Zhang Bin

Position Wanted: Computer Vision Researcher/Engineer

CNV A5406, No. 9 Mozhou Road, Jiangning District, Nanjing

z-bingo@outlook.com | 13912981928 | zb824524275

z-bingo | 知 Columns about Image/Video Enhancement | Home Page (More Details)

Education

Southeast University Sep. 2018 - Present

M.S. in Electronic and Communication Engineering, recommended postgraduate. Research interest: multivariate signal processing, image processing and computer vision.

Hohai University Sep. 2014 - June 2018

B.S. in Communication Engineering, GPA: 4.82/5.00, chair of student science and technology association.

Internship

Tencent Youtu Lab. **Computer Vision Intern** Nov. 2019 - Present

- Do the research about network architecture search (NAS) especially its applications on object detection.
- The proposed face detector ASFD ranked No.1 on WIDER Face and FDDB, two popular datasets for face detection.
- The proposed cartoon face detector ACFD ranked No.1 at IJCAI-iCartoon Face Challenge (Detection Track).

ArcSoft Inc. **Computer Vision Intern** Mar. 2019 - Sep. 2019

- Study the commonly used noise models of camera, managed to generate sRGB training dataset as similar as possible to the real-world data, and develop a software to yield noisy images for image denoising and super resolution tasks.
- Investigate and study the learning based image denoising algorithms especially for burst of images and videos, and improve them for the better generalization to real-world noisy images.
- Propose an attention mechanisam enhanced kernel prediction network with superior performance than KPN. (accepted by ICASSP 2020)

GoVion Co. Ltd. **Image Processing Intern** June 2018 - Jan. 2019

- Based on the traditional image processing algorithms, designed and developed the core algorithms for detecting defects of LCD and OLED panels, such as MURA, bubble, color edges, and so on.
- Some algorithms have been applied on the products.

Projects

End-to-End Searched Detctor - (*NAS & Detection*)

- Observe that utilizing the backbone of image classification into detection tasks is sub-optimal, especially that they focus on single-scale output rather than multi-scale output, and channel number redundancy.
- Build a supernet with shrinkable convolution layers and channels, and train it in a one-shot way.
- Take inference time as the constraint directly and utilize generic algotithm to discover different detectors. As a result, COCO AP can **improve about 2 points** under the same inference time.

ACFD: Asymmetric Cartoon Face Detector - (*IJCAI-iCartoonFace Challenge, Detection Track*)

- A one-stage anchor-based face detector is proposed for detecting the cartoon faces, in which asymmetric convolution blocks are employed to enhance the ability of network to generate features with more diverse receptive fields handling hard faces furtherly and one-shot NAS is utilized to refine the dilation of backbone network; two-step match strategy is utilized to match enough high-quality anchors for each gt face; margin loss is used to enhance the power of discrimination of classification branch; TensorRT is adopted to accelerate the final model.
- **Ranked No.1** on the leaderboard.

ASFD: Automatic and Scalable Face Detector - (*NAS & Detection*)

- Formulate the gap between genetic object detection and face detection by applying NAS method of genetic object detection on face detectors; based on the differentiable NAS, a feature module, AutoFAE, is searched with the specific search space; the proposed ASFD achieves a better tradeoff between effectiveness and efficiency, ASFD-D0 **runs more than 320FPS** with VGA-resolution images on V100 GPU without any optimization, ASFD-D6 achieves the state-of-the-art performance.
- **Ranked No.1** on the popular WIDER Face and FDDB.

EMD and Its Applications on Image Processing - (*Undergraduate and Postgraduate Thesis*)

- Empirical mode decomposition (EMD) is a fully data-driven technique for non-linear and non-stationary signal.
- Proposed a novel bidimensional multivariate EMD (BMEMD), and found its application on multi-scale image fusion.
- Developed the fast version based on order statistic filters.

Algorithms of Pedestrian Detection and Action Recognition - (*JSCVC 2018*)

- Pedestrian detection: Faster-RCNN and YOLO, the ResNet50 backbone.
- Action recognition: Temporal Segment Networks with backbone of ResNet50, and 3D Convolutional Networks.

