Lecture slides available at

http://goo.gl/MdA6vi

台灣人工智慧學校技術領袖培訓班

Transfer Learning: Introduction to Transfer Learning

Yu-Chiang Frank Wang 王鈺強, Associate Professor

Graduate Inst. Comm. Engineering & Dept. Electrical Engineering

National Taiwan University

2018/02/26

About Myself

- **Current Position**
  - Associate Professor 2017 – present
    GICE & EE, National Taiwan University
- **Research Interests**
  - Computer Vision, Machine Learning, Image & Video Processing
- **Education**
  - PhD in ECE, Carnegie Mellon University, 2004-2009
    - Advisor: David Casasent (1942-2015)
  - MS in ECE, Carnegie Mellon University, 2002-2004
    - Advisor: David Casasent (1942-2015)
  - BS in EE, National Taiwan University, 1997-2001

About Myself (cont'd)

- **Honors & Awards**
  - TAAI & IPPR MS Thesis Award 2017
    (advisee: Wei-Yu Chen; co-advisor: Prof. Ming-Syan Chen)
  - Nominated for Best Paper Awards
    IEEE AVSS 2015, IEEE ICME 2013, IAPR MVA 2013, IEEE ICIP 2010
  - Postdoctoral Researcher Award (advisee: Dr. Chia-Po Wei)
    Ministry of Science & Technology, 2015
  - Honorable Mention (advisee: Hui-Tang Chang; co-advisor: Prof. Ming-Syan Chen)
    IPPR PhD/MS Thesis Award, 2015
  - Outstanding Young Researcher
    National Science Council, 2013-2015
  - First Place Award
    Taiwan Tech Trek, National Science Council, 2011

About Myself (cont'd)

- **Academic/Professional Services**
  - Invited Conference/Tutorial Speaker

    TWSIAM 2017, HDBDA 2015, ACML 2014, AI Forum 2012, WCF 2011
  - Conference Chair

    Program Co-Chair (IEEE AVSS 2019, CVGIP 2016, IEMV 2014)

    Local Arrangement Co-Chair (ACCV 2016)

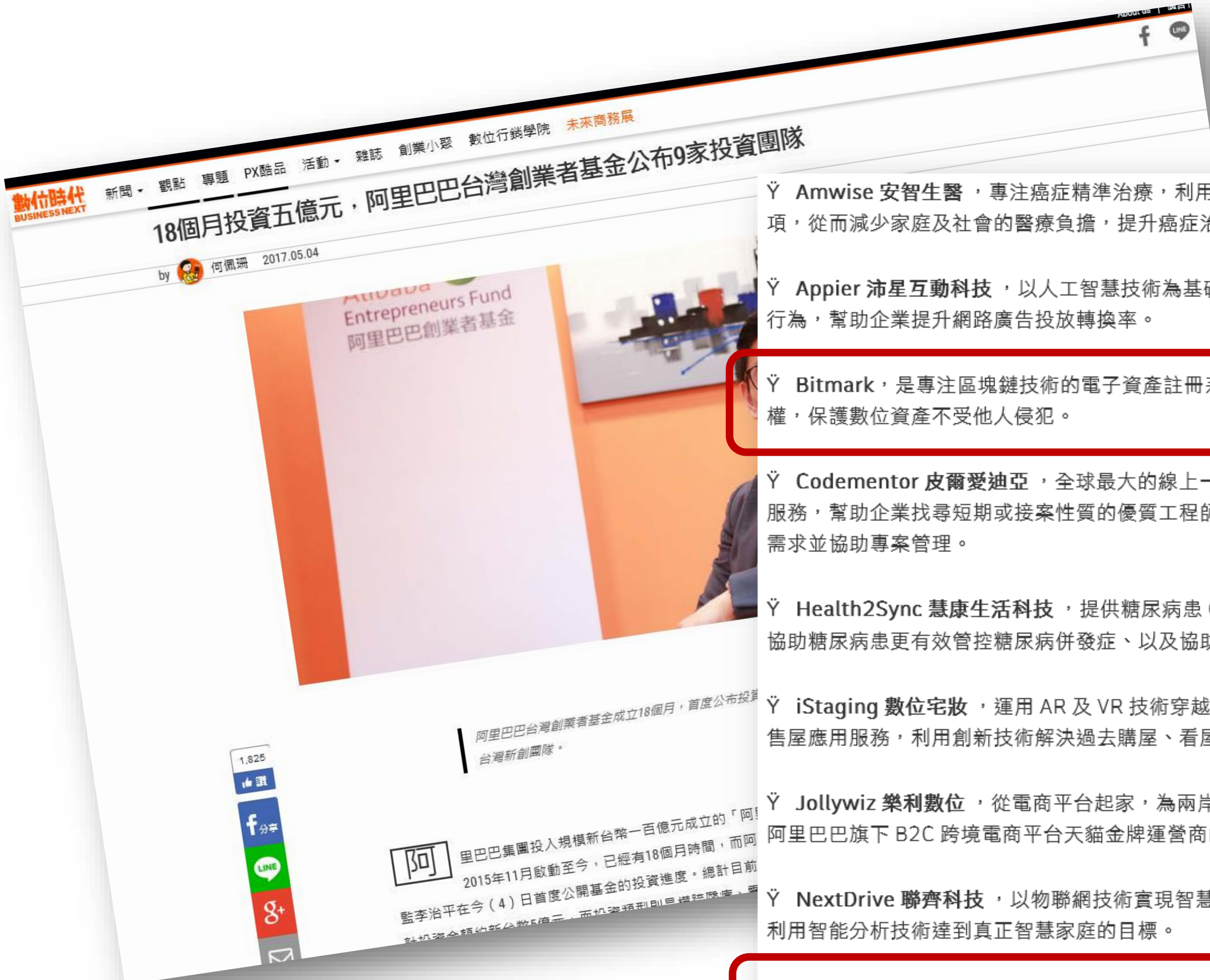
    Publicity Co-Chair (IMV 2013)
  - TPC/PC/Area Chair

    IEEE CVPR, ICCV, ICME, AVSS, WACV, FG, IAPR MVA, IJCAI, ACML, etc.

About Myself (cont'd)

- **Work Experience**
  - Deputy Director 2015 – 2017
    Research Center for IT Innovation (CITI), Academia Sinica
  - Associate Research Fellow 2013 – 2017
    CITI, Academia Sinica
  - Assistant Research Fellow 2009 – 2013
    CITI, Academia Sinica
  - Research Assistant, NHRI/NTU 2001 – 2002
- **Industrial Experience**
  - Advisor/Consultant/Collaborator
    - TSMC, ITRI, Viscovery, UmboCV, Digital Drift, BitMark, Theia, etc.

Alibaba Entrepreneurs Fund [2017/05/04]



- Amwise 安智生醫，專注癌症精準治療，利用基因檢測，協助醫師及患者找出最佳的治療選項，從而減少家庭及社會的醫療負擔，提升癌症治療品質，創造最大的醫療效益。
- Appier 沛星互動科技，以人工智慧技術為基礎，幫助企業洞察消費者在不同裝置的跨螢轉換行為，幫助企業提升網路廣告投放轉換率。
- Bitmark，是專注區塊鏈技術的電子資產註冊系統，使用者能將他們的數位原創內容註記所有權，保護數位資產不受他人侵犯。

