Presentation

2022
深度学习基础

“The potential benefits of artificial intelligence are huge, so are the dangers.” - Dave Waters.

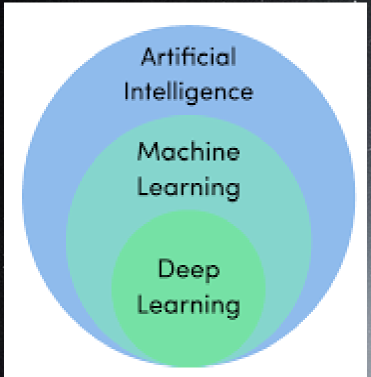
组员：林子超、谷京兆、王嘉楠、唐菲遥、赵盈盈

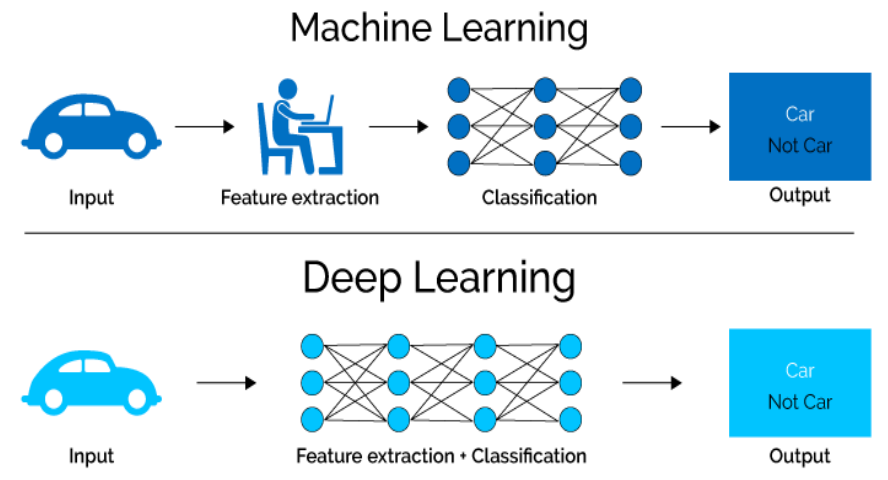
CONTENT

01

深度学习简介

Introduction





深度学习是机器学习的分支，是一种以人工神经网络为架构，对资料进行表征学习的算法。

至今已有数种深度学习框架，如深度神经网络、卷积神经网络和深度置信网络和循环神经网络已被应用在计算机视觉、语音识别、自然语言处理、音频识别与生物信息学等领域并获取了极好的效果。

02

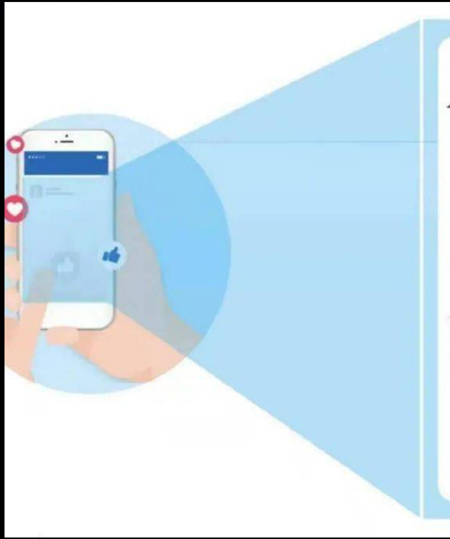
深度学习发展

History and future


1943
McCulloch Pitts Neuron – Beginning
1957
Frank Rosenblatt creates Perceptron
1960
The first Backpropagation Model
1962
Backpropagation with Chain Rule
1965
Birth of Multilayer Neural Network
1969
The Fall of Perceptron
1970
Backpropagation is computer coded
1971
Neural Network goes Deep
1980
Neocognitron – First CNN Architecture
1982
Hopfield Network – Early RNN
Proposal for Backpropagation in ANN
1985
Boltzmann Machine
1986
NetTalk – ANN Learns Speech
Implementation of Backpropagation
Restricted Boltzmann Machine
1989
CNN using Backpropagation
1989
Universal Approximators Theorem
1991
Vanishing Gradient Problem Appears
1997
The Milestone Of LSTM
2006
Deep Belief Network
2008
GPU Revolution Begins
2009
ImageNet is launched
2011
Combat for vanishing grandient
2012
AlexNet starts deep learning boom
2014
The birth of GANs
2016
AlphaGo beats human
2019
Trio win Turing Award
2020
The Lottery Ticket Hypothesis
The Top-Performing Model with Zero Training
2021
......
2022
2022 AI and machine learning trends
......


Stretched positon :
Analyzing System


我将全年无休为听障用户提供服务
I will be continuously providing services to people with hearing disabilities

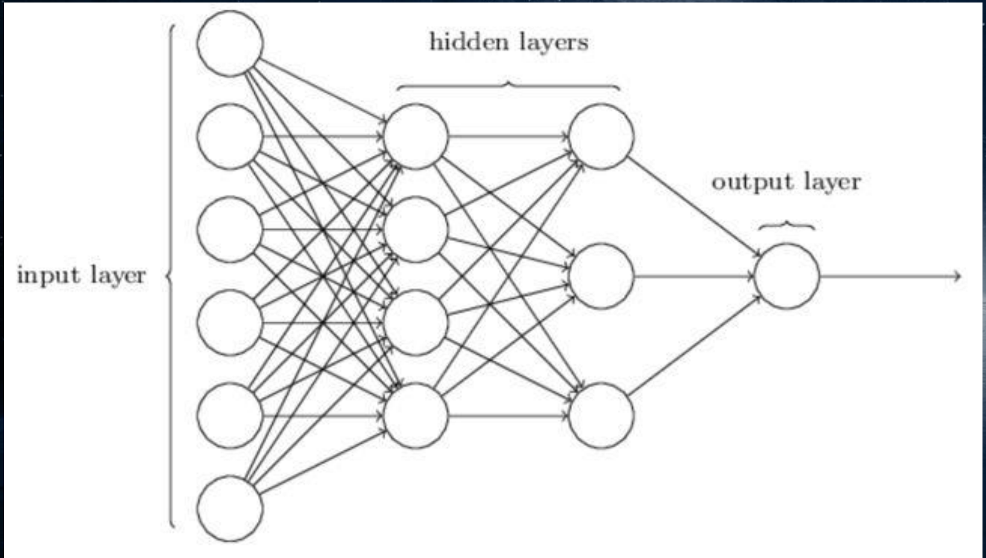


03

经典深度学习网络

Networks and architectures

3.1 FNN(Feedforward Neural Networks)

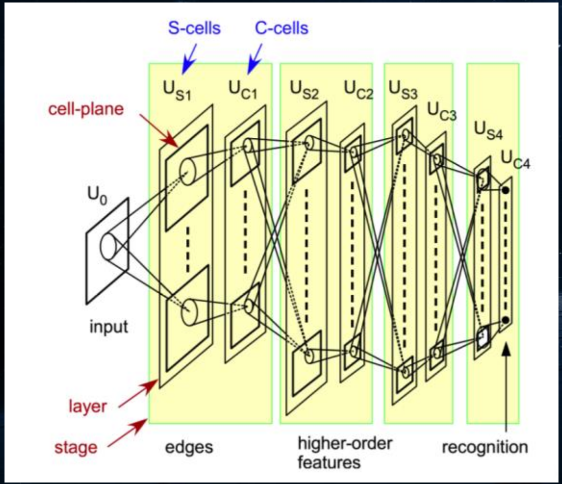


多层感知机（MLP）

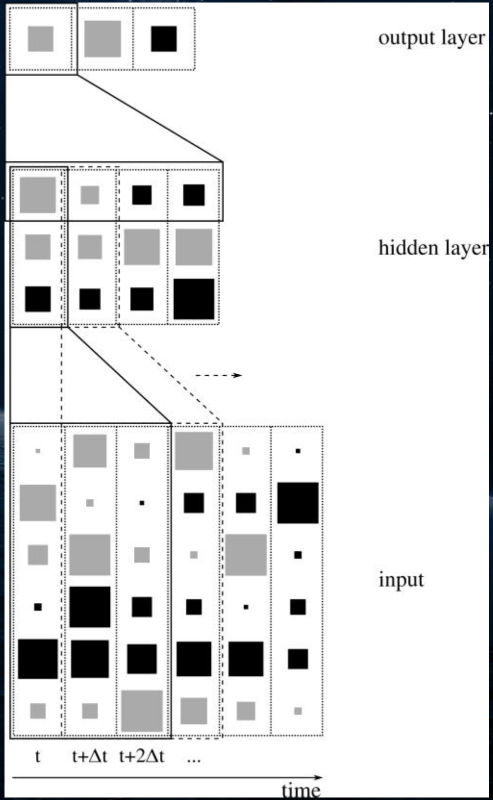
包含多个隐藏层的神经网络，层与层之间是全连接的。

前向传播、反向传播、梯度下降、激活函数、损失函数

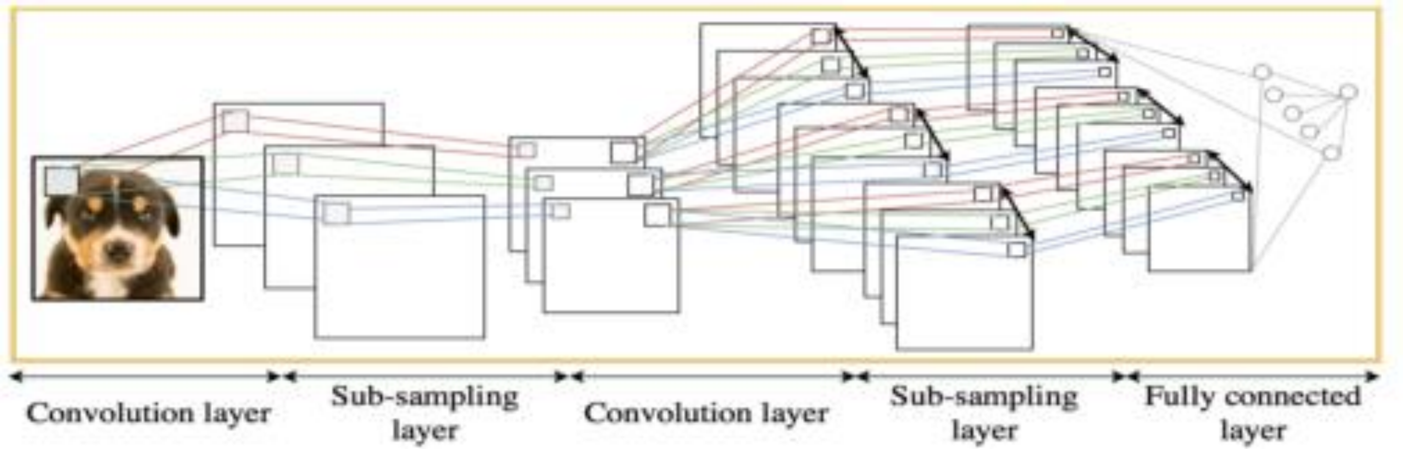
3.2 CNN(Convolutional Neural Networks)



网络	作者	年代
Neocognition	Kunihiko Fukushima	1980
TDNN	Alexander Waibel	1987
SIANN	Wei Zhang	1988
LeNet	Yann LeCun	1989
LeNet-5	Yann LeCun	1998
AlexNet	Hinton Alex Krizhevsky	2012
ZFNet	Matthew Zeiler Rob Fergus	2013
VGGNet	-	2014
GoogLeNet	Christian Szegedy	2014
ResNet	Kaiming He Xiangyu Zhang	2015



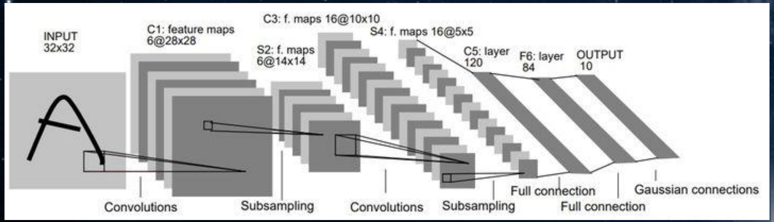
3.2 CNN(Convolutional Neural Networks)



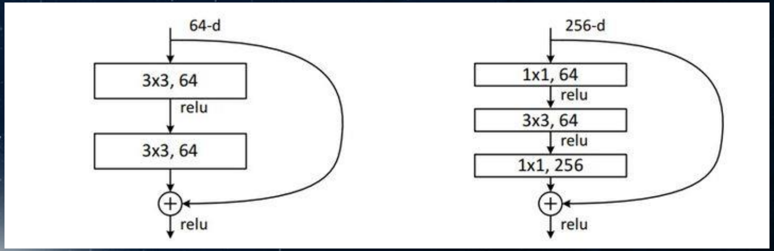
CNN是深度学习中一种流行且广泛使用的算法。它已被广泛应用于不同的应用领域，例如自然语言处理、语音处理和计算机视觉等等。

