

# Analysis of spatial information - multiplexed immunofluorescence (mIF) data

Bartosz Skowronek

## 1 Introduction

Using mIF data retrieved from biopsies on multiple patients, such as location coordinates of cells and their phenotypes, my task was to provide an overview of the data and attempt to detect TLS (tertiary lymphoid structure) candidates using neighborhood graphs constructed using the provided data. The data was divided into three panels, IF1, IF2 and IF3, each file in a panel for a different patient (“0197\_IF1” etc).

## 2 Overview

I made overview plots for each of the panels using the plotly package in Python, employing the color map provided in “coloring\_celltypes\_code\_snippet.py” and using similar color dictionaries for the other two panels made from scratch. Double-clicking a file name in the legend allows the viewer to browse a single patient’s biopsy data, and mousing over a cell provides additional information. The plot can be zoomed in by drawing a square over the data to highlight an interesting fragment:

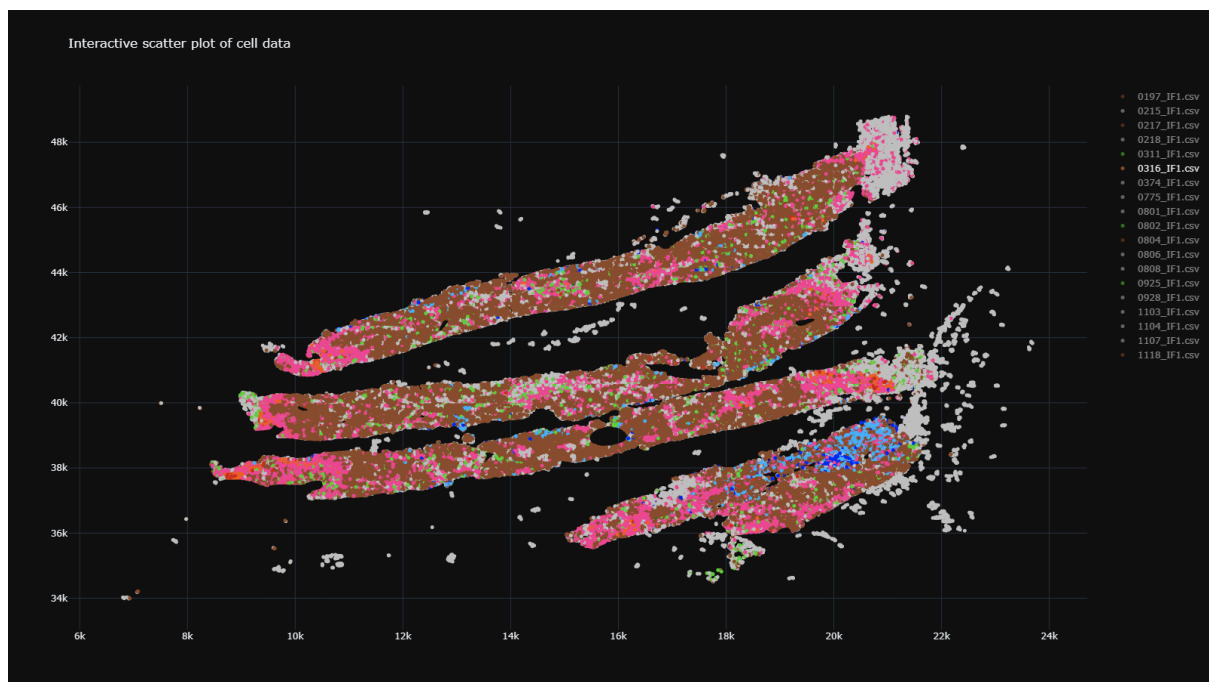


fig.1 overview of data for patient 0316 from the IF1 panel

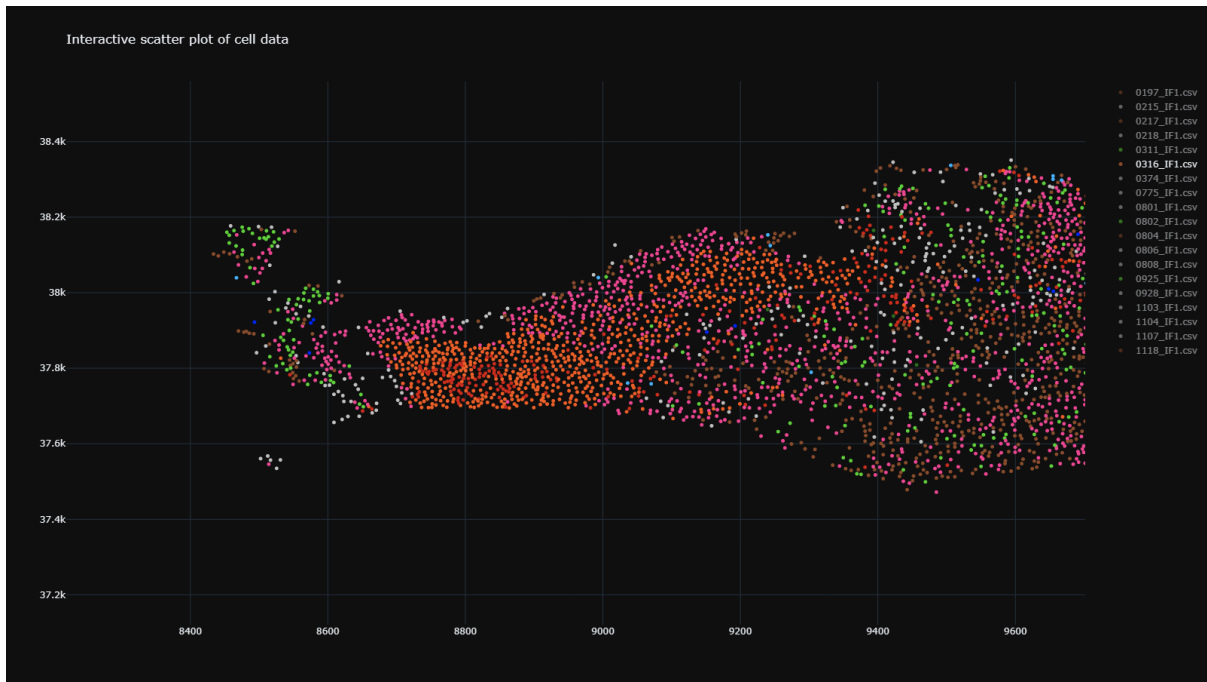
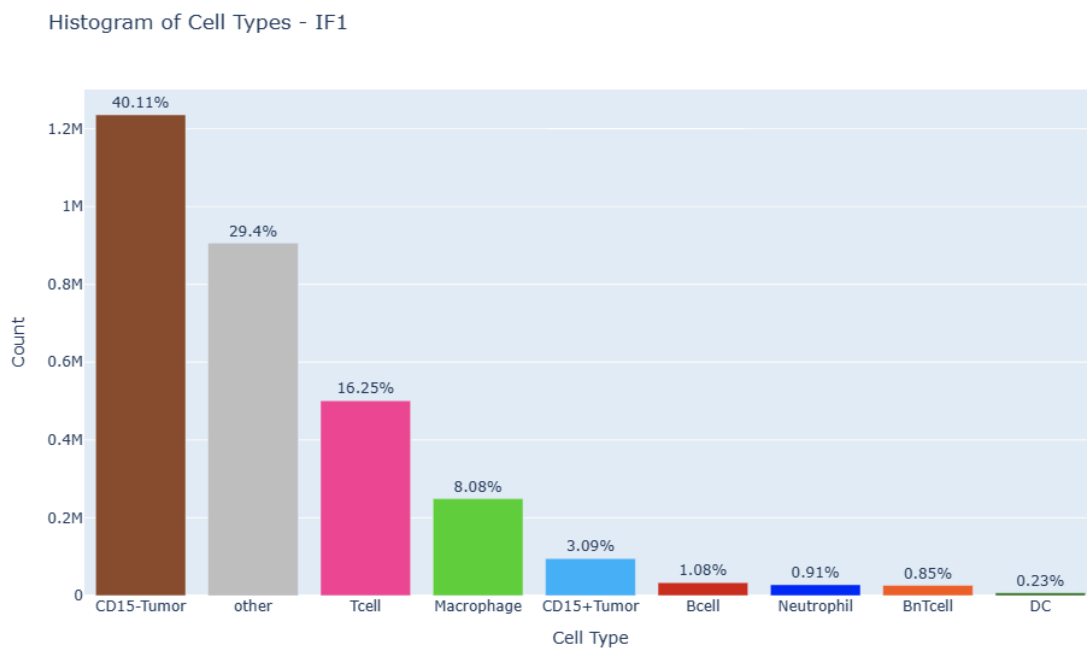
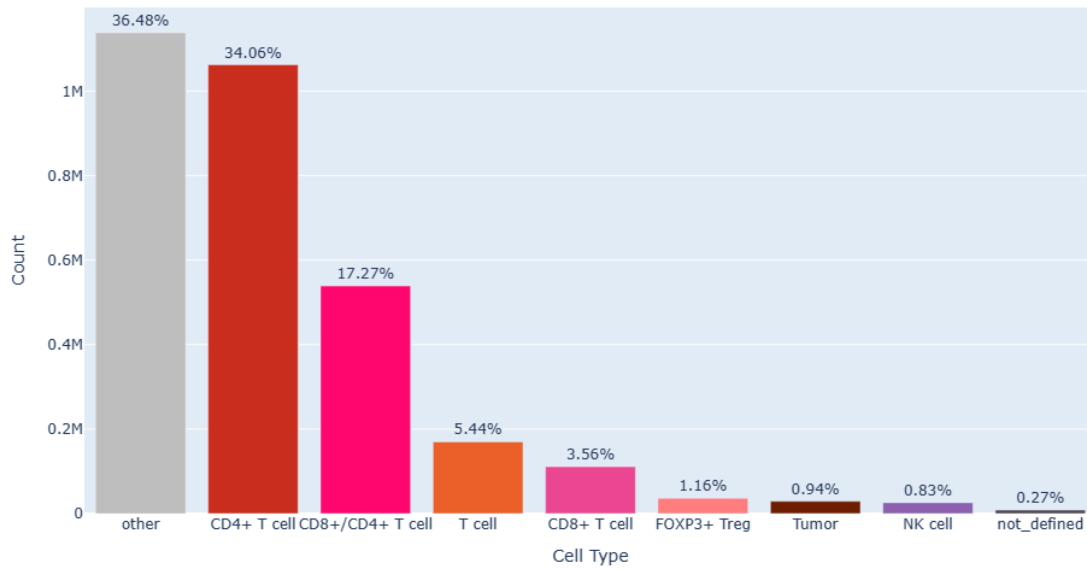


