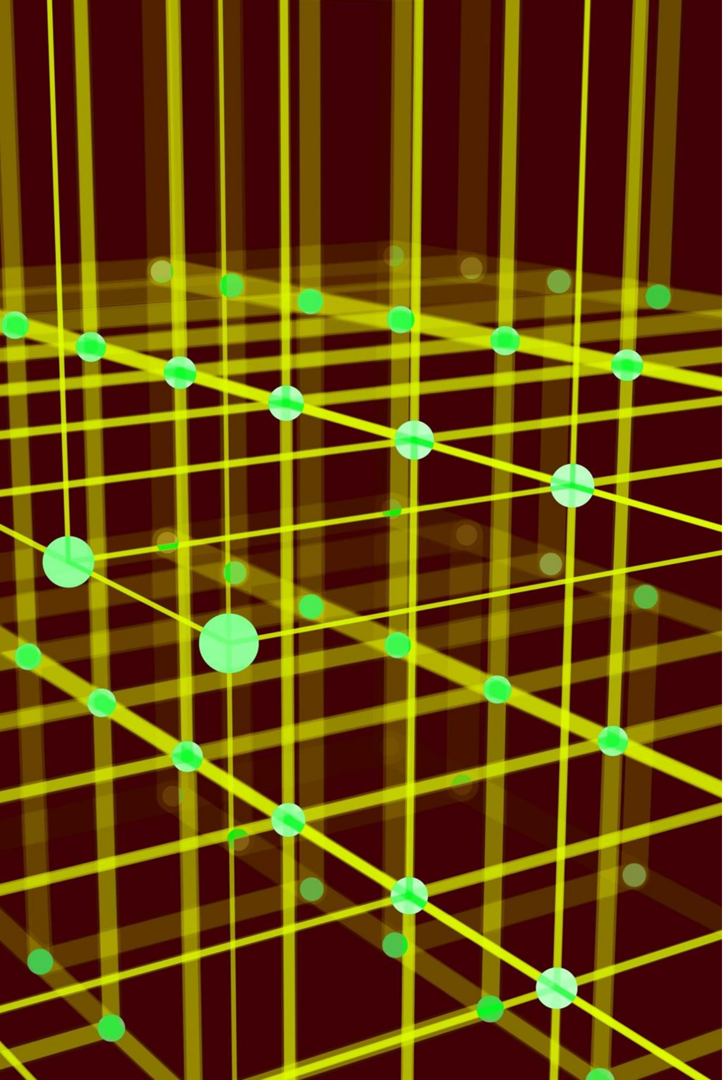


Deep Learning
Domain Adaptation

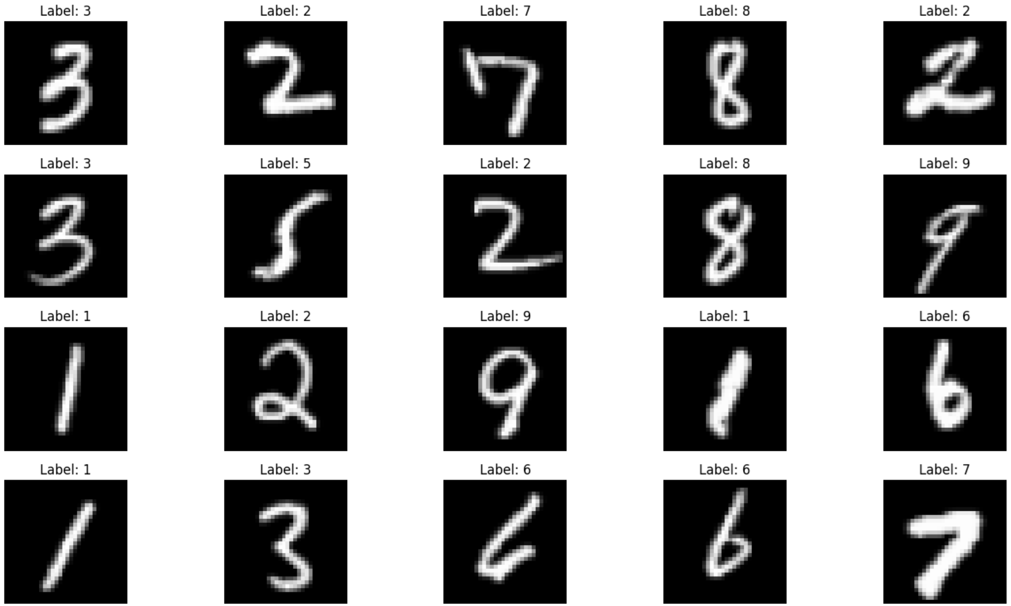
Vasileios Ilias Drouzas
Giannis Trantalidis

