Jiangtian Pan

[jiangtianpan.github.io] . [panjiangtian@hotmail.com]

Education

The Ohio State University (08/2017-05/2019)

- M.S. in Electrical and Computer Engineering.
- Coursework in Computer Vision, Machine Learning and Image Processing.

Wuhan University (09/2013-06/2017)

- B.S. in Automation, GPA 85/100 [Top 15%]
- AAA Graduate Student

Scholarships and Awards

- 2018: AI Star Intern in the JD AI Research [10/50]
- 2018: Ranked 2nd in the CVPR 2018 Workshop, LIP Contest
- 2017: AAA Graduate Student [10/450]
- 2016: Interdisciplinary Contest in Modeling (Honorable Mention)
- 2015: WHU Outstanding Student Leader [2/60]
- 2014: Avery Dennison Foundation Spirit of Invention Scholarship [10/150]

Academic Publication

- Yunhan Huang, Li Ding, Yun Feng, Jiangtian Pan. "*Epidemic Spreading in Random Walkers with Heterogeneous Interaction Radius*", JSTAT 2016(10), 103501.

Research Experience

CV Researcher, Megvii (Face++) Research, *Advisor: Dr. Yuzhi Wang* (07/2021 - present)

> Real-Time Video Quality Enhancement for Mobile Devices:

- ***Proposed a CNN based model***: Incorporated an attention module, skip connections, residual modules, and mask modules to optimize video quality in real-time.
- ***Optimized Computational Efficiency***: From 330 GOPs to 12 GOPs by implementing ***Space2Depth*** modules, ***Distillation***, and ***Pruning***. Achieved inference time reduction from 100+ ms to 15 ms on Qualcomm Snapdragon 8475 platform.
- ***Noise Data Collection***: Proposed a unique approach to collect noise from real-world data, enhancing the training dataset for improved model robustness.
- ***Data Augmentation Techniques***: Innovated data augmentation strategies including ISP simulation, brightness adjustment, motion simulation, and random disturbance to improve model adaptability.

- ***Loss Function Design***: Formulated gradient loss, edge loss, and mask loss functions to ensure robust model performance under challenging conditions.
- ***Deployment***: Developed an efficient pipeline with pre-processing and post-processing stages running on GPU, and model inference on NPU. Integrated this algorithmic module into the XIAOMI 12 Pro smartphone.

> Real-World Blind Super-Resolution via Feature Matching:

- ***Constructed a VQGAN based model***: Incorporated Residual Modules, Swin Transformer Block, Codebook, and Feature Matching Module to blind enhance video resolution from 720p to 1440p.
- ***HR dataset build***: Build and augmented high-resolution (HR) dataset from DIV2K, Flickr2K, DIV8K and real-word datasets via texture filtering.
- ***Data Enhancement***: Used pretrained models FeMaSR Real-ESRGAN, BasicSR++ to further enhance high-frequency textures via image filtering and patch filtering.
- ***Loss Function Design***: Formulated L2 loss, perceptual loss, gan loss and codebook loss functions to ensure robust model performance in real-world.
- ***Model Pruning***: Pruned model from 1600G FLOPs to 320G FLOPs (input shape is 512x512x3) with almost the same performance.

Research Intern, JD AI Research, *Advisor: Dr. Tao Mei* (05/2018 - 09/2018)

> CVPR Workshop, LIP Contest (Ranked 2nd)

- **Data Augmentation**: Enhanced the training dataset by implementing image rotations, reversals, and applying Mask-RCNN for multi-human object detection.
- **Model Fine-Tuning and Fusion**: Customized and fine-tuned state-of-the-art networks including JPP-Net, SS-NAN, DenseNet, RefineNet, and UPerNet on the LIP training set. Employed fusion strategies and integrated key-point data for performance refinement.
- **Hard Class Mining and Performance Optimization**: Identified and targeted several hard-to-classify categories, thereby elevating the overall model performance to achieve an 87.42% pixel accuracy.

Professional Experience

CV Engineer, vivo AI Research, (07/2019 - 06/2021)

> Real-Time Frame Interposition and Motion Deblur in Mobile Devices:

- **CNN Construction**: Constructed a sophisticated two-stage model rooted in SuperSloMo architecture, focusing on optical flow prediction and frame restoration.
- **Computational Efficiency**: Reduced model FLOPs significantly by employing distillation and pruning techniques.
- **Unsupervised Learning Strategy**: Proposed an unsupervised learning approach, enabling effective model training on Adobe240fps and UCF101 datasets.
- **Parallel Computing Pipeline Built**: Engineered a high-performance computing pipeline utilizing OpenCL for GPU-based calculations and VCap for NPU operations. Achieved a substantial reduction

in pipeline latency, from 150 ms/frame to 25 ms/frame.

> Face Clustering in Digital Album:

- **CNN-based Facial Feature Extraction**: Engineered a Convolutional Neural Network (CNN) model specifically tailored for the extraction of human facial features.
- **Graph Clustering Algorithm (GCA)**: Devised a novel Graph Clustering Algorithm grounded in graph structure and K-Nearest Neighbors (KNN) theory.
- **Clustering Implementation**: Adapted the GCA and integrated it with cosine distance metrics for real-time facial clustering on smartphones.
- **Performance Metrics**: Elevated the Accuracy and Recall rates from 0.92/0.79 (without GCA) to 0.99/0.94 (with GCA), showcasing the efficacy of the proposed algorithm.
- **Deployment and Scale**: Successfully implemented the resultant algorithmic module in over 1 million vivo smartphones, highlighting both scalability and real-world applicability.

> Real-Time Human Segmentation in Mobile Devices.

- **CNN Construction**: Developed an Encoder-Decoder model utilizing UNet and Depthwise Separable Convolution techniques, implemented in TensorFlow.
- **Data Augmentation**: Expanded the PFCN dataset (18,698 images) and the vivo dataset (5,000 images) by applying a series of augmentation techniques including cropping, flipping, shifting, zooming, and alpha blending, resulting in over 60,000 augmented images.
- **Inter-Frame Consistency**: Introduced an alpha blending technique to maintain inter-frame consistency during the processing stage.
- **Computational Efficiency**: Improved model efficiency by applying distillation and pruning strategies, reducing inference time from 40 ms/frame to 13 ms/frame on Qualcomm Snapdragon 855 platforms.

Skills

Programming: Python, Java, MATLAB, C++

Deep Learning Framework: PyTorch, TensorFlow, TFLite, MegEngine

Computer Vision Tools: OpenCV, Numpy, BasicSR

Professional Fields: Low-Level Vision, Image/Video Quality Enhancement, Real-Time CNN for Mobile Device