Automated STED microscopy for cell-biological high-throughput assays

Dissertation

for the award of the degree

Doctor rerum naturalium

within the doctoral program

Biomolecules: Structure - Function - Dynamics

of the Georg-August University School of Science (GAUSS)

submitted by

Sebastian Bierbaum

born in Würzburg

Göttingen, 2022

Members of the Thesis Committee

Prof. Dr. Stefan Jakobs — Department of Neurology
University Medical Center Göttingen

Prof. Dr. Andreas Janshoff — Institute for Physical Chemistry,
Georg-August-University Göttingen

Prof. Dr. Stefan W. Hell — Department of NanoBiophotonics,
Max Planck Institute for Multidisciplinary
Sciences, Göttingen

Members of the Examination Board

Prof. Dr. Stefan Jakobs (1st Referee)

Prof. Dr. Andreas Janshoff (2nd Referee)

Further members of the Examination Board

Prof. Dr. Stefan W. Hell

Prof. Dr. Sarah Köster — Institute for X-Ray Physics,
Georg-August-University Göttingen

Apl. Prof. Dr. Alexander Egner — Institute for Nanophotonics Göttingen e.V.,
Göttingen

Dr. Helmut Berg — Unit Head Enabling Technologies,
Bayer AG, Leverkusen

Oral examination: October 17, 2022

Abstract

In the past decade, super-resolution fluorescence microscopy has revolutionized biological research by enabling the study of biological functions down to the molecular scale. Despite profound technological developments, the image acquisition process is, however, still a mainly manual task. Robust automated microscopy techniques and streamlined acquisition schemes will be key to increase the throughput of super-resolution methods to meet the requirements for industrial applications, e.g., in pharmaceutical drug screenings.

In this thesis, I established an automated stimulated emission depletion (STED) nanoscopy platform with special attention to the attainable imaging speeds. By implementing camera-based widefield imaging for both cell and feature detection, the throughput of the imaging platform could be considerably increased compared to an approach based solely on laser scanning. A use case from current pharmaceutical research addressing centrosomal clusters in cancer cells was successfully optimized for STED microscopy. Mitotic cells of interest and the centrosomes they contain were identified and localized by means of machine learning and classical image analyses. Extracted widefield feature coordinates were then reproducibly transformed into the confocal image coordinate system for subsequent STED image acquisitions.

The feasibility of the proposed acquisition scheme was demonstrated in first automated proof of concept measurements with ten potentially active ingredients. STED images of individual centrosomes revealed the centriole orientation and protein localization with superior accuracy. Automated STED microscopy thus offers the possibility to observe drug-mediated responses with unprecedented detail and can be a valuable additional decision tool for refined mode of action elucidation and target validation in future drug screening campaigns.

Contents

List of Abbreviations

2D	**Two-dimensional**
3D	**Three-dimensional**
ADME	**Absorption, Distribution, Metabolism and Excretion**
APD	**Avalanche Photodiode**
CA	**Centrosome Amplification**
CC	**Centrosomal Clustering**
CIN	**Centrosomal Instability**
CNN	**Convolutional Neural Network**
DM	**Deformable Mirror**
DMSO	**Dimethyl Sulfoxide**
DNA	**Desoxyribonucleic Acid**
DOG	**Difference of Gaussians**
FA	**Formaldehyde**
FN	**Field Number**
FOV	**Field of View**
FWHM	**Full Width at Half Maximum**
GA	**Glutaraldehyde**
HCS	**High-Content Screening**
LOG	**Laplacian of Gaussian**
MINFLUX	**Minimal Photon Fluxes**
ML	**Machine Learning**
MOA	**Mechanism of Action**
MTOC	**Microtubule-Organizing Center**
NA	**Numerical Aperture**
NGS	**Next-Generation Sequencing**
PALM	**Photo-Activated Localization Microscopy**
PCI	**Peripheral Component Interconnect**
PFA	**Paraformaldehyde**
POC	**Proof of Concept**
ROI	**Region of Interest**
SAC	**Spindle Assembly Checkpoint**
SAR	**Structure-Activity Relationship**
sCMOS	**Scientific Complementary Metal-Oxide-Semiconductor**
SIM	**Structured Illumination Microscopy**
SLM	**Spatial Light Modulator**

SMLM	**Single-Molecule Localization Microscopy**
SNR	**Signal-to-Noise Ratio**
SR	**Super-resolution**
STED	**Stimulated Emission Depletion**
STORM	**Stochastic Optical Reconstruction Microscopy**

Chapter 1

Introduction

Finding effective treatments for previously uncurable diseases is one of the main drivers for scientific innovations, both in academic research and life science industries. The development of groundbreaking high-throughput technologies like next-generation DNA sequencing (NGS) and thereof arising "-omics" approaches (e.g., genomics, transcriptomics, proteomics, metabolomics) permits the capture of previously unprecedented amounts of biological data (Hasin, Seldin, and Lusis, 2017). Beyond that, advances in microscopic imaging, mass spectrometry and magnetic resonance as well as their proceeding automation enables large-scale acquisition of more and more detailed and information-rich datasets today. Together with progress in the fields of computational biology and bioinformatics, these methodologies pave the way towards the ongoing transition of cell biology from a descriptive to a fully quantitative science (Mogilner, 2016).

Despite this steep rise in relevant data, the number of known molecular targets that can be addressed by drug treatments remains in the range of only about 700 human biomolecules, showing only a minor increase over the past 20 years (Drews, 2000; Santos et al., 2017). Although the amount of newly approved molecular entities per year has approximately doubled over the last decade (Mullard, 2022), the overall failure rate during clinical evaluations stays consistently high, with launch probabilities below 10 % for drug candidates entering clinical phase I studies (Dowden and Munro, 2019). Typical development periods of 10-15 years and extremely high costs during late-stage clinical trials are calling for efforts that focus on improved success rates and shorter cycle times (Paul et al., 2010). Accordingly, novel decision tools that help selecting the most promising compounds and target structures for subsequent optimization, already in the early development phase, are highly desirable.

To this end, the work presented in this thesis comprises theoretical considerations and the resulting practical implementation of an automated STED microscopy platform intended for cell-based drug screening. By revealing the cellular response to an applied drug candidate with an unprecedented level of detail, this approach has great potential to narrow down the pool of metabolically active compounds quite early in the screening process, leading to a more streamlined drug discovery pipeline.

Hit generation ▶   Lead validation/ ▶   Preclinical ▶   Clinical studies
                          optimization   development

Primary screening   Secondary screening   Animal studies   Phase I trials
(HTS)   Orthogonal assays   - ADME   Phase II trials
                          Animal disease models   - toxicity   Phase III trials
                          Complementary methods   - dosage levels

Hits ▶   Lead compounds ▶   Clinical candidate ▶   Approved drug

        Chemical   Experimental
        modification   testing

                    Lead profile ◀

                    - SAR
                    - Selectivity
                    - Physicochemistry
                    - Safety measures
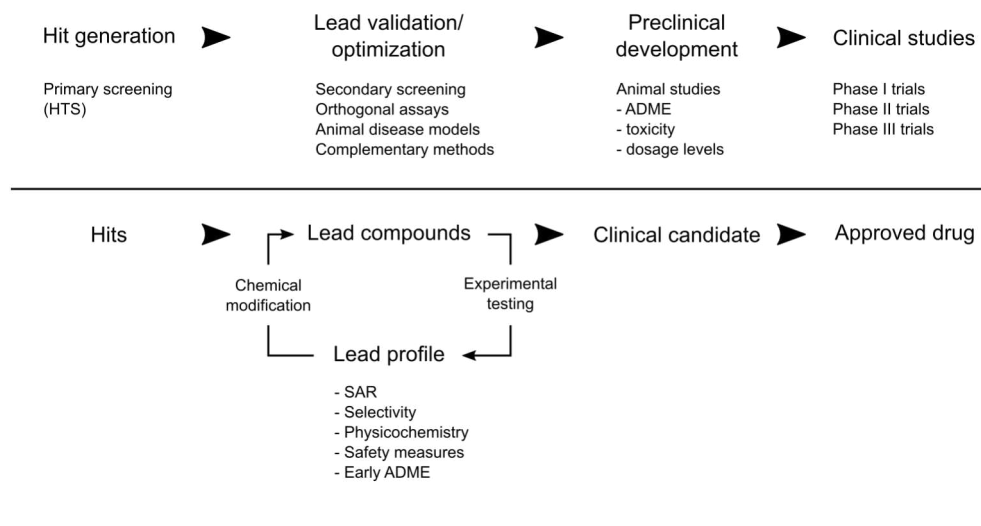   - Early ADME

Figure 1.1: Typical drug discovery workflow as implemented in industrial research and development departments. Major steps are hit generation, followed by lead validation and optimization, preclinical development and final evaluation in clinical studies. Comprehensive and multiparametric optimization in early phases helps minimizing attrition rates at late stages of the development process. HTS: high-throughput screening, ADME: absorption, distribution, metabolism, extraction; SAR: structure-activity relationship.

1.1 Principles and workflows of drug discovery

In pharmaceutical industries, drug discovery for a specific disease usually starts with a primary high-throughput screening, once an appropriate molecular target or disease model has been identified (Figure 1.1). Conduct of the assay is commonly performed in 384- or 1536-well microtiter plates and results in hits from the applied compound library that show assay activity above a certain threshold (Bleicher et al., 2003). These hits are then validated in secondary screenings using more sophisticated and orthogonal assays to identify several lead compounds for successive improvement. During lead optimization, chemical lead structures are continuously modified based on complementary experimental and computational techniques, to optimize their lead profiles. Iterative refinement of important molecular characteristics like the structure-activity-relationship (SAR), the target selectivity and specificity, and various physicochemical properties like solubility, polarity and lipophilicity increase the prospect for future success. In this development phase, also molecular safety measures like cytotoxicity and chemical stability as well as early absorption, distribution, metabolism and extraction (ADME) considerations are already taken into account. Afterwards, only the most promising lead compounds enter preclinical development to obtain full ADME and toxicity profiles and appropriate dosage levels in animal studies. As a result, solely a small number of compounds is selected as clinical candidates for final evaluation in clinical trials.

Although in practice the exact workflow depends on the actual use case at hand,

gathering valuable information already in the early discovery phase is key to avoid costly failures at later stages (Hughes et al., 2011). To this end, applied assays must be carefully selected and adequately validated to allow for accurate decisions on compounds before they enter the next phase. In general, two different assay types are routinely used at various stages of the development pipeline (Figure 1.2). Target-based assays rely on a known and well-validated molecular target that is intended to be selectively affected by a new molecular entity. Often, a protein of interest is purified for an *in vitro* assay or overexpressed in a genetically modified cell line. Proper target validation by linking the target itself with an underlying disease is therefore necessary already at the beginning of the development phase. Afterwards, suitable biochemical or cellular assays must be tested and optimized before performing the screening campaign. Typical compound libraries can comprise up to a few million molecules (Schamberger et al., 2011) and thus necessitate fully automated high-throughput screening capabilities. Detected hit compounds feature a significant impact on the specific molecular target, but only little to nothing is known about potential off-target effects and toxicity levels. The validation of hit or lead compounds and the elucidation of their presumed mechanism of action (MOA)
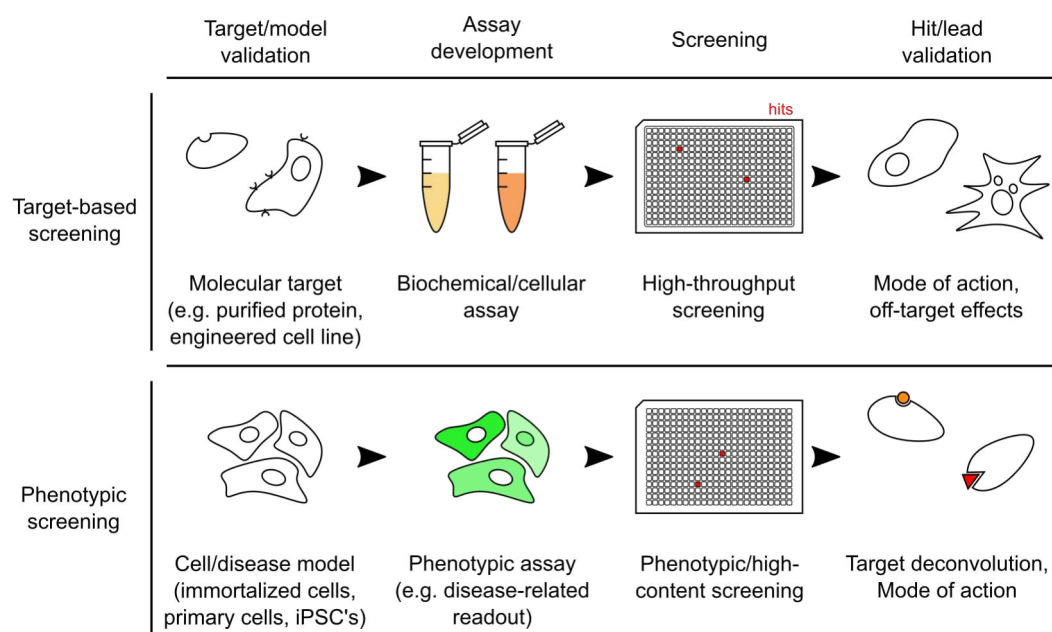
Figure 1.2: Screening workflows in drug discovery. Target-based screening approaches address a validated cellular target, either in purified form for biochemical or overexpressed in cellular assays. Single-parameter readout allows for testing large compound libraries in high-throughput screens. In phenotypic assays, a model system specific for a particular disease is established and then screened for morphological changes or cellular perturbations in response to applied drug candidates. Multiparametric readout on the single-cell level enables profiling, especially for high-content imaging approaches.

hence requires more elaborate cellular assays in subsequent downstream activities.

In alternative phenotypic screenings, initially only an appropriate cellular model of the explored disease is specified and validated in-depth (Zheng, Thorne, and McKew, 2013). Typical samples are immortalized cell lines, primary cells or induced pluripotent stem cells (iPSC's), ideally from both real patients and healthy test persons forming the control group. During subsequent assay development, meaningful disease-related cellular readouts must be established that can be assigned to either the healthy or diseased state. The availability of known reference compounds that serve as positive or negative controls can further simplify this validation step. In subsequent phenotypic screenings, potential drugs are tested with respect to their impact on diseased cells irrespective of their affected cellular target (Lang et al., 2006). As deployed drugs are acting on whole cells, their entire complexity is preserved and comprises all cellular abnormalities related to a given disorder. Hit compounds are in the following identified by predefined cellular responses, like targeted cell death, changes in a signaling pathway, or repression of a disease-related cellular feature (Zheng, Thorne, and McKew, 2013). As the underlying molecular targets are still unknown, target deconvolution must be carried out afterwards before MOAs can be deduced (Figure 1.2).

Target-based high-throughput screening approaches thus enable testing of ample chemical space, but usually neglect existing sample heterogeneity (Boutros, Heigwer, and Laufer, 2015). By averaging over all biomolecular or cellular entities present in a given well, only a single readout parameter is obtained for every applied substance and experimental condition. In contrast, phenotypic approaches deliver more comprehensive contextual information. This is especially true if cellular phenotypes are analyzed on a single-cell basis, thereby revealing cell-to-cell differences across the entire cell population. In this context, high-content screening (HCS) serves as the current state of the art technique for image-based profiling of new drug candidates and will be explained in greater detail in the next section.

