AI & Deep Learning

Von-Wun Soo

Prof. Department of Computer Science

National Tsing Hua University

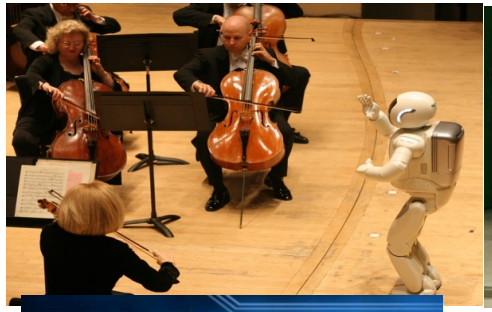
What is AI?

- The major concerns of AI:
  - Can machines think?
  - Can machines learn?
  - Can machines have common sense?
  - Can machines understand natural language? (semantics, intention, etc.)
  - Can machines understand images? (recognize objects, activities, etc.)
  - Can machine possess emotion or moral senses? (the value systems as human) ...

- Two major AI techniques:
  - Heuristic Search and Knowledge representation
- How to evaluate intelligence of a system?
  - Turing Test

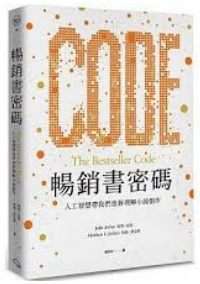
Turing Test

- 人類是有智慧的生物，但是人類所製造出來的機器是否也能有同等的智慧，甚至超越人類？
- 早在1950年英國科學家涂林(Turing)就提出了涂林測試概念來比較機器與人類的智慧。
- 通過涂林測試可以推測機器與人類具備至少相同的智慧。




TOYOTA



GLOBMEDIA
NEWS
Channel
WALL ST

CODE
The Bestseller Code
暢銷書密碼

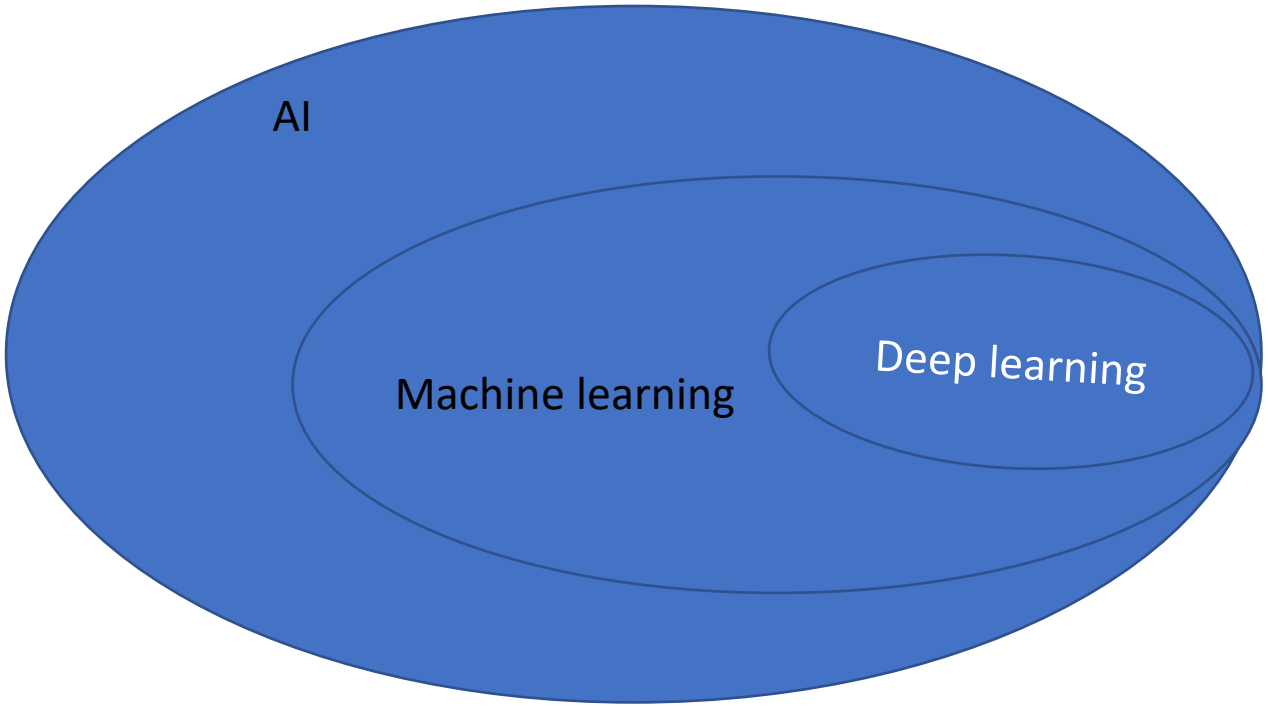



Google DeepMind
Challenge Match
8 - 15 March 2016
AlphaGo


THINK
DENKEN
$1,000
$2,000
$400
KEN
WATSON
BRAD
Michael Caine 77%
Major Critical 11%
Marine Corps 10%

KUKA


Deep learning & AI



What is deep Learning?

- A merge of traditional machine learning, neural networks, and statistical learning.
- Boosted by big data analysis and GPU hardware of Nvidia.
- Initiated around late 2000s, and by "learning for *deep belief nets" paper of Hinton, Osindero and Teh in 2006.*

- Pro: Show many successful and break thru in speech recognition, image recognition, natural language processing, drug discovery, etc.
- Con: Lack of accountability. Theory, ad hoc. No explanation on logical inference and causality. No guarantee converge, speed, and approximation accuracy.

Universal approximation theorem (Hornik et al., 1989; Cybenko, 1989)

- A feedforward network with a linear output layer and at least one hidden layer with any "squashing" activation function (such as the logistic sigmoid activation function) can approximate any Borel measurable function from one finite-dimensional space to another with any desired non-zero amount of error, provided that the network is given enough hidden units .

- However, we are not guaranteed that the training algorithm will be able to learn that function.

Why deep is better?

- Single layer NN is theoretically enough in universally
- approximately learning any function…(universal approximation theorem)

- Why multi-layers?
- Divide and conquer?
- Functional differentiation? (分層分工)
  - Easy to differential functions at different layers and share the features at different level of abstraction.
- More efficient in learning?
  - Reducing computational cost of representing some function.
- Allow feature sharing and transfer at different level of abstraction in different tasks and domains.

Why does deep learning tick?

- End-to-end training and learning without human intervention. (just feed big data)
- Allow to deal with big data with stochastic gradient descent SGD weight update. (Of course rely on GPU/TPU)
- Allow transfer learning by abstracted feature-sharing at some layer of neural networks.
- Competitive with human performance in many tasks that human are good at.

