RESEARCH ARTICLE

ADVANCED
INTELLIGENT
SYSTEMS
Open Access

www.advintellsyst.com

Computer Vision for Substrate Detection in High-Throughput Biomaterial Screens Using Bright-Field Microscopy

Robert Owen,* Aishah Nasir, Mahetab H. Amer, Chenxue Nie, Xuan Xue, Laurence Burroughs, Chris Denning, Ricky D. Wildman, Faraz A. Khan, Morgan R. Alexander, and Felicity R. A. J. Rose

High-throughput screening (HTS) can be used when *ab initio* information is unavailable for rational design of new materials, generating data on properties such as chemistry and topography that control cell behavior. Biomaterial screens are typically fabricated as microarrays or "chips," seeded with the cell type of interest, then phenotyped using immunocytochemistry and high-content imaging, generating vast quantities of image data. Typically, analysis is only performed on fluorescent cell images as it is relatively simple to automate through intensity thresholding of cellular features. Automated analysis of bright-field images is rarely performed as it presents an automation challenge as segmentation thresholds that work in all images cannot be defined. This limits the biological insight as cell response cannot be correlated to specifics of the biomaterial feature (e.g., shape, size) as these features are not visible on fluorescence images. Computer Vision aims to digitize tasks humans do by sight, such as identify objects by their shape. Herein, two case studies demonstrate how open-source approaches, (region-based convolutional neural network and algorithmic [OpenCV]), can be integrated into cell-biomaterial HTS analysis to automate bright-field segmentation across thousands of images, allowing rapid, spatial definition of biomaterial features during cell analysis for the first time.

1. Introduction

Discovery of new biomaterials is essential for future healthcare approaches as they are integral to the creation of new tissue engineering and regenerative medicine strategies,[1] for the development of implants and medical devices,[2,3] and in the next generation of pharmacological delivery.[4] It has become increasingly recognized that the inherent material properties of biomaterials, such as their surface chemistry,[5,6] topography,[7–9] and mechanical properties,[10,11] offer structural, mechanical, and compositional signals that direct complex cell behaviors without the need for exogenous biological molecules.[12] However, there is insufficient understanding of the role that materials play in directing cell phenotype to develop novel biomaterials in a systematic, reliable, predictable manner. Screening of large libraries has therefore been employed to comprehensively cover the biomaterial

R. Owen, M. H. Amer, X. Xue, L. Burroughs, M. R. Alexander,
F. R. A. J. Rose
School of Pharmacy
Faculty of Science
University of Nottingham
Nottingham NG7 2RD, UK
E-mail: robert.owen@nottingham.ac.uk

R. Owen, A. Nasir, C. Denning, F. R. A. J. Rose
Biodiscovery Institute
University of Nottingham
Nottingham NG7 2RD, UK

A. Nasir, C. Denning
Division of Cancer & Stem Cells
Faculty of Medicine & Health Sciences
University of Nottingham
Nottingham NG7 2RD, UK

M. H. Amer
Division of Cell Matrix Biology and Regenerative Medicine
School of Biological Sciences
Faculty of Biology
Medicine and Health
University of Manchester
Manchester M13 9NT, UK

C. Nie
School of Computer Science
Faculty of Science
University of Nottingham
Nottingham NG8 1BB, UK

X. Xue
Department of Chemistry
Xi'an Jiaotong-Liverpool University
Suzhou Industrial Park 215123, China

design space and proved an effective means to identify novel materials and build comprehension.[13,14]

The advent of new manufacturing and computational capabilities has produced high-throughput screening (HTS) platforms for biomaterials to discover these properties.[15] These include systems to screen the effects of chemical (e.g., composition,[16] wettability[17]), and physical (e.g., topography,[18–20] 3D architecture,[21] stiffness[22]) properties on cell behavior (e.g., attachment, proliferation, differentiation) individually or in combination,[23] as recently reviewed by Yang et al.[1] These are powerful tools which allow thousands of conditions and experiments to be conducted simultaneously on a single platform at relatively low cost. HTS platforms for soluble drug discovery involves the assay of hundreds of individual conditions which are physically separated by microwell walls in a well plate, allowing individual conditions and cell populations to be assayed without cross-communication of cells and soluble factors.[24] In contrast, biomaterial HTS generally relies on substrates of interest being deposited into a single unit, such as a microarray or "chip," in a manner analogous to pathology tissue microarrays.[25]

The resulting lack of physical separation between substrates results in analysis of cell performance on biomaterial HTS arrays being largely confined to high-content imaging (HCI), the acquisition of multiple images at each location resulting in vast datasets.[26] These are typically bright-field images of the biomaterial location and multi-channel fluorescent images of the cells. User friendly, accessible, image analysis software programs such as CellProfiler,[27] ImageJ/FIJI[28,29] (open source), and Imaris (Oxford Instruments)[30] present a straightforward way to automate the assessment of cell response through the segmentation and quantification of the fluorescent images. These software packages utilize a model-based approach where fluorescent staining can be used to locate and segment cells on each substrate (e.g., by nuclei and cytoskeleton), and immunocytochemistry can be used to quantify the expression of protein(s) of interest on a per cell and per substrate basis by correlating with fluorescence intensity. These HCI approaches have facilitated the discovery of many new biomaterials, for example, chemistries for the expansion and differentiation of stem cells,[5,31] and topographies that enhance in vivo osseointegration[32] and maintain tenocyte phenotype.[33]

While these studies use established, robust tools for automatically segmenting and analyzing fluorescent cell images captured by HCI, none automate the extraction of data from bright-field images captured of the biomaterials they are studying. This is because the lower contrast bright-field images that are used to locate the material region of interest (ROI) (e.g., polymer spot, manufactured topography) present a much larger automation challenge compared to fluorescence segmentation. The model-based approaches used for fluorescent images do not translate to large bright-field datasets as it is not readily possible to define segmentation thresholds that work across all images. Therefore, unless done manually by the researcher (a laborious task), these data are often not used, which reduces the accuracy of the process and does not extract all available data. If bright-field segmentation could be automated across the vast image datasets generated by HTS, it would be possible to mask the automated fluorescent analysis to the biomaterial region, allowing more accurate characterization, e.g., normalization of cell markers to substrate surface area and determination of whether cells are on top or between topographical features.

Computer Vision is a field of artificial intelligence and machine learning that digitizes tasks humans do by sight, such as identifying objects by their shape.[34] Its popularity has surged in recent years across all disciplines that process large quantities of image data, from autonomous vehicles[35] to the healthcare sector in the automation of diagnosis from medical images.[36] By detecting biomaterial features by their overall visual appearance, a Computer Vision approach to segmentation of bright-field biomaterial images has the potential to overcome the current barriers to their inclusion in HTS analyses. Here, we present two Computer Vision approaches that permit automated bright-field segmentation in both chemical and topographical biomaterial screening arrays and demonstrate how they can be incorporated into HCI analyses to improve mechanistic biological insight and achieve significant time savings.

2. Results

2.1. Case Study 1: Computer Vision for High-Throughput Polymer Microarray Chemistry Screens

To demonstrate how Computer Vision can be used to automatically segment materials from high-throughput chemistry screens, here we use a polymer microarray screen of primary human human-bone-marrow-derived mesenchymal stromal cells (hMSCs) as an exemplar. Primary hMSCs from three donors were cultured on polymer microarrays formed from 300 commercially available monomers deposited in triplicate on each cell repellent poly(2-hydroxyethyl methacrylate) (polyHEMA)-coated glass slide (900 locations total, Figure 2). To quantify primary hMSC number on these materials at different timepoints, cells were fixed at days 5, 7, and 14 (9 arrays total), and nuclei (4′,6-diamidino-2-phenylindole [DAPI]) and cytoskeleton (α-tubulin) were fluorescently stained. Data was captured through HCI, with each polymer spot (8100 total) imaged using an automated fluorescence microscope, capturing two fluorescent images and a bright-field image at each location (24 300 images, **Figure 1**B).

2.1.1. The Challenge of Segmenting Polymer Microarray Spots

Variability in the bright-field images captured from the 8100 polymer spots is illustrated by the three spots presented in Figure 1C, making obvious the difficulties with edge threshold-based automated segmentation. Despite robotic deposition of polymers and imaging, tolerances in both these processes, different material chemistries, and the imaging process itself

