Detection and Characterisation of Fibronectin Structures in the Tumour Extracellular Matrix*

Faisal Jayousi, Xavier Descombes, Emmanuel Bouilhol, Anne Sudaka, Ellen van Obberghen-Schilling, Laure Blanc-Féraud

Detection and Characterisation of Fibronectin Structures in the Tumour Extracellular Matrix*

Faisal Jayousi
Université Côte d'Azur
CNRS, INRIA, I3S
Sophia Antipolis, France
faisal.jayousi@cnrs.fr

Xavier Descombes
Université Côte d'Azur
INRIA, CNRS, I3S
Sophia Antipolis, France
xavier.descombes@inria.fr

Emmanuel Bouilhol
Université Côte d'Azur
INSERM, CNRS, IBV
Nice, France
emmanuel.bouilhol@univ-cotedazur.fr

Anne Sudaka
Anatomopathology Department and CRB
Centre Antoine Lacassagne
Nice, France
anne.sudaka-bahadoran@nice.unicancer.fr

Ellen Van Obberghen-Schilling
Université Côte d'Azur
INSERM, CNRS, IBV
Nice, France
ellen.van-obberghen@univ-cotedazur.fr

Laure Blanc-Féraud
Université Côte d'Azur
CNRS, INRIA, I3S
Sophia Antipolis, France
blancf@i3s.unice.fr

***Abstract*—Characterising the tumour extracellular matrix (ECM) holds promise for identifying predictive biomarkers, particularly in assessing patient response to immunotherapy. While the cellular components of the tumour microenvironment have been extensively characterised, the non-cellular elements of this ecosystem remain underexplored. In this study, we investigated the geometry of Fibronectin (FN) in immunofluorescence images, a key ECM protein, in head and neck tumours. Our analysis identified two primary classes associated with FN structure: (a) aligned fibres, and (b) reticular fibre-like, which exist on a spectrum of structural variation. We proposed an approach leveraging Voronoi diagrams as adaptive windows. Circular statistics were then used to capture the spatial organisation of FN, providing insights into its structural heterogeneity within the tumour microenvironment.**

***Index Terms*—Fluorescence microscopy, Bioimage analysis, Circular statistics, Spatial tessellations, Extracellular matrix**

I. Introduction

The tumour microenvironment represents a complex and continuously evolving entity comprising cancer cells and a diverse array of immune cells, stromal cells, blood vessels, neurons and extracellular matrix (ECM). The ECM, consisting of an intricate meshwork of proteins and glycosaminoglycans, serves as the architectural scaffold of tissues and organs, providing a dynamic milieu for physical and biochemical signalling in cells. Throughout the course of tumour progression, the ECM undergoes pronounced topological and biophysical alterations, distinguishing it from normal tissue.

While numerous studies have focused on the cellular constituents of the tumour microenvironment and their spatial distribution, less attention has been paid to the non-cellular ECM components. In present study we investigate the geometry of Fibronectin (FN) in head and neck tumours. FN plays a crucial role in scaffolding collagen assembly, prior to being remodelled and replaced by dense collagen bundles associated to treatment resistance. Our ultimate goal is to gain insights into disease progression through a deeper understanding of the structural features of the ECM.

To do so, we analyse FN immunofluorescence images from head and neck cancer patients. The images are preprocessed to eliminate non-fibrillar objects. The associated graph representing the fibrillar components is then extracted, wherein the nodes serve as the generators of the corresponding Voronoi diagram. In Section III, we detail our proposed methodology for the detection and characterisation of FN fibres. We explain the rationale behind our approach. In Section IV, we present the experimental results and findings of our study.

II. Problem Statement

The data consist of immunofluorescence images of head and neck tumour tissue. Each image has dimensions of (1872×1404) pixels, with a resolution of 2.013 pixels per micrometer (px/µm). We focus our analyses on fluorescence images of stromal Fibronectin (FN). Two distinct forms of fibrillar FN are singled out based on the biological significance of fibrillar FN assemblies and their role in ECM maturation and remodelling [1]. Our primary aim is to develop an algorithmic pipeline capable of characterising fibrillar FN found in our images. In tumour tissue, these manifestations lie on a continuum between the following two states (Fig. 1):

1) Reticular fibre-like structures: Fibres forming a network-like arrangement.
2) Aligned fibres: Fibres exhibiting a certain orientation.

