Abstract—Inspection of tissues using a light microscope is the primary method of diagnosing many diseases, notably cancer. Highly multiplexed tissue imaging builds on this foundation, enabling the collection of up to 60 channels of molecular information plus cell and tissue morphology using antibody staining. This provides unique insight into disease biology and promises to help with the design of patient-specific therapies. However, a substantial gap remains with respect to visualizing the resulting multivariate image data and effectively supporting pathology workflows in digital environments on screen. We, therefore, developed Scope2Screen, a scalable software system for focus+context exploration and annotation of whole-slide, high-plex, tissue images. Our approach scales to analyzing 100GB images of $10^9$ or more pixels per channel, containing millions of individual cells. A multidisciplinary team of visualization experts, microscopists, and pathologists identified key image exploration and annotation tasks involving finding, magnifying, quantifying, and organizing regions of interest (ROIs) in an intuitive and cohesive manner. Building on a scope-to-screen metaphor, we present interactive lensing techniques that operate at single-cell and tissue levels. Lenses are equipped with task-specific functionality and descriptive statistics, making it possible to analyze image features, cell types, and spatial arrangements (neighborhoods) across image channels and scales. A fast sliding-window search guides users to regions similar to those under the lens; these regions can be analyzed and considered either separately or as part of a larger image collection. A novel snapshot method enables linked lens configurations and image statistics to be saved, restored, and shared with these regions. We validate our designs with domain experts and apply Scope2Screen in two case studies involving lung and colorectal cancers to discover cancer-relevant image features.

Index Terms—Histopathology, Focus+Context, Image Analysis

1 INTRODUCTION

Since the end of the 19th century, the diagnosis of many diseases - cancer in particular - has involved human inspection of stained tissue sections using a simple light microscope [19]. Histopathology in both research and clinical settings still involves microscopy-based inspection of physical slides but a rapid shift to digital instruments and computational analysis (scope to screen) is now underway [19]. Digital pathology [74] in a clinical setting focuses on the analysis of tissues stained with colorimetric dyes (primarily hematoxylin and eosin, H&E [67]) supplemented by single-color immunohistochemistry methods that use antibodies to detect molecular features of interest [52]. In research settings, recently developed high-plex imaging methods such as CyCIF [35, 36], CODEX [26], and mxIF [25] to measure the levels and sub-cellular localization of 20-60 proteins, providing single-cell information on cell identities and states in a preserved tissue.
environment. The resulting data are complex, involving multi-channel gigapixel images having $10^6$ or more cells. Underdevelopment of analytical and visualization methods is a barrier to progress in digital pathology, explaining the continuing dominance of physical slides.

Machine learning on high-plex tissue images has shown promise, particularly with respect to automated classification of cell types [16, 37], tissue morphology [65], and cellular neighborhoods [40]. However, such data-driven approaches do not leverage hard-won information known primarily to anatomic pathologists on which cell and tissue morphologies are significantly associated with disease outcome or response to therapy. A hundred years of clinical pathology has investigated many striking and recurrent image features whose significance still remains unknown. A critical need, therefore, exists for new software tools that optimally leverage human-machine collaboration in ways that are not supported by existing interfaces [11, 62].

Pathologists are very efficient at extracting actionable information from physical slides, frequently panning across a specimen while switching between low and high magnifications. They record key observations in notes and by placing dots on slides next to the key features. Digital software needs to reproduce this efficiency and functionality (including a ‘dotting’ function) while using visual metaphors to present associated data and using machine learning to find similar and dissimilar visual fields. Designing scalable visual interfaces that will work in the context of high-volume clinical workflows [47] and to high-dimensional research data represents a substantial challenge.

We addressed these challenges as a team of visualization researchers, pathologists, and cell biologists via a process of goal specification, iterative testing and design, and real-world implementation in a biomedical research laboratory. We make three primary contributions. (1) We demonstrate task-tailored, lens-centric focus+context technique, which enables intuitive interaction with large (ca. 100 GB) multi-channel images and linked multivariate data (Fig. 1). The lensing technique allows users to focus on different aspects of a region for close-up analysis while maintaining the surrounding context. We design novel domain-specific encodings in which features computed from the image (spatial cross-correlation or cell identity) can be accessed in conjunction with the image. (2) We integrate interactive real-time spatial histogram similarity search algorithms able to identify recurrent patterns across gigapixel multi-channel images at different resolutions. Integrated into the lens, this search guides analysts to regions similar to the one in focus, enabling exploratory analysis at scale. (3) We present a scalable system that combines lens and search features with interactive annotation tools, enabling a smooth transition from exploration to knowledge externalization. Analysts can save, filter, and restore regions of interest (ROIs) within the image space (along with underlying statistics of the filtered single-cell data, channel identities, and color settings) and export them for continued study. Two use-cases demonstrate the applicability of our approach to patient-facing (translational) cancer research and point to future applications in diagnosis and patient care.

2 RELATED WORK

The related work is three-fold. We first discuss large-scale image viewers as an enabler for our approach. We then summarize focus+context techniques in comparison to overview+detail and pan&zoom, with a focus on image data. Lastly, we compare ROI annotation approaches.

