北大医学-教育论坛 2018

人工智能(AI)与医学

赵屹 M.D. Ph.D.
中国科学院计算技术研究所
中科信息产业研究院精准医学中心

目录

Contents

01

个人介绍

简介

赵屹，中科院计算所生物信息P.I. 博导；中科信息产业研究院精准医学研究所所长

北京大学医学部/清华大学

主要从事多组学生物信息分析研究、数据挖掘/机器学习/人工智能算法在医学应用研究

近十年在*Cell stem cell, Cell metabolism, Nature structural and molecular biology*，*Journal of Clinical Investigation, Journal of Hepatology, Genome Research, Trends in genetics, Nucleic Acids Research*等国际著名期刊发表论文40余篇，

其中以第一作者及通信作者发表论文40篇，总SCI引用过2000次，单篇引最高345次，超过百次引用的文章7篇。

European Research Council非编码RNA领域基金评审人；国际RNA联盟**RNAcentral专家**成员。**Frontier in Genetics编委**；高校教材《分子诊断学》(中国医药科技出版社)第3版编委。

中国生物工程学会计算生物学与生物信息学专业委员，北京医学遗传学会委员，中华医学会心血管病学分会精准心血管病学学组委员，世中联计算医学委员会委员

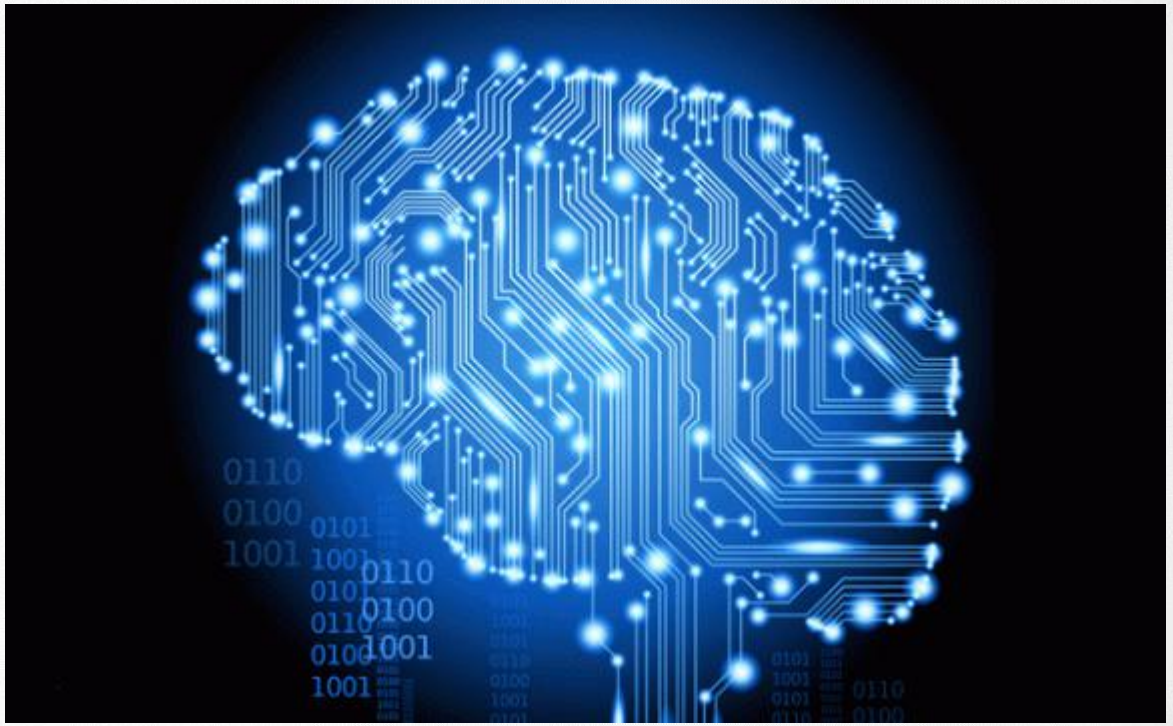
02

人工智能概念

什么是人工智能

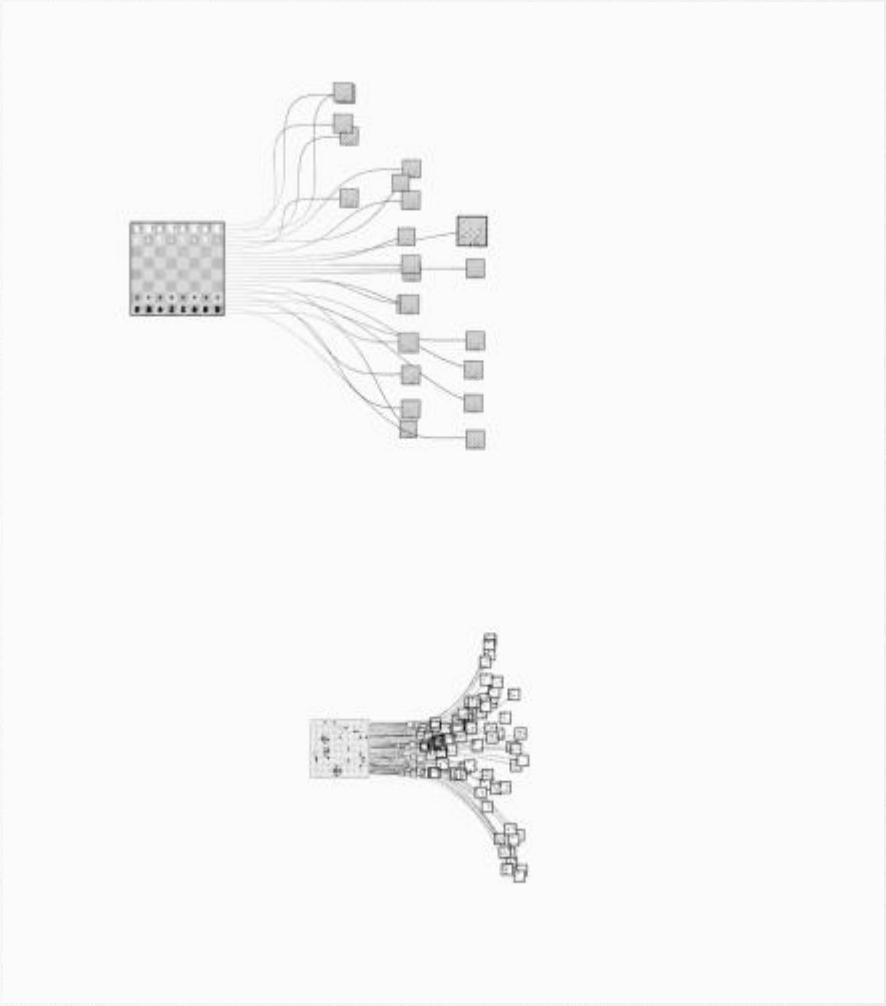
人工智能（**AI**. Artificial Intelligence）亦称机器智能，是指有人工制造出来的系统所表现出来的智能。

