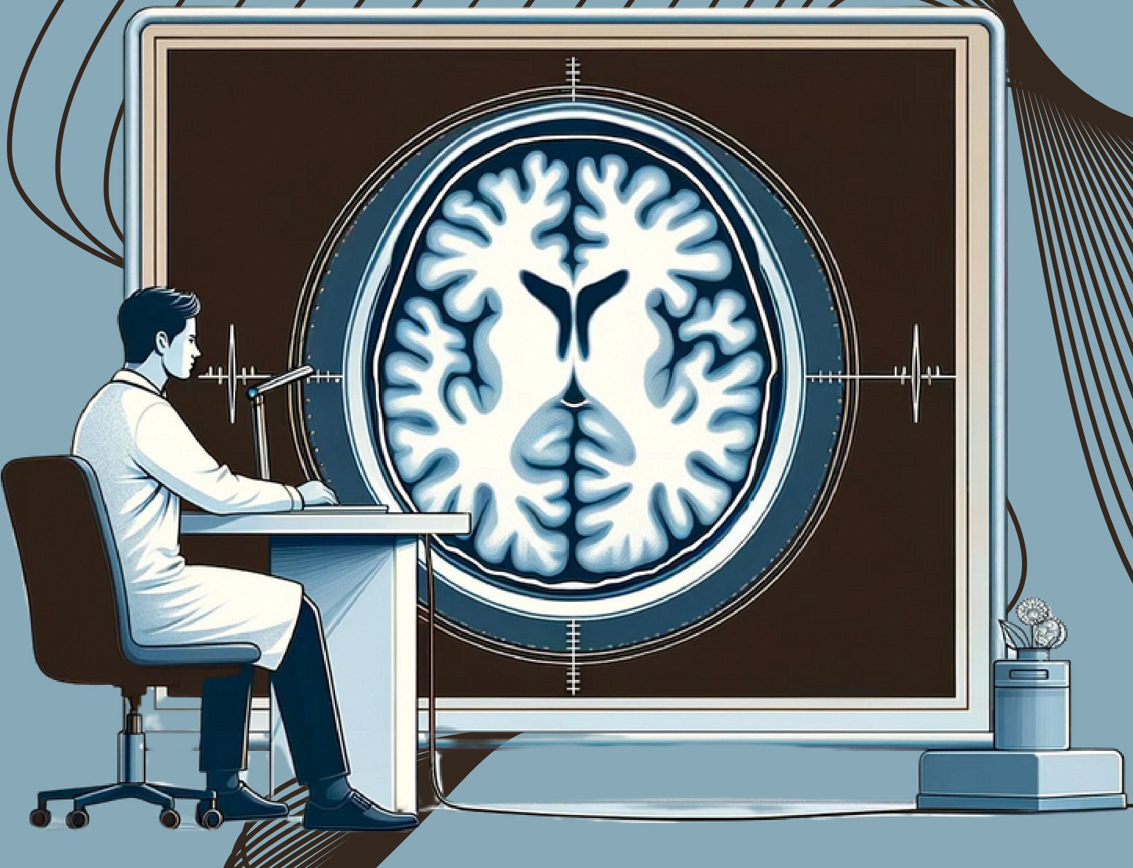
MEDICAL IMAGING SYMPOSIUM FOR PHDS AND POSTDOCS

MISP² 2024

schedule

keynote speaker

book of abstracts



17 January 2024 | Netherlands Cancer Institute | Amsterdam

SCHEDULE

8:30 - 9:00	Registration
9:00 - 9:05	Opening
9:05 - 10:15	Keynote talk
10:15 - 10:30	Coffee break
10:30 - 12:00	Session 1: CT & Pathology
12:00 - 12:45	Lunch
12:45 - 14:30	Session 2: MRI
14:30 - 15:00	Coffee break
15:00 - 16:00	Session 3.1: Other modalities
16:00 - 16:10	Break
16:10 - 17:00	Session 3.2: Other modalities
17:00 - 18:00	Closing borrel

KEYNOTE SPEAKER

Efstratios Gavves | UvA and Ellogon.AI

Dr. Efstratios Gavves is an Associate Professor at the University of Amsterdam, Lecturer of Deep Learning at the MSc in AI, Scholar ELLIS European Network of Excellent in AI, and co-founder of Ellogon.AI. His core scientific hypothesis, and research focus, is the computational understanding of time, dynamics, and causality, which is believed to be fundamental for AI that generalizes and extrapolates to open, embodied, interactive "visual data worlds". On this, he pioneered dynamical video representations in visual spatiotemporal data in 2014, while in 2016, he introduced the paradigm-shifting Siamese algorithms in visual object tracking, challenging 30 years of practice, now adopted by virtually all state-of-the-art. Since 2021, he has spearheaded research on Causal Representation Learning for algorithms that learn autonomously by integrating cause-and-effect in their core. To research the computational learning of time, he was awarded the ERC Career Starting Grant 2020, the NWO VIDI, and became director of the QUVA Deep Vision Lab with Qualcomm, and the POP-AART Lab with the Netherlands Cancer Institute and Elekta. Efstratios currently supervises a team of 15 doctoral students, on theory and applications of Deep Learning and Dynamics, and Computer Vision. He has also been advancing AI algorithms for identifying cancer patients that benefit from immunotherapy.

