DEEPHEALTH

D1.1 Use cases and applications specifications and requirements

Project ref. no.	**H2020-ICT-11-2018-2019 GA No. 825111**
Project title	Deep-Learning and HPC to Boost Biomedical Applications for Health
Duration of the project	1-01-2019 – 31-12-2021 (36 months)
WP/Task:	WP1/ T1.1, T1.2
Dissemination level:	PUBLIC (with CONFIDENTIAL version)
Document due Date:	31/05/2019 (M5)
Actual date of delivery	07/06/2019 (M6)
Leader of this deliverable	UNITO
Author (s)	Marco Grangetto, Ynse Hoornenborg, with contributions of all partners
Version	v1.0

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 82511

Document history

Version	Date	Document history/approvals
0.1	8/04/2019	Requests to use-case partners to draft requirements and interactions with platform providers using a common template
0.2	12/04/2019	Consolidation of templates for all Use Cases
0.3	16/04/2019	Table of content and draft text
0.4	17/04/2019	Updates to ToC
0.5	17/05/2019	Merged in the Application Platform requirements
0.6	21/05/2019	Merged all contributions and sent out for review by all partners
0.7	27/05/2019	Deadline for comments and final editing
0.8	29/05/2019	Consolidated document provided to all partners for confidentiality check and for final internal peer review
1.0	05/06/2019	Final quality check after peer review ended
1.0	07/06/2019	Submitted document to EC

Table of contents

Executive summary

This deliverable reports on the activities carried out in WP1 of the DeepHealth project for the first 5 months targeting the definition of the 14 use cases. These will be used to demonstrate the project outcomes with deployment on the 5 platforms contributed by the project consortium. This deliverable precisely delineates the requirements of each use case in terms of objectives, methods, dataset preparation and model architecture for the Deep Learning tools. Particular attention is devoted to the description of privacy requirements of medical data to secure proper data retrieval, annotation, anonymization and exploitation for machine learning. Finally, the deliverable describes how use case requirements will be matched with the features of the different platforms.

1. The DeepHealth approach

1.1. Introduction

There are great expectations to unleash disruptive innovation across the health sector thanks to better exploitations of health big data [1] and Artificial Intelligence (AI)/Machine Learning (ML) [2]. The DeepHealth project aims at contributing to such epochal change exploiting one of the "biggest big data", i.e. biomedical multidimensional images and state of the art AI based on Deep Learning (DL) methods [3][4][5]. DL is obtaining astonishing results in many sectors, with healthcare making no exception. The DL approach is based on training deep and complex Neural Networks – more specifically, Convolutional Neural Networks (CNNs) which are well suited to images. The training of neural networks from large datasets of images requires significant amounts of computing resources and this is why one of the main goals of *DeepHealth* is to combine and optimize the computing power of HPC systems with Big Data software for efficiently training DL algorithms. Moreover, in order to develop industry-ready solutions, it is of utmost importance to consider the flexibility, scalability and efficiency of these solutions. Within this context, *DeepHealth* aims at providing not only the maximum raw performance, but also at designing a framework capable of automatically selecting the most efficient setups, balancing performance, energy and accuracy.

Figure 1 shows a typical scenario where image processing is needed for helping in diagnosis. Doctors and other medical personnel are the end users. Expert users, typically computer scientists or engineers, are service providers for end users. End users have the knowledge to label images and provide metadata in relation to images. Expert users are in charge of processing the images provided by end users, mainly for training predictive models. These models are operated by end users via software platforms in order to get clues that can help them in the diagnosis. The training environment in Figure 1 represents where expert users work with large datasets of images for training predictive models. The production environment in the same figure represents where end users use a software platform or application to get knowledge from an image or a set of images. Trained predictive models are integrated within software platforms.

Currently, expert users must perform complex and time-consuming tasks in order to combine different toolkits to provide complete solutions to end users. Despite availability of many solutions to address the distinct problems related to the processing of biomedical images, expert users face the following difficulties: (1) different partial solutions are difficult to be combined in a complete solution because they were developed independently, by different developers and using different programming languages and distinct criteria; (2) some of the available toolkits/libraries were developed for a particular operating system and only a few of them are independent of the operating system; (3) the installation of each toolkit/library requires a great deal of effort for system administrators; (4) expert users have to spend many hours for designing and running sequences of operations on images; (5) most of partial solutions are not designed to run on parallel computers featuring accelerator devices such as GPUs, FPGAs or many-cores, so it is difficult to take advantage of HPC systems or Big Data distributed computing environments.

Nowadays, it is possible to construct efficient solutions to boost the processing of biomedical images thanks to the current level of development of (i) HPC Systems, (ii) Big Data middleware for processing data distributed on a computer cluster, and (iii) Deep Learning techniques for designing predictive models.

The ideal and desired scenario is one in which the time and effort required to process large datasets of biomedical images is the minimum possible, while meeting accuracy constraints. In order to reach such an ideal scenario, two tasks regularly carried out by expert users should be improved: a) manipulation of images (pre-processing, transformations, feature extractions, quality assessment, etc.), steps b.1 and b.2 in Figure 1, and b) the training of predictive models, steps b.3 and b.4 in the same figure. The availability of a toolkit that integrates all the functionalities needed by expert users will increase their **productivity** measured with proper KPIs, i.e. time-of-pre-processing-images

(**toppi**), measured in **person-hours**, and time-of-training-models (**totm**), measured in **hours**, are reduced significantly.

Solutions already exist for an efficient distribution of Deep Learning operations [6][7]. But focusing on the innovation in Europe, one of the main goals of *Deep*Health is the development of a European Distributed Deep Learning Library (EDDLL) coupled with advanced programming models optimized for parallel execution: On one hand, the internals algorithms of EDDLL will be adapted to exploit the performance of advanced hardware accelerators (CPUs, GPUs, FPGAs, etc.), and on the other hand, the procedure for training predictive models will be efficiently distributed on Hybrid and Heterogeneous HPC + Big Data architectures. The algorithms of ECVL will also be adapted to hardware accelerators.

The *Deep*Health toolkit includes both libraries, ECVL and EDDLL, plus the front-end exposed to expert users for an efficient usage all the functionalities of these libraries. This toolkit will be free and open-source software. Proprietary software platforms will be adapted to use EDDLL and ECVL.

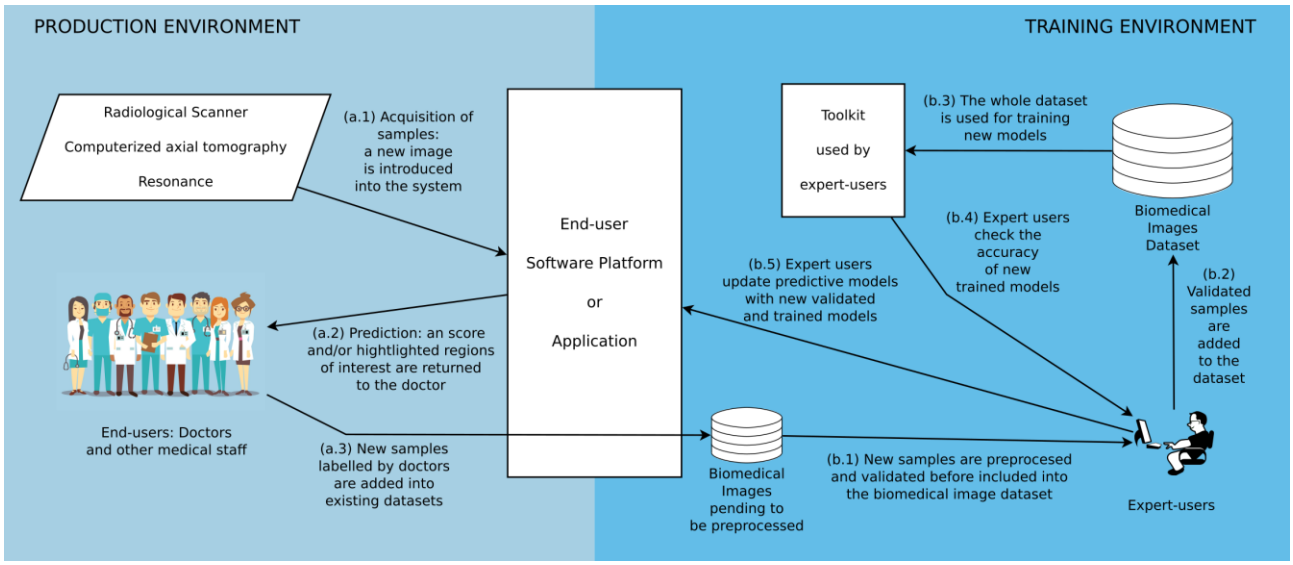

Figure 1. A typical scenario where tasks related to biomedical image processing are carried out. Steps a.1, a.2 and a.3 are done in production environments on a daily basis. Steps b.1 to b.5 in training environments are carried out when necessary.

In order to evaluate the new capabilities offered by the EDDLL and the ECVL, DeepHealth will consider 14 case studies that cover particularly the main diseases and pathologies in the fields of (1) neurological diseases, (2) tumor detection and early cancer prediction, and (3) digital analysis of heart and brain pathologies and automated image annotation, that will be deployed on 7 proprietary platforms available within the project consortium. Each use case is based on one of the following real-world medical tasks and their corresponding datasets:

- Migraine & Seizures prediction (UC1)
- UNITOPath dataset and use-case (UC2)
- UNITOBrain dataset and use-case (UC3)
- CITTA DELLA SALUTE Chest dataset and use case (UC4)
- Deep Image Annotation (UC5)
- Promort (UC6)
- Major Depression (UC7)
- Dementia (UC8)
- Study of structural changes in lumbar spine pathology (UC9)
- Predictive and Populational Model for Alzheimer's Disease (AD) using Structural Neuroimaging (UC10)
- Image Analysis and prediction for Urology (UC11)

- Skin cancer melanoma detection (UC12)
- Epileptic seizure detection (UC13)
- Prediction of multiple sclerosis patient outcome (UC14)

1.2. Use case structure

All use cases in the DeepHealth project are supported by multidisciplinary teams. In order to manage the large number of use cases and platforms involved in the workplan a common structure of the roles and activities required in each use case has been agreed and defined.

The following table provides a description of the 3 main roles identified, namely End user, Machine Learning expert and Platform provider.

Role	Description of activities
End user	Medical staff, responsible of defining use case goals, select dataset, provide labelled data (usually with the help of hospital ICT staff)
Machine Learning expert	Staff in charge of preparing a reference implementation for the use case (if not yet available): select proper ML model, CNN structure, provide a training reference software, measure inference performance with proper metrics
Platform provider	Staff in charge of deploying/porting the reference software on the platform

To manage and easy collaboration within the consortium this structure has been used to allocate responsibilities for each use case. The activities of the first 5 months of the project in WP1 allowed to delineate the requirements of each use case and plan the integration with the corresponding hosting platforms as detailed in the following Sect. 2 and Sect. **¡Error! No se encuentra el origen de la referencia.**.

2. Use case requirements

In this section, every use case is described in detail with particular attention to the definition of the goals to be achieved and the requirements to be taken into account for the development of the DeepHealth toolkit and final deployment on every hosting platform.

2.1. UC1 - Migraine & Seizures prediction

In this use case, automatic prediction of migraine neurological disorder will be considered. Personalized migraine predictions have been already achieved by the powerful mechanisms of MigraineNet platform [87] (PF2), combining user's information, Artificial Intelligence (AI) algorithms and Deep Learning (DL) techniques. The main target of this use case is the reduction of the required model training and evaluation times through the HPC technologies offered by the project. Additionally, the quickly processing data thanks to simultaneously developing/training and evaluating prediction models for more migraineurs is expected to enhance the scalability of the MigraineNet system. Last but not least, expansions of the platform's implementation in other health issues such as epileptic seizures are going to robust platform's potential.

2.1.1. Use case description and objectives

2.1.1.1. Migraine Predictions

Migraine is a neurological disorder expressed through intense headache attacks, affecting almost 15% of the world population, and it is considered among the 8 most disabling diseases based on World Health Organization (WHO) [1], but still is not adequately treated. This is highly related to the fact that there are 7 categories of migraine triggers that may apply in different combinations in the same person. Moreover, they are usually underestimated, and thus 50% of people who suffer from

migraines prefer self-treatment (e.g., over-the-counter medication) than visit a doctor [9]. However, they do not have the necessary experience or tools. Moreover, the diversity of the triggering factors and their combinations make it difficult to predict the next migraine attack and apply personalized treatment. Predicting the next migraine incident is a feature not included in any of the applications we are aware of. However, this is an important feature since it a) may help migraineurs schedule their activities based on their foreseen condition (limiting the need for re-scheduling during or after the migraine) or b) allow them to avoid the next incident (or at least limit the intensity of the next incident) by avoiding the triggering factors of their migraines and receiving their medication on time, i.e., at the beginning of the migraine incidents. According to the doctors, the sooner you receive medication during a migraine, the better its effectiveness. In this direction, solutions based on algorithms and DL techniques can constitute a powerful mechanism for addressing the diversity of the factors that can cause a migraine and allow for the prediction/forecasting of the next migraine incident of an individual, based on previous incidents, overall medical history, etc.

2.1.1.2. Epilepsy

Epilepsy is the most common serious neurological condition and affects more that 50 million people worldwide [10], while there are also recordings of 65 million patients around the world [11]. Despite the advances in epilepsy treatment, only 70% of patients are adequately treated from anti-epileptic drugs [12]. The unprovoked epileptic seizures affect negatively the patients' health status and their quality of life. For this reason, there is always the need for seizures logging systems improvements which will continuously record the biomedical signals of the patients' body.

