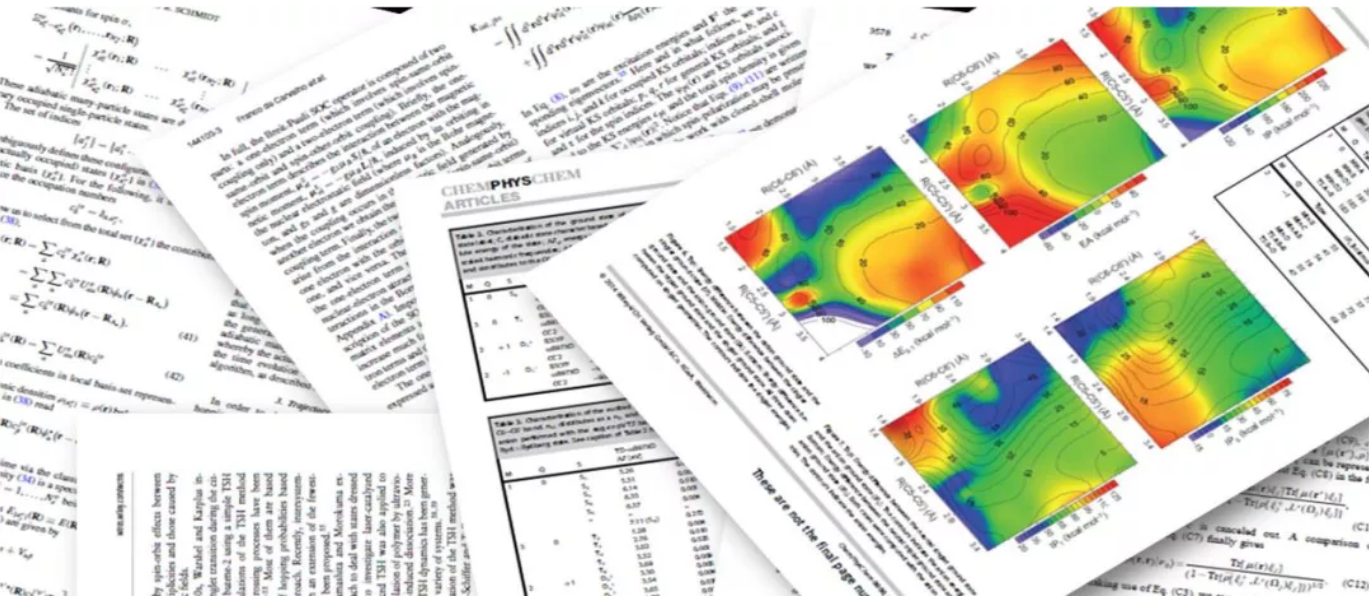


讀論文七步走！CV老司機萬字長文：一篇論文需要讀4遍

視學算法 2021-12-27 16:53



視學算法報導

編輯：LRS

【新智元導讀】讀論文對於AI新手和工程師來說可能是一件比較難的事。最近一位從業超5年的CV老司機發布了一篇萬字長文，講述了讀論文七步法，從找論文到總結，每篇論文由淺入深都需要讀4遍。

人工智能從業者需要讀論文嗎？



機器學習的各種模型發展日新月異，想要走在技術的最前沿，不看論文，光看博客、技術文檔是遠遠不夠的。

但隨著AI研究從業者逐漸增多，各類科普、教程做的實在是太好了，一些本科生甚至初、高中生都能利用工具來訓練自己的模型到現實中的應用。

但對於那些缺乏學術背景、沒有系統學習過如何寫論文的人來說，讀一篇學術論文可能比訓練一個模型要難太多了。



並且如果在訓練模型過程中有了一些創新的想法，想要發表一篇論文，但讀論文都讀不好的話，那寫論文肯定也會出問題。

最近Nvidia developer上就有一篇博客用萬字長文，七步教會你如何「務實地」讀論文。



博客作者Richmond Alake是一名機器學習和計算機視覺工程師，主要工作就是幫助各種初創公司開發深度學習模型來解決商業應用中的計算機視覺任務，技術領域的從業時間超過5年。

