~~Continual Learning &~~
Domain Generalisation

COMPSCI 760
2024 Semester 1

Olivier Graffeuille

ogra439@aucklanduni.ac.nz

Domain Generalisation

Outline

1. Overview & Definitions
2. Techniques for Domain Generalisation
    1. Data-Driven approaches
    2. Representation Learning approaches
    3. Learning-Strategy approaches

Outline

1. ***Overview & Definitions***
2. Techniques for Domain Generalisation
    1. Data-Driven approaches
    2. Representation Learning approaches
    3. Learning-Strategy approaches

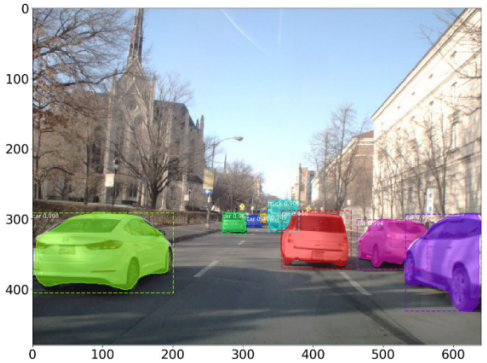
Motivation

1. Often we want our model to work in settings it was not trained for
    1. Mostly computer vision:
        1. Self-driving cars in uncommon weather conditions
        2. Person identification
    2. But also:
        1. Time series forecasting
        2. Speech utterance recognition

Motivation

- Humans are good at this, but models often fail!

"Domain Shift"

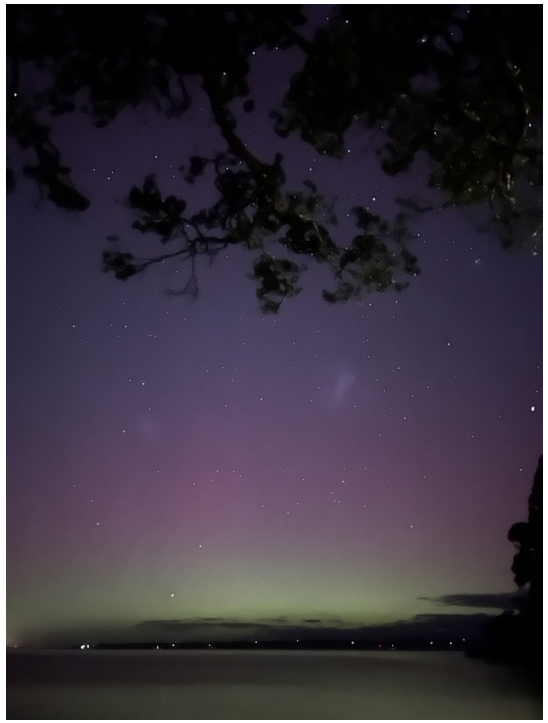


?

Fu, C., Hu, P., Dong, C., Mertz, C., & Dolan, J. M. (2018, November). Camera-based semantic enhanced vehicle segmentation for planar lidar. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (pp. 3805-3810). IEEE.

Motivation

- Getting rich data from EVERY domain is generally impossible
  - Setting: urban, rural, wild
  - Time: day time, night time, sunset
  - Weather: sunny, foggy, Aurora Australis…?



Domain Generalisation

The goal:

To learn a model from one or more related domains (i.e., diverse training datasets) that will generalize well on *unseen* testing domains

Motivation

This is different to continual learning!