1.2   High-content screening for drug discovery

High-content screening (HCS) offers the possibility to map cellular structures on the single-cell or even sub-cellular level (Lang et al., 2006). Assay specific target structures are specifically stained with fluorescent probes and then monitored using automated widefield or confocal microscopy techniques. As fluorescent labels can often be applied in parallel, so-called multiplexing methods allow for the simultaneous measurement of several factors for every individual cell. Routinely used readout parameters include the intensity level, morphology (i.e., the size and shape) as well as texture parameters for every labeled structure (Zanella, Lorens, and Link, 2010). The combination of these features results in a cellular fingerprint reflecting each cell's individual phenotype. Accordingly, changes in the expression level of certain proteins or in the molecular composition of specific cellular complexes or organelles can be

detected and related to deployed hit or lead compounds (Lin et al., 2020). Single-cell analysis furthermore allows for the detection of cellular sub-populations that show either a varying response to a given treatment, or are not affected at all. The possibility to perform live-cell assays further extends the accessible parameter space to the temporal domain. Generating huge datasets of highly-resolved images also enables the use of more objective analysis strategies, mainly driven by recent progress in machine learning (ML) and deep learning-based image processing (Chandrasekaran et al., 2021). Giving these algorithms access to the full variance of cellular responses to drug treatments can hence support the selection of promising lead structures and important molecular parameters for subsequent optimization.

From a technical perspective, most advanced HCS systems currently available are confocal imaging devices that run on multiwell plates in a fully automated fashion (Li and Xia, 2019). They mainly utilize a spinning-disk architecture and camera-based detection unit for highly parallelized and fast acquisition schemes. Multiple excitation laser wavelengths and detection filters allow for multi-color experiments, either in sequential or parallel mode. During measurements, either single-plane images or Z-stacks can be acquired on all available or only predefined well positions. Routinely utilized objective lenses are air or high-numerical aperture (NA) water immersion objectives that involve no compromise regarding their image quality and resolution as compared to standalone confocal microscopes.

Although these screening systems have proved to be highly useful, their obtainable resolution is limited by light diffraction, meaning that cellular structures closer than about 200 nm cannot be distinguished. Pioneering work in the past 20 years in order to surpass this inherent physical limit in biological applications, however, has led to super-resolution microscopy approaches that are nowadays routinely used in academic research (Sahl, Hell, and Jakobs, 2017). Resolution improvements by at least one order of magnitude now permit studies of cell biological processes with unprecedented molecular detail, both in fixed and living cells. Nevertheless, commercial setups that also provide high-throughput capabilities are still lacking. In the following section, different concepts of super-resolution microscopy are shortly introduced and assessed regarding their suitability for automated imaging workflows.

1.3 Automated high-throughput super-resolution imaging - which technique to choose?

The two most influential super-resolution fluorescence microscopy techniques are stimulated emission depletion (STED) microscopy (Hell and Wichmann, 1994; Klar et al., 2000; Vicidomini, Bianchini, and Diaspro, 2018), and single-molecule localization microscopy (Lelek et al., 2021), the latter commonly represented by the PALM (Betzig et al., 2006) and STORM approach (Rust, Bates, and Zhuang, 2006). Both methods achieve typical resolutions down to 20 nm and rely on the fact that only a limited amount of fluorescent probes being attached to a target structure is allowed

to fluoresce. Simultaneously, surrounding labels are prevented from contributing to the signal (Hell, 2007). In either case, the underlying fundamental concept is based on switching between two discriminate molecular states, usually a fluorescent *on* and a non-fluorescent *off*-state. However, the way how the switching process is experimentally realized differs between both methods and has substantial impact on the attainable spatio-temporal resolution.

In localization-based methods, switching between molecular states is achieved stochastically (Lelek et al., 2021). All fluorescent probes initially reside in a non-fluorescent off-state, such that they cannot be excited by an excitation light source. Afterwards, only a small sub-population of probes is randomly switched on and then imaged by widefield illumination on a camera detector. Their coordinates can be deduced with superior accuracy by 2D fitting as long as the signals of individual fluorophores are not spatially overlapping. As a result, every acquired image provides localizations of only a subset of fluorophores which are then combined to achieve the final super-resolved image.

This procedure has several drawbacks regarding its potential for automated high-throughput imaging: First, often thousands to ten-thousands of images must be acquired to obtain a sufficient sampling rate for typical structures of interest and labeling densities. This regularly leads to acquisition times between several minutes and almost an hour for one single field of view (FOV). Second, the ideal switching kinetic strongly depends on the underlying labeling density and must be carefully optimized for efficient imaging. In practice, this is often difficult to achieve and lacks real-time adaptability. Lastly, extensive post-processing is required which prolongs the process from acquisition to final results.

In STED microscopy, molecular switching is realized in a coordinate-targeted fashion by applying a spatially modulated light pattern. A diffraction-limited excitation laser and a modified depletion laser featuring an intensity minimum in its center (the so-called STED beam) are coordinately scanned over the sample. At every position, only the fluorescent probes in the center of the STED laser beam predominantly remain in their excited state and thus can emit fluorescence. As the attainable resolution only scales with the power of the STED laser, it is independent of the number of fluorophores being attached to the sample. The structure's labeling density therefore only has an impact on the resulting intensity level that can be counterbalanced by appropriate excitation laser settings. The super-resolved image is then directly created through raster-scanning the sample pixel by pixel, without the need for multiple image acquisitions or additional post-processing steps. Moreover, the imaging FOV can be confined solely to the desired sample region by setting an appropriate scan range. This reduces both the image acquisition time and overall amount of data in favor of a streamlined data management.

More recent innovations in the field, namely MINFLUX (Balzarotti et al., 2017; Gwosch et al., 2020) and MINSTED (Weber et al., 2021), have further enhanced the concept of super-resolution microscopy by combining both methodologies: After

isolating individual emitters stochastically as in SMLM, they become localized by utilizing a donut-shaped beam with precisely known position of the central intensity minimum. Attainable resolutions were reported to range between 1 nm to 3 nm and hence correspond to the size of a single biomolecule, with a localization precision that has only recently entered even the Ångström range (Weber et al., 2022). To achieve this, however, every attached fluorophore must be processed in successive order, thereby elongating the required acquisition time for a typical region of interest (ROI) to up to several hours (Schmidt et al., 2021). These methods are therefore not yet feasible for use in an automated imaging platform in view of the efficiency and speed requirements requested by drug screening.

As an additional technique, super-resolution structured illumination microscopy (SR-SIM) provides an up to twofold resolution improvement and fast imaging speeds (Heintzmann and Huser, 2017). In a recent implementation using an eightfold parallelized multiplane approach, video-rate acquisition in 2D and a frame rate higher than 1 frames per second for volumetric imaging were achieved (Descloux et al., 2020). However, due to the fact that this method requires sophisticated data postprocessing and the obtained resolution is still limited to about 100 nm, it was not further considered in the course of the present thesis.

In consequence, STED microscopy was chosen for an evaluation of automated super-resolution imaging with increased throughput. In the past, several important aspects regarding the automation of STED microscopy have already been established and - in part - also commercialized: adaptively changing the illumination pattern in dependence of local sample and staining properties helps to reduce photobleaching artefacts while stabilizing intensity levels (Staudt et al., 2011; Heine et al., 2017). Auto-alignment procedures ensure optimal performance also over extended acquisition times (Gould et al., 2013). As a precise overlap of the excitation and STED laser beams is essential for optimal STED performance, this is a central prerequisite for the intended automated imaging platform running under unsupervised conditions. The adjustment of optimal acquisition parameters by real-time, artificial intelligence-driven image analyses further helps to generate an uniform image quality, independent of particular labeling conditions (Durand et al., 2018).

Fully automated STED microscopy without manual user intervention has, to my knowledge, not been realized to date. In this work, I am going to explain the considerations and steps I conducted during the implementation of such a system for phenotypic screens of cell-based assays. Exceptional focus was set on the attainable image acquisition speed which is essential for achievable high-throughput capabilities. The study was carried out as part of the collaborative research project *Automated STED-Nanoscopy for Cell-Biological High-Throughput Analyses* between Bayer AG and Abberior Instruments GmbH. Within the project, Abberior Instruments was responsible for the development of an automated STED microscope prototype, whereas Bayer AG contributed necessary classical and machine learning-based algorithms for feature detection and STED image analysis. The model use case I utilized in

the course of the project was likewise determined by Bayer AG and dealt with centrosomal clusters in human cancer cells. Importantly, all data, analyses and results presented throughout this thesis were acquired, conceptualized and carried out by myself, if not otherwise specified. Details about the biological background and the addressed research question of the model use case will be further described in the next section.

1.4 Model use case: Centrosomal clustering in cancer cells

Accurate segregation of replicated DNA into daughter cells is key for regular cell division and viable progeny (Alberts et al., 2008). To this end, a bipolar spindle apparatus forms during mitosis, consisting of microtubules that bind to individual chromosomes. After alignment in a so-called equatorial plate, sister chromatides split up and get pulled apart to opposite sides of the dividing cell (Figure 1.3a). The microtubule network itself is controlled by the centrosome, a cellular organelle that comprises many different proteins and is usually referred to as the microtubule organizing center (MTOC; Bettencourt-Dias and Glover, 2007). Just as for DNA, the number of centrosomes is accurately maintained during the cell cycle in healthy cells, meaning that the centrosome duplicates exactly once and only once per individual cycle (Nigg and Holland, 2018).

In cancer cells and tumor tissue, however, the copy numbers of centrosomes are often found to be overamplified (Sabat-Pośpiech et al., 2019). In this case denoted as centrosome amplification (CA), the formation of multipolar spindles frequently leads to improper allocation of DNA into daughter cells, regularly followed by cell death (Ganem, Godinho, and Pellman, 2009). To circumvent lethal cell divisions, in many cancer cells these supernumerary centrosomes appear to be organized in so-called centrosomal clusters (Krämer, Maier, and Bartek, 2011). As a result, a pseudo-bipolar spindle apparatus can emerge that facilitates equal chromosome separation and survival of daughter cells.

Due to persisting imperfections, amplified centrosomes are often linked with centrosomal instability (CIN), a process typical for many cancer types and associated with increased tumor progression (Ganem, Godinho, and Pellman, 2009). Breaking of centrosomal clusters hence offers a promising therapeutic strategy to specifically target and suppress tumor cells (Ogden, Rida, and Aneja, 2012). Consequently, potential drug candidates are intended to de-cluster centrosomal aggregates and to promote the formation of multipolar mitotic spindles, thereby forcing only cancerous cells with higher centrosome copy numbers into cell death (Pannu et al., 2014).

A fully mature centrosome consists of two cylindrical centrioles, one parent centriole and one closely attached procentriole (Nigg and Holland, 2018). During each cell cycle, the procentriole detaches from its mother and becomes an immature daughter centriole at the beginning of G1 phase (Figure 1.3b). Both centrioles stay linked by a flexible tether, but only the mother centriole is fully mature and carries distal
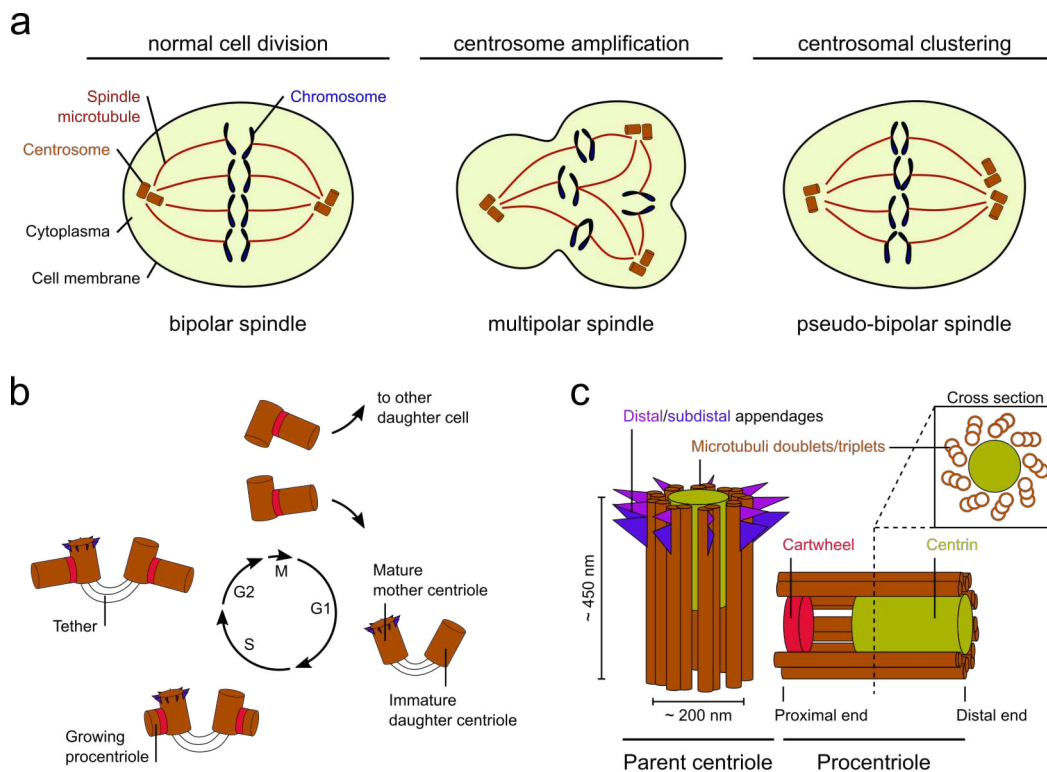
Figure 1.3: Centrosomal structure and their effect on cell division. (a) In healthy cells, two centrosomes organize a bipolar spindle apparatus during mitotis (left). Presence of supernumerary centrosomes can lead to multipolar spindles and subsequent cell death (center). In cancer cells, excess centrosomes are frequently observed in clusters to ensure pseudo-bipolar mitotic spindle and cell viability of daughter cells (right). (b) Centrosome genesis in the course of the cell cycle. After cell division, the procentriole detaches from its parent and stays connected by a flexible tether. During S-phase, new procentrioles emerge on both parent centrioles, mediated by a cartwheel protein assembly. Distal and subdistal appendages are only present at the previous parental centriole before start of mitosis. (c) A fully mature centrosome consists of one parental centriole and a firmly attached procentriole. Centrin protein assembles into the lumen of both centrioles, preferentially at their distal ends. G1/G2: gap 1/gap 2 phase, S: synthesis phase, M: mitosis. Inspired by and adapted from (Nigg and Holland, 2018).

and subdistal appendages. In S-phase, new procentrioles start to grow on both centrioles, mediated by a cartwheel structure serving as a scaffold. By late G2 phase, elongation of procentrioles is completed. The two parent centrioles separate and become two fully mature centrosomes required for cell division.

Each individual centriole consists of a microtubule scaffold with nine-fold symmetry (Brito, Gouveia, and Bettencourt-Dias, 2012). While these microtubules are arranged in triplets between the center and proximal end, they show only doublet configuration at the centriole's distal end (Figure 1.3c). At their distal and subdistal appendages, the mother centriole attaches ciliary vesicles and nucleates microtubules of the mitotic spindle, respectively (Uzbekov and Alieva, 2018). The protein centrin

assembles into the lumen of both centrioles, predominantly at their distal ends, and is thus an appropriate target for their quantification. The cartwheel structure mainly consists of SAS6 protein and solely arises at the proximal end of procentrioles. Under physiologically intact conditions, both centrioles measure approximately 450 nm in length and 200 nm in diameter, i.e., their size is in the range of the diffraction limit (Marteil et al., 2018). For various cancer cell lines, however, also over-elongated centrioles were repeatedly observed with influence on both centrosomal activity and chromosome segregation (Marteil et al., 2018).

Most studies focusing on centrosomal content were conducted by means of confocal fluorescence microscopy. With regard to super-resolution microscopy, mainly structured illumination microscopy (SIM) and 3D-SIM were applied to elucidate specific protein localizations or molecular composition (Fu and Zhang, 2019). However, in a few cases STED microscopy was used to resolve the distribution of protein clusters at the centriole-cilium interphase (Lee et al., 2014; Tony Yang et al., 2015) and at the centrosome linker (Vlijm et al., 2018). It is therefore assumed that STED microscopy is suitable to precisely determine the number of centrioles, even though an exact quantification may be challenging for densely clustered structures. By applying automated STED microscopy, this work intends to reveal the exact number of centrioles and their orientation in mitotic cells of a cancer cell line.