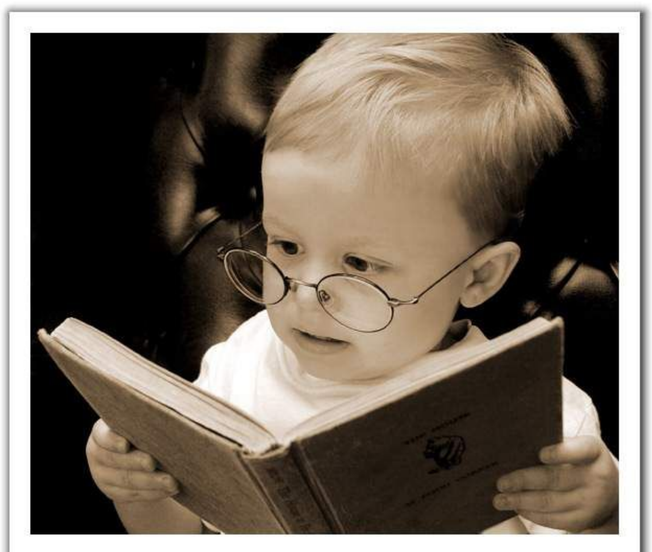
2015

- Codementor 皮爾愛迪亞，全球最大的線上一對一程式語言教育平台，同時提供工程師媒合服務，幫助企業找尋短期或接案性質的優質工程師，滿足企業在尋求外部軟體工程人才資源上的需求並協助專案管理。
- Health2Sync 慧康生活科技，提供糖尿病患 O2O 線上數據結合線下的全天照護創新模式，協助糖尿病患更有效管控糖尿病併發症、以及協助醫護人員提升管理效率。
- iStaging 數位宅妝，運用 AR 及 VR 技術穿越實境、雲端科技、空間設計等三大特色的專業售屋應用服務，利用創新技術解決過去購屋、看屋與裝潢設計遇到的不便。
- Jollywiz 樂利數位，從電商平台起家，為兩岸同步電子商務及品牌營運團隊，並為首家獲得阿里巴巴旗下 B2C 跨境電商平台天貓金牌運營商的台資企業。
- NextDrive 聯齊科技，以物聯網技術實現智慧家庭及智慧節能的目標，透過整合家中數據，利用智能分析技術達到真正智慧家庭的目標。
- Viscovery 創意引睛，以深度學習及電腦視覺為基礎，透過影音辨識及圖像搜索能力，提供企業廣告投放效能評估，大幅提升投資效益。

2017

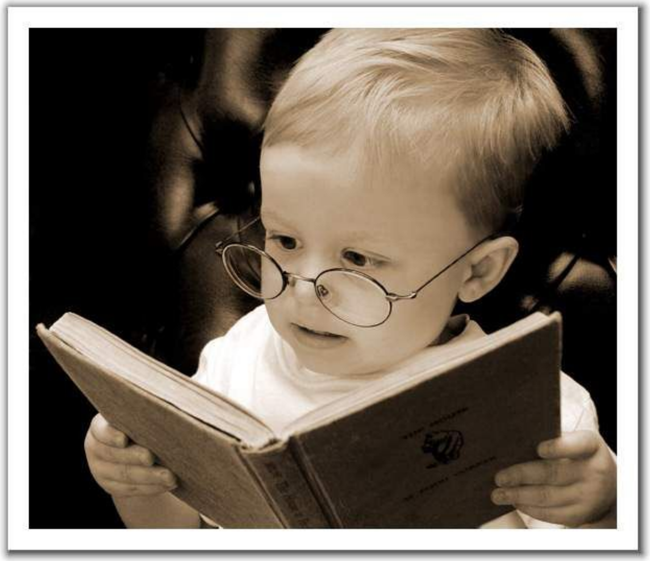
What Will Be Covered in Today's Lecture?

- Transfer Learning
  - Introduction to Transfer Learning (TL)
  - Challenges in Transfer Learning
  - TL for Visual Analysis
  - TL for Visual Synthesis and Manipulation

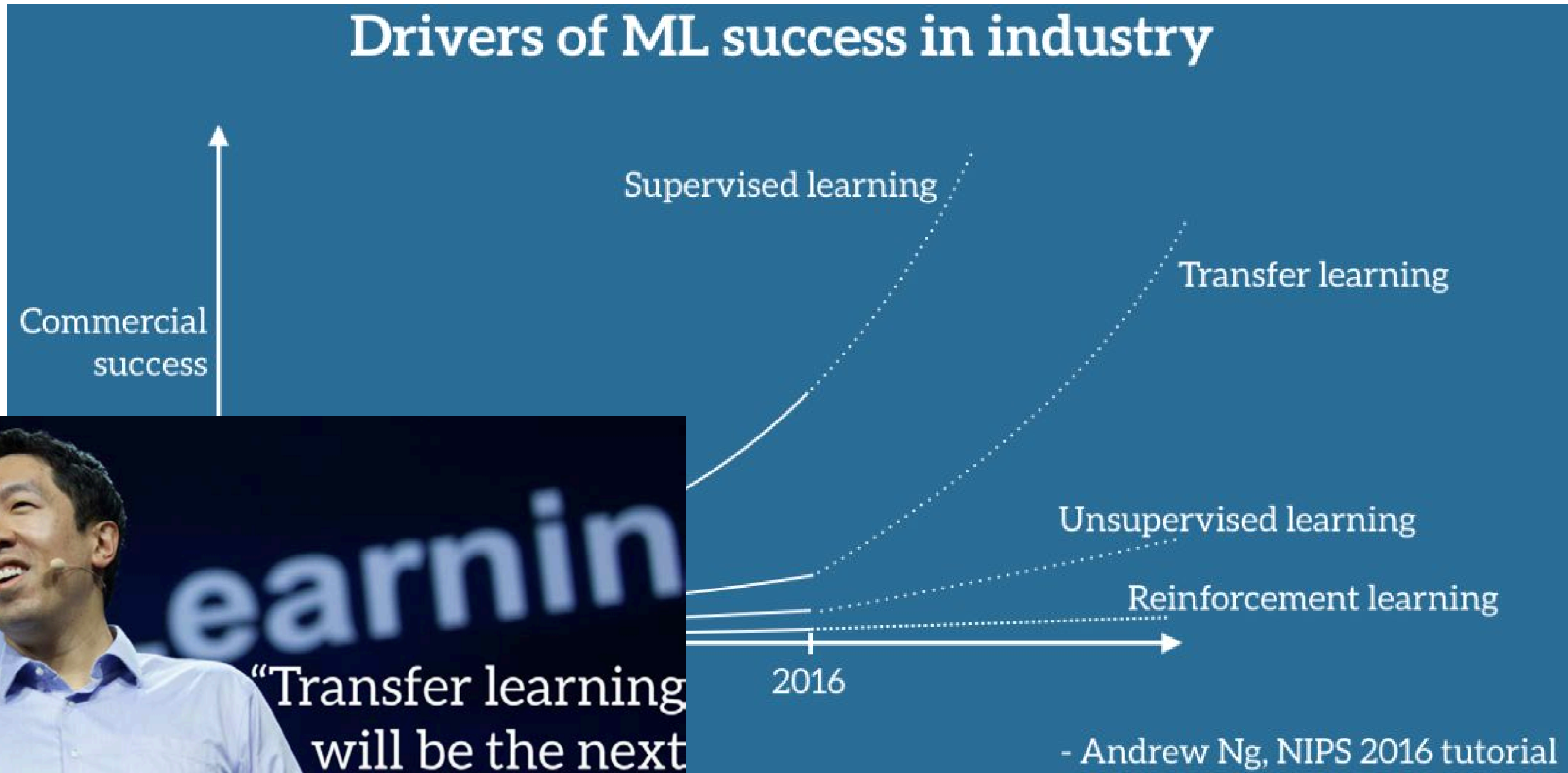


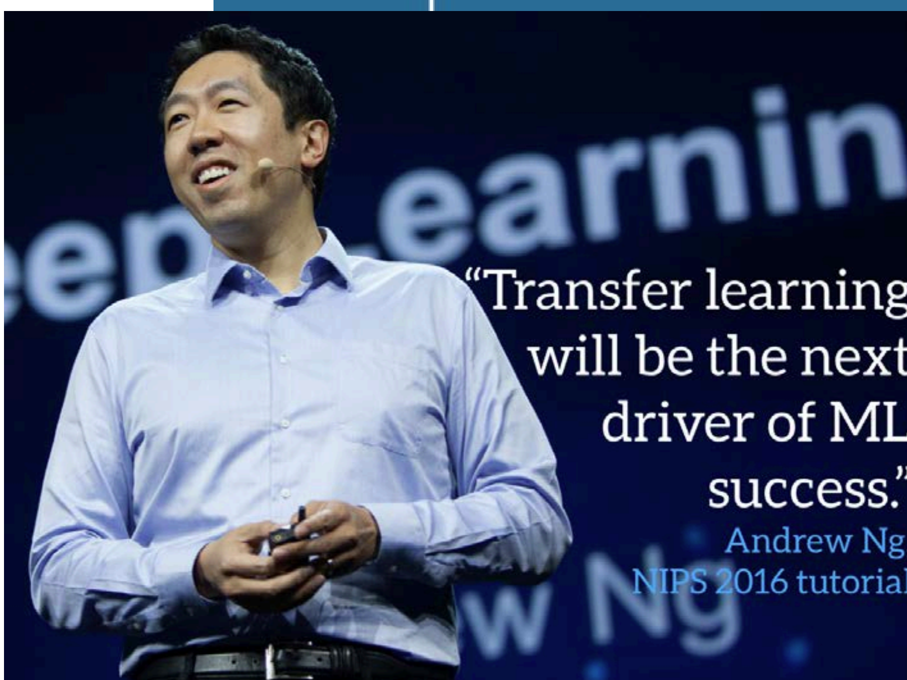
Topic #1 (09:30~10:30)

