课程大作业

重庆大学人工智能导论课程

类脑感知与普适智能研究院

微电子与通信工程学院

周喜川 教授

2020.9

完成内容

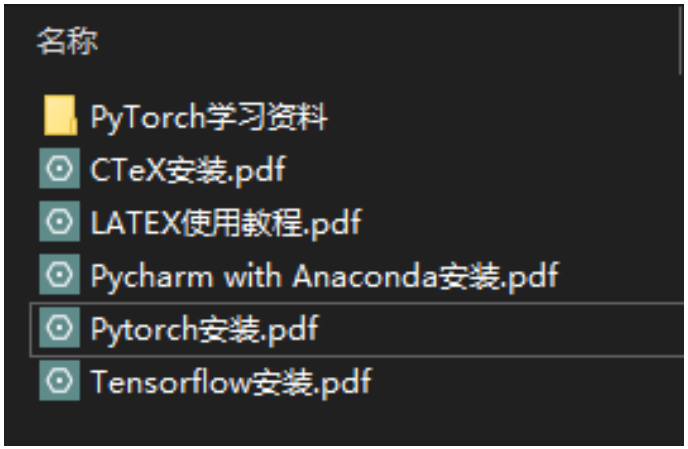
- 根据提供选题，完成一篇**课程论文**（占总成绩70%）
- **选题**：
  - 综述性文献（可选领域）
    1. 领域一：基于FPGA的深度学习加速
    2. 领域二：深度学习遥感图像处理
    3. 领域三：机器视觉应用
  - 应用型论文
    a. 红外目标检测
    b. 遥感图像处理
    c. 高光谱图像分类

考核细则

知识点与考核项	分数
论文使用CTex编排（提供源文件）	+ 5
应用型论文提供源代码及说明文档	+ 5
在arXiv上发表	+ 5
在最终提交时间考试周（19周）之前提交	+ 0.5 / 周

文档说明

工具类



a) PyTorch 学习资料（PyTorch 的相关学习材料）

b) CTeX 安装流程（CTeX 是一个论文排版编译工具）

c) LATEX 使用教程

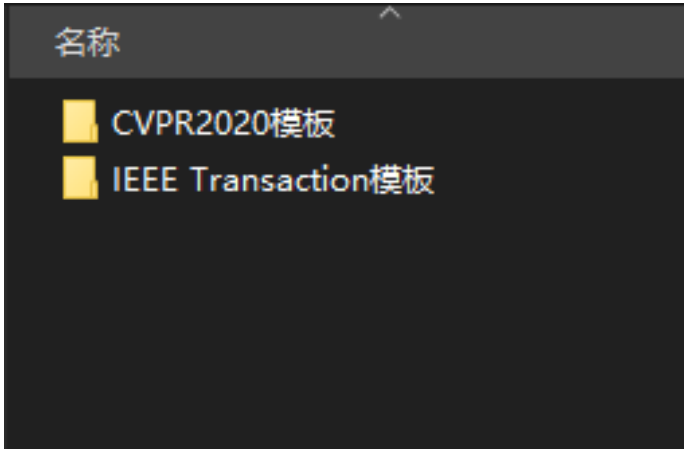
d) PyCharm with Anaconda 安装流程（搭建 Python 编译器）