The most prominent means to investigate epilepsy is the electroencephalogram which captures the changes of the brain electrical activity and can declare a possible seizure occurrence [13]. This diagnostic modality requires scalp electrodes attachment which is an uncomfortable method for continuous health monitoring. Except from the abnormalities in the EEG signal recordings, clinical symptoms, such as increases in oxygen availability, cerebral blood flow, blood-oxygen-level-dependent signals and changes in heart rate, are also presented before the seizure onset [14]. Non-EEG methods have been developed using combinations of a variety of sensing methods and wearable technologies. The aforementioned technologies are non-invasive, safe and due to their user-friendliness nature are appropriate health logging systems at the patients' home environment [15].

The main use case objectives are reported in the following table.

Objectives of the use case (migraine prediction)	
Type of problem	**Objective**
Prediction performance	Enhancement of the migraine prediction performance while limiting the false-positives.
Training time	The reduction of the required model training and evaluation times in order to enhance the platforms benefits.

The major use case objectives for the epilepsy-related extension are reported in the following table.

Objectives of the use case (seizure prediction)	
Type of problem	**Objective**
Prediction for seizures	The prediction of an upcoming seizure attack thanks to the improvements of the capabilities of the platform.

Classification for seizures	The features classification using bio-signals from wearable devices and insights from digital questionnaires leading to the direct recognition/warning of a seizure attack.
Training time	Evaluate the performance of the mechanisms in terms of training time without and with the DeepHealth kit. The aim is to exploit the DeepHealth toolkit in such a way so as to reduce the required model training and evaluation times in order to enhance the platforms benefits.

The use case outcome can be measured according to many metrics available in the state of the art; in the following tables, we mention the KPIs most suited to the use case for both migraine and seizure aspects.

Known KPIs (seizure prediction)	
KPI 2.1	Enhancement of the reliability of seizure identification.
KPI 2.2	Seizure prediction accuracy (in percentage).
KPI 2.3	Reduction of the training and the evaluation times of the seizure predictions.

Known KPIs (migraine prediction)	
KPI 1.1	Reduction of training and evaluation times of the migraine predictions.
KPI 1.2	Percent of predicted migraine incidents comparing to actual incidents, i.e., number migraine attacks that were predicted before occurring.
KPI 1.3	Percent of false-positive predictions, i.e., percentage of alerting a patient that a migraine incident may happen but the patient never experiencing it.

2.1.2. Partner roles and contributions

This use case is under the responsibility of WINGS that will provide all expertise according to the following table. The use case will be deployed on PF2. Given that the MigraineNet platform has been developed over a long period of time and under the supervision/directions of medical experts/doctors, the medical knowledge/background regarding the involved data, platform's use and development is already available to the WINGS members. In the event of additional medical knowledge being needed with respect to the modifications of the platform's medical aspects and structure, WINGS may engage doctors specialized in migraine and/or seizure issues so as to acquire the respective knowledge and address any new requirements.

Roles and actions		
User	**Partner**	**Action**
End user (data analytics experts)	WINGS	Pre-processing and analysis of dataset.

Machine learning experts	WINGS	Design/select and development of the machine learning techniques used for training and testing purposes.
Platform provider	WINGS	Deploy experiment on PF2.

2.1.3. Technical requirements

2.1.3.1. Dataset description

Migraine Prediction

There is already a dataset available from the first pilot round which consisting of:

a) Information related to the patients' profile (e.g. individual's sex, age, habits and other medical/environmental/working conditions related to) and contact information.

b) questionnaires with patients' feedback regarding migraine incidents.

c) the daily forms being questionnaires related to the patients' everyday activities and habits that may trigger a migraine incident.

Moreover, WINGS has already contacted with neurologists in order to start anew pilot rounds engaging new and past pilot users for testing and validation purposes. The new dataset will consist of all the above modalities as well as data from wearable devices.

Type and format	CSV/JSON files format
Annotation metadata	No manual annotation is needed (neither for the already collected data nor for the new dataset) – the patient apart from the daily forms, includes information if he/she does experience a migraine and thus, the "annotation" is automatically provided.
Number of samples (retrospective/perspective)	Digital daily questionnaires related to approximately 60 pilot users for personal information from the first pilot rounds which lasted approximately three months (retrospective data) / the new pilot round (prospective data) is expected to last for at least 3 months and similarly collect digital questionnaires on a daily basis.
Repository/sharing method	Data stored in WINGS infrastructure and available only to WINGS.
Other information	To avoid diffusing sensitive data, informed consent and any other required GDPR processes will be taken into account for both retrospective and prospective data.

Epilepsy

The epilepsy data to be investigated in the context of UC1 for predicting seizures will comprise:

- Digital questionnaires (user profile information and daily form information) in csv files;
- Bio-signals from wearables exported to csv files;
- Digital daily questionnaires for personal information e.g. individual's sex, age, habits and other medical/environmental/working conditions related to migraine.

In particular, after investigating that a similar approach can be followed for seizures as well, the data will be collected from a mobile application and wearables through a pilot involving real users. The table below summarizes the main characteristics of the dataset to be collected.

Type and format	CSV/JSON files format

Annotation metadata	No manual annotation will be needed– the patient apart from the daily forms, will also include information if he/she did experience a seizure and thus, the "annotation" will be automatically provided.
Number of samples (retrospective/perspective)	The pilot is envisioned to last for at least 3 months and collect digital questionnaires on a daily basis (prospective data).
Repository/sharing method	The data will be stored in WINGS infrastructure and will be available only to WINGS members.
Other information	To avoid diffusing sensitive data, informed consent and any other required GDPR processes will be taken into account.

2.1.3.2. Pre-processing and features extraction

Migraine Prediction

The data used for migraine prediction was collected through a mobile app and directly stored in CSV formats exploitable from the respective training mechanisms. The data to be collected in the new pilot round will accordingly be collected through a mobile app and directly stored in CSV formats exploitable from the respective training mechanisms.

Epilepsy

Depending on the outcome of the research investigation, the data to be used in this case will need to be collected from start. In this case, the mobile app will be adjusted to collect seizure-related data which will be directly stored in CSV/JSON formats exploitable from the respective training mechanisms. The dataset will accordingly comprise of:

- Data collected from patient's digital questionnaires: evaluation of the reliability of the answers in order to support the future statistical analysis.
- Data collected from wearables: noise filtering and signal quality estimation (it depends on the sensors nature; each signal will be individually pre-processed).

2.1.3.3. Model design and testing

Migraine Prediction

- Mechanisms based on the unsupervised machine learning technique called Self-Organizing Maps (SOM) are used for patient classification based on their profiles.
- Deep Learning / Long Short-Term Memory (LSTM) is used for training mechanisms that aim to identify the migraine patterns of the patient, i.e., the model that describes how migraine triggers affect his migraine attacks. The user classification may be used or not for defining the data to be used when building the migraine model of the patient

Epilepsy

Depending on the outcome of the research investigation, regarding the applicability or not of the same approach in epilepsy, the mechanisms applied to migraine prediction will be adjusted to serve seizure predictions.

2.2. UC2 - UNITOPath dataset and use-case

This use case will leverage on a set of whole-slide histological images of colon biopsies provided by the Pathology Unit of Department of Medical Science in UNITO. Methods for classification of biopsy images through DeepHealth software tools and interdisciplinary team deep learning methods will be investigated.

2.2.1. Use case description and objectives

The use-case will focus on gastrointestinal pathology, specifically colon biopsies: these particular samples represent a cornerstone activity for any surgical pathology laboratory. Differential diagnosis includes a limited number of entities, mostly neoplastic (i.e. adenomas) and more rarely inflammatory. Histopathological characterization of colorectal polyps by pathologists is the major tool for deciding the following clinical/therapeutic management of patients. The histological slides contain histological sections of the biological specimens stained with hematoxylin and eosin (H&E); H&E are chemical substances used to achieve visible colour contrast, allowing morphological diagnosis based on pattern recognition and assessment of specific features.

Deep-learning is being investigated by the scientific community as a possible tool to cut the overall laboratory workload, but also to improve the diagnostic and prognostic efficacy of histological examination. In this use case, some recent results in colorectal polyps classification [17][18] will be taken into account and, if possible, used as a benchmark and independently validated.

The medical experts participating to the use case selected six different classes for automatic classification which represent the most common diagnoses and lead to different patients' management:

- Normal tissue (NORM)
- Hyperplastic polyp (HP)
- Tubular Adenoma, high-grade dysplasia (TA.HG)
- Tubular Adenoma, low-grade dysplasia (TA.LG)
- Tubulovillous Adenoma, high-grade dysplasia (TVA.HG)
- Tubulovillous Adenoma, low-grade dysplasia (TVA.LG)

The Pathology Unit of UNITO will provide a labelled set of whole-slide biopsy images that will be exploited by machine learning experts in the team for models training and testing on DeepHealth ODH platform (PF5).

The major use case objectives are reported in the following tables, referring to objective and performance metrics that will be used.

Objectives of the use case	
Type of problem	**Objective**
Classification.	Classification of adenoma type and dysplasia grade in square patch
Classification.	Classification of adenoma type and dysplasia grade in whole slide
Training time	Reduction of the required model training and evaluation times.

Known KPIs	
KPI 1	Classification accuracy: confusion matrix, precision, recall, F-score.
KPI 2	Classification accuracy: precision, recall, F-score.
KPI 3	Time to train/time to test.
KPI 4	Time to diagnosis (including slide preparation and digitalization).

2.2.2. Partner roles and contributions

This use case is under the responsibility of UNITO that will provide all expertise according to the following table. The use case will be deployed on PF5.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	UNITO	Select images to be included in the dataset based on the diagnosis, provide labels
End user (ICT expert)	UNITO	Preprocess images (color correction, cropping etc.)
Machine learning expert	UNITO	Design classification network, provide reference software for training and testing
Platform provider	UNITO	Deploy experiment on PF5
End user (medical personnel/ expert user)	UNITO	Discuss achieved KPI

2.2.3. Technical requirements

2.2.3.1. Dataset description

Whole H&E slides, prepared for routine diagnostic assessment, will be scanned on a high-throughput Hamamatsu NanoZoomer S210 scanner, which can acquire up to 210 slides in a single batch at 20X or 40X. Before acquisition, patient's identification data on the whole-slide will be hidden, while a specific, progressive ID will be assigned to the output file of each case together with the diagnostic classification. No other clinical data will be collected and stored. The NDP.scan software, provided with the Hamamatsu NanoZoomer scanner, will be used for whole-slides digital scan: optimized H&E parameters provided by the software and automatic focus points acquisition will be used. The dataset will be constituted by high resolution microscopy images: RGB, optical microscopic image in '.ndpi' file format, multiresolution pyramid, up to 50000x50000 pixel resolution at the highest level.

Each image will be accompanied by the ancillary classification information, namely:

- classification label
- Region of Interest (ROI) (part of the image used by the pathologist to classify) in NDPA file format or ORBIT.BIO format

The medical staff is confident to select a dataset of 3000 such images. The images will be completely anonymized and accessible through UNITO internal OMERO repository [88], that allow sharing only within UNITO network with proper credentials.

2.2.3.2. Pre-processing

Digital slides will then be labelled according to the histological diagnosis and representative areas will be annotated by end users and used in the following processing.

A number of pre-processing steps are required before using images for training machine learning algorithms. In particular the following tools are necessary:

1. RGB to grayscale conversion
2. Contrast enhancement
3. Image normalization
4. Segmentation, e.g. Otsu method

5. Connected components
6. Vertices to convex polygons (ROI)
7. Rescale images
8. Extraction of fixed resolution image patches

2.2.3.3. Model design and testing

Many well-known convolutional neural network models originally designed for image classification can be applied to this use case. In particular, within DeepHealth we foresee to exploit model architecture inspired by VGG [19] and ResNet [20]. Classical Adam and Stochastic Gradient Descent learning algorithms will be used to train the models.

2.3. UC3 - UNITOBrain dataset and use-case

This Use Case is based on data provided by Interventional Neuroradiology Unit of Città della Salute University Hospital (CDSS). The objective of this use case is to synthetize a Computer Tomography (CT) Perfusion map of the brain, informative enough to be used for ordinary clinical purposes, from a reduced data set.

2.3.1. Use case description and objectives

Occlusion of a cerebral vessel causes a sudden decrease of blood perfusion of the corresponding vascular territory. Blood flow deficit is greater at the center of this area, being the most peripheral regions furnished by collateral circulation. Since brain resistance to ischemia is limited, a lesion develops in a few minutes or hours, starting from the center of the ischemic territory, and progressively enlarging up to the most peripheral territory. The irreversible brain lesion at a given time is defined “ischemic core”, the surrounding at risk brain, not dead yet, is defined “penumbra”. Distinction between core and penumbra is critical for efficiently cure the patient. CT perfusion has been demonstrated to allow to identify both regions [21].

CT perfusion (CTP) measures, with a sample time of roughly 1 Hz, the passage of a contrast medium bolus into the brain, on a pixel-by-pixel basis. Serial low-dose scans are acquired; time-density curves, corresponding to the contrast media passage in brain tissue, are calculated; parametric maps are calculated. The most relevant parameters used in clinical practice are Cerebral Blood Volume and Cerebral Blood Flow, (CBV and CBF) corresponding to the integral and the tangent of the curve, respectively [22]. In clinical settings, they approximate core and penumbra areas [23].