fig.2 zoomed in data; orange, red and pink cells are respectively B, BnT and T cells

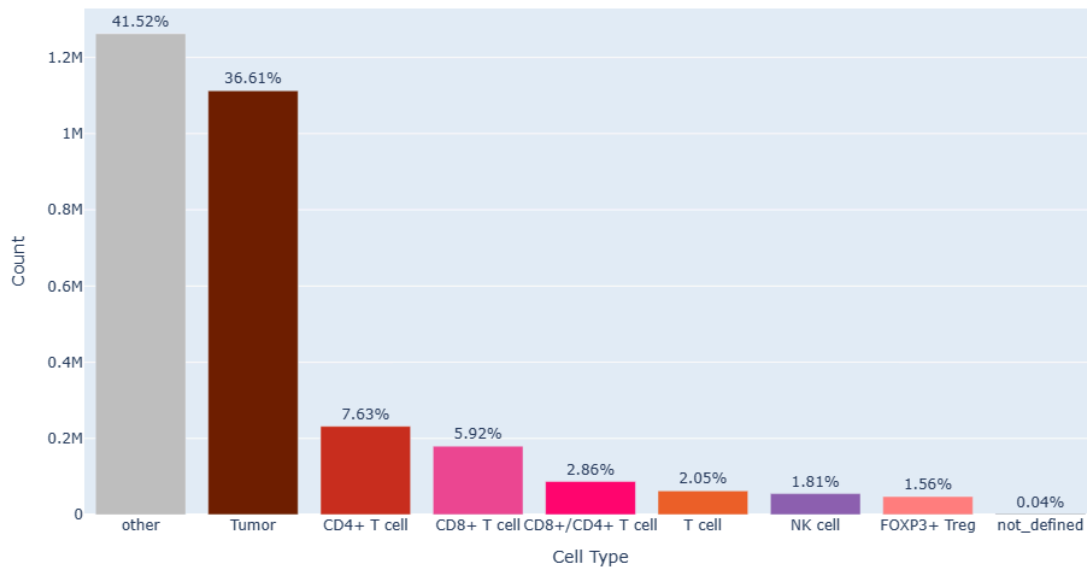
Due to the size of the provided dataset their rendering was rather slow, so i extracted them to standalone html files, those files (“plotIF1.html” etc) are available for download from the repository. In addition to those three general plots, i also made histograms of cell type frequency for each panel:



Histogram of Cell Types - IF2



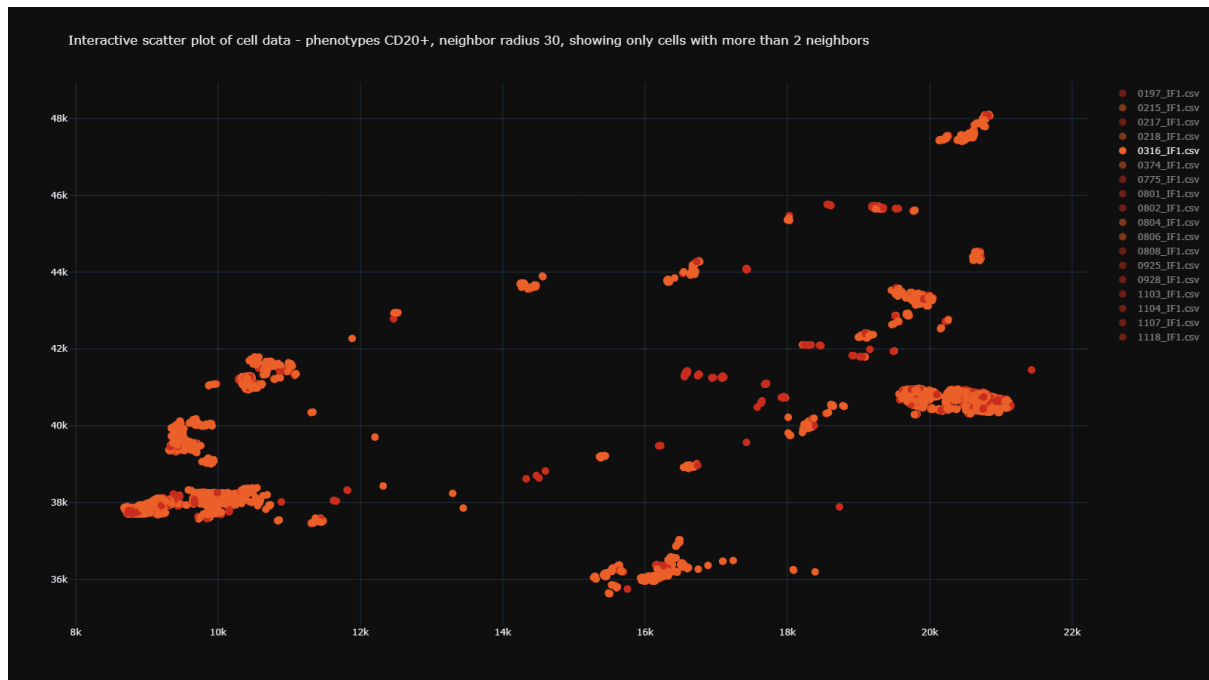
Histogram of Cell Types - IF3



As we can see, while panels 1 and 3 have similar proportions of tumor cells, panel 2 has way less and mostly consists of various T cells. Most cell types are used in more than one panel, whereas some other types like macrophages and neutrophils appear only in the first.

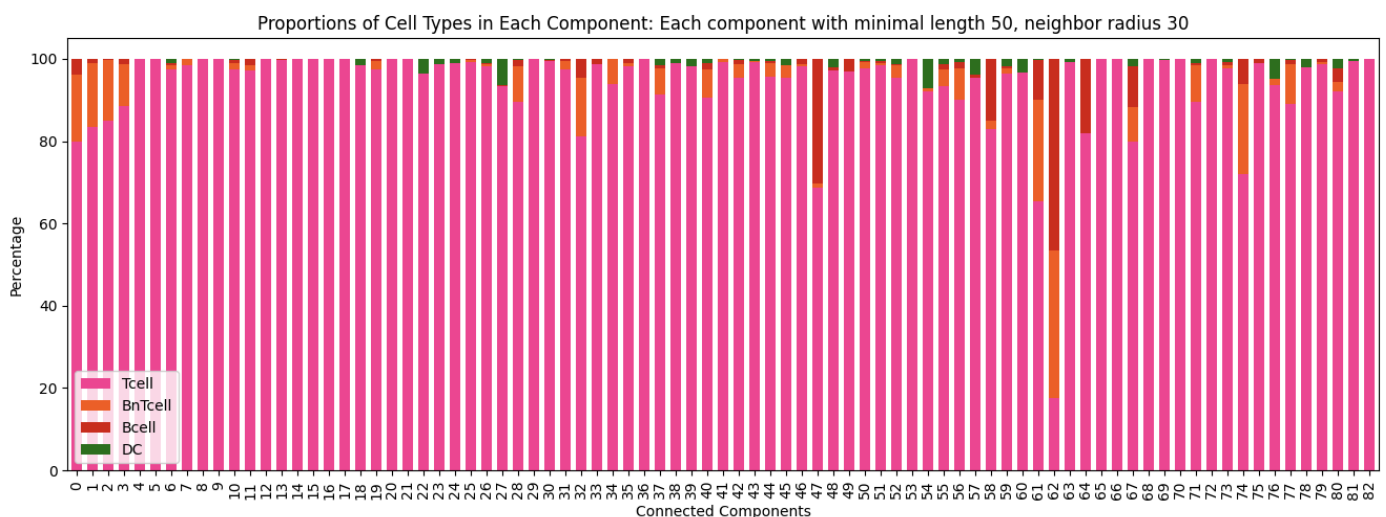
### 3 Using graphs and searching for TLS