1. We are not re-training a deployed model
2. Only focused on domains, not tasks. **Domain:**

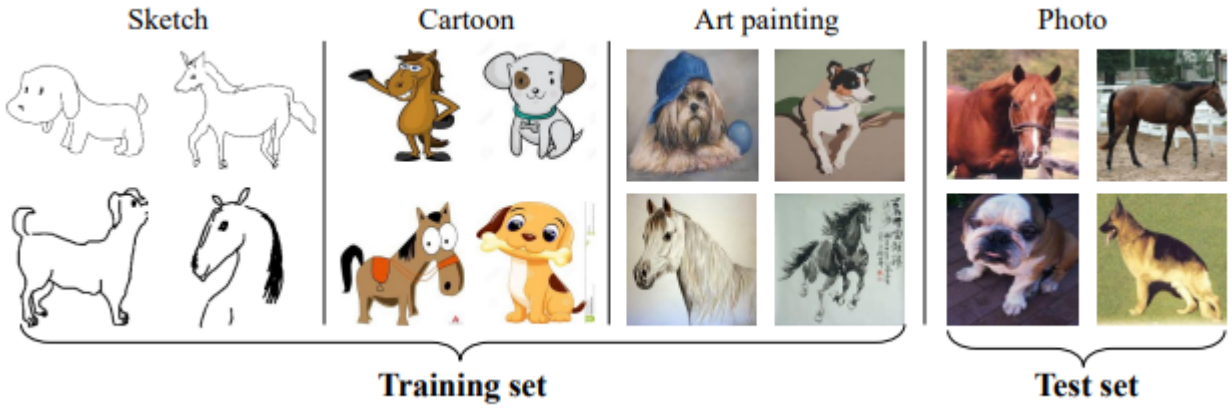
   A feature space and marginal probability density function. $D = \{\mathcal{X}, p(X)\}$

   E.g. Sunny photos of pets, painting of pets, audio recording of pet.

Is there such thing as "Task Generalisation"?

Domain Generalisation

In practice:



Note: the task is still the same!

Wang, J., Lan, C., Liu, C., Ouyang, Y., Qin, T., Lu, W., ... & Yu, P. (2022). Generalizing to unseen domains: A survey on domain generalization. IEEE Transactions on Knowledge and Data Engineering.

Domain Generalisation

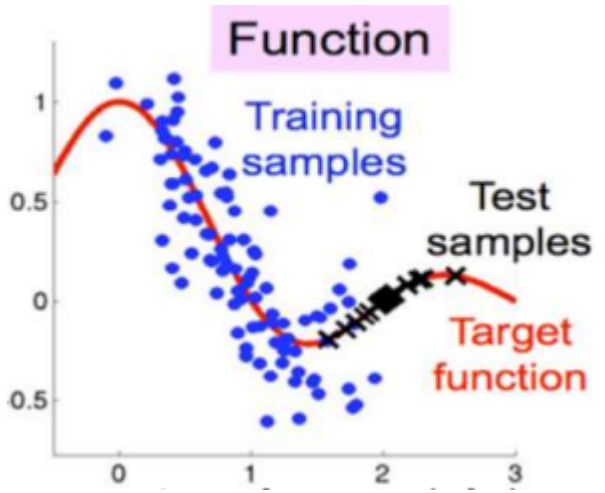
1. Domain Generalisation:
    1. We are given **multiple** source domains
    2. We must generalise to **unseen** domains
2. Domain Adaptation:
    1. We are given **one** source domain
    2. We must generalise to **known** domains (seen during training)

Outline

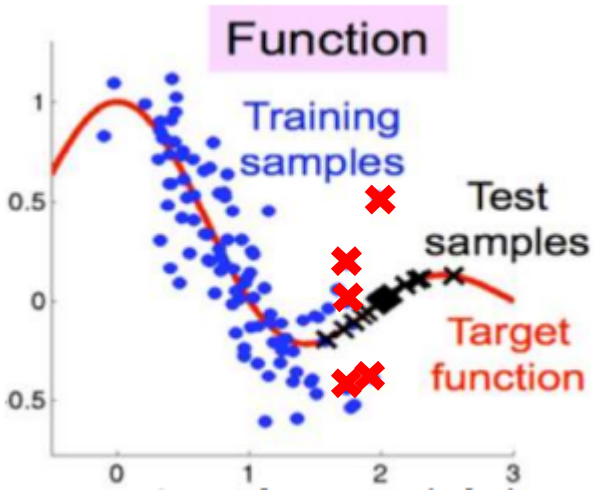
1. Overview & Definitions
2. Techniques for Domain Generalisation
    1. ***Data-Driven approaches***
    2. Representation Learning approaches
    3. Learning-Strategy approaches