R. D. Wildman
Centre for Additive Manufacturing
Faculty of Engineering
University of Nottingham
Nottingham NG8 1BB, UK

F. A. Khan
Digital and Technology Services (DTS)
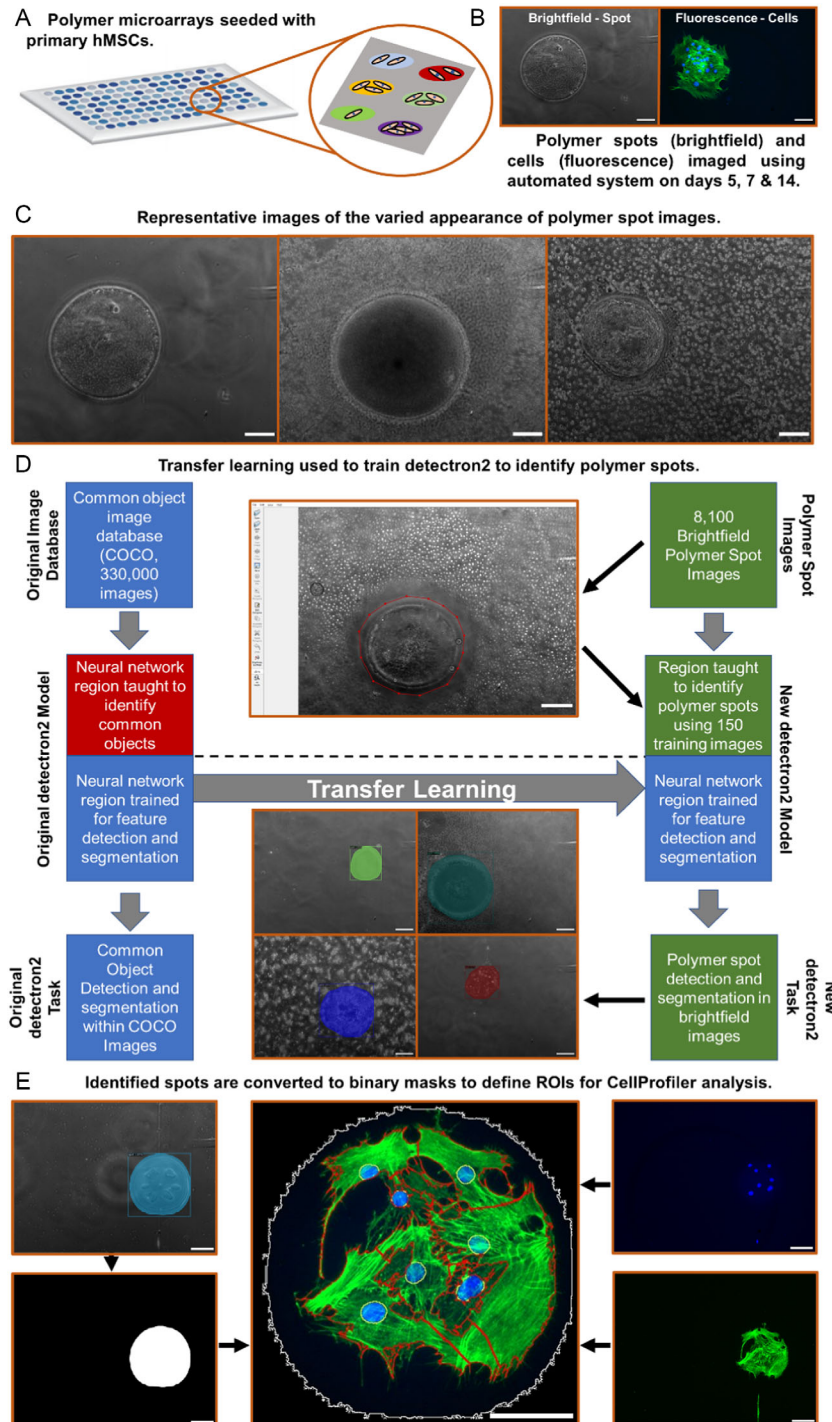University of Nottingham
Nottingham NG7 2RD, UK

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

Figure 1. A) Schematic of a polymer microarray containing 300 polymers in triplicate (900 spots per array), seeded with primary human mesenchymal stromal cells (3 donors, 3 time points = 9 arrays = 8100 spots). B) At each timepoint, every spot imaged in bright-field (polymer spot) and fluorescent channels (cells, blue: nuclei [DAPI], green: cytoskeleton [α-tubulin]). C) Representative variability in the brightness, contrast, and noise of bright-field images captured from the 8100 polymer spots. D) Flow chart showing how transfer learning modified the original *detectron2* model to identify polymer spots. The original model (left) has two parts: an underlying neural network trained to automatically detect and segment image features (blue), and a targeting region which identifies specific common objects (red). Through transfer learning, we retarget the neural network to identify features of interest (green) by manually labeling 150 bright-field images. E) The retargeted model can then detect and segment polymer spots in the entire 8100 image dataset. This can be binarized to create a mask to define an ROI for use in CellProfiler analyses. All scale bars—100 μm.

(e.g., differences in illumination) mean that in each image the polymer spots have different locations, size, morphology, brightness, contrast, and background noise. These combined make it impossible to set a singular intensity threshold that can be applied to the whole dataset that will robustly segment the spots. However, despite each image being unique, they all contain similar circular shapes. When viewing these images as humans, it is readily possible to identify the spot regardless of the other differences. Therefore, our hypothesis is that a Computer Vision that uses shape recognition approaches is likely to be well suited to this challenge.

2.1.2. Region-Based Convolutional Neural Networks Can Identify Polymer Spots via Transfer Learning

Region-based convolutional neural networks (R-CNNs) are machine learning models used for object detection within images.[37] While creating and training an R-CNN *ab initio* is a complicated undertaking, adapting an open-source, pretrained R-CNN is much more accessible (Figure 1D). Here, we use *detectron2* provided by FacebookAI.[38]

The original R-CNN was trained on the 'Common Image in Context' (COCO) dataset of ≈300 000 images of complex everyday scenes containing common objects in their natural context.[39] This developed a neural network capable of effective object segmentation. Through transfer learning, here we easily retargeted this neural network to detect polymer spots using a much smaller dataset. A total of 150 polymer spots (1.85% of the whole bright-field dataset) were manually outlined using LabelMe,[40] and this information was used to iteratively retrain the R-CNN to find a single polymer spot within each image. When all 8100 spot images were inputted into the retrained model, where present, polymer spots were identified regardless of image diversity or the number of cells present, as confirmed by manual inspection of a representative sample. By taking the output images from the R-CNN and converting them into a binary mask, an ROI could be incorporated into the CellProfiler analysis pipeline to delineate the area covered by the polymer spot (Figure 1E).

2.1.3. "Hit" Chemistries for the Attachment and Growth of Primary hMSCs

The final analysis of "hit" chemistries for hMSCs adherence and growth on the polymer spots incorporated the ability to determine spot presence and normalize cell area to spot area by ranking materials according to two metrics: the number of nuclei per 10 000 spot pixels and the percentage spot coverage by cytoskeleton over days of culture and per donor (**Figure 2**A,B). For materials that support hMSC attachment and growth, they will proliferate on these substrates increasing their coverage overtime. These data allow the expression of cell number per spot area and total coverage of the surface by cells, respectively, with the best performing substrates allowing cells to reach confluence (100% coverage) by the study endpoint. To rank materials, a score of 1 was awarded for each spot where the polymer exceeded thresholds of 1.5 nuclei per 10 000 spot pixels and 50% spot coverage (purple dashed lines) as this awarded ≈50% of polymers a

point at the first timepoint for each donor. The scores for each parameter were equally weighted, averaged, and then 0–1 normalized (0 min, 1 max) to give the final ranking and identify hits (score >0.75). By ranking materials by their ability to support both high cell numbers and high total cell coverage, a deeper insight into hMSC substrate preference is achieved than just considering nuclei number alone (Figure 2C–E). The final ranking of all materials is available in Supporting Information. In comparison to unmasked analyses, this allowed cell number to be expressed as a function of area to account for differences in spot size.

To demonstrate that more insightful analysis is possible by masking to the polymer spot, a comparison of the ranking by nuclei number to nuclei number per 10 000 spot pixels was made (Figure 2F). Although there is agreement in large parts of the dataset where cells are only present on the polymer spot and the polymer spot is average in size (black), the importance of taking into consideration the size and presence of a spot is revealed in the remainder of the dataset. Where polymer spots had delaminated (red) or overgrown the spot perimeter (green) they achieve an accurate, lower rank versus the false positive returned from the unmasked analysis. Conversely, where polymer spots are below average in size but supported high numbers of cells (blue), they are ranked higher when masked due to the normalization to surface area. Importantly for biomaterials discovery, some of these are the final "hit" materials and in unmasked analyses would have been discarded.

2.1.4. Neural Network Approaches Are Successful Across Image Acquisition Modalities

To ensure the transferability of the R-CNN segmentation approach to other chemistry screen datasets where a different image acquisition platform is used (Zeiss Axio Observer Z1; see methods), the methodology was applied to a disparate dataset of a polymer microarray screen analyzing the adherence of pseudo-virus particles (PPs). Two polymer microarrays (as used for the mammalian cell experiments: 1800 spots total) were used to screen PP attachment, each imaged in bright-field (polymer spot) and a fluorescent channel (attachment at 4 h of AlexaFluor-647 tagged pseudo-virus). Using the previously trained R-CNN generated from the primary hMSC dataset for the PP screen resulted in 31% of the microarray spots being successfully identified, as determined by manual inspection. Labeling a subset of images and retraining the R-CNN resulted in a model that identified all polymer spots across the whole dataset, demonstrating the transfer-learning process should be repeated for each image-acquisition method.

