

Mohamed H. Elhoseiny

*PhD Student–Fall-2011-present, Rutgers University,
Department of Computer Science*

A: 110 Frelinghuysen Rd
Piscataway, NJ, USA, 08854
E: m.elhoseiny@cs.rutgers.edu

W: paul.rutgers.edu/~mhe19 (include details)

Research interests: Language & Vision, Computer Vision, Deep Learning, Machine Learning, and MultiMedia

Profile

- Involvement in Artificial Intelligence and Computer Vision communities as a reviewer and workshop organizer (e.g., I organized the first workshop on “Closing the Loop Between Vision and Language” in ICCV 2015, Chile)
- Research and Hands-on Experience on several projects in Computer Vision, Vision & Language, Deep Learning., Machine Learning, Multimedia, Natural Language Processing.
- Mentorship and Participation on various Artificial Intelligence Projects.
- Eight years of industry experience as a Research and Development Software Engineer.
- Extensive programming, problem solving, integration and multidisciplinary engineering skills.

Education

PhD in Computer Science, Rutgers University – Supervisor: Prof. Ahmed Elgammal –**2016** expected– GPA 4.00

Thesis: My PhD thesis in the intersection between Vision and Language.

MSc in Computer Science, Rutgers University – Supervisor: Prof. Ahmed Elgammal –**2014** –GPA 4.0.

Report: “Write a Classifier: Zero Shot learning using Purely Textual Descriptions”, published at ICCV’13.

MSc in Computer Systems, Ain Shams University (ASU) – Supervisor: Prof. Taymour Nazmy – **2010** – courses GPA 3.55

Thesis: “High Performance Activity Monitoring for Scenes including Multiple Agents”, – research GPA 4.00

BSc in Computer Science, Ain Shams University (ASU) – Supervisor: Prof. Said ElGhoneimy – **2006** – GPA 3.96.

Thesis: “Compiler Construction ToolKit”

Employment

Research Intern at Computer Vision Group, Adobe Research, San Jose–under: Dr. Scott Cohen and Dr. Brian Price – **Summer 2015–Now**

Research Intern at Vision&Learning Group, SRI International, Princeton–under: Dr. Hui Cheng and Dr. Harpreet Sawhney – **Summer 2014**

Computer Vision Research Intern at IPPLEX Corporation, LA, CA– under: Dr. Jeremi Sudol and David McKinnon– **Summer 2012**

Research & Development Associate at Future Business Development, Cairo, Egypt– under Dr. Adnan Faqeeh – **Summer 2011**

Senior R&D Software Engineer (part-time) at House of Invention (HOI) GROUP, Cairo, Egypt– **2007-2011**

Undergrad-Research Mentor & Teaching Assistant at Computer Science Department, ASU, Cairo, Egypt–**2007-2011**

Honors and Awards

- Best Research-Intern Award (silver), Vision and Learning Group, Center of Vision Technologies, SRI International, Summer 2014
- CVPR16 Doctorial Consortium award, ICLR 2016 Travel award.
- Nominated by the department of Computer Science at Rutgers University for Google and Microsoft fellowships 2014.
- 8/14 of the projects, I mentored from 2007 to 2011, received awards from Information Technology Industry Development Agency (ITIDA), Egypt.
- Research-MSc fellowship by the Faculty of Computer and Information Sciences, Ain Shams university , 2007-2010.
- Ranked 3rd in the Computer Systems Bachelor Program and received the Honor degree, as indicated in my transcripts, in 2006.
- Ranked 1st in my High School class, in 2002 and received an honor-shield for my rank.
- Won Travel awards for my ICIP publications by Rutgers University 2013 and 2014, also for ICCV 2013 from PAMI-TC.

