

Stage M2

Remy Dubois

April 2017 - Aug 2017

Contents

List of Figures

1 Biological and Mathematical context

1.1 Biological context

1.1.1 Gene Expression

Gene expression is the process by which the DNA material is transformed into useful products for the cell. This process follows a well known and established process along which DNA components (pairs of nucleotides) are successively read, torn apart in smaller groups, then transformed group by group one at a time. Although well mastered for decades, this phenomena still raises important question than divide the scientific community. Among them is the selection process, which decides whether a given portion of the DNA will - or will not - go through this expression process and actually deliver its information.

Gene Expression process The Gene expression process can be decomposed into three substeps: DNA Transcription - RNA splicing - mRNA Translation. See fig. ?? for illustration.

1. DNA Transcription: This first step focuses on turning the DNA code into another nucleic acid called RNA. All along the cell's life, DNA remains cloistered into the nuclei, and RNA will actually move the information out of the cell nuclei, where it is actually needed by the cell (along the cell membrane for instance, for the production of reparative proteins).
2. RNA splicing: Once the RNA molecule has been created, it is then stripped and split by group of nucleotide pairs, each group representing an individual gene with its proper function. Each group can then be seen as a container of one gene's information and are thereafter called mRNA (m stands for *messenger*).
3. mRNA Translation: This third step take as input the mRNA molecule described above and turns it into the proper mean of action of genes, which mainly takes to form of a protein. Once produced in the nuclei, each mRNA molecule moves out of the nuclei and will be read by the ??, in order to produce the intended protein.

Figure 1: Global overview of the whole gene expression process

Expression Level

RNA Localization

1.1.2 Expression patterns

Localization patterns

1.1.3 Place of this project

1.2 Mathematical Context

1.2.1 Image analysis

Segmentation

Spot detection

1.2.2 Computer vision

Recent history: a convolutional-based approach

A very quick introduction to filtering and pattern recognition

Pros and cons

Applications to biomedical case study

1.2.3 Learning and Domain Adaptation

Learning Phase

Domain Adaptation

2 Natural Data observation

2.1 Spot detection

2.1.1 Gaussian Kernels

2.2 Segmentation

2.2.1 Thresholding segmentation

2.2.2 Single threshold-based methods

GMM

KMeans (Otsu)

2.2.3 Multi threshold-based methods

GMM

Otsu

2.2.4 Graph based segmentation

Oversegmentation

Region adjacency graphs

Hierarchical graph merge

3 Generative Model

3.1 Feature engineering

3.2 Image generation process

3.2.1 Morphological structure

4 Transfer Learning strategies

4.1 Model choice

4.1.1 Inception

Architecture

Strategy

4.1.2 Squeezenet

Architecture

Strategy

4.1.3 Comparison

Versatility

Scalability (deployment)

Accuracy

Final word

4.2 Off-the-shelf training