Mohammed Jaseel Kunnathodika

Nationality : Indian
Date of birth: 27/06/1998
University of Stuttgart
Address: Stuttgart-Vaihingen, Allmandring 1 8A - 22D, 70569, Stuttgart, Germany (Home)

+49 15739399740
jaseelkt1official@gmail.com
st191717@stud.uni-stuttgart.de
GitHub
LinkedIn

Education

•National Institute of Technology, Calicut *2017-2021*
Bachelor of Technology in Electrical and Electronics Engineering CGPA: 8.01

•University of Stuttgart, Germany *2024-*
Master of Science in Electrical Engineering

Experience

•Research on Conditional Video Generation (University of Stuttgart, Research Project)
September 2025 – Present

- Investigating controllability in video diffusion models, focusing on learning how to guide generation using structured control signals rather than unconstrained sampling
- Implementing and extending a Ctrl-V–inspired two-stage setup: first synthesize control frames (e.g., bbox, trajectories, semantic channels), then generate temporally consistent videos
- Exploring multi-modal controls—bounding boxes, trajectories, and physics-related signals (speed/acceleration)—to enforce coherence and physics consistency in autonomous-driving scenes.
- Experimenting with either ControlNet-style or SPADE-style conditioning modules (used alternatively, not simultaneously) to compare spatial/temporal control and compositional fidelity
- Fine-tuned pretrained diffusion backbones on the KITTI-360 dataset,

•Speech-to-Speech Translation System (Sony Europe, Working Student) *May 2025 – Present*
Developed and trained a Transformer-based speech generation model from scratch for multilingual audio synthesis and cross-language alignment.
Stuttgart, Germany

- Developed a Transformer-based Text-to-Speech model with dedicated encoder-decoder architecture, training the system entirely from scratch
- Utilized high-performance H200 GPUs for large-scale parallel distributed training, optimizing multilingual speech generation.
- Designed and implemented adapter layers and architectural modifications to efficiently map text embeddings to quantized audio token space, improving cross-modal alignment.
- Applied advanced finetuning and transfer techniques to the pretrained text encoder, adapting its representations to the audio token space for better cross-modal alignment.
- Designed and prototyped a full speech translation pipeline enabling seamless language switching in audio while preserving original characteristics, delivered via a custom web application.
- Engineered phonetic alignment, speaker diarization, and force-aligned subtitles to synchronize dubbed audio precisely and maintain high naturalness across languages..
- Expanded system capabilities to real-time, high-quality audio replacement and cross-speaker consistency, ensuring robust performance in production settings.

•Software Engineer - BOSCH Global Software Technologies Pvt. Ltd
23/07/2021 – 11/03/2024 Coimbatore, India

- Led the implementation of the full CHADEMO charging communication stack for Japanese-market EVs, ensuring robust integration with AUTOSAR architecture and power electronics (DCDC converter, onboard charger.
- Developed automated simulation frameworks for hardware testing (CAPL scripting, Vector Canoe), streamlining validation of CAN protocols and charging functionalities.
- Expertise in Embedded C programming.
- Built advanced vehicle diagnostics using DFC modules for real-time fault detection, freeze-frame logging, and diagnostic trouble code (DTC) handling on CHADEMO CAN networks..
- Designed and deployed custom network management protocols, supporting charger connection detection, ECU wake-up, and safety failure state handling for compliant, reliable operation in real vehicles.

•Student Assistant - Web Application for University Student Evaluation (University of Stuttgart)
10/01/2025 – 31-04-2025 Stuttgart, Germany

– Developed front-end and back-end for a university evaluation web app with dynamic content delivery (Flask).
– Integrated core modules such as 2D image to 3D conversion for student visualization.
– Built evaluation dashboard for streamlined results analysis.
– Developed an AI-based documentation agent for industrial supply chain, enabling intelligent retrieval (RAG) and summarization of heterogeneous datasets (CAD, BOM, datasheets).

Personal Projects

–YouTube Learning Companion - STARTUP *Feb 2025 – June 2025*

Designed an AI-powered educational webapp that turns long-form videos into structured, interactive study experiences—enabling students to learn deeply without watching entire videos.