3.2 CNN(Convolutional Neural Networks)

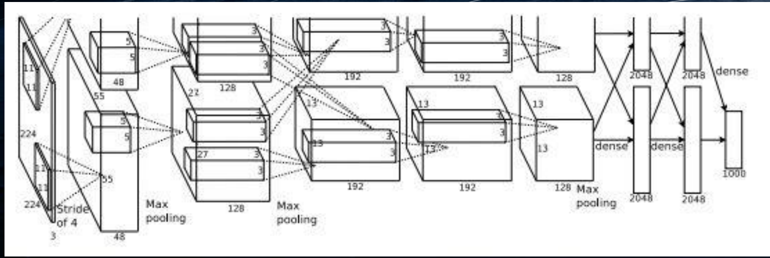
LeNet-5



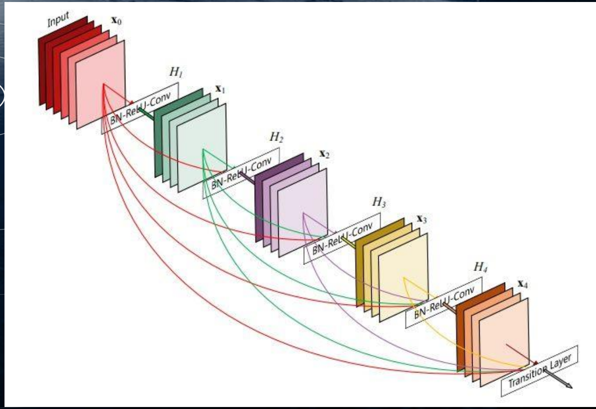
ResNet



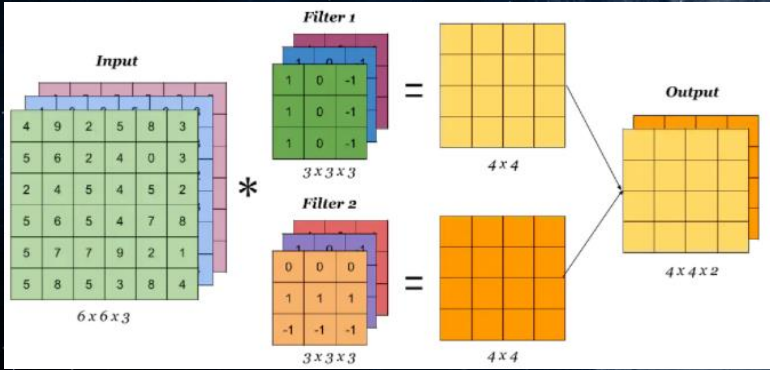
AlexNet



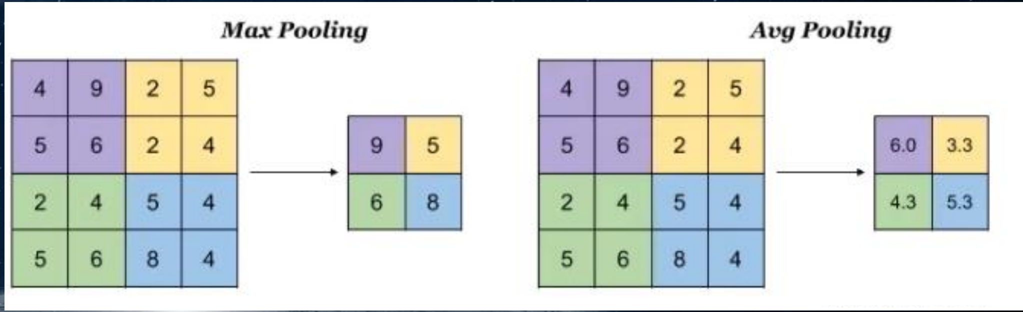
DenseNet



3.2 CNN(Convolutional Neural Networks)

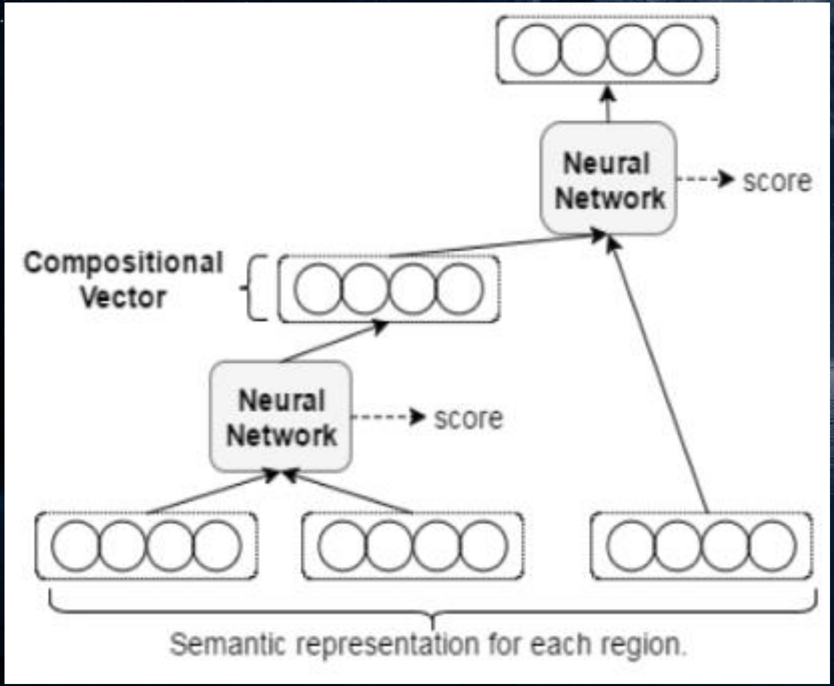


卷积层：通过卷积操作对输入图像进行降维和特征抽取。



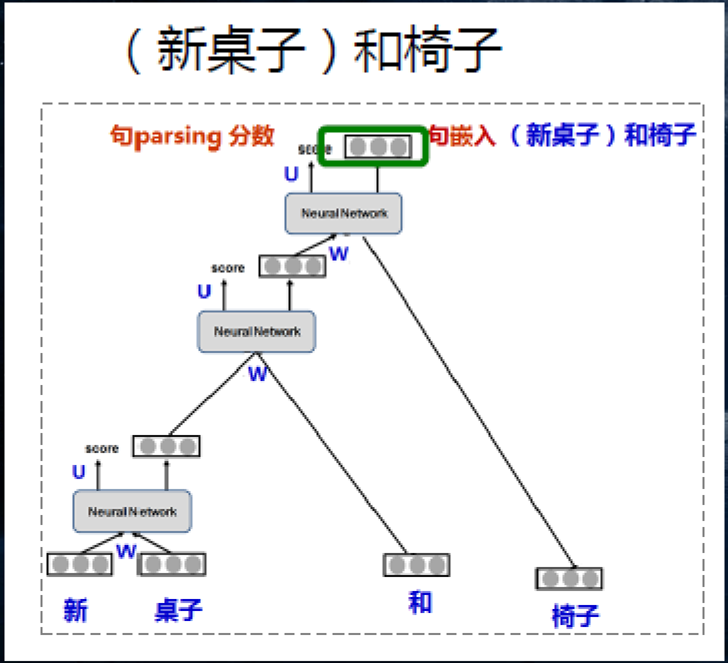
池化层：降维。

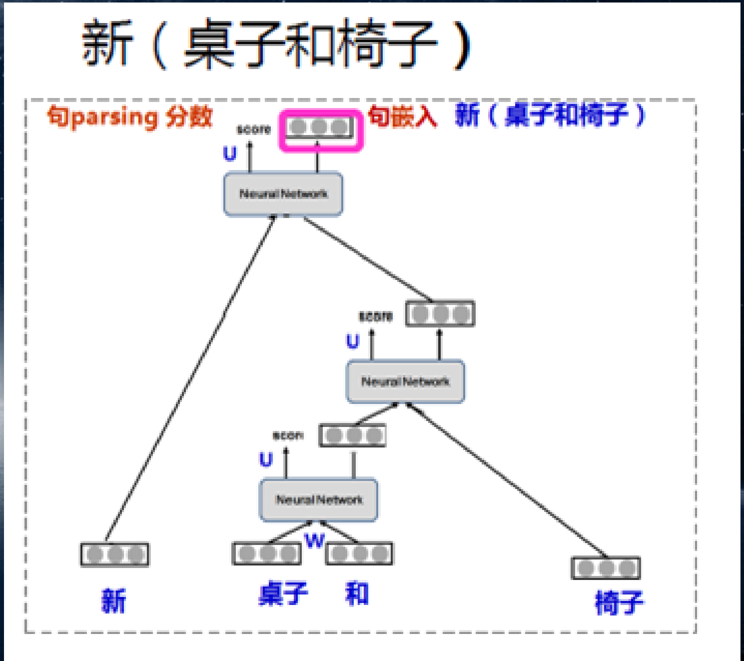
3.3 RvNN(Recursive Neural Network)



基本思想：将处理问题在结构上分解为一系列相同的“单元”，单元的神经网络可以在结构上展开，且能沿展开方向传递信息。

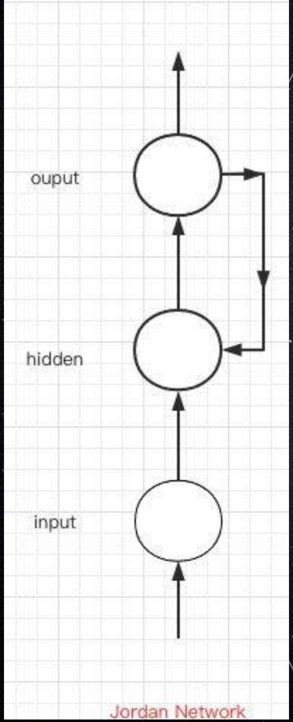
3.3 RvNN(Recursive Neural Networks)



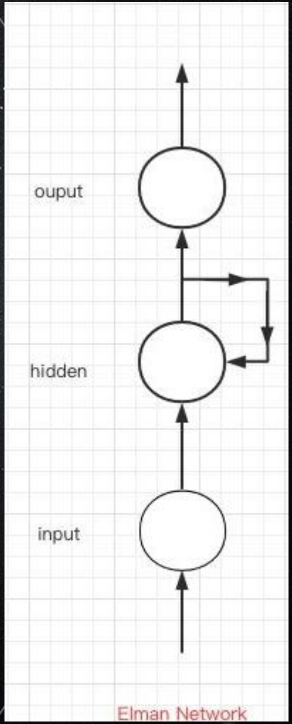


不同的句法结构所对应的句嵌入不同，分数也不同。

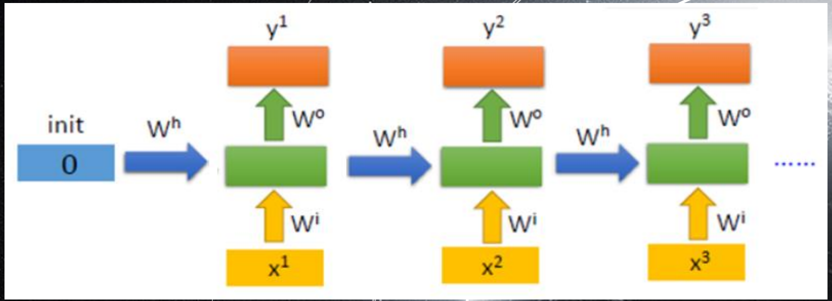
3.4 RNN(Recurrent Neural Networks)



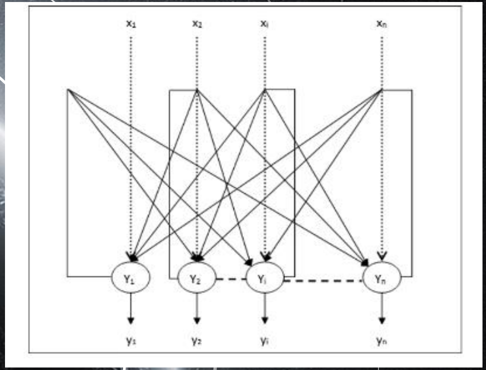
网络	作者	年代
Hopfield Network	John Hopfield	1982
Jordan Network	Michael I. Jordan	1986
Elman Network	Jeffery Elman	1990
NHC	Jurgen Schmidhuber	1992
LSTM	Jurgen Schmidhuber	1997
Bidirectional RNN	M. Schuster K. Paliwal	1997
GRU	K. Cho	2014