——维基百科



人工智能的热潮——智能围棋

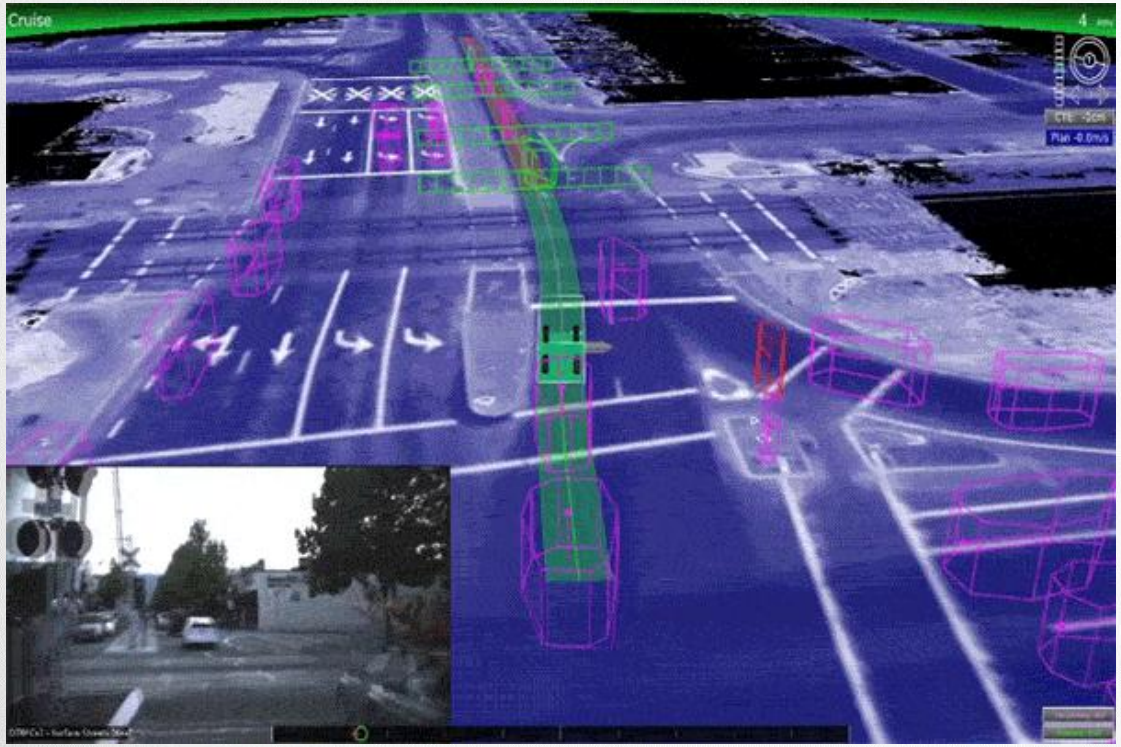
AlphaGo



人工智能的热潮——自动驾驶

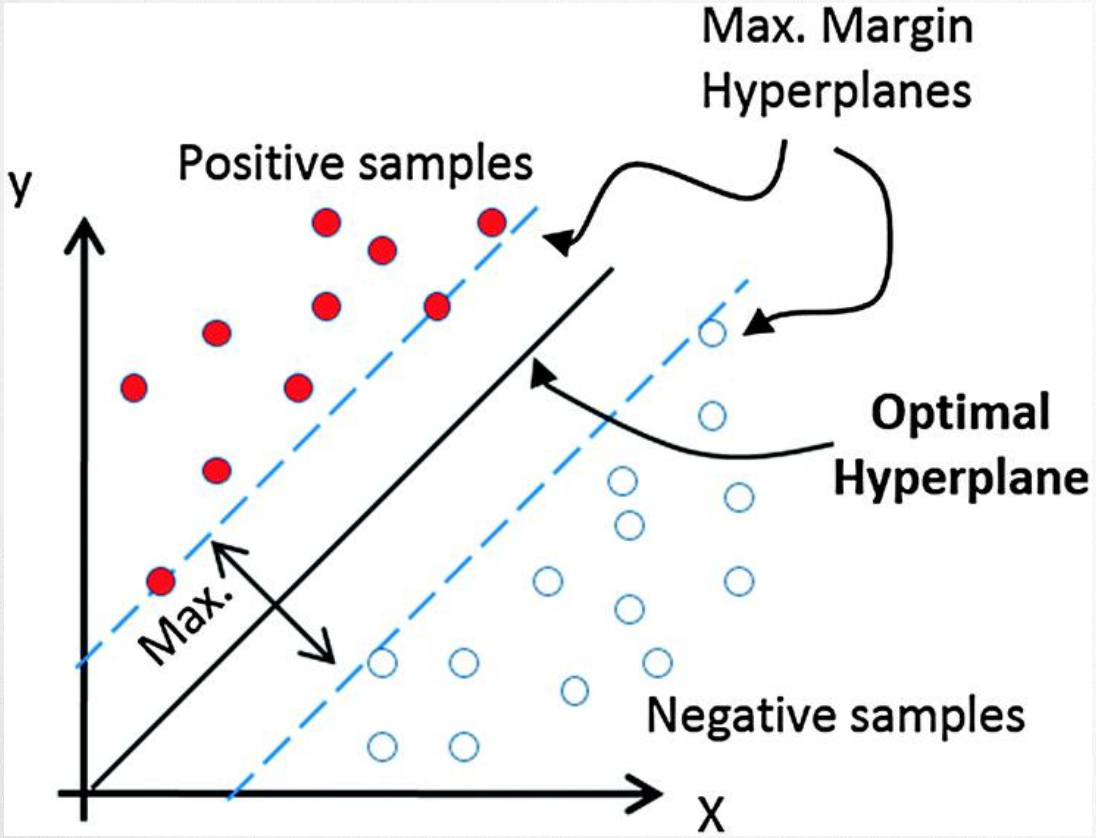
Google

TESLA

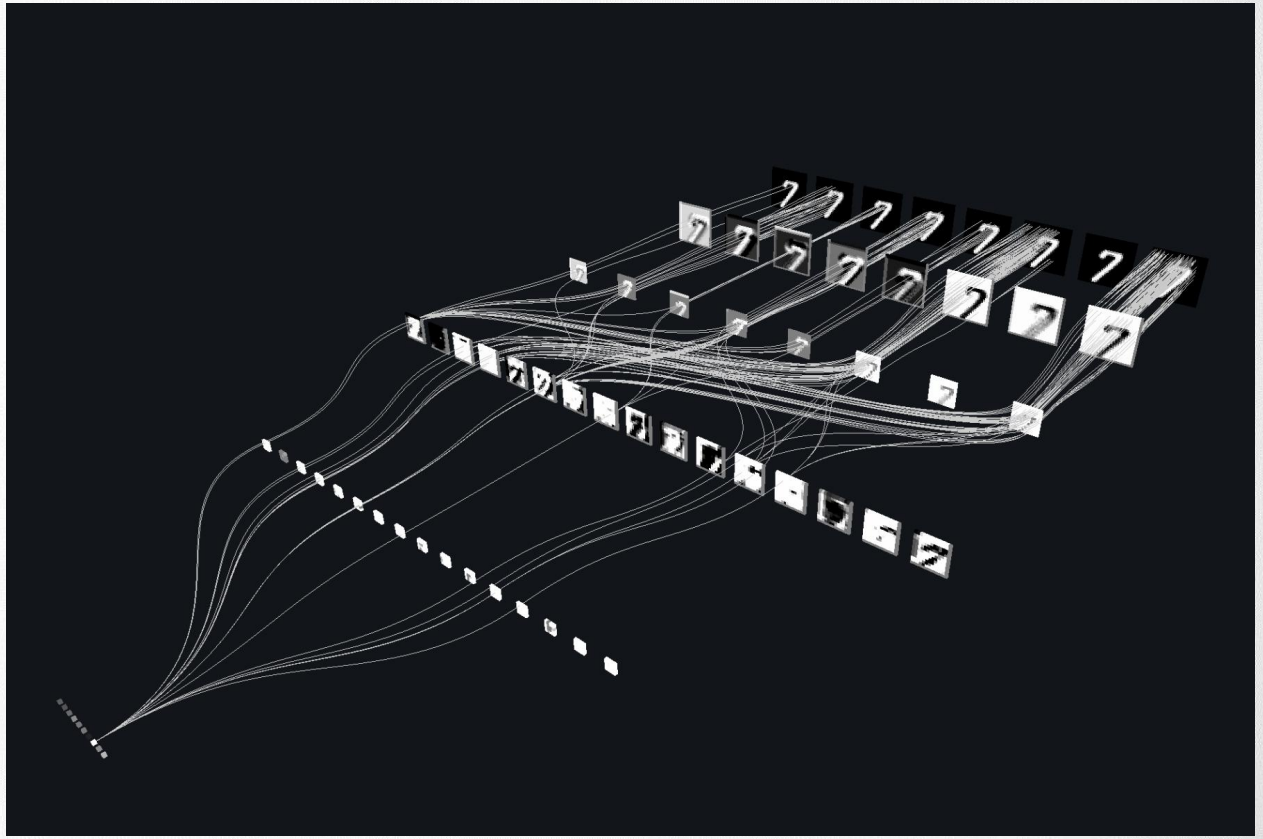


人工智能关键算法：深度学习

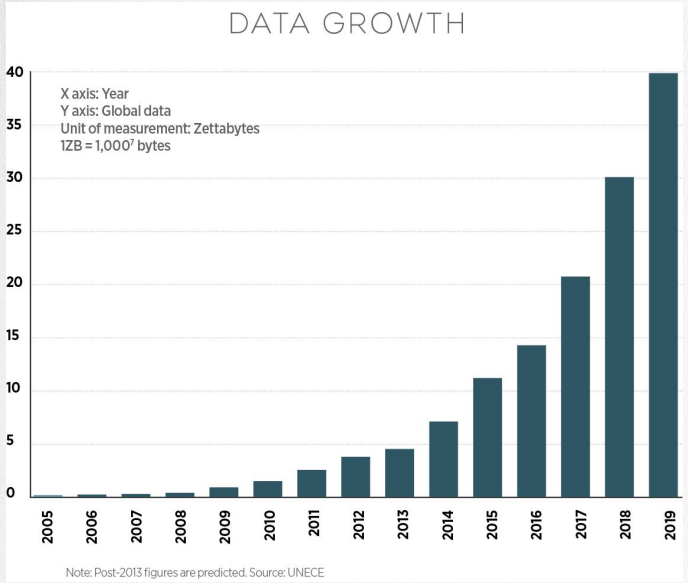
机器学习
参数N：$N < 10^2$

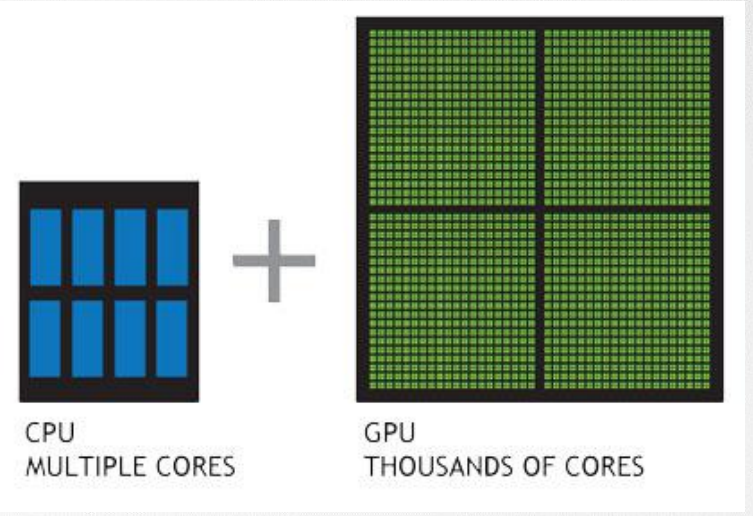


