

Yasser Hifny Abdelhaleem Essa

Email: yhifny@yahoo.com

OBJECTIVE

I am keen to embark on a full time career within the field of speech and language processing. My main interests are related to speech recognition and synthesis, natural language processing, statistical pattern recognition, and information theory. The long term goal of my research is to develop solid and practical solutions for the human machine interface.

EDUCATION

- Ph.D., Computer Science, March 2007.
Thesis: Conditional Random Fields for Continuous Speech Recognition.
University of Sheffield, UK.
- M.S., Electronics and Communication Engineering, November 2000.
Thesis: Online Arabic Handwriting Recognition.
University of Cairo, Egypt.
- B.S., Electronics and Communication Engineering, June 1995.
University of Cairo, Egypt.

RESEARCH AND EXPERIENCE

Professor

**University of Helwan
Egypt**

August 2022 – Present

I am an academic staff member at the faculty of computers and artificial intelligence. I am working on the following projects:

1. 140-Online virtual assistance using large-scale Arabic speech recognition.
2. Medical code prediction using deep learning methods.
3. Medical data anonymization using NLP methods.
4. Arabic diacritics restoration using Maximum Entropy language models (done).
5. Retriever augmented generation (RAG) for Quran.
6. Expanding Arabic thesauruses by examples using dense retriever methods.
7. Ibn-Ginni: an improved morphological analyzer for Arabic (done).
8. Arabic word sense disambiguation using sensegram (done).
9. ALNASIKH: an Arabic OCR system based on transformers (done).
10. Arabic spelling error correction using machine learning methods.
11. Autism detection using deep learning methods.

Senior Research Scientist

Cerence

Montreal, Canada until December 2020;
Remote from Cairo, Egypt since January 2021

January 2020 – August 2024

I was a Natural Language Understanding (NLU) R&D group member. I worked on the following projects:

1. Semantic parsing using the transformer model.
2. General knowledge question answering system.
3. Sequence to sequence rejection modeling using white lists.
4. Domain arbitration using LSTMs.
5. Car manual question answering.
6. Question answering using LLMs.

7. Speech recognition using pre-trained models.
8. Speech recognition using advanced fine-tuning methods.
9. Speech recognition error correction model using transformers.
10. Semantic parsing using LLMs and Lora fine-tuning.
11. Retriever augmented rescoring for speech recognition error correction.

Associate Professor

**University of Helwan
Egypt**

November 2016 – July 2022

I was an academic staff member at the faculty of computers and artificial intelligence. I was on leave from 2016 until 2020. I was working on the following projects:

1. Arabic syntactic diacritics restoration using BERT models (done).
2. Efficient Arabic emotion recognition using deep neural networks (done).
3. Open vocabulary Arabic diacritics restoration (done).
4. Arabic code-switching speech recognition using monolingual data (done).
5. Detection of distributed denial of service attacks in VoIP networks using deep learning methods (done).

Senior Research Scientist
Montreal, Canada

**Nuance Communications, Inc
August 2013 – December 2019**

I was a member of langdev acoustic modeling group where I developed state of the art deep acoustic models (from August 2013 to April 2014, I worked remotely from Egypt). Since April 2016, I am a member of the Natural Language Understanding (NLU) R&D group. I worked for the following projects:

1. Monolingual acoustic modeling using deep neural networks (several languages).
2. Multilingual acoustic modeling using deep neural networks.
3. Stacked deep neural networks acoustic modeling.
4. Efficient deep neural networks training using RMSPROP.
5. Large margin estimation of CNNs for sentence classification.
6. Neural medical document classification using Keras.
7. Medical diagnosis justification using Keras.
8. Medical document classification using multitask learning.

Assistant Professor

**University of Helwan
Egypt**

August 2009 – October 2016

I was an academic staff member at the Faculty of Computer Sciences and Information. I was on leave from 2013 until 2016.

*Post-Doc Fellow (speech to speech
translation group)*

**IBM T.J. Watson Research Center
NY, USA**

April 2007 – May 2008

I was a member of the speech to speech translation project over Skype channels. Moreover, I introduced the Trusted Expectation-Maximization (TEM), a new discriminative training scheme, for pattern recognition applications. In particular, the TEM algorithm may be used for Hidden Markov Models (HMMs) based discriminative training.

