Lluis Gomez-Bigorda

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

Phone: +34 93 581 18 28
Email: lgomez@cvc.uab.es
Homepage: https://lluisgomez.github.io/
GitHub: https://github.com/lluisgomez/
ORCID: 0000-0003-1408-9803
Google Scholar: U5DQ99QAAAAJ

Summary

Publications: 10 journal articles indexed in JCR (8 in Q1); over 50 peer-reviewed conference papers, 10 in CORE A* venues and 25 in CORE A venues; 1 book (Ph.D. thesis) and 5 book chapters.

Supervision: 3 Ph.D. theses completed, 2 ongoing; 8 M.Sc. theses.

Impact (Google Scholar): h-index 29; 7,091 total citations; 957 citations per year on average over the last 4 years.

Employment

Universitat Autònoma de Barcelona (Dept. of Computer Science) and Computer Vision Center, Ramón y Cajal Research Fellow, March 2022–present.

Computer Vision Center, Postdoctoral Researcher, 2020–2022.

Osaka Prefecture University, Dept. of Computer Science and Intelligent Systems, Visiting Researcher, 2018–2022.

Computer Vision Center and Media Integration and Communication Center (University of Florence), TECNIOspring+ Postdoctoral Research Fellow (Marie Skłodowska-Curie COFUND), 2018–2020.

Computer Vision Center, Postdoctoral Researcher, 2016–2018.

Computer Vision Center – Universitat Autònoma de Barcelona, Research and Teaching Assistant, 2011–2016.

Hangar.org, C/C++ Senior Developer and Principal Software Engineer, 2005–2011.

Pangea.org – Internet Service Provider, Senior Developer and System Administrator, 2001–2005.

Drac Telemàtic – Internet Service Provider, Junior Developer, 1999–2001.

Steria Computing Services, Junior Developer, 1999.

Education

Ph.D. Computer Science, *Cum Laude*, Universitat Autònoma de Barcelona, 2016. Supervisor: Prof. Dimosthenis Karatzas.

M.Sc. Computer Vision and Artificial Intelligence, Universitat Autònoma de Barcelona, 2012.

M.Sc. (interdisciplinary) Free Software, Universitat Oberta de Catalunya, 2011.

B.Sc. 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, Responsible AI (bias, fairness and privacy in multimodal models), Document Analysis, and Digital Humanities.

Peer-Reviewed Journal Articles

1. Kang, L., Fu, X., Gomez, L., Fornés, A., Valveny, E., Karatzas, D. (2026). Preserving privacy without compromising accuracy: Machine unlearning for handwritten text recognition. *Pattern Recognition*, 172, 112411.
2. Kang, L., Fu, X., Souibgui, M.A., Barsky, A., Gomez, L., Vazquez-Corral, J., Fornés, A., Valveny, E., Karatzas, D. (2025). A Benchmark for Symbolic Reasoning from Pixel Sequences: Grid-Level Visual Completion and Correction. *Mathematics*, 13.
3. Net, F., Folia, M., Casals, P., Bagdanov, A.D., Gomez, L. (2025). EUFCC-340K: A faceted hierarchical dataset for metadata annotation in GLAM collections. *Multimedia Tools and Applications*.
4. Gomez, L., Biten, A.F., Tito, R., Mafla, A., Rusiñol, M., Valveny, E., Karatzas, D. (2021). Multimodal grid features and cell pointers for scene text visual question answering. *Pattern Recognition Letters*.
5. Mathew, M., Gomez, L., Karatzas, D., Jawahar, C.V. (2021). Asking Questions on Handwritten Document Collections. *International Journal on Document Analysis and Recognition*.
6. Mafla, A., Tito, R., Dey, S., Gomez, L., Rusiñol, M., Valveny, E., Karatzas, D. (2021). Real-time lexicon-free scene text retrieval. *Pattern Recognition*.
7. Bazazian, D., Gomez, R., Nicolaou, A., Gomez, L., Karatzas, D., Bagdanov, A. (2017). FAST: Facilitated and Accurate Scene Text Proposals through FCN Guided Pruning. *Pattern Recognition Letters*.
8. Gomez, L., Nicolaou, A., Karatzas, D. (2017). Improving Patch-based Scene Text Script Identification with Ensembles of Conjoined Networks. *Pattern Recognition*.
9. Gomez, L., Karatzas, D. (2017). TextProposals: a Text-specific Selective Search Algorithm for Word Spotting in the Wild. *Pattern Recognition*.
10. Gomez, L., Karatzas, D. (2016). A Fast Hierarchical Method for Multi-script and Arbitrary Oriented Scene Text Extraction. *International Journal on Document Analysis and Recognition*.