The following results chapter is divided into four parts: in the first section, the assay development for the centrosomal clustering use case will be explained in full detail. Afterwards, both boundary conditions and technological limits that culminated into a specific design for an automated STED image acquisition scheme are discussed. Subsequently, the individual steps towards implementation of microscope automation are introduced. Besides embedding of technical components, this includes the development of robust and precise image analysis algorithms for feature identification and localization. In the final section, the results of a first proof of concept compound screening with ten deployed drug substances is presented.

Chapter 2

Experiments and Results

2.1   Assay development of the model use case

To evaluate the potentials and limitations of an automated screening STED microscopy platform, centrosomal clusters in cancer cells were identified as a promising model use case by Bayer AG. According to literature, many cancer cell lines exhibit an overamplified centrosomal content that frequently accumulates in so-called centrosomal clusters (Marteil et al., 2018). As their individual components, the centrioles, are in the size range of the diffraction limit of light, signals from different centrioles are likely overlapping when imaged with classical confocal or widefield microscopy. This may hinder the precise quantification of the underlying number of centrioles per cluster, especially when dense clusters are of particular interest.

As pointed out in the introductory part, this is a relevant obstacle in the evaluation of potentially anti-cancerous drugs in a cell-based assay. A potent active ingredient may efficiently breakup centrosomal clusters and stimulate cell death by improper DNA segregation during cell division. Hence, a target protein for the characterization of centrosomal clusters should ideally have several properties: First, it must localize to individual centrioles, instead of, e.g., the surrounding pericentriolar matrix, to allow for precise quantification of copy numbers. Furthermore, it must be present at both the mother and the daughter centriole, irrespective of the current cell cycle stage. Finally, enough proteins must accumulate at the target structure and be accessible for labeling to yield sufficient signal intensities. It is the last aspect that is of particular importance for STED microscopy due to commonly observed lower photon counts and likewise reduced signal-to-noise ratios. Taking into consideration the aforementioned points, centrin was identified as a suitable primary target protein for the intended screening purpose.

Due to the fact that in the final drug screening assay mitotic cells are of specific interest, additional secondary targets were specified to support cell cycle classification. α-tubulin is a key element of microtubules that form the mitotic spindle apparatus during cell division. Being also directly connected to centrosomes, it also reveals whether a particular centrosome is localized at a spindle pole and thus actively involved in the cell division process. In addition, DNA was chosen as innate marker for chromosome segregation and elucidation of the cell cycle stage. U2OS

osteosarcoma bone cancer cells were initially defined as cellular model system by cooperation partners within Bayer AG and used for optimizing the sample preparation protocols. Later in the project, I encouraged the switch to the BT549 mammary gland breast cancer cell line due to its higher yield of cells showing centrosome amplification. The detailed route towards establishment of the final staining protocol and cellular target system is outlined in the following section.

2.1.1   Optimization of cell fixation and labeling protocols

Screening enormous numbers of potential active ingredients in automated imaging pipelines involves specific requirements regarding not only the microscopy instrumentation, but also the sample preparation. This is of special importance in the case of fluorescence microscopy assays as protocols used in academia are often cumbersome. Especially for super-resolving microscopy techniques like STED, the demands on sample preparation are even higher than for standard confocal microscopy, as the susceptibility to artifacts increases significantly with resolution. In the industrial environment, labeling protocols must additionally be highly automatable and scalable to the microtiter plate format. Beyond that, protocols must be as simple as possible, preferentially only rely on cheap and broadly available chemicals and should preferably not depend on exact temperature or time constraints. To this end, the sample preparation protocol for the centrosomal clustering assay was optimized for ease of use while reducing fixation and labeling artifacts as much as possible.

While icecold methanol is often used for microtubule fixation in academic literature (Whelan and Bell, 2015; Hua and Ferland, 2017), this approach is usually not feasible for large scale sample preparation in industrial scales. Usually, methanol must be cooled down to $-20\,°C$, kept at this temperature during fixation and the incubation time needs to be accurately maintained. Moreover, cells tend to collapse when fixed with methanol which is of major importance when cellular targets shall be accurately measured in mitotic, i.e., spherical cells. Hence, for cell fixation only cross-linking fixatives like formaldehyde (FA) and glutaraldehyde (GA) were considered. Likewise, cross-linking reagents are less prone to introduce artifacts due to variations in the incubation time[1]. Liquids were always removed completely before new solutions were added to a well, as it is commonly done by a pipetting robot or inverted centrifuge during automated sample preparation. Elaborate fixation procedures in which the fixative is added while the old cell medium is simultaneously removed were not considered.

As a starting point, a standard fixation protocol with $4\,\%$ FA was used for microtubule fixation. In the resulting images, however, undesirable artifacts were observed (Appendix Figure A.1a). Frequent gaps along individual microtubules suggested that either the target structure has not been labeled continuously, or that the microtubules ruptured during the fixation process. Furthermore, various dark spots

---

[1]Protocols in literature vary strongly in that sense, with reported incubation times ranging from 5 min (Wurm et al., 2010) to 20 min (Li et al., 2017).
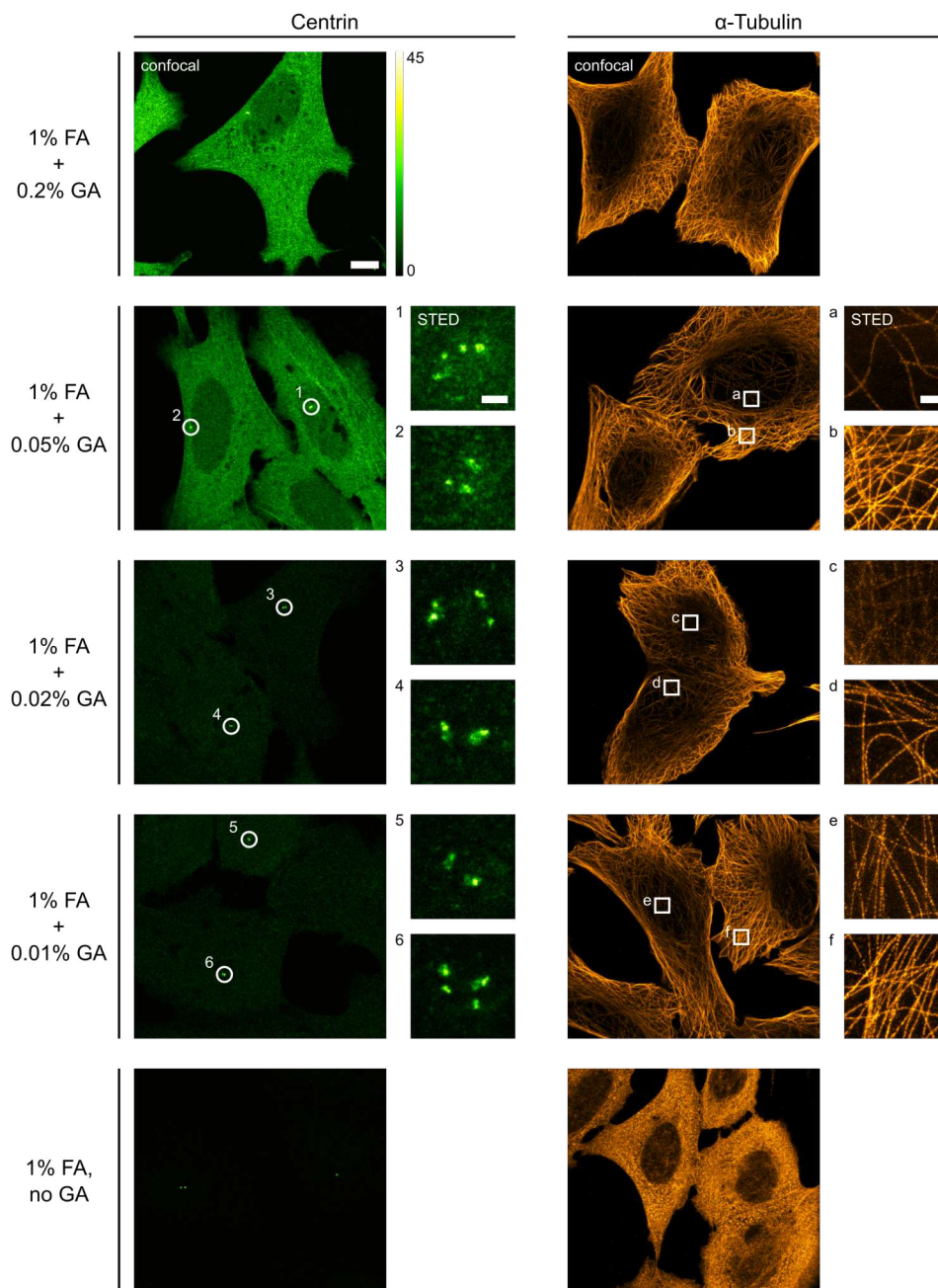
Figure 2.1: Optimizing cell fixation for centrin and α-tubulin imaging. Representative immunofluorescence images for different fixative concentrations in U2OS *wt* cells. Cells were stained with centrin (left) and α-tubulin (right) primary antibodies, secondary antibodies were coupled to Abberior STAR RED and Abberior STAR ORANGE dyes, respectively. Stainings were performed on individual samples to eliminate any cross-talk effects. For centrin labeling, confocal overview images were scaled equally to illustrate the varying background levels for different glutaraldehyde (GA) concentrations. Insets show exemplary STED images of both target structures (pixel size: 25 nm). Scale bar 10 µm, centrin/α-tubulin STED insets 500 nm/1 µm.

close to the substrate indicate that the plasma membrane detached during fixation, a phenomenon commonly denoted as membrane blebbing. When together with FA also 0.2 % glutaraldehyde (GA) were used during fixation, the microtubule structure preservation improved substantially (Appendix Figure A.1b). Membrane blebbing artifacts appeared to be dependent on the applied FA concentration, as both the size and frequency of blebs decreased when the concentration was reduced to 1 % (Appendix Figure A.1c).

The dependence of membrane blebbing on the FA concentration was further evaluated in cell fixation experiments. U2OS *wt* cells were live-stained with a membrane stain and showed no indentations before fixation (data not shown). Subsequently, cells were fixed with 0.2 % GA together with different FA concentrations ranging from 1 to 8 % (Appendix Figure A.2). The membrane structures appeared to be more intact the lower the FA concentration was, resulting in only minor sample impairment at a concentration of 1 % FA and 0.2 % GA.

Finally, the influence of the fixation conditions on the primary target centrin was evaluated. Surprisingly, U2OS cells exhibited substantial unspecific background signal when the protocol optimized for α-tubulin labeling and membrane structure preservation was applied for centrin staining (Figure 2.1). This could impede both target identification in the detection process and data analysis of the final STED acquisitions. However, by reducing the GA concentration to 0.02 %, the background signal could be reduced to manageable levels while still satisfactorily preserving microtubule structure. Further GA reduction lead to no additional improvement, whereas complete omission of GA during fixation resulted in a destroyed microtubule cytoskeleton (Figure 2.1). Control experiments revealed that the primary anti-centrin antibody is responsible for the background signal, independent of its applied concentration (data not shown). The final fixation protocol with 1 % FA and 0.02 % GA thus represented a satisfactory compromise that worked for both centrin and α-tubulin labeling and was used in all further experiments.

2.1.2   Optimization of target labeling

For applying STED microscopy, target labeling with organic fluorophores is generally preferred over fluorescent proteins due to their superior photostability and higher molecular brightness. These properties also result in an improved imaging performance, i.e., better resolution. For that reason, immunolabeling protocols using primary and secondary antibodies were optimized for the target proteins centrin and α-tubulin. As the final imaging workflow is intended to work autonomously without manual intervention, the labeling must be robust and reproducible for a large number of different cells and varying incubation conditions. Moreover, signals must be specific and resulting signal-to-noise ratios (SNRs) high enough to reliably identify protein localizations. Therefore, a suitable dye pair must be selected to avoid substantial cross-talk between both detection channels. These requirements can be challenging in the case of STED imaging as the number of suitable organic
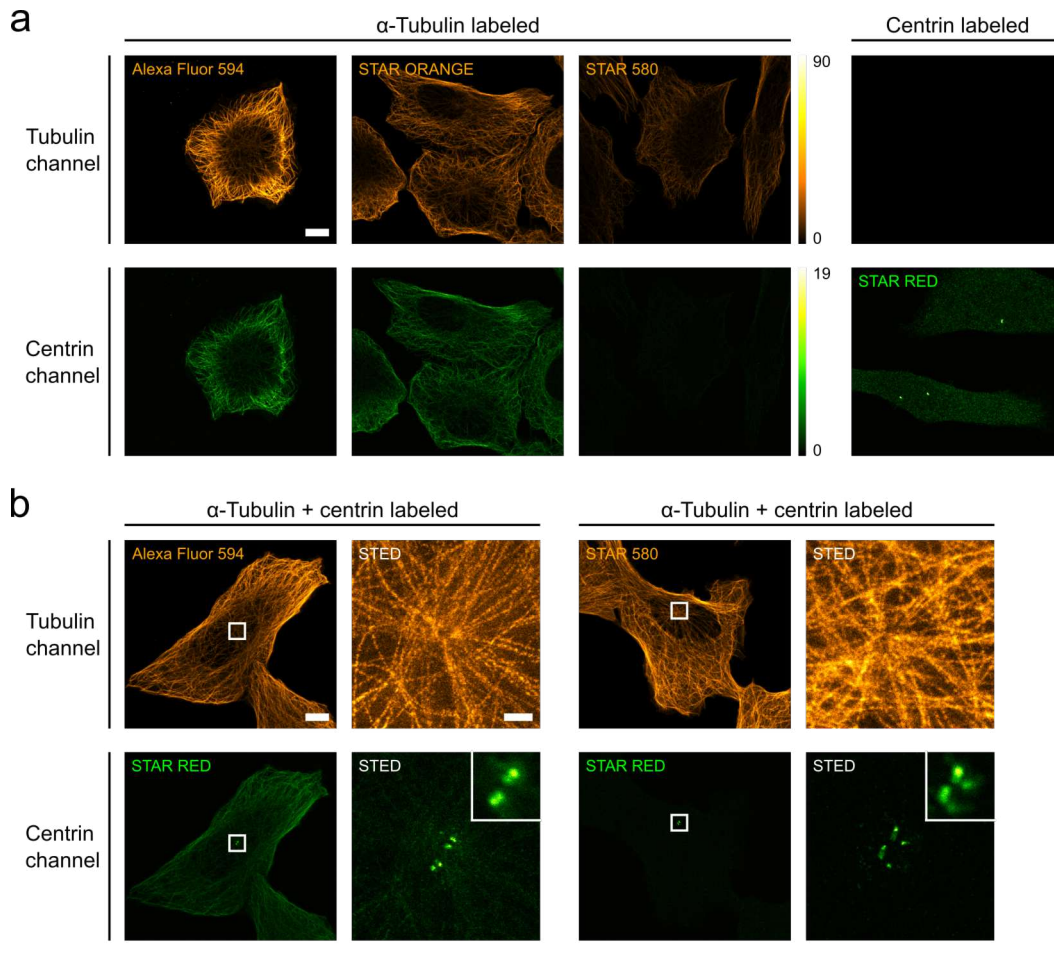
Figure 2.2: Optimizing fluorophore labels for immunofluorescence imaging of U2OS cancer cells. (a) α-tubulin was labeled with secondary antibodies coupled to Alexa Fluor 594, Abberior STAR ORANGE and Abberior STAR 580 and imaged both in the α-tubulin (561 nm exc., 580-625 nm det.) and centrin (640 nm exc., 650-700 nm det.) detection channel (left). For comparison, centrin was labeled by Abberior STAR RED (right). Confocal images were acquired with identical parameter sets for each channel and identically scaled, as denoted by the colorbar. Scale bar 10 µm. (b) Dual-color confocal imaging using either Alexa Fluor 594 or Abberior STAR 580 for labeling α-tubulin, Abberior STAR RED for centrin, and DAPI for DNA (not shown). Both centrin overview images feature identical identical scaling. Boxed regions around centrosomal features were subsequently imaged in STED mode (insets). For centrin, insets also show the cooresponding confocal image. Scale bar 10 µm, insets 1 µm.

fluorophores is limited and must also match the available excitation wavelengths and detection filter settings.

