

Yannis Kalantidis

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

Education

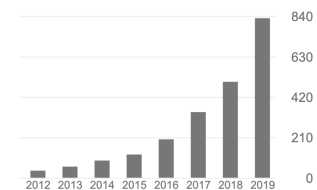
- 2009–2014 **Ph.D. in Computer Science**, *School of Electrical and Computer Engineering*, National Technical University of Athens, Greece.
Ph.D. Thesis: Clustering and nearest neighbor methods for visual search
- 2002–2009 **Diploma/M.Eng. in Electrical and Computer Engineering**, *School of Electrical and Computer Engineering*, National Technical University of Athens, Athens, Greece.
Diploma Thesis: Content- and geometry-based image retrieval

Working experience

- Feb 2017 – Sept 2019 **Research Scientist**, *Facebook AI*, Menlo Park, CA,
Research and development on representation learning, video understanding, vision and language, multi-modal learning.
- Jan 2016 – Dec 2016 **Research Scientist**, *Yahoo Labs*, San Francisco, CA,
Research and development on multi-modal classification, personal media search, vision and language.
- Jan 2015 – Dec 2015 **Postdoctoral Researcher**, *Yahoo Research*, San Francisco, CA,
Research and development on web-scale visual search, image & video representations, web-scale classification and clustering.
- 2009–2014 (5 years) **Image Video and Multimedia Laboratory, School of Electrical and Computer Engineering, National Technical University of Athens**, *Research and Development in FP7 European Projects*, Imagination (2009, under ICCS/NTUA), WEKNOWIT (2009-2011 under CERTH/ITI).

Scientific Publication, Patent & Citation Records

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



Google Scholar Citations graph (December 2019)

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, AAAI 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, NeurIPS 2016]

Teaching experience

- Fall 2016 **Guest lecturer at CS131 - Computer Vision: Foundations and Applications**, *Computer Science Department*, Stanford University, Course taught by Prof. Fei-Fei Li and Juan Carlos Niebles.
I was invited to give three lectures on basic computer vision principles (motion, gestalt and segmentation).
- June 2019 **Visual Search and Question Answering**, *Tutorial at the International Conference on Multimedia Expo (ICME)*, 2019, Together with Lu Jiang (Google AI) and Liangliang Cao (UMass).
- October 2019 **Image Representations and Fine-grained Recognition**, *Tutorial at Data Science Africa (DSA) Summer School*, 2019, Accra, Ghana.
- 2009–2014 **Image and Video Analysis, Lab assistant/lecturer**, *School of Electrical and Computer Engineering*, National Technical University of Athens, Greece.
- 2010–2014 **Programming Techniques, Teaching Assistant**, *School of Electrical and Computer Engineering*, National Technical University of Athens, Greece.
- 2009–2010 **Introduction to Programming, Teaching Assistant**, *School of Electrical and Computer Engineering*, National Technical University of Athens, Greece.

Awards & Scholarships

- 2017 **Outstanding Reviewer Award**, *International Conference on Computer Vision*, (ICCV).
- 2016 **Best Paper Award**, *NIPS Large-Scale Computer Vision Systems Workshop*.
- 2015 **Outstanding Reviewer Award**, *International Conference on Computer Vision*, (ICCV).
- 2013 **Thomaidion Award for Scientific Publication**, *Getting the Look: Clothing Recognition and Segmentation for Automatic Clothing Suggestions in Everyday Photos*, National Technical University of Athens.
- 2012 **Thomaidion Award for Scientific Publication**, *VIRaL: Visual Image Retrieval and Localization*, National Technical University of Athens.
- 2011 **Thomaidion Award for Scientific Publication**, *Retrieving landmark and non-landmark images from community photo collections*, National Technical University of Athens.
- 2005 **Erasmus Scholarship**, *Exchange Studies at Lund University*, Lund, Sweden.

