

Lluís Gomez-Bigorda

Computer Vision Center
Universitat Autònoma de Barcelona
Edificio O, Campus UAB
08193 Bellaterra (Cerdanyola)
Barcelona, Spain

Phone: +34 93 581 18 28
Fax: +34 93 581 16 70
Email: lgomez@cvc.uab.es
Homepage: <http://lluisgomez.github.io/>
GitHub: <http://github.com/lluisgomez/>

Employment

Universitat Autònoma de Barcelona, (Ramón y Cajal) Research Fellow, March 2022–present.
Computer Vision Center, Postdoctoral Researcher, April 2016–April 2022.
Computer Vision Center – Universitat Autònoma de Barcelona, Research and Teaching Assistant, 2011–2016.
Hanger.org, C/C++ Senior Developer and Principal Software Engineer, 2005–2011.
Pangea.org – Internet Service Provider, Senior Developer and System Administrator, 2001–2005.
Drac Telematic – Internet Service Provider, Junior Developer, 1999–2001.
Steria Computing Services, Junior Developer, 1999.

Education

Ph.D. Computer Vision, *Cum Laude*, Universitat Autònoma de Barcelona, 2016.
M.Sc. Computer Vision and Artificial Intelligence, Universitat Autònoma de Barcelona, 2012.
M.Sc. Computer Science, Universitat Oberta de Catalunya, 2011.
B.S. Computer Science, Universitat Oberta de Catalunya, 2010.

Languages

Catalan – mother tongue.
Spanish – mother tongue.
English – C1 level of the Common European Framework of Reference for Languages (CEFR). TOEFL certified.
Italian – basic conversation skills.

Research

Research Interests: Machine Learning, Computer Vision, Document Analysis, and Digital Humanities.

Peer-Reviewed Journal Articles

1. Gomez, L., Biten, A.F., Tito, R., Mafla, A., Rusiñol, M., Valveny, E., and Karatzas, D. (2021). Multimodal grid features and cell pointers for scene text visual question answering. *Pattern Recognition Letters*.
2. Mathew, M., Gomez, L., Karatzas, D., and Jawahar, C.V. (2021). Asking Questions on Handwritten Document Collections. *International Journal on Document Analysis and Recognition*.
3. Mafla, A., Tito, R., Dey, S., Gomez, L., Rusiñol, M., Valveny, E., and Karatzas, D. (2021). Real-time lexicon-free scene text retrieval. *Pattern Recognition*.

4. Bazazian, D., Gombru, R., Nicolaou, A., Gomez, L., Karatzas, D., and Bagdanov, A. (2017). FAST: Facilitated and Accurate Scene Text Proposals through FCN Guided Pruning. *Pattern Recognition Letters*.
5. Gomez, L., Nicolaou, A., and Karatzas, D. (2017). Improving Patch-based Scene Text Script Identification with Ensembles of Conjoined Networks. *Pattern Recognition*.
6. Gomez, L. and Karatzas, D. (2017). TextProposals: a Text-specific Selective Search Algorithm for Word Spotting in the Wild. *Pattern Recognition* (In press).
7. Gomez, L. and Karatzas, D. (2016). A Fast Hierarchical Method for Multi-script and Arbitrary Oriented Scene Text Extraction. *International Journal on Document Analysis and Recognition*.

Book Chapters

8. Gomez, R., Gomez, L., Gibert, J., Karatzas, D. (2019). Self-Supervised Learning from Web Data for Multimodal Retrieval. In *Multimodal Scene Understanding 1st Edition: Algorithms, Applications and Deep Learning (2019)*. Eds: M. Y. Yang, B. Rosenhahn, V. Murino. ISBN: 9780128173596. Elsevier / Academic Press.

