Deep Learning in Remote Sensing

Challenges:

1. Deep learning needs large amount of data to train
2. Remote sensing typically has small amount of annotated data

Research Questions:

1. Is it possible to transfer features from CV to the remote sensing domain?
2. Do transferred features more effective than fully-trained?
3. How to better exploit deep learning in RS data?

Reference

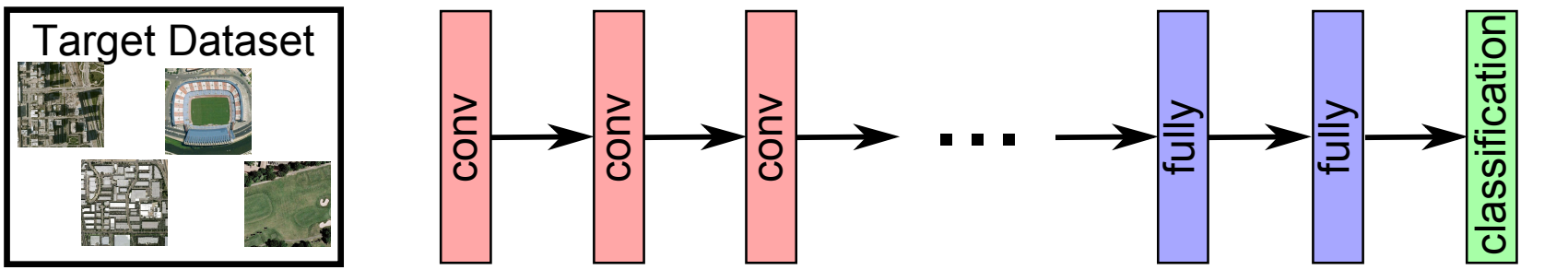
K. Nogueira, O. A. B. Penatti and J. A. dos Santos. Towards better exploiting convolutional neural networks for remote sensing scene classification. Pattern Recognition, 61, 539-556, 2017.

