

Master Informatique
Parcours Intelligence Artificielle : Fondements et
Applications (IAFA)
Projets Chef d'oeuvre 2024-2025

Sommaire

1 Identifier des marqueurs langagiers de démence	3
2 Automatic analysis of text messages for disasters management	4
3 Restauration & segmentation d'images par Apprentissage automatique fondée sur des modèles	6
4 Optimisation de requêtes dans un système d'intégration de données	7
5 Restauration d'image par Apprentissage fondée sur des modèles	9
6 APPLICATION DE LA MÉTHODE DRPCA UNFOLDÉ (UNFOLDED) BASÉE SUR ADMM AUX JEUX DES DONNEES	10
7 Check That! Détection automatique de fake news - apprentissage automatique et chatGPT	12
8 ELOQUENT - Détection de texte générés automatiquement	13
9 Développement d'une Approche Explicable pour les Modèles d'Apprentissage par Ensemble	14
10 Étude temporelle de caractéristiques phonétiques sur des patients atteints de syndromes parkinsoniens	15
11 Analyse sensorielle et extraction d'opinions	16
12 Building Roof Reconstruction from Airborne LiDAR Point Clouds	18
13 Rendu haute qualité de nuages de points	19
14 Vers l'unification du rendu réaliste et du rendu stylisé	20
15 Arbres de Décision Multivariés	21
16 Clonage de voix	22

17 Visual Analytics of Book Content through LLM representations	23
18 Hekzam : marqueurs sobres	24
19 Hekzam : compromis précision/performance de l'OCR	26

1 Identifier des marqueurs langagiers de démence

Encadrants

Braud Chloé (chloe.braud@irit.fr) CNRS et Metheniti Eleni (Metheniti.Eleni@irit.fr)

Mots clés/keywords

TAL ; Alzheimer detection

Contexte/Context

Les difficultés cognitives accompagnant des troubles comme la maladie d'Alzheimer et maladies apparentées se manifestent dans l'expression langagière des patients. Le développement de systèmes automatiques à partir de données langagières de patients permet d'une part de mieux comprendre certains symptômes de ces maladies et leur impact sur la communication, et pourrait également un jour servir d'aide au diagnostic, en s'ajoutant à la série de tests déjà effectués par les patients. Etre capable de détecter des patients souffrant de tels troubles au cours d'une conversation pourrait permettre d'adapter la communication, par exemple dans le cadre d'un centre d'appel ou pour l'usage d'un chatbot.

Ce projet a pour but de construire un système automatique identifiant une personne souffrant de la maladie d'Alzheimer à partir d'un extrait audio. Le corpus comporte des descriptions courtes d'une image faites par des patients souffrant de démence ou des patients contrôle (en Anglais). Il est demandé de comparer l'efficacité d'un système de détection quand sont considérées uniquement les 10 premières secondes de l'extrait, ou la moitié, ou l'extrait entier, afin de déterminer la rapidité avec laquelle une identification pourrait être faite. Le système se basera sur des informations issues d'une transcription textuelle de l'extrait audio avec la possibilité de tester différents types d'architectures en se basant sur la littérature existante (des informations provenant de l'audio pourront être utilisées éventuellement).

Objectifs/Objectives

Les étudiants devront fournir un état de l'art sur la tâche. Ils devront fournir un pré-traitement des données incluant a minima une transcription audio vers texte. Ils devront enfin développer un système de classification en comparant la prise en compte en entrée de documents complets ou partiels. Le code devra être déposé dans un git et correctement documenté.

Références/References

- Luz, S., Haider, F., de la Fuente, S., Fromm, D., & MacWhinney, B. (2021). [Detecting cognitive decline using speech only: The ADReSSo Challenge](#). arXiv preprint arXiv:2104.09356.
- Cui, X., Gamage, A., Hanley, T., Mu, T. (2021). [Identifying Indicators of Vulnerability from Short Speech Segments Using Acoustic and Textual Features](#). In Interspeech (pp. 1569-1573).

2 Automatic analysis of text messages for disasters management

Encadrants

Braud Chloé (chloe.braud@irit.fr) CNRS et Metheniti Eleni (Metheniti.Eleni@irit.fr)

Mots clés/keywords

NLP ; Tweet analysis ; Crisis

Contexte/Context

Recently, Twitter has been widely used to generate valuable information in crisis situations. All types of crises are concerned including ecological crises (flood, hurricanes, earthquakes, etc.), sudden crises (Marseille building collapse, Notre Dame burns, the Lubrizol chemical plant in Rouen, etc.), security crises (Gilets Jaunes demonstration, etc.) and sanitary crises (e.g. SRAS, Ebola, Covid-19). During each crisis, a huge amount of messages are posted, some of them could be useful for population and rescue departments. It is important to note that such useful messages do not always require an urgent and rapid action from rescue teams: messages about affected people, or infrastructure damages can be seen as more urgent compared to others. Also, both urgent and non urgent messages are drowned in a deluge of off-topic or personal messages that may contain crisis-related keywords. Therefore, new tools are needed to early access useful information that will allow the emergency units to anticipate actions, coordinate efforts, and share information.

