

Proposition de stage 2025

Exploitation d'un grand modèle de langage et d'une reconstruction 3D neurale sémantique pour la navigation autonome d'un robot

Contact : steve.bourgeois@cea.fr

Contexte du stage

Si la robotique autonome connaît actuellement de grandes avancées, les intelligences artificielles qui guident ces robots se limitent généralement à une planification à court terme, sur l'environnement immédiat du robot, et pour des domaines de requêtes prédéfinies. Pour aller vers plus de polyvalence, il est nécessaire de développer des Intelligences Artificielles (IA) capable de planifier, sur la base d'une requête en langage naturel, une série d'actions sur des horizons temporels lointain, impliquant des lieux ou des éléments qu'ils ont observés par le passé mais qui sont hors du champ de perception actuel du robot.

Or, l'essor récent des grands modèles de langage (LLM) a démontré la capacité de ces IA à exploiter leur connaissance générale du monde pour planifier des séries d'actions (chaînes de pensées ou *chain-of-thought*) sur des domaines variés. Pour exploiter ces connaissances générales dans le cas d'un robot spécifique dans un environnement spécifique, il devient nécessaire de connecter le LLM aux capacités du robot et au monde réel qui l'entoure (encrage ou *grounding*). Pour tirer parti au maximum des capacités de planification à horizon lointain des LLM, cette connexion ne doit pas se limiter au seule environnement perceptible à l'instant courant par le robot, mais à une représentation 3D de l'environnement qu'il a pu précédemment exploré.

Objectifs du stage

Dans le cadre de ce stage, nous proposons d'étudier le domaine du *Vision-Language Navigation* [1,2,3], domaine de recherche interdisciplinaire à la frontière du traitement naturel du langage, de la vision par ordinateur et de la robotique. L'étudiant aura pour objectif de mettre en place un démonstrateur de navigation robotique autonome exploitant un LLM pour la compréhension de la requête utilisateur et la planification d'actions, ainsi qu'une représentation neurale 3D sémantisée de la scène et d'outils d'analyse de cette dernière pour connecter le LLM au monde réel.

En raison de cet aspect interdisciplinaire, ce stage impliquera à la fois un laboratoire de vision par ordinateur (LVML) et un laboratoire de robotique (LCSR) du CEA.

L'étudiant aura pour charge de :

- Prendre en main les représentations neurales 3D développées au laboratoire LVML et mettre en place les outils d'analyse adaptés à la tâche de navigation (recherche de la destination, détection des obstacles...).
- Mettre en place un LLM en mode REACT pour analyser la requête utilisateur, réaliser la planification de tâches de navigation et invoquer les outils nécessaires (analyse de scène 3D, calcul et suivi de trajectoire...).
- Tester et évaluer la méthode dans un simulateur (eg. Habitat)
- Mettre en place le démonstrateur sur une plateforme robotique réelle (intégration des algorithmes de localisation, de contrôle-commande, etc., dans le framework robotique ROS).

- [1] Liu, R., Wang, W., & Yang, Y. (2024). Volumetric Environment Representation for Vision-Language Navigation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 16317-16328).
- [2] Wang, Z., Li, X., Yang, J., Liu, Y., Hu, J., Jiang, M., & Jiang, S. (2024). Lookahead Exploration with Neural Radiance Representation for Continuous Vision-Language Navigation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13753-13762).
- [3] Gu, J., Stefani, E., Wu, Q., Thomason, J., & Wang, X. E. (2022). Vision-and-language navigation: A survey of tasks, methods, and future directions. *arXiv preprint arXiv:2203.12667*.

Compétences développées au cours du stage

Ce stage permettra à l'étudiant de découvrir le domaine de l'intelligence artificielle pour la robotique, domaine à l'intersection des dernières avancées en terme de reconstruction 3D (Neural Fields, segmentation en vocabulaire ouvert...), du contrôle robotique, et des modèles de langage. L'étudiant découvrira aussi les aspects plus appliqués, avec l'usage de framework largement employés dans la communauté (ROS, NeRFStudio...). L'étudiant aura l'opportunité de travailler avec une équipe de chercheurs seniors mais aussi de doctorants.

Compétences souhaitées

Le candidat devra disposer d'une bonne maîtrise de python et des connaissances en robotique. Une expérience sur Pytorch et/ou ROS sera appréciée.

Informations générales

Formation / Niveau d'étude	Ingénieur, Master 2 / Bac+5
Possibilité poursuite	Oui, en thèse ou CDD selon profil.
Durée	6 mois
Lieu	Palaiseau (91) – Centre d'intégration de Nano-INNOV
Indemnités de stage	Entre 700 € et 1400 € suivant formation. Aide au logement / transport / restauration.

Candidatures

- Joindre CV + lettre de motivation à steve.bourgeois@cea.fr avec le nom du stage auquel vous postulez
- Ne pas hésiter à détailler les projets ou cours auxquels vous avez participé
- Indiquer les dates de début/fin de stage envisagées.
- Ce stage pourra prendre une orientation recherche ou industrie en fonction du profil du candidat

CEA Tech LIST

Les activités de recherche du CEA Tech LIST sont centrées sur les systèmes à logiciel prépondérant. Ces activités s'articulent autour de trois thématiques: les Systèmes Embarqués (architectures et conception de systèmes, méthodes et outils pour la sûreté des logiciels et des systèmes, systèmes de vision intelligents), les Systèmes Interactifs (ingénierie de la connaissance, robotique, réalité virtuelle et interfaces sensorielles) et les Capteurs et le traitement du signal (instrumentation et métrologie des rayonnements ionisants, capteurs à fibre optique, contrôle non destructif).