1 - Fully Training

Training from scratch

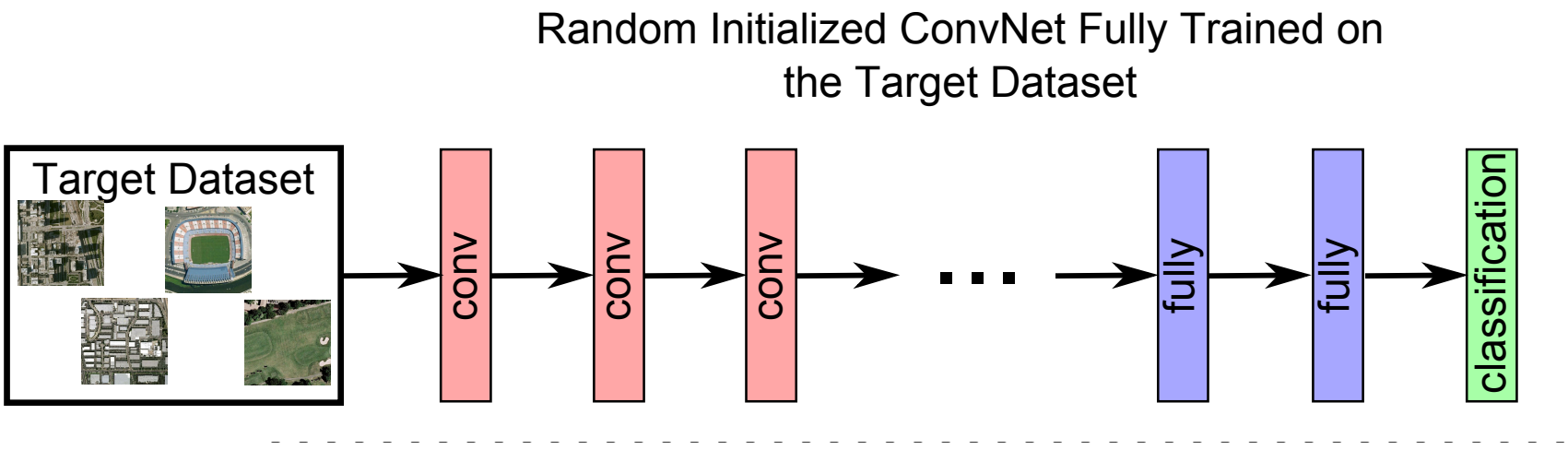
1 - Fully Training

Training from scratch

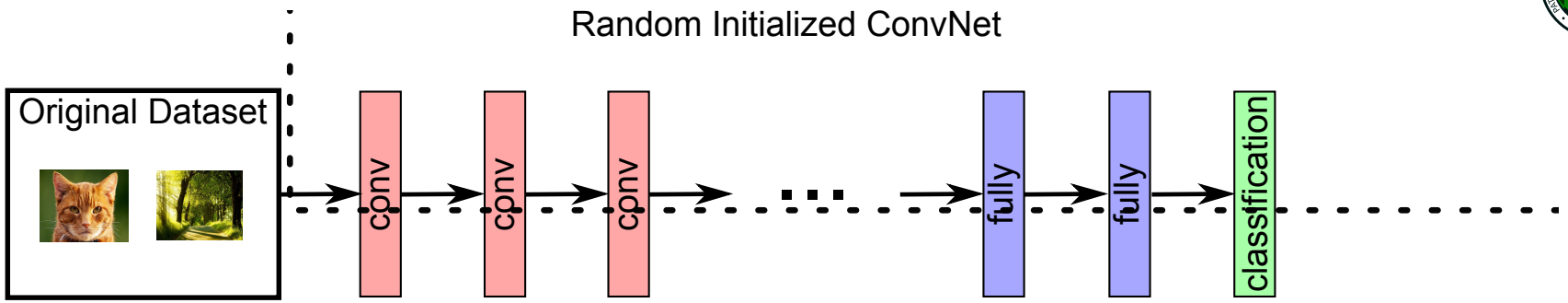


1 - Fully Training

