

# 简历

姓名：黄世华  
性别：男  
出生日期：1995 年 10 月 04 日  
最高学历：学士，物联网工程  
电话：13377439253（微信同号）  
邮箱：[shihuahuang95@gmail.com](mailto:shihuahuang95@gmail.com)  
个人网页：<http://www.shihuahuang.cn/>  
求职岗位：全职计算机视觉算法研究/工程



## 个人简介

黄世华于 2018 年在东北大学获得物联网工程学士学位；现在美国密歇根州立大学计算机与工程系攻读博士学位，师从 Kalyanmoy Deb 教授（ACM Fellow, IEEE Fellow, ASME Fellow, 印度科学院及工程院院士）和 Vishnu Boddeti 助理教授。黄世华的研究方向为基于神经网络的表征学习，近三年共发表 10 篇学术论文，其中多篇发表在计算机视觉顶级会议 IEEE ICCV 及 CVPR 和计算机顶级期刊 IEEE TCYB（影响因子：19.118）和 TNNLS（影响因子：14.255）；累计 Google 学术引用 224 次。此外，黄世华还多次参加科研和工业界组织的算法大赛，并获得了第五届“四维图新”杯自动驾驶赛道冠军、IJCAI2019 阿里巴巴 AI 对抗防御赛道和 NTIRE2020 视频超分赛道的测评第一名。

## 主要经历

**2021/09—至今**                      香港理工大学&密歇根州立大学（美国）                      博士研究生

—— 导师：Prof. Kaychen Tan, Prof. Kalyanmoy Deb 和 Prof. Vishnu Boddeti

—— 研究方向：多目标神经网络架构搜索、实时语义分割和鲁棒模型设计。

—— 成果产出：一作一篇 CVPR 和 TITS 文章（在投）、一个大规模鲁棒模型 benchmark。

**2018/07—2021/07**                      南方科技大学                      全职研究助理

—— 工作内容：1. 完成独立和交叉的科研任务，包括生成对抗网络、密集图像预测、神经网络架构搜索和 Vision Transformer；2. 协助实验室负责人指导本科生及部分研究生开展深度学习科研工作；3. 管理实验室科研服务器，包括但不限于 DGX2 和 80 张 A100s 平台。

—— 成果产出：一作一篇 ICCV 和 TAI 文章、其他多篇包括但不限于 TCYB 和 TNNLS 的高水平论文和多个挑战赛的冠军。

**2014/09—2018/06**                      东北大学                      全日制学士

—— 成果产出：一作一篇目标检测的 PRCV 和密集图像预测多任务的 arXiv 文章。

## 代表论文

(相关代码均已开源)

1. **Shihua Huang, Zhichao Lu, Kalyanmoy Deb, and Vishnu Boddeti.** Revisiting Residual Networks for Adversarial Robustness: An Architectural Perspective. *IEEE CVPR, 2023.*

**简介：**系统性地从**神经网络模块**（拓扑结构、卷积核大小、激活函数和归一化函数）到**结构**（深度和宽度）去研究神经网络架构对鲁棒性的影响。最终提出鲁棒残差模块和发现深并瘦的鲁棒缩放规律，基于此两点发现提出鲁棒残差网络，相较于 WRN，在**普遍 Autoattack 提升~3%情况下，节省了接近一半的参数量**，有望成为对抗训练的新网络架构。同时复现了 DeepMind 所提出的高阶对抗训练框架（从 JAX 到 Pytorch），在该框架

下我们模型击败了 DeepMind，取得了目前全球最好的表现。

2. **Shihua Huang**, Zhichao Lu, Ran Cheng, and Cheng He. FaPN: Feature-aligned Pyramid Network for Dense Image Prediction. *IEEE ICCV*, 2021.

**简介：**为解决在密集图像预测任务中多次下采样导致高层特征严重丢失细节信息，进而在上采样特征融合过程中产生语义特征错位的问题，我们提出了在上采样中引入特征对齐模块，以分辨率较大及细节信息丰富的特征作为位置参考，逐渐恢复高层特征的位置信息，该方法在四种密集图像预测任务上相较于强基准算法均取得了显著的提升。同时所提出的 FaPN 网络曾辅助 **Facebook** 研究人员取得了语义分割 ADE20k 测评的第一名，截止 2022 年，前五名有三个算法均使用了其 FaPN。

3. **Shihua Huang\***, Zhichao Lu\*, Ran Cheng, Kaychen Tan, Changxiao Qiu, and Fan Yang. Modularized and Automated Design of Feature Pyramid Networks for Real-time Semantic Segmentation. Submitted, *IEEE TITS*, 2022. (\*同等贡献)

**简介：**为实现高性能的实时语义分割，提出通过两步进程：1. 新型的分割顶部的模块化方法，根据其功能性将其分解成 5 个部分，并通过穷举每个部分已知的操作得到性能最佳的分割顶部组合方案；2. 在确定分割顶部情况下，使用双目标（耗时 vs 性能）进化搜索算法定制在不同耗时情况下的一组最佳网络架构。该算法在三个标准数据集上相较于当前手工设计和自动化搜索的模型均在性能和精度上获得了优越性。

4. **Shihua Huang**, Cheng He, and Ran Cheng. Multimodal Image-to-Image Translation via a Single Generative Adversarial Network. *IEEE TAI*, 2022.