Objectifs/Objectives

"The students have to read and summarize related work on the task. They have to propose a system able to take a set of tweets related to a crisis as input, and that will output an ordered set of extracted information that could be given, for example, to emergency services. In order to build such a system, we will investigate:

- Automatic classification : used to build a system able to identify important messages and classify them according to some categories.
- Adapting to new language and/or domain : ideally, we'd like to have a system able to tackle a new crisis or which we have no, or only a few amount of annotated data. What can we do if we have all these data about storms, but we have to analyze data for a hurricane? Ideally, we'd like to be able to analyze data for all languages, especially when facing an international crisis. How can we adapt to a new language?
- Summarization : the final part of the project is to extract the most important information and order the messages, in order to provide an useful and readable summary. This can be done in several ways, and the students will have to propose a solution.

Références/References

- Imran, M., Castillo, C., Diaz, F., and Vieweg, S. (2018). [Processing social media messages in mass emergency: Survey summary](#). In Companion Proceedings of the The Web Conference 2018, pages 507–511.
- Imran, M., Mitra, P., and Castillo, C. (2016). [Twitter as a lifeline: Human-annotated twitter corpora for nlp of crisis-related messages](#). In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), Paris, France. European Language Resources Association (ELRA).
- Kedzie, C., McKeown, K., Diaz, F. (2015, July). [Predicting salient updates for disaster summarization](#). In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) (pp. 1608-1617).

3 Restauration & segmentation d'images par Apprentissage automatique fondée sur des modèles

Encadrants

Denis KOUAME (kouame@irit.fr) IRIT

Mots clés/keywords

Apprentissage, Segmentation, restauration d'image, artéfact

Contexte/Context

"De nombreux travaux ont été réalisés dans le domaine de l'apprentissage avec de nombreux inconvénients parmi lesquels, on trouve l'apprentissage supervisés 'aveugle' (ou end-to-end). Pour rendre le processus d'apprentissage (et les résultats) explicables, on a embarqué de plus des modèles dans l'apprentissage. L'un des procédés le plus connu est celui dit de déroulement (unfolding). C'est dans ce cadre se situe ce projet où on s'intéresse à la restauration d'image et à la segmentation dans les images.

Objectifs/Objectives

L'objectif de ce travail dans un premier à se familiariser au procédé d'apprentissage de l'unfolding, de comparer des méthodes existantes, en les implémentant en les enrichissant et en les testant sur des images standards et médicales.

Références/References

- Rapport de stage, Jessy KHAFIF, M2 IAFA, Université Paul Sabatier, 2022-2023

4 Optimisation de requêtes dans un système d'intégration de données

Encadrants

YIN Shaoyi (shaoyi.yin@irit.fr) IRIT

Mots clés/keywords

Optimisation de requêtes, intégration de données

Contexte/Context

L'intégration de données hétérogènes a été un sujet actif depuis de nombreuses années. Il y avait déjà des fondements théoriques [5] et des applications industrielles [4] au début des années 2000. Avec de plus en plus d'applications « Big Data » et de tâches « Science de données », les limites de ces théories et systèmes ont été révélées et nous sommes confrontés à de nouveaux défis sur l'intégration de données [2][3][6]. Veuillez d'abord prendre connaissance de ces fondements et nouveaux défis à travers des articles de recherche, et ensuite développer un mini-prototype d'un système d'intégration de données avec l'approche Médiateur/Adaptateurs [7], afin de donner votre propre avis via des expérimentations. Attention, nous ne pouvons pas traiter tous les aspects sur l'intégration de données dans ce projet.

Objectifs/Objectives

En accord avec les enseignements effectués en parcours IAFA, vous allez vous focaliser sur l'aspect optimisation de requêtes [1]. Voici une feuille de route pour la réalisation du projet : (1) lecture d'articles ; (2) choix de sources de données à intégrer ; (3) conception de l'architecture du système ; (4) développement et validation du prototype ; (5) rédaction du rapport.

