LULEÅ TEKNISKA UNIVERSITET

INTERNATIONAL NEURAL NETWORK SOCIETY
IJCNN2023
INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS
18-23 JUNE 2023 | GOLD COAST CONVENTION AND EXHIBITION CENTRE, QUEENSLAND, AUSTRALIA

Chandigarh College of Engineering and Technology

Domain Adaptable Self-supervised Representation Learning on Remote Sensing Satellite Imagery

Muskaan Chopra[1,*], Prakash Chandra Chhipa[2,*], Gopal Mengi[1,*], Varun Gupta[1] and Marcus Liwicki[2]

[1]Chandigarh College of Engineering and Technology, Punjab University, Chandigarh, India
[2]Machine Learning Group, EISLAB, Lulea˚ Tekniska Universitet, Lulea˚, Sweden

*first – co-authors with equal contributions

Self-supervised Learning on Natural Scenes

- ✓ No direct human supervision
- ✓ Effective representations
- ✓ Improved downstream tasks performance

<u>But</u>

- o Need large data
- o Learn invariance by geometric (crop, flip, etc.) and color (contrast, saturate, etc.) transformations

Common Schema of Self-supervised Representation Learning[1,2,3] (SSL) Approach

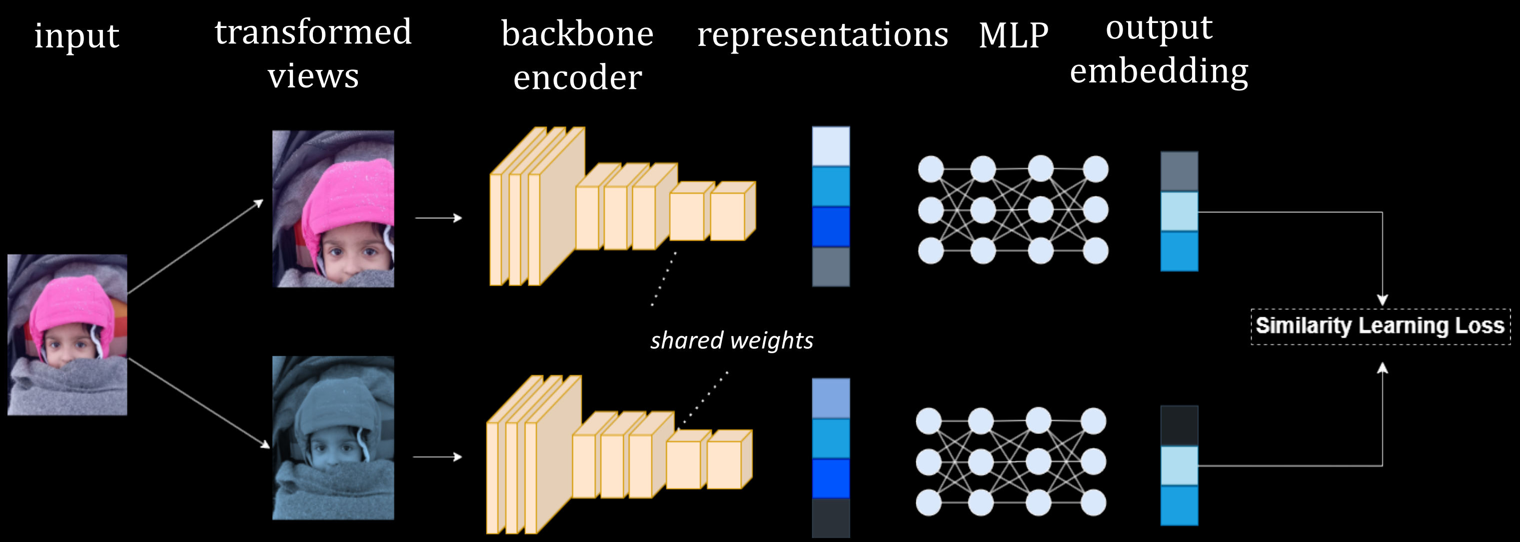


Figure Inspired : Chhipa, Prakash Chandra. "Self-supervised Representation Learning for Visual Domains Beyond Natural Scenes." Licentiate Thesis, Luleå tekniska universitet (2023).

[1]**Contrastive** - Chen, Ting, et al. "A simple framework for contrastive learning of visual representations." International conference on machine learning. PMLR, 2020

[2]**Distillation**- Grill, Jean-Bastien, et al. "Bootstrap your own latent-a new approach to self-supervised learning." Advances in neural information processing systems 33 (2020).