Le CEA Tech LIST a de nombreux partenariats avec les grands acteurs industriels du nucléaire, de l'automobile, de l'aéronautique, de la défense et du médical pour étudier et développer des solutions innovantes adaptées à leurs besoins. Il réalise une recherche qui va du concept de système jusqu'au démonstrateur, contribuant au transfert de technologies et à l'innovation par l'émergence de nouvelles entreprises.

Laboratoire Vision pour la Modélisation et la Localisation (LVML)

Laboratoire Vision pour la Modélisation et la Localisation (LVML) du CEA Tech LIST mène des recherches en vision par ordinateur et intelligence artificielle. Nous adressons en particulier les problématiques suivantes :

- Géolocalisation et cartographie d'environnement par vision et fusion de capteurs (robotique mobile, drones...)
- Systèmes et de vision pour la robotique : préhension, manipulation, assemblage d'objets...
- Contrôle de conformité, détection de défauts géométriques, colorimétriques, etc...
- Analyses hyperspectrales : détection de matériaux, tri,
- Correction, amélioration d'images et vidéos (superrésolution, upframing, ...)
- Compression de réseaux de neurones
- ...

Laboratoire Contrôle et Supervision Robotique (LCSR)

Le Laboratoire Contrôle et Supervision Robotique (LCSR) du CEA Tech LIST travaille à une interaction optimale dans un environnement où coexistent Hommes et Robots Mobiles. C'est dans ce contexte que s'inscriront les travaux de ce stage, au regard de l'état de l'existant et du savoir-faire du laboratoire, développé notamment lors de projets industriels de transitique et de projets de recherche de conduite de véhicules. En particulier, le LCSR proposera l'exploitation d'une application logicielle de navigation autonome d'un robot mobile dans le cadre de ce stage.

Proposition de stage 2025

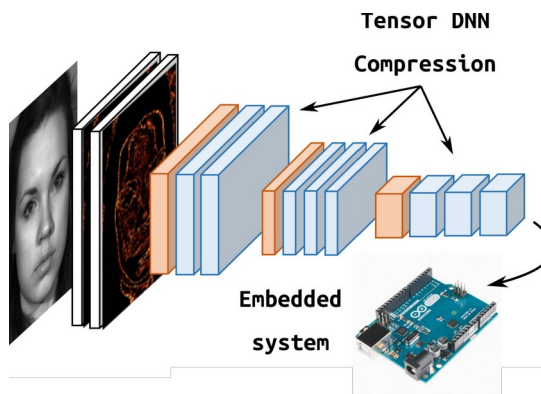
Compression des architectures d'IA

Optimisation des modèles d'apprentissage profond via la combinaison de méthodes de compression (tensorielles, quantification, parcimonie)

Contact : mohamed-oumar.ouerfelli@cea.fr

Contexte du stage

Les réseaux de neurones (NN) ont révolutionné notre façon de vivre : les médias sociaux, les systèmes de recommandation, le commerce électronique sont tous basés sur des architectures NN complexes. Cependant, l'immensité des tâches que nous associons aux grandes entreprises mondiales nous fait souvent tenir pour acquis les objets du quotidien. Un smartphone doit régler rapidement les couleurs pour obtenir une image parfaite. Une voiture autonome doit être capable de prédire une collision avec une précision d'une fraction de seconde. Non seulement la plupart des produits de haute technologie utilisent des modèles DL pour être décisifs dans le monde réel, mais ils présentent en plus le défi majeur de devoir être des systèmes rapides, petits, embarqués et portables, avec des capacités de calcul limitées, consacrées à l'efficacité et à la rapidité. Selon les applications, ces systèmes ont la particularité supplémentaire de nécessiter un certain degré de confiance : la décision d'une voiture autonome de freiner devant un mur doit être digne de confiance ! Des modèles DL plus petits et optimisés, plutôt qu'énormes et inefficaces, ont ce potentiel. Ils peuvent être rapidement interprétés et certifiés, en n'ayant à contrôler que les paramètres absolument essentiels.



Objectifs du stage

Le stage permettra d'acquérir une expérience pratique pour tester et développer différentes méthodes (au-delà de l'état de l'art (SOTA) de compression et d'élagage NN, basées sur l'optimisation du cœur des modèles d'IA via des manipulations tensorielles des architectures DL et des approches de quantification et de parcimonie. Les techniques, inspirées d'approches traditionnelles et nouvelles [3], ouvriront la possibilité d'embarquer des modèles complexes sur des appareils portables.

Le stagiaire explorera les différentes méthodes de compression, et s'intéressera en particulier à leur fusion à travers une approche d'implémentation algorithmique et/ou une approche plus fondamentale via une modélisation mathématique. Les modèles compressés devront être efficaces et conserver le plus possible la précision des architectures d'origine. Dans le cadre du stage, le candidat explorera également de nouveaux outils développés au CEA LIST [4], basés sur l'algèbre tensorielle, capables de tirer parti des performances de compression avec une compréhension plus fondamentale des structures géométriques sous-jacentes. Le développement de ces outils pourrait être suivi d'une thèse ultérieure, au sein du projet HOLIGRAIL du PEPR AI [5], consacrée à la mise en place d'un nouveau SOTA en compression NN et au développement d'outils et de logiciels prêts à être déployés.

[1] https://twitter.com/ylecun/status/1574233818298466304?t=0gqX_98O5YFC32WXssGkhw&s=09

[2] <https://github.com/nebuly-ai/exploring-AI-optimization>

[3] <https://towardsdatascience.com/neural-network-pruning-101-af816aaea61>

[4] M. Ouerfelli, M. Tamaazousti, V. Rivasseau. "Random tensor theory for tensor decomposition." In AAAI 2022.