除此之外，他還是一個博客專家，寫過100多篇關於人工智能和機器學習主題的文章，總瀏覽量超過一百萬。



天才第一步：選個好題目

機器學習和數據科學領域有大量可供研究的主題領域，但這並不一定意味著每個研究課題都是好的選擇。

對於初入AI行業的人來說，可能更關注於機器學習的各種應用場景，但如果考慮到長期的職業前景的話，那些專業術語又讓AI新手感到迷茫。

所以確定一個研究主題是很困難的，對每個人來說，好題目的定義都是不一樣的，不過有一點是相同的，要選擇一個感興趣的或者已經有經驗的機器學習領域。

對於博客作者來說，深度學習就是他的興趣。

Richmond表示他當時「誤入」計算機視覺的原因就是因為對姿態估計、動作分類、手勢識別這些課題特別感興趣，開發深度學習模型解決CV問題也會讓他覺得很有成就感。

不同領域的工程師也需要不同的專業技能。

如果你的目標是自然語言處理工程師，那你就需要研究Transformer，RNN，語言模型，情感分類等等。

對於數據科學家來說，需要更關注數據增強，數據可視化，模式識別的技術。

Computer Vision Engineer	Convolutional Neural Network, Real-Time Object Detection, Generative Adversarial Networks, Pose Estimation
Natural Language Processing Engineer	Transformers, Recurrent Neural Networks, LSTM, Language Modelling, Computational Linguistics, Sentiment analysis, Text Classification, Word Embedding
Data Engineer	Data Pipelines, Data Modelling, Security and Privacy
Data Scientist	Data Pipelines Optimization, Data Augmentation, Data Visualisation approaches, Pattern Recognition
Deep Learning Engineer	Single Shot Detection, Transformers, Feature Extraction
Robotics Engineer	Inverse Kinematics, Automation, Map Planning, Reinforcement Learning
AI Ethicist	Ethical AI, AI Explainability

在這篇博客中，作者選擇他比較熟悉的姿態估計（Pose Estimation）為例來介紹。

第二步：找論文

讀論文，最重要的就是找到主題相關的、優質的論文。

PapersWithCode是一個比較常用的工具，能夠幫助研究人員快速找到機器學習相關的論文，並且還會把論文相關的數據集、代碼和其他資源都一起呈現出來。

Pose Estimation

705 papers with code • 19 benchmarks • 77 datasets

Pose Estimation is a general problem in Computer Vision where we detect the position and orientation of an object.

(Image credit: [Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose](#))

Benchmarks

Add a Result

These leaderboards are used to track progress in Pose Estimation

Trend	Dataset	Best Model	Paper	Code	Compare
	MPII Human Pose	🏆 Soft-gated Skip Connections			See all
	COCO test-dev	🏆 UDP-Pose-PSA(384x288)			See all
	Leeds Sports Poses	🏆 OmniPose			See all
	OCHuman	🏆 MIPNet (HRNet-W48)			See all
	J-HMDB	🏆 LSTM PM			See all

在網站中搜索Pose Estimation後，就會把相關的數據集、性能最佳的模型等列出來。

例如當你鎖定一篇論文後，就可以點到詳情頁，看到相關信息。

DensePose: Dense Human Pose Estimation In The Wild

CVPR 2018 • Riza Alp Güler, Natalia Neverova, Iasonas Kokkinos • [Edit social preview](#)

In this work, we establish dense correspondences between RGB image and a surface-based representation of the human body, a task we refer to as dense human pose estimation. We first gather dense correspondences for 50K persons appearing in the COCO dataset by introducing an efficient annotation pipeline... [read more](#)

PDF

Abstract

CVPR 2018 PDF

CVPR 2018 Abstract

Code

facebookresearch/detectron

★ 24,868

PyTorch

facebookresearch/detectron2

★ 19,250

PyTorch

facebookresearch/DensePose

★ 6,168

Caffe2

UBC-Computer-Vision-Group/DwNet

★ 20

PyTorch

ubc-vision/DwNet

★ 20

PyTorch

[See all 34 implementations](#)

