Identification of rare cell phenotypes in Sarcoma H&E by a combination of deep learning and image processing approaches

Constantino Carlos Reyes-Aldasoro[1,2], Priya Lakshmi Narayanan[2,3], Tatiany L. Silveira[2,3], Tom Lund[2,3], Manuel Salto-Tellez [2,3,4]
1 City St George's, University of London
2 The Institute of Cancer Research, UK
3 The Royal Marsden Hospital, UK
4 Queen's University Belfast, Northern Ireland

Background

Rare cells, i.e., those with a lower abundance within a population, play an important part in the conditions of health and disease of an organism. In Sarcoma, the presence of these cells can reveal the state of immune response or angiogenesis among other conditions [1]. Thus, its identification is crucial and one way to identify rare cells is through its phenotype [2]. Pooling of rare cells from large whole slide images (WSI) for training is tedious and time-consuming (**Fig. 1**). Deep learning techniques [3, 4] have been widely used for cellular segmentation and identification. However, an important pre-requisite of deep learning mandates large number of manually annotated training data, which is not always available.

Traditional image processing algorithms can help finding rare cells, especially when there is limited training data to train deep learning models

Materials

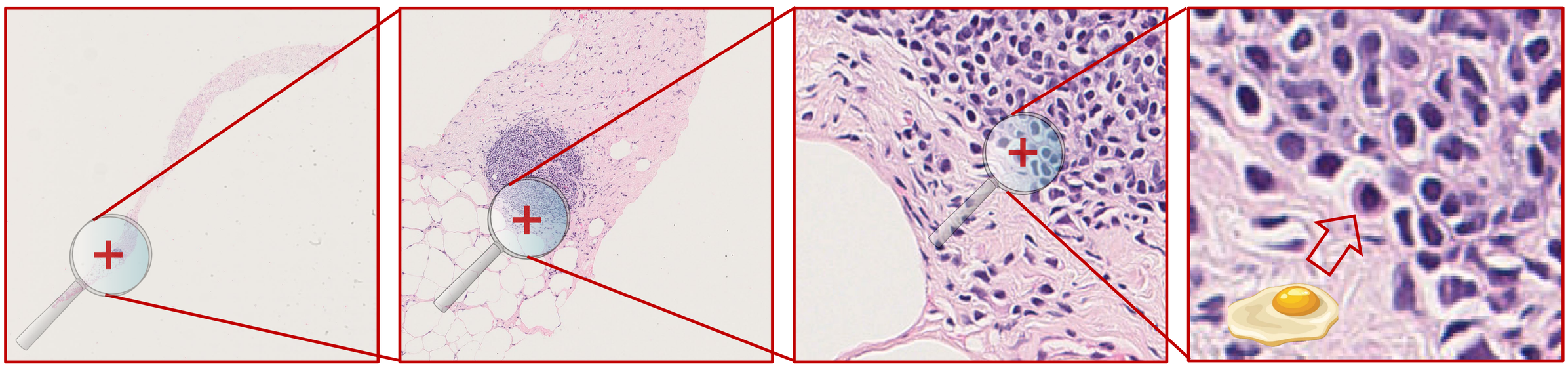

Fig 1. Whole slide images of a sarcoma stained with haematoxylin and eosin (H&E). Regions of interest illustrate the difficulty of finding one of the rare cells of interest. The chromatic characteristics of the cell; dark round nuclei surrounded by a lighter round cytoplasm (like a fried egg) can be exploited to refine results when there are limited samples to train deep learning models.

Methods

Unified segmentation and classification model based on m-DRDIN [5] was trained with *endothelial*, *mesenchymal*, *lymphocyte*, and *ignore phenotype*. *Ignore phenotype* constituted cells part of out of focus and artifact region such as tissue folding, and tissue edges with stained with dye colors.

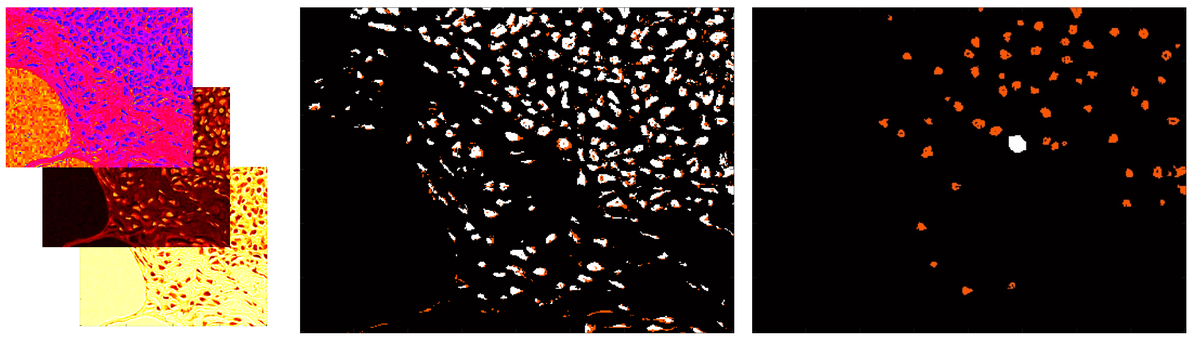

Fig 2. *Conversion of H&E image to HSV colour space, identification of "white and yolk" and classification of a rare cell.*

Classification of mesenchymal cells and lymphocytes was refined using image analysis algorithm based, which exploited the "fried egg" phenotype [6]. The algorithm identified nuclei (the *yolk*) and cytoplasm (the *white*) based on a combination the intensity of hue and saturation in the HSV space and morphological characteristics based on [7,8] to discard large or elongated cells (tumour cells) and those with little contrast between nuclei and cytoplasm (**Fig. 2**).

