

LINFENG ZHANG

(+86) 138-4030-1165 zhanglinfeng1997@outlook.com [Google Scholar](#) Born in 1997.10

EDUCATION & AWARDS

Tsinghua University (Ph.D. Student) 09/2019 - 06.2024 (Expectation)

Institute for Interdisciplinary Information Sciences, Supervised by Kaisheng Ma

Research Direction: Knowledge Distillation, Computer Vision, Efficient Neural Network

MSRA Fellowship, 2022 (Top-12 in Asia)

Jiangnanxiang Scholarship, 2022 (Top-20 in Tsinghua University)

National Scholarship, 2020 (2%), Tang Lixin Scholarship (First Prize, 1%), 2021

First Class Scholarship in Tsinghua University (2%), 2023

Northeastern University (Bachelor) 08/2015 - 06/2019

School of Computer Science And Engineering

Outstanding Graduates (<2%), Outstanding Undergraduate Thesis (<2%), The First Prize Scholarship(<3%)

Outstanding Winner in Microsoft Student Club Practice Space, 2018

The First Prize Winner of ACM and MCM in NEU

ENACTUS China Regional Competition First Prize, 2018

ENACTUS China Regional Innovation Competition Gold Award, 2018

INTERNSHIP

DIDI Global, Cargo Autonomous Driving 04/2022 - 04.2023

Acceleration of Point Cloud Detection and Multi-view Detection: We have proposed two knowledge distillation methods for acceleration of **point cloud-based and multi-view images-based 3D detection**. Detectors trained with these methods have been **deployed on autonomous driving trucks** and tested in the streets of Beijing. The corresponding paper has been published in **CVPR2023**.

Kuaishou Technology, Y-tech Lab 04/2021 - 10.2021

Acceleration of GAN for Cartoon Style Transfer: Cartoon style transfer is an essential AI function in the **live stream and video editing** in Kuaishou APP. I have proposed a novel **knowledge distillation** method for GAN compression and successfully compressed GANs from **5G FLOPs to 200M FLOPs**. This model has been used in business and the corresponding paper has been published in **CVPR2022**.

Institute for Interdisciplinary Information Core Technology 06/2019 - 12.2021

AI Model Optimization for AI Acceleration Chips: I optimized commonly used neural network models such as ResNet and Faster RCNN to meet the computational and storage requirements of Qiming 910/920 AI acceleration chips. In addition to this, I led the recruitment and training of talented individuals in the algorithm direction and established an engineering team of approximately ten members.

PUBLICATION

First Author × 18, Citation 1500+, H-index: 13

1. **Linfeng Zhang**, Jiebo Song, Anni Gao, Jingwei Chen, Chenglong Bao, and Kaisheng Ma. Be Your Own Teacher: Improve the Performance of Convolutional Neural Networks via Self Distillation. IEEE International Conference on Computer Vision (ICCV2019).

2. **Linfeng Zhang**, Zhanhong Tan, Jiebo Song, Jingwei Chen, Chenglong Bao, and Kaisheng Ma. SCAN: A Scalable Neural Networks Framework Towards Compact and Efficient Models. Neural Information Processing Systems (NeurIPS2019).
3. **Linfeng Zhang**, Muzhou Yu, Tong Chen, Zuoqiang Shi, Chenglong Bao, and Kaisheng Ma. Auxiliary Training: Towards Accurate and Robust Models. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR2020).
4. **Linfeng Zhang**, Yukang Shi, Zuoqiang Shi, Kaisheng Ma, and Chenglong Bao. Task-Oriented Feature Distillation. Neural Information Processing Systems (NeurIPS2020).
5. **Linfeng Zhang**, Chenglong Bao, and Kaisheng Ma. Self-Distillation: Towards Efficient and Compact Neural Networks. IEEE Transactions of Pattern Analysis and Machine Intelligence (IEEE TPAMI).
6. **Linfeng Zhang**, and Kaisheng Ma. Improve Object Detection with Feature-based Knowledge Distillation: Towards Accurate and Efficient Detectors. The International Conference on Learning Representations (ICLR2021).
7. **Linfeng Zhang**, Xin Chen, Xiaobing Tu, Pengfei Wan, Ning Xu, Kaisheng Ma. Wavelet Knowledge Distillation: Towards Efficient Image-to-Image Translation. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR2022).
8. **Linfeng Zhang**, Xin Chen, Junbo Zhang, Runpei Dong, Kaisheng Ma. Contrastive Deep Supervision. European Conference on Computer Vision (ECCV2022 **Oral Presentation**).
9. **Linfeng Zhang**, Runpei Dong, Huang-Shuo Tai Kaisheng Ma. Pointdistiller: structured knowledge distillation towards efficient and compact 3d detection. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR2023).
10. **Linfeng Zhang**, Kaisheng Ma. A Good Data Augmentation Policy Is Not All You Need: A Multi-Task Learning Perspective” in IEEE Transactions on Circuits and Systems for Video Technology (IEEE TCSVT),
11. **Linfeng Zhang**, Kaisheng Ma. Structured Knowledge Distillation Towards Efficient Object Detection. IEEE Transactions of Pattern Analysis and Machine Intelligence (IEEE TPAMI).
12. **Linfeng Zhang**, Kaisheng Ma. Tiny Updater: Towards Efficient Neural Network-Driven Software Updating. IEEE International Conference on Computer Vision (ICCV2023, **Oral Presentation**).
13. Fei Li*, **Linfeng Zhang*** (**co-first author**), Zhenbo Li. Multi-Frequency Representation with Privilege Information for Video Super-Resolution. IEEE International Conference on Computer Vision (ICCV2023).
14. **Linfeng Zhang**, Runpei Dong, Xin Chen, Kaisheng Ma. ReKo: Region-aware Knowledge Distillation Towards Efficient Image-to-Image Translation. The 34th British Machine Vision Conference 2023 (BMVC2023).
15. **Linfeng Zhang**, Yukang Shi, Ke Wang, Hung-shuo Tai, Yuan He, Kaisheng Ma. Structrued Knowledge Distillation Towards Multi-view 3D Detection. The 34th British Machine Vision Conference 2023 (BMVC2023 **Oral Presentation**).

