Marc'Aurelio Ranzato

Google DeepMind
London, England, United Kingdom
email: ranzato@google.com
web: https://ranzato.github.io/

RESEARCH INTERESTS

My primary research interests are in the area of machine learning, with applications to computer vision, natural language processing and speech recognition. In particular, I am interested in approaches to enable efficient learning and inference. I have been working on distributed large-scale systems, modular architectures, self-supervised and continual learning methods. I have designed algorithms for various applications, such as object recognition in natural images, machine translation, summarization and acoustic modeling for speech recognition.

CURRENT APPOINTMENT

Google DeepMind, London, England - United Kingdom
Position: *Senior Director of Research*
Period: August 2021 to present
- Leading group of about 30 Research Scientists and Research Engineers working on topics ranging from agentic self-improvement to distributed optimization, modular architectures, and continual learning.
- Co-leading Gemini area responsible for bringing research innovations from across Google DeepMind and Google Research into Gemini.
- Co-Leading strategic research initiatives in Frontier AI on topics related to architecture innovation and self-improvement.
- Managing a team of Research Scientists and Research Engineers across 3 sites in the U.K. and U.S.A.

PAST WORK EXPERIENCE

Facebook Inc., New York, NY - USA
Last Held Position: *Director, Research Scientist*
Period: September 2013 to June 2021
- Co-founded Facebook AI Research lab, took critical role in hiring and setting initial objectives and processes.
- Took leadership position to improve the organization; e.g., revise hiring process and define career expectations for Research Scientists.
- Managed a team of Research Scientists working on NLP.
- Significantly contributed as well as led several research projects on language modeling, low resource machine translation, continual learning, face recognition, etc.
- Contributed to product launch that improved translation services.

Google Inc., Mountain View, CA - USA
Last Held Position: *Research Scientist* at Google Brain Team
Period: October 2011 to September 2013
- Contributed to distbelief framework for training large-scale neural networks.
- Worked on various research projects in vision and audio modeling.

Yahoo! Research, Mission College, Santa Clara, CA, U.S.A.

Position: *Summer Intern*
Host: Kilian Weinberger, Malcolm Slaney, Olivier Chapelle, Kishore Papineni
Period: Summer 2008
Project: Learning sparse and locally shift-invariant feature hierarchies of images

Microsoft Research LTD., Cambridge, U.K.
Position: *Summer Intern*
Host: Martin Szummer
Period: Summer 2007
Project: Learning semantic representations of text documents from partially labeled collections

EDUCATION

New York University, The Courant Institute of Mathematical Sciences, New York, NY, U.S.A.
Ph.D. degree in Computer Science, May 2009
Advisor: Prof. Yann LeCun
Thesis: Unsupervised learning of feature hierarchies
Committee: Rob Fergus, Geoffrey Hinton, Yann LeCun, Sebastian Seung, Eero Simoncelli

University of Padova, Padova, Italy
Laurea in Electronics Engineering, April 2004 110/110 cum laude
Advisor: Prof. Pietro Perona and Prof. Ruggero Frezza
Thesis: Automatic recognition of biological particles in microscopic images

Conservatory "Frescobaldi", Ferrara, Italy
Diploma of music in violin, June 1999 9/10

ACADEMIC EXPERIENCE

University of Toronto, Department of Computer Science, Toronto, ON - CANADA
Position: *Post-Doctoral Fellow*
Supervisor: Prof. Geoffrey Hinton
Period: July 2009 to June 2011
Project: "Learning gated MRF's"
California Institute of Technology, Computer Vision Laboratory, Pasadena, CA, U.S.A.
Position: *Graduate Research Assistant*
Supervisor: Prof. Pietro Perona
Period: August 2003 - March 2004, July 2004 - August 2004
Project: "Automatic Visual Recognition of Biological Particles"

TEACHING EXPERIENCE

Teaching Assistant, New York University
Graduate Course on Machine Learning, Fall 2007. Instructor: Prof. LeCun
Guest Lectures
- Lecture at the Deep Learning: Theory, Algorithms and Applications Summer School, "Modular Continual Learning", Trento, 2023
- Guest Lecture on "Continual Learning in the Era of Very Large Models". Stanford University MSE 338, May 2023 - Instructor: Benjamin Van Roy
- Lecture at the Deep Learning Theory Summer School on"Learning from Non-Stationary Datasets",

