AAAI 2020

Lab Seminar (2020.02.20)

WooJu Lee
dnwn24@kaist.ac.kr

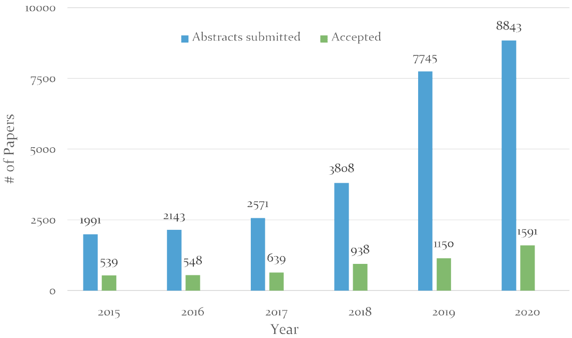
Torkamani, MohamadAli, et al. "Differential Equation Units: Learning Functional Forms of Activation Functions from Data." AAAI (2020)
Diangang Li et al. “Infrared-Visible Cross-Modal Person Re-Identification with an X Modality.” AAAI (2020)

Outline

- Overview
- Invited Talk
  - Bengio, Hinton, Lecun
- Featured paper
  - Differential Equation Units: Learning Functional Forms of Activation Functions from Data
  - Infrared-Visible Cross-Modal Person Re-Identification with an X Modality
- Conclusion

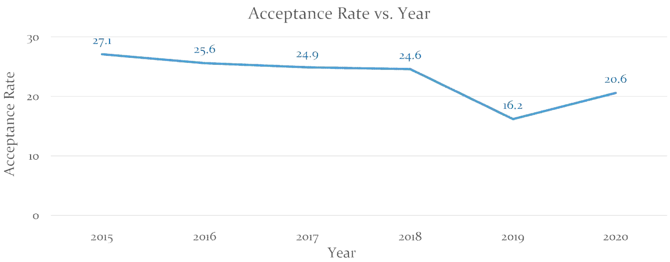
Overview

- Abstracts submitted: 8843



Overview

- Accepted: 1591 (20.56%)
  - 454 oral presentation papers (20 minute)
  - 1137 poster papers (2 minute)



Invited Talk

- Hinton, Bengio, and Lecun
- Current challenges of deep learning
  - Hinton – Invariance → Equivariance
  - Bengio – Low-level → High level cognition
  - Lecun – Labelling data costs a lot → Self-supervised learning

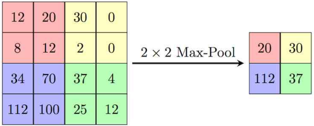
Invited Talk- Hinton

- New version of Capsule Network?
- Translational invariance and equivariance

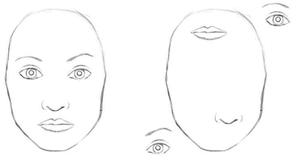
Invariance	Human	Human	Human
Equivariance	Human on the left	Human on the middle	Human on the right

Invited Talk- Hinton

- Typical Convolution Neural Network
  - Max pooling leads to translation invariance



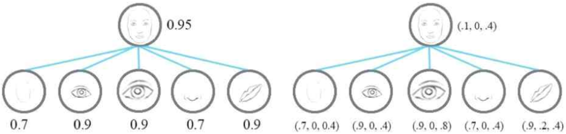
  - Max pooling leads to lose spatial hierarchy in the features.



Invited Talk- Hinton

- Capsule Network – equivariance

	Typical CNN	Capsule Network
Basic Unit	Neuron(scalar)	Capsule(vector)



Invited Talk – Yann Lecun

- Deep supervised learning works well for perception
  - When labeled data is abundant
  - Problem: Producing a dataset with clean labels is expensive

