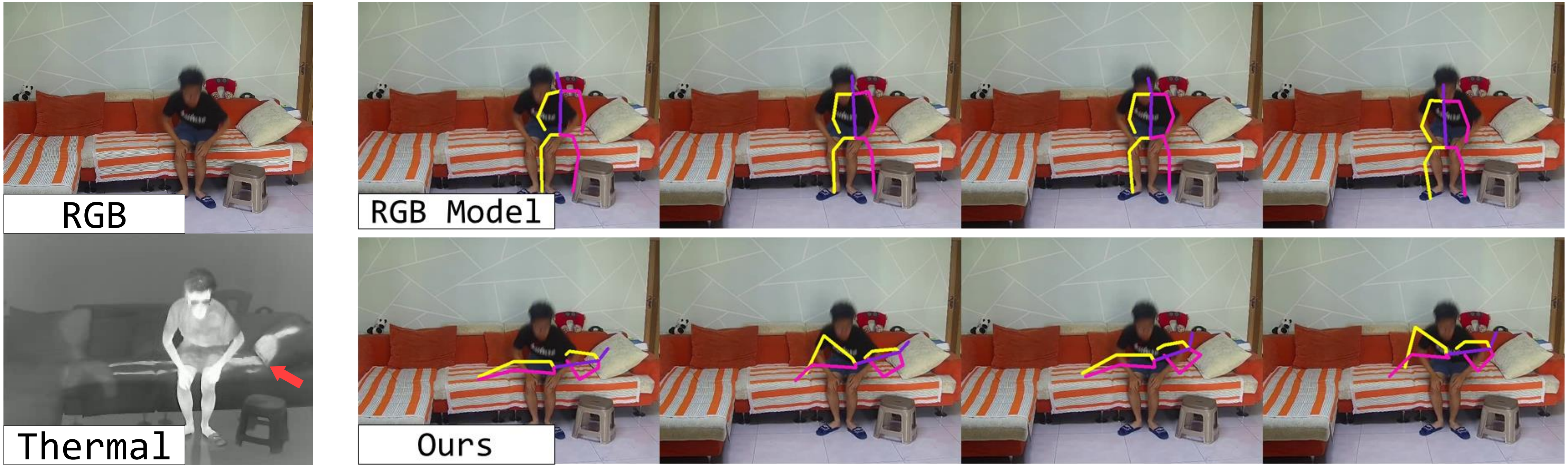
All reasonable

Results

Compare with different input modalities

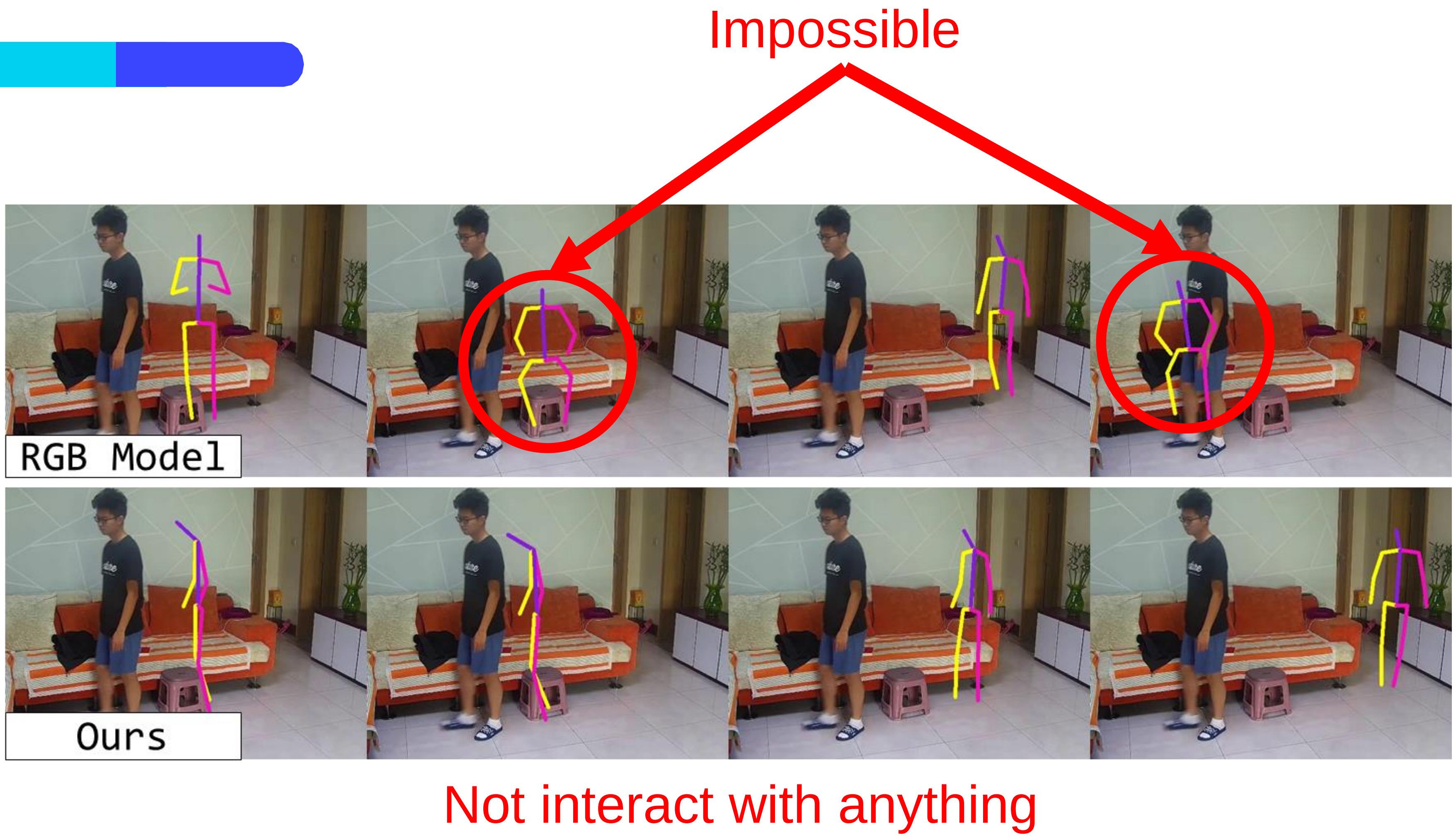
Input	MPJPE			NLL	Semantic Score(%)
	Top 1	Top 3	Top 5		
RGB	22.06	27.21	31.12	105.03	**87.56**
Thermal	**18.33**	**22.25**	**25.25**	**103.75**	82.11
T w/o pose	19.62	24.00	27.27	104.38	80.55