2.1.5. Neural Network Segmentation Is Significantly Faster and as Robust as Manual Segmentation

A comparison to manual segmentation was performed to calculate time-savings and validate the robustness of the automated approach. Manual segmentation of polymer spots and subsequent fluorescent intensity measurements on two microarrays using the polygon tool in ImageJ took two researchers 20 working hours: 40 s per polymer spot (**Figure 3**A), with each
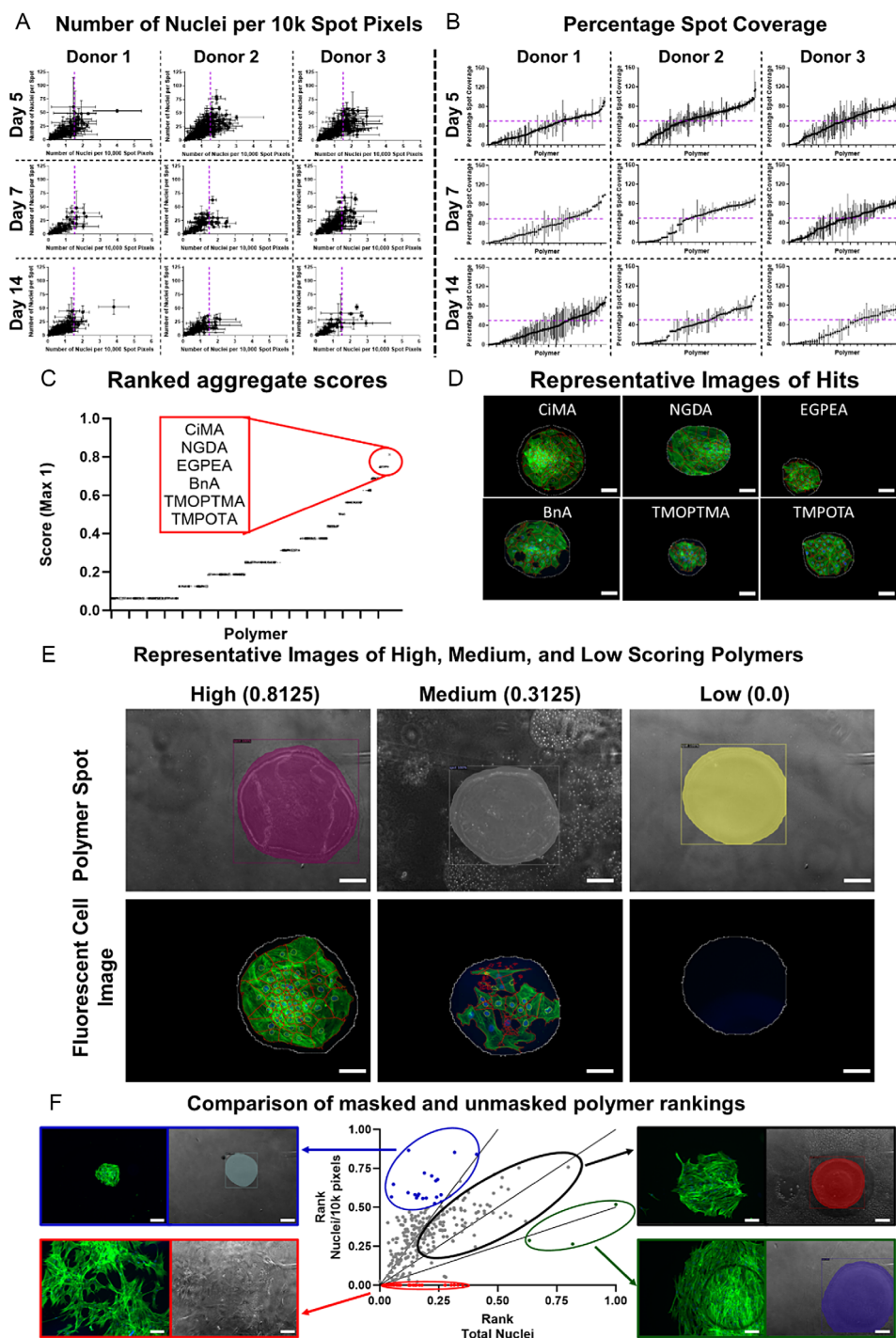
Figure 2. Overview of the ranking pipeline to identify hit materials for hMSC attachment and growth on HTS polymer microarrays. A) The relationship between mean number of nuclei and mean nuclei per 10 000 polymer spot pixels for each polymer spot on each microarray. Each point represents a triplicate of one polymer. Threshold for "hit": 1.5 nuclei per 10 000 spot pixels (purple dashed line). B) The high-low ranked percentage coverage of each polymer spot by cytoskeleton on each microarray. Each point represents one polymer. Threshold for "hit": 50% coverage (purple dashed line). All error bars $\pm$ SD. C) Final ranking and identification hits (0.75+) and D) corresponding images. E) Representative images of high, medium, and low scoring polymers showing both the segmented spots and CellProfiler output image. F) Comparison of material rankings by unmasked (total nuclei) and masked (nuclei/10 000 spot pixels). Both rankings 0–1 normalized with grey curves $y = 0.5x$, $y = x$, and $y = 2x$ for visualization purposes. Regions and difference to masked analyses: Black—agree, blue—underestimated, green—overestimated, and red—undetected false positives. All scale bars—100 µm.
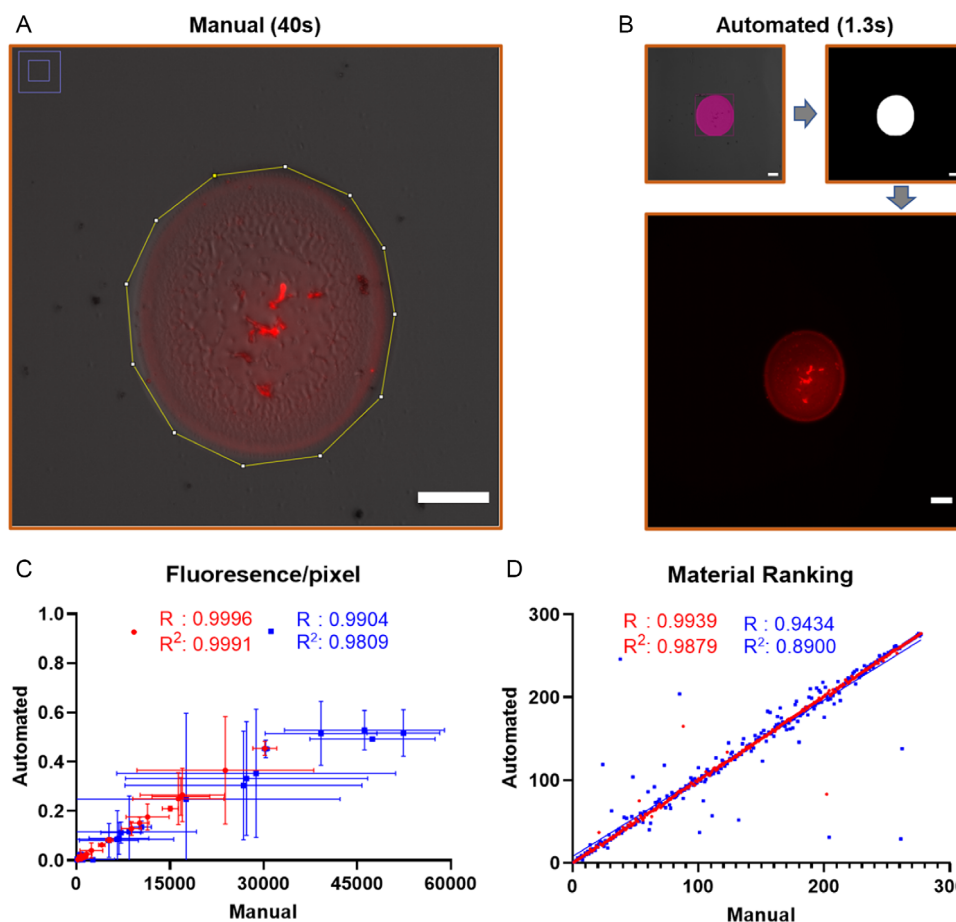
**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**

www.advintellsyst.com

Figure 3. Workflows of the A) manual and B) automated approaches of quantifying AlexaFluor-647 tagged pseudo-virus particle adherence (4 h) to HTS polymer microarrays. Manual analysis takes 40s per spot in ImageJ to define the polymer area using the polygon tool and measure the fluorescence intensity. Automated analysis averages 1.3s per spot, following retraining of the R-CNN identification of the polymer producing a binary mask for ROI analysis, and fluorescence intensity per pixel calculated within using CellProfiler. C) Raw fluorescence per pixel values shows positive correlation between automated and manual approaches across two arrays segmented by two users (R and R^2 values, red: array 1, blue: array 2) resulting in an almost perfect correlation in D) overall material rankings, indicating no loss of analysis quality through automation. All scale bars—100 µm.

researcher segmenting one microarray. After retraining the R-CNN (50 min), using a binary mask generated from the R-CNN output to define an ROI and calculating the total fluorescence within using CellProfiler, each array took 20 min to analyze: 1.3 s per spot (Figure 3B).

Time savings afforded by automated analysis are only valuable if the outputs are true. To validate them against manual processing, the correlation coefficients (R, R^2) for fluorescence per pixel values (Figure 3C) and the material ranking (Figure 3D) from the manual and automated analyses was calculated for the two arrays. For all four comparisons, R ranged from 0.94 to 0.99 and R^2 from 0.89 to 0.99, with three arrays exceeding 0.98, indicating strong positive correlation and validating the automated approach.

2.2. Case Study 2—Computer Vision for Topography Screens

To demonstrate how Computer Vision can be used to automatically segment materials from high-throughput topography screens, here we use a reference bright-field image dataset of

the TopoChip to develop a Computer Vision segmentation approach. The "TopoChip," a previously developed topography screening platform, contains a library of 2176 distinct, randomly designed surface topographies termed "TopoUnits." These are created using mathematical algorithms to combine three primitive shapes into topography "features" either 10, 20, or 28 µm in size (**Figure 4**A).[18] Once a computer vision approach was created using a reference TopoChip, this was validated using an experimental dataset to assess the spatial response of human-induced pluripotent-derived cardiomyocytes (hiPSC-CMs) to the topographies through staining of nuclei (DAPI) and cytoskeleton (cardiac troponin), assessing whether we can now segment cells by their location, e.g., on top of or in between topography features (Figure 4B).