Princeton, 2021
- Guest lecture on "Low-Resource Machine Translation", New York University Spring 2021 - Instructor: Yann LeCun
- Lecture at the Mediterrenean Machine Learning Summer School on "Low-Resource Machine Translation", Milano 2021
- Guest lecture on "Low-Resource Machine Translation", Stanford CS224N 2020 - Instructor: Christopher Manning
- Guest lecture on "Energy-Based Models for Structured Prediction - A case study: Machine Translation", Center for Data Science, New York University, Spring 2018 - Instructor: Prof. Yann LeCun
- Guest lexture on "Analyzing and Tackling Challenges in NMT" at Harvard University CS 287, Spring 2018 - Instructor: Prof. Alexander Rush
- Guest lecture at New York University, Spring 2017 - Instructor: Prof. Yann LeCun.
- Guest lecture on "Deep Learning for Vision" at Stanford CS 231A, Winter 2014 = Instructor: Prof. Silvio Savarese.
- UCLA IPAM Summer School 2012 on "Deep Learning".
- Guest lecture at Graduate Course on Advanced Machine Learning, Spring 2010. University of Toronto - Instructor: Prof. Zemel.
- Lecture at CIFAR Summer School, 2010.

Tutorials
- Tutorial on "Unsupervised Deep Learning" with Alex Graves, at the Conference on Neural Information Processing Systems, 2018
- Tutorial on "Challenges in Machine Translation" at the Advanced Course on Data Science and Machine Learning, Pontignano, 2018
- Tutorial on "Deep Learning for Vision, Text and Sequences" at DeepLearn Summer School, Bilbao 2017.
- Tutorial on "Deep Learning" at the International Conference on Computer Vision and Pattern Recognition 2014
- Tutorial on "Large Scale Visual Recognition" at the International Conference on Computer Vision and Pattern Recognition 2013
- Tutorial on "Deep Learning" with Y. LeCun at the International Conference on Machine Learning 2013
- Tutorial on "Deep Learning" at the International Conference on Computer Vision and Pattern Recognition 2012

Students Advised
- Zeyu Wang (summer 2025, internship at Google DeepMind)
- Bo Liu (summer 2023, internship at Google DeepMind)
- Adam Fisch (summer 2022, internship at DeepMind)
- Massimo Caccia (summer 2022, internship at DeepMind)
- Lucas Caccia (spring 2021)
- Yuntian Deng (summer 2019, internship at Facebook)
- Tom Veniat (summer 2019 - winter 2020, internship at Facebook, Ph.D. co-advisor)
- Lajanugen Logeswaran (summer 2018, internship at Facebook)
- Junxian He (summer 2019, internship at Facebook)
- Sandeep Subramanian (summer intern 2019 and 2019, internship at Facebook)
- Senthil Purushwalkam (summer 2018, internship at Facebook)
- Arslan Chaudhry (summer 2018, internship at Facebook)
- Tianxiao Shen (summer 2018, internship at Facebook)

- Guillaume Lample (Fall 2016 - Fall 2019, Ph.D. co-advisor)
- Yoon Kim (summer 2017, internship at Facebook)
- Sam Wiseman (summer 2016 and summer 2017, internship at Facebook)
- Sebastien Jean (summer 2015, internship at Facebook)
- Gregoire Mesnil (summer 2014 internship at Facebook)
- Babak Shakibi (winter 2014 internship at Facebook)
- David Eigen (summer 2013 internship at Google)
- Matthew D. Zeiler (summer 2012 internship at Google)
- Sida Wang (University of Toronto)

PROFESSIONAL ACTIVITIES

General Chair: Conference on Neural Information Processing systems 2021

Program Chair:

- Conference on Neural Information Processing systems 2020
- International Conference on Learning Representations 2017
- International Conference on Learning Representations 2018

Board Member:

- International Conference on Learning Representations 2018 - 2019
- Conference on Neural Information Processing systems 2021 till present
- Faculty of the Doctoral Program in Brain, Mind and Computer Science at the University of Padua, Italy.