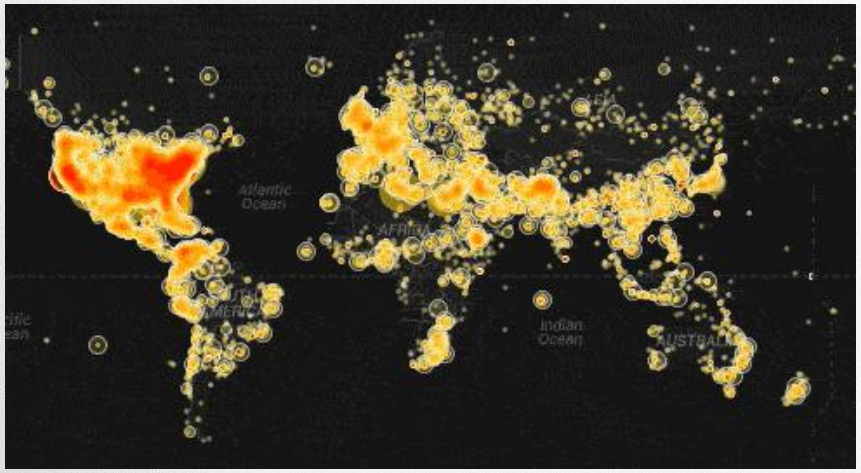
深度学习
参数N：$10^9 \leq N$

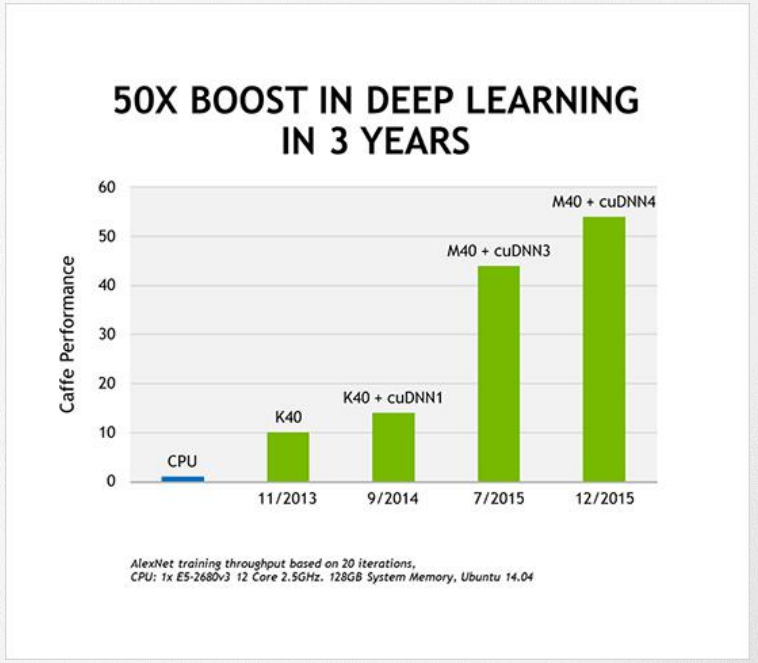


大数据和计算能力提升引爆这一次AI浪潮









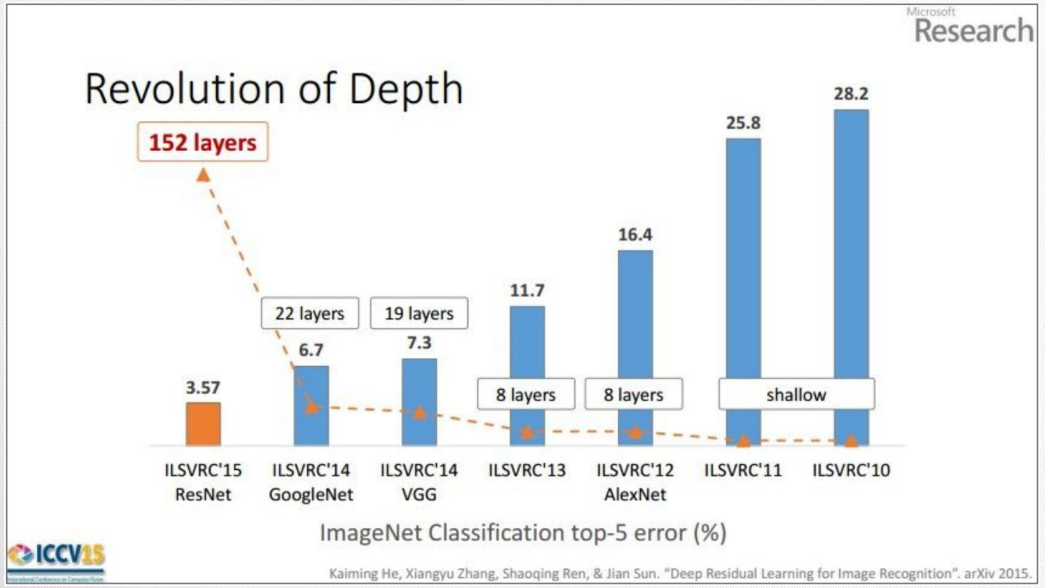
人工智能的应用—图片分类

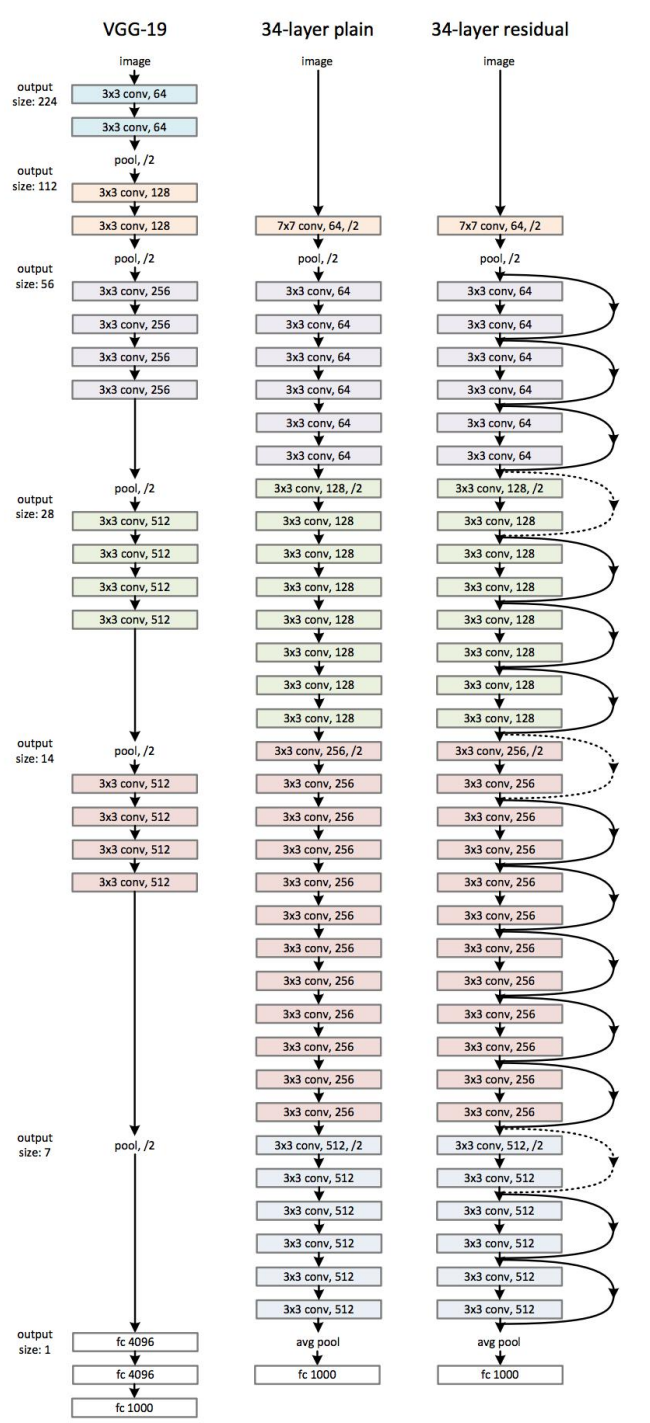
ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
  - 1.2 M train
  - 100k test.