2.1 Scalable Image Viewers For Digital Pathology

Many biomedical visualization systems focus on the display of large 2D imaging data and apply multi-resolution techniques such as image pyramids [10] to handle large data sizes at interactive rates. DeepZoom [44] hierarchically divides images into tile pyramids and delivers pieces as required by the viewer. Zarr [9, 45], a file format and library, abstracts this concept by providing storage of chunked, compressed, N-dimensional arrays. Viewers such as OpenSeadragon [3] and Viv [41] leverage these libraries and add GPU-accelerated rendering capabilities. On top of that, many solutions offer data-management, atlas, and analysis capabilities. OMERO PathViewer [11] is a widely used web-based viewer for multiplexed image data. As an extension to the data management platform OMERO, it supports a variety of microscope file formats. Online cancer atlases such as Pancreatlas [57] and Pan-Cancer [73] support data exploration with storytelling capabilities. Minerva Story [28, 55] is a new tool used to create atlases for the Human Tumor Atlas Network [56]. Other solutions focus on combining image visualization with analytics. Nappi [62] is a fast and light-weight multi-dimension viewer designed for browsing, annotating, and analyzing large multi-dimensional images. Written in Python, it can be extended with analytic functionality, e.g., in combination with SciMap [51]. Other analytical tools focus on end-users, such as the open-source solutions histoCAT [58] and Facetto [33], and commercial tools such as Halo [30] and Visiopharm’s TissueAlign [6] supporting split-screen comparison for serial sections. Screenit [21] presents a design to analyze smaller histology images at multiple hierarchy levels. Similarly, ParaGlyer [50] is an analysis approach for multiparametric medical images that permits analysis of associated feature values and comparisons of volumetric ROIs by voxel subtraction. Somarakis et al. [63] offer comparison views with a focus on spatially-resolved omics data in a standard viewer. These tools feature multiple linked views for overview+detail exploration. In comparison, our solution focuses on interactive focus+context and rich annotation with contextual details displayed near the ROI and supports a neighborhood-aware similarity search on top of local image pixel and feature value comparison. Most viewers operate on much smaller datasets. Our viewer builds on Facetto [33] and Minerva [28, 55] and supports multi-channel and cell-based rendering with linked data at a scale few other solutions support. The main contribution of this paper, however, is the embedded interactive lensing technique for multivariate image data and its task-tailored features supporting the digital pathology workflow.

2.2 Focus+Context-based Image Exploration

Cockburn et al. [20] categorize interaction techniques to work at multiple levels of detail into focus+context (F+C), overview+detail (O+D), zooming, and cue-based views. F+C minimizes the seam between views by displaying the focus within the context, O+D uses spatial separation, zooming temporal separation [71], and cue-based methods selectively highlight or suppress items. Comparative studies show that F+C techniques are often preferred and allow for efficient and effective target acquisition [61] and steering tasks [27] in multi-scale scenarios. A common F+C technique is the lens [14], a generic see-through interface that lies between the application and the cursor. Tominski et al. [68, 69] present a conceptual pipeline for lensing consisting of selection (what data), the lens-function (filters, analysis), and a join operation with the underlying visualization (mapping, rendering). They further categorize into lens properties (shape, position, size, orientation), and into data tasks, e.g., (geo)spatial analysis. Different lenses to magnify, select, filter, color, and analyze image data were proposed: Carpendale et al. [17] present a categorization of 1-3D distortion techniques to magnify in 2D uniform grids. Focusing on lens-based selection, MoleView [29] selects spatial and attribute-related data ranges in spatial embeddings and Trapp et al. present a technique for filtering multi-layer GIS data for city planning [70]. Similarly, Vollmer et al. propose a lens to aggregate ROIs in a geospatial scene to reduce information overload [72]. Flowlens [23] features a lens for biomedical application: to minimize visual clutter and occlusions in cerebral aneurysms. There are a few tools in digital histopathology with lensing capabilities. Vitesse [24], positions linked views around an image viewer [41] and includes a lens to show a predefined set of channels. However, by design, they do not focus on supporting a specific pathology process, nor does the lens support magnification, feature augmentation, comparison, or search.

2.3 Handling and Visualization of ROI Annotations

Different techniques exist to mark, visualize, and extract ROIs in images, but only a small subset is used in the digital pathology domain. QuPath [13], an extensible software platform, allows annotating histology images with free form selection tools and more advanced selection options like pixel-based nearest neighbors and magic wand, extending from the clicked pixel to neighboring areas with a threshold. Visiopharm’s viewer [7] similarly provides different geometric shapes to
annotate images. In Orbit [66] and Halo [30] users can define inclusion and exclusion annotations, organize them in groups, and train a classifier. Going beyond manual annotation, Quick-Annnotator [43] leverages a deep-learning approach to search and suggests regions similar to a given example. Similarly, Ip and Varnesh [31] narrow down and cull out ROIs of high conformity and allow users to interactively identify the exceptional ROIs that merit further attention on multiple scales. We incorporate these ideas but instead apply a fast neighborhood-based histogram search running on multiple image channels in real-time to guide the user to similar areas in the viewport.

When large images are annotated on different scales it becomes challenging to navigate in an increasingly cluttered space. Some features might even be too small to be identifiable at certain zoom levels. Scalable Insets [34] is a cue-based technique that lays out regions of interest as magnified thumbnail views and clusters them by location and type. TrailMaps [75] proposes an algorithm to automatically create such insets (here bookmarks) based on user interaction and previously viewed locations. They also offer timeline- and category-based groupings for a better overview and faster navigation. We choose a more familiar design to cater to the application domain needs and conventions but enhance our approach by supporting rich annotations that store not only the ROI but also linked single-cell data and descriptive statistics. Closely related to our approach is the work by Minde et al. [46] that proposes annotations linked to contextual information so that they remain meaningful during the analysis and possible state changes. We extend this idea with overview, search, and restoring capabilities integrated into focus-context navigation in large-scale multivariate images.