- Transfer Learning
  - Introduction to Transfer Learning (TL)
  - Challenges in Transfer Learning
  - Transfer Learning for Visual Analysis
  - Transfer Learning for Visual Synthesis

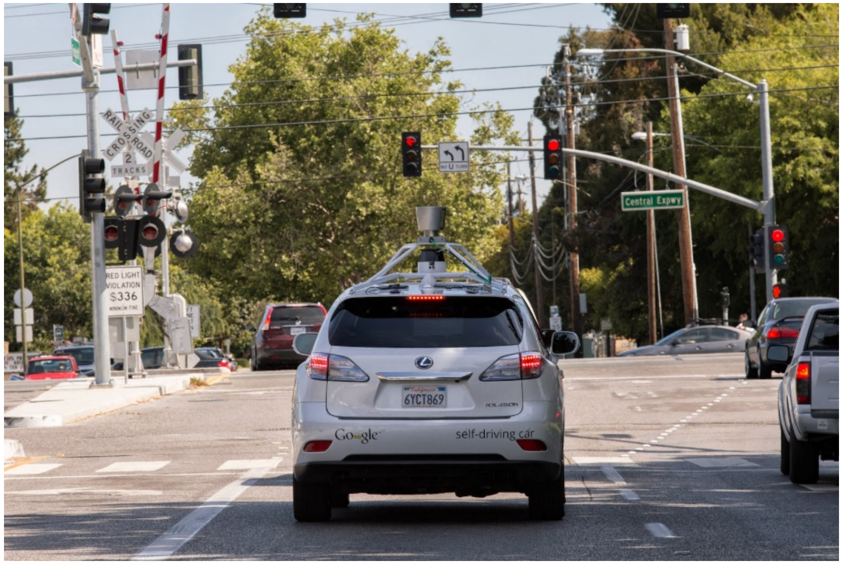


Why You Should Know Transfer Learning?





Why You Should Know Transfer Learning?



https://techcrunch.com/2017/02/08/udacity-open-sources-its-self-driving-car-simulator-for-anyone-to-use/
https://googleblog.blogspot.tw/2014/04/the-latest-chapter-for-self-driving-car.html

Why You Should Know Transfer Learning?

- Let's see some keywords first
  - Machine learning
  - Big data (or data science, etc.)
  - Deep learning
- Later we will talk about when you really need it (or not).
- Yes, we will highlight a number of deep learning techniques for TL.
- No, we will not detail learning algorithms for each TL model, deep or not.

Transfer Learning: What, When, and Why?

- What is Transfer Learning?
  - "Transfer learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem." – Wikipedia

  - English please? (maybe plus some examples…)

Transfer Learning: What, When, and Why? (cont'd)

- Let's start from some examples first…

vs.

Transfer Learning: What, When, and Why? (cont'd)

- Examples #2

vs.

Transfer Learning: What, When, and Why? (cont’d)

- Let’s see some practical examples…
  - Semantic analysis

★★★★☆ **pattern recognition in engineering**
By Michael R. Chernick on February 8, 2008
Format: Hardcover
Fukunaga is a standard source for pattern recognition methods often cited in the engineering literature. Covers parametric (particularly linear and quadratic discriminant algorithms) and nonparametric methods (density estimation). It is designed for and popular with engineers. When I was working at Nichols Research Corporation Fukunaga's papers and this book (earlier edition) were often cited as sources to justify the algorithms we used for discrimination problems. In fact Fukunaga had been a consultant to the company (used primarily by the Boston branch of the company where the KENN algorithms were developed). It is a reputable source. I still like Duda and Hart (1973) for good explanations of the fundamental concepts. The second edition that was recent ly published with Stark as a third author is also highly recommended. For statisticians McLachlan's book is now far and away the best source.
Comment | 31 people found this helpful. Was this review helpful to you? Yes No Report abuse

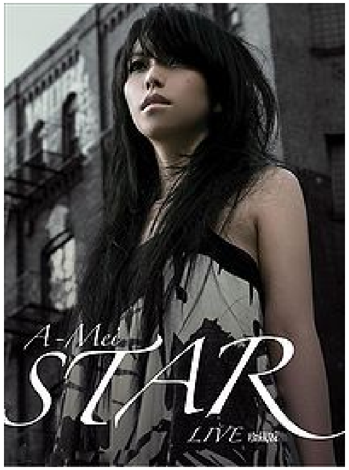
While perfectly serviceable as an hour-and-a-half of shocks and scares, there's substance missing to The Boy that prevents it from truly coming to life.
Full Review… | March 17, 2016

Source Domain

Target Domain

Transfer Learning: What, When, and Why? (cont'd)

- Let's see some practical examples…
  - Information Retrieval



Source Domain

Target Domain

Transfer Learning: What, When, and Why? (cont'd)

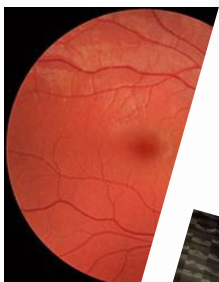
- Let's see some practical examples…
  - Image Classification

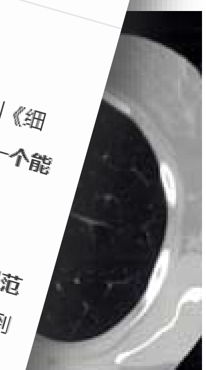
Source Domain

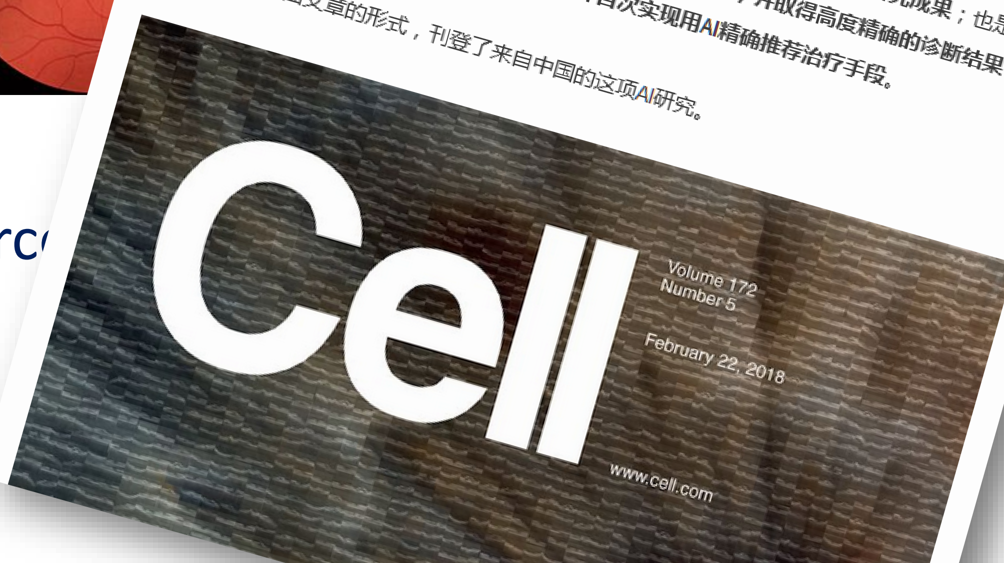
Target Domain

Transfer Learning: What, When, and Why? (cont'd)

- Let's see some pr
  - Medical Image







《细胞》重磅：中国科学家研发的AI影像诊断系统首登顶级期刊，可诊断眼病和肺炎两大类疾病，准确性匹敌顶尖医生 | 科学大发现

原创 2018-02-23 BioTalker 奇点网

今天，由张康教授领导的广州妇女儿童医疗中心和加州大学圣迭戈分校团队，在顶级期刊《细胞》上发表了一篇人工智能（AI）在医疗领域应用的重磅研究成果：**基于深度学习开发出一个能诊断眼病和肺炎两大类疾病的AI系统**[1]，**准确性匹敌顶尖医生。**

这不仅是**中国研究团队首次在顶级生物医学杂志发表有关医学人工智能的研究成果**；也是**世界范围内首次使用如此庞大的标注好的高质量数据进行迁移学习，并取得高度精确的诊断结果**，达到匹敌甚至超越人类医生的准确性；还是**全世界首次实现用AI精确推荐治疗手段。**

《细胞》以封面文章的形式，刊登了来自中国的这项AI研究。

Sourc

omain

Transfer Learning: What, When, and Why?