keynote talk:
Causal and Dynamical Computer Vision towards Cyberphysical AI

ABSTRACTS

CT & PATHOLOGY

1. A New Challenge for Universal Lesion Segmentation in CT: ULS23

Max de Grauw, Radboudumc

To address recent increases in radiologists' workload, particularly in the context of a projected surge in global cancer cases, we have launched a medical challenge for the task of automated universal lesion segmentation in CT. Current practices for quantifying disease progression in CT scans, such as the Response Evaluation Criteria In Solid Tumors (RECIST) guidelines, rely on manual measurement of lesions. This is often a time-consuming process, with significant inter-observer variability, prompting the exploration of automated segmentation models to alleviate these problems. ULS23 hopes to leverage the strides made in AI-based segmentation models for specific tumor types but emphasizes the need for models adept at handling the diversity of lesions in the thorax-abdomen area. We have released novel fully annotated 3D training data and compiled existing datasets to facilitate model training. We also curated a diverse multi-centre test set for this task to be used in a type II challenge format. With the release of our baseline ULS models, and those from the challenge participants, our goal is to encourage collaboration and promote progress in the field.

---

2. CT-Based Ventilation Imaging Using Anatomically Constrained Image Registration

Paris Tzitzimpasis, UMC Utrecht

CT-ventilation imaging (CTVI) is an image processing modality that utilizes phase-resolved CT images and employs a series of post-processing steps in order to provide surrogate lung function maps. Although its potential in avoiding the irradiation of highly ventilated pulmonary regions during radiotherapy has been demonstrated, it remains a challenging task due to the need to interpret observed anatomical changes as functional information of ventilation. Here we introduce a novel deformable image registration (DIR) - based solution for accurate CTVI generation, integrating the physical characteristics of lung deformations in its design. The proposed method consists of two steps: 1) A DIR step via a novel solution specifically tailored for the lungs, which models the estimated deformations as a set of local contractions/expansions and rotations and 2) Generation of a ventilation map using a combination of motion and tissue density information. We tested our method using a publicly available dataset comprising inhale/exhale CT scans paired with ground truth ventilation images (RefVI's) obtained with reference nuclear medicine methods. The estimated CTVIs were tested against the RefVIs using the Spearman correlation coefficient and Dice overlap of low- and high-function lung defined as the 25% least and most functional regions respectively. The results indicate that our framework can consistently generate ventilation maps with high fidelity compared to reference approaches from nuclear medicine and achieve an overall performance that considerably outperforms other state-of-the-art image based methods.

3. Automated Identification of Normal/Abnormal Findings in CT Radiology Reports

Philipp Lena, Radboudumc

A significant portion of CT scans in medical imaging do not reveal clinically actionable findings. The overall goal of this project is to develop a method for automatically identifying and labelling "normal/unremarkable" organs or regions in CT scans, thereby enhancing the efficiency of radiologists and medical professionals. Central to this project is the creation of a dataset containing CT scans, both with and without findings in specific organs or regions. We utilize the Vicuna 13B Large Language Model for the initial labelling of this dataset. The model processes the corresponding radiology reports, classifying organs as “normal” or “abnormal.” In a test set focused on kidneys (n=80), Vicuna 13B demonstrated robust performance, achieving an F1-score of 0.93 (with recall at 0.98 and precision at 0.88) for identifying “normal” cases, and an F1-score of 0.87 (recall 0.79, precision 0.96) for “abnormal” cases. These metrics illustrate the model's capability to accurately classify findings from complex and unstructured medical texts. The labelled dataset (with labels for multiple organs) will serve as the basis to pretrain an image encoder, which can be further refined for other classification tasks.

---

4. Prostate DeepReg: Unsupervised Deep Learning for Image Registration of Prostate Camera Images to H&E Images

Ryan Schoop, The Netherlands Cancer Institute

Fluorescent lifetime imaging (FLIm) and diffuse reflectance spectroscopy (DRS) are two optical techniques that are currently being investigated for prostate tissue characterization. For this research, measurements from these optical techniques of both tumor and healthy tissue are required. However, obtaining a dataset in which the optical measurements of tissue specimens are labelled by tissue type is challenging. To label the optical measurements of a prostate specimen, an annotated hematoxylin-eosin (HE) stained image is required. Accompanying the optical measurements is a snapshot camera image of the same prostate surface of which an HE image is made. So, the task of correlating the optical measurements to tissue type information can be translated to an image registration task between the camera image and the HE image. Tissue deformation that occurs during processing from the fresh specimen to the HE image complicates this registration. In this work, an unsupervised neural network approach is investigated to achieve this multimodal image registration of camera images and HE images. Camera images acquired from prostate specimens obtained directly after prostatectomy, and their associated HE images, obtained after histopathologic processing, are used for training, testing, and evaluation. Extensive pre-processing is required to prepare the images for training after which different training experiments are conducted, aimed at comparing different modes of data augmentation, different loss functions, and different amounts of regularization. Altering these training parameters affects the network performance in nuanced ways, which do not necessarily improve registration accuracy. The biggest improvement in registration accuracy is achieved by image pre-processing.