Research

Publications and Patents

Computer Vision, Multi-view learning, Vision&Language

1. Mohamed Elhoseiny, Tarek-El-Gaaly, Amr Bakry, Ahmed Elgammal, “A Comparative Analysis and Study of Multiview Convolutional Neural Network Models for Joint Object Categorization and Pose Estimation”, **ICML**, 2016, acceptance rate 24.5%.
2. * **Means Co-first authors** Amr Bakry*, Mohamed Elhoseiny*, Tarek-El-Gaaly, * Ahmed Elgammal, “Digging Deep into the layers of CNNs: In Search of How CNNs Achieve View Invariance”, **ICLR**, 2016 (average review scores is 7/10), acceptance rate 25%.
3. Mohamed Elhoseiny, Tarek-El-Gaaly, Amr Bakry, Ahmed Elgammal, “Convolutional Models for Joint Object Categorization and Pose Estimation”, (ICLR16 Workshop presentation)
4. Mohamed Elhoseiny, Jingen Liu, Hui Cheng, Harpreet Sawhney, Ahmed Elgammal, “Zero Shot Event Detection by Multimodal Distributional Semantic Embedding of Videos”, **AAAI**, 2016 (oral), acceptance rate 25.6%.
5. Han Zhang, Tao Xu, Mohamed Elhoseiny, Xiaolei Huang, Shaoting Zhang, Ahmed Elgammal, Dimitris Metaxas, “SPDA-CNN: Unifying Semantic Part Detection and Abstraction for Fine-grained Recognition”, **CVPR**, 2016 , acceptance rate 29.9%.
6. Amr Bakry, Tarek El-Gaaly, Mohamed Elhoseiny, Ahmed Elgammal, “Joint Object Recognition and Pose Estimation using a Nonlinear View-invariant Latent Generative Mode”, **WACV**, 2016 (to appear), algorithms track acceptance rate 30%.

7. Mohamed Elhoseiny, Ahmed Elgammal, "Overlapping Domain Cover for Scalable and Accurate Kernel Regression Machines", **BMVC (oral)**, 2015, **oral** acceptance rate **7%**.
8. Sheng Huang, Mohamed Elhoseiny, Ahmed Elgammal, "Learning Hypergraph-regularized Attribute Predictors", **CVPR**, 2015 (28.4%)
9. Mohamed Elhoseiny, Ahmed Elgammal, "Generalized Twin Gaussian Processes using Sharma-Mittal Divergence", the "**Machine Learning journal**", **ECML-PKDD** special issue, 2015, also **orally** presented in **ECML-PKDD** conference, ~10% acceptance rate
10. Mohamed Elhoseiny, Sheng Huang, Ahmed Elgammal, "Weather Classification with deep Convolutional Neural Networks", **ICIP**, 2015 (**oral**).
11. Mohamed Elhoseiny, Ahmed Elgammal, Babak Saleh, "Tell and Predict: Kernel Classifier Prediction for Unseen Visual Classes from Unstructured Text Descriptions", Arxiv, 2015. presented in **CVPR15** Workshop on Language and Vision and **EMNLP15** workshop on Language and Vision. This work together with my ICCV13 work +Additional experiments is consolidated in my TPAMI submission.
12. Sheng Huang, Mohamed Elhoseiny, Ahmed Elgammal, "Improving Non-Negative Matrix Factorization via Ranking Its Bases", **ICIP**, 2014. (44%)
13. Mohamed Elhoseiny, Babak Saleh, Ahmed Elgammal, "Write a Classifier: Zero Shot Learning Using Purely Textual Descriptions", **ICCV**, 2013 (27.8%)
14. Mohamed Elhoseiny, Bing Song, Jeremi Sudol, David McKinnon. "Low-Bitrate Benefits of JPEG Compression on SIFT Recognition", **ICIP**, 2013 (44%).
15. Mohamed Elhoseiny, Amr Bakry, Ahmed Elgammal., "MultiClass Object Classification in Video Surveillance Systems -Experimental Study", SISM Workshop, **CVPR**, 2013 (oral, [demo](#)).
16. Mohamed Elhoseiny, Babak Saleh, Ahmed Elgammal, "Heterogeneous Domain Adaptation: Learning Visual Classifiers from Textual Description", VisDA Workshop, **ICCV**, 2013 (presentation)
17. Marwa Abdelmonem, Mohamed Elhoseiny, Asmaa Ali, Karim Emara, Habiba Abdel Hafez, Asmaa Gamal, "Dynamic Optical Braille Recognition (OBR) System". **IPCV**, Las Vegas, Nevada, USA, 2009