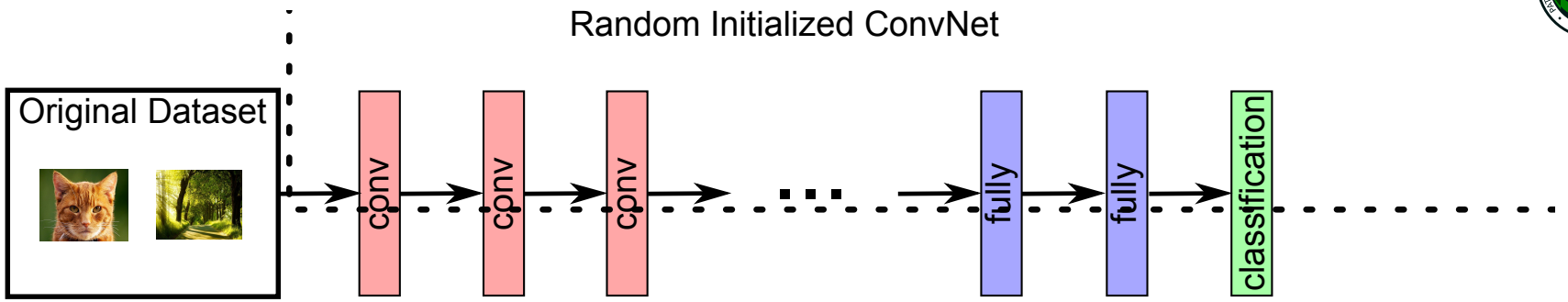
Training from scratch



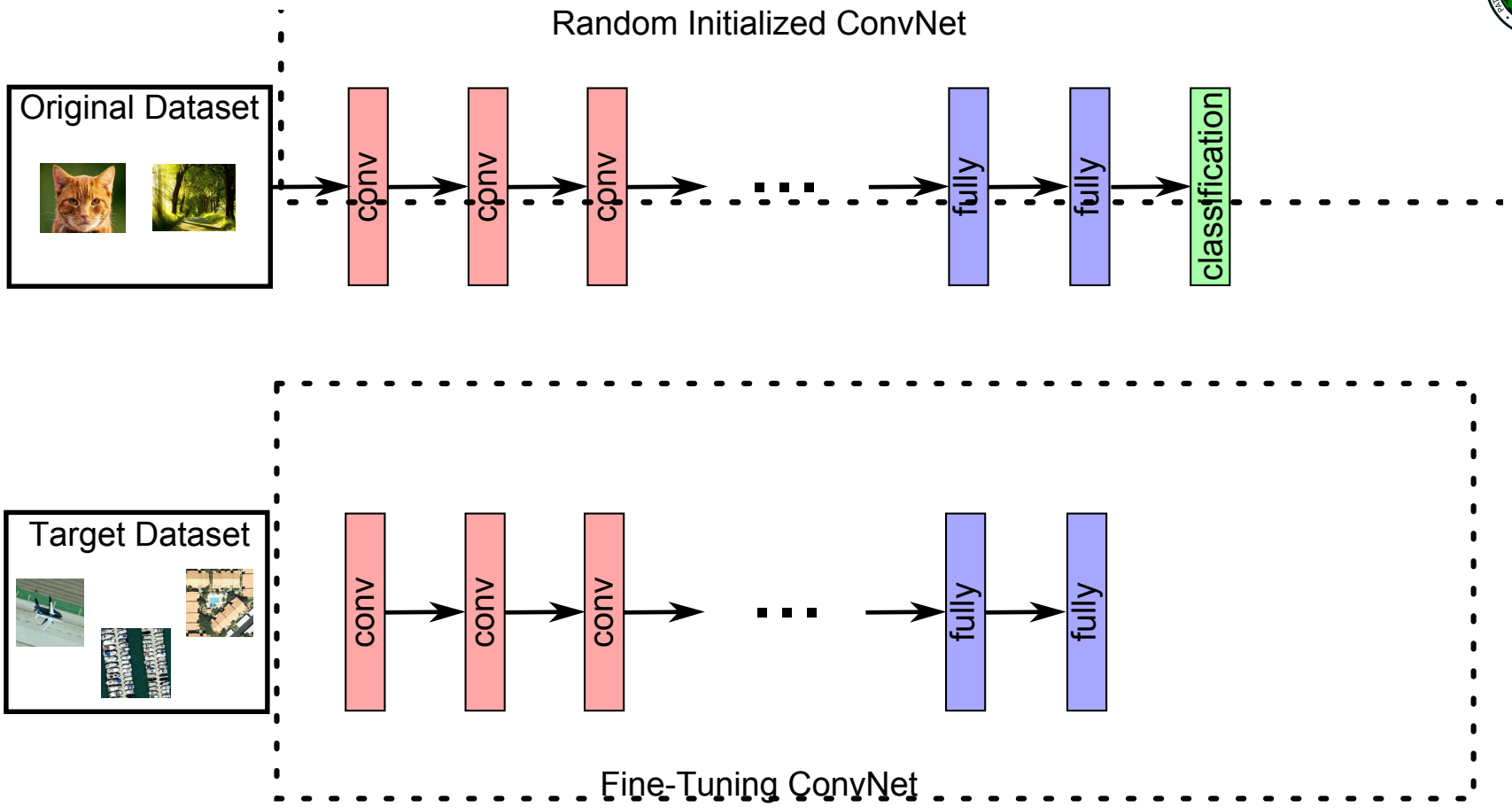
2 - Fine-Tuning



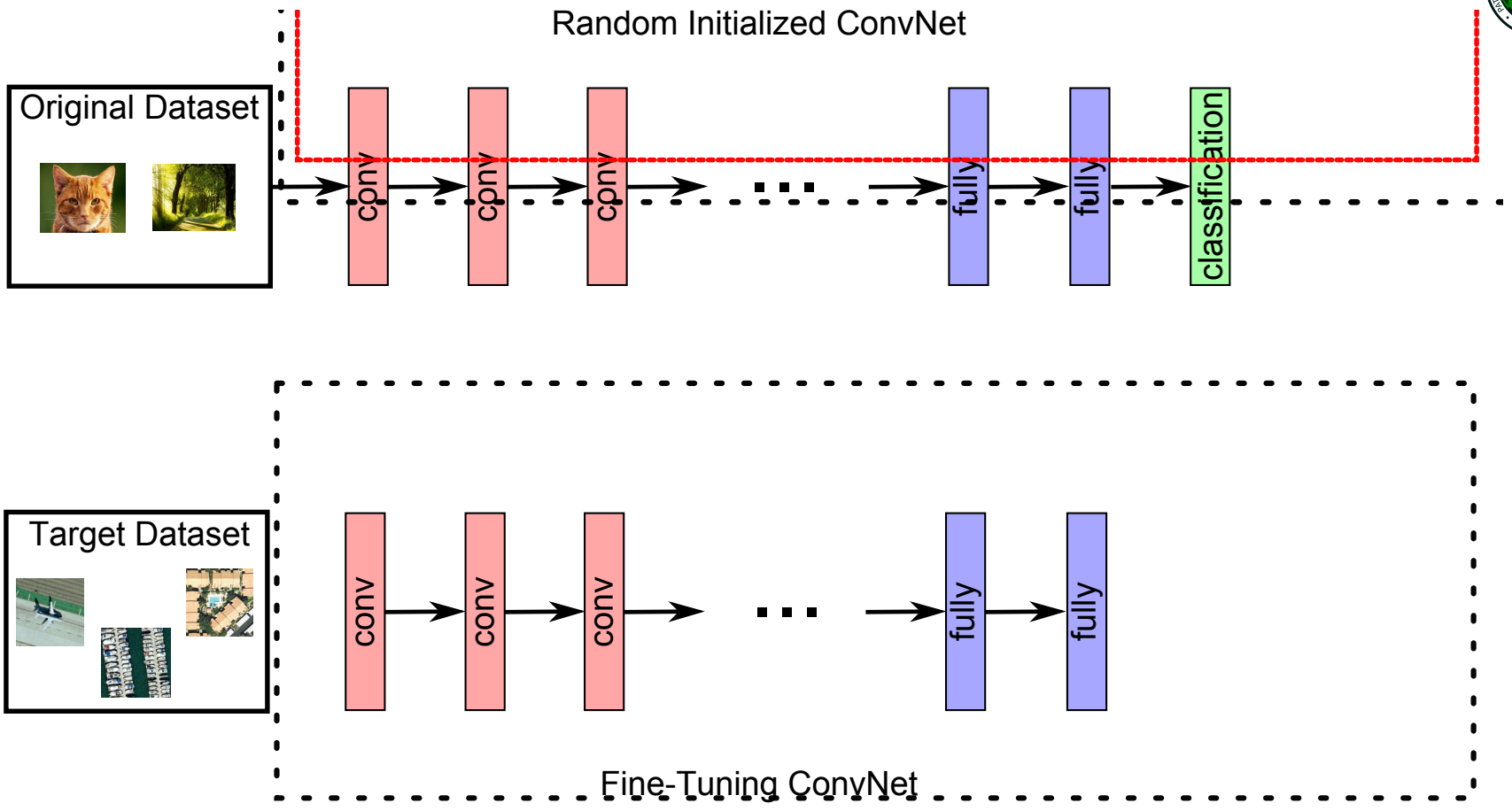
2 - Fine-Tuning



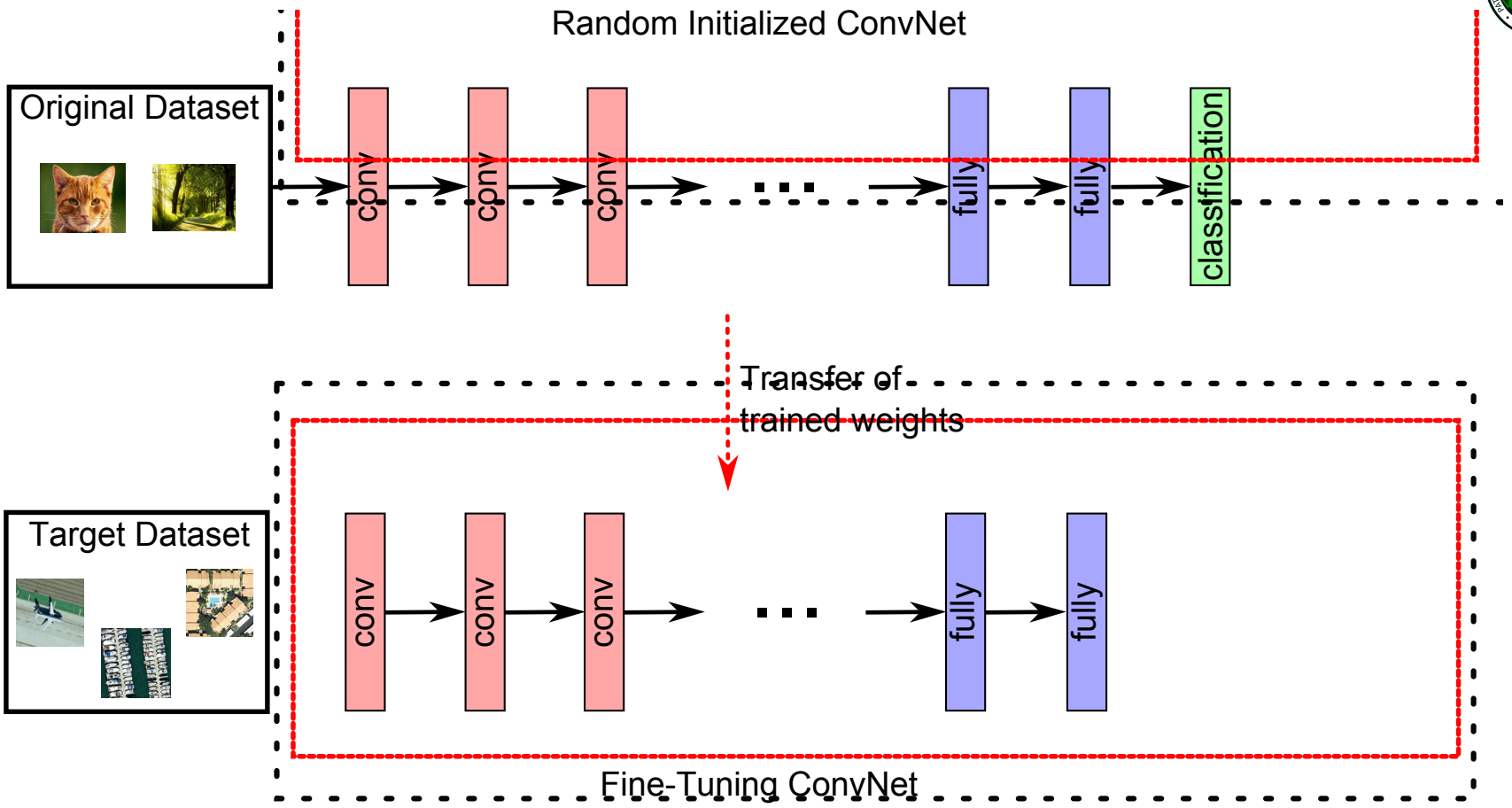