- What is Transfer Learning?
  - "Transfer learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem." – Wikipedia

- What is the common assumption in Machine Learning?
  - Training data (typically annotated) would be available.
  - Training and test data are drawn from the same feature space and with the same distribution.

(Traditional) Machine Learning vs. Transfer Learning

- Machine Learning
  - Collecting/annotating data is typically expensive.

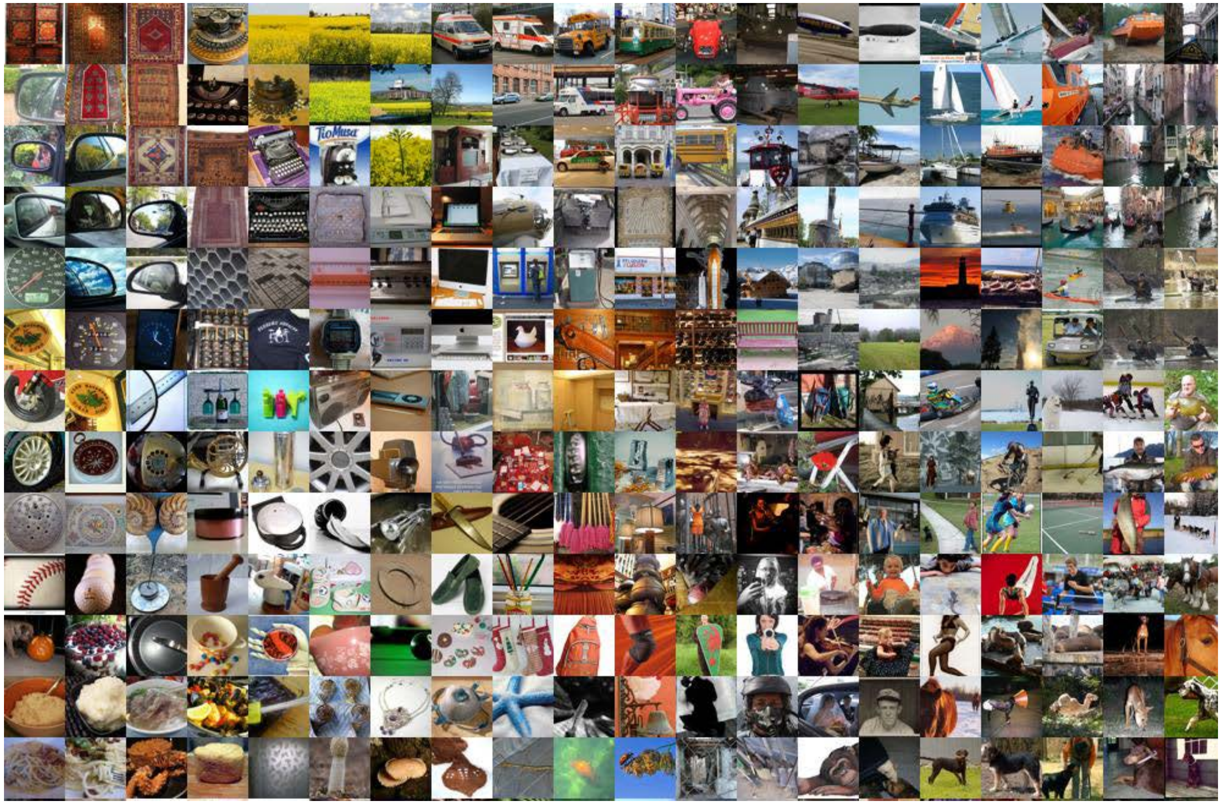
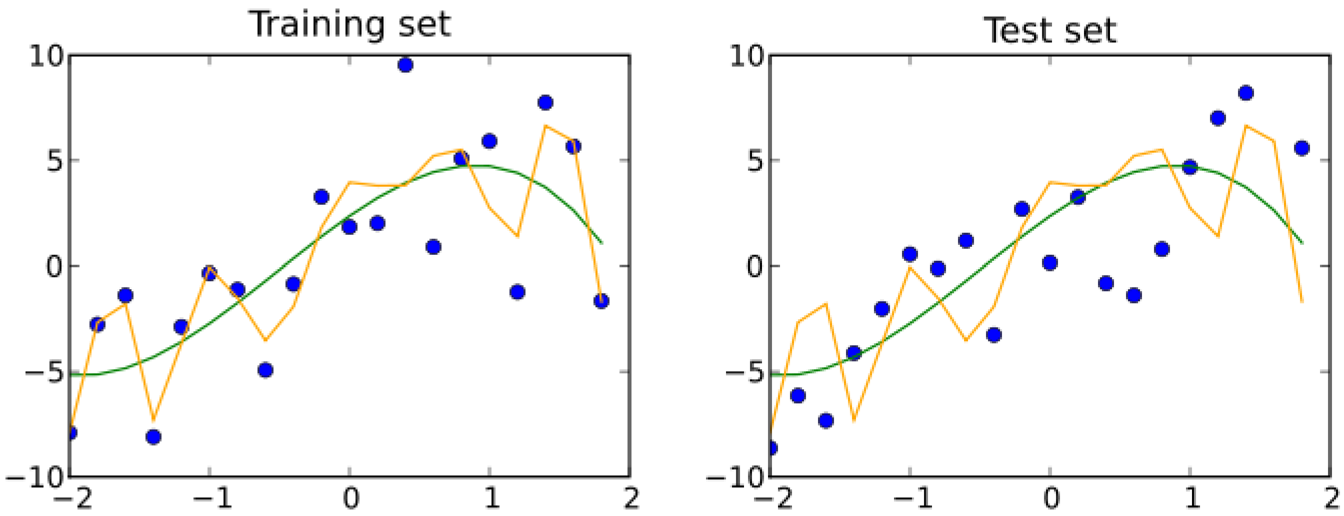


*Image Courtesy: Andrej Karpathy

(Traditional) Machine Learning vs. Transfer Learning (cont'd)

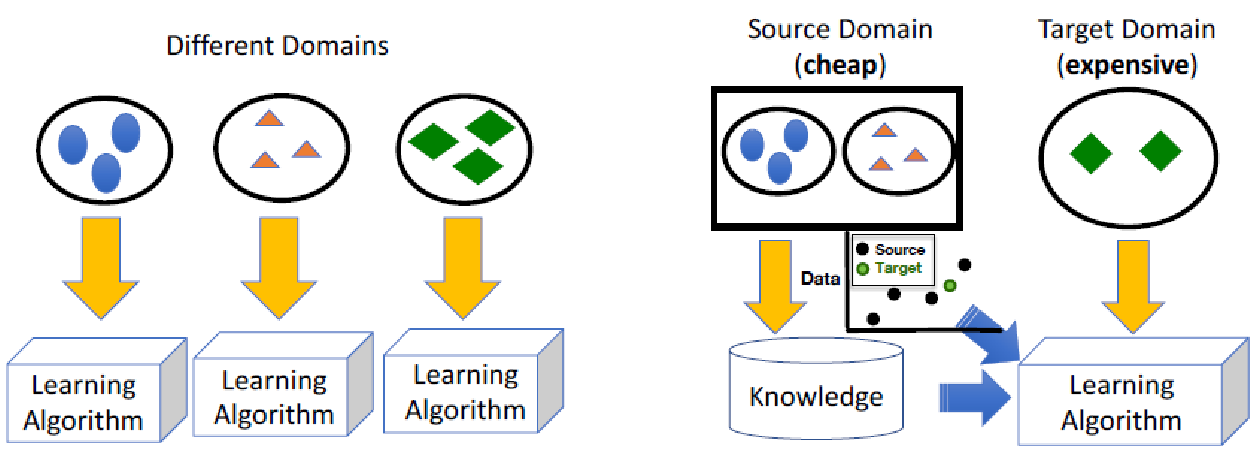
- Machine Learning
  - Collecting/annotating data is typically expensive.
  - Assuming same training/test data distributions might not practical.