Natural Language Processing and MultiMedia

1. Mohamed Elhoseiny, Ahmed Elgammal, "Text 2 Multilevel MindMap: A New Way for Interactive Visualization and Summarization of Natural Language Text, Springer International Journal of Multimedia Tools and Applications (**MTAP**), 2015
2. Mohamed Elhoseiny, Ahmed Elgammal, "English2MindMap: Automated system for Mind Map generation from Text", International Symposium on Multimedia (**ISM**), 2012 (**oral presentation**)
3. Asmaa Hamdy, Mohamed H. Elhoseiny, Radwa Elsahn, Eslam Kamal, "Mind Map Automation (MMA) System". **SWWS**, Las Vegas, Nevada, USA, 2009

Action Recognition and GPU

1. Mohamed H. Elhoseiny, "High Performance Activity Monitoring for Scenes including Multi-Agents ". MSc thesis, Faculty of Computer and Information Sciences, **Ain Shams University**, July 2010.
2. Mohamed H. Elhoseiny, H.M.Faheem, Eman Shaaban, T.M. Nazmy, "GPU framework for teamwork Action Recognition". Faculty of Computer and Information Sciences, **Ain Shams University**, 2010

Arxiv papers / submitted articles

2015-present

1. Mohamed Elhoseiny, Scott Cohen, Walter Chang, Brian Price, Ahmed Elgammal, "Sherlock: Scalable Fact Learning in Images", Arxiv, 2015
2. Mohamed Elhoseiny, Ahmed Elgammal, Babak Saleh, "Write a Classifier: Predicting Visual Classifiers from Unstructured Text Descriptions", (submitted to **TPAMI**), 31 Dec 2015

Talks and Presentations

- Talk: "Teaching Machines to See with Language", **UC Berkeley**, Berkeley, CA April 2016
- Talk: "Zero Shot Event Detection by Multimodal Distributional Semantic Embedding of Videos" at **AAAI**, Pheinx, AZ Feb. 2016
- Talk: "Weather Classification with deep Convolutional Neural Networks" at **ICIP** 2015, Monterial, Canada Sept. 2015
- Talk: "Overlapping Domain Cover for Scalable and Accurate Kernel Regression Machines" at **BMVC** 2015, Swansea, UK Sept. 2015
- Talk: "Generalized Twin Gaussian Processes using Sharma-Mittal Divergence" at **ECML-PKDD**, Portugal, Portugal Sept. 2015
- Poster: "Tell and Predict: Kernel Classifier Prediction for Visual Classes from Text Descriptions" **CVPR W**, Boston, MA June 2015
- Poster: "Learning Hypergraph-regularized Attribute Predictors", at **CVPR W**, Boston, MA June. 2015
- Poster: "Write a Classifier: Zero Shot Learning Using Purely Textual Descriptions", at **ICCV**, Sydney, AU Dec. 2013
- Poster: "Heterogeneous Domain Adaptation", at **ICCV W**, Sydney, AU Dec. 2013

Service

- Workshop Organizer for the "Closing the Loop between Vision and Language", ICCV, 2015
- Reviewer for TPAMI, TIP, CVIU, TMM, MVA journals
- Reviewer for BMVCW
- Reviewer for CVPR, ICCV, ECCV papers (2012-Present) (external)

Work Experience

Computer Vision Research Intern (on Language&Vision) Adobe Research, San Jose, CA

Summer 2015

In this position, I worked on defining a novel method for gaining visual knowledge in a structured way. We name our setting "Sherlock Mohamed Elhoseiny

Problem” of computer vision. I also proposed a deep Learning representation method to gain such structured visual knowledge. The outcome of this project was submitted to one of the premier conferences on Machine Learning.

