Hani Alomari

+1-540-449-7806 | hani@vt.edu | LinkedIn | Google Scholar

Blacksburg, Virginia - 24060, USA

Research Interests

I work at the intersection of computer vision and natural language processing, focusing on multimodal learning and vision-language models. My research centers on building robust, semantically aligned representations across modalities to support retrieval, reasoning, and interpretability. Specific interests:

- Cross-modal retrieval across images, text, video, and audio.
- Learning diverse and semantically meaningful embeddings for multimodal alignment.
- Structured information extraction and representation from multimodal data.
- Knowledge structures and reasoning in vision-language models.

Education

- **Virginia Tech** — *January 2023 - (Expected) January 2028*
  PhD in Computer Science - GPA: 4.00/4.00 - Advisor: Dr. Chris Thomas — Blacksburg, VA, USA
- **Jordan University of Science and Technology (JUST)** — *February 2020 - June 2022*
  M.S. in Data Science - GPA: 4.26/4.30 - Advisor: Prof. Rehab Duwairi — Irbid, Jordan
- **Jordan University of Science and Technology (JUST)** — *February 2016 - January 2020*
  B.S. in Computer Science - GPA: 4.04/4.20 - Advisor: Dr. Malak Abdullah — Irbid, Jordan

Publications

C=Conference, J=Journal, S=In Submission

[S.1] Hammad Ayyubi, Junzhang Liu, Ali Asgarov, Zaber Ibn Abdul Hakim, Najibul Haque Sarker, Zhecan Wang, Chia-Wei Tang, **Hani Alomari**, et al. **ENTER (v2): Event Based Interpretable Reasoning for VideoQA**, Under review AAAI 2026

[C.1] **Hani Alomari**, Anushka Sivakumar, Andrew Zhang, Chris Thomas. **Maximal Matching Matters: Preventing Representation Collapse for Robust Cross-Modal Retrieval**, *ACL 2025 Main*

[C.2] Md Atabuzzaman, Gino DiMatteo, **Hani Alomari**, et al. **Real-Time Ultra-Fine-Grained Surgical Instrument Classification**, *CVPR 2025 FGVC Workshop*

[C.3] Zhecan Wang, Junzhang Liu, Chia-Wei Tang, **Hani Al-Omari**, et al. **JourneyBench: A Challenging One-Stop Vision-Language Understanding Benchmark of Generated Images**. *NeurIPS 2024*.

[C.4] Hammad Ayyubi, Junzhang Liu, Zhecan Wang, **Hani Al-Omari**, et al. **ENTER: Event Based Interpretable Reasoning for VideoQA**. *NeurIPS 2024 MAR Workshop* - Spotlight -.

[C.5] **Hani Al-Omari**, Rehab Duwairi, et al. **DLJUST at SemEval-2021 Task 7: Hahackathon - Linking Humor and Offense**. SemEval-2021 workshop.

[C.6] **Hani Al-Omari**, Malak Abdullah, Samira Shaikh. **EmoDet2: Emotion Detection in English Textual Dialogue Using BERT and BiLSTM Models**. ICICS 2020, IEEE.

[C.7] **Hani Al-Omari**, Malak Abdullah, Ola AlTiti, Samira Shaikh. **JUSTDeep at NLP4IF 2019 Task 1: Propaganda Detection Using Ensemble Deep Learning Models**. NLP4IF 2019 workshop.

[C.8] Ayat Abedalla, Ali Fadel, Ibraheem Tuffaha, **Hani Al-Omari**, Mohammad Omari, Malak Abdullah, Mahmoud Al-Ayyoub. **MTRECS-DLT: Multi-Modal Transport Recommender System Using Deep Learning and Tree Models**. SNAMS 2019 IEEE.

[C.9] **Hani Al-Omari**, Malak Abdullah, Bassam Nabeel. **EmoDet at SemEval-2019 Task 3: Emotion Detection in Text Using Deep Learning**. SemEval-2019 workshop.

[J.1] **Hani Al-Omari**, Rehab Duwairi (2023). **So2al-wa-Gwab: A New Arabic Question-Answering Dataset Trained on Answer Extraction Models**. *TALLIP, ACM*.