In general, best STED resolution is obtained in the far-red wavelength regime, using an excitation wavelength of 640 nm and a STED laser at 775 nm. As the optimal resolution is desired for the primary target centrin, Abberior STAR RED and Abberior STAR 635P, two fluorophores specifically developed for STED applications, were tested for centrin staining. I found no significant difference between both dyes and

decided to use Abberior STAR RED for centrin labeling (data not shown). A commonly used dye pair for two-color STED imaging consists of Abberior STAR RED and Alexa Fluor 594. In control experiments, however, substantial cross-talk was observed from the α-tubulin labeling into the centrin channel (Figure 2.2a). This finding was surprising as this dye combination is routinely used in the field. However, in the present use case the brightness of the two target proteins was very different, and a relatively high excitation laser power had to be applied for centrin imaging. As the emission spectra of Abberior STAR RED and Alexa Fluor 594 largely overlap and channel separation is mostly achieved by different excitation spectra, this principle works less efficiently in the present case. Slightly less, but still detrimental cross-talk was found when α-tubulin was labeled with Abberior STAR ORANGE. Only with Abberior STAR 580, a fluorophore with a spectrum shifted to shorter wavelengths, the cross-talk could be reduced to negligible levels (Figure 2.2a).

In concluding dual-color experiments, the microtubule signal bleeding into the centrin channel was indeed obstructive when α-tubulin was labeled with Alexa Fluor 594 (Figure 2.2b). The underlying filamentous structure could be clearly observed in the centrin channel and reached intensity values of up to half of the maximum intensity values observed for centrin. Although this cross-talk might be less impeding in subsequent STED images, it may hinder the automated feature detection process. In contrast, use of Abberior STAR 580 as the α-tubulin label decreased the cross-talk drastically (Figure 2.2b). Therefore, in all further experiments Abberior STAR RED was used to label centrin and Abberior STAR 580 was used to label α-tubulin.

2.1.3   Optimization of the cellular model system

Initially, the U2OS human bone cancer cell line was considered to be an appropriate model system for studying the effect of drugs on centrosomal clusters. Previous findings showed that U2OS cells often contain supernumerary centrosomes, although their indicated cellular fractions varied between studies: reported values include 2-3 % for an EGFP-PLK4-U2OS cell line (Konotop et al., 2016) and a range between 7 % (Wong et al., 2015) and 20 % for U2OS *wt* cells (Ganem, Godinho, and Pellman, 2009; Kwon et al., 2008). Furthermore, U2OS cells are easy to culture, rather flat and show an epithelial growth behavior (Westrate et al., 2014) which makes them a reasonable choice for super-resolution imaging applications.

U2OS cells were prepared with the previously optimized fixation protocol using 1 % FA and 0.02 % glutaraldehyde. Centrin was labeled with Abberior STAR RED and α-tubulin was labeled with Abberior STAR 580. DNA was stained with DAPI. Unexpectedly, only very few centrosomal clusters were found in hundreds of manually inspected cells during protocol optimization. Instead, centrosomal structures almost exclusively consisted of either a single or two centriole pairs, each composed of one mother and one daughter centriole (Figure 2.3a). This observation is consistent with the fact that centrosome duplication occurs during S phase in the cell cycle, meaning that the adherent cells were imaged either before or after duplication. This
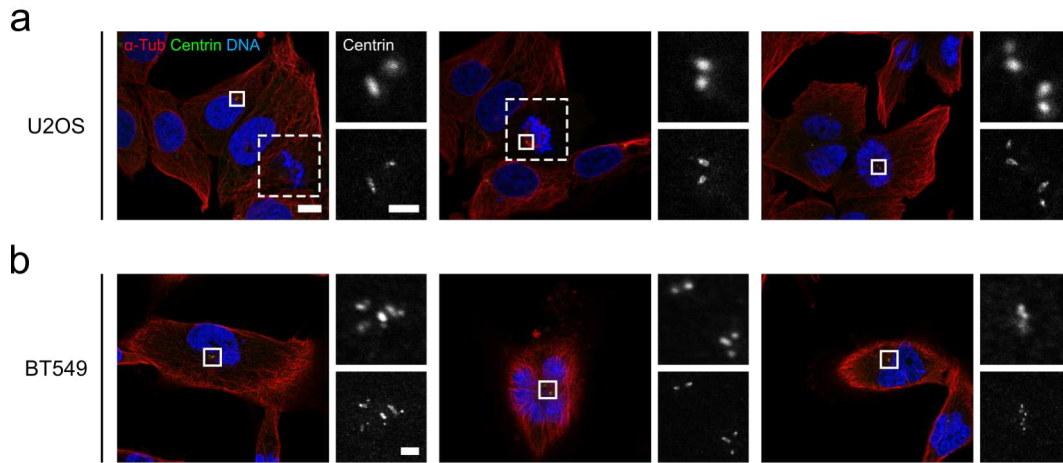
Figure 2.3: Centrosomal content differs in cancer cell lines. Representative fluorescent images of immunolabeled U2OS wt and BT549 *wt* human cancer cell lines stained for α-tubulin, centrin and DNA (labeled with DAPI). Identified centrin spots in confocal overview images (boxed regions) were subsequently imaged in confocal and STED mode with identical pixel sizes (upper and lower insets). U2OS cells predominantly exhibited normal centrosomal content with either one or two centriole pairs per cell. The dashed boxes indicate a rare mitotic cell that was refocused and then investigated (middle overview). BT549 cells contained abnormal centrosomal features more frequently, but not in every cell. Images were acquired manually with a STEDYCON microscope. Scale bar 10 μm, insets 1 μm.

finding was crucial for the intended assay that was mainly targeted at mitotic cells, which normally make up only about 5 % of all cells (Schorl and Sedivy, 2007). Since a large fraction of cells that feature an abnormal centrosomal content is essential for reasonable throughput in the final drug screening application, an alternative cellular model system had to be identified.

In literature, the murine cell line N1E-N115 is frequently used in declustering studies as it features supernumerary centrosomes in about 80 % of cells (Wong et al., 2015). However, since potential novel drug candidates are aimed at targeting human cancer cells, this option was not feasible. Therefore, with BT549 and MDA-MB-231 two mammary gland human breast cancer cell lines were evaluated as alternatives for which higher fractions of cells with centrosome amplification have been previously reported: for the BT549 cell line, fractions between 19 % (Wong et al., 2015) and 45 % (Kwon et al., 2008) of cells were published. For MDA-MB-231 cells, corresponding values of 16 % (Wong et al., 2015) and 45 % (Ganem, Godinho, and Pellman, 2009) were reported. When brought into culture, however, the MDA-MB-231 cell line showed a challenging growth behavior with overly stretched cells growing on top of each other that was therefore excluded (Appendix Figure A.3). In contrast, BT549 cells exhibited feasible cell morphologies and a good surface coverage and were thus further evaluated.

After applying the identical labeling protocol as for U2OS, BT549 cells indeed
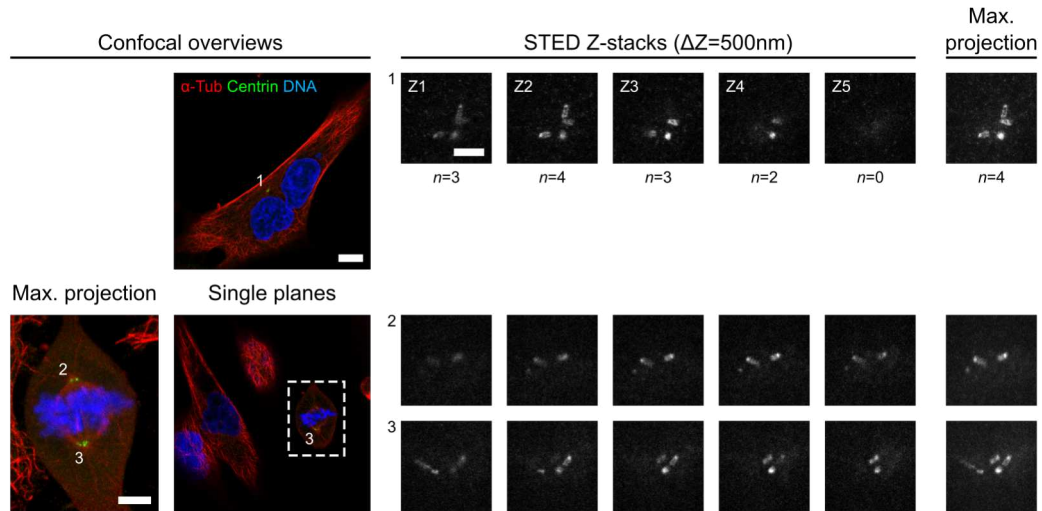
Figure 2.4: Accurate centriole quantification requires Z-stack image acquisition. Fixed BT549 *wt* cells were labeled for α-tubulin, centrin and DNA (DAPI). Centrosomal features were localized in confocal overview images by either manual focusing through cells of interest (top-left) or by acquisition of confocal Z-stacks (bottom-left). Note that for the mitotic cell (dashed box) both features of interest became only visible in the maximum intensity projection whereas in the single image plane only a single feature could be identified. Centrosomal regions of interest (ROIs) were subsequently imaged as Z-stacks in STED mode (ΔZ = 500 nm, total Z-range = 4 μm). Depicted images (Z1-Z5) stem from consecutive image planes and were then combined in maximum intensity projections. The observable number of centrioles *n* varied between individual image planes, but could be correctly quantified in the projection, as illustrated for the agglomerate in the interphase cell (top panel). Scale bar confocal overview 10 μm, confocal inset 5 μm, STED insets 1 μm.

featured a higher percentage of cells of interest in manual fluorescence imaging experiments (Figure 2.3b). Although exact quantification was challenging in view of required imaging times, both dense clusters and individual centrosomes with abnormal numbers of centrioles were regularly observed. By comparing confocal and STED images of identical centrosomal features, the STED images revealed superior level of detail, especially for dense structures (Figure 2.3b, insets). The projected orientation of the cylinder-shaped centrioles and their underlying quantity could be better resolved in STED mode.

For the evaluation of potential declustering agents that are capable of breaking up centrosomal aggregations, exact knowledge of the true centriole number is essential in order to weight individual compounds based on their declustering power. A thorough analysis hence requires cells which show a high degree of centrosomal clustering, which was not the case in the tested U2OS cell line. For the evaluation of automated STED microscopy for screening applications, therefore the BT549 cell line was considered as a better suited model system and used in all succeeding implementation steps.

2.1.4 Manual evaluation of the mitotic cell assay

After optimizing fixation and labeling conditions and the cellular model system, BT549 cells were prepared and manually imaged to specify suitable imaging conditions for the intended assay. Centrosomal features of interest were identified by manually focusing through cells. Subsequently, only these relevant regions-of-interest (ROIs) were imaged in STED mode as Z-stacks (Figure 2.4). In this procedure, single layers are imaged successively over a specified Z-range and a defined slice distance ΔZ. Both in interphase and mitotic cells, the particular focal plane proved to be paramount for exact quantification. Positional deviations in the order of only 500 nm could lead to differing counts of centrioles between consecutive images, depending on their respective orientation in 3D space[2] (STED insets, Figure 2.4). Hence, for a full picture of centrosomal clusters, Z-stack acquisitions are required. These Z-stacks can be summarized into maximum intensity projections along the Z-axis. Manual investigation of acquired Z-stack data suggested that the number of centrioles can be determined with sufficient accuracy in projected image data. However, the deduction of more extensive parameters, like, e.g., the exact spatial orientation of centrioles and their angles of intersection, would require an analysis of the Z-stacks as a whole.

Z-stack acquisition proved to be mandatory not only for super-resolution imaging of centrosomal clusters, but also for their detection when performing confocal imaging on the cellular level. Centrosomes can be arbitrarily located in the cytoplasm, especially in mitotic cells where the orientation of the spindle apparatus is not *a priori* defined. As all ROIs must be identified before consecutive STED imaging, cells of interest ought to be screened completely to not miss any centrosomal feature of interest (Figure 2.5). While selecting and imaging dozens of mitotic BT549 cells by hand, a rather strong cell-to-cell variability was observed. Detected mitotic spindles showed either a bipolar or multipolar morphology. Interestingly, provided that supernumerary centrosomes were present in a cell, the latter was occasionally also detected in non-treated cells. Assuming an overall high cell heterogeneity also in a drug screening scenario, the necessary amount of data easily exceeds the quantity that can still be acquired by hand. The required sample sizes for significant conclusions thus further demonstrate - aside from other aspects like data consistency, objectivity, or costs - the need for an automated STED imaging platform.

For the actual implementation of such a platform, several factors have to be considered, with the particular research question and the applied sample type (cell layer, spheroid, or tissue slice) being only the two most apparent ones. Any image acquisition scheme furthermore depends on the required imaging modalities (widefield, confocal, STED), the use of specific objective lenses (magnification, depth of focus) and trade-offs, e.g., between the desired imaging speed and localization accuracy.

---

[2]Note that throughout this thesis, solely 2D STED microscopy was performed with no PSF contraction along the Z-axis. If in future applications 3D STED is intended to be used, the need for Z-stack acquisitions becomes even more significant as the PSF shape in Z-direction then reduces substantially from $\approx 1\,\mu m$ to only $\approx 100\,nm$.
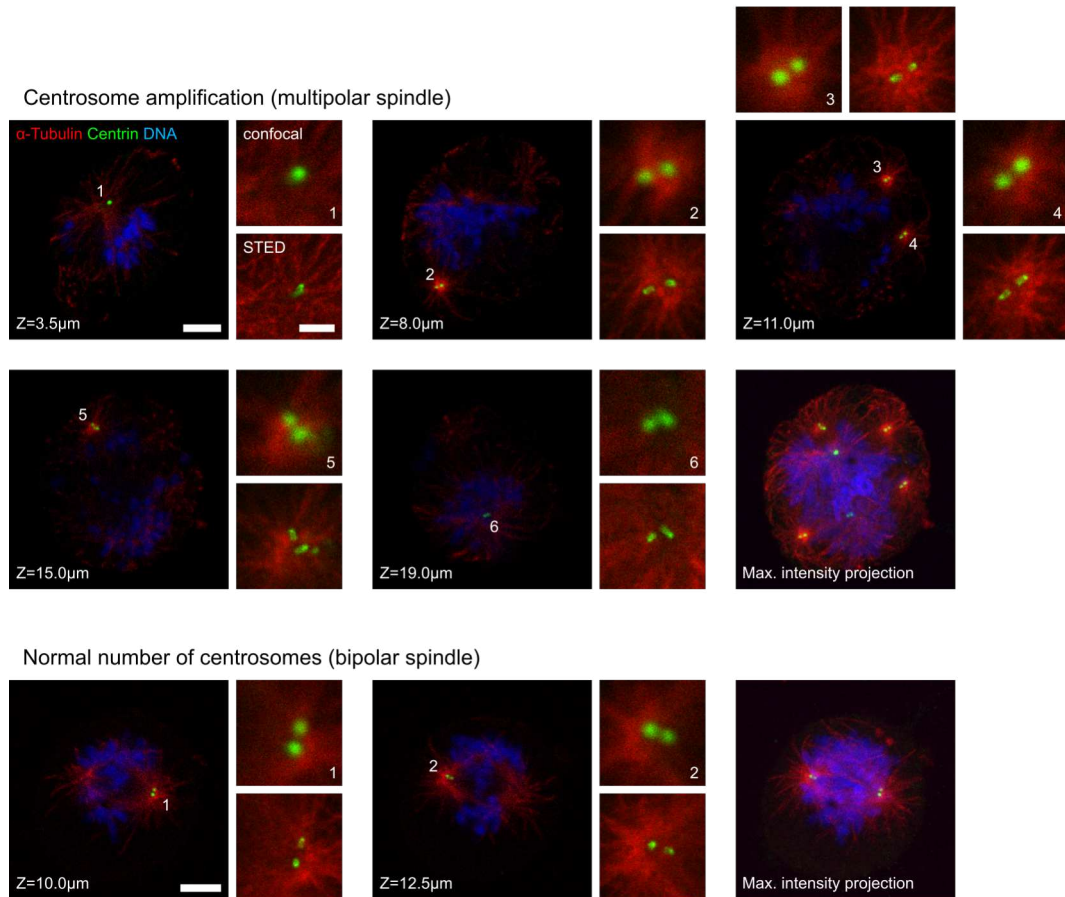
Figure 2.5: Mitotic BT549 cells show high variability regarding centrosome amplification and polarity of spindle apparatus. Exemplary immunofluorescence images of manually picked mitotic cells with supernumerary (top) and normal centrosome content (bottom). Confocal overview images were acquired as Z-stacks (ΔZ = 500 nm, total Z-range = 26 μm and 20 μm, respectively) and then combined to maximum intensity projections. Individual centrosomal features appear at distinct Z-heights inside the stack and were manually re-focused before acquiring single-plane confocal and STED images (insets). Samples were stained by α-tubulin and centrin antibodies and DAPI (DNA). Scale bar 5 μm, insets 1 μm.