Workshop Organizer

- Deep Learning Sympsium
  Conference: Neural Information Processing Systems (NIPS) 2015
- Tutorial on Deep Learning for Computer Vision
  Conference: Computer Vision and Pattern Recognition (CVPR) 2014
- Representation Learning
  Conference: International Conference of Machine Learning (ICML) 2012
  Co-oranizers: Aaron Courville, Hugo Larochelle, Yoshua Bengio
- Challenges in Learning Hierarchical Models: Transfer Learning and Optimization
  Conference: Neural Information Processing Systems (NIPS) 2011
  Co-oranizers: Quoc V. Le, Ruslan Salakhutdinov, Andrew Ng, Josh Tenenbaum
- Deep Learning and Unsupervised Feature Learning
  Conference: Neural Information Processing Systems (NIPS) 2010
  Co-organizers: Honglak Lee, Yoshua Bengio, Geoffrey Hinton, Yann LeCun, Andrew Y. Ng

Area Chair

- International Conference of Machine Learning
- Neural Information Processing Systems

- Computer Vision and Pattern Recognition
- International Conference of Computer Vision
- Computer Vision and Pattern Recognition
- ACM Multimedia
- Conference on Uncertainty in Artificial Intelligence

Reviewer

- Neural Computation
- Journal of Machine Learning
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- International Journal of Computer Vision
- International Journal of Machine Learning and Cybernetics
- Neural Information Processing Systems
- International Conference on Learning Representations
- International Conference of Machine Learning
- Artificial Intelligence and Statistics
- Computer Vision and Pattern Recognition
- Eurepean Conference of Computer Vision
- International Conference of Computer Vision
- Conference on Knowledge Discovery and Data Mining
- AAAI Conference on Artificial Intelligence
- Conference on Uncertainty in Artificial Intelligence
- International Joint Conference on Artificial Intelligence

Guest Editor for the International Journal of Computer Vision Special Issue on "Deep Learning", jointly with G.E. Hinton and Y. LeCun

HONORS

– Top-100 most influential scholars in Machine Learning for the year of 2016 according to AMiner.
– NYU Dean's dissertation fellowship, 2008-2009
– "Henry M. MacCracken" Award Scholarship, 2004-2008
– "Ing. Aldo Gini" Award for Italian researchers abroad, 2004

AWARDS

– Best paper award at EMNLP 2018
– Best resource paper award at EMNLP 2019
– Test of time award at ICML 2022

CONFERENCE PAPERS

- A. Douillard, Y. Donchev, J.K. Rush, S. Kale, Z. Charles, G. Teston, Z. Garrett, J. Shen, R. McIlroy, D. Lacey, A. Rame, A. Szlam, **M. Ranzato**, P.R. Barham *Streaming DiLoCo with overlapping communication.* Second Conference on Language Modeling (CoLM) 2025

- Y. Chen, X. Song, C. Lee, Z. Wang, Q. Zhang, D. Dohan, K. Kawakami, G. Kochanski, A. Doucet, **M. Ranzato**, S. Perel, N. de Freitas *Towards Learning Universal Hyperparameter Optimizers with Transformers.* Neural Information Processing Systems (NeurIPS) 2022

- L. Caccia, J. Xu, M. Ott, **M. Ranzato**, L. Denoyer. *On Anytime Learning at Macroscale.* Conference on Lifelong Learning Agents (COLLA) 2022

- A. Lee, M. Auli, **M. Ranzato** *Discriminative Reranking for Neural Machine Translation.* Association for Computational Linguistics (ACL) 2021

- T. Veniat, L. Denoyer, **M. Ranzato**. *Efficient Continual Learning with Modular Networks and Task-Driven Priors.* International Conference on Learning Representations (ICLR) 2021

- J. Shen, P.J. Chen, M. Le, J. He, J. Gu, M. Ott, M. Auli, **M. Ranzato**. *The Source-Target Domain Mismatch Problem in Machine Translation.* Conference of the European Chapter of the Association for Computational Linguistics (EACL) 2021

- Y. Deng, A. Bakhtin, M. Ott, A. Szlam, **M. Ranzato**. *Residual Energy-Based Models for Text Generation.* International Conference in Learning Representations (ICLR) 2020

- J. He, J. Gu, J. Shen, **M. Ranzato**. *Revisiting Self-Training for Neural Sequence Generation.* International Conference on Learning Represetations (ICLR) 2020

- P.J. Chen, J. Shen, M. Le, V. Chaudhary, A. El-Kishky, G. Wenzek, M. Ott, **M. Ranzato**. *Facebook AI's WAT19 Myanmar-English Translation Task Submission.* **Winner of the English-Burmese machine translation competition** at the Workshop on Asian Translation at EMNLP 2019

- S. Edunov, M. Ott, **M. Ranzato**, M. Auli. *On The Evaluation of Machine Translation Systems Trained With Back-Translation.* Association for Computational Linguistics (ACL) 2020

- G. Lample, A. Sablayrolles, **M. Ranzato**, L. Denoyer, H. Jegou *Large Memory Layers with Product Keys.* Conference on Neural Information Processing Systems (NeurIPS) 2019