e) PyTorch 安装（PyTorch 是一个常用的深度学习框架）

f) TensorFlow安装（TensorFlow同样是一个常用的深度学习框架）

文档说明

论文支持类

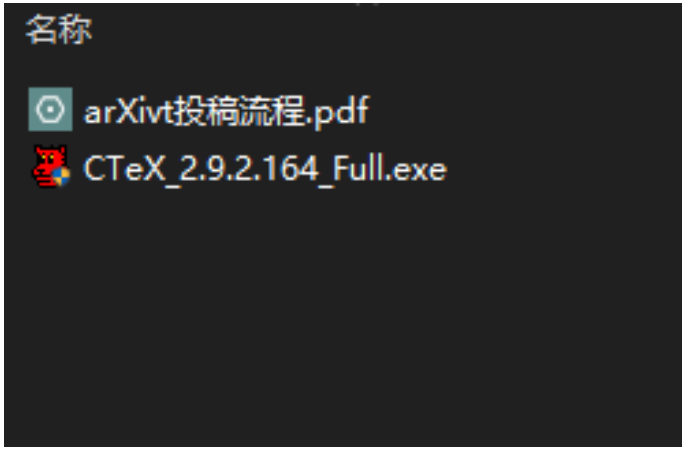


a) CVPR2020 模板 （包括LaTex版本和Word版本）

b) IEEE Transaction 模板（包括LaTex版本和Word版本）

文档说明

资源类

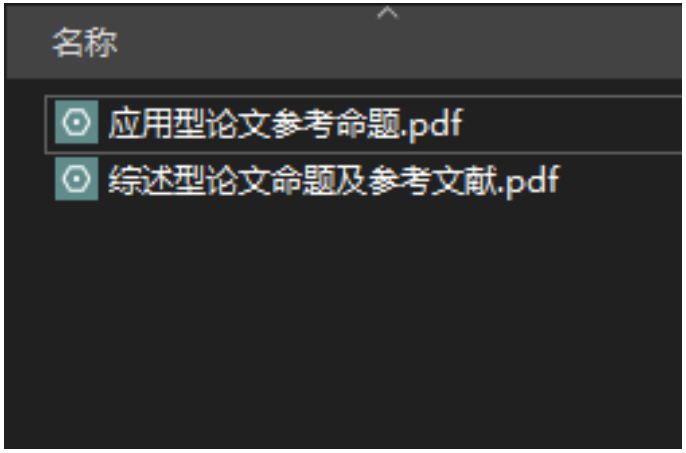


a) arXiv 上传流程（最终的课程论文要上传到 arXiv 上）

b) CTeX.exe（CTex可执行文件）

文档说明

论文参考类

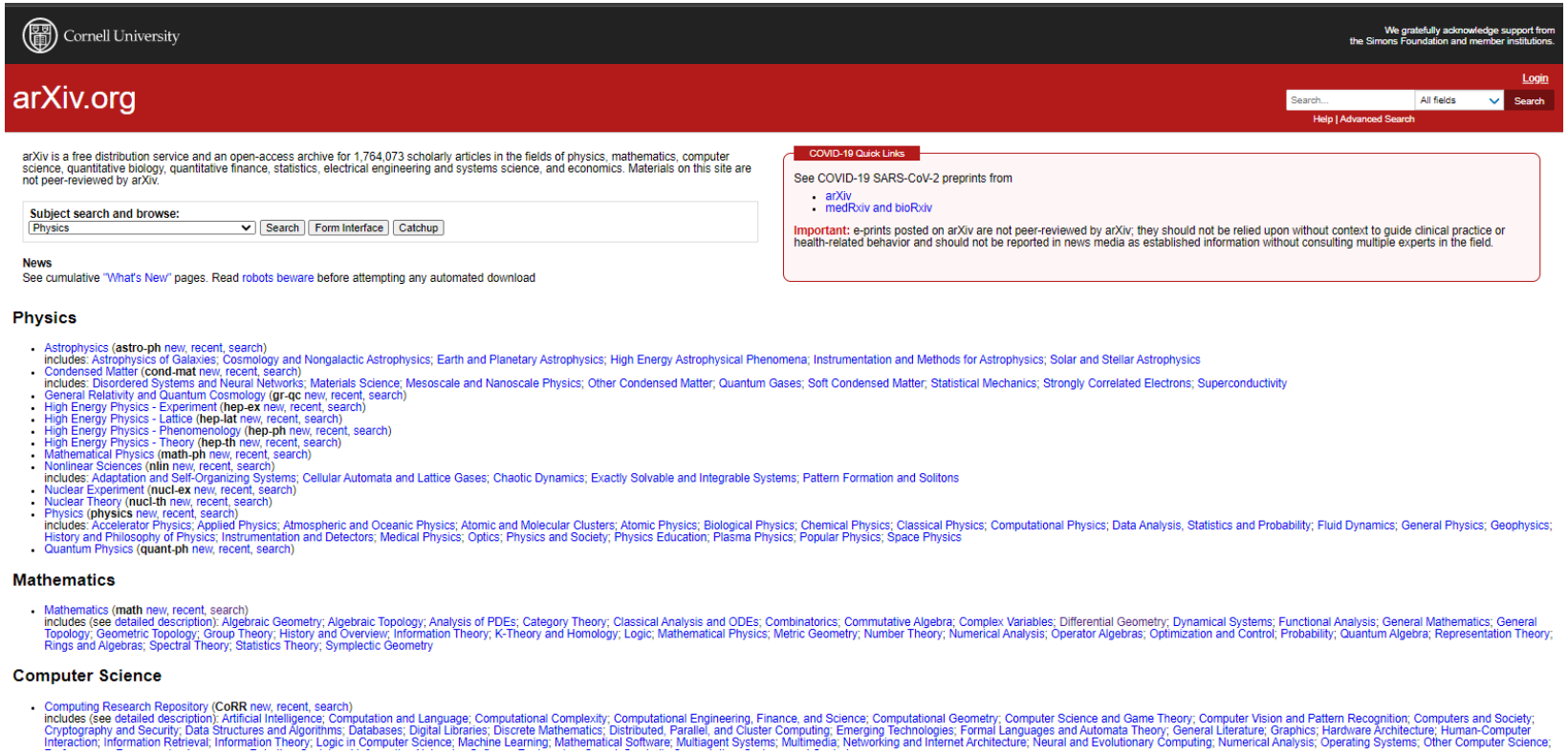


a) 应用型论文参考命题

b) 综述型论文命题及参考命题

基本要求

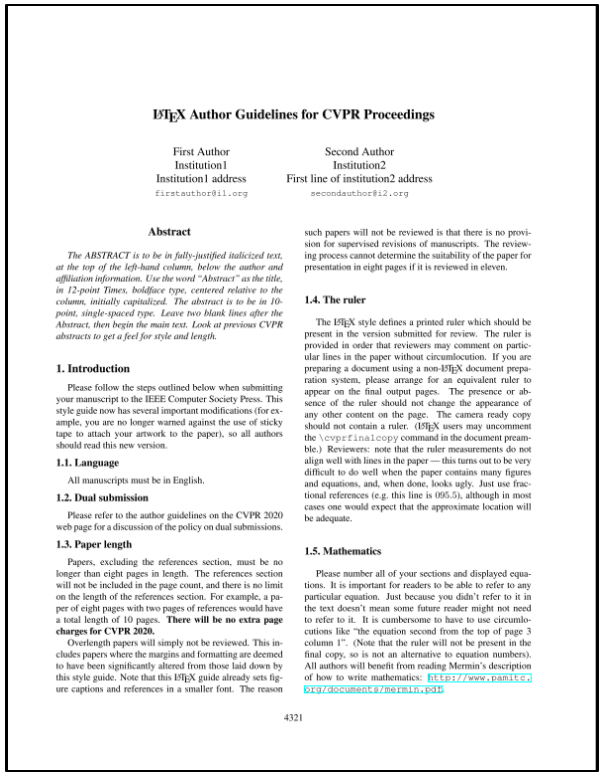
- **提交方式**：课程论文需要提交到在 $arXiv$[1] 上，最终评分将根据 arXiv 上投稿的课程论文进行评分。



[1] https://arxiv.org/

基本要求

- **篇幅及文档格式**：6页及以上，全英文，PDF 或 LaTex[1]源文档。论文风格参考CVPR的模板或IEEE模板。



LaTeX Author Guidelines for CVPR Proceedings

First Author
Institution1
Institution1 address
firstauthor@i1.org

Second Author
Institution2
First line of institution2 address
secondauthor@i2.org

Abstract

The ABSTRACT is to be in fully-justified italicized text, at the top of the left-hand column, below the author and affiliation information. Use the word "Abstract" as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type. Leave two blank lines after the Abstract, then begin the main text. Look at previous CVPR abstracts to get a feel for style and length.

1. Introduction

Please follow the steps outlined below when submitting your manuscript to the IEEE Computer Society Press. This style guide now has several important modifications (for example, you are no longer warned against the use of sticky tape to attach your artwork to the paper), so all authors should read this new version.

1.1. Language

All manuscripts must be in English.

1.2. Dual submission

Please refer to the author guidelines on the CVPR 2020 web page for a discussion of the policy on dual submissions.

1.3. Paper length

Papers, excluding the references section, must be no longer than eight pages in length. The references section will not be included in the page count, and there is no limit on the length of the references section. For example, a paper of eight pages with two pages of references would have a total length of 10 pages. **There will be no extra page charges for CVPR 2020.**

Overlength papers will simply not be reviewed. This includes papers where the margins and formatting are deemed to have been significantly altered from those laid down by this style guide. Note that this LaTeX guide already sets figure captions and references in a smaller font. The reason such papers will not be reviewed is that there is no provision for supervised revisions of manuscripts. The reviewing process cannot determine the suitability of the paper for presentation in eight pages if it is reviewed in eleven.

1.4. The ruler

The LaTeX style defines a printed ruler which should be present in the version submitted for review. The ruler is provided in order that reviewers may comment on particular lines in the paper without circumlocution. If you are preparing a document using a non-LaTeX document preparation system, please arrange for an equivalent ruler to appear on the final output pages. The presence or absence of the ruler should not change the appearance of any other content on the page. The camera ready copy should not contain a ruler. (LaTeX users may uncomment the `\cvprfinalcopy` command in the document preamble.) Reviewers: note that the ruler measurements do not align well with lines in the paper — this turns out to be very difficult to do well when the paper contains many figures and equations, and, when done, looks ugly. Just use fractional references (e.g. this line is 095.5), although in most cases one would expect that the approximate location will be adequate.

1.5. Mathematics

Please number all of your sections and displayed equations. It is important for readers to be able to refer to any particular equation. Just because you didn't refer to it in the text doesn't mean some future reader might not need to refer to it. It is cumbersome to have to use circumlocutions like "the equation second from the top of page 3 column 1". (Note that the ruler will not be present in the final copy, so is not an alternative to equation numbers). All authors will benefit from reading Mermin's description of how to write mathematics: `http://www.pamitc.org/documents/mermin.pdf`.