3.4 RNN(Recurrent Neural Networks)



雏形

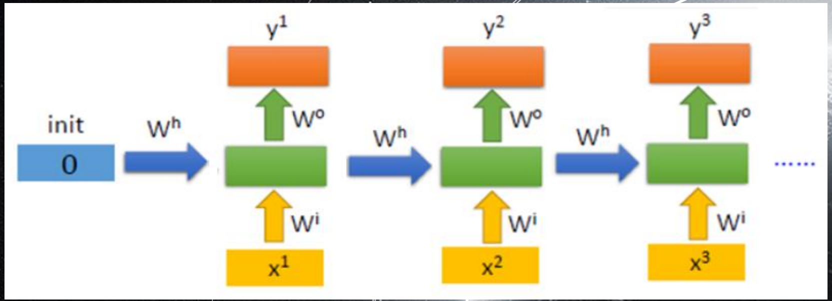


Hopfield Network

RNN是在深度学习中广泛使用和流行的算法，尤其是在NLP和语音处理中。与传统的神经网络不同，RNN利用了网络中的序列信息。

首先利用一个完整的pattern训练网络，以后只输入pattern的一部分数据时，可以通过达到最小的E(x)来复原完整的pattern。

3.4 RNN(Recurrent Neural Networks)



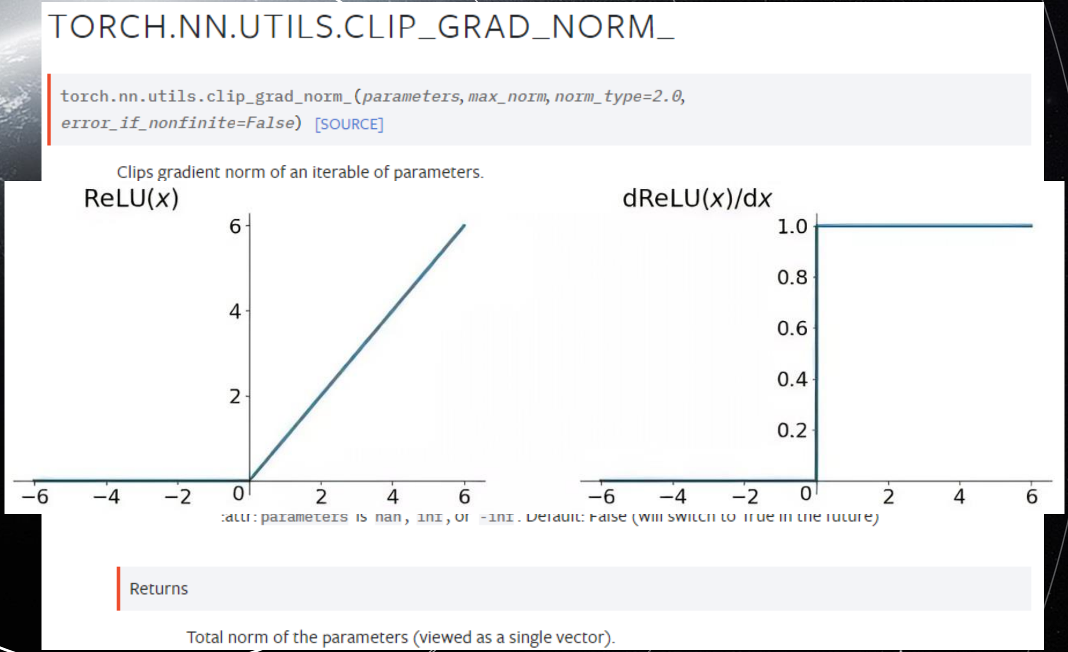
RNN是在深度学习中广泛使用和流行的算法，尤其是在NLP和语音处理中。与传统的神经网络不同，RNN利用了网络中的序列信息。

为什么现在没有很多人去用这种最基本的RNN呢?

梯度爆炸和梯度消失。这导致训练时梯度不能在较长的序列中一直传递下去，从而使RNN无法捕捉到长距离的影响。

3.4 RNN(Recurrent Neural Networks)

解决方法：
1.设定一个梯度阈值，更新的时候，如果梯度超过这个阈值范围就会被强制限制在这个阈值之内。
2.使用ReLU、LeakyReLU、eLU等激活函数。


TORCH.NN.UTILS.CLIP_GRAD_NORM_

```
torch.nn.utils.clip_grad_norm_(parameters, max_norm, norm_type=2.0, error_if_nonfinite=False)
```
[SOURCE]

Clips gradient norm of an iterable of parameters.

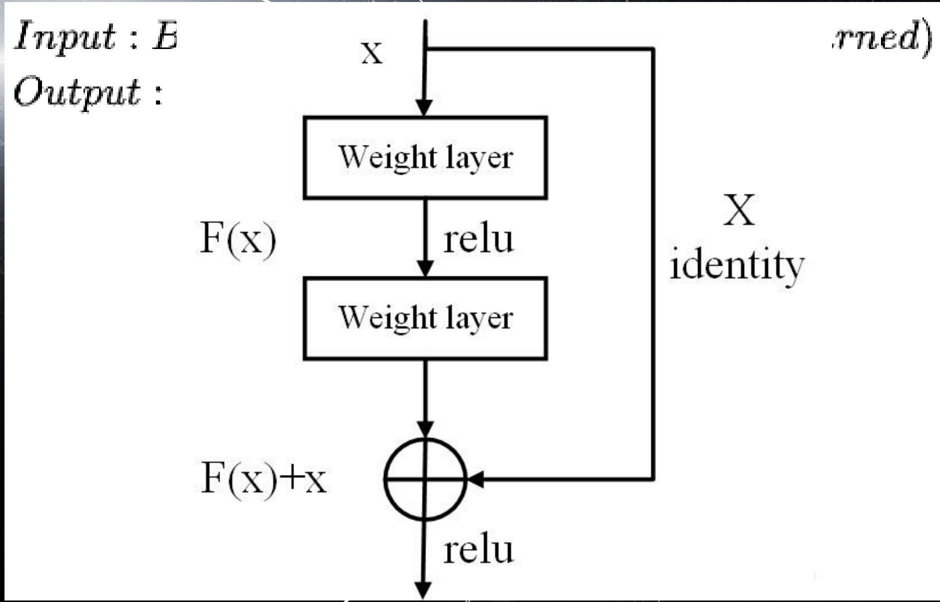
:attr:`parameters` is nan, inf, or -inf. Default: False (will switch to True in the future)

Returns

Total norm of the parameters (viewed as a single vector).

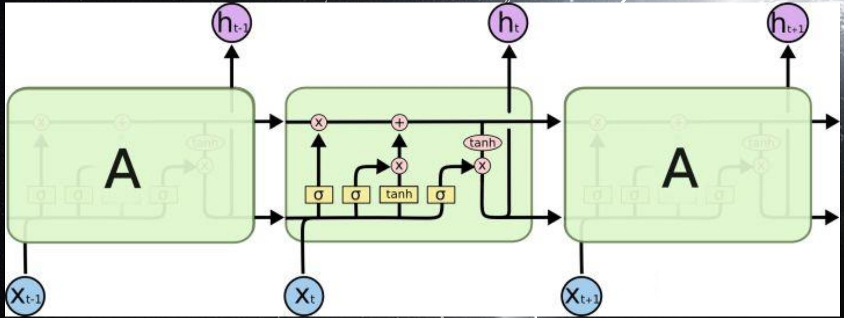
3.4 RNN(Recurrent Neural Networks)

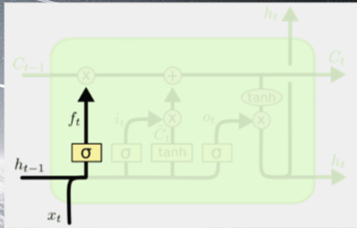
解决方法:
3.batch-normalization:
这样做可以将数据规整到统一区间，减少数据的发散程度，缓和了权重参数放大缩小带来的影响，在一定程度上解决了梯度消失和爆炸的问题。
4.残差连接（residual connection)



3.4 RNN(Recurrent Neural Networks)

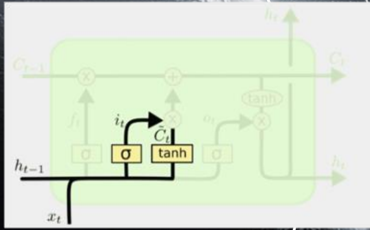
解决方法：
5. LSTM & GRU





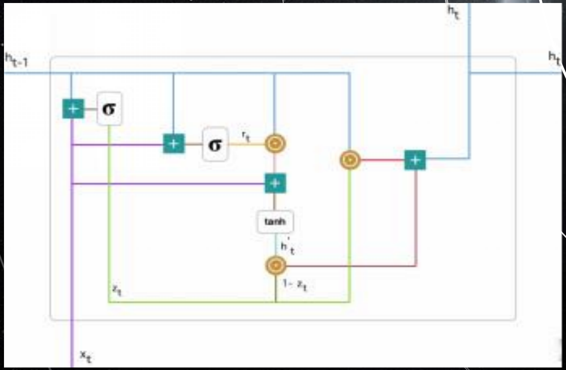
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \ + \ b_f\right)$$

遗忘门

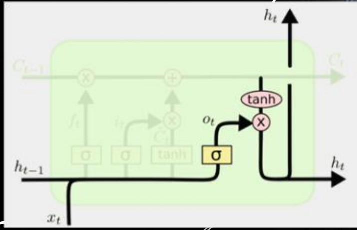


$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \ + \ b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \ + \ b_C)$$

输入门



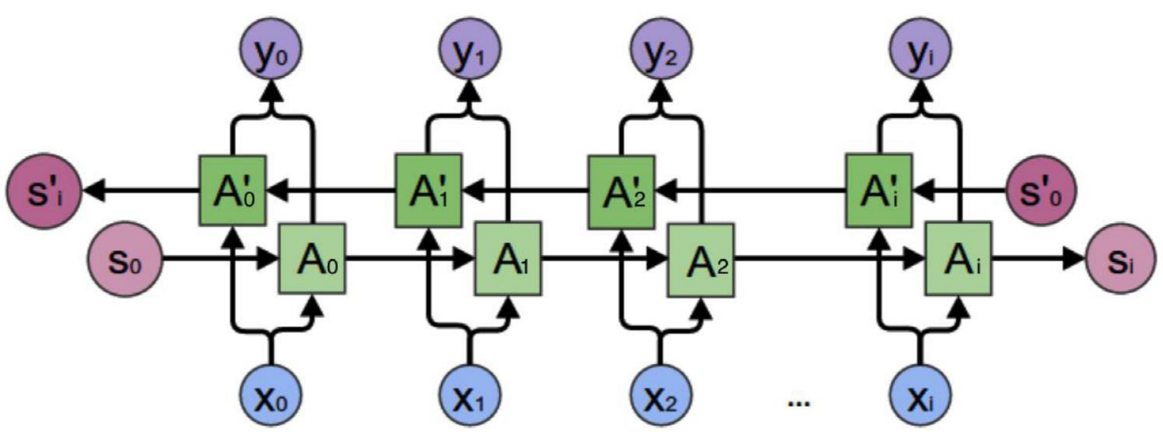
重置门、
更新门



$$o_t = \sigma\left(W_o \ [h_{t-1}, x_t] \ + \ b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

输出门

3.4 RNN(Recurrent Neural Networks)

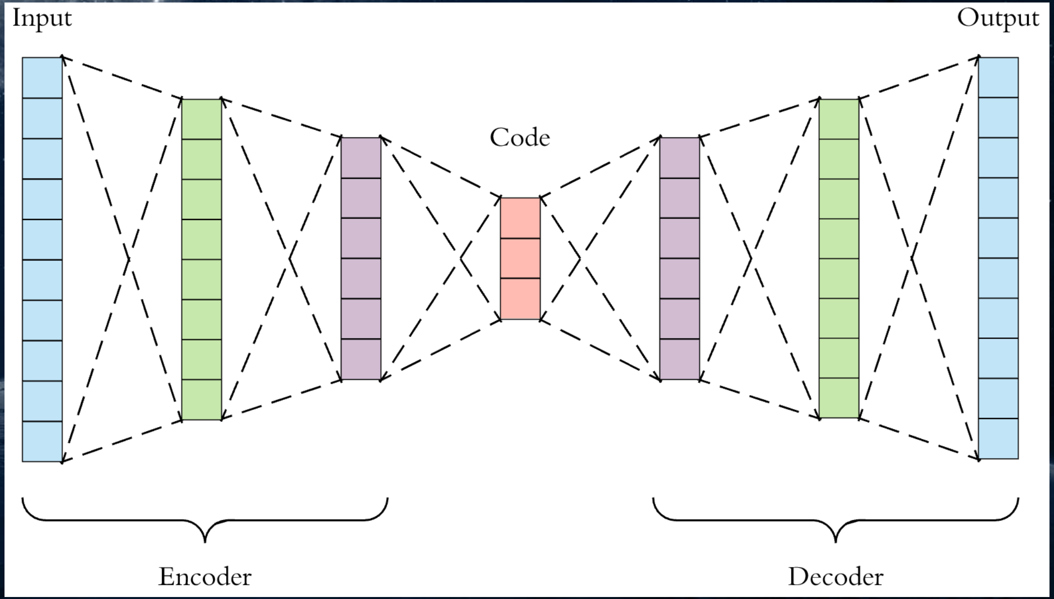


Bidirectional RNN

RNN和LSTM都只能依据之前时刻的时序信息来预测下一时刻的输出，但在有些问题中，当前时刻的输出不仅和之前的状态有关，还可能和未来的状态有关系。比如预测一句话中缺失的单词不仅需要根据前文来判断，还需要考虑它后面的内容，真正做到基于上下文判断。

