AIMH Lab for Healthcare and Wellbeing

Marco Di Benedetto, Fabio Carrara, Luca Ciampi, Fabrizio Falchi, Claudio Gennaro, Giuseppe Amato

Artificial Intelligence for Media and Humanities laboratory
Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", CNR
<name.surname>@isti.cnr.it

Abstract

In this work we report the activities of the Artificial Intelligence for Media and Humanities (AIMH) laboratory of the ISTI-CNR related to Healthcare and Wellbeing. By exploiting the advances of recent machine learning methods and the compute power of desktop and mobile platforms, we will show how artificial intelligence tools can be used to improve healthcare systems in various parts of disease treatment. In particular we will see how deep neural networks can assist doctors from diagnosis (e.g., cell counting, pupil and brain analysis) to communication to patients with Augmented Reality.

1 Introduction

With the advent of the first impressive results coming from Deep Learning methodologies and tools, the possibility to embrace AI technologies in real life scenarios started to become a practical opportunity for both research and industrial communities. That was the first point in which AI-enabled applications appeared in nowadays work and entertaining life events, like smart recommendation systems or image editing smartphone apps. At the same time, we witnessed dramatic improvements in both software and embedded *system-on-a-chip* (SoC) hardware components, allowing powerful computer graphics capabilities as well as object tracking on mobile devices. In general, advancements were so impressive that even complex decision-making frameworks as in a healthcare systems could benefit from machine-generated solutions. More in particular, the opportunity to analyze microscopic biological images, make diagnosis on several diseases based on magnetic resonance data, or macroscopic phenotypes visual inspection, are now reaching the medical community and considered to be integrated into standard wellbeing processes.

Given the above, what follows is a gentle introduction to our efforts to let AI help us in a research and industry-ready healthcare connection.

2 Research for Wellbeing

Computer-assisted healthcare and wellbeing are part of our main interest since the AIMH lab was created. Apart from *curiosity-driven* initiatives, several governative projects have been involved, including:

- **AI4EU**: A European AI On Demand Platform and Ecosystem. Project activities include the design of a European AI on-demand platform to support this ecosystem and share AI resources produced in European projects.
- **AI4Media**: A Centre of Excellence delivering next generation AI research and training at the service of Media, Society and Democracy.
- **AI-MAP**: ALS-related pathologies and Precision Medicine register.

Our practical commitments have included several topics, among which the most important are shown in the following.

2.1 Cell Counting in Microscopy Images

Exploiting well-labeled training sets has led deep learning models to astonishing results for counting biological structures in microscopy images [Balakrishnan e Thilagavathi, 2013]. However, dealing with weak annotations, i.e., when multiple human raters disagree due to non-trivial patterns, remains a rather unexplored problem. Stronger labels can be obtained by aggregating and averaging the decisions given by several raters to the same data [Arteta *et al.*, 2016], but the scale of the counting task and the limited budget for labeling prohibit this. As a result, it is crucial to make the most with small quantities of multi-rater data. To this end, we propose a two stage counting strategy in a weakly labeled data scenario [Ciampi *et al.*, 2021] [Ciampi *et al.*, 2022]. First, we detect and count the biological structures, and in the second step, we refine the predictions, increasing the correlation between the scores assigned to the samples and the agreement of the raters on the annotations. We assess our methodology on a novel dataset comprising fluorescence microscopy images of mice brains containing extracellular matrix aggregates named perineuronal nets. We demonstrate that we significantly enhance counting performance, improving confidence calibration by taking advantage of the redundant information characterizing the small sets of available weak labels.