Domain Generalisation

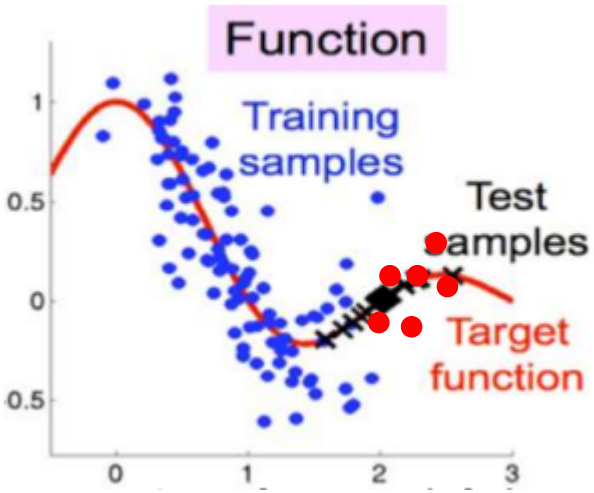
1. In a simple regression task:



Original Problem



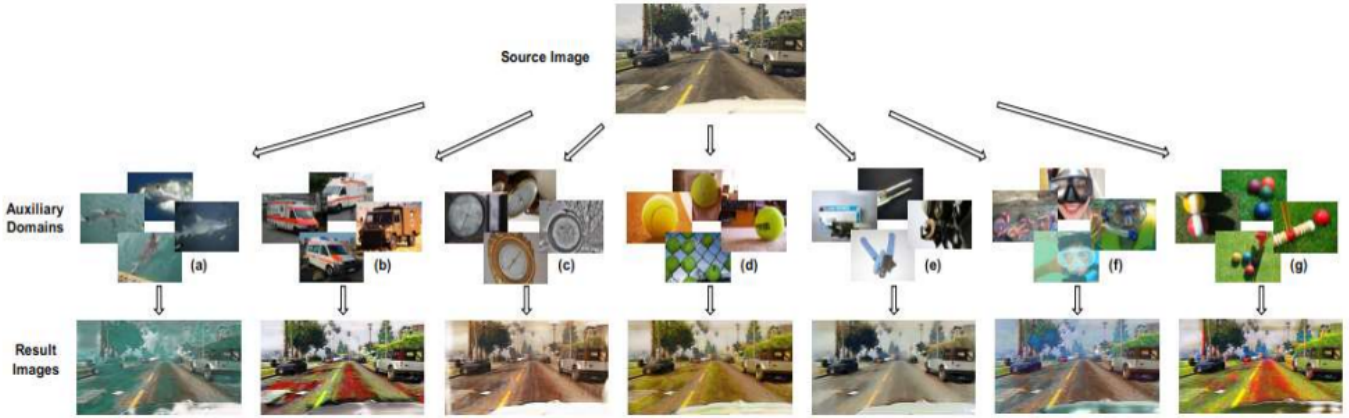
Data Re-weighting



Data Generation

Data Augmentation

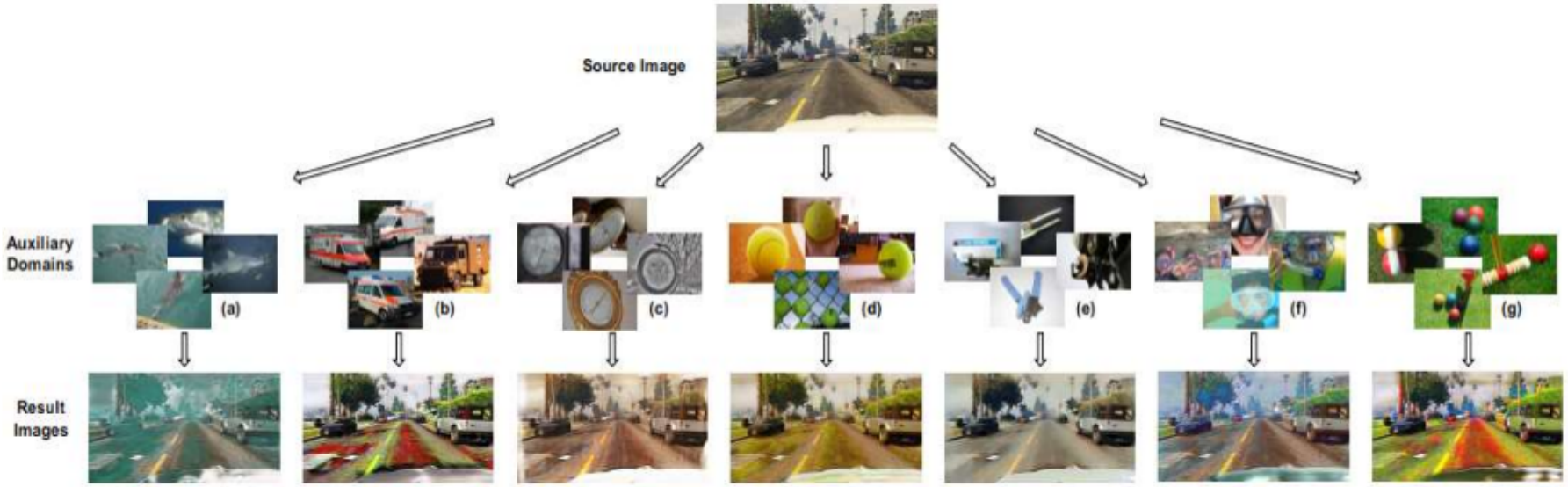
- Methods of augmenting input data in some way to imitate different domains, to improve generalisability
  - Typical augmentation: rotate, add noise…