https://upload.wikimedia.org/wikipedia/commons/thumb/0/0e/Traintest.svg/700px-Traintest.svg.png

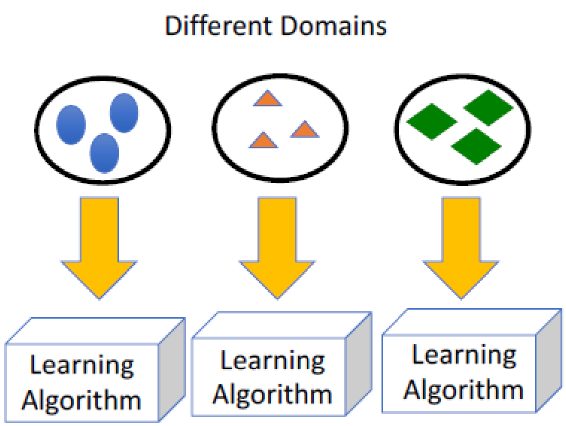
(Traditional) Machine Learning vs. Transfer Learning (cont'd)

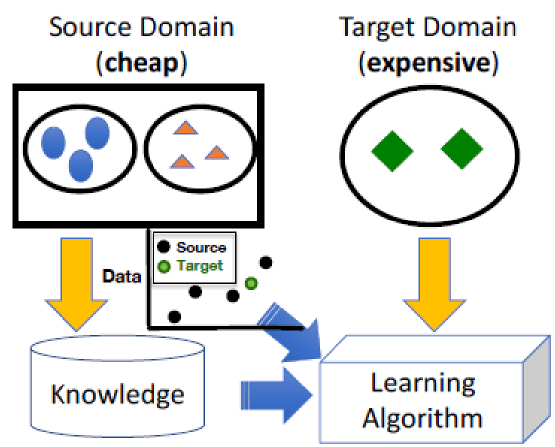
- Transfer Learning
  - Collecting/annotating data is typically expensive.
  - Assuming same training/test data distributions might not practical.
  - Improved learning & understanding in the (target) domain of interest by leveraging knowledge from a different source domain.



(Traditional) Machine Learning vs. Transfer Learning (cont'd)

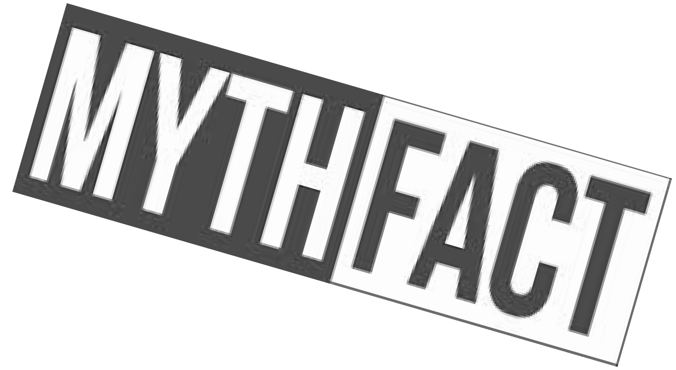
Models	Settings	Source & Target Domains	Source & Target Tasks
(Traditional) Machine Learning		Same	Same
Transfer Learning	Inductive (歸納)	Same	Different but related
	Transductive (直推)	Different but related	Same
	Unsupervised	Different but related	Same/different but related





S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE TKDE, 2010.

Myth or Fact #1



- 舉一反三 or 東施效顰?
  - When/how would TL be preferable?

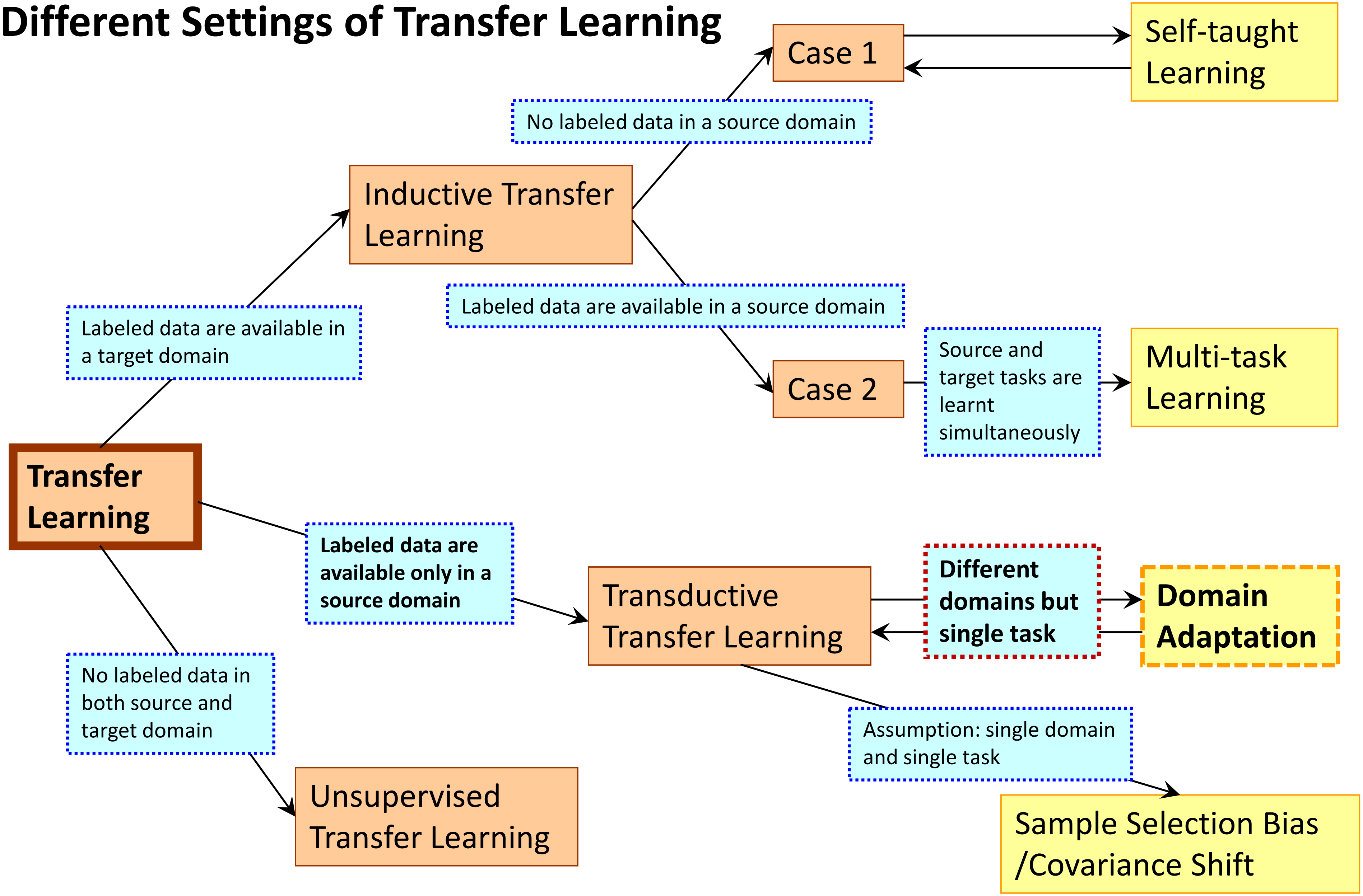
https://pic.pimg.tw/fscholars/1329718097-1051865464_n.jpg
https://i.ytimg.com/vi/INcuNyPgheo/maxresdefault.jpg

