→**Terms that I haven't searched yet** :

→Image feature for alignment :

→DNA-stained nuclei serve as an excellent image feature for alignment

→File Formats that are actually used in this internship (can they be integrated into this tool ?)

→High-plex CyCIF images

→MCMICRO , the Nextflow-based image processing pipeline developed by HTAN

→The fundamental reason for **cycles** in high-plex imaging is to **overcome the spectral overlap limitations of fluorophores**

→You can't just put 60 different fluorescent antibodies on the tissue at once, because you only have, say, 5 distinct color "channels" you can detect. Many of those 60 fluorophores would emit light that overlaps into other channels, creating a messy, uninterpretable signal.

→(3 markers/cycle * 20 cycles = 60 markers).

→**Motivations** :

→Recent approaches to highly multiplexed imaging generate high-plex data from sequential rounds of lower-plex imaging.

→Attaining mosaic images with single-cell accuracy requires robust image stitching and image registration capabilities that are not met by existing methods.

→Results :

→Generating accurate whole-slide mosaics from individual multiplexed images

→**Requirements** :

→ Development using python

→ASHLAR reads image formats from most commercial microscopes and slide scanners

→ASHLAR outputs standard OME-TIFF (Ready for analysis by exiting images pipelines)

→Open Source : https://github.com/labsyspharm/ashlar

→Well documented : Home | ASHLAR (installation guides , test data , example usage)

→Offers a command line oriented interface

→Offers an Application programming interface (API) for integration with other tools

→ASHLAR is available as a Docker container

→ASHLAR can directly process any image format supported by the widely used Open Microscopy Environment (OME)

→OME-TIFF (.ome.tif or .ome.tiff)

→OME-NGFF / OME-Zarr (.ome.zarr or .zarr)

→Has been incorporated into MCMICRO

→Has been tested with several hundred CyCIF, CODEX and mxIF images acquired from 12 types of mouse and human tissues at seven different institutions on five different microscopes and slide scanner platforms

→Example datasets : Example datasets | MCMICRO

→Useful tutorial : Tutorial | MCMICRO

→Experimental background | MCMICRO

→Ashlar requires unstitched individual “tile” images as input - it is not suitable for microscopes or slide scanners that only provide pre-stitched images.

→Introduction :

→ **20-60 plex Sub-cellular resolution microscopy** : The resulting data can be processed to determine the molecular and physical relationships of cells

within the tissue to each other etc… . Research has shown that spatial profiling by highly multiplexed microscopy can reveal features of normal and diseased tissues and their responses to therapy that cannot be discerned in other ways

→**Tools used for that** :

→Conventional fluorescence and brightfield microscopes

→Are cyclic approaches in which high-plex data are generated by the repeated acquisition of lower-plex images, each of which has two to six channels of information.each of which has two to six channels of information. Each channel represents an image acquired with excitation and emission filters matching one antibody or oligo-coupled fluorophore

→MxIF
→CyCIF
→CODEX
→4i
→mIHC

→Vaporize specimens with ion beams or lasers followed by atomic mass spectrometry

→Multiplexed Ion Beam Imaging (MIBI)
→Imaging Mass Cytometry (IMC)

→Existing imaging methods differ in resolution, field of view and number of distinct antigens or genes that can be detected (the assay ‘plex’)

→**Problems** :

→**Fluorescence microscopes** :

→Can acquire data from up to 6 different channels
→At resolutions down 0.25 µm (laterally)
→which makes a detailed analysis of intracellular structures possible

→**Slide Scanners** :

→Are microscopes equipped with rigid slide holders that move in X and Y

→At resolutions sufficient for subcellular imaging, collecting data from a whole slide involves acquiring an array of multiple image ‘tiles’.(10^3 or more for large sepcimens of 6 cm2)

→Each tile is a multi-wavelength megapixel-scale :
→Wavelength = channels
→megapixel-scale = This refers to resolution and data size
→Each tile has a different lateral (x, y) stage position
→This means that each unique tile image corresponds to a slightly different physical area of the tissue specimen. The scanner precisely moves the slide from one (x,y) coordinate to the next, taking a picture at each stop, until the entire area of interest on the slide has been covered.

→**Différence** :

→The number of wavelengths in each tile, the number of tiles and the number of imaging ‘cycles’ (each of which involves the acquisition of a full set of tiles), differs with the microscope and the multiplexing technology.

→**What is universally true** ? :

→it is universally true that tiles from all cycles must be merged accurately into a single high-plex ‘mosaic’ image.

→**The Data level concept** :

→High-plex mosaic images represent the key ‘Level 2 or 3’ data type for all subsequent visualization and quantitative data analysis.

→Level 1 : corresponding to single, raw image tiles
→Level 2 : data to stitched, illumination corrected mosaics
→Level 3 : to mosaic images that have also been subjected to manual or automated quality control to improve interpretability and accuracy

→**What is the greatest challenge ?** :

→It is increasingly clear that the greatest challenges in the acquisition and analysis of high-plex image data lies not in image acquisition *per se*, but in the subsequent image processing steps.