Professor: Themos Stafylakis

Contents

Datasets - MNIST

- MNIST
- Collection of handwritten digits
- 28 x 28 pixels
- 60000 examples
- Centered digits
- Grayscale images

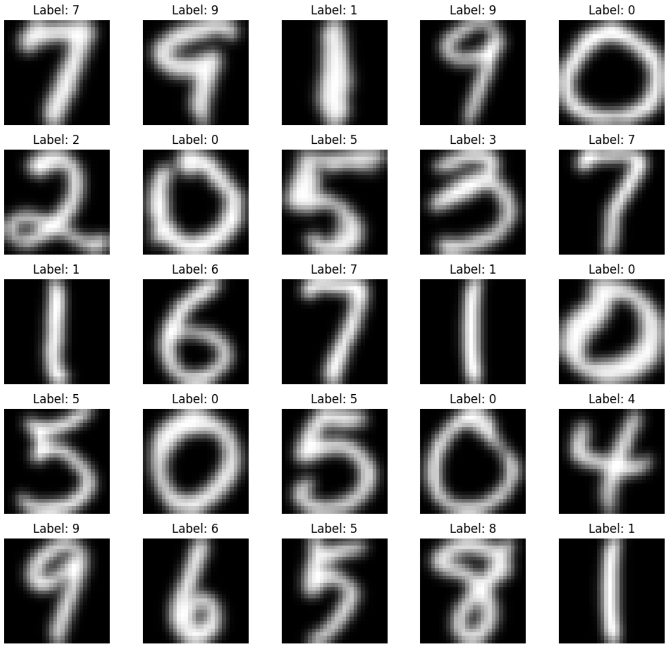


Datasets - SVHN

- SVHN (Street View House Numbers)
- House Numbers
- 32 x 32 pixels RGB
- 73257 examples
- Non-centered digits
- 3 RGB channels

Datasets - USPS

- ❖U.S. Postal Service
- ❖Digit dataset automatically scanned from envelopes
- ❖16×16 pixel grayscale images
- ❖9,298 samples



Pre - processing

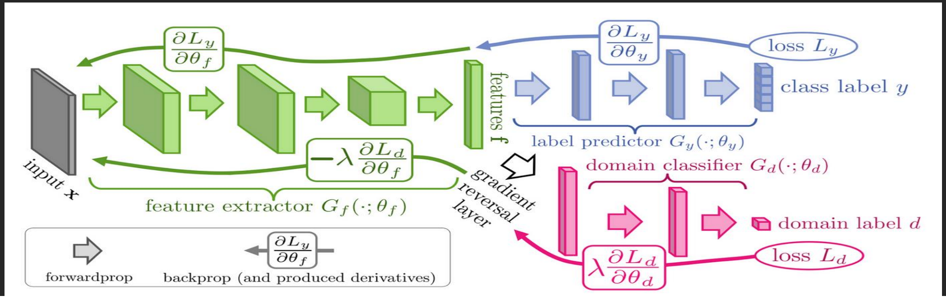
- Padded the MNIST images to 32x32 in order to match SVHN's dimensions
- MNIST: Repeat the single grayscale channel three times to create a 3-channel image
- Normalize to [-1,1]

Algorithms

Simple CNN (Source only)