Myth or Fact #2

- As a data scientist/machine learner/AIer,
  you need *transfer learning* as a solution to your problem.

https://media.licdn.com/mpr/mpr/shrinknp_800_800/p/6/005/09b/03a/24ed5d2.jpg

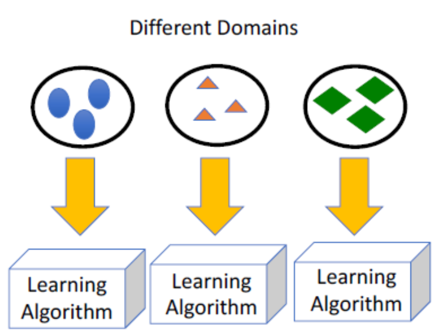
Different Settings of Transfer Learning

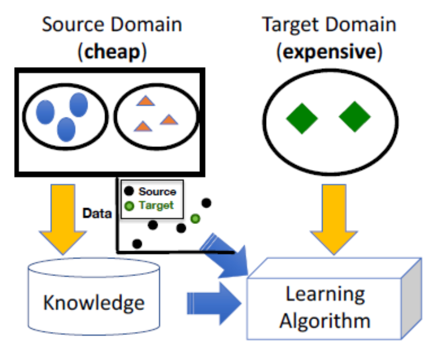


S. J. Pan and Q. Yang, “A survey on transfer learning,” IEEE TKDE, 2010.

Different Settings for Transfer Learning (cont'd)

Settings	Related Areas	Source Domain Labels	Target Domain Labels	Tasks
Inductive Transfer Learning	Multi-Tasks Learning	O	O	Regression, Classification
	Self-Taught Learning	X	O	Regression, Classification
Transductive Transfer Learning	Domain Adaptation, etc.	O	X	Regression, Classification
Unsupervised Transfer Learning		X	X	Clustering, Dimension Reduction





S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE TKDE, 2010.

Approaches for Transfer Learning

- Instance Transfer
  - Re-weight source-domain label instances for adaptation

Sugiyama *et al*. NIPS'07

Bickel *et al*. ICML'07

Kanamori *et al*. JMLR'09

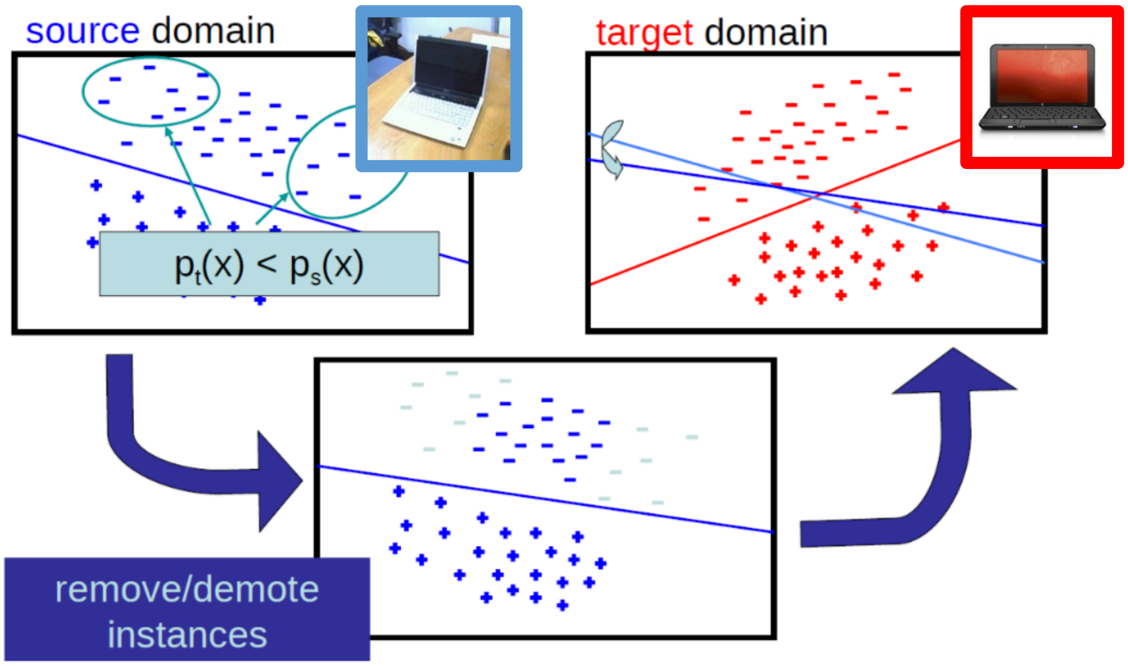
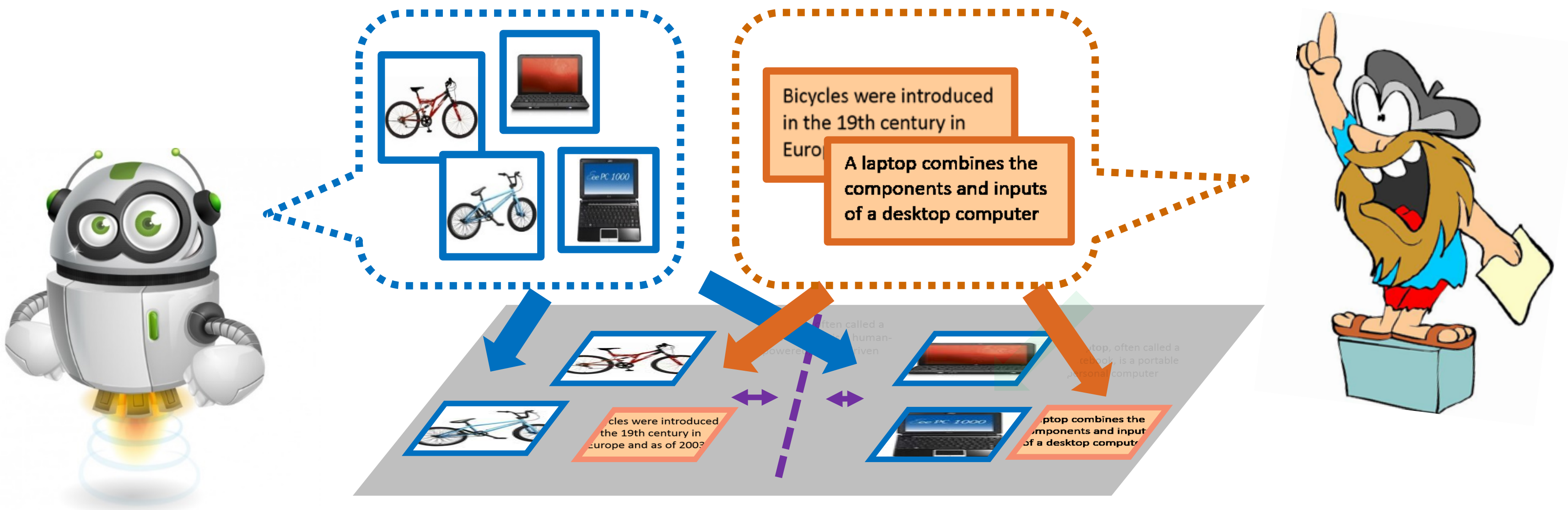


Image: Courtesy to Ming-Wei Chang.

Approaches for Transfer Learning

- Instance Transfer
  - Re-weight source-domain label instances for adaptation
- Feature Transfer (i.e., Common Feature Space)
  - Derive common feature representation for describing cross-domain data



Approaches for Transfer Learning

- Instance Transfer
  - Re-weight source-domain label instances for adaptation
- Feature Transfer
  - Derive common feature representation for describing cross-domain data
- Parameter Transfer
  - Discover shared learning model parameters for cross-domain data
  - E.g., Domain-adaptive SVM (IEEE PAMI'10)

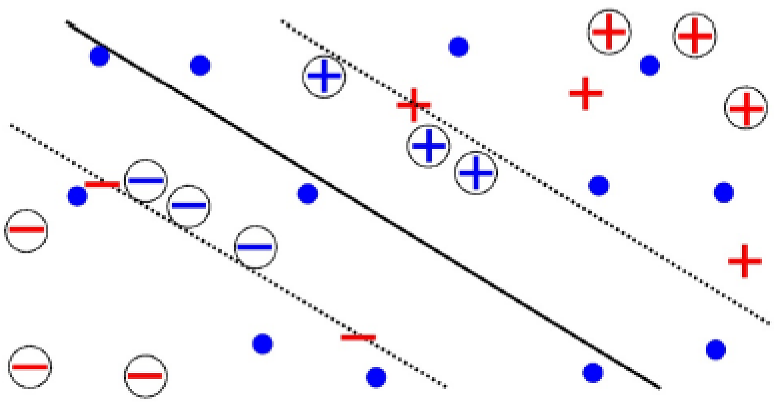


Image: Courtesy to Amaury Habrard.