Tasks

Pose Estimation

第三步：讀論文

確定完主題、找到合適的論文後，我們終於要開始讀了。

每一篇論文的格式實際上都是「八股文」，採用固定的格式進行編寫，這種格式既能幫助論文作者梳理思路，也能幫助讀者快速定位文章中的創新點。

對於研究新手來說，拿到一篇論文的第一個想法就是從頭到尾地把論文讀一遍，然後開始做筆記，這種想法雖然無可厚非，但肯定不是最高效讀論文的一種方式。

更實用的一種閱讀方法是先理解論文研究的背景，「標題」、「摘要」和「結論」部分是快速理解論文內容的三個關鍵部分。

在第一遍讀完論文後需要達成三個目標：

- 1、確認這篇論文和自己的研究內容是相關的
- 2、讀完論文的內容、方法和結論後，能夠對論文的研究背景和意義有所了解
- 3、認識到文章的主要創新點，總結論文作者的目標、方法和主要貢獻

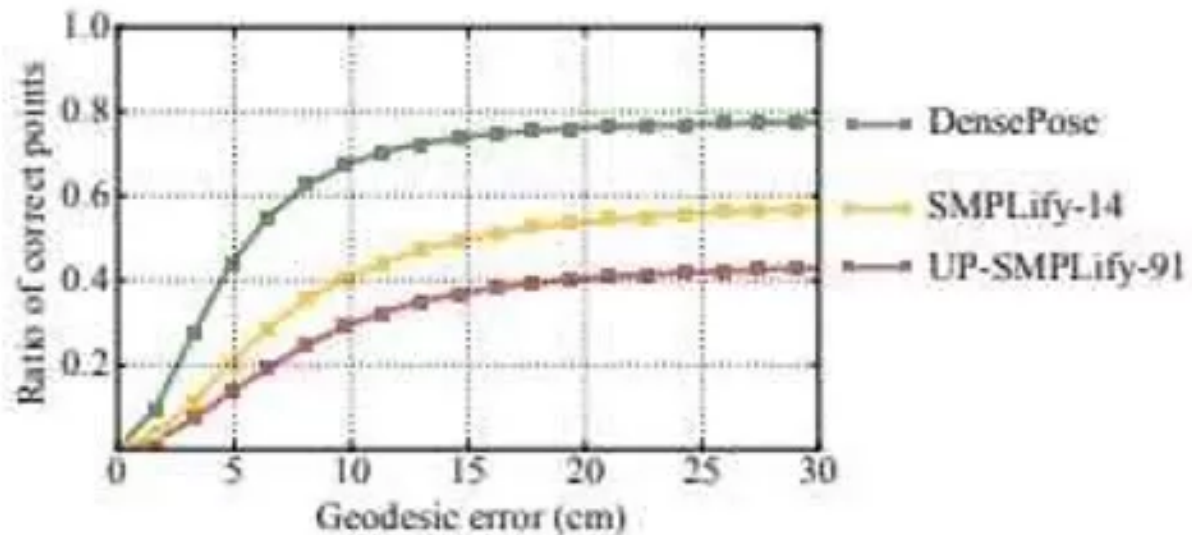
第四步：再讀一遍論文

第二遍讀論文的主要目的就是熟悉論文的內容。和讀第一遍一樣，不要直接讀論文的核心內容，只有充分了解論文的背景、研究出發點才能更好地理解論文。