A thorough evaluation of these factors was performed for the use case addressing centrosomal clustering and is detailed in the next section.

2.2 Automated STED microscopy - a technical evaluation

Automated microscopy platforms in general require stable instrumentation, reliable algorithms and robust data handling. Respective devices are designed to ensure permanent and unsupervised operation without any user intervention[3]. This requires both fail-safe and stable illumination sources, optics, detectors and electronics. Apart from hardware, nowadays also the applied image processing algorithms are of paramount importance, as they are often fully integrated and decide, for example, where to image and which settings to use (Durand et al., 2018).

Concerning automated STED microscopy, additional constraints must be considered with respect to the attainable imaging acquisition speed. Due to the super-resolving power of STED microscopy, smaller pixel sizes must be applied in comparison to standard confocal or widefield imaging approaches (Table 2.1). According to the Nyquist-Shannon sampling theorem (Shannon, 1949), the pixel size must equal at least half the size of the expected image resolution. Typical STED resolutions are in the range of 30-40 nm for biological specimens, calling for pixel sizes in the order of 15 to 20 nm. Furthermore, STED images are frequently darker when compared to confocal images with, apart from that, identical parameter settings. Typical reasons for this include the usage of time-gated detection schemes for commonly applied pulsed STED imaging (Moffitt, Osseforth, and Michaelis, 2011). More importantly, also imperfections in the root of the STED doughnut point spread function (PSF) lead to reduced fluorescence emission (Jahr, Velicky, and Danzl, 2020). To compensate for lower intensity levels, one usually applies more excitation laser power during STED acquisitions with an at the same time overall extended pixel dwell time[4]

In summary, both effects lead to drastically increased acquisition times for STED imaging in comparison with diffraction-limited, i.e., widefield or confocal, imaging modalities (Table 2.1). As a consequence, the total size of the regions of interest (ROIs) where STED imaging shall be performed must be reduced to a minimum in the intended screening application[5]. Moreover, also the fast and precise localization of ROIs in prior imaging steps is a key factor to obtain optimal acquisition speeds in any automated super-resolution imaging platform.

In a typical scenario, the detection of potential features of interest is a two-step process: first, designated cells of interest must be identified within the entire cell

---

[3]In industrial high-content analysis screening pipelines, the entire sample handling is often fully automated by robot lines with integrated picking robots. In this project, however, only automated STED microscopy on multiwell plates was addressed while sample preparation was still performed manually.

[4]When imaging biological specimens, it is often beneficial to acquire multiple line accumulations with shorter pixel dwell times rather than a single line with an identically long exposure time. This approach allows for triplet-state relaxation of applied fluorophores and can reduce photobleaching (Donnert et al., 2006). The number of line accumulations used for STED imaging is typically between 3 and 10, with pixel dwell times in the range of 5 to 10 µs, depending on the sample.

[5]Confocal imaging can be further accelerated using parallelized spinning-disk acquisition schemes. Similar approaches for parallelized STED have been successfully established in academia. Such systems are, however, not yet commercially available and were therefore not further considered in this work. Please refer to the discussions section for details on future perspectives.

Imaging mode	Pixel size	FOV size	Dwell time	Image acquisition time
Widefield	65 nm	$166 \times 140 \,\mu m^2$	$\sim 100 \, ms$	$\sim 100 \, ms$
Confocal	100 nm	$75 \times 75 \,\mu m^2$	$\sim 10 \,\mu s$	$\sim 5.6 \, s$
STED	20 nm	$75 \times 75 \,\mu m^2$	$\sim 3 \times 10 \,\mu s$	$\sim 7 \, min$

Table 2.1: Acquisition times for different imaging modalities. Given numbers are rule of thumb estimates for camera-based widefield and laser-scanning confocal microscopy (no spinning-disk). For all imaging modes, a 100X objective lens was assumed that is regularly used in the final STED imaging step. Indicated field of view (FOV) sizes reveal specifications of the microscopes and cameras used in the present study. High-level details like scanner flyback times for confocal microscopy were omitted for reasons of simplicity.

population. This may include cells of a specific phenotype, a particular cell cycle stage or cells that feature a certain expression level of some specific protein of interest. Consecutively, features of interest must be localized inside these cells to determine their (x, y, z)-coordinates for subsequent STED imaging. Considering the present use case that addresses centrosomal clusters only in mitotic cells, both procedures were incorporated into an intended general image acquisition scheme as preview and overview imaging steps, respectively (Figure 2.6a). Mitotic cell detection was expected to be attainable within single-plane preview images. For feature detection in cellular overviews and final STED imaging of centrioles, however, Z-stack acquisitions were shown to be necessary. In between these acquisition steps, fast and robust automated image analyses must be applied to elucidate respective spatial coordinates of cells and centrosomal features.

In the course of the underlying collaborative *Screening-STED* research project, two approaches for automated STED microscopy were identified, established and evaluated: In the first approach, the project partner Abberior Instruments GmbH developed a prototype microscope solely based on the confocal laser-scanning imaging modality (Figure 2.6b). In the past three years, I critically accompanied the development of this device and conducted both performance tests and first automated measurements. Due to its collaborative nature, however, large parts of the project were processed by internal and external partners beyond my direct influence[6].

In an alternative approach, I therefore independently established automated STED imaging of centrosomal features by upgrading an existing Abberior Instruments ExpertLine STED microscope in the research group of Prof. Stefan Jakobs. Therein, I combined fast camera-based widefield preview and overview imaging for cell and centrosome detection with subsequent laser-scanning acquisition solely on found features of interest (Figure 2.6c). Hereinafter, I will at first introduce both the prior considerations and the final design concept of the Abberior Instruments automated prototype device. Potential speed benefits associated with widefield imaging

---

[6]Especially automated sample preparation on multiwell plates and machine learning-based cell and feature detection algorithms for the automated Abberior Instruments prototype device were processed by other persons.
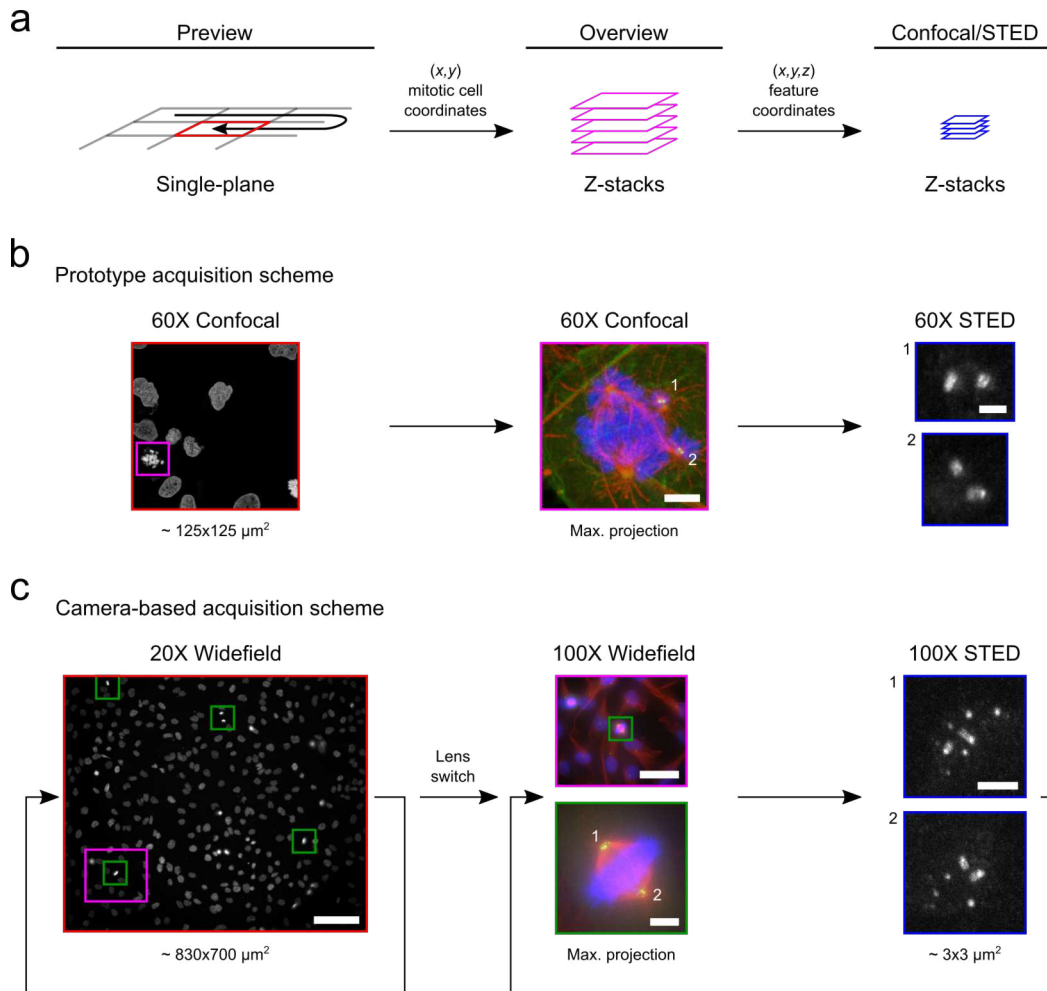
Figure 2.6: Intended acquisition schemes for automated STED imaging. (a) The general pipeline consists of consecutive preview, overview and confocal/STED acquisition steps for cell detection, feature detection and final feature imaging, respectively. (b) Acquisition scheme as realized in the *Screening-STED* prototype device developed by Abberior Instruments. Scale bar confocal overview 5 μm, STED insets 500 nm. (c) Alternative scheme using camera-based widefield imaging in prior detection steps to maximize attainable screening speeds. This procedure was implemented by a custom upgrade of a present Abberior Instruments ExpertLine setup. Scale bar widefield preview 125 μm, widefield overview 50 μm, overview inset 5 μm, STED insets 1 μm.

will be afterwards described in the subsequent section.

2.2.1 Abberior Instruments automated STED prototype design

The prototype device developed by Abberior Instruments was based on a pure confocal raster-scanning imaging unit, i.e., a focused laser beam is scanned across the sample line by line. In comparison with standard widefield imaging, the axial resolution is significantly increased as out-of-focus light is rejected at a pinhole positioned in the detection path. To avoid aberration artifacts especially during STED

imaging, usually two galvanometric mirrors for each the x and y scan direction are used. As a drawback, this so-called quad-scanner design leads to considerably reduced scan rates. When scanning the entire field of view (FOV) of a given objective lens, the maximum line frequency is limited to approximately $500\,\text{Hz}$. At the same time, the size of galvanometric mirrors is constrained by their respective moment of inertia which also restricts accessible FOV sizes. The maximum confocal FOV of a 100X oil immersion objective lens typically used for STED imaging yields only about $75 \times 75\,\mu\text{m}^2$. For widefield imaging, however, the same lens allows a quadratic FOV size as large as approximately $155 \times 155\,\mu\text{m}^2$, assuming a field number (FN) of 22. The resulting FOV size in the sample plane thus differs by more than a factor of four between both imaging modalities.

Throughout the initial project phase, several optimizing parameters were identified in joint discussions with Abberior Instruments to obtain maximal imaging speed despite these constraints. These included the magnification factor of the applied objective lenses and the adjusted pixel dwell times and pixel sizes for the particular imaging steps. As a starting point, image acquisition could be accelerated by the use of objective lenses with less magnification during preview and overview imaging. The FOV size A of a given objective is proportional to the inverse square of its magnification M, i.e.,

$$A \sim \frac{1}{M^2}.$$

An objective with just half the magnification thus allows for a quadruple of the acquisition area, assuming otherwise identical scan parameters. As low magnification objectives usually also feature smaller numerical apertures, smaller pixel sizes can be used for imaging without diminishing their resolving power. This property predestines the usage of these objectives especially for speeding up the preview imaging step where precise structural information is often not necessarily required. By using an objective with a larger FOV, less image acquisitions are required for the screening of a given sample area, which reduces the number of required stage movements and settling times. Moreover, a larger single image features less edge regions than a mosaic of smaller ones, which diminishes potential complications of the feature detection process on image boundaries.

The overall image acquisition time T furthermore scales linearly with the selected pixel dwell time t, i.e.,

$$T \sim t.$$

A reduction of the pixel dwell time to the lowest possible end results in a linear speed up of the entire acquisition workflow. Whereas for STED imaging the necessary dwell time may depend on various external factors like the labeling density, boundary conditions are often less strict for preview and overview acquisitions.

Not least, the attainable acquisition speed can be further enhanced by increasing the selected pixel size. Neglecting restrictions caused by the scan unit itself, the

Magnification / NA (FOV size)	Pixel size	Pixel per line scan	Min. dwell time
100X / 1.45	100 nm†	750	2.7 µs
($75 \times 75\,\mu m^2$)	200 nm	375	5.3 µs
	500 nm	150	13.3 µs
60X / 1.42	100 nm†	1250	1.6 µs
($125 \times 125\,\mu m^2$)	200 nm	625	3.2 µs
	500 nm	250	8.0 µs
20X / 0.85	180 nm†	2083	1.0 µs
($375 \times 375\,\mu m^2$)	500 nm	750	2.7 µs
	1000 nm	375	5.3 µs

Table 2.2: Attainable confocal pixel dwell times for different objectives and pixel sizes. In all scenarios, a maximum line frequency of 500 Hz was considered which meets the typical limit of the quad-scanner design. Approximate Nyquist pixel sizes for a wavelength of 600 nm are indicated by †. Effects like scanner flyback that would further slow down the overall scan frequency were not considered.

image acquisition time decreases quadratically with the chosen pixel size *s*, i.e.,

$$T \sim \frac{1}{s^2}.$$

This means that by doubling the pixel size, the identical FOV could be acquired in just a quarter of the measurement period. In reality, however, the magnification factor, pixel size and dwell time are interconnected through the maximum line frequency of the used scanner (Table 2.2). Assuming a line frequency of 500 Hz, the dwell times cannot be independently shortened for at the same time increasing pixel sizes. Whereas for Nyquist pixel sizes this dependency is of no practical importance, it becomes relevant under so-called undersampling conditions that proved to be preferable in the preview imaging step. When the shortest possible pixel dwell time is pursued, a further increase in the pixel size subsequently only leads to a linear speed increase by reducing the number of line scans along the slow scan axis. In contrast, for the fast scan axis, i.e., along the line itself, no acceleration is observed.