5. Multi-centric validation of an AI-based sTIL% scoring model for breast cancer H&E whole-slide images proves to be prognostic

Yoni Schirris, The Netherlands Cancer Institute

Tumor-infiltrating lymphocytes (TILs) have reached evidence as a prognostic biomarker in Breast Cancer (BC), which are scored by pathologists on Haematoxylin and Eosin (H&E)-stained tumor tissue. However, TILs are not used in daily clinical practice. Although the International TIL Working Group (TILWG) provides a guideline for pathologists, most are not comfortable scoring TILs. Hence, we develop an Artificial Intelligence (AI) model to predict the stromal TILs (sTIL%) score directly from H&E whole-slide images (WSIs) in an external validation setting from patients across multiple centers. Across 6 cohorts, only patients with an H&E WSI, a sTIL% scored by a trained pathologist following the TILWG guidelines, and all relevant clinical data were included (N≈2500). We train a multiple instance learning regression model to predict the sTIL% score directly from the H&E WSIs, with varying training and validation setups across 6 retrospective cohorts. We observe a linear relationship between the predicted and the pathologist's scores (Pearson's r 0.54-0.75) across all cohorts when training on ~200 open-source images. If we split the ground truth into the clinically relevant sTIL-low (<30%) and high (>=30%) groups, our model achieves an AUROC of 0.80-0.94 across cohorts, and it is prognostic for young N0 TNBC patients. We envision that the AI-TIL scores will be used for clinical decision-making on top of well-known prognostic factors, allowing clinicians to estimate a patient's prognosis more accurately and tailor (neo)adjuvant systemic therapies.

---

6. Bridging the Clinical Integration Gap of AI-Based Algorithms in Histopathology

Khrystyna Faryna, Radboudumc

AI-based algorithms demonstrate the potential to augment pathologists' capabilities, enhancing diagnostic accuracy, and expediting patient care. However, the clinical integration of these algorithms faces significant challenges. In our research, we address several aspects that hinder the clinical integration of deep learning in histopathology. Firstly, to enhance the adaptability of AI algorithms across diverse clinical environments, a crucial aspect involves addressing domain generalization challenges. Our first project explores the integration of augmentation techniques to expose algorithms to a broader spectrum of histopathological image variations. By simulating diverse imaging conditions, algorithms can better generalize and perform consistently across a range of clinical settings, promoting their robustness and reliability. Secondly, ensuring the reliability and generalizability of public and commercial AI algorithms is paramount for their clinical integration. Our second project investigates independent validation using crowdsourced data, providing a diverse and representative dataset for algorithm evaluation. Such validations contribute to transparency, instil confidence among clinicians, and pave the way for regulatory approvals, fostering a smoother transition of AI algorithms into routine clinical practice. Thirdly, in our next project, we explore the potential of leveraging deep learning to bypass traditional cancer grading schemes towards data-driven outcome prediction approaches. By leveraging AI algorithms to analyze complex patterns and subtle nuances in histopathological images, a more nuanced understanding of disease progression and patient outcomes can be achieved.

7. Artificial intelligence-based morphometric signature to identify ductal carcinoma in situ with low risk of progression to invasive breast cancer

Marcelo Sobral-Leite, The Netherlands Cancer Institute

Ductal carcinoma in situ (DCIS) may progress to ipsilateral invasive breast cancer (iIBC), but the majority never will. Because DCIS is treated as early breast cancer, many women with harmless DCIS face overtreatment. To identify these women that may forego treatment, we hypothesized that DCIS morphometric features relate to the risk of subsequent iIBC. We developed an artificial intelligence-based DCIS morphometric analysis pipeline (AIDmap) to detect DCIS as a pathologist and measure morphological structures in hematoxylin-eosin-stained (H&E) tissue sections. These were from a case-control study of patients diagnosed with primary DCIS, treated by breast-conserving surgery without radiotherapy. We analyzed 689 WSIs of DCIS of which 226 were diagnosed with subsequent iIBC (cases) and 463 were not (controls). The distribution of 15 duct morphological measurements in each H&E was summarized in 55 morphometric variables. A ridge regression classifier with cross-validation predicted 5 years free iIBC with an area under the curve of 0.65 (95% CI 0.55–0.76). A morphometric signature based on the 30 variables most associated with outcome, identified lesions containing small-sized ducts, low number of cells and low DCIS/stroma area ratio. This signature was associated with lower iIBC risk in a multivariate regression model including grade, ER, HER2 and COX-2 expression (HR = 0.56; 95% CI 0.28–0.78). AIDmap has the potential to identify harmless DCIS that may not need treatment.

