

# Pose for Everything: Towards Category-Agnostic Pose Estimation

## – Supplementary Material –

Lumin Xu, Sheng Jin, Wang Zeng, Wentao Liu, Chen Qian  
Wanli Ouyang, Ping Luo, and Xiaogang Wang

### 1 Potential Application

We envision that our proposed Category-Agnostic Pose Estimation (CAPE) may make many positive broader impacts and have some potential applications, *e.g.* pose data (pseudo-)labeling, keypoint-based object tracking, contour-based instance segmentation, and graph matching.



Fig. 1: Example of pose data labeling. Each row shows one support image with manually annotated keypoint definition, and four query images with pseudo-labeled poses predicted by our CAPE model. Our CAPE model generates pose predictions in accordance with the annotations of the support images and fits in all kinds of keypoint definitions.

We apply our CAPE model on the pose data labeling task and show that the CAPE model may work as a pseudo-annotator. Semi-automatic labeling approach has been widely applied for data labeling [1,5], where images are first

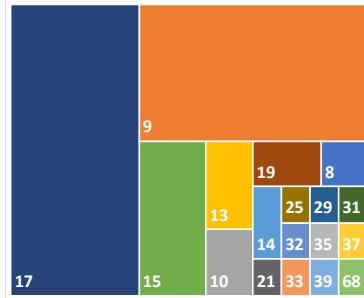


Fig. 2: Number of keypoints are diverse across different categories in the MP-100 dataset, ranging from 8 to 68. In the diagram, the keypoint number is shown in each block and the block size represents the proportion of categories.

labeled by a pre-trained pose model and then corrected by human annotators. In this way, the cost of the laborious labeling process is largely reduced. However, previously each pose model only focus on one single category with the specific pose definition (or keypoint configuration). For facial keypoint localization, popular definitions include 21-kpt definition (AFLW [2]), 68-kpt definition (300W [3]) and 98-kpt definition (WFLW [4]). And previous methods have to train separate models to handle different pose definitions for each object category. In Fig. 1, we show our pre-trained CAPE model is promising in handling various pose definitions. In each row, we show one support image with manually annotated keypoint definition, and four query images with pseudo-labeled poses predicted by our CAPE model. We observe that our CAPE model generates pose predictions in accordance with the pose definitions of the support images without re-training or fine-tuning, which makes the pseudo-labeling process easy and flexible.

## 2 Multi-category Pose (MP-100) Dataset

### 2.1 Number of keypoints

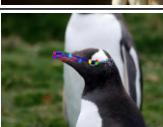
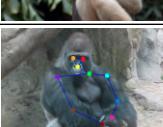
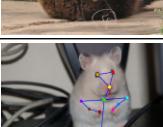
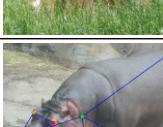
In Fig. 2, we analyze the number of keypoints for each category in the MP-100 dataset. In the diagram, the block size represents the proportion of categories with the keypoint number shown in the block. We notice that the number of keypoints are diverse across different categories, ranging from 8 to 68.

### 2.2 Example of Categories

In Table 1, we show the name of each category, the number of instances of that category, and an image example with keypoint annotations in the MP-100 dataset.

Table 1: Examples of 100 categories in the MP-100 dataset. ‘#Inst’ indicates the number of instances of the category. ‘kpt’ means keypoint.

Category (#Inst)	Example	Category (#Inst)	Example	Category (#Inst)	Example
17-kpt human body <b>(218)</b>		19-kpt human face <b>(217)</b>		21-kpt human hand <b>(217)</b>	
68-kpt human face <b>(217)</b>		alpaca face <b>(187)</b>		antelope body <b>(217)</b>	
arctic wolf face <b>(150)</b>		beaver body <b>(197)</b>		bed <b>(217)</b>	
bighorn sheep face <b>(217)</b>		bison body <b>(217)</b>		blackbuck face <b>(172)</b>	
bobcat body <b>(217)</b>		bonobo face <b>(200)</b>		bus <b>(216)</b>	
california sea lion face <b>(163)</b>		camel face <b>(140)</b>		cape buf-falo face <b>(164)</b>	
capybara face <b>(165)</b>		cat body <b>(217)</b>		chair <b>(218)</b>	
cheetah body <b>(212)</b>		chipmunk face <b>(207)</b>		common warthog face <b>(154)</b>	