Work Experience

- **Virginia Tech** *March 2023 - Present*
  Graduate Research Assistant - Multimedia Lab Blacksburg, VA, USA
  - Currently leading two ongoing projects: developing a deep learning framework for 3D room layout generation from Room Impulse Responses, and researching multi-manifold representation learning to enhance cross-modal alignment by preserving modality-specific semantics and shared information.
  - Developed novel Maximal Pair Assignment Similarity function with specialized loss functions to prevent set collapse and enhance semantic diversity, achieving state-of-the-art results (6.9% RSUM improvement).
  - Led the development of a dataset using Vision-LLM to generate dynamic, multi-perspective image descriptions, addressing MS-COCO limitations through automated caption generation.
  - Supported JourneyBench development through dataset annotation and conducted comprehensive benchmarking of cross-modal retrieval models on the new dataset.
  - Contributed to hospital medical image classification projects using advanced deep learning architectures.
- **Jordan University of Science and Technology** *February 2020 - November 2022*
  Graduate Research Assistant Irbid, Jordan
  - Designed and implemented machine learning and deep learning models, including transformer-based architectures for NLP and computer vision tasks.
- **Jordan University of Science and Technology** *June 2019 - September 2019*
  Research Assistant Internship Irbid, Jordan
  - Participated in competitive research projects such as NLP4IF 2019, focusing on deep learning for NLP.
  - Developed and optimized architectures, debugging and resolving issues to improve model performance.

Projects Experience

- **Medical Image Classification for Hospital Applications** *May 2024 - August 2024*
  Tools: Python, PyTorch, CNN Architectures, Medical Imaging Frameworks
  - Collaborated on projects utilizing state-of-the-art deep learning models for various medical tasks, including medical instruments classification and anomaly detection.
  - Enhanced accuracy and reliability of medical image classification systems through model fine-tuning and optimization.
- **So2al-wa-Gwab: Arabic Question Answering System** *Feb 2020 - Jun 2022*
  Python, TensorFlow, PyTorch, BERT, QANet, BiDAF
  - Created a novel Arabic QA dataset and benchmark, addressing key limitations in existing datasets including translation errors and context size.
  - Implemented and evaluated multiple deep learning architectures (BERT, BiDAF, QANet) across 7 Arabic QA datasets.
  - Demonstrated performance improvement using human-annotated data versus machine translation approaches.
- **EmoDet: Emotion Detection in Text** *February 2019 - January 2020*
  Tools: Python, PyTorch, GloVe, BERT, BiLSTM, Psycholinguistic Features
  - Developed a deep learning-based emotion detection system that combine different features from different sources, such as psycholinguistic features, pretrained word embedding features.
  - Achived 16% improvement in F1-score compared to baseline models.
- **Step Tracker Android Application** *March 2019 - May 2019*
  Android Studio, Kotlin, Firebase
  - Developed a step-tracking Android app using phone sensors, featuring real-time step counting and caloric tracking.
  - Utilized sensor data (accelerometer/gyroscope) to develop real-time tracking algorithms, demonstrating expertise in raw data preprocessing and time-series analysis.
  - Implemented data visualization with calendar integration for activity trend analysis.
  - Integrated Firebase backend for secure user data storage and authentication.

Skills

- **Programming Languages :** Python, C++, Java, Bash
- **Machine Learning & AI :** PyTorch, TensorFlow, Keras, Scikit-learn, Hugging Face Transformers
- **Data Processing :** Pandas, NumPy, NLTK, SpaCy, OpenCV, Matlab
- **Development Tools :** Docker, MySQL, Git
- **Data Visualization :** Matplotlib, Seaborn, Tableau

Honors & Awards

- **Outstanding Graduate Student Award** *February 2020 - June 2022*
  Awarded a full scholarship to pursue a Master's degree for outstanding academic performance.
- **Outstanding Undergraduate Student Award** *February 2020*
  Achieved the top rank in the graduating class (62 students), earning recognition on the Honors List.