第二遍主要從「引言」和「圖表」入手。

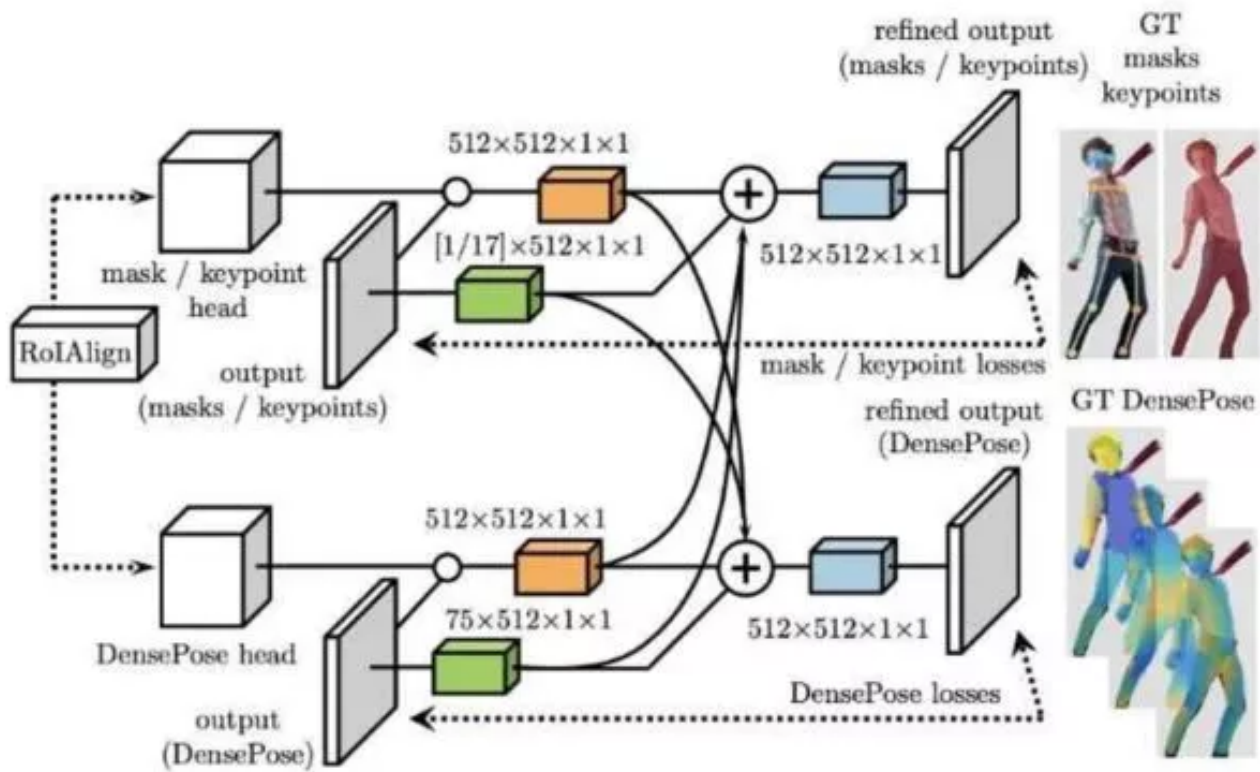
引言部分概述了研究工作的目標，會解釋論文中的問題域、研究範圍、先前的研究工作和方法。能夠幫助讀者找到該領域過去的相似研究工作，為讀者拓寬了探索思路，引言部分也為讀者提供了理解論文所需的必要知識。

圖標的話屬於論文中的說明性材料，能夠幫助讀者理解論文中的研究問題，並對文中所提出的方法進行解釋。表格也能夠對論文中涉及到的相關方法進行量化分析。



<i>Method</i>	AUC₁₀	AUC₃₀
<i>Full-body images</i>		
UP-SMPLify-91	0.155	0.306
SMPLify-14	0.226	0.416
DensePose	0.429	0.630
<i>All images</i>		
SMPLify-14	0.099	0.19
DensePose	0.378	0.614
Human Performance	0.563	0.835