  - In continual learning: change the domain (style, setting, etc)



Yue et al. Domain Randomization and Pyramid Consistency: Simulation-to-Real Generalization without Accessing Target Domain Data. ICCV,2019.

Data Augmentation

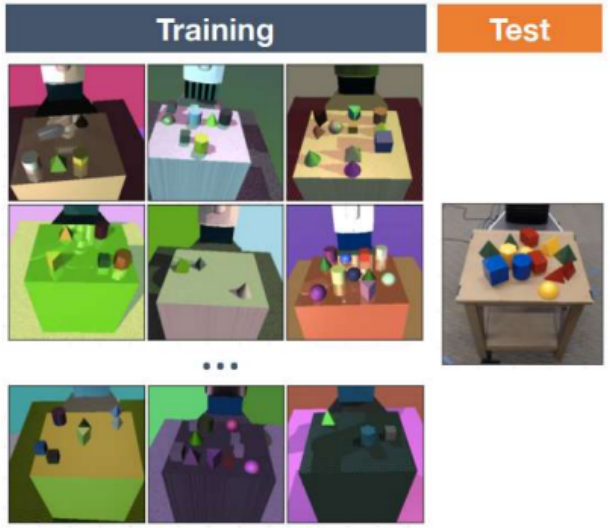
- Methods of augmenting input data in some way to imitate different domains, to improve generalisability