Skills

Programming Languages: Python, C/C++, Matlab

Libraries and Others: PyTorch, OpenCV, Ubuntu, git, docker, LaTeX, TensorFlow

Publications

- **B. Zhang***, J. Li*, Y. Wang, and *et. al.*, "ACFD: Asymmetric Cartoon Face Detector", *IJCAI 2020 Workshops*.
- **B. Zhang***, J. Li*, Y. Wang, and *et. al.*, "ASFD: Automatic and Scalable Face Detector", *arXiv*.
- **B. Zhang**, S. Jin, Y. Xia, and *et. al.*, "Attention Mechanism Enhanced Kernel Prediction Networks for Burst Image Denoising", *ICASSP*, 2020. (**CCF-B, oral**, *arXiv*)
- Y. Xia, **B. Zhang**, W. Pei, and D. P. Mandic, "Bidimensional Multivariate Empirical Mode Decomposition with Applications in Multi-Scale Image Fusion", *IEEE Access*, 2019. (**IF=4.098, Q1**, *Open Access*)

Awards

- Nov. 2016, **National Undergraduate Scholarship**.
- Nov. 2016, **Outstanding Science Volunteer**, Changzhou.
- Jan. 2016, **Excellent Student Cadre**, Hohai University.
- Nov. 2018, **Outstanding Winner**, Jiangsu Postgraduate Computer Vision Innovation & Practice Competition.
- Sep. 2017, **First Prise**, National Undergraduate Electronics Design Contest.
- Aug. 2016, **First Prize**, College Student Smart Internet Innovation Application Design Contest.

张　斌

意向职位: 计算机视觉算法研究员/工程师

南京市江宁区秣周东路 9 号中国无线谷 A5406

z-bingo@outlook.com | 13912981928 | zb824524275

z-bingo | 知 深度图像/视频增强专栏 | 个人主页（查看更多）

教育经历

东南大学 2018 年 9 月 - 至今

信息科学与工程学院，电子与通信工程在读硕士，推免研究生。研究方向为：多维信号处理、数字图像处理以及计算机视觉等。

河海大学 2014 年 9 月 - 2018 年 6 月

物联网工程学院，通信工程学士学位。在校期间，GPA：4.82/5.00，综合排名：2/118，学院学生科学与技术协会主席。

实习经历

腾讯优图实验室 计算机视觉见习研究员 2019 年 11 月 - 至今

- 从事目标检测、人脸检测相关算法的研究，尤其是神经网络架构搜索（NAS）在其中的应用。
- 提出的 ASFD 人脸检测器在 WIDER Face 和 FDDB 两个数据集上位居第一。
- 提出的 ACFD 卡通脸检测器在 2020 IJCAI-iCartoon Face Challenge (Detection Track) 中名列第一。

虹软（南京）多媒体技术股份有限公司 计算机视觉算法实习生 2019 年 3 月 - 2019 年 9 月

- 调研了常用的相机噪声模型，通过此类模型生成尽可能与真实含噪图像相似的训练集；并开发了相应的 demo，可为图像去噪、超分辨率等任务生成含噪声的训练集。
- 学习了基于 CNN 的图像去噪、去模糊以及增强等算法，尤其是基于深度学习的以对多帧图像或视频为输入的相关方法，致力于提高此类方法对真实场景高强度噪声的泛化能力。
- 提出了一种基于注意力机制的核预测多帧去噪网络，其性能优于 KPN 等方法，论文发表于 ICASSP 2020(oral)。

高视（南京）智能科技有限公司 图像处理算法实习生 2018 年 6 月 - 2019 年 9 月

- 基于传统图像处理算法，设计并完成了用于 LCD 和 OLED 屏幕的缺陷检测算法，如 MURA、bubble 以及 color edges 等缺陷。部分算法已应用于公司产品中。

项目经历

通用目标检测器端到端搜索 - (NAS&Detection)

