

# Yannis Kalantidis

4100 Manila Ave  
94609, Oakland, CA  
US Green Card holder  
☎ +1-415-254-1235  
✉ ykalant@image.ntua.gr  
🌐 www.skamalas.com

## Working Experience

- Feb 2017 – Sept 2019 **Research Scientist, Facebook AI**, Menlo Park, CA,  
Research and development on representation learning, video understanding and vision & language.
- Jan 2015 – Dec 2016 **Research Scientist, Yahoo Research**, San Francisco, CA,  
Research and development on web-scale visual search, classification from noisy data, video representation, adversarial & unsupervised learning. Collaborated with Stanford on the Visual Genome project [IJCV, 2016] (<https://visualgenome.org/>).

## Education

- 2009–2014 **Ph.D. in Computer Science**, *School of Electrical and Computer Engineering*, National Technical University of Athens, Greece.
- 2002–2009 **Diploma/M.Eng. in Electrical and Computer Engineering**, National Technical University of Athens, Athens, Greece.

## Scientific Publication, Patent & Citation Records

- Publications Papers published in CVPR, NeurIPS, ICCV, ECCV, ICLR, TPAMI, IJCV, AACL, CVIU, WSDM, CHI and other.
- Citations Google Scholar citations (profile link): **2256 citations**, **h-index: 18**, **i10-index: 25** (December 2019)
- Patents 3 US Patents, 4 US Patent Applications & 2 Defensive Publications.

## Selected Recent Publications

- B. Kang, S. Xie, M. Rohrbach, Z. Yan, A. Gordo, J. Feng, **Y. Kalantidis**. *Decoupling Representation and Classifier for Long-Tailed Recognition*. **ICLR, 2020**.
- Y. Chen, H. Fan, B. Xu, Z. Yan, **Y. Kalantidis**, M. Rohrbach, Y. Shuicheng, J. Feng. *Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution*. **ICCV, 2019**.
- Y. Chen, M. Rohrbach, Z. Yan, Y. Shuicheng, J. Feng, **Y. Kalantidis**. *Graph-Based Global Reasoning Networks*. **CVPR, 2019**.
- L. Zhou, **Y. Kalantidis**, X. Chen, J. Corso, M. Rohrbach. *Grounded Video Description*. **CVPR (oral), 2019**.
- B. Xiong, **Y. Kalantidis**, D. Ghadiyaram, K. Grauman. *Less is More: Learning Highlight Detection from Video Duration*. **CVPR 2019**.
- Z. Shou, Z. Yan, **Y. Kalantidis**, L. Sevilla-Lara, M. Rohrbach, X. Lin, S.-F. Chang. *DMC-Net: Generating Discriminative Motion Cues for Fast Compressed Video Action Recognition*. **CVPR 2019**.
- J. Liang, L. Jiang, L. Cao, **Y. Kalantidis**, L.J. Li and A. Hauptmann. *Focal Visual-Text Attention for Memex Question Answering*. *Transactions on Pattern Analysis and Machine Intelligence* **TPAMI, 2019**.
- J. Zhang, **Y. Kalantidis**, M. Rohrbach, M. Paluri, M. Elhoseiny. *Large-Scale Visual Relationship Understanding*. **AAAI, 2019**.
- Y. Chen, **Y. Kalantidis**, J. Li, Y. Shuicheng, J. Feng. *A<sup>2</sup>-Nets: Double Attention Networks*. **NeurIPS, 2018**.
- Y. Chen, **Y. Kalantidis**, J. Li, Y. Shuicheng, J. Feng. *Multi-Fiber Networks for Video Recognition*. **ECCV, 2018**.
- R. Krishna, ..., **Y. Kalantidis et al.** *Visual Genome: Connecting Language and Vision Using Crowdsourced Annotations*. **IJCV 2017**.
- S. Chancellor, **Y. Kalantidis**, J.A. Pater, M. De Choudhury, D.A. Shamma *Multimodal Classification of Moderated Online Pro-Eating Disorder Content*, **CHI, 2017**.
- P. Garrigues, S. Farfadi, H. Izadinia, K. Boakye and **Y. Kalantidis**. *Tag Prediction at Flickr: a View from the Darkroom*. **NeurIPS LSCVS Workshop (best paper award), 2016**.
- Y. Kalantidis**, C. Mellina and S. Osindero. *Cross-dimensional Weighting for Aggregated Deep Convolutional Features*. **ECCVw, 2016**.
- Y. Avrithis, **Y. Kalantidis**, E. Anagnostopoulos and I. Z. Emiris. *Web-scale image clustering revisited*. **ICCV (oral), 2015**.
- Y. Kalantidis** and Y. Avrithis. *Locally Optimized Product Quantization for Approximate Nearest Neighbor Search*. **CVPR, 2014**.

## Research Interests

- Summary **Representation learning in constrained environments & multi-modal learning**.  
Learning with limited computation, data and other resources, edge computing.
- Representation learning for image and video**.  
[ECCV 2016, ECCV 2018, NeurIPS 2018, CVPR 2019a, ICCV 2019]
- Vision and language, multi-modal search**.  
[IJCV 2017, WSDM 2017, AACL 2019, CVPR 2019b, TPAMI 2019]
- Video understanding**.  
Description and grounding [CVPR 2019b, Arxiv 2019], summarization [CVPR 2019c], motion [CVPR 2019d]
- Web-scale nearest neighbor search, clustering and classification**.  
[ECCV 2012, CVPR 2014, ICCV 2015, NeurIPSw 2016]

## Open-source Development Projects & Programming Skills

- VIRaL Principal developer of the Visual Image Retrieval and Localization tool: [viral.image.ntua.gr](http://viral.image.ntua.gr).
- LOPQ & CroW Principal developer of the Locally Optimized Product Quantization ([ykalant/lopq](https://github.com/ykalant/lopq)) and CroW ([ykalant/crow](https://github.com/ykalant/crow)) github repos.
- Facebook AI Contributor to a number of open-source Facebook AI github repos, including: [facebookresearch/GloRe](https://github.com/facebookresearch/GloRe), [facebookresearch/OctConv](https://github.com/facebookresearch/OctConv).
- Programming Python, C/C++, Matlab, Caffe2, pyTorch.