[3]**Information Maximization**- Zbontar, Jure, et al. "Barlow twins: Self-supervised learning via redundancy reduction." International Conference on Machine Learning. PMLR, 2021.

LULEÅ TEKNISKA UNIVERSITET
INTERNATIONAL NEURAL NETWORK SOCIETY
IJCNN2023
INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS
18–23 JUNE 2023 | GOLD COAST CONVENTION AND EXHIBITION CENTRE, QUEENSLAND, AUSTRALIA
Chandigarh College of Engineering and Technology

Visual Concepts in Remote Sensing Imagery

dense residential

coastal mansions

difference is obvious and discriminative in nature

difference is neither significant nor easily explainable

Learning without (or least) human supervision should be favorable

Hypothesis

Self-supervised representation learning (SSL) can be adapted to remote sensing imagery without much customizations.

Questions

- SSL often needs large data to learn on natural scenes (Ref: ImageNet), does it require massive amount in remote sensing also?
- Can same set of transformations be used to learn invariant features in remote sensing?
- How effective domain adaptation and knowledge transfer it can deliver?

Investigation Protocol

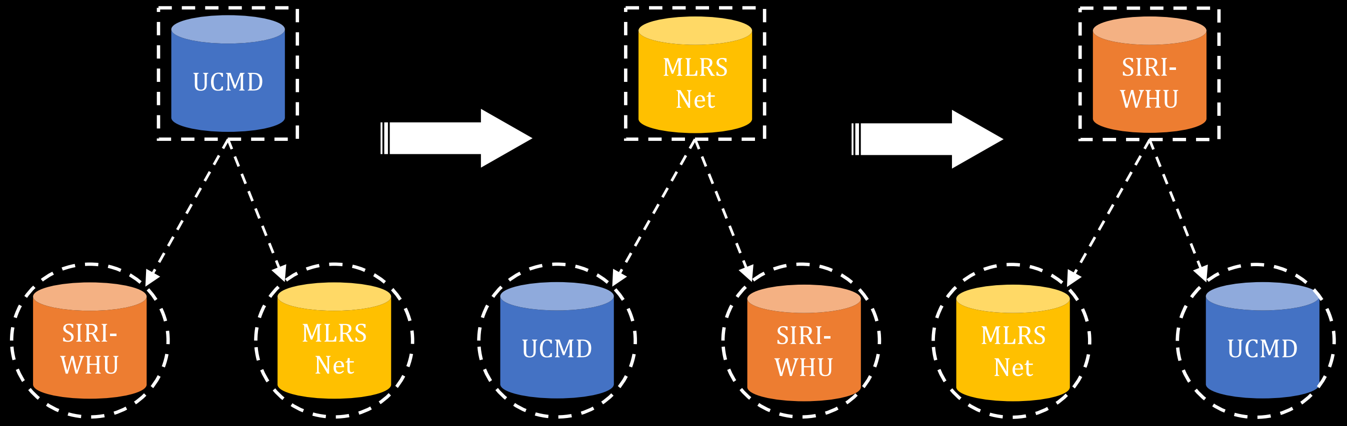
1. Identify multiple remote sensing imagery datasets
2. Chose self-supervised learning method
3. Examine SSL domain adaptation in round-robin fashion
   - ✓ perform self-supervised pretraining on one (source) dataset
   - ✓ evaluate knowledge transfer through downstream task on remaining (target) datasets
   - ✓ repeat until round-robin finishes

Investigation Protocol

three remote sensing datasets

contrastive SSL pretraining

downstream task



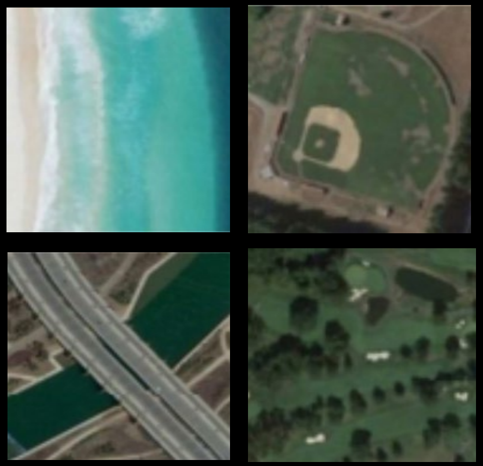
contrastive learning and dataset description come soon

Datasets

MLSRNet[1] Dataset

MLRS Net

46 classes and 109,161 examples



Classes - beach, baseball diamond, bridge, golf course, freeway, industrial area, forest, etc.

