ZEESHAN KHAN (+91)9983621955 zeeshan.khan@research.iiit.ac.in

EDUCATION

International Institute of Information Technology, Hyderabad (IIIT-H) *(August'21 - March'23)*
- **Master of Science in Computer Science Engineering (MS CSE)**

Institute of Technology Nirma University, Ahmedabad *(July'14 - May'18)*
- **Bachelor of Technology, Instrumentation and Control Engineering (B. Tech IC)**

PUBLICATIONS

- **Grounded Video Situation Recognition**
  Zeeshan Khan, CV Jawahar, Makarand Tapaswi
  Neural Information Processing Systems (NeurIPS 2022) [Project Page]
- **More Parameters? No Thanks!**
  Zeeshan Khan, Sukruth Kartheek, Vinay Namboodiri, CV Jawahar
  Association for Computational Linguistics (ACL Findings 2021) : [Paper][Code]
- **Deep Weakly-Supervised High Speed High Dynamic Range Video Generation**
  Zeeshan Khan, Mukul Khanna, Parth Shettiwar, Shanmuganathan Raman.
  International Conference on Pattern Recognition (ICPR 2022) [Oral presentation][Paper] video]
- **Appearance Consistent Human Pose Transfer Transfer via Dynamic Feature selection**
  Ashish Tiwari, **Zeeshan Khan**, Shanmuganathan Raman.
  International Conference on Pattern Recognition (ICPR 2022) [Paper]
- **Exploring Pair-wise NMT for Indian Languages.**
  Sukruth Kartheek, Sreedhar Rajpurohit, Sai Himal Allu, Aman Singhal, **Zeeshan Khan**, Vinay Namboodiri, CV Jawahar
  International Conference on Natural Language Processing, (ICON 2020) : [Paper]
- **FHDR: HDR Image Reconstruction from a Single LDR Image using Feedback Network**
  Zeeshan Khan, Mukul Khanna, Shanmuganathan Raman.
  IEEE ***Global Conference in Signal and Information Processing****, 2019* : [Oral Presentation] [Paper] [Code]

RESEARCH EXPERIENCE

CVIT, IIIT Hyderabad *Research Fellow*
Mentors: Prof. C.V. Jawahar, Prof. Makarand Tapaswi, Prof Vinay Namboodiri *(July'20 - Present)*

- **Grounded Video Situation Recognition:**
  - Proposed a new framework that combines Visual Grounding with Semantic Role Labelling in a weakly supervised setting, for Situation Recognition in videos. Achieved State-of-the-art results for SRL on a large scale movie dataset (VidSitu).
  - Designed a novel 3-stage Transformer model called *VideoWhisperer* that enables joint Action Recognition, Semantic role labelling to answer questions like (*Who did what, to whom, with what, where etc*), and also allows for visual grounding that localises the agents/patients/tools taking part in the action, in the spatio-temporal domain.

- **Multilingual Neural Machine Translation(MNMT):**
  - Studied the long standing problem of negative interference in MNMT, and proposed to learn language dependent submodules to overcome it. Achieved State-of-the-art performance on 8+ languages on the multilingual TED talks dataset.
  - Proposed a multi-task learning scheme through iterative pruning and retraining, to find language dependent parameters from the existing multilingual parameters, without using any new language dependent adaptive layers.

- **Exploring Pair-wise NMT for Indian Languages :**
  - Developed State-of-the-art NMTs for 11 low resource Indian Languages. Achieved rank #1 on the workshop on Asian Translation (WAT) leaderboard for multiple language pairs.
  - Proposed a back-translation filtering mechanism, to filter the noisy back-translated monolingual corpora and retain only the high quality training data to finetune an MNMT to a single language pair.