Invited Talks

- Universita Politecnica Catalunya**, *Learning efficient representations for image and video understanding*, Deep Learning for Artificial Intelligence, November 21th 2019, Barcelona, Spain.
- Data Science Africa (DSA)**, *Computer Vision For Global Challenges: Learning and Outcomes*, Invited talk, October 2019, Accra, Ghana.
- Stanford University**, *Efficient spatio-temporal reasoning for video understanding*, Invited to present to the Vision Group by J.C. Niebles, February 2019, Stanford, CA, USA.
- Computer Vision Center (CVC)**, *Learning Spatiotemporal Video Representations*, Invited talk, September 18th 2018, Barcelona, Spain.
- University of California Berkeley**, *Learning Video Representations*, Guest Lecturer at CS 189, August 8th 2018, Berkeley, CA, USA.
- University of San Francisco**, *Pushing the Frontiers of Video Understanding*, Seminar Series in Data Science, April 20th 2018, San Francisco, CA, USA.

Universita Politecnica Catalunya, *Computer Vision at Facebook research*, Deep Learning Open Lecture, November 20th 2017, Barcelona, Spain.

Columbia University, Digital Video and Multimedia Laboratory, invited by Prof. Shih-Fu Chang, August 19th 2014, New York, NY, USA.

Stanford University, Image, Video and Multimedia Systems Group, invited by Prof. Bernd Girod, August 1st 2014, Stanford, CA, USA.

FXPAL, *Nearest Neighbor Search and Clustering for Large Scale Visual Search*, Fuji-Xerox Palo Alto Laboratory, July 29th 2014, Palo Alto, CA, USA.

Stanford University, Stanford Vision Lab, invited by Prof. Fei-Fei Li, July 28th 2014, Stanford, CA, USA.

Reviewer/Editor in journals

- 2016–today **Transactions on Pattern Analysis and Machine Intelligence**, *Reviewer*, [PAMI], IEEE.
- 2016–today **Computer Vision and Image Understanding**, *Reviewer*, [CVIU], Elsevier.
- 2019 **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, *Guest Editor for the Special Issue on Computer Vision-based Approaches for Earth Observation*, [JSTARS], IEEE.
- 2015–today **Signal Processing Letters**, *Reviewer*, [SPL], IEEE.
- 2015–today **Transactions on Knowledge and Data Engineering**, *Reviewer*, [TKDE], IEEE.
- 2015–today **The VLDB Journal**, *Reviewer*, [VLDBJ], Springer.
- 2015–today **Transactions on Image Processing**, *Reviewer*, [TIP], IEEE.
- 2010–2016 **Multimedia Tools and Applications**, *Reviewer*, [MTAP], Springer.
- 2012–2016 **Signal Processing: Image Communication**, *Reviewer*, [SPIC], Elsevier.
- 2014–2016 **Transactions on Circuits and Systems for Video Technology**, *Reviewer*, [CSVT], IEEE.
- 2015–2016 **Transactions on Intelligent Systems and Technology**, *Reviewer*, [TIST], ACM.
- 2014–today **Transactions on Multimedia**, [TMM], IEEE.

Organizer/Committee member in conferences & workshops

- 2020 **Workshop on Computer Vision for Agriculture, ICLR 2020**, *Lead Organizer*.
- 2019 **Workshop on Computer Vision for Global Challenges, CVPR 2019**, *Lead Organizer*.
- 2016–today **IEEE Conference on Computer Vision and Pattern Recognition**, *Technical Program Committee Member*, [CVPR].
- 2016–today **IEEE European Conference on Computer Vision**, *Technical Program Committee Member*, [ECCV].
- 2015–today **IEEE International Conference on Computer Vision**, *Technical Program Committee Member*, [ICCV], (**Outstanding Reviewer Award** in 2015 and 2017).
- 2013–today **ACM Multimedia Conference**, *Technical Program Committee Member*, [ACM-MM].
- 2011 **IEEE Workshop on Content-based Multimedia Indexing**, *Technical Program Committee Member*, [CBMI].
- 2009 **IEEE Workshop on Content-based Multimedia Indexing**, *Local Chair Member*, [CBMI].

Open-source development projects

- VIRaL Principal developer in the Visual Image Retrieval and Localization tool (On-line demo: <http://viral.image.ntua.gr>).
- LOPQ Principal developer in the open source code for Locally Optimized Product Quantization (Available on Yahoo's github: <https://github.com/yahoo/lopq>).