16. **Linfeng Zhang**, Xiaoman Zhang, Chenglong Bao, Kaisheng Ma. Wavelet J-Net: A Frequency Perceptive on Convolutional Neural Networks. International Joint Conference on Neural Networks (IJCNN2021).
17. Yuni Lai*, **Linfeng Zhang*** (co-first author), Donghong Han, Rui Zhou, Guoren Wang, Fine-grained emotion classification of Chinese microblogs based on graph convolution networks. World Wide Web Journal (WWW J).
18. Muzhou Yu, **Linfeng Zhang (co-first Author)**, Kaisheng Ma, Revisiting Data Augmentation in Model Compression: An Empirical and Comprehensive Study, International Joint Conference on Neural Networks 2023.
19. Tianchen Zhao, Xuefei Ning, Ke Hong, Zhongyuan Qiu, Lu Pu, **Linfeng Zhang**, Yali Zhao, Lipu Zhou, Guohao Dai, Huazhong ang, Yu Wang. Ada3D: Exploiting the Spatial Redundancy with Adaptive Inference for Efficient 3D Object Detection. IEEE International Conference on Computer Vision (ICCV2023).
20. Runpei Dong, Zhanhong Tan, Mengdi Wu, **Linfeng Zhang**, Kaisheng Ma. Finding the Task-Optimal Low-Bit Sub-Distribution in Deep Neural Networks. International Conference on Machine Learning (ICML2022 **Spotlight**).
21. Runpei Dong, Zekun Qi, **Linfeng Zhang**, Junbo Zhang, Jianjian Sun, Zheng Ge, Li Yi, Kaisheng Ma. Autoencoders as Cross-Modal Teachers: Can Pretrained 2D Image Transformers Help 3D Representation Learning? International Conference on Learning Representation (ICLR2023)
22. Xiaolong Ma, Sheng Lin, Shaokai Ye, Zhezhi He, **Linfeng Zhang**, Geng Yuan, Sia Huat, Tan, Zhengang Li, Deliang Fan, Xuehai Qian, Xue Lin, Kaisheng Ma, Yanzhi Wang. Non-Structured DNN Weight Pruning – Is It Beneficial in Any Platform? IEEE Transactions on Neural Networks and Learning Systems (IEEE TNNLS).
23. Zimo Liao, Zhicheng Luo, Qianyi Huang, **Linfeng Zhang**, Fan Wu, Qian Zhang, Yi Wang. SMART: screen-based gesture recognition on commodity mobile devices. Annual International Conference on Mobile Computing and Networking (MobiCom21 **Oral Presentation**).
24. Muzhou Yu, Sia Huat Tan, Kailu Wu, Runpei Dong, **Linfeng Zhang**, Kaisheng Ma, CORSD: Class-Oriented Relational Self Distillation. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2023).
25. Zimo Liao, Zimo Liao, Zhicheng Luo, Qianyi Huang, **Linfeng Zhang**, Fan Wu, Qian Zhang, Guihai Chen, Gesture Recognition Using Visible Light on Mobile Devices. IEEE Transaction on Networking (IEEE TON).