This continuum of configurations poses challenges in discriminating and segmenting them due to their often blended nature. This complexity renders patch-based approaches sub-optimal. Furthermore, patch-based methods depend on the *local neighbourhood*, where the optimal size and shape of the neighbourhoods vary according to the data.

* This work was supported by the French government through the 3IA Côte d'Azur Investments ANR-19-P3IA-0002 managed by the National Research Agency, Unicancer (TOPNIVO-ORL09) and the Fondation ARC-Unicancer (TOP'UP).

The aim of this work is to translate this continuum into a continuous metric that enables quantification of fibre alignment. The key discriminator between these arrangements is the diversity of local orientations. However, defining the term local poses an additional hurdle, as previously discussed. In this paper, we introduce a pipeline that detects FN fibres and that makes use of Voronoi triangulation to construct adaptive windows within which we compute a directional statistic that measures fibre alignment. Voronoi cells serve as adaptive local windows that dynamically adjust their shape and size according to the underlying data. This adaptability improves robustness, particularly in our case where traditional fixed-size windows are suboptimal.

In this study, four annotated images serve as a comparison baseline in section III, and we provide statistical justification for the existence of these fiber arrangements in section IV.

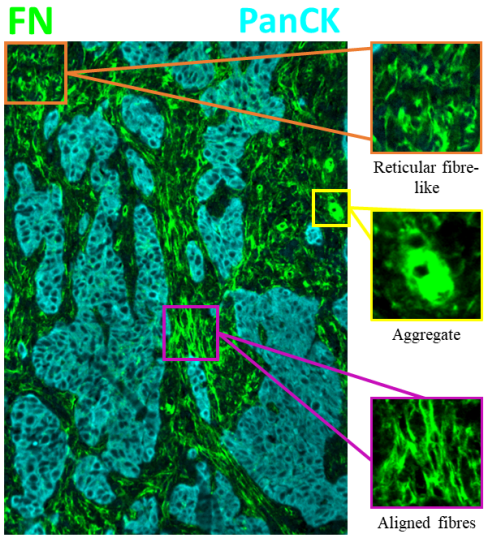


Fig. 1: Fluorescence microscopy image of FN distribution and tumour stained by PanCytokeratine (PanCK) in an invasive squamous cell carcinoma of the tongue. Resolution: 2.013 px/µm. Stained using Opal technology and acquired using the Vectra Polaris (Akoya Biosciences) Multispectral Imaging System.

III. Method

A. Preprocessing

Non-fibrillar FN objects, namely the aggregates (Fig. 1), need to be removed as part of a preprocessing step. Simple thresholding is often inadequate for their detection, as FN fibres can also exhibit significant luminosity. Nonetheless, these aggregates are distinguishable by their distinctive geometry, which closely resembles that of ellipses. In this section, we introduce a method designed to detect geometric objects within images that admit a representation through marks [2].

We consider grayscale images represented by $u \in \mathcal{U} \subset \mathbb{R}^{p\times q}$. A mark in $\mathbb{R}^d$ encodes geometric descriptors of the object of interest. Let $\mathscr{M}$ denote the set of marks. For ellipses, a mark $m = (a, b, \theta)$ represents the semi-major and semi-minor axes lengths and orientation, the support $\mathcal{P}_m$ of the ellipse is defined as $\{\boldsymbol{x} = (x, y) \in \mathbb{R}^2 : (\frac{x\cos\theta+y\sin\theta}{a})^2 + (\frac{x\sin\theta-y\cos\theta}{b})^2 \leq 1\}$.

Consider an arbitrary set $\mathscr{X}$ in $\mathbb{R}^2$ representing the positions of the centres of geometric objects. The set of (unordered) n-point configurations is formally defined as the set $\Omega_n = \{\omega = \{\omega_1, \ldots, \omega_n\} \in (\mathscr{X} \times \mathscr{M})^n, \omega_i = (\boldsymbol{x}_i, m_i), i = 1, \ldots, n\}$, where $\boldsymbol{x}_i \in \mathscr{X}$ is the position of the centre of the i-th geometric object, and m_i is the associated mark. The configuration set Ω is then defined as the union of all finite sets of configurations:

$$\Omega = \bigcup_{n\in\mathbb{N}} \Omega_n$$

For each $\omega \in \Omega$, we define a marked point process using the Gibbs density as follows:

$$\mathrm{d}\mathbb{P}(\omega) \propto \exp[-\mathcal{J}(\omega)]\mathrm{d}\mathbb{P}_0(\omega) \tag{1}$$

where $\mathbb{P}_0$ is the measure of a Poisson process and $\mathcal{J}(\cdot)$ is an energy function. Let

$$\Omega \ni \omega \mapsto \mathcal{J}(\omega) := \sum_{i=1}^{n} \mathcal{J}_D(\omega_i) + \sum_{i\sim j} \mathcal{J}_R(\omega_i, \omega_j) \tag{2}$$

where $\mathcal{J}_D$ is a data fidelity term and $\mathcal{J}_R$ is a regularisation term. Finding an optimal configuration implies finding the configuration with the highest probability given the data. Equivalently, this means minimising the energy function $\mathcal{J}$ over Ω. The data fidelity term $\mathcal{J}_D$ is taken as the convolution between the image u and kernels defined by the shapes of $\omega \in \Omega$. The regularisation term prevents object overlap by assigning an infinitely large value when the shapes defined by ω_i and ω_j intersect, and maintaining a value of zero otherwise.

It follows that

$$\mathcal{J}_D(\omega_i) = -\frac{1}{\#\xi_i} \sum_{(k,l)\in\xi_i} u(k,l) + \frac{1}{\#\partial\xi_i} \sum_{(k,l)\in\partial\xi_i} u(k,l) + \tau_{\min} \tag{3}$$

where $\#A$ is the cardinality of set A, and

$$\begin{cases} \xi_i := \mathcal{P}_{(a_i,b_i,\theta)} \cap \mathbb{Z}^2 \\ \partial\xi_i := \mathcal{P}_{(a_i+1,b_i+1,\theta)} \cap (\mathbb{R}^2 \setminus \mathcal{P}_{(a_i,b_i,\theta)}) \cap \mathbb{Z}^2 \end{cases}$$

The function $\mathcal{J}_D(\cdot)$ selects objects that exhibit a contrast surpassing $\tau_{\min}$ with their boundaries. In other words, any admissible configuration $\omega \in \Omega$ will be such that $\mathcal{J}_D(\omega) \leq 0$.

Traditionally, estimating the global optimum for the energy function involves incorporating simulated annealing schemes into a Reversible-jump Markov chain Monte Carlo (RJMCMC) or a Multiple Births and Deaths sampler [2]. This process requires careful tuning of annealing parameters, resulting in significant computational time. In this work, we introduce a more efficient approach—a two-step greedy algorithm (Algorithm 1)—for estimating a local optimum of the energy

function. During the first step, we choose a candidate for each pixel, opting for the shape that minimises the negative data term. Subsequently, we sort these candidates with respect to their energy and eliminate objects as needed to satisfy the non-overlap constraint (Fig. 3c).

Algorithm 1: Object Detection Algorithm

Data: Image $u \in [0,1]^{p\times q}$, Set of marks $\mathscr{M}$, $\tau_{\min}$
Result: Binary mask bm

Initialise empty arrays v, c, and bm, all of the same size as u;
foreach $(k,l) \in \{1,\dots,p\} \times \{1,\dots,q\}$ **do**
  $v(k,l) \leftarrow \operatorname{argmin}_i \mathcal{J}_D(\omega_i)$;
  if $\mathcal{J}_D(\omega_{v(k,l)}) \leq 0$ **then**
    $c(k,l) \leftarrow v(k,l)$;
foreach j *in* $\operatorname{argsort}(\boldsymbol{vec}(c))$; // in descending order
do
  $k_j \leftarrow \frac{1}{\#\xi_j}\mathbb{1}_{\xi_j} - \frac{1}{\#\partial\xi_j}\mathbb{1}_{\partial\xi_j}$;
  if $\max \boldsymbol{vec}(k_j + \mathrm{bm}) \leq 1$; // if disjoint
  then
    $\mathrm{bm} \leftarrow \mathrm{bm} + k_j$;
return bm;