MRI

8. Medical traumatic stress and its effects on white matter in pediatric brain tumor patients

Anne Leenders, Prinses Máxima Centrum

Pediatric brain tumor patients (PBTP) face the risk of neurocognitive impairments and related white matter changes. While post-traumatic stress symptoms (PTSS) have been known to affect cognition and white matter in other populations, their impact on PBTP remains unclear. This study investigates the influence of PTSS on neurocognitive functioning and limbic white matter in PBTP. Sixty-six PBTP (6-16 years) underwent neuropsychological assessments and brain MRI one year post-diagnosis. Parents completed PTSS proxy questionnaires (CRIES-13; 1-3 months and one year post-diagnosis). Z-scores and the percentage of impairments (>1SD) for attention, processing speed, executive functioning, and memory were compared to norms using (t-tests, chi-square tests). Multi-shell diffusion MRI data were analyzed for white matter tractography (FA/AD). Linear regression models, with age at diagnosis, treatment intensity, and tumor location as covariates, explored the effects of PTSS on neurocognition and white matter. Neurocognition and limbic white matter associations were explored with correlations. Attention (M=-.49, 33% impaired; $P<.05$) and processing speed (M=-.57, 34% impaired; $P<.05$) were significantly lower than healthy peers. PTSS was linked to poorer processing speed (β=-0.64, $P<.01$). Limbic white matter metrics were associated with treatment intensity, age at diagnosis, and tumor location, but not with PTSS. No significant associations were found between neurocognition and white matter metrics. Elevated PTSS correlated with poorer processing speed in PBTP, underscoring the importance of monitoring and timely referrals to enhance psychological well-being and neurocognitive functioning. Subsequent research should concentrate on longitudinal follow-up and explore the impact of PTSS interventions on neurocognitive performance.

---

9. Learning Dynamic Undersampling and Reconstruction for Accelerated Cardiac MRI

George Yiasemis, The Netherlands Cancer Institute

Undersampling the k-space is pivotal in accelerating MRI acquisitions. Existing methods often use 1D equidistant or random trajectories, training deep learning models for reconstruction. In our study, we introduce a novel approach: jointly learning dynamic undersampling schemes and reconstruction for cardiac cine MRI data. Our method learns to harness shared knowledge between timeframes, significantly improving reconstruction metrics compared to networks trained on random or equidistant subsampled data.

10. Time-Resolved Biomechanics using Spectro-Dynamic MRI: Proof of Principle in the Muscles of the Thigh

Max van Riel, UMC Utrecht

Measuring biomechanical quantities in a dynamic setting is important in studying the functioning of the musculoskeletal system, and can help in diagnosing pathologies. Spectro-Dynamic MRI is a recently developed method that allows for time-resolved identification of biomechanical tissue properties by working directly from k-space data. We show a proof-of-principle application of time-resolved Spectro-Dynamic MRI in the human thigh with a straightforward and inexpensive acquisition and reconstruction setup. An inflatable pressure cuff was placed around the thigh of a volunteer. The pressure cuff inflated during the scan, deforming the muscles. The time-resolved non-rigid displacement fields and k-space data were reconstructed by solving an optimization problem. The first principal strain value and direction were calculated using the eigendecomposition of the strain tensor using the plain strain assumption. The estimated time-resolved images and motion field showed a clear deformation in the left leg caused by the pressure of the inflatable cuff, while the right leg remained stationary. An abrupt change in the direction of the first principal strain can be seen at the border between different muscle groups. We expect this change to be related to the underlying anatomy of these muscles. This suggests that time-resolved measurements of biomechanical tissue properties during dynamic loads using Spectro-Dynamic MRI is possible with a straightforward experimental setup. Future extensions of this work include a 3D implementation, the use of muscle segmentations to deal with sliding motion, and the addition of a constitutive relation describing the biomechanical behavior of the muscle tissue.