Other Peer-Reviewed Conference and Workshop Articles

9. Nguyen, K., Biten, A.F., Mafla, A., Gomez, L., Karatzas, D. (2023). Show, Interpret and Tell: Entity-aware Contextualised Image Captioning in Wikipedia. In *Proc. 37th AAAI Conference on Artificial Intelligence (AAAI-23)*.
10. Souibgui, M.A., Biswas, S., Mafla, A., Biten, A.F., Fornés, A., Kessentini, Y., Lladós, J., Gomez, L., Karatzas, D. (2023). Text-DIAE: A Self-Supervised Degradation Invariant Autoencoders for Text Recognition and Document Enhancement. In *Proc. 37th AAAI Conference on Artificial Intelligence (AAAI-23)*.
11. Net, F., Folia, M., Casals-Puig, P., Gomez, L. (2023). Transductive Learning for Near-Duplicate Image Detection in Scanned Photo Collections. In *Proc. International Conference on Document Analysis and Recognition (ICDAR 2023)*.
12. Molina, A., Gomez, L., Ramos Terrades, O., Lladós, J. (2022). A Generic Image Retrieval Method for Date Estimation of Historical Document Collections. In *Proc. IAPR Workshop on Document Analysis Systems (DAS)*.
13. Vivoli, E., Biten, A.F., Mafla, A., Karatzas, D., Gomez, L. (2022). MUST-VQA: MULTilingual Scene-Text VQA. In *Proc. European Conference on Computer Vision (ECCV 2022) Workshops*.
14. Brugués i Pujolràs, J., Gomez, L., Karatzas, D. (2022). A Multilingual Approach to Scene Text Visual Question Answering. In *Proc. IAPR Workshop on Document Analysis Systems (DAS)*.
15. Biten, A.F., Tito, R., Gomez, L., Valveny, E., Karatzas, D. (2022). OCR-IDL: OCR Annotations for Industry Document Library Dataset. In *Proc. European Conference on Computer Vision (ECCV 2022) Workshops*.
16. Biten, A.F., Mafla, A., Gomez, L., Karatzas, D. (2022). Is An Image Worth Five Sentences? A New Look into Semantics for Image-Text Matching. In *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.
17. Biten, A.F., Gomez, L., Karatzas, D. (2022). Let there be a clock on the beach: Reducing Object Hallucination in Image Captioning. In *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.
18. Souibgui, M.A., Biten, A.F., Dey, S., Fornés, A., Kessentini, Y., Gomez, L., Karatzas, D., Lladós, J. (2022). One-shot Compositional Data Generation for Low Resource Handwritten Text Recognition. In *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.
19. Molina, A., Riba, P., Gomez, L., Ramos-Terrades, O., Lladós, J. (2021). Date Estimation in the Wild of Scanned Historical Photos: An Image Retrieval Approach. In *Proc. International Conference on Document Analysis and Recognition (ICDAR), 2021*.