3 BACKGROUND: MULTIPLEX TISSUE IMAGING

We analyze multiplexed tissue imaging data generated with CyCIF [35] but our visualization approach can be applied to images acquired using other technologies such as CODEX [26]. Images are segmented and signal intensity is measured at a single-cell level. Here we provide a brief overview of the process and data (Fig. 2).

Acquisition. Multiplexed tissue imaging allows to analyze human tissue specimens obtained from patients for pathologic diagnosis. The approach used by the investigators, as described in our previous work [33], involves iterative immunofluorescence labeling with 3-4 antibodies to specific proteins followed by imaging with a high-resolution optical microscope in successive cycles. This results in 16-bit four-channel image datasets for up to 60 proteins of interest (60 images), 30k x 30k in resolution, and often greater than 100GB in size, allowing for extensive characterization and correlation of markers of interest in large tissue areas at sub-cellular resolution.

Processing. High-resolution optical microscopes have limited fields of view, so large samples are imaged using a series of individual fields which are then stitched together computationally to form a complete mosaic image using software such as ASHLAR [48, 49]. A nonrigid (B-spline) method [32, 42] is applied to register microscopy histology mosaics from different imaging processes [15], e.g., CyCIF and H&E. CyCIF mosaics can be up to 50,000 pixels in each dimension and contain as many as 60 channels, each depicting a different marker. Mosaic images are then classified pixel-by-pixel to discriminate cells using, e.g., a random forest [64], then individual cells are segmented [38]. Segmentation information is stored in 32-bit masks that define the cell ID for each pixel in a multi-channel image stack. Next, per-cell mean intensity values (arithmetic mean) for each cell. The tabular data is linked via cell ID and X,Y position.

segmentation mask labels individual cells in each tissue specimen with a unique cell ID. Similar to each image channel, the mask is stored in pyramid form. A CSV file stores single-cell features (columns) for each cell (row). These features consist of extracted mean intensity values per image channel for that cell, x and y position of the cell in image space, and its cell ID.

4 DOMAIN GOALS AND TASKS

This project is rooted in a collaboration with physicians and researchers in the Department of Pathology and the Laboratory of Systems Pharmacology (LSP) at Harvard Medical School. Four experts in the domain of digital histopathology participated in the project. The team consists of two pathologists, two computational biologists, and four computer scientists. The overall goal of our collaborators is to characterize the features of tumors including cell types & states, their interactions, and their morphological organization in the tumor microenvironment.

Pathologists are physicians who diagnose diseases by analyzing samples acquired from patients. Anatomic pathologists specialize in the gross and microscopic examination of tissue specimens. They characterize cell and tissue morphology using light microscopy, and molecular features using immunohistochemistry and immunofluorescence. The pathologists involved in this project engage in research and have expertise in imaging, computational biology, and defining the role of diverse cell states in shaping and regulating the tumor microenvironment.

Computational and Cell Biologists complement expertise in biomedical science with skills in technical fields including mathematics, computer science, and physics. Multiplexed immunofluorescence experiments involve collection of primary imaging data which is used for a wide variety of complex computational tasks including image registration and segmentation, and extraction of numerical feature data, as well as downstream analyses of cell states, spatial statistics, and other phenotypes. Biologists interpret aspects of cell morphology and marker expression, but pathologists complement these analyses with greater depth of experience with human tissue morphology and disease states.

Visualization Experts. By contrast, the computer scientists provide expertise in visualization and visual data analysis. They work in close collaboration with the aforementioned investigators to provide novel analytics prototypes that perform a variety of analysis tasks and can be integrated into research studies and laboratory IT infrastructure. To understand domain goals, the visualization experts in this study participated in weekly meetings focused on image processing, biomedical topics, and on iterative goal-and-task analyses for the proposed approach. The collaboration with the LSP started in 2018 with Facetto [33]. In Fall 2020, this team began working together to develop advanced tools that cater to visual exploration, inspection, and annotation. We followed the design study methodology by Sedlmair et al. [60]. This methodology describes a setting in which visualization experts analyze a domain-specific problem, design a visualization approach to solving it, validate their design, and reflect on lessons learned.

4.1 Tasks and Challenges

In weekly sessions, we identified a workflow (Fig. 3) of consecutive tasks (T1-T6) leading from image exploration and close-up inspection of regions to annotation and extraction of patterns.

T1. Explore Multimodal, Highplexed Image and Feature Data in Combined Setting: A pivotal task is rapid navigation and visualization of multi-channel images. Pathologists normally operate by moving slides physically on a microscope stage and switching between view...
magnifications levels. They depend on a seamless visual experience to diagnose diseases or conditions. Challenge: Image analysis must not only provide seamless pan & zoom, but also switching between channels of different image modalities. Existing solutions do not scale beyond 4 to 5 channels. They also lack on-demand rendering, blending of multiple channels, and ways to highlight and recall ROIs.