- We start with a very simple convolutional neural network with:
  - A convolutional layer: input 3 channels, output 16 channels, 3x3 kernel
  - Max - pooling layer
  - Fully connected layer
- CrossEntropyLoss
- Adam optimizer (learning rate: 0.001)

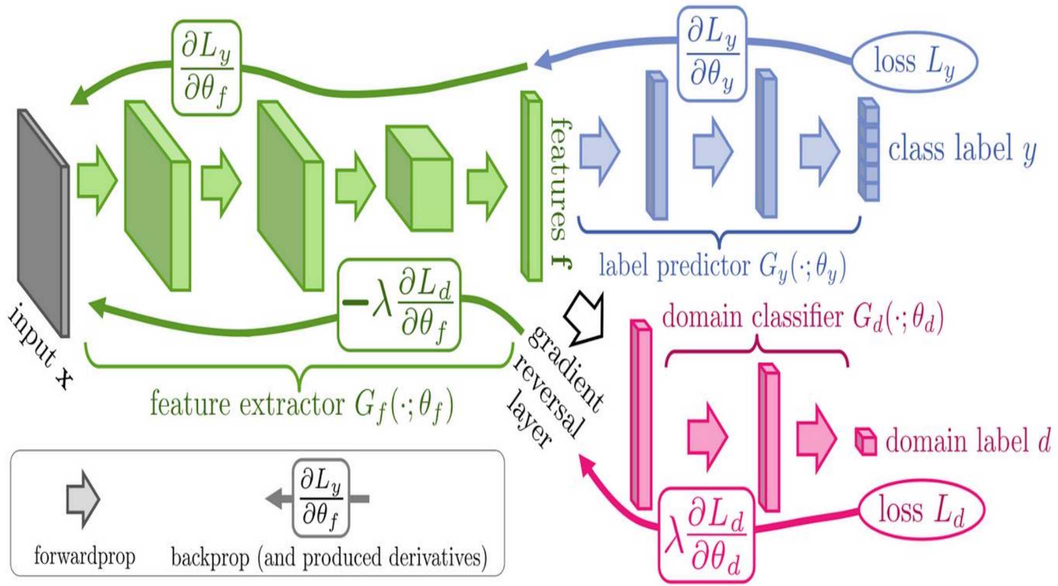
DANN (2014)



- Learn domain-invariant features
- Feature Extractor extracts features from input data.
- Domain classifier that encourages the feature extractor to produce features that are indistinguishable between source and target domains. (Using Gradient Reversal Layer after the feature extractor)

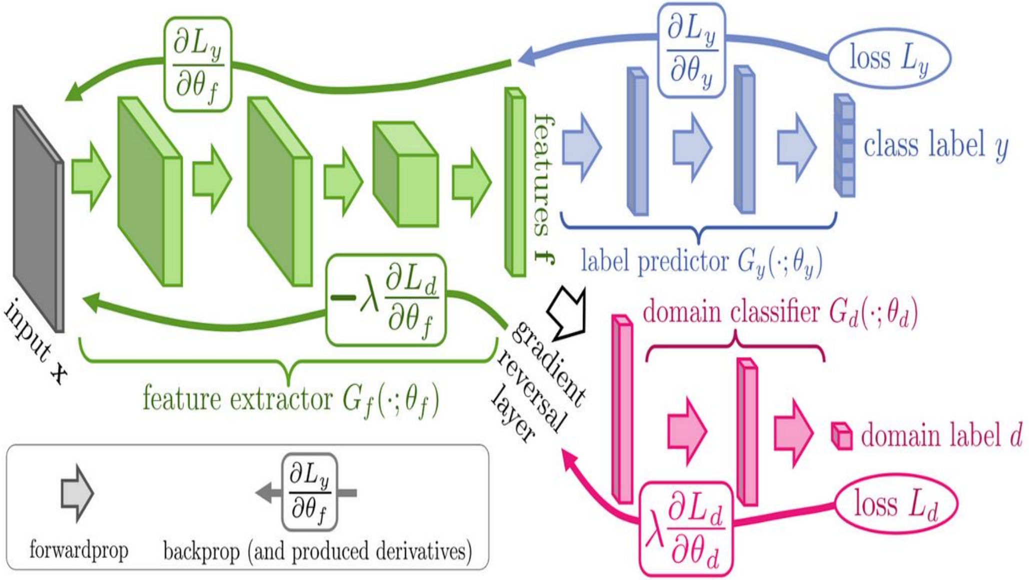
DANN (1/2)