PATENTS

1. **Linfeng Zhang**, Image Classification Processing Method of Neural Networks Based on Frequency Domain Wavelet Basis Processing, CN111382795B
2. **Linfeng Zhang**, An Enhanced Data Augmentation Method to Improve the Accuracy and Robustness of Convolutional Neural Networks, CN111368977B
3. **Linfeng Zhang**, Donghong Han, Yunni Lai, Xu Wang, Shuai Zhao, Fine-Grained Classification Method of Social Network Text Sentiment Based on Graph Convolutional Networks, CN110472003B

4. **Linfeng Zhang**, Image Object Detection Method and Detector Based on Knowledge Distillation and Its Training Method, CN112164054A
5. **Linfeng Zhang**, Task-Driven Feature Distillation Deep Neural Network Learning Training Method and System, Readable Storage Medium, CN112132268A
6. **Linfeng Zhang**, An Auxiliary Training Method to Simultaneously Improve the Accuracy and Robustness of Neural Networks, CN111275192B
7. **Linfeng Zhang**, A Self-Distillation Training Method and Scalable Dynamic Prediction Method for Convolutional Neural Networks, CN110472730A
8. **Linfeng Zhang**, Xin Chen, Xiaobing Tu, Training Method and Apparatus for Image Processing Model, Electronic Device, CN114492731A
9. Kaisheng Ma, **Linfeng Zhang**, A Method to Accelerate Diffusion Models through Multi-Student Knowledge Distillation
10. Muzhou Yu, **Linfeng Zhang**, Kaisheng Ma, Neural Network Model Compression Method and System, CN117151189A

MAIN PROJECTS

(1) Self-Distillation: Towards Efficient and Compact Models 2019 - 2020

Project Leader, First Author, **ICCV2019**×1, **NeurIPS2019**×1, **TPAMI**×1

- **Motivation:** Existing knowledge distillation methods need a pre-trained teacher model, which leads to unstable student performance and costs long training time.
- **Solution:** We propose the first teacher-free knowledge distillation method named self-distillation, which regards the deep layers of a model as the teacher and the shallow layers of a model as the student. Besides, self-distillation enables the student to work in a dynamic manner depending on the current input data. Self-distillation has already been used in Qiming AI chips, Huawei phones and Taobao recommendation system. It is considered as the pioneer and the most representative method in the domain of teacher-free knowledge distillation by many papers and surveys. Abundant researchers have proposed incremental work based on self-distillation in its theory and application on the other tasks, such as semantic segmentation and pre-trained language models.

(2) Auxiliary Training: Towards Accurate and Compact Models October 2020

Project Leader, First Author, **CVPR2020**

- **Motivation:** Deep neural networks are sensitive to the natural image corruption (e.g. noise, blurring). Naive data augmentation can alleviate this problem but leads to accuracy drop.
- **Solution:** We propose auxiliary training, which formulates data augmentation learning as a multi-task learning task, and solves it by adding several auxiliary classifiers for different types of augmentation. Besides, we introduce input-aware self distillation and selective batch normalization for further performance improvements. Our method first improves model accuracy and robustness at the same time.

(3) Task-Oriented Feature Distillation June 2021

Project Leader, Co-first Author, **NeurIPS2021**

- **Motivation:** We study the question: which kind of information of teachers are crucial for knowledge distillation and which kind of transformation function can extract these information.

- **Solution:** We assume that the task-oriented information are the most crucial information for knowledge distillation. Then, we apply several auxiliary classifiers to extract these task-oriented information and distill them from teachers to students. Besides, we also propose an orthogonal regularization method for better distillation performance. Experiments demonstrate the performance of our method on both RGB images and point cloud data.

(4) Improve Object Detection with Feature-based Knowledge Distillation

October 2021

Project Leader, First Author, ICLR2021

- **Motivation:** Most of knowledge distillation methods are designed for classification and usually fail in the more challenging task such as object detection
- **Solution:** we design two novel knowledge distillation methods (attention-guided distillation and non-local distillation) for object detection, which show excellent improvements on both one-stage and two-stage detectors, both anchor-based and anchor-free method, both object detection and instance segmentation. For example, it leads to 4.1 AP improvements on Faster RCNN on COCO2017.

