Utility of High-Throughput Imaging Mass Cytometry for Cancer Research: A feasibility study

Sindhura Thirumal*, Amoon Jamzad*, Tiziana Cotechini*, Charles T. Hindmarch*, Céline Hardy*, Nathalia Kim*, Amber Simpson*, Charles H. Graham*, David M. Berman*, D. Robert Siemens*, Parvin Mousavi, *SM IEEE**
*Queen's University, Kingston, Canada
Email: sindhura.thirumal@queensu.ca

***Abstract*—The Hyperion^{TM} Imaging System is a novel technology that uses imaging mass cytometry (IMC) to improve upon current methods of tissue imaging, enabling sub-cellular spatial resolution and acquisition of up to 37 proteins on a single tissue slide. The technology is fairly new, and thus we want to explore the types of analysis possible with these data. Here, we introduce an analysis pipeline to utilize machine learning-based analysis for IMC data using data from a muscle invasive bladder cancer patient cohort. We also propose a novel augmentation method to handle the challenge of low number of tissue samples from IMC studies. Our augmentation method was validated and shown to perform better than when only using the original data. Both our pipeline and augmentation method show promise for applications in future research studies and clinical evaluation of this technology. Our results indicate the feasibility of using the proposed framework with a more robust data set to identify prognostic features, which is an important foundation for further clinical research.**

***Index Terms*—imaging mass cytometry (IMC), Hyperion, machine learning, data augmentation**

I. INTRODUCTION

Both tumour cells and the tumor microenvironemnt (TME) are important considerations when evaluating various cancers at the tissue level [1]. The TME refers to the complex network of blood and lymphatic vessels, extracellular matrix and various cells surrounding a tumor including stromal cells, such as fibroblasts and smooth muscle cells, and immune cells. During the early stages of tumor formation, transformed cancer cells can be eliminated by immune cells. However, with time, cancer cells evolve and are able to evade destruction by cells of the immune system, resulting in the formation of a tumor [2]. Thus, it is important to obtain a better understanding of the various cell types within the TME, since they may contribute to tumor progression. This can be achieved with the use of multiplexed tissue imaging systems such as the Hyperion^{TM} imaging system.

Hyperion^{TM} is an imaging mass cytometry (IMC) technology capable of high-throughput imaging of cells and proteins in situ. It produces multi-channel sub-cellular images per tissue sample, where each channel highlights the expression level of a given protein. Based on these expression levels, it is possible to identify a cell type in the sample. Unlike traditional tissue imaging methods, IMC enables detection of up to 37 protein markers at once [3]. This increased detection is an exciting improvement to typical imaging methods, since the fusion of this information will allow for evaluating rare, diverse immune cell types. Considering the sub-cellular spatial resolution of this imaging technique along with the number of available imaging channels, machine learning (ML) based approaches are considered to be a good candidate for analysis of these high-dimensional data. IMC data have been used in cancer analysis studies in the past [4], [5], however since it is a relatively new technology, these studies are limited and use regular statistical approaches rather than ML to estimate clinical outcomes [6]. Hence, there is a much needed urgency to leverage the quantitative nature of this technology.

ML-based analyses are generally *data hungry* in nature. With medical data specifically, it has been a longstanding challenge to acquire data from large numbers of subjects. While ML requires data points in the thousands, most patient cohorts are only in the tens. Therefore, proper data augmentation becomes crucial to enable appropriate learning from the data. In IMC images, the protein expressions are measured for each *individual cell* in the image. For each cell, the averaged pixel intensities within the cell corresponds to its protein expression. Therefore, normal image augmenting techniques like rotations and flips would not work. Though a rotation or flip would change the spatial location of the pixels in relation to the whole image, the pixel intensities within each cell would remain unchanged. Thus, a proper IMC data augmentation technique is necessary.

In this paper, we propose an end-to-end pipeline for ML-based analysis of IMC multi-channel data including a novel cell-level augmentation approach to increase data diversity, hence the generalizability of ML models. To demonstrate the utility of the proposed pipeline, we studied the association of the TME with various clinical parameters in a cohort of patients with muscle invasive bladder cancer (MIBC). The TME of MIBC contains a number of different immune cells, making it an ideal candidate for IMC analysis. We designed several experiments with signatures of IMC data and classification approaches and demonstrated that the combination of cell types found within the TME show promise as a foundation for future clinical research in this field.