Références/References

- 1 S. Adali, K. S. Candan, Y. Papakonstantinou, V. S. Subrahmanian. Query Caching and Optimization in Distributed Mediator Systems. SIGMOD Conference 1996: 137-148.
- 2 A. Doan, P. Konda, et. al. Toward a System Building Agenda for Data Integration (and Data Science). IEEE Data Engineering Bulletin, Volume 41, Number 2, June 2018, pp.35-46.
- 3 B. Golshan, A. Halevy, G. Mihaila, W-C. Tan. Data Integration: After the Teenage Years. PODS 2017: 101-106.
- 4 A.Y., Halevy, A., Rajaraman, J. J., Ordille. Data Integration: The Teenage Years. VLDB 2006: 9-16.
- 5 M., Lenzerini. Data Integration: A Theoretical Perspective. PODS 2002: 233-246.
- 6 M. Stonebraker and I. F. Ilyas. Data Integration: The Current Status and the Way Forward. IEEE Data Engineering Bulletin, Volume 41, Number 2, June 2018, pp.3-9.

- 7 G. Wiederhold. Mediators in the Architecture of Future Information Systems. Computer 25(3): 38-49 (1992).

5 Restauration d'image par Apprentissage fondée sur des modèles

Encadrants

Denis KOUAME (kouame@irit.fr) IRIT

Mots clés/keywords

Apprentissage, JPEG, restauration d'image, artéfact

Contexte/Context

De nombreux travaux ont été réalisés dans le domaine de l'apprentissage avec de nombreux inconvénients parmi lesquels, on trouve l'apprentissage supervisés 'aveugle' (ou end-to-end). Pour rendre le processus d'apprentissage (et les résultats) explicables, on a embarqué de plus des modèles dans l'apprentissage. L'un des procédés le plus connu est celui dit de déroulement (unfolding). C'est dans ce cadre se situe ce projet où on s'intéresse à l'élimination d'artéfacts dans les images.

Objectifs/Objectives

L'objectif de ce travail est de comparer des méthodes existantes, en les implémentant et en les testant sur des images standards et médicales.

Références/References

- Xueyang Fu , Menglu Wang, Xiangyong Cao , Xinghao Ding , and Zheng-Jun Zha , ""A Model-Driven Deep Unfolding Method for JPEG Artifacts Removal"", IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS
- Ming-Yu Liu, Thomas Breuel, Jan Kautz, [Unsupervised Image-to-Image Translation Networks](#) arxiv.

6 APPLICATION DE LA MÉTHODE DRPCA UNFOLDÉ (UNFOLDED) BASÉE SUR ADMM AUX JEUX DES DONNEES

Encadrants

PHAM Duong Hung; KOUAME Denis (duong-hung.pham@irit.fr); IRIT

Mots clés/keywords

Blind Deconvolution (BD), Blood Flow, Clutter Separation, Robust Principal Component Analysis (RPCA), Power Doppler, High Resolution

Contexte/Context

Dans l'imagerie de contraste ultrasonore par ultrasons (en anglais : contrast-enhanced ultrasound imaging CEUS), le problème de la reconstruction des vaisseaux sanguins et de la suppression des tissus des données brutes a également suscité un grand intérêt de la part de la communauté. Par exemple, les techniques RPCA itératives consistent à formuler un modèle des données et puis à chercher une solution en optimisant un problème inverse associé à ce modèle. Ce modèle permet une application de régularisations, par exemple, la parcimonie du sang ou le rang faible des tissus, pour utiliser de la connaissance préalable des données d'entrée pour une reconstruction efficace du flux sanguin. Afin d'améliorer encore la qualité de la reconstruction, une étape de traitement de déconvolution a été récemment introduite dans le cadre RPCA [1]. Cette approche est appelée RPCA déconvoluée (DPRCA).

Cependant, l'applicabilité de cet algorithme est encore limitée en raison du réglage manuel des hyperparamètres. Pour traiter ce problème, l'approche déroulement (unfolding) est de plus en plus utilisée car elle permet d'exprimer des algorithmes itératifs sous forme de réseaux de neurones. Plus précisément, chaque étape d'itération d'un algorithme itératif est convertie en une couche du réseau et finalement l'ensemble de ces couches forme un CNN profond. Inspiré par cette idée, nous avons récemment développé un algorithme CNN unfoldé basé sur le déroulement de l'algorithme DRPCA composé d'étapes ADMM (appelé DPCANet par la suite). Cet algorithme proposé a non seulement montré une amélioration significative de l'estimation du flux sanguin, mais a également fourni une estimation automatique des hyperparamètres [2].

Objectifs/Objectives

Les objectifs du travail :

- Lisez les 2 articles pour comprendre la méthode DRPCANet et des codes associés en Python.
- Appliquer cette méthode à autres jeux des données (simulées et réelles).

Références/References

- 1 D. H. Pham, A. Basarab, I. Zemmoura, J. -P. Remenieras and D. Kouam e," Joint Blind Deconvolution and Robust Principal Component Analysis for Blood Flow Estimation in

Medical Ultrasound Imaging,” in IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 68, no. 4, pp. 969-978, April 2021.