- Crow Principal developer in the open source code for Cross-dimensional feature Weighting for convolutional features (Available on Yahoo's github: <https://github.com/yahoo/crow>).
- Facebook AI Contributor to a number of open-source Facebook AI repos, including:
- Global Reasoning Networks: [facebookresearch/GloRe](https://github.com/facebookresearch/GloRe)
 - Octave Convolution: [facebookresearch/OctConv](https://github.com/facebookresearch/OctConv).

Patents & Defensive Publications

- US Patent **Visual Clothing Retrieval**, *Lyndon Kennedy, Li-Jia Li, Ioannis Kalantidis*, US Patent 9,460,518, Pub. Date: 4 October 2016.
- US Patent **Locally optimized feature space encoding of digital data and retrieval using such encoding**, *Lyndon Kennedy, Ioannis Kalantidis, David Shamma*, US Patent 9,659,214, Pub. Date: 2017/5/23.
- US Patent **Computerized system and method for modifying a media file by automatically applying security features to select portions of media file content**, *Bart Thomee, Ioannis Kalantidis, Lyndon Kennedy*, US Patent 10,467,387, Pub. Date: 2019/11/5.
- US Patent Application **Method and system for selecting supplemental content using visual appearance**, *Ioannis Kalantidis, Lyndon Kennedy, Ayman Farahat, Ricardo Baeza-Yates, Simon Osindero and David A. Shamma*, US Patent App. 15/097,658, 2017/10/19.
- US Patent Application **Systems and/or methods for constructing online ads using media objects created by generative processes**, *Clayton Mellina, Ioannis Kalantidis, Tobias Baumgartner*, US Patent App. 15857540, 2019/7/4.
- US Patent Application **Content sharing recommendations**, *Bart Thomee, Andrew Stadlen, Austin Shoemaker, Ioannis Kalantidis, Dan Tasse and David A. Shamma*, US Patent App. 15/134,633, 2017/10/26.
- US Patent Application **Selective content insertion into areas of media objects**, *Bart Thom  le, Ioannis Kalantidis, Clayton Ellis Mellina, Simon Osindero*, US Patent App. 15/204,450, 2018/1/11.
- Defensive Publication **Scalable Triangulation-based Logo Recognition**, *Yannis Kalantidis, Michele Trevisiol, Llu  s Garcia Pueyo and Roelof van Zwol*, Defensive Publication, Docket No.: ID11-7243, Pub. Date: April 29 2011.
- Defensive Publication **Cross-dimensional Weighting for Aggregated Deep Convolutional Features**, *Yannis Kalantidis, Clayton Mellina, Simon Osindero, David A. Shamma and Gerry Pesavento*, Defensive Publication, Docket No.: ID15-11464, Pub. Date: TBD.

Programming skills

- Programming Python, C/C++, Matlab. Familiar with PIG, HIVE, PHP.
- Other Pytorch, Caffe2, Hadoop, Spark, OpenCV, L  T  X, Pgfplots

Professional Memberships

- 2009–today **Institute of Electrical and Electronics Engineers**, [IEEE].
- 2009–today **Technical Chamber of Greece**, [GTC].

Languages

- Greek Native
- English Fluent
- German Basic

Certificate of Proficiency in English
Zertifikat Deutsch als Fremdsprache, Test DaF (grade 3)

