SANGHYUK CHUN

sanghyuk.chun@gmail.com
http://sanghyukchun.github.io/home

RESEARCH INTERESTS

My research focuses on understanding leaned representations by machine learning models. Especially, I'm interested in the robustness and the uncertainty estimation of deep models on high-dimensional sensory data such as computer vision, speech enhancement and music information retrieval.

CONFERENCE AND WORKSHOP PAPERS

CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features
Sangdoo Yun, Dongyoon Han, Seongjoon Oh, **Sanghyuk Chun**, Junseok Choi, Youngjoon Yoo
Accepted at *International Conference on Computer Vision (ICCV) 2019*
Oral presentation (187/4303=4.3%)
Paper: https://arxiv.org/abs/1905.04899
Github: https://github.com/ClovaAI/CutMix

Photorealistic Style Transfer via Wavelet Transforms
Sanghyuk Chun*, Jaejun Yoo*, Youngjung Uh*, Byungkyu Kang, Jung-woo Ha
Accepted at *International Conference on Computer Vision (ICCV) 2019*
(*authors contributed equally)
Paper: https://arxiv.org/abs/1903.09760
Github: https://github.com/ClovaAI/WCT2

Visualizing and Understanding Self-attention based Music Tagging
Minz Won, **Sanghyuk Chun**, Xavier Serra
Presented at *International Conference on Machine Learning (ICML) 2019*
Machine Learning for Music Discovery Workshop.
Oral presentation
Paper: https://drive.google.com/open?id=1mYU1fjXkrcQBpTyzuCszyceBm2yNC9_0

An Empirical Evaluation on Robustness and Uncertainty of Regularization methods
Sanghyuk Chun, Seongjoon Oh, Sangdoo Yun, Dongyoon Han, Junsuk Choe Youngjoon Yoo
Presented at *International Conference on Machine Learning (ICML) 2019*
Uncertainty & Robustness in Deep Learning Workshop.
Paper: http://www.gatsby.ucl.ac.uk/~balaji/udl2019/accepted-papers/UDL2019-paper-21.pdf

Where To Be Adversarial Perturbations Added? Investigating and Manipulating Pixel Robustness Using Input Gradients
Sanghyuk Chun*, Jisung Hwang*, Younghoon Kim*, Jaejun Yoo, Jihoon Kim, Dongyoon Han
Presented at *International Conference on Learning Representations (ICLR) 2019*
Debugging Machine Learning Models Workshop. URL: https://debug-ml-iclr2019.github.io/
Paper: https://debug-ml-iclr2019.github.io/cameraready/DebugML-19_paper_4.pdf
(*authors contributed equally)

PREPRINTS

Toward Interpretable Music Tagging with Self-attention
Minz Won, **Sanghyuk Chun**, Xavier Serra
Paper: https://arxiv.org/abs/1906.04972
Github: https://github.com/minzwon/self-attention-music-tagging

Multi-Domain Processing via Hybrid Denoising Networks for Speech Enhancement
JangHyun Kim*, Jaejun Yoo*, **Sanghyuk Chun**, Adrian Kim, Jung-woo Ha
Paper: `https://arxiv.org/abs/1812.08914`
Demo: `http://mdphdnet.github.io/`
(*authors contributed equally)

WORK EXPERIENCES

Naver, Search & Clova Feb 2018 - Now
Machine Learning Researcher at CLAIR (Clova AI Research) *Seongnam, Korea*

- Developed and researched a cross-domain image retrieval system between human faces and emojis. Now deployed in LINE Japan and achieved 66% improvement of the before highest CTR.
- Participated in various research projects targeted to major machine learning conferences.
- **Research area:** unsupervised learning, cross-modal learning, generative model, robustness, uncertainty estimation

Kakao corp. Feb 2016 - Feb 2018
Research Engineer at ART (Advanced Recommendation Technology) *Seongnam, Korea*

- Developed and maintained a large-scale real-time recommender system (Toros) for various services and data domain in Daum and Kakao including news (Daum News similar article), cross-domain (between News articles and video clips), blog (Brunch similar article recommendation), music streaming (KakaoMini and Melon personalized recommendation and similar song recommendation), community service (Daum Cafe hit item recommendation), Web comic service (Daum Webtoon and Kakao Page) personalized item application push service in Daum app and Kakao Page.
- Researched online CTR optimization methods using Multi-armed Bandit (MAB) (Thompson Sampling). Developed and maintained the MAB system and achieved 20% improvement of CTR.
- Developed and maintained the online user clustering algorithm based on the streaming K-means for music streaming service using music features extracted from pre-trained deep networks.
- Researched and developed the text categorization deep learning system based on LSTM and continuous bag-of-words for Shopping service. About 60M items are automatically tagged using this system among more than 100M untagged items.
- **Keywords:** Multi-armed Bandit, user embedding, user clustering, online algorithm, matrix factorization, text embedding, text classification, cross-modal recommendation, similar-item recommendation, contents-based recommendation.

Naver, Naver Labs Aug 2015 - Dec 2015
Research Internship *Seongnam, Korea*

- Implemented Batch Normalization (BN) using C++ Caffe Framework and tested AlexNet, Inception v2 and VGG architectures at ImageNet datasets. Code: `http://github.com/SanghyukChun/caffe`
- Researched new normalization technique for RNN using Lua Torch Framework.

KAIST, ALIN-LAB (Algorithmic Intelligence Lab) Sep 2013 - Feb 2016
Research Internship and Master Student *Deajeon, Korea*

- Researched a new algorithm for multiple centered robust PCA, *i.e.*, subspace clustering with robust subspaces. Proposed an efficient and scalable algorithm to solve the proposed problem and an efficient initialization technique for the proposed algorithm. Also proved the convergence property of the proposed algorithm and the property of the proposed initialization algorithm.
  Master Thesis `http://library.kaist.ac.kr/thesis02/2016/2016M020143583_S1.pdf`
  Python Code: `http://github.com/SanghyukChun/SC_SI`
- Participated in a research project to develop a robust algorithm for ECG Authentication in a noisy environment. (based on noise reduction using PCA)

KAIST, NDSL (Networked and Distributed Computing System Lab) Jul 2013 - Oct 2013
Research Internship *Deajeon, Korea*

- Designed and developed high-performance and storage-efficient index structure for high-speed traffic (> 40G) monitoring and storing system using the Pthread and the asynchronous event-driven I/O.
- See section 4 of the published paper (not participated as an author) for details: *USENIC ATC* 2015. FloSIS: A Highly Scalable Network Flow Capture Systemfor Fast Retrieval and Storage Efficiency

IUM-SOCIUS Jun 2012 - Jan 2013
Internship, Software Engineer *Seoul, Korea*

- Developed and maintained the internal statistic and admin systems (Python Django, MongoDB).
- Developed and maintained the main service system (Java Spring, Java Spring Batch).
- Participated in the new service as main developer (Ruby on Rails).

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea *Feb 2016*
Master in Electrical Engineering
Adviser: Jinwoo Shin
Main research interest: robust PCA, subspace clustering, machine learning, algorithm
Master Thesis: **Scalable Iterative Algorithm for Robust Subspace Clustering: Convergence and Initialization**
Paper: `http://library.kaist.ac.kr/thesis02/2016/2016M020143583_S1.pdf`
Github: `http://github.com/SanghyukChun/SC_SI`

Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea *Feb 2014*
B.S. in Electrical Engineering and Management Science (double major)