UCMD[2] Dataset

UCMD

21 classes and 2100 examples

Classes - airplane, dense residential, harbor, freeway, parking lot, forest, river, etc.

SIRI-WHU[3] Dataset

SIRI-WHU

12 classes and 2400 examples

Classes - agriculture, commercial, harbor, idle land, industrial, meadow, overpass, park, pond, residential, river, and water

[1] *MLSRNet, https://data.mendeley.com/datasets/7j9bv9vwsx/2*

[2] *UCMD, http://weegee.vision.ucmerced.edu/datasets/landuse.html*

[3] *SIRI-WHU, ttp://www.lmars.whu.edu.cn/prof_web/zhongyanfei/e-code.html*

Adapting SSL on Diabetic Retinopathy

✓Contrastive Self-Supervision
- SimCLR[1]
- learn similarity for positive pair
- learn dissimilarity otherwise
- ResNet-50 backbone

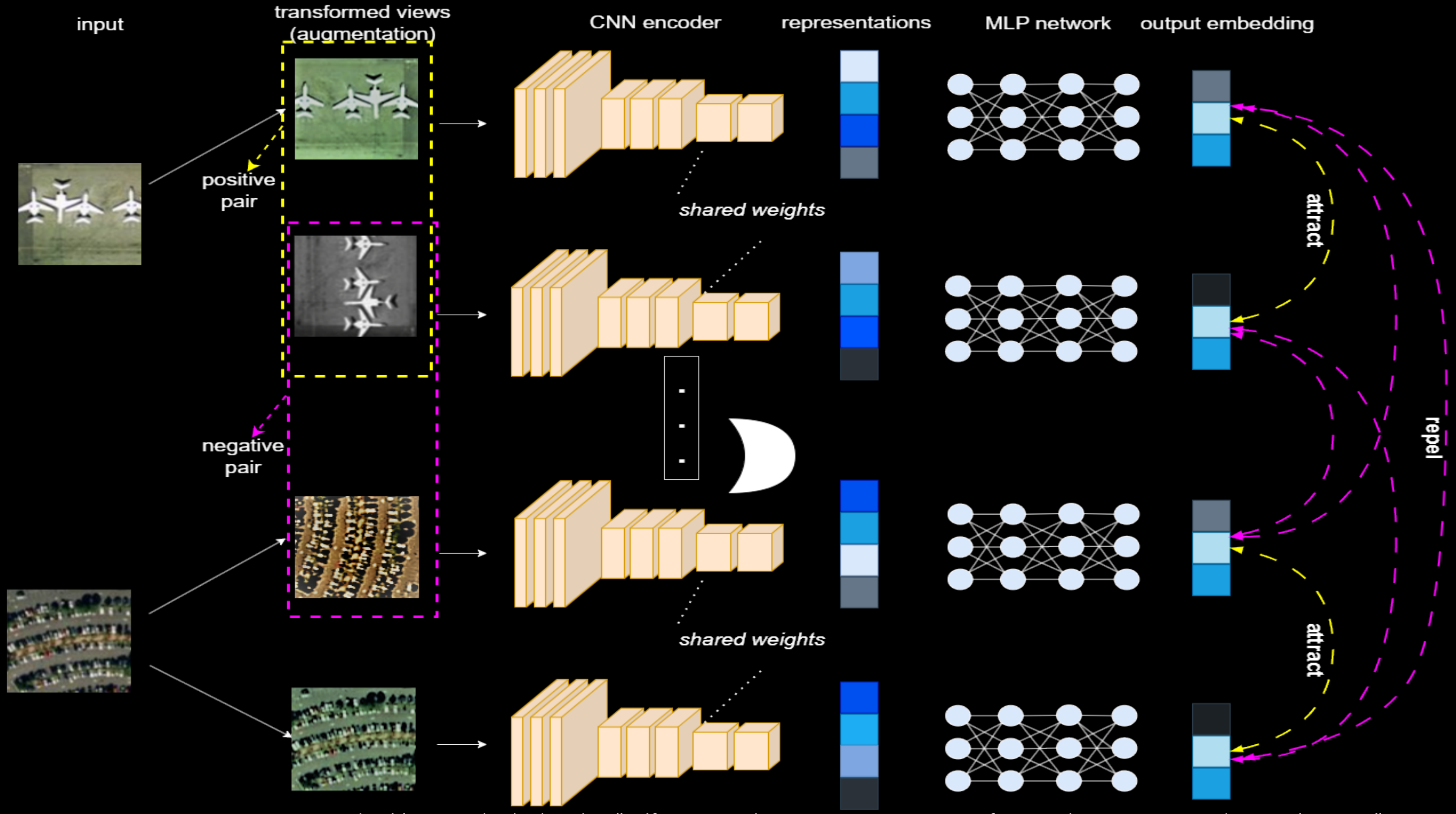


Figure Inspired : Chhipa, Prakash Chandra. "Self-supervised Representation Learning for Visual Domains Beyond Natural Scenes." Licentiate Thesis, Luleå tekniska universitet (2023).