- Feature Extractor
  - Convolutional layers
  - ReLU activations
  - Max pooling layers
  - Batch Normalization
- Label Classifier
  - Fully Connected Layers (classification)
  - ReLU activations
  - Batch Normalization
  - Output layer

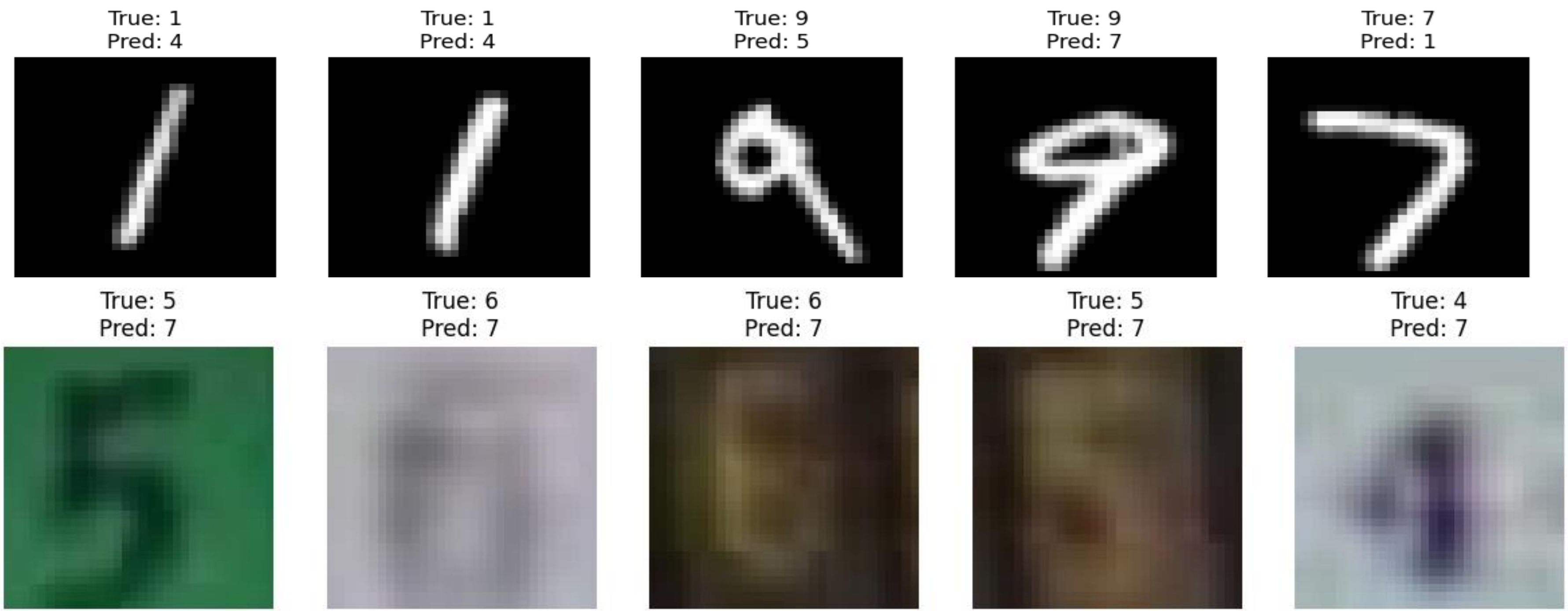


DANN (2/2)

- Domain Classifier
  - Gradient Reversal layer
  - Fully Connected layers
  - Sigmoid Activation function
  - Batch Normalization