A relevant problem of CTP is the high dose of x-ray deployed to the patients.

Information contained in conventional CTP acquisition must be kept redundant – i.e. patient must be over-scanned - because of the relevant noise of the images, including small movements (pixels are sub millimetric), large vessels accounting for large signals and anatomical variations. This noise is reduced in conventional elaborations by smoothing, filtering, segmenting, resampling, rigid transformations, considering arterial input functions and, mainly, by oversampling.

Our aim is to demonstrate that NN may generate synthetic CBV and CBF maps with reduced and/or noisy sets of data, since they perform well in such conditions.

In order to do it, we will train a NN to generate synthetic CBV and CBF maps based on full dataset, using as metadata maps generated by ordinary algorithms. Then, we will test the performance of the NN using reduced dataset (from 90 to 10% of the original). We will evaluate the clinically relevant information delivered in CTP synthetic maps by comparison with original calculated maps.

The major use case objectives are reported in the following table.

Objectives of the use case	
The objective of UC3 is to generate the brain CT Perfusion map, used for clinical purposes, i.e. blood volume, blood flow, using a reduced or noisy dataset.	
Type of problem	**Objective**
Inference/Generation of image	Generation of Blood Flow perfusion map
Inference/Generation of image	Generation of Blood Volume perfusion map

The use case outcome can be measured according to many metrics available in the state of the art; in the following table we mention the KPI most suited to the use case.

Known KPIs	
KPI 1	Number of CTAs to estimate perfusion. Current state of the art is 20 to 40 sequential low-dose CT acquisition. DeepHealth target is 10.
KPI 2	KPI 2 Impact of patient movements (artifact) during acquisition. Movement larger than 5 mm make completely unreliable the CT perfusion maps generation. Target of DeepHealth is 10 mm.

2.3.2. Partner roles and contributions

This use case is under the responsibility of UNITO that will provide all expertise according to the following table. The use case will be deployed on PF5.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	UNITO	Select images to be included in the dataset.
End user (expert user)	UNITO	Preprocess images, provide software for perfusion map generation to build the training/testing set.
Machine learning expert	UNITO	Design classification network, provide reference software for training and testing.
Platform provider	UNITO	Deploy experiment on PF5.
End user (medical personnel / expert user)	UNITO	Discuss achieved KPI.

2.3.3. Technical requirements

2.3.3.1. Dataset description

The dataset for training will include CTP set. Each sample will include 25 to 35 sequential blocks of 16-to-64 images each (CTPI) and CBF + CBV maps generated from CTPI by state-of–the art algorithms.

Type and format	512x512, 16-bit depth gray scale CT scans, 16-to-64 slices x 20 acquisitions; DICOM format, anonymized.
Annotation metadata	CBF and CBV Perfusion maps calculated based on CTPI. Age, stroke severity, side of lesion.

Number of samples (retrospective/perspective)	200 (retrospective) 100 (prospective) of 320 to 1280 images each.
Repository/sharing method	Hospital Picture archiving and communication system (PACS) sharable only on UNITO internal network.
Other information	To avoid diffusing sensitive data, images will be anonymized before use. Informed consent will be obtained for prospective cases; reasonable attempt to obtain consent will be done for retrospective cases.

2.3.3.2. Pre-processing

The dataset must be processed before using it to train the ML models. The following pre-processing steps are applied:

- DICOM reading
- Generation of perfusion map
- Simple segmentation tool, e.g. Otsu method
- Image spatial registration

2.3.3.3. Model design and testing

The following machine learning tools will be exploited to obtain the expected objectives:

Network architecture	Autoencoder.
Training method	Adam, SGD.
Other	Loss function to be defined, likely to be MSE with respect to target perfusion.

2.4. UC4 - Chest dataset and use case

In this use case automatic analysis of chest Computer Tomography (CT) scans will be considered. Lung nodules are small focal lesions in lung parenchyma, can be solitary or multiple and in many cases are accidentally found in CT scans: their identification is time consuming in current clinical activity for the radiologist and, since these small lesions are difficult to “label”, patients often need to perform follow-up CT scans in order to assess their benignity/malignancy, resulting in increased radiation exposure and anxiety for the patient and increased work amount for doctors.

The goal of UC4 will be to train AI systems to recognize lung nodules using chest CT scans, providing radiologists an efficient tool for daily activity.

2.4.1. Use case description and objectives

Lung nodules are quite common incidental findings in CT (computed tomography) scans and can be defined as small focal lesions (ranging from 5 to 30 mm) that can be solitary or multiple.

CT scans allows, thanks to ionizing radiations, the visualization of lungs, pleura and thoracic vascular structures; from raw data acquired with the scan a data set consisting of some hundred images can be reconstructed and examined by the radiologist; image slice thickness and gap between slices can be planned in reconstruction phase, in order to allow visualization of smaller structures and obtain multiplanar image reformats that could be useful in clinical activity.

Most lung nodules found during image analysis are benign, but in some cases they can be an early manifestation of an evolutive disease, therefore requiring follow-up CT scans according to current guidelines [24]: this generates increased work volume in CT units and implies further exposition to ionizing radiations and anxiety for the patient.

Identification of these lesions is often difficult and time consuming in daily activity, in particular solid nodules must be differentiated from pseudo nodular findings such as intraparenchymal lymph nodes, obliterated bronchiectasis, atelectatic lung etc.

Some recently published studies show promising results about AI (artificial intelligence) applications for detection and characterization of lung nodular lesions [25][26].

The aim of CDSS is to create a multidisciplinary team to retrospectively review CT scans performed in our Institution and to create a database of anonymized cases to train and validate neural networks in lung nodules identification, in order to obtain a potentially helpful aid for daily activity of thoracic radiologist.

Objectives of the use case	
The objective of UC4 is to detect pulmonary nodules, to be used for clinical purposes.	
Type of problem	**Objective**
Segmentation	Detection of pulmonary nodules.

The use case outcome can be measured through the following metrics.

Known KPIs	
KPI 1	Time-to-train model.
KPI 2	Statistical accuracy (F-score, intersection over union for segmented nodule area/ identification of the areas, false negative-false positive rate, ROC curve).

2.4.2. Partner roles and contributions

This use case is under the responsibility of CDSS that will provide all expertise according to the following table. The use case will be deployed on PF5.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	CDSS	Select records to be included in the dataset based on the diagnosis, provide manual segmentation.
End user (medical personnel)	CDSS	Evaluate results in terms of KPI (e.g. time to training including manual segmentation effort).
Machine learning expert	UNITO	Design/select neural network for image segmentation, provide reference software for training, validation and testing.
Platform provider	UNITO	Deploy experiment on PF5.

2.4.3. Technical requirements

2.4.3.1. Dataset description

The dataset for training has the following characteristics.

Type and format	512x512, 16-bit depth gray-scale CT scans (approx. 200 per exam), DICOM format, anonymized.

Annotation metadata	Nodule localization (slice number and position in pixel matrix, including nodule ray and diameter).
Number of samples (retrospective/ prospective)	2000 nodules (single or multiple, depending on Patient) retrospectively identified in CT scans available in hospital archives, performed between 2012 and 2018. Each data set consists of a variable number of images (between 180 and 400 depending on scan length), all of which are reconstructed with a thin slice thickness (1.25-2 mm).
Repository/sharing method	Hospital Picture archiving and communication system (PACS), sharable with UNITO internal network according to agreements done.
Other information	To avoid diffusing sensitive data, images will be anonymized before use.

2.4.3.2. Pre-processing

The following image pre-processing steps will be applied:

- DICOM reading
- Histogram Equalization methods

2.4.3.3. Model design and testing

The following network architectures will be used: U-Net, X-Net or a CNN-based architecture in general. The training methods foreseen are Adam and SGD. A loss function will be applied, e.g. binary cross-entropy.

2.5. UC5 - Deep Image Annotation

This Use Case will deal with automatic annotation of image-based medical examinations with text in natural language.

2.5.1. Use case description and objectives

Various official statistics show that the number of medical examinations relying on images has increased dramatically over the past years [75]. Not only has the number of examinations increased, but also, due to higher resolution scanners, the information provided by each imaging modality. Indeed, a recent study shows that radiologists need to review one image every three to four seconds to meet the demand of their medical centers [29]. In order to reduce the overall time, they often complete the textual reports by means of “copy and paste” sentences from previous texts. Even if retyping the information for each single report could lead to more errors, “copy and paste” is a major source of medical errors, leading to unnecessary and sometimes expensive follow-ups [33]. This use case aims at implementing and evaluating machine learning models for the automatic generation of medical reports, learning from datasets composed by image examinations and related reports.

Until a few years ago, approaches tackling this problem were mainly focused on extracting disease names from the textual collection and attaching them to the image examinations as labels. Recently, developments in Deep Learning models allowed to automatically generate natural text from features extracted with a multimodal approach, where images and related text are combined in a joint semantic description.

[31] presents a deep learning model that, given a training dataset composed by medical examinations and related textual reports, annotates new examinations with the disease name, its location, and other information. The model is composed by two cascaded deep neural networks: image information is extracted by a CNN, trained to classify an image in a set of categories corresponding to frequent co-occurring MeSH (Medical Subject Headings) tags. The final embeddings are then used for training a recurrent neural network to generate 5-words long sentences. Eventually, the CNN embeddings and the Recurrent Neural Networks (RNN) context vectors are joined to form a new set of labels that, when used in a new training stage, leads to better results. The system is evaluated using BLEU scores

[30]. The authors tested four different neural networks: two CNNs (NiN, GoogLeNet), and two RNNs based on LSTM and GRU units. One of the most interesting results of this paper is the re-training strategy, where labels for the CNNs are joint image-text vectors. The paper uses the Indiana University Chest X-Ray Collection [27], that includes 7470 X-Ray images and 3955 associated reports (OpenI collections [76]).

[28] have recently published a model based on the same CNN-RNN combination as (Shin et al., 2016), but with a hierarchical LSTM endowed with attention mechanisms, that have proven very effective in image captioning tasks. Their model is able to generate texts of arbitrary lengths. First, the CNN (VGG-19 model, [32]) is trained for performing multi-label classification, using the tags specified in the textual reports as output labels. The CNN embedding and the most likely tags are then fed to a sequence of LSTM-based networks. The first LSTM network is trained for generating sentences: based on the input visual embedding and the tags, it establishes if more data is to be fed to the second LSTM. This network includes an attentional mechanism processing both the visual and the semantic embeddings. The second LSTM-based network generates words (word embeddings) based on the input received by the sentence LSTM. The final number of words is not fixed, as in [31], but depends on a deep output layer in the first LSTM. Such layer is used to stop the generation of words in the current sentence and or the generation of new sentences. The authors used several metrics to evaluate their new model (BLEU, METEOR, ROUGE, CIDER, see the original papers for references). They experimented their model on two different datasets: the Indiana University Chest X-Ray Collection [27] and a subset of the Pathology Education Informational Resource (PEIR) digital library [77], where the text associated to images is only one-sentence long. In the experimental comparison with different neural architectures, including a variant of the proposed CNN-RNN combinations without attentional mechanisms, the full attention-based model obtained the best quantitative and qualitative scores.

In this use case we will implement the neural architectures described in [31] and [28] and experiment them on the datasets (composed by image-based medical exams and associated textual reports) provided by the partners.

Two public datasets available for developing and testing the software in UC5 are:

- Indiana U. Chest X-Rays [78].
- Chest dataset of the PERI collection [79].

The first dataset contains longer texts than the second.

The major use case objectives are reported in the following tables, referring to objective and performance metrics that will be used.

Objectives of the use case	
Automatically annotate medical examinations with sentences in natural language that describe the relevant visual content.	
Type of problem	**Objective**
Text generation	Annotate medical images

Known KPIs	
KPI 1	Bilingual evaluation understudy **(**BLEU).
KPI 2	Crowd-based human evaluation (online platforms).

2.5.2. Partner roles and contributions

This use case is under the responsibility of UNITO that will provide all expertise according to the following table. The use case will be deployed on PF5.

Roles and actions		
User	**Partner**	**Action**
End user (expert user)	UNITO	Provide UC5 compliant dataset using publicly available data.
Machine learning expert	UNITO	Design text generation network, provide reference software for training and testing.
Platform provider	UNITO PHILIPS	Deploy experiment on PF5. Test CNN model on PF1.
End user (medical personnel/ expert user)	UNITO	Discuss achieved KPI.

2.5.3. Technical requirements

2.5.3.1. Dataset description

We expect to receive a dataset that includes image-based medical examinations with the associated textual reports. Each dataset must be composed of images and reports from a single 2D modality.

Type and format	Images: RGB format (with original bit depth) from 2D image-based examinations. Text: UTF-8 plain text
Annotation metadata	Images: No metadata or annotation required. Text: No specific format required. If the reports have a structured format, the semantics of each section should be clearly specified.
Number of samples (retrospective/prospective)	In the order of 10k (retrospective).

2.5.3.2. Pre-processing

The pre-processing will deal with both textual and image data.

Textual reports will undergo the following steps:

1. Tokenization
2. Computation of word embeddings, in the language used in the reports, if no other pre-computed embeddings are available (word embeddings must be available or compute, with a word2vec model, before textual data can be processed).
3. Images will be processed with the following pipeline:
4. RGB to grayscale conversion, if needed
5. Equalization and thresholding
6. Rescaling

2.5.3.3. Model design and testing

The model will be deployed and tested according to the following network architectures:

Network architecture	VGG-19.
Training method	SGD (and any of the standard optimizers).
Other	Learning rate decay, dropout.
Network architecture	LSTM, GRU (sequence-to-one, sequence-to-sequence) with and without attention.