4



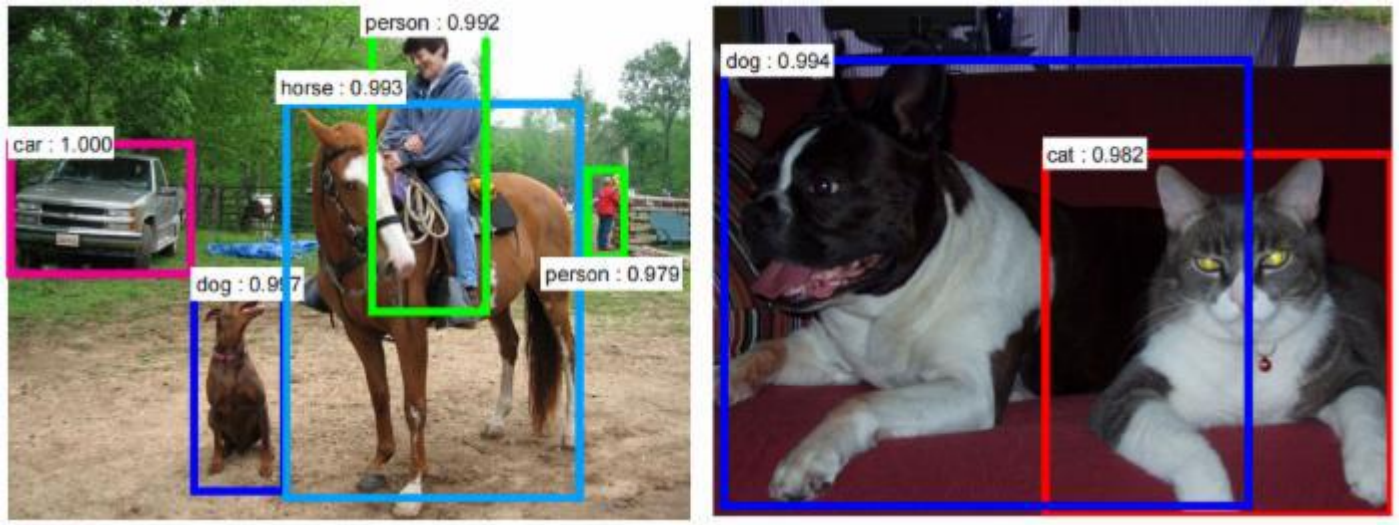


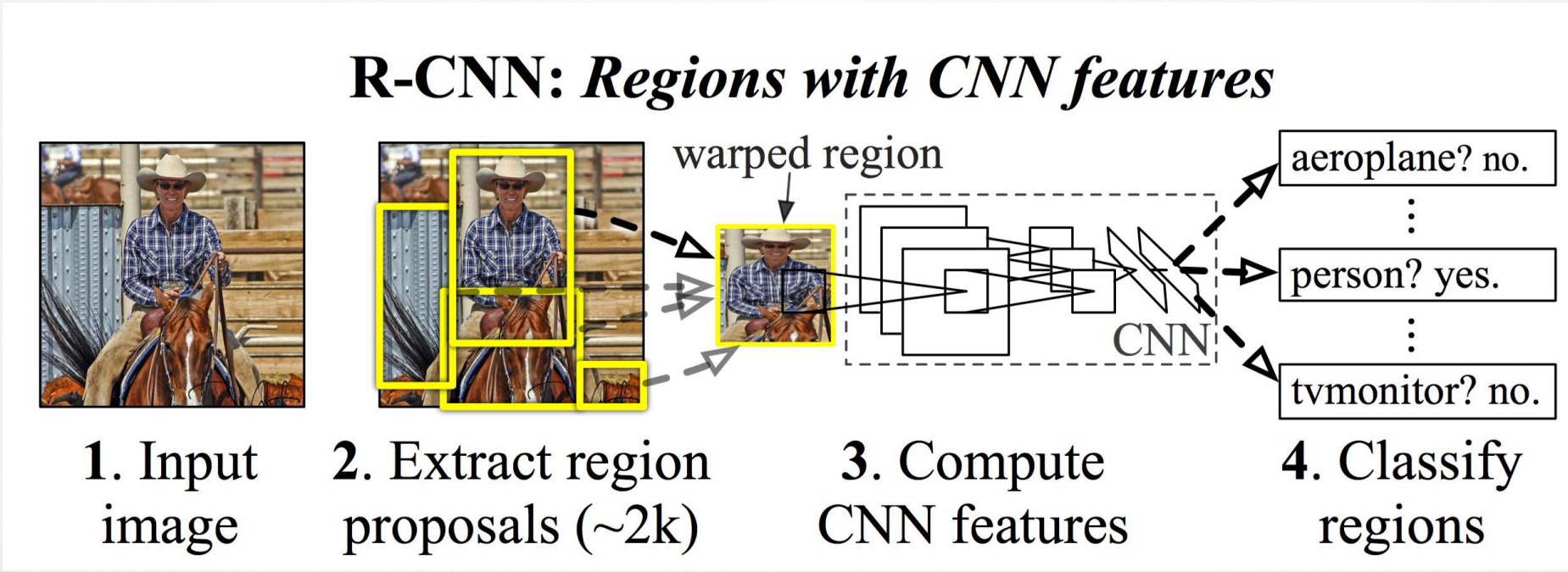
人类对自然图片的分类准确率大约为95%

卷积神经网络的分类准确2015年大约为96%

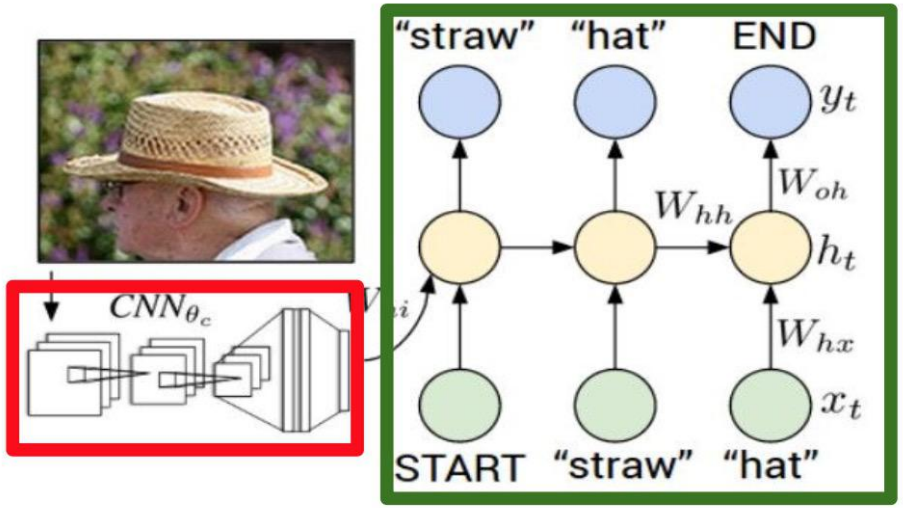
深度学习在自然图片识别方面达到了人类水平

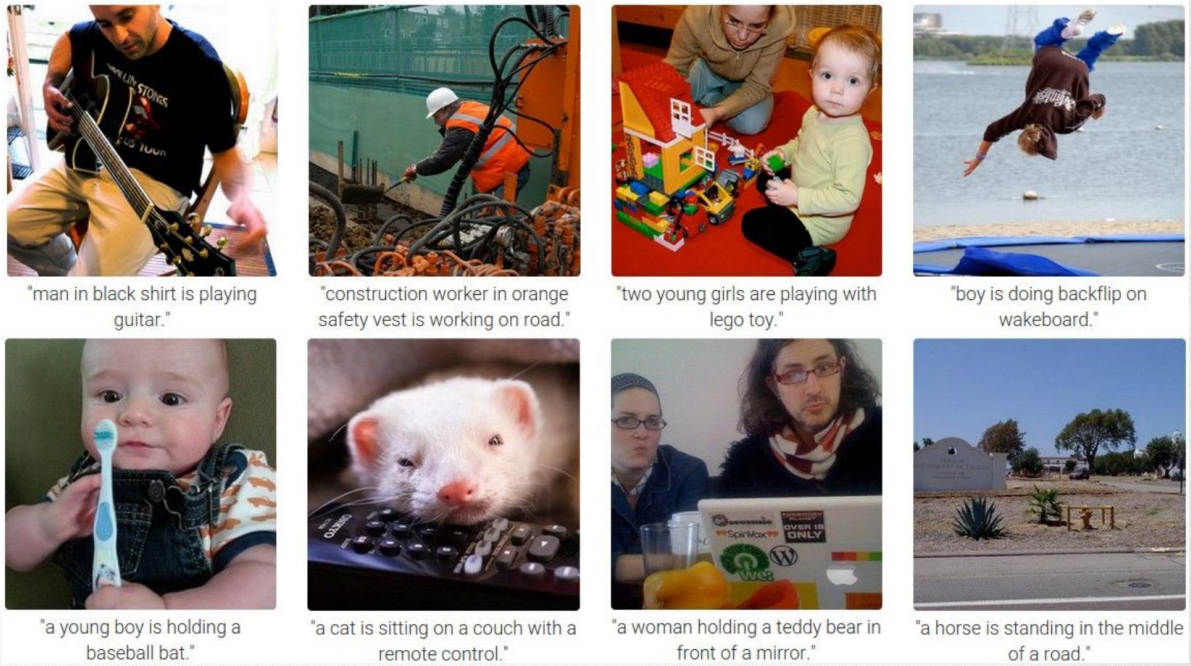
人工智能应用—目标检测


person : 0.992
horse : 0.993
car : 1.000
dog : 0.9
person : 0.979
dog : 0.994
cat : 0.982


R-CNN: Regions with CNN features
warped region
CNN
aeroplane? no.
person? yes.
tvmonitor? no.
1. Input image
2. Extract region proposals (~2k)
3. Compute CNN features
4. Classify regions

人工智能应用—看图说话


Recurrent Neural Network
"straw"
"hat"
END
START
"straw"
"hat"
y_t
h_t
x_t
W_{oh}
W_{hh}
W_{hx}
CNN_{θ_c}
Convolutional Neural Network