[1]*Chen, Ting, et al. "A simple framework for contrastive learning of visual representations." International conference on machine learning. PMLR, 2020*

Evaluations on Domain Adaptable SSL and Knowledge Transfer

1. Downstream Tasks

- Multi-class classification
- SSL domain adaptation
  - Distinct pretraining and downstream RSI datasets
- SSL same domain/dataset
  - Same pretraining and downstream RSI datasets

2. Label Efficiency Vs. SSL Domain Adaptation

- 10%, 50%, 100% labels
- Same RSI domain Vs. cross RSI domain SSL
- Supervised ImageNet knowledge transfer

3. Qualitative Analysis

- Class Activations

RSI - Remote Sensing Satellite Imagery

Multiclass classification Task

- Compares result with previous works
- 46 classes

Dataset - MLRSNet

Author	Method	Accuracy
[30]	DenseNet201-SR-Net	87.87
[30]	ResNet101-SR-Net	87.48
[17]	Self-Supervised Learning	96.00
Our Results	Self-supervised Domain Adaptation[1]	**96.34**
	Self-supervised Domain Adaptation[2]	**97.87**
	Self-supervised Same Domain[3]	**96.59**

[1] *Pretrained on UCMD and downstream on MLRSNet*

[2] *Pretrained on SIRI-WHU and downstream on MLRSNet*

[2] *Pretrained and downstream on MLRSNet*

LULEÅ TEKNISKA UNIVERSITET

INTERNATIONAL NEURAL NETWORK SOCIETY IJCNN2023 INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS 18-23 JUNE 2023 GOLD COAST CONVENTION AND EXHIBITION CENTRE, QUEENSLAND, AUSTRALIA

Chandigarh College of Engineering and Technology

Multiclass classification Task

- Compares result with previous works
- 25 classes

Dataset - UCMD

Author	Method	Accuracy
[25]	ResNet 50	98.00
[26]	DCNN	93.48
[23]	GoogleNet	97.10
[14]	Semisupenised ensemble projection	66.49
Our Results	Self-supervised Domain Adaptation[1]	**98.50**
	Self-supervised Domain Adaptation[2]	**98.75**
	Self-supervised Same Domain[3]	**99.68**