20. Riba, P., Molina, A., Gomez, L., Ramos-Terrades, O., Lladós, J. (2021). Learning to Rank Words: Optimizing Ranking Metrics for Word Spotting. *In Proc. International Conference on Document Analysis and Recognition (ICDAR)*, 2021.
21. Mafla, A., Rezende, R.S., Gomez, L., Larlus, D., Karatzas, D. (2021). StacMR: Scene-Text Aware Cross-Modal Retrieval. *In Proc. Winter Conference on Applications of Computer Vision (WACV)*, 2021.
22. Mafla, A., Dey, S., Biten, A.F., Gomez, L., Karatzas, D. (2021). Multi-modal reasoning graph for scene-text based fine-grained image classification and retrieval. *In Proc. Winter Conference on Applications of Computer Vision (WACV)*, 2021.
23. Janouskova, K., Matas, J., Gomez, L., Karatzas, K. (2020). Text recognition-real world data and where to find them. *In Proc. International Conference on Pattern Recognition (ICPR)*, 2020.
24. Reddy, S., Mathew, M., Gomez, L., Rusiñol, M., Karatzas, D., Jawahar, C.V. (2020). RoadText-1K: Text Detection and Recognition Dataset for Driving Videos. *In Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2020.
25. Mafla, A., Dey, S., Biten, A.F., Gomez, L., Karatzas, D. (2020). Fine-grained Image Classification and Retrieval by Combining Visual and Locally Pooled Textual Features. *In Proc. IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2020.
26. Gomez, R., Gibert, J., Gomez, L., Karatzas, D. (2020). Exploring Hate Speech Detection in Multimodal Publications. *In Proc. IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2020.
27. Biten, A.F., Tito, R., Mafla, A., Gomez, L., Rusiñol, M., Valveny, E., Jawahar, C.V., Karatzas, D. (2019) Scene Text Visual Question Answering. *International Conference on Computer Vision (ICCV)*, 2019.
28. Rusiñol, M., Gomez, L., Landman, A., Silva-Constenla, M., Karatzas, D. Automatic Structured Text Reading for License Plates and Utility Meters. *BMVC Workshop on Visual AI and Entrepreneurship (VAIE)*, 2019.
29. Patel, Y., Gomez, L., Rusiñol, M., Karatzas, D., Jawahar, C.V. Self-Supervised Visual Representations for Cross-Modal Retrieval. *International Conference on Multimedia Retrieval (ICMR)*, 2019.
30. Gomez, R., Biten, A.F., Gomez, L., Gibert, J., Rusiñol, M., Karatzas, D. (2019). Selective Style Transfer for Text. *International Conference on Document Analysis and Recognition (ICDAR)*, 2019.
31. Biten, A.F., Tito, R., Mafla, A., Gomez, L., Rusiñol, M., Mathew, M., Jawahar, C.V., Valveny, E., Karatzas, D. (2019). ICDAR 2019 Competition on Scene Text Visual Question Answering. *International Conference on Document Analysis and Recognition (ICDAR)*, 2019.
32. Biten, A.F., Gomez, L., Rusiñol, M., Karatzas, D. (2019). Good News, Everyone! Context driven entity-aware captioning for news images. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
33. Gomez, L., Mafla, A., Rusiñol, M., Karatzas, D. (2018). Single Shot Scene Text Retrieval. *European Conference on Computer Vision (ECCV)*, IEEE, 2018.
34. Gomez, R., Gomez, L., Gibert, J., Karatzas, D. (2018). Learning to Learn from Web Data through Deep Semantic Embeddings. *Multimodal Learning and Applications Workshop, ECCV Workshops 2018*.
35. Gomez, R., Gomez, L., Gibert, J., Karatzas, D. (2018). Learning from #Barcelona Instagram data. *Multimodal Learning and Applications Workshop, ECCV Workshops 2018*.
36. Gomez, L., Rusiñol, M., Karatzas, D. (2018). Cutting Sayre's Knot: Reading Scene Text without Segmentation. Application to Utility Meters. *Document Analysis Systems (DAS), 13th IAPR International Workshop*.
37. Karatzas, D., Gomez, L., Rusiñol, M., Nicolaou, A. (2018). The Robust Reading Competition Annotation and Evaluation Platform. *Document Analysis Systems (DAS), 13th IAPR International Workshop*.
38. Gomez, L., Rusiñol, M., Karatzas, D. (2017). LSDE: Levenshtein Space Deep Embedding for Query-by-string Word Spotting. *Document Analysis and Recognition (ICDAR), 14th International Conference on. IEEE*, 2017.