機器會下棋嗎? (Question before 2016)

- 自從1997年IBM研發出會下西洋棋的超級電腦「深藍」(DeepBlue)，戰勝了當年的國際西洋棋世界冠軍卡斯帕羅夫。西洋棋棋賽方面深藍可以說已經通過了涂林測試。

- 過去中國圍棋由於搜索空間大（branching factor)與search depth (敵我各約 200 多步）電腦尚無法打敗圍棋高手。

- 電腦圍棋進展　2010 4Dan ; 2011 5 Dan ; 2012 6 Dan
- 　2013 圍棋程式 Z en, 在日本9段高手讓三子之下已經打敗了 9 段高手。

2016 March 15 李世石9段 vs AlphaGo

2017 ALPHAGO zero 打敗 ALPHAGO Master

- 基於自己與自己比賽 不需仰賴人類圍棋知識
- Start from scratch
- 40 天訓練用了4 顆 google TPU
- 打敗所有人類高手與圍棋程式

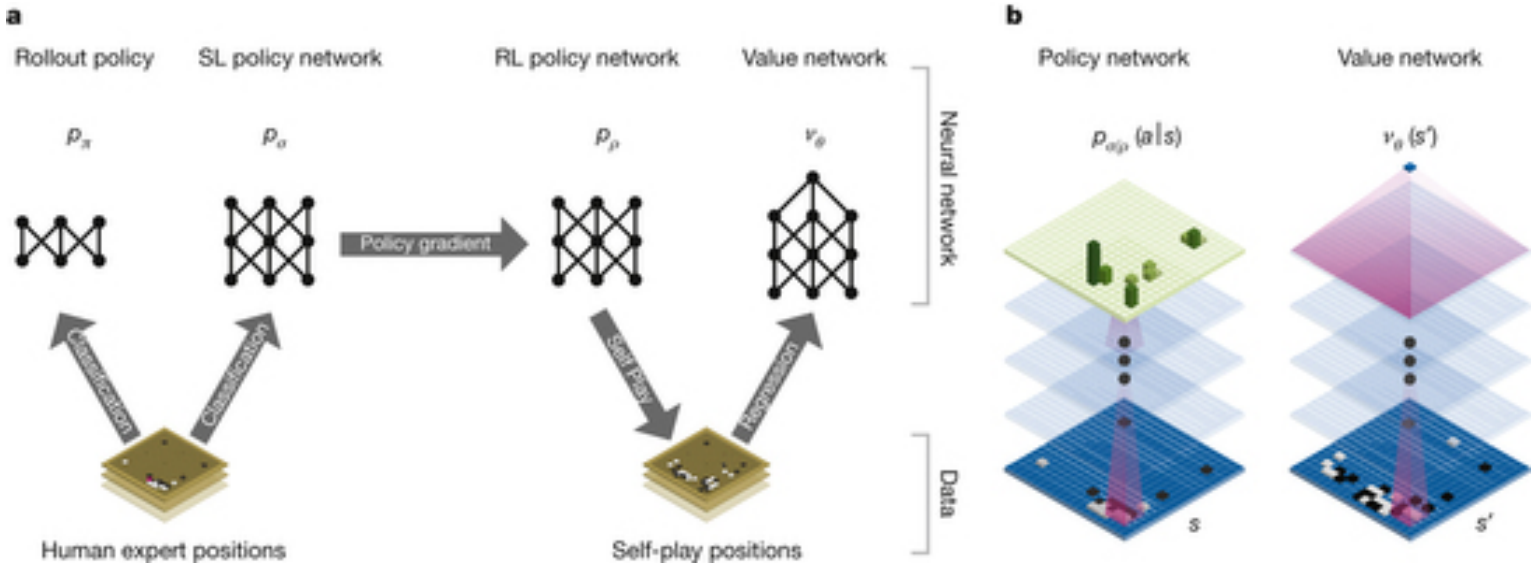
蒙特卡羅樹搜索法(Monte Carlo Tree Search)

- 圍棋程式採用了蒙特卡羅樹隨機搜索策略（Monte Carlo tree search) 使得圍棋程式達到與人類高手一較高下的程度
- Stochastic search
- Balance Exploration and Exploitation
- with limited computation resource
-

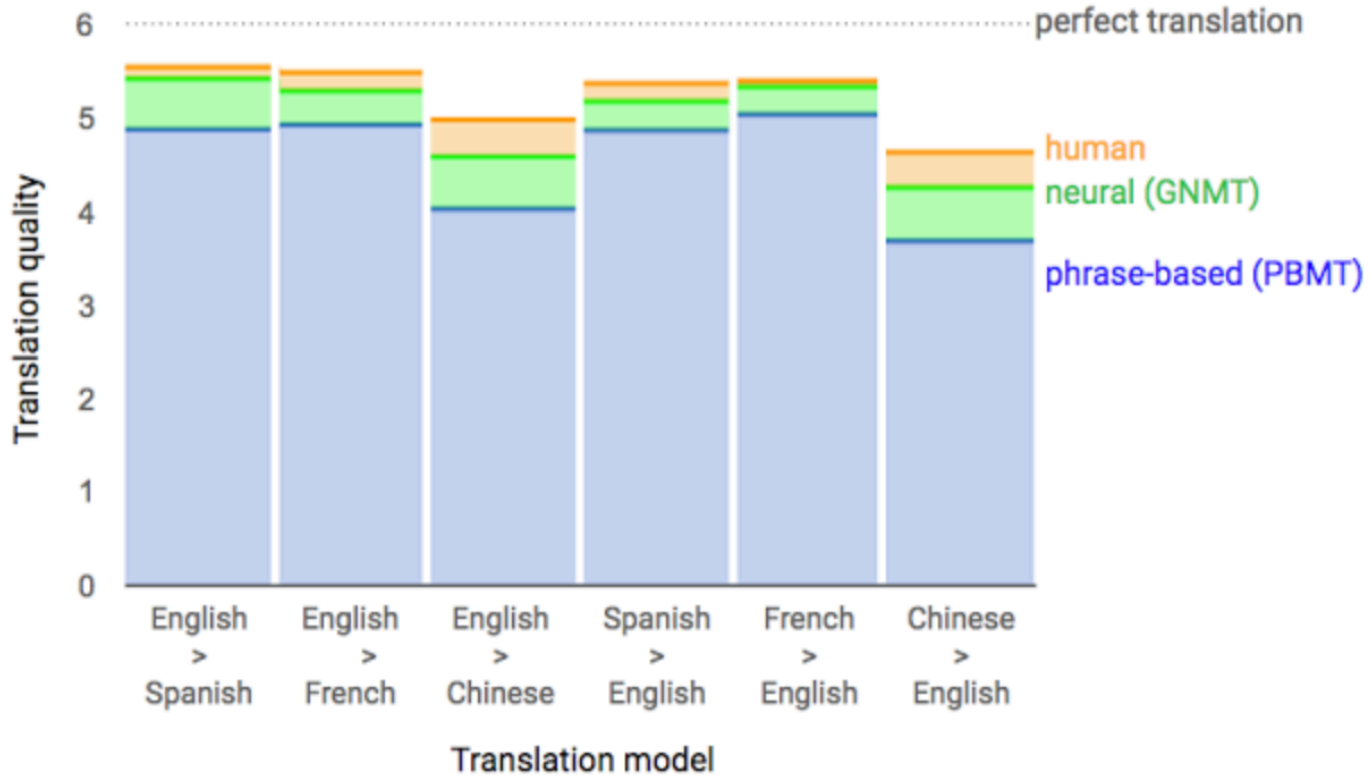
$$\mathrm{UCB} = (1-\alpha)\mathrm{S} + \alpha\sqrt{\frac{\ln v}{n_v}}$$