4321

CVPR会议论文模板

Conference Paper Title*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

This document is a model and instructions for LaTeX. Please observe the conference page limits.

II. Ease of Use

A. Maintaining the Integrity of the Specifications

The IEEEtran class file is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections III-A–III-E below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads—LaTeX will do that for you.

Identify applicable funding agency here. If none, delete this.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as "3.5-inch disk drive".
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: "Wh/m^2" or "webers per square meter", not "webers/m^2". Spell out units when they appear in text: ". . . a few henries", not ". . . a few H".
- Use a zero before decimal points: "0.25", not ".25". Use "cm^3", not "cc".)

C. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \quad (1)$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(1)", not

IEEE会议论文模板

[1] https://www.latex-project.org/

基本要求

重庆大学 CHONGQING UNIVERSITY

论文内容：5段式，包括

- Introduction
- Related Work
- Math
- Experiment
- Conclusion

1 Introduction

Deep neural networks (DNN) have shown significant improvements in several application domains including computer vision and speech recognition. In computer vision, a

2 Related Work

Deep neural networks often suffer from over-parametrization and large amounts of redundancy in their models. This typically results in inefficient computation and memory

3 Binary Convolutional Neural Network

We represent an L-layer CNN architecture with a triplet $\langle \mathcal{I}, \mathcal{W}, * \rangle$. $\mathcal{I}$ is a set of tensors, where each element $\mathbf{I} = \mathcal{I}_{l(l=1,\ldots,L)}$ is the input tensor for the l^{th} layer of CNN

4 Experiments

We evaluate our method by analyzing its efficiency and accuracy. We measure the efficiency by computing the computational speedup (in terms of number of high preci-

5 Conclusion

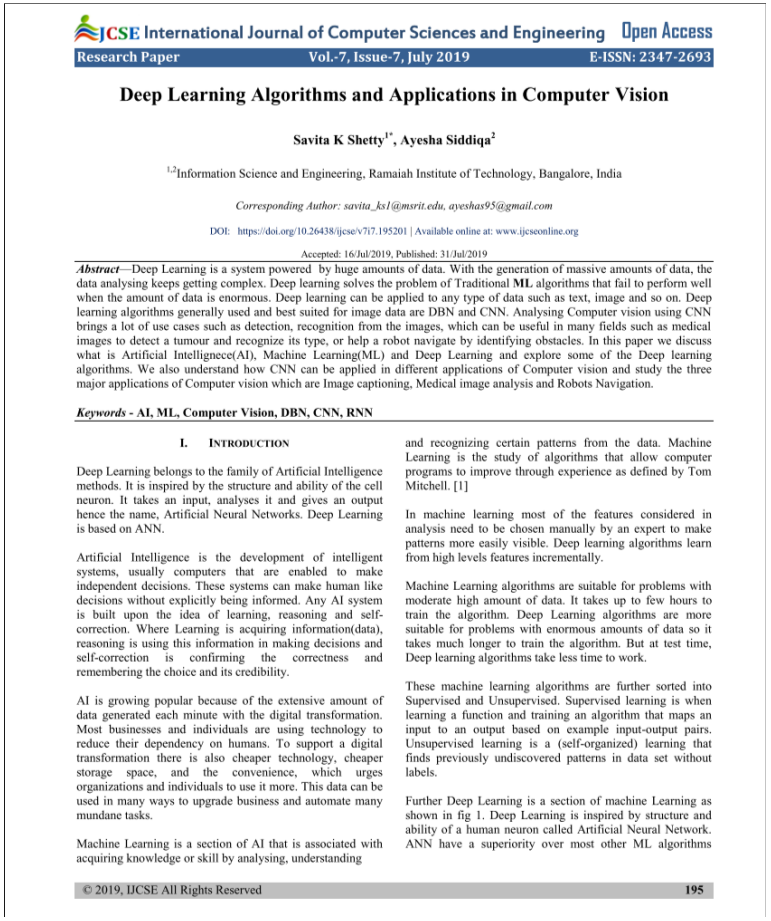
We introduce simple, efficient, and accurate binary approximations for neural networks. We train a neural network that learns to find binary values for weights, which reduces

[Rastegari M , Ordonez V , Redmon J , et al. XNOR-Net: ImageNet Classification Using Binary Convolutional Neural Networks[J]. 2016.]

Deep Learning Algorithms and Applications in Computer Vision

Introduction

导言部分引导读者从一个一般的主题领域到一个特定的研究领域。



IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-7, July 2019 E-ISSN: 2347-2693

Deep Learning Algorithms and Applications in Computer Vision

Savita K Shetty[1*], Ayesha Siddiqa[2]

[1,2]Information Science and Engineering, Ramaiah Institute of Technology, Bangalore, India

Corresponding Author: savita_ks1@msrit.edu, ayeshas95@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7i7.195201 | Available online at: www.ijcseonline.org

Accepted: 16/Jul/2019, Published: 31/Jul/2019

***Abstract*—Deep Learning is a system powered by huge amounts of data. With the generation of massive amounts of data, the data analysing keeps getting complex. Deep learning solves the problem of Traditional ML algorithms that fail to perform well when the amount of data is enormous. Deep learning can be applied to any type of data such as text, image and so on. Deep learning algorithms generally used and best suited for image data are DBN and CNN. Analysing Computer vision using CNN brings a lot of use cases such as detection, recognition from the images, which can be useful in many fields such as medical images to detect a tumour and recognize its type, or help a robot navigate by identifying obstacles. In this paper we discuss what is Artificial Intelligence(AI), Machine Learning(ML) and Deep Learning and explore some of the Deep learning algorithms. We also understand how CNN can be applied in different applications of Computer vision and study the three major applications of Computer vision which are Image captioning, Medical image analysis and Robots Navigation.**

***Keywords* - AI, ML, Computer Vision, DBN, CNN, RNN**

I. INTRODUCTION

Deep Learning belongs to the family of Artificial Intelligence methods. It is inspired by the structure and ability of the cell neuron. It takes an input, analyses it and gives an output hence the name, Artificial Neural Networks. Deep Learning is based on ANN.

Artificial Intelligence is the development of intelligent systems, usually computers that are enabled to make independent decisions. These systems can make human like decisions without explicitly being informed. Any AI system is built upon the idea of learning, reasoning and self-correction. Where Learning is acquiring information(data), reasoning is using this information in making decisions and self-correction is confirming the correctness and remembering the choice and its credibility.

AI is growing popular because of the extensive amount of data generated each minute with the digital transformation. Most businesses and individuals are using technology to reduce their dependency on humans. To support a digital transformation there is also cheaper technology, cheaper storage space, and the convenience, which urges organizations and individuals to use it more. This data can be used in many ways to upgrade business and automate many mundane tasks.