* Pioneered a *multi-expert architecture* that breaks down long form videos and delivers lay-friendly explanations, technical deep dives or code-centric views on demand.
* Integrated automatic diagram generation via Mermaid js, and deployed a Desmos-based visualization module to render math equations from video transcripts as interactive graphs—dramatically improving clarity for abstract topics.
* Supported multi-stage video breakdown: lengthy videos are segmented and explained part-by-part with context persistence, enabling deeper, coherent follow-up questions across multiple sessions.
* Built a hybrid similarity recommender that surfaces thematically adjacent videos, guiding learners to the next skill with one click.
* System is prepared for commercial rollout with performance-tuned backend (async pipelines, streaming, faiss caching) and a planned subscription-based launch.

–Perplexity-Style Web Search and Summarization Assistant Agent *Feb 2025 – Present*

Designed and deployed an intelligent web search integrated into a Discord bot for real-time query answering.

* Designed and implemented a hybrid retrieval system combining multiple external search APIs to dynamically search and extract information from both web articles and video transcripts for open-domain queries.
* Developed a custom multi-factor ranking pipeline starting with sentence embedding and cosine similarity, followed by BM25 relevance scoring (softmax-normalized) for efficient CPU-based matching.
* Incorporated additional ranking factors including content recency decay, domain/channel reputation scoring, snippet richness evaluation, and manual bonus terms to prioritize high-quality and trustworthy sources.
* Built robust preprocessing and normalization pipelines involving query refinement, ISO 8601 date parsing, lightweight summarization from YouTube transcripts, and graceful fallback mechanisms against missing metadata or blocked API responses.
* Integrated the complete search, ranking, and reranking modules into a Discord bot, enabling users to receive real-time, citation-supported responses optimized for reliability and diversity across a wide range of topics.
* Optimized the end-to-end retrieval-to-response system for low latency, CPU efficiency, and high recall, aiming to deliver a user experience similar to modern AI search platforms like Perplexity AI.

–Human Activity Recognition (HAR) using Deep learning *Dec 2024 – Feb 2025*

Developed a deep learning-based system to classify physical activities from wearable sensor data.

* Utilized gyroscope and accelerometer data for activity recognition, applying extensive time-series preprocessing methods like sliding window, and evaluated on Sequence to Label vs Sequence to Sequence labeling methods.
* Implemented oversampling techniques to address the class imbalance and compared the performance of GRU and LSTM architectures for activity classification.
* Designed a custom loss function incorporating PolyLoss to penalize low-confidence predictions and integrated Regularized dropout (R-drop) to enhance model generalization.
* Leveraged Bayesian optimization to fine tune model hyperparameters for optimal performance.

–Diabetic Retinopathy Detection using Deep Learning *Oct 2024 – Dec 2024*

Developed a medical image classification pipeline to detect diabetic retinopathy from retinal fundus images.

* Applied background trimming and aspect-ratio preserved resizing using the LANCZOS filter to preprocess high-resolution retinal images for efficient CNN processing.
* Fine-tuned VGGNet, MobileNet, and InceptionV2 models using transfer learning for diabetic retinopathy classification on a small dataset.

* Implemented ensemble techniques to boost accuracy and used grid search optimization via WandB for hyperparameter tuning.
* Integrated Grad-CAM visualizations to identify key retinal regions influencing predictions, aiding debugging, model tuning, and clinical interpretability.

Technical Skills and Interests

AI/ML Frameworks and Tools: PyTorch, TensorFlow, Keras, Retrieval-Augmented Generation, MCP, Prompt Engineering
Deep Learning MLOps: Experience with HPC clusters, distributed training on Linux
Version Control: Git, GitHub
Web Development: Next.js, React.js, Node.js, HTML5
Cloud and Databases: MongoDB, PostgreSQL,SQL, Cloud Deployment (Render, Vercel, AWS)
Software and Tools:Matlab, Docker, Unity
Areas of Interest: Generative AI ,Reinforcement learning, Robotics

Achievements and Leadership

–**AUTOSAR COM Stack Owner** Bosch Global Software Technologies, India *2023 – 2024*
* Led development of AUTOSAR communication stack for EV ECUs, ensuring reliable signal handling and system integration.

–**Suspension and Steering Lead – Formula Student** Team Unwired, NIT Calicut *2018 – 2019*
* Designed and manufactured Suspension and steering system including upright; team ranked 24th nationally at SAE SUPRA 2019.

–Achieved All Kerala Rank 370 in KEAM 2017 (Graduate Engineering Entrance), out of more than 100,000 candidates.