- S. Purushwalkam, M. Nickel, A. Gupta, **M. Ranzato** *Task-Driven Modular Networks for Zero-Shot Compositional Learning* at the International Conference on Computer vision (ICCV) 2019

- F. Guzmán, P.J. Chen, M. Ott, J. Pino, G. Lample, P. Koehn, V. Chaudhary, **M. Ranzato**, *The FLoRes Evaluation Datasets for Low-Resource Machine Translation: Nepali-English and Sinhala-English.* **Best resource paper award** at the conference on Empirical Methods in Natural Language Processing (EMNLP) 2019

- T. Shen, M. Ott, M. Auli, **M. Ranzato**. *Mixture Models for Diverse Machine Translation: Tricks of the Trade.* International Conference on Machine Learning (ICML) 2019

- Arslan Chaudhry, **M. Ranzato**, M. Rohrbach, M. Elhoseiny *Efficient Lifelong Learning with A-GEM*. International Conference on Learning Representations (ICLR) 2019

- G. Lample, S. Subramanian, E. Smith, L. Denoyer, **M. Ranzato**, Y. Boureau. *Multiple-Attribute Text Rewriting*. International Conference on Learning Representations (ICLR) 2019

- G. Lample, M. Ott, A. Conneau, L. Denoyer and **M. Ranzato**. *Phrase-Based and Neural Unsupervised Machine Translation*. **best paper award** at the conference on Empirical Methods in Natural Language Processing (EMNLP) 2018

- M. Ott, M. Auli, D. Granger and **M. Ranzato**. *Analyzing Uncertainty in Neural Machine Translation*. International Conference on Machine Learning (ICML) 2018

- S. Edunov, M. Ott, M. Auli, D. Grangier and **M. Ranzato**, *Classical Structured Prediction Losses for Sequence to Sequence Learning*. North America Chapter of the Association for Computational Linguistics (NAACL) 2018

- G. Lample, A. Conneau, L. Denoyer and **M. Ranzato** *Unsupervised Machine Translation Using Monolingual Corpora Only*. International Conference on Learning Representations (ICLR) 2018

- A. Conneau, G. Lample, **M. Ranzato**, L. Denoyer and H. Jégou, *Word translation without parallel data*. International Conference on Learning Representations (ICLR) 2018

- S. Gross, **M. Ranzato** and A. Szlam, *Hard Mixtures of Experts for Large Scale Weakly Supervised Vision*. Computer Vision and Pattern Recognition (CVPR), 2017

- D. Lopez-Paz and **M. Ranzato**, *Gradient Episodic Memory for Continual Learning*. Neural Processing Informa tion (NIPS) 2017

- G. Lample, N. Zeghidour, N. Usunier, A. Bordes, L. Denoyer and **M. Ranzato**, *Fader networks: Manipulating images by sliding attributes*. Neural Processing Information (NIPS) 2017

- S. Chintala, **M. Ranzato**, A. Szlam, Y. Tian, M. Tygert and W. Zaremba *Scale-invariant learning and convolutional networks*. Applied and Computational Harmonic Analysis 42 (1), 154-166 2017

- J. Li, A.H. Miller, S. Chopra, **M. Ranzato** and J. Weston *Learning through dialogue interactions*. International Conference on Learning Representations (ICLR) 2017

- J. Li, A.H. Miller, S. Chopra, **M. Ranzato** and J. Weston *Dialogue learning with human-in-the-loop*. International Conference on Learning Representations (ICLR) 2017

- **M. Ranzato**, S. Chopra, M. Auli and W. Zaremba, *Sequence level training with recurrent neural networks*. International Conference on Learning Representations (ICLR) 2016

- Y. Taigman, M. Yang, **M. Ranzato** and L. Wolf, *Web-scale training for face identification*. Computer Vision and Pattern Recognition (CVPR), 2015

- T. Mikolov, A. Joulin, S. Chopra, M. Mathieu and **M. Ranzato**, *Learning Longer Memory in Recurrent Neural Networks*, in International Conference on Learning Representations (ICLR), 2015

- Y. Taigman, M. Yang, **M. Ranzato** and L. Wolf, *DeepFace: Closing the Gap to Human-Level Performance in Face Verification*, in IEEE Proc. of Computer Vision and Pattern Recognition Conference (CVPR), 2014