2.2.1. The Challenge of Segmenting Images of Cells Cultured on Repetitive, Identical Topographies

Each image of a TopoUnit can contain up to 784 identical features with no variation, and these topographies are different
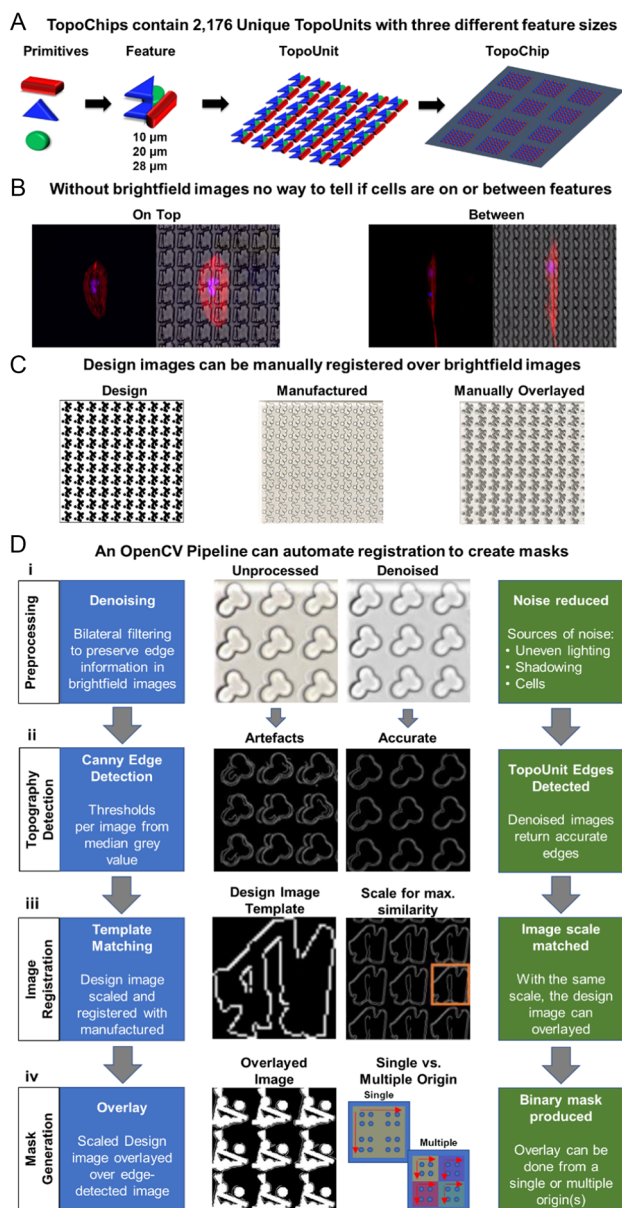
Figure 4. A) The TopoChip screening tool contains 2176 unique TopoUnits composed from topographical features formed from combinations of three primitive shapes. Features are in three sizes (10, 20–28 μm) resulting 100, 196, or 784 features per TopoUnit. B) Cells (hiPSC-CMs) may exhibit a spatial response to topography, residing either on top or in between features. This response can only be determined with the corresponding bright-field image. C) As each TopoUnit is designed, the original CAD image can be manually transformed and registered over images captured of the manufactured TopoUnit. D) Automating this process using Computer Vision allows us to create binary masks for subsequent cell analyses. D) i) First, bright-field manufactured images are denoised to improve accuracy of ii) *Canny edge detection*. iii) A template of the feature is then extracted from the design image and undergoes a series of transformation sweeps (rotate, flip, scale) until the process that achieves the maximum similarity with the edge detected image is identified. iv) This transformation is then applied to the design image and is then layered over the edge detected image either as a whole (single origin) or as segments from multiple origins.

at each of the 2176 locations in the dataset. Without variation, an R-CNN cannot be targeted to this application; therefore, a different Computer Vision approach is needed to automatically segment TopoUnit features.

To designate whether a region is a feature or space in the bright-field image requires prior knowledge of what is foreground (the topographical feature) and what is background (the space between topographical features). As the shapes of the topographies screened are designed (unlike the polymer microarray spots), this can be done with the input computer-aided design (CAD) file which specifies their manufacture. These input design images can be manually registered to overlay the output bright-field image (Figure 4C). Here, we use algorithmic Computer Vision to automate this process and generate binary masks that can be used to segment foreground and background.

2.2.2. Algorithmic Computer Vision Can Automate Feature Segmentation Within a TopoUnit

The algorithmic pipeline was written using an open-source Computer Vision library (OpenCV). First, images of each TopoUnit were preprocessed using bilateral filtering to denoise the images, eliminating artefacts that would otherwise be caused, e.g., by uneven lighting or the presence of cells in experimental conditions (Figure 4D). This allowed features to be accurately detected using *Canny edge detection*,[41] with thresholds calculated per image from the median grey value (Figure 4E). To account for differences in image size and resolution between design and output images, sweeps of different transformations (scale, rotation, flip) were performed on the design images to identify which caused the highest similarity with the edge detected output image using template matching (Figure 4F). With the necessary transformations elucidated, the design image can then be overlaid to create a binary mask of the TopoUnit feature locations. Overlaying from a single origin or multiple origins was compared (Figure 4G).

2.2.3. Time Savings from Algorithmic Computer Vision Make Topography Segmentation a Realistic Undertaking

Manual definition of the foreground and background in TopoUnit images took 8–35 min per image depending on the number/size of features that were present (100/28 μm—8 min, 196/20 μm 15 min, or 784/10 μm—35 mins). Due to the large time requirements, manual segmentation is currently done on an *ad hoc* basis downstream of automated cell segmentation. If manual segmentation of an entire TopoChip were undertaken, with 686, 748, and 742 of each feature number, respectively, across the 2176 TopoUnits per chip, it would take ≈29.5 days of continuous work (≈89 working days at 8 h day^{-1}). In contrast, the pipeline generates a mask every 1.2 s, meaning mask generation for a whole TopoChip takes only ≈70 min, which means that it is now realistic to incorporate automatic spatial segmentation of cell response to topography into cell analysis pipelines using this method.

2.2.4. Robustness of Automated OpenCV Segmentation

To compare single-origin- and multiple-origin-generated masks and validate their accuracy, regions identified as foreground in masks were segmented as objects using CellProfiler and overlaid over the original bright-field images (**Figure 5**A). These (150) were manually examined for each TopoUnit feature size (10/20/28 μm), determining whether the mask had accurately segmented foreground regions for each mask generation method by visual inspection. Masks were deemed unusable where segmented foreground regions did not perfectly overlay TopoUnit features (Figure 5B). This was most commonly caused by an error in the overlaying process (drift—single overlay only), an inability to match the extracted template (no match—both processes) or a failure to identify all feature coordinates (missing—multiple overlay only).

Comparing the two mask generation approaches, overlaying at multiple origins versus a single origin resulted in a higher percentage of useable masks for every TopoUnit feature size and was twice as successful in the 20 and 10 μm feature size TopoUnits where there are 2–8 times more features to segment per image (Figure 5C—left, 28 μm: 91% vs 83%, 20 μm: 85% vs 36%, 10 μm: 70% vs 36%). Despite being inferior to the multiple overlay approach overall, single overlay was able to create a usable mask of TopoUnits where multiple overlay was not 7.7% of the time, meaning if it was possible to automate the decision of whether a single- or multi-origin mask was better than the theoretical optimum, success rate would increase to 28 μm: 94%, 20 μm: 95%, and 10 μm: 80%.

By inverting the single-origin masks, "adding" them to the multiple origin masks to create an image of the difference between the masks and calculating the pixel intensity, differences between single- and multiple-origin masks can be quantified, with larger differences indicating a difference in what has been segmented as foreground. Visualizing these data (Figure 5C—right) reveals that where single origin outperforms multiple origin, the median difference exceeds 0. Therefore, automatically selecting single-origin masks if the calculated pixel intensity exceeds a threshold of 0.05 raises the success rate to 28 μm: 93%, 20 μm: 89%, and 10 μm: 70%. Note that the higher background noise in the 10 μm feature size results in a greater range of values when comparing masks meaning no threshold results in a greater success rate than multiple origin alone; therefore, for this feature size, a pure multiple origin approach is used. Overall, automating foreground segmentation offers significant benefit versus manual segmentation for the vast majority of TopoUnits and makes undertaking segmentation a realistic (in terms of time commitment) prospect.