Results

To validate the output, an experienced pathologist (TLS) manually annotated 2212 cells in 12 whole slide images.
Recall varied between 0.75 and 0.96, depending on the "strictness" of the morphological characteristics (**Fig. 3**). A higher recall would imply a lower precision, which was not measured.

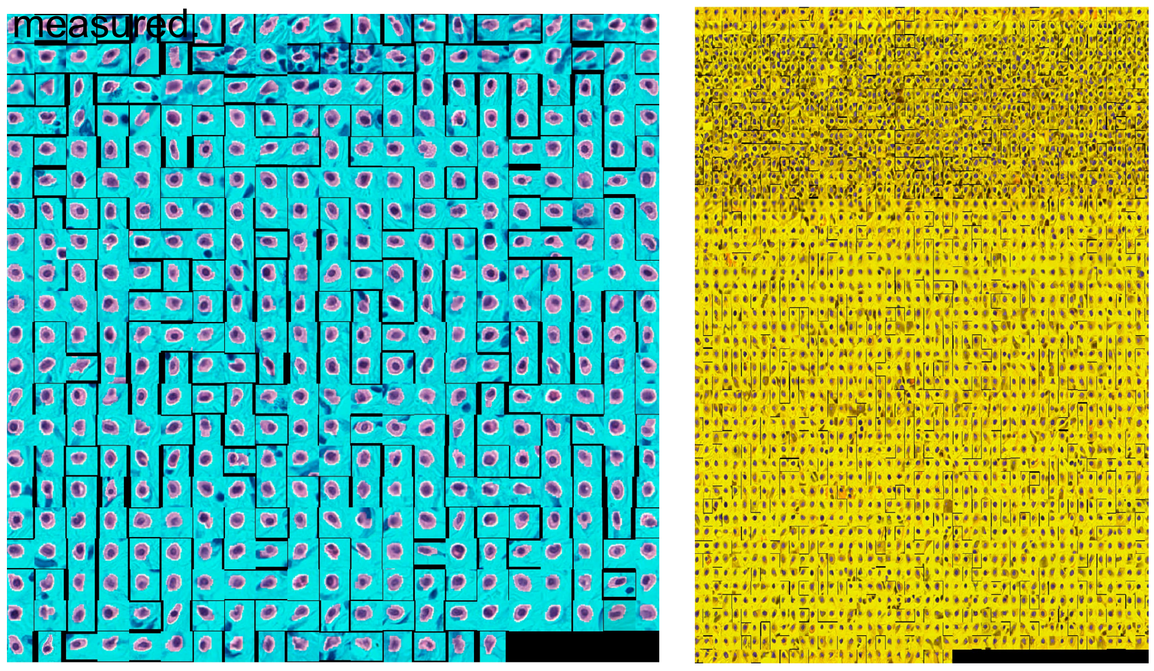

Fig 3. *Montages with all cells detected in the WSI. On the left, with blue background those that fit the phenotype. On the right with yellow background, all other cells.*

REFERENCES
1] Aashi Jindal, Prashant Gupta, Jayadeva, and Debarka Sengupta, "Discovery of rare cells from voluminous single cell expression data," Nature Communications, vol. 9, no. 1, Nov. 2018.
[2] Heba Z. Sailem, Julia E. Sero, and Chris Bakal, "Visualizing cellular imaging data using phenoplot," Nature Communications, vol. 6, no. 1, Jan. 2015.
[3] Kai Yao, Nash D. Rochman, and Sean X. Sun, "Cell type classification and unsupervised morphological phenotyping from low-resolution images using deep learning," Scientific Reports, vol. 9, no. 1, Sept. 2019.
[4] Alexander Hillsley, Javier E. Santos, and Adrianne M. Rosales, "A deep learning approach to identify and segment alpha-smooth muscle actin stress fiber positive cells," Scientific Reports, vol. 11, no. 1, Nov. 2021.
[5] Ruchika Verma and et al., "MoNuSAC2020: A multi-organ nuclei segmentation and classification challenge," IEEE Trans. Med. Imaging, vol. 40, no. 12, pp. 3413–3423, Dec. 2021.
[6] J Spergser, P DeSoye, W Ruppitsch, A Cabal Rosel, N Dinhopl, MP Szostak, I Loncaric, R Chopra-Dewasthaly, and HJ Busse, "Mycoplasma tauri sp. nov. isolated from the bovine genital tract," Syst App Microbiology, vol. 45, no. 1, pp. 126292, Jan. 2022
[7] CC Reyes-Aldasoro , L J Williams, S Akerman, C Kanthou, G M Tozer, An automatic algorithm for the segmentation and morphological analysis of microvessels in immunostained histological tumour sections J Microsc . 2011 Jun;242(3):262-78.
[8] CC Reyes-Aldasoro, , MA Björndahl, S Akerman, J Ibrahim, MK Griffiths, GM Tozer, Online chromatic and scale-space microvessel-tracing analysis for transmitted light optical images, Microvascular Research, V 84 (3), Nov 2012, pp 330-339

CONTACT
reyes@city.ac.uk