This work was supported by Bladder Cancer Canada.

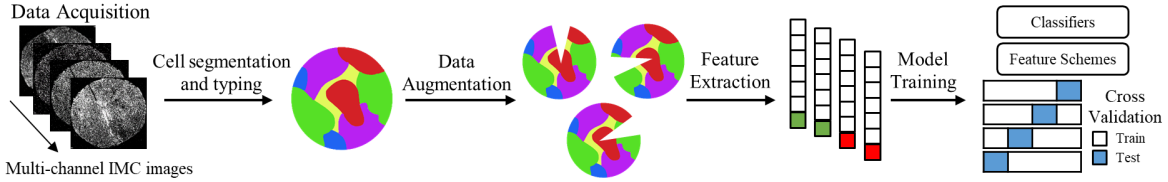

Fig. 1. Overview of methodology. Multi-channel images are acquired from tissue samples using IMC. Individual cells within samples are segmented and labelled as cell types. Data is augmented using the proposed sector elimination approach. Features are extracted based on cell proportion in the samples. The average performance in cancer recurrence estimation is evaluated for different feature subsets and classification models, using 4-fold cross validation.

II. MATERIALS AND METHODS

An overview of our proposed method is illustrated in Figure 1. After protein expression images are acquired, cells within the images are segmented and labelled with their corresponding cell types. The data are then augmented and the proportions of each cell type within the images are extracted as features for model training. Finally, the association of different clinical labels with the TME are estimated through the use of various models and features.

A. Data

We used 30 tissue samples collected from 15 MIBC patients who underwent cystectomy (removal of the bladder) at Kingston Health Sciences Center for this study. The study was approved by the Institutional Ethics Review Board. For each patient, 10 binary (yes/no) clinical parameters were also collected, including smoking, lymph node spread (LN), pathological lymphovascular invasion (LVI), carcinoma in situ (CIS), pathological margin, presence of tertiary lymphoid structures (TLS), recurrence, overall survival, neoadjuvant chemotherapy (NACT) prior to cystectomy, and prior treatment with Bacillus Calmette-Guérin (BCG) immunotherapy.

Tissue samples were biopsied, fixed on a slide, stained with 28 protein markers, and imaged using the Hyperion^{TM} system. For each tissue sample, a 28-channel image was generated using MCD Viewer (Fluidigm), where each channel represents a different protein marker. The pixel intensities of each channel correspond to the expression level of the given protein. The data was corrected for any signal spillover using the CATALYST R package. These high-dimensional data can then be analysed to identify the cell types within each sample, which will be explained in detail in *Data Pre-Processing*. Across all tissue samples, we defined 26 distinct cell types, such as CD8$^+$ T cells, CD163$^+$ macrophages, actin$^+$, and stromal cells. The list of all cell types is shown in Figure 4. The proportions of each cell type for a given sample were represented as a vector with 26 values and form the features that will be used for ML experiments.

B. Data Pre-Processing

To extract the protein expression of individual cells within the tissue images, the cells were first segmented. We used a watershed segmentation method to outline nuclei and create a mask. Objects either too small or large to be nuclei were excluded from the mask. Each object in the nucleus mask was then dilated and used as the cell mask. An expert annotator familiar with the presentation of cells visually verified the accuracy of the resulting cell masks. Given an image from a channel, the intensities of the pixels within each cell mask were averaged, representing the expression of the protein for that cell. Therefore, every cell was now represented by 28 protein expressions.

To annotate each cell with the mentioned 26 distinct types, we first used PhenoGraph clustering to identify groups of cells within a sample that have similar protein expressions [7]. With this, millions of cells can be represented within tens of clusters, making the manual cell annotation possible and convenient for a human expert. The protein expression data were log-transformed and normalized before clustering [7]. Next, manual annotation of cell types was performed based on the protein expressions of cells. After the cells were annotated with their types, the proportion of each cell type within a tissue sample was calculated and used as a feature for that sample, resulting in 26 features per each of the 30 samples.

C. Data Augmentation

As mentioned before, an important step in preparing data for ML analysis is ensuring there are enough data to properly train the model. We developed a novel cell-level spatial augmentation approach called *sector elimination*.