List of Publications

Selected publications

- B. Kang, S. Xie, M. Rohrbach, Z. Yan, A. Gordo, J. Feng, **Y. Kalantidis**. Decoupling Representation and Classifier for Long-Tailed Recognition. In *International Conference on Learning Representations (ICLR)* 2020.
- C.Y. Ma, **Y. Kalantidis**, G. AlRegib, P. Vajda, M. Rohrbach, Z. Kira. Learning to Generate Grounded Image Captions without Localization Supervision. In *arXiv:1906.00283*, 2019.
- 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. In *International Conference on Computer Vision (ICCV)* 2019.
- Y. Chen, M. Rohrbach, Z. Yan, Y. Shuicheng, J. Feng, **Y. Kalantidis**. Graph-Based Global Reasoning Networks. In *Computer Vision and Pattern Recognition (CVPR)*, 2019.
- L. Zhou, **Y. Kalantidis**, X. Chen, J. Corso, M. Rohrbach. Grounded Video Description. In *Computer Vision and Pattern Recognition (CVPR)*, 2019.
- B. Xiong, **Y. Kalantidis**, D. Ghadiyaram, K. Grauman. Less is More: Learning Highlight Detection from Video Duration. In *Computer Vision and Pattern Recognition (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. In *Computer Vision and Pattern 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 In *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2019.
- J. Zhang, **Y. Kalantidis**, M. Rohrbach, M. Paluri A. Elgammal, M. Elhoseiny. Large-Scale Visual Relationship Understanding. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2019.
- Y. Chen, **Y. Kalantidis**, J. Li, Y. Shuicheng, J. Feng. Double Attention Networks. In *Neural Information Processing (NeurIPS)*, 2018.
- Y. Chen, **Y. Kalantidis**, J. Li, Y. Shuicheng, J. Feng. Multi-Fiber Networks. In *European Conference on Computer Vision (ECCV)*, 2018.
- R. Krishna, Y. Zhu, O. Groth, J. Johnson, K. Hata, J. Kravitz, S. Chen, **Y. Kalantidis**, L.-J. Li, D.A. Shamma, M. Bernstein and L. Fei-Fei. Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations. In *International Journal of Computer Vision (IJCV)*, 2017.
- S. Chancellor, **Y. Kalantidis**, J.A. Pater, M. De Choudhury, D.A. Shamma Multimodal Classification of Moderated Online Pro-Eating Disorder Content. In *Conference on Human Factors in Computing Systems (CHI)*, 2017.
- L. Jiang, **Y. Kalantidis**, L. Cao, S. Farfadi, J. Tang and A. Hauptmann. Delving Deep into Personal Photo and Video Search. In *Web Search and Data Mining (WSDM)*, 2017.
- P. Garrigues, S. Farfadi, H. Izadinia, K. Boakye and **Y. Kalantidis**. Tag Prediction at Flickr: a View from the Darkroom. In *Large Scale Computer Vision Systems Workshop, NeurIPS*, 2016, *Best Paper Award Winner*.
- Y. Kalantidis**, C. Mellina and S. Osindero. Cross-dimensional Weighting for Aggregated Deep Convolutional Features. In *Web-scale Vision and Social Media (VSM) Workshop, ECCV*, 2016.
- Y. Kalantidis**, A. Farahat, L. Kennedy, R. Baeza-Yates and D.A. Shamma. Visual Congruent Ads for Image Search. In *International Conference on Pattern Recognition (ICPR)*, 2016.

Y. Kalantidis, L. Kennedy, H. Nguyen, C. Mellina and D.A. Shamma. LOH and behold: Web-scale visual search, recommendation and clustering using Locally Optimized Hashing. In *Web-scale Vision and Social Media (VSM) Workshop, ECCV*, 2016.

Y. Avrithis, **Y. Kalantidis**, E. Anagnostopoulos and I. Z. Emiris. Web-scale image clustering revisited. In *International Conference on Computer Vision (oral) (ICCV)*, Santiago, Chile, December 2015.

Y. Kalantidis and Y. Avrithis. Locally Optimized Product Quantization for Approximate Nearest Neighbor Search. In *Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, June 2014.

G. Tolias, **Y. Kalantidis**, and Y. Avrithis. Towards large-scale geometry indexing by feature selection. *Computer Vision and Image Understanding (CVIU)*, March 2014.

Y. Kalantidis, L. Kennedy and L.-J. Li. Getting the Look: Clothing Recognition and Segmentation for Automatic Clothing Suggestions in Everyday Photos. In *International Conference on Multimedia Retrieval (Oral paper) (ICMR)*, Dallas, TX, April 2013.