Training method	SGD (and any standard optimizer).
Other	Learning rate decay, dropout.
Network architecture	Word2vec (CBOW and Skip-Grams).
Training method	Standard.

2.6. UC6 - Promort

This use-case provides a vertical application to support digital pathology in the context of prostate tumor diagnosis.

2.6.1. Use case description and objectives

In this use case, a vertical application will be created that will use EDDLL-based models to automatically suggest annotations for prostate tumor slides. The application will present these suggestions to a pathologist for validation and, if necessary, correction. The DeepHealth libraries will be used to exploit the Promort image dataset to train effective models and then to integrate the prediction functionality into the new vertical application. The Karolinska Institute will test this application in a clinical research setting, where it will be inserted in a continuous model improvement cycle.

The main steps of a typical workflow for this application are: extraction, preprocessing, training and prediction. The problem of automatically annotating prostate tumor slides with machine learning methods has not received much attention from the scientific community to date. However, initial experiments show that the techniques used on breast cancer provide a good starting point. Our initial work is based on various state-of-the-art implementations available in the CAMELYON16 challenge results [34]. In that contest, several research teams worked to tackle the problem of recognizing breast cancer areas in histopathological slides. Different approaches were used both at the preprocessing level – e.g., using stain normalization of slides – and during training/prediction with the use of different neural network architectures and the image resolution. It is worth noting that both the best [35] and one of the worst performing models share the same neural network architecture (the GoogLeNet). This observation highlights the importance of working on the details of extraction, preprocessing, training and prediction to obtain a good performance.

Objectives of the use case	
Type of problem	**Objective**
Image segmentation	Automatically annotate probable tumor areas
Function approximation	Estimate Gleason Score

The performance can be measured according to the following KPIs.

Known KPIs	
KPI 1	Time-to-model-in-production (**ttmip**).
KPI 2	Time-of-training-models (**totm**).
KPI 3	Time to annotate slide, measured in person-hours (ttas)

2.6.2. Partner roles and contributions

UC6 is a joint responsibility of the Karolinska Institute (KI) and CRS4. In particular, KI will mainly provide medical expertise while CRS4 will be responsible of the development of the vertical application and of the model design.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	KI	Evaluate generated annotations, provide corrections. Provide feedback on UI and application functionality.
End user (expert user)	KI	Support integration in clinical research workflow. Support interaction between medical and technical roles.
Machine learning expert	CRS4	Model design and development.
Platform provider	CRS4	Development of vertical application.

2.6.3. Technical requirements

2.6.3.1. Dataset description

Type and format	Digital slides, MIRAX or NDPI format.
Annotation metadata	Tissue status classification, Gleason score.
Number of samples (retrospective/prospective)	5000 annotated slides.
Repository/sharing method	Images will be copied to CRS4.
Other information	Annotations acquired through Digital Pathology platform (PF6). Data format is not an issue.

2.6.3.2. Pre-processing

The WSI collection, made of digitalized slides, is produced using scanners. Slides must be pseudonymized (e.g., no labels on the slides, no reference to the patient in the filename) in order to be uploaded on the platform.

Beyond this point, a generation of slide patches and masks step is necessary. Here, tissue patches and related masks are produced using the annotations produced by the pathologists through the usage of the dedicated application; in particular:

- patches are extracted from digital slides at different zoom levels and sizes.
- patches are extracted producing multiple versions of each region of interest in order to find the best fitting inside the output image.
- each patch has a set of binary masks associated representing tissue classifications (normal, stressed, tumoral or unclassified).
- each patch/mask pair is completely anonymous and has no reference to its position within the digital slide.

2.6.3.3. Model design and testing

At this stage, support for a variety of topologies (e.g., GoogLeNet, VGG, U-Net) is still required to advance the work on semantic segmentation both at patch and pixel level. It is planned, in order to perform the training step, to use standard, state-of-the-art optimizers [63][64][65].

2.7. UC7 - Major Depression

In this use-case, the main target of the application is to perform classification of disordered vs. healthy states. To this end, state-of-the-art machine learning based algorithms will run detection algorithms

over big amounts of neuroimaging and biological marker data. A secondary aim is to establish predictors for a change in clinical state (e.g. prediction of therapy outcome under treatment or prediction of the possibly most successful treatment forms). The end user will take advantage of Hybrid and heterogeneous HPC and Big Data clusters with various machine learning based detection algorithms whose implementation will require no mathematical or IT-knowledge on the end user side.

2.7.1. Use case description and objectives

Major Depressive Disorder (MDD) affects 15% of the population. Many of them are not in adequate treatment because the medical system does not have sufficient capacity, or they are unaware of their disorder. In future, support systems will be able to help medical staff in diagnosing and monitoring disorders in cases where specialist services are not available.

MDD is a multi-factorial disorder with environmental, genetic and biological factors being relevant for vulnerability to and onset for the disease. Moreover, these factors influence the course of the disease, e.g. impacting on brain structure and function. A possible pathogenic mechanism of the regional volume reductions reported in the brain of depressed patients might be a glucocorticoid-mediated inhibition of proliferation and neurogenesis as a consequence of a chronic activation of the hypothalamus-pituitary stress axis (HPA) [53][54][55]. Interestingly, the resulting volume reductions can be partially reversed with antidepressant treatment [56]. In a sample of 958 MDD patients and 2078 healthy controls MDD patients taking antidepressants had larger volumes compared to patients without antidepressant medication [57]. Furthermore, preclinical studies suggest that the effect of HPA dysregulation on hippocampal volume reduction, neurogenesis and their reversal by antidepressant response vary across hippocampal regions and are especially present in CA1, CA3 (atrophy of dendrites) and dentate gyrus (neurogenesis) [53] [56]. The use case just will use biological, clinical and data from neuroimaging to help classifying MDD and predicting illness course.

UC7 will show the clinical applicability on improved and novel platforms. Also, specialist services and hospitals will benefit from data integration and diagnostic approaches including deep learning and artificial intelligence. Thus, a huge market exits in this field of health services. This can be extended to the preventive area, e.g. to physical exercise monitoring including psychological states.

The main objectives of this use-case are summarized in the following table.

Objectives of the use case	
Type of problem	**Objective**
Classification	Can patients and healthy controls be classified based on the existing data?
Prediction	Can the baseline measurement predict disease progression? Can a difference between baseline and follow-up measurements predict disease progression?
Categorisation	What categories (neurobiological entities) do the data cluster into and do these match ICD-10 disorders and how do they predict illness course variables?

In order to evaluate the final performance, a table summarizing the known KPIs follows.

Known KPIs	
KPI 1	Time-to-model-in-production (ttmip).
KPI 2	Time-of-training-models (totm).
KPI 3	Classification accuracy (error rate, F-score, etc.)
KPI 4	Prediction accuracy

2.7.2. Partner roles and contributions

For this use-case, OVGU (Universitätsklinik für Psychiatrie und Psychotherapie - KPSY) will be in charge of selecting data to be used for the neural network's training. Such an operation will be led by medical personnel. After data selection, expert users by OVGU will process data and prepare the dataset, as well as measure KPIs. Then, the development of the model will be OGVU's (Institut für Automatisierungstechnik - IFAT) and THALES's responsibility. Training and testing models will be responsibility of OVGU-IFAT, THALES, Philips and WINGS. The final evaluation of the achieved accuracy will be performed by the medical personnel at KPSY.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	OVGU (KPSY)	Select data to be included in the overall dataset based on the diagnosis and completeness of individual participant datasets, provide classification using taxonomy to be agreed. Clarify with the university's IT and data protection departments and ethics board how data can be shared with partners and how AI platforms can be installed in the department of psychiatry.
End user (expert user)	KPSY, OVGU (IFAT)	Export dataset using DICOM format plus data and metadata in CSV format, measure current toppi.
Machine learning expert	OVGU (IFAT)	Develop model, implement training and testing.
Platform provider	THALES Philips WINGS	Develop model, implement testing . Implement testing. Implement training and testing.
End user (medical personnel)	OVGU (KPSY)	Evaluate the achieved accuracy.

2.7.3. Technical requirements

2.7.3.1. Dataset description

Here, three datasets will be used for training: MRI Study (retrospective), Cambridge data set (retrospective) and prospective questionnaire data. A detailed technical description for them follows.

MRI Study Dataset	
Type and format	MRI-data in DICOM format. MRI segmentation values, blood values, demographic data, details about disorder, drug consumption, body measurements (total 447 variables) in CSV format 804kB.
Annotation metadata	Gender, age, education, handedness, marital status height, weight & BMI are available.
Number of samples (retrospective/ prospective)	150 with MDD, 200 controls, retrospective data from case-control study.

Repository/sharing method	To be determined in association with the university's IT and data protection departments and ethics board.
Other information	MRI data is available for all participants, however not every participant has all remaining data. Data was collected at 2 MRI scanning sites, each with the same MRI-scanner, but using different head coils (8 and 32 channel).
Cambridge data set	
Type and format	MRI-data in DICOM format, approximately 60MB per participant. Genetic markers, blood values, demographic data, details about disorder, drug consumption, body measurements in CSV format.
Annotation metadata	N/A
Number of samples (retrospective/ prospective)	Retrospective data from case-control study.
Repository/sharing method	To be determined in association with the university's IT and data protection departments and ethics board
Treatment monitoring data	
Type and format	Time series data from questionnaire based on the ICD-10 and and DSM 5 criteria of depression as well as additional variables such as medical history. The use of wearable devices is also being considered in order to minimize the information that needs to be provided by the patient and receive more objective measurements. All data will be stored in a CSV format.
Annotation metadata	N/A
Number of samples (retrospective /prospective)	To be determined (prospective, cohort study)
Repository/sharing method	Stored in WINGS infrastructure but shared among both OVGU and WINGS partners upon patients consent.

2.7.3.2. Pre-processing

In order to pre-process data, segmentations software, such as FreeSurfer will be used. In particular, the following features will be extracted from the data:

- Region of interest (ROI) volumes
- Average cortical thickness of ROIs

2.7.3.3. Model design and testing

It is here planned to use a convolutional neural network model to solve the task. In particular the following models proposed in the scientific literature will be applied in the present use case:

- U-Net [49] for segmentation task;
- ResNet [20] for classification task.

In the case of the treatment monitoring data, machine learning classification mechanisms are considered in order to identify patients with similar profiles. Mechanisms based on deep learning will be trained for building the models that "describe" the evolution of the depression and the response of the patient to the suggested treatment.

2.8. UC8 - Dementia

In this use-case, the main target is to perform classification of healthy states vs. several forms of dementia (Alzheimer, vascular, fronto-temporal, mixed) and mild cognitive impairment. Also, an aim is to find predictors for disease progression that can be helpful for clinical guidance or in the decision for the necessary support that the subject needs in the near future. To this end, state-of-the-art machine learning based algorithms will run detection algorithms over biological markers as well as neuroimaging data. The end user will take advantage of hybrid and heterogeneous HPC + Big Data clusters with various machine learning based detection algorithms whose implementation will require no mathematical or IT-knowledge on the medical side.

2.8.1. Use case description and objectives

Dementia is an age-related disorder affecting 46 Million people [81]. Therapy and enhancing of cognitive reserve can be most effective when accurate diagnosis and very early diagnosis can be achieved. The diagnostic approaches being involved are interdisciplinary, including disciplines like psychiatry, neurology, neuroradiology, and clinical chemistry. Decision support systems will be most effective when they are running on fast platforms. This system could be used by all general practitioners, neurologists, psychiatrists and psychologists. Today in the hospital the IT systems are too slow that they can be of bigger benefit to the health care providers, but in the future those techniques fulfilling the requirements will win the market.

The main goals of this use-case are summarized in the following table.

Objectives of the use case	
Objective	**Type of problem**
Classification	Can forms of dementia (Alzheimer's disease, frontotemporal dementia, vascular dementia) and healthy controls be classified based on the existing data?
Prediction	Can the baseline measurement predict disease progression? Can a difference between baseline and follow-up measurements predict disease progression?
Diagnostic value	Can subgroups of patients be defined based on neurobiological markers using AI models that better represent the disease course than current diagnostic systems?

Here follows a table listing all the identified KPIs for the current use-case.

Known KPIs	
KPI 1	Time-to-model-in-production (**ttmip**).
KPI 2	Time-of-pre-processing-images (**toppi**)
KPI 3	Classification accuracy (error rate, F-score, etc.)
KPI 4	Prediction accuracy

2.8.2. Partner roles and contributions

For this use-case, OVGU (Universitätsklinik für Psychiatrie und Psychotherapie - KPSY) will be in charge of selecting data and establishing a biomarker databank to be used for the neural network's training. Such an operation will be led by medical personnel. After data selection, expert users by OVGU will export the dataset. Then, the implementation of the model, as well as training and testing, will be OVGU (Institut für Automatisierungstechnik – IFAT), THALES, Philips and WINGS

responsibility. The final evaluation of the achieved accuracy will be performed by the medical personnel at OGVU.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	OVGU (KPSY)	Select images to be included in the dataset based on the diagnosis, provide classification using taxonomy to be agreed. Establish a biomarker databank in XLS format that can be used with machine learning algorithms.
End user (expert user)	OVGU (KPSY)	Export dataset using DICOM format plus XML metadata.
Machine learning expert	OVGU	Implements model, training and testing.
Platform provider	THALES Philips	Implement testing. Implement testing.
End user (medical personnel)	OVGU (KPSY)	Evaluate the achieved accuracy.

