Leveraging past knowledge with deep transfer learning strategies

Zaharah Bukhsh

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Outline

- Typical machine learning systems
- Quick introduction to Transfer Learning
- Relevance of Transfer learning to practice
- Applications of Transfer learning (2)
- Summary and outlook

EINDHOVEN UNIVERSITY OF TECHNOLOGY

How typical machine learning systems learn?

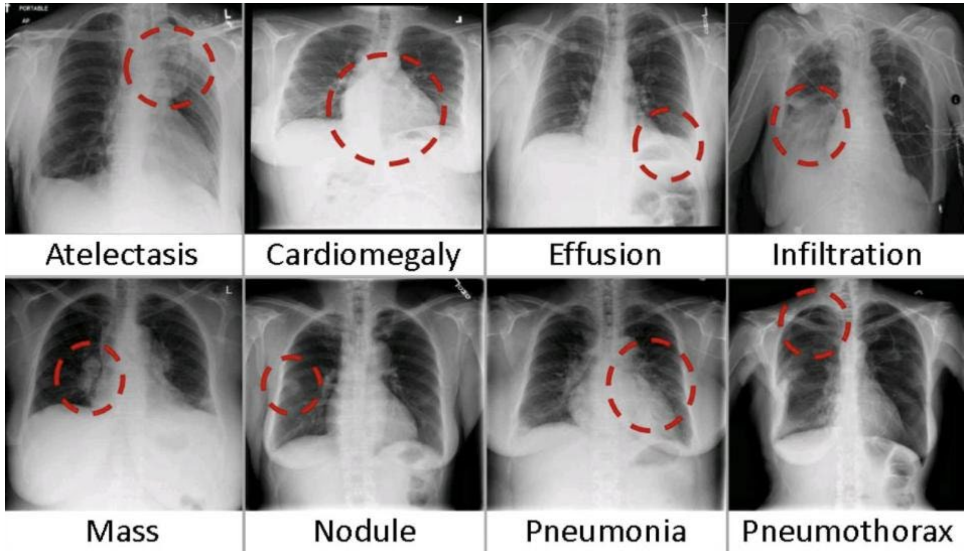
- Isolated single learning tasks
  - Learning is performed for individual tasks
  - Knowledge is not retained or accumulated

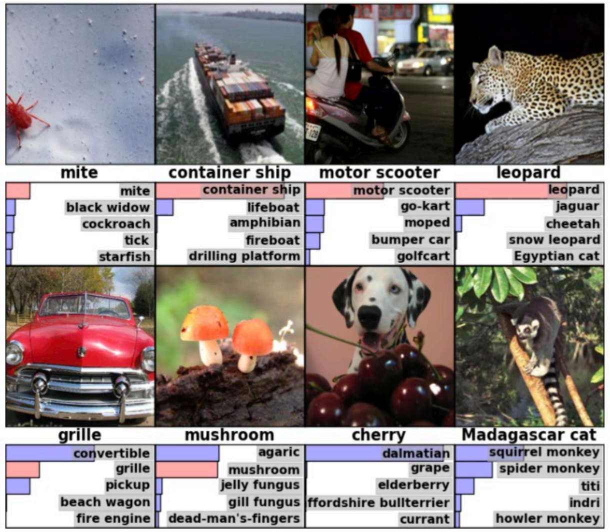
Dataset 1 → Learning system for task 1

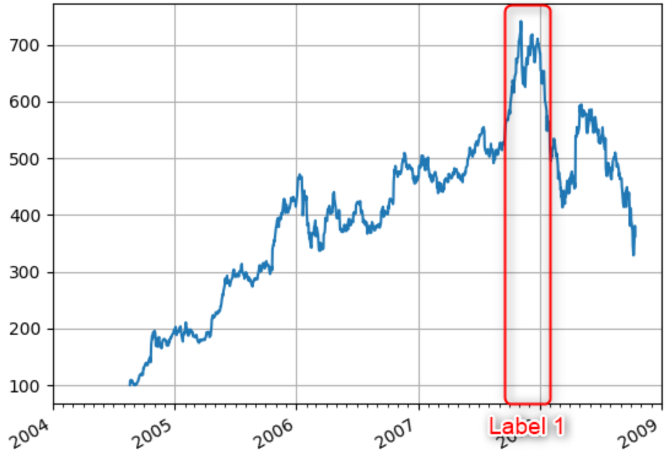
Dataset 2 → Learning system for task 2

Rarest animals

Labeled data?







