1. **Description of the assignment of the project and the focus of your work**

Introduction to the problem

*In digital pathology a diagnosis is carried out by analysing histopathological samples which are pieces of human tissues extracted via surgical operation. Specimens are typically stained with H&E (haematoxylin and eosin) so that different structures come in different shades between blue (haematoxylin binds to cell nuclei as they are negatively charged) and pink (eosin binds to extracellular matrix and cytoplasm as they are positively charged) for better distinguish between them [2].*

*The traditional approach consists in the visual examination of samples with the aim of detecting abnormalities and carrying out a certain diagnosis i.e. if a tissue is cancerous or not. Visual examination is time consuming, prone to inter-reader and intra-reader variability and non-reproducible: making available to the pathologists a tool that supports them during the evaluation can help overcoming these issues [3].*

# *Nowadays digital slides [4] are obtained by scanning specimens placed on conventional glass slides: such multi-resolution slides are called WSI (Whole Slide Images) and can be elaborated numerically with many related advantages enabling, among all, telepathology that is a key requirement for second opinions on cases and remote consults. [1]*

*The focus of the project and clinical insights*

*The focus of the project is developing a software able to produce an attention-map for cancer detection that drives pathologists’ attention to certain areas of the slice that might be pathological.*

*(image)*

*Uncertainty is a crucial aspect of such application and it is modelled with deep neural networks, more specifically with a Bayesian CNNs trained with WSI images representing AD tissue( short for adenoma, a benign tumour of epithelial tissue with glandular origin or characteristics), AC tissue (short for adenocarcinoma, a malign tumor) and healthy tissue.*

*(images)*

1. **Theory about CNNs focused on the application outline, providing references**

*CNNs*

*Neural networks (NNs) are a machine learning approach that do not consist in one single computational unit but that relies on several units, called neurons, differently interconnected via weights: NNs take inspiration from the way the brain is organised. The knowledge of the network is preserved in the weights that stand for biological synapses and the behaviour is related to its architecture: the learning process consists in adjusting the weights, extracting a mathematical model that fits training data.*

*Among all the networks we have CNNs (short for Convolutional Neural Networks) that emulate the visual cortex especially as regards the local connectivity between multiple layers and the hierarchical representation of the input: they consist in a input later, a variable number of hidden layers and a output layer. Each hidden layer is dedicated to identifying a multiple feature of the input: low-level features are condensed in the deepest layers while problem-specific features belong to last layers, features that are user-independent. Such characteristics allow the network to be more flexible when extracting different combination of small patterns when talking about image recognition tasks, eventually combining them for the aim of the network.*

*(image)*

*Classification tasks*

*The aim of CNNs is give a output that is a class or a prediction of the input and this is accomplished via several different layers that can be differently organized: the 3 characteristic layers are convolutional layers, non-linear layers and pooling layers. As hidden units are connected to local receptive fields and share weights, the input can have a high dimension without resulting in a large number of parameters, these parameters are learned during the training via backpropagation.*

*The backpropagation algorithm consists in the update of the weights, initially random initialised, basing on a error term that is computed with the output given by the network and the desired output.*

*(details about backprop…? Bullet points)*

*The core of CNNs is the convolutional layers that is composed by a set of filters: each filter activates when a specific feature is detected in the input.*

*(details about the layers…?)*

*One of the challenges of image analysis such as automatic histological assessment of cancer is not only the variability of the morphology of the tissue related to the pathology, but also the variability between patients that has always made tedious to find handcrafted features that could make the system accurate and reliable.*

*Bayesian CNNs*

*Typical drawbacks of conventional CNNs are, firstly, the risk of overfitting when the network is not trained on very large datasets and, secondly, the absence of a measure of uncertainty related to the prediction: Bayesian CNNs where thought to overcome these problems.*

*…*

1. **Detailed description of the method you used (describe most salient aspects of your code too)**

* Pipeline schematica
* Motivazione scelta dell’architettura della rete
* Processing delle immagini
* validazione

1. Detailed description of the results (provide graphs, tables, etc.);
2. Results discussion;
3. Future development

Possibilità di rendere disponibile il tool a diversi patologi per concordare sulla diagnosi in base alla mappa proposta, permettendo di annotare direttamente l'immagine prima di inviarla per un consulto 🡪 tele-istopatologia

**References**:

* **[1]** Image analysis and machine learning in digital pathology: Challenges and opportunities <https://www.sciencedirect.com/science/article/abs/pii/S1361841516301141>
* **[2]** Chan JK (2014). [*"The wonderful colors of the hematoxylin-eosin stain in diagnostic surgical pathology"*](https://www.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&tool=sumsearch.org/cite&retmode=ref&cmd=prlinks&id=24406626). Int J Surg Pathol. **22** (1): 12–32. [*doi*](https://en.wikipedia.org/wiki/Digital_object_identifier):[*10.1177/1066896913517939*](https://doi.org/10.1177%2F1066896913517939). [*PMID*](https://en.wikipedia.org/wiki/PubMed_Identifier) [*24406626*](https://www.ncbi.nlm.nih.gov/pubmed/24406626)

# [3] Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous Digital Pathology Images <https://www.sciencedirect.com/science/article/pii/S2352396417305078>

# [4] <https://www.sciencedirect.com/topics/medicine-and-dentistry/digital-pathology>

# [5] Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology [Harshita Sharmaa](https://www.sciencedirect.com/science/article/abs/pii/S0895611117300502#!) <https://www.sciencedirect.com/science/article/abs/pii/S0895611117300502>

# Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases [Andrew Janowczyk](https://www.ncbi.nlm.nih.gov/pubmed/?term=Janowczyk%20A%5BAuthor%5D&cauthor=true&cauthor_uid=27563488)1 and [Anant Madabhushi](https://www.ncbi.nlm.nih.gov/pubmed/?term=Madabhushi%20A%5BAuthor%5D&cauthor=true&cauthor_uid=27563488)

# Receptive fields, binocular interaction and functional architecture in the cat's visual cortex [D. H. Hubel](https://www.ncbi.nlm.nih.gov/pubmed/?term=Hubel%20DH%5BAuthor%5D&cauthor=true&cauthor_uid=14449617) and [T. N. Wiesel](https://www.ncbi.nlm.nih.gov/pubmed/?term=Wiesel%20TN%5BAuthor%5D&cauthor=true&cauthor_uid=14449617)

# Bayesian Neural Network Series Post 1: Need for Bayesian Neural Networks <https://medium.com/neuralspace/bayesian-neural-network-series-post-1-need-for-bayesian-networks-e209e66b70b2>

# CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more… <https://medium.com/analytics-vidhya/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5>

* Deep Bayesian Active Learning with Image Data <https://arxiv.org/pdf/1703.02910.pdf>

# Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology <https://www.sciencedirect.com/science/article/abs/pii/S0895611117300502>