cow body <b>(217)</b>		dassie face <b>(181)</b>		deer body <b>(216)</b>	
dog body <b>(217)</b>		elephant body <b>(216)</b>		fallow deer face <b>(157)</b>	
fennec fox face <b>(182)</b>		ferret face <b>(217)</b>		fox body <b>(217)</b>	
gentoo penguin face <b>(212)</b>		gerbil face <b>(164)</b>		german shepherd dog face <b>(149)</b>	
gibbons face <b>(188)</b>		giraffe body <b>(217)</b>		golden retriever face <b>(140)</b>	
gorilla body <b>(231)</b>		grebe body <b>(217)</b>		grey seal face <b>(180)</b>	
grizzly bear face <b>(159)</b>		guanaco face <b>(145)</b>		gull body <b>(217)</b>	
hamster body <b>(231)</b>		hippo body <b>(217)</b>		horse body <b>(217)</b>	
kingfisher body <b>(216)</b>		klipspringer face <b>(142)</b>		leopard body <b>(207)</b>	



skirt <b>(217)</b>		skunk body <b>(217)</b>		sling <b>(217)</b>	
sling dress <b>(218)</b>		sofa <b>(217)</b>		sparrow body <b>(217)</b>	
spider monkey body <b>(217)</b>		sport utility vehicle <b>(204)</b>		squirrel body <b>(203)</b>	
swivel chair <b>(217)</b>		table <b>(217)</b>		tern body <b>(217)</b>	
trousers <b>(218)</b>		vest <b>(218)</b>		vest dress <b>(218)</b>	
vinegar fly body <b>(232)</b>		warbler body <b>(217)</b>		weasel body <b>(232)</b>	
wolf body <b>(216)</b>		woodpecker body <b>(217)</b>		wren body <b>(217)</b>	
zebra body <b>(216)</b>					

### 2.3 Splits

We provide five dataset splits (Split1, Split2, ..., Split5) for the MP-100 dataset. In each split, we have 70 categories for train, 10 for val, and 20 for test. We make the test sets of these five dataset splits non-overlapping. And the object categories in train/val/test set are also mutually exclusive. The details of each dataset split are shown in Table 2, Table 3, Table 4, Table 5, and Table 6.

Table 2: The categories contained in train/val/test set of MP-100 dataset Split1.

Set	Category		
Train	17-kpt human body	gerbil face	quokka face
	19-kpt human face	german shepherd dog face	raccoon body
	21-kpt human hand	gibbons face	rat body
	68-kpt human face	giraffe body	rhino body
	antelope body	grebe body	short sleeved outwear
	arctic wolf face	grey seal face	short sleeved shirt
	blackbuck face	grizzly bear face	shorts
	bobcat body	gull body	skirt
	bonobo face	hippo body	sling
	bus	kingfisher body	sling dress
	california sea lion face	leopard body	sofa
	camel face	locust body	sparrow body
	cape buffalo face	long sleeved dress	spider monkey body
	capybara face	long sleeved shirt	sport utility vehicle
	cat body	olive baboon face	table
	chair	onager face	trousers
	cheetah body	otter body	vest
	chipmunk face	pademelon face	vest dress
	cow body	panda body	warbler body
	deer body	panther body	wolf body
	elephant body	passenger car	wren body
	fallow deer face	pig body	zebra body
	fennec fox face	polar bear body	
	ferret face	proboscis monkey face	
Val	beaver body	guanaco face	vinegar fly body
	gentoo penguin face	hamster body	weasel body
	golden retriever face	macaque body	
	gorilla body	przewalski horse face	
Test	alpaca face	fox body	short sleeved dress
	bed	horse body	skunk body
	bighorn sheep face	klipspringer face	squirrel body
	bison body	lion body	swivel chair
	common warthog face	long sleeved outwear	tern body
	dassie face	rabbit body	woodpecker body
	dog body	sheep body	

Table 3: The categories contained in train/val/test set of MP-100 dataset Split2.