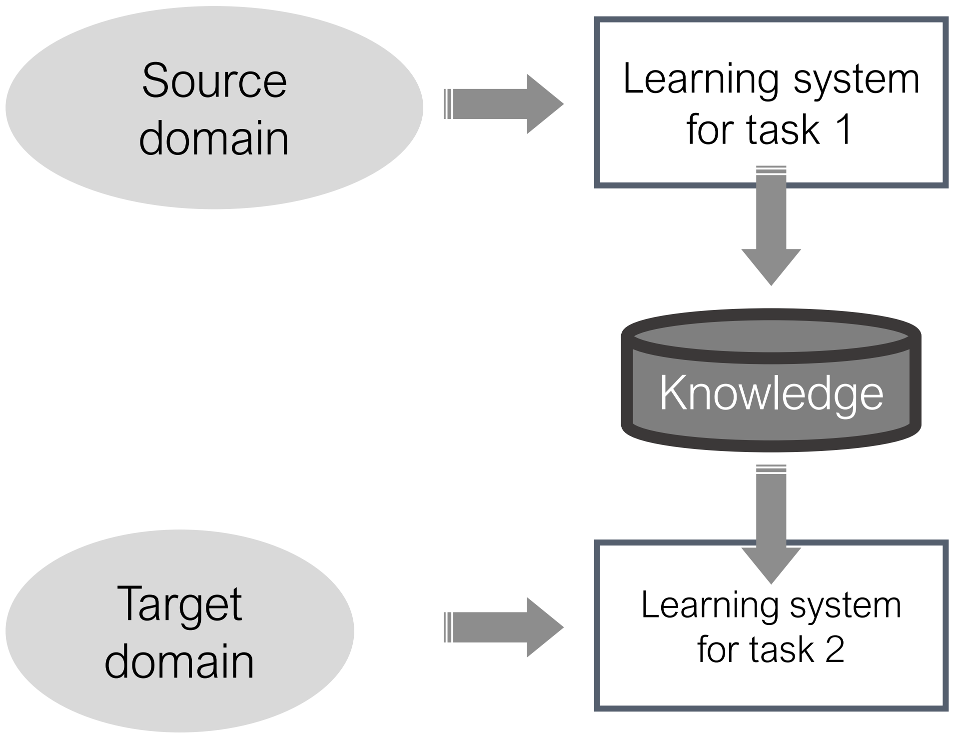
Major limitation of deep learning

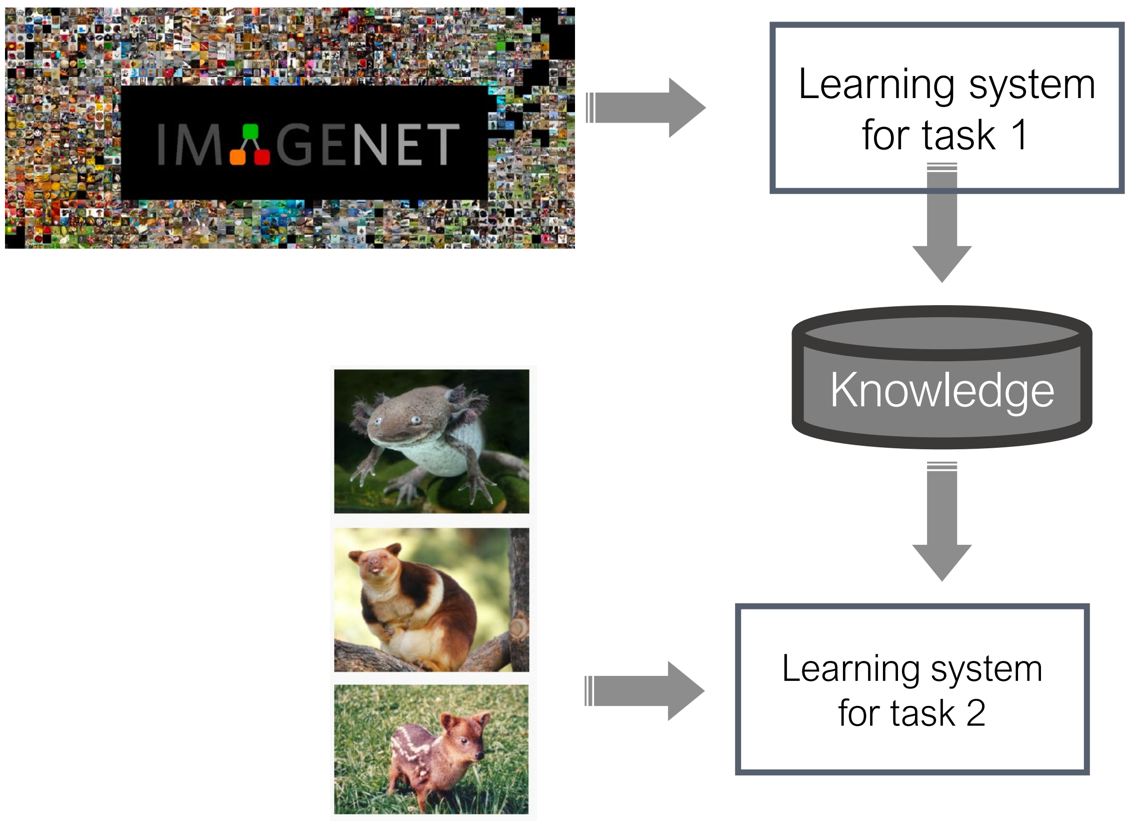
- Not data efficient: Learning requires millions of labeled examples

  models do not generalize well to new domains; not like humans!