T2. Close-up ROI Analysis: Once a region of interest is found in the tissue specimen, experts focus and zoom in on the area for close-up inspection and measurement, without losing the spatial context of e.g., a tumor region’s surrounding immune cells. Challenge: In addition to interactive rendering of different resolution levels in a combined space, experts need to focus and measure without losing proportions and larger context. Panning and zooming between overview+detail and individual marker channels requires a large amount of mental effort as either context or details are lost [34].

T3. Regional Comparison of Image Markers: For a region in focus, cell biologists need to relate and compare between different marker expressions (e.g., DNA, CD45, Keratin) of different image modalities (CyCIF [35, 36], H&E [67], CODEX [26], mIF [25], etc.), encoded in individual image channels. Challenge: Whole-slide switching between channels can lead to losing focus and change blindness due to different morphological structures.

T4. Relate to Spatially Referenced Single-Cell Expressions: Besides looking at the raw image, experts analyze extracted single-cell marker values and their spatial statistics in (A) image and (B) high-dimensional feature space. Of special interest is cell density in the tissue, counts and spatial arrangements of cell-types, and distributions of marker intensities. For each of these descriptive statistics, it is important to relate regional phenomena to statistics of the whole image. Challenge: Providing complex spatially referenced information in proximity while dealing with a dense cellular image space and catering to highly specific domain conventions.

T5. Find Similar Regions: Analyzing a whole slide image is time-consuming. Often the cancer micro-environment consists of repetitive patterns of cell-cell interactions and morphological structures across channels that pathologists annotate and compare to each other. Challenge: Finding and guiding users to such structures in an interactive fashion on different spatial scales and across image dimensions.

T6. ROI Annotation and Summaries: A common pathology task concerns the manual delineation of tumor mass and other structures on the digitized tissue slide, known as region annotation. These annotated regions need to be extracted, semantically grouped, and summarized in a structured way for collaboration and examination. Challenge: The annotation process must be integrated seamlessly with the analysis so that experts can extract, group, and refine patterns along the way.

5 Approach

We used the tasks of Section 4 to guide the design and implementation of Scope2Screen, playing the translator role put forth in the design study methodology by Sedmair et al. [60]. Fig. 1 offers an overview of the user interface. After importing the data, analysts can explore (Fig. 1) the image by activating a set of different channels (A) with distinct color configurations and adjustable intensity ranges. The selections are then rendered in a combined view (B) for interactive panning and zooming (T1). Using a virtual lens (C), users can focus, magnify, and measure regions of interest for close-up analysis (T2). By toggling image channel combinations inside and outside of the lenses, one can regionally compare different combinations of marker expressions (T3). Other lens filters link to underlying single-cell data, offering descriptive statistics about marker distributions and cell counts and types (T4). Using a search-by-example approach, the tool guides users to regions with similar patterns as those in scope (T5). To save and extract a region, analysts can take snapshots that save the ROI together with relevant notes (Fig. 1 D), current settings, and interior single-cell data for the region (T6). Annotated areas can be filtered and exported to share with collaborators and to recall for further examination. In the following sections, we introduce the corresponding techniques and features (F1-F6) and discuss design decisions to enable these tasks.

5.1 Scalable Image Exploration (F1)

We designed Scope2Screen for interactive exploration and multi-scale visualization (T1) of resection specimens (Sec. 3). To make the viewer scalable for high-resolution data, we decided to leverage image pyramids [10] to load only sections of the image for a given viewport and zoom level. To enable flexible exploration of image channels and to visualize data in the viewport, the viewer operates in two multi-resolution rendering modes: channel-based and cell-based. Channel-based rendering maps intensity values of selected image channels to color. It then computes a mixed color value for each pixel in the viewport. Cell-based rendering leverages a layered segmentation mask that indexes each pixel to a cell. This way, each cell can be colored individually to visually express selections, cell types, etc. Both rendering modes operate at interactive rates, allowing users to pan & zoom, select regions, and to color and mix channels in real-time. These rendering modes cater to our experts’ needs to analyze on tissue and single-cell level. Sec. 6 gives details on the implementation of our system.

5.2 Lensing Features for Multivariate Image Data

Conveniently, the principal technology of the microscope, the lens, is highly adaptable. To enable close-up analysis (T2) of ROIs and to connect the optical and digital experiences for users, we introduced a digital lens designed to imitate the familiar experience of inspecting through an eyepiece. To support the requirements of our collaborators, we equip the lens with features (Fig. 4,5), ranging from magnification and channel filters (F3) to descriptive single-cell statistics (F4).

5.2.1 Magnification and Measuring (F2)

Our users rely on being able to toggle quickly between high and low magnification powers as part of an established workflow for considering a region of interest up-close, as a localized arrangement of cells, and in-context, as part of a larger tissue sample (T2). Within the virtual viewer, this zoom-level interchange can be challenging to control. Constraining magnification to the lens’s boundary (Fig. 5) while maintaining the contextual overview then becomes a convenient strategy for handling simultaneous focal analysis. Because the magnifying lens (Fig. 5 A), when active, occludes part of the image space, we experimented with a common spatial manipulation to create a faux-spherical representation: the fisheye (B). However, distortion is a troubling approach for experts who make evaluations based on morphology, leading us to introduce a hybrid plateau model (C) that maintains the original composition within the central area of the lens using a standard zoom and only compresses the periphery to seamlessly transfer into the context without occlusion.
5.2.2 Channel Analysis (F3)

To address regional comparison of channels (T3) that represent data from the same modality (e.g., CyCIF), other modalities (e.g., H&E [67]), and across different planes (e.g., slide sectioning), we iterated over different designs and finally settled on the following features.