Set	Category		
Train	17-kpt human body	ferret face	passenger car
	19-kpt human face	fox body	pig body
	21-kpt human hand	gentoo penguin face	przewalski horse face
	68-kpt human face	german shepherd dog face	rabbit body
	alpaca face	gibbons face	raccoon body
	antelope body	gorilla body	rat body
	arctic wolf face	grebe body	rhino body
	bed	grey seal face	sheep body
	bighorn sheep face	guanaco face	short sleeved dress
	bison body	hamster body	short sleeved outwear
	blackbuck face	hippo body	shorts
	bobcat body	horse body	skunk body
	bus	kingfisher body	sofa
	california sea lion face	klipspringer face	sparrow body
	cat body	leopard body	sport utility vehicle
	cheetah body	lion body	squirrel body
	chipmunk face	locust body	swivel chair
	common warthog face	long sleeved dress	vinegar fly body
	cow body	macaque body	wolf body
	dassie face	olive baboon face	woodpecker body
	dog body	onager face	wren body
	elephant body	pademelon face	zebra body
	fallow deer face	panda body	
	fennec fox face	panther body	
Val	bonobo face	otter body	vest dress
	deer body	quokka face	weasel body
	gerbil face	skirt	
	long sleeved outwear	tern body	
Test	beaver body	grizzly bear face	sling dress
	camel face	gull body	spider monkey body
	cape buffalo face	long sleeved shirt	table
	capybara face	polar bear body	trousers
	chair	proboscis monkey face	vest
	giraffe body	short sleeved shirt	warbler body
	golden retriever face	sling	

Table 4: The categories contained in train/val/test set of MP-100 dataset Split3.

Set	Category		
Train	19-kpt human face	gibbons face	quokka face
	21-kpt human hand	giraffe body	rabbit body
	68-kpt human face	golden retriever face	rat body
	alpaca face	gorilla body	sheep body
	antelope body	grebe body	short sleeved dress
	bighorn sheep face	grizzly bear face	short sleeved outwear
	blackbuck face	gull body	short sleeved shirt
	bonobo face	hamster body	shorts
	bus	horse body	skirt
	california sea lion face	kingfisher body	sling
	camel face	klipspringer face	sling dress
	capybara face	lion body	sparrow body
	cat body	long sleeved dress	spider monkey body
	chair	long sleeved outwear	sport utility vehicle
	chipmunk face	long sleeved shirt	table
	common warthog face	macaque body	tern body
	dassie face	onager face	trousers
	deer body	otter body	vest
	dog body	pademelon face	warbler body
	fallow deer face	panda body	weasel body
	fennec fox face	passenger car	woodpecker body
	fox body	pig body	zebra body
	gerbil face	polar bear body	
	german shepherd dog face	proboscis monkey face	
Val	arctic wolf face	bobcat body	squirrel body
	beaver body	cape buffalo face	swivel chair
	bed	elephant body	
	bison body	skunk body	
Test	17-kpt human body	hippo body	rhino body
	cheetah body	leopard body	sofa
	cow body	locust body	vest dress
	ferret face	olive baboon face	vinegar fly body
	gentoo penguin face	panther body	wolf body
	grey seal face	przewalski horse face	wren body
	guanaco face	raccoon body	

Table 5: The categories contained in train/val/test set of MP-100 dataset Split4.