2 - Fine-Tuning



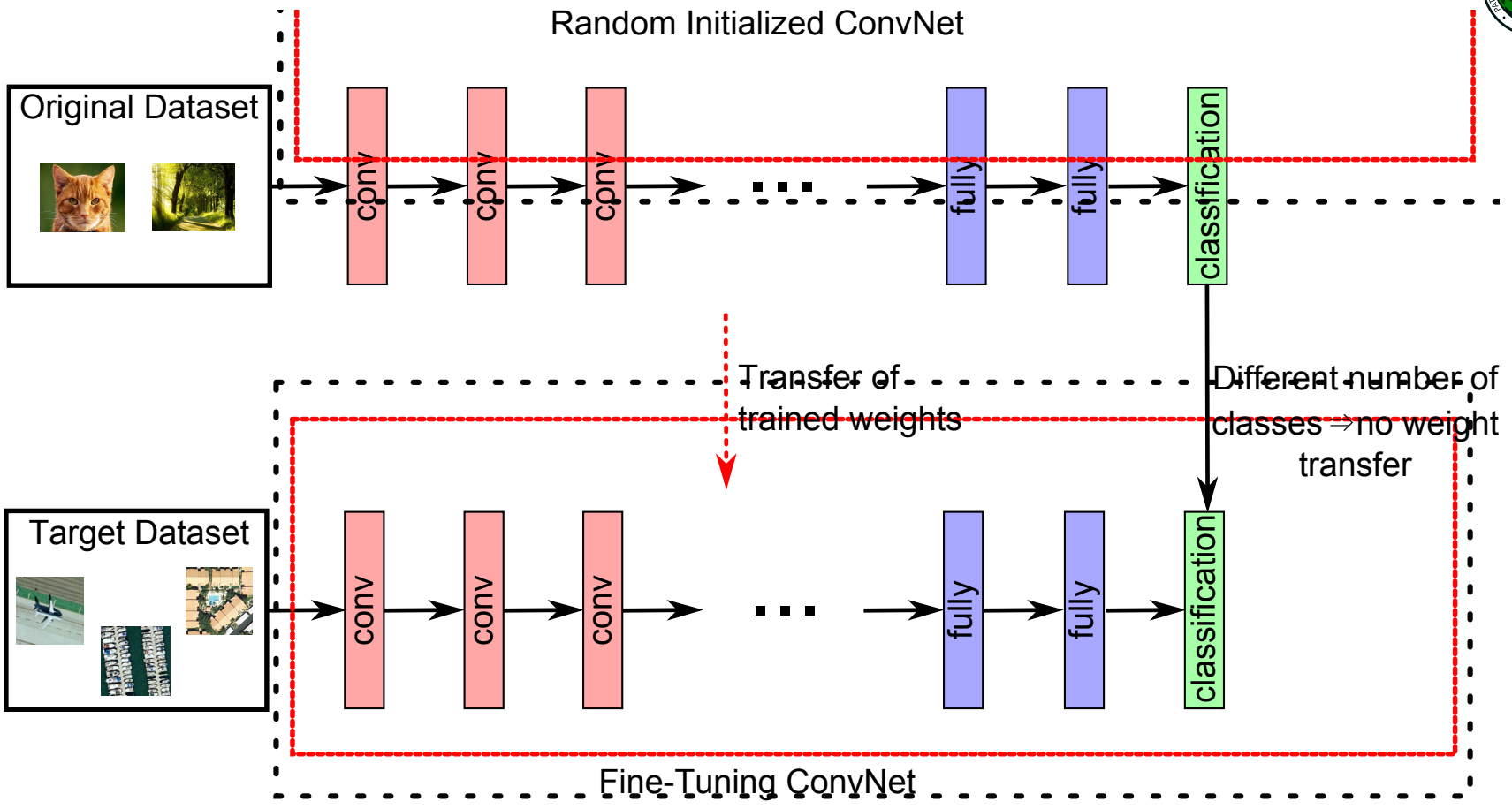
2 - Fine-Tuning



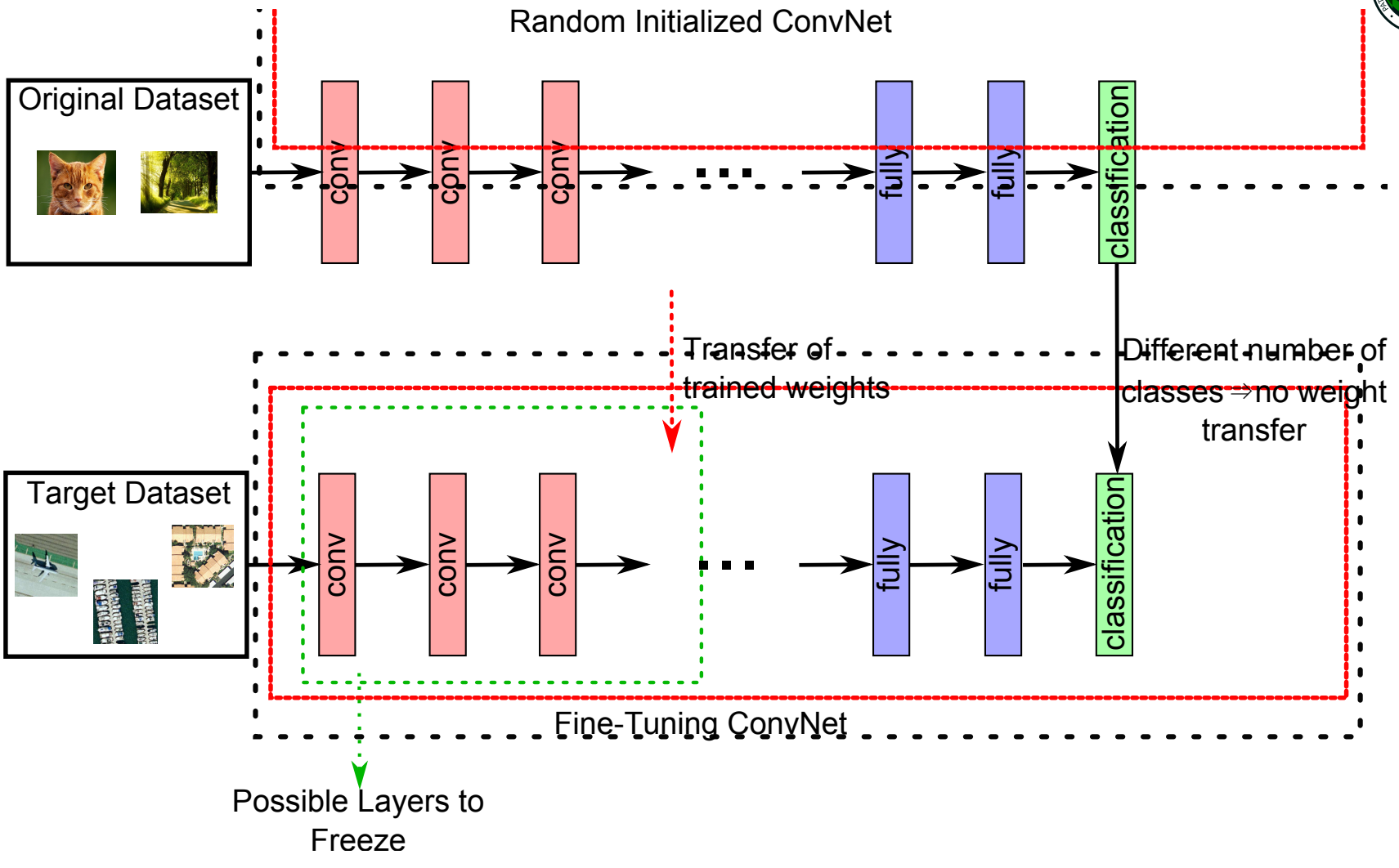
2 - Fine-Tuning



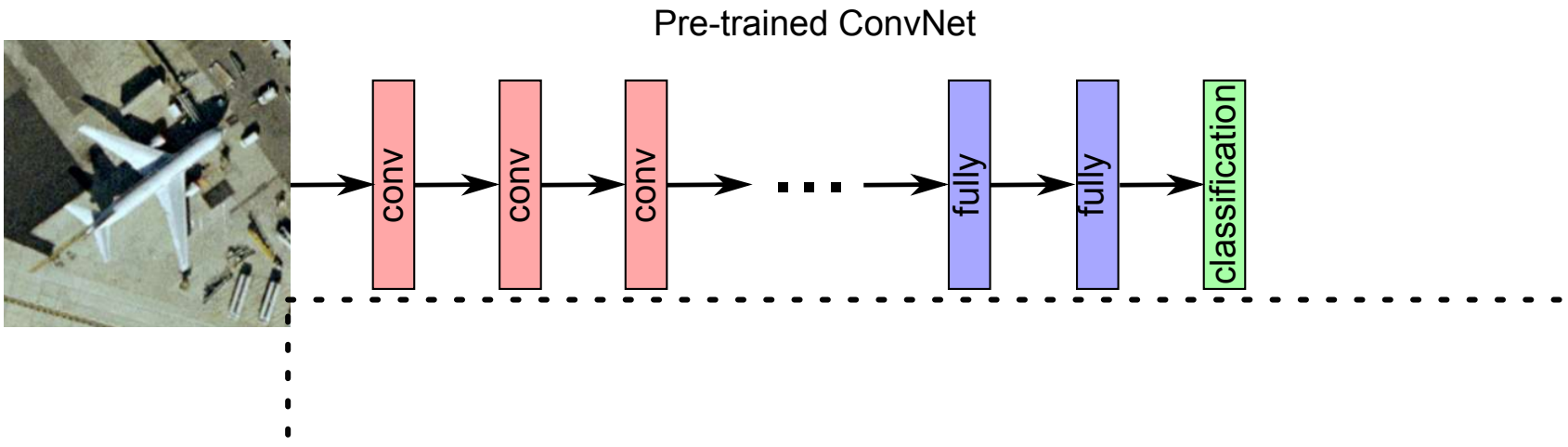
2 - Fine-Tuning