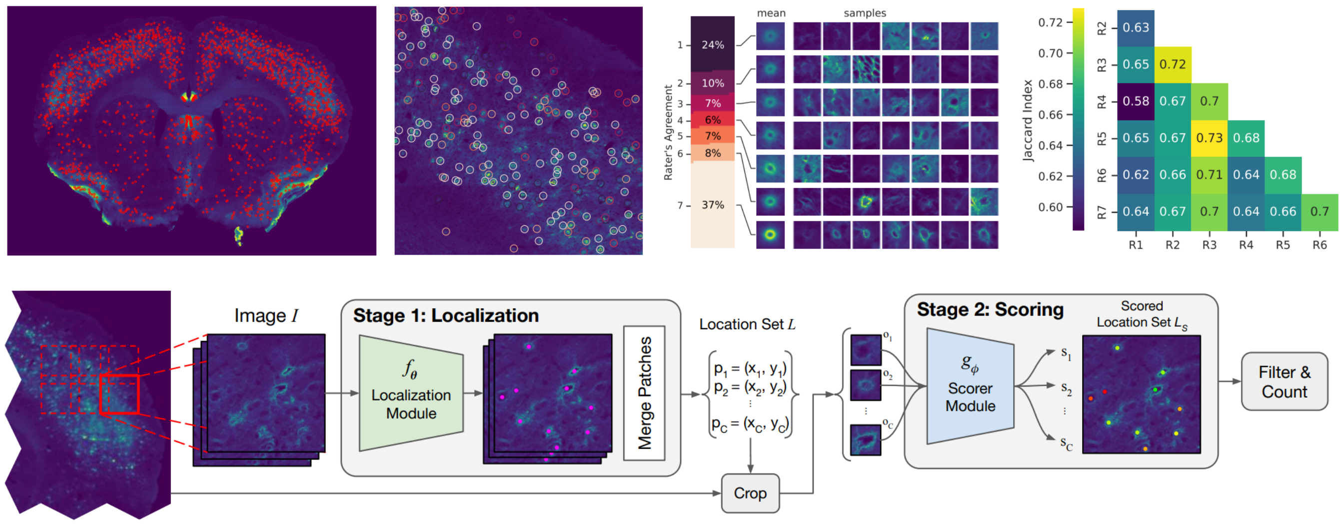

Figure 1: **Counting dataset and pipeline.** *Top row*: a sample from the multi-rater subset (PNN-MR, labeled by 7 raters), the color-encoded raters and the corresponding Jaccard Index between the PNN sets found by each rater. *Bottom row*: we model the task as a two-stage process. In the first one, we detect the objects exploiting a localization model f_θ, previously trained on a large collection of dot-annotated images that may have weak labels. In the second stage, we employ a scorer model g_θ that assigns to the objects localized in the previous step an "objectness" score, which we correlate with the pattern uncertainty quantified by the agreement's level. Image Courtesy of [Ciampi *et al.*, 2021].

2.2 Detect Frontotemporal Dementia

Behavioral variant frontotemporal dementia (bvFTD) is a neurodegenerative syndrome whose clinical diagnosis remains a challenging task especially in the early stage of the disease. Currently, the presence of frontal and anterior temporal lobe atrophies on magnetic resonance imaging (MRI) is part of the diagnostic criteria for bvFTD. However, MRI data processing is usually dependent on the acquisition device and mostly require human-assisted crafting of feature extraction [McCarthy *et al.*, 2018] [Möller *et al.*, 2016]. Following the impressive improvements of deep architectures, in [Di Benedetto *et al.*, 2021] we report on bvFTD identification using various classes of artificial neural networks (see Figure 2), and present the results we achieved on classification accuracy and obliviousness on acquisition devices using extensive hyperparameter search. In particular, we will demonstrate the stability and generalization of different deep networks based on the attention mechanism, where data intra-mixing confers models the ability to identify the disorder even on MRI data in inter-device settings, i.e., on data produced by different acquisition devices and without model fine tuning, as shown from the very encouraging performance evaluations that dramatically reach and overcome the 91.0% value on the AuROC and balanced accuracy metrics.