The first lens feature allows users to quickly isolate each selected channel individually for improved views of a distinct channel (Fig. 4 A) in focus while keeping the multi-channel setting in the context. The second feature combines multiple channels in the scope (Fig. 4 B, Fig. 9) while the context can keep a distinct setting (Fig. 1). We specifically designed this setting for semantically dependent channels such as RGB images for H&E staining. It also addresses our experts’ needs to analyze spatial relations of a set of independent channels, for example, from specific immune markers in a region of interest, while keeping globally a different set of more general cancer and stromal markers. Addressing early feedback from our experts, we added the capability to equip the lens with multiple sets of such channel combinations (Fig. 9) in advance or add them during analysis. These sets can be quickly toggled during exploration to investigate different biological questions, e.g., focusing on immune reactions, or tissue architecture. To overcome occlusion, we chose to offer two solutions leveraging temporal and spatial separation. Firstly, we introduce adjustable interpolation controls allowing to blend seamlessly from the overlaying lens-image to the underlying global channel combination. This transition helps to visually align and keep track of often very different structures in different channel combinations. To further reduce change blindness, a split-screen lens juxtaposes a second lens instance in proximity, which displays (copies) the occluded part in the original global channel setting (Fig. 4 C). This allows for side-by-side comparison.

5.2.3 Feature Augmentation (F4)

Our multiplex image data (Sec. 3), often comprises up to 60 channels. While three channels can guarantee non-overlapping (RGB) colors, adding even more channels makes visual decoding for analysts mentally challenging and increasingly inaccurate. Additionally, color encodings of quantitative data are often hard to gauge [39]. Instead, to enable quantitative analysis of selected regions, we chose to augment the image space with descriptive statistics of the extracted single-cell data (T4) using more abstract visual encodings (Fig. 4 D-F). We developed three task-tailored lens settings showing marker distributions, density reports, and cell types and counts. With every update of the lens position on screen, the back-end queries the in-memory CSV table (Sec. 3) for cells in the lens’s area and returns cell Id values along with the requested statistics, which are then processed and rendered into different charts. To speed this up, we execute spatial range queries with a ball-tree index structure [22]. This yields a run time complexity of $O(n \log n)$.

**Single-Cell Histograms.** To analyze marker distribution in a selected region, we compute binned histograms of the single-cell aggregates (see CSV table, Sec. 3) for selected channels. We present these histograms (Fig. 4 D) in proximity to the focus area for quick look-ups. We decided to arrange the histograms in a vertical layout to ease comparison.

According to our domain collaborators, absolute comparison of individual markers is statistically not meaningful as the signal-to-noise ratio changes per cycle and stain in the imaging process (Sec. 3). Instead, they are interested in a relative comparison of the distributions. We use a $\log_{10}$ transformation to make these skewed marker distributions comparable, followed by a cut-off of the 1st and 99th percentile to remove outliers. Fig. 1 (C) shows the lens operating in this setting. To further analyze where cells lie in that spectrum, we chose to offer a brush functionality. The user can filter a range (min-max) in the brush to remove outliers.

**Radial Chart.** Additionally to histograms, we provide an overview of all markers by arranging their mean values in a radial layout in proximity to the lens (Fig. 4 E). This decision allows for a more com-
To find similar regions, we chose to provide a method operating directly on the whole image to align as close as possible with the visual perception of cell-type classification and clustering [16, 37] and set them into the full histogram computation. This is then compared to the lens histogram using Chi-square distance (see Eq. 1) for two arrays X and Y with N dimensions across all channels C. This leads to a similarity map with a similarity value for each pixel in the image (Fig. 6, Step 3).

\[
\frac{1}{C} \sum_{c=1}^{C} \sum_{i=1}^{N} \frac{(x_{c,i} - y_{c,i})^2}{(x_{c,i} + y_{c,i})}
\]  

Step 4: We use marching squares [53] to detect contours in the similarity map and extract these contours as polygons that we render on screen (Fig. 6, Step 4).

Step 2 and 3 can be computed for multiple channels. To not lose information, we compute each channel’s similarities separately and then aggregate them to a combined similarity map. Similar to our multi-resolution rendering strategy (Sec. 5.1), we execute the histogram search algorithm on the fly on the respective TIFF-file layer (Sec. 3) that aligns with the current zoom level. Fully zoomed out, the aggregation level is higher; while zoomed in, we process the highest resolution but for a smaller image area. The approach can also be employed in full resolution to the whole image (Sec. 5.4).

### 5.4 Descriptive ROI Summaries (F6)

When using a light microscope, a pathologist often ‘dots’ whole slides with a pen to indicate ROIs for later examination. We introduced a digital “dotter” to represent this approach but with extended support for annotation (T6) during exploration (Fig. 7). The lens functions as a camera lens that can snapshots whenever an ROI is in focus.

Our collaborators currently narrate image annotations to data stories for examination, teaching and outreach of their research [55]. The link to analysis results is often lost as annotations must be manually recreated in the used tools. To maintain analytic provenance, we developed a novel rich snapshot method that not only saves the image area under the lens but also stores all associated data: a list of active image channels in focus and in the peripheral context, channel colors, and range settings, the linked single-cell data in scope, and textual annotation such as title and a more detailed description pathologists and cell biologists can add.