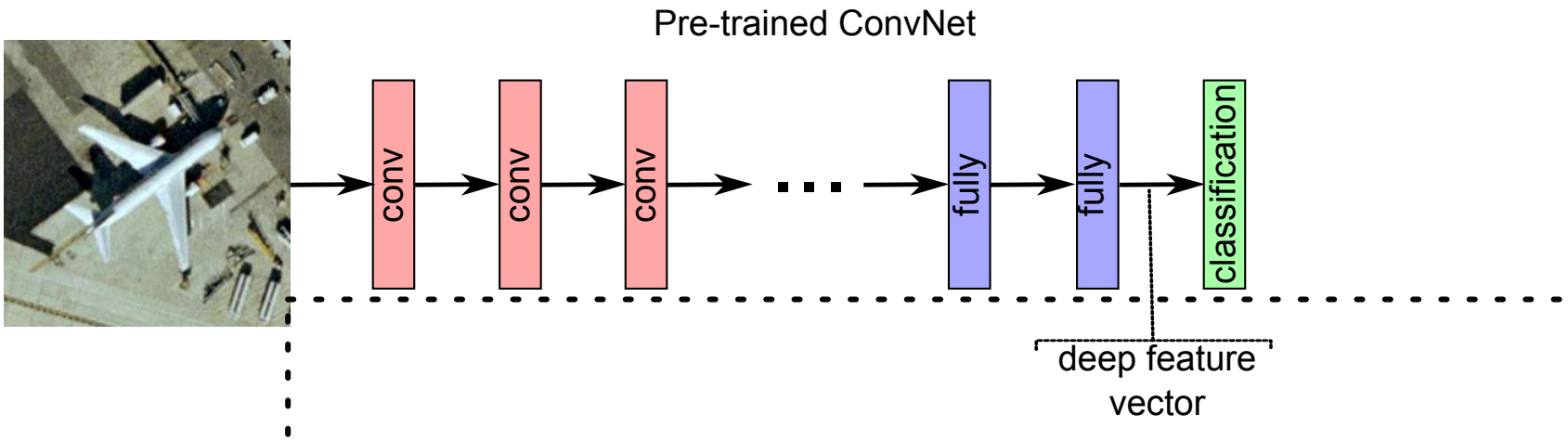
2 - Fine-Tuning



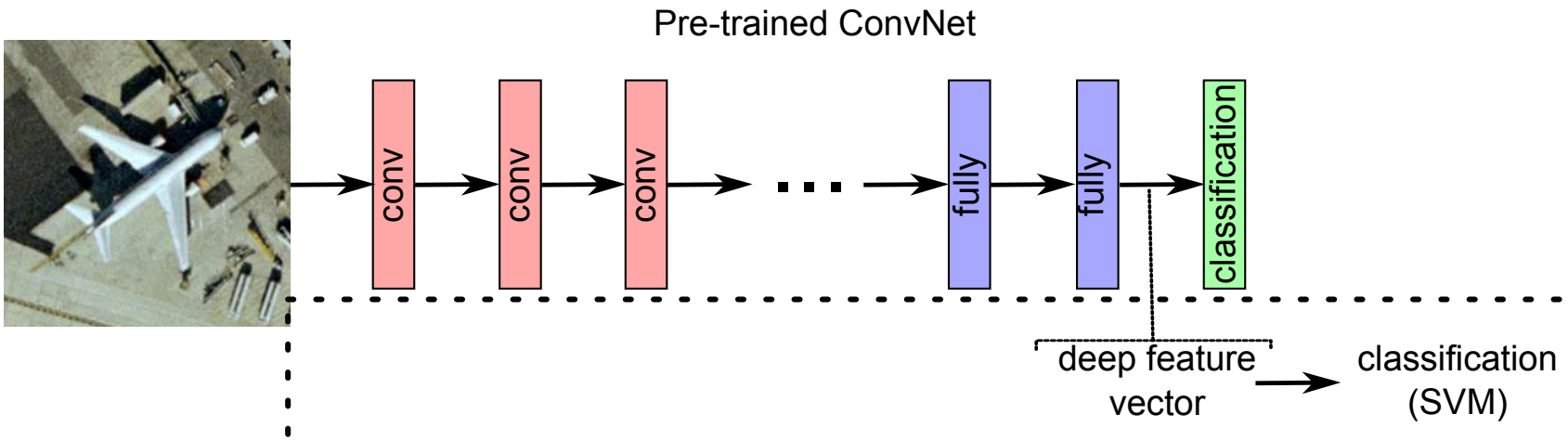
3 - Feature Extractor



3 - Feature Extractor



3 - Feature Extractor



Datasets

Statistics

- Three public datasets:

	UC-Merced	RS19	Coffee
Type	Land-use	Land-use	Agricultural
Dataset Size	2,100	1,005	2,876
Image Dimensions (px.)	256×256	600×2600	64×64
Band Composition	R-G-B	R-G-B	NIR-R-G
Spatial Resolution	0.3m	0.5m	2.5m
Number of Classes	21	19	2
Average Distribution per Class	100	52	1,438

Datasets

Examples

UCMerced land-use Dataset

(a) Agricultural

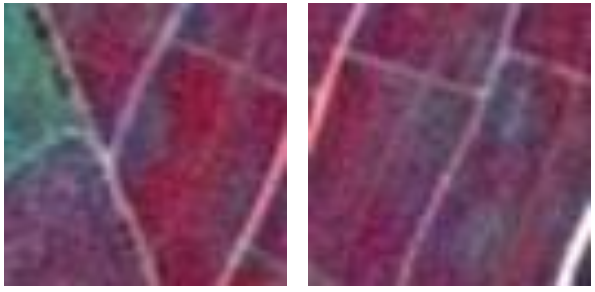
(b) Dense Residential

RS19 Dataset

(c) Beach

(d) Football Field

Brazilian Coffee Scenes Dataset



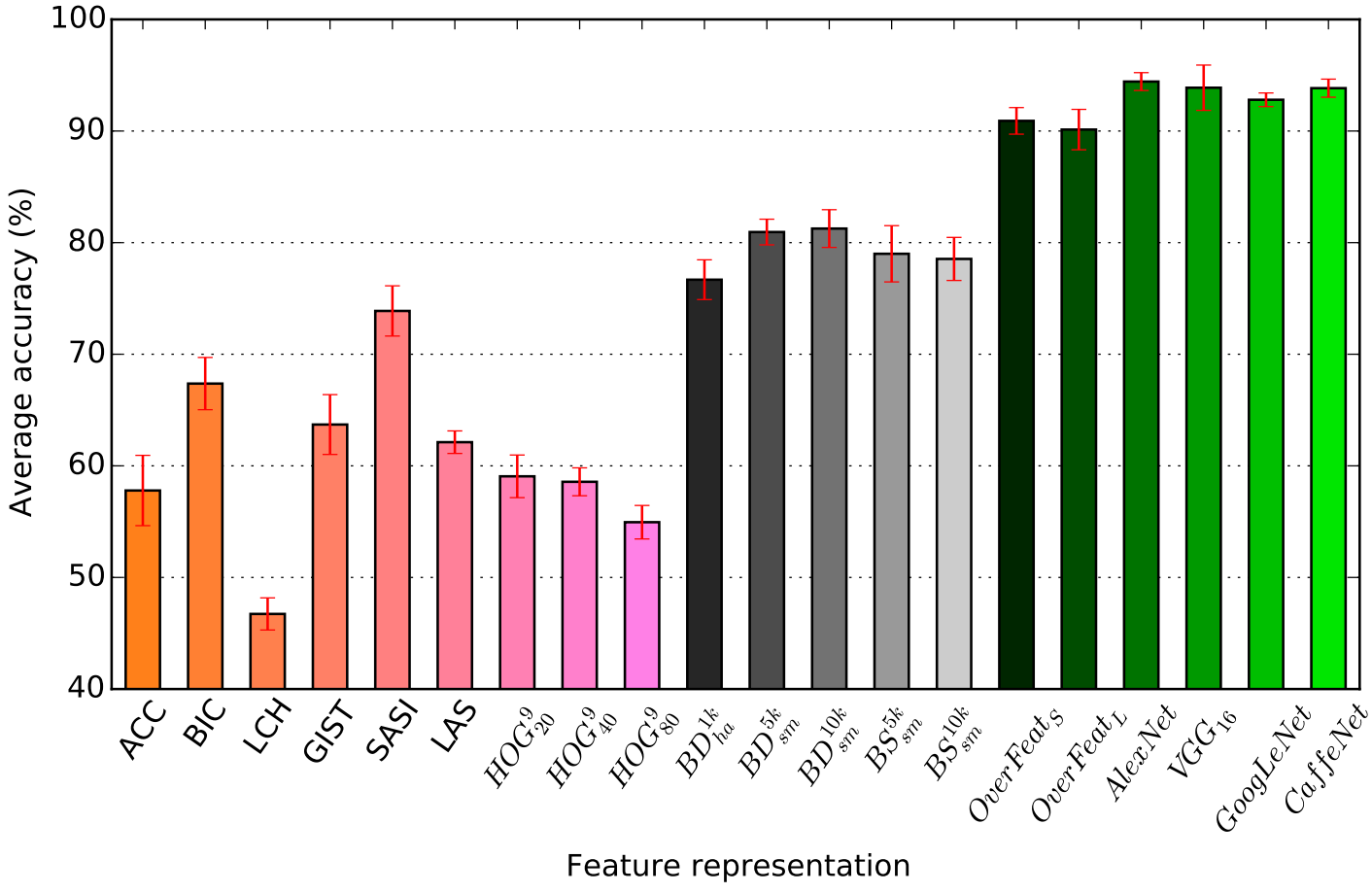
(e) Coffee

(f) Non-coffee

Some Experiments

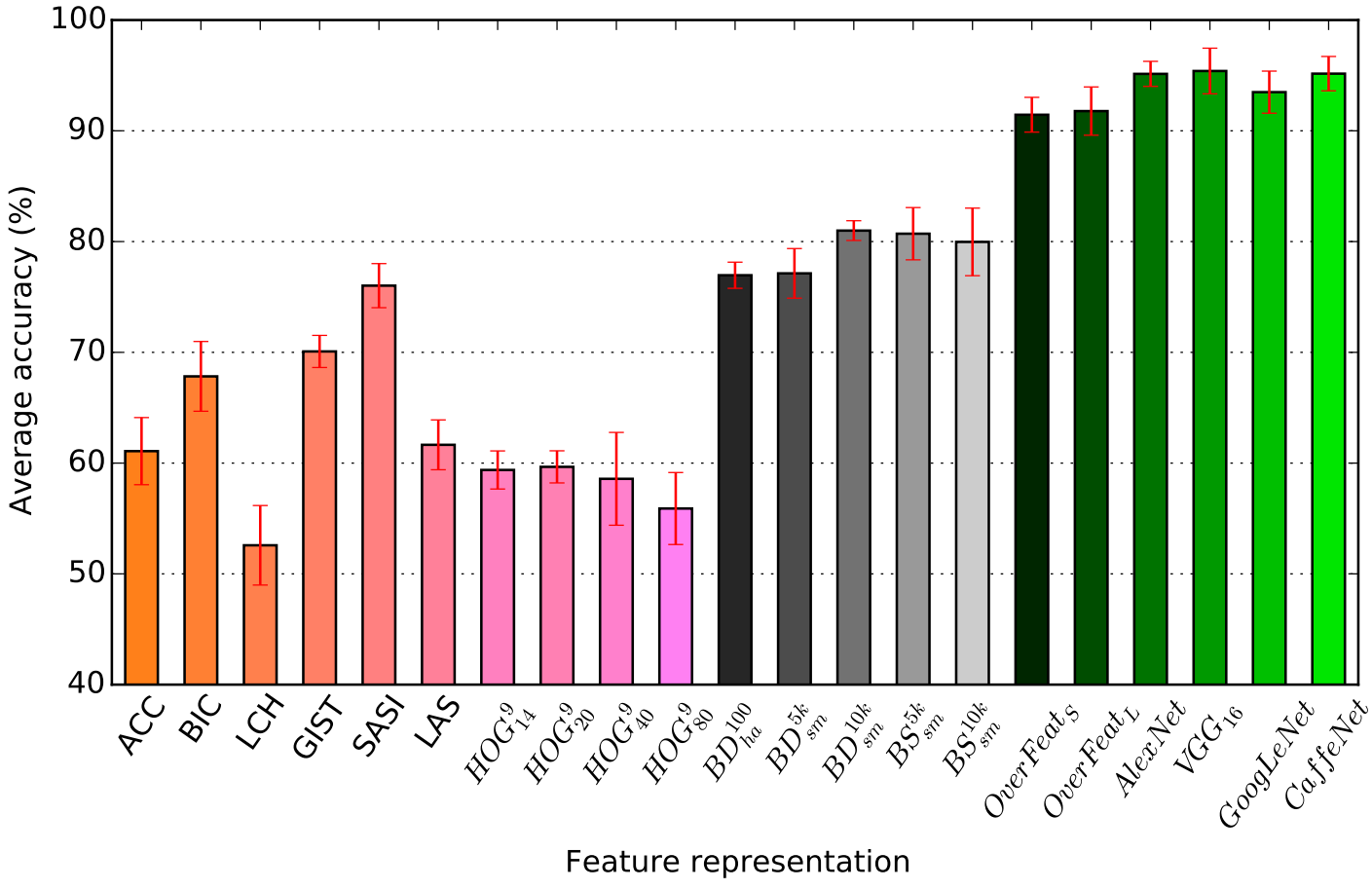
- A - Generalization Power Evaluation
- B - Comparison of ConvNets Strategies
- C - Comparison with Baselines