→**Introduction of uncertainty** :

→For example, even the best microscopes require computational alignment of tiles to form a mosaic, since mechanical tolerances and imperfect calibration introduce uncertainty into recorded tile positions.

→**Solution** :

→To enable the assembly of a mosaic, tiles are slightly overlapped during acquisition so that each pair of adjacent tiles contains some identical cells. Image features in these cells are then used as reference points for ‘stitching’ adjacent tiles into a seamless mosaic

→**What happens in cyclic imaging** ? :

→In cyclic imaging all tiles from the second and subsequent cycles must also be aligned to the mosaic through ‘registration’ of image features across corresponding tiles.

→**Problems with the tools that exist** :

→Multiple tools exist for registering image stacks and stitching image tiles

→and some are available in common image analysis software such as ImageJ

→However, we have found that open-source tools currently available for stitching and registering whole-slide images are unsatisfactory when applied to high-plex cyclic images with respect to :

→Speed
→Reliability
→Accuracy

→Some commercial instruments have also integrated stitching routines, but we have found that these methods are only sufficient for visual review and are generally not accurate enough for quantitative single-cell analysis.

→This means that the stitched images produced by commercial software might *look* fine to the human eye. You can scroll around the large image and get a general sense of the tissue morphology and marker distribution. For a quick qualitative assessment, it might be acceptable

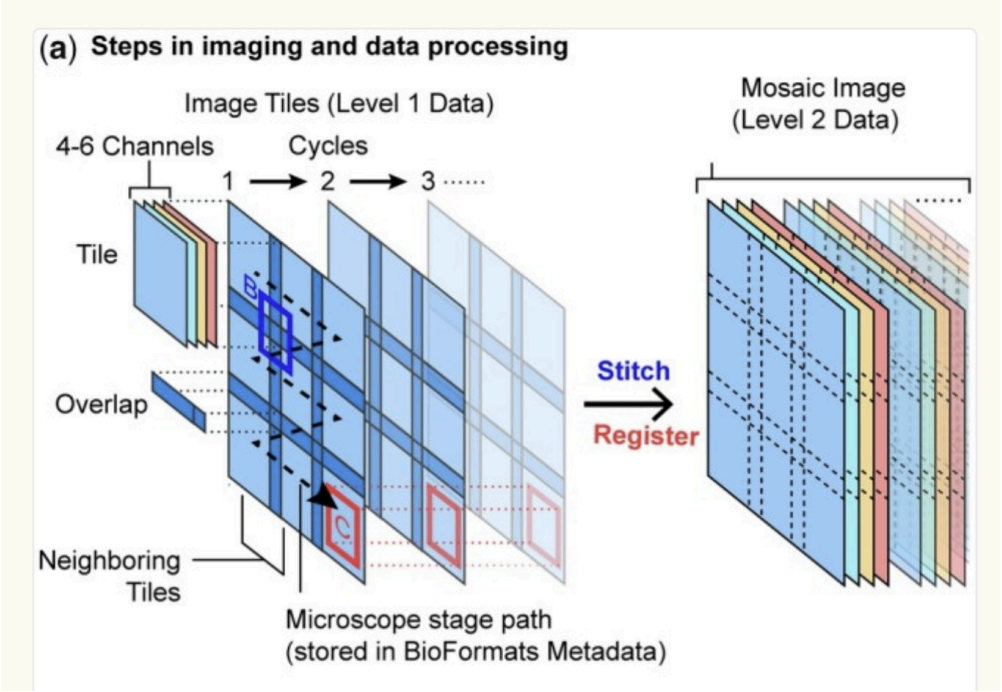
→ You can visually scan the whole stitched image, recognize the overall tissue structure, and see where different cell types generally are. It might look "fine" because small misalignments or distortions of a few pixels are not easily detectable by the human eye at a macroscopic level, and they don't hinder the general qualitative assessment.