For each tissue sample, we suggest to exclude all cells within a randomly selected 30° spatial pie section of the images. Augmented partial images of the tissue were created as a result. The approach is illustrated in Figure 1. We performed augmentation prior to calculation of cell proportions in a tissue sample (i.e. our features extraction) and expected this approach to be more realistic than typical feature-level augmentation approaches. This process was randomly repeated 50 times for each of the 30 tissue samples. As a result, there are 50 augmented instances of each of the 30 tissue samples, generating a total of 1,500 augmented data samples. Using the cell type annotations, the proportion of cell types within each sample was then calculated as the feature vector (26 cell type proportions $\rightarrow$ 26 features).

D. Experiments

We used different classification models to evaluate TME association with various clinical labels, shown in Table 1. The 26 features representing the proportion of different cell types were ranked using an ANOVA F-score statistic. We

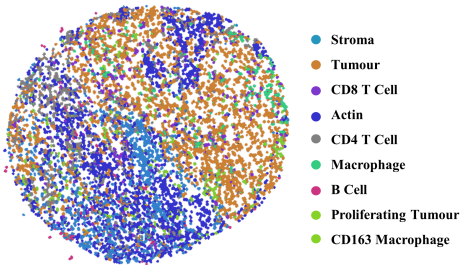

Fig. 2. Sample result of cell type identification for one of the tissues. Cell clusters are annotated by a human expert. The frequency distribution of each cell type determined from these result will generate the feature vectors for prediction models.

used logistic regression (LR), random forest (RF), decision tree (DT), k-nearest neighbour (KNN), and an ensemble of all four models, in a 4-fold cross validation configuration, to validate the performance of varying numbers of top ranked features. For each classifier, the default parameters were used. During fold generation, care was taken to ensure that all data samples from a single patient remained in the same fold. We also used Synthetic Minority Over-sampling Technique (SMOTE) to balance the data, such that each fold contained an equivalent amount of the label we were approximating. The cross validation for each classifier was repeated 50 times - each time with different, randomly generated folds - and the average accuracy of the classifiers on the test fold data in all runs was reported. Based on the results of the experiments, we decided on the the optimal number of ranked features, along with the top performing classifier, for each label. Additionally, we repeated the ML analyses using the original data without any augmentation to compare with our proposed augmentation approach.

III. RESULTS & DISCUSSION

The sample result of cell type identification through segmentation and clustering for one of the samples is illustrated in Figure 2. The cell mask in Figure 2 was coloured and labeled based on the cell types annotated by the expert. Before extracting the cell proportions, the tissue images were augmented as explained. Several augmented cell masks from a sample tissue are visualized in the top part of Figure 3.

To validate the proposed augmentation method, we projected features representing the proportion of cells that we extracted from the original tissue images along with those from their augmented partial versions to a 2D t-SNE space, shown in Figure 3. The plot is labeled and coloured by tissue samples and the features from original images are marked with "X". As can be seen in Figure 3B, the augmented data contains enough variability from the original tissue sample, while remaining similar enough to not compromise the strength of the labels. This lends itself well for the proposed augmentation approach to boost model training.

Although more patients would be preferred in order to draw any clinical conclusion with high significance, our augmentation method will at least allow for ML analysis on small patient

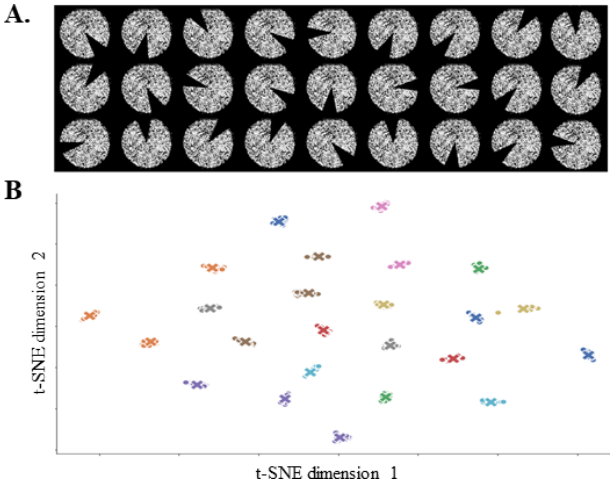

Fig. 3. Demonstration of the proposed Sector Elimination augmentation. A. Visualization of proposed random sector elimination approach for augmentation, showing the augmented partial images. B. The t-SNE plot of feature vectors of the generated dataset coloured by tissue samples, where "x" represents the original data and "○" represents the augmented data.

cohort data at some capacity. This is a notable step forward since until now, there has not been a standardized method of augmenting IMC data. Importantly, the proposed augmentation is not limited to the data or task used in this paper and can be applied to augment any kind of multi-channel sub-cellular imaging data for different types of cancers.