These rich snapshots are available as thumbnails in the Dotter panel (Fig. 7 C) and are interactively linked to overlays within the viewport. Inside the panel, the user can save, delete and load ROIs from a database. Names and descriptions can be edited and referenced as tags for filtering results. It is important to our users to be able to quickly recall and restore the snapshots for further analysis and fine tuning, but also to return to ROIs in their original context. Thus, by clicking on thumbnails, the viewer navigates back to the coordinates and zoom level of the overlay. By clicking on the overlay’s marker icon, we fully restore the lens, along with its global view setting, i.e. active channels, range and color mappings for the context (Fig. 7 D). This workflow facilitates hand-offs between colleagues who benefit from shared evaluation.

To extend the search for an ROI in the Dotter panel, users can use HistoSearch (Sec. 5.3) to find regions alike in the whole image space (other than the viewport during interactive analysis). We render these
annotations as image-overlays but as soon as the user picks up a region, we restore the lens, and the user can alter the region as needed.

6 Architecture and Implementation
Scope2Screen is an open-source web-application (available here: [41]) with a back-end Python server built on Flask and a Javascript frontend. The server’s restful API allows to retrieve image and feature data and to steer analytics and is easily extendable with new methods and corresponding API endpoints. The frontend is built using Bootstrap, D3.js, and OpenSeadragon (OSD) [3], a web-based zoomable image viewer, which we extend significantly.

We developed a lensing library that supports a subset of the interactive features of Scope2Screen as a generic plugin [1] for OSD. The library relies on a hidden viewer that provides access to both in-view and out-of-view image data to support controlled rendering within the lens, along with other complementary features. Our application utilizes and extended lensing’s logic with additional features (Sec. 5.2).

Scope2Screen also builds on Facetto [33] but makes improvements to the architecture. Instead of preprocessing OME-TIFF [2] image stacks to DeepZoom [44], we now read chunks (cropped 2D arrays of the respective layer in the image channel/mask) on-the-fly from layered OME-TIFFs to render it in the viewport at multiple resolution levels, depending on the current zoom setting. We rely on Zarr [45], a format for the storage and handling of chunked, compressed, N-dimensional arrays. The client-side rendering is hardware accelerated. It relies on WebGL [8] shaders from Minerva [28, 55] and supports Facetto’s channel and cell-based rendering modes (Sec. 5.1) for both the lens (focus) and the whole viewport (context). To access and filter the linked single cell feature (CSV) data more dynamically and at scale, we moved data processing and ball-tree [22] indexed querying (Section 5.2.3) to the back-end so that the client only loads requested pieces (in lens or viewport), aligning with our multi-resolution rendering strategy.

7 Use Cases
We applied Scope2Screen to study two cancer datasets that we acquired from sections of lung and colon cancer using CyCIF [35]. Immediately adjacent sections were H&E stained [67] and used to evaluate tissue morphology. We carried out the analysis together with our collaborators over zoom using a Pair-Analytics approach [12]; we steered the tool guided by the domain collaborators. This method is advantageous because it pairs Subject Matter Experts (SME) with expertise in a domain with Visual Analytics Experts (VAE) who have the technical expertise in the operation of VA tools but lack contextual knowledge.

7.1 Use Case 1: Colorectal Cancer
Tumors are complex ecosystems of numerous cell types arranged into recurrent 3D structures. However, the patterning of specific tumor and immune cell-states is poorly understood due to the difficulty of mapping high-dimensional data onto large tissue sections. In use case 1, two pathologists and one cell biologist analyzed a human colorectal carcinoma (CRC1) from the Human Tumor Atlas Network (HTAN) (PMID 32302568) to explore tumor and immune cell interactions.

Data: CRC1 is a human colorectal adenocarcinoma that was serially sectioned at 5μm intervals into 106 sections. 24-marker CyCIF was performed on every 5th section. Every adjacent section was stained with H&E. CyCIF and H&E images were registered and stitched, and single-cell segmentation and fluorescence intensity quantification were performed using MCMMicro [59]. Cell types were defined by expression of lineage- & state-specific markers. Here, we analyze one CyCIF and an adjacent H&E image. The data included 40 CyCIF channels and 3 (RGB) H&E channels encompassing 1.28 mio cells in a field 26,139 x 27,120 pixels (8,495 x 8,814 microns) in size (87.66 GB).

Analysis: We began the analysis with a low magnification overview of the CyCIF images using DAPI (blue), keratin (white), and aSMA (red) channels in the whole viewport. In combination, these channels illuminated pathologically-relevant structures of the tissue, including the morphology of nuclei (DAPI), abnormal epithelial cells (keratin), and muscular layers (aSMA) (Fig. 8 A, B). A second channel defined immune populations including CD4+ helper T cells (red), CD8+ cytotoxic T cells (green), and CD20+ B cells (white) to analyze immune interactions with tumor and adjacent normal regions (Fig. 8 C).

For each marker, we defined an upper and lower color mapping range using the channel range sliders (Fig. 1 A). The low-magnification view revealed a small region of tumor budding cells (≤1mm²), a phenomenon in which infiltrating single tumor cells or small clusters of cells (≤4) appear to “bud” from the tumor-stromal interface at the invasive margin, correlating with poor clinical outcomes (Fig. 8 D).