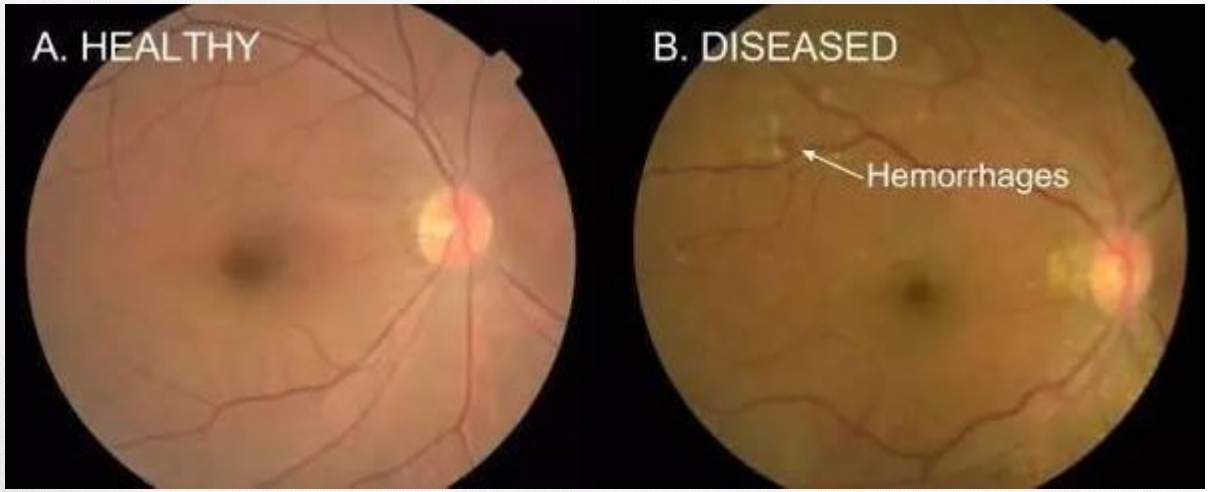

"man in black shirt is playing guitar."
"construction worker in orange safety vest is working on road."
"two young girls are playing with lego toy."
"boy is doing backflip on wakeboard."
"a young boy is holding a baseball bat."
"a cat is sitting on a couch with a remote control."
"a woman holding a teddy bear in front of a mirror."
"a horse is standing in the middle of a road."