**简介：**使用单个轻量型的生成对抗网络完成多域之间的多模态变换（例如输入一只猫图片，可得到不同长相的老虎或豹子，反之亦然），同时迁移的图片风格可控。该算法被上海航空研究者用于翼型和流场之间的可控转换，减少翼型设计中风洞实验次数。

5. Cheng He, **Shihua Huang**, Ran Cheng, Kay Chen Tan, and Yaochu Jin. Evolutionary Multi-objective Optimization Driven by Generative Adversarial Networks (GANs). *IEEE TCYB*, 2021.

**简介：**传统的进化计算是通过在父代上不断地交叉变异过程产生新的子代，考虑到生成对抗网络的强大生成能力、采用适当训练方法后的稳定性及生成样本的多样性，使用生成对抗网络产生进化过程中的子代；并利用生成对抗网络中的判别器对于样本的优异有一定的判别能力，提前筛选得到的子代，可减少昂贵的适应值评估次数。

## 项目

1. **Vision Transformer 探索**

**简介：**1. 利用 Transformer 中多头自注意力机制改进上采样：将要上采样的语义特征作为 Key 和 Value，然后未曾下采样的细节特征作为 Query，然后 Query 中的每一个位置都会通过动态的方式（attention map）去调整将上采样后的特征组合，跟 FaPN 中的 FAM 作用一致，但是更具全局性，同时性能更优；2. 探讨 Vision Transformer 架构中哪些模块是其成功的主要原因。保持住 Transformer 自有模块，研究 FFN（MLP）和自注意力机制的顺序，和使用不可学习的层（多尺度 Pooling，Local binary 还有高斯分布权重的卷积）取代自注意机制。发现 Transformer 架构本身相较于自注意力机制更有效。

2. **第五届“四维图新”杯自动驾驶挑战赛**

**简介：**基于 **Cascade R-CNN+ResNet101** 架构，针对目标检测加入了 **ROI Align** 来提升模型定位性能和可变形卷积来提高模型对物体的建模能力；并针对物体分割问题，将特征金字塔中的多尺度特征经过 SE 模块和双线性插值方式叠加融合，最后和由感受野更小的 Spatial Path 得到的特征再次融合；引入不同空洞率卷积组成的 **Atrous Spatial Pyramid 提取不同尺寸的物体特征**。为了惩罚模型容易错误分类的样本，从而提升模型的泛化性能，

我们是使用了难例挖掘交叉熵作为我们的损失函数。我们同时也引入了辅助损失函数用来增强模型中间层的特征提取能力，这样也能防止模型过深容易造成梯度消失问题。最后我们还采用了 GroupNorm 和 多尺度测试技巧来增强模型的鲁棒性能。

### 3. IJCAI2019 workshop —— 阿里巴巴人工智能对抗算法竞赛

**简介：**输入图像大小是 224\*224，并且单次线上测试时间不超过 25 分钟，我们采用较为轻量化的 ResNet50，同时使用 PGD 和 当时最有效的 TRADES loss 对其进行有效的对抗训练。在 ResNet50 的基础上我们引入了部分空洞和可变型卷积，使其有更强的全局特征提取能力；同时测试样本由五种攻击模型生成和缓解模型对抗训练难度，采用了分阶段逐步增大 PGD 的迭代次数和扰动范围；考虑到线上对抗算法生成的样本跟线下用于训练的样本在数据域上可能存在比较大的差别，我们通过一些巧妙设计的实验发现，在测试时将 BatchNorm 的模式由 eval 改成 train 以后（即测试时所用的均值和标准差不是由训练集统计而来，而是当前批数据统计而来），线上测试精度大幅度提升了 5%。

### 4. CVPR2020 workshop —— NTIRE2020 视频超分挑战赛

**简介：**采用了 EDVR 视频超分算法中使用前后多帧输入对齐的思路，提出了一个新颖的视频超分框架，主要由 Spatial Attention、Pre-Deblur 和 PAFU 模块。其中 Spatial Attention 是由一些 non-local 操作算子组成，主要是为了在多帧对齐之前对提取到全局特征提升对齐效果；Pre-Deblur 是 EDVR 提出来的模块；而 PAFU 是由一组串联（Progressive）的 AFUs（Alignment, Fusion 和 Update）模块组成，用于逐步完善特征对齐效果，其中 Alignment 和 Fusion 使用的分别是 EDVR 中提出来的 PCD 和 TSA 模块组成。最后为了提升模型的精度，使用了缩放和翻转输入的测试增强技术。

## · 竞赛 ·

**国际** IJCAI2019 workshop: IJCAI-19 Alibaba Adversarial AI Challenge on Defense, 1/2519;

CVPR2020 workshop: NTIRE 2020 Challenge on Video Deblurring, 1/7; MICCAI2021 workshop: GAMMA 2021, 6/566。

**国内** 第五届“四维图新”杯自动驾驶赛项冠军, 2019; 数智重庆·全球产业赋能创新大赛【赛场一】5/2990, 2020。

## · 其他 ·

**自我评价：**有超六年的深度学习算法研究经历，曾从事包括但不限于目标检测、语义分割、生成对抗网络、神经网络架构搜索、鲁棒网络设计等方向研究，可灵活快速完成各种深度学习项目需求。无障碍阅读英文文献，及时跟进工作方向的全球技术迭代，提升能力水平。

**综合素养：**自我驱动力强、对人工智能领域充满激情、与人和善、对工作环境适应能力强；既喜欢挑战超前的困难领域，也执着于深耕熟知的领域。

**专业技能：**熟练掌握 Python, Pytorch, Linux。

**期刊审稿：**IEEE TIP、TNNLS、TMM 和 TCDS; Elsevier Neural Network 和 Applied Soft Computing; Springer Complex & Intelligent System。