Credit: Kate Saenko talk, Boston University

Transfer learning to the rescue

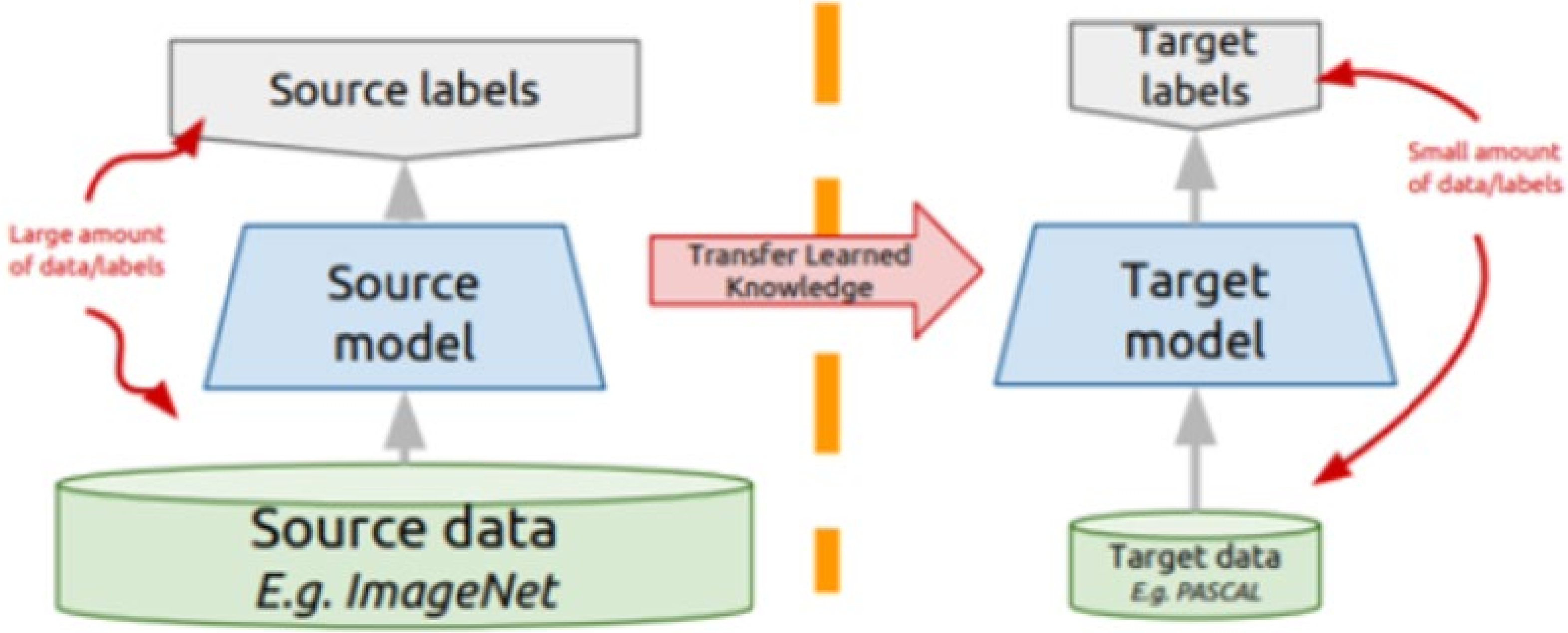




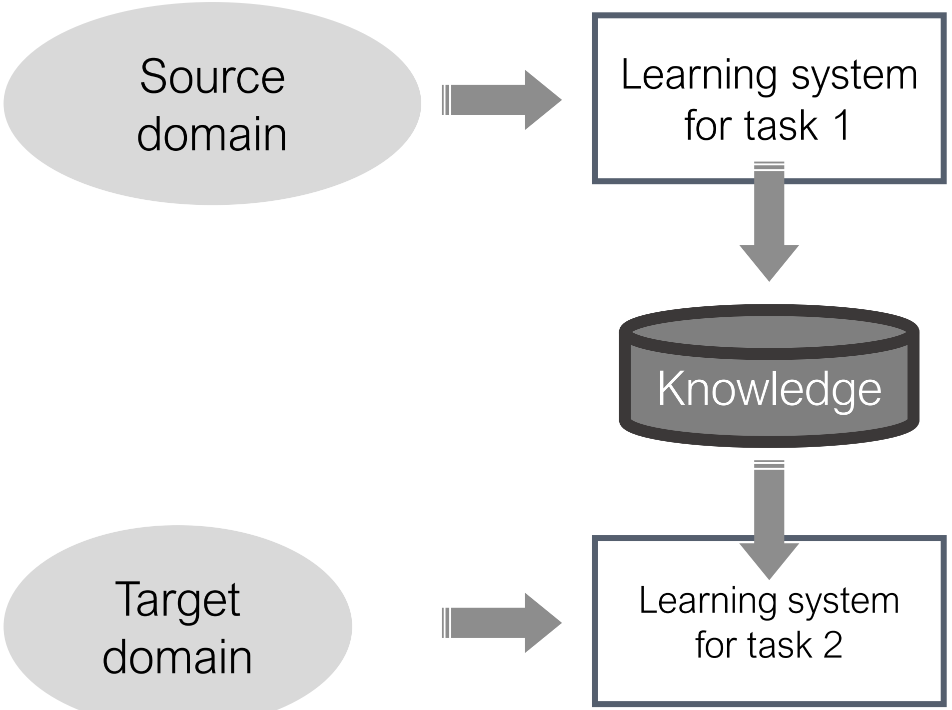
Transfer learning to the rescue

Instead of training a network from scratch:
- Take a network trained on a different domain for a different **source task**
- Adapt it for your domain and your **target task**

Variations
- Same domain, different task.
- Different domain, same task.



Transfer learning to the rescue



- Learning of new task relies on previously learned task
- Does not assume that future data must be in the same feature space and have the same distribution
- Learning process can be faster, more accurate and/or need less training data

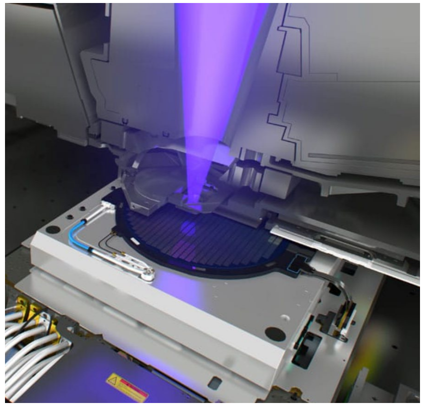
> *Situation where what has been learned in one setting is exploited to improve generalization in another setting.*

Goodfellow et al, Deep Learning book

Not a new concept! Goes by many names e.g. Learning to Learn, Knowledge Consolidation, Domain Adaption

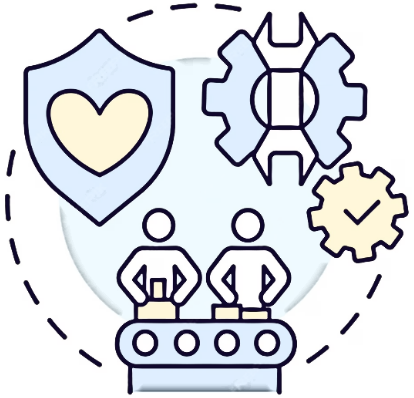
Relevance of transfer learning in practice

- Data efficiency
  - Small datasets
  - Limited annotation
- Model (training) efficiency
- Utility across domains
- Faster development cycles
- Embedded and edge development



