Functional Knowledge Transfer with Self-supervised Representation Learning

Prakash Chandra Chhipa[1,*], Muskaan Chopra[2], Gopal Mengi[2], Varun Gupta[2], Richa Upadhyay[1] , Meenakshi Subhash Chippa[1] , Kanjar De[1] , Rajkumar Saini[1] , Seiichi Uchida[3] and Marcus Liwicki[1]

[1]Lulea˚Tekniska Universitet, Lulea˚, Sweden, [2]Punjab University, Chandigarh, India, [3]Kyushu University, Fukuoka, Japan

Motivation

Recent joint-embedding architecture and method based self-supervised representation learning (SSL) approaches are capable to learning efficient representations without the need of direct human supervision. But it requires:

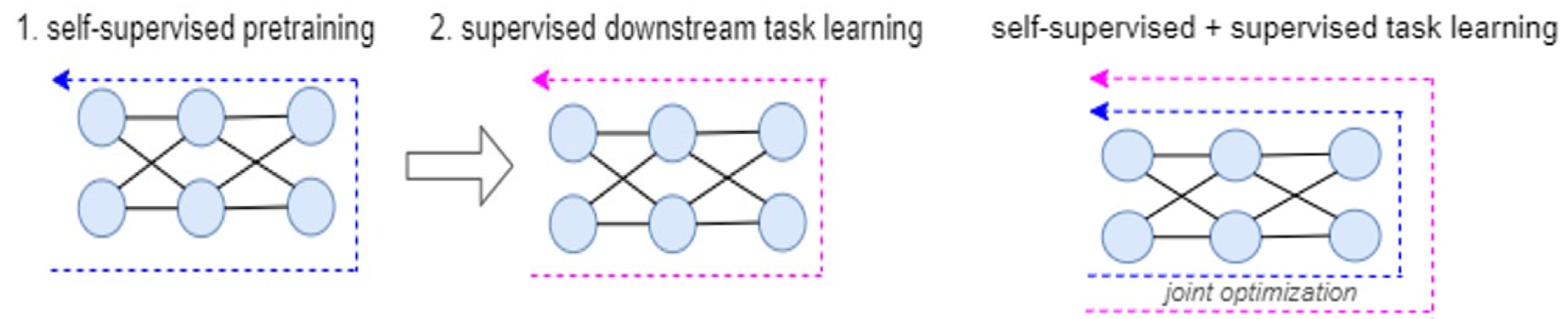
- Massive amount of training data (typical choice –ImageNet having 1.2 millions images)
- Heavy computation resource due to large batch (1024 so on)

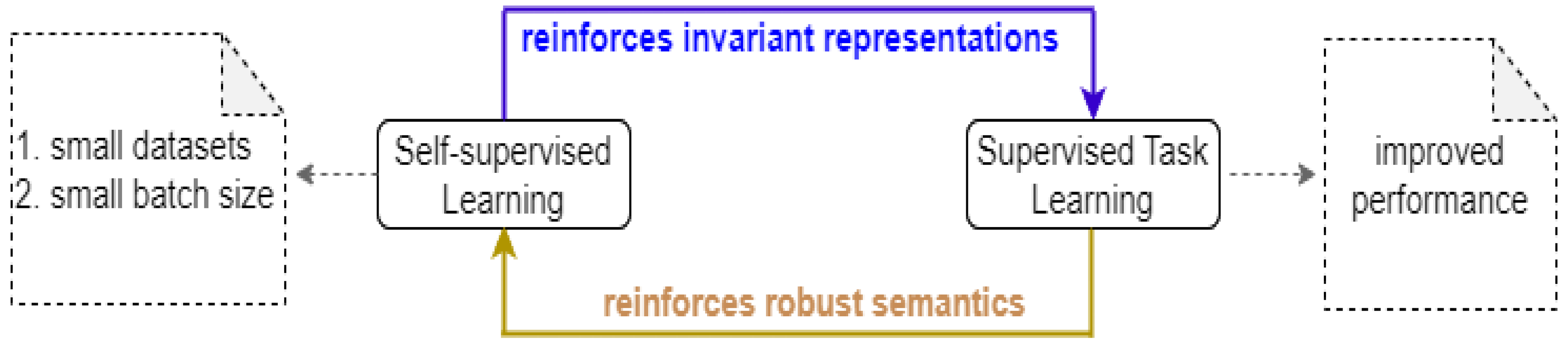
These requirements makes SSL inaccessible to small-scale data and having lack of computing resources

Hypothesis

Shifting the representational knowledge transfer paradigm to functional knowledge transfer can enable the learning of efficient self-supervised representations for small-scale data.