2.2.5. Validation of Automated Topography Segmentation

To validate the ability of this OpenCV pipeline to automatically segment TopoUnits, this algorithmic approach was applied to a different, more challenging dataset of a TopoChip cell screen. Rather than a reference dataset of a "clean" TopoChip, this experimental dataset screens the spatial distribution of hiPSC-CMs on the topographies, increasing the bright-field image noise. A different automated microscope was used to capture images to validate the process across imaging systems, and a comparison to manual segmentation was performed to validate the robustness of the automated approach.
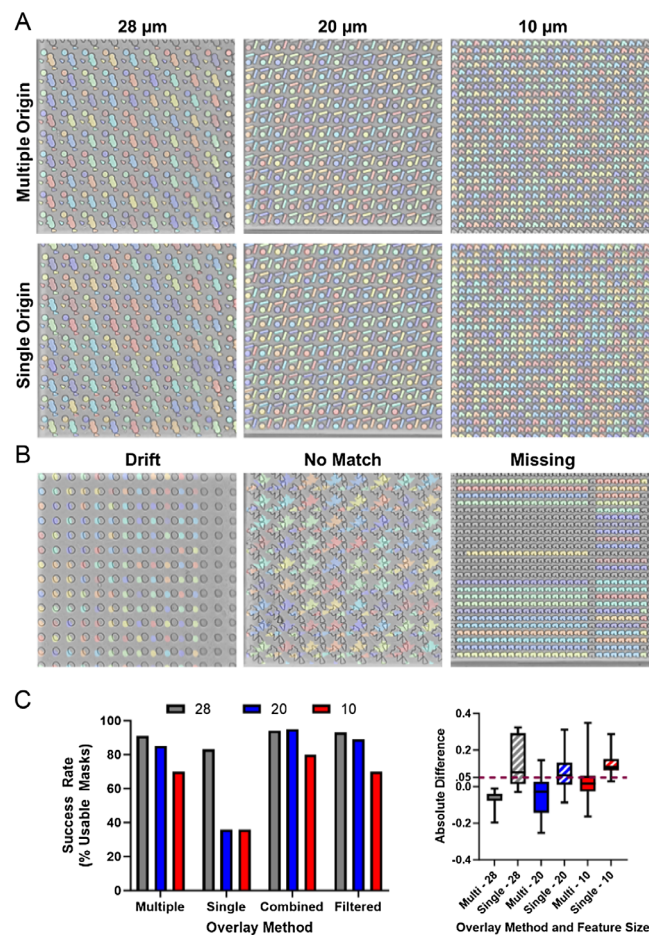

Figure 5. A) Representative images of successful TopoUnit segmentations using the multiple- and single-origin processes for each TopoUnit feature size. B) Where masks were unusable, it was predominantly due to drift (single origin only), inability to match the template (both), or failure to identify every feature coordinate (multiple origin only). For A&B, binary masks are false colored and overlaid over bright-field to visualize successful/failed segmentation. C) Overlaying from multiple origins outperformed single origin at every feature size overall, yet there were some TopoUnits where single origin was superior. By quantifying differences in masks generated by the two processes, a threshold (purple dashed line) could be set above which single-origin masks should be used, bringing the overall success rate closes to the theoretical combined maximum. D) Representative output of how masks from the OpenCV pipeline can be incorporated into CellProfiler analyses. The bright-field image of the TopoUnit (Manufactured) is processed through the pipeline to generate a mask. When this along with fluorescent images of hiPSC-CMs nuclei (blue—DAPI) and cytoskeleton (red—cardiac troponin) are used as inputs in CellProfiler cells can be automatically segmented and overlaid on the original bright-field image (Cells—all). By using the mask to define an ROI, segmentation can be limited to cells on top of features (Cells–top), or by inverting the mask, cells in between features (cells–between), allowing automated differential spatial analysis of cells on TopoChip screens.
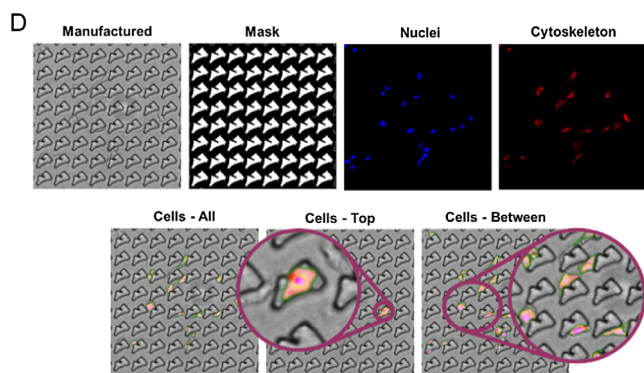
Figure 5. Continued.

2.2.6. Applicability of Algorithmic Computer Vision Between Datasets Using Mammalian Cells

Unmodified, the pipeline was unable to create masks from the new dataset due to the difference in bright-field image size and resolution between the reference and experimental dataset. However, by modifying the upper and lower limits on the transformation sweeps (Figure 4D-iii) performed before matching the similarity between the design and bright-field image, this was overcome with no other changes to the pipeline required, despite the increase in image noise from cells. As a result, binary masks could be generated using the pipeline for the experimental dataset (Figure 5D). When incorporated into a CellProfiler pipeline that detects nuclei and cytoskeleton, this allowed masking of fluorescent cell images that permitted automated differential spatial segmentation of cell location on TopoChips for the first time.

3. Discussion

This study demonstrates how automated, open-source Computer Vision approaches can be implemented into HTS of biomaterials to improve the insight from analyses by allowing spatial segmentation of cells according to the biomaterial features (chemistry or topography). Previously, analysis of cell response on these screening tools was largely confined to HCI of fluorescent cells only,[5,31–33] often disregarding bright-field biomaterial images due to the difficulties with automated segmentation of these image types. The limitation of an approach that only analyses fluorescent cell images is that it assumes substrates are either all identical (in the case of chemical screens) or that substates are manufactured in perfect accordance with the design file (in the case of topographical screens), and in both cases it gives no information about the spatial distribution of cells on the biomaterial substrate.

The HTS platforms used here (chemistry/topography) presented two disparate segmentation challenges requiring different Computer Vision approaches. For the chemistry screens, hundreds to thousands of materials as sessile drops of monomer solution are deposited, forming spots that have a similar overall morphology and a singular occurrence per image. Each location has a different substrate morphology due to the difference in material properties (e.g., hydrophobicity and/or viscosity

affecting spreading prior to polymerization). This means that cell number for a material is also affected by substrate surface area, but this cannot be accounted for without defining an ROI based on the substrate morphology. Furthermore, within each screen, there may be some locations where either the substrate has detached from the substrate surface leaving just the cell repellent coating, or the coating has also delaminated exposing the cell-permissive substrate underneath. In these scenarios, if substrate morphology is not accounted for, automated cell image analysis would give a false-negative or false-positive result, respectively, for the chemistry, requiring manual check of each image.

Segmenting 2D fluorescent cell images is relatively facile using pixel intensity thresholding within software packages such as CellProfiler and ImageJ, as excitation only occurs where a fluorophore is present. The challenge of automating segmentation of large bright-field datasets stems from the impossibility of defining a singular intensity threshold across the dataset for the boundary of the substrate as every pixel in every image has an intensity. Despite this, it is easy for a human to identify the substrate(s) within the image as we do so by their shape rather than their edge intensity. As Computer Vision aims to train computers to replicate tasks humans do by sight, it offers a promising solution within the field of image analysis. Within the latest versions of Imaris (9.6+), a popular commercial microscopy image analysis software package,[30] the ability to define foreground and background within an image set using machine learning has begun to be introduced. Here, they directly link to LABKIT, a FIJI[29] plugin that uses a random forest machine learning approach to classify pixels for foreground/background[42] rather than the object detection approach pursued here.

The variation in the similar morphologies of chemical screens makes them ideal for identification by R-CNNs as this can be used in the training of the neural network to teach the software exactly what morphology we are segmenting for. Interest in automating object detection has surged across all fields of research, with *detectron2* previously being applied to applications from improving detection of road damage[43] to assessing fall risk from posture.[44] Within biological research, use of R-CNNs has primarily focused on the segmentation of cells or subcellular organelles, for example in the LIVEcell project that uses deep learning to facilitate stain-free, live-cell segmentation,[45] or in histologically stained tissues sections in pathology.[46,47] Here, we demonstrated how across two disparate datasets with images acquired by different users operating different HCI systems, an R-CNN can be retrained to quickly and accurately segment a feature of interest from the entire bright-field dataset through manually labeling only 150 images. While the manual training and retraining through transfer learning has a time cost (≈50 min using a typical desktop computer for these examples), the benefits become clear almost immediately when applied to the remainder of the dataset due to the increase in segmentation speed. The trained R-CNN can also be used in all subsequent analyses where the spot images have been captured using the same acquisition method. With manual segmentation, other than operator familiarity with the protocol, previous analyses offer no opportunity to expedite subsequent analyses.

By incorporating this approach into an HTS of polymer chemistries, we were able to discover "hit" materials for human MSC

attachment and growth that would not have been discovered in unmasked analyses. By masking cell images to an ROI defined by the automated bright-field segmentation, cell number normalized to substrate area and percentage substrate coverage could be used as metrics allowing us to account for differences in substrate area and morphology for the first time in this type of screen and automatically exclude false negatives and false positives. Using a combination of the R-CNN for binary mask generation and a subsequent CellProfiler analysis pipeline to quantify fluorescence, analysis was over 30 times faster than manual processing with no compromise on robustness. When comparing approaches on a fluorescence per pixel basis, error bars are consistently equal in the x and y directions, indicating that where variation in fluorescent intensity occurred within a polymer triplicate on the microarray it was correctly detected by both approaches. Although greater concordance in overall material ranking between manual and automated approaches was observed in array 1 than array 2, this is likely due to greater experimental variability in array 2 rather than disagreement in the analysis as correlation coefficients remain very high (R: 0.94, R^2: 0.89). There is also approximately equal spread of outliers above and below the trendline in the material ranking, indicating that neither the manual nor the automated approach was biased to giving higher or lower rankings versus the other.