B. Fibre detection & characterisation

We characterise the fibres with respect to their orientations. To do so, we use a bank of Gabor kernels. The Gabor kernel is a directional filter. It is a product of an elliptical Gaussian and a sinusoidal plane wave. It was first used to model receptive fields of simple cells in the visual cortex [3]. Mathematically, the 2D Gabor kernel is expressed as:

$$g_{(\lambda,\theta,\psi,\sigma,\gamma)}(x,y) = \exp\left[-\frac{x'^2+\gamma y'^2}{2\sigma^2}\right]\cos\left(2\pi\frac{x'}{\lambda}+\psi\right) \tag{4}$$

where $x' = x\cos\theta + y\sin\theta$ and $y' = -x\sin\theta + y\cos\theta$. The parameters controlling the kernel's properties are the wavelength of the sinusoidal factor λ, the orientation of the normal to the parallel strips θ, the phase offset ψ (set to zero hereafter), the standard deviation of the Gaussian envelope σ, and the spatial aspect ratio γ. Let $\mathscr{G}$ be a set of Gabor kernels with different parameters allowing for a versatile analysis of orientation and spatial features. The most prominent orientation information encoded by the various Gabor kernels in $\mathscr{G}$ at pixel (x,y) in the image u is given by (Fig. 3.b):

$$\tilde{u}(x,y) = \max_{g\in\mathscr{G}} g * u(x,y) \tag{5}$$

However, this operation is also prone to detecting edges, which is undesirable in the context of fibre detection. This issue is addressed by dividing $g \in \mathscr{G}$ into two overlapping halves, with the central Gaussian envelope in each half precisely aligned. Each convolution is subsequently replaced with two convolutions, and the *pixel-wise* minimum value is taken between them. Next, we proceed to construct the graph of the fibres and subsequently delineate Voronoi cells. The graph's edges are determined by thresholding the Gabor response and thinning the resulting image using Zhang-Suen's algorithm [4]. Nodes in the graph are then sequentially introduced, classified into two distinct types:

- End nodes: These nodes denote the terminal points of the skeleton, corresponding to pixels with only one neighboring pixel.
- Bifurcation points: These nodes are locations where the skeleton branches or merges.

The Voronoi diagram is then constructed using the set of nodes as generators. To quantify the degree of alignment among the fibers within these cells, we compute the local variance of the orientation $\boldsymbol{\theta}^*_{\mathcal{V}} \in \operatorname{argmax}_{g\in\mathscr{G}} g * u|_{\mathcal{V}}$ within each Voronoi cell. Given the cyclic nature of angular data, classical central tendency measures can be misleading by treating elements of the quotient set $\mathbb{R}/k\mathbb{Z}$, for $k > 0$, as distinct entities. Circular statistics adjust for this periodicity. We will focus on 2π-periodic data, for simplicity, noting that $\mathbb{R}/2\pi\mathbb{Z}$ is homeomorphic to $\mathbb{R}/k\mathbb{Z}$, allowing for equivalent analysis with different perdiocities by scaling angles. Let $\boldsymbol{\theta} = (\theta_1,\dots,\theta_d) \in [0,2\pi]^d$, the circular variance of $\boldsymbol{\theta}$ is given by

$$\mathbb{V}_{2\pi}[\boldsymbol{\theta}] = 1 - \rho \tag{6}$$

where $\rho = d^{-1}\sqrt{\left(\sum_{i=1}^d \cos\theta_i\right)^2 + \left(\sum_{i=1}^d \sin\theta_i\right)^2}$.

The lower the value of $\mathbb{V}_k[\cdot]$, the more concentrated the distribution. However, unlike ordinary variance, circular variance is bounded above by 1. Specifically, lower values of $\mathbb{V}_\pi[\cdot]$ indicate that the fibres in question are more aligned. To ensure that stronger Gabor responses contribute more to the statistic, the angles are weighted by their corresponding Gabor responses. If $\mathbf{G} = (G_1,\dots,G_d)$ is the Gabor response map, then

$$\rho_w = \frac{\sqrt{\left(\sum_{i=1}^d G_i\cos\theta_i\right)^2 + \left(\sum_{i=1}^d G_i\sin\theta_i\right)^2}}{\sum_{i=1}^d G_i} \tag{7}$$

IV. Results

Table II presents the results obtained from our aggregate detection using Leave-One-Out Cross-Validation, compared against two state-of-the-art methods. To validate our fibre segmentation approach, we generate fibre images along with their corresponding ground truth masks using Algorithm 2. We randomly perturb the vertices of a hexagonal lattice. Its edges are then randomly removed with probability p. We generate a realisation of the Potts model [5] to represent a finite number of states in the image and smooth the image accordingly using a family of Gaussian kernels to create varied textures. To replicate the variability in photon counts that occurs during image acquisition, Poisson noise is added. For the Potts model, we use $s = 4$ states and a coupling parameter of $\beta = 5$. We benchmark our method against the U-Net architecture trained for 100 epochs (Fig. 2). The results are reported in Table I.