- 2 P. Vassili et al, «Deep Unfolding RPCA for High Resolution Flow Estimation», IUS 2022, Venice, Italy.

7 Check That! Détection automatique de fake news - apprentissage automatique et chatGPT

Encadrants

Josiane Mothe, josiane.mothe@irit.fr, IRIT

Mots clés/keywords

fake news; detection; NLP

Contexte/Context

Les fausses informations mélangent généralement des informations réelles et des informations fausses pour tromper les lecteurs. D'autre part, les producteurs de fausses nouvelles utilisent couramment le langage subjectif.

Objectifs/Objectives

Ce projet vise à développer des modèles d'apprentissage automatique pour indiquer si une phrase issue d'un article de presse exprime un point de vue subjectif ou objectif de l'auteur. Les données sont en différentes langues (anglais, espagnol, arabe, italien, ...). Les modèles utilisés pourront s'appuyer sur des technologies de type ChatGPT (large language models), LangChain ou/et sur des technologies à base de dictionnaires et de modèles comme les forêts aléatoires. Les résultats de différents modèles ou différentes configurations de modèles seront présentés (évaluation de l'efficacité) sur plusieurs jeux de données. Les données fournies sont annotées.

Références/References

- 1 Alberto Barrón-Cedeño et al. Overview of the CLEF-2024 CheckThat! Lab: Check-Worthiness, Subjectivity, Persuasion, Roles, Authorities, and Adversarial Robustness https://clef2024.clef-initiative.eu/publications_lncs/637930_1_En_Online_Combine_p2.pdf#page.28
- 2 Galassi, A., Ruggeri, F., Barrón-Cedeño, A., Alam, F., Caselli, T., Kutlu, M., ... & Zaghouani, W. (2023). Overview of the CLEF-2023 CheckThat! lab task 2 on subjectivity in news articles. Working Notes of CLEF. <https://www.dei.unipd.it/~faggioli/temp/CLEF2023-proceedings/paper-020.pdf>
- 3 Géraud Faye et al. HYBRINFOX at CheckThat! 2024 - Task 2: Enriching BERT Models with the Expert System VAGO for Subjectivity Detection. <https://ceur-ws.org/Vol-3740/paper-37.pdf>
- 4 Pooja Premnath et al. SSN-NLP at CheckThat! 2024: From Feature-based Algorithms to Transformers: A Study on Detecting Subjectivity. <https://ceur-ws.org/Vol-3740/paper-52.pdf>

8 ELOQUENT - Détection de texte générés automatiquement

Encadrants

Josiane Mothe, josiane.mothe@irit.fr, IRIT

Mots clés/keywords

chat GPT, IA générative, détection automatique

Contexte/Context

De nombreux texte sont générés par des IA. Comment détecter qu'un texte a été écrit par un humain ou par une machine? L'étude portera sur des textes en anglais et en français et potentiellement d'autres langues, en fonction du groupe d'étudiants sélectionné.

Objectifs/Objectives

Proposer un (ou plusieurs) modèle d'apprentissage automatique testé sur des collections afin de détecter si un texte a été écrit par un humain ou une machine.

Références/References

- 1 <https://eloquent-lab.github.io/>
- 2 Evaluating Poro-34B-Chat and Mistral-7B-Instruct-v0.1: LLM System Description for ELOQUENT at CLEF 2024708-711 Vasumathi Neralla, Sander Bijl de Vroe <https://ceur-ws.org/Vol-3740/paper-68.pdf>
- 3 ELOQUENT CLEF Shared Tasks for Evaluation of Generative Language Model Quality , Karlgren, Jussi et al. <https://swepub.kb.se/bib/swepub:oai:DiVA.org:ri-72876?tab2=abs&language=en>

9 Développement d'une Approche Explicable pour les Modèles d'Apprentissage par Ensemble

Encadrants

Moncef GAROUANI, moncef.garouani@irit.fr, IRIT

Mots clés/keywords

Machine Learning, Ensemble Learning, Explicabilité de l'IA (XAI)

Contexte/Context

L'apprentissage par ensemble (ensemble learning) est une technique puissante en machine learning, qui combine plusieurs modèles de base (comme les arbres de décision, SVM, ou réseaux de neurones) pour améliorer les performances prédictives. Cependant, malgré les gains en précision, ces modèles d'ensemble deviennent souvent opaques et difficiles à interpréter, ce qui pose un défi majeur pour leur adoption dans des domaines sensibles tels que la médecine, la finance ou le droit.

Ce projet vise à explorer des techniques d'explicabilité (XAI) afin de rendre les modèles d'ensemble (stacking) plus transparents et compréhensibles, tout en maintenant leur performance.