39. Gomez, R., Shi, B., Gomez, L., Numann, L., Veit, A., Matas, J., Belongie, S., Karatzas, D. (2017). ICDAR2017 Robust Reading Challenge on COCO-Text. *Document Analysis and Recognition (ICDAR), 14th International Conference on. IEEE, 2017.*
40. Iwamura, M., Morimoto, N., Tainaka, K., Bazazian, D., Gomez, L., Karatzas, D. (2017). ICDAR2017 Robust Reading Challenge on Omnidirectional Video. *Document Analysis and Recognition (ICDAR), 14th International Conference on. IEEE, 2017.*
41. Gomez, L., Patel, Y., Rusiñol, M., Karatzas, D., Jawahar, C.V. (2017). Self-supervised learning of visual features through embedding images into text topic spaces. *International Conference on Computer Vision and Pattern Recognition (CVPR), 2017.*
42. Karatzas, D., Gomez, L., Rusiñol, M. (2017). The Robust Reading Competition Annotation and Evaluation Platform. *1st International Workshop on Open Services and Tools for Document Analysis.*
43. Patel, Y., Gomez, L., Rusiñol, M., and Karatzas, D. (2016). Dynamic Lexicon Generation for Natural Scene Images. *2nd International Workshop on Robust Reading, ECCV Workshops 2016.*
44. Bazazian, D., Gombru, R., Nicolaou, A., Gomez, L., Karatzas, D., and Bagdanov, A. Improving Text Proposals for Scene Images with Fully Convolutional Networks. *1st Workshop on Deep Learning for Pattern Recognition, ICPR Workshops 2016.*
45. Gomez, L., and Karatzas, D. (2016). A Fine Grained Classification Approach to Scene Text Script Identification. *Document Analysis Systems (DAS), 2016 12th IAPR International Workshop on. IEEE, 2016.*
46. Nicolaou, A., Gomez, L., and Karatzas, D. (2016). Visual Script and Language Identification. *Document Analysis Systems (DAS), 2016 12th IAPR International Workshop on. IEEE, 2016.*
47. Gomez, L., and Karatzas, D. (2015). Object Proposals for Text Extraction in the Wild. *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on. IEEE, 2015.*
48. Karatzas, D., Gomez, L., Nicolaou, A., Ghosh, S., Bagdanov, A., Iwamura, M., Matas, J., Neumann, L., Ramaseshan Chandrasekhar, V., Lu, S., Shafait, F., Uchida, S., Valveny, E. (2015). ICDAR 2015 Robust Reading Competition. *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on. IEEE, 2015.*
49. Ghosh, S., Gomez, L., Karatzas, D., Valveny, E. (2015). Efficient indexing for Query By String text retrieval. *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on. IEEE, 2015.*
50. Gomez, L., and Karatzas, D. (2014). Scene Text Recognition: No Country for Old Men?. *Computer Vision- ACCV 2014 Workshops. Springer International Publishing, 2014.*
51. Gomez, L., and Karatzas, D. (2014). MSER-Based Real-Time Text Detection and Tracking. *Pattern Recognition (ICPR), 2014 22nd International Conference on. IEEE, 2014.*
52. Karatzas, D., Robles, S., and Gomez, L. (2014). An on-line platform for ground truthing and performance evaluation of text extraction systems. *Document Analysis Systems (DAS), 2014 11th IAPR International Workshop on. IEEE, 2014.*
53. Gomez, L., and Karatzas, D. (2013). Multi-script text extraction from natural scenes. *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on. IEEE, 2013.*
54. Karatzas, D., Shafait, F., Uchida, S., Iwamura, M., Gomez, L., Robles Mestre, S., Mas, J., Fernandez, D., Almazan, J., de las Heras, L. P. (2013). ICDAR 2013 Competition on Robust Reading. *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on. IEEE, 2013.*

Invited Talks, Conference Tutorials and Demonstrations

2023: Gomez, L. Deep Archive: Exploring the past with artificial intelligence. *Invited Talk at the Experiment.AI event.*