- N. Zhang, M. Paluri, **M. Ranzato**, T. Darrell and L. Bourdev, *PANDA: Pose Aligned Networks for Deep Attribute Modeling*, in IEEE Proc. of Computer Vision and Pattern Recognition Conference (CVPR), 2014

- **M. Ranzato**, J. Susskind, V. Mnih and G.E. Hinton, *On Deep Generative Models with Applications to Recognition*, in IEEE Proc. of Computer Vision and Pattern Recognition Conference (CVPR), 2011

- M. Denil, B. Shakibi, L. Dinh, **M. Ranzato**, N. de Freitas, *Predicting Parameters in Deep Learning*, Advances in Neural Information Processing Systems (NIPS), MIT Press, 2013

- A. Frome, G. Corrado, J. Shlens, S. Bengio, J. Dean, **M. Ranzato**, T. Mikolov, *DeViSE: A Deep Visual-Semantic Embedding Model*, Advances in Neural Information Processing Systems (NIPS), MIT Press, 2013

- M.D. Zeiler, **M. Ranzato**, R. Monga, M. Mao, K. Yang, Q.V. Le, P. Nguyen, A. Senior, V. Vanhoucke, J. Dean, G.E. Hinton, *On Rectified Linear Units for Speech Processing*, International Conference on Acoustic, Speech and Signal Processing (ICASSP), 2013

- A. Senior, G. Heigold, **M. Ranzato**, K. Yang, *An Empirical Study of Learning Rates in Deep Neural Networks for Speech Recognition*, International Conference on Acoustic, Speech and Signal Processing (ICASSP), 2013

- G. Heigold, V. Vanhoucke, A. Senior, P. Nguyen, **M. Ranzato**, M. Devin, J. Dean, *Multilingual Acoustic Models using Distributed Neep Neural Networks*, International Conference on Acoustic, Speech and Signal Processing (ICASSP), 2013

- J. Dean, G. Corrado, R. Monga, K. Chen, M. Devin, Q.V. Le, M. Mao,**M. Ranzato**, A. Senior, P. Tucker, K. Yang, A.Y. Ng, *Large Scale Distributed Deep Networks*, Advances in Neural Information Processing Systems (NIPS), MIT Press, 2012

- Q.V. Le, **M. Ranzato**, R. Monga, M. Devin, G. Corrado, K. Chen, J. Dean, A.Y. Ng, *Building High-Level Features Using Large Scale Unsupervised Learning*, International Conference of Machine Learning (ICML), 2012.

- K. Swersky, **M. Ranzato**, D. Buchman, B.M. Marlin, N. de Freitas, *On Autoencoders and Score Matching for Energy Based Models*, International Conference of Machine Learning (ICML), 2011.

- **M. Ranzato**, J. Susskind, V. Mnih and G.E. Hinton, *On Deep Generative Models with Applications to Recognition*, in IEEE Proc. of Computer Vision and Pattern Recognition Conference (CVPR), 2011

- **M. Ranzato**, V. Mnih and G.E. Hinton, *Generating More Realistic Images Using Gated MRF's*, Advances in Neural Information Processing Systems (NIPS), MIT Press, 2010

- G. Dahl, **M. Ranzato**, A. Mohamed and G.E. Hinton, *Phone Recognition with the Mean-Covariance Restricted Boltzmann Machine*, Advances in Neural Information Processing Systems (NIPS), MIT Press, 2010

- **M. Ranzato** and G.E. Hinton, *Modeling Pixel Means and Covariances Using Factorized Third-Order Boltzmann Machines*, in IEEE Proc. of Computer Vision and Pattern Recognition Conference (CVPR), 2010

- **M. Ranzato**, A. Krizhevsky and G.E. Hinton, *Factored 3-Way Restricted Boltzmann Machines for Modeling Natural Images*, in Proc. of the 13-th International Workshop on Artificial Intelligence and Statistics (AISTATS), 2010

- K. Jarrett, K. Kavukcuoglu, **M. Ranzato** and Y. LeCun, *What is the Best Multi-Stage Architecture for Object Recognition?*, in IEEE Proc. of International Conference on Computer Vision (ICCV), 2009

- K. Kavukcuoglu, **M. Ranzato**, R. Fergus, Y. LeCun, *Learning Invariant Features through Topographic Filter Maps*, in IEEE Proc. of Computer Vision and Pattern Recognition Conference (CVPR), 2009

- **M. Ranzato**, M. Szummer, *Semi-supervised Learning of Compact Document representations with Deep Networks*, International Conference of Machine Learning (ICML), 2008.

