

Unsupervised Pre-training for Person Re-identification

In this supplementary material, we will provide results of pre-training on the existing dataset and share more detailed results for our “small-scale” and “few-shot” experiments.

1. Pre-training on existing dataset

In this section, we conduct unsupervised pre-training with existing person Re-ID datasets. Specifically, we use all the training images in **CUHK03**, **Market1501**, **DukeMTMC** and **MSMT17** and refer it as **CMDM** to train MoCoV2, Table 1 shows the details of the training images. We use $[0.3256, 0.3058, 0.3162]$, $[0.2603, 0.2489, 0.2391]$ as mean and std, which are calculated from all these training images. As there are only 69447 images, the default setting $queue_size = 65536$ in MoCoV2 is not suitable anymore for generating negative pairs, because this may lead to too many false negative samples. We try different $queue_size \in \{2048, 4096, 8192, 16384\}$ and find $queue_size = 4096$ is a good choice. We use the same data augmentation and temperature described in the main text.

Item	CUHK03	Market1501	DukeMTMC	MSMT17	total
#ids	767	751	702	1041	3261
#images	7368	12936	16522	32621	69447

Table 1: The details of training images we used for pre-training on existing dataset.

Table 2 shows the performance comparison for using model pre-trained on “CMDM”. It can be seen that, “CMDM unsup.” model is less competitive than ImageNet supervised model, and is much weaker than our model which is pre-trained on LUPerson. This further proves the effectiveness and necessity of our LUPerson dataset.

2. More results for small-scale and few-shot

As introduced in the main text, we introduce two new settings: small-scale and few-shot, where the target dataset has limited annotated samples. Supplemented to Table 6 in the main text, we provide all the details and the left results in Table 3.

Table 3 shows that our model can always perform better than ImageNet pre-trained models, and the advantage is more obvious when the training data is less, which is very

pre-train	CUHK03	Market1501	DukeMTMC	MSMT17
IN sup.	70.5/71.2	87.5/95.1	79.4/89.0	63.7/85.1
CMDM unsup.	51.7/52.7	81.4/92.8	74.4/84.8	48.7/72.6
LUP unsup.	74.7/75.4	91.0/96.4	82.1/91.0	65.7/85.5

Table 2: Performance comparison for using different pre-trained models on MGN baseline. “CMDM unsup.” and “LUP unsup.” refer to unsupervised pre-trained model on CMDM and LUPerson, “IN sup.” refers to supervised pre-trained model on ImageNet. The first number is mAP and the second is $cmcl$.

meaningful for person Re-ID in practical application.

scale	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
#id	75	150	225	300	375	450	525	600	675	751
#images	1,170	2,643	3,962	5,226	6,408	7,814	9,120	11,417	11,727	12,936
IN sup.	53.1/76.9	67.7/86.8	75.2/90.8	79.1/92.5	81.5/93.5	81.5/93.5	84.8/94.5	85.9/95.2	86.9/95.2	87.5/95.1
IN unsup.	58.4/81.7	70.2/89.1	76.6/91.9	80.0/93.0	82.0/94.1	83.7/94.3	85.4/94.5	86.4/95.0	87.4/95.5	88.2/95.3
LUP unsup.	64.6/85.5	76.9/92.1	81.9/93.7	84.1/94.4	85.8/94.9	87.8/95.8	88.8/95.9	89.8/96.2	90.5/96.4	91.0/96.4

(a) Market1501 small-scale

scale	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
#id	751	751	751	751	751	751	751	751	751	751
#images	1,293	2,587	3,880	5,174	6,468	7,758	9,055	10,348	11,642	12,936
IN sup.	21.1/41.8	53.2/75.1	68.1/87.6	75.4/90.4	80.2/92.8	83.0/93.6	84.2/94.0	86.3/94.7	86.7/94.6	87.5/95.1
IN unsup.	18.6/36.1	56.5/77.5	69.3/87.8	78.8/88.3	78.3/90.9	81.7/93.3	84.4/94.1	86.4/95.0	87.1/95.2	88.2/95.3
LUP unsup.	26.4/47.5	63.5/83.0	78.3/92.1	80.3/92.7	84.2/93.9	86.7/94.7	88.4/95.5	89.8/96.0	90.4/96.3	91.0/96.4

