Mohamed Ashraf Abdelsalam

Toronto, Canada
✉ mabdelsalam944@gmail.com
🌐 mohamedabdelsalam9.github.io
in linkedin.com/in/mabdelsalam9

Education

2019–2021   **MSc in Machine Learning**, *Mila – Quebec AI Institute, Université de Montréal*, Canada
      Adviser   Sarath Chandar
      Thesis   Incremental Implicitly-Refined Classification for Lifelong Learning

2013–2018   **BSc in Aerospace and Communication**, *Zewail University of Science and Technology*, Egypt
      Adviser   Elsayed Hemayed
      Thesis   Attribute-based Face Generation Using Progressive GANs

Experience

July 2021 – Present   **ML Research Engineer → Senior ML Research Engineer**, *Samsung AI Center*, Toronto, Canada
Led and contributed to key initiatives in multimodal AI research and development.

Image Gallery Search   Multimodal Retrieval, Vision-Language Models, Large Language Models
- Designed a novel framework for situation-aware text-to-image retrieval, leveraging contextual cues from neighboring images in structured personal galleries to enable context-driven search beyond explicitly describing visual elements in target photos.
- Created a scalable annotation pipeline utilizing Vision-Language Models (VLMs) and Large Language Models (LLMs), enabling efficient generation of a large-scale situational retrieval dataset. (Manuscript in Progress)
- Led the creation of a modular benchmarking repository for text-to-image retrieval models, streamlining evaluation and accelerating iterative improvements in gallery search capabilities.

Structured Image Captioning   Vision-Language Models, Semantic Parsing, Controllable Image Captioning
- Proposed *Visual Abstract Meaning Representation Graphs (Visual AMRs)*, a linguistically informed alternative to scene graphs, for abstracting image content into structured representations capable of capturing high-level semantics. (CoNLL 2022)
- Introduced *Structured Semantic Augmentation*, leveraging Visual AMRs to diversify and condition captions based on user controls, setting new benchmarks in caption diversity and controllability. (ECCV 2024)

Procedural Video Understanding   Video-Language Understanding, Zero-shot Transfer, Generative Modeling
- Developed *GePSAn*, a generative model for anticipating next steps in procedural activities (e.g., cooking videos), leveraging text-based procedural corpora for pretraining and transferring knowledge to videos in a zero-shot manner. Achieved state-of-the-art results on YouCookII, addressing challenges in modeling diverse and plausible future steps. (ICCV 2023)

2019 – 2021   **Postgraduate Research Assistant**, *Mila – Quebec AI Institute, Université de Montréal*, Canada

Incremental and Lifelong Learning, Image Classification, Natural Language Generation
- Proposed the *IIRC setup*, a benchmark for evaluating lifelong learning models in dynamic scenarios with hierarchical labels (e.g., "bear" and "polar bear"), requiring models to refine knowledge and deduce relationships incrementally. (CVPR 2021)
- Demonstrated that incorporating a semantic loss objective during training improves diversity in dialogue response generation, particularly for smaller datasets. Evaluated the effectiveness of large language model embeddings for semantic loss objectives. (SIGDial 2021)

May 2019 – Sep 2019   **Machine Learning Intern**, *National Bank of Canada*, Montreal, Canada

Anomaly Detection, Large-scale Data Processing (Apache Spark, Hadoop)
- Implemented learning based techniques for anomaly detection in a large distributed dataset of SQL queries

June 2016 – Sep 2016   **Undergraduate Research Intern**, *ETH Zurich*, Switzerland

Super-Resolution, Dictionary Learning, Sparse Coding
- Implemented and evaluated image super-resolution methods using learned dictionaries and sparse coding techniques, supervised by Radu Timofte.

Publications

ECCV 2024   Controllable Image Captioning with Structured Semantic Augmentation.
Kalliopi Basioti, **M**ohamed A. Abdelsalam, Federico Fancellu, Vladimir Pavlovic, Afsaneh Fazly

ICCV 2023	GePSAn: Generative Procedure Step Anticipation in Cooking Videos. **M**ohamed A. Abdelsalam, Samrudhdhi Rangrej, Isma Hadji, Nikita Dvornik, Konstantinos Derpanis, Afsaneh Fazly
CoNLL 2022	Visual Semantic Parsing: From Images to Abstract Meaning Representation. **M**ohamed A. Abdelsalam, Zhan Shi, Federico Fancellu, Kalliopi Basioti, Dhaivat Bhatt, Vladimir Pavlovic, Afsaneh Fazly
CVPR 2021	IIRC: Incremental Implicitly-Refined Classification. **M**ohamed A. Abdelsalam, Mojtaba Faramarzi, Shagun Sodhani, Sarath Chandar
SIGDial 2021	A Brief Study on the Effects of Training Dialogue Models with a Semantic loss. Prasanna Parthasarathi*, **M**ohamed A. Abdelsalam*, Joelle Pineau, Sarath Chandar

Primers

arXiv 2022	An Introduction to Lifelong Supervised Learning. Shagun Sodhani, Mojtaba Faramarzi, Sanket Vaibhav Mehta, Pranshu Malviya, **M**ohamed A. Abdelsalam, Janarthanan Janarthanan, Sarath Chandar

Languages

Arabic (Native), English (Fluent), French (Intermediate)