[5] <https://www.pepr-ia.fr/>

Compétences

Le candidat devra disposer d'une bonne maîtrise de Python et une connaissance de base en réseau de neurones pour lui permettre de développer en autonomie des architectures et tester différentes méthodes de compression. Une forte connaissance de l'algèbre linéaire, *machine learning* et statistique est conseillée (décomposition SVD, *Maximum Likelihood Estimation*, stratégies d'optimisation en DL, etc.).

CEA Tech LIST

Les activités de recherche du CEA Tech LIST sont centrées sur les systèmes à logiciel prépondérant. Ces activités s'articulent autour de trois thématiques: les Systèmes Embarqués (architectures et conception de systèmes, méthodes et outils pour la sûreté des logiciels et des systèmes, systèmes de vision intelligents), les Systèmes Interactifs (ingénierie de la connaissance, robotique, réalité virtuelle et interfaces sensorielles) et les Capteurs et le traitement du signal (instrumentation et métrologie des rayonnements ionisants, capteurs à fibre optique, contrôle non destructif).

Le CEA Tech LIST a de nombreux partenariats avec les grands acteurs industriels du nucléaire, de l'automobile, de l'aéronautique, de la défense et du médical pour étudier et développer des solutions innovantes adaptées à leurs besoins. Il réalise une recherche qui va du concept de système jusqu'au démonstrateur, contribuant au transfert de technologies et à l'innovation par l'émergence de nouvelles entreprises.

Laboratoire Vision pour la Modélisation et la Localisation (LVML)

Laboratoire Vision pour la Modélisation et la Localisation (LVML) du CEA Tech LIST mène des recherches en vision par ordinateur et intelligence artificielle. Nous adressons en particulier les problématiques suivantes :

- Géolocalisation et cartographie d'environnement par vision et fusion de capteurs (robotique mobile, drones...)
- Systèmes et de vision pour la robotique : préhension, manipulation, assemblage d'objets...
- Contrôle de conformité, détection de défauts géométriques, colorimétriques, etc...
- Analyses hyperspectrales : détection de matériaux, tri,
- Correction, amélioration d'images et vidéos (superrésolution, upframing, ...)
- Compression de réseaux de neurones
- ...

Informations générales

Formation / Niveau d'étude	Ingénieur, Master 2 / Bac+5
Possibilité poursuite	Oui, en thèse ou CDD selon profil.
Durée	6 mois
Lieu	Palaiseau (91) – Centre d'intégration de Nano-INNOV
Indemnités de stage	Entre 700 € et 1400 € suivant formation. Aide au logement / transport / restauration.



Département Intelligence Ambiante et Systèmes Interactifs
Laboratoire Vision et Ingénierie des Contenus
3D & Mobilité

Candidatures

- Joindre CV + lettre de motivation à **mohamed-oumar.ouerfelli@cea.fr** avec le nom du stage auquel vous postulez
- Ne pas hésiter à détailler les projets ou cours auxquels vous avez participé
- Indiquer les dates de début/fin de stage envisagées.
- Ce stage pourra prendre une orientation recherche ou industrie en fonction du profil du candidat

Proposition de stage 2024-2025

Création d'un jeu de données pour la localisation sur des objets flexibles

Contact : fabrice.mayran-de-chamisso@cea.fr

Contexte du stage

Le CEA LVML a une longue expérience du recalage 3D d'objets rigides, de la détection d'anomalies, de la description d'images, de la reconstruction 3D, du rendu 3D, etc. Nous disposons également d'un bras robot (UR10e) qui nous a permis de collecter un jeu de données pour la localisation et la segmentation d'objets rigides.

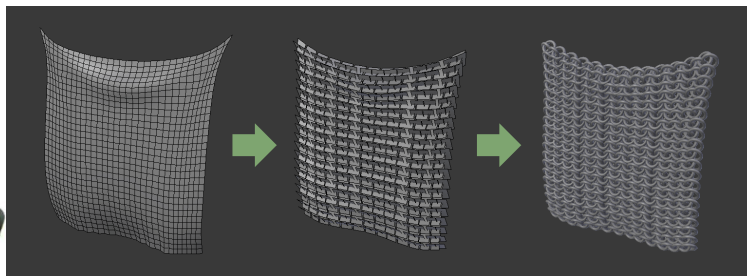
Beaucoup d'objets de la « vraie vie » ne sont pas totalement rigides : articulés, déformables, mous, etc. Afin d'entraîner des réseaux de neurones à caractériser la déformation de ces objets, il est nécessaire de disposer de jeux de données synthétiques et si possible réelles (cf. <https://dataverse.orc.gmu.edu/dataset.xhtml?persistentId=doi:10.13021/orc2020/XXLVXM>).