Books

11. Gomez, L. (2016). Exploiting similarity hierarchies for multi-script scene text understanding. Ph.D. Thesis, Universitat Autònoma de Barcelona. ISBN: 978-84-943427-9-0.

Book Chapters

12. Ruiz, S., Gomez, L. (2025). Charting Pathways: An Intersectional Impact Assessment for Vision and Language Foundation Models. In *Women, Technology, and Power – Unmasking (and dealing with) digital disparities in the times of the platforms*. Eds: M. Gutiérrez, A. Moreno Cano. Tirant lo Blanch.

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

14. Gomez, L. (2012). Web semántica y gestión de archivos: una introducción. In *Gestión de la innovación y nuevas estrategias de investigación y difusión del fondo documental artístico*. ISBN: 978-84-9704-993-3.

Peer-Reviewed Conference and Workshop Articles

15. El Aichouni, M., Kang, L., Gomez, L. (2027). Revisiting Fusion for Low-Data Multimodal Adaptation. *Conference of the European Chapter of the Association for Computational Linguistics (EACL 2027)*. Accepted, to appear.

16. Net, F., Molina, A., Llacer-Caro, S., Gomez, L. (2026). Revisiting how we access to historical archives: Auditing Gender Stereotypes and the Division of Labour in the Analysis of Historical Photography Collections. In Proc. *20th International Conference on Document Analysis and Recognition (ICDAR 2026)*.

17. Selvas-Sala, C., Kang, L., Gomez, L. (2026). SALMUBench: A Benchmark for Sensitive Association-Level Multimodal Unlearning. In Proc. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2026)*.

18. Gómez Grabowska, A., Gonzàlez, J., Gomez, L. (2026). Debiasing CLIP with Neural Interventions. In Proc. *48th European Conference on Information Retrieval (ECIR 2026)*.

19. Gomez, L. (2026). Rethinking Membership Inference Attacks for CLIP. In Proc. *40th AAAI Conference on Artificial Intelligence (AAAI-26)*.

20. Gomez, L. (2025). Measuring Text-Image Retrieval Fairness with Synthetic Data. In Proc. *48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2025)*.

21. Gomez, L. (2025). Over the Top-1: Uncertainty-Aware Cross-Modal Retrieval with CLIP. In Proc. *41st Conference on Uncertainty in Artificial Intelligence (UAI 2025)*.

22. El Aichouni, M., Gomez, L., Kang, L. (2025). Mitigating Distribution Bias in Multimodal Datasets via Clustering-Based Curation. In Proc. *Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA 2025)*.

23. Net, F., Gomez, L. (2024). EUFCC-CIR: A Composed Image Retrieval Dataset for GLAM Collections. In *European Conference on Computer Vision (ECCV 2024) Workshops*.

24. Kang, L., Yang, F., Wang, K., Souibgui, M.A., Gomez, L., Fornés, A., Valveny, E., Karatzas, D. (2024). GRIF-DM: Generation of Rich Impression Fonts using Diffusion Models. In Proc. *27th European Conference on Artificial Intelligence (ECAI 2024)*.

25. Kang, L., Souibgui, M.A., Yang, F., Gomez, L., Valveny, E., Karatzas, D. (2024). Machine Unlearning for Document Classification. In Proc. *International Conference on Document Analysis and Recognition (ICDAR 2024)*.

26. Net, F., Hernández, N., Molina, A., Gomez, L. (2024). A Transformer-Based Object-Centric Approach for Date Estimation of Historical Photographs. In Proc. *46th European Conference on Information Retrieval (ECIR 2024)*.

27. 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).

28. 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)*.

29. 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)*.

30. 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)*.

31. Vivoli, E., Biten, A.F., Mafla, A., Karatzas, D., Gomez, L. (2022). MUST-VQA: MUltilingual Scene-Text VQA. In *European Conference on Computer Vision (ECCV 2022) Workshops*.

32. 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)*.

33. Biten, A.F., Tito, R., Gomez, L., Valveny, E., Karatzas, D. (2022). OCR-IDL: OCR Annotations for Industry Document Library Dataset. In *European Conference on Computer Vision (ECCV 2022) Workshops*.

34. 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. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.

35. Biten, A.F., Gomez, L., Karatzas, D. (2022). Let there be a clock on the beach: Reducing Object Hallucination in Image Captioning. In Proc. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.