A - Generalization Power Evaluation



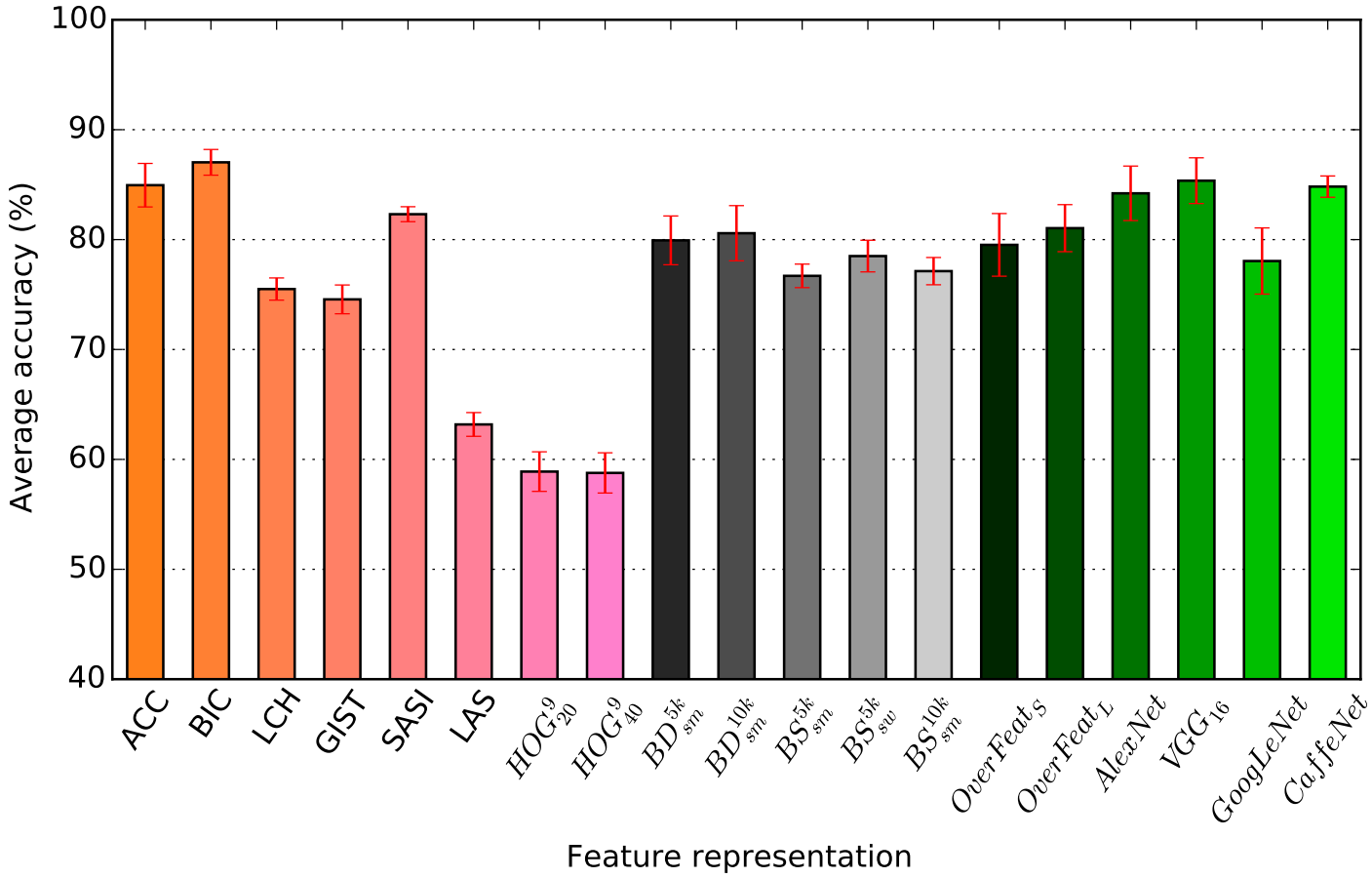
UCMerced Land-use Dataset

A - Generalization Power Evaluation



RS19 Dataset

A - Generalization Power Evaluation



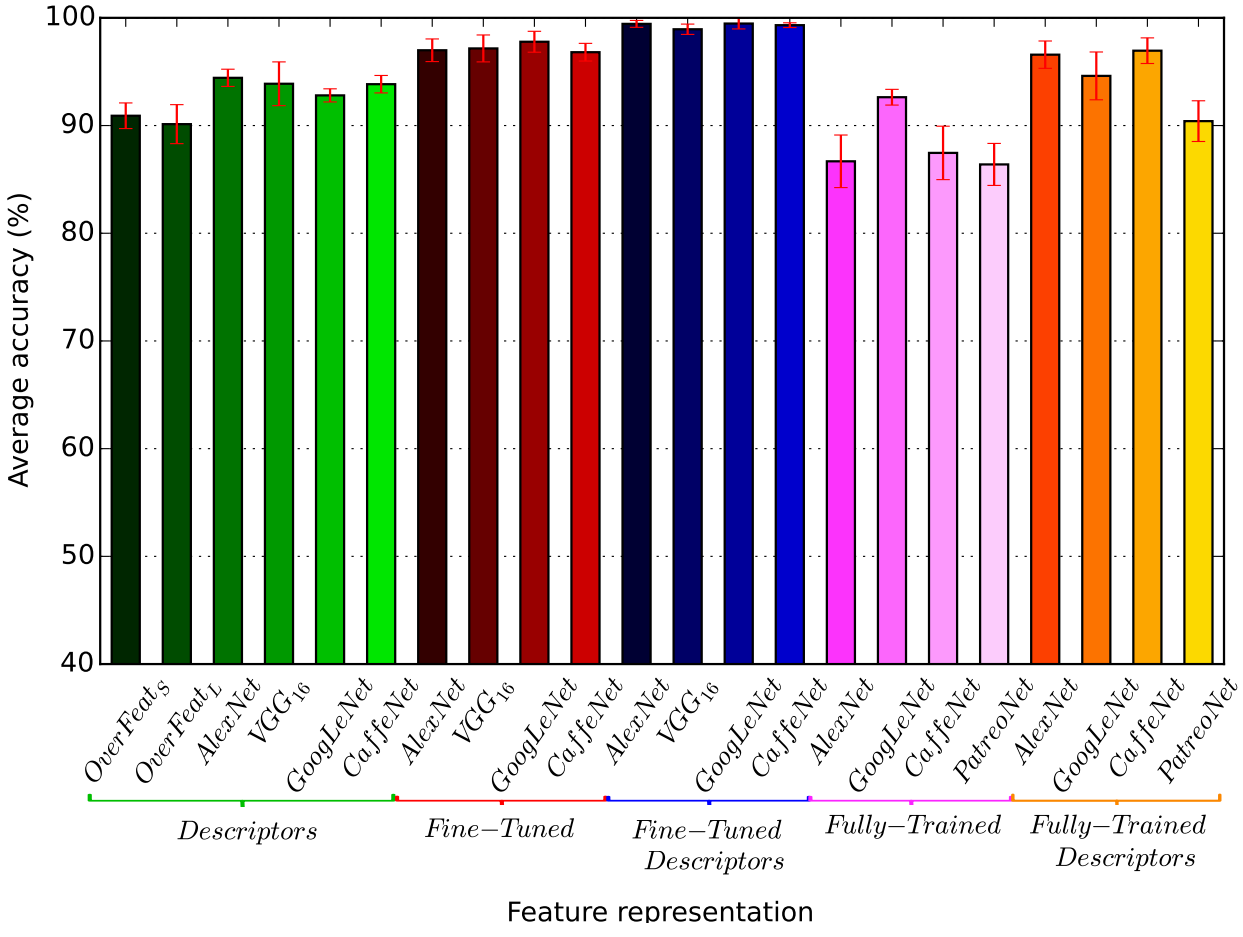
Brazilian Coffee Scenes Dataset

A - Generalization Power Evaluation

Conclusions

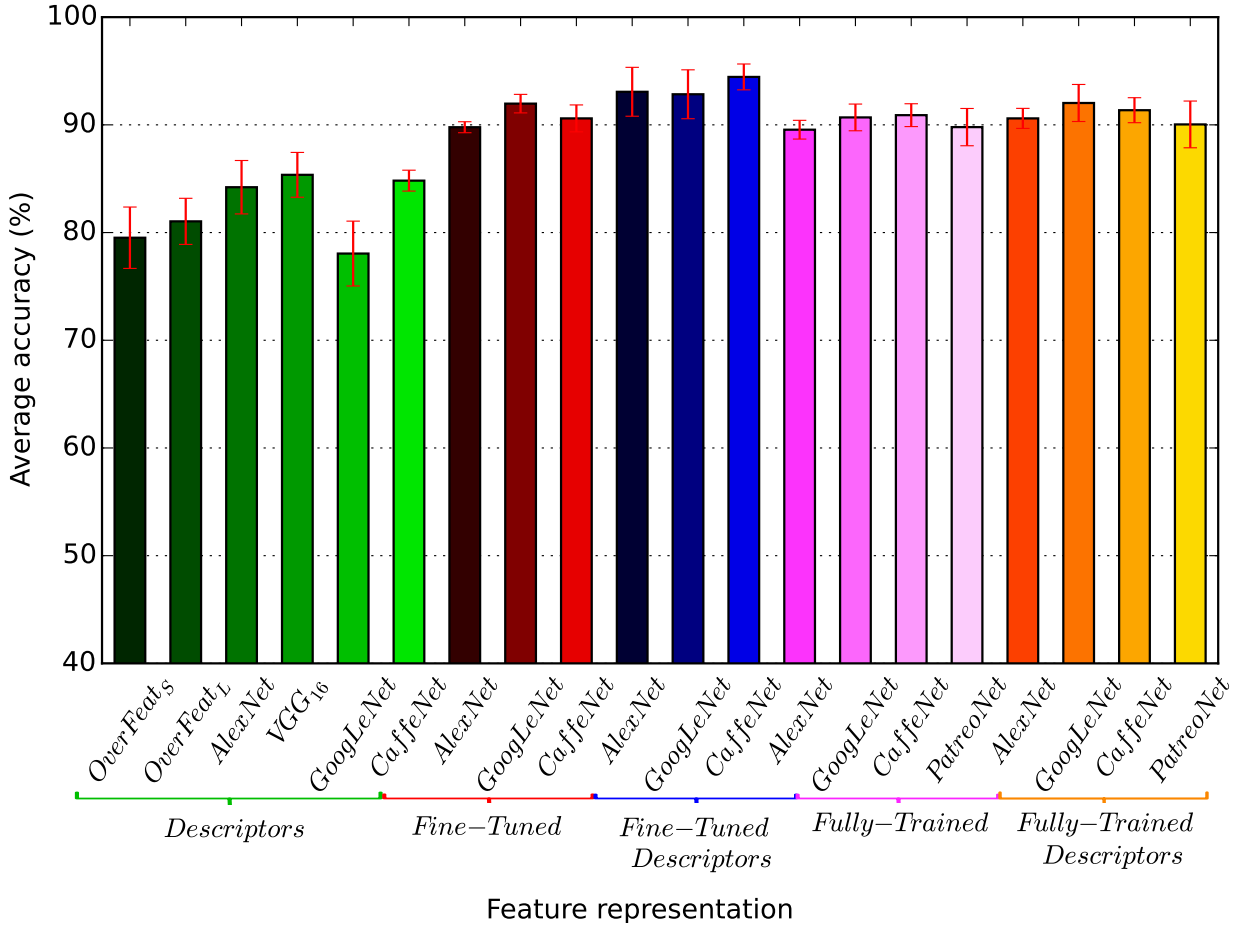
- It is possible to exploit feature representation learned in computer vision datasets into the remote sensing scenario
  - Deep features generalize better to aerial dataset than to agricultural ones
  - Agricultural images are composed of finer and more homogeneous textures/color

B - Comparison of ConvNets Strategies



UCMerced Land-use Dataset

B - Comparison of ConvNets Strategies

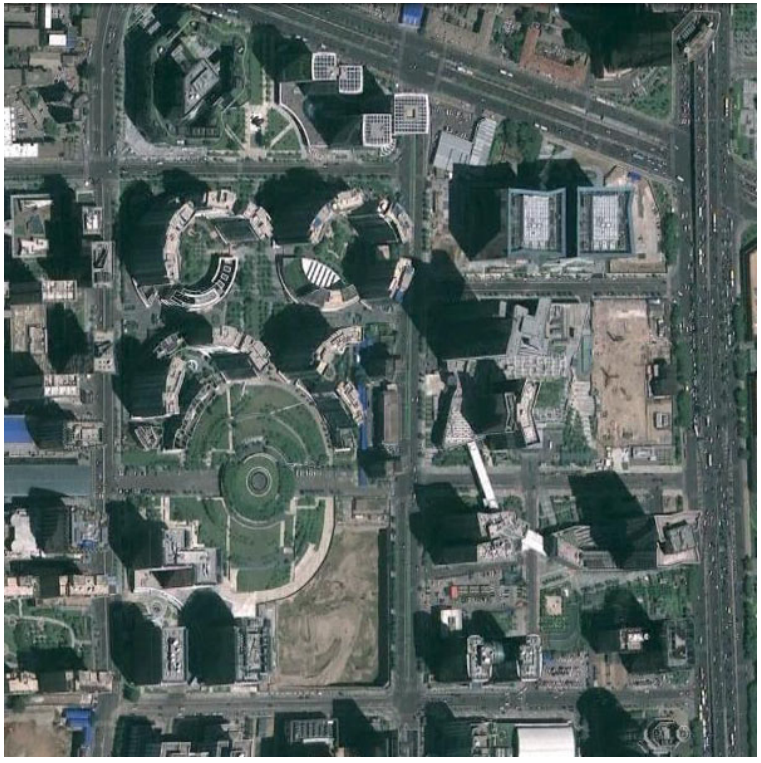


Brazilian Coffee Scenes Dataset

B - Comparison of ConvNets Strategies

Medium Residential $\xrightarrow[\text{into}]{\text{missclassified}}$ Dense Residential