Machine Learning is a section of AI that is associated with acquiring knowledge or skill by analysing, understanding and recognizing certain patterns from the data. Machine Learning is the study of algorithms that allow computer programs to improve through experience as defined by Tom Mitchell. [1]

In machine learning most of the features considered in analysis need to be chosen manually by an expert to make patterns more easily visible. Deep learning algorithms learn from high levels features incrementally.

Machine Learning algorithms are suitable for problems with moderate high amount of data. It takes up to few hours to train the algorithm. Deep Learning algorithms are more suitable for problems with enormous amounts of data so it takes much longer to train the algorithm. But at test time, Deep learning algorithms take less time to work.

These machine learning algorithms are further sorted into Supervised and Unsupervised. Supervised learning is when learning a function and training an algorithm that maps an input to an output based on example input-output pairs. Unsupervised learning is a (self-organized) learning that finds previously undiscovered patterns in data set without labels.

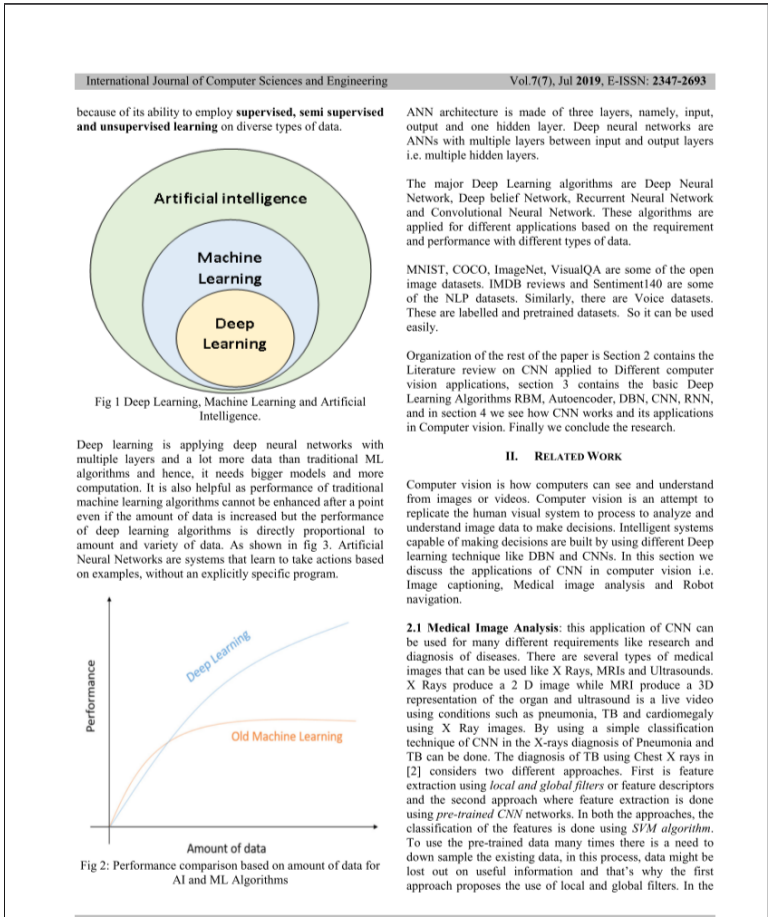
Further Deep Learning is a section of machine Learning as shown in fig 1. Deep Learning is inspired by structure and ability of a human neuron called Artificial Neural Network. ANN have a superiority over most other ML algorithms

[Shetty S K , Siddiqa A . Deep Learning Algorithms and Applications in Computer Vision[J]. INTERNATIONAL JOURNAL OF COMPUTER ENCES AND ENGINEERING, 2019, 7(7):195-201..]

Deep Learning Algorithms and Applications in Computer Vision

Related Work

对相关领域研究进行综述。阐述目前的关键问题，如何解决问题，评估研究成果。



International Journal of Computer Sciences and Engineering Vol.7(7), Jul 2019, E-ISSN: 2347-2693

because of its ability to employ **supervised, semi supervised and unsupervised learning** on diverse types of data.

Fig 1 Deep Learning, Machine Learning and Artificial Intelligence.

Deep learning is applying deep neural networks with multiple layers and a lot more data than traditional ML algorithms and hence, it needs bigger models and more computation. It is also helpful as performance of traditional machine learning algorithms cannot be enhanced after a point even if the amount of data is increased but the performance of deep learning algorithms is directly proportional to amount and variety of data. As shown in fig 3. Artificial Neural Networks are systems that learn to take actions based on examples, without an explicitly specific program.

Fig 2: Performance comparison based on amount of data for AI and ML Algorithms

ANN architecture is made of three layers, namely, input, output and one hidden layer. Deep neural networks are ANNs with multiple layers between input and output layers i.e. multiple hidden layers.

The major Deep Learning algorithms are Deep Neural Network, Deep belief Network, Recurrent Neural Network and Convolutional Neural Network. These algorithms are applied for different applications based on the requirement and performance with different types of data.

MNIST, COCO, ImageNet, VisualQA are some of the open image datasets. IMDB reviews and Sentiment140 are some of the NLP datasets. Similarly, there are Voice datasets. These are labelled and pretrained datasets. So it can be used easily.

Organization of the rest of the paper is Section 2 contains the Literature review on CNN applied to Different computer vision applications, section 3 contains the basic Deep Learning Algorithms RBM, Autoencoder, DBN, CNN, RNN, and in section 4 we see how CNN works and its applications in Computer vision. Finally we conclude the research.

II. Related Work

Computer vision is how computers can see and understand from images or videos. Computer vision is an attempt to replicate the human visual system to process to analyze and understand image data to make decisions. Intelligent systems capable of making decisions are built by using different Deep learning technique like DBN and CNNs. In this section we discuss the applications of CNN in computer vision i.e. Image captioning, Medical image analysis and Robot navigation.

2.1 Medical Image Analysis: this application of CNN can be used for many different requirements like research and diagnosis of diseases. There are several types of medical images that can be used like X Rays, MRIs and Ultrasounds. X Rays produce a 2 D image while MRI produce a 3D representation of the organ and ultrasound is a live video using conditions such as pneumonia, TB and cardiomegaly using X Ray images. By using a simple classification technique of CNN in the X-rays diagnosis of Pneumonia and TB can be done. The diagnosis of TB using Chest X rays in [2] considers two different approaches. First is feature extraction using *local and global filters* or feature descriptors and the second approach where feature extraction is done using *pre-trained CNN* networks. In both the approaches, the classification of the features is done using *SVM algorithm*. To use the pre-trained data many times there is a need to down sample the existing data, in this process, data might be lost out on useful information and that's why the first approach proposes the use of local and global filters. In the

[Shetty S K , Siddiqa A . Deep Learning Algorithms and Applications in Computer Vision[J]. INTERNATIONAL JOURNAL OF COMPUTER ENCES AND ENGINEERING, 2019, 7(7):195-201..]

Deep Learning Algorithms and Applications in Computer Vision

Math

阐述研究方法的具体细节、背后的数学原理以及公式推导等。一般会采取“总-分”的形式。

result image dataset. The next step is prostate segmentation or delineation. The segmentation process is based on three features, *appearance, shape prior and spatial relationship*, for more efficiency. The third step is *feature extraction* differentiating features are identified from the images to classify them into malignant or benign. The final step is classification into benign or malignant. This is done using a CNN and the features extracted to differentiation in earlier steps.