36. 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. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.

37. 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)*.

38. 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)*.

39. Mafla, A., Rezende, R.S., Gomez, L., Larlus, D., Karatzas, D. (2021). StacMR: Scene-Text Aware Cross-Modal Retrieval. In Proc. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.

40. 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. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.

41. Gomez, R., Gibert, J., Gomez, L., Karatzas, D. (2020). Location Sensitive Image Retrieval and Tagging. In Proc. *European Conference on Computer Vision (ECCV)*.

42. Janouskova, K., Matas, J., Gomez, L., Karatzas, D. (2020). Text recognition – real world data and where to find them. In Proc. *International Conference on Pattern Recognition (ICPR)*.

43. 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)*.

44. 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)*.

45. 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)*.

46. Biten, A.F., Tito, R., Mafla, A., Gomez, L., Rusiñol, M., Valveny, E., Jawahar, C.V., Karatzas, D. (2019). Scene Text Visual Question Answering. In Proc. *International Conference on Computer Vision (ICCV)*.

47. Rusiñol, M., Gomez, L., Landman, A., Silva-Constenla, M., Karatzas, D. (2019). Automatic Structured Text Reading for License Plates and Utility Meters. *BMVC Workshop on Visual AI and Entrepreneurship (VAIE)*.

48. Patel, Y., Gomez, L., Rusiñol, M., Karatzas, D., Jawahar, C.V. (2019). Self-Supervised Visual Representations for Cross-Modal Retrieval. In Proc. *International Conference on Multimedia Retrieval (ICMR)*.

49. Gomez, R., Biten, A.F., Gomez, L., Gibert, J., Rusiñol, M., Karatzas, D. (2019). Selective Style Transfer for Text. In Proc. *International Conference on Document Analysis and Recognition (ICDAR)*.

50. 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. In Proc. *International Conference on Document Analysis and Recognition (ICDAR)*.

51. Biten, A.F., Gomez, L., Rusiñol, M., Karatzas, D. (2019). Good News, Everyone! Context driven entity-aware captioning for news images. In Proc. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

52. Gomez, L., Mafla, A., Rusiñol, M., Karatzas, D. (2018). Single Shot Scene Text Retrieval. In Proc. *European Conference on Computer Vision (ECCV)*.

53. 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*.

54. Gomez, R., Gomez, L., Gibert, J., Karatzas, D. (2018). Learning from #Barcelona Instagram data. *Multimodal Learning and Applications Workshop, ECCV Workshops*.

55. Gomez, L., Rusiñol, M., Karatzas, D. (2018). Cutting Sayre's Knot: Reading Scene Text without Segmentation. Application to Utility Meters. In Proc. *13th IAPR International Workshop on Document Analysis Systems (DAS)*.

56. Karatzas, D., Gomez, L., Rusiñol, M., Nicolaou, A. (2018). The Robust Reading Competition Annotation and Evaluation Platform. In Proc. *13th IAPR International Workshop on Document Analysis Systems (DAS)*.

57. Gomez, L., Rusiñol, M., Karatzas, D. (2017). LSDE: Levenshtein Space Deep Embedding for Query-by-string Word Spotting. In Proc. *14th International Conference on Document Analysis and Recognition (ICDAR)*.

58. Gomez, R., Shi, B., Gomez, L., Neumann, L., Veit, A., Matas, J., Belongie, S., Karatzas, D. (2017). ICDAR2017 Robust Reading Challenge on COCO-Text. In Proc. *14th International Conference on Document Analysis and Recognition (ICDAR)*.

59. Iwamura, M., Morimoto, N., Tainaka, K., Bazazian, D., Gomez, L., Karatzas, D. (2017). ICDAR2017 Robust Reading Challenge on Omnidirectional Video. In Proc. *14th International Conference on Document Analysis and Recognition (ICDAR)*.

60. 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. In Proc. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

61. 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*.

62. Patel, Y., Gomez, L., Rusiñol, M., Karatzas, D. (2016). Dynamic Lexicon Generation for Natural Scene Images. *2nd International Workshop on Robust Reading, ECCV Workshops*.

63. Bazazian, D., Gomez, R., Nicolaou, A., Gomez, L., Karatzas, D., Bagdanov, A. (2016). Improving Text Proposals for Scene Images with Fully Convolutional Networks. *1st Workshop on Deep Learning for Pattern Recognition, ICPR Workshops*.

64. Gomez, L., Karatzas, D. (2016). A Fine Grained Classification Approach to Scene Text Script Identification. In Proc. *12th IAPR International Workshop on Document Analysis Systems (DAS)*.

