

论文	作者	会议	年份
Makeup-Go: Blind Reversion of Portrait Edit	Ying-Cong Chen ; Xiaoyong Shen ; Jiaya Jia (IEEE Fellow) 等	IEEE International Conference on Computer Vision (ICCV), 全球三大计算机视觉顶会之一	2017
Detail-Revealing Deep Video Super-Resolution	Xin Tao ; Hongyun Gao ; Renjie Liao ; Jue Wang ; Jiaya Jia (IEEE Fellow) 等	IEEE International Conference on Computer Vision (ICCV)	2017
Skeleton-Based Action Recognition with Spatial Reasoning and Temporal Stack Learning	Chenyang Si, Ya Jing, Wei Wang, Liang Wang, Tieniu Tan	The European Conference on Computer Vision (ECCV), 全球计算机视觉三大顶会之一	2018
A Deeply-initialized Coarse-to-fine Ensemble of Regression Trees for Face Alignment	Roberto Valle, Jose M. Buenaposada, Antonio Valdes, Luis Baumela	The European Conference on Computer Vision (ECCV)	2018
Defense Against Adversarial Images Using Web-Scale Nearest-Neighbor Search	Abhimanyu Dubey, Laurens van der Maaten, Zeki Yalniz, Yixuan Li, and Dhruv Mahajan	2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 全球计算机视觉三大顶会之一	2019
Extreme Relative Pose Estimation for RGB-D Scans via Scene Completion	Zhenpei Yang, Jeffrey Z. Pan, Linjie Luo, Xiaowei Zhou, Kristen Grauman, and Qixing Huang	2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)	2019
Self-Supervised Adaptation of High-Fidelity Face Models for Monocular Performance Tracking	Jae Shin Yoon, Takaaki Shiratori, Shoou-I Yu, and Hyun Soo Park	2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)	2019
Adaptively Connected Neural Networks	Guangrun Wang, Keze Wang, Liang Lin	2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)	2019
Weakly-Supervised Discovery of Geometry-Aware Representation for 3D Human Pose Estimation	Xipeng Chen, Kwan-Yee Lin, Wentao Liu, Chen Qian, Liang Lin	2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)	2019
Spatial-Aware Graph Relation Network for Large-Scale Object Detection	Hang Xu, Chenhan Jiang, Xiaodan Liang, Zhenguo Li	2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)	2019

1. Makeup-Go: Blind Reversion of Portrait Edit
 2. Detail-Revealing Deep Video Super-Resolution
 3. Defense Against Adversarial Images Using Web-Scale Nearest-Neighbor Search
 4. Extreme Relative Pose Estimation for RGB-D Scans via Scene Completion
 5. Self-Supervised Adaptation of High-Fidelity Face Models for Monocular Performance Tracking
 6. Adaptively Connected Neural Networks
 7. Weakly-Supervised Discovery of Geometry-Aware Representation for 3D Human Pose Estimation
 8. Spatial-Aware Graph Relation Network for Large-Scale Object Detection
 9. Skeleton-Based Action Recognition with Spatial Reasoning and Temporal Stack Learning
 10. A Deeply-initialized Coarse-to-fine Ensemble of Regression Trees for Face Alignment
-

1. Makeup-Go: Blind Reversion of Portrait Edit

论文: Makeup-Go: Blind Reversion of Portrait Edit

作者: Ying-Cong Chen ; Xiaoyong Shen ; **Jiaya Jia (IEEE Fellow)** 等

会议: 2017 IEEE International Conference on Computer Vision (ICCV), 全球三大计算机视觉顶级会议之一

体会:

腾讯优图杰出科学家贾佳亚博士作为国际图像处理的顶尖专家，本次第一个检索的就是这位新生代IEEE Fellow署名参与的文章。

这篇论文发表于2017年的IEEE计算机视觉国际会议（ICCV）中，主要介绍了一种基于CRN组件回归网络处理面部细微细节和颜色变化的方法，论文很有启发性。论文中所介绍的方法在社交领域的图片“卸妆”前后效果明显，论文中提到了已经被修饰美化过的照片（Edited Image）可以通过一个CNN网络处理得到一个Output图片，通过论文的Figure 1可以看到，该方法的输出图片与原始未经处理的图片比较，人脸的细节恢复度得到了一个不错的效果。深挖作者团队所发现的一篇独家报道也提出，这个网络并不是对所有案例都是有效，不过提出的component regression network，不仅在图像修复领域，在很多领域有非常强的应用。

这种技术被论文团队主要用在了社交领域的“一键卸妆”上，但进一步反思过来，论文所提出的CRN网络若用在模糊照片的修复上，也许可能会有意外的效果，对于特殊场景下，如低分辨率的监控画面的人脸表情的分析，可能也是一个不错的方向。

2. Detail-Revealing Deep Video Super-Resolution

论文: Detail-Revealing Deep Video Super-Resolution

作者: Xin Tao ; Hongyun Gao ; Renjie Liao ; Jue Wang ; **Jiaya Jia (IEEE Fellow)** 等

会议： 2017 IEEE International Conference on Computer Vision (ICCV), 全球三大计算机视觉顶级会议之一

体会：

这篇文章也是发表于2017年IEEE的ICCV文章，同样也产自于腾讯优图实验室，业界前列的实验室，一个神奇的地方。

这篇论文与上篇论文不同，研究的是一种超分辨率下的视频方法。做视频超分辨率的 **motivation**，传统而基础的问题，但应用场景非常多，比如在监控级别的场景下对车牌号进行分析识别。

论文中提到了该领域的两个主要难点，一个是怎样去得到一个effective的网络，二是模型问题。研究团队提出的这个 **CNN-based** 网络框架，如论文中 **Figure 2** 所描述的方法，设置了 **Fully convolutional**，对 **SPMC**层不做任何参数，此外还有 **Conv LSTM**。对于这个网络，这篇长达9页的论文也提到，实验中输入三个相同帧，效果不是很好，但输入三个连续帧时，可以得到比较好的效果（如论文 **Figure 5** 所示，以对运动的汽车识别其车牌号举例，处理还原后的清晰度很高）。

研究团队在一场专访报告中也指出，落地良好并且在工业级别也取得了令人兴奋的效果，**提出的 SPMC layer 在效果和速度方面也优于 BayesSR、MFSR、DESR、VSRNet 这些前辈**，可见论文所提出的方法适应性很好，如果在视频流的人脸分析的去模糊处理中用到此模型，不知适应性会怎样。

3. Defense Against Adversarial Images Using Web-Scale Nearest-Neighbor Search

论文： Defense Against Adversarial Images Using Web-Scale Nearest-Neighbor Search

作者： Abhimanyu Dubey, Laurens van der Maaten, Zeki Yalniz, Yixuan Li, and Dhruv Mahajan

会议： 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 全球计算机视觉三大顶会之一

体会： 从这篇文章，我开始尝试阅读本领域内最新的顶级文章，不再局限于论文的知名度和关注度，看一下最新的技术都在撸啥。

这篇文章我找到了刚刚发表不久，于今年6月16~20日 在美国洛杉矶举办的全球计算机视觉的三大顶会之一的 **CVPR 2019** (IEEE Conference on Computer Vision and Pattern Recognition)，计算机领域顶会第一次进入 **Top20** 的行列的会议。论文刚出炉，正热乎，拿来文献查新。