Whereas in the polymer microarray screen, the feature to be segmented occurs once per image and is similar in macro-morphology across the dataset, topography screens presented a different challenge. They are precisely manufactured using embossing or printing, resulting in many repeating, identical morphologies that need to be segmented in each image, but a unique macro-morphology in each image across the dataset. Each image of a TopoUnit can contain up to 784 identical features with no variation, and these topographies are dissimilar at each of the 2176 locations in the dataset. This uniqueness means R-CNNs are no more efficient at segmentation than manual approaches. Within a dataset, there is a singular image of each location, meaning targeting and identification have the same input image, and as such an R-CNN only detects what is labeled. Simply put, to train an R-CNN to identify every feature in every TopoUnit requires manually labeling every feature in every image—*de facto* manual segmentation.

To address this challenge, approaches available through the OpenCV library were investigated, a toolkit which has previously been used for purposes spanning automated detection of face-mask compliance[48] to understanding sign language.[49] Within the field of biomedicine, researchers have devised ways of integrating OpenCV's capabilities into ImageJ to streamline their combined capabilities in image analysis.[50] The edge detection approach pursued in OpenCV to overcome this problem is similar in principle to how fluorescent images are segmented; pixel intensity is used to define the edge of the segmentation target. However, only an outline is produced which is insufficient to designate foreground and background in convoluted TopoUnit images. The designed nature of each topography within the screen means that there are CAD files for each image location that depict what is foreground (topography) and what is background (space).

Here, we showed how a second Computer Vision approach can systematically transform these design files until they have a high degree of similarity with the edge-detected image and can subsequently be used as a mask that defines foreground/background. Although this algorithmic approach could not produce a usable mask for every TopoUnit, a success rate of up to 90% makes spatial segmentation of cells on topography screens a realistic proposition for the first time; where successful, it is ≈1000× faster than manual approaches, and where unsuccessful, it is possible that modification of the parameters used to extract, scale, and match templates could permit generation of masks for this subset in a semiautomated manner for the remaining TopoUnits, meaning that the laborious task of defining each topography is still avoided.

4. Summary

We have demonstrated how two different open-source, freely accessible, Computer Vision approaches can be used to automate feature segmentation from large bright-field datasets, and how this can integrate into existing fluorescent image analysis pipelines for high-throughput biomaterial chemistry and topography discovery. They offer significant time savings without compromising accuracy and allow us to perform more insightful data analysis. These approaches are not limited to biomaterial discovery datasets; depending on the nature of the bright-field feature(s) to be segmented, the two methodologies here can be used as a framework to automate bright-field image analysis in the wider field of biomedical research to identify objects in otherwise difficult-to-segment bright-field images where description of boundaries by pixel intensity is not practical, e.g., by rapidly retargeting an R-CNN on an ad hoc basis. This technology is now at a point where it is accessible and usable to researchers whose primary focus is not Computer Vision research. Researchers already accustomed to (or familiar with) HCI and downstream analyses are likely to possess the required basic familiarity with programming languages such as Python to immediately utilize these tools, although if not, such knowledge is available within most research institutes.

5. Experimental Section

All Materials Sourced from Sigma Aldrich/Merck Unless Otherwise Stated.

Polymer Microarray—Fabrication: Polymer microarrays were fabricated as described previously.[16] Briefly, epoxy-coated glass slides (Genetix) that had been dip coated in polyHEMA (4 w/v% in 95 v/v% ethanol/water) were used as substrates for all arrays. Monomer solutions (Sigma Aldrich, Scientific Polymers, and Polyscience) were formulated for all 300 materials by mixing at 50 v/v% with dimethylformamide and addition of a photo-initiator (2,2-dimethoxy-2-phenyl acetophenone) at 1 w/v%. These were spotted onto substrates using a XYZ3200 dispensing station (Biodot) and quilled metal pins (946MP3B, Arrayit) under an inert argon atmosphere (O2 < 2000 ppm, 25 °C, 35% humidity) polymerized with UV light. Three replicates of each of the 300 monomers (Figure S3, Supporting Information) were printed on each slide, yielding a total of 900 polymer spots per array. Validation of these polymer chemistries via time-of-flight secondary ion mass spectrometry has previously been demonstrated.[51]

Polymer Microarray—MSC Culture: Three donors of primary hMSCs were obtained from Lonza with certification of mycoplasma-free status and maintained in basal MSC growth medium (MSCGM) supplemented with 10% (v/v) fetal bovine serum, 2% (w/v) L-glutamine and 0.1% (w/v), gentamicin–amphotericin (#PT-3001; Lonza, Germany). For all

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

experiments, cells were used between the third and sixth passages and cultured as individual donor stocks.

In preparation for cell culture, polymer microarrays were sterilized by UV irradiation for 20 min before incubation in MSCGM for 1 h prior to seeding with 1×10^5 hMSCs per microarray.

Polymer Microarray—Virus Particle Adsorption: A PP adsorption assay modified from a previous study was used.[52] Briefly, AlexaFluor-647 tagged PPs or bovine serum albumin stock solution was resuspended in Milli-Q water or Dulbecco's buffered saline to obtain a $5 \, mg \, mL^{-1}$ of PP assay solution. The polymer microarray was repeatedly washed with 10 mL Milli-Q water and then placed immediately into 20 mL of assay solution, incubated, and rocked gently in the dark at room temperature (RT) under ambient conditions (20–22 °C, ≈50% RH) for 4 h. Then, the array slide was repeatedly washed again with Milli-Q water to remove poorly attached PPs.

Polymer Microarray—Imaging: For MSC screening, assessment of cell number was carried out 5, 7, and 14 days post-seeding by fixation with 4% paraformaldehyde for 20 min at RT. Cells were washed and permeabilized using 0.1% (v/v) Triton X-100 in phosphate buffer saline (PBS), and then stained with DAPI for nuclei (blue; Invitrogen) and α-tubulin as a cytoskeletal marker (green; #PA1-20988, Invitrogen). Microarrays were then imaged using an automated fluorescence microscope (IMSTAR), capturing each position in bright-field, and the two fluorescent channels.

For virus particle adsorption screening, microarrays were imaged using an automated fluorescence microscope (Zeiss Axio Observer Z1), capturing each position in bright-field and with a Cy5 filter.

Polymer Microarray—Feature Segmentation by R-CNN: Detectron2 was used for all R-CNN segmentation, with training annotations made using LabelMe. Both were downloaded from GitHub.[38,40]

A new model was needed for each image acquisition platform. In both cases, 150 images were taken from the whole dataset and converted from.-tiff to.jpg. This image series was opened in LabelMe and in each image the spot defined using the "Create Polygons" tool. Each spot within a dataset was given the same label, identifying it as the same object, and each label was saved as a.json file within the same directory as the training images. The provided *detectron2* model was retrained using these data for 3000 iterations on a local machine (CPU: Intel i7-7820HK, GPU: Nvidia GeForce 1080Ti, RAM: 32 GB). The python code used for training is shown in Section S1, Supporting Information.

To run the trained model, the model.pth file was placed in a base directory. Images to be analyzed were added to a subfolder "newimages." Using the python script in Section S2, Supporting Information, a "for" loop was used to iterate over all bright-field images and presented them to the trained *detectron2* model for segmentation. Identified polymer spots were outputted as false-colored images to an "outputs" folder, and binary masks where a pixel value of "1" represents the ROI (polymer spot), and a pixel value of "0" represents the background to a "masks" folder with an "_M" suffix. The filenames, mask names, area, and area ration were then exported as a csv file.

Polymer Microarray—Automated Cell Analysis: CellProfiler v2.2.0 was used for all polymer microarray analyses.[27]

For hMSC screens, three image sources were imported; binary spot masks from *detectron2*, fluorescent nuclei images, and fluorescent cytoskeleton images. Polymer spots were detected in the binary images and used to mask the nuclei and cytoskeleton images. Nuclei were then detected as primary objects using pixel threshold and size constraints, and associated cytoskeleton was identified as secondary objects propagating from these. Intensity of detected objects was measured and exported, along with images outlining the polymer spot and segmented cells to allow corroboration of outputs.

For virus particle screens, binary masks and fluorescent PP images were imported. Polymer spots were detected in the binary images and used to mask the PP images. Image intensity was measured in the masked images and the data outputted as a spreadsheet.

Polymer Microarray—Ranking Polymers for MSC Attachment and Growth: From the CellProfiler outputs, materials shown to support attachment of primary hMSCs were scored as followed for each array. On a per spot basis, the number of nuclei, total cytoskeleton area, and spot area was extracted. The nuclei per 10 000 spot pixels and percentage spot coverage

were then calculated and averaged across the triplicate. If delamination occurred (a spot size of 0 pixels) in more than two spots in a triplicate, the polymer on that array was excluded from the analysis; only duplicates and triplicates were considered. Nuclei density and percentage coverage graphs were visualized for each of the nine arrays and a scoring threshold established at 1.5 nuclei/10 000 pixels and 50% spot coverage. For each array where a polymer exceeded these thresholds, they were scored 1 point. To establish a final ranking that equally weights the two scoring metrics, the scores for each metric were averaged and min/max normalization between 0 and 1, retaining their relative weighting. These values were the final scores, with 1 being the maximum score. The full ranking is available in Section S3, Supporting Information.

Polymer Microarray—Validation of R-CNN Segmentation: Manual quantification of PP adherence was performed in ImageJ. Multichannel images containing the bright-field and fluorescent data were loaded as a stack and the edge of the polymer spot manually defined using the polygon selection tool. Fluorescence intensity per pixel was measured using the measure tool to record values for "area" and "integrated density" and exported to excel. Mean fluorescence per pixel values were calculated and averaged across the triplicate of polymers and expressed as mean ± standard deviation (SD). For the automated quantification, from the data outputted from CellProfiler, spot area and total fluorescence were extracted to calculate fluorescence per pixel and averaged across the triplicate, as per the manual approach.