---

11. Quantification and optimization of imaging quality for small bowel tracking during MR-guided radiotherapy

Saskia Damen, UMC Utrecht

MR-guided radiotherapy in the pelvis aims to deliver the prescribed radiation dose in as few treatment sessions as possible. As a highly radiosensitive organ-at-risk, this makes monitoring the small bowel during therapy particularly important. Due to the complex anatomical environment, it is however challenging to visualize and track the bowel with MRI. In this work we investigate two metrics, acutance and inverse consistency (IC), to quantify the effect of MRI protocol changes. Three variations of a balanced SSFP sequence (α=35°(FA35), α=55°(FA55) and α=65° (FA65)) on a 1.5T MR-Linac system (Unity, Elekta AB) were compared. The three sequences were acquired for ~30 seconds on 11 patients, over up to 10 fractions, resulting in 57 datasets.
Motion tracking was performed using an optical flow algorithm, with acutance (image quality) and IC (tracking quality) being evaluated on the bowel bag. A paired Mann-Whitney test showed a significant difference in acutance between FA35 and the other two protocols, in which the latter performed better. Between FA55 and FA65 no significant difference was observed. For IC significant differences between all three protocols were observed, with FA65 performing best and FA35 worst. The acutance and IC show the differences in quality between the protocols in an objective matter, with respect to motion tracking. It was also noted that the protocol with the higher acutance results in a lower inverse consistency. Based on these metrics the protocol with the flip angle of α=65° performed best in both image quality and registration results.

12. Conditional neural fields with shift modulation for multi-sequence MRI translation

Yunjie Chen, Leiden University Medical Center

Multi-sequence magnetic resonance imaging (MRI) has found wide applications in both modern clinical studies and deep learning research. However, in clinical practice, it frequently occurs that one or more of the MRI sequences are missing due to different image acquisition protocols or contrast agent contraindications of patients, limiting the utilization of deep learning models trained on multi-sequence data. One promising approach is to leverage generative models to synthesize the missing sequences, which can serve as a surrogate acquisition. State-of-the-art methods tackling this problem are based on convolutional neural networks (CNN) which usually suffer from spectral biases, resulting in poor reconstruction of high-frequency fine details. To tackle this problem, we propose Conditional Neural fields with Shift modulation (CoNeS), a model that takes voxel coordinates as input and learns a representation of the target images for multi-sequence MRI translation. The proposed model uses a multi-layer perceptron (MLP) as the decoder for pixel-to-pixel mapping. Hence, each target image is represented as a neural field that is conditioned on the source image via shift modulation with a learned latent code. Experiments showed that the proposed method outperformed state-of-the-art methods for multi-sequence MRI translation both visually and quantitatively. Moreover, we conducted spectral analysis, showing that CoNeS was able to overcome the spectral bias issue common in conventional CNN models. We further tested a segmentation network using the synthesized images at inference. The results showed that CoNeS improved the segmentation performance when some MRI sequences were missing and outperformed other synthesis models.

---

13. Artificial Intelligence and Radiologists at Prostate Cancer Detection in MRI: The PI-CAI Challenge

Joeran Bosma, Radboudumc

The PI-CAI challenge aims to rigorously evaluate the diagnostic performance of artificial intelligence (AI) algorithms compared to radiologists in identifying clinically significant prostate cancer (csPCa) using MRI. In this multi-center, retrospective study, over 10,000 prostate MRI exams from 9,129 patients (2012-2021) were curated from four European tertiary care centers. All patients were men suspected of harboring prostate cancer, without a history of treatment or prior csPCa findings. The reference standard for evaluation was set by histopathology and follow-up (≥ 3 years). Radiologists and AI algorithms had to localize csPCa lesions and assess the case-level likelihood of csPCa, utilizing bpMRI exams alongside clinical variables such as age, PSA levels, and scanner model. The study protocol was established in conjunction with 16 experts across prostate radiology, urology, and AI. 62 radiologists participated and over 500 algorithms were developed by more than 1200 individuals across 50 countries. Once developed, 22 algorithms were independently tested in a fully-blinded setting. The top five AI algorithms, when trained on 9107 cases and combined, achieved a diagnostic performance characterized by an AUROC of 0.90, across 400 cases (including external data). Preliminary results from the first 14 radiologists with 2–15 years (median: 9) of experience show a mean AUROC of 0.86, on the same cases. These findings indicate a promising role for AI in enhancing csPCa diagnostics.

14. Federated Learning for MRI-Based Dementia Diagnosis

Kaouther Mouheb, Erasmus MC

Diagnosing dementia is challenging for clinicians due to its complexity and overlapping etiologies. Artificial intelligence (AI) offers a potential solution, especially with the growing availability of radiological data. However, the sensitive nature of medical imaging data often prevents its sharing, presenting a significant obstacle to training effective AI models. To address this issue, federated learning emerges as a promising paradigm for training models in a decentralized manner without sharing data. To investigate its efficiency in the task of brain MRI-based diagnosis, we conducted a simulated study using the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, focusing on diagnosing Alzheimer's Disease from T1-weighted MRI scans. Each ADNI cohort (ADNI1, ADNI2, ADNI3) is treated as a separate client. We trained a DenseNet model in three settings: centralized (simulating data sharing), isolated (simulating no collaboration), and federated learning using the vanilla FedAvg algorithm (simulating a federated collaboration). The results indicate that the federated model with a test accuracy of 89% is comparable to the centralized model (91%), while outperforming the isolated models (76%, 85%, 78% for ADNI1, ADNI2, ADNI3 respectively). While our findings are promising, a closer look reveals bias in the federated model, with 92% accuracy in ADNI3 compared to 87% in ADNI1 and ADNI2. This bias could escalate in more realistic scenarios where data heterogeneity is more pronounced, stemming from factors such as scanner vendors, image resolution, and specialities across centers. Future work involves inspecting and improving fair federated learning techniques to address data heterogeneity in real-world clinical datasets for dementia diagnosis.