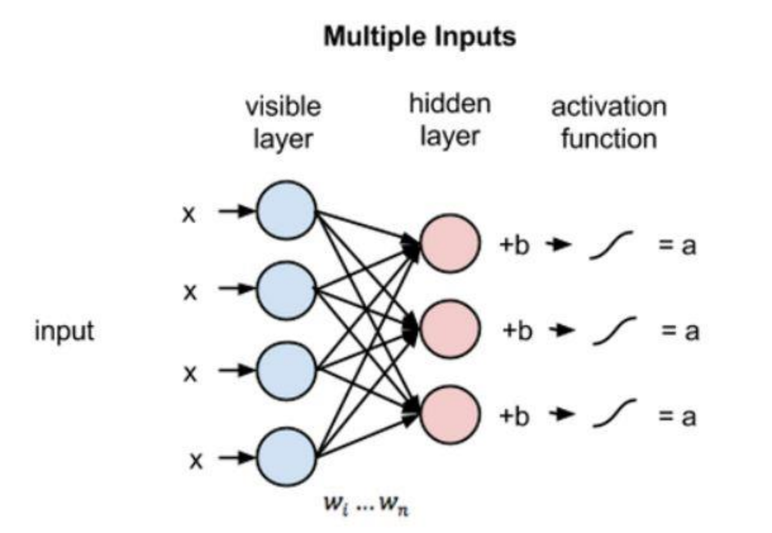
3.5 AE(AutoEncoders)

自动编码器通常是作为一个非监督算法使用，并且主要应用在降维和压缩。它们的技巧就是尝试让输出等于输入。

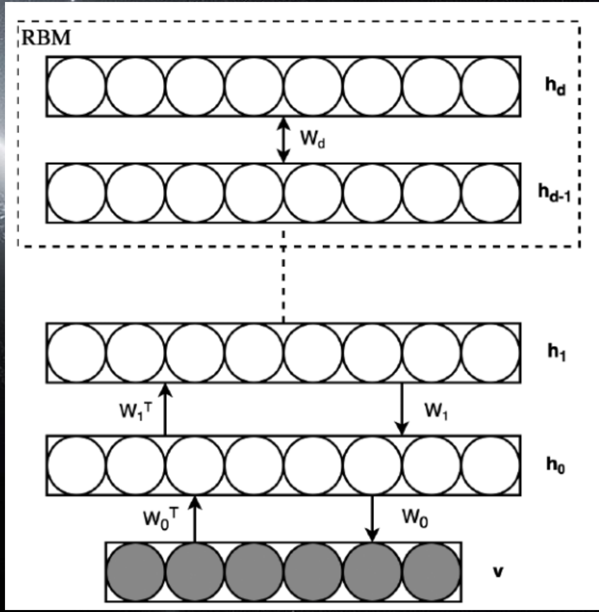


另外，也可以根据这个思路来重新得到略微有些不同的输入数据，甚至是更好的数据，这可以用于训练数据的增强，数据的去噪等。

3.6 DGN(Deep Generative Networks)

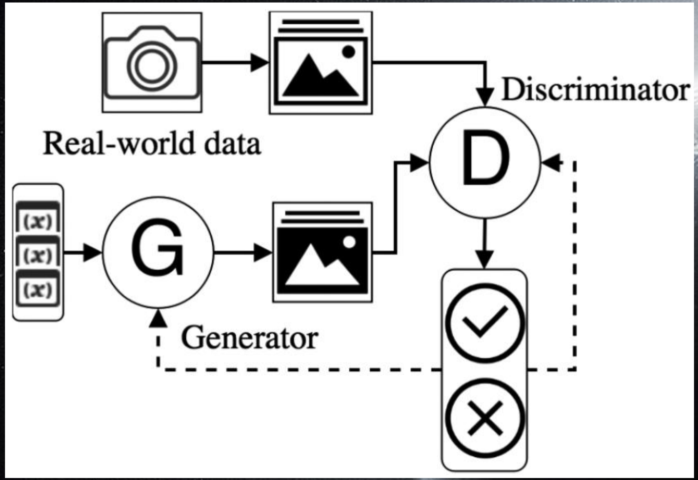


RBM(Restricted Boltzmann Machines) 是带有生成能力的随机神经网络，也就是它可以通过输入来学习到一个概率分布。相比其他网络，它的最大特点就是只有输入和隐藏层，不包含输出。



DBN(Deep Belief Networks)

3.6 DGN(Deep Generative Networks)



GAN(Generative Adversarial Networks)

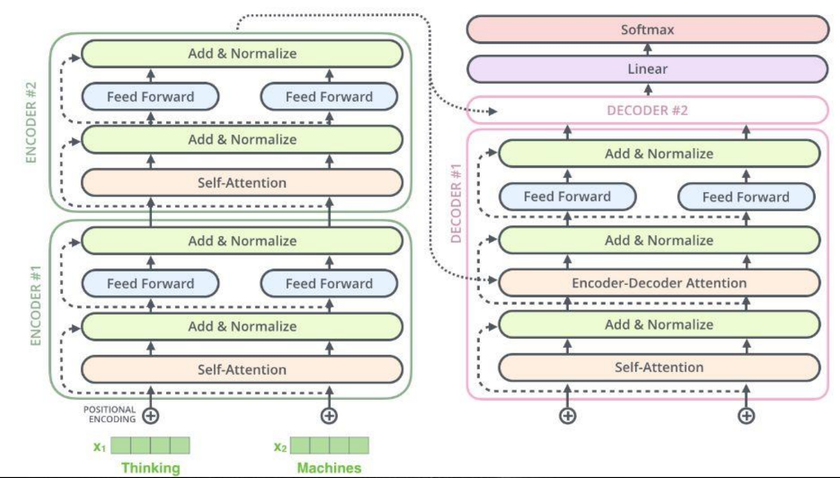
GAN 由一个生成模型G和一个判别模型D组成。G负责接受一个随机的噪声，通过这个噪声生成图片。D负责判别这张图片是真实的还是生成的。

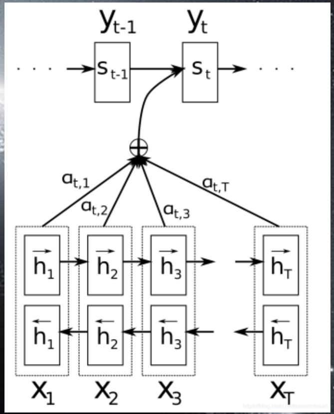
博弈体现在哪里？

DCGAN、InfoGAN、CycleGAN、WGAN、Self-Attention GAN、BigGAN……

3.7 Transformer

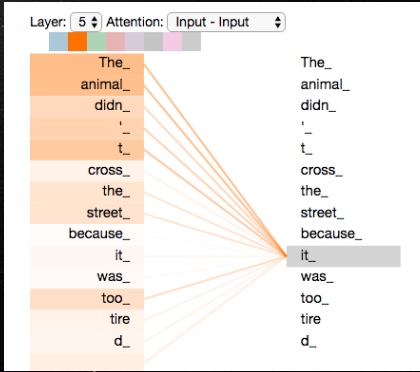
Attention机制：让网络关注特定的数据点。



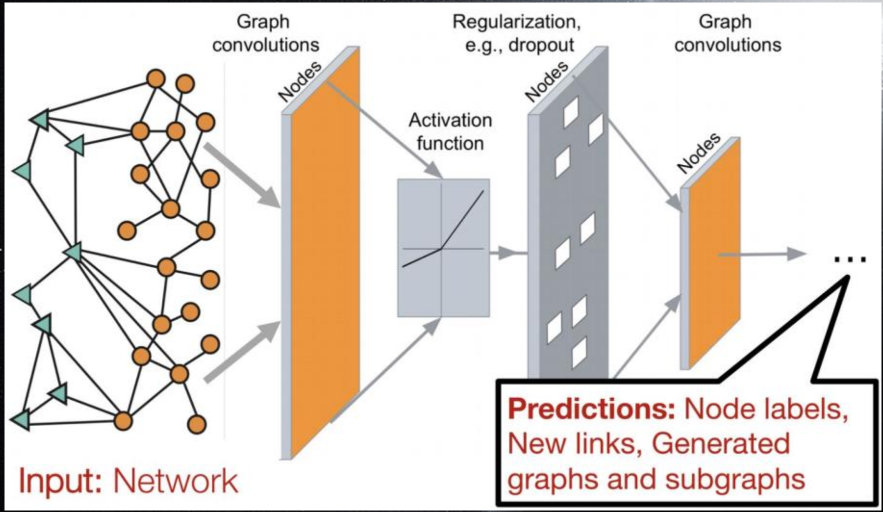


相比于拥有复杂的LSTM单元，Attention机制是根据输入数据不同部分的重要性来赋予权重。

Self-Attention：



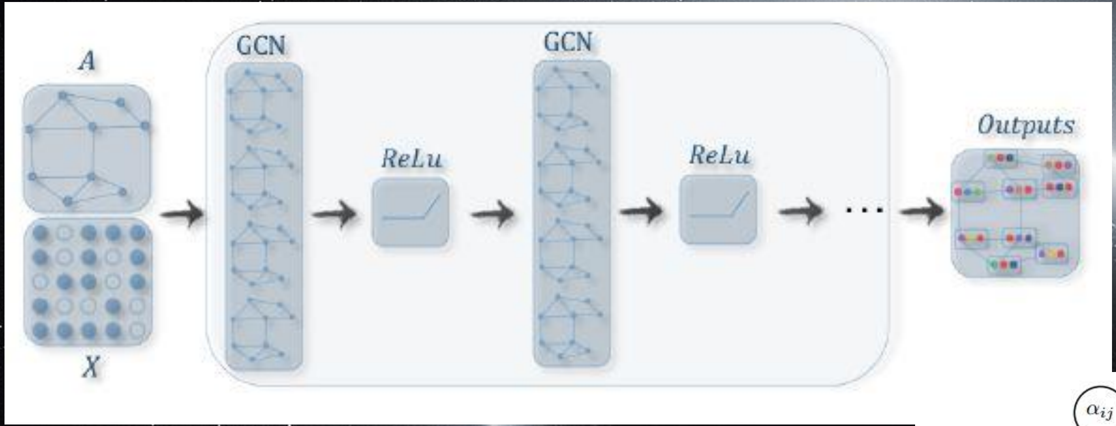
3.8 GNN(Graph Neural Networks)



GNN的目标就是建模图数据，也就是可以识别到一个图里结点之间的关系并生成一个数值型的表征数据，类似于一个嵌入向量（embedding）。

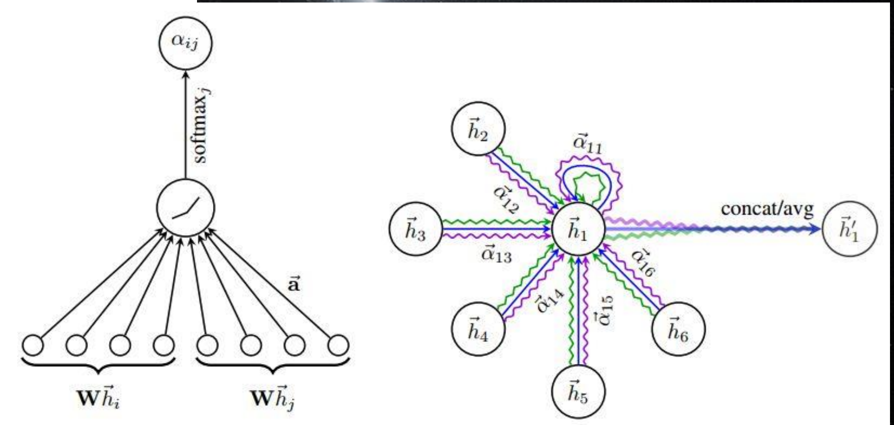
社交网络、化学混合物、知识图谱、空间数据等。

3.8 GNN(Graph Neural Networks)



GCN

GAT



04

深度学习技术及应用

Techniques and applications

基于深度学习的自然语言处理

词嵌入

Word2Vec　Glove　FastText　Contexual Word Embeddings　Attention

2022

序列模型

Seq2Seq模型。其包括一个Encoder和一个Decoder，Encoder将序列作为输入，然后将输入在隐空间的表示作为输出，这个表示会输入到Decoder中，并输出一个新的序列。

常用的有CNN、LSTM、Transformers等。

基于深度学习的自然语言处理 Demo

hypothesis

使用经预训练的DeBERTa模型完成NLI任务

推断一个句子中蕴含某种给定关系的概率。

Input:
template: [CLS]premise[SEP]hypothesis[SEP]