Invited Talk – Yann Lecun

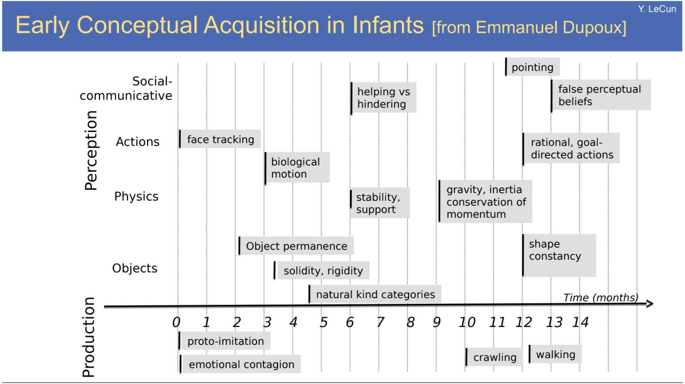
- How do humans and animals learn?
  - Observation
  - Prediction

KAIST

URL URBAN ROBOTICS LAB

Invited Talk – Yann Lecun

- How do humans and animals learn?



Invited Talk – Yann Lecun

- Self-supervised learning
  - Training data is autonomously labelled
  - It is still supervised-learning

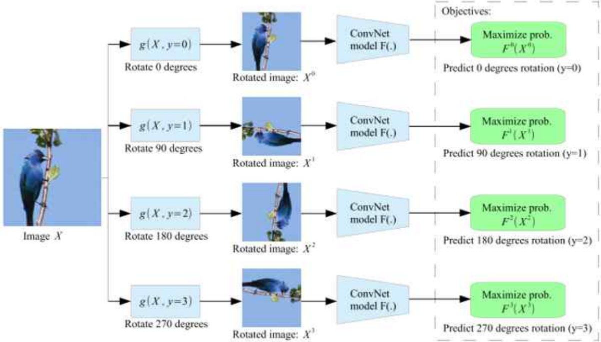
	Supervised Learning	Semi-supervised Learning	Unsupervised Learning	Self-supervised Learning
Data	Labelled data	Combination of labeled and unlabeled data	Unlabelled data	Unlabelled data
Type of problems	Regression, Classification		Association, Clustering	Prediction, Filling in the blank

KAIST

URL URBAN ROBOTICS LAB

Invited Talk – Yann Lecun

- Self-supervised learning example
  - Unsupervised representation learning by prediction image rotations
  - RotNet, 2018, ICLR

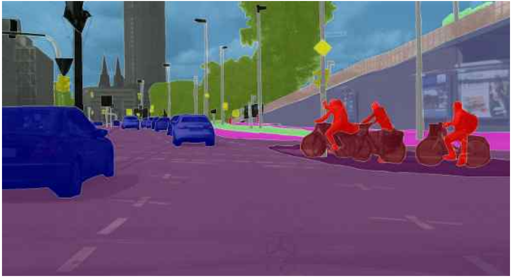


KAIST

URBAN ROBOTICS LAB

Invited Talk – Yann Lecun

- Self-supervised learning
  - Future: It may be possible to predict paths of other agents



Featured paper 1 - DEU

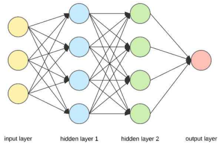
- Title: Differential Equation Units: Learning Functional Forms of Activation Functions from Data
- Problem: Functional form of the activation function is fixed regardless of dataset
- Contribution: It enables each neuron to learn a particular nonlinear activation function.

Featured paper 1 - DEU

- Main idea: Finding parameters of an ordinary differential equation(ODE) for each neuron

- Each neuron has own ODE

→ $ay''(t)+by'(t)+cy(t) = u(t)$, where $u(t) = \begin{cases} 0 & t \leq 0 \\ 1 & t > 0 \end{cases}$



  - a,b,c: scalars used to parameterize the ODE
  - c1,c2: initial conditions of ODE's solution
  - u(t): Heaviside(unit) step function

- a,b,c,c1,c2 parameters are treated like biases of the neuron and trained

Featured paper 1 - DEU

- Main idea:
  - Solution of ODE w.r.t (a,b,c,c1,c2) parameters

$$ay''(t) + by'(t) + cy(t) = u(t),$$
$$\text{where } u(t) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