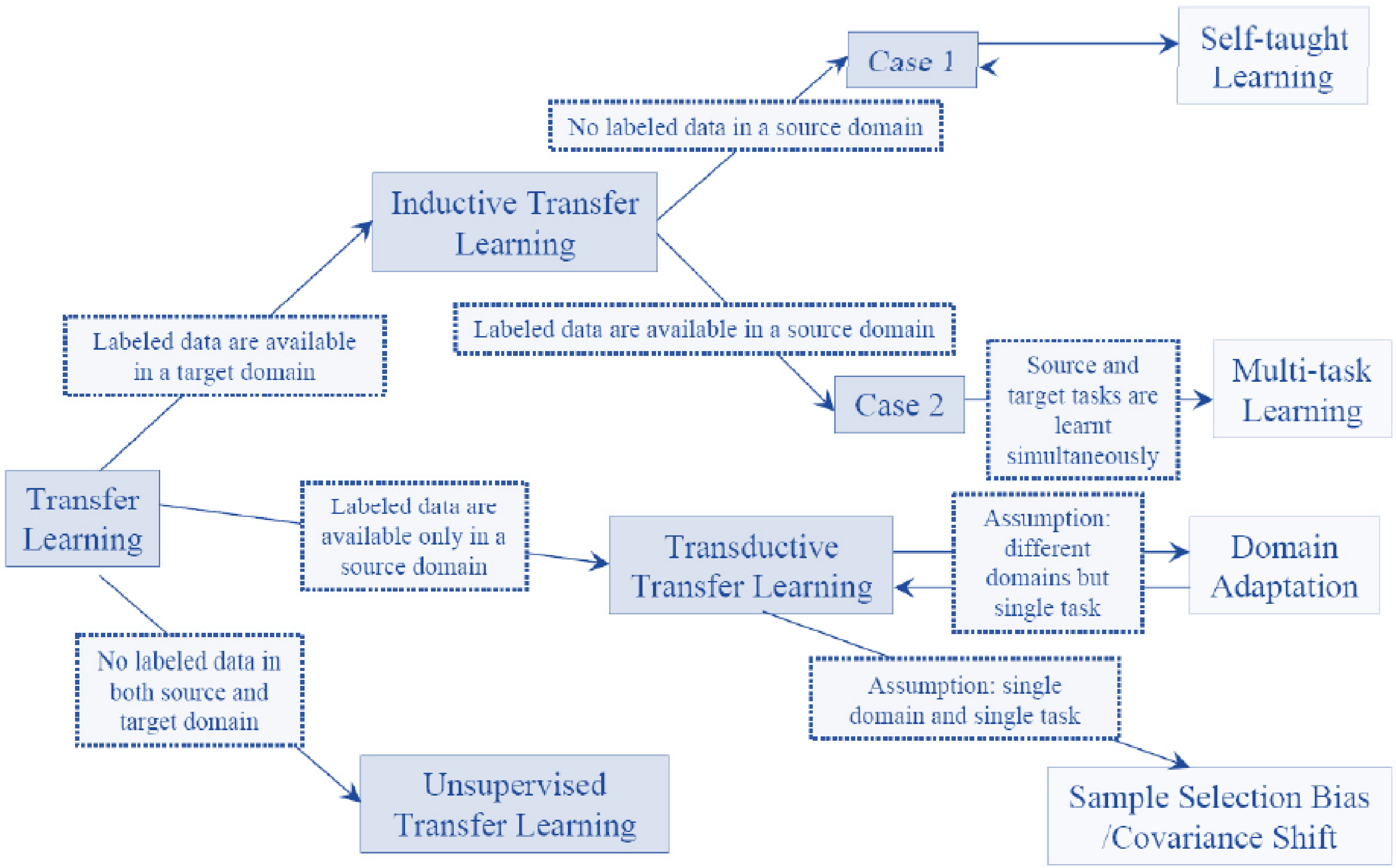
New component

New product



Operating condition

Different transfer strategies



Source: Pan and Yang, 2010

Transfer learning applications

In-domain and Cross-domain Transferability for Damage Detection
Which source dataset provides most generic features?

Remaining useful lifetime prediction via deep domain adaptation
How to learn from labeled source data for unlabeled target tasks?

Transfer learning applications

In-domain and Cross-domain Transferability for Damage Detection
Which source dataset provides most generic features?

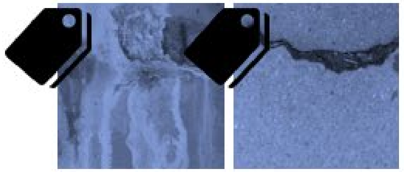
Bukhsh, Z.A., Jansen, N. & Saeed, A. Damage detection using in-domain and cross-domain transfer learning. *Neural Comput & Applic* **33**, 16921–16936 (2021). https://doi.org/10.1007/s00521-021-06279-x

Remaining useful lifetime prediction via deep domain adaptation
How to learn from labeled source data for unlabeled target tasks?

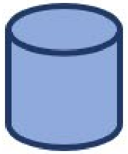
In-domain and Cross-domain TL
Background and objective