DANN: Examples of wrong predictions

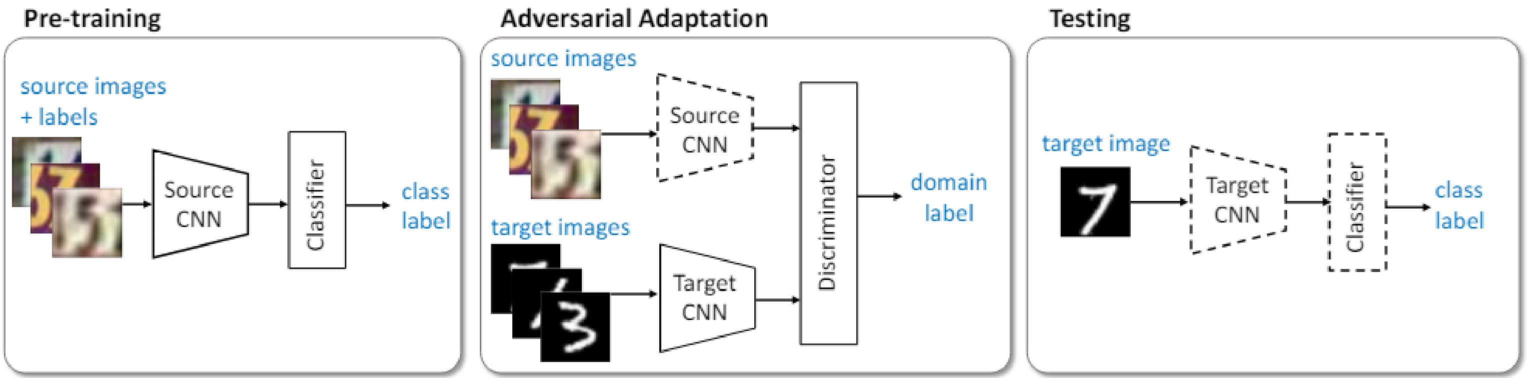


CORAL

- Deep CORAL loss (minimizes domain shift)
- Aligns covariances of source & target domains
- Squared frobenius norm between batch covariances
- Normalize the loss

$$\ell_{CORAL} = \frac{1}{4d^2}\|C_S - C_T\|_F^2$$

ADDA (2017)



1 . Pre-train a source encoder CNN

2. Adversarial adaptation by learning a target encoder CNN such that a discriminator cant distinguish source vs target encoding

Discriminator trained to distinguish source vs target features - target encoder updates weights so that Discriminator cant distinguish the features from the 2 domains.

Benchmarking - Results

Algorithm	MNIST → SVHN	SVHN → MNIST	MNIST → USPS	USPS → MNIST
Source Only	20.96%	51.63%	58.15%	32.50%
DANN	**23.91%**	**66.71%**	**87.74%**	**76.66%**
CORAL	22.87%	62.96%	86.35%	62.82**%**

State-of-the-art scores

Methods	Source / Target	MNIST / USPS	USPS / MNIST	SVHN / MNIST	MNIST / SVHN
Source Only		78.9	57.1±1.7	60.1±1.1	20.23±1.8
w/o augmentation					
CORAL [43]		81.7	-	63.1	-
MMD [48]		81.1	-	71.1	-
DANN [10]		85.1	73.0±2.0	73.9	35.7
DSN [2]		91.3	-	82.7	-
CoGAN [25]		91.2	89.1±0.8	-	-
ADDA [49]		89.4±0.2	90.1±0.8	76.0±1.8	-
DRCN [11]		91.8±0.1	73.7±0.1	82.0±0.2	40.1±0.1
ATT [37]		-	-	86.20	52.8
ADA [13]		-	-	97.6	-
AutoDIAL [3]		97.96	97.51	89.12	10.78
SBADA-GAN [35]		97.6	95.0	76.1	**61.1**
GAM [16]		95.7±0.5	98.0±0.5	74.6±1.1	-
MECA [32]		-	-	95.2	-
DWT		**99.09**±0.09	**98.79**±0.05	**97.75**±0.10	28.92 ±1.9
Target Only		96.5	99.2	99.5	96.7

Source: https://arxiv.org/pdf/1903.03215v2