Objectifs/Objectives

- Étudier les méthodes actuelles de XAI : Examiner les techniques courantes comme LIME, SHAP, pour comprendre comment elles peuvent être appliquées à des modèles d'ensemble.
- Développer une méthodologie d'explicabilité : Proposer une approche qui permet d'expliquer les prédictions d'un modèle d'ensemble (stacking) tout en maintenant ses performances.
- Implémentation et comparaison : Implémenter plusieurs algorithmes d'apprentissage par ensemble et tester différentes méthodes d'explicabilité sur ces modèles. Comparer la précision des prédictions et le niveau d'interprétabilité.

Références/References

- 1 <https://www.baeldung.com/cs/bagging-boosting-stacking-ml-ensemble-models>
- 2 <https://www.javatpoint.com/stacking-in-machine-learning>
- 3 <https://medium.com/ai-for-tomorrow/lintelligence-artificielle-explicable-e3c3a9mystifier-les-bo%C3%A0tes-noires-e0c2a7c8a036>

10 Étude temporelle de caractéristiques phonétiques sur des patients atteints de syndromes parkinsoniens

Encadrants

Mauclair Julie, julie.mauclair@irit.fr, IRIT

Mots clés/keywords

Traitement Automatique de la parole, parole pathologique

Contexte/Context

Projet ANR Voice4PD pour la différenciation précoce entre les maladie de Parkinson et l'atrophie multi systématisée.

Objectifs/Objectives

Étudier l'évolution de paramètres audio au cours d'un enregistrement de patients.

Références/References

- 1 <https://hal.science/hal-02798552v2/file/58.pdf>

11 Analyse sensorielle et extraction d'opinions

Encadrants

Isabelle Ferrané, isabelle.ferrane@irit.fr, IRIT

Mots clés/keywords

Traitement automatique de la parole et du langage ; Compréhension de la parole (NLU/SLU) ; Analyse de sentiments ; adaptation de modèles à un domaine de spécialité ; classification non supervisée

Contexte/Context

Projet régional RESENTI (Variétés RESistantes Analyses SENSorielles eT Innovations ouvertes), projet transdisciplinaire en lien avec les domaines de l'agronomie, de la viticulture (cépages, viticulture, ...), de l'oenologie (qualité/défaut des vins dégustés) et du marketing (arguments de vente, ...). La contribution de l'IRIT est de traiter les enregistrements audio des focus group (enregistrements de conversations lors de sessions de dégustation de vins) pour extraire les arguments pouvant constituer des leviers et les freins à l'adoption des variétés résistantes.

Objectifs/Objectives

Intégrer des méthodes d'analyse de sentiments et de classification non supervisée pour caractériser les locuteurs et les opinions formulées lors de focus groups, en termes d'opinions positives, négatives ou neutre. Il s'agit aussi bien de travailler sur la phase amont (génération des transcriptions) et sur la phase aval (analyse des tours de parole transcrits) en proposant/adaptant/testant des méthodes et modèles existants. L'objectif est d'une part d'améliorer la qualité des transcriptions automatiques (robustesse vis à vis des domaines de spécialité, des environnements bruités, ...) et d'adapter des modèles existants au domaine de spécialité (viticulture, agronomie, commerce) et d'autre part d'appliquer des méthodes de clustering pour regrouper les arguments en faveur ou pas de l'adoption des variétés résistantes (cépages spécifiques). On pourra éventuellement explorer des méthodes d'analyse de sentiments et/ou d'argument mining" ou encore des approches E2E pour le SLU (Spoken Language Understanding).

Références/References

- <http://argumentationmining.disi.unibo.it/publications/ArgMin2018b.pdf>
- <https://doi.org/10.1146/annurev-statistics-030718-105242https://dl.acm.org/doi/pdf/10.1145/2436256.2436274>
- <https://cdn.openai.com/papers/whisper.pdf>
- <https://huggingface.co/models?other=sentiment-analysishttps://scikit-learn.org/stable/modules/generated/sklearn.cluster.SpectralCluster.html>

- [https://www.sciencedirect.com/science/article/abs/pii/S0885230822000134?
via%3Dihub](https://www.sciencedirect.com/science/article/abs/pii/S0885230822000134?via%3Dihub)

12 Building Roof Reconstruction from Airborne LiDAR Point Clouds

Encadrants

Adrien Nautré; Loic Barthe (adrien.nautre@rollingdot.ai; loic.barthe@irit.fr), RollingDot; IRIT

Mots clés/keywords

Machine Learning, 3D Geometry, Point Cloud

Contexte/Context

The Institut national de l'information Géographique et forestière (IGN) has carried out an airborne LiDAR round acquisition campaign in France. These data cover a large proportion of towns and cities with many buildings. The points clouds acquired have already been classified by IGN by hand. Many applications require the detection of building roofs and the calculation of their surface areas.