Visual inspection of several bridges

Massive visual data

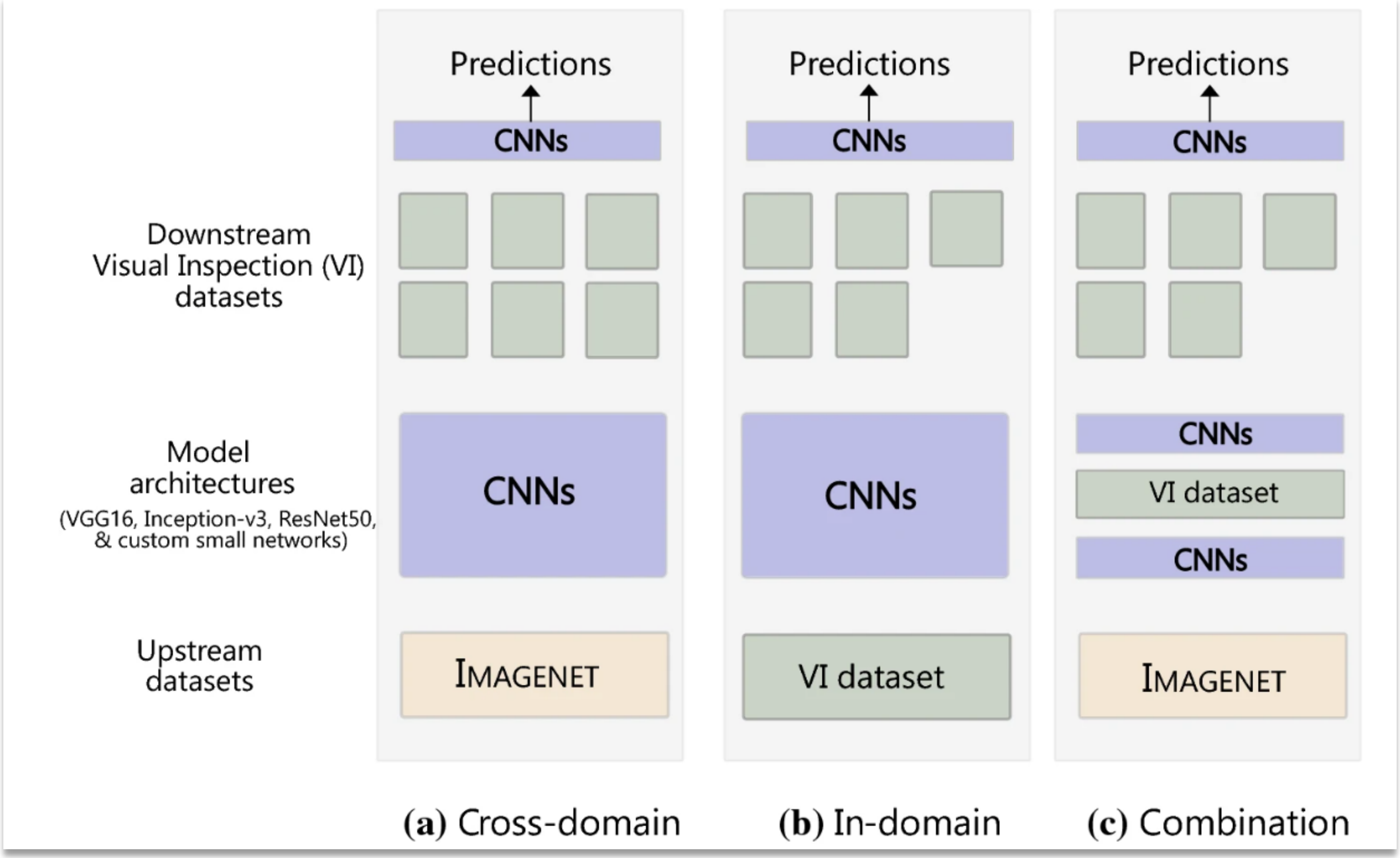


Manual damage identification

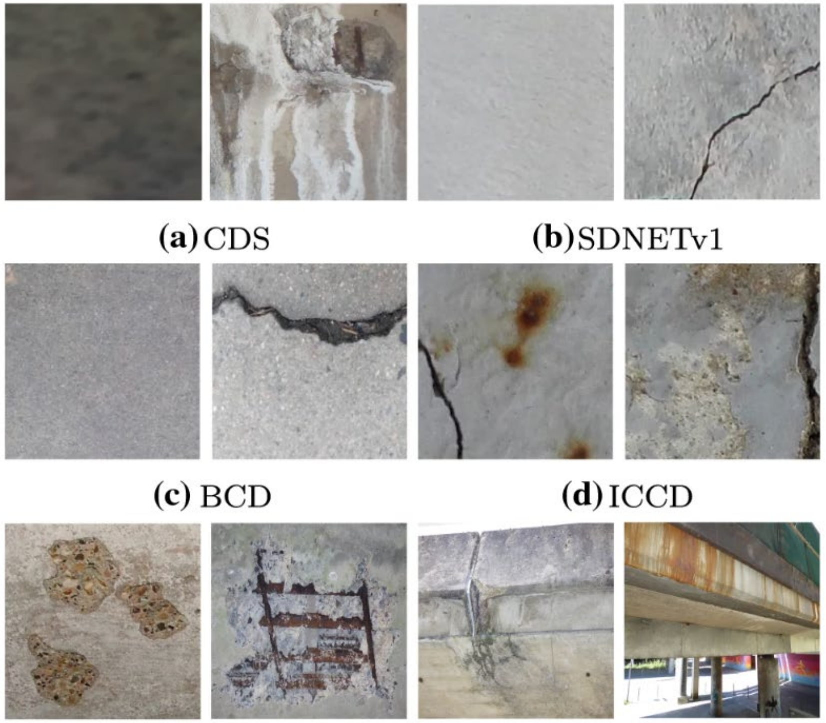


Small-size carefully labelled data

In-domain and Cross-domain TL
Methodology



In-domain and Cross-domain TL
Datasets



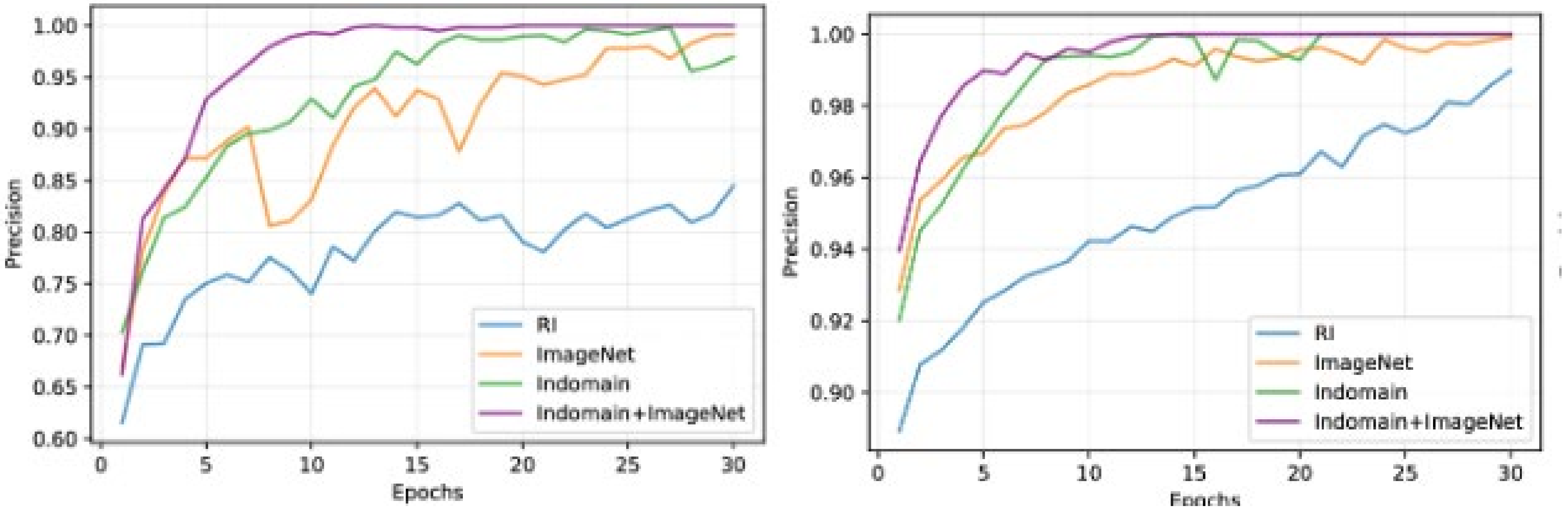
(a) CDS (b) SDNETv1 (c) BCD (d) ICCD (e) MCDS (f) CODEBRIM

Examples from the bridge inspection datasets. From left to right: (a) complex lighting with no crack, Crack with efflorescence, and exposed bars. (b) Intact concrete, minor crack, (c) Intact concrete, large crack on deck, (d) crack with rust stains, crack with minor scaling, (e) concrete scaling, corrosion with exposed bars, (f) spallation and efflorescence, rust stains

Table 2 Overview of (bridge) visual inspection datasets

Dataset	Instances	Classes	Problem
CDS [26]	1,027	2	Binary
SDNETv1 [11]	13,620	2	Binary
BCD [67]	5,390	2	Binary
ICCD [37]	60,010	2	Binary
MCDS [28]	2,411	10	Multi-label
CODEBRIM [42]	8,304	6	Multi-label

	CDS	SDNETv1	BCD	ICCD	MCDS	CODEBRIM
	Random Initialization					
CBR Tiny	0.79 ± 0.02	0.74 ± 0.01	0.97 ± 0.00	0.93 ± 0.13	0.67 ± 0.02	0.80 ± 0.04
CBR Small	0.78 ± 0.02	0.72 ± 0.01	0.96 ± 0.0	0.98 ± 0.0	0.67 ± 0.02	0.80 ± 0.03
CBR LargeW	0.75 ± 0.03	0.74 ± 0.01	0.97 ± 0.01	0.96 ± 0.02	0.67 ± 0.01	0.80 ± 0.04
CBR LargeT						
VGG16						
Inception-v3						
ResNet50						
VGG16						
Inception-v3						
ResNet50						
VGG16						
Inception-v3						
ResNet50						

