

Hisham Barakat

Lead AI Engineer — Scalable ML Solutions — 95% Inference Optimization

Hishambarakat16@gmail.com | +86 166 0152 0270 | +973 3302 5065 | linkedin.com/in/hishambarakat
[hishambarakat16.github.io](https://github.com/hishambarakat16)

Education

Peking University <i>Master in Computer Science</i>	Sept 2023 – Present <i>Beijing, China</i>
<ul style="list-style-type: none">• Courses: Computational Vision, Reinforcement Learning, Computer Networking, Advanced Systems.• Optimize Jittor backend for TensorLayerX's multi-backend AI framework.	
Bahrain Polytechnic <i>Bachelor of ICT</i>	Sept 2017 – May 2021 <i>Isa Town, Bahrain</i>
<ul style="list-style-type: none">• Courses: Data Center Mgmt, Advanced Networking, Programming, IT Project Mgmt.	

Skills

Languages: Python, C++, SQL, YAML

Technologies: PyTorch, TensorRT, Diffusion, LLMs, Flask, Docker, AWS, Kubernetes, Power BI, Git, CUDA.

Experience

Jiangsu Huaxiang Digital Technology Group Co.,Ltd <i>Lead AI Engineer (Collaborative Project for Jiangsu Huaxiang)</i>	Sept 2024 – Present <i>Remote</i>
<ul style="list-style-type: none">• Architect a scalable AI image generation app, expanding from 3 to 30 models (text-to-image with 21 styles, upscaling, face/style transfer) for local and API-based deployment.• Achieve 95% inference time reduction (from 20s to less than 1s) by optimizing upscaling processes with TensorRT for GPU acceleration, enhancing user experience.• Design independent pipelines for diverse model types (SDXL, SD1.5, custom variants), optimizing resource allocation and request handling on constrained hardware.• Lead a team of 3 junior engineers, providing mentorship and skill-building sessions to ensure effective contributions and project continuity.• Spearhead transition to cloud deployment, building CUDA-ready Docker containers to enable seamless production scaling and operational efficiency.	
Huawei Technologies Bahrain W.L.L <i>IP Engineer</i>	Sept 2022 – Aug 2023 <i>Seef, Bahrain</i>
<ul style="list-style-type: none">• Automated configuration scripts, boosting efficiency by 50%.• Ensured 100% network connectivity through rapid troubleshooting.	
Tatweer Petroleum <i>Data Science & AI Engineer</i>	Sept 2021 – Aug 2022 <i>Awali, Bahrain</i>
<ul style="list-style-type: none">• Cleaned 1M data points from 6 systems, enabling diagnostics for 300+ wells.• Developed ML models to enhance oil recovery across 300+ wells.	
Amazon Web Services <i>Solutions Architect intern</i>	Feb 2021 – Jun 2021 <i>Manama, Bahrain</i>
<ul style="list-style-type: none">• Built a serverless Face-Detection Dashboard with 90%+ accuracy, reducing costs by 95%.• Managed 5 developers to integrate Lambda with S3/DynamoDB.	