In the practical application, obtained intensity values must merely ensure proper target validation. In the present use case where centrosomal clusters were solely addressed in mitotic cells, this limit was already reached in the cell detection task. Test images with a 20X oil immersion objective revealed that a pixel dwell time of 1.5 µs is likely sufficient to identify mitotic cells based on DNA stained with DAPI. This minimum dwell time, though, was only accessible when a pixel size of 180 nm, i.e., near the Nyquist limit, was used. Considering the texture and morphology of condensed chromatin structures in dividing cells, this level of detail was however not required.

Nevertheless, an objective lens switch was initially considered as a feasible strategy to combine fast target validation and super-resolved STED imaging: preview imaging was envisaged to take place with the 20X objective to identify mitotic cells

of interest solely based on the DNA signal. After switching to a 60X objective, the intracellular features of interest were meant to be localized within confocal Z-stacks. Since the favored 60X/NA 1.42 oil immersion lens featured a state of the art numerical aperture, the same objective could be applied for final STED imaging without impairing resolution. At the same time, the accessible FOV is approximately 2.7 times larger as for a 100X objective. Thereby, larger cell assemblies of adeherent cells can be imaged within one FOV, which may be of importance in future applications. A potential pitfall for this approach could however be the occurrence of air droplets in the immersion oil during the lens change. For robust automated measurements, prior extensive testing would be required. Abberior Instruments chose a different approach and decided to exclusively use the 60X oil lens in the entire imaging process (Figure 2.6b). Features of interest were targeted by adjusting the mirror deflections of the quad-scanner. Hence, the sample did not need to be physically moved in between the overview and STED imaging step.

Beyond that, Abberior Instruments identified several additional cornerstones with respect to the final prototype design: for unsupervised measurements in multiwell plates, an oil dispenser served as a critical component for continuous oil supply without manual intervention. Moreover, a fast autofocus system was considered a key element to ensure stability over the expected measurement time of several hours to up to a few days per well plate. Unexpectedly, the settling time between consecutive image planes was in the order of one second when Z-stacks were acquired with the Olympus autofocus system in continuous mode[7]. This turned out to be a bottleneck with respect to the estimated acquisition times. Abberior Instruments hence developed an alternative focusing method that can be used simultaneously with the continuous autofocus system provided by Olympus. By using a deformable mirror (DM), the axial focus position could be shifted within the specimen while neither moving the stage nor the autofocus optics. This resulted in tremendously increased Z-stack acquisition speed which was decoupled from autofocus interference. On the downside, the accessible axial travel range of the DM was limited to $\pm 6\,\mu m$, resulting in a total available Z-range of only $12\,\mu m$ in the sample plane. For certain applications, this may restrict the general applicability of the system. Mitotic cells, for example, often already exhibit cell diameters of approximately $20\,\mu m$.

2.2.2 Abberior Instruments prototype test measurements

In confocal test imaging, the DM revealed no recognizable negative effects with regard to image quality or signal-to-noise ratio. For STED acquisitions, however, image formation is much more susceptible to aberrations evoked by the utilized optical elements (Booth et al., 2015). In this context, two effects are of particular interest: first, chromatic aberrations can lead to a reduced spatial overlap between the excitation and STED PSF. This frequently results in reduced intensities, especially in

---

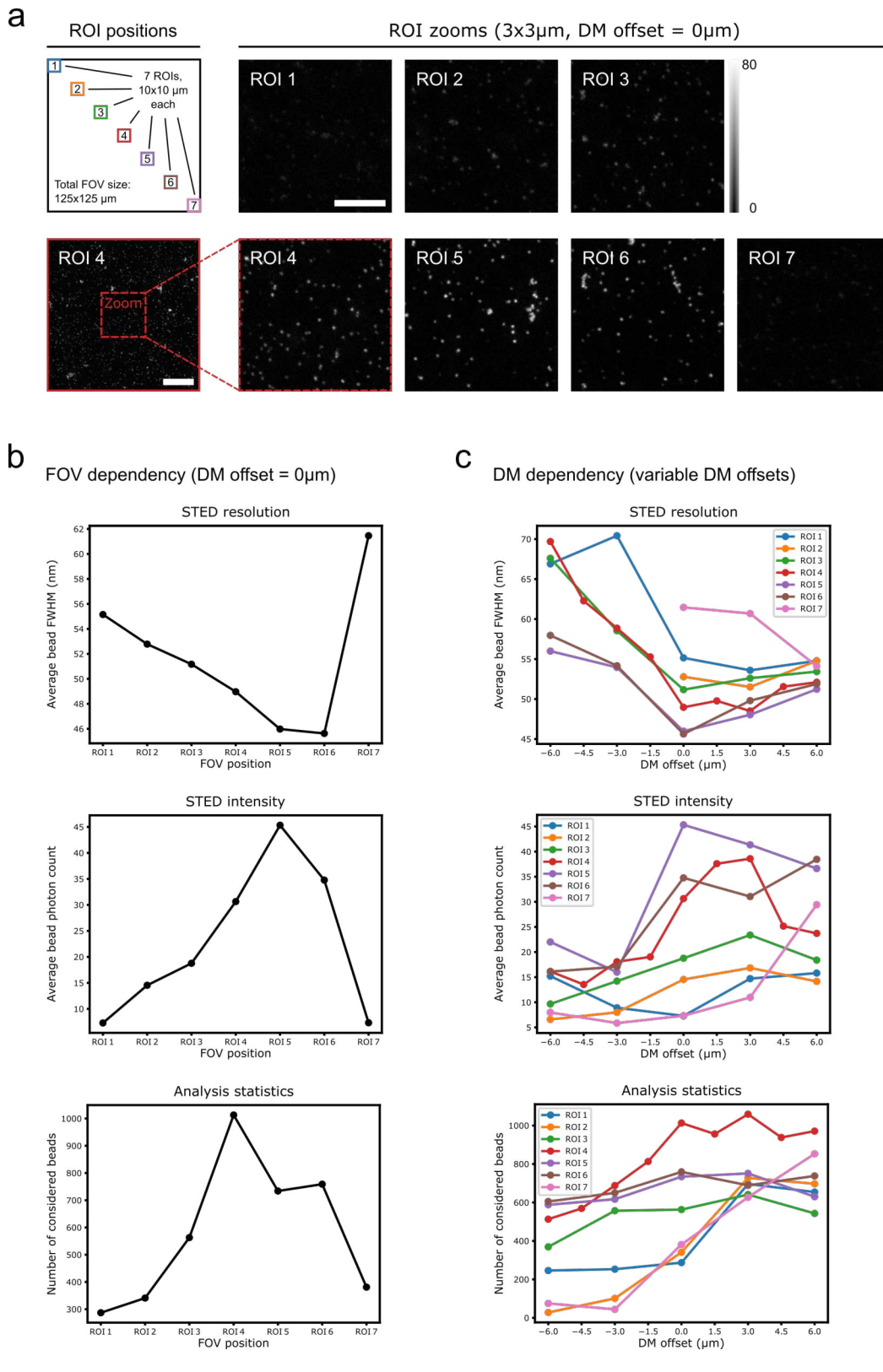[7]Value communicated by Abberior Instruments

Figure 2.7: Performance measurements of the Abberior Instruments automated STED prototype device. (a) Illustrative STED image quality at seven regions of interest (ROIs), measured by 40 nm fluorescent beads. (*continued on the following page*)

the outer regions of the FOV, where objectives often are less well corrected. Second, spherical and higher order aberrations like astigmatism could substantially impair the root of the STED PSF itself. Consequently, the zero of the STED PSF exhibits a substantial intensity level, thereby also unintentionally de-exciting fluorophores in its very center. Respective STED images tend to be dim and blurred with drastically reduced STED resolution. To compensate for this, usually adaptive optics like DMs and spatial light modulators (SLMs) are applied to ensure both optimal beam shaping and overlap. The prototype device incorporated both components, but utilized the DM for focusing instead of correcting aberrations. To benchmark the prototype's optical performance, STED images of features with a known size below the diffraction limit were acquired and analyzed[8].

Fluorescent beads with a nominal diameter of 40 nm were selected as an appropriate sample due to their high brightness and photostability. Images were acquired in seven regions of interest (ROIs) arranged along the diagonal of the accessible FOV (Figure 2.7a). Even for zero deflection - for which the DM was calibrated - huge discrepancies with respect to obtained intensity levels and resolutions were observed. Whereas beads imaged within central ROIs were crisp with sharp edges, those depicted in the outer regions of the FOV looked fuzzy and dim. The average full width at half maximum (FWHM) differed by up to 15 nm depending on the ROI position (Figure 2.7b). Observed intensity levels likewise dropped from about 45 counts near the center to less than ten counts at the FOV corners, i.e., by more than 80 %. Additionally, for all seven ROI positions the influence of selected DM offsets, i.e., the axial focus position z, was examined (Figure 2.7c). For all ROIs considered, an impairment of the imaging performance was observed, especially in the negative direction. Remarkably, also at the calibration position (ROI 4), the measured resolutions (FWHM increase by 20 nm) and intensity levels (average count rate decrease by 60 %) were substantially decreased for large deflections.

When using the prototype device, not only the accessible focus range was limited, but the imaging performance was compromised as well due to aberrations.

---

[8]Images were acquired in a joint measurement session between Abberior Instruments and me.

Figure 2.7: At all ROI positions, a $10 \times 10\,\mu m^2$ sized area was imaged with maximum STED laser intensity and varying deformable mirror (DM) offsets. Shown example images correspond to zero DM offset, i.e., no deflection from the calibration position, at all ROI positions. Scale bar 2 μm, zoomed inset 1 μm. (b) Measured average full width at half maximum (FWHM; top) and fluorescence intensity values (middle) at the seven specified ROI positions for zero DM offset. FWHM and intensity values were obtained by fitting a 2D-Gaussian function to each individual bead. In outer ROIs, separating beads from noise was challenging and lead to a reduced number of fits incorporated into the analysis (bottom). (c) The identical analysis performed for variable DM offsets ranging from $-6.0\,\mu m$ to $6.0\,\mu m$ in steps of $3.0\,\mu m$. In the center region (ROI 4), a finer sampling in steps of $1.5\,\mu m$ was performed. FOV: field of view.

The latter led to a non-isotropic optical performance, meaning that the size and signal strength of an imaged feature depend on its particular location inside the cell. Moreover, the expected acquisition times of the entirely confocal system exceeded the requirements laid down for an industrial application[9]. In concluding measurements, the prototype was evaluated by means of the centrosomal clustering use case at an external partner site. Herein, single-plane preview images were finally acquired with a pixel size of 450 nm and a dwell time of 4 µs to detect mitotic cells. This resulted in an acquisition time of 0.5 s per preview image. Assuming the screening of an entire 96-well plate, which is a plausible scenario due to the rare occurrence of mitotic cells, almost 200k preview images would be necessary to capture the entire well plate. In consequence, already the preview imaging step would take more than 27 hours while completely neglecting any delays caused by stage movements, settling times and machine learning-based cell detection. Confocal overview Z-stacks of mitotic cells were subsequently acquired by an external project partner using a pixel size of 130 nm, a pixel dwell time of 5 µs and two line accumulations. Assuming a FOV size of $25 \times 25 \, \mu m^2$ that usually fits a mitotic cell in lateral dimensions, acquisition of a single overview image already took about 370 ms. The recording of 26 slices with an interslice distance of 420 nm resulted in a total acquisition time of approximately 10 s for the Z-stack of a single cell. Estimating that the screening of at least 50 cells might be necessary for each applied condition, the sole feature extraction step would accumulate to more than 13 h of measurement time per well plate.

2.2.3 Custom camera-based acquisition scheme

As an alternative, I conceived an acquisition scheme in which widefield fluorescence imaging is applied for all tasks that do not call for finely resolved structural details (Figure 2.6c). By incorporating a fast camera-based detection unit into the periphery of an existing STED microscope, measurement times of all feature extraction steps can be drastically reduced. The usage of a 20X air objective allows for an approximately 38-times larger FOV during preview imaging compared to the prototype design[10]. At the same time, widefield acquisitions can be obtained already at exposure times in the order of 100 ms. The duration of preview imaging of an entire 96-well plate thus decreases to only 8.5 min, a striking speed increase by a factor of approximately 190. By applying a 100X oil immersion objective for overview Z-stacks, the still larger FOV permits scanning of multiple cells within a single image at likewise shorter acquisition times. Not least, since a fast piezo was considered for

---

[9]A maximum acquisition time of 24 h per 96-well plate was specified by Bayer AG. In reality, the attainable imaging speed significantly depends on the respective use case, e.g., on the occurrence of cells of interest, the average number of features being present in each cell, and their typical feature size. An exact *a priori* estimate is therefore often difficult.

[10]Assuming a field number of 22 with a corresponding quadratic FOV of $775 \times 775 \, \mu m^2$ in the sample plane.

Z-stacking, the travel range along the *z*-direction is practically unlimited under real-life conditions. Again assuming an acquisition time of 100 ms per single overview image and 20 slices per stack with a slice distance of 1 µm, the scan of a mitotic cell would last only about 2 s. This corresponds to a speed increase by a factor of five compared to the confocal approach. At the same time, features can be detected inside the entire cell irrespective of their axial position. The steps that were followed while implementing this camera-based approach are described in detail in the next section.

2.3 Automated STED imaging with camera-based feature detection for improved imaging speeds

2.3.1 Custom upgrade of a commercial STED-microscope for automated image acquisitions

To implement my custom camera-based acquisition scheme, I upgraded a commercial STED microscope in the laboratory of Prof. Stefan Jakobs at the University in Göttingen. The existing setup was an ExpertLine device from Abberior Instruments GmbH with four excitation lasers and a 3 W high-power 775 nm STED laser for confocal and STED imaging. It was equipped with a motorized XY stage and a LED illumination source for widefield fluorescence imaging. An available quad-band beam splitter matched the LED wavelengths and thus allowed sequential illumination without the need to switch between individual beam splitters in multi-color experiments, an otherwise very time consuming step. The LED source was not integrated into the Imspector software platform used for confocal and STED imaging, and support for systematic measurements of predetermined sample regions was lacking. However, the Imspector software provided a Python interface with which all motorized components of the microscope could be controlled electronically. For the present Olympus IX 83 body, this included the XY stage, the integrated autofocus system, the filter cube turret and the deflection unit - the latter guiding the optical beam path either to the microscope's tube or its side port. These conditions were indispensable prerequisites for the below-mentioned embedding of a camera detection unit for automated STED imaging.