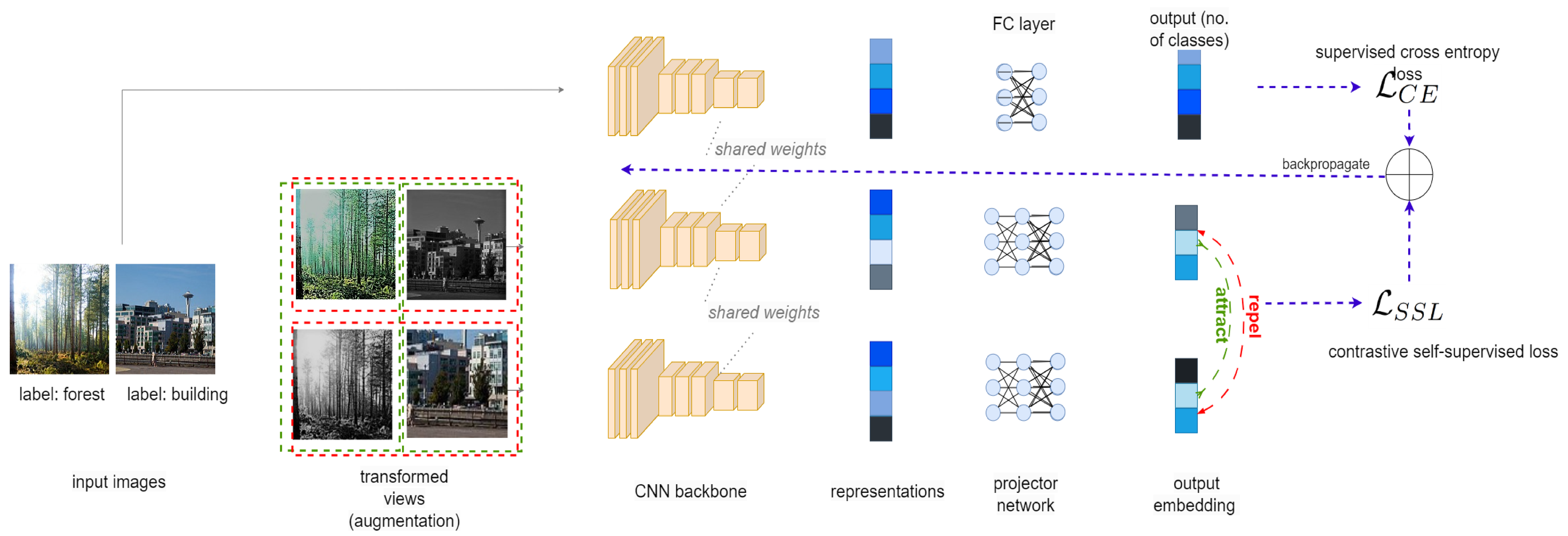
- By using joint optimization of SSL and supervised task





Method

- Jointly optimize the self-supervised and supervised learning where SSL encourages invariant representation while being supported semantics by label being used in supervised setting.
- ResNet-50 backbone and batch size of 256



$$\mathcal{L}_{SSL} = \sum_{(x',x'')\in\mathcal{T}(X)} -\log \frac{e^{\mathcal{A}}}{\sum_{k=1}^{2|X|} 1_{[k\neq x']} e^{\mathcal{B}}}$$

$$\mathcal{A} = (sim(g(\Theta_g; f(\Theta_f; x')), g(\Theta_g; f(\Theta_f; x'')))/\tau)$$

$$\mathcal{B} = (sim(g(\Theta_g; f(\Theta_f; x')), g(\Theta_g; f(\Theta_f; x^k)))/\tau)$$

$$\mathcal{L}_{CE} = -\frac{1}{|D|} \sum_{(x,y)\in D} \sum_{c\in\mathcal{C}} y_c \log(f(\Theta; x_c))$$

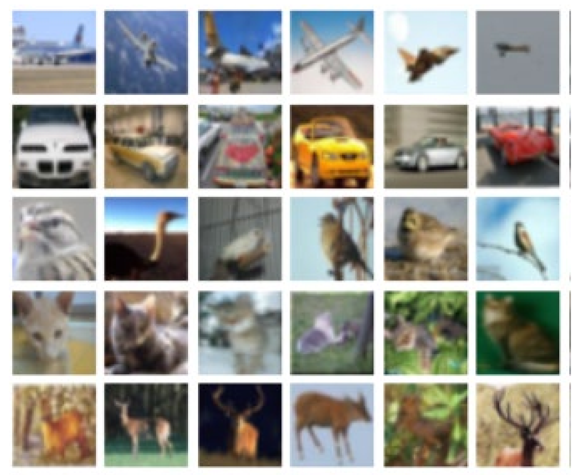
$$\mathcal{L}_{FKT} = \mathcal{L}_{CE} + \lambda\, \mathcal{L}_{SSL}$$

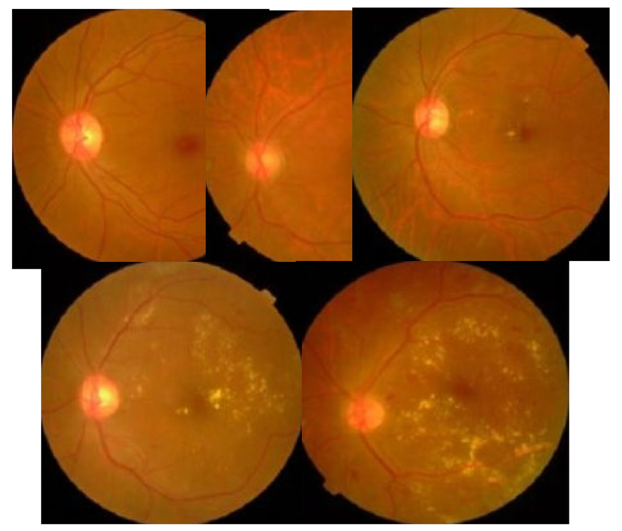
Datasets

CIFAR10 (60k images)

Aptos – Eye fundus (3.6k images)

Intel Images (25k images)





Results

Dataset	Method	Accuracy	Precision	Recall
CIFAR10	Representational Transfer $	92.20±0.11	92.18±0.10	92.21±0.10
	Functional Transfer	**93.60±0.10**	**93.62±0.13**	**93.59±0.11**
Intel Image	Representational Transfer	93.18±0.15	93.15±0.18	93.17±0.20
	Functional Transfer	**93.70±0.13**	**93.33±0.11**	**93.31±0.11**
Aptos 2019	Representational Transfer	83.10±0.10	83.05±0.09	**83.05±0.12**
	Functional Transfer	**83.32±0.11**	**83.14±0.10**	83.04±0.10