2.2 Navigation for Robots or autonomous vehicles is implemented as a combination of GPS system and image analysis. This can be done using Image processing techniques such as edge detection for identifying the lane/ or the free path to move in or by using better Deep learning algorithms since they don't need explicit training with all types of data, they perform better with inputs not familiar with.

YOLO is an object detection technique used in real time. YOLO uses a single CNN network to detect and localize the object in the image. YOLO is used in real time as it is believed to be very fast. [8]

Memory can also be used for generating the textual sentence.

III. Deep Learning Algorithms

Deep neural networks are not easy to train with back propagation due to the problem of vanishing gradient which impacts the time taken for training and reduces accuracy. Artificial Neural Networks calculate cost function based on the net difference between the Neural Network's predicted output and actual output in the training data. Based on the cost, weights and biases are altered after each process. Till the cost is as little as possible. Gradient is the rate at which cost will change based on weights and biases.

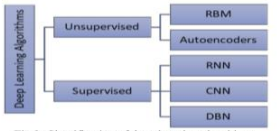


Fig 3: Classification of deep learning algorithms.

Table 1: Summary of Different applications of Deep learning in computer vision.

Problem	Method (Description)	References
Object detection	• YOLO is a technique based on CNN used for real time object detection which can be used in Robot navigation systems.	[8,9]
Object recognition	• Face recognition systems • Traffic signal recognition for self-driving cars.	[6] [10]
Medical image analysis	• xRays can be classified using CNN to detect TB and pneumonia. • Segmenting Lesion using an MRI Scan • Detecting Prostate cancer and predicting if its Benign or malignant.	[10] [11] [5]
Image captioning	• CNN and LSTM are used in feature extraction and sentence generation. • Extension of this application by visual question answering.	[11] [7]

This is the reason for late bloom of Deep nets. The problem of vanishing gradient can be avoided by using Deep Learning techniques. That is why Deep learning algorithms perform best with problems with huge data set.

3.2 Autoencoders

Autoencoder is a specialized Artificial Neural Network which learns a representation (encoding) for a set of data, by training the network to ignore signal noise. It also tries to re-generate the initial input from the reduced encoding a representation. The process of re generation of the input helps with dimensionality reduction as the system learns to ignore noise. An autoencoder may have any number of hidden layers.

Both RBMs and Autoencoders support unsupervised learning and are used in generative models because both techniques attempt to recreate the input.

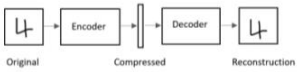


Fig 5: Autoencoder

3.3 Deep Belief Networks (DBN)

DBNs may be defined as a simple combination of unsupervised learning algorithms suck as RBMs and autoencoders.

[Shetty S K , Siddiqa A . Deep Learning Algorithms and Applications in Computer Vision[J]. INTERNATIONAL JOURNAL OF COMPUTER ENCES AND ENGINEERING, 2019, 7(7):195-201..]

Deep Learning Algorithms and Applications in Computer Vision

Experiment

这部分要详细介绍与实验相关的具体细节。包括介绍实验数据、评测标准和比较方法等基本信息。

Vol.7(7), Jul 2019, E-ISSN: 2347-2693

Pooling layer is the second layer and it helps in reducing the number of parameters since the images may be large. There are different types of pooling techniques based on requirement but most commonly used is Max Pooling which only considers the highest concentrated element of the obtained feature map.

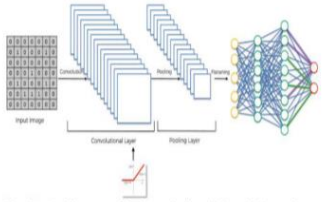


Fig 9: Architecture of a Convolutional Neural Network.

After flattening of the derived feature map, 1D array if fed into the Fully connected network. This functions just as any other neural network and after some back propagation most appropriate classification is presented.

Other than image analysis CNN is also used in Optical Character recognition to convert hand written document into digital text, which is a part of Natural Language Processing.

IV. Applications of Deep Learning in Computer Vision

Convolutional Neural Networks are specialized fully connected neural networks with a shared weight architecture. It uses to small simple patterns to identify and analyse larger complex patterns thus reducing the risk of overfitting that may be caused by full connection.

A human eye tries to identify certain features in an image to analyse, classify and recognize the image. Similarly, CNNs use features through feature map vectors to identify the feature that can be used to classify images. Input to a CNN is image or video stream and the output is class the image belongs to.

Table 2 : COMPARISON OF THE DL ALGORITHMS

Parameter	RBM	Autoencoder	CNN	RNN	DBN
Type of learning	Unsupervised	Unsupervised	Supervised	Supervised	Supervised
Generative / Discriminative	Generative	Generative	Discriminative	Discriminative / Generative	Generative
Input data	Any type of data	Any type of data	3-D structured data, like Voice, Images	Mainly, Textual data,	Text, image
Output	Reconstructed input	Reconstructed input	Classified, predicted output	Sequence Prediction	Classified, predicted output
Application	Dimensionality Reduction/ Classification	Dimensionality Reduction	Image and voice analysis, classification, detection, recognition	NLP, Speech recognition	NLP, dimensionality reduction [6]

Some of the other major applications of CNN are in (a) Image recognition system like face recognition in smartphones use CNN and analysis of medical images to find tumors and classify it. (b) Video analysis since videos are like images with a temporal dimension it can be applied on videos. (c) NLP: CNNs are used for sentence retrieval, classification, prediction and other NLP tasks. (d) Drug Discovery AtomNet is a specialized CNN architecture used to discover chemical features like bonds between elements.

4.1 Medical Image analysis: is one of the most used applications of CNN. The main idea is to use medical images to extract information that may be missed because of human error or just automate the feature extraction to further help in more effective clinical diagnosis. Image analysis is useful for different purposes like segmentation, abnormality detection, disease classification, or computer aided diagnosis where images from medical imaging techniques such as X-ray or MRI are interpreted by systems.

4.2 Robot Navigation: Navigation of an autonomous Robot may happen through sensors, GPS or vision. Vision based systems to navigate through the way of a robot's motion can be implemented using Convolutional Neural Network. This implementation requires high computation and dataset with perfect labels. This application requires object detection and lane detection techniques.

4.3 Image Captioning: is a more specific application of CNN. Basic architecture of an image captioning architecture

[Shetty S K , Siddiqa A . Deep Learning Algorithms and Applications in Computer Vision[J]. INTERNATIONAL JOURNAL OF COMPUTER ENCES AND ENGINEERING, 2019, 7(7):195-201..]