Output:
Entailment: ....%
Contradiction: ....%
Neutral: ....%

```
verbalizer = {
    "/people/person/nationality": "{subj}'s nationality is {obj} .",
    "/time/event/locations": "{subj} happens in {obj} .",
    "/people/person/children": "{obj} is the children of {subj} .",
    "/business/company/advisors": "{obj} is the advisors of {subj} .",
    "/business/location": "{obj} is the location of {subj} .",
    "/business/company/majorshareholders": "{obj} is the major shareholders of {subj} .",
    "/people/person/place_lived": "{subj} lives in {obj} .",
    "NA": "There's no relationship between {obj} and {subj} .",
    "/business/company/place_founded": "{obj} is the founded place of {subj} .",
    "/location/neighborhood/neighborhood_of": "{obj} is the neighborhood of {subj} .",
    "/people/deceasedperson/place_of_death": "{obj} is the place of death of {subj} .",
    "/film/film/featured_film_locations": "{obj} is the featured film location of {subj} .",
    "/location/region/capital": "{obj} is the capital of {subj} .",
    "/business/company/founders": "{obj} is the founder of {subj} .",
    "/people/ethnicity/geographic_distribution": "{obj} is the geographic distribution of {subj} .",
    "/location/country/administrative_divisions": "{subj} is the administrative division of {obj} .",
    "/people/deceasedperson/place_of_burial": "{obj} is the place of burial of {subj} .",
    "/location/country/capital": "{obj} is the capital of {subj} .",
    "/business/person/company": "{subj} works in {obj} .",
    "/location/location/contains": "{obj} is located in {subj} .",
    "/location/administrative_division/country": "{subj} is the administrative division of {obj} .",
    "/location/us_county/county_seat": "{obj} is the country seat of {subj} in America .",
    "/people/person/religion": "{subj} believes in {obj} .",
    "/people/person/place_of_birth": "{obj} is the birthplace of {subj} .",
    "/people/person/ethnicity": "{obj} is the ethnicity of {subj} ."
}
```

https://github.com/Caesar-s1mple/NLI

基于深度学习的自然语言处理 Demo

数据集：nyt10m

1	premise	hypothesis	score	relation	h	t
2	One of the newest auditoriums in New York , built into the bedrock b	Carnegie Hall is located in N	0.959800363	/location/location/contains	{'pos': [33, 41], 'id': 'm.059rby'	{'pos': [73, 86], 'id': 'm.016p8t', 'name': 'Carnegie Hall'}
3	A lot of people wo n't go across the river to Hoboken because , ` Oh	Hoboken is located in New	0.87095207	/location/location/contains	{'pos': [88, 98], 'id': 'm.05fjf', 'n	{'pos': [46, 53], 'id': 'm.0xn7b', 'name': 'Hoboken'}
4	The Rev. Charles Hewawasam , a Roman Catholic priest who lost a nu	Matara is located in Sri Lank	0.963463485	/location/location/contains	{'pos': [115, 124], 'id': 'm.06m	{'pos': [106, 112], 'id': 'm.0bl81f', 'name': 'Matara'}
5	The two men met at a Methodist church in downtown New Haven , v	Yale University is located in	0.875556707	/location/location/contains	{'pos': [50, 59], 'id': 'm.0f2nf', '	{'pos': [120, 135], 'id': 'm.08815', 'name': 'Yale University'}
6	He said he just hit two men , '' Ms. Treasure said at their home in Rose	Rosedale is located in Quee	0.970679998	/location/location/contains	{'pos': [79, 85], 'id': 'm.0ccvx',	{'pos': [68, 76], 'id': 'm.0434dk', 'name': 'Rosedale'}
7	His father , Heinz Kreckel of Iserlohn-Letmathe , Germany , retired as	Iserlohn is located in Germa	0.926600218	/location/location/contains	{'pos': [50, 57], 'id': 'm.0345h',	{'pos': [30, 38], 'id': 'm.018z3k', 'name': 'Iserlohn'}
8	Mr. Zawahiri said that Iraq had become '' the place for the greatest b	Iraq is located in Middle Eas	0.674619734	/location/location/contains	{'pos': [242, 253], 'id': 'm.04ws	{'pos': [23, 27], 'id': 'm.0d05q4', 'name': 'Iraq'}
9	Nowhere is the concern deeper than in western Michigan , where the	Coopersville is located in M	0.803336143	/location/location/contains	{'pos': [46, 54], 'id': 'm.04rrx', '	{'pos': [218, 230], 'id': 'm.0xc60', 'name': 'Coopersville'}
10	After Mr. Bremer left Iraq , and after a short tour by John D. Negropo	Baghdad is located in Iraq .	0.788466156	/location/location/contains	{'pos': [22, 26], 'id': 'm.0d05q4	{'pos': [136, 143], 'id': 'm.01fqm', 'name': 'Baghdad'}
11	28 Germany to Ban Smoking Partly Germany , Europe 's biggest toba	Germany is located in Europ	0.808117807	/location/location/contains	{'pos': [43, 49], 'id': 'm.02j9z', '	{'pos': [3, 10], 'id': 'm.0345h', 'name': 'Germany'}
12	The meeting , in the house of Sheik Hamid Turki al-Shawka , a promi	Baghdad is located in Iraq .	0.798232615	/location/location/contains	{'pos': [204, 208], 'id': 'm.0d05	{'pos': [233, 240], 'id': 'm.01fqm', 'name': 'Baghdad'}
13	They have never heard of Papua or Indonesia .	Papua is located in Indones	0.166268274	/location/location/contains	{'pos': [34, 43], 'id': 'm.03ryn',	{'pos': [25, 30], 'id': 'm.03xrq', 'name': 'Papua'}
14	The quality and richness of the Roaring Lions frescoes , which have es	Rome is located in Italy .	0.747110426	/location/location/contains	{'pos': [261, 266], 'id': 'm.03rjj'	{'pos': [232, 236], 'id': 'm.06c62', 'name': 'Rome'}
15	Perhaps the time has come for even more creative measures , like the	Rajahmundry is located in I	0.970831335	/location/location/contains	{'pos': [141, 146], 'id': 'm.03rk(	{'pos': [108, 119], 'id': 'm.048pp7', 'name': 'Rajahmundry'}
16	At 6:30 p.m. , McCarren Park Pool , Lorimer Street , between Driggs A	Greenpoint is located in Brc	0.985110939	/location/location/contains	{'pos': [108, 116], 'id': 'm.0cr3	{'pos': [95, 105], 'id': 'm.02372m', 'name': 'Greenpoint'}
17	G.M. plans to close assembly plants in Oklahoma City ; Lansing , Mich.	Oshawa is located in Ontari	0.973645687	/location/location/contains	{'pos': [104, 111], 'id': 'm.05kr_	{'pos': [95, 101], 'id': 'm.018n1k', 'name': 'Oshawa'}
18	Correction : September 19 , 2006 , Tuesday An article on Friday abou	Plainfield is located in New .	0.521991611	/location/location/contains	{'pos': [121, 131], 'id': 'm.05fjf'	{'pos': [330, 340], 'id': 'm.0h6l4', 'name': 'Plainfield'}
19	The president said that he had received assurances from Moscow tha	Moscow is located in Russia	0.438936323	/location/location/contains	{'pos': [68, 74], 'id': 'm.06bnz',	{'pos': [56, 62], 'id': 'm.04swd', 'name': 'Moscow'}
20	The division of Latin America 's largest country into two evenly match	Brazil is located in Latin Am	0.885988176	/location/location/contains	{'pos': [16, 29], 'id': 'm.04pnx',	{'pos': [205, 211], 'id': 'm.015fr', 'name': 'Brazil'}
21	Airbus has 344 planes in service in China , Hong Kong and Macao , bu	Hong Kong is located in Chi	0.38832432	/location/location/contains	{'pos': [36, 41], 'id': 'm.0d05w3	{'pos': [44, 53], 'id': 'm.03h64', 'name': 'Hong Kong'}
22	Dale Eastlund , head of the air consulting group , said that Lagos and	Bangalore is located in India	0.959626317	/location/location/contains	{'pos': [138, 143], 'id': 'm.03rk(	{'pos': [71, 80], 'id': 'm.09c17', 'name': 'Bangalore'}
23	Drawing on a business plan he devised three years ago as a senior in	Dallas is located in Texas .	0.747941196	/location/location/contains	{'pos': [69, 74], 'id': 'm.07b_l', '	{'pos': [202, 208], 'id': 'm.0f2rq', 'name': 'Dallas'}
24	He has a serious menu of dishes like smoked sunchoke soup with bra	Carroll Gardens is located ir	0.9781425	/location/location/contains	{'pos': [221, 229], 'id': 'm.0cr3	{'pos': [203, 218], 'id': 'm.026j4b', 'name': 'Carroll Gardens'}
25	Because the story that follows is really about Joseph Simmons -LRB-	Hollis is located in Queens .	0.97966671	/location/location/contains	{'pos': [169, 175], 'id': 'm.0ccv:	{'pos': [160, 166], 'id': 'm.02vrh8', 'name': 'Hollis'}
26	When the N.H.L. 's 30 general managers meet before the entry draft	Vancouver is located in Briti	0.982167959	/location/location/contains	{'pos': [93, 109], 'id': 'm.015jr',	{'pos': [81, 90], 'id': 'm.080h2', 'name': 'Vancouver'}
27	The results scared the pants off Mr. Sondheim when he caught an ea	Newbury is located in Engla	0.95806396	/location/location/contains	{'pos': [197, 204], 'id': 'm.02jx1	{'pos': [187, 194], 'id': 'm.0xnvx', 'name': 'Newbury'}
28	The relative vernacular cosmopolitanism of a region like Princeton als	Princeton is located in New	0.843381166	/location/location/contains	{'pos': [80, 90], 'id': 'm.05fjf', 'n	{'pos': [57, 66], 'id': 'm.0ljsz', 'name': 'Princeton'}
29	-LRB- WPP and VNU are partners in a joint venture , AGB Nielsen Me	Hong Kong is located in Chi	0.067435063	/location/location/contains	{'pos': [168, 173], 'id': 'm.0d05	{'pos': [176, 185], 'id': 'm.03h64', 'name': 'Hong Kong'}
30	Mr. Arif was born in Oran , Algeria , in 1965 and was a lieutenant in th	Oran is located in Algeria .	0.979435325	/location/location/contains	{'pos': [28, 35], 'id': 'm.0h3y', '	{'pos': [21, 25], 'id': 'm.012gg6', 'name': 'Oran'}
31	DEYARMIN , Daniel N. Jr. , 22 , Lance Cpl. , Marine Forces Reserve ; Ta	Tallmadge is located in Ohi	0.962878168	/location/location/contains	{'pos': [81, 85], 'id': 'm.05kkh',	{'pos': [69, 78], 'id': 'm.013k26', 'name': 'Tallmadge'}
32	She is a daughter of Barbara C. Weiler of Manhattan and the late Gec	Manhattan is located in Nev	0.678708732	/location/location/contains	{'pos': [138, 151], 'id': 'm.02_2	{'pos': [42, 51], 'id': 'm.0cc56', 'name': 'Manhattan'}
33	This is what collectors live for , '' said Robert H. Blumenfield , the chair	Christie is located in Califorr	0.093204737	/location/location/contains	{'pos': [109, 119], 'id': 'm.01n7	{'pos': [209, 217], 'id': 'm.03wfw08', 'name': 'Christie'}