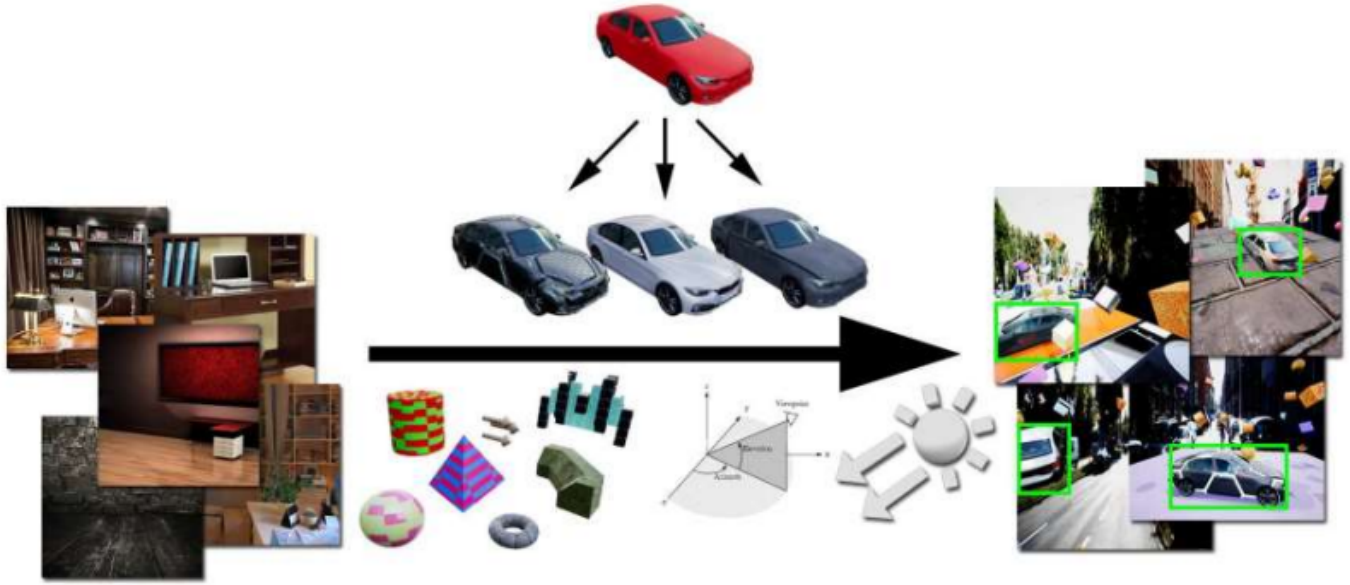


Yue et al. Domain Randomization and Pyramid Consistency: Simulation-to-Real Generalization without Accessing Target Domain Data. ICCV,2019.

Data Randomisation

- Use simulated environments to replicate real problems
- Randomise the simulation to generate diverse domains





Tobin, et al. Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World. IROS 2017.

Structured Data Randomisation

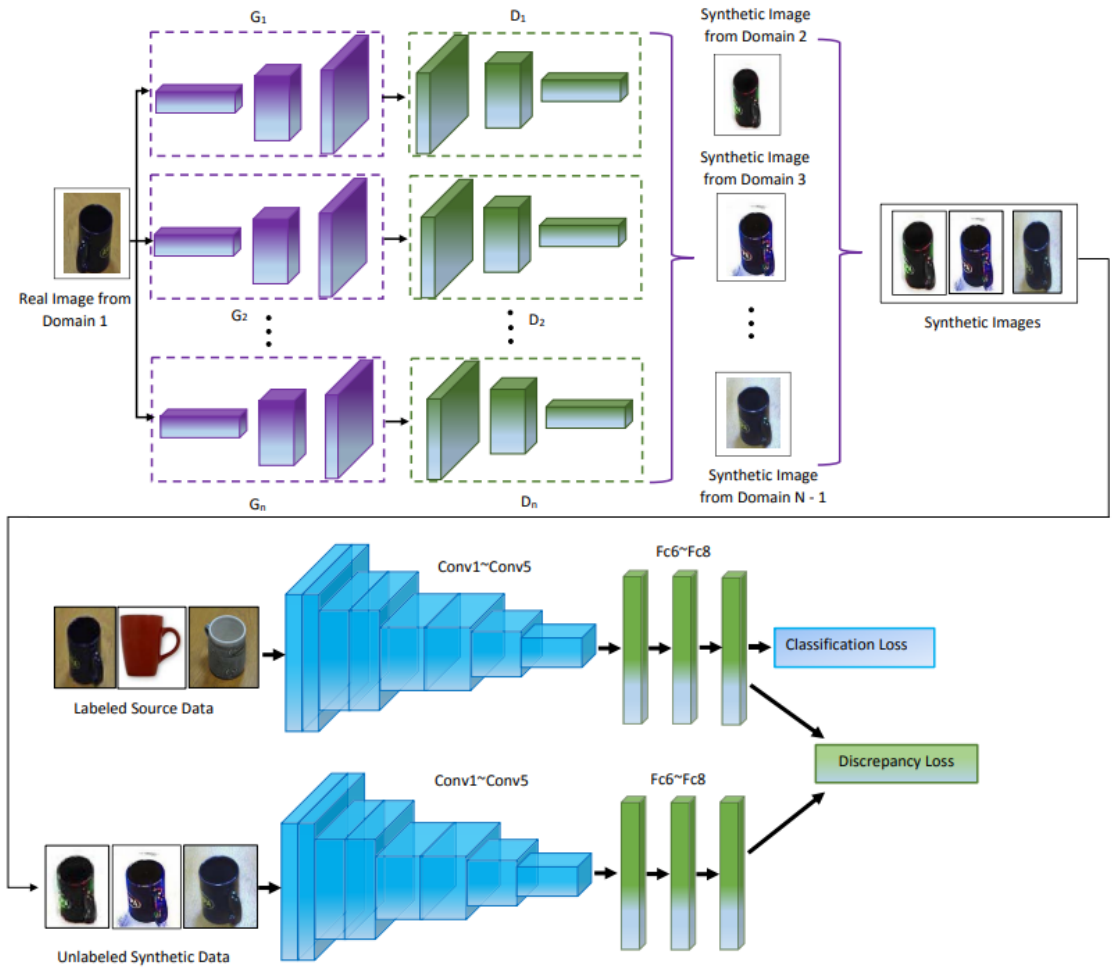
- Build a knowledge graph of object association in the real world
- Generate scenes using this knowledge
  - Data contains class relationships and allows for learning from context

