Andrew Kae

CONTACT INFORMATION

Andrew Kae
New York, NY USA

E-mail: andrew.kae@gmail.com
Site: `https://www.andrewkae.com`

SUMMARY

Senior Machine Learning and Computer Vision research engineer with a Ph.D. in Computer Science and extensive industry experience developing and deploying on-device ML models for image and video understanding. Research background includes visual search, image synthesis, domain adaptation, semantic segmentation, action recognition, image inpainting, and image denoising.

EDUCATION

University of Massachusetts Amherst, Amherst, MA

Ph.D. in Computer Science, May 2014
- Advisor: Erik Learned-Miller

Cornell University, Ithaca, NY

M.Eng in Computer Science, December 2004
B.A. in Computer Science, May 2003

WORK EXPERIENCE

Google, New York, NY

Senior Software Engineer, Advanced Camera Algorithm Team **December 2022 - Present**
- Research and implement state-of-the-art machine learning models for image denoising in the HDR+ imaging pipeline used by the Google Camera app.
- Survey latest research in neural network and Transformer architectures, training methodologies, and loss functions that are suitable for on-device deployment.
- Implement, train, evaluate, and deploy new models to balance image quality within on-device constraints for compute, memory, and latency.
- Continually improve training data and infra, including identifying new data sources and integrating them into the training pipeline.
- Collaborate with partner teams to evaluate, tune, and integrate denoising models.

Apple, Cupertino, CA

Research Engineer in System Intelligence and Machine Learning **April 2019 - November 2022**
I worked on the following tasks for human action understanding:
- Researched ML models for human action understanding in personal video collections, enabling the Photos app to identify activities and organize videos into larger themes (e.g. outdoor activities) that can later be surfaced to the user.
- Designed, trained, and evaluated ML models for human action recognition with an emphasis on on-device deployment.
- Collaborated with engineering and product teams to integrate models into on-device systems and user experiences.
- Balanced model accuracy and robustness within on-device compute, memory, and latency constraints.

I also worked on these tasks for generative image inpainting:
- Researched generative image inpainting approaches for an on-device Photos feature designed to remove unwanted objects or people from user photos.
- Identified and developed appropriate training and evaluation data for image inpainting.
- Implemented, trained, and evaluated candidate models, including development of a working on-device demo.

Oath, New York, NY
Research Scientist in Video Intelligence Group **August 2017 - March 2019**

- Developed research prototypes for video understanding applications and investigated new ML approaches for production use.
- Developed image-to-video domain adaptation techniques to leverage web-supervised image data for video classification using a modified GAN loss.
- Helped mentor a project implementing image synthesis techniques for prototyping new advertising formats.
- Published research from these works at WACV and CVPR.

Curalate, New York, NY

Research Engineer **November 2014 - August 2017**

- Designed and implemented a visual fashion search framework that detected clothing items in user images and retrieved visually similar products from large product catalogs.
- Applied state-of-the-art object detection, visual representation, and image retrieval techniques to real-world product search.

Yahoo! Labs, Santa Clara, CA

Intern in Display Advertising Group **Summer 2010**

- Developed an approach for classifying display advertisements into a taxonomy using OCR features extracted from ad images together with additional ad-level features.
- Used Hadoop to process and analyze millions of advertisements.
- Research from this work resulted in a KDD workshop publication and patent.

Academic Research Experience

University of Massachusetts Amherst, Amherst, MA

- Conducted research in computer vision and machine learning with a focus on semantic image and video labeling.
- Developed models incorporating global shape priors into discriminative frameworks for semantic face labeling.
- Developed a conditional restricted Boltzmann machine approach to jointly model temporal and shape dependencies for semantic labeling of faces in YouTube videos.
- Published these works at CVPR 2013 and CVPR 2014.
- Conducted research in document-specific OCR, including methods for learning document-specific character models and exploiting language statistics to improve recognition.
- Published OCR research at CVPR, JMLR, IJDAR, and ICDAR.

Publications

Andrew Kae and Yale Song. *Image to Video Domain Adaptation Using Web Supervision.* Winter Conference on Applications of Computer Vision (WACV), 2020.

Bor-Chun Chen and Andrew Kae. *Toward Realistic Image Compositing with Adversarial Learning.* Computer Vision and Pattern Recognition (CVPR), 2019.

Andrew Kae, Benjamin Marlin, and Erik Learned-Miller. *The Shape-Time Random Field for Semantic Video Labeling.* Computer Vision and Pattern Recognition (CVPR), 2014.

Andrew Kae*, Kihyuk Sohn*, Honglak Lee, and Erik Learned-Miller. *Augmenting CRFs with Boltzmann Machine Shape Priors for Image Labeling.* Computer Vision and Pattern Recognition (CVPR), 2013.
*The first and second authors made equal contributions and should be considered co-first authors.

Gary B. Huang, Andrew Kae, Carl Doersch, and Erik Learned-Miller. *Bounding the Probability of Error for High Precision Optical Character Recognition.* Journal of Machine Learning Research

(JMLR), 2012.

Andrew Kae, Kin Kan, Vijay K Narayanan, Dragomir Yankov. *Categorization of Display Ads using Image and Landing Page Features*. The Third Workshop on Large-scale Data Mining: Theory and Applications'11 (LDMTA'11), in conjunction with SIGKDD2011.

Andrew Kae, David A. Smith, Erik Learned-Miller. *Learning on the Fly: A font-free approach towards multilingual OCR*. IJDAR, 2011.

Andrew Kae, Gary Huang, Erik Learned-Miller, Carl Doersch. *Improving State-of-the-Art OCR through High-Precision Document-Specific Modeling*. Computer Vision and Pattern Recognition (CVPR), 2010.

Andrew Kae, Erik Learned-Miller. 2009. *Learning on the Fly: Font free approaches to difficult OCR problems*. International Conference on Document Analysis and Recognition (ICDAR), 2009.

Skills

- **Languages**: Python, C++, Java, Scala
- **Machine Learning / Computer Vision**: PyTorch, JAX, TensorFlow, NumPy, OpenCV, MATLAB
- **Infrastructure**: AWS

Patents

- Andrew Kae, Kin Fai Kan, Vijay K. Narayanan. *Automatic classification of display ads using ad images and landing pages*. Publication number US8732014 B2.

Awards

- NSF East Asia and Pacific Summer Institute (EAPSI) Fellowship 2011.
- Computer Vision and Pattern Recognition (CVPR) Doctoral Consortium 2014.

Links

- `https://techcrunch.com/2014/05/04/whichbeers-wants-to-help-you-find-the-right-beer/`
- `https://www.meetup.com/deeplearn/events/238950120/`

Professional Service

Reviewer for

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- International Conference on Computer Vision (ICCV).