---

15. Time-Resolved Cardiac Function: Myocardium Strain and First-Pass Perfusion Using MR-MOTUS

Thomas Olausson, UMC Utrecht

Cardiac strain analysis and first-pass myocardial perfusion MR imaging are pivotal indicators in diagnosing and managing coronary artery disease. These examinations involve several dynamics including cardiac motion, respiratory motion, bulk motion, and contrast inflow. To mitigate the motion artifacts patients are instructed to perform difficult breath holds and equipped with ECGs to trigger the acquisition per heartbeat. There are drawbacks to using these types of motion mitigation methods. One is the effectiveness of ECGs to extract a cardiac phase surrogate is patient-dependent and introduces a workflow hiccup for technicians. On top of this, patients may have arrhythmias or induce bulk motion which disturbs the periodicity assumption of the acquisition leading to corrupted data points. Additionally, patients may be non-compliant with performing breath holds. We showcased our MR-MOTUS framework for combined time-resolved motion and image reconstruction in time-resolved contrast-enhanced cardiac MRI in patients. This enables motion-corrected, time resolved first pass perfusion imaging of the myocardium without the need for ECG and breath holds. The explicit modelling of motion enables separation of motion and enhancement allowing spatial alignment of all contrast-enhanced dynamics and using all the data more efficiently. In addition, our framework outputs the separation of bulk, cardiac, and respiratory motion which further provides extra diagnostic value through strain quantification.

OTHER MODALITIES

16. Generalising Artery Centreline Tracking using Scale- and Rotational Symmetries

Dieuwertje Alblas, University of Twente

Centrelines are fundamental for visualization, segmentation and model building of arteries in 3D image data. Automated extraction of centrelines from image data can be achieved using iterative tracking. Here, a tracker is initialized at a seed point inside the artery lumen and traverses the centreline while iteratively assessing the local artery orientation. Recent works used a convolutional neural network (CNN) to estimate this orientation. However, CNNs are not robust against variations in artery size and tortuosity. We present a scale-invariant, rotation-equivariant method (SIRE) to estimate local artery orientations. Preserving scale- and rotational symmetries form the key contribution of our work and allow SIRE to estimate local orientation for arteries of any calibre or tortuosity. SIRE consists of a gauge equivariant mesh CNN (GEM-CNN), operating in parallel on multiple nested spherical meshes of varying sizes. Each mesh contains a projection of the image intensities within the corresponding sphere. Rotation equivariance is obtained by the equivariant properties of the GEM-CNN processing these image features on the spherical meshes. Scale-invariance is the result of weight sharing across these meshes and aggregating their results using a maximum operator. We used two datasets to show that SIRE can be broadly applied. First, a dataset containing small, tortuous coronary arteries and second, a dataset containing large, relatively straight abdominal aortic aneurysms (AAAs). We found that tracking the performance of coronaries using SIRE trained on AAAs is comparable to a SIRE trained on coronary arteries and vice-versa. This implies SIRE generalises to unseen arteries due to its symmetry preservations.

17. Pixel Tracks and Pseudo-Depth Maps From Monocular Laparoscopic Video Clips Using Implicit Neural Representations

Beerend Gerats, Meander Medisch Centrum

The reconstruction of surgical scenes from laparoscopic video has many potential applications, from education to intra-operative context awareness. Recently, the use of implicit neural representations (INRs) was proposed for the reconstruction of surgical scenes. However, depth maps from stereoscopic imaging are required, limiting the application of these methods to robotic surgery only. In this research, we propose the application of an INR-based method to monocular video clips for the reconstruction of surgical scenes and the generation of pseudo-depth maps. We use "OmniMotion", a novel method for tracking pixels through a video clip by reconstruction of a 3D virtual scene with INRs. We evaluate its applicability on monocular laparoscopic videos by assessing pixel tracking accuracy and pseudo-depth map correctness. For our experiments, we use video clips from the SCARED dataset that was specifically designed for depth estimation in laparoscopic video. We find that OmniMotion is able to provide accurate pixel tracks through short monocular laparoscopic video clips. The method performs particularly well on anatomy-related pixels, while its performance on dynamic tool-related pixels is fragile. In video clips that provide a clear view of the abdomen and involve camera movement, the method is able to generate useful pseudo-depth maps. Our results show the potential for INR-based methods on monocular laparoscopic video clips for the reconstruction of surgical scenes, pixel tracking and the generation of pseudo-depth maps. Further development of the method is necessary to make its application computationally efficient and reliable for 3D tracking of surgical tools.