To establish camera-based detection, the present binocular tube was exchanged for a trinocular tube exhibiting an additional camera port. A scientific complementary metal-oxide-semiconductor (sCMOS) camera was mounted to this port and controlled by means of a Python interface. Noteworthy, the camera had a switchable cooling fan, which proved essential to eliminate vibrations and thus image artifacts while imaging in STED mode. The XY stage was further upgraded with a piezo focusing system to enable fast and precise image stacking when the autofocus is disabled. The LED source was controlled via serial communication using USB from Python as well. All needed functionalities were joined in a custom graphical user
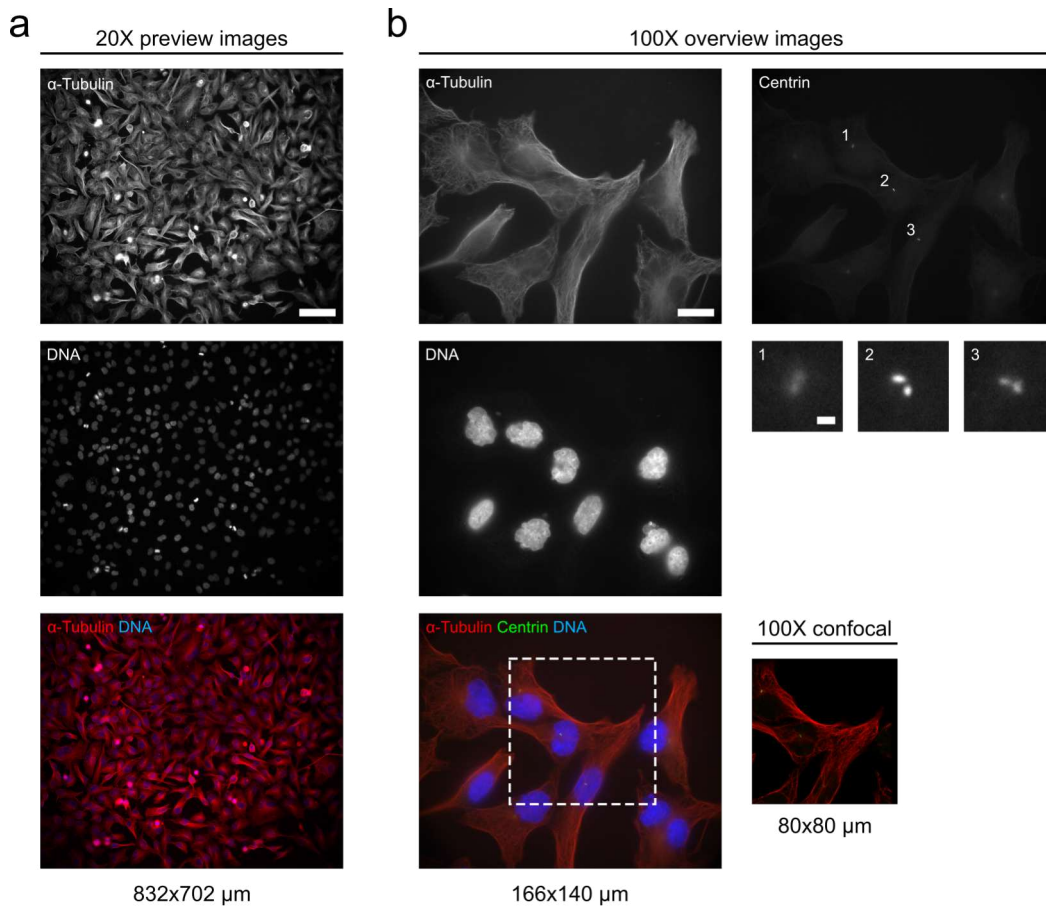
Figure 2.8: Widefield imaging enables robust cell and feature detection in preview and overview images. (a) Representative dual color preview images acquired with a 20X air objective lens. Mitotic cells are easily distinguishable over the entire field-of-view, despite reduced intensity levels near the edges. Scale bar 100 µm. (b) Three-color overview images acquired with a 100X high-NA oil objective lens. Centrosomal features can be reliably identified in the centrin channel when in focus (insets). For comparison, the same sample region was also imaged in confocal mode (indicated by the dashed box). Scale bar 20 µm, inset 1 µm. Samples were fixed BT549 cells stained for α-tubulin, centrin and DNA. Different color channels were sequentially acquired with a single quadband filter cube. Denoted image sizes are maximum accessible field-of-views for the respective objective magnification (full-frame acquisitions).

interface (GUI) in the Python programming language that comprised different imaging modalities and online image analysis. The sample wells, the accessible field of views (FOVs) and the color channels could be therein specified for widefield imaging. The latter involved also the respective excitation wavelengths, illumination intensities and exposure times of the camera. For preview and overview imaging, the lens magnification, autofocus setting and imaging type (i.e., Z-stack or not) could be set individually.

The widefield image quality regarding attainable intensity levels and sharpness

of object edges was evaluated in first test measurements. Acquired images confirmed that mitotic cells can be reliably detected in preview images using a 20X air objective lens (Figure 2.8a). Due to its large depth of field, both adherent interphase cells and spherical mitotic cells could be sharply resolved in a single image plane when the focus was set close to the substrate. In this scenario, the autofocus system can be used in continuous mode to eliminate focal drift. At the outer regions of the accessible camera field of view, a reduced illumination intensity was observed. This edge shading resulted from the mounted light source. For multicolor image acquisitions, different LED excitation wavelengths could be rapidly switched on and off in the microsecond range while using a single quadband beam splitter cube. For the present LED light source, acceptable exposure times were in the range of only 10 ms for the DNA and 100 ms for the α-tubulin channel. As a result, in combination with the large field of view, rapid preview imaging for cell detection was possible.

Widefield imaging was also evaluated concerning intracellular feature detection using a 100X oil objective lens (Figure 2.8b). This high-NA lens is a common choice for STED image acquisition and was initially also envisaged for preceding confocal cell imaging. Against expectations, centrosomal features could be clearly identified in the centrin channel if located in focus. Widefield imaging hence could be used for target identification, which led to a further speed improvement in comparison to confocal overviews. The required feature detection steps depend on the applied use case and were specifically designed and optimized for centrosomal clusters as described in the next sections.

2.3.2 Mitotic cell detection in widefield preview images

Narrowing down the total sample area to only a valuable subset of target regions is an important step for acquisition efficiency. In the present use case, only mitotic cells were of particular interest for the final screening application. This implies that both cell detection and classification are necessary to extract the mitotic subpopulation of cells. In the past years, deep learning algorithms evolved to be standard methods for many computer vision and biological image analysis problems (Moen et al., 2019). Modern algorithms are routinely trained end-to-end, meaning that raw data can be directly fed into the analysis pipeline without manually selecting specific image features. By training with datasets from different measurement campaigns, applied algorithms are in principle also able to cope with sample variabilities and inherent data inhomogeneities. Resulting feature extraction processes therefore have the potential to continuously improve over time. If generalization on newly acquired data can be achieved, this also obviates the need for manually adapting classical detection algorithms, which is often cumbersome and highly subjective.

For mitotic cell detection, cells of interest can be located at every position inside preview images without spatial restrictions. For such tasks convolutional neural networks (CNNs) are a common choice due to their invariance with respect to the
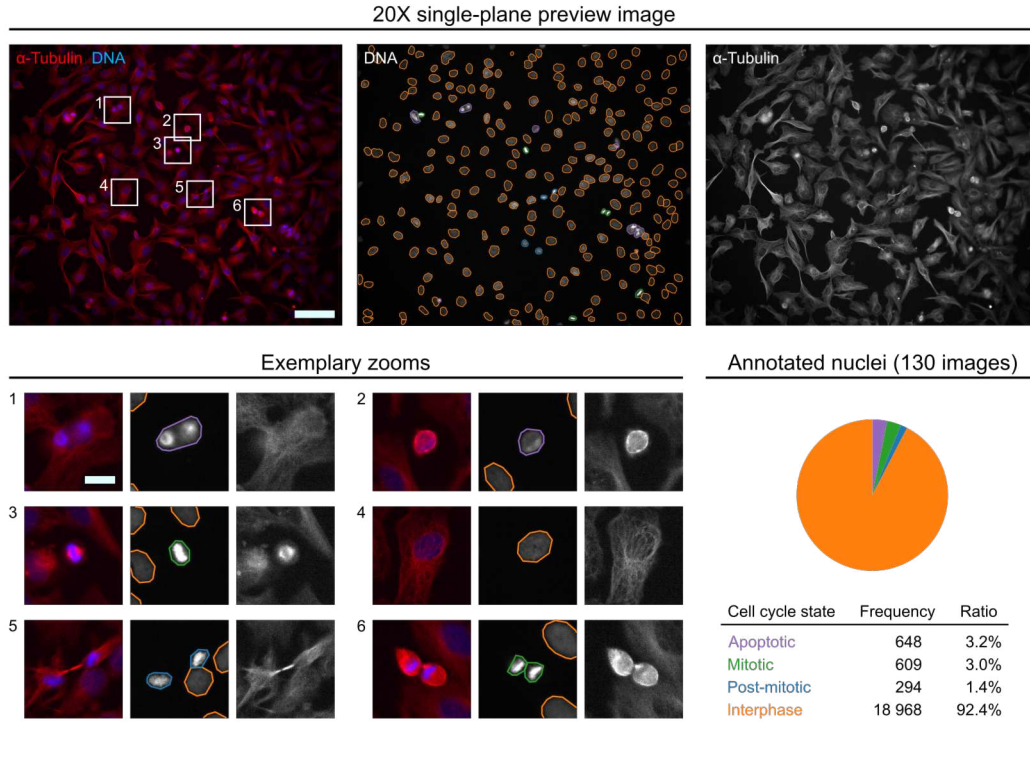
Figure 2.9: Preview widefield imaging for cell detection. Exemplary preview images of the *α*-tubulin and DNA channel were acquired with a 20X objective lens. Cell nuclei were segmented by classical image processing and then manually refined and annotated. Polygon contours shown in the DNA images are colorcoded in correspondence to their annotated cell cycle class (purple = apoptotic, green = mitotic, blue = post-mitotic, orange = interphase). In total, 130 preview images were prepared for AI training featuring substantial class imbalance (pie chart). Scale bar 100 µm, insets 20 µm.

position of the object. For this reason, RetinaNet - a state-of-the-art single-stage object detection algorithm - was applied for mitotic cell detection (Lin et al., 2017). A dataset of 130 single-plane preview images of untreated BT549 *wt* cells was acquired using a 20X air objective lens. The dataset comprised both the DNA and *α*-tubulin acquisition channel and was subsequently split 60-20-20 into training, validation and test datasets. Cell nuclei were automatically segmented and represented as enclosing polygons by a tailor-made, classical image analysis workflow, except for those reaching the image borders (Figure 2.9a). All polygons were then manually checked and corrected where required. Each nucleus was finally assigned to one of four predefined classes representing different cell cycle states, namely *interphase*, *mitotic*, *post-mitotic* and *apoptotic* (Figure 2.9b). Interphase cells showed mostly flat shapes with rather homogeneous intensity profiles between nuclei. In contrast, mitotic cells were characterized by a spherical morphology and condensed chromatin structures, the latter resulting in much higher intensity levels in the DNA channel. Presence of a spindle apparatus and alignment of chromosomes in a so-called metaphase plate

were unequivocal attributes for mitotic classification (Figure 2.9b, inset 3). Assignment to the apoptotic class was mainly driven by spherical cell morphologies without a clearly condensed DNA substructure or missing mitotic spindle (Figure 2.9b, insets 1,2). Further indicators were cellular blebbing, a typical sign of apoptosis, and high DNA intensity levels in larger blob-like patterns. Importantly, apoptotic annotations were not driven by a specific counterstaining, and in some cases, distinction between mitotic and apoptotic was challenging. In case of doubt, assignment as being apoptotic was preferred to maintain purity of the mitotic class and avoid excessive false-positive detection. Not least, in the post-mitotic class many chromatin structures were still condensed despite the already existing presence of two separated daughter cells. The DNA intensity levels were often comparable to the mitotic class. Decisive criterion was a bright α-tubulin signal at the so-called cleavage furrow typical for cytokinesis (Figure 2.9, inset 5).

In the first place, the RetinaNet was trained end-to-end with full class information of the interphase, mitotic and apoptotic class. To this end, polygon annotations at the pixel level were converted to bounding boxes that enclosed the particular nucleus. The post-mitotic class was omitted during training due to its smaller sample size and inherent class ambiguity. Training was performed in a transfer learning scenario using a ResNet50 backbone model with frozen weights that were pre-trained on the ImageNet dataset (Deng et al., 2009). For this, 16-bit raw images from the camera had to be mapped to 8-bit intensity range. Image normalization was performed in relation to average intensities of the interphase class (Appendix Table C.1). More precisely, for every annotated nucleus first its mean intensity was calculated by considering only its associated annotated pixels. These intensity values were then averaged for every class to obtain an estimate of typical intensity levels per class. Finally, the average intensities for the interphase class were multiplied by arbitrarily chosen factors of 20 and 10 to determine the maximum intensity values for the DNA and α-tubulin channel, respectively. The multiplication factors were set such that the resulting maximum intensity levels still exceed those of all other classes. Since the camera delivered a constant offset value of 100 counts, this value was chosen as the minimum threshold. In conclusion, image normalization thresholds $[I_{\min}, I_{\max}]$ of color channels were calculated relative to the respective means μ

of the interphase nuclei class according to

$$\text{DNA:} \qquad \mu_{\text{interphase}} = 143$$
$$[I_{\text{min}}, I_{\text{max}}] = [100, 100 + 43 * 20]$$
$$= [100, 960] \xrightarrow[\text{mapping}]{\text{8-bit}} [0, 255]$$

$$\alpha\text{-Tubulin:} \qquad \mu_{\text{interphase}} = 120$$
$$[I_{\text{min}}, I_{\text{max}}] = [100, 100 + 20 * 10]$$
$$= [100, 300] \xrightarrow[\text{mapping}]{\text{8-bit}} [0, 255].$$

In the final application, one then has to estimate the average interphase intensities in some test images for proper scaling. An alternative approach in which the 1 % and 99 % percentils of the entire training dataset were used for image normalization did not generalize on unseen data and was therefore not further investigated.

During training with full class information, the RetinaNet loss function was balanced by different class weights w in dependence of their respective sample size n_{class} according to

$$w_{\text{class}} = \frac{1}{n_{\text{class}}} \cdot \frac{n_{\text{tot}}}{N_{\text{classes}}}$$

Herein, N_{classes} is the number of considered classes and n_{tot} their total number of instances, leading to

$$w_{\text{interphase}} = \frac{1}{18968} \cdot \frac{20225}{3} \approx 0.36,$$

$$w_{\text{mitotic}} = \frac{1}{609} \cdot \frac{20225}{3} \approx 11.07,$$

$$w_{\text{apoptotic}} = \frac{1}{648} \cdot \frac{20225}{3} \approx 10.40.$$

Metrics used for evaluating model results were precision, recall and F1-score that are defined as

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}},$$

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}},$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}.$$

The F1-Score simply is the harmonic mean of precision and recall. For RetinaNet, these metrics were always evaluated at an intersection-over-union (IoU) value of 0.5, i.e., annotated and estimated bounding boxes must overlap by at least 50 % to count

Class	Metric	Test dataset
Interphase	Instances	3614
	$Prec_{IoU=.50}$	0.958
	$Rec_{IoU=.50}$	0.908
	$F1\text{-}Score_{IoU=.50}$	0.932
	$AP_{IoU=.50-0.95}$	0.692
Mitotic	Instances	140
	$Prec_{IoU=.50}$	0.588
	$Rec_{IoU=.50}$	0.714
	$F1\text{-}Score_{IoU=.50}$	0.645
	$AP_{IoU=.50-0.95}$	0.354
Apoptotic	Instances	158
	$Prec_{IoU=.50}$	0.190
	$Rec_{IoU=.50}$	0.437
	$F1\text{-}Score_{IoU=.50}$	0.264
	$AP_{IoU=.50-0.95}$	0.084

Table 2.3: Test results for RetinaNet trained with full class information. Training was performed on the α-tubulin and DNA imaging channels considering the interphase, mitotic and apoptotic nuclei classes while the post-mitotic class was omitted. The test dataset comprised 26 20X preview images. Abbreviations: Prec = precision, Rec = recall, AP = average precision, IoU = intersection-over-union.

for true positives. In addition, the average precision (AP) metric was considered to evaluate object detections also for different IoU values.