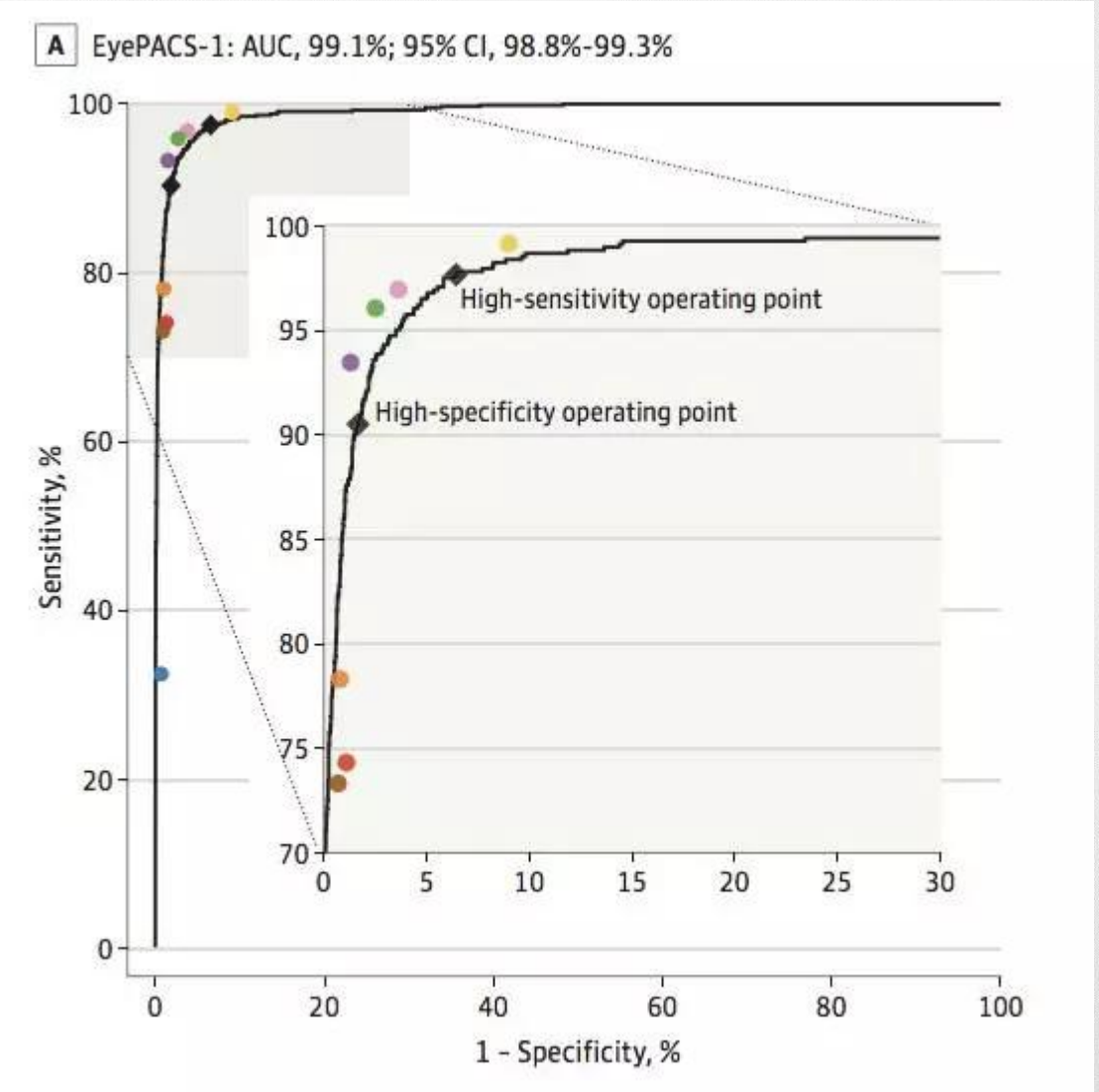
03

人工智能与医学影像

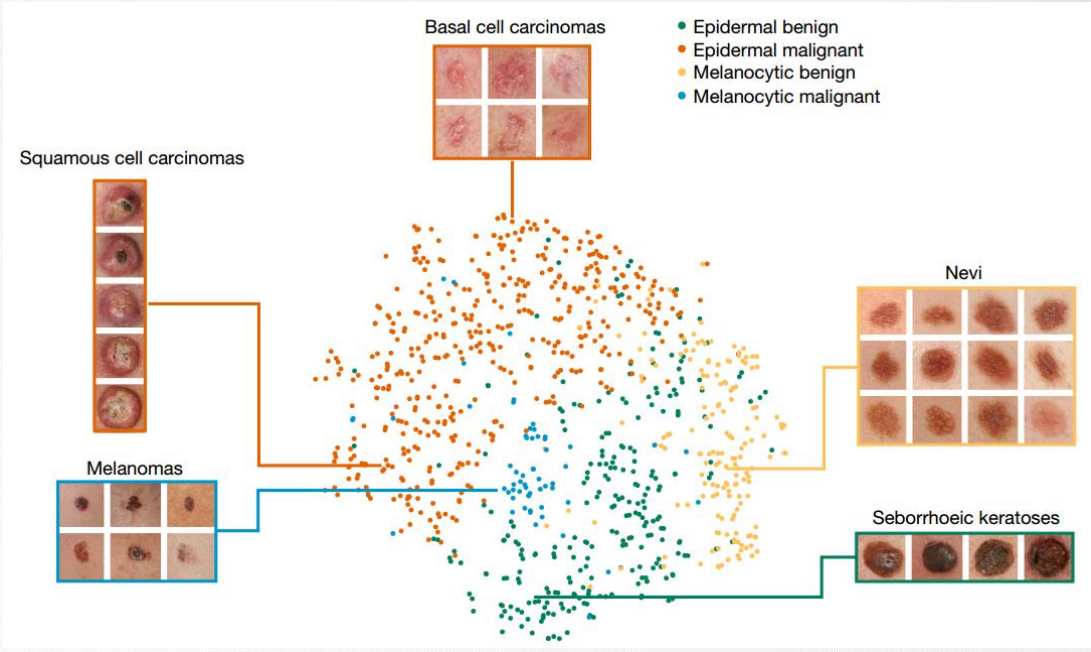
JAMA:检测视网膜眼底照片中糖尿病性视网膜病变的深度学习算法

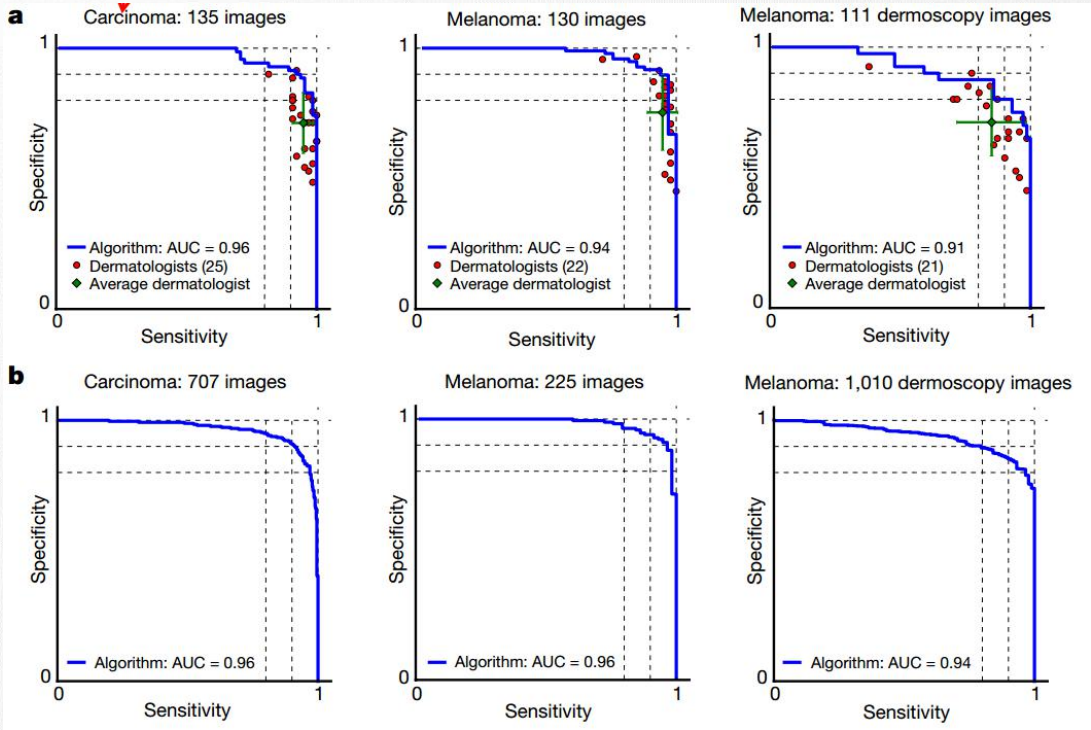
使用深度卷积神经网络的专为图像分类而优化过的神经网络模型，该网络使用 128175 张视网膜图像的数据集进行了训练，其中的每一张图像都针对糖尿病性视网膜病变、糖尿病性黄斑水肿和图像等级进行了 3 到 7 次评估。所得到的算法使用两个互相独立的数据集进行了验证，其中的每张图像测试所参考的标准是一个 7 或 8 人的美国认证眼科医生中大多数人的意见。





Nature:图片对皮肤癌检测





Nature: 基于大脑核磁图像的自闭症风险预测

LETTER

doi:10.1038/nature21369

Early brain development in infants at high risk for autism spectrum disorder

Heather Cody Hazlett[1,2], Hongbin Gu[1], Brent C. Munsell[3], Sun Hyung Kim[1], Martin Styner[1], Jason J. Wolff[4], Jed T. Elison[5], Meghan R. Swanson[2], Hongtu Zhu[6], Kelly N. Botteron[7], D. Louis Collins[11], John N. Constantino[7], Stephen R. Dager[8,9], Annette M. Estes[9,10], Alan C. Evans[11], Vladimir S. Fonov[11], Guido Gerig[12], Penelope Kostopoulos[11], Robert C. McKinstry[13], Juhi Pandey[14], Sarah Paterson[15], John R. Pruett Jr[7], Robert T. Schultz[14], Dennis W. Shaw[8,9], Lonnie Zwaigenbaum[16], Joseph Piven[1,2] & the IBIS Network*

Brain enlargement has been observed in children with autism spectrum disorder (ASD), but the timing of this phenomenon, and the relationship between ASD and the appearance of behavioural symptoms, are unknown. Retrospective head circumference and longitudinal brain volume studies of two-year olds followed up at four years of age have provided evidence that increased brain volume may emerge early in development[1,2]. Studies of infants at high familial risk of autism can provide insight into the early development of autism and have shown that characteristic social deficits in ASD emerge during the latter part of the first and in the second year of life[3,4]. These observations suggest that prospective brain-imaging studies of infants at high familial risk of ASD might identify early postnatal changes in brain volume that occur before an ASD diagnosis. In this prospective neuroimaging study of 106 infants at high familial risk of ASD and 42 low-risk infants, we show that hyperexpansion of the cortical surface area between 6 and 12 months of age precedes brain volume overgrowth observed between 12 and 24 months in 15 high-risk infants who were diagnosed with autism at 24 months. Brain volume overgrowth was linked to the emergence and severity of autistic social deficits. A deep-learning algorithm that primarily uses surface area information from magnetic resonance imaging of the brain of 6–12-month-old individuals predicted the diagnosis of autism in individual high-risk children at 24 months (with a positive predictive value of 81% and a sensitivity of 88%). These findings demonstrate that early brain changes occur during the period in which autistic behaviours are first emerging.

We first reported increased brain volume in adolescents and adults with ASD over twenty years ago[5]. Subsequent reports suggested that brain overgrowth in ASD may be most apparent during early childhood[6–8]. A study of infants at risk for ASD (33 high-risk and 22 low-risk infants), scanned from 6 to 24 months of age, found enlarged brain volume present at 12 and 24 months in the 10 infants that were later diagnosed with autism at 24 months of age or later[9] (mean age, 32.5 months).

In the present study, we examined data from a subset of individuals from a longitudinal study comprising 318 infants at high familial risk for ASD (HR), of which 70 met clinical best-estimate criteria for ASD (HR-ASD) and 248 that did not meet the criteria for ASD (HR-neg) at 24 months of age, and 117 infants at low familial risk (LR) for ASD, who also did not meet the criteria for ASD at 24 months (see Methods for diagnostic and exclusion criteria). The three groups were comparable in (mean) race/ethnicity (85% white), family income, maternal age at birth (33 years old), infant birth weight (8 lb), and gestational age at birth (39 weeks). The HR-ASD group had more males than the other two groups (83% of the HR-ASD group was male compared to 59% and 57% of the LR and HR-neg groups, respectively) and mothers in the LR group had a higher education level (Extended Data Table 1).

Infants were evaluated at 6, 12 and 24 months of age, which included detailed behavioural assessments and high-resolution magnetic resonance imaging (MRI) of the brain, to prospectively investigate brain and behavioural trajectories during infancy. The analyses described below were conducted on a subset of 106 high-risk ($n = 15$ HR-ASD; $n = 91$ HR-neg) and 42 low-risk infants for whom all three MRI scans were successfully obtained. On the basis of our previous findings at 2–4 years of age[2], we hypothesized that brain overgrowth in ASD begins before 24 months of age; that overgrowth is associated with hyperexpansion of the cortical surface area; and that these early brain changes are temporally linked to the emergence of the defining behaviours of ASD. We also investigated whether differences in the development of brain characteristics might suggest early biomarkers (that is, occurring before the onset of the defining behaviours of ASD) for the detection of ASD.

We first examined group differences in the trajectories of brain growth rate (Fig. 1). The growth rate of the total brain volume (TBV) did not differ between groups from 6 to 12 months of age. However, pairwise comparisons at 24 months showed large effect sizes for HR-ASD compared to LR and HR-ASD compared to HR-neg. The HR-ASD group showed a significantly increased TBV growth rate in the second year compared to both the LR and HR-neg groups (Extended Data Table 2). In addition, the HR-ASD group showed a significantly increased surface area growth rate from 6 to 12 months of age compared to both the HR-neg and LR groups, with the most robust increases observed in the left/right middle occipital gyrus, right cuneus and right lingual gyrus area (see Fig. 2). No group differences were observed in cortical thickness. We observed a significant correlation between surface area growth rate of 6–12 months and enlargement in TBV at 24 months of age in all subjects ($r_{192} = 0.59$, $P < 0.001$), as well as in the combined HR subgroup ($r_{139} = 0.63$, $P < 0.001$). Raw means, standard deviations and effect sizes for the group comparisons of TBV and surface area are provided in Extended Data Table 3. Regional differences in surface area change rate (6–12 months) were observed in the HR-ASD group (Fig. 2).

[1]Department of Psychiatry, University of North Carolina, Chapel Hill, North Carolina 27599, USA. [2]Carolina Institute for Developmental Disabilities, Chapel Hill, North Carolina 27599, USA. [3]College of Charleston, Charleston, South Carolina 29424, USA. [4]Department of Educational Psychology, University of Minnesota, Minneapolis, Minnesota 55455, USA. [5]Institute of Child Development, University of Minnesota, Minneapolis, Minnesota 55455, USA. [6]Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina 27599, USA. [7]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri 63110, USA. [8]Department of Radiology, University of Washington, Seattle, Washington 98105, USA. [9]Center on Human Development and Disability, University of Washington, Seattle, Washington 98105, USA. [10]Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington 98105, USA. [11]Montreal Neurological Institute, McGill University, Montreal, Quebec H3A 0G4, Canada. [12]Tandon School of Engineering, New York University, New York, New York 10003, USA. [13]Mallinckrodt Institute of Radiology, Washington University, St. Louis, Missouri 63110, USA. [14]Center for Autism Research, The Children's Hospital of Philadelphia and University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [15]Department of Psychology, Temple University, Philadelphia, Pennsylvania 19122, USA. [16]Department of Pediatrics, University of Alberta, Edmonton, Alberta T6G 2R3, Canada. *A list of participants and their affiliations appears at the end of the paper.