Deep Learning Algorithms and Applications in Computer Vision

Conclusion

在论文最后会有总结展望，一般用一段来再次总结和强调本文的创新思路和实验结果，然后说明未来建议的研究方向和开放问题

Sequence Prediction	Classified, predicted output
NLP, Speech recognition	NLP, dimensionality reduction [6]

ns of CNN are in (a) face recognition in ' medical images to find alysis since videos are on it can be applied on for sentence retrieval, NLP tasks. (d) Drug CNN architecture used ds between elements.

one of the most used s to use medical images ssed because of human action to further help in ge analysis is useful for abnormality detection, aided diagnosis where iques such as X-ray or

f an autonomous Robot or vision. Vision based of a robot's motion can Neural Network. This tation and dataset with es object detection and

specific application of captioning architecture

V. Conclusion and Future Scope

Deep Learning is a part of artificial intelligence that is based on artificial neural networks. Deep Learning algorithms more suitable for problems with huge datasets, other problems with smaller datasets may be solved simply by using Machine Learning. We compare the different models used in different problems such as object detection, object recognition, captioning and so on.

Some of the major deep learning algorithms are briefly studied such as RBM and Autoencoder that use unsupervised learning and CNN, DBN and RNN that use supervised learning. The algorithms are compared based on their inputs, outputs and basic working. We compare these algorithms based on parameters such as inputs data, output data and applications.

Using CNN can reduce a lot of computation because it doesn't need to visit the image pixel by pixel instead CNN uses filters. We discuss some of the CNN applications in computer vision such as in Medical image analysis of digital medical images such as EEG, ECG, X-ray and MRI scan reports to find any anomalies or unusual growths. CNN in robot navigation to help robots or other autonomous systems to move without any human intervention. CNN is also used in combination with RNN for image caption generation.

Based on this study, it can be concluded that CNN can accomplish the desired result in deep learning problems with image inputs. However, CNN's require high computational costs since they require a GPU and in absence of GPU they are very slow to train since they need lot of training data. In many cases, this drawback can be overcome by using pre trained models by fine tuning based on requirements.

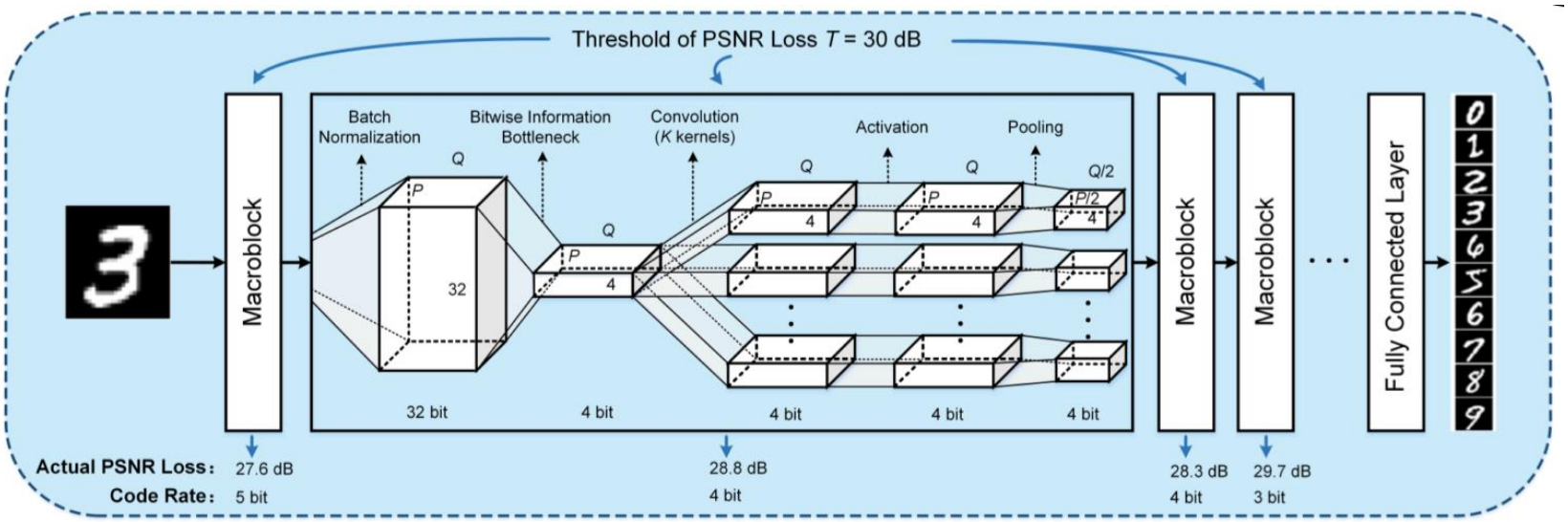
REFERENCES

[1] T. M. Mitchell, "Machine Learning", McGraw Hill Education; First edition, New York, USA

[Shetty S K , Siddiqa A . Deep Learning Algorithms and Applications in Computer Vision[J]. INTERNATIONAL JOURNAL OF COMPUTER ENCES AND ENGINEERING, 2019, 7(7):195-201..]

量化神经网络

Neural Network Activation Quantization with Bitwise Information Bottlenecks

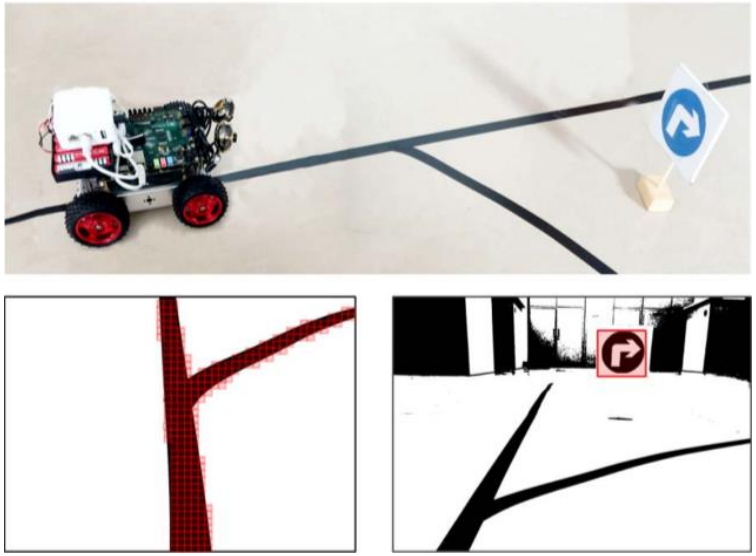


具有位信息瓶颈层的深度神经网络

[Zhou X , Liu K , Shi C , et al. Neural Network Activation Quantization with Bitwise Information Bottlenecks[J]. 2020.]