基于深度学习的计算机视觉

目标检测

R-CNN　Fast R-CNN　Faster R-CNN　YOLO　SSD　Retina-Net　M2Det

语义分割

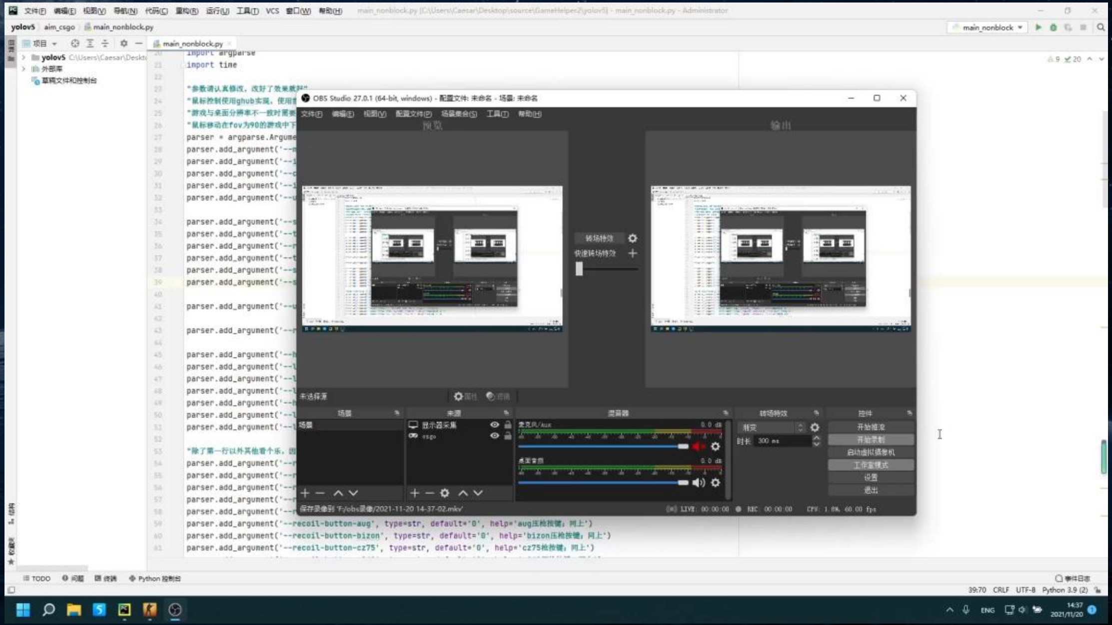
姿势估计

2022

基于深度学习的计算机视觉 Demo

将目标检测技术运用于游戏……

YOLOv5



https://github.com/Caesar-s1mple/csgo-yolov5

TECHNIQUES

无监督学习

GAN

迁移学习

BERT

知识蒸馏

分布式系统

优化方法

SGD、AdaGrad、Adadelta、Adam、AdaMax、NAdam

深度学习框架

APPLICATIONS

自然语言处理

情感分析、机器翻译、文本匹配、摘要、问答……

视觉数据处理

图像分类、目标检测和语义分割、视频处理、数据集……

语音和音频处理

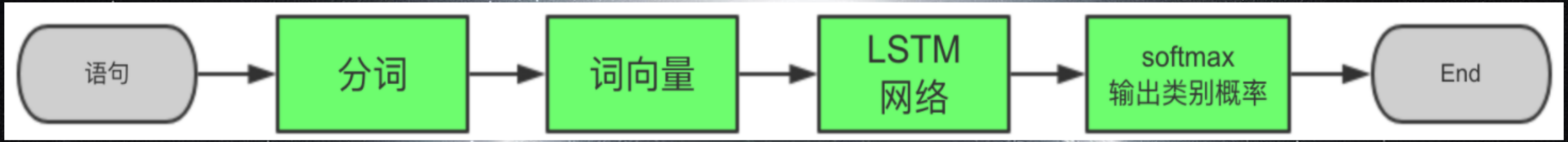
语音情感识别、语音增强、语音识别……

其他应用

社交网络分析、信息检索、交通预测、自动驾驶、生物医学、灾害管理系统……

自然语言处理

情感分析



传统方法：基于情感词典
优点：简单易行，通用性有保障
缺点：
1.精度不高。简单的线性叠加会造成很大的精度损失。词语权重同样不是一成不变的，而且也难以做到准确。
2.新词发现。对于新的情感词，词典不一定能够覆盖
3.词典构建难。基于词典的情感分类，核心在于情感词典。而情感词典的构建需要有较强的背景知识，需要对语言有较深刻的理解，在分析外语方面会有很大限制。

深度学习
优点：精度高，通用性强，不需要情感词典
难点：
1.语句太长
2.新词和口语化词的大量出现。需要在分词方面和词向量方面进行改进。

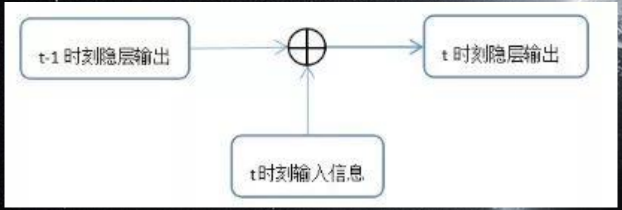
自然语言处理

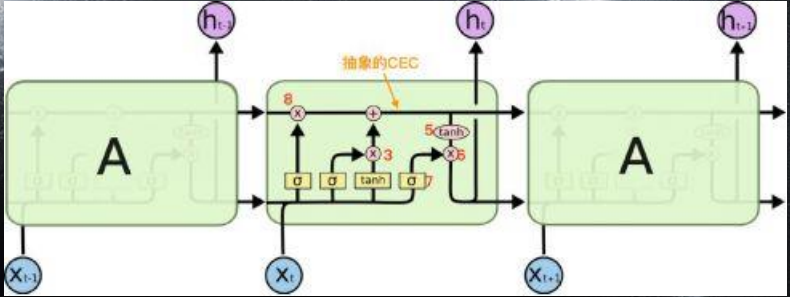
机器翻译

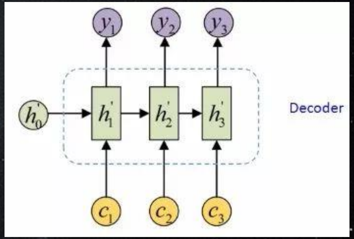
2014
Sutskever et.al.
Seq2Seq

1980s　　2012年

基于规则的机器翻译　　统计机器翻译　　神经网络机器翻译







优点：

1.超高的翻译表现。神经网络在翻译上的表现超越了过去 20 年来的一切，甚至学会了协调不同语言的性和格

2.直接翻译，无需中转语言。

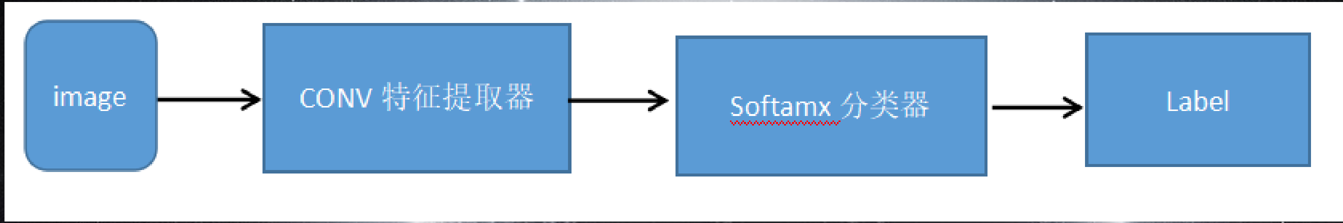
缺点：

翻译短句时出现问题。没有上下文或者是在训练数据中出现的较少，这时应使用简单的统计翻译。

视觉数据处理

图像分类

——图像中待识别物体单—



常用的标准网络模型：
LeNet、AlexNet、VGG系列、ResNet系列、Inception系列、DenseNet系列、GoogLeNet、NasBet、Xception、SeBet(state-of-art)
轻量化网络模型：
MobileNet v1,v2、ShuffleNet v1,v2,SqueezeNet
常用的数据集：
ImageNet、CIFAR10/100、Pascal VOC

目前轻量化模型在具体项目应用时用的比较广泛：
优点：
1.参数模型小，方便部署
2.计算量小，速度快
缺点：
轻量化模型在精度上没有ResNet系列、Inception系列、DenseNet系列、SeNet的accuracy高

视觉数据处理

目标检测和语义分割

——图像中含有多个目标物

常用数据集：VOC2012,MSCOCO

FCN

SegNet

Dilated Convolutions

DeepLab (v1 & v2)

RefineNet

PSPNet

Large Kernel Matters

逐像素分割

去掉了末端的全连接层，加入了对池化层的处理：使用编码-译码架构或是采用膨胀卷积。

模型	分数 (VOC2012)
FCN	67.2
SegNet	59.9
Dilated Convolutions	75.3
DeepLab (v1 & v2)	79.7
RefineNet	84.2
PSPNet	85.4
Large Kernel Matters	83.6
DeepLab v3	85.7

视觉数据处理

PSPNET

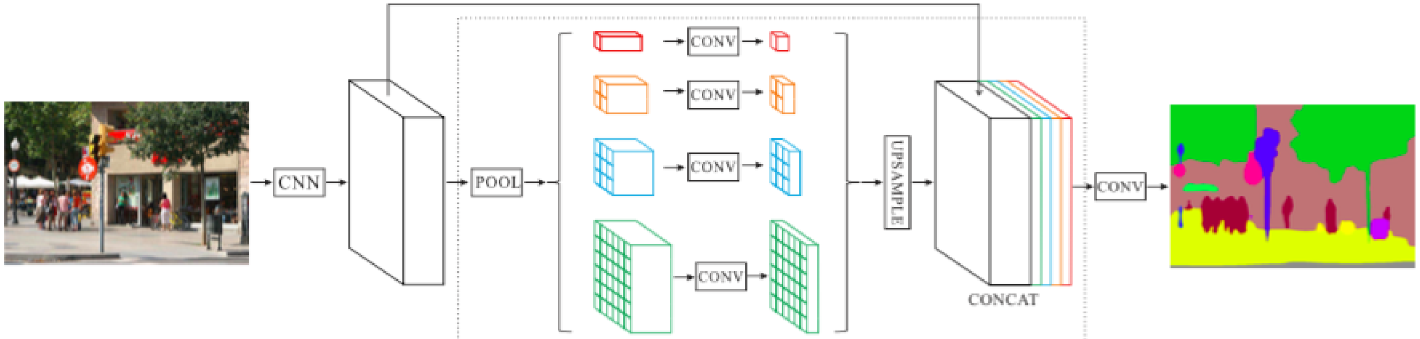
PSPNet 2016年

Pyramid Scene Parsing Network 金字塔场景解析网络

主要贡献：

提出了金字塔池化模块来聚合图片信息

使用附加的损失函数

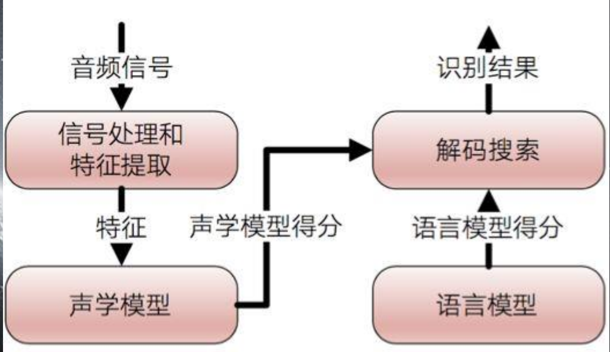


(a) Input Image (b) Feature Map (c) Pyramid Pooling Module (d) Final Prediction

语音和音频处理

语音识别

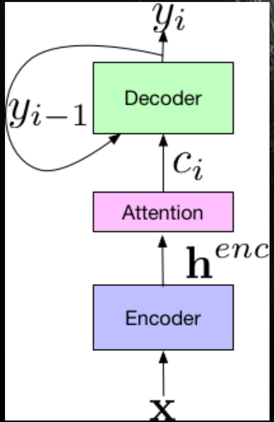
本质：序列识别问题
关键：信号处理预处理技术和声学模型部分



经典模型：HMM,GMM

传统模型存在音频信号表征的低效问题，因此考虑加入DNN进行处理，结合后得到了DNN-HMM模型
优点：训练成本不高，有相对较高的识别概率

循环神经网络、LSTM、编码-解码框架、注意力机制等基于深度学习的声学模型



谷歌的 Listen-Attend-Spell (LAS) 端到端语音识别系统

其他应用

社交网络分析

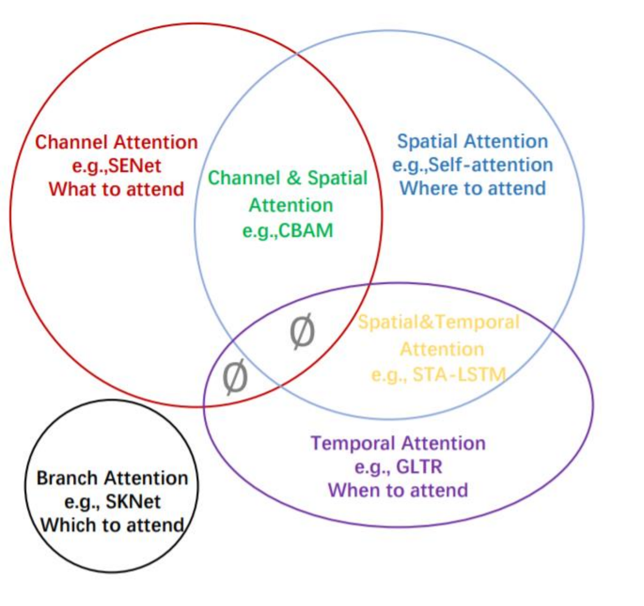
本质：一个由节点（人）和边（社交关系）组成的图
特性：小世界现象、无标度特性

1.社交网络结构分析
2.社交网络群体行为形成与互动规律
3.社交网络信息传播与演化机理
应用：社交推荐，舆情分析，隐私保护，用户画像，谣言检测等等

1.1 社区静态发现算法
Mark Newman 提出了针对模块度的最大化的贪心算法FN
1.2 社区动态发现算法
Palla, Gergely, et al.的派系过滤算法
1.3 社区演化分析
Hopcroft, John, et al.基于相邻时刻相似度直接比较的演化虚拟社区发现
2.1用户行为分析
2.2 社交网络情感分析
2.3 个体影响力分析
3.1在线社交网络信息检索
3.2 社交网络信息传播规律
3.3 话题发现和演化
3.4 影响力最大化