Deep Learning with neural networks

- Value networks vs. policy networks
- 利用深度類神經網路來沽算目前棋局黑白的輸贏
- 利用深度類神經網路來計算目前棋局的下一步



Machine translation achievement



Google Sequence-to-Sequence based model performance. Source

World big companies compete in AI technologies

- Google DeepMind: AlphaGO
- IBM Watson: jeopardy game champion (Q/A)
- Facebook FAIR：self-taught chatbots, machine translation
- Microsoft: **Cortana personal assistant,** draw picture from text
- Amazone: drones
- Apple: automatic driving
- Baidu百度: NLP、speech、auto-driving
- Alibaba: reading comprehension test Q/A achieves(slightly over) human performance

- Failure cases：
- IBM collaborated with MD Anderson cancer center for making Watson become a cancer doctor but failed.
- Facebook: Chatbot out of control, chatting bots create their own languages that cannot be understood by human even they are speaking English. Sometimes it learned to cheat.

Autonomous driving car

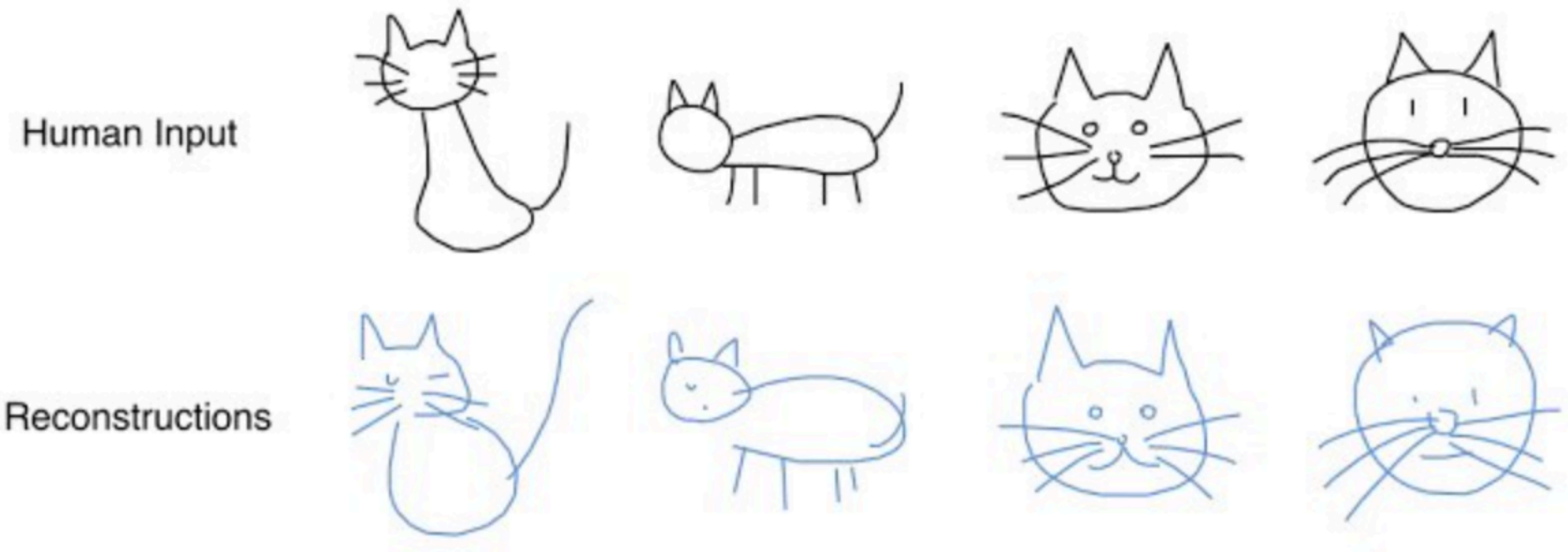
Lip reading

- Google Deepmind, in collaboration with Oxford University
- trained on a television dataset, was able to surpass the professional lip reader from the BBC channel.