Our method outperforms the benchmark in segmenting fibres and offers the additional benefit of interpretability without

Algorithm 2: Synthetic image generation

Data: Number of states s, Coupling constant $\beta > 0$, Vector of standard deviations for s Gaussian kernels $\sigma = (\sigma_1, \ldots, \sigma_s) > \mathbf{0}$, Deletion probability $p \in [0, 1]$, Maximal displacement $k \geq 0$

Initialise zero arrays $u, v \in \mathbb{R}^{p \times q}$;
Generate $\boldsymbol{P} \sim \text{Potts}(\beta, s)$;
foreach $x \in \{k, \ldots, p-k\}, y \in \{k, \ldots, q-k\}$ **do**
  if $x + y \in 2\mathbb{Z}$ **then**
    $(\Delta_x, \Delta_y) \sim \mathcal{U}(\{-k, \ldots, 0, \ldots, k\}^2)$;
    $u(x + \Delta_x, y + \Delta_y) \leftarrow 1$;
foreach $v \in \text{Delaunay}(\text{supp}\, u)$ **do**
  Remove v with probability p;
foreach $i \in \{1, \ldots, s\}$ **do**
  Construct Gaussian kernel g_i with mean 0 and variance σ_i^2;
  $u_i \leftarrow u * g_i$;
  $v \leftarrow v + u_i \mathbb{1}_{\{\boldsymbol{P}=i\}}$;
$w \leftarrow \text{Otsu}(v)$; // mask
Add Poisson noise to v;
return (v, w);

the need for labelled samples. However, effective use of the Gabor filter bank requires some prior knowledge of fibre thickness.

The pipeline overview is illustrated in Figure 3. Using the annotated FN images, we analyse the circular variance distribution relatively to the labels (aligned or reticular fibre-like). Two key observations warrant mention. Firstly, the Kolmogorov-Smirnov and Mann-Whitney U tests (Fig. 4) demonstrate significant differences in the circular variance values between each of the two classes. This finding supports the delineation of the two classes made by the experts. Secondly, while the annotated images assign regions to one of the two classes, many regions exhibit characteristics of both to varying degrees, leading to ambiguity. This presents a challenge in precisely validating the classification based on computed circular variance of orientations.

Method	Metric				
	Accuracy	F1	Jaccard	Recall	Precision
U-Net	0.85	0.72	0.56	0.56	**0.99**
std	±0.04	±0.02	±0.03	±0.03	±0.01
Ours	**0.93**	**0.88**	**0.79**	**0.84**	0.94
std	±0.02	±0.01	±0.01	±0.02	±0.01

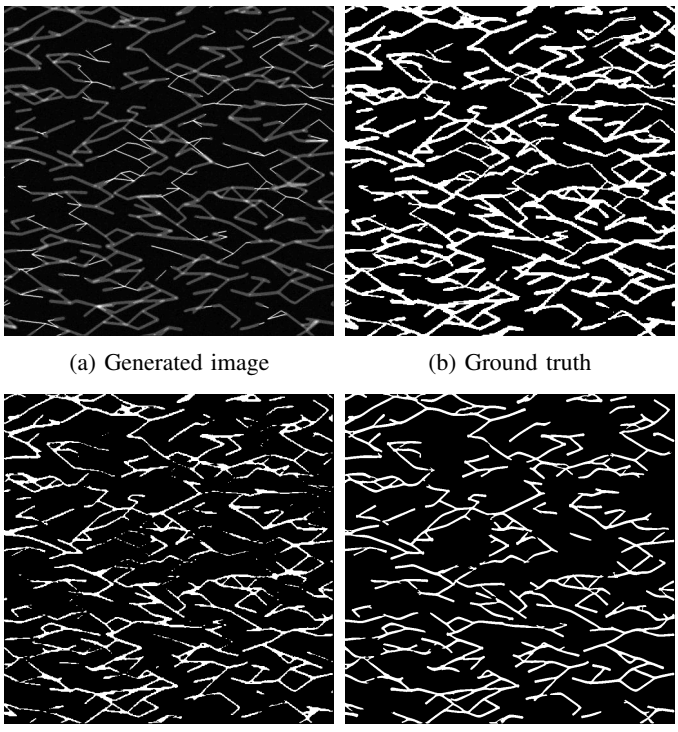
TABLE I: Performance metrics for fibre segmentation on synthetic images