Y. Avrithis and **Y. Kalantidis**. Approximate gaussian mixtures for large scale vocabularies. In *European Conference on Computer Vision (ECCV)*, Florence, Italy, October 2012.

G. Tolias, **Y. Kalantidis**, and Y. Avrithis. Symcity: Feature selection by symmetry for large scale image retrieval. In *ACM Multimedia (Oral paper) (MM 2012)*, Nara, Japan, October 2012. ACM.

Y. Avrithis, **Y. Kalantidis**, G. Tolias, and E. Spyrou. Retrieving landmark and non-landmark images from community photo collections. In *ACM Multimedia (Oral paper) (ACM MM 2010)*, Firenze, Italy, October 2010.

Y. Avrithis, G. Tolias, and **Y. Kalantidis**. Feature map hashing: Sub-linear indexing of appearance and global geometry. In *ACM Multimedia (Oral paper) (ACM MM 2010)*, Firenze, Italy, October 2010.

Y. Kalantidis, LG. Pueyo, M. Trevisiol, R. van Zwol, and Y. Avrithis. Scalable triangulation-based logo recognition. In *International Conference on Multimedia Retrieval*, Trento, Italy, April 2011.

Y. Kalantidis, G. Tolias, Y. Avrithis, M. Phinikettos, E. Spyrou, P. Mylonas, and S. Kollias. Viral: Visual image retrieval and localization. *Multimedia Tools and Applications (MTAP)*, 2011.

Other publications

L Jiang, LL Cao, Y Kalantidis, S Farfadi, AG Hauptmann. Visual Memory QA: Your Personal Photo and Video Search Agent In *AAAI*, demo presentation, 2017.

Y Kalantidis, C Xu, C Mellina Do Silhouettes Dream? *Creativity & Cognition*, Art Track, 2017.

I. Kollia, Y. Kalantidis, K. Rapantzikos, and A. Stafylopatis. Improving semantic search in digital libraries using multimedia analysis. In *Journal of Multimedia*, vol. 7, no 2, pp. 193-204, 2012.

I. Kollia, Y. Kalantidis, K. Rapantzikos, and S. Kollias. Visual analysis for semantic search in digital libraries. In *Constantinides International Workshop on Signal Processing (CIWSP 2013)*, London, 2013.

E. Spyrou, Y. Kalantidis, G. Tolias, Ph. Mylonas, and S. Kollias. Intelligent content retrieval using a visual vocabulary and geometric constraints. In *FUZZ-IEEE 2010, in conjunction with The 2010 IEEE World Congress on Computational Intelligence (WCCI 2010)*, July 2010.

- S. Papadopoulos, C. Zigkolis, G. Tolia, Y. Kalantidis, P. Mylonas, Y. Kompatsiaris, and A. Vakali. Image clustering through community detection on hybrid image similarity graphs. In *International Conference on Image Processing (ICIP 2010)*, Hong Kong, September 2010.
- E. Spyrou, Y. Kalantidis, and Ph. Mylonas. Exploiting a region-based visual vocabulary towards efficient concept retrieval. In *in Proceedings of Recognising and tracking events on the Web and in real life, in conjunction with SETN 2010 (EVENTS 2010)*, May 2010.
- Y. Kalantidis, E. Spyrou, Ph. Mylonas, and S. Kollias. Using a region and visual word approach towards semantic image retrieval. In *Workshop on Semantic Media Adaptation and Personalization (SMAP 2010)*, Limassol, Cyprus, December 2010.
- Y. Kalantidis, G. Tolia, E. Spyrou, Ph. Mylonas, Y. Avrithis, and S. Kollias. Visual image retrieval and localization. In *Greek Electrical and Computer Engineering Student Conference*, Thessaloniki, Greece, April 2009.
- Y. Kalantidis, G. Tolia, E. Spyrou, Ph. Mylonas, and Y. Avrithis. Visual image retrieval and localization. In *International Workshop on Content-Based Multimedia Indexing (CBMI 2009)*, Chania, Greece, June 2009.
- Y. Kalantidis. Content- and geometry-based image search. *Diploma Thesis (National Technical University of Athens, Athens, Greece, February 2009)*.