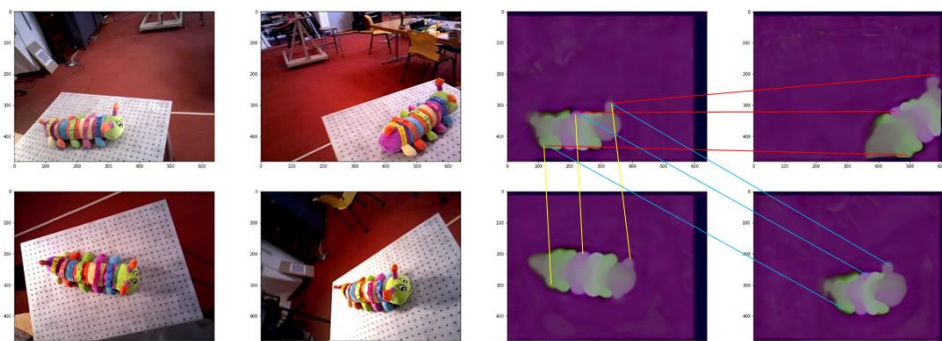
Objectifs du stage

Le stage consiste en

- 1) L'appropriation par le candidat des efforts existants de l'état de l'art sur les jeux de données d'objets flexibles et leur traitement
- 2) La constitution d'un dataset de synthèse photoréaliste d'objets déformables en utilisant un pipeline de type BlenderProc basé sur la génération procédurale via Blender (https://github.com/DLR-RM/BlenderProc/tree/main/examples/datasets/bop_challenge)
- 3) Si possible, la constitution d'un dataset réel, en utilisant des techniques novatrices développées durant le stage
- 4) L'utilisation d'outils de vision par ordinateur (réseaux de neurones notamment) pour effectuer des tâches de suivi de points, reconstruction de déformation, etc. sur les objets déformables.

Selon les résultats obtenus et le profil du candidat, une poursuite en thèse pourra être envisagée.





Compétences

Le candidat devra maîtriser deux technologies clés : le **maniement de Blender (UV mapping, compositing, rigging, animation, ...)**, et la **programmation en python (des connaissances en python sous Blender sont un plus)**. Des connaissances en apprentissage profond en **python/Pytorch** sont préférables de même que des connaissances en vision par ordinateur (géométrie projective, etc.). Des connaissances en robotique sous ROS(2) sont un plus, de même que des connaissances en C/C++.

Une forte autonomie et un esprit d'initiative sont attendus de la part du candidat.

Informations générales

Formation / Niveau d'étude	Ingénieur, Master 1 ou Master 2 / Bac+5
Possibilité poursuite	Oui, en thèse ou CDD selon profil.
Durée	4-6 mois
Lieu	Palaiseau (91) – Centre d'intégration de Nano-INNOV
Indemnités de stage	Entre 700 € et 1400 € suivant formation (grille fixe). Aide au logement / transport / restauration.

Candidatures

- Joindre CV + lettre de motivation + bulletins de notes à fabrice.mayran-de-chamisso@cea.fr avec le nom du stage auquel vous postulez
- Ne pas hésiter à détailler les projets ou cours auxquels vous avez participé
- Indiquer les dates de début/fin de stage envisagées.
- En raison du **grand nombre de candidatures reçues**, nous ne pouvons pas répondre individuellement à toutes les candidatures. Les candidats présélectionnés seront recontactés sous 3 semaines à partir de la réception de la candidature.

Proposition de stage 2024-2025

Intégration des contraintes épipolaires dans le rendu d'images depuis un nouveau point de vue

Contact : romain.dupont@cea.fr

Contexte du stage

Le CEA LIST développe une méthode de rendu de nouvelles images observée depuis un nouveau point de vue virtuel à partir d'une seule image fournie en entrée. Un réseau de neurones a ainsi été développé au LVML permettant de mieux encoder la pose relative de la caméra, en exploitant des concepts liés à la 3D tels que la contrainte épipolaire [1]. Il a été observé quelques résultats incohérents avec certaines propriétés du changement de repère. On souhaite ainsi lors de ce stage analyser ces incohérences et développer une nouvelle méthode d'apprentissage permettant d'y remédier.

[1] EpipolarNVS : exploiter la géométrie épipolaire pour la synthèse de nouvelles vues à image unique. Gaétan Landreau, Mohamed Tamaazousti (BMVC 2022)

(<https://arxiv.org/abs/2210.13077>)

Objectifs du stage

Le stage consiste en

- 1) Prendre en main le réseau de neurones générant de nouvelles images
- 2) Analyser les résultats et développer une nouvelle méthode d'apprentissage
- 3) Évaluer les résultats

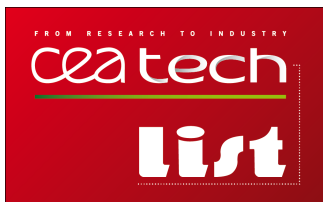
Selon les résultats obtenus et le profil du candidat, une poursuite en thèse pourra être envisagée.

Compétences

Maîtriser Python et le C++ est indispensable, ainsi qu'une expérience avec PyTorch. Des connaissances en vision par ordinateur (traitement d'images, modélisation 3D) sont un gros plus.

Informations générales

Formation / Niveau d'étude	Ingénieur, Master 1 ou Master 2 / Bac+5
Possibilité poursuite	Oui, en thèse ou CDD selon profil.
Nationalité	Européenne (contrainte projet)
Durée	4-6 mois
Lieu	Palaiseau (91) – Centre d'intégration de Nano-INNOV Accès via RER B (arrêt Le Guichet ou Massy-Palaiseau)
Indemnités de stage	Entre 700 € et 1400 € suivant formation (grille fixe). Aide au logement / transport / restauration.



Candidatures

- Joindre CV + bulletins de notes + rapports de projet à romain.dupont@cea.fr avec le nom du stage auquel vous postulez. Lettre de motivation bienvenue.
- Indiquer les dates de début/fin de stage envisagées.

CEA Tech LIST

Les activités de recherche du CEA Tech LIST sont centrées sur les systèmes à logiciel prépondérant. Ces activités s'articulent autour de trois thématiques: les Systèmes Embarqués (architectures et conception de systèmes, méthodes et outils pour la sûreté des logiciels et des systèmes, systèmes de vision intelligents), les Systèmes Interactifs (ingénierie de la connaissance, robotique, réalité virtuelle et interfaces sensorielles) et les Capteurs et le traitement du signal (instrumentation et métrologie des rayonnements ionisants, capteurs à fibre optique, contrôle non destructif).