*PhD student,
Speech and Hearing (SpandH) Group*

**University of Sheffield
Sheffield, UK
Feb. 2002 — Dec. 2006**

Based on the Maximum Entropy (MaxEnt) principle, the thesis aims to provide a non parametric framework for the acoustic modeling problem in continuous speech recognition systems.

- Developed Augmented Conditional Random Fields (ACRFs) based acoustic models that outperformed most published results on the TIMIT database.
- Developed MaxEnt Kernel machine that showed similar performance to Support Vector Machine (SVMs).
- Developed discriminant feature selection algorithm based on the MaxEnt principle.

*Speech Department Manager,
RDI speech Group*

**Research & Development
International (RDI)
Egypt
Feb.2000 — Jan. 2002**

Managed projects in Text To Speech (TTS), speech recognition, limited domain speech compression, and speech verification. Designed and developed different algorithms for RDI ArabTalk Text To Speech system.

- Developed the timing stage for the TTS system, which predicted the duration of the target phonemes for the input text. Phonology to acoustical modeling was achieved via a BP neural network. Moreover, the same model was used to predict the energy for the incoming phoneme sequence.
- Developed a model to predict and generate the global intonation contour that enhanced the quality of synthesized speech. A recurrent neural network was used to model the phonology to acoustical process.
- Implemented a real time unit selection algorithm, which handled certain properties of the Arabic language. This algorithm searched four hours of speech data to select the acoustic units that match the input targets.
- Designed and supervised a speech compression system, which utilized linguistic information from limited domain speech data.
- Analyzed requirements of a limited domain pronunciation verification system. Managed a team of engineers and helped in the design of the system, which used speech recognition technology, to discover the recitation errors of any reader.

*Senior Software Engineer,
Software development division*

**ISIS "Information Systems
International Services" Software
Egypt
Sep.1998 – Jan. 2000**

Participated in the analysis, design, and development of international projects. Contributed in writing system functionality, design, and functionality verification documents. I also wrote the company software protection engine. Development Environments: Java/SAS/C++/Delphi.

Software Engineer

**Motahida Software
Egypt
March 1997 – Aug. 1998**

Participated in the analysis, design, and development of Arabic based multimedia programs and generic interactive tutorial applications. Development Environments: C++/Delphi.

PROFESSIONAL
ACTIVITIES

Reviewed papers for ICASSP 2005, IEEE Signal Processing Letters journal, and IEEE Transactions on Audio, Speech and Language Processing.

AWARDS	<ul style="list-style-type: none"> • Parmajan challenge award for Arabic language processing 2023. • Parmajan challenge award for Arabic language processing 2022. • Motorola studentship for PhD study, 2002. • The software award (1000 L.E.) from the National Arabic Authority for manufacturing, 2001. • National research center scholarship for postgraduate study, 1996. • Excellence awards from the University of Cairo in 1991, 1994, and 1995.
COMPUTER SKILLS	<ul style="list-style-type: none"> • OS: Linux, MS Windows; Programming languages: C++ (g++, visual c++, Borland), Java, Unix shell, Python, Delphi; Applications: Paradox, MS Visio, MS Office, Matlab, L^AT_EX.
TEACHING EXPERIENCE	<ul style="list-style-type: none"> • Assistant professor/associate professor/ professor (2009-2024) at the University of Helwan. Lectured courses on pattern recognition, digital signal processing, speech processing, computer organization, computer networking I, computer networking II, Python programming, multimedia, and C programming. • Guest professor (2009) at the Arab Academy for Science, Technology and Maritime Transport (College of Computing and Information Technology). Lectured a course on statistical machine learning. • Teaching assistant (2003-2006) at the University of Sheffield. Helped students with their problems in tutorial and laboratory sessions. In addition, group projects were supervised. • Lectured a course on <i>Programming Windows</i> at the University of Cairo in the summer of 1995.
PUBLICATIONS	See Google scholar page.
REFERENCES	<p>Prof. Mohsen Rashwan University of Cairo Cairo, Egypt mohsen_rashwan@rdi-eg.com</p> <p>Dr. Mohamed Afify Microsoft Research Cairo, Egypt mohamed_afify2001@yahoo.com</p>