(b) Market1501 few-shot

scale	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
#id	70	140	210	280	351	421	491	561	631	702
#images	1,670	3,192	5,530	6,924	8,723	10,197	11,939	13,500	15,111	16,522
IN sup.	45.1/65.3	58.3/75.4	64.7/80.2	68.5/83.0	71.8/84.6	74.1/85.6	75.5/86.8	76.8/87.3	78.0/88.3	79.4/89.0
IN unsup.	48.1/66.9	60.6/76.6	65.8/80.2	69.5/82.9	72.5/84.4	75.0/86.2	76.3/86.9	77.4/87.3	78.5/88.7	79.5/89.1
LUP unsup.	53.5/72.0	65.0/78.9	69.4/81.9	72.8/84.7	75.6/86.7	77.6/87.1	78.9/88.2	80.2/89.2	81.1/90.0	82.1/91.0

(c) DukeMTMC small-scale

scale	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
#id	702	702	702	702	702	702	702	702	702	702
#images	1,679	3,321	4,938	6,599	8,278	9,923	11,564	13,201	14,860	16,522
IN sup.	31.5/47.1	56.2/72.1	65.4/79.8	71.0/83.9	73.9/85.7	75.8/86.6	77.2/87.8	78.3/88.6	79.1/88.8	79.4/89.0
IN unsup.	32.4/48.0	56.4/72.2	65.3/80.2	70.2/83.4	73.7/85.1	75.8/86.7	77.7/88.2	78.7/88.7	79.4/89.0	79.5/89.1
LUP unsup.	35.8/50.2	61.0/74.9	72.3/83.8	75.2/86.8	77.7/87.4	79.4/88.4	80.8/89.2	81.7/90.3	82.0/90.6	82.1/91.0

(d) DukeMTMC few-shot

scale	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
#id	104	208	312	416	520	624	728	832	936	1,041
#images	3,659	6,471	9,787	13,006	15,917	19,672	22,680	26,335	29,529	32,621
IN sup.	23.2/50.2	34.6/64.0	41.9/70.8	46.7/74.5	50.3/76.9	53.9/79.4	56.9/81.2	59.6/82.4	61.9/84.2	63.7/85.1
IN unsup.	22.6/48.8	32.7/60.9	40.4/68.7	45.1/72.2	49.0/75.0	52.7/78.0	55.7/79.9	58.6/82.0	60.9/83.0	62.7/84.3
LUP unsup.	25.5/51.1	36.0/62.6	44.6/71.4	49.2/74.9	53.0/77.7	56.4/79.7	59.5/81.8	61.9/83.6	63.7/85.0	65.7/85.5

(e) MSMT17 small-scale

scale	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
#id	1,041	1,041	1,041	1,041	1,041	1,041	1,041	1,041	1,041	1,041
#images	3,262	6,524	9,786	13,048	16,310	19,572	22,834	26,096	29,358	32,621
IN sup.	14.7/34.1	35.6/61.4	44.5/71.1	52.0/76.9	56.2/79.5	58.8/81.7	60.9/82.8	62.5/84.2	63.4/84.5	63.7/85.1
IN unsup.	13.2/29.2	33.5/58.6	41.4/67.1	47.7/72.7	53.3/77.6	56.5/79.6	59.1/81.5	60.9/82.3	62.4/83.8	62.7/84.3
LUP unsup.	17.0/36.0	37.4/61.4	49.0/73.6	53.9/78.5	57.4/80.5	60.0/82.1	62.9/83.5	64.2/84.5	65.0/85.1	65.7/85.5

(f) MSMT17 few-shot

Table 3: Performance for small-scale and few-shot setting with MGN method for Market1501, DukeMTMC and MSMT17. “IN sup.” and “IN unsup.” refer to supervised and unsupervised pre-trained model on ImageNet, “LUP unsup.” refers to unsupervised pre-trained model on LUPerson. The first number is *mAP* and second is *cmcl*.