The result of ANOVA F-scores ranking of cell proportion features with respect to different clinical labels is visualized as a heat map in Figure 4. The heat map is grouped by the level of significance of features for the different labels. The scores are normalized per label for better visualization of their distributions. A higher score shows higher association of the specific cell types in the TME with the given label. Although the patient cohort is limited for direct clinical interpretation, there are still important observations that can build the foundation for future clinical research. The figure shows that B cells were relevant to the presence of TLS. Notably, TLS are primarily composed of B cells and thus, this association appears to have biological relevance [8]. In addition, $CD8^+$ T and $actin^+$ cells were recognized to be relevant to recurrence risk. In a past study, increased frequency of $CD8^+$ T cells at the invasive margin in MIBC was found to be associated with prolonged overall survival [1]. $Actin^+$ has shown to play a role in cancer progression through modulation of gene expression [9]. It is typically found mostly in muscle cells, including the thick muscular wall of the bladder, so these results are not entirely unexpected given that the definition of MIBC is tumor spread into the muscle layer of the bladder wall. Finally, we found that $CD163^+$ macrophages were relevant to prior BCG exposure. BCG is an immunotherapy given to patients with non-muscle invasive bladder cancer (NMIBC), often a precursor to MIBC. There is evidence that $CD163^+$ macrophages in NMIBC are associated with BCG failure [10], and it is notable that patients in our cohort with prior BCG

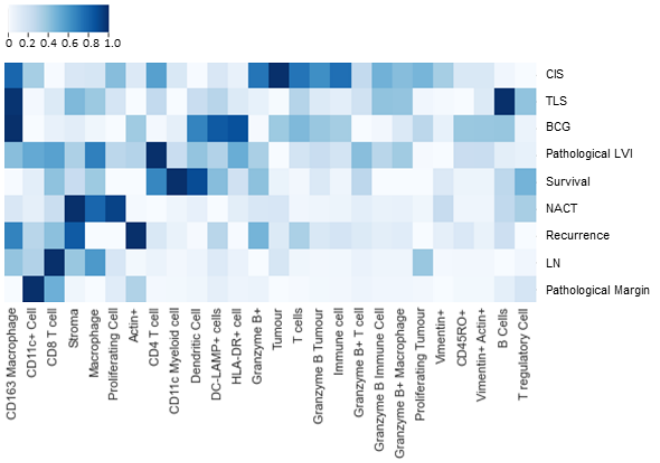

Fig. 4. Heat map depicting the significance (ANOVA F-scores) of the various cell proportions (x-axis) as features for the given clinical parameter (y-axis). Values are normalized between 0 and 1.

exposure for their NMIBC diagnosis experienced BCG failure since they progressed to MIBC.

The grid search results, to identify the optimal classifier and number of features as described in *Experiments*, is summarized in Table 1, shown as percentages. These results illustrate the potential for ML to be used in a predictive setting. In addition, we evaluated the performance of these models on the original data without any augmentation and overall, the performance was lower. For instance, the accuracy when estimating smoker status, LN, and NACT were 56.7±5.4%, 46.5±6.7%, and 54.4±4.5 respectively. For these labels, we could see that augmentation improved the model performance in a statistically significant way ($p<0.001$). For the other labels where performance was not improved significantly, we observed that standard deviation was much higher when the model was run without the augmented data. This indicates that without the augmented data, the model is more sensitive to the distribution of data within folds. Thus, our proposed augmentation seems to positively affects the training of the ML classifiers. This, in addition to the results in Figure 3, show promise for sector elimination augmentation to be further researched and eventually applied in larger clinical studies.