We used the standard lens magnifier to focus analysis on the budding region while maintaining spatial context.

To further explore spatial patterns of marker expression, we activated the “single-cell radial chart”. It provides a dynamic display of the mean single-cell expression levels of all CyCIF markers within the lens and the global mean of the markers across the entire image for comparison. This enabled the experts to see that tumor and immune cells in the several regions, including the budding region, were positive for PD-L1, a protein that suppresses the activity of cytotoxic CD8 T cells, which is often clinically targeted by antibody therapies. To capture images and associated single-cell data of these ROIs for subsequent review of immune interactions and tumor features, we took snapshots and annotated the areas ’PD-L1 rich region’ (Fig. 8 E) for later analysis.

We next used the split-screen lens (Fig. 4 C) to view H&E and CyCIF images side-by-side. Using this tool, pathologists validated the alignment of H&E and CyCIF channels acquired from adjacent sections to compare histologic and molecular features. They identified areas in the H&E images with marked chronic inflammation (lymphocytes and macrophages) in the peri-tumoral stroma for further evaluation. Direct comparison of the H&E and CyCIF data in these tissue regions using the split-screen lens allowed the pathologists to characterize lymphocyte aggregates with predominantly cytotoxic CD8+ T cells in peri-tumoral stroma (Fig. 8 E), as well as clusters of CD20+ B cells...
and CD4+ helper T cells forming ‘tertiary lymphoid structures’ (TLS). Direct comparison of H&E and CyCIF data in this manner allowed for targeted molecular characterization of immune populations such as lymphocytes which is not possible with H&E alone (Fig. 8 F). To compare the marker intensity value distributions more precisely, we switched to the ‘lens histogram’ (Fig. 4 A) and compared the results with the ‘cell-type’ lens within each region to assess marker expression with the results of the automated cell type classifier.

Based on review of the images, the pathologists used the Dotter’s ‘snapshot’ function to annotate three immunologically distinct regions (described above), three tumor regions with distinctive histomorphology (glandular, solid, and mucinous regions), and adjacent normal colonic mucosa. Using the ‘similarity search’ on the normal mucosal region (Fig. 6, top), we identified areas with similar histologic features, confirmed by pathologist inspection, embedded within the tumor mass which were not readily apparent on initial low-magnification review of the H&E images, validating the utility of the search method. We saved these findings to our database for subsequent retrieval.

Feedback: Although we worked with our collaborators in weekly sessions over several months, we received additional comments on design improvements and future features within the 2-hour analysis session. They found the normal lens magnification the most useful. The fish eye lens was problematic for review of tissue images due to distortion of cell morphology and tissue architecture, which may complicate the interpretation of important pathologic features such as nuclear shape. This aligns with earlier feedback motivating the design of the plateau lens, which they viewed as a helpful improvement. One user was also suggested to equip the lens with different predefined filter settings to be able to quickly toggle for different analysis tasks. While both the radial chart and histograms were well received, they asked to highlight under- or over-represented marker values more prominently. The histograms were described as easier to read and should be extended to represent non-active image channels as well. They also asked for a) a heatmap that color codes marker correlation, and b) the ability to define channel combinations inside the lens on-the-fly, not only in pre-configuration. A recurring piece of feedback during the session was to add more descriptive statistics such as bin sizes in the histogram, ratios of cell populations, and ways to precisely define intensity ranges by value. The proximity of single-cell data to the inspected area and the H&E channel comparison, which is limited or non-existent in other tools, were described as particularly useful. The ability to simultaneously inspect and mark different regions of the tumor was perceived as a promising area for further development.

7.2 Use Case 2: Lung Adenocarcinoma

Lung adenocarcinoma is a common subset of lung cancer that, in later stages, often does not respond to chemotherapy. Immunotherapy has shown great promise, but patient response varies according to each tumor’s specific microenvironment. To assess why only certain patients respond, one needs an in-depth understanding of the tumor and immune landscape. Together with two biologists, we apply Scope2Screen to explore the immune reactions in lymphocyte structures.

Data: Using t-CyCIF, we have prepared a dataset of human lung adenocarcinoma wedge resection. The image data consists of 38 channels, each with a resolution of 39,843 x 29,227 pixels (12,949 x 9,499 microns) and 118.43 GB in size containing 534,677 segmented cells. The data is enriched with cell-type classification from a deep immune profiling of the tumor microenvironment.

Analysis: We began the analysis with the whole tissue in the viewpoint and activated the DNA channel for a general overview of the tissue architecture. Addressing the feedback from Use Case 1, we equipped the lens with a set of four predefined channel combinations, designed to investigate the tumor’s immune response from a high level to detail.

The first channel combination (Fig. 9 A) consisted of basic cell markers and was designed to assess overall tissue composition. Lung cells are marked by TTF-1, which enabled the detection of the aberrant tumor morphology in the upper region of the sample. The lens magnifier allowed the biologists to see that the tumor is heavily infiltrated with immune cells (marked by CD45 and Vimentin). We then switched to a second channel combination for immune cell markers (Fig. 9 B) to further investigate the types of immune cells infiltrating the tumor. This revealed dense aggregates of immune cells, mainly composed of a core of B cells, surrounded by T cells, which are known tumor-associated tertiary lymphoid structures (TLS). Lens magnification allowed our experts to quickly detect, mark and further characterize the large number (10-12) of TLS’s in this lung specimen. By switching to a third channel combination customized for TLS markers (Fig. 9 C), we inspected each TLS more closely (see Sec. 7.1 for a closer description).