- **M. Ranzato**, Y. Boureau, Y. LeCun, *Sparse Feature Learning for Deep Belief Networks*, Advances in Neural Information Processing Systems (NIPS), MIT Press, 2007

- **M. Ranzato**, Y. LeCun, *A Sparse and Locally Shift Invariant Feature Extractor Applied to Document Images*, International Conference on Document Analysis and Recognition (ICDAR), 2007.

- Y. LeCun, S. Chopra, **M. Ranzato**, F.J. Huang, *Energy-Based Models in Document Recognition and Computer Vision*, International Conference on Document Analysis and Recognition (ICDAR), 2007.

- **M. Ranzato**, F.J. Huang, Y. Boureau, Y. LeCun, *Unsupervised Learning of Invariant Feature Hierarchies with Applications to Object Recognition*, in IEEE Proc. of Computer Vision and Pattern Recognition Conference (CVPR), 2007

- **M. Ranzato**, Y. Boureau, S. Chopra, Y. LeCun, *A Unified Energy-Based Framework for Unsupervised Learning*, In Proc. of the 11-th International Workshop on Artificial Intelligence and Statistics (AISTATS), 2007

- **M. Ranzato**, C.S. Poultney, S. Chopra, Y. LeCun, *Efficient Learning of Sparse Representations with an Energy-Based Model*, Advances in Neural Information Processing Systems (NIPS), MIT Press, 2006

JOURNAL PAPERS

- R. Cusack, C. O'Doherty, C.J. Charvet, **M. Ranzato** *Defending the foundation model view of infant development* Trends in Cognitive Sciences 2025.

- R. Cusack, **M. Ranzato**, C.J. Charvet *Helpless infants are learning a foundation model.* Trends in Cognitive Sciences 2024.

- J. Bornschein, A. Galashov, R. Hemsley, A. Rannen-Triki, Y. Chen, A. Chaudhry, X. He, A. Douillard, M. Caccia, Q. Feng, J. Shen, S. Rebuffi, K. Stacpoole, D. de las Casas, W. Hawkins, A. Lazaridou, Y.W. Teh, A.A. Rusu, R. Pascanu, **M. Ranzato** *NEVIS'22: A Stream of 100 Tasks Sampled from 30 Years of Computer Vision Research.* Journal of Machine Learning Research 2023.

- N. Goyal, C. Gao, V. Chaudhary, P.J. Chen, G. Wenzek, D. Ju, S. Krishnan, **M. Ranzato**, F. Guzman, A. Fan. *The FLORES-101 Evaluation Benchmark for Low-Resource and Multilingual Machine Translation.* Transactions of the Association for Computational Linguistics (TACL) 2022

- A. Bakhtin, Y. Deng, S. Gross, M. Ott, **M. Ranzato**, A. Szlam. *Residual Energy-Based Models for Text.* Journal of Machine Learning Research (JMLR) 2021.

- **M. Ranzato**, V. Mnih, J. Susskind, G.E. Hinton, *Modeling Natural Images Using Gated MRFs.* IEEE Trans. Pattern Analysis and Machine Intelligence, 2013.

- **M. Ranzato**, P.E. Taylor, J.M. House, R.C. Flagan, Y. LeCun, P. Perona, *Automatic recognition of biological particles in microscopic images.* Pattern Recognition Letters, Vol. 28, Issue 1, 1 Jan. 2007, pp. 31-39.

OTHER PUBLICATIONS AND TECHNICAL REPORTS

- Z. Huang, A. Kuncoro, Q. Feng, J. Shen, L. Dery, A. Szlam, **M. Ranzato**. *Context Training with Active Information Seeking.* arXiv:2605.13050 2026

- A. Douillard, K. Rush, Y. Donchev, Z. Charles, N. Fallen, A. Dubey, I. Gog, J. Dean, B. Woodworth, Z. Garrett, N. Keating, J. Bishop, H. Prior, E. Yvinec, A. Szlam, **M. Ranzato**, J. Dean. *Decoupled DiLoCo for Resilient Distributed Pre-training.* arXiv:2604.21428 2026

- A. Douillard, Y. Donchev, K. Rush, S. Kale, Z. Charles, Z. Garrett, G. Teston, D. Lacey, R. McIlroy, J. Shen, A. Ramé, A. Szlam, **M. Ranzato**, P. Barham. *Streaming DiLoCo with overlapping communication: Towards a Distributed Free Lunch.* arXiv 2501.18512 2025

- Gemini Team. *Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities.* arXiv:2507.06261 2025