2.3 Diagnose by Pupil Analysis

Pupil dynamics alterations have been found in patients affected by a variety of neuropsychiatric conditions, including autism [Nyström *et al.*, 2018]. Studies in mouse models have used pupillometry for phenotypic assessment and as a proxy for arousal [Artoni *et al.*, 2019]. Both in mice and humans, pupillometry is noninvasive and allows for longitudinal experiments supporting temporal specificity; however, its measure requires dedicated setups. In [Mazziotti *et al.*, 2021], we introduce a convolutional neural network that performs online pupillometry in both mice and humans in a web app format (see Figure 3). This solution dramatically simplifies the usage of the tool for the nonspecialist and nontechnical operators. Because a modern web browser is the only software requirement, this choice is of great interest given its easy deployment and setup time reduction. The tested model performances indicate that the tool is sensitive enough to detect both locomotor-induced and stimulus-evoked pupillary changes, and its output is comparable to state-of-the-art commercial devices.

2.4 Human Body Tracking

We know that a good communication from our own doctor to us is crucial in understanding a particular illness that is affecting the patient. For most localized diseases it is important to identify the part of the body that is involved and, with the help of medical imaging, picture out the lesion itself. This is relatively simple with 2D images like X-rays (e.g., bone injury), but it became very complicated whenever 3D parts have been acquired: 2D *slices* of the volume are shown, making it difficult for the patient to understand the whole figure.

In this context, we believe that showing the whole reconstructed 3D model (e.g., bones or internal organs) can help people understand their condition. We push the experience by visualizing the model *anchored* to patient's body that is tracked in real time: by using Augmented Reality in both desktop or mobile devices, patients can then explore their body condition and directly look at the disease, even during its evolution in time. In addition, we exploit recent browser capabilities to run neural network computations to empower both doctors and patients to execute complex analysis on