[1] *Pretrained on SIRI-WHU and downstream on UCMD*

[2] *Pretrained on MLRSNet and downstream on UCMD*

[3] *Pretrained and downstream on UCMD*

Multiclass classification Task

- Compares result with previous works
- 12 classes

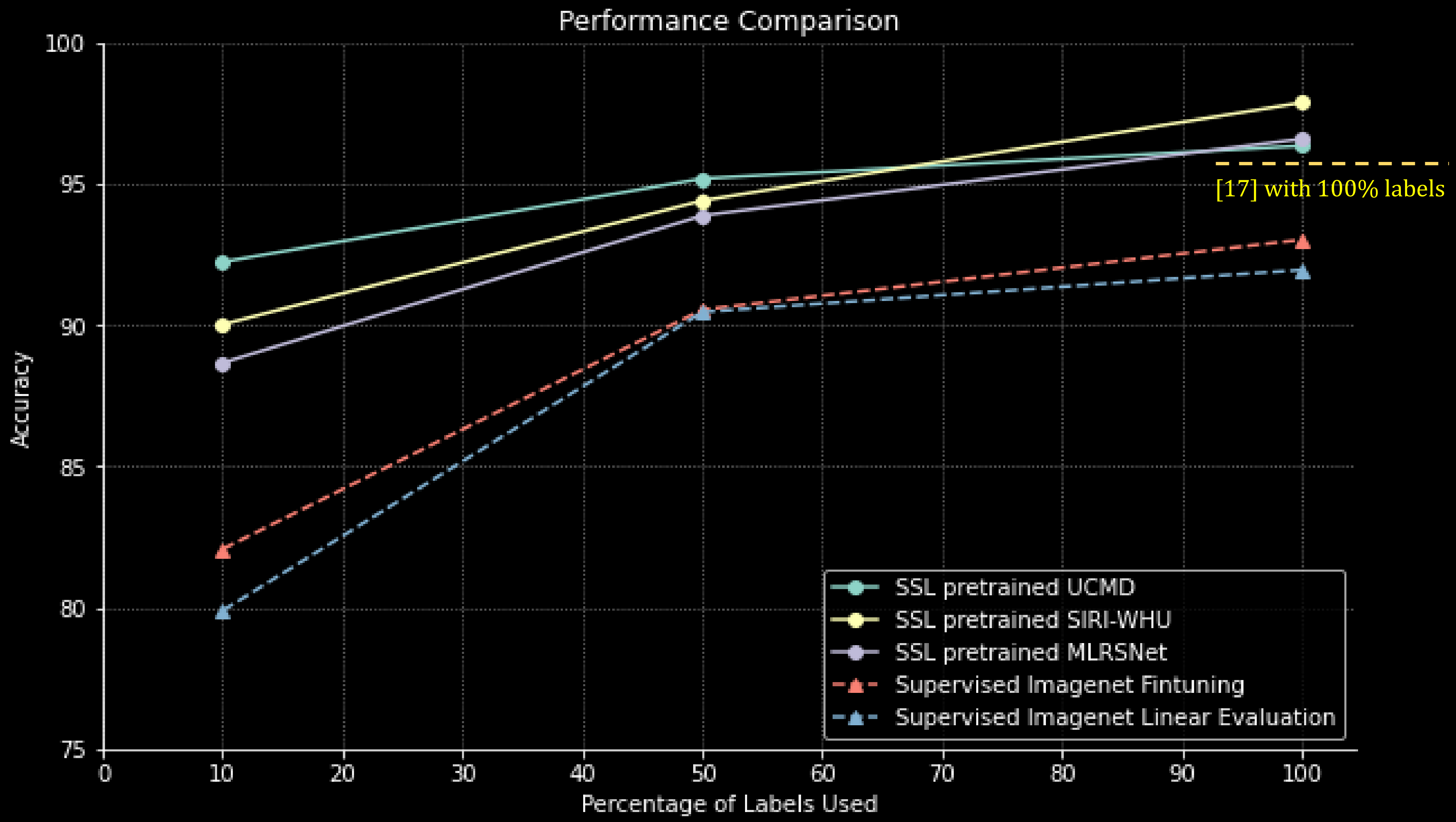
Dataset - SIRI-WHU

Author	Method	Accuracy
[27]	AlexNet SPP SS	95.07
[28]	MCNN	93.75
[29]	Inception-LSTM	**99.73**
Our Results	Self-supervised Domain Adaptation[1]	96.87
	Self-supervised Domain Adaptation[2]	97.50
	Self -supervised Same Domain[3]	**99.68**

[1]*Pretrained on UCMD and downstream on SIRI-WHU* [2]*Pretrained on MLRSNet and downstream on SIRI-WHU* [3]*Pretrained and downstream on SIRI-WHU*

Label Efficiency Vs. SSL Domain Adaptation

- Classification performance on MLRSNet Dataset
- SSL pretrained models consistently outperforms
- Knowledge transfer in SSL cross domain and SSL same domain are comparable