论文的研究团队来自Facebook AI研究院，工业界的实力干将之一，今年产出论文中共有 37 篇论文被CVPR收录，其中包括 15 篇 Oral 论文。

Defense Against Adversarial Images Using Web-Scale Nearest-Neighbor Search，中文译名为《使用网络规模近邻搜索的对抗性图像的防御》，论文太热乎，目前网路上尚没有此篇论文的笔记，读起来很吃劲。该论文的简述中表明，卷积网络对 **adversarial images** 不具有鲁棒性，这些图像随数据分布而扰动，使损失达到最大化。论文演示了一种 **最近邻的对抗性图片的防御手段** (**nearestneighbor defenses**)。研究团队在ImageNet这个庞大且开源的数据集上做了测试，在论文的第3部分，介绍了一个攻击模型 (**Attack Model**) 与攻击方法 (**Adversarial Attack Methods**)，并在第4部分，介绍了通过最近邻进行对抗防御的原理 (**Adversarial Defenses via Nearest Neighbors**)。第5部分，论文的研究团队在 **ImageNet** 上进行了利用实验模型进行精度防御的结果，也就是 **Table 1** 中的实验数据值。**Figure 5** 中对灰盒和黑盒中的数据量由 10^6 到 10^9 这几个数量级的增加中，利用 **CBW-D** 对 **PGD** 对抗图像的分类在 **ImageNet** 数据集上精度的变化。

4. Extreme Relative Pose Estimation for RGB-D Scans via Scene Completion

论文: Extreme Relative Pose Estimation for RGB-D Scans via Scene Completion

作者: Zhenpei Yang, Jeffrey Z. Pan, Linjie Luo, Xiaowei Zhou, Kristen Grauman, and Qixing Huang

会议: 2019 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 全球计算机视觉三大顶会之一

体会:

Extreme Relative Pose Estimation for RGB-D Scans via Scene Completion, 中文译名为《基于场景补全的RGB-D扫描的极端相对姿态估计》。来自世界不同地方的几个该领域的大拿共著了这篇论文, 同样发表在刚刚结束不久的2019CVPR上。这篇论文主要介绍的是研究团队所引入的新的研究方法对极端相对姿态估计与场景补全。

研究团队在 SUNCG, Matterport, SanNet 这三个数据集上用不同的图片角度做了大量实验(见论文中的 Table 1), 关键思想是在3D场景下进行RGB-D扫描补全以得到一个基础的几何形状, 然后计算相对位姿。实验结果显示在**极端不重叠之间的相对位姿之间所采取的处理方法**取得了encouraging的效果。

5. Self-Supervised Adaptation of High-Fidelity Face Models for Monocular Performance Tracking

论文: Self-Supervised Adaptation of High-Fidelity Face Models for Monocular Performance Tracking

作者: Jae Shin Yoon, Takaaki Shiratori, Shoou-I Yu, and Hyun Soo Park

会议: 2019 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 全球计算机视觉三大顶会之一

体会:

Self-Supervised Adaptation of High-Fidelity Face Models for Monocular Performance Tracking, 中文译名为《用于单目性能跟踪的高保真人脸模型的自监督适应》, 来自Facebook这个顶级研究团队, 发表于IEEE的2019 CVPR。

论文的研究队伍提出了一种**基于「连续帧纹理一致性」进行自监督域适应的方法**, 能够将普通相机拍出的2D图片转化为一个3D格式(效果见论文的 Figure 3)。

论文研究团队首先通过训练一个可以直接从单个二维图像驱动人脸模型的新网络来规避对特殊输入数据的需求(drive a face model just from a single 2D image); 然后, 在假设人脸在连续帧上的外观是一致的前提下, 基于「连续帧纹理一致性」进行自监督域适应。实验表明, 在不需要任何来自新领域的标记数据的情况下, 能够让手机摄像头中的高保真人脸模型执行复杂的面部运动(drive a high-fidelity face model to perform complex facial motion from a cellphone camera without requiring any labeled data from the new domain)。

6. Adaptively Connected Neural Networks

论文: Adaptively Connected Neural Networks

作者: Guangrun Wang, Keze Wang, Liang Lin

会议: 2019 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 全球计算机视觉三大顶会之一

介绍:

中文译名为《自适应连接神经网络》。论文研究人员引入一种新的自适应连接神经网络 (ACNet), 从两方面改进了传统的卷积神经网络 (CNN)。一是 ACNet 可以自适应地决定神经元连接属于全局连接抑或局部连接, 从而进行自适应局部推断或全局推断。论文实验可以证明, **现有的卷积神经网络(CNN)、经典的多层感知器 (MLP) 和最近提出的非局域网络 (NLN) 都是 ACNet 的特例。**二是 ACNet 不仅可以适用于传统的欧氏数据 (例如图像、音频等), 也可以适用于非欧氏数据 (graph data)。实验表明, ACNet 在 ImageNet-1K/CIFAR 图像分类、COCO 2017 目标检测和分割、CUHK03 行人重识别以及 CORA 文档分类等任务中达到了 State-of-the-art 效果。

具体来说, 研究人员首先使用自变换操作 (Self Trans 模块) 提取像素级特征、卷积操作 (CNN 模块) 提取局部特征、多层感知器操作 (MLP 模块) 提取全局特征, 然后使用自适应连接神经网络(ACNET 模块)融合三者, 得到局部与全局自适应的特征, 这样 ACNet 既有自变换操作和卷积操作所具有的局部推断能力, 又具有多层感知器操作所具有的全局推断能力。

7. Weakly-Supervised Discovery of Geometry-Aware Representation for 3D Human Pose Estimation

论文: Weakly-Supervised Discovery of Geometry-Aware Representation for 3D Human Pose Estimation

作者: Xipeng Chen, Kwan-Yee Lin, Wentao Liu, Chen Qian, Liang Lin

会议: 2019 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 全球计算机视觉三大顶会之一

介绍:

中文译名为《基于结构表征的弱监督3D人体姿态估计》。论文的研究团队提出一种解决方法, 拟从大量多视角图像中提取额外的 3D 人体结构信息, 使用额外信息辅助单张图像的 3D 人体姿态估计任务。在提取额外信息的过程中, 只使用带有 2D 标注的多视角图像作为训练集, 选取编解码器作为主干网络, 训练编解码器实现不同视角下 2D 人体信息的相互转换。为了让转换仅仅基于人体结构, 选取 2D 人体骨架作为本文方法的 2D 人体信息, 而没有使用原始图像。进一步加入了对 3D 结构的一致性约束, 使得抽取到的额外信息的 3D 结构更加稳定。因为抽取的额外信息蕴含了人体的 3D 结构信息, 所以将它映射到 3D 关键点坐标将会比直接利用 2D 图像或者 2D 坐标更为容易。继而验证了**仅仅使用简单的两层线性全连接层, 可以从额外信息中解码出相对合理的 3D 人体姿态。**

经过实验验证, 本文提取的额外信息可以作为对 3D 人体姿态信息的补充, 简单灵活的融合到现有的 3D 人体姿态估计方法中, 得到更加准确的预测结果。在标准的大型 3D 人体数据库 Human3.6M 上, 本文提取的额外信息对三种不同的 3D 人体姿态估计方法都有较大提升。对于现有最好的开源 3D 人体姿态估计方法, 在标准的数据划分下使用评估指标 MPJPE, 本文提出的方法仍然有 7% 的提升, **在现有的方法中达到最好的效果。**

8. Spatial-Aware Graph Relation Network for Large-Scale Object Detection

论文: Spatial-Aware Graph Relation Network for Large-Scale Object Detection

作者: Hang Xu, Chenhan Jiang, Xiaodan Liang, Zhenguo Li

会议: 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 全球计算机视觉三大顶会之一

介绍:

中文译名为《空间感知的图关系网络及在大规模目标检测的应用》。本论文的研究团队的工作目标是设计一个基于图卷积神经网络的检测框架，它可以同时**利用语义和空间关系，直接从训练集中有效地学习到可解释的稀疏图结构，并根据学到的图结构进行推理和特征传播，增强小目标、罕见类和模糊遮挡目标的特征相应提高检测结果。**