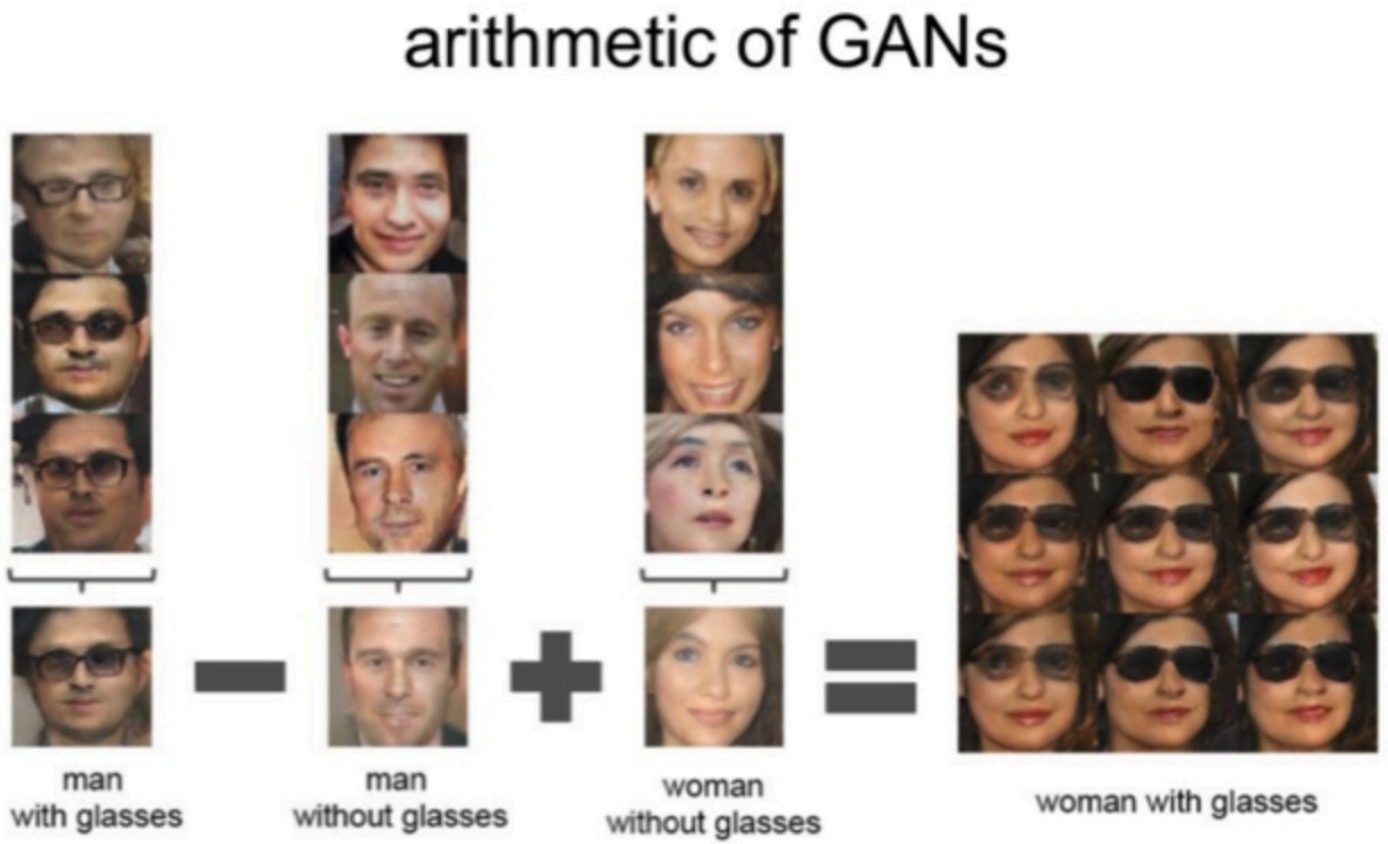
OCR: Google Maps and Street View

- Google Brain Team reported on how they introduced a new OCR (Optical Character Recognition) engine into its Maps, through which street signs and store signs are recognized.

SketchRNN: teaching machines to draw

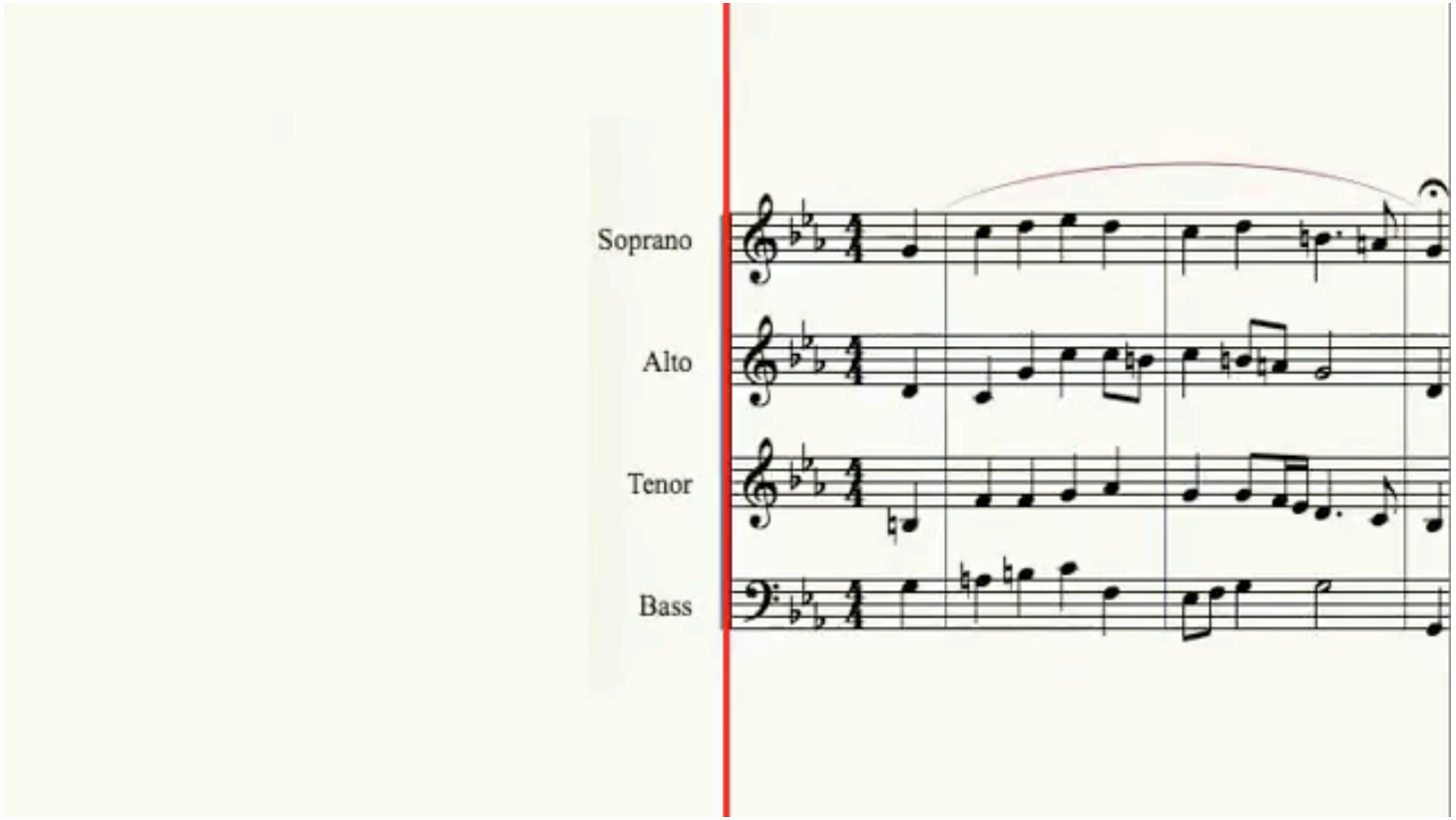


Arithmetic logic on images based on GAN



Learning Art Style transfer

Can computer compose music?
Compose bach using deep learning



WaveNet plays piano at master level

Can machines compose poems? TAAI2016 Von-Wun Soo et.al.