---

18. Medical mesh quality optimization with implicit neural representation

Patryk Rygiel, University of Twente

Automatic medical image segmentation methods typically construct voxel masks, from which triangular meshes can be extracted using algorithms such as marching cubes. Such two-step approaches are widely used to obtain meshes from medical images but come with drawbacks. First, meshes may come with topological errors such as multiple components, holes or cavities due to the uneven contrast or occlusions in the source image. Second, due to the discrete nature of the underlying image grid, the obtained meshes often lack smooth surfaces. This limits the applicability of triangular meshes in downstream tasks such as computational fluid dynamics (CFD). In this work, we explore the use of implicit neural representations (INRs) to fix topological errors and control the smoothness of output surface meshes. INRs are optimized to extract surfaces from topologically incorrect binary segmentation masks of abdominal aortic aneurysms (AAAs). We perform a quantitative analysis of the mesh quality using the Euler characteristic, and a qualitative analysis by assessing the correctness of the reconstruction and smoothing grades of the mesh that can be controlled with network configuration. We find that hyperparameter settings for the INR are dependent on the quality and fine detail of the input segmentation mask and that the Euler characteristic provides a strong quantitative criterion for mesh quality.

19. Using 3D point cloud and graph based neural networks to improve the estimation of pulmonary function tests

Jingnan Jia, Leiden University Medical Center

Pulmonary function tests are an important clinical metric to measure the severity of interstitial lung disease for systemic sclerosis patients. Deep neural networks have been previously proposed for this task which takes as input either the CT scan or the binary mask of its pulmonary vessels. Such input formats either lack structural information (i.e. CT) or are very basic (i.e. binary masks) in context of pulmonary vessels. We instead propose to use 3D point clouds and graphs that can better encode pulmonary vessel connectivity and hierarchy. In this paper, we proposed to use point cloud network and graph network to estimate PFTs from CT scans. We first extracted centerlines of pulmonary vessels from a chest CT scan using a pre-trained network. After that, we concerted vessel centerlines into a point cloud, with voxels being represented by a list of points with coordinates and radius as the corresponding features. We further converted vessel centerlines to graph, with voxels and voxel interactions being represented by nodes and edges, respectively. Based on the point cloud dataset graph dataset, we proposed a point neural network on point cloud data (PNN-Vessel) and a graph neural network on graph data for the estimation of PFTs.

---

20. Medical imaging as a means of understanding severe acute respiratory syndrome coronavirus 2 infection in non-human primates

Marieke Stammes, Biomedical Primate Research Centre

A hallmark of coronavirus disease 2019 (COVID-19) is lower respiratory tract infection and viral-induced pneumonia. Nonhuman primates have shown to mirror mild-to-moderate human disease determined with concordant immunologic, virologic, and lung histopathologic findings in combination with imaging abnormalities. Particularly in the lower respiratory tract, medical imaging of pulmonary disease contributes uniquely to the ability to understand and measure the consequences of infection. Advanced medical imaging of the lungs of SARS-CoV-2-exposed nonhuman primates has shown great promise in detecting and longitudinally evaluating disease in a non-invasive manner. By now it is known that infection with SARS-CoV-2 may impact multiple anatomical sites. This is found in both human as non-human primates for a longer period of time up to 100 days post infection. However, it is still unknown whether this is directly due to the virus infection itself or that they are secondary effects caused by the activation of certain immune cell mechanisms. ImmunoPET is a powerful tool that can provide insights into this question. In this context, we evaluated five different PET tracers directed against multiple immune-related processes but also the spike protein of the virus. We proved, with that specific tracer, that the virus resides in the nasal mucosa and SARS-CoV-2 initiated lung lesions. These results but also from the other tracers showed that ImmunoPET enables tracking of the immune response, which opens a gateway for monitoring the quality and quantity of the response during infections.

21. Advancing Surgical Phase Recognition: Harnessing Spatio-Temporal Information and Self-Supervised Learning

Yiping Li, Eindhoven University of Technology

Identifying different phases of surgical procedures is crucial for improving surgical assistance, allowing real-time monitoring to help surgeons anticipate and adjust to critical transitions. The use of deep learning techniques has demonstrated considerable potential in fulfilling this role. The state-of-the-art method emphasises the use of spatio-temporal information. The initial implementation has been done in Robotic-assisted thoracic surgery (RATS) during RATS lobectomy/segmentectomy procedures, yielding initial results. These methods will be expanded to thoracic phases in Robotic Assisted Minimal Invasive Esophagectomy (RAMIE) procedures, as part of the NWO-funded INTRA-SURGE project for the future of surgery. Additionally, we plan to develop a network that incorporates anatomical landmarks to improve performance. Our current initiatives involve annotating and processing data, as well as exploring self-supervised learning methods. We focus on leveraging contrastive and generative methods, with a strategic emphasis on integrating temporal information into self-supervised training. This is an innovative endeavor within the surgical domain. We will extensively test different self-supervised learning frameworks to improve surgical phase recognition while addressing the limitations of annotated data.