- 2020: Gomez, L. Scene Text Understanding beyond OCR. *Invited Talk at the BCN.ai Winter event*.
- 2019: Gomez, L. Scene Text Visual Question Answering. *Invited Talk at the Deep Learning Barcelona Symposium (DLBCN 2019)*.
- 2019: Karatzas, D., Rusiñol, M., Gomez, L., Gomez, R., Gibert, J., Patel, Y. Vision and Language: the text modality in Computer Vision. *ICDAR-Tutorials. Document Analysis and Recognition (ICDAR), 2019 15th International Conference on. IEEE, 2019*.
- 2018: Gomez, L. Self-supervised learning of visual features through embedding images into text topic spaces. *Invited Talk at the Deep Learning Barcelona Symposium (DLBCN 2018)*.
- 2018: Gomez, L. Exploring large document collections by visualizing multi-modal Deep Semantic Embeddings. *Invited Talk at the 2018 meeting of the Institute for Document Analysis Knowledge Science (IDAKS)*.
- 2018: Gomez, L. Text and Vision: from scene text understanding to deep semantic embeddings. *Seminar at The Media Integration and Communication Center (MICC), Florence, 2018*.
- 2016: Gordo, A., and Gomez, L. Scene-Text Localization, Recognition, and Understanding. *DAS-Tutorials. Document Analysis Systems (DAS), 2016 12th IAPR International Workshop on. IEEE, 2016*.
- 2013: Gomez, L., and Karatzas, D. Demonstration of a Human Perception Inspired System for Text Extraction from Natural Scenes. *CBDAR-Demos. Document Analysis and Recognition (ICDAR), 2013 12th International Conference on. IEEE, 2013*.

Grants, Fellowships, & Awards

- Ramón y Cajal Research Fellowship (RYC2020-030777-I), Spanish Government Plan for Scientific Research and Innovation 2020.
- H2020 Marie Skłodowska-Curie actions of the European Union, grant agreement No 712949 (TECNIOspring PLUS), 2018.
- NVIDIA GPU Grant program recipient, 2018.
- Catalan University and Research Agency (AGAUR), Dissertation Completion Fellowship 2016.
- Google Summer of Code program recipient, 2014.
- Japan Student Services Organization (JASSO), Student Exchange Support Program Scholarship, 2013.
- Google Summer of Code program recipient, 2013.
- Catalan University and Research Agency (AGAUR), Research Fellowship, 2012-2016.
- Computer Vision Center, Research Scholarship 2012.
- Universitat Oberta de Catalunya, M.Sc. Thesis Excellence Distinction, 2011.
- Spanish Ministry of Science and Education Scholarship, 2010.
- Catalan University and Research Agency (AGAUR), Enginycat Scholarship, 2009.
- V2 Institute for the Unstable Media, exStream Residency Program finalist, 2004.
- Provincial Government of Barcelona, Special Jury Mention Best Digital Publishing Award, 1999.
- Universitat Autònoma de Barcelona, Steria-UAB University-Enterprise Fellowship, 1999.

Participation in R&D Projects

- 2022–Present: Coeli-IA - from Text to Media, paradigm shift in management/dissemination of heritage content (ACE034/21/000084). Principal Investigator: Lluís Gomez. Founded by ACCIÓ - Agency for Business Competitiveness, Catalan Government (145,000 €).
- 2021–Present: ReadQA - Reading Systems for Visual Question Answering (PID2020-116298GB-I00). Principal

Investigators: Dimosthenis Karatzas and Lluís Gomez. Founded by Spanish Ministry of Science, Innovation and Universities (90,000 €).

2021–Present: ReadQA - Reading Systems for Visual Question Answering (PID2020-116298GB-I00). Principal Investigators: Dimosthenis Karatzas and Lluís Gomez. Founded by Spanish Ministry of Science, Innovation and Universities (90,000 €).

2015–2018: RAW - Reading in the Wild (TIN2014-52072P). Principal Investigators: Dimosthenis Karatzas and Andrew D. Bagdanov. Founded by Spanish Ministry of Economy (110,000 €).