Le CEA Tech LIST a de nombreux partenariats avec les grands acteurs industriels du nucléaire, de l'automobile, de l'aéronautique, de la défense et du médical pour étudier et développer des solutions innovantes adaptées à leurs besoins. Il réalise une recherche qui va du concept de système jusqu'au démonstrateur, contribuant au transfert de technologies et à l'innovation par l'émergence de nouvelles entreprises.

Laboratoire Vision pour la Modélisation et la Localisation (LVML)

Laboratoire Vision pour la Modélisation et la Localisation (LVML) du CEA Tech LIST mène des recherches en vision par ordinateur et intelligence artificielle. Nous adressons en particulier les problématiques suivantes :

- Géolocalisation et cartographie d'environnement par vision et fusion de capteurs (robotique mobile, drones...)
- Systèmes et de vision pour la robotique : préhension, manipulation, assemblage d'objets...
- Contrôle de conformité, détection de défauts géométriques, colorimétriques, etc...
- Analyses hyperspectrales : détection de matériaux, tri,
- Correction, amélioration d'images et vidéos (superrésolution, upframing, ...)
- Compression de réseaux de neurones
- ...

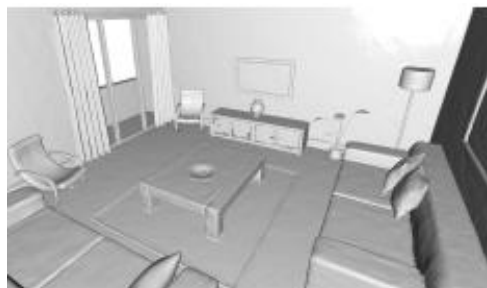
Proposition de stage 2025

Apprentissage profond pour l'interprétation de représentations 3D implicites

Contact : doriand.petit@cea.fr

Contexte du stage

Le développement de robots autonomes, qu'il s'agisse de robots terrestres comme volants, nécessite la mise en place d'intelligences artificielles (IA) capables d'interpréter l'environnement observé par ce dernier. Pour pouvoir raisonner au-delà du seul environnement observable par ses capteurs à l'instant courant, le robot doit construire une carte 3D l'environnement qu'il a observé au cours de ses déplacements et disposer de moyens pour interpréter cette carte 3D.



Exemple de reconstruction 3D basée sur une représentation neurale

Actuellement, les nuages de points 3D correspondent à la représentation 3D la plus utilisée pour cette tâche, et de nombreuses solutions de Deep Learning ont été proposées pour leur interprétation. Plus récemment, des représentations 3D exploitant des réseaux de neurones ont été introduites [1,2]. Ces représentations, généralement désignées sous le terme de Neural Fields, ont démontré des bonnes performances en terme de reconstruction, mais très peu de travaux ont investigué la manière d'interpréter directement ces représentations à l'aide d'un réseau de neurones.

[1]: Representing Scenes as Neural Radiance Fields for View Synthesis, Ben Mildenhall and Pratul P. Srinivasan and Matthew Tancik and Jonathan T. Barron and Ravi Ramamoorthi and Ren N, 2020, ECCV

[2] Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. ACM Transactions on Graphics (ToG), 41(4), 1-15.

Objectifs du stage

Ce stage a pour objectif de proposer une solution d'interprétation par réseaux de neurones de scènes 3D représentées sous forme de Neural Fields.

Le stage aura pour objectifs de :

- Etudier l'usage des architectures classiques (CNN, Transformers...) dans le cadre de l'interprétation de Neural Fields
- Etudier le co-design d'une représentation Neural Fields et de sa méthode d'interprétation afin d'obtenir le meilleur compromis qualité de reconstruction 3D/qualité d'interprétation possible.
- Evaluer les performances de la méthode proposée

Compétences souhaitées

Le candidat devra disposer d'une bonne maîtrise de python et d'une forte connaissance en réseau de neurones. Une expérience sur Pytorch sera appréciée.

Compétences développées au cours du stage

Ce stage permettra à l'étudiant de découvrir d'une part le domaine extrêmement dynamique des représentations 3D neurales, et d'autre part celui de l'interprétation de scène 3D.

Informations générales

Formation / Niveau d'étude	Ingénieur, Master 2 / Bac+5
Possibilité poursuite	Oui, en thèse ou CDD selon profil.
Durée	6 mois
Lieu	Palaiseau (91) – Centre d'intégration de Nano-INNOV
Indemnités de stage	Entre 700 € et 1400 € suivant formation. Aide au logement / transport / restauration.

Candidatures

- Joindre CV + lettre de motivation à doriand.petit@cea.fr avec le nom du stage auquel vous postulez
- Ne pas hésiter à détailler les projets ou cours auxquels vous avez participé
- Indiquer les dates de début/fin de stage envisagées.
- Ce stage pourra prendre une orientation recherche ou industrie en fonction du profil du candidat

CEA Tech LIST

Les activités de recherche du CEA Tech LIST sont centrées sur les systèmes à logiciel prépondérant. Ces activités s'articulent autour de trois thématiques: les Systèmes Embarqués (architectures et conception de systèmes, méthodes et outils pour la sûreté des logiciels et des systèmes, systèmes de vision intelligents), les Systèmes Interactifs (ingénierie de la connaissance, robotique, réalité virtuelle et interfaces sensorielles) et les Capteurs et le traitement du signal (instrumentation et métrologie des rayonnements ionisants, capteurs à fibre optique, contrôle non destructif).