- 用于分类任务的骨干网络对于检测任务是次优的，如：注重于单尺度输出而非多尺度输出、通道数冗余等。
- 构建了层数和通道数可变的检测器超网络，基于 OneShot-NAS，通过渐进的方式对其进行训练。
- 基于遗传算法，直接使用推理耗时作为约束，搜索得到不同耗时下的检测器。在相同耗时下，**COCO AP 可稳定提升约 2 个点**。

非对称卡通人脸检测器 (ACFD) - (IJCAI-卡通人脸检测挑战赛)

- 提出了一种基于 anchor 的一阶段检测器，结合非对称卷积模块，使得模型具有提取多样化感受野特征的能力，增强对于困难脸的检测能力；使用两步 anchor 匹配策略，为每个脸匹配足够多的高质量 anchor；采用 margin loss，增强分类分支对于困难样本的辨识能力；使用 TensorRT 对最终的模型进行加速。
- 检测赛道榜单**排名第一**。

自动搜索和协同缩放人脸检测器 (ASFD) - (NAS&Detection)

- 发现了通用目标检测和人脸检测在使用 NAS 搜索部分模块时的差异；针对人脸检测任务，对搜索空间和策略进行了特定的设计；基于可微分的 NAS 方法，搜索了特征模块 AutoFAE，用于融合并增强多尺度特征；提出的 ASFD 实现了很好的性能与速度之间的平衡，在没有任何工程优化的前提下，ASFD-D0 可在 V100 GPU 上使用 VGA 分辨率图像**运行超过 320FPS(3.1ms)**，ASFD-D6 可以实现最优的性能。
- 在 WIDER Face 和 FDDB 两大数据集上**排名第一**。

经验模式分解及其在图像处理中的应用 - (本科学位论文、硕士毕业论文)

- 经验模式分解（EMD）是一种全数据驱动的、无需预设参数、适用于非线性非平稳信号处理的时频分析算法。
- 提出了一种二维多元 EMD（BMEMD）算法，可用于多尺度图像融合及纹理分析；并且基于顺序统计滤波器提出了 BMEMD 算法的快速版本。

行人检测和视频动作识别算法设计 - (2018 年江苏省研究生计算机视觉大赛)

- 行人检测：考虑速度要求，选用单检测检测器 YOLOv1，并以 ResNet50 为骨干网络。
- 视频动作识别：选用 TSN，将 RGB 个 RGB-diff 作为两个 stream 的输入，通过 RGB-diff 引入运动特征。

★ 个人技能

编程语言：Python, C/C++, Matlab

其他：PyTorch, OpenCV, Ubuntu, git, docker, LaTeX, TensorFlow

科研成果

- **B. Zhang***, J. Li*, Y. Wang, and *et. al.*, "ACFD: Asymmetric Cartoon Face Detector", *IJCAI 2020 Workshops*.
- **B. Zhang***, J. Li*, Y. Wang, and *et. al.*, "ASFD: Automatic and Scalable Face Detector", *arXiv*.
- **B. Zhang**, S. Jin, Y. Xia, and *et. al.*, "Attention Mechanism Enhanced Kernel Prediction Networks for Burst Image Denoising", *ICASSP*, 2020. (**CCF-B, oral**, *arXiv*)
- Y. Xia, **B. Zhang**, W. Pei, and D. P. Mandic, "Bidimensional Multivariate Empirical Mode Decomposition with Applications in Multi-Scale Image Fusion", *IEEE Access*, 2019. (**IF=4.098, Q1**, *Open Access*)
- 一种用于多图像融合的二维多元经验模态分解算法.（专利）
- 一种快速二维多元经验模态分解算法.（专利）
- 面部检测、图像检测神经网络训练方法、装置和设备.（专利）

荣誉及奖项

- 2016 年 11 月，**国家奖学金**；
- 2016 年 11 月，**优秀科普志愿者**，常州市；
- 2016 年 1 月，**优秀学生干部**，河海大学；
- 2018 年 11 月，**特等奖**，江苏省研究生计算机视觉大赛；
- 2017 年 9 月，**一等奖**，大学生电子设计竞赛；
- 2016 年 8 月，**一等奖**，全国大学生智能互联创新大赛；