论文提出的 SGRN 框架由两个模块组成：一个稀疏关系图学习模块 (Relation Learner) 和一个空间感知图推理模块 (Spatial Graph Reasoning)。关系图学习模块首先从视觉特征中学习一个稀疏邻接矩阵，它保持了最相关的 T 个连接关系。然后，收集前一个分类器的权重，并将其映射到每个目标上，从而成为每个目标的视觉向量。目标之间的相对空间信息 (距离、角度) 被用来学习高斯核参数，以确定图形卷积的模式。在空间感知图推理模块中，根据稀疏邻接矩阵和高斯核对不同区域的视觉嵌入进行演化和传播。空间图推理模块的输出与原始区域特征相连接，以改进分类和定位。

9. Skeleton-Based Action Recognition with Spatial Reasoning and Temporal Stack Learning

论文: Skeleton-Based Action Recognition with Spatial Reasoning and Temporal Stack Learning

作者: Chenyang Si, Ya Jing, Wei Wang, Liang Wang, Tieniu Tan

会议: 2018 The European Conference on Computer Vision (ECCV), 全球计算机视觉三大顶会之一

介绍:

中文译名为《行为识别论文笔记之多纤维网络》。本论文的研究团队文章的核心 motivation 就是认为目前的 sota 的 3D 网络 (比如 I3D 以及 R(2+1)D-34 网络) 的计算量 FLOPs 都太高了。常用的 2D 卷积网络如 resnet-152 或是 vgg-16 网络大概是 10+ 的 GFLOPs, 而刚刚提到的两种 3D 卷积网络则达到了 100+ GFLOPs。作者认为, 当计算量相近的时候, 由于 3D 网络模型能额外的学习到时空信息, clip-based 的模型 (即指 3D 网络) 应该要比 frame-based 的模型 (即指 2D 网络) 有更好的效果。所以, 这篇文章的目标就是在**保持现有 sota 的 3D 模型的效果的同时, 大大提高其网络效率。**

文章主要是进一步优化了 Multi-Path 模块的结构, 并将其用于了 3D 卷积网络, 从而大大提高 3D 卷积网络的效率。在效率大大提高后, 其实也更有利于我们继续将网络做的更复杂更有效, 像之前的 I3D 的效率实在太差了, 很难进一步增加复杂度。一方面通过引入网络加速技巧对模型速度进行优化, 一方面通过增加网络对时序建模的能力来对模型效果进行提高, 应该是未来 3D 网络研究更平衡的一种发展道路吧。

10. A Deeply-initialized Coarse-to-fine Ensemble of Regression Trees for Face Alignment

论文: A Deeply-initialized Coarse-to-fine Ensemble of Regression Trees for Face Alignment

作者: Roberto Valle, Jose M. Buenaposada, Antonio Valdes, Luis Baumela

会议: 2018 The European Conference on Computer Vision (ECCV), 全球计算机视觉三大顶会之一

介绍:

本论文的作者认为, 出现人脸特征点距离真实位置偏移过大, 是因为算法初始化时的特征点不够鲁棒, 于是提出一种**使用深度卷积网络粗略估计特征点位置, 结合3D人脸姿态估计与重投影确定特征点初始位置, 然后使用经典的回归树集成 (Ensemble of Regression Trees, ERT) 方法提精位置。**

论文的创新性在于, 作者结合深度学习方法与传统方法, 将深度学习方法得到的结果用于传统方法的特征点初始化, 作者认为深度学习方法得到的特征点位置更加鲁棒。速度上, 该文在NVidia GeForce GTX 1080 (8GB) GPU 与 Intel Xeon E5-1650 3.50GHz (6 cores/12 threads, 32 GB of RAM)机器上可以达到实时 (32fps) ,比 清华&商汤 开源 CVPR2018 超高精度人脸对齐算法 LAB 要快, 但 LAB 比该文的精度要高。

@IDforYIT

本文件内容由**Markdown (.md)** 编写, 导出为本**PDF**。