Le CEA Tech LIST a de nombreux partenariats avec les grands acteurs industriels du nucléaire, de l'automobile, de l'aéronautique, de la défense et du médical pour étudier et développer des solutions innovantes adaptées à leurs besoins. Il réalise une recherche qui va du concept de système jusqu'au démonstrateur, contribuant au transfert de technologies et à l'innovation par l'émergence de nouvelles entreprises.

Laboratoire Vision pour la Modélisation et la Localisation (LVML)

Laboratoire Vision pour la Modélisation et la Localisation (LVML) du CEA Tech LIST mène des recherches en vision par ordinateur et intelligence artificielle. Nous adressons en particulier les problématiques suivantes :

- Géolocalisation et cartographie d'environnement par vision et fusion de capteurs (robotique mobile, drones...)
- Systèmes et de vision pour la robotique : préhension, manipulation, assemblage d'objets...
- Contrôle de conformité, détection de défauts géométriques, colorimétriques, etc...
- Analyses hyperspectrales : détection de matériaux, tri,
- Correction, amélioration d'images et vidéos (superrésolution, upframing, ...)
- Compression de réseaux de neurones
- ...

Proposition de stage 2025

Localisation 3D d'objets déformables dans une image par Deep Learning

Contact : boris.meden@cea.fr

Contexte du stage

La localisation 3D d'un objet dans une image consiste à estimer la position 3D et l'orientation 3D dudit objet par rapport à la caméra qui l'observe (eg. [ZebraPose]). Cette tâche est impliquée dans d'innombrables domaines d'applications (robotique autonome, Réalité Augmentée, contrôle de conformité, suivi de chantier...). Si cette problématique est étudiée depuis de nombreuses années, les solutions développées restent peu utilisées en raison de leurs contraintes de mise en œuvre, parmi lesquelles la nécessité de disposer d'un modèle 3D précis de l'objet pour entraîner la méthode ainsi que leur restriction au cas d'objets rigides.



Exemple de localisation d'objet 3D

Objectifs du stage

Ce stage aura pour objectif de mettre au point une méthode de localisation 3D d'objets déformables par Deep Learning qui soit à la fois performante et facile à mettre en œuvre.

Plus précisément, ce stage explorera l'utilisation de méthodes de reconstruction 3D de type *Neural Fields* pour permettre, d'une part, l'entraînement de méthodes de localisation 3D d'objet lorsque le modèle 3D n'est pas initialement disponible, et, d'autre part, d'étendre ces méthodes de localisation 3D aux cas des objets déformables. En effet, ces méthodes de reconstructions 3D ont révolutionné le domaine de la reconstruction 3D au cours des dernières années en raison de leur simplicité de mise en œuvre (simple caméra) [InstantNGP], y compris dans le cas d'objets déformables [Nerfies, shapeOfMotion].

Le stage aura donc pour objectifs de :

- mettre en place un processus de reconstruction 3D basée sur une représentation Neural Fields et l'exploiter pour l'entraînement d'une méthode de localisation 3D d'objet;
- étendre cette méthode de reconstruction, d'entraînement et de localisation au cas des objets déformables.

Pour y parvenir, l'étudiant bénéficiera des briques technologiques du laboratoire en termes de reconstruction 3D et de localisation 3D, ainsi que l'expertise des membres du laboratoire sur ces domaines.

[ZebraPose] Su, Y., Saleh, M., Fetzner, T., Rambach, J., Navab, N., Busam, B., ... & Tombari, F. (2022). ZebraPose: Coarse to fine surface encoding for 6dof object pose estimation, CVPR 2022.

[InstantNGP] Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (ToG)*, 41(4), 1-15.

[Nerfies] Park, K., Sinha, U., Barron, J. T., Bouaziz, S., Goldman, D. B., Seitz, S. M., & Martin-Brualla, R. (2021). Nerfies: Deformable neural radiance fields. ICCV 2021

[shapeOfMotion] Wang, Qianqian et al. "Shape of Motion: 4D Reconstruction from a Single Video." *ArXiv abs/2407.13764* (2024)

Compétences développées au cours du stage

Ce stage permettra à l'étudiant de développer ses compétences en Deep Learning et de découvrir/approfondir ses connaissances sur la localisation 3D d'objet et la reconstruction 3D. L'étudiant aura l'opportunité de travailler avec une équipe de chercheurs seniors mais aussi de doctorants.

Compétences souhaitées

Le candidat devra disposer d'une bonne maîtrise de python, de connaissances en Vision par Ordinateur, et d'une forte connaissance en réseau de neurones. Une expérience sur Pytorch sera appréciée.

Informations générales

Formation / Niveau d'étude	Ingénieur, Master 2 / Bac+5
Possibilité poursuite	Oui, en thèse ou CDD selon profil.
Durée	6 mois
Lieu	Palaiseau (91) – Centre d'intégration de Nano-INNOV
Indemnités de stage	Entre 700 € et 1400 € suivant formation. Aide au logement / transport / restauration.

Candidatures

- Joindre CV + lettre de motivation à boris.meden@cea.fr avec le nom du stage auquel vous postulez
- Ne pas hésiter à détailler les projets ou cours auxquels vous avez participé
- Indiquer les dates de début/fin de stage envisagées.
- Ce stage pourra prendre une orientation recherche ou industrie en fonction du profil du candidat

CEA Tech LIST

Les activités de recherche du CEA Tech LIST sont centrées sur les systèmes à logiciel prépondérant. Ces activités s'articulent autour de trois thématiques: les Systèmes Embarqués (architectures et conception de systèmes, méthodes et outils pour la sûreté des logiciels et des systèmes, systèmes de vision intelligents), les Systèmes Interactifs (ingénierie de la connaissance, robotique, réalité virtuelle et interfaces sensorielles) et les Capteurs et le traitement du signal (instrumentation et métrologie des rayonnements ionisants, capteurs à fibre optique, contrôle non destructif).