電腦	由機械組成，會讓你上癮 生活的部分，有趣也先進 使你玩遊戲，使你看新聞 連接線製成，不喜歡灰塵	**手**	可以開瓶蓋，用來抱抱枕 有能力煮菜，能用來關門 雙手會剪紙，四肢的部分 雙手能煮菜，人體的部分
哭	太傷心釀成，被甩了釀成 想要被安慰，離別所釀成 哀傷所導致，失敗所造成 會想被安慰，孤獨所造成	**錢**	能夠買文具，用來追女生 一種必需品，因此要節省 經濟的部分，貨幣的象徵 印鈔機製成，能用來旅行
狗	喜歡到處跑，痛恨陌生人 會讓你高興，不想要壞人 狗能一起玩，狗會當家人 是一種守衛，很想要主人	**跑步**	奔跑會很喘，跑步會很累 會讓你變瘦，會想要喝水 你將會流汗，有能力很累 使你喝飲料，讓你想喝水
食物	好吃的也香，拿來吃也香 代表必需品，懼怕被浪費 由白米製成，會讓你花錢 用麵粉製成，有能力腐爛	**老師**	是一種職稱，有能力學習 會須要課本，會用到粉筆 代表啟發者，喜歡問問題 討厭催作業，想要出考題
朋友	讓你想信賴，會引起共鳴 是一種良藥，有能力慶生 不喜歡背叛，有辦法傾聽 是一種知己，很喜歡友情	**水**	是一種溶劑，有能力載舟 氫和氧組成，能用來爬山 會帶來淹水，能拿來洗手 海的一部分，水會往下流

Can machines have common sense?
用矩陣分解法學習常識聯想

- Location -- Activity Association
- Learning Association between L and A given a set of observed association examples X

$$\text{Min}_{\arg H} \|X - L\,H\,A\|^2 + \alpha \|L\,H\,A_{xproj} - L\|^2 + \beta \|L\,H\,A_{yproj} - A\|^2 + \gamma \|H\|$$

- $\mathbb{X}$: observation matrix L; location matrix A: Activity matrix
- We extract 139 locations and 436 activities and 667 location-activity pairs from ConceptNet as training data.
- 667/(139x436) = 667/60604 = 1.1%

Predict location activity association

- amphitheatre # listen_to_concert,
- broadcast_studio # produce_tv_weather_report,
- classroom # teach_student,
- concert_hall # listen_to_concert,
- department_store # buy_furniture,
- factory # create_product,
- farm # produce_food,
- french_restaurant # buy_french_food
- garden # clean_garden,
- hair_salon # get_hair_color_change,
- hair_salon # have_hair_cut,
- health_food_store # buy_healthy_cookie,
- japanese_restaurant # eat_sushi,
- museum # study_historical_artifact,
- ocean # swim_in,
- post_office # send_letter,
- restaurant # eat_food,
- salon #do_hair, salon # have_hair_cut,
- television_studio # produce_television_program,
- ticket_office # meet_friend,
- university # learn_something,
- zoo # see_animal

Results of common sense association learning

The performance is evaluated in two aspects:

1) without counting unsure results

- Precision = 81 / 106 = 0.764,
- F-score = 0.478;

2) counting unsure results as correct

- Precision = 85 / 106 = 0.801,
- F-score = 0.485.

Music recommendation & prediction
Pan Lei and Von-Wun Soo

- LFM-1b Dataset包含了英國音樂網站Last.fm中(類似KKBOX)，120,322活躍用戶的聆聽紀錄約十億筆（2013年1月 至2014年八月)
- 用使用者聆聽音樂點閱紀錄的大數據 LFM-1b Dataset 與Million song dataset 的交集 dataset
- 訓練Neural Collaborative Filtering model,1gpu 7hr/epoch,
- 在一個曾點閱與99個未點閱的曲子混合的預測下, 在前10名出現的正確率HR 達 93%（高於前文獻88% 約 5% 以上), 前五名出現HR 達88.2 %

Can machines recognize jokes? Peng-Chen Yu and Von-Wun Soo NAACL2018

Table 1: Statistics of four datasets

Dataset	#Pos	#Neg	Type	Lang
16000 One-Liners	16000	16002	One-liner	EN
Pun of the Day	2423	2403	Pun	EN
Short Jokes	231657	231657	All	EN
PTT Jokes	1425	2551	Political	CH

Table 4: Example Sentences

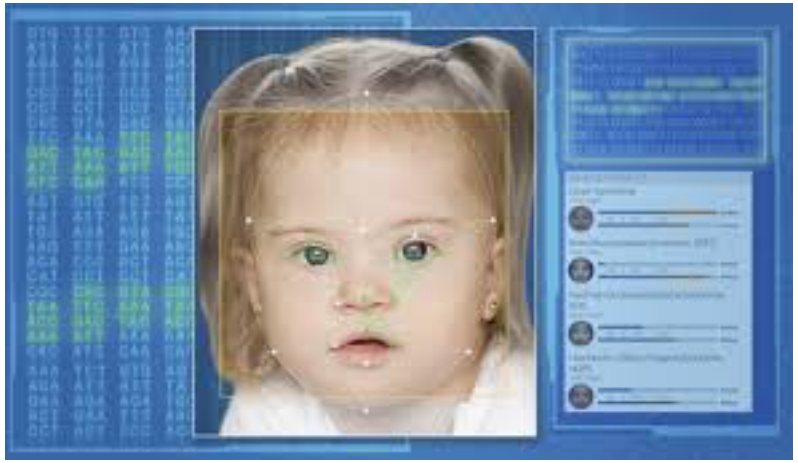
	Sentence
TP	when he gave his wife a necklace he got a chain reaction
TN	the barking of a dog does not disturb the man on a camel
FP	rats know the way of rats
FN	it's a fact taller people sleep longer in bed

Table 2: Comparison of Different Methods of Humor Recognition

		16000 One-Liners				Pun of the Day			
		Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
Previous Work									
	Word2Vec+HCF	0.854	0.834	0.888	0.859	0.797	0.776	0.836	0.705
	CNN					0.861	0.857	0.864	0.864
Our Methods									
	CNN	0.877	**0.899**	0.856	0.877	0.867	0.880	0.859	0.869
	CNN+F	0.892	0.896	0.928	0.898	0.892	0.886	0.907	0.896
	CNN+HN	0.885	0.877	0.902	0.889	0.892	**0.889**	0.903	0.896
	CNN+F+HN	**0.897**	0.872	**0.936**	**0.903**	**0.894**	0.866	**0.940**	**0.901**

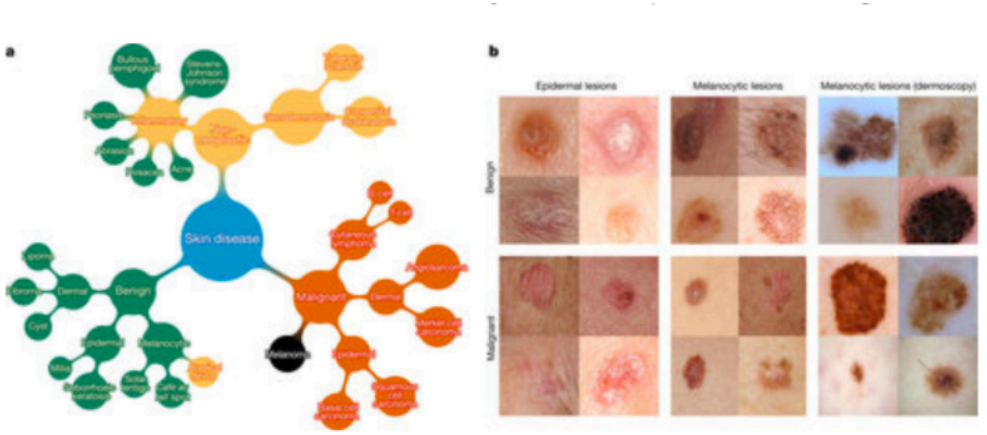
Face2Gene

- Using face recognition to associate phenotype features of faces with in rare genetic disorders.