2.8.3. Technical requirements

2.8.3.1. Dataset description

Type and format	Dementia biobank
Annotation metadata	N/A
Number of samples (retrospective/ prospective)	200 patients with different types of dementia and 100 healthy controls. Furthermore, there are prospective longitudinal data available from these patients. Retrospective data.
Repository/sharing method	Based on data protection guidelines these data cannot be shared openly, they can only be commonly used together with partners for data analysis purposes, machine learning.
Other information	A data collaboration contract needs to be set up.

2.8.3.2. Pre-processing

Please refer to section 2.7.3.2.

2.8.3.3. Model design and testing

Please refer to section 2.7.3.3.

2.9. UC9 - Study of structural changes in lumbar spine pathology

The present use case aims to develop an infrastructure with massive storage capacity and intensive computational modelling of lumbar magnetic resonance images together with clinical data of patients with low backpain to investigate possible correlations between pathology and structural alterations.

2.9.1. Use case description and objectives

The analysis of massive data of medical images combined with digitalized clinical information is a powerful tool in biomedical research because it allows for the definition of associations between different pathologies and their image biomarkers with reliability. To this end, an initial sample of 8,597 patients, with a total of 42,436 MRI Scans (Axial-T2, Sagittal-T1, Sagittal- T2 and Sagittal-STIR) from the Public Health Service of the Valencian Region was used to create a population database of medical imaging associated with Electronic Health Records, anonymized and interoperable between the Digital Medical Image Management (GIMD, from its Spanish acronym) project and the Medical Imaging Databank of the Valencian Region (BIMCV). In the implementation of the image database, the following parameters were evaluated: anonymization methods, description of the acquisition protocols in the DICOM headers, correction of inhomogeneity, methodology associated with the anatomical segmentation of lumbar spine structures, based mainly in the Expectation-Maximization (EM) algorithm for estimating Gaussian mixture models (GMM), and the application of a matching-squares (MS) algorithm to determine anatomical contours, degenerative alterations and Modic changes. GMMs allow for the detection of the vertebral structures, which are refined with the MS algorithm. It is possible to identify and computationally label the anatomical contours of bodies, facets, ligaments, discs, foramina, spinal canal and nerve structures. These data are suitable for the application of the Machine Learning techniques and for training Deep Neural Networks that model the complex variations in the Magnetic Resonance Image (MRI) in an automated way.

The following tables show the objectives established for this use case and the metrics employed to measure the achieved results.

Objectives of the use case	
Type of problem	**Objective**
Data Curation	Provide reliable knowledge of the anatomical changes in MRI.
Classification	Establishing the indicators and their structural correlations with each other and with clinical data.
Prediction	Predict the physio pathological process of aging and the evolution of low back pain.

Known KPIs	
KPI 1	Enhancement of the reliability of lumbar spine identification.
KPI 2	Accuracy for the segmentation of lumbar spine structures.
KPI 3	Time-of-training-models (totm). Reduction of training and evaluation times of brain images segmentation and classification.
KPI 4	Time-of-pre-processing-images (toppi). Reduction of the time to transform and feature extraction from images in order to prepare them as input to DL models.
KPI 5	Time-to-model-in-production (ttmip). Reduction of the time for integrating the developed models into a software platform for making possible to doctors the use of the models in the process of diagnose.

2.9.2. Partner roles and contributions

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel & data curators)	FISABIO	Select records to be included in the dataset based on the diagnosis and provide labels for each image.
End user (data curators & expert user)	FISABIO	Export dataset using Medical Imaging Data Structure (MIDS).
Machine learning expert	UPV & FISABIO	Design/select the Deep learning techniques and DNN topologies to be used for the models.
Platform Provider	CEA UNITO Philips THALES EVERIS	Deploy training and testing on PF3 Deploy training and testing on PF5 Testing on PF1 Testing on PF4 Deploy training and testing on PF7
End user (medical personnel)	FISABIO	Evaluate the achieved accuracy.

2.9.3. Technical requirements

2.9.3.1. Dataset description

This use case will exploit a dataset of 42,436 MRI Scans (Axial-T2, Sagittal-T1, Sagittal-T2 and Sagittal-STIR) available from the Public Health Service of the Valencian Region. The technical details of the provided data are specified in the following table.

Type and format	DICOM and NifTI on MIDS [82]
Annotation metadata	Integrated in MIDS too
Number of samples (retrospective/prospective)	+8,597 patients from BIMCV database Retrospective
Repository/sharing method	File care/ Not to be shared for privacy reasons

2.9.3.2. Pre-processing

1. Correction of the manual segmentation dataset (46 segmented sessions will be corrected and new tags will be added in nifty format).
2. Segmentation with multi-power inputs (T1, T2, STIR)
3. The Unet network will be modified to receive several image modalities at the same time. It will be trained with the T1 and T2 sequences.
4. Anonymization exploiting the modules defined in the DICOM format.

2.9.3.3. Model design and testing

As the size of the dataset is large enough and images are samples in a high dimensional space, the needs for training and testing different DNN topologies require computational power. So, the main requirement is hardware, mainly on premise, for training the models.

- Different topologies of convolutional neural networks need to be tested in order to compare and select the best models. Both widely used topologies as VGG16, VGG19 [32] and ResNet [20] by one hand, and ad-hoc topologies designed by FISABIO and UPV will be tested.

- Statistical analysis by identifying different areas in the brain in order to compute the value of different parameters.

2.10. UC10 - Predictive and Populational Model for Alzheimer's Disease (AD) using Structural Neuroimaging

In UC 10 we aim at correlating brain identified surface areas according to different criteria (age, gender, pathologies, and geographical area) in the population of the Valencian region. To achieve this purpose, the abovementioned MRI images will be used. This kind of images make possible to quantify all the structures of interest, in order to compare which areas of the brain produce the possible variability between healthy and pathological subjects, specifically in relation to Alzheimer's disease.

2.10.1. Use case description and objectives

The Valencia Regional Ministry for Health needs to make the best use of the vast amount of health data of its population which is acquired and stored in the framework of its public service. Among those datasets, biomedical images contained in the Imaging Biobank of the Valencia Region are the ones offering a higher potential for investigating new avenues for advanced public health service. The Big Data in Brain Imaging aims at exploiting the rich information of MRI images from more than 10,000 individuals. This dataset provides detailed anatomical and morphological data of the brain and is continuously subject to a process of Quality Control to ensure that the Biobank provides a reliable and robust information repository for advanced research and health services.

In order to progress towards this objective, it is essential to discriminate relevant brain structures. In this respect, the detection of substructures beyond the three main brain tissues (WM, GM, and CFS) is a complex subject because they are defined by weakly visible limits. The algorithms are not only based on MRI, but they are guided by previous information in relation to anatomical structures, patterns. One of our main objectives is to correlate three measures (volume, area and cortical thickness) of the brain identified surface areas according to several criteria (age, gender, pathologies and geographical area) by using MRI images from the population of the Valencian area. This kind of images make possible to quantify all the structures of interest, in order to compare which areas of the brain produce the possible variability between healthy and pathological subjects, specifically in relation to Alzheimer's disease.

The following tables show the objectives established for this use case and the metrics employed to measure the achieved results.

Objectives of the use case	
Type of problem	**Objective**
Classification	Design and develop populational models of the volumetric degeneration of the gray matter due to Alzheimer's disease.
Classification	Establishing the indicators of degeneration. The indicators will be extracted from the developed models according to their relevance in the models.
Prediction	Assess the health status of an individual with respect to that model.

Known KPIs	
KPI 1	Enhancement of the reliability in the identification of brain areas.

KPI 2	Segmentation of brain structures.
KPI 3	Time-of-training-models (**totm**). Reduction of the time for training and evaluating models to segment and classify brain images.
KPI 4	Time-of-pre-processing-images (**toppi**). Reduction of the time to transform and feature extraction from images in order to prepare them as input to DL models.
KPI 5	Time-to-model-in-production (**ttmip**). Reduction of the time for integrating the developed models into a software platform for making possible to doctors the use of the models in the process of diagnose.

2.10.2. Partner roles and contributions

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel & data curators)	FISABIO	Select records to be included in the dataset based on the diagnosis and provide labels for each image.
End user (data curators & expert users)	FISABIO	Export dataset using MIDS format.
Machine learning experts	FISABIO & UPV	Design/select the Deep learning techniques and DNN topologies to be used for the models.
Platform Provider	UNITO Philips EVERIS	Deploy training and testing on PF5 Testing on PF1 Deploy training and testing on PF7
End user (medical personnel)	FISABIO	Evaluate the achieved accuracy.

2.10.3. Technical requirements

2.10.3.1. Dataset description

The present use case aims at exploiting the rich information of MRI images from more than 10,000 individuals. This dataset provides detailed anatomical and morphological data of the brain, and it is composed by 2D clinical images corresponding to an average of 24 slices per session, and 6 sessions per patient in average, so the total number of 2D images is more than 1.4 Million of images. When reconstructed into 3D images more than 60,000 samples will be available.

Type and format	DICOM and NifTI on MIDS [83][84]
Annotation metadata	Integrated in MIDS too
Number of samples (retrospective/prospective)	+10,000 individual records from BIMCV database, consisting in +1,400,000 2D images. Retrospective.
Repository/sharing method	File care / Not to be shared for privacy reasons.

2.10.3.2. Pre-processing, Model design and testing

The same aspects and research approaches described for the previous use case (see Sect. 2.9.3.2 and Sect. 2.9.3.3) apply.

2.11. UC11 - Image Analysis and prediction for Urology

Nowadays in urology, the image analysis is of utmost importance for a correct diagnosis. Urologic Radiology includes several ways to carefully look inside organs like the kidneys, ureters, bladder, and reproductive organs. Imaging tools can be basic like x-rays to more complex tools. Generally, these imaging tests are detailed, often pain-free, and quick. The various existing imaging technologies have meant greater diagnostic power and lead to an important increase in the number and quality of images obtained and viewed by practicing clinicians, mainly in the field of urology, where static and dynamic images are fundamental to the diagnosis and treatment of almost all conditions. Our focus will be on the imagistic diagnosis of kidney tumours (benign and malignant) and also adrenal tumours, for which the use of artificial intelligence for diagnosis purposes is rarely described in the specialty literature. A main target of this use case is developing a data library of cases with anonymized imaging data obtained from the patients with renal and adrenal tumours together with their radiological description and diagnosis. Using the DeepHealth toolkit, we will investigate if the application of machine learning and artificial intelligence can improve the diagnostic capabilities, possibly leading to changes in the standard of care for these conditions and expansions in the coverage for new diagnostic imaging modalities.

2.11.1. Use case description and objectives

Kidney parenchymal cancer represents 2-3% of all cancers, presenting with the highest incidence in Western countries. Over the last two decades the incidence of renal cell carcinoma (RCC) increased by about 2% worldwide. In Europe, in 2012, there were approximately 84,400 new cases of renal cell carcinoma and 34,700 kidney-cancer-related deaths in the European Union [41][42][43]. In Europe, overall mortality rates for RCC increased up to the early 1990s, before stabilising or declining thereafter. Mortality has decreased since the 1980s in Scandinavian countries and since the early 1990s in some countries like France, Germany, Austria, the Netherlands, and Italy. However, in other European countries like Croatia, Estonia, Greece, Ireland and Slovakia, mortality rates still show an increasing trend. Data from the United States also show an increased incidence [41][44]. In the latest decades, the improvement and the spreading of the imagistic diagnosis techniques has brought its contribution to the diagnosis and treatment of these tumours. Due to the increased detection of tumours by ultrasound (US) and computed tomography (CT), the number of incidentally diagnosed RCCs has increased, therefore these tumours are usually smaller and of lower stage, permitting in several cases a treatment with curative intent and prolonging the cancer specific survival [41][45].

Nowadays, worldwide the computed tomography is used to characterise renal and adrenal masses, with the imaging performed before, and after, administration of intravenous contrast material to demonstrate enhancement. In CT imaging, enhancement in renal masses is determined by comparing Hounsfield units (HUs) before, and after, contrast administration. A change of fifteen, or more, HUs demonstrates enhancement. Computed tomography usually allows accurate diagnosis of RCC (Renal Cell Carcinoma) but cannot reliably distinguish other types of benign or borderline renal tumours, like oncocytoma and fat-free angiomyolipoma, from malignant renal neoplasms. Abdominal CT provides information on function and morphology of the contralateral kidney, primary tumour extension, venous involvement, enlargement of loco-regional lymph nodes, condition of the adrenal glands and other solid organs, aspects that are important for the diagnosis, staging and treatment method [41][46][47][48].

Renal tumour segmentation and analysis is a very important step for doctors in deciding the stage of cancer and determining the treatment method of choice. In this direction, solutions based on Artificial Intelligence (AI) algorithms and DL (Deep Learning) techniques can constitute a powerful mechanism for enhancing the diagnosis and staging abilities of renal and adrenal tumours, in the way that a computer can be “taught” to recognize a certain pathology after a certain number of cases (pathological and normal cases described by the radiologist) is defined in a data library. Compared to traditional regression statistics, artificial intelligence methods appear to be accurate and more explorative for analysing large data cohorts

This use case will deliver a novel application accompanied by a cloud-based platform that will provide urologists, radiologists and oncologists in our hospital and maybe other hospitals and medical departments with better diagnostic capacities, allowing them for a better staging and treatment method for the respective patient category.

The main use case tasks and goals are detailed below.