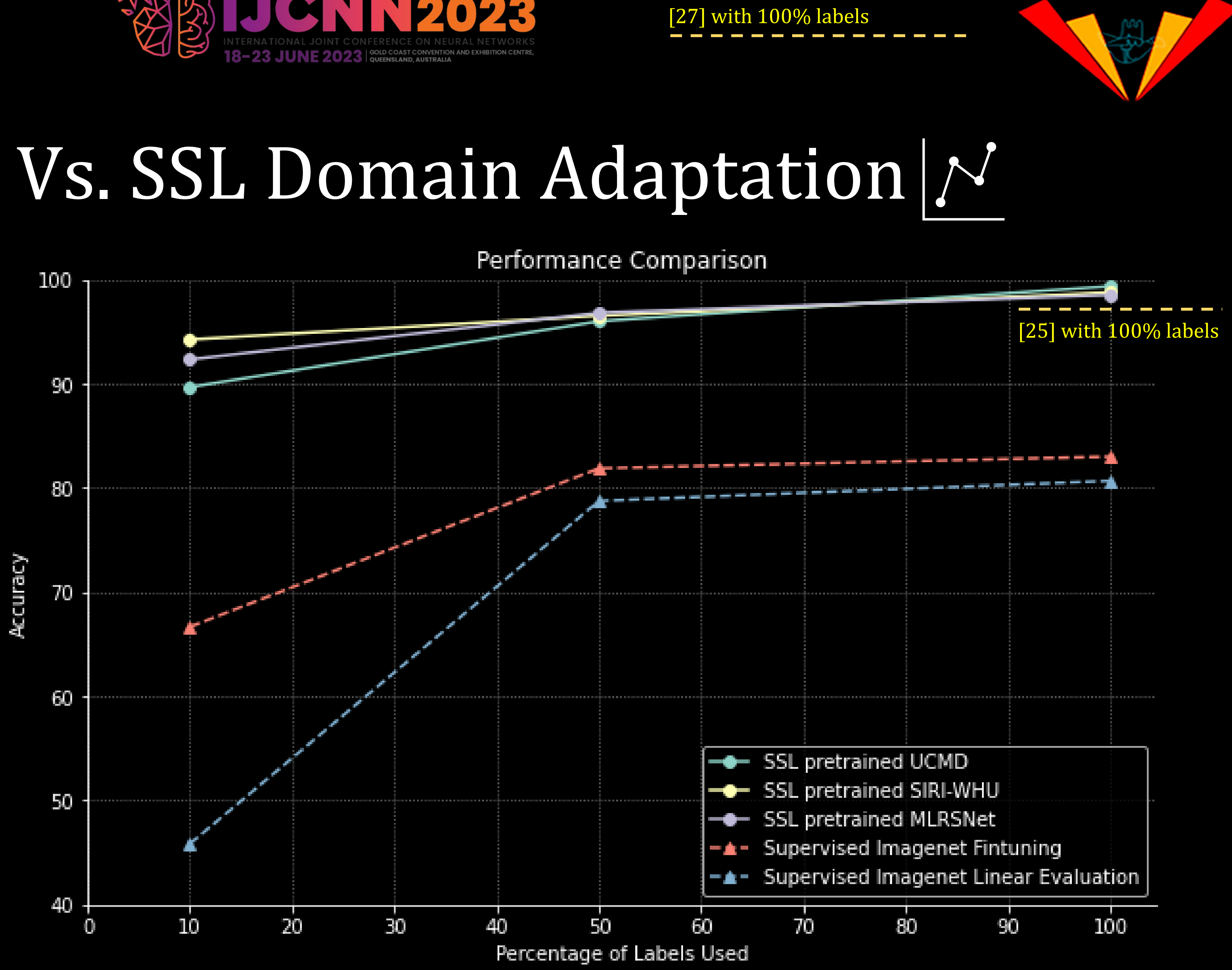
LULEÅ TEKNISKA UNIVERSITET

INTERNATIONAL NEURAL NETWORK SOCIETY
IJCNN 2023
INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS
18-23 JUNE 2023 | GOLD COAST CONVENTION AND EXHIBITION CENTRE, QUEENSLAND, AUSTRALIA