- A. Douillard, Q. Feng, A. Rusu, A. Kuncoro, Y. Donchev, R. Chhaparia, I. Gog, M. **Ranzato**, J. Shen, A. Szlam. *DiPaCo: Distributed Path Composition.* arXiv 2403.10616 2024

- B. Liu, R. Chhaparia, A. Douillard, S. Kale, A. Rusu, J. Shen, A. Szlam, **M. Ranzato**. *Asynchronous Local-SGD Training for Language Modeling.* arXiv 2401.09135 2024

- A. Douillard, Q. Feng, A. Rusu, R. Chhaparia, Y. Donchev, A. Kuncoro, **M. Ranzato**, A. Szlam, J. Shen. *DiLoCo: Distributed Low-Communication Training of Language Models.* arXiv 2311.08105 2023

- A. Fisch, A. Rannen-Triki, R. Pascanu, J. Bornschein, A. Lazaridou, E. Gribovskaya, **M. Ranzato** *Towards robust and efficient continual language learning.* arXiv 2307.05741 2023

- M. Caccia, A. Galashov, A. Douillard, A. Rannen-Triki, D. Rao, M. Paganini, L. Charlin, **M. Ranzato**, R. Pascanu. *Towards compute-optimal transfer learning.* arXiv 2304.13164 2023

- L. Caccia, J. Xu, M. Ott, **M. Ranzato**, L. Denoyer *On Anytime Learning at Macroscale.* arXiv:2106.09563 2021
- N. Goyal, C. Gao, V. Chaudhary, P.J. Chen, G. Wenzek, D. Ju, S. Krishnan, **M. Ranzato**, F. Guzman, A. Fan. *The FLORES-101 Evaluation Benchmark for Low-Resource and Multilingual Machine Translation.* arXiv:2106.03193 2021
- S. Subramanian, R. Collobert, **M. Ranzato**, Y. Boureau. *Multi-scale Transformer Language Models.* arXiv:2005.00581 2020
- A. Bakhtin, S. Gross, M. Ott, Y. Deng, **M. Ranzato**, A. Szlam *Real or Fake? Learning to Discriminate Machine from Human Generated Text.* arXiv:1906.03351 2019
- A. Chaudhry, M. Rohrbach, M. Elhoseiny, T. Ajanthan, P.K. Dokania, P.H.S. Torr, **M. Ranzato** *On Tiny Episodic Memories in Continual Learning.* arXiv:1902.10486 2019
- S. Wiseman, S. Chopra, **M. Ranzato**, A. Szlam, R. Sun, S. Chintala and N. Vasilache, *Training Language Models Using Target-Propagation.* arXiv:1702.047702017
- J. Van Amersfoort, A. Kannan, **M. Ranzato**, A. Szlam, D. Tran and S. Chintala, *Transformation-based models of video sequences.* arXiv:1701.08435 2016
- M. Tygert, A. Szlam, S. Chintala, **M. Ranzato**, Y. Tian, W. Zaremba, *Convolutional networks and learning invariant to homogeneous multiplicative scalings.* arXiv:1506.08230 2015
- **M. Ranzato**, A. Szlam, J. Bruna, M. Mathieu, R. Collobert and S. Chopra, *Video (Language) Modeling: A Baseline for Generative Models of Natural Videos*, arXiv:1412.6604 2015
- **M. Ranzato** *On Learning Where To Look*, ArXiv:1405.5488 2014
- K. Kavukcuoglu, **M. Ranzato**, Y. LeCun, *Fast Inference in Sparse Coding Algorithms with Applications to Object Recognition*, CBLL Technical Report December 2008, arXiv 1010.3467
- Y. LeCun, S. Chopra, R. Hadsell, **M. Ranzato**, F.J. Huang, *A Tutorial on Energy-Based Learning*, in Bakir et al. (eds) "Predicting Structured Outputs", MIT Press 2006

WORKSHOP AND DEMONSTRATION PAPERS

- L. Logeswaran, A. Lee, M. Ott, H. Lee, **M. Ranzato**, A. Szlam. *Few-shot Sequence Learning with Transformers.* Meta-Learning Workshop at NeurIPS 2020
- G. Mesnil, T. Mikolov, **M. Ranzato**, Y. Bengio, *Ensemble of Generative and Discriminative Techniques for Sentiment Analysis of Movie Reviews*, workshop at the International Conference on Representation Learning (ICLR), 2015
- D. Eigen, I. Sutskever, **M. Ranzato**, *Learning Factored Representations in a Deep Mixture of Experts*, workshop at the International Conference on Representation Learning (ICLR), 2014
- O. Yadan, K. Adams, Y. Taigman, **M. Ranzato** *Multi-GPU Training of ConvNets*, workshop at the International Conference on Representation Learning (ICLR), 2014
- E. Horster, M. Slaney, **M. Ranzato**, K. Weinberger, *Unsupervised Image Ranking*, Proc. of the first ACM workshop on Large-scale multimedia retriaval and mining, Beijing, China, 2009