Objectifs/Objectives

In this context, the aim is to study the feasibility of automating the classification of building roofs and the calculation of their surfaces from airborne LiDAR data. The expected outcome of the project is a description of the processing of point clouds and a benchmark of the methods used.

Références/References

- Cartographie 3D du sol et du sursol de la France : <https://geoservices.ign.fr/lidarhd>
- PCEDNet : A Lightweight Neural Network for Fast and Interactive Edge Detection in 3D Point Clouds : <https://www.irit.fr/recherches/STORM/MelladoNicolas/publications-pcednet/>
- Machine learning-based segmentation of aerial LiDAR point cloud data on building roof : <https://www.tandfonline.com/doi/full/10.1080/22797254.2023.2210745#d1e909>
- Point2Roof: End-to-end 3D building roof modeling from airborne LiDAR point clouds : <https://www.sciencedirect.com/science/article/abs/pii/S0924271622002362>
- Seed point set-based building roof extraction from airborne LiDAR point clouds using a top-down strategy : <https://www.sciencedirect.com/science/article/abs/pii/S0926580521001114>

13 Rendu haute qualité de nuages de points

Encadrants

VANDERHAEGHE David david.vanderhaeghe@irit.fr IRIT

Mots clés/keywords

informatique graphique, nuage de points, rendu, gpu

Contexte/Context

Les nuages de points sont la représentation classique acquise par les scanners 3D (LiDAR par exemple). L’affichage direct des points par de simples primitives ponctuelles présentes de nombreux artefacts (trous, discontinuité de couleurs). Nous souhaitons explorer les techniques de rendu de nuages de points permettant de limiter ces artefacts, sans passer par une représentation maillée.

Objectifs/Objectives

Après un état de l’art sur les récentes avancées dans le domaine du rendu de nuages de points, l’objectif est d’implémenter un démonstrateur d’une ou plusieurs techniques de rendu. L’environnement pourra être au choix Unreal engine, ou un démonstrateur ad hoc en c++/opengl. L’accent devra être mis sur l’apparence du rendu. Les questions de gestion de niveau de détails et de structures d’accélération pourront être mises de côté.

Références/References

- "Kerbl, B., Kopanas, G., Leimkuehler, T., & Drettakis, G. (2023). 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Transactions on Graphics, 42(4), 1–14. <https://doi.org/10.1145/3592433><https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/>
- Schütz, M., Herzberger, L., & Wimmer, M. (2024). SimLOD: Simultaneous LOD Generation and Rendering for Point Clouds. Proceedings of the ACM on Computer Graphics and Interactive Techniques, 7(1), 1–20. <https://doi.org/10.1145/3651287><https://www.cg.tuwien.ac.at/research/publications/2024/SCHUETZ-2024-SIMLOD/SCHUETZ-2024-SIMLOD-paper.pdf>

14 Vers l'unification du rendu réaliste et du rendu stylisé

Encadrants

Mathias Paulin Mathias.Paulin@irit.fr IRIT-STORM

Mots clés/keywords

Méthodes de Monte-Carlo, estimateurs de l'équation du rendu, fonctions de style, fonctions de l'espérance d'une variable aléatoire

Contexte/Context

Ce projet se positionne dans le domaine du rendu en informatique graphique. Les algorithmes de rendu sont généralement séparés en deux grandes familles : le rendu réaliste, reposant sur la méthode de Monte Carlo pour calculer un estimateur de l'équation des transferts radiatifs pour assurer une réalité physique à l'image et le rendu expressif reposant sur l'évaluation de fonctions de style, indépendamment de toute réalité physique, pour se concentrer sur l'aspect visuel et esthétique de l'image. L'objectif de ce projet est d'étudier et de reproduire les résultats publiés à siggraph 2024 et proposant une unification de ces deux familles dans un seul modèle de calcul pour produire des images mêlant réalité physique et style esthétique.

Objectifs/Objectives

Les attendus sont une analyse de l'état de l'art du domaine à partir de l'article "Stylized Rendering as a Function of Expectation, R. West and S. Mukherjee, ACM Transactions on Graphics (SIGGRAPH 2024)", et la production d'un prototype logiciel, fondé par exemple sur pbrt-4 en C++, permettant de reproduire les résultats de l'article

Références/References

- Stylized Rendering as a Function of Expectation, R. West and S. Mukherjee, ACM Transactions on Graphics (SIGGRAPH 2024), <https://dl.acm.org/doi/10.1145/3658161> et <http://cv.rexwe.st/>

15 Arbres de Décision Multivariés

Encadrants

COOPER Martin cooper@irit.fr IRIT

Mots clés/keywords

Machine Learning, Explicabilité

Contexte/Context

Construction d'algorithmes efficaces pour expliquer des décisions prises par un arbre de décision multivarié.