Objectives of the use case	
Type of problem	**Objective**
Creating a data library	Creating a data library of imagistic acquisitions on the PF1 platform with renal tumour cases (RCC, renal cysts, oncocytoma, etc.), adrenal tumour cases and normal cases (no pathological modifications).
Describing the regions of interest	Providing description and radiological diagnosis for the regions of interest for the introduced cases.
Training the model	Train the AI to do classification of the images using pre-trained CNN model using newly introduced patient images to acquire the diagnosis in the regions of interest for the physician
Future exploitation	Development of an application to be used in our department and maybe other departments by urologists, radiologists and oncologists for radiological diagnosis in renal and adrenal tumours

The metrics used to rank the achieved results are listed in the following table.

Known KPIs	
KPI 1	Time-to-model-in-production (**ttmip**).
KPI 2	Time-of-training-models (**totm**).
KPI 3	Time-of-pre-processing-images (**toppi**).
KPI 4	Kidney segmentation
KPI 5	Adrenal gland segmentation
KPI 6	Classification of Renal Cell Carcinoma (RCC)

2.11.2. Partner roles and contributions

The use case will be worked out with the contributions and roles summarized in the following table.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	SCTH	Evaluate results in terms of KPI (e.g. time to training including manual segmentation effort).
End user (medical personnel)	SCTH	Evaluate generated annotations, provide corrections. Provide feedback on UI and application functionality.

End user (expert user)	SCTH	Support integration in clinical research workflow. Support interaction between medical and technical roles.
Machine learning experts	SIVECO	Model design, implementation and training.
Machine learning experts	SIVECO	Development of application.
Platform provider	Philips	Testing on PF1

2.11.3. Technical requirements

2.11.3.1. Dataset description

A dataset comprising at least 500 cases of renal and adrenal tumours imaging and 500 cases of normal kidneys and adrenals will be collected.

The dataset sample will be constituted by:

- CT scan data before and after contrast administration (DICOM images) – approx. 1 GB of information for each patient.
- Radiological description (user profile information and daily form information): DICOM images
- Radiological diagnosis for the specific cases (DICOM images).
- Histopathological diagnoses for the adrenal tumours data in order to enhance the diagnostic sensitivity of the newly introduced case in the application.

2.11.3.2. Pre-processing

Different kind of pre-processing will be applied to the data sample as detailed below.

- CT imaging data: time series epoching/segmentation, filtering, artefact detection/rejection.
- Radiological description of the areas of interest: the areas of interest will be highlighted, and radiological description will be added.
- Radiological diagnosis: radiological diagnosis will be provided for each data sample.

2.11.3.3. Model design and testing

This use-case will use a CNN model for image classification and segmentation in the process of diagnose of renal tumours. The model will help physicians to improve detection and classification.

Deep neural networks will be trained on CT data and radiological description and diagnosis. The architecture of the model will contain the best Deep Learning techniques that will be detailed at a later date.

2.12. UC12 - Skin cancer melanoma detection

This Use Case will consider skin cancer, which is the most common cancer in humans, and it has an increasing trend. It affects mainly Caucasians and accounts for a very high probability in life, especially in relation to increased life-span. Even if its mortality rate remains low, its morbidity is relevant in the population, demanding for surgery frequently occurring on the face with disfiguring consequences. This UC is based on data available from a public online archive and an internal database which will be organized and made public within the ethical requirements. The aim is to provide a useful tool from the segmentation of medical images and automated the melanoma detection, from dermoscopic images.

2.12.1. Use case description and objectives

Among skin cancers, melanoma accounts for a relevant risk of tumor dissemination and metastatization and consequent high risk of death. Overall, the major problem related with skin cancer

is delayed diagnosis. For epithelial skin cancer, such as basal cell carcinoma and squamous cell carcinoma, delayed diagnosis is mainly responsible for a larger and more disfiguring surgery, which could also affect relevant functional structures (e.g. mouth, ears, eyelids, nose). For melanoma a delayed diagnosis may mean death due to potential tumor aggressiveness. Early diagnosis represents the ideal and cheap solution to fight against the skin tumor consequences. However, this requires a high expertise, linked with dermatology specialty and availability of technologies. Dermoscopy is reported to improve diagnostic accuracy of melanoma by 5-30% compared with simple clinical observation. However, dermoscopic techniques require formal training and skill in image interpretation through so-called pattern analysis [36][37][38]. Whereas dermoscopy, which provides magnified skin images of the tumor and visualization of sub-surface structures, is largely used in Europe, in vivo confocal microscopy, which provides nearly histologic information without tissue excision, is only available in the main Research and University hospitals requiring a specific expertise for image interpretation. At UNIMORE the world largest dataset of images of skin cancer is available, completed of clinical, dermoscopy and confocal microscopy images, annotated with conclusive diagnosis (histologic or clinically confirmed), and relevant patient's data.

Segmentation of images and extraction of features is able to lead to a rapid and automatic identification of diagnostic clues, which can facilitate image interpretation and diffusion of technologies among other doctors.

State-of-the-art approaches on this field have proved once again the effectiveness of deep learning algorithms, as a matter of fact, Convolutional Neural Networks (CNNs) are currently the cornerstone of medical images analysis. Segmentation CNNs are able to extract features through a contracting path and exploit them to generate a segmentation mask across the expanding path; the size of feature maps decreases progressively in the former, whereas in the latter it increases back to the input resolution thanks to up-sampling operators and fractionally-strided convolutions, thus producing an encoder-decoder architecture [39].

Our aim is to exploit CNN to address both the problem of skin lesion image segmentation and classification, taking into account the trade-off between training a model from scratch and employing images of a very different nature through transfer learning.

The major use case objectives are reported in the following tables, referring to objective and performance metrics that will be used.

Objectives of the use case	
Type of problem	**Objective**
Segmentation and classification	Identify potentially dangerous skin lesions.
Explainability and visualization	Identify which medical visual patterns contribute to the final decision of the algorithms.
Confidence calibration	Ensuring that the output of the algorithms provide probability estimates representative of the true correctness likelihood.
Classification.	Discriminate whether to remove the skin lesion reflecting the opinion of medical experts. This choice should be independent from the final, post biopsy, diagnosis (benign or malignant).

Known KPIs	
KPI 1	Segmentation quality measured by a threshold Jaccard index metric [40]. Reference values obtained from the international ISIC competition [85].
KPI 2	Classification accuracy with respect to Ground Truth annotations. Reference values obtained from the international ISIC competition [85].
KPI 3	Qualitative evaluation of the informative value of the additional explanation on the outcomes. This can be obtained with questionnaires and intra medical personnel agreement.

2.12.2. Partner roles and contributions

This use case is under the responsibility of UNIMORE that will provide all expertise according to the following table. The use case will be deployed on PF1, PF3, PF7 (training/testing).

Roles and actions		
User	**Partner**	**Action**
End user (Medical personnel)	UNIMORE	Select images to be included in the database and verify ground truth correctness and evaluate the achieved results
Machine learning expert	UNIMORE	Model selection, training and testing method selection
Platform provider	Philips EVERIS CEA	Testing on PF1 Training and testing on PF7 Testing and training on PF3

2.12.3. Technical requirements

2.12.3.1. Dataset description

Type and format	Dermoscopic images, JPEG. Confocal images, JPEG.
Annotation metadata	Segmentation masks in binary PNG images; approximate age, general anatomic site, type of diagnosis, sex available.
Number of samples (retrospective/prospective)	In the order of 20k (retrospective).
Repository/sharing method	Public archive available online. Internal database which will be organized and made public within the ethical requirements.
Other information	Images size is between 640 x 480 and 6668 x 4339 pixels. Not every type of metadata is available for every image.

2.12.3.2. Pre-processing

The pre-processing consists of Random rotation, random horizontal flipping, random vertical flipping, random shifting, random shearing, random scaling, random color jitter, image normalization, connecter components.

2.12.3.3. Model design and testing

The convolutional neural network models applied to this use case are convolutional Neural Networks for classification (e.g. ResNet). Convolutional Neural Networks for segmentation (e.g. U-Net, SegNet, DeepLabv3+). Adam optimizer, SGD optimizer.

2.13. UC13 - Epileptic seizure detection

The main target of the use case is to perform detection of epileptic seizures on ECG and EEG signals, with the possibility to explore also seizure prediction. To this end, state-of-the-art machine learning based algorithms will be trained and will need to be constantly running the epilepsy detection algorithms over a very large amount of data.

2.13.1. Use case description and objectives

This use case will target epileptic seizure detection and classification from ECG and EEG signals, using a private database provided by CHUV. The preprocessing and classification algorithm will be initially the most relevant ones in the state-of-the-art that use deep learning techniques for population-wide analysis [50][51][52].

The use case objectives are reported in the following tables, referring to objective and performance metrics that will be used.

Objectives of the use case	
Type of problem	**Objective**
Classification	Detection of seizure type on EEG
Classification	Seizure detection

Known KPIs	
KPI 1	Time-to-model-in-production (**ttmip**).
KPI 2	Time-of-training models (**totm**).
KPI 3	Classification performance. State of the art for seizure detection [50]: Accuracy 88.67%, Specificity 90.00%, Sensitivity 95.00%.
KPI 4	Enhancement of the reliability of seizure identification.
KPI 5	Seizure prediction accuracy.
KPI 6	Reduction of the training and the evaluation times of the seizure predictions.

2.13.2. Partner roles and contributions

This use case is under the responsibility of CHUV that will provide all expertise according to the following table. The use case will be deployed on PF2.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	CHUV	Select records to be included in the dataset based on the diagnosis, provide labels.
End user (expert user)	CHUV	Export dataset using EDF format plus and annotations.
Machine learning expert	CHUV EPFL WINGS	Model selection, training and testing method selection.
Platform provider	WINGS	Deploy on PF2
End user (medical personnel)	CHUV	Evaluate the achieved accuracy.

2.13.3. Technical requirements

2.13.3.1. Dataset description

During the development of the platform, public Electroencephalogram (EEG) data (downloaded from the internet [https://physionet.org/pn6/chbmit/]) in edf/csv file formats will be used for testing mechanisms that will train the model, and the predictive capabilities of the proposed approach. As soon as the mechanisms become more stable, the mechanisms developed by WINGS will be set-up in the EPFL infrastructure so as to run the final experiments and refinements using CHUV data but still maintaining the data privacy as defined by the ethics committee. CHUV dataset is described in the table below.

Type and format	EDF plus.
Annotation metadata	Integrated in EDF plus.
Number of samples (retrospective/prospective)	2000 records from Repomse database. Retrospective.
Repository/sharing method	File care.

2.13.4. Pre-processing

For seizure detection, the pre-processing will be performed by anonymizing (export from database) and normalizing each EEG signal with Z-score normalization, zero mean and standard deviation of 1 before feeding into the 1-D deep convolutional network (CNN) for training and testing. The sampling rate of the EEG signal is set at 173.61 Hz.

For seizure prediction, the following pre-processing and feature extraction steps are foreseen before training the ML models:

- 1st approach: time series epoching/segmentation, filtering, artifact detection/rejection, averaging and re-referencing;
- 2nd approach: converting multiple channels of EEG signals into a single surrogate channel and filtering to increase SNR. Empirical Mode Decomposition (EMD) is also performed for further SNR enhancement.

The features from signals will be extracted in both time (statistical features) and frequency (spectral features) domains.

2.13.5. Model design and testing

For the seizure detection problem, on EPFL's side, a state-of-the-art CNN architecture will be adopted. The following table shows the details of one of the best-performing architectures in the bibliography [50], that we will take as a main reference.

Table 1
The details of CNN structure used in this research.

Layers	Type	Number of neurons (output layer)	Kernel size for each output feature map	Stride
0–1	Convolution	4092 × 4	6	1
1–2	Max-pooling	2046 × 4	2	2
2–3	Convolution	2042 × 4	5	1
3–4	Max-pooling	1021 × 4	2	2
4–5	Convolution	1018 × 10	4	1
5–6	Max-pooling	509 × 10	2	2
6–7	Convolution	506 × 10	4	1
7–8	Max-pooling	253 × 10	2	2
8–9	Convolution	250 × 15	4	1
9–10	Max-pooling	125 × 15	2	2
10–11	Fully-connected	50	–	–
11–12	Fully-connected	20	–	–
12–13	Fully-connected	3	–	–

After every convolutional layer, a leaky rectifier linear unit (LeakyRelu) is used as activation function. Finally, a Softmax function is used to compute the final probability for each output class. Overall, the functions required in this module are the conventional ones required by CNNs.

For seizures prediction, on WINGs side, Long Short-Term Memory (LSTM) networks will be employed.

2.14. UC14 - Prediction of the multiple sclerosis patients' outcome

The main target of the use case is to assess fatigue using an app and a wearable device and correlate it with the brain lesions of processed MRI images of the patients. To this end, MRI Images and questionnaires will be used for a set of 20 to 50 patients. MRI images will be processed to find the appropriate segmentation method and calculate the number and volume of MS lesions to enable the correlation with fatigue parameters.

2.14.1. Use case description and objectives

The use case can be divided into two main objectives: 1) the segmentation of MRI images using state-of-the-art deep learning techniques to extract the lesion and set an initial baseline for the patient, and 2) the association between the parameters obtained from different bio-signals captured by wearable devices with a fatigue index for the patient, with the ultimate aim of obtaining objective indicators of fatigue. A prospective dataset will be acquired for this use case with data from 20 to 50 patients. While the first objective is fully aligned with the DeepHealth project goals and will take advantage of the new

framework to improve the state-of-the-art techniques [58][59][61], the second objective has not yet been explored and there are no yet deep learning algorithms validated for this purpose. Therefore, it will not be considered in the KPIs used to evaluate the use case.