Type	# lanes	Median	Parking lanes	total
Urban	2, 4	trees, concrete, none	none	6
Urban	2	none	parallel, 45-degree	2
Urban	3	none	90-degree	1
Suburban	2, 4	trees, none	none	4
Suburban	2	none	parallel	1
Rural	1, 2, 4	none	none	3
TOTAL				17

Prakash, A., Boochoon, S., Brophy, M., Acuna, D., Cameracci, E., State, G., ... & Birchfield, S. (2019, May). Structured domain randomization: Bridging the reality gap by context-aware synthetic data. In *2019 International Conference on Robotics and Automation (ICRA)* (pp. 7249-7255). IEEE.

Generative Models

- Use generative models to generate additional data in different domains
- Align representations of real and synthetic data by regularisation



Rahman, M. M., Fookes, C., Baktashmotlagh, M., & Sridharan, S. (2019, January). Multi-component image translation for deep domain generalization. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 579-588). IEEE.

Outline

1. Overview & Definitions
2. Techniques for Domain Generalisation
    1. Data-Driven approaches
    2. ***Representation Learning approaches***
    3. Learning-Strategy approaches

Representation Learning Approaches

- Encourage models to learn domain-invariant features
  - Features that are meaningful and consistent across domains
- Again, humans are good at this
  - The concepts we use to identify a cat on grass or in snow are the same

Domain Generalization with Adversarial Feature Learning

The idea:

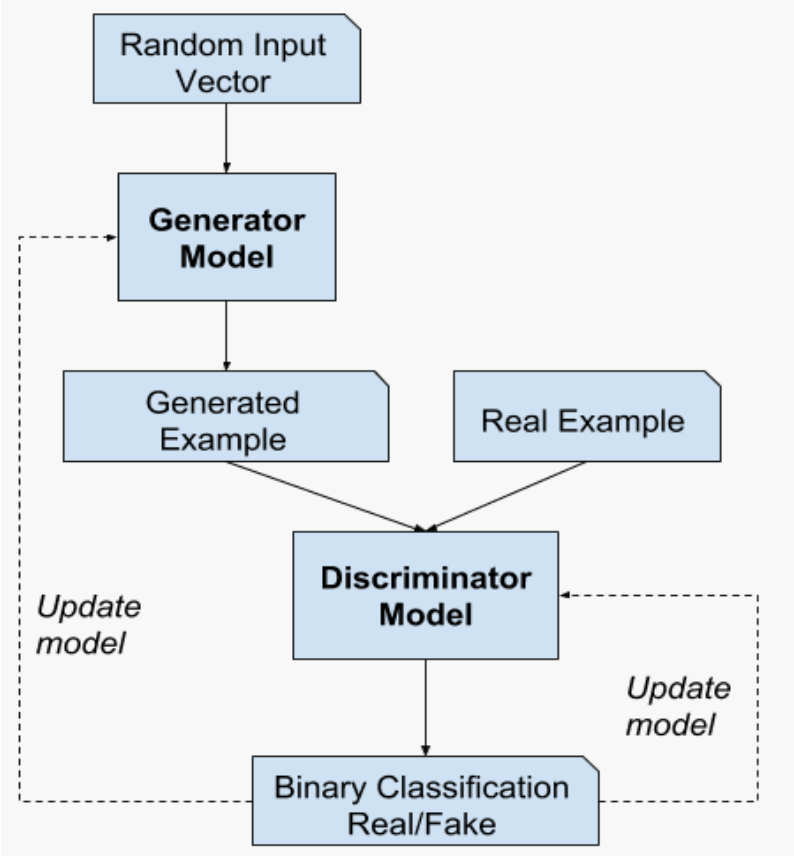
Use adversarial learning to force the model to learn domain-invariant features

Domain Generalization with Adversarial Feature Learning

- In GANs (Generative Adversarial Network), an adversarial discriminator tries to identify generated images
    - The discriminator forces the image generator to generate realistic images (or at least, ones indistinguishable to the discriminator)

Generative Adversarial Networks (GANs)

- Generator
  - Generate fake samples, trying to fool the discriminator
  - Updated to generate better fake samples
- Discriminator
  - Discriminating real and fake samples
  - Updated to better identify fake vs real
- Two Person, Zero-Sum game
  - Winner remain unchanged, loser updates



Generative Adversarial Networks (GANs)

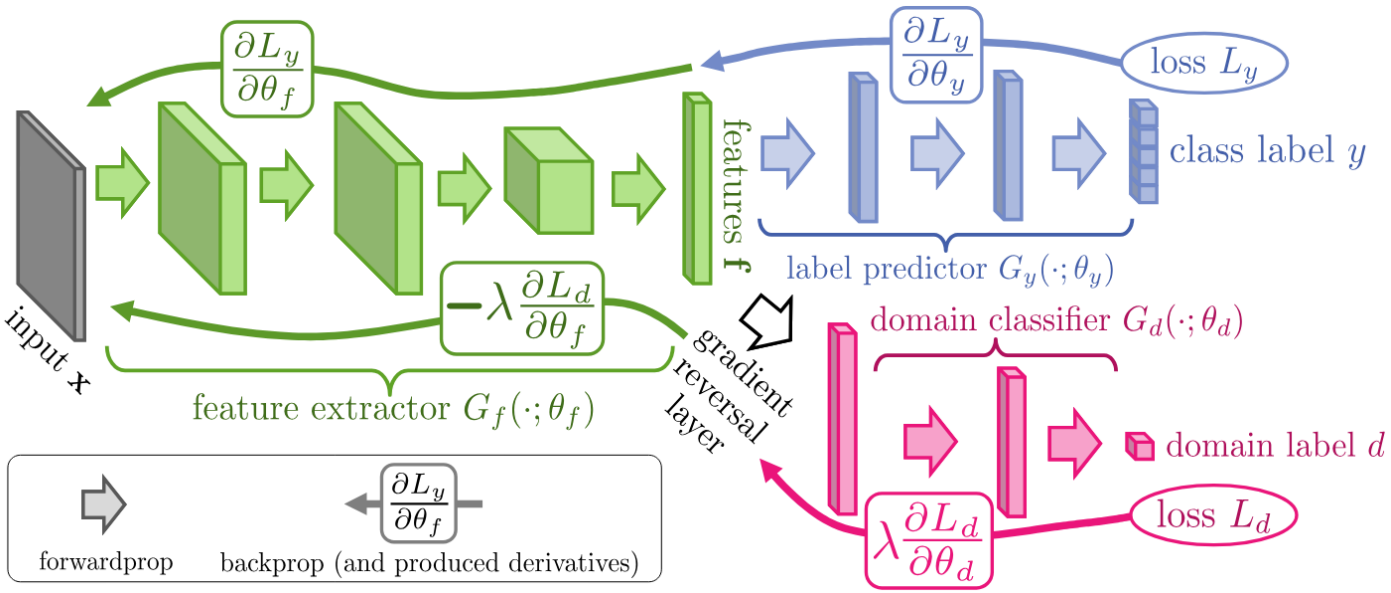
- Result:
  - The generator can generate extremely convincing fake samples (that can fool the discriminator)
  - Useful for image generation, encryption, Super-resolution…
- Challenges
  - Difficult to train, unstable
  - Mode collapse