Chandigarh College of Engineering and Technology

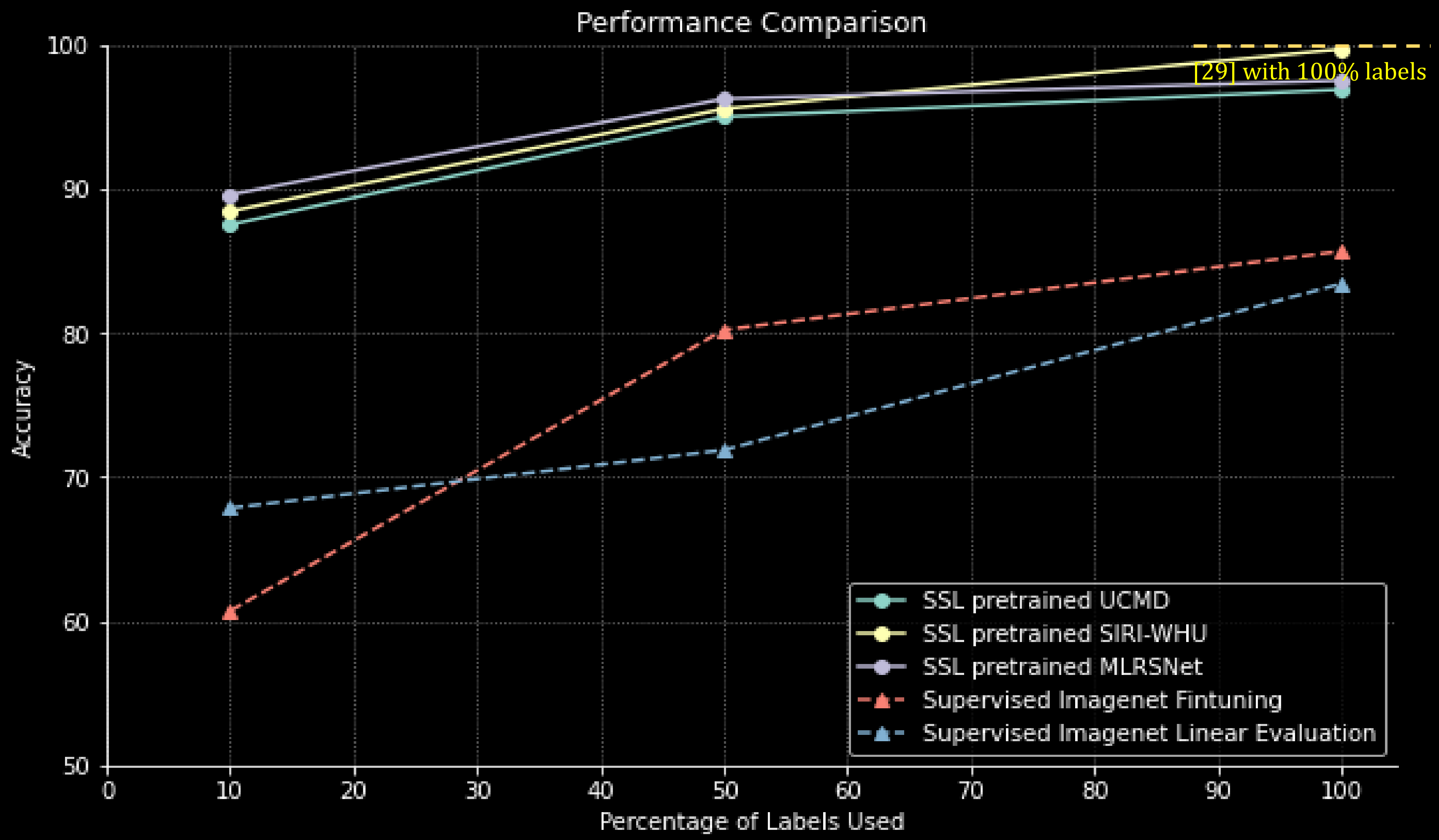
Label Efficiency Vs. SSL Domain Adaptation

- Classification performance on UCMD Dataset
- SSL pretrained models consistently outperforms over supervised ImageNet
- Knowledge transfer in SSL cross domain and SSL same domain are close



Label Efficiency Vs. SSL Domain Adaptation

- Classification performance on SIRI-WHU Dataset
- SSL pretrained models consistently outperforms over supervised ImageNet
- Knowledge transfer in SSL cross domain and SSL same domain are close

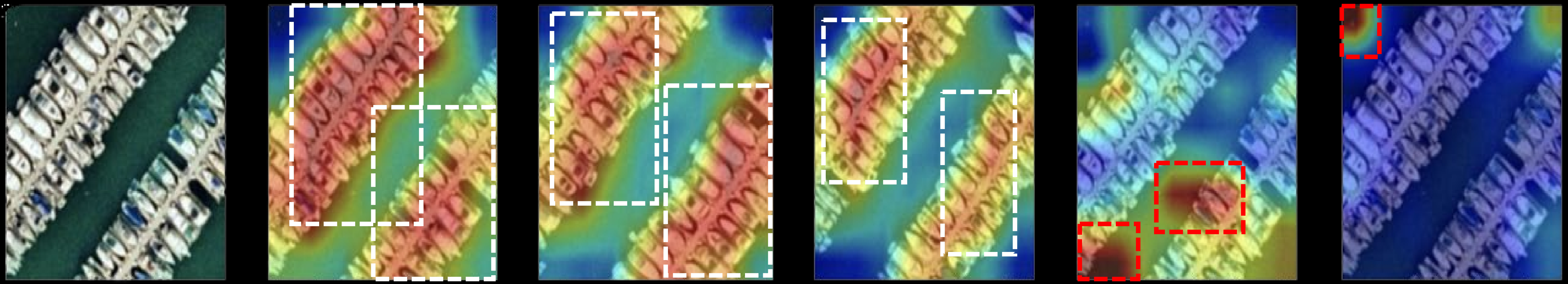


Qualitative Analysis

- SSL representations are robust and attend the visual concept rightly

Example from UCMD dataset

original image | UCMD pretrained | SIRI-WHU pretrained | MLRSNet pretrained | ImageNet + finetuned | ImageNet + linear evaluation



