

Wenda Qin

+1 (617)-383-1146 • wdqin@bu.edu • Department of Computer Science, Boston University, Boston MA 02215

RESEARCH INTERESTS

- **Computer Vision**
 - **Natural Language Processing**
 - **Vision-and-language Multimodal Understanding**
-

EDUCATION

- **Boston University, Boston, MA** Sept. 2018 - present
GPA: 4.00
Image & Video Computing (IVC) Group
Ph.D. - Computer Science, advised by Margrit Betke
 - **Boston University, Boston, MA** Sept. 2016 - Jul. 2018
GPA: 3.64
M.Sc. - Computer Science
 - **South China University of Technology, Guangdong, China** Sept. 2012 - Jul. 2016
GPA: 3.62
B.E. - Software Engineering
-

PUBLICATIONS

- **“Extracting text from scanned Arabic books: a large-scale benchmark dataset and a fine-tuned Faster-R-CNN model”** Randa Elanwar, **Wenda Qin**, Margrit Betke, Derry Wijaya. International Journal on Document Analysis and Recognition (IJ DAR), 2021, 14 pages.
 - **“Text and metadata extraction from scanned Arabic documents using support vector machines”** **Wenda Qin**, Randa Elanwar, and Margrit Betke. Journal of Information Science, 2020, 12 pages.
 - **“Extracting text from scanned Arabic books: A large-scale benchmark dataset and a fined-tuned Faster-R-CNN model”**
Randa Elanwar, **Wenda Qin**, and Margrit Betke. Submitted to the International Journal on Document Analysis and Recognition (IJ DAR), 2020, 10 pages.
 - **“LAL: Linguistically Aware Learning for Scene Text Recognition.”**
Yi Zheng, **Wenda Qin**, Derry Wijaya, and Margrit Betke. ACM Multimedia Conference (ACM 2020). 2020, 9 pages.
 - **“Project SEARCH (Scanning EARs for Child Health): Validating an Ear Biometric Tool for Patient Identification in Zambia”**
Lauren Etter, Alinani Simukanga, **Wenda Qin**, Rachel Pieciak, Chris Gill, Lawrence Mwananyanda, Caroline Carbo, Margrit Betke, Jackson Phiri. Gates Open Research, 2020, 14 pages.
 - **“Making Scanned Arabic Documents Machine-accessible Using an Ensemble of SVM Classifiers”**
Randa Elanwar, **Wenda Qin**, and Margrit Betke. International Journal on Document Analysis and Recognition (IJ DAR), 2017, 25 pages
 - **“LABA: Logical Layout Analysis of Book Page Images in Arabic Using Multiple Support Vector Machines”**
Wenda Qin, Randa Elanwar, and Margrit Betke. 2nd IEEE Int. Workshop on Arabic and derived Script Analysis and Recognition (ASAR 2018). 2017, 5 pages.
 - **“Analysis of Stanford University’s Human-computer Interaction (HCI) Course System”**
Liannan Lin, Chuan Long, and **Wenda Qin**. 11th International Conference on Computer Science & Education 2016 (ICCSE 2016). 2016, 6 pages.
-

RESEARCH EXPERIENCE

- **Emotion Analysis on News Title and Images of Gun Violence** Jul. 2021 - present
Collaborators: Sejin Paik, Ge Gao, Carley Reardon, Derry Wijaya, Margrit Betke, Lei Guo
Proposed a dataset of reader emotions on titles and images of gun violence news. Developed baseline models

to analyze the current challenges of vision-and-language models to predict and reason readers' emotions.

- **Vision-and-language Navigation** Jan. 2021 - present
Collaborators: Teruhisa Misu, Derry Wijaya
Proposed a new mechanic to allow the navigation model explainable on its focus of words in the instruction. Utilized a set of snapshots obtained during the training of the navigation model to achieve a new state-of-the-art performance in the R2R dataset.
- **Text Recognition in Natural Scene** Feb. 2020 - Nov. 2020
Collaborators: Yi Zheng, Derry Wijaya, Margrit Betke
Proposed a scene text recognition system that utilizes visual and linguistic features from the given image and a large text corpus.
- **Ear Study of Infants in Zambia** Jan. 2019 - present
Collaborators: School of Public Health at Boston University, Department of Computer Science at University of Zambia
Built an ear matching system based on traditional and deep learning algorithms/models. Then deployed it on the Android phone. Evaluated its accuracy with volunteers in Boston and Lusaka, Zambia.
- **Arabic Document Analysis** Jun. 2017 - Jul. 2021
Collaborators: Randa Elanwar, Margrit Betke
Proposed a layout analysis system that detects, classifies, and segments Arabic document pages into text/non-text areas. Built an Arabic-specific document image dataset for evaluation.

INDUSTRY EXPERIENCE

- **Honda Research Institute USA, Inc.** Jan. 2021 - May. 2021
Research Intern
 - Built and analyzed the existing Vision-and-Language Navigation (VLN) models on interior environments navigation, i.e., the R2R dataset.
 - Developed a more explainable model to understand the connections between AI's predictions and textual instructions.
 - Explored different approaches to improve existing models. Achieved a new SOTA performance based on a novel observation of the model behaviors in the past.
- **Guangzhou Wu Xiang Technology Co. Ltd** Jul. 2015 - Aug. 2015
Software Engineering Intern
 - Prototyped a mobile application for team efficiency management using Axure.
 - Implemented a server interface for data transfer of the app in C# and ASP.Net, then deployed it on a cloud server platform.
 - Worked with team developers for data transfer between the client-side (mobile app) with the server.

SKILLS

- **Programming Languages:** Python, C++, Java, C#, Matlab, LaTeX
- **Deep Learning, ML and CV Tools:** Pytorch, Tensorflow, OpenCV, Scipy, Scikit-learn, OpenGL
- **Web Development:** HTML, Javascript

TEACHING EXPERIENCE

- **Boston University:**
 - 2018 & 2020 & 2021 Fall: CS 480/680 Computer Graphics, Teaching Assistant
 - 2019 Spring & Fall: CS 440/640 Artificial Intelligence, Teaching Assistant

PROFESSIONAL ACTIVITIES

- Organized layout analysis competition challenge: physical layout analysis of scanned Arabic books in 2nd International Workshop on Arabic and derived Script Analysis and Recognition (ASAR2018)

AWARDS

- University Second Class Scholarship, South China University of Technology
- Third Place in SCUT "Challenge Cup" Undergraduate Science and Technology Works