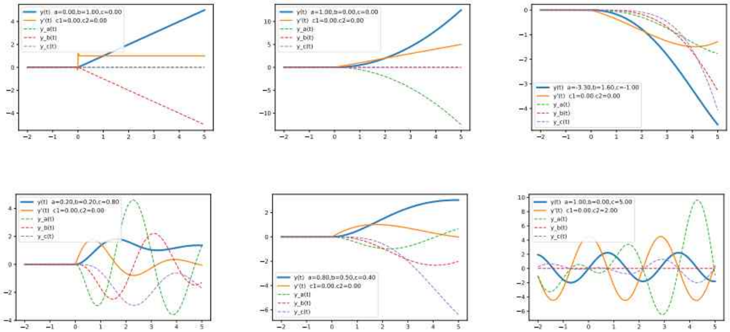
Subspace	Solution
$a = 0, b = 0, c \neq 0$	$\sigma(t)/c$ (i.e., sigmoid when $c = 1$)
$a = 0, b \neq 0, c = 0$	$xu(x)/b + c_1$ (i.e., ReLU when $b = 1$ and $c_1 = 0$)
$a = 0, b \neq 0, c \neq 0$	$c_1 e^{-(cx)/b} - u(x)e^{-(cx)/b}/c + u(x)/c$
$a \neq 0, b = 0, c = 0$	$\frac{x^2 u(x)}{2a} + c_2 x + c_1$
$a < 0, b = 0, c < 0$ or $a > 0, b = 0, c > 0$	$c_2 \sin(\sqrt{\frac{c}{a}}x) + c_1 \cos(\sqrt{\frac{c}{a}}x) - \frac{u(x)}{c}\left(\cos(\sqrt{\frac{c}{a}}x) - 1\right)$
$a < 0, b = 0, c > 0$ or $a > 0, b = 0, c < 0$	$c_1 e^{\sqrt{-\frac{c}{a}}x} + c_2 e^{-\sqrt{-\frac{c}{a}}x} + \frac{u(x)e^{-\sqrt{-\frac{c}{a}}x}}{2c} + \frac{u(x)e^{\sqrt{-\frac{c}{a}}x}}{2c} - \frac{u(x)}{c}$
$a \neq 0, b \neq 0, c = 0$	$\frac{u(x)a}{b^2} e^{-\frac{bx}{a}} - \frac{u(x)a}{b^2} - c_1 \frac{a}{b} e^{-\frac{bx}{a}} + \frac{u(x)x}{b} + c_2$

KAIST

URL URBAN ROBOTICS LAB

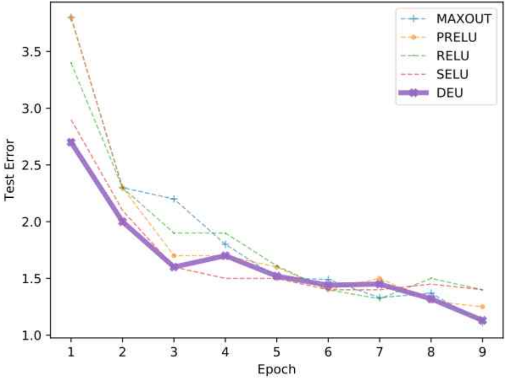
Featured paper 1 - DEU

- Result



Featured paper 1 - DEU

- Result



Featured paper 2 – Fusion strategy

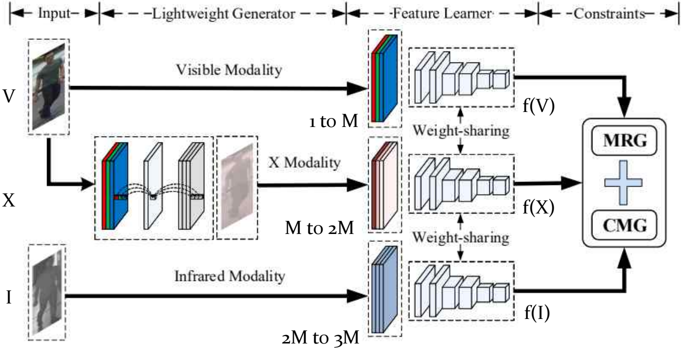
- Title: Infrared-Visible Cross-Modal Person Re-Identification with an X Modality

- Problem: Significant gap between the RGB and infrared images.