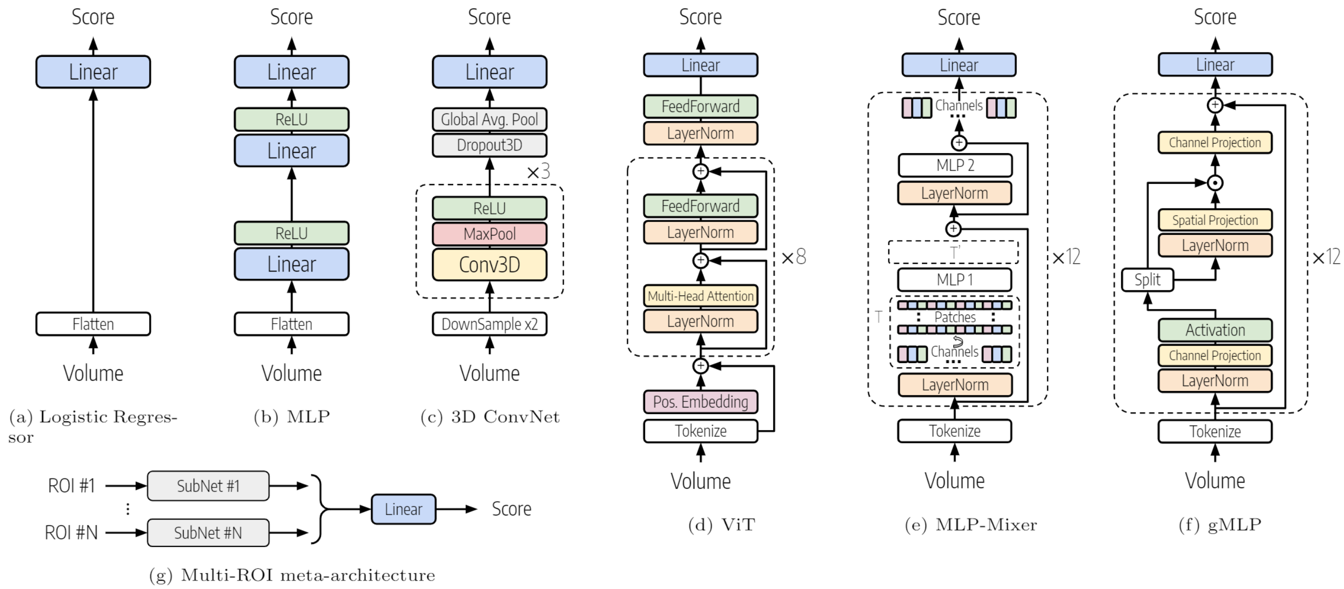

Figure 2: **Architectures of the evaluated networks for bvFTD identification.** Starting from the simple *Logistic Regressor* (a), we explored various neural models, considering the *Multi-Layer Perceptron* (b), *3D Convolution* (c), and the more recent *Visual Transformer* (d), *MLP-mixer* (e), and *gMLP* (f). In each case, the input 3D medical image is flattened or tokenized before entering the actual network. For region-based classification, each network is replicated (with an independent set of weights) after region extraction, and their output is then linearly processed for the final classification label (g). Image Courtesy of [Di Benedetto *et al.*, 2021].

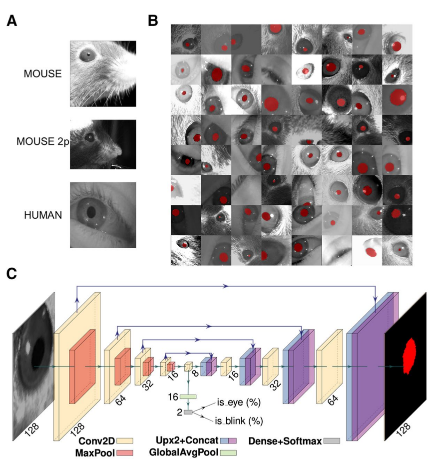

Figure 3: **Pupillometry dataset, CNN architecture, and performances.** *A*, Examples of images taken from the dataset. The first image depicts a head-fixed mouse with dark pupils, the second one is a head-fixed mouse with a bright pupil, during two-photon microscope sessions. The last image is a human eye taken during experiments wearing virtual reality goggles. *B*, The 64 examples of data augmentation fed to CNN. The images are randomly rotated, cropped, flipped (horizontally or vertically), and changed in brightness/contrast/sharpness. *C*, CNN architecture with an encoder–decoder "hourglass" shape. The encoder part comprises a sequence of convolutional layers. Starting from the last encoder output, the decoder part iteratively upsamples and fuses feature maps with corresponding encoder maps to produce the output pixel map. The pixel probability map and eye/blink probabilities are computed by applying the sigmoid activation to the network outputs in an element-wise manner. Image Courtesy of [Mazziotti *et al.*, 2021].

medical images (see Numel.AI in the following Section).

3 Applications towards Industry and Society

We concentrated our efforts in providing practical usages of our research results by providing open-source projects, as described in the following. To our hope, these frameworks will involve both research and industry communities to cooperate in a shared playground.

MEYE

Alteration of pupil dynamics is an important biomarker that can be measured noninvasively and across different species. *MEYE* (https://www.pupillometry.it) is an open-source web app that, through deep learning, can perform real-time pupil size measurements in both humans and mice, with accuracy similar to commercial-grade eye trackers. The tool requires no installation, and pupil images can be captured using infrared webcams, opening the possibility of performing pupillometry widely, cost-effectively, and in a high-throughput manner.

Numel.AI - *Communicate with Care*

We built a web application called Numel.AI (https://numel.ai) [Di Benedetto, 2021] that, at the time of writing, performs face tracking and volumetric rendering to visualize head scans. The app uses a dedicated neural network to identify brain lesions and show them to the user (see Figure 4). Given its *browser-only* compute functionalities and its REST API flexibility, the framework exhibit low-cost maintenance and can be locally installed to personal websites or cloud providers.