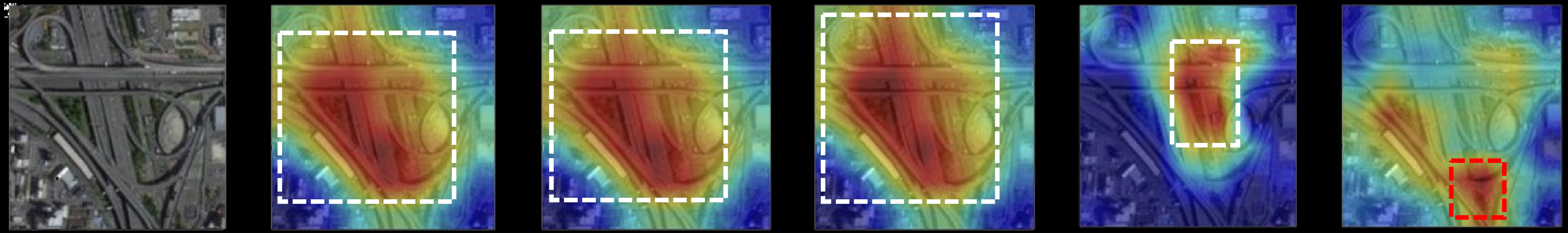
SSL same dataset | *SSL domain adaption* | *Supervised knowledge transfer*

Qualitative Analysis

- SSL representations are robust and attend the visual concept rightly

Example from SIRI-WHU dataset

original image	UCMD pretrained	SIRI-WHU pretrained	MLRSNet pretrained	ImageNet + finetuned	ImageNet + linear evaluation



SSL same dataset

SSL domain adaption

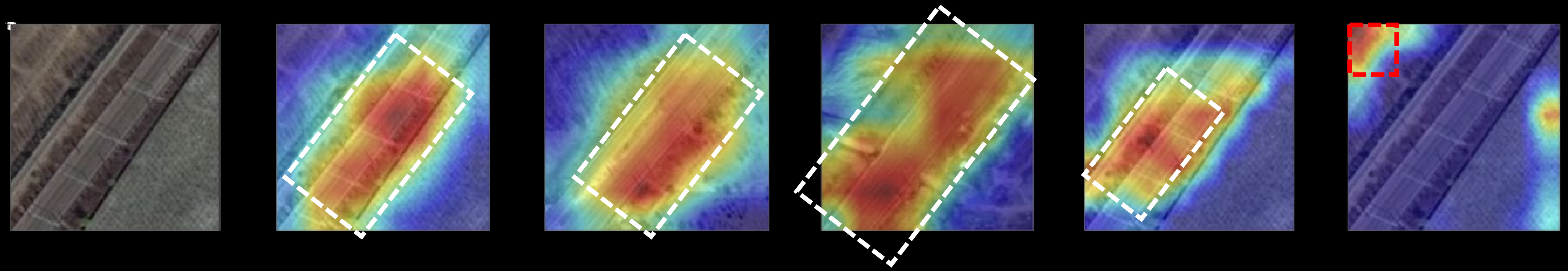
Supervised knowledge transfer

Qualitative Analysis

- SSL representations are robust and attend the visual concept rightly

Example from MLRSNet dataset

original image	UCMD pretrained	SIRI-WHU pretrained	MLRSNet pretrained	ImageNet + finetuned	ImageNet + linear evaluation



SSL domain adaption (UCMD pretrained, SIRI-WHU pretrained)

SSL same dataset (MLRSNet pretrained)

Supervised knowledge transfer (ImageNet + finetuned, ImageNet + linear evaluation)

Conclusions

Contribution

Adapting contrastive self-supervised representation learning to remote sensing satellite imagery domain and verifying the domain adaptation by formulating and examining robust hypothesis

Achievements

Achieved efficient domain adaptable knowledge transfer and shown performance improvement in downstream tasks over supervised knowledge transfer, supported by qualitative analysis

Future Work

Further investigation on adapting other SSL approaches and evaluating challenging downstream tasks including segmentations and detections

LULEÅ TEKNISKA UNIVERSITET

INTERNATIONAL NEURAL NETWORK SOCIETY
IJCNN2023
INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS
18-23 JUNE 2023 | GOLD COAST CONVENTION AND EXHIBITION CENTRE, QUEENSLAND, AUSTRALIA

Chandigarh College of Engineering and Technology

Thank you

prakash.chandra.chhipa@ltu.se

GitHub