Objectifs/Objectives

Un logiciel pour la construction d'arbres de décision multivariés (la condition à chaque noeud peut dépendre de plusieurs variables) à partir d'un dataset. Ce logiciel sera paramétré par le type de condition associée aux noeuds. Dans un deuxième temps, un logiciel qui retourne une explication d'une décision prise par un tel arbre de décision.

Références/References

- Tractable Explaining of Multivariate Decision Trees, Carbonnel, Cooper, Marques-Silva, KR 2023, <https://doi.org/10.24963/kr.2023/13>

16 Clonage de voix

Encadrants

Pellegrini Thomas thomas.pellegrini@irit.fr IRIT

Mots clés/keywords

parole, voix, synthèse de la parole, traitement automatique de la parole

Contexte/Context

Pour entraîner des modèles d'analyse de parole, par exemple de détection de mots clés, d'émotions, de transcription parole vers texte, nous avons besoin d'enregistrements sonores, qui parfois, n'existent pas. C'est le cas par exemple de données de parole d'enfants, en français en particulier. Les systèmes de synthèse de parole (TTS pour Text-To-Speech) sont a priori assez performants de nos jours pour tenter de répondre à ce besoin.

Ce projet vise à étudier l'état de l'art en TTS, en particulier lorsqu'on utilise les systèmes en mode "clonage de voix" (voice cloning), pour générer des données de parole synthétique. Nous utiliserons plusieurs datasets de parole réelle, et nous comparerons l'utilisation de données de parole réelle versus parole synthétique clonée pour entraîner des modèles de classification (réseaux de neurones) pour des tâches d'analyse de parole données.

Objectifs/Objectives

- État de l'art sur les systèmes TTS avec clonage
 - Génération de données de parole synthétique avec un ou deux systèmes TTS open-source (par exemple WhisperSpeech et XTTS),
 - Mettre en place des filtres pour rejeter les mauvaises générations (hallucinations possibles...),
 - Entraîner des classifieurs sur des données synthétiques,
 - Caractériser les différences entre parole réelle et synthétique (analyse de représentations auto-supervisées, par exemple avec WavLM), avec de l'analyse en composante principale ou autre.

Références/References

- Coqui TTS - a library for advanced text-to-speech generation, <https://github.com/coqui-ai/TTS>, 2023
- WhisperSpeech, <https://github.com/collabora/WhisperSpeech>
- K. C. Yuen, L. Haoyang and C. E. Siong, "ASR Model Adaptation for Rare Words Using Synthetic Data Generated by Multiple Text-To-Speech Systems," In Proc. APSIPA ASC, Taipei, 2023, pp. 1771-1778
- A. Werchaniak, et al, "Exploring the application of synthetic audio in training keyword spotters," in Proc. ICASSP, 2021, <https://assets.amazon.science/20/8b/696fd11d4d99a69e2adfc4602ccd/exploring-the-application-of-synthetic-audio-pdf>

17 Visual Analytics of Book Content through LLM representations

Encadrants

Hurter Christophe (christophe.hurter@enac.fr) & Muller Philippe (philippe.muller@irit.fr), ENAC + IRIT

Mots clés/keywords

NLP, data visualization, clustering

Contexte/Context

This project aims to visualize the content of a book through the lens of tokenization or semantic vectors. Large Language Models (LLMs) transform text into tokens, which are essentially high-dimensional vectors. By processing text, such as a few pages of a book, the content can be converted into a set of high-dimensional vectors. These vectors can then be mapped onto a two-dimensional space using dimensionality reduction techniques (ie t-sne, umap). The resulting visualization will reveal clusters of similar semantics.

This process can be repeated using a sliding window approach, where overlapping sections of the book are tokenized sequentially. The goal is to create a series of images that represent the evolving content of the book through its tokens. The objective of this project is to generate these images, explore ways to minimize changes between consecutive visualizations, analyze the images, and connect them with the underlying narrative of the book. Additionally, data processing techniques such as edge bundling can be applied to cluster similar tokens and track their evolution, providing insights for visual analytics.

Objectifs/Objectives

Literature survey, analytics tools, experiments

Références/References

- Riehmann, P., Hanfler, M., & Froehlich, B. (2005, October). Interactive sankey diagrams. In IEEE Symposium on Information Visualization, 2005. INFOVIS 2005. (pp. 233-240). IEEE.

18 Hekzam : marqueurs sobres

Hekzam : marqueurs sobres

Contexte

Hekzam est un ensemble de logiciels en cours de conception qui a pour objectif de permettre l'écriture d'examens papiers qui peuvent être corrigés automatiquement ou pseudo-automatiquement. Voici la manière typique dont Hekzam est utilisé pour faire et corriger un examen.