前沿研究调研——方法

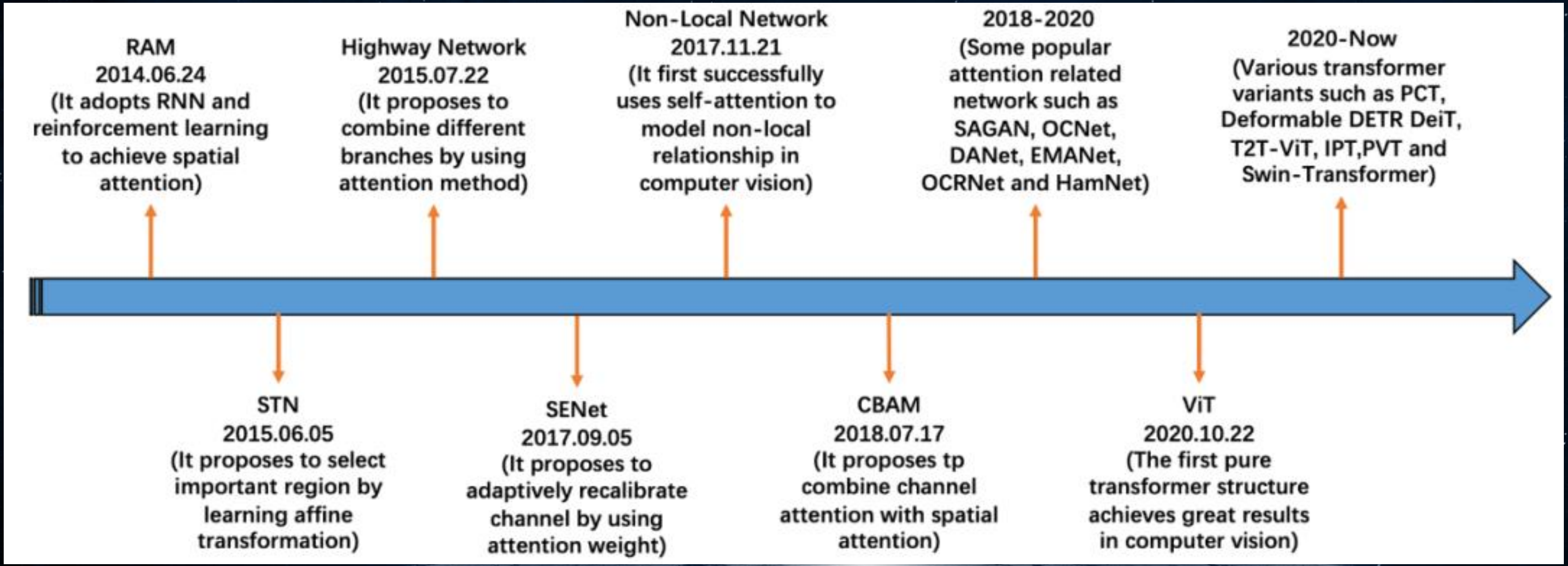
Attention



· 在过去的十年中，注意机制在计算机视觉中发挥着越来越重要的作用。

· 注意力机制可以根据数据域进行分类。其中包括通道注意力、空间注意力、时间注意力和分支注意力四大类基本注意力，以及通道&空间注意力和空间&时间注意力相结合的两大类混合注意力。

· 将不同的注意力方法根据其操作领域而不是应用任务进行分组。

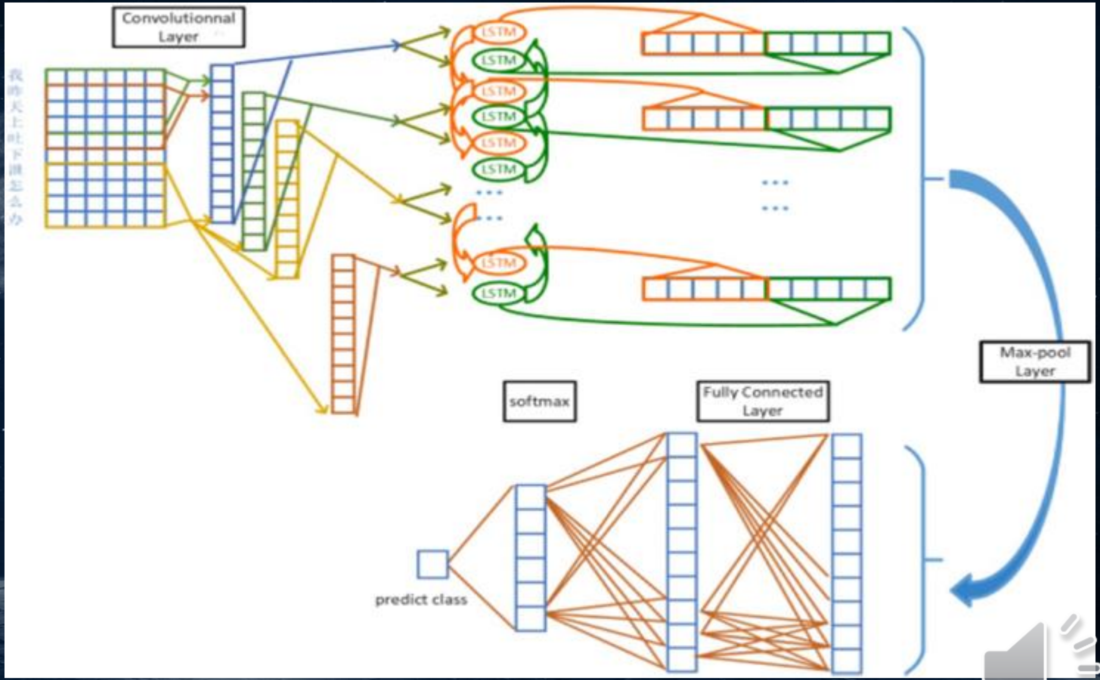


七个潜在的研究方向:

1. 注意力机制的充分必要条件
2. 更加通用的注意力模块
3. 注意力机制的可解释性
4. 注意力机制中的稀疏激活
5. 基于注意力机制的预训练模型
6. 适用于注意力机制的优化方法
7. 部署注意力机制的模型

前沿研究调研——应用

虚假新闻检测（NLP）



Hybrid deep learning model

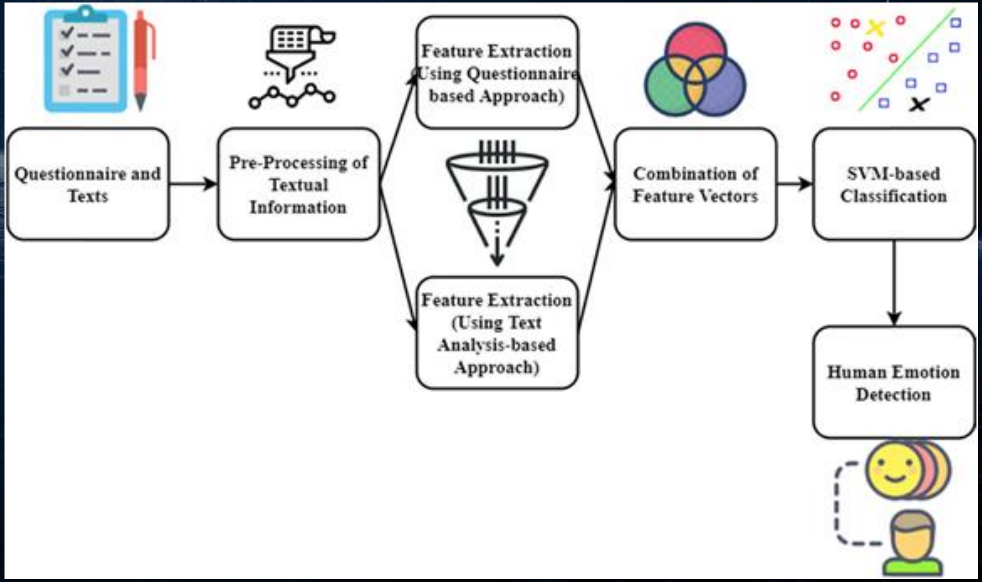
(i) data pre-processing
(ii) data augmentation using ACGAN
(iii) classifcation using hybrid CNN and RNN.

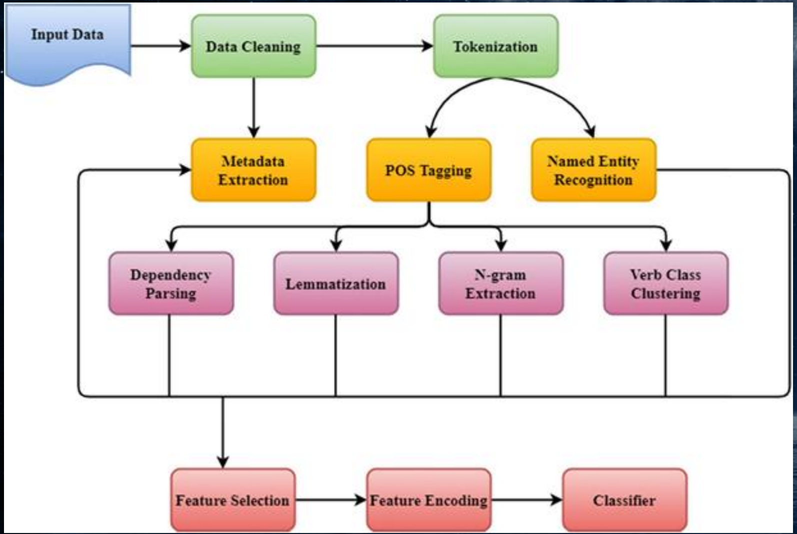
前沿研究调研——应用

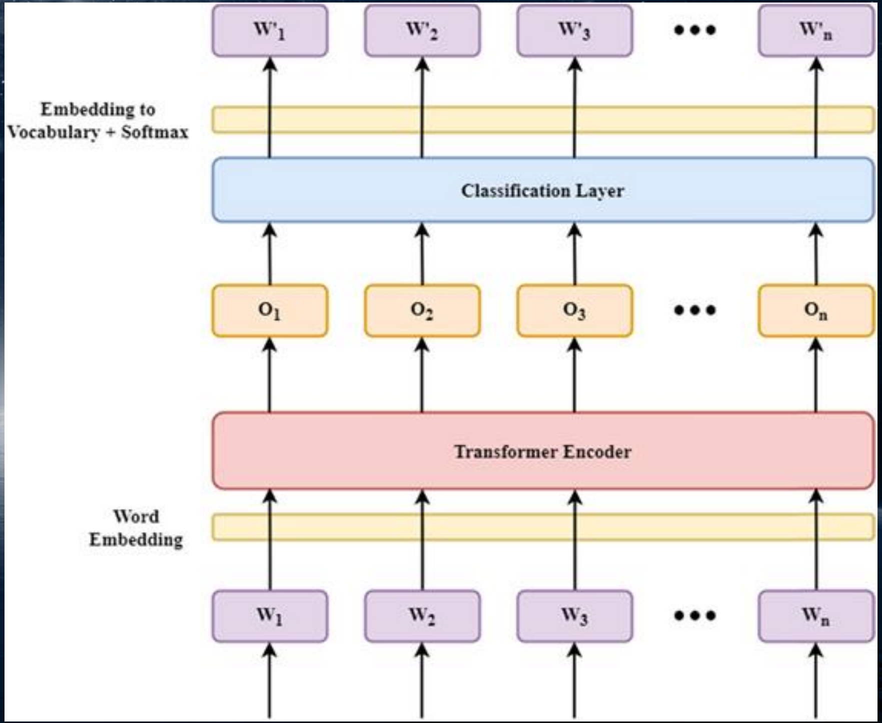
大数据中进行人类情感检测分析（NLP）

深度学习辅助语义文本分析，DLSTA

RNN&CNN：数据预处理，文本标注







Word2Vec（skip-grams）：
将数据映射至特征向量

前沿研究调研——应用

X射线检测新冠（医疗诊断，CV）

Multiscale Deep CNN Architecture（多尺度深度CNN）

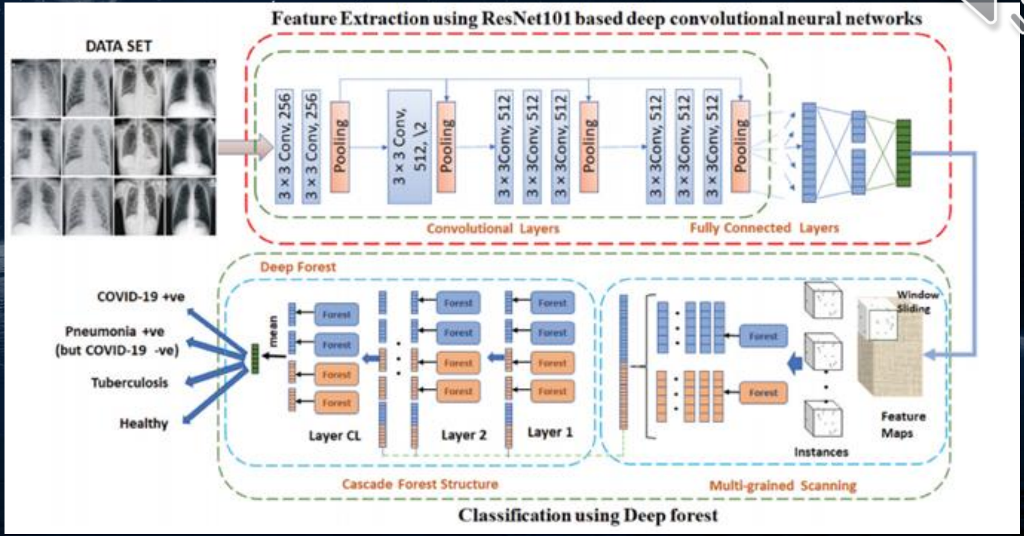
将图像分解为7种模式，每种模式都包含彼此唯一的信息。
此外，随机选择各种模态组合作为多尺度深度CNN的输入。
来自不同模式的信息聚合可以有效提高网络的预测能力。