V. Conclusion

In conclusion, we proposed an interpretable pipeline that characterises fibronectin (FN) structures within the tumour extracellular matrix (ECM) by utilising a bank of Gabor kernels to analyse their orientations. The Gabor kernel, originally designed to model receptive fields of simple cells in the visual cortex, proved effective in capturing directional information within the ECM fibres. By constructing a graph of the fibres and delineating Voronoi cells based on the graph's nodes, we were able to construct adaptive windows of different shapes and sizes that take into account the underlying data.

By utilising circular statistics, we quantified the alignment of FN fibres within Voronoi cells, accounting for the cyclic nature of angular data. Overall, our pipeline serves as a robust quantifier of the organisation and spatial arrangement of FN structures within the tumor ECM.

This work constitutes a first step towards identifying potential biomarkers in the tumour extracellular matrix (ECM) with predictive value for patient outcomes, including their response to immunotherapy. Ongoing studies will involve correlating geometric features of FN in the context of a clinical immunotherapy trial.



(a) Generated image (b) Ground truth

(c) U-Net output (d) Our method

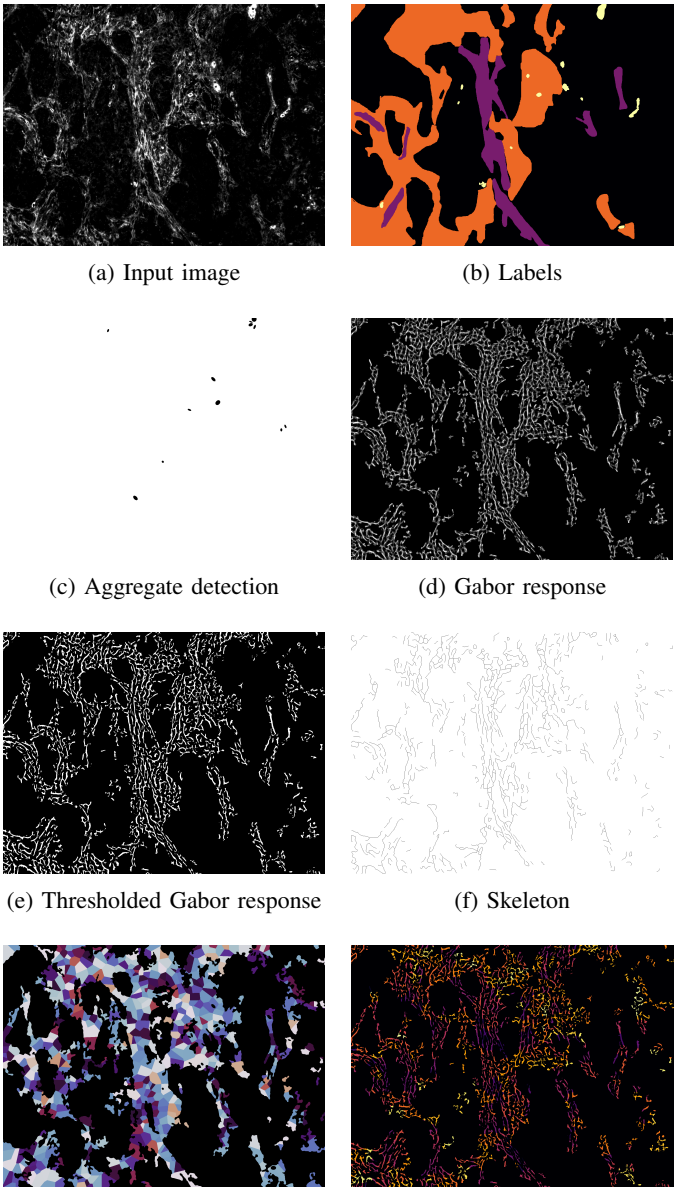
Fig. 2: Fibre segmentation comparison between U-Net trained for 100 epochs and our method.