Computer Vision Lab, IIT, Gandhinagar *Research Assistant*
Mentor: Prof. Shanmuganathan Raman *(February'19 - June'20)*

- **Deep High Speed High Dynamic Range Video Generation from Off-The-Shelf Cameras**
  - Proposed to generate high FPS HDR video, from a sequence of low FPS alternating exposure LDR frames for the first time.
  - Used pre-trained video frame interpolation models to generate multiple high and low exposure frames in the LDR domain, and merged them at each timestep to generate a high speed HDR video at arbitrarily high frame rates.
- **Appearance Consistent Human Pose Transfer via Dynamic Feature Selection:**
  - Proposed a Novel 3-way GAN consisting of image, pose, and appearance pathways. That progressively transforms a human image from a source to a target pose. Used deformable convolutions, non-local attention, and adaptive instance normalisation for appearance transfer. Achieved State-of-the-art results on Fashion dataset.
- **FHDR: HDR Image Reconstruction from A Single Exposure LDR Image using Feedback Network**
  - Proposed a novel Feedback CNN, for HDR image generation from a single exposure LDR image. Achieved State-of-the-art performance, with significant improvement over prior methods.
  - Designed a novel Dense Feedback Block using hidden states of RNN, to transfer high-level information to low-level features. LDR to HDR representations are learned in multiple iterations via a feedback loop.

Raxter @ https://raxter.io, Gandhinagar *Research Assistant*
Mentor: Prof. Sourish Dasgupta~ CEO at Raxter (Asst. Professor @ DAIICT) *(Jan'18 - Jan'19)*

- **RAxBot: Reinforcement Learning(Q-learning) based Scholarly Article Recommendation Engine:** [Pre-print]
  - Designed and implemented RAxBot: The core Personalised Scholarly Article Recommendation Engine behind **raxter.io**, providing scalable content delivery and adaptive recommendations.
  - Proposed a novel Q-learning based framework that exploits the user selection pattern history and takes action to predict the user behaviour. Based on the prediction a query is generated and sent to the Elasticsearch server, for content retrieval.
- **Scholarly Article Recommendation Engine using a Query Augmentation Framework:**
  - Implemented a novel personalised Recommendation Engine using Content Based Filtering and a Query Augmentation technique. Achieved **A+** grade (10/10) in the final semester.
  - Designed a complex 3-level Elasticsearch query. i) Topic-level using LDA topic models ii) Word Embedding level using Word2vec model, and iii) Noun Phrase level. Performed extensive experimentation to determine the query weights.

CONSULTANCY PROJECTS

Azure Knowledge Corporation, Ahmedabad *Technical Consultant*
Project Coordinator : Abhilash Mankad ~ COO at Azure *(Jul'20 - Aug'20)*

- **Automatic Advertisement and Signboard Detection in Mobility**
  - Designed and developed a deep learning based tool to automate the filtering of advertisements and signboards on streets in the wild. Trained a YOLO-V3 object detection algorithm for the given task.
  - Further developed a re-trainable tool using transfer learning, which allows it to adapt to new datasets and classes.

Honeywell International India Pvt. Ltd., Bangalore *Research Consultant, representative from IITGN*
Project Coordinator : Manjuprakash Rama Rao ~Director, Architecture and Innovation *(Oct'19 - Mar'20)*

- **Synthetic Data Generation for Person Intrusion Detection Using Human Pose Transfer** [Architecture-diagram] **[Code]**
  - Proposed a 2-stream GAN framework for Human Pose Transfer, involving foreground and background image generation. FG path deals with pose transfer and BG path inpaints the background for generating consistent target FG and BG images.
  - Extended the Progressive Attention Transfer Network(PATN)(CVPR-2019) by proposing FG and BG losses for self-supervised generation of target FG and BG masks, allowing to separately model the FG and the BG paths.

INTERNSHIPS

Google Summer of Code with Sugarlabs (GSoC'17) *Intern*
Mentor: Walter Bender ~ Founder Sugarlabs, ex-Executive Director of The MIT Media Lab *(May'17- July'17)*

- **Say No To GTK2: Graphical User Interface Refactoring :** [Blog with Code]
  - Completed the Google Summer of Code program'17 and was rewarded a sum of **2400$.**
  - Refactored the GUI and migrated 9 major applications of Sugarlabs from GTK2 to GTK3 toolkit in Python, enabling further growth and development of the applications. Also, Ported from GST 0.10 to GST 1.0.