Figure 4: **Augmented Reality and AI for Medical Imaging.** Numel.AI is a web app that uses Augmented Reality to track human faces and overlays medical imaging over the patient's body. By using a dedicated neural network, the REST API-enabled framework allows for in-browser computation on 3D data. Image Courtesy of https://numel.ai

4 Our Technology-Transfer Perspective

We plan to extend our counting framework by providing a single model, still trained in two separate stages, to deliver the same counting performance while reducing the overall computation by sharing the parameters. We would like to improve and expand our systems further to support and evaluate novel telemedicine protocols. Our hope is to create a community that refines and consolidates pupillometric performances to produce a tool that can be applied easily in different environments. Also, we target to integrate hands tracking and add bone fracture augmented reality plugin to our web apps. As a new venture, we will try to integrate our Reinforcement Learning framework to work with diagnose plan prediction.

References

[Arteta *et al.*, 2016] C. Arteta, V. S. Lempitsky, e A. Zisserman. Counting in the wild. In *ECCV*, 2016.

[Artoni *et al.*, 2019] P. Artoni, A. Piffer, V. Vinci, J. LeBlanc, C. Nelson, T. Hensch, e M. Fagiolini. Deep learning of spontaneous arousal fluctuations detects early cholinergic defects across neurodevelopmental mouse models and patients. *Proceedings of the National Academy of Sciences*, 117:201820847, 07 2019.

[Balakrishnan e Thilagavathi, 2013] V. Balakrishnan e K. Thilagavathi. Automatic red blood cell counting using hough transform. pages 267–271, 04 2013.

[Ciampi *et al.*, 2021] L. Ciampi, F. Carrara, V. Totaro, R. Mazziotti, L. Lupori, C. Santiago, G. Amato, T. Pizzorusso, e C. Gennaro. Learning to count biological structures with raters' uncertainty. *Under Review.*, 2021.

[Ciampi *et al.*, 2022] L. Ciampi, F. Carrara, G. Amato, e C. Gennaro. Counting or localizing? evaluating cell counting and detection in microscopy images. In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022)*, 2022.

[Di Benedetto *et al.*, 2021] M. Di Benedetto, F. Carrara, B. Tafuri, S. Nigro, R. De Blasi, F. Falchi, C. Gennaro, G. Giglio, G. Logroscino, e G. Amato. Deep networks for behavioral variant frontotemporal dementia identification from multiple acquisition sources. *Computers in Biology and Medicine.*, 2021. Under Review.

[Di Benedetto, 2021] M. Di Benedetto. Numel.AI - Communicate with Care. https://numel.ai, 2021.

[Mazziotti *et al.*, 2021] R. M. Mazziotti, F. Carrara, A. Viglione, L. Lupori, L. L. Verde, A. Benedetto, G. Ricci, G. Sagona, G. Amato, e T. Pizzorusso. Meye: Web-app for translational and real-time pupillometry. *bioRxiv*, 2021.

[McCarthy *et al.*, 2018] J. McCarthy, D. L. Collins, e S. Ducharme. Morphometric mri as a diagnostic biomarker of frontotemporal dementia: A systematic review to determine clinical applicability. *NeuroImage: Clinical*, 20:685–696, 2018.

[Möller *et al.*, 2016] C. Möller, Y. A. Pijnenburg, W. M. van der Flier, A. Versteeg, B. Tijms, J. C. de Munck, A. Hafkemeijer, S. A. Rombouts, J. van der Grond, J. van Swieten, et al. Alzheimer disease and behavioral variant frontotemporal dementia: automatic classification based on cortical atrophy for single-subject diagnosis. *Radiology*, 279(3):838–848, 2016.

[Nyström *et al.*, 2018] P. Nyström, T. Gliga, E. Jobs, G. Gredebäck, T. Charman, M. Johnson, S. Bölte, e T. Falck-Ytter. Enhanced pupillary light reflex in infancy is associated with autism diagnosis in toddlerhood. *Nature Communications*, 9, 05 2018.