65. Nicolaou, A., Gomez, L., Karatzas, D. (2016). Visual Script and Language Identification. In Proc. *12th IAPR International Workshop on Document Analysis Systems (DAS)*.

66. Gomez, L., Karatzas, D. (2015). Object Proposals for Text Extraction in the Wild. In Proc. *13th International Conference on Document Analysis and Recognition (ICDAR)*.

67. 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. In Proc. *13th International Conference on Document Analysis and Recognition (ICDAR)*.

68. Ghosh, S., Gomez, L., Karatzas, D., Valveny, E. (2015). Efficient indexing for Query By String text retrieval. In Proc. *13th International Conference on Document Analysis and Recognition (ICDAR)*.

69. Gomez, L., Karatzas, D. (2014). Scene Text Recognition: No Country for Old Men? *Computer Vision – ACCV 2014 Workshops*. Springer.

70. Gomez, L., Karatzas, D. (2014). MSER-Based Real-Time Text Detection and Tracking. In Proc. *22nd International Conference on Pattern Recognition (ICPR)*.

71. Karatzas, D., Robles, S., Gomez, L. (2014). An on-line platform for ground truthing and performance evaluation of text extraction systems. In Proc. *11th IAPR International Workshop on Document Analysis Systems (DAS)*.

72. Gomez, L., Karatzas, D. (2013). Multi-script text extraction from natural scenes. In Proc. *12th International Conference on Document Analysis and Recognition (ICDAR)*.

73. 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. In Proc. *12th International Conference on Document Analysis and Recognition (ICDAR)*.

Invited Talks, Conference Tutorials and Demonstrations

2025: Gomez, L. Invited Talk at Hangar.org AIAIAI.

2025: Gomez, L. Invited Talk at the Robust IR Workshop, SIGIR 2025.

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. Tutorial at the 15th International Conference on Document Analysis and Recognition (ICDAR 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 and 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.

2016: Gordo, A., Gomez, L. Scene-Text Localization, Recognition, and Understanding. Tutorial at the 12th IAPR International Workshop on Document Analysis Systems (DAS 2016).

2013: Gomez, L., Karatzas, D. Demonstration of a Human Perception Inspired System for Text Extraction from Natural Scenes. Demo at the 12th International Conference on Document Analysis and Recognition (ICDAR 2013).

Grants, Fellowships, & Awards

EuroHPC AI & Data-Intensive Applications Access Call (EHPC-AI-2024A02-040), 32,000 node hours on MareNostrum5, 2024.

Google Research Credits Program recipient, 2023.

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

As Principal Investigator

2025–2026: InterVisions – Participatory AI for Intersectional Bias Auditing (CERV 101214711). Co-PI (CVC): Lluis Gomez. Coordinated by ALIA – Associació Cultural de Dones per a la Recerca i l'Acció. Funded by the European Union, CERV programme (245,417 €).

2024–2025: FairCLIP – Training a Fair CLIP Model with Hybrid Real and Synthetic Data (EHPC-AI-2024A02-040). PI: Lluis Gomez. Funded by EuroHPC AI & Data-Intensive Applications Access Call (32,000 node hours on MareNostrum5).

2022–2024: Coeli-IA – From Text to Media: Paradigm Shift in the Management and Dissemination of Heritage Content. PI: Lluis Gomez. Funded by ACCIÓ, Catalan Government (144,542 €).

2021–2023: ReadQA – Reading Systems for Visual Question Answering (PID2020-116298GB-I00). Co-PIs: Dimosthenis Karatzas and Lluis Gomez. Funded by Spanish Ministry of Science and Innovation (89,419 €).

2020–2021: BeARS – Beyond Automatic Reading Systems. Co-PIs: Marçal Rusiñol and Lluis Gomez. Funded by AGAUR, Catalan Government (97,000 €).

2018–2020: DeepPhotoArchive. PI: Lluis Gomez. Funded by TECNIOspring PLUS (H2020 MSCA COFUND) / ACCIÓ (113,339 €).

As Research Team Member

2022–2025: ELSA – European Lighthouse on Secure and Safe AI. European Commission flagship initiative; member of the CVC research team within a consortium of 28 international institutions.

2018–2020: READS – Reading the Scene (TIN2017-89779P). Principal Investigators: Dimosthenis Karatzas and Ernest Valveny. Funded by Spanish Ministry of Economy, Industry and Competitiveness (81,554 €).