Initially i attempted to see what i could find while utilizing a very simple technique: the file “graph\_simple.py” picks only cells with a specified phenotype and then uses the `radius_neighbors_graph` method from `sklearn.neighbors` to only show cells that have at least n neighbors within a specified radius. Here’s an example for the data shown in the Overview section :



While there are some sites that somewhat convincingly (due to visible B/T cell clusters) seem like potential TLS candidates in data from several patients, this isn’t a very precise method and in some cases like for 0197\_IF1 the results are not that good and the start and end points of the potential TLS are not always obvious. This interactive plot can be viewed in the “plotIF1\_neighbors” file.

#### 3.1 Connected components



Using `networkx.connected_components` and `sklearn.radius_neighbors_graph`, I received the above plot for the file “0316\_IF”. Each bar represents a neighboring graph of continuously connected components that has a length specified by a parameter. Plotting these small networks doesn’t give that much information for the example “0316\_IF.csv”, but with slightly modified parameters it provides new insight into the example file “0197\_IF1.csv” that wasn’t clearly explained using “graph\_simple.py”:

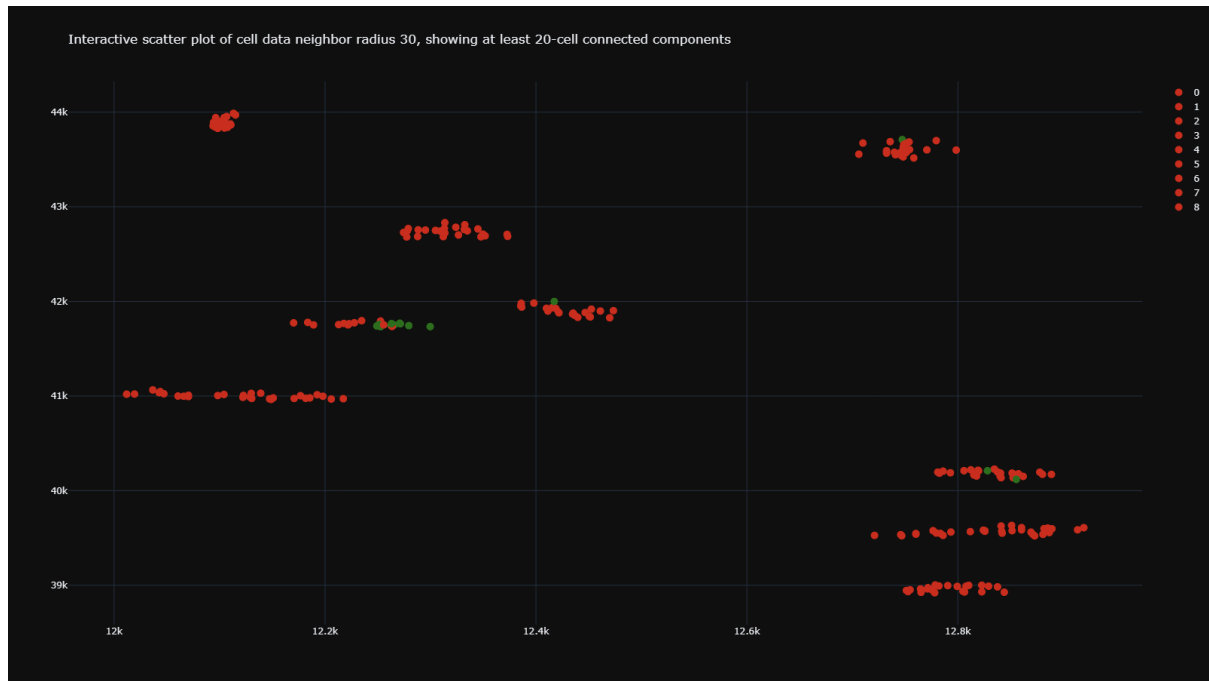


fig. 8: TLS candidates in “0197\_IF1” visible using “graph\_connected\_components.py”

The visible networks of lymphatic cells seem like good candidates for TLS sites. The interactive plot is saved to the file “plot\_clustered.html”.

## 4 Some conclusions

While those visualizations can be clearly improved upon methodically, they showcase that even without complex clustering algorithms, some interesting discoveries can be made analyzing mIF spatial data.