The major use case objectives are reported in the following tables, referring to objective and performance metrics that will be used.

Objectives of the use case	
Type of problem	**Objective**
Segmentation	Lesion extraction
Classification	Tissue separation (Gray, white matter, CSF)

Known KPIs	
KPI 1	Time-to-model-in-production (**ttmip**).
KPI 2	Time-of-training-models (**totm**).
KPI 3	Time-of-pre-processing-images (**toppi**).
KPI 4	Dice similarity coefficient (**DSC**) [58].
KPI 5	Lesion True Positive Rate (**LTPR**) [58].
KPI 6	Lesion False Positive Rate (**LFPR**) [58].

2.14.2. Partner roles and contributions

This use case is under the responsibility of CHUV that will provide all expertise according to the following table. The use case will be deployed on PF1.

Roles and actions		
User	**Partner**	**Action**
End user (medical personnel)	CHUV	Select images to be included in the dataset based on the diagnosis.
End user (expert user)	CHUV	Export dataset using DICOM format.
Machine learning expert	CHUV EPFL	Model selection, training and testing method selection.
End user (medical personnel)	CHUV	Evaluate the achieved accuracy.
Platform provider	Philips	Testing on PF1
End user (expert user)	CHUV	Label the images (selection of lesion ROI).

2.14.3. Technical requirements

2.14.3.1. Dataset description

Type and format	DICOM and NRRD [86].
Annotation metadata	Region of the manually segmented lesion.
Number of samples (retrospective/ prospective)	• Private CHUV dataset. 20-50 prospective MRI images from MS patients. • Public MICCAI 2008 MS lesion segmentation challenge dataset. 43 T1-weighted (T1w), T2-weighted (T2w), and FLAIR MRIs. • Public ISBI 2015 longitudinal MS lesion segmentation challenge dataset. 21 visit sets, each with T1w, T2w, proton density-weighted (PDw), and FLAIR MRIs.
Repository/sharing method	Private data will be stored at CHUV, and it will be physically copied after anonymization to the EPFL servers according to a specific Data Management Plan between both institutions. After this, an SFTP server over a VPN will be set up at EPFL to allow remote access from CHUV. For the public datasets, no specific actions are required.

2.14.3.2. Pre-processing

The pre-processing will be performed by applying the following steps:

1. Skull-strip of all images using the Brain Extraction Tool [60].
2. Intensity normalization and a 6 degree-of-freedom intra-subject registration using one of the 3 mm scans as the target image to align the different modalities.
3. Crop all the images to a 164 206 52 voxel sub-volume with the brain roughly centred.

2.14.3.3. Model design and testing

The first approach to be used is the one described in [58]. The network architecture is a 7-layer convolutional encoder network with shortcut. Pre-training is performed on the input images using a stack of convolutional **Restricted Boltzmann Machines** (RBM). The pre-trained weights and bias terms are used to initialize a convolutional encoder network, which is fine-tuned on pairs of input images (x), and segmentations (y). This is illustrated in the image below.

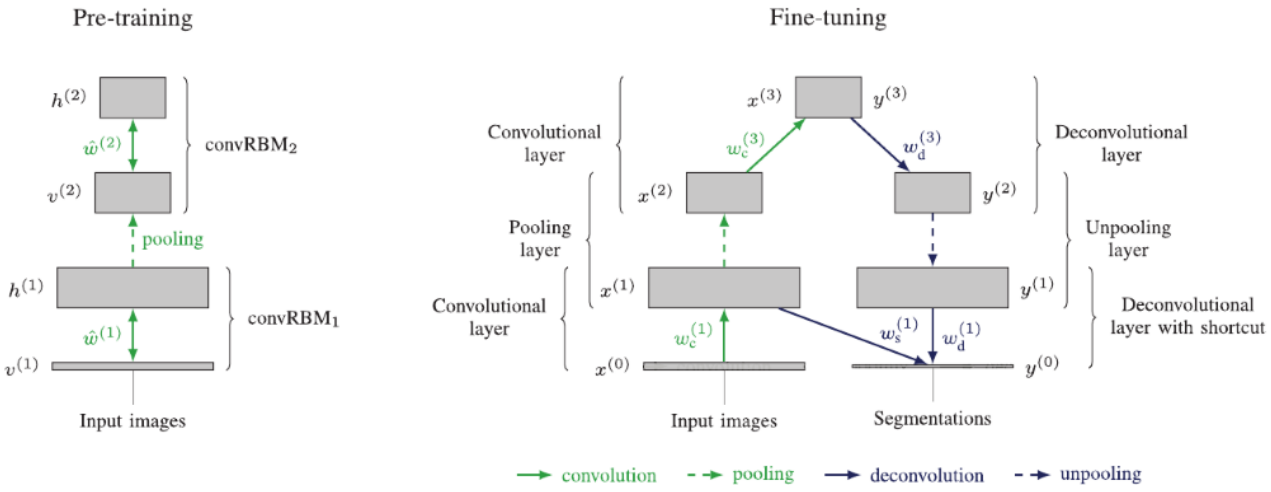

Figure 2 Network architecture

3. Data privacy requirements

For each use case and platform pair the requirements needed to **manage the data privacy.** The requirements needed for ensuring data privacy when cooperating between the use case provider and the application platform provider will be described in this section.

The full guidelines to handle data privacy within DeepHealth are to be specified in Deliverable D1.6 GDPR strategy and ethical guidelines. The solutions presented below will be compatible with the guidelines to be specified in D1.6.

3.1. Data privacy in UC1 Migraine & seizures prediction on PF2 MigraineNet

3.1.1. Migraine Predictions

Both the data to be used in UC1 and the MigraineNet platform are maintained and handled by WINGS. In particular, MigraineNet platform's enhancements in the context of migraine predictions will be achieved using historical data which have already been collected in a past pilot conducted from WINGS (prior to GDPR). These data were both collected and stored in WINGS infrastructure and thus data privacy is maintained as it was when WINGS initially collected them. Any further actions needed for maintaining compliance with the GDPR rules will be handled from WINGS.

3.1.2. Seizure Predictions

Depending on the outcome of the seizure aspects to be researched, WINGS may turn to doctors with expertise in seizures and ask for their guidance with respect to questions that would allow the prediction of a seizure. If this is applicable, a pilot round with real users providing their data through questionnaires and wearable devices will be organized. During this process, informed consent and any other process related to GDPR will be followed from WINGS members participating in DeepHealth.

3.2. Data privacy in UC2 UNITOPath on PF5 Open-DeepHealth

The dataset contains only RGB images, whose privacy is granted by anonymizing them before moving the data from the medical premises through the UNITO internal network or temporary copies on external drives.

3.3. Data privacy in UC3 UNITOBrain on PF5 Open-DeepHealth

The dataset contains only DICOM images, whose privacy is granted by anonymizing them before moving the data from the medical premises through the UNITO internal network temporary copies on external drives.

3.4. Data privacy in UC4 Chest dataset and use case on PF5 Open-DeepHealth

The dataset contains only CT scan, whose privacy is granted by anonymizing them before moving the data from the medical premises to the ODH storage space.

3.5. Data privacy in UC5 Deep Image Annotation

This use cases uses publicly available data, anonymised for research purposes.

3.6. Data privacy in UC6 Promort on PF6 Digital Pathology

UC6 on PF6 will only use pseudonimized images. The application needs to be able to operate in a private hospital network with only tightly controlled external communication channels from KI to CRS4 to support transmission of anonymized model feedback data, including anonymized image

annotations and segments. Data transferred between KI and CRS4 must be encrypted during transit and server authenticity must be verified prior to establishing connectivity.

The web platform used by pathologists to perform reviews on automatically annotated slides will provide a further level of anonymization by applying an automatically generated UUID4 label for each slide in order to mask pseudonimized codes generated during the slides acquisition process.

3.7. Data privacy in UC7 Major Depression

3.7.1. On PF1 Open Innovation Platform

Deployment of the Open Innovation platform is done inside the OVGU premise. No connection is being made to PHILIPS or any third party.

3.7.2. On PF2 MigraineNet

In the context of UC7's use/implementations on MigraineNet platform, digital questionnaires, diaries and/or wearable devices will be used for the data collection. The information will be collected through pilot tests during the DeepHealth project. During these pilots, the users will be asked for their informed consent for their data to be collected, stored and processed in WINGS infrastructure and for their data to be accessed and processed by both the WINGS and OVGU teams. Therefore, no data privacy aspects are expected to arise.

3.7.3. On PF4 PIAF

An NDA will be signed for data sharing and Each Party expressly agrees that any exploitation or commercialization of the health data that may be shared between the Parties in the frame of the discussions covered by this non-disclosure agreement is strictly forbidden.

3.8. Data privacy in UC8 Dementia

3.8.1. On PF1 Open Innovation Platform

Deployment of the Open Innovation platform is done inside the OVGU premise. No connection is being made to PHILIPS or any third party.

3.8.2. On PF4 PIAF

An NDA will be signed for data sharing and Each Party expressly agrees that any exploitation or commercialization of the health data that may be shared between the Parties in the frame of the discussions covered by this non-disclosure agreement is strictly forbidden.

3.9. Data privacy in UC9 Lumbar Spine Pathology

3.9.1. On PF1 Open innovation Platform

Deployment of the Open Innovation platform is done inside the FISABIO premise. No connection is being made to PHILIPS or any third party.

3.9.2. On PF3 ExpressIF

Anonymized datasets will be uploaded to one of CEA's secured server during the training and test phases according to FISABIO directives. CEA's secured servers are not accessible from the internet.

3.9.3. On PF5 Open-DeepHealth

Anonymized datasets sets will be copied on the PF5 infrastructure at UNITO. Access to datasets will be granted according with FISABIO directives (UC9 data owner). Data will be protected according with best practices in the field.

3.9.4. On PF7 Lumen

Anonymized datasets will be uploaded to Lumen storage secured server during the training and test phases according to FISABIO directives.

3.10. Data privacy in UC10 Alzheimer Disease

3.10.1. On PF1 Open Innovation Platform

Deployment of the Open Innovation platform is done inside the FISABIO premise. No connection is being made to PHILIPS or any third party.

3.10.2. On PF5 Open-DeepHealth

Anonymized datasets sets will be copied on the PF5 infrastructure at UNITO. Access to datasets will be granted according with FISABIO directives (UC9 data owner). Data will be protected according with best practices in the field.

3.10.3. On PF7 Lumen

Anonymized datasets will be uploaded to Lumen storage secured server during the training and test phases according to FISABIO directives.

3.11. Data privacy in UC11 Urology on PF1 Open innovation Platform

Deployment of the Open Innovation platform is done inside the SIVECO premise. No connection is being made to PHILIPS or any third party.

3.12. Data privacy in UC12 Skin Cancer Melanoma Detection on PF1 Open Innovation Platform

The use case use publicly available data.

3.13. Data privacy in UC13 Epilieptic Seizure Detection on PF2 MigraineNET

The data to be used during the development phase of the MigraineNet functionality so as to also serve seizure related issues are publicly available EEG datasets. Data from open source databases are protected in advance and therefore it is not required to perform any further data privacy procedures. Following the development phase, during the refinement of the mechanisms, the MigraineNet platform will use CHUV dataset. Due to the restriction of the data to be maintained either in the CHUV or in the EPFL premises, during this period, MigraineNet mechanisms will be set-up within the EPFL infrastructure (upon the necessary NDAs).

3.14. Data privacy in UC14 Multiple Sclerosis on PF1 Open Innovation Platform

Deployment of the Open Innovation platform is done inside the EPFL/ CHUV premise. No connection is being made to PHILIPS or any third party.

4. Conclusions

In this deliverable we provide a high-level view of all use cases foreseen in the DeepHealth project in order to summarize the main contributions expected by all the experimental efforts and multidisciplinary interaction among the involved partners. The document addresses many aspects including the goals, roles, used dataset and expected machine learning approach per each use case. Moreover, the major features of all platforms used to support the use cases are described (in the confidential version of this deliverable) along with plans to guarantee privacy when medical data needs to be shared or moved.

The heterogeneity and complexity of the challenges undertaken within the project is summarized in the following table. Here we briefly recall all use cases goals, the medical data that we leverage on and the integration efforts of DeepHealth platforms.