Set	Category		
Train	17-kpt human body	gibbons face	raccoon body
	alpaca face	giraffe body	rhino body
	antelope body	golden retriever face	sheep body
	arctic wolf face	grebe body	short sleeved dress
	beaver body	grey seal face	shorts
	bed	grizzly bear face	skunk body
	bighorn sheep face	guanaco face	sling
	bison body	gull body	sling dress
	bobcat body	hippo body	sofa
	bonobo face	horse body	sparrow body
	camel face	klipspringer face	spider monkey body
	cape buffalo face	leopard body	squirrel body
	capybara face	locust body	swivel chair
	chair	long sleeved dress	table
	cheetah body	long sleeved outwear	tern body
	common warthog face	long sleeved shirt	trousers
	cow body	olive baboon face	vest
	dassie face	panda body	vest dress
	dog body	panther body	vinegar fly body
	ferret face	passenger car	warbler body
	fox body	polar bear body	woodpecker body
	gentoo penguin face	proboscis monkey face	wren body
	gerbil face	quokka face	
	german shepherd dog face	rabbit body	
Val	19-kpt human face	onager face	sport utility vehicle
	elephant body	przewalski horse face	wolf body
	hamster body	short sleeved shirt	
	lion body	skirt	
Test	21-kpt human hand	deer body	pademelon face
	68-kpt human face	fallow deer face	pig body
	blackbuck face	fennec fox face	rat body
	bus	gorilla body	short sleeved outwear
	california sea lion face	kingfisher body	weasel body
	cat body	macaque body	zebra body
	chipmunk face	otter body	

Table 6: The categories contained in train/val/test set of MP-100 dataset Split5.

Set	Category		
Train	17-kpt human body	ferret face	raccoon body
	21-kpt human hand	fox body	rat body
	68-kpt human face	gentoo penguin face	sheep body
	alpaca face	giraffe body	short sleeved dress
	beaver body	golden retriever face	short sleeved outwear
	bed	gorilla body	short sleeved shirt
	bighorn sheep face	grey seal face	skunk body
	bison body	grizzly bear face	sling dress
	blackbuck face	guanaco face	sofa
	bus	gull body	spider monkey body
	california sea lion face	hippo body	squirrel body
	camel face	klipspringer face	swivel chair
	cape buffalo face	lion body	table
	capybara face	locust body	trousers
	cat body	long sleeved outwear	vest
	chair	long sleeved shirt	vest dress
	cheetah body	macaque body	vinegar fly body
	chipmunk face	olive baboon face	warbler body
	cow body	pademelon face	weasel body
	dassie face	pig body	wolf body
	deer body	polar bear body	wren body
	dog body	proboscis monkey face	zebra body
	fallow deer face	przewalski horse face	
	fennec fox face	rabbit body	
Val	common warthog face	otter body	tern body
	horse body	panther body	woodpecker body
	kingfisher body	rhino body	
	leopard body	sling	
Test	19-kpt human face	german shepherd dog face	passenger car
	antelope body	gibbons face	quokka face
	arctic wolf face	grebe body	shorts
	bobcat body	hamster body	skirt
	bonobo face	long sleeved dress	sparrow body
	elephant body	onager face	sport utility vehicle
	gerbil face	panda body	

## References

1. Jin, S., Xu, L., Xu, J., Wang, C., Liu, W., Qian, C., Ouyang, W., Luo, P.: Whole-body human pose estimation in the wild. In: Eur. Conf. Comput. Vis. (2020) [1](#)
2. Kostinger, M., Wohlhart, P., Roth, P., Bischof, H.: Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In: Int. Conf. Comput. Vis. Worksh. (2011) [2](#)
3. Sagonas, C., Antonakos, E., Tzimiropoulos, G., Zafeiriou, S., Pantic, M.: 300 faces in-the-wild challenge: database and results. Image and Vision Computing (2016) [2](#)
4. Wu, W., Qian, C., Yang, S., Wang, Q., Cai, Y., Zhou, Q.: Look at boundary: A boundary-aware face alignment algorithm. In: IEEE Conf. Comput. Vis. Pattern Recog. (2018) [2](#)
5. Zafeiriou, S., Trigeorgis, G., Chrysos, G., Deng, J., Shen, J.: The menpo facial landmark localisation challenge: A step towards the solution. In: IEEE Conf. Comput. Vis. Pattern Recog. Worksh. (2017) [1](#)