---

22. Image-based deep learning system for coronary artery tortuosity diagnosis in coronary angiography

Miriam Cobo Cano, The Netherlands Cancer Institute

Coronary artery tortuosity is usually an undetected condition in patients undergoing coronary angiography. This condition requires a longer examination by the specialist to be detected. Yet, detailed knowledge of the morphology of coronary arteries is essential for planning any interventional treatment, such as stenting. We analyzed coronary artery tortuosity in coronary angiography with artificial intelligence techniques to develop an algorithm capable of automatically detecting this condition in patients. We used convolutional neural networks to classify patients into tortuous or non-tortuous based on their coronary angiography and performed a saliency maps examination to visually verify the results. The developed model was trained both on left (Spider) and right (45°/0°) coronary angiographies following a fivefold cross-validation procedure. A total of 658 coronary angiographies were included. Experimental results demonstrated satisfactory performance of our image-based tortuosity detection system, with a mean area under the curve of 0.96 ± 0.03 over the test sets. The accuracy, sensitivity, specificity, positive predictive values, and negative predictive values of the deep learning model were (87 ± 6)%, (87 ± 10)%, (88 ± 10)%, (89 ± 8)%, and (88 ± 9)%, respectively. The proposed system has comparable sensitivity and specificity in coronary artery tortuosity detection with independent expert radiological visual examination for a conservative threshold of 0.5. By adapting the threshold of our algorithm, it can serve as a first screening to predict a patient's likelihood of being diagnosed with coronary artery tortuosity, providing additional assistance to specialist cardiologists in their work. These findings have promising applications in the field of cardiology and medical imaging.

23. Open-Ended Medical Visual Question Answering Through Prefix Tuning of Language Models

Tom van Sonsbeek, University of Amsterdam

Medical Visual Question Answering (VQA) is an important challenge, as it would lead to faster and more accurate diagnoses and treatment decisions. Most existing methods approach it as a multi-class classification problem, which restricts the outcome to a predefined closed set of curated answers. We focus on open-ended VQA and motivated by the recent advances in language models consider it as a generative task. Leveraging pre-trained language models, we introduce a novel method particularly suited for small, domain-specific, medical datasets. To properly communicate the medical images to the language model, we develop a network that maps the extracted visual features to a set of learnable tokens. Then, alongside the question, these learnable tokens directly prompt the language model. We explore recent parameter-efficient fine-tuning strategies for language models, which allow for resource- and data-efficient fine-tuning. We evaluate our approach on the prime medical VQA benchmarks, namely, Slake, OVQA and PathVQA. The results demonstrate that our approach outperforms existing methods across various training settings while also being computationally efficient.

---

24. Impact of Interrater Segmentation Variability on Radiomics Feature Distributions

Daan Boeke, The Netherlands Cancer Institute

Radiomics research has gained significant attention in recent decades, providing treatment prediction models with promising results on internal validation. However, the robustness and reproducibility of these models are limited for multi-center data. A major issue is the interrater variability characterizing the segmentation process, which introduces bias into the radiomics model. Past studies are often limited by obtaining only a few varying segmentations from radiologists. An algorithm was developed to generate realistically altered segmentations from a ground-truth segmentation and its associated radiological image. Here, we utilized a probability map to select likely alternative segmentations randomly. N=1000 segmentations were generated for both a lung and liver tumor slice extracted from the Medical Segmentation Decathlon dataset. From each segmentation, all n=102 PyRadiomics library features were extracted for analysis. The alternative segmentation results showed high median Dice Similarity Coefficients for the lung (0.950, CI: 0.948-0.950) and liver tumor (0.888, CI: 0.885-0.891), indicating minor differences in segmentation. Although some features remained stable with median Quartile Coefficients of Dispersion of 2.0% (CI: 1.2%-2.9%) and 2.6% (CI: 2.3%-2.7%), several radiomic features exhibited instability: 19 liver features exceeded 5% up to even 26.3% (CI: 26.2%-26.3%). Furthermore, 32.4% (CI: 22.5%-41.2%) of lung features displayed strong evidence of multiple distinct modes in their distribution (Hartigan, $p<0.05$). The simulated feature distributions revealed notable feature variability due to minor segmentation modifications. This highlights the challenges in validating radiomics models and underscores concerns regarding their practical utility in clinical settings. The developed algorithm could be used to simulate interrater variability in future segmentations.

MISP^2 | 17 January 2024 | Netherlands Cancer Institute | Amsterdam