Le CEA Tech LIST a de nombreux partenariats avec les grands acteurs industriels du nucléaire, de l'automobile, de l'aéronautique, de la défense et du médical pour étudier et développer des solutions innovantes adaptées à leurs besoins. Il réalise une recherche qui va du concept de système jusqu'au démonstrateur, contribuant au transfert de technologies et à l'innovation par l'émergence de nouvelles entreprises.

Laboratoire Vision pour la Modélisation et la Localisation (LVML)

Laboratoire Vision pour la Modélisation et la Localisation (LVML) du CEA Tech LIST mène des recherches en vision par ordinateur et intelligence artificielle. Nous adressons en particulier les problématiques suivantes :

- Géolocalisation et cartographie d'environnement par vision et fusion de capteurs (robotique mobile, drones...)
- Systèmes et de vision pour la robotique : préhension, manipulation, assemblage d'objets...
- Contrôle de conformité, détection de défauts géométriques, colorimétriques, etc...
- Analyses hyperspectrales : détection de matériaux, tri,
- Correction, amélioration d'images et vidéos (superrésolution, upframing, ...)
- Compression de réseaux de neurones
- ...

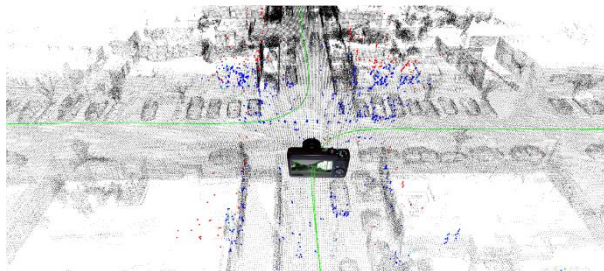
Proposition de stage 2025

Localisation d'un robot mobile dans une carte LIDAR à partir d'une camera RGB

Contact : vincent.gay-bellile@cea.fr

Contexte du stage

La localisation précise d'un robot mobile est une condition essentielle pour la navigation autonome. Une approche couramment utilisée consiste à faire correspondre les données des capteurs embarqués sur le robot à une carte préalablement établie. De nombreuses méthodes existantes reposent sur l'utilisation du même type de capteur (LiDAR ou



caméras) pour la cartographie et la localisation. Les capteurs LiDAR offrent des mesures de distance très précises et les solutions permettant de créer des cartes LIDAR à grande échelle sont largement répandues. Toutefois, ces capteurs sont souvent coûteux et encombrants, ce qui les rend difficilement intégrables sur des robots ou des drones de tailles limitées. Les caméras sont quant-à-elles peu coûteuses, légères et largement disponibles, mais il existe peu de solutions robustes pour la création de cartes visuelles à grande échelle. De plus, l'appariement entre une image et une carte visuelle est souvent sensible aux changements d'illumination et aux variations de point de vue. L'objectif de ce stage est de tirer parti des avantages de ces deux types de capteurs en utilisant un LiDAR pour la cartographie et une caméra pour la localisation.

Objectifs du stage

L'objectif de ce stage est de développer deux approches complémentaires pour la localisation d'une caméra dans une carte LiDAR :

Suivi de pose 6-DoF : La première méthode consistera à suivre en continu la pose 6 degrés de liberté (6-DoF) d'une caméra dans une carte 3D issue d'un LiDAR. Cette solution devra permettre une localisation précise et temps réel sous l'hypothèse qu'une estimation initiale approximative de la pose du robot est disponible.

Relocalisation vision/LiDAR : La seconde approche visera à estimer la localisation globale du robot dans une carte LiDAR, sans aucune information préalable sur sa position initiale. Elle a pour objectif de servir d'initialisation à la méthode de suivi et, par conséquent, ne nécessite pas de fonctionner en temps réel.

Compétences souhaitées

Le candidat devra disposer d'une bonne maîtrise du C++ et de python et d'une forte connaissance en vision par ordinateur.

Compétences développées au cours du stage

Ce stage permettra à l'étudiant de découvrir différents domaines de recherche en intelligence artificielle : la localisation, la reconstruction 3D, ... L'étudiant aura l'opportunité de travailler avec une équipe de chercheurs seniors mais aussi de doctorants.

CEA Tech LIST

Les activités de recherche du CEA Tech LIST sont centrées sur les systèmes à logiciel prépondérant. Ces activités s'articulent autour de trois thématiques: les Systèmes Embarqués (architectures et conception de systèmes, méthodes et outils pour la sûreté des logiciels et des systèmes, systèmes de vision intelligents), les Systèmes Interactifs (ingénierie de la connaissance, robotique, réalité virtuelle et interfaces sensorielles) et les Capteurs et le traitement du signal (instrumentation et métrologie des rayonnements ionisants, capteurs à fibre optique, contrôle non destructif).

Le CEA Tech LIST a de nombreux partenariats avec les grands acteurs industriels du nucléaire, de l'automobile, de l'aéronautique, de la défense et du médical pour étudier et développer des solutions innovantes adaptées à leurs besoins. Il réalise une recherche qui va du concept de système jusqu'au démonstrateur, contribuant au transfert de technologies et à l'innovation par l'émergence de nouvelles entreprises.