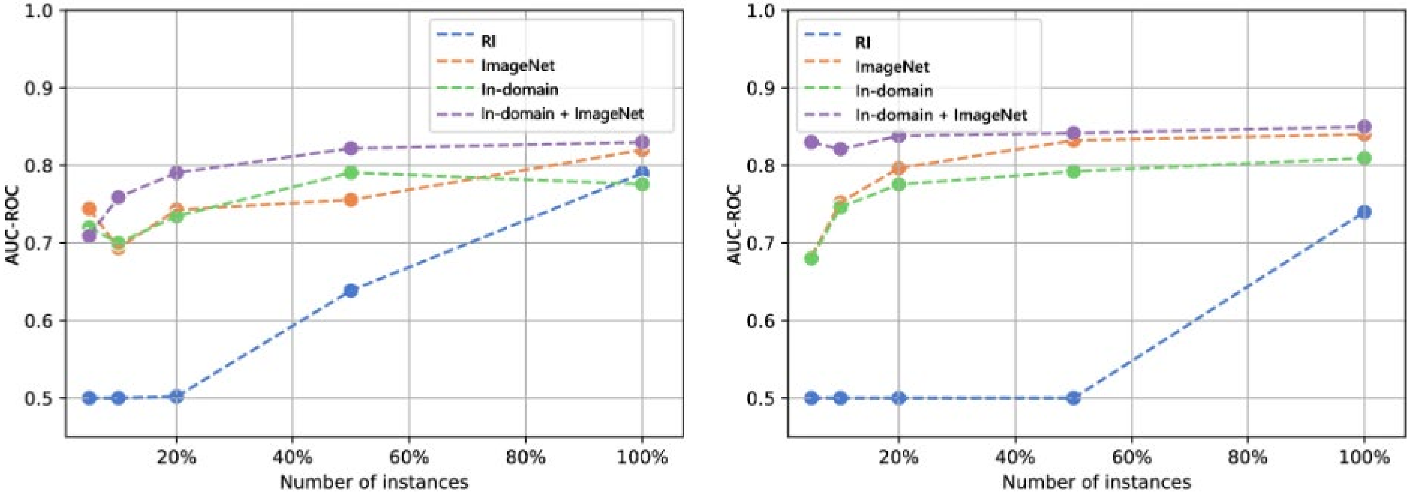


(a) CDS (b) SDNETv1

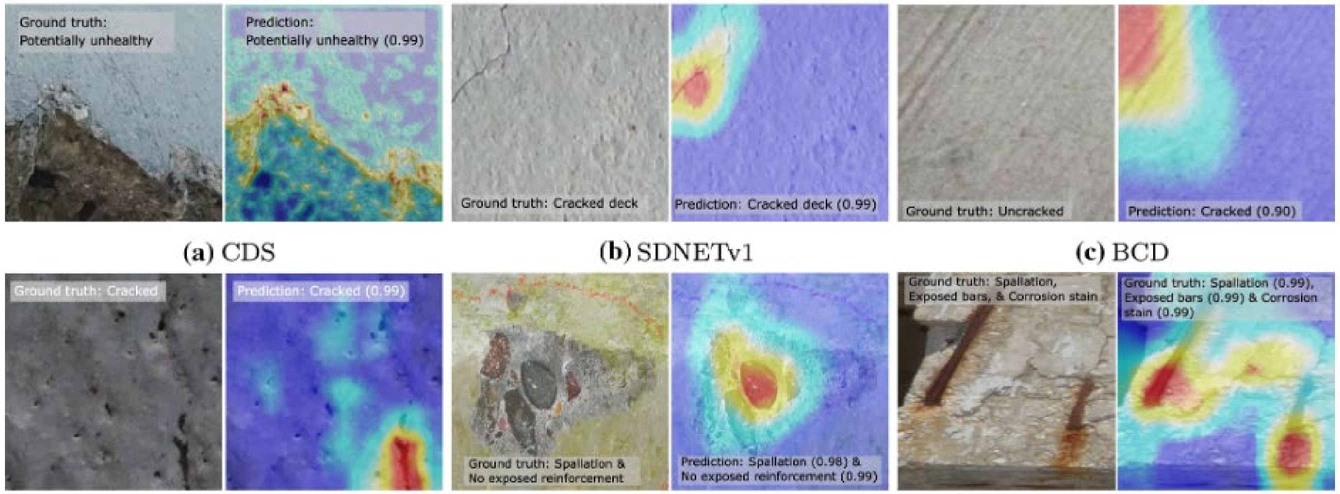
In-domain and Cross-domain TL Results

In-domain and Cross-domain TL

Results



(a) CDS **(b)** SDNETv1



(a) CDS **(b)** SDNETv1 **(c)** BCD

(d) ICCD **(e)** MCDS **(f)** CODEBRIM

Visual explanations for classification and localization of VGG16 models for all six datasets. Grad-CAM highlights class discriminative features. The heatmap localizes the specific damage regions

Transfer learning applications

In-domain and Cross-domain Transferability for Damage Detection
Which source dataset provides most generic features?

Remaining useful lifetime prediction via deep domain adaptation
How to learn from labeled source data for unlabeled target tasks?

da Costa, P. R. D. O., Akçay, A., Zhang, Y., & Kaymak, U. (2020). Remaining useful lifetime prediction via deep domain adaptation. *Reliability Engineering & System Safety*, *195*, 106682.

Deep domain adaptation
Background and objective

- Remaining Useful Lifetime Predictions can be made when enough run-to-failure data exists.
- How can we predict failures for similar components without observing failure information?