對於深度學習來說，模型架構圖也是必不可少的，有時候讀一篇論文只看架構圖就能理解了文章的主要工作。



第五步：再再讀一遍

第三遍要更加深入地讀論文，但文章中的公式、不熟悉的術語都可以跳過。

從摘要到結論重新看一遍，在讀完每個章節後都休息一下，將所有想法、見解、收穫和不熟悉的概念一起記下來。

並且讀論文時要有節奏，一個比較有效的學習方法是學50分鐘，然後休息15分鐘。如此學習兩次以後，可以休息30分鐘以上來干點別的事。

如果還不熟悉這種節奏的話，可以從學25分鐘休息5分鐘開始練習。

第六步：再再再讀一遍

這次是最後一遍讀論文了，但也是最難的、最考驗思維和學習能力的一遍。

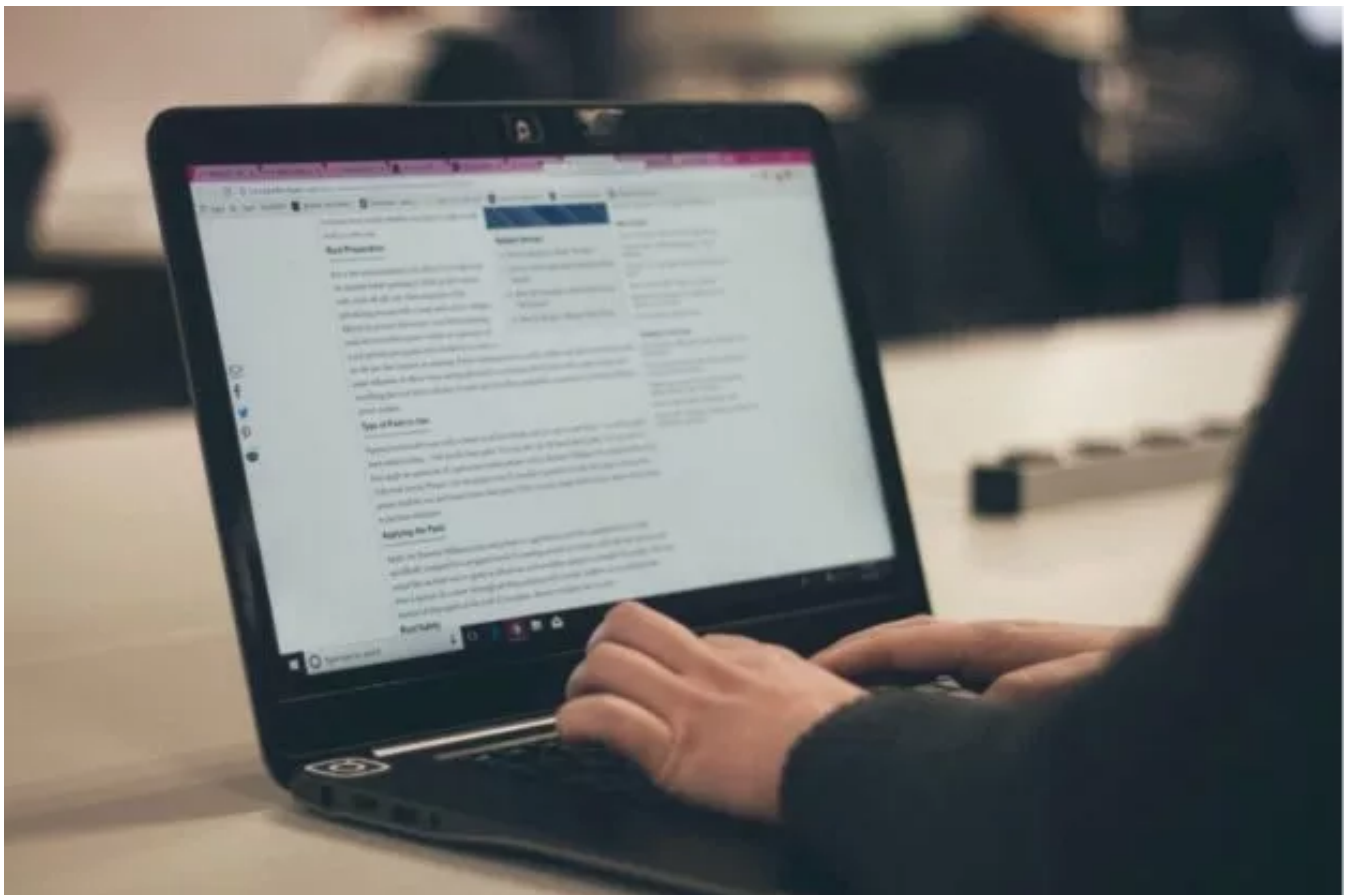
你需要將前面遇到的所有不理解的、不熟悉的術語、公式、名詞、概念、算法都進行深入研究。最重要的就是利用外部資源來幫助理解論文，例如相關文獻、搜索引擎、同學的討論等等。

最後一遍讀論文所需時間通常是不固定的，幾個小時、幾天甚至幾週都有可能。並且你在搜索引擎中可能也找不到論文的相關討論，這就需要你來思考、實踐或者聯繫作者來尋找答案。

第七步：總結

把論文中的內容用自己的話重新總結出來能夠幫助你對論文內容加深理解，確保學到的知識可以記得更長遠一點。

不管是手寫還是打字總結，都能夠幫助加強記憶。



毋庸置疑，AI從業人員必須要學會讀論文，而初學者往往很難入手。通過這樣一遍遍、有規則地讀論文，能幫你快速掌握、理解論文中的內容。

參考資料：

<https://developer.nvidia.com/blog/how-to-read-research-papers-a-pragmatic-approach-for-ml-practitioners/>



视学算法

一个AI的地方

地球不爆炸，算法不放假

點個 在看paper不斷！



喜歡此內容的人還喜歡

捨棄譜歸一化，這篇ICCV'21論文用梯度歸一化訓練GAN，效果極好
機器之心