IV. CONCLUSIONS

IMC is a powerful imaging technology that can be used to provide an in-depth profile of numerous cell types in situ. We have introduced a robust ML pipeline for analysis of high-throughput IMC data, and validated the use of it through cystectomy tissues from a pilot cohort of MIBC patients. We utilized the proposed pipeline to demonstrate how the TME may be associated with different clinical parameters, which can be considered as an important basis for further clinical investigations. We also introduced a novel method for augmenting IMC data at the cell level, which is essential for any type of AI-related study considering the limitation in clinical data collection. The performance of ML classifiers when using our augmented data compared to only the original

TABLE I
ASSOCIATION OF TME WITH DIFFERENT CLINICAL FEATURES (%)

Label	Model	# Feat.	Accuracy	Sensitivity	Specificity
Smoker	KNN	17	64.1±3.5	75.2±5.2	52.7±4.9
LN	LR	6	75.5±7.4	69.2±13.1	81.8±3.6
LVI	KNN	11	64.5±7.3	64±12.3	64.9±7.5
TLS	LR	9	63.9±14.1	71.3±15.3	56.5±16.3
CIS	LR	1	64.3±4.2	61.2±6.7	67.4±3.5
BCG	DT	4	72.7±4.5	61.1±7.8	84.4±3.8
Margin	LR	2	80.4±1.6	75.2±7.9	85.5±8
NACT	DT	1	60.8±2.4	26.3±3.5	95.4±3.1
Survival	LR	3	65.3±3.6	85.5±5	45±6.2
Recurrence	KNN	5	78.2±5.3	80.6±5.5	75.8±8.4

data is significantly higher, which further emphasizes the importance of our proposed augmentation method and the value of developing it further. With a more robust data set, this pipeline has the potential to be used for clinical outcome prediction and biomarker identification. In the future, we hope to delve deeper into the analysis of rich IMC data with larger subsets of patients. This paper only scratches the surface of what is possible with Hyperion^{TM}, but it is a promising start. The large amount of data generated by Hyperion^{TM} is ideal for deep learning analyses, specifically, and is an exciting application for further studies with this MIBC data. Furthermore, we plan to investigate the use of cell-based protein expressions before converting them to cell type proportions of whole tissue.

REFERENCES

[1] A. Yu, et al. "Presence of lymphocytic infiltrate cytotoxic T lymphocyte CD3+, CD8+, and immunoscore as prognostic marker in patients after radical cystectomy," *PLoS One,* vol. 13, no. 10., October 2018.

[2] M. Wang, et al. "Role of tumor microenvironment in tumorigenesis," *Journal of Cancer,* vol. 8, no. 5, pp. 761, February 2017.

[3] Q. Chang, O. I. Ornatsky, I. Siddiqui, A. Loboda, V. I. Baranov, D. W. Hedley. "Imaging mass cytometry," *Cytometry part A,* vol. 91, no. 2, pp. 160-169, February 2017.

[4] S. Martinez-Morilla, et al. "Biomarker Discovery in Patients with Immunotherapy-Treated Melanoma with Imaging Mass Cytometry," *Clinical Cancer Research,* vol. 27, no. 7, pp. 1987-1996, April 2021.

[5] R. Elaldi, et al. "High Dimensional Imaging Mass Cytometry Panel to Visualize the Tumor Immune Microenvironment Contexture," *Frontiers in immunology,* vol. 12, pp. 1254, April 2021.

[6] Y. Zhu, et al. "SIO: A Spatioimageomics Pipeline to Identify Prognostic Biomarkers Associated with the Ovarian Tumor Microenvironment," *Cancers,* vol. 13, no. 8, pp. 1777, April 2021.

[7] J. H. Levine, et al. "Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis," *Cell,* vol. 162, no. 1, pp. 184-197, July 2015.

[8] M. Koti, et al. "Tertiary lymphoid structures associate with tumour stage in urothelial bladder cancer," *Bladder Cancer,* vol. 3, no. 4, pp. 259-267, October 2017.

[9] M. Izdebska, W. Zielińska, D. Grzanka, M. Gagat. "The role of actin dynamics and actin-binding proteins expression in epithelial-to-mesenchymal transition and its association with cancer progression and evaluation of possible therapeutic targets," *BioMed research international,* vol. 2018, January 2018.

[10] R. Pichler, J. Fritz, C. Zavadil, G. Schäfer, Z. Culig, A. Brunner. "Tumor-infiltrating immune cell subpopulations influence the oncologic outcome after intravesical Bacillus Calmette-Guérin therapy in bladder cancer," *Oncotarget,* vol. 7, no. 26, pp. 39916, June 2016.