Data was filtered so only duplicates and triplicates remained, then the raw fluorescence per pixel values for the manual and automated approaches were plotted against each other for each array to determine correlation. In addition to this, the outright rankings were also compared by high-low sorting of the outputs for the two approaches and assigning a number from 1 to 300, with 300 being the highest, and the correlation determined.

Activity time estimates were generated from file information, e.g., difference in timestamps between first and last files produced in a process.

TopoChip—Fabrication: TopoChips were fabricated as previously described.[18] A reference bright-field and design image dataset of each location was created and kindly provided by Jan de Boer.

TopoChip—Algorithmic Feature Segmentation: The "Open Computer Vision Library" (OpenCV)[53] was used for computer vision processes in combination with the "NumPy," "pandas," "os," and "re" libraries for the necessary mathematical operations, data analysis, operating system interfacing, and use of regular expressions, respectively.

For both single- and multiple-origin overlay approaches, initial steps were the same. Overlay script(s) (Single/Multiple_Origin_Overlay.py) and a reference spreadsheet "TopoChipFeatures.xlsx" were placed in the root directory, and design and bright-field images were placed in two subdirectories: "design" and "bf," respectively. When the script was run, anticipated feature size and spacing were extracted from the reference spreadsheet for each TopoUnit, bright-field images were denoised, design images were transformed (crop, rotate, flip), and *Canny edge detection* was applied to both image sets with the output cropped to the contours.

For single-origin overlay, a template of a single topography was extracted from the bright-field image and scaled to achieve the best match with a singular corresponding topography in the design image. This scaling was then applied to the design image, resized to the same image size as the bright-field image, and exported as a binary mask. For multiple origin overlays, the extracted template from the design image was scaled and matched to all occurrences of the topography in the edge-detected bright-field and the locations recorded. A cropped and scaled topography from the design image was then overlaid at each location and exported as a binary mask. Scripts for single and multiple overlay mask generation can be found in Section S4 and S5, Supporting Information, respectively. Time calculations for manual segmentation were made by manually segmenting each feature five times in ImageJ.

TopoChip—Validation of Algorithmic Segmentation: To determine whether masks were "usable," masks generated from the single and multiple overlay processes along with the original bright-field image were analyzed using CellProfiler v4.2.6. To eliminate noise, masks underwent

Gaussian filtering before identifying segmented foreground regions as primary objects. These were false-colored and overlaid over the original bright-field image. The first 150 of each feature size were then manually inspected and deemed usable or unusable to calculate the success rate. All image math calculations to determine filter thresholds for when single- or multiple-origin masks should be used were performed within the same pipeline.

TopoChip—hiPSC-CM Culture: The hiPSC cell line, ReBl-PAT, was derived, cultured, and differentiated to CMs as previously described.[54] CM medium (RPMI 1640 basal medium [Gibco, #21875034] with B27 supplement [LifeTechnologies, #0080085-SA], 10 μM Y-27632 [ROCKi, Tocris Bioscience, #0080085-SA], and 1% penicillin–streptomycin [P0781, Sigma-Aldrich]) was used for all experiments.

In preparation for cell culture, TopoChips were treated with oxygen plasma for 30 s (Zepto, Diener Electronic) and placed in a six-well tissue culture-treated polystyrene plate (Thermo Scientific Nunc). TopoChips were sterilized via submersion in 70% ethanol for 20 min followed by 15 min UV light exposure, washing three times in sterile Ca^{2+}/Mg^{2+}-free PBS (Gibco), and incubation overnight. Immediately before use, TopoChips were incubated in CM for 1 h at 37 °C prior to seeding with hiPSC-CMs (0.5×10^5/TopoChip).

TopoChip—hiPSC-CM Imaging: hiPSC-CMs (seeded at 15 days) were fixed after 15 days of culture (hiPSC-CMs ≈30 days) in 4% paraformaldehyde (Bio-Rad) for 15 min at RT and stained with anti-cardiac troponin I–1:100 (ab47003) as previously described.[54] Briefly, samples were washed thrice in 0.1% Tween-20 (Fisher Scientific) in PBS, permeabilized with 0.1% Triton-X (Sigma) in PBS for 15 min at RT, and incubated with 4% goat serum in 0.1% Triton-X (blocking solution) for 1 h at RT. Subsequently, primary antibody incubation was performed overnight at 4 °C in blocking solution and incubated with Alexa Fluor 647 (Life Technologies #A-21244) 1:400 followed by counterstaining with 0.5 μg mL^{-1} DAPI (Sigma #D9542) and HCS Cell Mask (1:500; H32714) for 30 min each. Samples were mounted with anti-fade medium Pro-Long Glass (ThermoFisher #P36984) and fluorescent images were acquired using a Zeiss Axio Observer Z1 microscope (Carl Zeiss, Germany) equipped with a Hamamatsu Flash 4.0 CMOS camera and motorized stage for automated acquisition using EC Plan-Neofluar 20×/0.50 Ph 2 to provide sufficient resolution while enabling the use of the autofocus function. Images were cropped to a smaller field of view (280 × 280 μm) that did not include the walls of the TopoUnits to improve the autofocus function and subsequent data analysis.

Captured images of the CM TopoChip were of different resolution and aspect ratio to the reference dataset. To account for this, the scale used for template matching (line 116 in multiple origin overlay script) was altered from (2.1, 2.3, 100) to (1.0, 2.0, 100) to allow a greater range of scales to be compared for matching the experimental image to the CAD image.

TopoChip—Automated Cell Analysis: To spatially segment cells, masks, fluorescent cell nuclei, and cytoskeleton images and the original bright-field images were analyzed using CellProfiler v2.2.0. Nuclei were detected as primary objects using pixel threshold and size constraints, and associated cytoskeleton was detected as secondary objects propagating from these. Segmented cells were overlaid over the original bright-field image either as detected, or after masking using as generated or inverted masks to visualize whether cells were present on topographical features or in the spaces between.

Statistics: All data visualizations and statistics were performed in GraphPad Prism 10.2.0. Correlation between manual and automated segmentation approaches was assessed via simple linear regression to determine correlation coefficients (R—Pearson correlation coefficient, R^2—coefficient of determination) with a two-tailed P value.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

[1] L. Yang, S. Pijuan-Galito, H. S. Rho, A. S. Vasilevich, A. D. Eren, L. Ge, P. Habibović, M. R. Alexander, J. De Boer, A. Carlier, P. Van Rijn, Q. Zhou, *Chem. Rev.* **2021**, *121*, 4561.

[2] J.-F. Dubern, A. L. Hook, A. M. Carabelli, C.-Y. Chang, C. A. Lewis-Lloyd, J. C. Luckett, L. Burroughs, A. A. Dundas, D. J. Humes, D. J. Irvine, M. R. Alexander, P. Williams, *Sci. Adv.* **2023**, *9*, eadd7474.

[3] J. Kohn, *Nat. Mater.* **2004**, *3*, 745.

[4] X. Han, A. Alu, H. Liu, Y. Shi, X. Wei, L. Cai, Y. Wei, *Bioact. Mater.* **2022**, *17*, 29.

[5] A. Nasir, J. Thorpe, L. Burroughs, J. Meurs, S. Pijuan-Galito, D. J. Irvine, M. R. Alexander, C. Denning, *Adv. Healthc Mater.* **2021**, *10*, e2001448.

[6] A. Latif, L. E. Fisher, A. A. Dundas, V. Cuzzucoli Crucitti, Z. Imir, K. Lawler, F. Pappalardo, B. W. Muir, R. Wildman, D. J. Irvine, M. R. Alexander, A. M. Ghaemmaghami, *Adv. Mater.* **2022**, e2208364, https://onlinelibrary.wiley.com/doi/10.1002/adma.202208364.

[7] M. H. Amer, M. Alvarez-Paino, J. Mclaren, F. Pappalardo, S. Trujillo, J. Q. Wong, S. Shrestha, S. Abdelrazig, L. A. Stevens, J. B. Lee, D.-H. Kim, C. González-García, D. Needham, M. Salmerón-Sánchez, K. M. Shakesheff, M. R. Alexander, C. Alexander, F. R. Rose, *Biomaterials* **2021**, *266*, 120450.

[8] S. Vermeulen, B. Van Puyvelde, L. Bengtsson Del Barrio, R. Almey, B. K. Van Der Veer, D. Deforce, M. Dhaenens, J. De Boer, *Adv. Sci.* **2023**, *10*, e2203880.

[9] M. J. Dalby, N. Gadegaard, R. O. C. Oreffo, *Nat. Mater.* **2014**, *13*, 558.

[10] R. Owen, C. Sherborne, T. Paterson, N. H. Green, G. C. Reilly, F. Claeyssens, *J. Mech. Behav. Biomed. Mater.* **2016**, *54*, 159.

[11] S.-B. Han, J.-K. Kim, G. Lee, D.-H. Kim, *Adv. Biosyst.* **2020**, *4*, e2000247.

[12] H. Donnelly, S. Vermeulen, M. Tsimbouri, M. J. Dalby, in *Tissue Engineering* (Eds: J. De Boer, C. A. Van Blittersswijk), Academic Press, Cambridge **2023**, pp. 261–292.