基于FPGA的神经网络设计

DANoC: An Efficient Algorithm and Hardware Codesign of Deep Neural Networks on Chip



配备了DANoC硬件原型的自动驾驶机器人汽车

Visible unit	Length	Start Address
1	0	-
2	1	0x00
3	0	-
4	2	0x01
5	0	-
6	0	-
7	0	-
8	1	0x03
9	1	0x04
Events	Address mapping table	

Figure legends:

1 1 Visible unit 1 is active

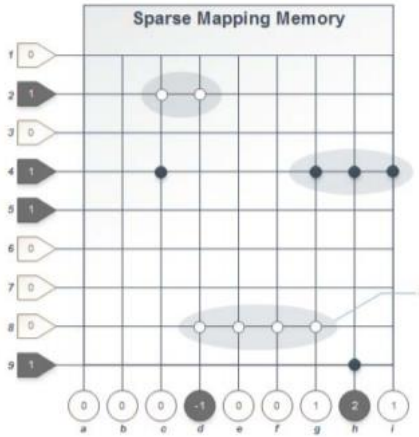
2 0 Visible unit 2 is inactive

2 h Neuron-h has a state of 2 and spikes

1 i Neuron-i has a state of 1 and keep static

	Hidden unit	Weight	Group Size
0x00	c	-1	2
0x01	c	1	1
	g	1	3
0x03	d	-1	4
0x04	h	1	1

Weight table

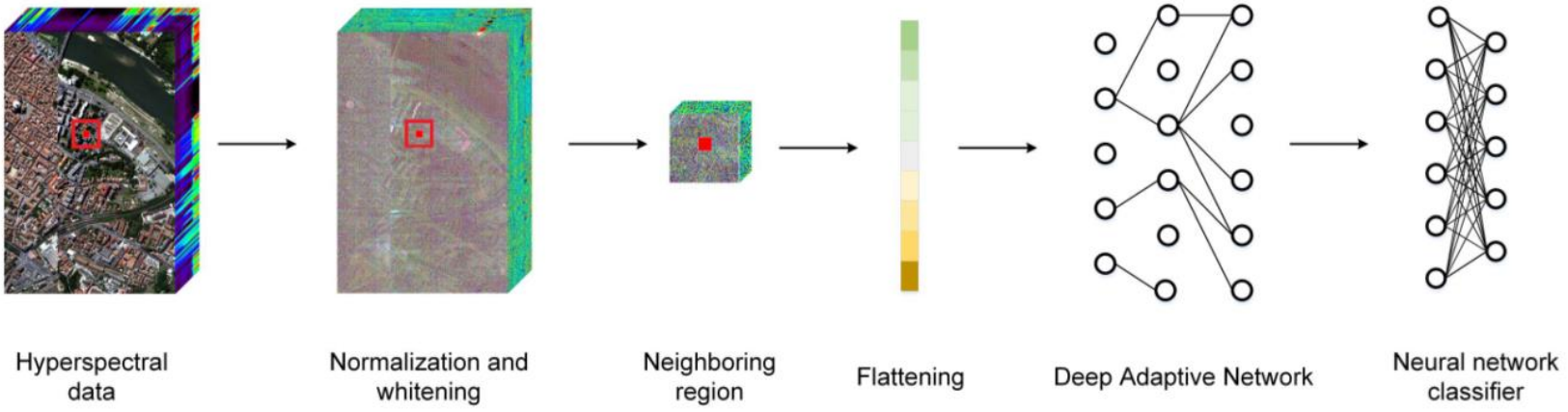


SNC的SMM

[Zhou X , Li S , Tang F , et al. DANoC: An Efficient Algorithm and Hardware Codesign of Deep Neural Networks on Chip[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, PP(7):1-12.]

高光谱遥感图像处理

Deep Adaptive Network: An Efficient Deep Neural Network with Sparse Binary Connections

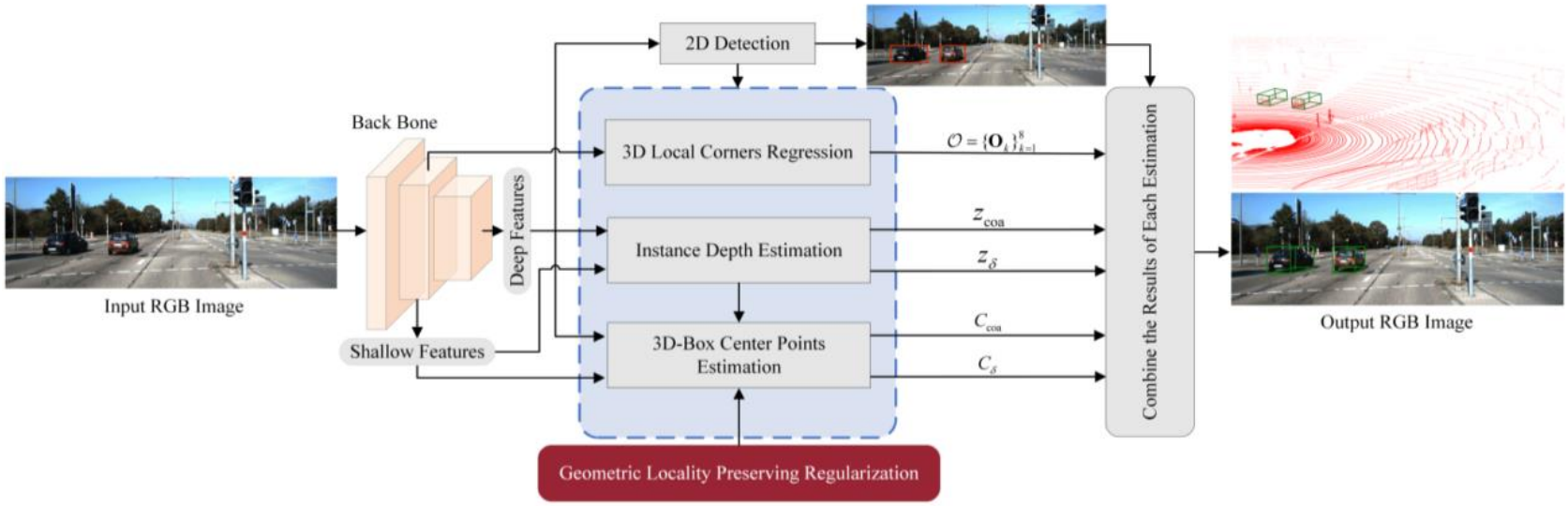


利用所提出的DAN对高光谱数据进行空间光谱分类

[Zhou X , Li S , Qin K , et al. Deep Adaptive Network: An Efficient Deep Neural Network with Sparse Binary Connections[J]. 2016.]