前沿研究调研——应用

新冠早期诊断（CV）

使用ResNet101从胸部x射线图像中提取特征

使用深度森林模型预测感染者



- Requires a small number of hyper-parameters
- Trained on a multi-class dataset

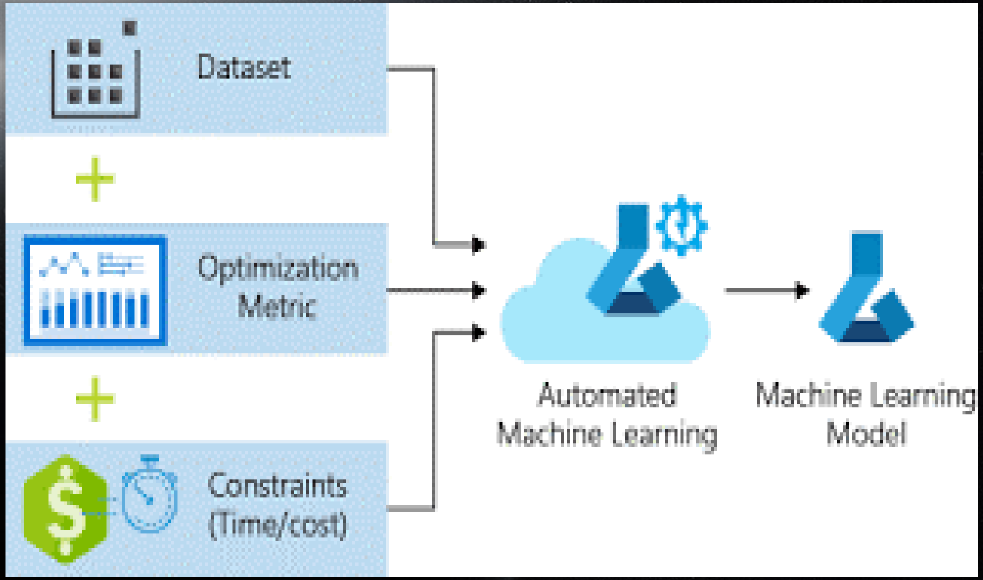
05 未来发展

- Automated Machine Learning (AutoML)
- Multi-modal Learning
- Tiny ML
- Quantum ML
- ……

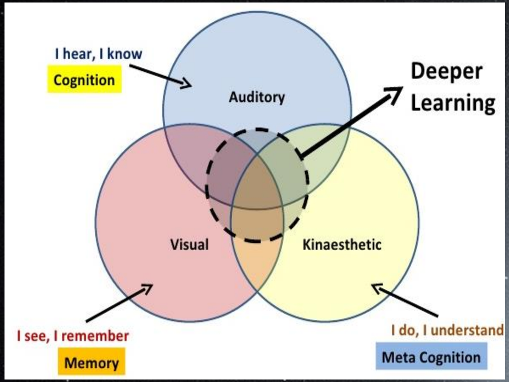
“Predicting the future isn’t magic, it’s artificial intelligence.” - Dave Waters.

Automated Machine Learning (AutoML)

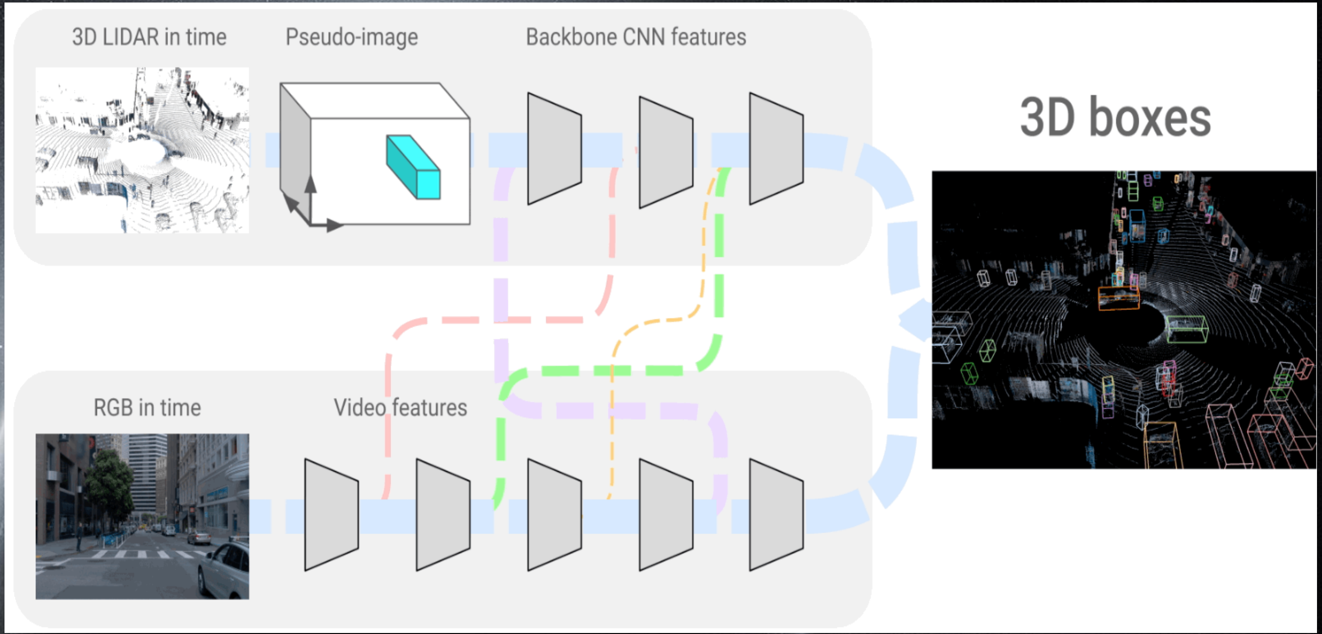
- 半监督学习和自我监督学习的改进
- 降低人工标注数据的成本
- 将选择和调整神经网络模型的工作自动化



Multi-model Learning



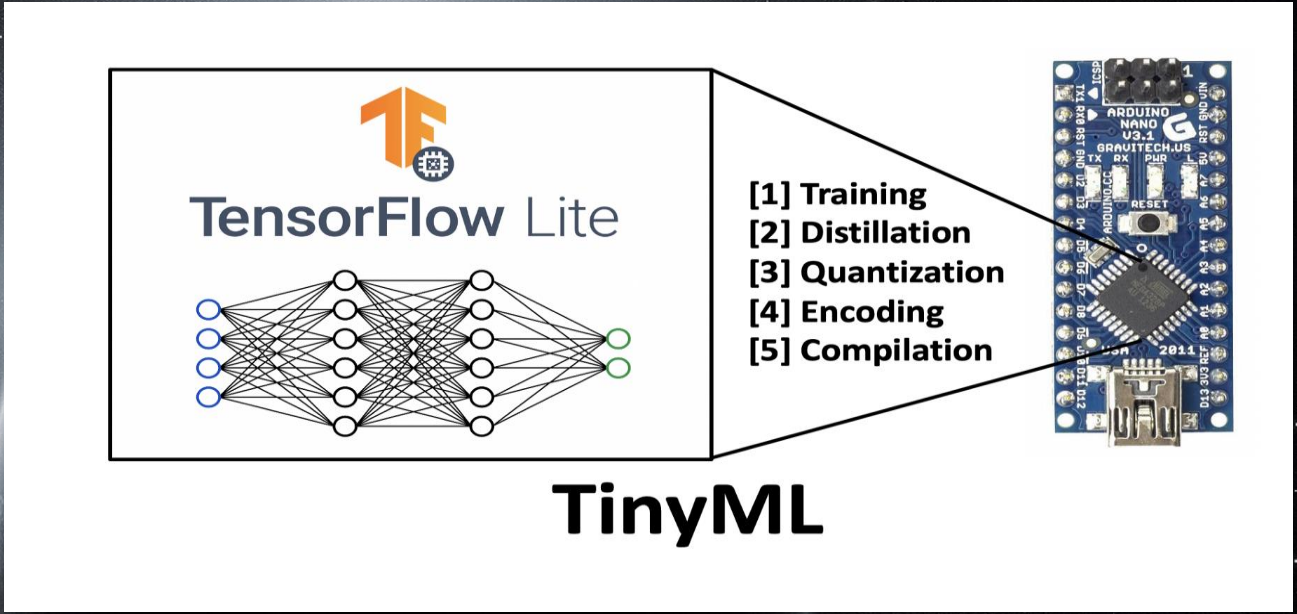
单个ML模型中支持多种模式，如文本、视觉、语音和物联网传感器数据。



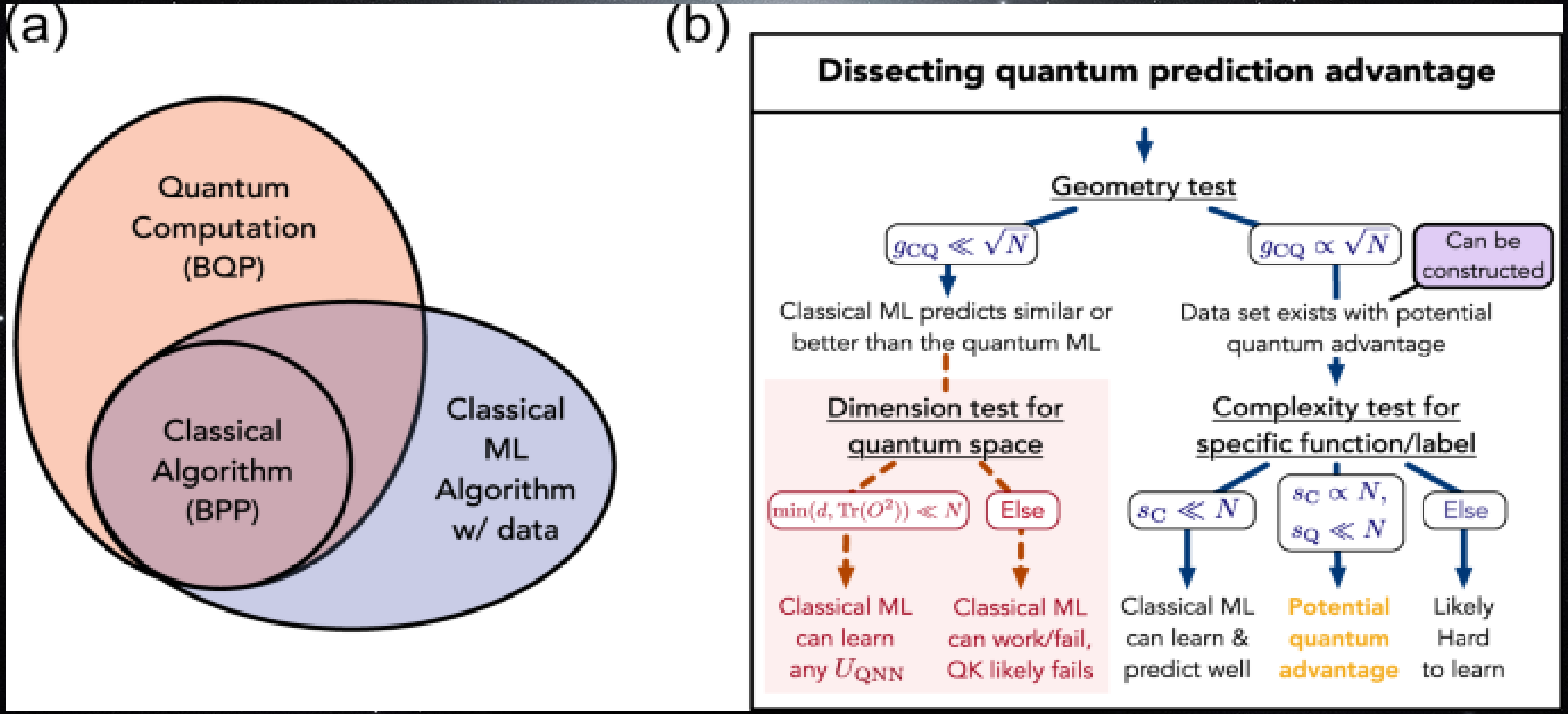
上图展示了一项自动驾驶汽车技术，有效地将3D激光雷达点云与RGB图像及时结合，也以视频的形式及时流化，学习不同传感器之间的连接及其特征表示。

Tiny ML

用于开发在硬件受限设备上运行的人工智能和ML模型，如为汽车、冰箱和电表供电的微控制器。



Quantum ML



(a)
Quantum Computation (BQP)
Classical Algorithm (BPP)
Classical ML Algorithm w/ data
(b)
Dissecting quantum prediction advantage
Geometry test
$g_{CQ} \ll \sqrt{N}$
$g_{CQ} \propto \sqrt{N}$
Can be constructed
Classical ML predicts similar or better than the quantum ML
Data set exists with potential quantum advantage
Dimension test for quantum space
Complexity test for specific function/label
$\min(d, \mathrm{Tr}(O^2)) \ll N$
Else
$s_C \ll N$
$s_C \propto N$, $s_Q \ll N$
Else
Classical ML can learn any U_{QNN}
Classical ML can work/fail, QK likely fails
Classical ML can learn & predict well
Potential quantum advantage
Likely Hard to learn

Presentation

Thanks

深度学习基础