[13] Y. Mei, K. Saha, S. R. Bogatyrev, J. Yang, A. L. Hook, Z. I. Kalcioglu, S.-W. Cho, M. Mitalipova, N. Pyzocha, F. Rojas, K. J. Van Vliet, M. C. Davies, M. R. Alexander, R. Langer, R. Jaenisch, D. G. Anderson, *Nat. Mater.* **2010**, *9*, 768.

[14] D. G. Anderson, S. Levenberg, R. Langer, *Nat. Biotechnol.* **2004**, *22*, 863.

[15] J. Seo, J.-Y. Shin, J. Leijten, O. Jeon, G. Camci-Unal, A. D. Dikina, K. Brinegar, A. M. Ghaemmaghami, E. Alsberg, A. Khademhosseini, *Biomaterials* **2018**, *153*, 85.

[16] A. L. Hook, C.-Y. Chang, J. Yang, D. J. Scurr, R. Langer, D. G. Anderson, S. Atkinson, P. Williams, M. C. Davies, M. R. Alexander, *J. Vis. Exp.* **2012**, *2012*, e3636.

[17] M. S. Kim, G. Khang, H. B. Lee, *Prog. Polym. Sci.* **2008**, *33*, 138.

[18] H. V. Unadkat, M. Hulsman, K. Cornelissen, B. J. Papenburg, R. K. Truckenmüller, A. E. Carpenter, M. Wessling, G. F. Post, M. Uetz, M. J. T. Reinders, D. Stamatialis, C. A. Van Blittersswijk, J. De Boer, *Proc. Natl. Acad. Sci. U S A* **2011**, *108*, 16565.

[19] U. Tuvshindorj, V. Trouillet, A. Vasilevich, B. Koch, S. Vermeulen, A. Carlier, M. R. Alexander, S. Giselbrecht, R. Truckenmüller, J. De Boer, *Small* **2022**, *18*, e2105704.

[20] F. F. B. Hulshof, Y. Zhao, A. Vasilevich, N. R. M. Beijer, M. De Boer, B. J. Papenburg, C. Van Blittersswijk, D. Stamatialis, J. De Boer, *Acta Biomater.* **2017**, *62*, 188.

[21] M. Vassey, L. Ma, L. Kämmerling, C. Mbadugha, G. F. Trindade, G. P. Figueredo, F. Pappalardo, J. Hutchinson, R. Markus, S. Rajani, Q. Hu, D. A. Winkler, D. J. Irvine, R. Hague, A. M. Ghaemmaghami, R. Wildman, M. R. Alexander, *Matter* **2023**, *6*, 887.

[22] A. D. Rape, M. Zibinsky, N. Murthy, S. Kumar, *Nat. Commun.* **2015**, *6*, 8129.

[23] L. Burroughs, M. H. Amer, M. Vassey, B. Koch, G. P. Figueredo, B. Mukonoweshuro, P. Mikulskis, A. Vasilevich, S. Vermeulen, I. L. Dryden, D. A. Winkler, A. M. Ghaemmaghami, F. R. Rose, J. De Boer, M. R. Alexander, *Biomaterials* **2021**, *271*, 120740.

[24] V. Blay, B. Tolani, S. P. Ho, M. R. Arkin, *Drug Discov. Today* **2020**, *25*, 1807.

[25] N. M. T. Jawhar, *Ann. Saudi. Med.* **2009**, *29*, 123.

[26] W. Buchser, M. Collins, T. Garyantes, R. Guha, S. Haney, V. Lemmon, Z. Li, O. J. Trask, in *Assay Guidance Manual* (Eds: S. Markossian, A. Grossman, M. Arkin, D. Auld, C. Austin, J. Baell, K. Brimacombe, T. D. Y. Chung, N. P. Coussens, J. L. Dahlin, V. Devanarayan, T. L. Foley, M. Glicksman, K. Gorshkov, J. V. Haas, M. D. Hall, S. Hoare, J. Inglese, P. W. Iversen, Z. Li, J. R. Manro, J. McGee, O. McManus, M. Pearson, T. Riss, P. Saradjian, G. S. Sittampalam, M. Tarselli, O. J. Trask Jr., et al.), Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda, MD **2004**.

[27] D. R. Stirling, M. J. Swain-Bowden, A. M. Lucas, A. E. Carpenter, B. A. Cimini, A. Goodman, *BMC Bioinformatics* **2021**, *22*, 433.

[28] C. A. Schneider, W. S. Rasband, K. W. Eliceiri, *Nat. Methods* **2012**, *9*, 671.

[29] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, *Nat. Methods* **2012**, *9*, 676.

[30] O. Instruments, *Imaris Microscopy Image Analysis Software* **2023**, https://imaris.oxinst.com/ (accessed: October 2023).

[31] A. D. Celiz, J. G. W. Smith, A. K. Patel, A. L. Hook, D. Rajamohan, V. T. George, L. Flatt, M. J. Patel, V. C. Epa, T. Singh, R. Langer, D. G. Anderson, N. D. Allen, D. C. Hay, D. A. Winkler, D. A. Barrett, M. C. Davies, L. E. Young, C. Denning, M. R. Alexander, *Adv. Mater.* **2015**, *27*, 4006.

[32] F. F. B. Hulshof, B. Papenburg, A. Vasilevich, M. Hulsman, Y. Zhao, M. Levers, N. Fekete, M. De Boer, H. Yuan, S. Singh, N. Beijer, M.-A. Bray, D. J. Logan, M. Reinders, A. E. Carpenter, C. Van Blittersswijk, D. Stamatialis, J. De Boer, *Biomaterials* **2017**, *137*, 49.

[33] S. Vermeulen, A. Vasilevich, D. Tsiapalis, N. Roumans, P. Vroemen, N. R. M. Beijer, A. Dede Eren, D. Zeugolis, J. De Boer, *Acta Biomater* **2019**, *83*, 277.

[34] R. Szeliski, in *Computer Vision* (Ed: R. Szeliski), Springer International Publishing, Cham **2022**, pp. 1–26.

[35] J. Janai, F. Güney, A. Behl, A. Geiger, *Found. Trends. Comput. Graphics Vision* **2020**, *12*, 1.

[36] A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, Y. Liu, E. Topol, J. Dean, R. Socher, *NPJ Digit. Med.* **2021**, *4*, 5.

[37] R. Szeliski, in *Computer Vision* (Ed: R. Szeliski), Springer International Publishing, Cham **2022**, pp. 187–271.

[38] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, R. Girshick, *Detectron2* **2019**, https://github.com/facebookresearch/detectron2.

[39] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, P. Dollár, in *Computer Vision–ECCV 2014: 13th European Conf., Zurich, Switzerland, September 6-12, 2014, Proc., Part V 13*, Springer, Berlin **2014**.

[40] K. Wada, *Labelme: Image Polygonal Annotation with Python*.

[41] J. Canny, *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *PAMI-8*, 679.

[42] M. Arzt, J. Deschamps, C. Schmied, T. Pietzsch, D. Schmidt, P. Tomancak, R. Haase, F. Jug, *Front. Comput. Sci.* **2022**, *4*.

[43] V. Pham, C. Pham, T. Dang, in *2020 IEEE Int. Conf. on Big Data (Big Data)*, IEEE, Piscataway, NJ **2020**.

[44] R. Divya, J. D. Peter, *Complex Intell. Syst.* **2022**, *8*, 3021.

[45] C. Edlund, T. R. Jackson, N. Khalid, N. Bevan, T. Dale, A. Dengel, S. Ahmed, J. Trygg, R. Sjögren, *Nat. Methods* **2021**, *18*, 1038.

[46] A. V. Matias, A. Cerentini, L. A. B. Macarini, J. G. A. Amorim, F. P. Daltoé, A. Von Wangenheim, *SN Comput. Sci.* **2021**, *2*, 285.

[47] A. Felsen, Y. Yuan, N. Burzynski, D. Reitano, Z. Wang, K. A. Sethi, F. Lu, K. Chiu, in *2021 IEEE Int. Conf. on Big Data (Big Data)*, IEEE, Piscataway, NJ **2021**.

[48] H. Adusumalli, D. Kalyani, R. Krishna Sri, M. Pratapteja, P. V. R. D. Prasada Rao, in *2021 Third Int. Conf. on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Tirunelveli, India, February **2021**.

[49] A. Pradeep, M. Asrorov, M. Quronboyeva, in *2023 7th Int. Multi-Topic ICT Conf. (IMTIC)*, Jamshoro, Pakistan, May **2023**.

[50] C. Domínguez, J. Heras, V. Pascual, *Comput. Biol. Med.* **2017**, *84*, 189.

[51] A. L. Hook, D. J. Scurr, *Surf. Interface Anal.* **2016**, *48*, 226.

[52] A. J. Blok, P. Gurnani, A. Xenopoulos, L. Burroughs, J. Duncan, R. A. Urbanowicz, T. Tsoleridis, H. Müller-Kräuter, T. Strecker, J. K. Ball, C. Alexander, M. R. Alexander, *Biointerphases* **2020**, *15*, 061005.

[53] Itseez, *Open Source Computer Vision Library* **2023**, https://github.com/itseez/opencv.

[54] D. Mosqueira, I. Mannhardt, J. R. Bhagwan, K. Lis-Slimak, P. Katili, E. Scott, M. Hassan, M. Prondzynski, S. C. Harmer, A. Tinker, J. G. W. Smith, L. Carrier, P. M. Williams, D. Gaffney, T. Eschenhagen, A. Hansen, C. Denning, *Eur. Heart J.* **2018**, *39*, 3879.