单目机器视觉

MoNet3D: Towards Accurate Monocular 3D Object Localization in Real Time

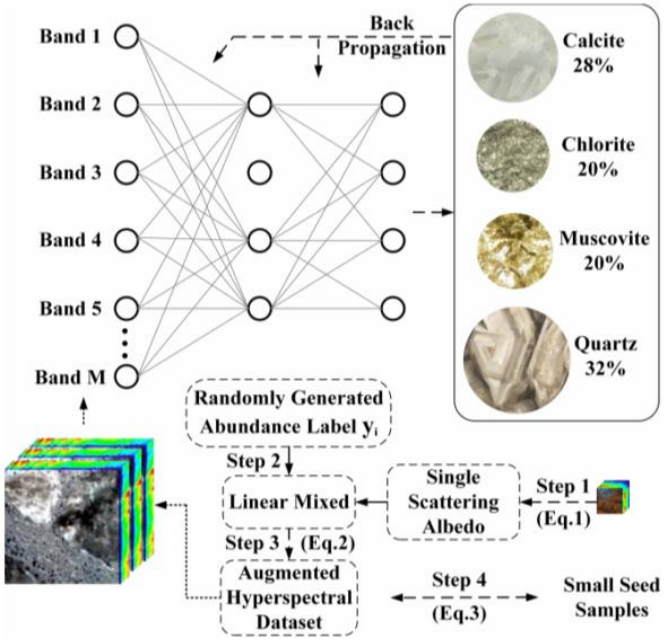


MoNet3D从单目RGB图像中提取特征，用于三维目标定位

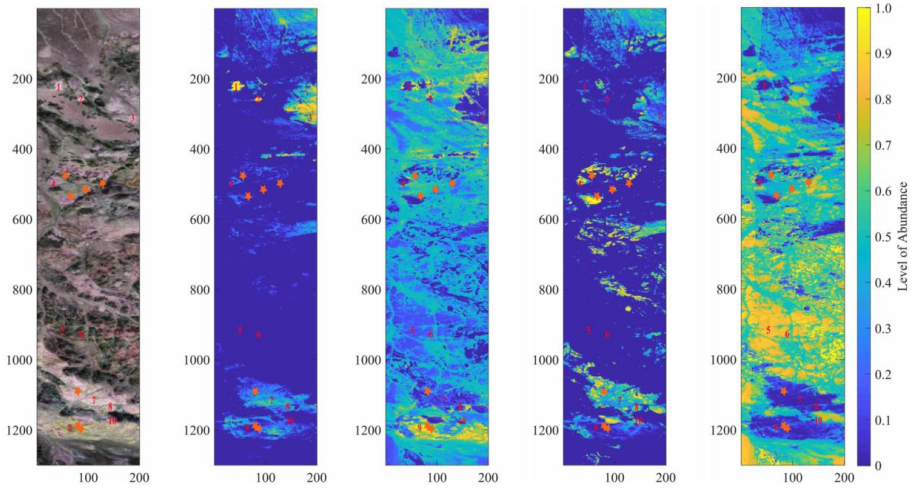
[Zhou X , Li S , Qin K , et al. MoNet3D: Towards Accurate Monocular 3D Object Localization in Real Time[C].ICML. 2020.]

小样本学习

Hapke Data Augmentation for Deep Learning-Based Hyperspectral Data Analysis With Limited Samples



用Hapke数据增强方法训练SDNN的例子



利用基于航空高光谱数据的Hapke增强深度学习方法预测矿物丰度图

[K. Qin et al., "Hapke Data Augmentation for Deep Learning-Based Hyperspectral Data Analysis With Limited Samples," in IEEE Geoscience and Remote Sensing Letters, doi: 10.1109/LGRS.2020.2989796.]

应用型论文

基本要求

- 具有一定编程能力，对已有的工作进行迁移学习，基于现有的开源代码进行training。
- 参考的论文或成果最好是2018年之后的最新成果（ICML、AAAI、CVPR、ICLR等会议上找）

应用型论文(可选题目参考)

- **红外目标检测**

[1] Choi S , Lee S , Kim Y , et al. Hi-CMD: Hierarchical Cross-Modality Disentanglement for Visible-Infrared Person Re-Identification[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.

（Github Code： https://github.com/bismex/HiCMD）

- **遥感图像语义分割**

[1] Chen L C , Papandreou G , Schroff F , et al. Rethinking Atrous Convolution for Semantic Image Segmentation[J]. 2017.

（Github Code：https://github.com/lcylmhlcy/Semantic-segmentation）

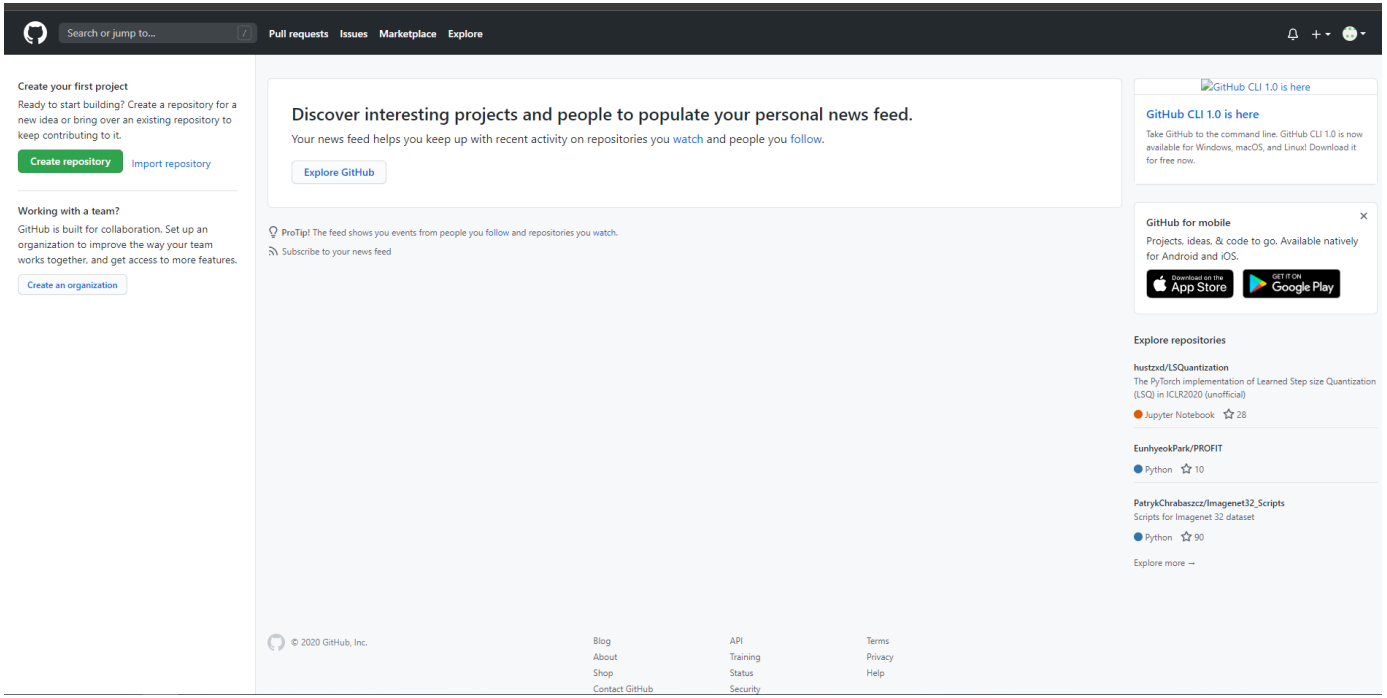
- **高光谱图像分类**

[1] Roy S K , Krishna G , Dubey S R , et al. HybridSN: Exploring 3D-2D CNN Feature Hierarchy for Hyperspectral Image Classification[J]. IEEE Geoence and Remote Sensing Letters, 2020, 17(2):277-281.

（Github Code：https://github.com/gokriznastic/HybridSN）

Github简介

$GitHub$[1]是最大的开源代码托管平台，旨在促进在一个共同项目上工作的个人之间的代码托管、版本控制和协作。通过该平台，无论何时何地，都可以对项目进行操作（托管和审查代码，管理项目和与世界各地的其他开发者共同开发软件。



[1] https://github.com/