Luong Phat Nguyen

Docteur en Vision par Ordinateur | AI Research Scientist — LLM multimodaux & IA vocale

Paris, France luongphatnguyen@gmail.com +33 6 95 32 32 97
lphatnguyen.github.io linkedin.com/in/nguyen-luong-phat github.com/lphatnguyen

PROFIL

Docteur en vision par ordinateur, avec plus de 7 ans d'expérience en R&D en IA, de la recherche académique au déploiement industriel. Auteur de quatre publications internationales, j'ai conçu et optimisé des modèles de vision, des LLM multimodaux et des systèmes vocaux temps réel, avec des résultats concrets en précision et en latence. Je souhaite mettre cette double expertise scientifique et industrielle au service d'un poste de Research Scientist/Engineer, pour développer des approches innovantes en IA multimodale et les transformer en solutions performantes à fort impact.

EXPÉRIENCE PROFESSIONNELLE

AI Research Scientist | **Rizlum AI** Mars 2025 – Aujourd'hui | Paris, France

- Conçu des modèles multimodaux d'analyse de documents d'assurance et de mutualité (factures de santé, constats amiables) : **plus de 95 % d'extraction correcte**, en quelques secondes par document.
- Généré des données synthétiques et anonymisé les données d'identité (secteur régulé) : performance équivalente aux données réelles, sans exposition de données personnelles.
- Développé les modèles propriétaires Speech-to-Text et Text-to-Speech de Rizlum AI (**latence < 200 ms**) : conversation à faible latence et fin de la dépendance aux API vocales tierces.
- Co-développé un agent vocal speech-to-speech temps réel pour un client assureur (revue d'architecture, STT/TTS, évaluation) : **80 % des adhésions téléphoniques automatisées**.

Technologies : LLM multimodaux, TTS, ASR, PyTorch, Hugging Face, RAG, DeepEval, inférence temps réel

Chercheur post-doctoral en LLM | **Toyota Motor Europe & LISSI** Juil. 2024 – Mars 2025 | Paris, France

- Conçu et entraîné des LLM multimodaux vision-langage (architectures Mamba et Transformers).
- Industrialisé les pipelines d'entraînement multimodaux à grande échelle (VQA, Conceptual Captions).
- Créé un agent conversationnel émotionnellement intelligent ; travaux connexes publiés à **IEEE ICIP 2025**.

Technologies : LLM multimodaux, Mamba, reconnaissance d'émotions faciales, PyTorch

AI Research Scientist | **Trax Retail** Déc. 2021 – Avril 2024 | Paris, France

- Développé des modèles de reconnaissance de produits en rayon (**plus de 95 % de précision**, plus de 500 références) : audit photo de quelques secondes au lieu d'un relevé manuel en magasin.
- Embarqué des algorithmes temps réel en Swift dans un SDK iOS (CoreML, ARKit, SceneKit) : inférence sur le téléphone, sans dépendance réseau ni coût d'inférence serveur.
- Optimisé et généralisé des modèles pré-entraînés sans ré-annotation massive : déploiement plus rapide et moins coûteux sur de nouvelles enseignes et de nouveaux pays.
- Calculé en temps réel des indicateurs métier (*share-of-shelf*) à **< 1 % d'erreur et < 200 ms de latence**, exploitables par les marques pendant la visite magasin.

Technologies : PyTorch, CoreML, ARKit, Swift, OpenCV, vision 3D, réalité augmentée

Doctorant en Vision par Ordinateur | **LIFAT, Université de Tours** Nov. 2018 – Déc. 2021 | Tours, France

- Mené une thèse sur les motifs spatio-temporels dans les vidéos : **2 publications** (ICPR, VISAPP).
- Conçu des architectures légères de classification vidéo (R(2+1)D, R3D, LSTM, ViT) pour peu de données annotées ; exploité le flux optique (FlowNet2, PWC-Net) et la fouille de motifs séquentiels.

Technologies : PyTorch, Keras, OpenCV, Python, C++, fouille de données

COMPÉTENCES CLÉS

IA & Deep Learning	LLM & VLM multimodaux, Text-to-Speech, Speech-to-Text, Speech-to-Speech, agents conversationnels, Transformers, CNN, LSTM, modèles à espace d'états (Mamba), fine-tuning (LoRA / QLoRA / PEFT), optimisation et distillation de modèles, RAG, DeepEval
Vision par ordinateur	Reconnaissance et classification d'images et de vidéos, estimation de flux optique, réalité augmentée, vision 3D, extraction de motifs spatio-temporels
Mise en production	Inférence temps réel et faible latence, modèles embarqués (CoreML), industrialisation de pipelines d'entraînement, anonymisation des données d'identité et génération de données synthétiques (secteur régulé)
Langages	Python, C/C++, Swift, Bash, JavaScript, Matlab, LaTeX
Frameworks & outils	PyTorch, Hugging Face, vLLM, Keras, OpenCV, CoreML, ARKit, SceneKit, Docker, Git, Linux, Google Cloud, FastAPI
Transverses	Recherche appliquée, transfert recherche $\rightarrow$ produit, publication scientifique, enseignement universitaire, traitement de données à grande échelle

FORMATION

Doctorat en Vision par Ordinateur | **Université de Tours — Laboratoire LIFAT** 2018 – 2021

Sujet : *Extraction et caractérisation de motifs spatio-temporels dans les vidéos*.

Diplôme d'Ingénieur | **INSA Centre Val de Loire** 2015 – 2018

Spécialité : *Systèmes automatisés et informatique industrielle*.

PUBLICATIONS

L. P. Nguyen, H. Abdelkawy, A. Othmani. *EmoMamba: Advancing Dynamic Facial Expression Recognition with Visual and Textual Fusion*. IEEE International Conference on Image Processing (**ICIP**), 2025, pp. 1432–1437.

L. P. Nguyen, J. Mille, D. H. Li, D. Conte, N. Ragot. *Efficient Dynamic Texture Classification with Probabilistic Motifs*. International Conference on Pattern Recognition (**ICPR**), 2022, pp. 564–570.

L. P. Nguyen, J. Mille, D. Li, D. Conte, N. Ragot. *Trajectory Extraction and Deep Features for Classification of Liquid-gas Flow under the Context of Forced Oscillation*. International Conference on Computer Vision Theory and Applications (**VISAPP**), 2020, pp. 17–26.

L. P. Nguyen, M. Saleh, R. Le Bouquin Jeannès. *An Efficient Design of a Machine Learning-based Elderly Fall Detector*. International Conference on IoT Technologies for HealthCare (**HealthyIoT**), Springer, 2017, pp. 34–41.

ENSEIGNEMENT

- **Réseaux de neurones — Deep Learning** (cours magistraux et travaux pratiques) — 5^{e} année, École Polytechnique de l'Université de Tours, 2020.
- **Programmation orientée objet — C++** (travaux pratiques) — 3^{e} année, INSA Centre Val de Loire (GSI), 2019–2020.
- **Introduction au Deep Learning** (travaux pratiques) — Master 2 Big Data Management and Analytics, 2018–2019.

DISTINCTIONS

- Bourse d'excellence *Rencontres du Vietnam* — Odon Vallet, 2016–2018.
- Médaille d'argent, Olympiades nationales de physique (Vietnam), 2014.

LANGUES

Français	Courant (professionnel)
Anglais	Courant (avancé)
Vietnamien	Langue maternelle