Towards Labeling-free Fine-grained Animal Pose Estimation

Dan Zeng, Yu Zhu, Shuiwang Li, Qijun Zhao, Qiaomu Shen, Bo Tang

Project website: https://github.com/yzrs/FreeNet.

Introduction

Our FreeNet learns fine-grained APE with free full annotation labels.

Dataset A   Dataset B   Dataset C → New problem setting in APE → Fine-grained Animal Pose
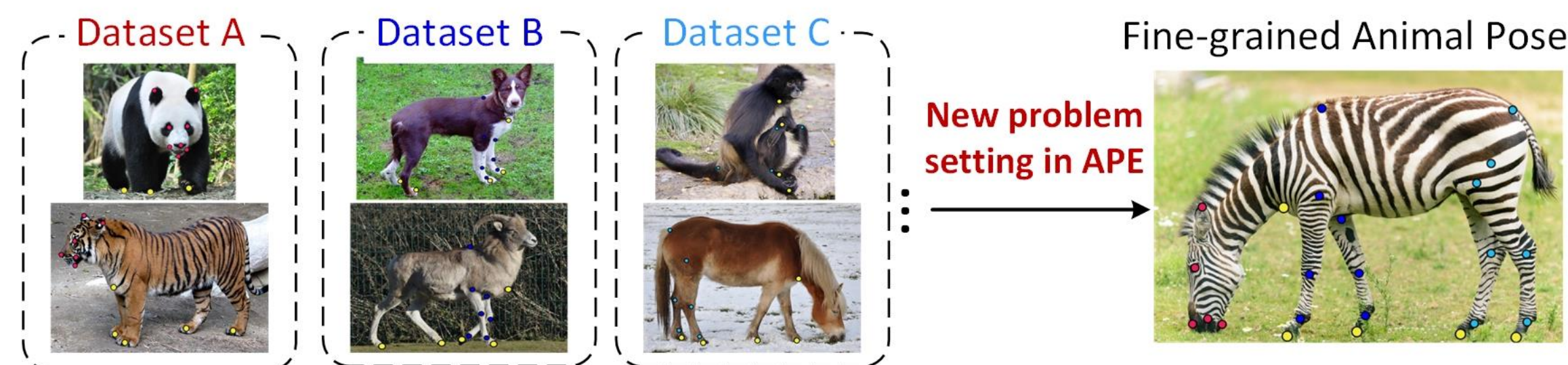
Challenges

➢ Lack standardized joint definitions across APE datasets

➢ Shared joints in different animal body parts exhibit different learning difficulties

➢ Unannotated joints aggravate the skewed joint distribution in the combined training data


AnimalPose, 20 joints   AP-10k, 17 joints
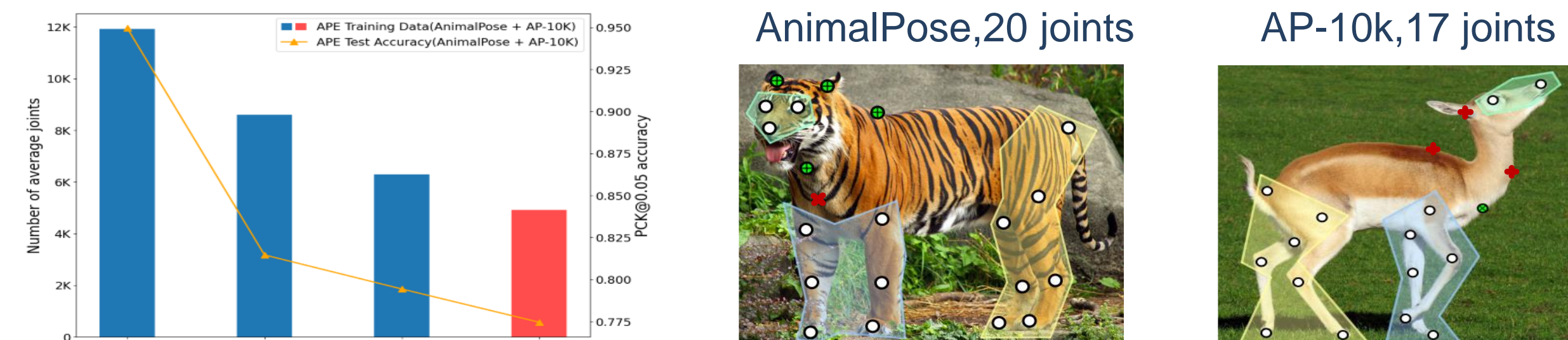
Motivation