- Contribution: Introducing an auxiliary X modality to reduce the gap

Featured paper 2 – Fusion strategy

- Result

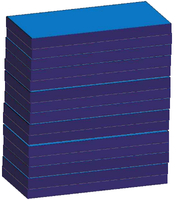


Featured paper 2 – Fusion strategy

- Loss = CMG(Cross Modality Gap) + MRG(Modality Respective Gap)
  - CMG: $\mathcal{L}_C = \mathcal{L}_{cross}^{\mathbf{I},\mathbf{X}} + \mathcal{L}_{cross}^{\mathbf{I},\mathbf{V}}.$

    $\mathcal{L}_{cross}^{\mathbf{I},\mathbf{X}} = \frac{1}{M}(\mathcal{L}_{\mathbf{I}-\mathbf{X}} + \mathcal{L}_{\mathbf{X}-\mathbf{I}}),$

D: Euclidean distance
y: target
I: Infrared
X: X-modality
V: Visibility



$$\mathcal{L}_{\mathbf{I}-\mathbf{X}} = \sum_{i=1}^{M}[\alpha_1 + \max_{\substack{j=M+1,\cdots,2M \\ y_i=y_j}} D(f(\mathbf{I}_i), f(\mathbf{X}_j)) - \min_{\substack{k=M+1,\cdots,2M \\ y_i \neq y_k}} D(f(\mathbf{I}_i), f(\mathbf{X}_k))]_+,$$

$$\mathcal{L}_{\mathbf{X}-\mathbf{I}} = \sum_{i=M+1}^{2M}[\alpha_1 + \max_{\substack{j=1,\cdots,M \\ y_i=y_j}} D(f(\mathbf{X}_i), f(\mathbf{I}_j)) - \min_{\substack{k=1,\cdots,M \\ y_i \neq y_k}} D(f(\mathbf{X}_i), f(\mathbf{I}_k))]_+.$$

  - MRG: similar to typical cross entropy loss

Featured paper 2 – Fusion strategy

- Result

Table 1: Comparison results (%) with the state-of-the-art IV-ReID methods on RegDB and SYSU-MM01 datasets.

Approach	RegDB				SYSU-MM01			
	$r=1$	$r=10$	$r=20$	mAP	$r=1$	$r=10$	$r=20$	mAP
One-stream (Wu et al. 2017)	13.11	32.98	42.51	14.02	12.04	49.68	66.74	13.67
Two-stream (Wu et al. 2017)	12.43	30.36	40.96	13.42	11.65	47.99	65.50	12.85
Zero-Padding (Wu et al. 2017)	17.75	34.21	44.35	18.90	14.80	54.12	71.33	15.95
TONE (Ye et al. 2018a)	16.87	34.03	44.10	14.92	12.52	50.72	68.60	14.42
HCML (Ye et al. 2018a)	24.44	47.53	56.78	20.80	14.32	53.16	69.17	16.16
BDTR (Ye et al. 2018b)	33.47	58.42	67.52	31.83	17.01	55.43	71.96	19.66
D-HSME (Hao et al. 2019)	50.85	73.36	81.66	47.00	20.68	62.74	77.95	23.12
cmGAN (Dai et al. 2018)	-	-	-	-	26.97	67.51	80.56	27.80
D^2RL (Wang et al. 2019c)	43.40	66.10	76.30	44.10	28.90	70.60	82.40	29.20
MSR (Feng, Lai, and Xie 2019)	48.43	70.32	79.95	48.67	37.35	83.40	93.34	38.11
AlignGAN (Wang et al. 2019a)	57.90	-	-	53.60	42.40	85.00	93.70	40.70
Our method	**62.21**	**83.13**	**91.72**	**60.18**	**49.92**	**89.79**	**95.96**	**50.73**

KAIST

URBAN ROBOTICS LAB

Thank you for your attention