Shubhajit Basak

Email - shubhajit.basak@yahoo.com | LinkedIn - shubhajitbasak | GitHub - shubhajitbasak | shubhajitbasak.github.io

Machine Learning Engineer & Computer Vision Researcher with a PhD in Computer Vision & Machine Learning and extensive experience building and deploying AI systems. Skilled in designing, optimizing, and productionizing models for edge/embedded devices, with expertise in object detection, face analytics, 3D reconstruction, and model compression. Strong background in end-to-end ML pipelines, model deployment, and performance optimization, with earlier career experience in data engineering and application support.

Core Competencies

- **Programming & Languages:** Python, C++, C#, Bash, SQL.
- **ML & DL Frameworks:** PyTorch, TensorFlow, ONNX, TensorRT, MediaPipe, YOLO, OpenCV, Open3D, PyTorch3D.
- **Deployment & MLOps:** Docker, Kubernetes, MLflow, model quantization, CI/CD for ML, NPU/FPGA optimization
- **Specialized Domains:** Face recognition, face analytics, object detection, network quantization, generative AI (diffusion models), adversarial learning, sensor fusion (thermal/millimeter-wave), synthetic data generation, 3D face reconstruction
- **Data Engineering:** Azure SQL Data Warehouse, Netezza, Microsoft SQL Server, ETL pipelines
- **Tools:** Jira, Bitbucket, Confluence, Blender, photogrammetry pipelines

Professional Experience

FotoNation Ltd. (previously part of Xperi) — *Machine Learning Engineer* May 2023 - Till Date

- Designed a **YOLOv5-based moving object detection** system for FPGA deployment, reducing inference latency by 25% compared to baseline.
- Developed **face detection and head pose estimation** for a **driver monitoring system** (DMS) and made it compatible on Novatek NPU boards with 2ms per frame as inference time.
- Created a **thermal super-resolution model** integrated into **object detection** pipelines, optimized for Hailo NPU with 20% model size reduction via model quantization.
- Collaborated in an **Agile** team of 6 engineers to deliver an end-to-end solution for the **Driver and Occupant Monitoring solution (DMS/OMS)** for a leading German OEM manufacturer, enabling real-time **face recognition, face activity detection, age/gender classification, drowsiness detection, and occlusion handling**.
- Implemented **camera distortion and perspective correction** pipelines and converted PyTorch models to **Qualcomm QNN & AIMET** formats through ONNX for **embedded deployment**.

University of Galway (sponsored by Xperi) — *PhD Researcher* Sep 2019 - Aug 2023

- Developed an **audio-conditioned diffusion model** for speech-driven talking head synthesis, improving realism scores by 3.4% over state-of-the-art baselines. Project: DiffusionVideoEditing
- Built a lightweight **3D dense facial landmark estimation** network, trained on synthetic datasets generated via **3DMM parametric modeling**. Project: Dense3DFaceLandmarks
- Designed weakly-supervised **3D face reconstruction** from single images using **hierarchical transformers**, achieving 0.95mm MAE in 3D face geometry.
- Generated accurate synthetic facial depth and head pose data from 3D Models through the CGI Tool Blender. Proposed an adversarial domain adaptation framework to transfer head pose learning from the synthetic datasets to real-world scenarios. Project: BlenderDataGeneration
- Published **7+ peer-reviewed papers** in Neural Networks, IEEE Access, Image and Vision Computing, and other venues (full list: Google Scholar)
- **Teaching**: As part of the PhD curriculum, worked as a teaching assistant on the following courses - Programming for Data Analytics (R Programming), Tools and Techniques for Large Scale Data Analytics (Java with Apache Spark), Computer Architecture and Operating Systems

Cognizant Tech Soln. Ltd. — *Data Engineer* Sep 2014 - July 2018

- Maintained and developed data warehouse solutions using **Azure Data Warehouse and Netezza**
- Designed **ETL pipelines** for a retail supply chain business using **Microsoft ETL stack**, improving processing throughput by approximately 20%.
- Developed **optimized SQL queries and indexing strategies** that reduced reporting query time from minutes to seconds.
- Built data models in collaboration with Front End/Back End teams for retail ordering systems

Cognizant Tech Soln. Ltd. — *Application Support Engineer* Sep 2011 - Aug 2014

- Led a 4-member support team maintaining business-critical applications with 100% SLA compliance for more than a year
- Delivered bug fixes and performance enhancements for ASP.Net and SQL Server-based solutions in the fuel and retail domain

Education

Ph.D. — ***Computer Science*** Sep 2019 - Sep 2023

- University of Galway, Ireland
  Area: Computer Vision and Machine Learning
  Thesis: "Contributions towards 3D synthetic facial data generation and 3D face analysis with weak supervision"
  Advisor: Michael Schukat (UoG), Rachel McDonnell (TCD), Peter Corcoran (UoG)

Masters — ***Computer Science (Data Analytics)*** Sep 2018 - Aug 2019

- University of Galway, Ireland
  Thesis: "Enriching the training data with an external knowledge graph database like WordNet, DBPedia, for text classification task"
  Advisor: Matthias Nickles (UoG)

Bachelors — ***Electronics and Communication*** Sep 2007 - Aug 2011

- West Bengal University of Technology, India
  Thesis: "Hand Recognition as a Biometric with Stereo Images"
  Advisor: Biswa Nath Chatterji (WBUT)

Selected Publications

- Bigioi D, **Basak S**, Stypułkowski M, Zieba M, Jordan H, McDonnell R, Corcoran P. Speech driven video editing via an audio-conditioned diffusion model. **Image and Vision Computing**. 2024
- **Basak S**, Mangapuram S, Costache G, McDonnell R, Schukat M. A lightweight 3D dense facial landmark estimation model from position map data. **IMVIP**. 2023
- **Basak S**, Khan F, Javidnia H, Corcoran P, McDonnell R, Schukat M. C3I-SynFace: A synthetic head pose and facial depth dataset using seed virtual human models. **Data in Brief**. 2023
- Basak S, Corcoran P, McDonnell R, Schukat M. 3D face-model reconstruction from a single image: A feature aggregation approach using hierarchical transformer with weak supervision. Neural Networks. 2022
- **Basak S**, Corcoran P, Khan F, Mcdonnell R, Schukat M. Learning 3D head pose from synthetic data: A semi-supervised approach. **IEEE Access**. 2021
- Khan F, Hussain S, **Basak S**, Lemley J, Corcoran P. An efficient encoder–decoder model for portrait depth estimation from single images trained on pixel-accurate synthetic data. **Neural Networks**. 2021