➢ Body part-aware learning balances the learning of shared joints among different datasets

➢ The circuit feedback mechanism improves the base network's predictions on unannotated joints

Contributions

➢ Address the **non-standardized annotation** problem, a **new** and significant challenge in fine-grained APE

➢ Propose **FreeNet** method to effectively utilize annotation data, which can predict denser joints with **free** full annotation labels

➢ Extensive experiments on non-standard datasets demonstrate the superiority of our method for fine-grained APE

Method: FreeNet
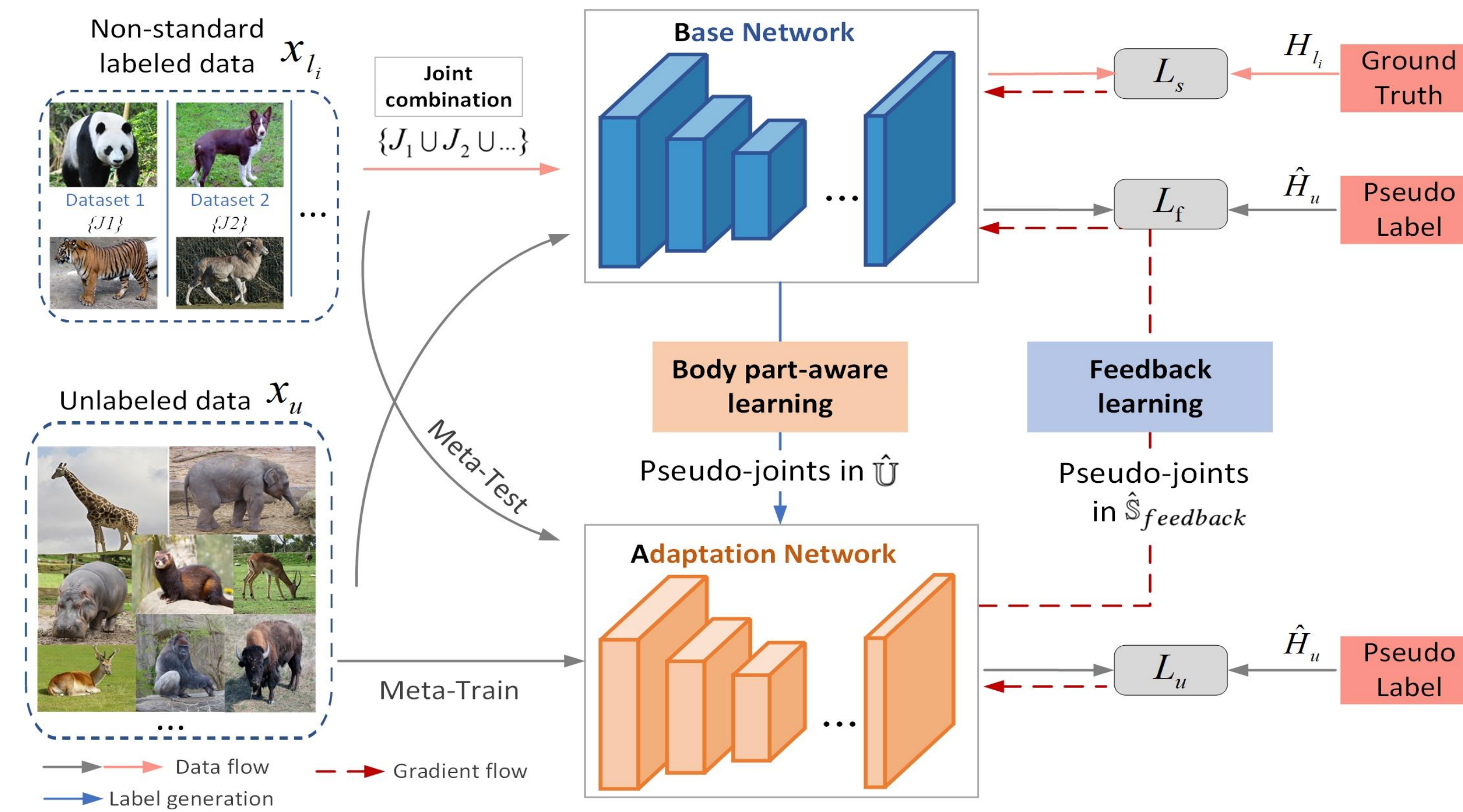
➢ Body part-aware learning
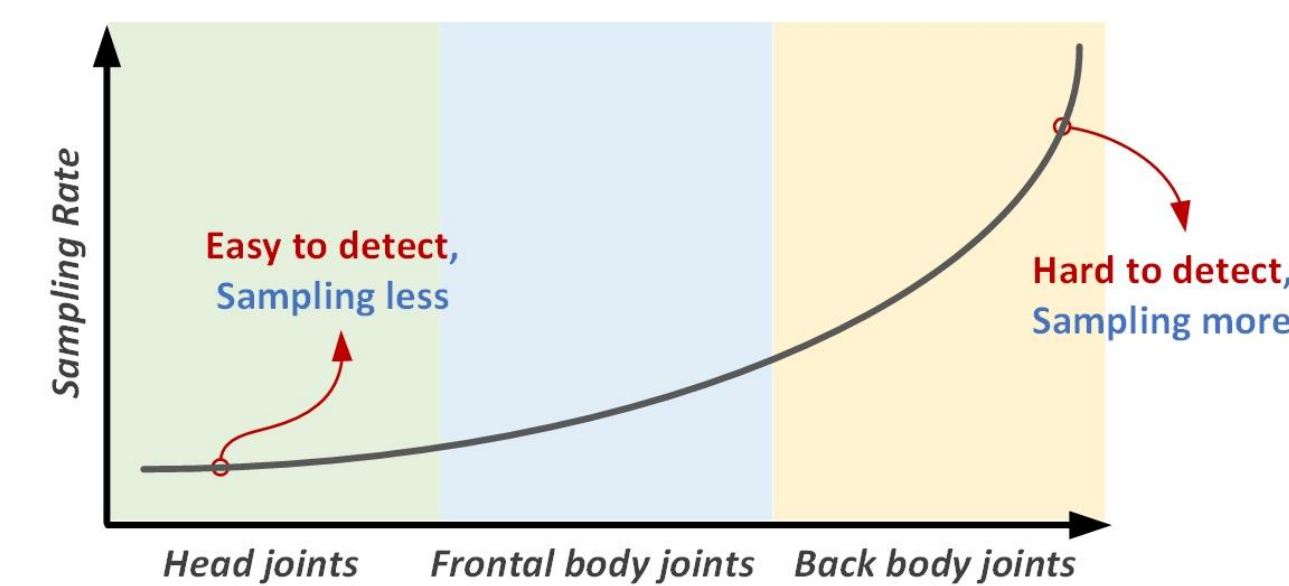
Enhances the adaptation network's tolerance to unannotated joints.