We used the Dotter’s snapshot capability and placed landmarks as we reviewed the TLS’s. We measured the TLS size using the ruler functionality (1,200 microns on average). We activated the cell type lens equipped with immune and cancer type classifications which showed B cells in the center (Fig. 4 F, Fig. 9 C) surrounded by epithelial lung cells. The cell biologists also used this setting to check the dataset segmentation, which they rated as very accurate. While focusing on one of these TLS, we activated HistoSearch to find areas with similar marker patterns, successfully identifying the other TLS areas. Surprisingly, in some TLS the HistoSearch detected the structure’s outer rim but not its center. Using the magnifier to zoom into these regions, we noted a distinctly higher level of B cell marker CD20. We focused the lens on this area and magnified for close-up inspection. To better understand the marker distribution in this region, we switched the lens function to the radial distribution chart (Fig. 4 B), revealing a high Ki67 and PCNA concentration, markers for cell proliferation and growth.

Subsequently, we switched to the fourth channel group to compare lymphocyte markers across a TLS (Fig. 9 D). We activated the histograms and moved the lens outside of a TLS towards the lung epithelial cells, recognizing TTF-1 positive tumor cells. Some of these were MHC-II positive. By shrinking the lens scope to single-cell size, we compared the TTF-1 cells, finding that cells with MHC-II expression are non-proliferative. This suggests that transient MHC-II expression is coupled with entry into a non-proliferative state.

Feedback: While users were able to make meaningful observations throughout the evaluation, their commentary indicated two categories of improvements. Channel views and statistical views could be merged so users do not have to rely on rapid memory recall required for toggling filters. For example, simultaneous access to the single-cell marker intensity distribution would have been helpful to monitor non-selected markers and explore cell lineages. Acknowledging that our tools do not support a wide range of unanticipated biological questions, the absence of tools for spatial analysis stalled certain leading lines of inquiry. Measuring the degree of cell proliferation around immune cell clusters would have been an important next step for our users, who recommended that we prioritize the introduction of supporting algorithms and visualizations for spatial correlation.

8 Hands-On User Study and Questionnaire

To further evaluate the intuitiveness and usefulness of Scope2Screen’s interactive interfaces and lens features, we conducted hands-on user studies in which three of our four domain collaborators (that were involved in goal specification, iterative testing and design, and use-cases), two biologists and one pathologist, steered the tool, gave thoughtful feedback, and additionally filled out a questionnaire.
9 Conclusion and Future Work

We present a design study aimed at supporting single-cell research into the composition, molecular states, and phenotypes of normal and diseased tissues, a rapidly growing area of basic and translational biomedical research, as well as pathologists studying human tissues for the purpose of diagnosis and disease management. Our Scope2Screen system supports fluid interactivity based on familiar microscopy metaphors while enabling multivariate exploration of quantitative and image data using a lens. Throughout the design process and expert feedback, we identified three key areas in which current work could be most usefully extended. **Combining Vision and Statistics:** According to our experts, visual needs tools for presenting numerical data alongside image data. In many cases, generation of the numerical data is not the problem: computational biologists are familiar with scripting and statistical tools for deriving single-cell data from images (via segmentation) and linking it to external sources (e.g., genomic data) but the information is most useful alongside the original images. Pathologists in particular need to combine deep knowledge of tissue architecture with quantitative data. However, most visual tools do not offer sufficient flexibility, and scripts or notebooks (e.g., Jupyter) lack the interactive visual exploration. We intend to extend Scope2Screen to support scripted statistical queries integrated with lensing. **High-dimensional Features on the Horizon:** Recent development in digital imaging such as the ability to measure spatial distribution of RNA expression will result in data with thousands, not dozens, of dimensions. Mapping such data into the image space while extracting relevant information will require dimensional reduction techniques and suitable visual representations of found features so that only the most relevant or explanatory data are presented. Very high-resolution 3D microscopy of tissues is also being integrated with the high-plex 2D data described here and this will require appropriate visual metaphors for moving between resolutions and data modalities. **Scalability across Datasets:** Our use cases demonstrate uses for Scope2Screen in the analysis of a single dataset stored locally. However, digital histology is expected to transition to the analysis of multiple datasets accessed via the cloud. While Scope2Screen scales to a large set of images, it does not yet support analysis and annotation of image collections or work interactively with Docker-based image analysis pipelines. Adding this functionality will close the gap from data exploration and analysis to generation of machine-assisted interpretative data reports for research and clinical applications including interactive publication via tools such as Minerva [28, 55].

10 Acknowledgements

This work is supported by the Ludwig Center at Harvard Medical School, and by NIH/NCI grants NCI U2C-CA233262, NCI U2C-CA233280 and NCI U54-CA225088.

11 Outside (Competing) Interests

PKS is a member of the SAB or Board of Directors of Genocean Software, Applied Biomath, and RareCyte Inc. and has equity in these companies; PKS is also a member of the SAB of NanoString. Sorger declares that none of these relationships have influenced the content of this manuscript. Other authors declare no competing interests.
10.1109/TDCG.2012.213