Medical image diagnosis

- Skin cancer diagnosis and a dermatologist inspects a lesion of interest with the assistance of a dermatoscope (handheld microscope).
- Stanford's deep learning algorithm was tested against 21 board-certified dermatologists who reviewed a reported 370 images.
- Results showed the deep algorithm had the same ability as the 21 dermatologists.

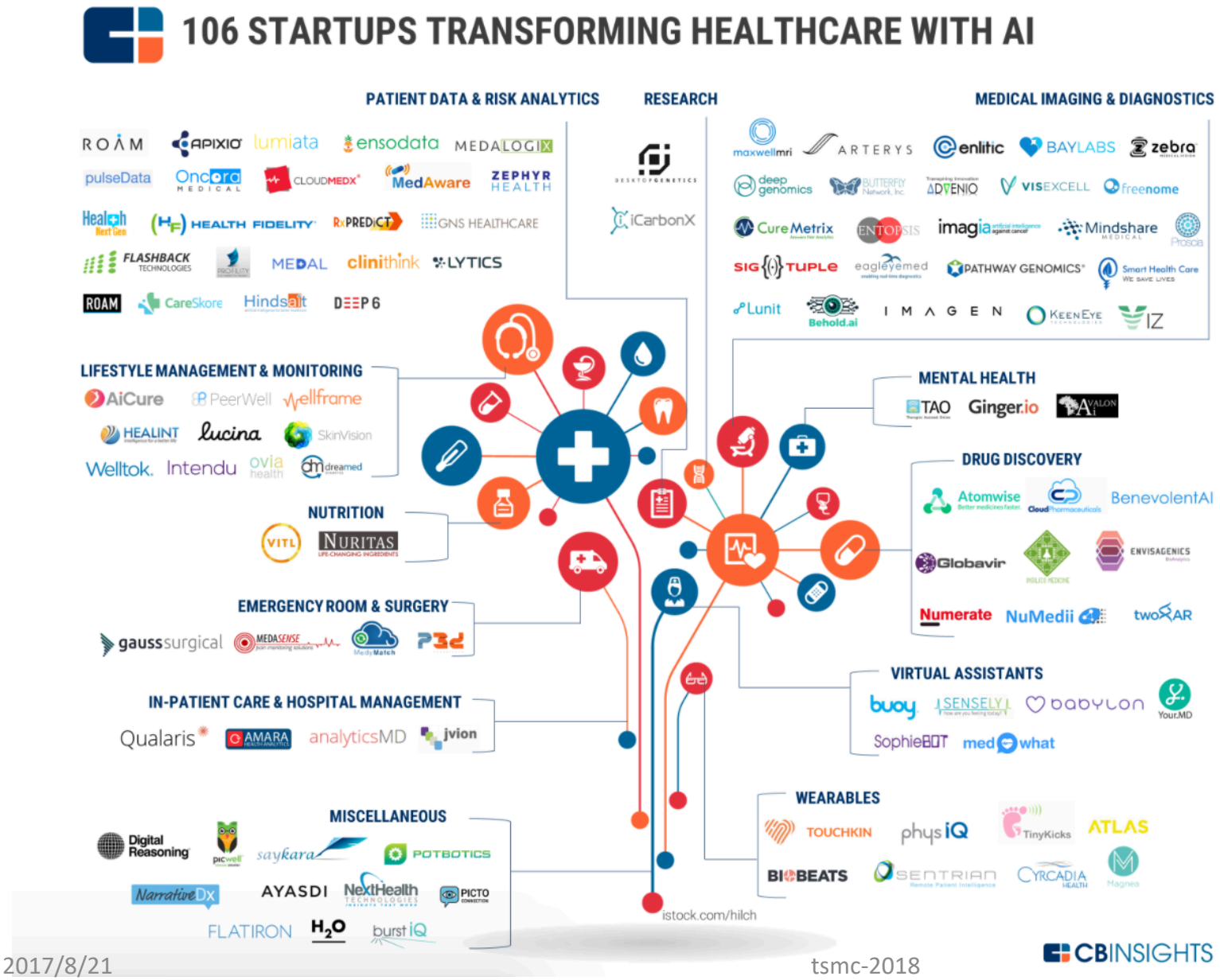


From the journal Nature: Dermatologist-level classification of skin cancer with deep neural networks

Market value on AI in healthcare

- Artificial Intelligence in healthcare market value ~ 80 億美元 by 2022

~ Estimated by Market&Markets May 03, 2017


106 STARTUPS TRANSFORMING HEALTHCARE WITH AI
PATIENT DATA & RISK ANALYTICS
RESEARCH
MEDICAL IMAGING & DIAGNOSTICS
ROAM
APIXIO
lumiata
ensodata
MEDALOGIX
pulseData
Oncora MEDICAL
CLOUDMEDX
MedAware
ZEPHYR HEALTH
Health Next Gen
HEALTH FIDELITY
RxPREDICT
GNS HEALTHCARE
FLASHBACK TECHNOLOGIES
MEDAL
clinithink
LYTICS
ROAM
CareSkore
Hindsait
DEEP6
DESKTOP GENETICS
iCarbonX
maxwellmri
ARTERYS
enlitic
BAYLABS
zebra
deep genomics
BUTTERFLY Network, Inc.
ADVENIO
VISEXCELL
freenome
CureMetrix
ENTOPSIS
imagia
Mindshare MEDICAL
Prosca
SIGTUPLE
eagleyemed
PATHWAY GENOMICS
Smart Health Care
Lunit
Behold.ai
IMAGEN
KEENEYE
IZ
LIFESTYLE MANAGEMENT & MONITORING
AiCure
PeerWell
wellframe
HEALINT
lucina
SkinVision
Welltok.
Intendu
ovia health
dreamed
NUTRITION
VITL
NURITAS
MENTAL HEALTH
TAO
Ginger.io
AVALON AI
DRUG DISCOVERY
Atomwise
CloudPharmaceuticals
BenevolentAI
Globavir
ENVISAGENICS
Numerate
NuMedii
twoXAR
EMERGENCY ROOM & SURGERY
gausssurgical
MEDASENSE
MedyMatch
P3d
VIRTUAL ASSISTANTS
buoy
SENSELY
babylon
YourMD
SophieBOT
medwhat
IN-PATIENT CARE & HOSPITAL MANAGEMENT
Qualaris
AMARA
analyticsMD
jvion
WEARABLES
TOUCHKIN
physiQ
TinyKicks
ATLAS
BIOBEATS
SENTRIAN
CYRCADIA HEALTH
Magnea
MISCELLANEOUS
Digital Reasoning
picwell
saykara
POTBOTICS
NarrativeDx
AYASDI
NextHealth TECHNOLOGIES
PICTO CONNECTION
FLATIRON
H2O
burst iQ
istock.com/hilch
CBINSIGHTS