| γ_{head}	$\gamma_{frontal}$	γ_{back}	mAP↑	PCK@0.05↑	std PCK↓
100%	100%	100%	56.9	**71.4**	6.6
90%	75%	60%	56.2	70.63	7.8
50%	60%	65%	**57.5**	71.28	6

➢ Feedback learning

• Loss design

$$\mathcal{L}_f = f \cdot \frac{\sum_{k=1}^{N_J} \left\{ \hat{H}_u^k \in \hat{\mathbb{S}}_{feedback} \right\}_1 \mathcal{L}_f^k}{\sum_{k=1}^{N_J} \left\{ \hat{H}_u^k \in \hat{\mathbb{S}}_{feedback} \right\}_1},$$

where $\mathcal{L}_f^k = \left\| \hat{H}_u^k - (\mathcal{B}(I_u; \theta_{\mathcal{B}}))^k \right\|^2.$

Feedback coefficient

$$f = \eta_{\mathcal{A}} \cdot (\nabla_{\theta_{\mathcal{A}}^{(t+1)}} \mathrm{MSE}(H_l, \mathcal{A}(x_l; \theta_{\mathcal{A}}^{(t+1)}))^\top .$$

$$\nabla_{\theta_{\mathcal{A}}} \mathrm{MSE}(\hat{H}_u, \mathcal{A}(x_u; \theta_{\mathcal{A}}^{(t)}))),$$