1. L'équipe enseignante crée un formulaire de sujet d'examen grâce aux logiciels fournis de Hekzam
2. Les apprenant-e-s passent leur examen en complétant le formulaire
3. L'équipe enseignante scanne les formulaires complétés
4. L'équipe enseignante utilise les logiciels fournis de Hekzam pour obtenir
 - Des notes (et analyses détaillées) pour les questions à correction complètement automatisée
 - Les saisies des apprenant-e-s (images, booléens, nombres...) pour les autres types de questions

Hekzam introduit deux types de marqueurs lors de la génération des formulaires de sujets d'examen. Le premier type de marqueur est placé sur chaque bord du formulaire, et a pour rôle d'aider au traitement des images scannées pour les recadrer au mieux, une feuille scannée n'étant jamais parfaitement alignée avec le sens initial. Le second type de marqueur a pour but d'identifier la page du formulaire : il faut pouvoir retrouver plus tard lors de l'analyse des scans le sujet dont le quel la page fait partie, et le numéro du page du sujet en question.

Objectifs du projet

Hekzam utilise actuellement des QR codes comme marqueurs des bords du formulaire, ces marqueurs servant à la fois de marqueurs de bord et d'identification de la page. Ces marqueurs sont robustes (ils sont presque toujours détectés et leur emplacement est plutôt bon) mais coûteux en encre utilisée, coûteux en espace requis sur chaque page de formulaire, et coûteux en terme de calcul pour leur génération et leur détection.

Le but du projet est d'optimiser les marqueurs utilisés par Hekzam pour les rendre plus sobres tout en gardant une très bonne robustesse. Une veille bibliographique et technologique sera réalisée sur les marqueurs utilisés en traitement d'image et sur leur pertinence pour ce cas d'utilisation. Différentes alternatives aux marqueurs actuels seront implémentés dans Hekzam (code de génération en Typst, code actuel de traitement d'images en C++ avec OpenCV et ZXing) et comparés sur les métriques de place occupée sur les copies, d'encre utilisée, de temps de calcul nécessaire à leur génération et à leur détection, de robustesse de détection des marqueurs, et de précision de détection de l'emplacement des marqueurs.

Candidature

Contactez-moi par mail avec les informations suivantes.

- Très courte motivation par rapport au sujet (2 phrases)
- Bulletins de notes (master et licence)
- CV court (parcours d'études, expériences professionnelles, compétences)

19 Hekzam : compromis précision/performance de l'OCR

Hekzam : compromis précision/performance de l'OCR

Contexte

Hekzam est un ensemble de logiciels en cours de conception qui a pour objectif de permettre l'écriture d'examens papiers qui peuvent être corrigés automatiquement ou pseudo-automatiquement. Voici la manière typique dont Hekzam est utilisé pour faire et corriger un examen.

1. L'équipe enseignante crée un formulaire de sujet d'examen grâce aux logiciels fournis de Hekzam
2. Les apprenant-e-s passent leur examen en complétant le formulaire
3. L'équipe enseignante scanne les formulaires complétés
4. L'équipe enseignante utilise les logiciels fournis de Hekzam pour obtenir
 - Des notes (et analyses détaillées) pour les questions à correction complètement automatisée
 - Les saisies des apprenant-e-s (images, booléens, nombres...) pour les autres types de questions

Hekzam permet de définir plusieurs types de champs dans un formulaire. Certains de ces champs nécessitent de faire de la reconnaissance optique de caractères (OCR) pour transformer la saisie de l'apprenant-e en donnée utilisable par un traitement automatisé. On va ici s'intéresser à deux champs.

- Champ de type chiffre unique. Par exemple une séquence de 8 champs pour un numéro étudiant.
- Champ de type suite de mots. Par exemple pour indiquer son prénom + nom.

Objectifs du projet

Le but principal du projet est d'évaluer les compromis entre précision et performance d'algorithmes et de technologies d'OCR utilisables par Hekzam. Il est à noter que Hekzam est conçu dans une démarche de sobriété numérique et a pour but de pouvoir être exécuté sur des ordinateurs très peu puissants.

Une étude bibliographique sera réalisée sur cette question, pour avoir un aperçu de l'état actuel du problème côté publications scientifiques. En parallèle, une étude expérimentale sera réalisée pour évaluer la précision et le coût de traitement de différentes technologies d'OCR libres et open source. Des modèles de machine learning *ad hoc* pourront être entraînés et évalués à condition que l'apprentissage soit très peu coûteux et que ces modèles soient utilisables sur un ordinateur très peu puissant.

Candidature

Contactez-moi par mail avec les informations suivantes.

- Très courte motivation par rapport au sujet (2 phrases)
- Bulletins de notes (master et licence)
- CV court (parcours d'études, expériences professionnelles, compétences)