Deep learning START-UPS in medicine

- **Google Deepmind in Health Care: a digital contact lens that could detect blood sugar level.**
- **IBM WatsonPaths h**elp physician to make decision and interpret from medical records.
- **Careskore:** start-up from Chicago, predict how likely a patient will be readmitted to a hospital, based on combination of clinical, labs, demographic and behavioral data**. In 2016, it got 4.3 million dollars fund.**
- **Freenome** diagnoses cancer from blood samples by finding correlations between cell-free DNA and cancers.

Medical image start-ups: X-ray, ECG, MRI, ultrasound, tomography

- **Enlitic**: Uses deep learning to analyze radiographs and CT and MRI scans. Outperformed four radiologists in detecting and classifying lung nodules as benign or malignant.
- Can interpret a medical image in milliseconds — up to 10,000 times faster than the average radiologist.
  - 50% better at classifying malignant tumors and had a false-negative rate (where a cancer is missed) of zero, compared with 7% for the humans.
- **Butterfly Network**: develop i-phone sized sanner that you could hold-up to a person's chest and see a vivid, moving, 3-D image of what's inside. Makes the imaging devices cheaper!

Medical image start-ups:
X-ray, ECG, MRI, ultrasound, tomography

- **Arterys:** The diagnostic software connects to a standard MRI machine to enable noninvasive quantification of cardiac blood flow. The deep learning analysis is based on a simple, 10-minute MRI scan. Raised **12 millions dollars fund from GE.**
- **Bay** uses deep learning technology, to help medical professionals to perform and interpret echocardiograms and may improve the treatment of heart diseases. **It got $5.5 million funding.**

Start-ups in Drug development

Speedup drug development from years to weeks

- **Atomwise: invested by Merck** uses deep learning to shorten the process of discovering new drugs, has raised **$51 million fund**. Examine 3D images from thousands of molecules and predict how molecules might act in the human body, including their potential efficacy as medication, toxicity and side effects.
- **Recursion Pharmaceuticals: raised $13M dollars**. The company identified novel uses for known drugs, bioactive compounds, and shelved pharma assets in rare genetic diseases.

Start-up in Human genome

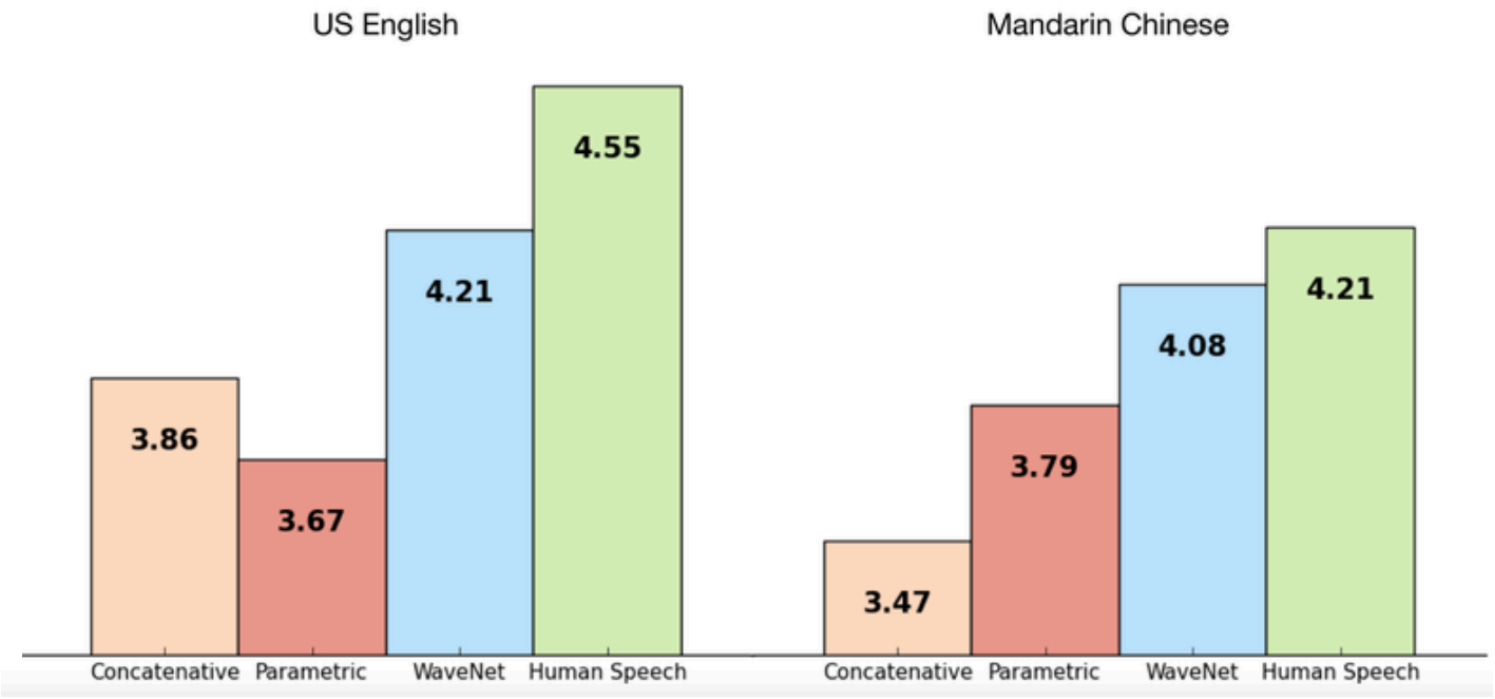
- **Deep Genomics**: uses deep learning to decode the meaning of the genome and predict the effects of a particular mutation based on hundreds of thousands of examples of other mutations.
- Turbine：adopt ‘systems biology’ to simulates the biological mechanisms of cancer and determines the most effective personalized therapies faster than any traditional healthcare service.