• Pseudo-joints selection criteria

Use the confidence score rankings from both networks to ensure unannotated joints are further improved.
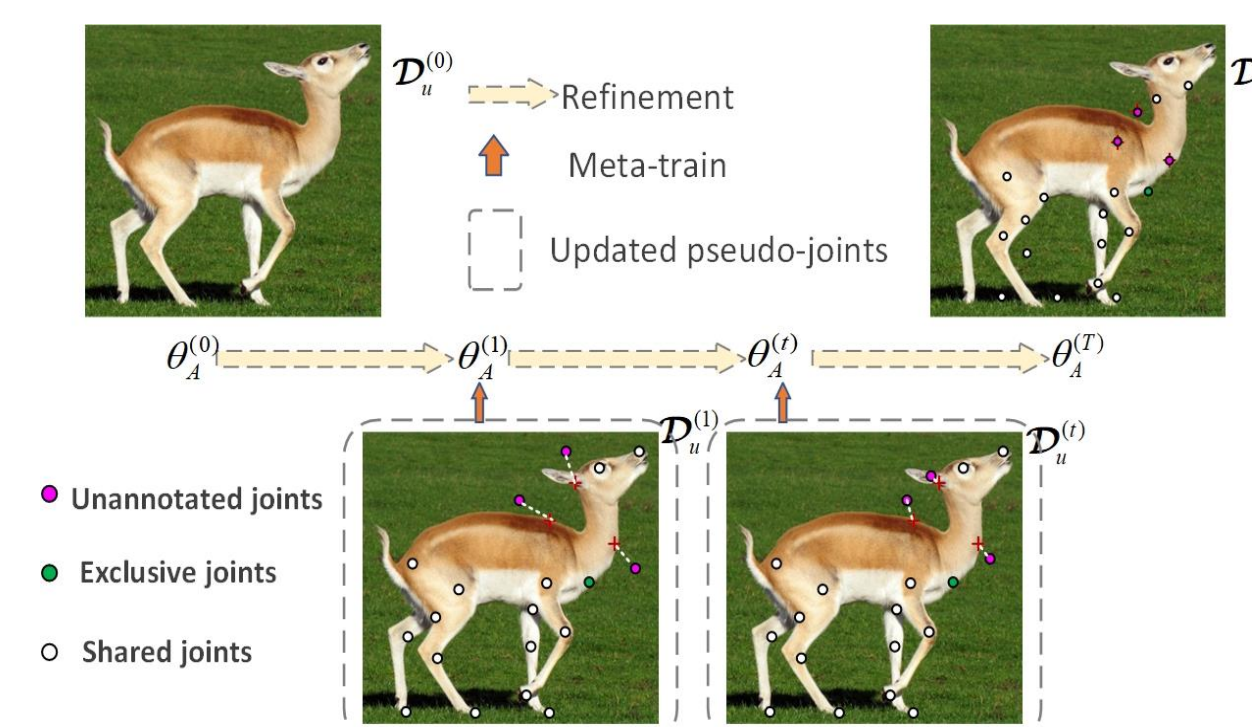
Experiments

➢ **Comparison with SOTA methods on nonstandard datasets**

Settings	Full Annot.	Methods	mAP↑	PCK@0.05↑	
5 ips	✓	ScarceNet	53.3	65.2	Scarce datasets with full annotations
25 ips	✓	ScarceNet	68.1	78.2	
3 synthetic datasets from 25 ips	✗	ScarceNet	55.04	66.26	Scarce datasets with zero full annotations
	✗	UDA	50.8	64.06	
	✗	FixMatch	43.8	57.56	
	✗	MPL	50.7	63.51	
	✗	Ours	**57.9**	**68.31**	

