Visual content analysis with Machine-Learning and Deep-Learning methods for biomedical and neuroscientific applications

PhD Candidate: Mattia Savardi

Email: m.savardi001@unibs.it

XXXII Cycle

Tutor: ing. Alberto Signoroni, ing. Sergio Benini

Background

The recent advancement in neural networks, also known as deep learning, has brought an epiphany in computer vision field, especially in fine grained object classification and detection. This breakthrough promises improvements in accuracy performances in many areas that must be investigated, such as **Visual content analysis with Machine-Learning and Deep-Learning methods for biomedical and neuroscientific applications.**

Objectives

Automatic systems for presumptive analysis in clinical microbiology; Relating video content to fMRI brain.

Methodologies

AUTOMATIC SYSTEMS FOR PRESUMPTIVE ANALYSIS IN CLINICAL MICROBIOLOGY: As a reference FLA system to create our experimental database we consider the WASPlab (Copan, Italy) which is capable to record high definition plate images at different incubation times and different lighting conditions, including combinations of front-light and back-light modes to produce naturally appearing images for specialists reading such images on diagnostic workstations. Traditional culturing process implies that for each plate a skilled microbiologist needs to handle and look the plate in order to give a result. Instead, FLA automatize this process, with more safety and contaminations guarantee, and finally acquire high-quality images of the plate with different setup (illumination, time of incubation, etc.). The clinical microbiologist can now look the plates from a monitor that opens new possibilities in terms of visualization, pre-processing and Computer aided AI systems as support of his decision. In this context the haemolysis, that is an effect caused by certain bacteria spaces that leads to the discoloration of the blood substrate surrounding the colony, is looked putting the plate backlit as similar by setting a proper light below the plate during the acquisition process. Here, thanks to the possibility to acquire high-quality Digital images, we address the challenge of a haemolysis detection tool using a single CNN, trained end-to-end from images directly, as support to clinical microbiologist. The proposed system requires no hand-crafted features; it is trained end-to-end directly from image labels and raw pixels, with a single network. In order to achieve this goal, we created a dataset which contains hemolytic and non-hemolytic colonies that actually representing the possible variety, including many hemolysis examples of different shape, intensity and dimension, carefully annotated with the help of a skilled clinical microbiologist. We collected a dataset of more that 1000 plates, produced by the inoculation on REMEL 5% sheep blood agar media of urine samples collected during routine lab tests.

RELATING VIDEO CONTENT TO FMRI BRAIN DATA: The main hypothesis of this research is to apply deep learning methods to video content in order to extract different visual data content representations. These representations are then linked with brain data with the aim to find relationships between stimuli and responses. As first step in order to investigate the presented and other research questions, we built an automatic framework for estimating SSD (Shot Scale Distribution). We train three DL architectures by using a portion of a dataset composed by 120 movies annotated with the help of prof. Kovacs reasearch group. The dataset includes the almost complete filmography by six different directors whose styles are consensually considered highly unique and distinguishable in film historiography of author cinema: Michelangelo Antonioni, Ingmar Bergman, Federico Fellini, Jean-Luc Godard, Martin Scorsese, and Bela Tarr. The challenges posed by such artistic video content lie in the variety of experimental aesthetic situations, the richness of the scene composition, and the presence of unconventional or highly symbolic content both in colour and B&W. During test we select movie, whenever possible, by uniformly sampling director's productions, and trying to balance B&W and colour movies. For each Convolutional Neural Network (CNN) we adopt four

Expected Results and Impact

AUTOMATIC SYSTEMS FOR PRESUMPTIVE ANALYSIS IN CLINICAL MICROBIOLOGY: The relevance of the proposed method about the identification of Haemolysis in blood agar plates and of the obtained results is twofold. On one hand this work is one among the first ones in the context of Digital Microbiology Imaging (DMI) and related to the emerging field of Full Laboratory Automation systems. On the other hand, it can be considered with a view to the clinical applications it may have most impact. The complexity of the task and the needed sensitivity, require as crucial a timely and accurate support for the clinical diagnosis. With the proposed system, despite some errors, the accuracy results are already highly satisfactory other than being produced at a speed compatible to the analysis and FLA workflow. We reach more than 98% in recall with only 2 false negative as can be seen in figure. The ultimate goal here is to accelerate and simplify clinical microbiological analysis and to reduce the turnaround time for the identification of the best antibiotic therapy, allowing to leave no stone unturned in the exploitation of all the information present on the culture plate. Promising results have been obtained but sill there are rooms for improvements.

RELATING VIDEO CONTENT TO FMRI BRAIN DATA: Up to now the work done was mainly focused to the creation of a freamework capable to extract information (as features) from movies . In the second phase (about which we do not have yet any result) we will try to map these features on fMRI brain data and viceversa. The three presented neural networks for the SSD classification have almost same performance when only the last layer, or all the fully-connected ones are trained, while a gain of +1.5 point is obtained on average by finetuning the whole net. Two different networks (per each architecture) are trained to deal with colours and B&W movies, respectively. The overall best model is VGG-16 trained with stochastic gradient descent with momentum 0.9 and weight decay 5e-7. Learning is carried on for 30 epochs with a base learning rate of 1e-5, divided by 10 in case of validation error plateaus. The best set-up leads to a precision of 80% with a recall of 77%. As a interesting test, we tried to use this feature in order to predict the author of a movie. For what concern the first two questions we found that, also with these simple formal features, is possible reach a good accuracy. Moreover, different authors generally use different shot features. The employed dataset includes the almost complete filmography by six different directors: Antonioni, Bergman, Fellini, Godard, Scorsese, and Bé la Tarr, for a total number of 144 movies from 1940 (black and white) up to 2010. Findings open up interesting lines of multi-disciplinary research across video analysis, cinema studies, psychology and neurosciences. In the next steps we will try to find clues of this features in fMRI in order to figure out how different multimedia content impact on human perception.