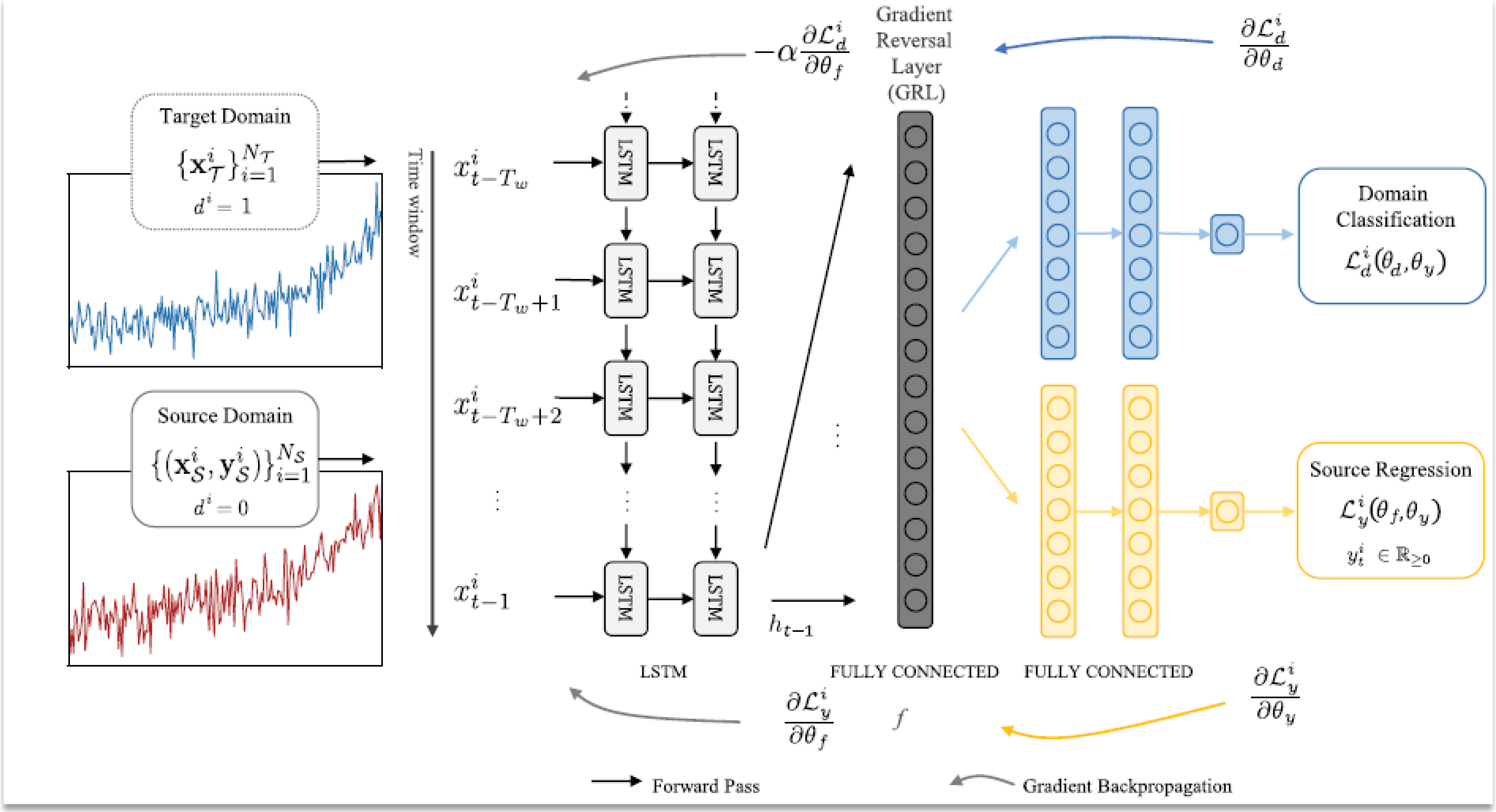
Cessna 152.

Cessna Skylane

Adapt from a labelled source domain to an unlabeled target domain serving as a component under different operating conditions with no observed failures.

Deep domain adaptation
LSTM-DANN

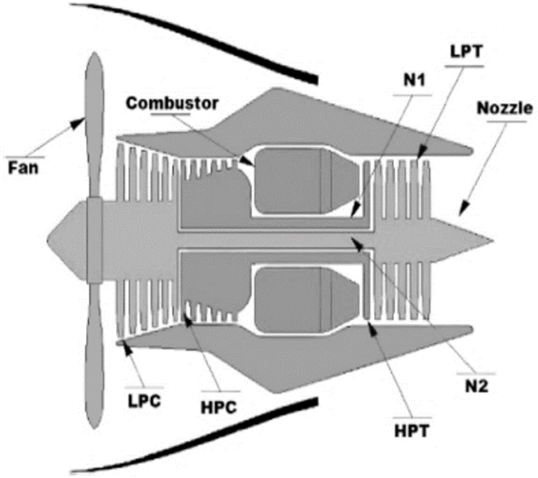
- Shared LSTM feature extractor
- Two feature mapping: Source Regression loss (yellow), Domain Classifier loss (blue)



Deep domain adaptation
C-MAPSS Datasets

Characteristics
- Turbofan engines
- Run-to-failure (Degradation) timeseries
- Multiple sensors (columns) per engine

Reasons to be used as surrogate datasets
- Data come from an aero-engine simulator
- Varying working conditions
- PHM08 (Proven to be a good prognostics dataset)

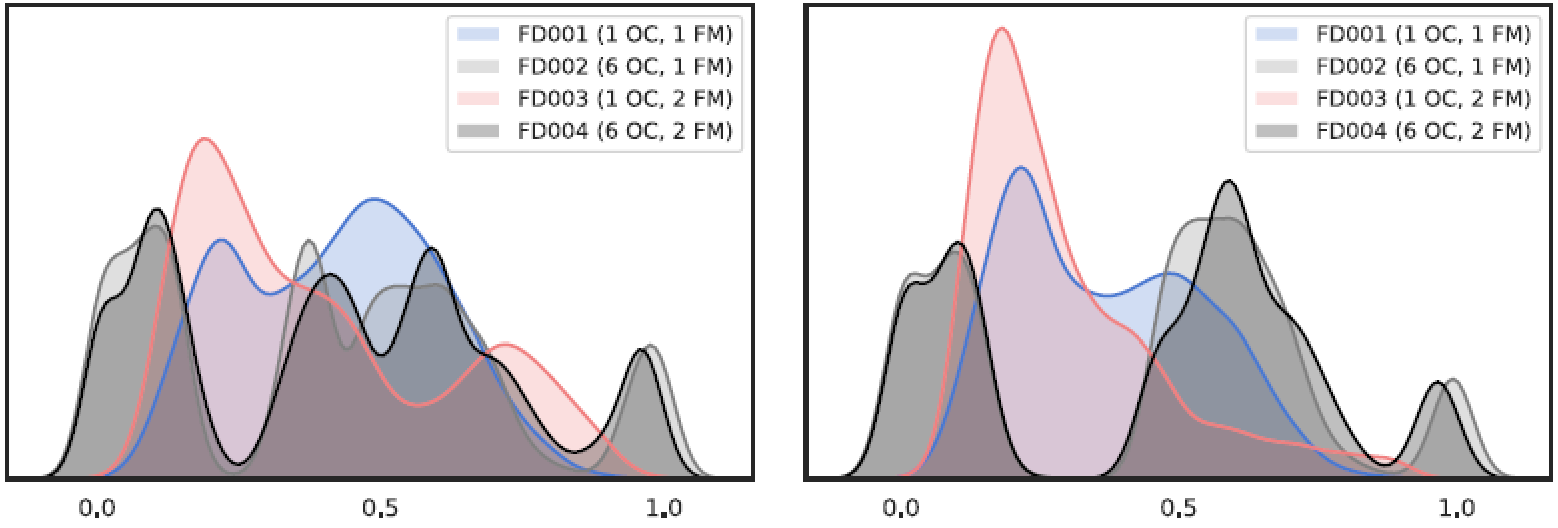


LPC: Low-pressure compressor
HPC High-pressure compressor
LPT: Low-pressure turbine
HPT: High-pressure turbine
N1: Turbine axis
N2: Turbine shaft

Dataset	FD001	FD002	FD003	FD004
Number of fault modes	1	1	2	2
Number of operating conditions	1	6	1	6
Number of training units	100	260	100	249
Number of testing units	100	259	100	248

Deep domain adaptation
C-MAPSS Datasets

Dataset	FD001	FD002	FD003	FD004
Number of fault modes	1	1	2	2
Number of operating conditions	1	6	1	6
Number of training units	100	260	100	249
Number of testing units	100	259	100	248



(a) Sensor 15 (b) Sensor 7

Fig. 4. Distribution of normalised sensor values between 100 and one time step before a failure. Sensor distributions are more similar between FD001 and FD003 and FD002 and FD004 pairs due to identical operating conditions.