2012–2014: Text and the City - Human Centred Scene Text Understanding (TIN2011-24631). Principal Investigator: Dimosthenis Karatzas. Founded by Spanish Ministry of Science and Innovation (78,000 €).

2011–2013: Knowledge Extraction from Documents Images with Heterogenous Contents (TIN2009-14633-C03-03). Principal Investigator: Josep Lladós Canet. Founded by Spanish Ministry of Science and Innovation.

Professional Activities

Reviewer for *CVPR*, *ICCV*, *ECCV*, *ICLR*, *International Journal on Computer Vision*, *Transactions on Pattern Analysis and Machine Intelligence*, *Pattern Recognition*, *Computer Vision and Image Understanding*, *International Journal on Document Analysis and Recognition*, *Neurocomputing*, *Transactions on Image Processing*, Packt Publishing.

Workshop Chair, *International Workshop on Camera Based Document Analysis* (2019).

Workshop Chair, *International Workshop on Robust Reading* (2018).

Programme Committee member, *International Workshop on Document Analysis systems* (2018).

Programme Committee member, *Workshop on Computer Vision for Fashion, Art and Design* (2018).

Area Chair, *International Conference on Document Analysis* (2017).

Workshop Chair, *International Workshop on Camera Based Document Analysis* (2017).

Programme Committee member, *International Workshop on Robust Reading* (2014, 2016).

Programme Committee member, *Workshop on Camera Based Document Analysis* (2013).

Member, Computer Vision Foundation, 2017–Present.

Member, International Association of Pattern Recognition, 2013–Present.

Teaching and Invited Seminars

Universitat Autònoma de Barcelona

Visual Recognition (M.Sc. Computer Vision), Invited Lecturer (2016–present).

Machine Learning for Computer Vision (M.Sc. Computer Vision), Invited Lecturer (2016–present).

Knowledge, Reasoning, and Uncertainty (B.Sc. Computer Science), Lecturer (2022–present).

Fundamentals of Natural Language (B.Sc. Artificial Intelligence), Lecturer (2022–2023).

Artificial Neural Networks and Deep Learning (B.Sc. Data Engineering), Lecturer (2021–2022).

Artificial Neural Networks and Deep Learning (B.Sc. Computational Mathematics and Data Analytics), Lecturer (2019–2021).

Object-Oriented Programming (B.Sc. Computational Mathematics and Data Analytics), Teaching Assistant (2018-19, 2019-20).

Advanced Programming (B.Sc. Data Engineering), Teaching Assistant (2018-19).

Pattern Recognition (M.Sc. Computer Science), Teaching Assistant (2014-15).

ENTI (School of New Interactive Technologies) - Universitat de Barcelona (UB)

Artificial Intelligence (B.S. Interactive Digital Content), Lecturer (2017-18, 2018-19, 2019-20).

Universitat Politècnica de Valencia

Computer Vision (M.Sc. Multimedia), Seminar (2011).

BALTAN Laboratories, Eindhoven

Computer Vision Seminar (2009).

Artfutura Festival, Barcelona

Pure Data Real-Time Programming Seminar (2004).

Research Stays

University of Florence, Italy

Media Integration and Communication Center, Invited Researcher (2018-19).

Osaka Prefecture University, Japan

Intelligent Media Processing Group, Invited Researcher (Jan. 2013 - Mar. 2013).

Programming Skills

Programming Languages: C, C++, Python, Java, Objective-C, Matlab, GNU Octave, L^AT_EX, C#, Perl, PHP, SQL, Pure-Data.

Software Libraries and Frameworks: TensorFlow, Caffe, PyTorch, OpenCV, NumPy, SciPy, Matplotlib, Boost, OpenGL, FFmpeg, Gstreamer, ARToolKit and ARToolKitPlus, TUIO, Qt, Tcl/Tk, GTK+, v4l and v4l2, dv1394 and dc1394, JavaME and Android SDK.

Knowledge on Linux kernel device module programming.

Last updated: June 23, 2023