2018–2019: Semantic Search in Digital Newspaper Libraries. Principal Investigator: Marçal Rusiñol. Funded by Fundación BBVA (74,526 €).

2015–2017: RAW – Reading in the Wild (TIN2014-52072P). Principal Investigator: Dimosthenis Karatzas. Funded by Spanish Ministry of Economy and Competitiveness (109,021 €).

2012–2014: Text and the City – Human Centred Scene Text Understanding (TIN2011-24631). Principal Investigator: Dimosthenis Karatzas. Funded by Spanish Ministry of Science and Innovation (78,045 €).

2010–2013: Knowledge Extraction from Document Images with Heterogeneous Contents (TIN2009-14633-C03-03). Principal Investigator: Josep Lladós. Funded by Spanish Ministry of Science and Innovation (195,000 €).

2009–2011: HuPerText – Human Perception Inspired Text Technologies (TIN2008-04998). Principal Investigator: Dimosthenis Karatzas. Funded by Spanish Ministry of Science and Innovation (49,610 €).

Technology Transfer

Co-founder (2019) of AllRead Machine Learning Technologies, a start-up commercialising research outcomes developed with academic partners. AllRead closed investment rounds of 450,000 € (2020) and 2.5 M€ (2022, pre-series A), and today employs more than 20 people and operates in 10 countries in Europe and Latin America.

Supervision

Ph.D. Theses

Francesc Net Barnes (in progress), Universitat Autònoma de Barcelona.

Sofia Llàcer Caro (in progress), Universitat Autònoma de Barcelona.

Ali Furkan Biten (2022). A Bitter-Sweet Symphony on Vision and Language: Bias and World Knowledge. Universitat Autònoma de Barcelona.

Andrés Patricio Mafla Delgado (2022). Leveraging Scene Text Information for Image Interpretation. Universitat Autònoma de Barcelona.

Raúl Álamo Gómez Bruballa (2020). Exploiting the Interplay between Visual and Textual Image Content for Scene Interpretation. Universitat Autònoma de Barcelona.

M.Sc. Theses

Saranne van der Linde (2024–25). Decoding Emotions in Motion: Multimodal Emotion Recognition Tool for Migrant Videos. M.Sc. Computer Vision, UAB.

Razvan-Florin Apatean (2022–23). Improving Zero-shot Composed Image Retrieval for images with multiple objects. M.Sc. Computer Vision, UAB.

Hicham El Muhandiz Aarab (2022–23). Empirical Study of Uncertainty Estimation Techniques for Zero-Shot Image Classification Pre-trained CLIP models. M.Sc. Computer Vision, UAB.

Joan Fontanals Martínez (2021–22). 2-phase cross-modal search. Combining dual encoders with Vision and Language Transformers. M.Sc. Computer Vision, UAB.

Albert Jiménez Tauste (2020–21). Automated Video Editing for Synchronized Mobile Recordings of Concerts. M.Sc. Computer Vision, UAB.

Josep Brugués i Pujolràs (2020–21). Scene Text Visual Question Answering and Visual Question Generation: A Multilingual Approach. M.Sc. Computer Vision, UAB.

Roger Casals Vilardell (2019–20). Conditional GAN for Synthetic Text Dataset Generation. M.Sc. Computer Vision, UAB.

Diego Ernesto Apestegui Quesada (2018–19). Creation of a Large-Scale Weakly-Annotated Dataset of Supermarket Grocery Products for Computer Vision. M.Sc. Computer Engineering, UAB.

Professional Activities

Board member (Vocal), Spanish Association for Pattern Recognition and Image Analysis (AERFAI).

Member, European Laboratory for Learning and Intelligent Systems (ELLIS).

Core member of the ICDAR Robust Reading Competition organizing team (2013–2019), an international initiative involving more than 10 institutions, whose benchmarks have accumulated over 3,000 citations.

Program Committee member of CVPR, ICCV, ECCV, and AAAI.

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

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

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

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

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

Member, Computer Vision Foundation, 2017–present.

Member, International Association for Pattern Recognition, 2013–present.

Teaching and Invited Seminars

Universitat Autònoma de Barcelona

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

Synthesis Project I (B.Sc. Artificial Intelligence), Project Tutor.

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

Machine Learning for Computer Vision (M.Sc. Computer Vision), Invited Lecturer (2016–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 (2018–2020).

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

Artificial Intelligence (B.Sc. Interactive Digital Content), Lecturer (2017–present).

Universitat Politècnica de València

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, LaTeX, 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 of Linux kernel device module programming.