	Goal	Data type	Training	Testing
UC1	Classification and prediction	Wearable bio-signals, plus questionnaire	PF2	PF2
UC2	Image Classification	High resolution RGB histopathological images	PF5	PF5
UC3	Image inference	CT brain scans	PF5	PF5
UC4	Image segmentation	Thoracic CT scans	PF5	PF5
UC5	Text generation	X-ray images with diagnosis report	PF5	PF1, PF5
UC6	Image segmentation and classification	High resolution RGB histopathological images	PF6	PF6
UC7	Image segmentation and classification	MRI images of brain	PF2	PF1, PF2, PF4
UC8	Image segmentation and classification	MRI images of brain, plus biomarkers		PF1, PF4
UC9	Image segmentation and classification	MRI images of spine	PF3, PF5	PF1, PF3, PF5, PF7
UC10	Image segmentation and classification	MRI images of brain	PF5	PF1, PF5, PF7
UC11	Image classification	CT images of abdomen		PF1
UC12	Image classification	Dermoscopic RGB images	PF3, PF7	PF1, PF7
UC13	Detection and classification	EEG signal	PF2	PF2
UC14	Image segmentation and prediction	MRI images of brain, plus wearable bio-signals		PF1

5. References

[1] European Commission. Redesigning health in Europe for 2020. 2012.
[2] European Commission, Communication Artificial Intelligence for Europe, 2018.
[3] Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning. MIT Press 2017.
[4] Z. Zhang, Y. Xie, F. Xing, M. McGough and L. Yang, "MDNet: A Semantically and Visually Interpretable Medical Image Diagnosis Network," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 3549-3557.
[5] Vivienne Sze, Yu-Hsin Chen, Tien-Ju Yang and Joel S. Emer. Efficient Processing of Deep Neural Networks: A Tutorial and Survey. Proc. IEEE, vol. 105, no. 12, Dec. 2017. DOI: 10.1109/JPROC.2017.2761740
[6] TensorFlow, https://www.tensorflow.org/
[7] Pytorch, http://pytorch.org/
[8] World Health Organization, Headache disorders Fact sheet, April 2016, available at http://www.who.int/mediacentre/factsheets/fs277/en/, last access 2018-04-05
[9] Medical News Today, “Headaches: Causes, Diagnosis and Treatments”, November 2017, available at http://www.medicalnewstoday.com/articles/73936.php, last access 2018-04-05
[10] World Health Organization, Epilepsy, February 2019, available at https://www.who.int/news-room/fact-sheets/detail/epilepsy, last access 2019-03-22.
[11] Epilepsy Action Australia, Epilepsy: The facts, available at https://www.epilepsy.org.au/about-epilepsy/, last access 2019-03-22.
[12] Carney, P. R., Myers, S., & Geyer, J. D. (2011). Seizure prediction: methods. *Epilepsy & behavior*, *22*, S94-S101.
[13] Ahammad, N., Fathima, T., & Joseph, P. (2014). Detection of epileptic seizure event and onset using EEG. *BioMed research international*, *2014*. Smith, S. J. M. (2005). EEG in the diagnosis, classification, and management of patients with epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(suppl 2), ii2-ii7.
[14] Moghim, N., & Corne, D. W. (2014). Predicting epileptic seizures in advance. *PloS one*, *9*(6), e99334.
[15] Van de Vel, A., Cuppens, K., Bonroy, B., Milosevic, M., Jansen, K., Van Huffel, S., ... & Ceulemans, B. (2013). Non-EEG seizure-detection systems and potential SUDEP prevention: state of the art. *Seizure*, *22*(5), 345-355.
[16] Rogawski, M. A. (2012). Migraine and epilepsy—shared mechanisms within the family of episodic disorders. In *Jasper's Basic Mechanisms of the Epilepsies [Internet]. 4th edition*. National Center for Biotechnology Information (US).
[17] Bruno Korbar, Andrea M Olofson, et. al. “Deep learning for classification of colorectal polyps on whole-slide images”, J Pathol Inform, 2017
[18] Bychkov, Dmitrii, et al. "Deep learning based tissue analysis predicts outcome in colorectal cancer." Scientific reports, 2018
[19] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition.", in proc. ICLR 2015.
[20] He, Kaiming, et al. "Deep Residual Learning for Image Recognition." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016
[21] Konstas AA, Goldmakher GV, Lee TY, Lev MH. Theoretic basis and technical implementations of CT perfusion in acute ischemic stroke, part 2: technical implementations. AJNR Am J Neuroradiol. 2009 May;30(5):885-92.
[22] Konstas AA, Goldmakher GV, Lee TY, Lev MH. Theoretic basis and technical implementations of CT perfusion in acute ischemic stroke, part 1: Theoretic basis. AJNR Am J Neuroradiol. 2009 Apr;30(4):662-8.
[23] Campbell BC, Christensen S, Levi CR, Desmond PM, Donnan GA, Davis SM, Parsons MW. Cerebral blood flow is the optimal CT perfusion parameter for assessing infarct core. Stroke. 2011 Dec;42(12):3435-40.
[24] MacMahon H et al.: Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017, Radiology. 2017 Jul;284(1):228-243.

[25] Nam JG et al: Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. Radiology. 2019 Jan;290(1):218-228.

[26] Raghu VK, Zhao W, Pu J, et al Feasibility of lung cancer prediction from low-dose CT scan and smoking factors using causal models Thorax Published Online First: 12 March 2019. doi: 10.1136/thoraxjnl-2018-21263

[27] Demner-Fushman, D., Kohli, M. D., Rosenman, M. B., Shooshan, S. E., Rodriguez, L., Antani, S., & McDonald, C. J. (2015). Preparing a collection of radiology examinations for distribution and retrieval. Journal of the American Medical Informatics Association, 23(2), 304-310.

[28] Jing, B., Xie, P., & Xing, E. (2018). On the automatic generation of medical imaging reports. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Long Papers), pages 2577–2586 Association for Computational Linguistics. arXiv preprint arXiv:1711.08195.

[29] McDonald, R. J., Schwartz, K. M., Eckel, L. J., Diehn, F. E., Hunt, C. H., Bartholmai, B. J., ... & Kallmes, D. F. (2015). The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. Academic radiology, 22(9), 1191-1198.

[30] Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). BLEU: a method for automatic evaluation of machine translation. In Proceedings of the 40th annual meeting on association for computational linguistics (pp. 311-318). Association for Computational Linguistics.

[31] Shin, H. C., Roberts, K., Lu, L., Demner-Fushman, D., Yao, J., & Summers, R. M. (2016). Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2497-2506).

[32] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In Proceedings of International Conference on Learning Representations.

[33] Tsou, A. Y., Lehmann, C. U., Michel, J., Solomon, R., Possanza, L., & Gandhi, T. (2017). Safe practices for copy and paste in the EHR. Applied clinical informatics, 26(01), 12-34.

[34] Ehteshami Bejnordi, Babak, Mitko Veta, Paul Johannes van Diest, Bram van Ginneken, Nico Karssemeijer, Geert Litjens, Jeroen A. W. M. van der Laak, et al. “Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer.” JAMA 318, no. 22 (December 12, 2017): 2199. https://doi.org/10.1001/jama.2017.14585.

[35] Wang, Dayong, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H. Beck. “Deep Learning for Identifying Metastatic Breast Cancer.” ArXiv:1606.05718 [Cs, q-Bio], June 18, 2016. http://arxiv.org/abs/1606.05718.

[36] Nachbar, F., Stolz, W., Merkle, T., Cognetta, A. B., Vogt, T., Landthaler, M., ... & Plewig, G. (1994). The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, 30(4), 551-559.

[37] Binder, M., Schwarz, M., Winkler, A., Steiner, A., Kaider, A., Wolff, K., & Pehamberger, H. (1995). Epiluminescence microscopy: a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. *Archives of dermatology*, 131(3), 286-291.

[38] Argenziano, G., Soyer, H. P., Chimenti, S., Talamini, R., Corona, R., Sera, F., ... & Hofmann-Wellenhof, R. (2003). Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet. Journal of the American Academy of Dermatology, 48(5), 679-693.

[39] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham

[40] Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., ... & Halpern, A. (2018, April). Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 168-172). IEEE.

[41] Ljungberg B, et al. Guidelines on Renal Cell Carcinoma. European Association of Urology Guidelines, 2019 at www.uroweb.org

[42] American Cancer Society, Cancer Facts & Figures 2016
[43] Ferlay, J., et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer, 2013. 49: 1374
[44] King, S.C., et al. Continued increase in incidence of renal cell carcinoma, especially in young patients and high grade disease: United States 2001 to 2010. J Urol, 2014. 191: 1665
[45] Kato, M., et al. Natural history of small renal cell carcinoma: evaluation of growth rate, histological grade, cell proliferation and apoptosis. J Urol, 2004. 172: 863
[46] Pedrosa, I., et al. MR imaging of renal masses: correlation with findings at surgery and pathologic analysis. Radiographics, 2008. 28: 985
[47] Israel, G.M., et al. How I do it: evaluating renal masses. Radiology, 2005. 236: 441.
[48] Israel, G.M., et al. Pitfalls in renal mass evaluation and how to avoid them. Radiographics, 2008. 28: 1325.
[49] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention 2015 Oct 5 (pp. 234-241). Springer, Cham
[50] Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adeli, H. (2018). Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Computers in Biology and Medicine, 100, 270–278. https://doi.org/10.1016/J.COMPBIOMED.2017.09.017
[51] Ullah, I., Hussain, M., Qazi, E., Aboalsamh, H. (2018). An Automated system for epilepsy detection using EEG brain signals based on Deep Learning Approach. https://arxiv.org/abs/1801.05412
[52] Yuan, Y., Xun, G., Jia, K., Zhang, A. (2019). A Multi-View Deep Learning Framework for EEG Seizure Detection. IEEE Journal of Biomedical and Health Informatics, 23 (1), 83-94
[53] Dranovsky, Alex, and E. David Leonardo. "Is there a role for young hippocampal neurons in adaptation to stress?." *Behavioural brain research* 227.2 (2012): 371-375.
[54] Duman, R. S. "Pathophysiology of depression: the concept of synaptic plasticity." *European psychiatry* 17 (2002): 306-310.
[55] Malykhin, N. V., and N. J. Coupland. "Hippocampal neuroplasticity in major depressive disorder." *Neuroscience* 309 (2015): 200-213.
[56] Samuels, Andrew. *The political psyche*. Routledge, 2015.
[57] Wise, T., et al. "Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: evidence from voxel-based meta-analysis." *Molecular psychiatry* 22.10 (2017): 1455.
[58] Brosch, T., Tang, L. Y. W., Yoo, Y., Li, D. K. B., Traboulsee, A., & Tam, R. (2016). Deep 3D Convolutional Encoder Networks With Shortcuts for Multiscale Feature Integration Applied to Multiple Sclerosis Lesion Segmentation. IEEE Transactions on Medical Imaging, 35(5), 1229–1239. https://doi.org/10.1109/TMI.2016.2528821
[59] Valverde, S., Cabezas, M., Roura, E., González-Villà, S., Pareto, D., Vilanova, J. C., Lladó, X. (2017). Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. NeuroImage, 155, 159–168. https://doi.org/10.1016/J.NEUROIMAGE.2017.04.034
[60] Jenkinson, Mark, Mickael Pechaud, and Stephen Smith. "BET2: MR-based estimation of brain, skull and scalp surfaces." Eleventh annual meeting of the organization for human brain mapping. Vol. 17. 2005.
[61] Roy, S., Butman, J. A., Reich, D. S., Calabresi, P. A., & Pham, D. L. (2018). Multiple Sclerosis Lesion Segmentation from Brain MRI via Fully Convolutional Neural Networks. ArXiv preprint http://arxiv.org/abs/1803.09172
[62] https://challenge2018.isic-archive.com/
[63] Zeiler, Mattew D., (2012). Adadelta: An adaptive learning rate method. ArXiv preprint http://arxiv.org/abs/1212.5701
[64] Kingma, Diederik P. and Ba, Jimmy (2014). Adam: A Method for Stochastic Optimization. ArXiv preprint http://arxiv.org/abs/1412.6980

[65] C. Duchi, John and Hazan, Elad and Singer, Yoram (2011). Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. Journal of Machine Learning Research, 12, 2121-2159
[66] M. Jenkinson, M. Pechaud, and S. Smith, “BET2: MR-based estimation of brain, skull and scalp surfaces,” in 11th Annu. Meet. Org. Human Brain Map., Toronto, ON, Canada, 2005, vol. 17.
[67] http://dicom.nema.org/medical/dicom/current/output/html/part04.html#PS3.4
[68] https://en.wikipedia.org/wiki/Picture_archiving_and_communication_system
[69] L. Richardson, M. Amundsen, S. Rudy, “RESTful Web APIs: Services for a Changing World”, ISBN: 978-1449358068
[70] L. Zadeh, Fuzzy sets, Information and Control, vol. 8, n° 13, pp. 338-353, 1965
[71] D. Gunning, Explainable Artificial Intelligence (XAI), 2016.
[72] A. Grivet Sébert and J.-P. Poli, Material Classification from Imprecise Chemical Composition : Probabilistic vs Possibilistic Approach, FuzzIEEE, 2018
[73] R. Pierrard, J.-P. Poli and C. Hudelot, Learning Fuzzy Relations and Properties for Explainable Artificial Intelligence, FuzzIEEE, 2018
[74] J.-P. Poli et L. Boudet, A Modular Fuzzy Expert System Architecture For Data and Event Streams Processing, IPMU, 2016
[75] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Healthcare_resource_statistics_-_technical_resources_and_medical_technology
[76] https://openi.nlm.nih.gov/
[77] http://peir.path.uab.edu/library/
[78] https://openi.nlm.nih.gov/gridquery?q=&it=xg&coll=cxr
[79] http://peir.path.uab.edu/library/index.php?/category/111
[80] https://surfer.nmr.mgh.harvard.edu/
[81] https://www.who.int/features/factfiles/dementia/en/
[82] https://github.com/BIMCV-CSUSP/MIDS/blob/master/README_proposal.md
[83] https://github.com/BIMCV-CSUSP/MIDS/blob/master/README_proposal.md
[84] https://docs.google.com/document/d/1chZv7vAPEebPDxMktfI9i1OkLNR2FELIfpVYsaZPr4/edit#
[85] https://challenge2018.isic-archive.com/
[86] http://teem.sourceforge.net/nrrd/format.html
[87] https://wings-ict-solutions.eu/solutions/health-issues-prediction-migrainenet
[88] https://www.openmicroscopy.org/omero/