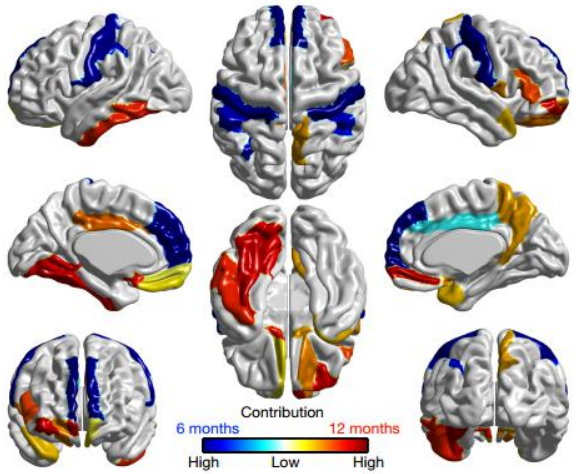


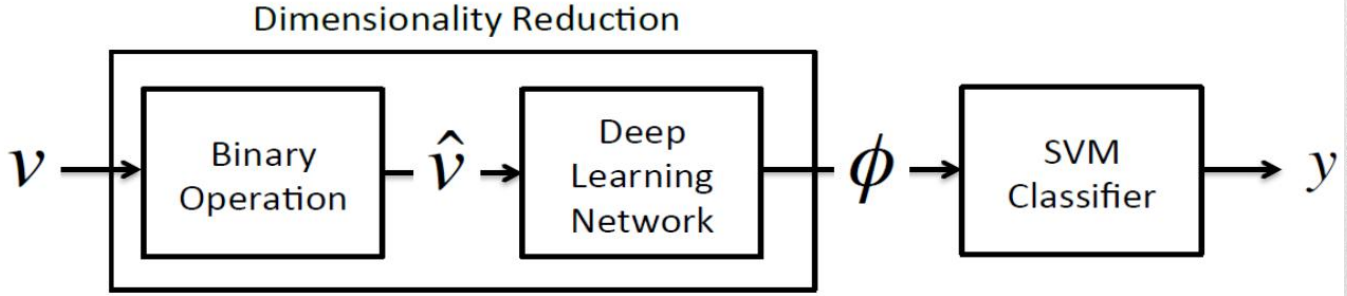
Figure 3 | Visualization of cortical regions with surface area measurements among the top 40 features contributing to the reduction in deep learning dimensionality. The cortical regions with surface area measurements that were among the top 40 features obtained from the nonlinear deep learning approach are visualized. The top 10 deep learning features observed include: surface area at 6 months in the right and left superior frontal gyrus, post-central gyrus, and inferior parietal gyri, and intracranial volume at 6 months. These features produced by the deep learning approach are highly consistent with those observed using an alternative approach (linear sparse learning) (Extended Data Fig. 1). Two tables listing the top 40 features from the deep learning approach and sparse learning are provided in Supplementary Tables 2 and 3.

预测2岁前的自闭症高危儿童（家族史）是否会在2岁之后被诊断为自闭症。

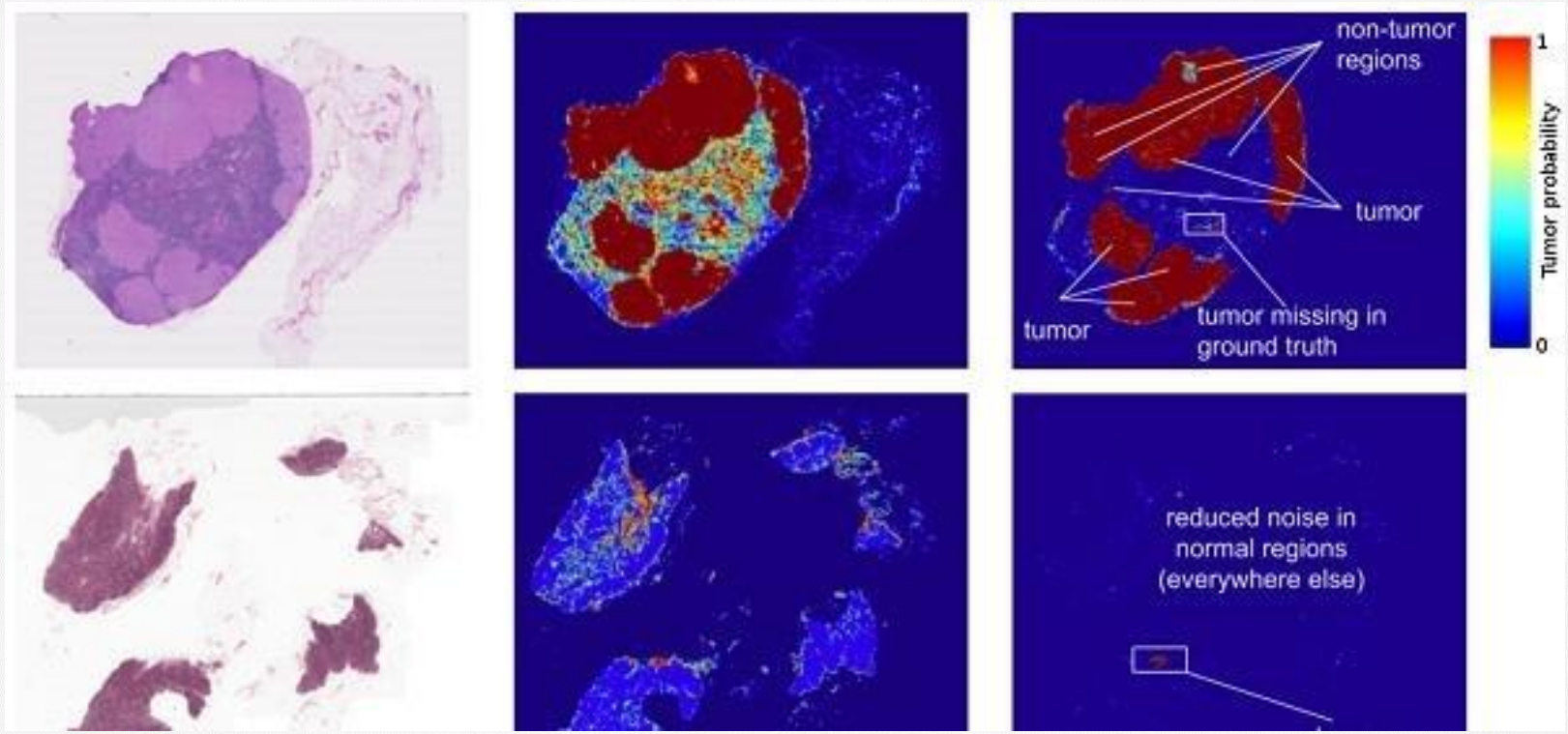
人工智能判别法(Hazlett et al. 2017) 准确度88%

传统行为问卷调查法(Charman 2014)准确度50%

Figure 2. Proposed Two-stage prediction pipeline that includes a non-linear dimension reduction step followed by a SVM classification step