Method	Metric	
	Precision	Recall
Stardist (Fluo)	0.40 ± 0.25	$\mathbf{0.65} \pm 0.12$
Cellpose (Cyto3)	0.13 ± 0.1	0.649 ± 0.13
Ours	$\mathbf{0.60} \pm 0.27$	0.43 ± 0.11

TABLE II: Performance statistics for aggregate detection using Leave-One-Out Cross-Validation

References

[1] G. Efthymiou, A. Saint, M. Ruff, Z. Rekad, D. Ciais, and E. Van Obberghen-Schilling, "Shaping up the tumor microenvironment with



(a) Input image (b) Labels

(c) Aggregate detection (d) Gabor response

(e) Thresholded Gabor response (f) Skeleton

(g) Voronoi cells (h) Final output

Fig. 3: Output of the image processing pipeline is illustrated as follows: (a) Input FN image. (b) Yellow: Aggregates, Purple: Aligned, Orange: Reticular fibre-like. (c) Detected aggregates represented by black objects. (d) Masked input image obtained using Eq. 5. (e) Thresholded image derived via Otsu's method. (f) Gabor response thinned using Zhang-Suen's algorithm. (g) Voronoi diagram generated using skeleton's nodes as generators. (h) Higher circular variance is reflected by a deeper shade of orange.

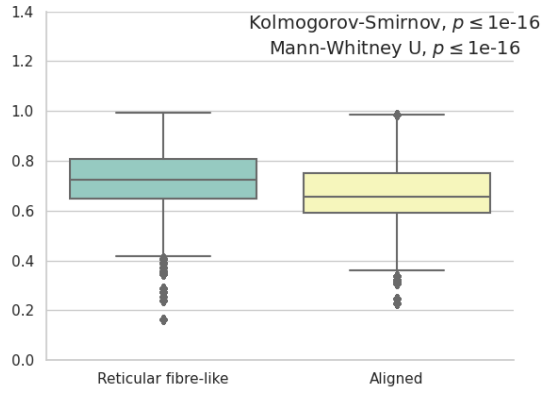


Fig. 4: Distribution of circular variance using annotated images: Kolmogorov-Smirnov and Mann-Whitney U tests reveal statistically significant difference between classes.

cellular fibronectin," *Frontiers in Oncology*, vol. 10, 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fonc.2020.00641

[2] X. Descombes, "Multiple objects detection in biological images using a marked point process framework," *Methods (San Diego, Calif.)*, vol. 115, pp. 2–8, 2017. [Online]. Available: https://doi.org/10.1016/j.ymeth.2016.09.009

[3] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Am. A*, vol. 2, no. 7, pp. 1160–1169, Jul 1985. [Online]. Available: https://opg.optica.org/josaa/abstract.cfm?URI=josaa-2-7-1160

[4] T. Y. Zhang and C. Y. Suen, "A fast parallel algorithm for thinning digital patterns," *Communications of the ACM*, vol. 27, no. 3, p. 236–239, Mar 1984.

[5] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, p. 721–741, Nov 1984.

[6] U. Schmidt, M. Weigert, C. Broaddus, and G. Myers, "Cell detection with star-convex polygons," in *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II*, 2018, pp. 265–273.

[7] M. Weigert, U. Schmidt, R. Haase, K. Sugawara, and G. Myers, "Star-convex polyhedra for 3d object detection and segmentation in microscopy," in *The IEEE Winter Conference on Applications of Computer Vision (WACV)*, March 2020.

[8] N. Petkov, "Biologically motivated computationally intensive approaches to image pattern recognition," *Future Generation Computer Systems*, vol. 11, no. 4, pp. 451–465, 1995, high Performance Computing and Networking. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0167739X9500015K

[9] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[10] C. Stringer and M. Pachitariu, "Cellpose3: one-click image restoration for improved cellular segmentation," *bioRxiv*, 2024. [Online]. Available: https://www.biorxiv.org/content/early/2024/02/12/2024.02.10.579780

[11] N. I. Fisher, *Statistical Analysis of Circular Data*. Cambridge University Press, 1993.

[12] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.