Laboratoire Vision pour la Modélisation et la Localisation (LVML)

Laboratoire Vision pour la Modélisation et la Localisation (LVML) du CEA Tech LIST mène des recherches en vision par ordinateur et intelligence artificielle. Nous adressons en particulier les problématiques suivantes :

- Géolocalisation et cartographie d'environnement par vision et fusion de capteurs (robotique mobile, drones...)
- Systèmes et de vision pour la robotique : préhension, manipulation, assemblage d'objets...
- Contrôle de conformité, détection de défauts géométriques, colorimétriques, etc...
- Analyses hyperspectrales : détection de matériaux, tri,
- Correction, amélioration d'images et vidéos (superrésolution, upframing, ...)
- Compression de réseaux de neurones
- ...

Informations générales

Formation / Niveau d'étude	Ingénieur, Master 2 / Bac+5
Possibilité poursuite	Oui, en thèse ou CDD selon profil.
Durée	6 mois
Lieu	Palaiseau (91) – Centre d'intégration de Nano-INNOV
Indemnités de stage	Entre 700 € et 1400 € suivant formation. Aide au logement / transport / restauration.

Candidatures

- Joindre CV + lettre de motivation à vincent.gay-bellile@cea.fr avec le nom du stage auquel vous postulez
- Ne pas hésiter à détailler les projets ou cours auxquels vous avez participé
- Indiquer les dates de début/fin de stage envisagées.
- Ce stage pourra prendre une orientation recherche ou industrie en fonction du profil du candidat

Proposition de stage 2024-2025

Intégration d'IA dans la reconstruction 3D

Contact : romain.dupont@cea.fr

Contexte du stage

Le CEA LIST développe une chaîne complète de reconstruction 3D dense depuis une séquence d'images ou vidéo. L'application principale est la cartographie 3D, qu'elle soit urbaine ou architecturale, à visée robotique, en usine par exemple, ou métrologique. Le stage s'intéressera ici à l'un des maillons faibles de cette chaîne, le calcul de la trajectoire des caméras et le calcul des points d'intérêts 3D éparses mais robustes. Cette partie appelée SFM (Structure-From-Motion) est le pilier d'une bonne reconstruction 3D et se doit d'être infaillible face aux différents contextes rencontrés : mouvements de caméra rapides, système mono ou multi-capteurs sans GPS ni IMU, environnements inhabituels ou hostiles, etc.

Malgré les récents développements en vision, cette brique reste fragile et ce stage aura pour objectif d'en robustifier ses performances en incorporant des briques basées sur l'IA, telles que détecteur de points et des algorithmes de matching basés sur l'apprentissage, et les LLM.

Objectifs du stage

Le stage consiste en

- 1) Prendre en main la chaîne algorithmique de la reconstruction 3D, notamment la brique SFM
- 2) Identifier les lacunes actuelles
- 3) Incorporer de nouveaux détecteurs de points d'intérêt basés sur l'IA
- 4) Incorporer un LLM pour améliorer les performances du SFM
- 5) Évaluer les résultats sur des grands jeux de données

Selon les résultats obtenus et le profil du candidat, une poursuite en thèse pourra être envisagée.

Compétences

Maîtriser Python et le C++ est indispensable, ainsi qu'une expérience avec PyTorch. Des connaissances en vision par ordinateur (traitement d'images, modélisation 3D, SLAM) sont un gros plus.

Informations générales

Formation / Niveau d'étude	Ingénieur, Master 1 ou Master 2 / Bac+5
Possibilité poursuite	Oui, en thèse ou CDD selon profil.
Nationalité	Européenne (contrainte projet)
Durée	4-6 mois
Lieu	Palaiseau (91) – Centre d'intégration de Nano-INNOV Accès via RER B (arrêt Le Guichet ou Massy-Palaiseau)
Indemnités de stage	Entre 700 € et 1400 € suivant formation (grille fixe). Aide au logement / transport / restauration.



Candidatures

- Joindre CV + bulletins de notes + rapports de projet à romain.dupont@cea.fr avec le nom du stage auquel vous postulez. Lettre de motivation bienvenue.
- Indiquer les dates de début/fin de stage envisagées.

CEA Tech LIST

Les activités de recherche du CEA Tech LIST sont centrées sur les systèmes à logiciel prépondérant. Ces activités s'articulent autour de trois thématiques: les Systèmes Embarqués (architectures et conception de systèmes, méthodes et outils pour la sûreté des logiciels et des systèmes, systèmes de vision intelligents), les Systèmes Interactifs (ingénierie de la connaissance, robotique, réalité virtuelle et interfaces sensorielles) et les Capteurs et le traitement du signal (instrumentation et métrologie des rayonnements ionisants, capteurs à fibre optique, contrôle non destructif).

Le CEA Tech LIST a de nombreux partenariats avec les grands acteurs industriels du nucléaire, de l'automobile, de l'aéronautique, de la défense et du médical pour étudier et développer des solutions innovantes adaptées à leurs besoins. Il réalise une recherche qui va du concept de système jusqu'au démonstrateur, contribuant au transfert de technologies et à l'innovation par l'émergence de nouvelles entreprises.

Laboratoire Vision pour la Modélisation et la Localisation (LVML)

Laboratoire Vision pour la Modélisation et la Localisation (LVML) du CEA Tech LIST mène des recherches en vision par ordinateur et intelligence artificielle. Nous adressons en particulier les problématiques suivantes :

- Géolocalisation et cartographie d'environnement par vision et fusion de capteurs (robotique mobile, drones...)
- Systèmes et de vision pour la robotique : préhension, manipulation, assemblage d'objets...
- Contrôle de conformité, détection de défauts géométriques, colorimétriques, etc...
- Analyses hyperspectrales : détection de matériaux, tri,
- Correction, amélioration d'images et vidéos (superrésolution, upframing, ...)
- Compression de réseaux de neurones
- ...