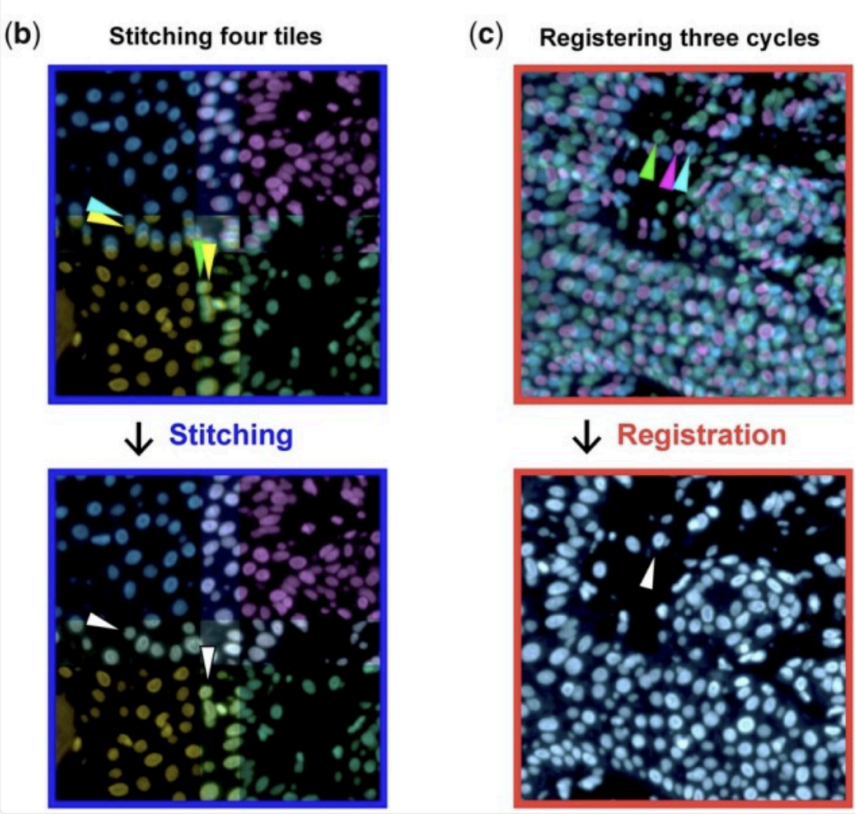
→This involves assigning numerical values to observations and features, allowing for precise measurements, statistical comparisons, and the identification of statistically significant differences or correlations.

→ "The mean fluorescence intensity of marker A in these cells is 1500 arbitrary units, significantly higher than the 500 units in control cells." (This can reflect protein expression levels)

→Existing tools also struggle with very large images and generally require substantial format conversion and file renaming, a non-trivial task when confronted with 100 GB of data contained in 10 ^4 megapixel-scale image tiles (a large 10-cycle whole-slide image)

→Overview of the full process :





→ASHLAR operates in three broad phases to convert a multi-cycle multi-tile (Level 1) dataset into a cohesive (Level 2) mosaic image.

→**Steps** :

→Tiles within the first imaging cycle are stitched;

→Tiles from the second and subsequent cycles are registered to corresponding tiles from the first cycle

→All tiles from all cycles are merged into a mosaic image.

→The output of stitching and registration is a list of new, corrected positions for all tiles in each cycle. Only in the final mosaic phase is the actual full-size many-channel mosaic image created.

→In many cases, the boundary of a tissue specimen is irregular, and a significant fraction of the tiles in a rectangular data collection grid contain few if any cells, posing a challenge for stitching as well as an opportunity to reduce data collection demands by creating irregular-shaped tile sets that closely follow the outlines of the tissue.

→Illustration explanation :

→a)-One cycle of whole-slide imaging (scanning) is achieved by moving the microscope stage along a controlled path and acquiring multichannel image tiles that overlap. Further cycles repeat the process after the specimens are re-labeled with new antibodies or other detection reagents.

→Note that the left-hand portion of this panel depicts just a single reference channel (blue) across three cycles for clarity—actual data contains multiple channels and an arbitrary number of cycles.

→b)-The corners of four neighboring tiles from one cycle are positioned using the recorded microscope stage positions (upper panel) and the corrected stitched positions (lower panel).

→Arrows indicate two individual cells in the tile overlap regions before and after stitching.

→c)-The centers of three image tiles from different cycles are positioned using recorded stage positions (upper panel) and post-ASHLAR registered positions (lower panel). Arrows indicate one cell before and after registration (A color version of this figure appears in the online version of this article.)

→ Generate an initial estimate of tile positions, ASHLAR uses data from image tiles (grids of pixels), recorded stage positions and physical pixel dimensions.

→ASHLAR leverages the Open Microscopy Environment Bio-Formats library to extract the necessary image data and metadata (stage position and pixel size) directly from native image files produced by the great majority of commercial microscopes, obviating the need for image format conversion and manual metadata extraction.

→What about Microscopes that do not support **BioFromats** ?

→For microscopes that do not support BioFormats, ASHLAR accepts a set of TIFF files using a configurable naming convention along with an explicit specification of tile overlap and acquisition order.

→We performed stitching and registration only on the reference image channel (typically Hoechst 33342-stained nuclei) and applied the resulting positional corrections to all other channels recorded within that cycle.

→This is sufficient because the chromatic aberration exhibited by research-grade wide-field microscopes is not a major contributor to image inaccuracy at resolutions typically used for tissue imaging (10–40× magnification, 0.3–0.95 NA air objectives).

→Because the different colors (wavelengths) of light are focused to nearly the same spot by the microscope's lenses, the green channel, red channel, blue channel (etc.) images captured in a single cycle will be inherently well-aligned with each other.

→**Contraint** :

→10x-40x magnification