

Le Tan Dang Khoa

+65 9157 0292
✉ letan.dangkhoa@gmail.com
📄 ltdk.me

Current Employment

- 2024–Now **Senior ML Engineer**, *Visenze*, Singapore.
- Oct 2018–2024 **Research Software Engineer**, *Visenze*, Singapore.

Previous Employment

- Mar 2016–2018 **Research Assistant**, *Singapore University of Technology and Design*, Singapore.
- May–Sep 2015 **Research Intern – Mitacs Program**, *University of Saskatchewan*, Saskatchewan, Canada.
- Feb–Apr 2014 **Research Intern**, *Japan Advanced Institute of Science and Technology*, Nomi, Japan.

Education

- 2023–Now **Master of Computing, Artificial Intelligence Specialisation**, National University of Singapore.
- 2011–2015 **Bachelor of Science in Information Technology, Honors Program**, Ho Chi Minh University of Science.

Projects

Visenze

- **MultiSearch Solution**
 - Evaluated 5+ vector search engines against in-house solution; proposed new architecture adopted as next-gen search infrastructure.
 - Engineered multimodal search algorithm fusing dense/sparse embeddings to optimize performance across **7 languages** and diverse domains (**fashion, furniture**).
 - Built end-to-end continuous improvement pipeline: LLM-powered training with customer data and evaluation framework for bad case analysis/visualization, boosting precision 5% quarterly.
 - Developed advanced search capabilities: reranking, query boosting, natural language filters, autocomplete, adaptive thresholding.
- **LLM Agents**
 - Architected LLM agent pipeline automating annotation (translation, labeling, query generation), eliminating vendor costs and manual effort.
 - Developed LLM-based product ingestion service improving search quality (**DCG ↑10%**).
- **Vector Search**
 - Resolved HNSW replica consistency issues and optimized parameters through systematic experimentation.
 - Developed novel distance metric reducing vector storage 25% and accelerating search 2× while maintaining recall.
- **Computer Vision**
 - Developed and deployed core product detection models powering search/tagging/recommendation systems: revamped annotation processes, built new training codebase, and extended API functionality.
 - Engineered duplicate product identification system: advanced visual embeddings with novel augmentations and rerank logic for customer catalogs.
- **Augmented Reality**
 - Architected end-to-end 3D labeling solution: mobile AR app for object capture/data collection, data refinement tool, and detection codebase (training/eval/visualization)
 - Engineered mobile AR features (virtual try-on, object tracking, gesture recognition) using DL models optimized for 30 FPS on-device inference via memory-efficient architectures
 - Deployed to Unity-based game achieving 10M+ Play Store downloads and top-30 iOS Education app ranking

Singapore University of Technology and Design

- **Urban-Area Scene-Based Localization** Spearheaded mobile implementation of the system. Led the porting of the inference framework from MATLAB to Android, focusing on performance optimization. Developed a high-performance SIFT feature extraction pipeline using OpenGL shader programs, maintaining real-time performance and accuracy on mobile devices [1].

- **Integer Optimization on Deep Learning** Developed novel deep learning methods for efficient image retrieval via hashing. Pioneered an end-to-end approach that jointly learns image features and optimal hash codes [2, 3]. Designed new feature aggregation schemes to enhance performance for both convolutional features [4] and traditional hand-crafted features [5], enabling effective large-scale supervised and unsupervised hashing [6].

National University of Singapore

- **HDB Rental Prediction:** Developed a regression model securing **#1 rank** on the competition's private leaderboard.
- **Self-ordering Menu Recommender:** Developed a deep reinforcement learning-based **recommendation system** trained on 11-month data from a F&B franchise in Singapore.
- **LightningCat (SaaS Platform):** Architected and developed the backend (cloud database, feature store, ML pipeline) powering **real-time updates** for Singapore's sporting facilities through an interactive map.
- **TemporalLens:** Developed **an NLP pipeline** using LLMs, unsupervised clustering, and weak supervision to construct event-centric knowledge graphs capturing narrative progression from news streams.

Others

- **Lymphocytes counting [Github]:** Developed a new method utilizing color differentiation and object contour analysis.
- **Feature learning:** Proposed a compression method for PCA features and applied to the facial recognition problem [7].

Technical Skills

- **Programming Languages:** Python, Go, C++, Java, C#, Obj-C, Julia, MATLAB.
- **Technologies:** VectorDB, LangChain, HuggingFace, AWS, OpenCV, Pandas, PyTorch, Tensorflow, NVIDIA Triton, Docker, MLFlow, Unity, Mediapipe, K8s, MLFlow.
- **Languages**
 - Vietnamese: Native proficiency.
 - English: Professional working proficiency. Academic IELTS score: 7.5.

Awards

- 2015 **Mitacs, Globalink Scholarship**, Canada.
- 2014 **Bank of Tokyo, Mitsubishi UFJ Scholarship**, Vietnam.
- 2014 **Sunflower Mission, Engineering and Technology Scholarship**, Vietnam.
- 2014 **IEEE, IEEE Student Travel Grant**, USA.
- 2014 **The National Foundation for Science and Technology Development, Research grant**, Vietnam.
- 2014 **Japan Student Services Organization, JASSO ASEAN scholarship**, Japan.

Selected Publications

- [1] N.-T. Tran, D.-K. Le Tan, A.-D. Doan, T.-T. Do, T.-A. Bui, M. Tan, and N.-M. Cheung, "On-device scalable image-based localization via prioritized cascade search and fast one-many ransac," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 1675–1690, 2019.
- [2] T.-T. Do, D.-K. Le Tan, T. T. Pham, and N.-M. Cheung, "Simultaneous feature aggregating and hashing for large-scale image search," in *CVPR*, July 2017.
- [3] D.-K. Le Tan, L. Huu, T. Hoang, T.-T. Do, and N.-M. Cheung, "Deepvq: A deep network architecture for vector quantization," in *CVPR Workshop on Learned Image Compression*, IEEE, 2018.
- [4] T. Hoang, T.-T. Do, D.-K. L. Tan, and N.-M. Cheung, "Selective deep convolutional features for image retrieval," *ACM Multimedia*, 2017.
- [5] T. Hoang, T.-T. Do, D.-K. L. Tan, and N.-M. Cheung, "Enhance feature discrimination for unsupervised hashing," *ICIP*, 2017.
- [6] D.-K. Le Tan, T.-T. Do, and N.-M. Cheung, "Supervised hashing with end-to-end binary deep neural network," in *ICIP*, IEEE, 2018.
- [7] D.-K. T. Le, H. P. Truong, and T. H. Le, "Facial expression recognition using statistical subspace," in *ICIP*, pp. 5981–5985, IEEE, 2014.
- [8] T.-T. Do, T. Hoang, D.-K. Le Tan, T. Pham, H. Le, N.-M. Cheung, and I. Reid, "Binary constrained deep hashing network for image retrieval without manual annotation," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 695–704, IEEE, 2019.