- K. Kavukcuoglu, **M. Ranzato**, Y. LeCun, *Fast Inference in Sparse Coding Algorithms with Applications to Object Recognition*, "Optimization for Machine Learning" workshop at Advances in Neural Information Processing Systems (NIPS), 2008
- D.R. Edgington, I. Kerkez, D.E. Cline, **M. Ranzato**, P. Perona, *Detecting, Tracking and Classifying Animals in Underwater Video*, IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), demonstration, New York, New York, 2006
- D.R. Edgington, I. Kerkez, D.E. Cline, D. Oliver, **M. Ranzato**, P. Perona, *Detecting, Tracking and Classifying Animals in Underwater Video*, IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), demonstration, San Diego, California, 2005

INVITED TALKS

- *Modular Foundation Models*
  - NeurIPS Workshop on Continual Learning for Lifelong Foundation Models, 2024
- *Learning and accruing knowledge over time using modular architectures*
  - Università Sapienza, Roma
  - International Symposium on Visual Computing 2021
- *Learning from Non-Stationary Datasets*
  - Deep Learning Theory Summer School at Princeton, 2021
  - Colloquia Patavina, 2021
- *Deep Learning for Vision*
  - workshop on Web-Scale Vision and Social Media at CVPR 2014
  - Workshop on Scene Understanding at CVPR 2014
  - Large Scale Visual Recognition tutorial at CVPR 2014
  - Large Scale Visual Recognition tutorial at CVPR 2013
  - Bay Area Vision Meeting 2013
- *On The Quest For Good Generative Models of Natural Images*
  - CIFAR NCAP Workshop, Vancouver, Canada, December 2010
  - CBLL seminar, New York University, New York, November 2010
- *Modeling Natural Images with Higher-Order Boltzmann Machines*
  - CIFAR Summer School, Toronto, Canada, August 2010
  - Redwood Center for Theoretical Neuroscience, June 2010
  - Computer Science Department, Stanford, June 2010
  - Google Research, Mountain View, June 2010
  - Department of Computer Science, University of California, San Diego, June 2010
  - Department of Computer Science, University of California, Irvine, June 2010

  - CIFAR NCAP Workshop, Vancouver, Canada, December 2009

- *High-Accuracy Object Recognition with a New Convolutional Net Architecture and Learning Algorithm*, Learning Workshop, Snowbird - Clearwater, April 2009

- *Unsupervised Learning of Sparse and Invariant Features Hierarchies*, Learning Workshop, Snowbird - Puerto Rico, March 2007

- *Efficient Learning of Sparse Representations with an Energy-Based Model*
  - Computational Vision Laboratory, California Institute of Technology, December 2006
  - Laboratory for Computational Vision, Center for Neuroscence - NYU, October 2006

- *Energy-Based Model for Unsupervised Learning of Sparse Overcomplete Representations*
  - Learning Workshop, Snowbird Utah, April 2006
  - CIAR NCAP Workshop, Vancouver Canada, December 2005

PATENTS

- A. Douillard, Q. Feng, A. Rusu, A. Kuncoro, Y. Donchev, R. Chhaparia, **M. Ranzato**, J. Shen, A. Szlam. Neural networks with distributed path composition. US Patent 18/885,242

- **M. Ranzato**, K.Q. Weinberger, E. Hoerster, M. Slaney. System and method for improved classification. US Patent 9,639,7802017

- J. Shlens, Q.V. Le, G.S. Corrado, **M. Ranzato**. Generating labeled images. US Patent 9,256,80712016

- A.W. Senior, **M. Ranzato**. Curriculum learning for speech recognition. US Patent 9,202,464

- **M. Ranzato**. Identifying objects in images. US Patent 9,129,190

- Y. Taigman, M. Yang, **M. Ranzato**. Systems and methods for facial representation. US Patent App. 14/530,585

- P. Perona, **M. Ranzato**, R. Flagan. Automatic visual recognition of biological particles. US Patent App. 11/122,575 US Patent 9,202,464