Computer Vision Research Intern (on Language&Vision) SRI International, Princeton, NJ Summer 2014

In this position, I was involved in designing a new ZSL method adopting on Google word2vec for Multimedia Event Detection (MED). I ended up with a novel Zero-Shot Event Detection method that outperforms the state-of-art based on distributional semantic embedding of multimodal information in videos (e.g. Actions, objects, ASR, OCR). I won Best Intern Award (Silver) for this work from SRI. The outcome of this internship got published in AAAI 2016.

Computer Vision Research Intern IPPLEX Corporation, Los Angeles, CA Summer 2012

I completely designed and implemented the following two projects during this internship.

- *Analytical study of SIFT Matching on various parameter settings:* resulted in my ICIP13 publication.
Tools: C++, Python, Evolution Robotics Libraries
- *Discriminative Image Classification: Co-segmentation Approach:* is a Graph-Cut approach for discriminative image classification.
Tools: C++, Python, OpenCV, Evolution Robotics Libraries

Research and Development (R&D) Associate FBD (Future Business Development), Egypt Summer 2011

- I solely built a MATLAB module for stock market prediction and visualization and I was accepted in this internship due to satisfaction of a proposed solution for the problem, that I submitted as a part of my application to this position.
Tools: MATLAB, C#, C++

Research Assistant Computer Science Dept., Rutgers University, NJ, USA 2014-present

Teaching Assistant Computer Science Dept., Rutgers University, NJ, USA 2011-2014

- CS110 Introduction to Computers and their applications (1 semester[sem]) • CS314 Principles of Programming Languages (4 sems)

Undergrad-Research Mentor& Teaching Assistant Computer Science Dept, ASU, Cairo, Egypt 2007-2011

Undergrad-Research Mentorship role

I participated in 14 Computer Vision and Natural Language Processing projects under this role, listed below. Eight of these project received ITIDA awards ([link](#)).

Computer Vision Projects

- Stereo-camera based object recognition system, 2011
- Intelligent Presentation Guru(IPG), 2010 ([demo](#)).
- Dynamic Optical Braille Recognition System, 2008 ([demo](#)).
- Auto Spelling Robot , 2007
- Intelligent Home Security system, 2011
- On-Chip Action Recognition System, 2010
- Medical Image Registration system (MIRS), 2008
- Smart class Room System (SMRS), 2010
- Gait Analysis for Human Identification (GAHI), 2009 ([demo](#)).
- Automatic Panoramic Image Stitching Engine (APISE), 2008

Natural Language Processing and Multimedia Projects

- AJEB (Question answering system for Arabic text), 2011
- Single level Mind-Map Generator ([demo](#)), 2009
- Story Building System (Text to Animation), 2011
- Intelligent Information Discovery & Processing System (IIDPS), 2009

Teaching Assistant role

- Peripheral Interfaces (1 sems)
- Distributed Processing (4 sems)
- Object Oriented Design Patterns (3 sems)
- Introduction to algorithms (3 sems)
- Computer Architecture (4 sems)
- Digital Signal Processing Applied (4 sems)
- Numerical computing methods (1 sem)

Senior R&D Software Engineer (part-time) House of Invention (HOI) GROUP, Cairo, Egypt 2007-2011

The following table describes the projects I worked on, my role and the tools used in each of them.