Speech interaction

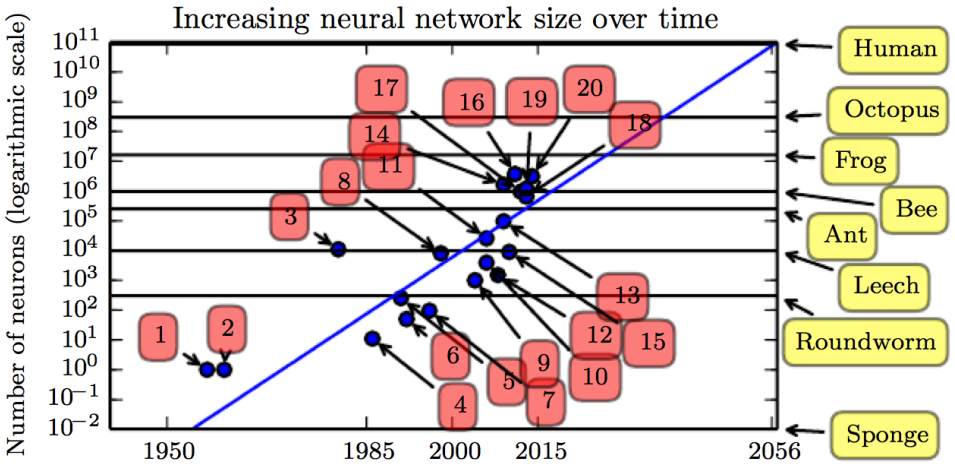
- Amazon’s Alexa, Apple’s Siri, Microsoft’s Cortana, or the many voice-responsive features of Google. Chinese search giant Baidu working on speech interface.
- Google use Chatbot to negotiate with a deal.

Text-2-speech generation

- End-to-end training on waveNet for Text-to-Speech

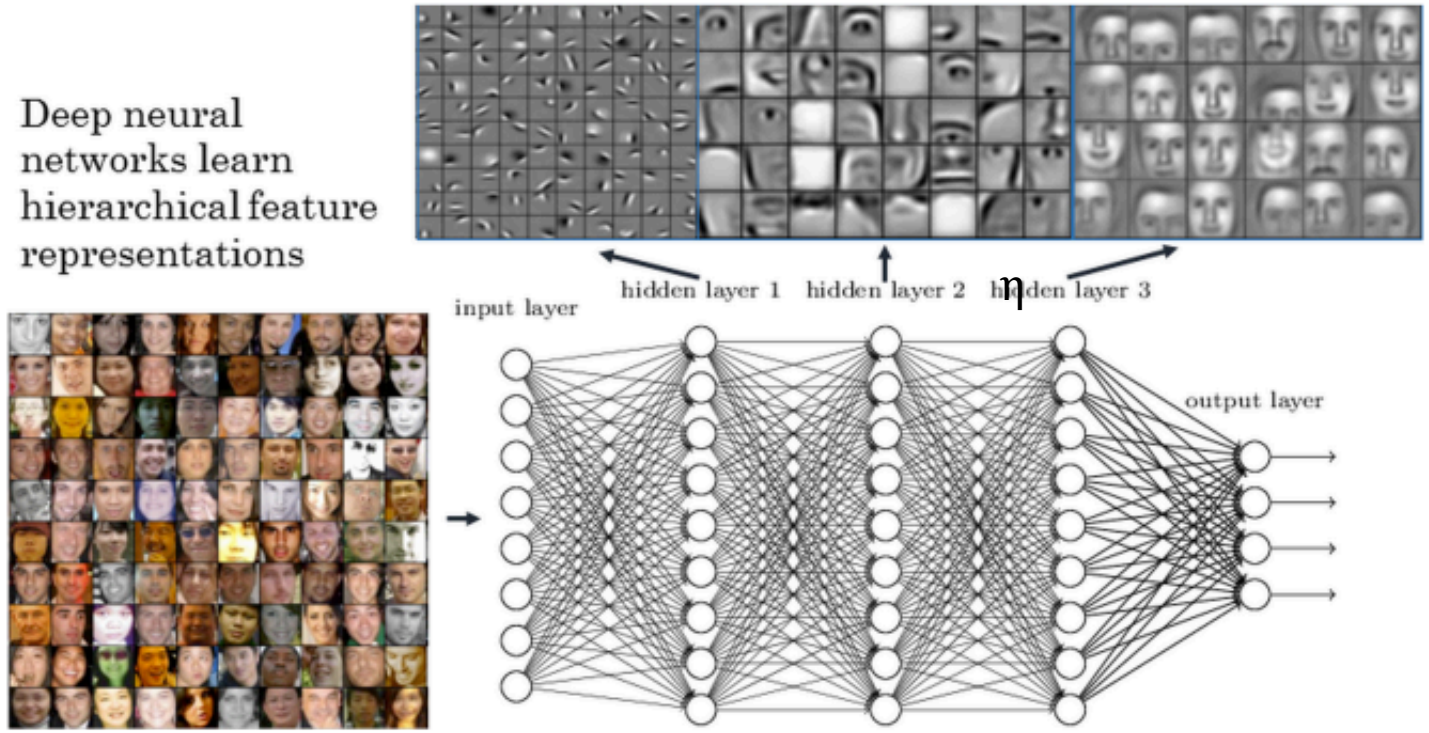


ARTIFICIAL neural networks vs BIOLOGICAL neural networks



1. Perceptron (Rosenblatt, 1958, 1962)
2. Adaptive linear element (Widrow and Hoff, 1960)
3. Neocognitron (Fukushima, 1980)
4. Early back-propagation network (Rumelhart et al., 1986b)
5. Recurrent neural network for speech recognition (Robinson and Fallside, 1991)
6. Multilayer perceptron for speech recognition (Bengio et al., 1991)
7. Mean field sigmoid belief network (Saul et al., 1996)
8. LeNet-5 (LeCun et al., 1998b)
9. Echo state network (Jaeger and Haas, 2004)
10. Deep belief network (Hinton et al., 2006)
11. GPU-accelerated convolutional network (Chellapilla et al., 2006)
12. Deep Boltzmann machine (Salakhutdinov and Hinton, 2009a)
13. GPU-accelerated deep belief network (Raina et al., 2009)
14. Unsupervised convolutional network (Jarrett et al., 2009)
15. GPU-accelerated multilayer perceptron (Ciresan et al., 2010)
16. OMP-1 network (Coates and Ng, 2011)
17. Distributed autoencoder (Le et al., 2012)
18. Multi-GPU convolutional network (Krizhevsky et al., 2012)
19. COTS HPC unsupervised convolutional network (Coates et al., 2013)
20. GoogLeNet (Szegedy et al., 2014a)

Deep neural networks



Deep learning models/architectures

- Autoencoder: feature extraction, dimension reduction, semantic encoding
- DCNN (deep convolutional neural network): hierarchical local feature extractions, receptive fields problems
- RNN (LSTM): handling long distance dependency problems in sequence information
- Boltzmann machines: Probabilistic computation on optimizing energy function over connected visible and hidden states
- Generative Adversary networks: Competing games between Discriminator and Generator