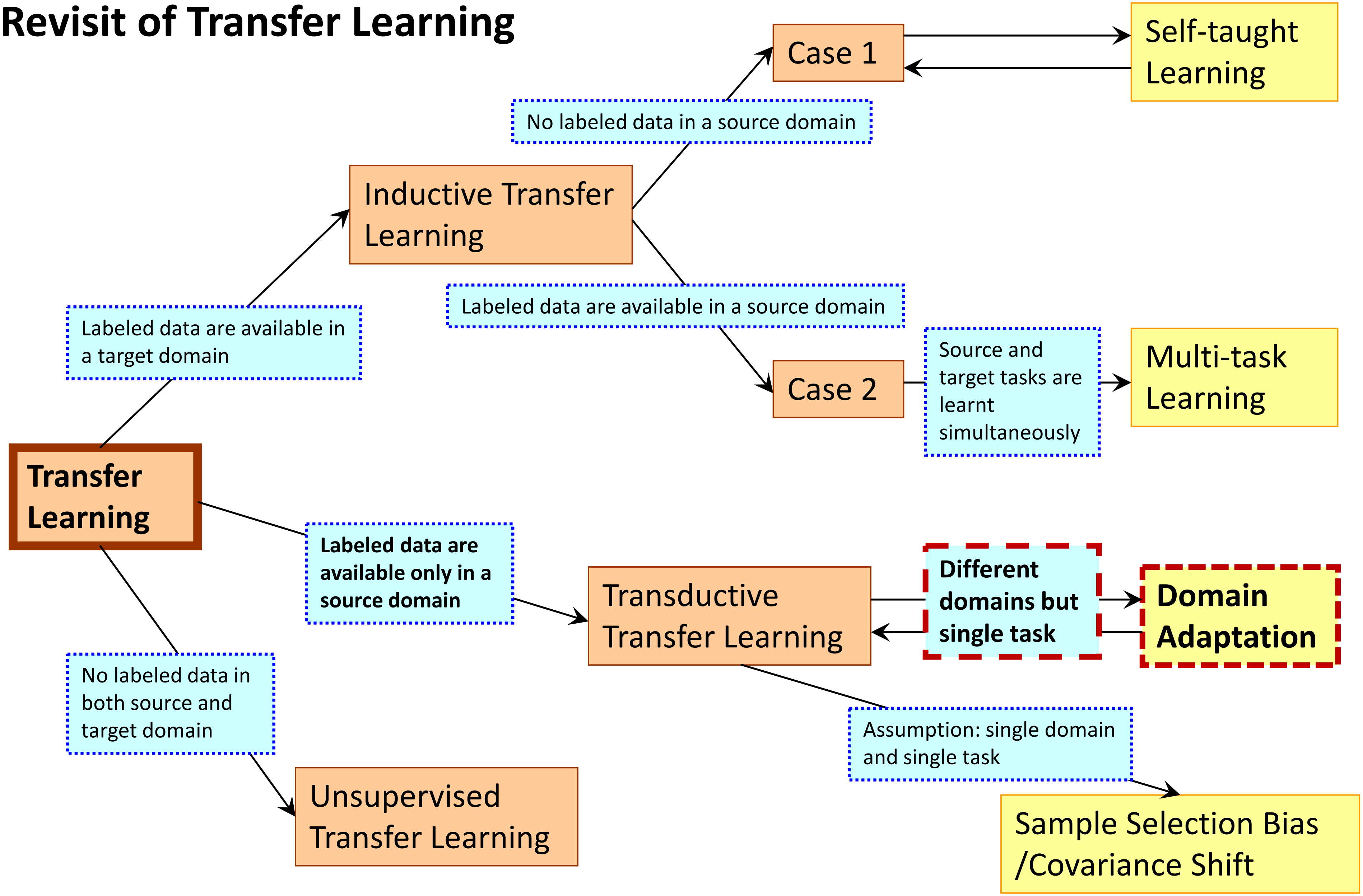
Approaches for Transfer Learning (cont'd)

- Instance Transfer
  - Re-weight source-domain label instances for adaptation
- Feature Transfer
  - Derive common feature representation for describing cross-domain data
- Parameter Transfer
  - Discover shared learning model parameters for cross-domain data
- Relational Knowledge Transfer
  - Build mapping of relational knowledge between cross-domain data

Methods	Inductive Transfer Learning	Transductive Transfer Learning	Unsupervised Transfer Learning
Instance Transfer (Instance Reweighting)	O	O	
Feature Transfer (Common Feature Representation)	O	O	O
Parameter/Model Transfer	O		
Relational knowledge Transfer	O		

S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE TKDE, 2010.

Revisit of Transfer Learning



S. J. Pan and Q. Yang, “A survey on transfer learning,” IEEE TKDE, 2010.

Domain Adaptation

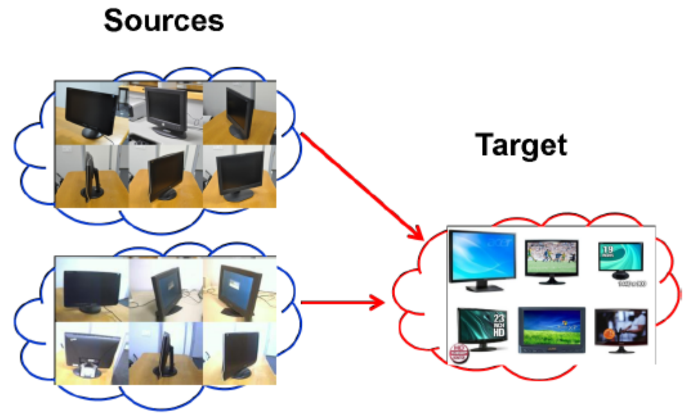
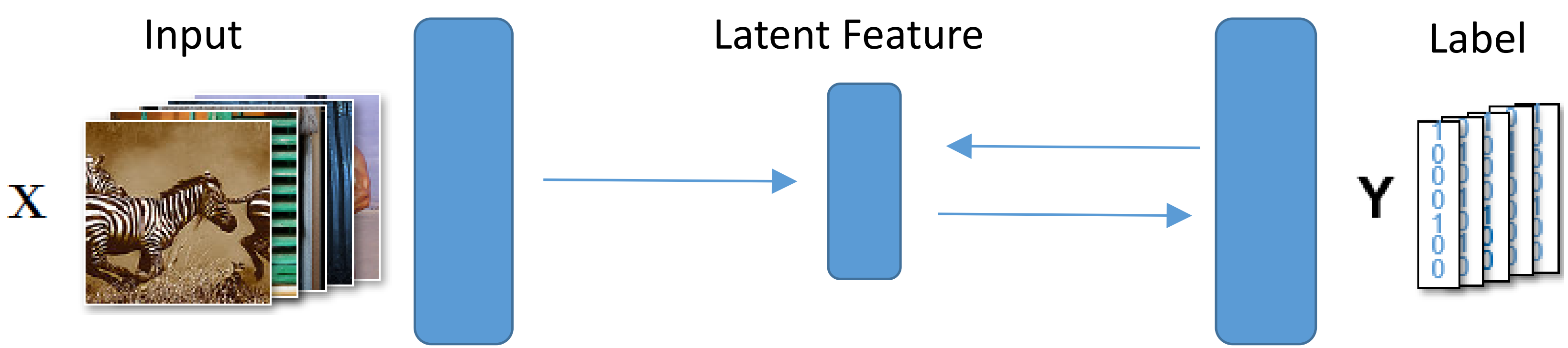