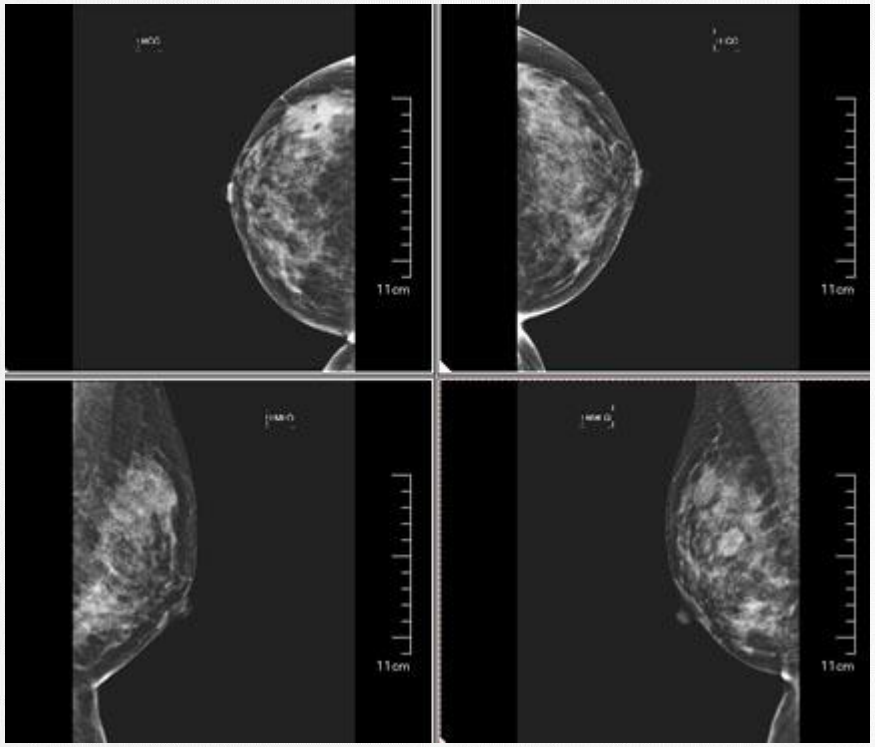


Google: 乳腺癌病理人工智能准确率已达89%



一位资深病理学家花了整整30个小时，仔仔细细分析了130张切片，依然以73.3%的准确率完败准确率达88.5%的人工智能。

乳腺癌影像学数据



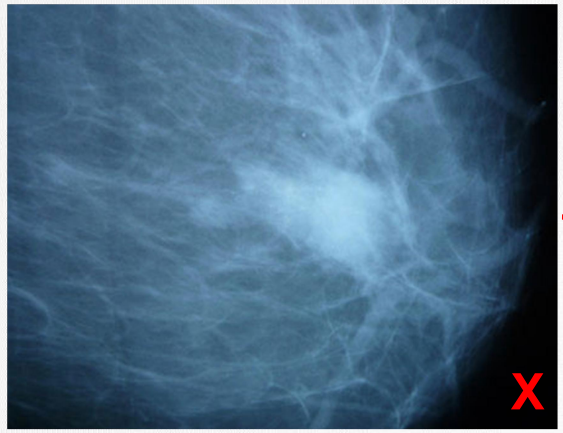
乳腺钼靶

乳腺B超

10,000例
有标注的

乳腺癌的影像学分类

Training



病理学报告：

……

病理学结论：

良/恶性 Y

Testing

模型

考虑为良性：20%

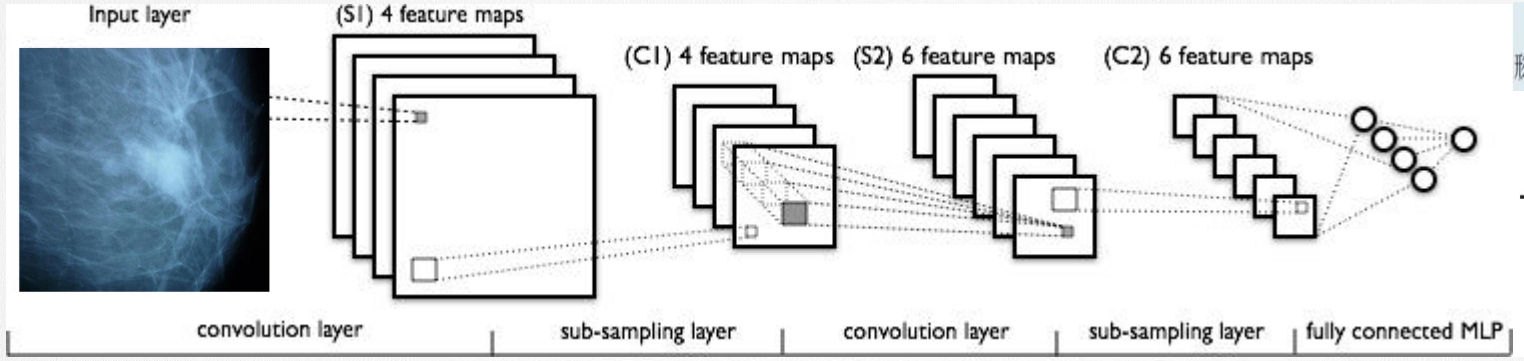
考虑为恶性：80%

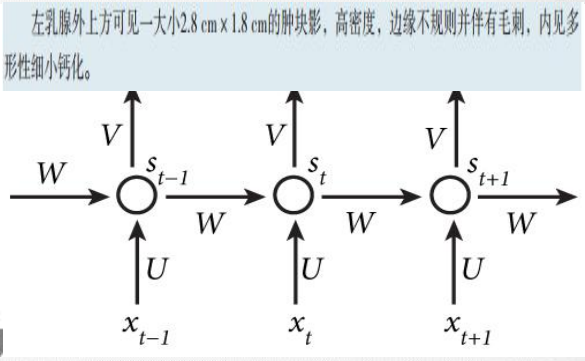
Dr. AI

2017.02.26

DrAI 影像辅助诊断系统

Training





Testing

模型

XXXXXX医院乳腺X线检查报告书

患者姓名：XXX　性别：女　年龄：51岁　放射学检查号码：12345678
门诊号：　住院号：123456　科室：乳腺外科　病区：

临床诊断：左乳肿块　检查日期：2010.10.10 前片 无

投照体位：
☑ 左侧：头足(轴)位、侧斜位　检查设备GE 2000D
☑ 右侧：头足(轴)位、侧斜位

影像学描述：
双侧乳腺致密腺体型。

左乳腺外上方可见一大小2.8 cm × 1.8 cm的肿块影，高密度，边缘不规则并伴有毛刺，内见多形性细小钙化。

右乳腺未见明显肿块与异常钙化。

双侧皮肤、乳头影正常。

双侧腋下可见小淋巴结，形态密度无异常。

影像学评估：
左乳外上病灶，考虑为恶性，BI-RADS：5。

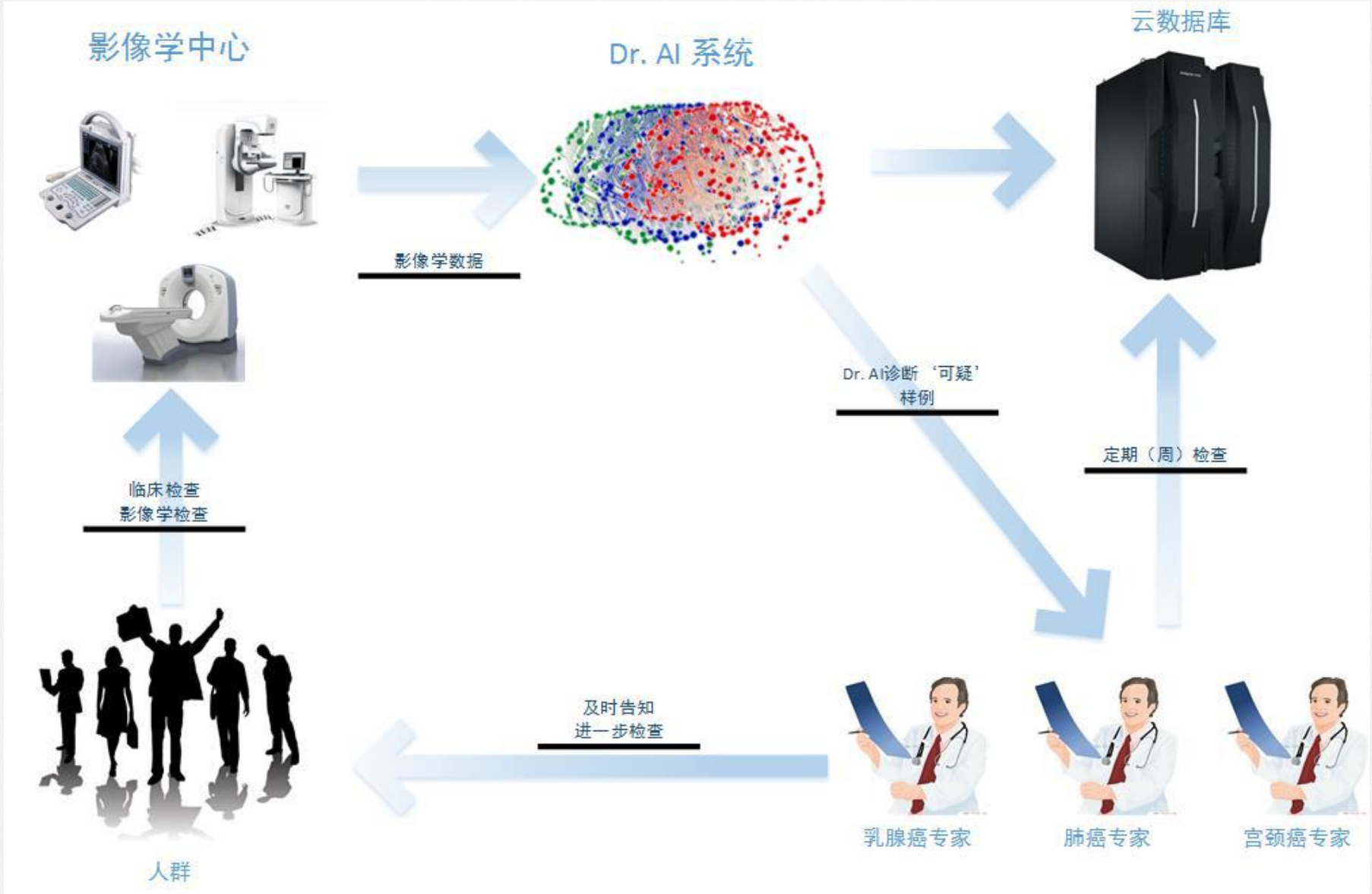
右乳未见异常发现，BI-RADS：1。

Dr. AI
2017.02.26

分级诊疗更需要人工智能辅助

- 分级诊疗要么让有限的专家资源下沉，要么复制更多的专家(AI)。
- 国家卫计委发布了2017版“人工智能辅助诊断技术管理规范”及“人工智能辅助诊断技术临床应用质量控制指标”。

DrAI.cn 助力分级诊疗



Dr. AI：愿景

医疗影像

鉴定分类疾病
精确度量敏感区域
提供影像学报告

辅助诊疗

辅助医生诊断疾病
辅助医生确定治疗方案
分级诊疗提高基层诊断水平

基因组学

寻找‘兴趣’位点
寻找基因型-表型关系

健康管理

构建个人健康档案
关注个人身体健康变化
预测疾病风险、降低风险
辅助慢性病人日常生活

感谢诸位聆听！欢迎合作！

微信：ebiomed