Project Description	My role	Tools
<i>Medical Image Manipulator:</i> is product that performs advanced image filtering and transformation tasks on medical images (e.g. chromosome segmentation for karyotyping automation).	Complete design and implementation.	C/C++ Delphi
<i>Softoria Poles (CAD System):</i> is a big software product, used by civil engineers, to perform structural design and analysis of lighting columns and signboards. The project is decomposed into many modules including safety, structural design, Windshield, and Graphics modules.	Software design and implementation of safety, structural design, Windshield and Graphics modules, which involves solution of computational geometry, numerical integration and computer graphics problems.	C/C++ C# OpenGL
<i>Softoria Monopole shop drawing (CAD System):</i> is a product that automates Shop Drawing, previously done manually by engineers.	Software design and implementation of the Graphics Module and the automatic generation of the shop	

The product gets monopole specs from the civil engineers and generates AUTOCAD shop drawings that can be directly used for monopole manufacturing monopole structures.	drawings of Openings, Antennas, and Antenna Platforms.	
<i>Softoria PalmTree Analyzer (CAD System)</i> : is a product whose function is to take specs of an artificial palm tree, used as a communication Tower. Then, it analyzes wind and weight forces on it considering wind shielding. Simultaneously a 3D model of the palm tree is rendered, where loads and analysis result are visualized on the rendered model.	Complete design and implementation.	

Software and HW Skills

- **Programming Languages** : C/C++ (9 yrs), C# (9 yrs), Python (5 yrs), Java(3 yrs), Assembly (3 yrs), Embedded C/C++ (5 yr), Delphi (5 yr), Prolog (5 yrs).
- **Proficient Programming Knowledge of MATLAB** (8 yrs), **MAPLE** (5 yrs), **Mathematica** (5 yrs)
- **Deep Learning** : Caffe and Overfeat
- **Optimization Tools** : CPLEX, NITRO, MOSEK, CVX, YAMLP, SEDUMI, GUROBI
- **Computer Vision libraries** : OpenCV, HTK, VTK, VLFEAT, FLANN
- **Machine Learning tools** : LIBSVM, LIBLINEAR, ANN.Net Library, HMM library, SVM.Net.
- **Embedded Vision** : LeanXCAM (standalone camera with DSP).
- **Parallel Programming and GPU** : MPI, OpenMP, NVidia CUDA SDK, OpenCL.
- **Computer Graphics** : OpenGL, DirectX.
- **GUI Tools** : MFC VC++, QT C++, Eclipse, And Visual Studio.Net.
- **Design Patterns** : MVC, Singleton, Factory, Command, Observer, Composite, Decorator, etc.
- **Web, DB and System Analysis Tools** : MS SQL, Oracle 10g, Rational Rose, ASP.Net.
- **Networking and Messaging** : Socket Programming, .Net Remoting, Message pipes.
- **OS Platforms** : Ubuntu, MacOS X, Windows
- **Programmable platforms** : FPGA, VHDL, DSPs, 8085, 8086, 8051, AT89C51.

PERSONAL INFORMATION & EXTRACURRICULAR ACTIVITIES

Languages : English (fluent), Arabic (fluent), French (fair)
Hobbies : Reading, Body Building, Playing soccer.
Undergrad Activities : ACM contests, building free DB systems for benevolent entities.
Postgrad Activities :

- Mentoring of Bachelor Projects 2007-2011
- Creation of CMS for Dept of Computer Science, ASU, based on Moodle to manage the courses 2010
- Participation in Quality Assurance and Accreditation Project for Faculty of Computer and Information Sciences (FCIS), ASU, in 2007.
- Participation in Continuous Improvement and Qualifying for Accreditation Project for FCIS, ASU, 2008-2011

References

Prof. Ahmed Elgammal, Professor of Computer Science at Rutgers University, elgammal@cs.rutgers.edu

Dr. Hui Cheng, Program Director at SRI International, hui.cheng@sri.com

Dr. Scott Cohen, Principal Scientist at Adobe Research, scohen@adobe.com

Dr. Jingen Liu, Senior Research Scientist at SRI International, jingen.liu@sri.com

Prof. Harpreet Sawhney, SRI fellow and CTO of the center of Vision Technologies at SRI International, harpreet.sawhney@sri.com