➢ **Effect of FreeNet design**

Model 1   Model 2   Ours   Ground Truth
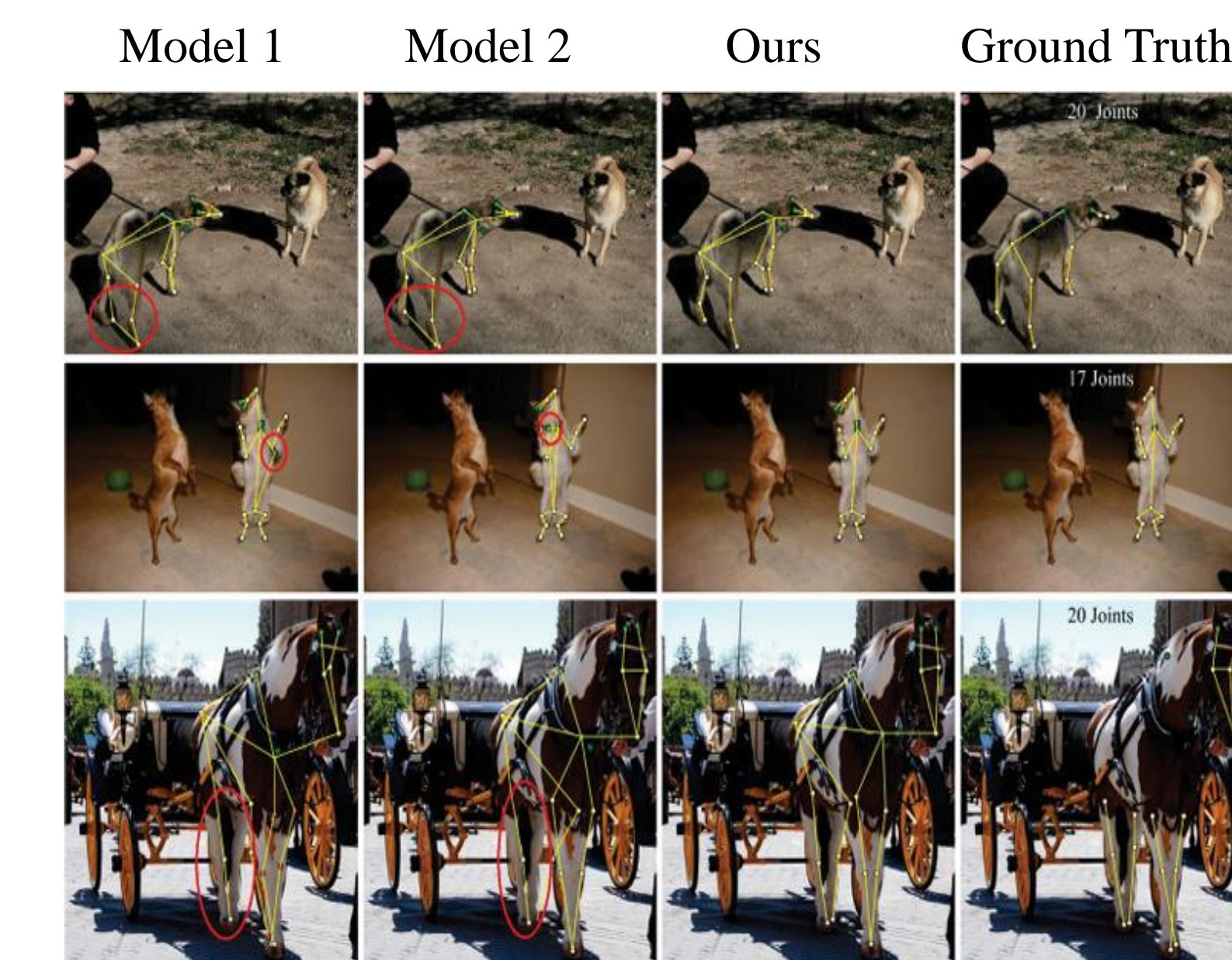
Results on combined real datasets, i.e., 10% AP-10k, and AnimalPose

Models	$\mathcal{L}_s$	$\mathcal{L}_u$	$\mathcal{L}_f$	mAP↑	PCK@0.05↑
1	✓	✗	✗	52.2	67.6
2	✓	✓	✗	56.2	70.63
Ours	✓	✓	✓	**57.26**	**71.36**

➢ **FreeNet can generate denser joints in real-world applications**

17 Joints   19 Joints   20 Joints   21 Joints   24 Joints   Ground Truth