Image: Courtesy to S.J. Pan

- What's DA?
  - Leveraging info from one or more source domains,
    so that the same learning task in the target domain can be addressed.
  - Typically all the source-domain data are labeled.
- Settings
  - Semi-supervised DA: few target-domain data are with labels.
  - Unsupervised DA: no label info available in the target-domain.
    (shall we address supervised DA?)
  - Imbalanced DA: fewer classes of interest in the target domain
  - Homogeneous vs. heterogeneous DA
  - Let's see some examples…

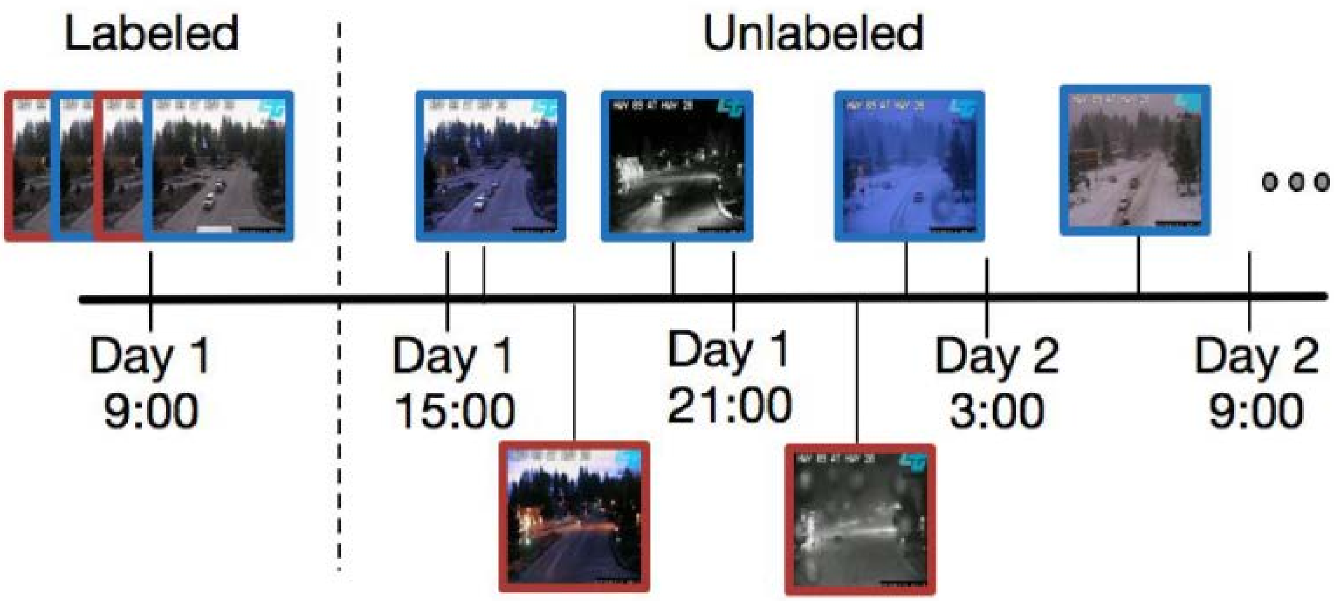
Revisit of Some Relevant Keywords

- Why we care about Transfer Learning?
  - Machine learning
  - Big data (or data science)
  - Deep learning
- How to train machine learning (or deep learning) models for solving transfer learning tasks under different settings?
  - E.g., semi-supervised or unsupervised TL via deep neural networks?



More TL Tasks

- Beyond standard classification, we might need to address continuous adaptation tasks.



- Beyond standard classification, we might need to address continuous adaptation tasks.

Input image at "blue hour" (just after sunset) | A database of time-lapse videos | Hallucinate at night

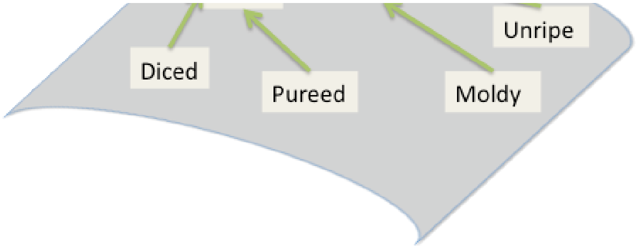
Shih et al., Data-driven Hallucination of Different Times of Day from a Single Outdoor Photo, SIGGRAPH 2013

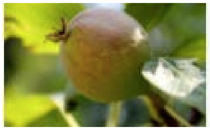
- Beyond standard classification, we might need to address continuous adaptation tasks.

Input image at "blue hour" (just after sunset)

A database of time-lapse videos

Hallucinate at night

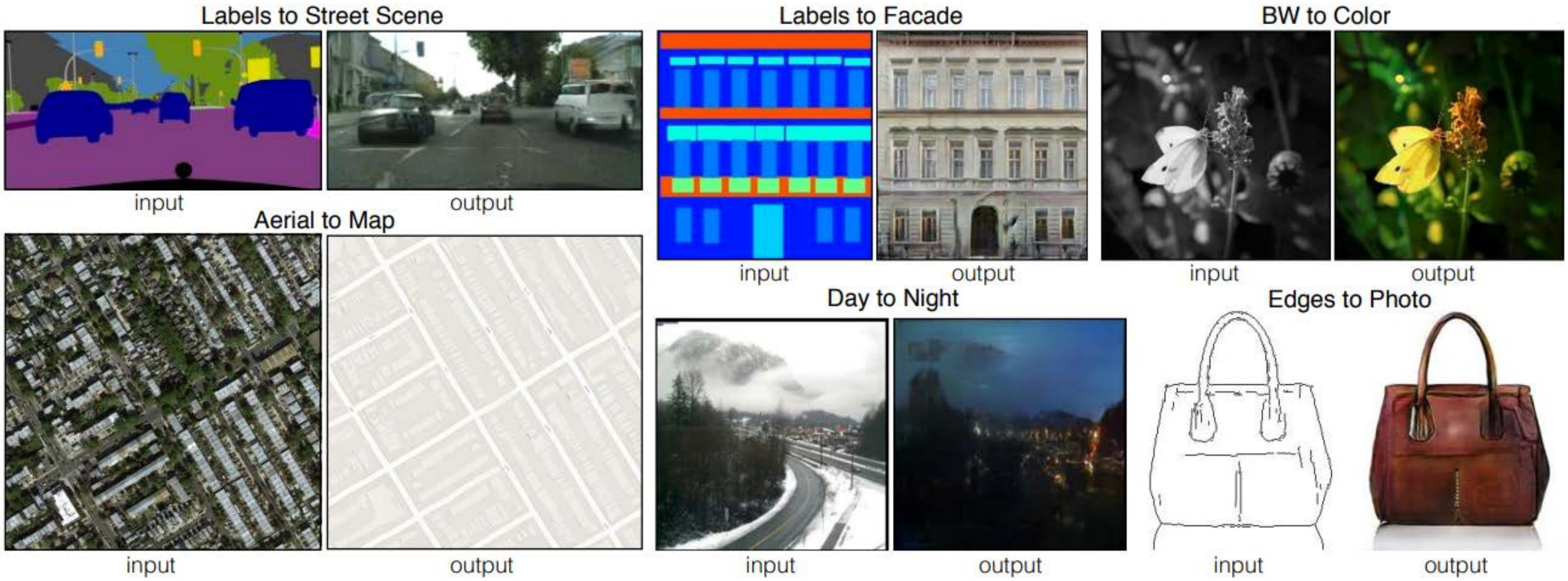




Lee, Wei, and Wang, Learning Grassmann Manifolds for Object State Discovery, IEEE ICASSP, 2017

More ML Tasks

- Beyond standard classification, we might need to address image translation/manipulation/style transfer tasks.



Zhu et al. "Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks." *CVPR* 2017.

- Beyond standard classification, we might need to address image translation/manipulation/style transfer tasks.

- Beyond standard classification, we might need to address image translation/manipulation/style transfer tasks.

Input

Let’s Take a Break…

- Transfer Learning
  - Introduction to Transfer Learning (TL)
  - Challenges in Transfer Learning
  - TL for Visual Recognition
  - TL for Visual Synthesis and Manipulation