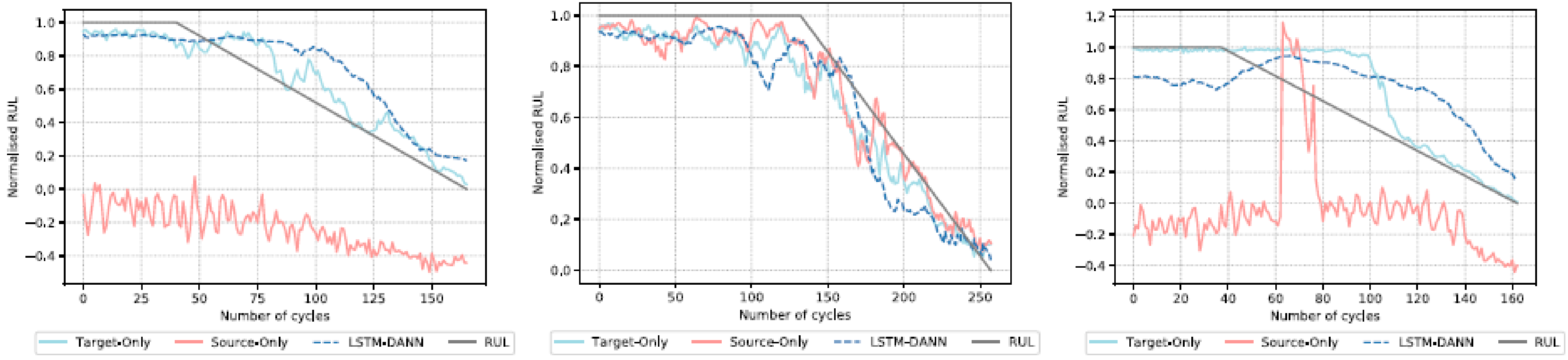
Results

Source: FD001	Target	SOURCE-ONLY	LSTM-DANN (Δ%)	TARGET-ONLY
-	FD002	71.70 ± 3.88	**48.62** (-32%) ± 6.83	17.76 ± 0.43
-	FD003	51.20 ± 3.39	**45.87** (-10%) ± 3.58	12.49 ± 0.29
-	FD004	73.88 ± 4.50	**43.82** (-41%) ± 4.15	21.30 ± 1.06
Source: FD002	Target	SOURCE-ONLY	LSTM-DANN (Δ%)	TARGET-ONLY
-	FD001	164.84 ± 23.00	**28.10** (-83%) ± 5.03	13.64 ± 0.80
-	FD003	154.04 ± 21.79	**37.46** (-76%) ± 1.54	12.49 ± 0.29
-	FD004	37.76 ± 2.17	**31.85** (-16%) ± 1.65	21.30 ± 1.06
Source: FD003	Target	SOURCE-ONLY	LSTM-DANN (Δ%)	TARGET-ONLY
-	FD001	49.94 ± 7.65	**31.74** (-36%) ± 9.37	13.64 ± 0.80
-	FD002	70.32 ± 4.02	**44.62** (-36%) ± 1.21	17.76 ± 0.43
-	FD004	69.28 ± 4.51	**47.94** (-31%) ± 5.78	21.30 ± 1.06
Source: FD004	Target	SOURCE-ONLY	LSTM-DANN (Δ%)	TARGET-ONLY
-	FD001	188.00 ± 25.95	**31.54** (-83%) ± 2.42	13.64 ± 0.80
-	FD002	**20.88** ± 1.66	24.93 (+19%) ± 1.82	17.76 ± 0.43
-	FD003	157.32 ± 20.37	**27.84** (-82%) ± 2.69	12.49 ± 0.29

Table 2: RMSE ± Standard Deviation - Comparison between SOURCE-ONLY, TARGET-ONLY and LSTM-DANN on the test datasets.

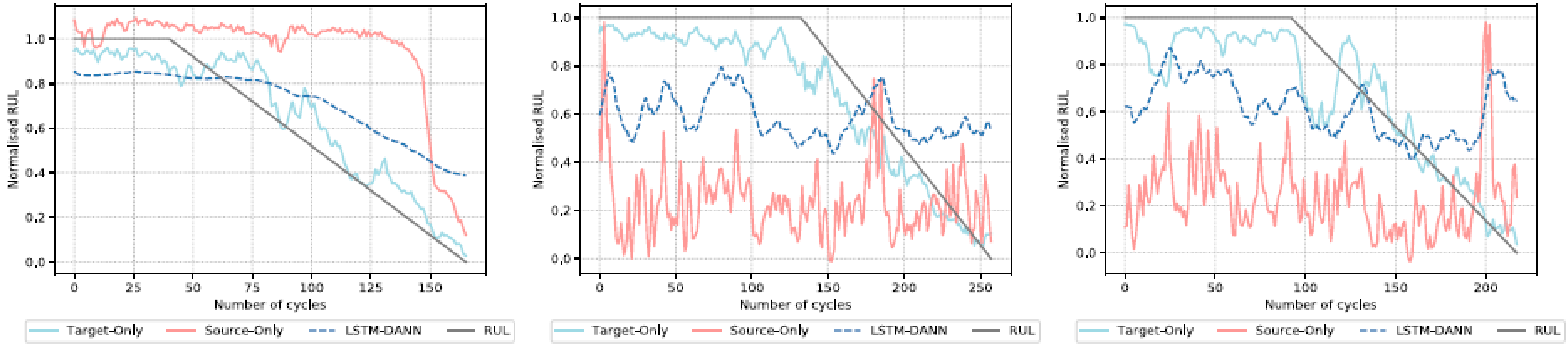
Results

Source: FD004



(a) Target: FD001 (b) Target: FD002 (c) Target: FD003

Source: FD003



(d) Target: FD001 (e) Target: FD002 (f) Target: FD004

Summary and outlook

- With different transfer learning strategies, deep models can:
  - learn from source datasets without (or limited) labels.
  - learn for target domain tasks without (or limited) labels.
  - Relevant for practice to adapt trained model to new fault types, operating conditions, and so on.

- Foundational models are becoming excellent tools for enabling domain-specific transfer learning towards practical applications.