Results



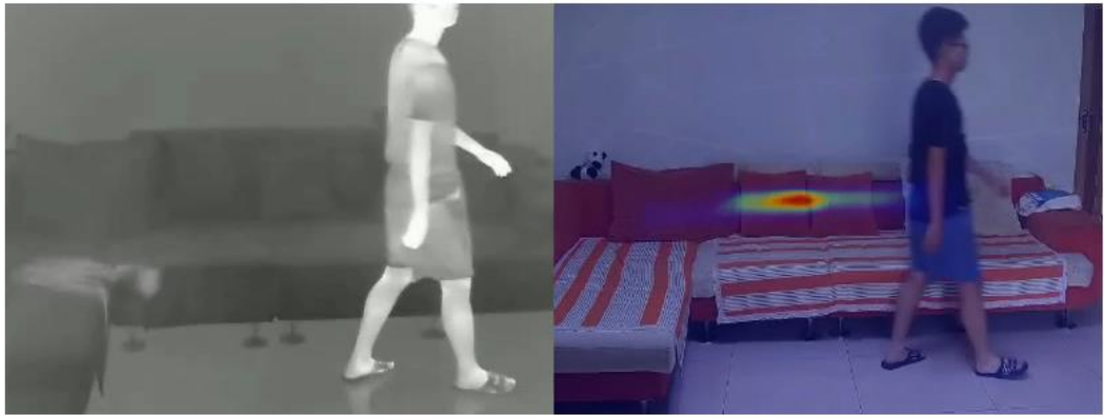
Results


RGB
Thermal


Impossible
RGB Model
Ours
Not interact with anything

Results

Correlation between thermal mark intensity and time



Results

Held-out data for generalization test

Results

Changed Factor	Modality	MPJPE Top 1	MPJPE Top 3	MPJPE Top 5	NLL	Semantic Score(%)
Arrangement	RGB	21.27	26.38	30.42	107.10	**93.69**
	Thermal	**20.41**	**25.10**	**28.36**	**105.37**	89.56
Background	RGB	25.07	30.02	33.47	111.67	**83.80**
	Thermal	**19.85**	**24.24**	**27.83**	**107.82**	81.49
Actor	RGB	**24.37**	29.20	32.77	114.87	**91.21**
	Thermal	24.60	**28.92**	**31.98**	**114.26**	81.33
Room	RGB	35.05	42.00	47.11	121.14	19.55
	Thermal	**23.05**	**27.59**	**31.16**	**112.84**	**36.88**

Conclusion

A novel task: past human pose estimation with thermal images

Thermal-IM dataset: RGB-Thermal-Depth videos about indoor human motion

A model tackling the task

- Outperforms the baselines
- Thermal imaging makes the problem easy
- Thermal model generalizes well across environment's appearance

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

What Happened 3 Seconds Ago? Inferring the Past with Thermal Imaging

THU-AM-060




Zitian Tang[1] Wenjie Ye[1] Wei-Chiu Ma[2] Hang Zhao[1,3]

[1]Tsinghua University, IIIS [2]MIT, CSAIL [3]Shanghai Qi Zhi Institute

MARS Lab