(5) Wavelet Knowledge Distillation Towards Efficient Image-to-Image Translation

June 2022

Project Leader, First Author, CVPR2022

- **Motivation:** GAN-based image-to-image translation models usually have a tremendous amount of parameters and thus suffer from low efficiency and bulky memory usage.
- **Solution:** In this project, we have analyzed the performance of GANs from a frequency perspective, which quantitatively shows that GAN, especially small GAN lacks the ability to generate high-quality high-frequency information in images. Based on the above observation, we propose wavelet knowledge distillation to address this issue by only distilling the high frequency information, instead of all the information from images generated by the teacher.
- Besides, we have studied the relation between discriminators and generators during model compression. It shows that compression on discriminators is necessary for maintaining its competition with compressed generators in adversarial learning, which further benefits the performance of generators.

(6) Contrast Deep Supervision

January 2021

Project Leader, First Author, ECCV2022 Oral Presentation

- **Motivation:** The traditional training method only supervises the neural network at its last layer and propagates the supervision layer-by-layer, which leads to hardship in optimizing the intermediate layers. Deep supervision is proposed to add auxiliary classifiers to the intermediate layers of deep neural networks. By optimizing these auxiliary classifiers with the supervised task loss, the supervision can be applied to the shallow layers directly. However, deep supervision conflicts with the well-known observation that the shallow layers learn low-level features instead of task-biased high-level semantic features.
- **Solution:** We propose a novel training framework named Contrastive Deep Supervision, which supervises the intermediate layers with augmentation-based contrastive learning. Experimental results on nine popular datasets with eleven models demonstrate its effects on general image classification, fine-grained image classification and object detection in supervised learning, semi-supervised learning and knowledge distillation.

(7) Point Distiller: Knowledge Distillation for Efficient 3D Objectors

May 2022

Project Leader, First Author, CVPR2023

- **Motivation:** Compared with image based 2D detectors, point clouds based 3D detectors usually require more computation, which has limited their deployment on edge devices.
- **Solution:** Based on the properties of point clouds, we propose local distillation to extract the local geometric structure information of point clouds, and then distill it from the teacher to the student. Besides, we introduce a reweighted learning strategy, which aims to only distill the points or voxels which make a more crucial influence for the prediction. Experimental results show that our method achieves 4.00× compression of PointPillars, SECOND, and PointRCNN detectors on KITTI without mAP drop.

(8) A Good Data Augmentation Policy Is Not All You Need

March 2022

Project Leader, First Author, IEEE TCSVT

- **Motivation:** Usually, the design of augmentation policies faces a diversity-difficulty trade-off. Automatic augmentation methods have been proposed to address this issue by searching the optimal data augmentation at the cost of expensive computation overhead.
- **Solution:** We propose to break the diversity-difficulty trade-off from a multi-task learning perspective. By formulating model learning on the augmented images and the original images as the auxiliary task and the primary task in multi-task learning, respectively, hard augmentation does not directly influence the training of the primary branch, and thus its negative influence can be alleviated. Hence, neural networks can learn valuable semantic information even with a totally random augmentation policy.

(9) Tiny Updater: Towards Efficient Neural Network-Driven Software Updating

July 2023

Project Leader, First Author, ICCV2023

- During the update of neural network-based software, users are required to download all the parameters of the neural network anew, which harms the user experience. To address this issue, we propose to apply knowledge distillation and neural network pruning methods to reduce the number of parameters that changed during model updating. Experimental results show that Tiny Updater can update the neural network-based software by only downloading a few parameters (10% to 20%) instead of all the parameters in the neural network.

(10) Multi-Frequency Representation with Privilege Information for Video Super-Resolution

July 2023

Project Leader, Co-First Author, ICCV2023

- CNN's limited receptive field restricts its ability to capture long-range spatial-temporal dependencies, leading to unsatisfactory performance in video super-resolution (VSR). To tackle this challenge, this paper presents a novel multi-frequency representation enhancement module (MFE) that performs spatial-temporal information aggregation in the frequency domain. Besides, we propose a novel VSR model training method referred to as privilege training to improve the performance of VSR.

PROFESSIONAL SERVICE

- I have been the reviewer for academic conferences including but not limited to ICCV, CVPR, ECCV, NeurIPS, ICLR, ICML, AAAI, IJCAI. Besides, I also review papers for journals including IEE TIP, IEE TMM, IEEE TPAMI, Informatuon Fusion, IEEE Signal Processing, IEEE TCSVT and so on.
- I have been the teaching assistant for *Advances in AI Acceleration Chips Design*, Tsinghua University in 2020 and 2023.
- I have helped my Ph.D. supervisor in mentoring more than seven Master's and Ph.D. candidates.