B - Comparison of ConvNets Strategies



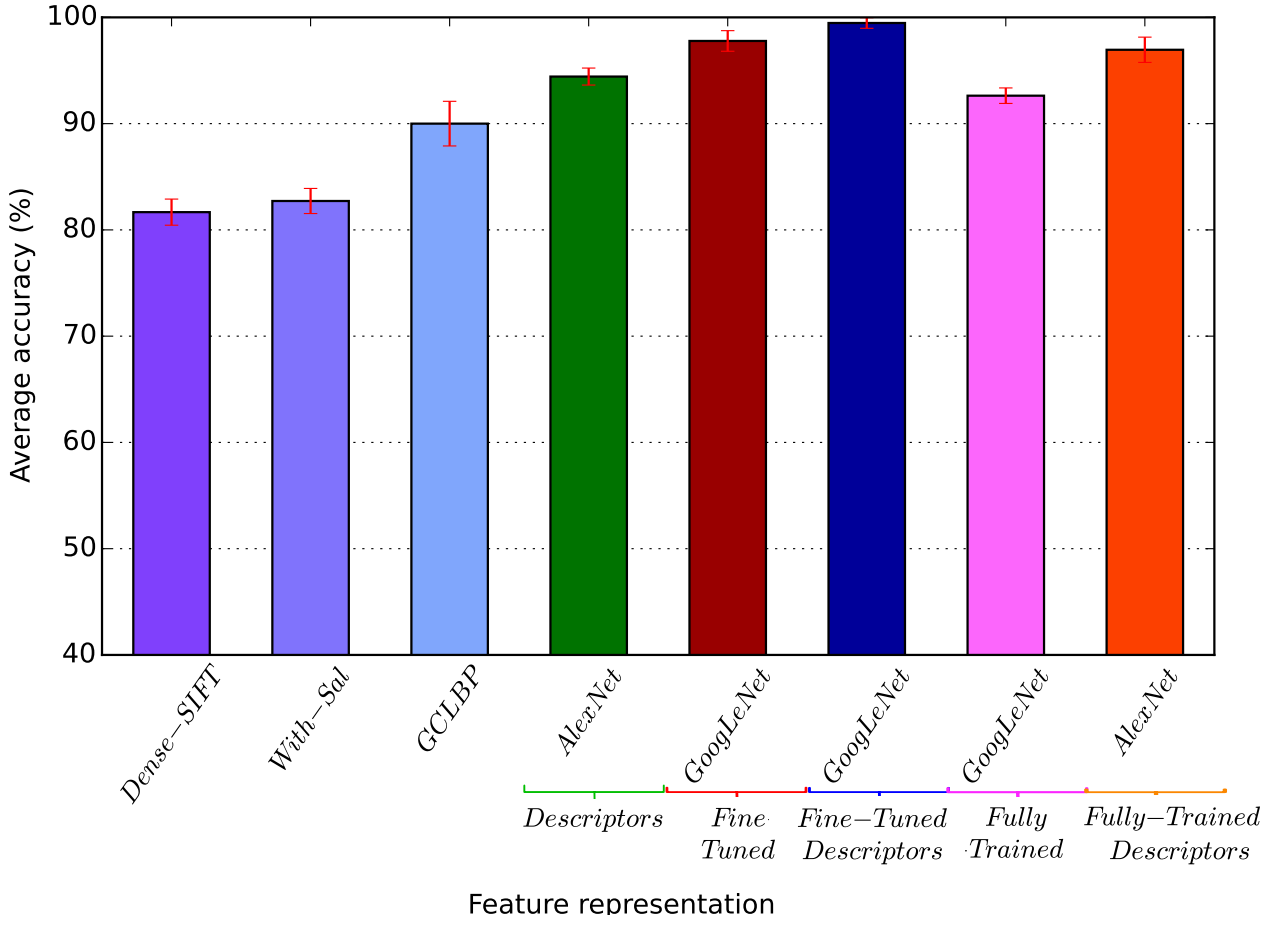
Commercial $\xrightarrow[\text{into}]{\text{missclassified}}$ Park

B - Comparison of ConvNets Strategies

Conclusions

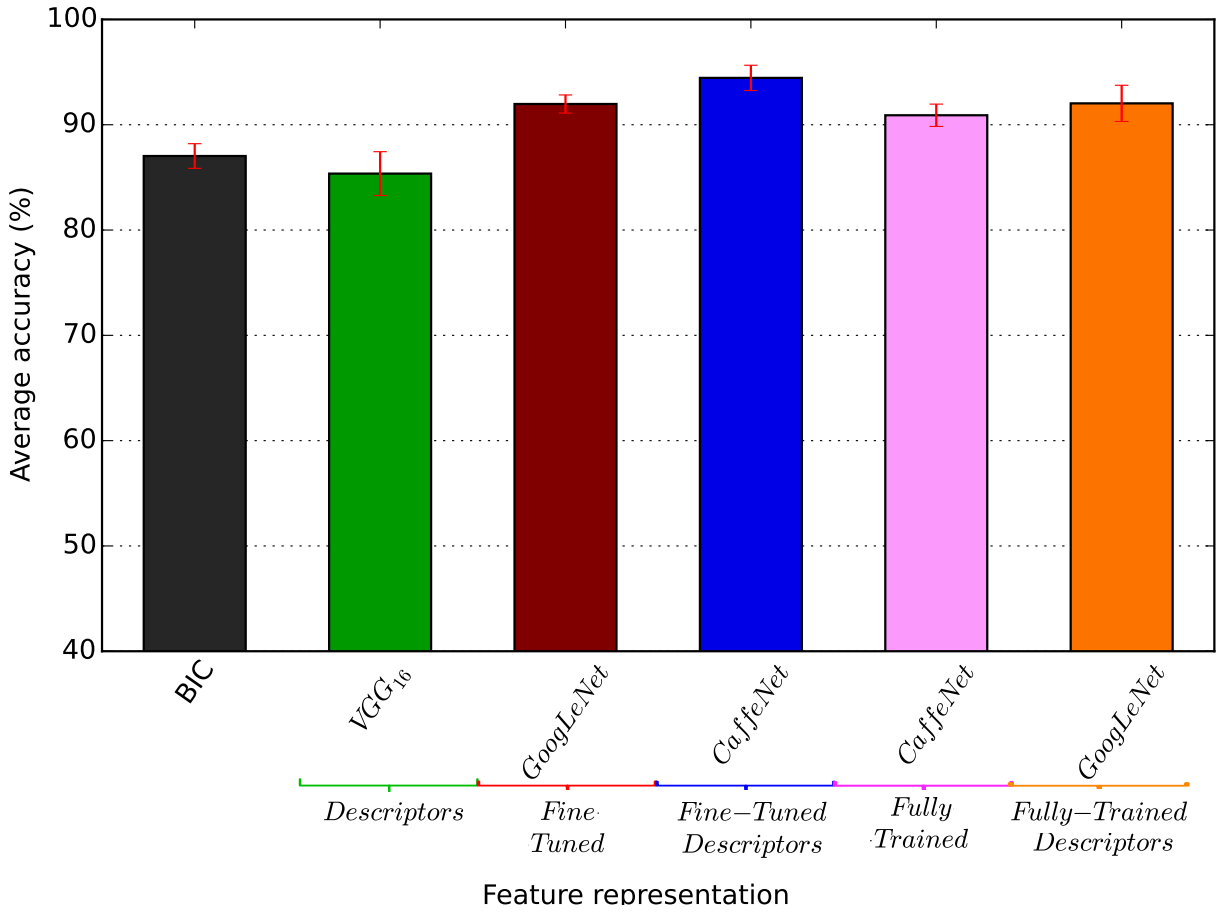
- Feature representation learned in everyday image datasets can be adjusted to the remote sensing domain
  - Fully training was not a good strategy maybe due to the small amount of labeled data available
  - Fine tuning is usually the best strategy
  - Replacing the last softmax layer by SVM was a better solution

C - Comparison with Baselines



UCMerced Land-use Dataset

C - Comparison with Baselines



Brazilian Coffee Scenes Dataset