https://thispersondoesnotexist.com

Domain Generalization with Adversarial Feature Learning

It tries to identify the domain of representations generated by the feature extractor

A good discriminator forces the feature generator to generate domain-invariant representations



Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., ... & Lempitsky, V. (2016). Domain-adversarial training of neural networks. *Journal of machine learning research*, *17*(59), 1-35.

Outline

1. Overview & Definitions
2. Techniques for Domain Generalisation
    1. Data-Driven approaches
    2. Representation Learning approaches
    3. ***Learning-Strategy approaches***

Learning-Strategy Approaches

- Instead of designing a model that generalises across tasks, design a *training strategy* that is model-agnostic
- "Meta-Learning" – Learning to learn

Learning to Generalize:
Meta-Learning for Domain Generalization

The idea:

Train models with a loss function that will perform well on new (test) domains

Li, D., Yang, Y., Song, Y. Z., & Hospedales, T. (2018, April). Learning to generalize: Meta-learning for domain generalization. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).

Learning to Generalize:
Meta-Learning for Domain Generalization

- During training, at every mini-batch, from the domains in our training data we assign "train" and "test" domains
- The "test" domains will just be used to ensure we are learning domain-invariant features
- During testing, we test the performance on actual test domains (which are unseen)

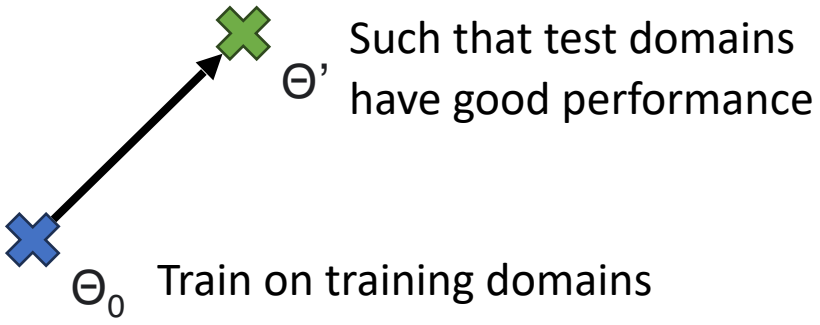
Li, D., Yang, Y., Song, Y. Z., & Hospedales, T. (2018, April). Learning to generalize: Meta-learning for domain generalization. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).

Learning to Generalize:
Meta-Learning for Domain Generalization

- Our loss function has two terms:
  1. Loss on training domains
  2. Loss on test domains, *after* step on training domains

$$\underset{\Theta}{\operatorname{argmin}} \ \mathcal{F}(\Theta)$$

Loss on training domains



Li, D., Yang, Y., Song, Y. Z., & Hospedales, T. (2018, April). Learning to generalize: Meta-learning for domain generalization. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).

Next lessons

- Wednesday: How to write a research paper in ML?
- Next Monday: Self-Supervised Learning

Li, D., Yang, Y., Song, Y. Z., & Hospedales, T. (2018, April). Learning to generalize: Meta-learning for domain generalization. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).