The results for a RetinaNet that was trained with full-class information showed a good performance with a precision of almost 96 % and a recall value above 90 % only for the interphase class (Table 2.3). However, precision and recall values for mitotic cell detection were not satisfactory. Especially a low precision value leads to time-consuming downstream STED imaging of misclassified cells. Interestingly, performance metrics for respective nuclei classes were comparable between training, validation and test datasets, indicating that the model appears to generalize (Appendix Table C.2). The corresponding loss curves were approaching a minimum plateau as expected and showed no sign of overfitting on the training data (Appendix Figure C.4a). The results for both underrepresented classes, i.e., the mitotic and apoptotic class, were underwhelming, even when considering class imbalances in the RetinaNet loss function. A possible reason for this is the rather small number of only 130 training images. Manual investigation of model estimates revealed that prediction errors typically arose during classification while basically all nuclei were detected irrespective of their underlying class. As a result, it appeared promising to split the problem into two independent tasks, namely general nucleus detection and subsequent classification. With this approach, the amount of training examples could be increased on the instance level despite the limited number of training images.

Class	Metric	Test dataset
Nucleus	Instances	3972
	$Prec_{IoU=.50}$	0.943
	$Rec_{IoU=.50}$	0.941
	$F1\text{-}Score_{IoU=.50}$	0.942
	$AP_{IoU=.50-0.95}$	0.715

Table 2.4: RetinaNet test results for nucleus detection. All nucleus annotations, i.e.,
interphase, mitotic, apoptotic and post-mitotic nuclei, were unified in a single nu-
cleus class. Training was performed on the α-tubulin and DNA imaging channels
with identical datasets as in the full-class scenario. Test data likewise comprised the
same 26 20X preview images as in the previous case. Abbreviations: Prec = preci-
sion, Rec = recall, AP = average precision, IoU = intersection-over-union.

Regarding nucleus detection, at first all annotated nuclei were summarized into
a single nucleus class. Then a RetinaNet model was trained on full-frame 20X pre-
view images to solely differentiate between the nucleus instances and unspecific
background. Both the α-tubulin and DNA detection channel were used for train-
ing. The corresponding results revealed high precision and recall values (Table 2.4)
and good agreement between training, validation and test datasets (Appendix Ta-
ble C.3). Loss curves likewise showed no discrepancy from the expected behavior
(Appendix Figure C.4b).

For nucleus classification, three established convolutional neural networks with
different architectures were tested, namely ResNet50 (He et al., 2016), MobileNetV2
(Sandler et al., 2018) and EfficientNetB0 (Tan and Le, 2019). Training was performed
only on the cropped nucleus parts of the preview images comprising both the α-
tubulin and DNA image channel. Nucleus crops were obtained by the smallest rect-
angles that fully enclosed the polygon annotations and therefore differed in size.
Hence, before training all crops had to be resized to the same target image size,
which acts as a hyperparameter. Only a fraction of interphase crops was used for
training to deal with class imbalances relative to the mitotic and apoptotic classes.
All models were initially loaded with weights pre-trained on the ImageNet dataset
with 1000 different classes. On top of the existing base model, a two-dimensional
global average pooling and a dense layer with three output neurons were added for
classifying between an interphase, mitotic or apoptotic cell nucleus. During an ini-
tial transfer learning period, only the top-layer network was trained for 30 epochs
with a specific learning rate. This learning rate is another hyperparameter and must
be carefully optimized for reliable training results. Afterwards the entire network
was optimized for another 20 epochs in a so-called fine-tuning period using a much
smaller learning rate of 1×10^{-5}. In ablation studies, optimal settings for both hy-
perparameters were established: a learning rate of 1×10^{-3} during transfer learning
proved suitable for all applied networks (Appendix Figure C.1). Remarkably, all
models reached similar performance levels after fine-tuning, regardless of previous

Class	Metric	Test dataset
Interphase	Instances	190
	Prec	0.93
	Rec	0.98
	F1-Score	0.95
Mitotic	Instances	125
	Prec	0.90
	Rec	0.85
	F1-Score	0.87
Apoptotic	Instances	136
	Prec	0.85
	Rec	0.82
	F1-Score	0.84

Table 2.5: ResNet50 test results for nucleus classification. ResNet50 was trained with an optimal target image size of 100×100 pixels and a learning rate of 1×10^{-3} in the transfer learning step. Only annotated nucleus crops from the interphase, mitotic and apoptotic class were used for training and comprised both the α-tubulin and DNA channel. Abbreviations: Prec = precision, Rec = recall.

variations due to distinct learning rates. A target image size of 100×100 pixels lead to most stable training results for all evaluated models (Appendix Figure C.2). In the direct comparison between models, the ResNet50 architecture performed best with respect to both the achieved loss and the accuracy levels (Appendix Figure C.3). Analysis of the test dataset revealed improved classification results for the optimal ResNet50 model when compared to the RetinaNet previously trained with full class information (Table 2.5). Precision and recall metrics for the apoptotic and, more importantly, also for the mitotic nucleus class were significantly higher. For interphase nuclei, comparable performance values were achieved despite their lower relative ratio in the training dataset. Comparison of the training, validation and test datasets showed only a minor decline in obtained precision and recall values between the training and validation/test results and indicates proper generalization (Appendix Table C.4). Analysis of ResNet50 class predictions for post-mitotic nuclei revealed that these nuclei were classified to all available classes with some preponderance towards the interphase class (Appendix Table C.5). This emphasizes the underlying heterogeneity within the post-mitotic nucleus class which was therefore not considered for training.

In conclusion, both models were combined into a joint nucleus detection and classification approach. Compared to the single RetinaNet model trained with full-class information, the combined approach shows overall better performance for the test dataset (Table 2.6). Especially the precision for mitotic nucleus detection improved by more than 30 %, whereas the recall decreased at the same time by only 7 %. In view of the time-consuming subsequent cellular and intracellular imaging steps, precise mitotic cell detection is the most crucial step for attainable imaging

Class	Metric	Test dataset
Interphase	Instances	3614
	$\text{Prec}_{\text{IoU}=.50}$	0.950
	$\text{Rec}_{\text{IoU}=.50}$	0.907
	$\text{F1-Score}_{\text{IoU}=.50}$	0.928
	$\text{AP}_{\text{IoU}=.50-0.95}$	0.681
Mitotic	Instances	140
	$\text{Prec}^{\text{IoU}=.50}$	0.899
	$\text{Rec}^{\text{IoU}=.50}$	0.636
	$\text{F1-Score}^{\text{IoU}=.50}$	0.745
	$\text{AP}^{\text{IoU}=.50-0.95}$	0.400
Apoptotic	Instances	158
	$\text{Prec}_{\text{IoU}=.50}$	0.418
	$\text{Rec}_{\text{IoU}=.50}$	0.405
	$\text{F1-Score}_{\text{IoU}=.50}$	0.412
	$\text{AP}_{\text{IoU}=.50-0.95}$	0.133

Table 2.6: Test results for nucleus detection and classification in a combined approach: a RetinaNet model trained on single nucleus class data was applied for object detection, followed by a ResNet50 model for nucleus classification. The ResNet50 was solely trained with nucleus crops from interphase, mitotic and apoptotic nucleus classes. The test dataset comprised 26 20X preview images and was identical to the RetinaNet training scenario with full class information. Abbreviations: Prec = precision, Rec = recall, AP = average precision, IoU = intersection-over-union.

speeds: false-positive detections waste the limited measurement time due to acquisition of futile imaging data during overview and STED acquisitions. False-negative events, in contrast, only reduce the available sample size from which cells of interest can be selected afterwards. In addition, the precision metric for apoptotic cell detection doubled without substantial impairment for the corresponding recall. For the interphase class, almost identical results could be obtained, once again stressing the point that the combined approach improved primarily underrepresented instance classes. When comparing training, validation and test data, the obtained metrics showed no major fluctuations between the datasets (Appendix Table C.6). This once more confirms the reliable training performance despite limited data basis.

To use the model in the final imaging application, newly acquired preview data must be appropriately normalized by the average intensity values of the interphase nucleus class. Based on this, robustness of model predictions was further evaluated with respect to deviating normalization thresholds. The nominal thresholds of 960 and 300 counts for the DNA and α-tubulin channel, respectively, were altered by 25 % either in the same or opposite direction. As expected, the resulting metrics showed considerably weaker model performance for differing normalizations (Appendix Table C.7). However, for all scenarios the achieved results were still better as for the single RetinaNet model. This highlights the need for accurate normalization

when the established cell detection approach will be used in first proof-of-concept measurements. For future perspective, however, a much larger and more diverse training dataset that also includes broader intensity ranges is highly desirable. In contrast to the bare RetinaNet model, the established approach is not a single-stage detection model anymore and therefore less efficient. Every detected nucleus must pass through the ResNet50 classification model which slows down the entire cell detection and thus also the imaging pipeline. Despite the limited number of available training images, it was nevertheless possible to apply the combined model approach in the final automated STED imaging pipeline for first proof-of-concept measurements.

2.3.3 Intracellular feature detection in widefield overview images

After identification and precise positioning of mitotic cells of interest, widefield Z-stacks were acquired utilizing a 100X high-NA objective lens for intracellular feature detection. Centrosomal features can be located almost everywhere inside a mitotic cell, thus the entire cell must be screened to detect all centrosomal features. When heading for the most universal implementation, again a CNN-based object detection and segmentation algorithm should be applied to specify individual target structures on the single pixel level, e.g., by training a Mask-RCNN model (He et al., 2017). For robust feature detection, however, sufficient training data and accurate annotations are absolutely mandatory. Direct comparison between flat interphase and spherical mitotic cells revealed substantially higher background levels in mitotic cells. The generation of training data, i.e., the acquisition of widefield Z-stacks solely of mitotic cells, served as a bottleneck due to time constraints such that CNN-based feature detection could not be realized under the given circumstances. However, in the case of centrosome detection CNN-based procedures were not at all required for precise feature extraction. Centrosomes appeared as dot-like structures in fluorescent images and could therefore easily be identified by classical spot detection algorithms.

In the present use case, widefield Z-stacks of mitotic cells were imaged with a slice distance of $1\,\mu m$ which was sufficient to not miss individual features. In the final procedure, Z-stack images were first combined to a single maximum intensity projection image (Figure 2.10). For centrosome detection, only the centrin channel was used and normalized by the maximum intensity[11]. The center coordinates (x, y) of individual centrosomes were then determined by a difference of Gaussians (DOG) blob detection algorithm, an approximation of the more general Laplacian of Gaussian (LOG) approach. The settings were manually chosen after analysis of ten widefield Z-stacks containing mitotic cells that served as exemplary test data. Blob

---

[11]This normalization strategy assumes that there is at least one centrosomal feature present in each widefield stack. In the given scenario, this was practically always the case for full-frame acquisitions using a 100X objective lens. Under differing conditions, more elaborate normalization strategies might be necessary.
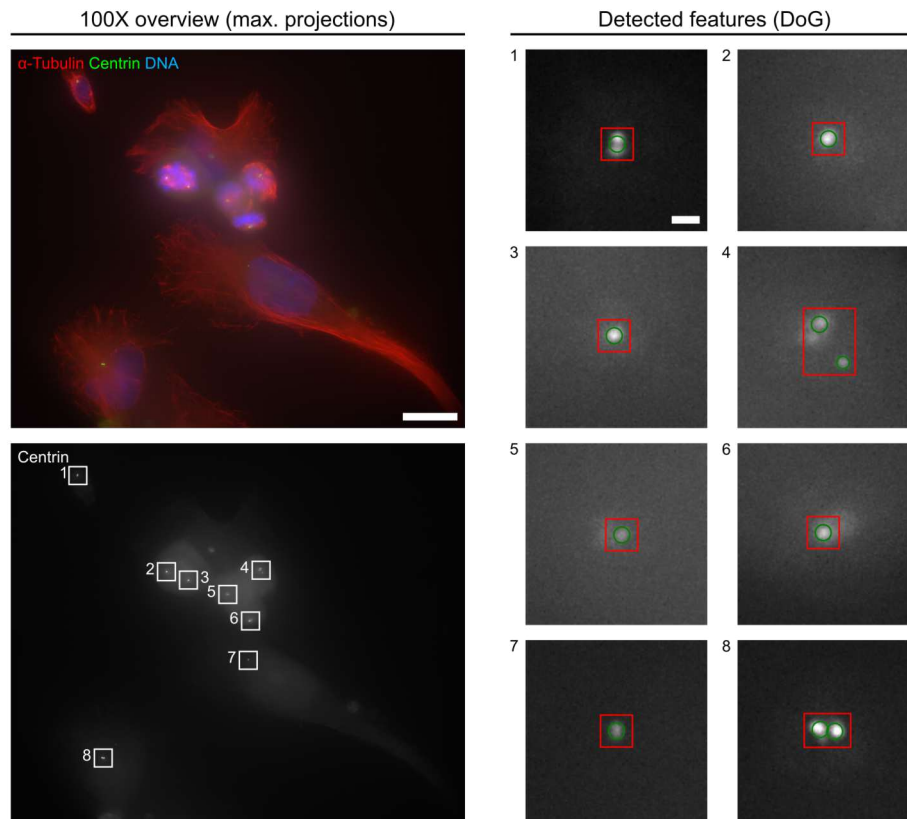
Figure 2.10: Centrosomal feature detection in widefield imaging Z-stacks. Widefield Z-stacks were acquired using a 100X oil objective lens and the projected into a single maximum intensity projection image (left). Centrosomal features were identified by a difference of Gaussians (DOG) blob detection algorithm on the normalized centrin projection image. Insets show found feature coordinates (green circles) and calculated bounding boxes (red squares, right). Features were merged into clusters if their separation distance was below a predefined threshold (insets 4 and 8). Note that with this procedure only feature coordinates (x, y) could be determined while for identification of corresponding z-coordinates the entire Z-stack must be analyzed. Scale bar 20 µm, insets 1 µm.

detection provides besides center coordinates also a radius measure which was used for bounding box generation (Figure 2.10). Identified spots were combined into clustered features if the separation distance between their center coordinates was below a certain threshold. In the present case dealing with centrosomes, a maximum distance of 2 µm proved reasonable by means of visual inspection. Established bounding boxes were afterwards propagated through the entire Z-stack while calculating their respective average intensity levels for all slices. The z-position of the slice at which the maximum intensity level was measured was set as z-coordinate for every feature individually. Derived (x, y, z)-coordinates could thereafter be used for subsequent STED image acquisition. For this, robust matching between widefield and confocal imaging coordinates is an essential prerequisite which was established as described in the next section.

2.3.4   Linking widefield and confocal/STED image coordinates

When preview or overview images are acquired in camera-based widefield mode, it is necessary to accurately position the identified features of interest in the confocal coordinate system for subsequent confocal or STED imaging. At the same time, the field of view size used for STED imaging should be set as small as possible as this step is by far the most time-consuming in the whole workflow. Both preconditions require that feature coordinates can be precisely transformed from the widefield to the confocal/STED image coordinate system. Although the identical 100X objective lens is used for both imaging modalities in the present case, their coordinate systems do usually not coincide. Possible reasons for this are manifold: Obviously, the camera system in use may be mounted with some rotation angle or lateral shift relative to confocal mode. Moreover, the origin of the confocal coordinate system can receive some lateral offset in the course of scanner calibration that must be taken into account.

For the implemented camera system, the resulting field of view was much larger for widefield than for confocal imaging. As a consequence, not all coordinates that are displayed in a widefield image can also be reached by laser scanning in confocal mode. For complete accessibility, features of interest that were found in widefield images must therefore be repositioned by means of lateral stage movements. Though this would not be necessary for all widefield coordinates, STED image acquisition at the center of the confocal field of view offers major advantages. The confocal center typically matches with the optical axis of the used objective lens where usually also the best optical performance is obtained. In addition, spatial light modulators (SLM) that compensate for optical abberations are also calibrated for optimal correction in the FOV center. The intended workflow therefore ensures STED image acquisition with optimal and identical image quality, i.e., resolving power, for all features of interest. Losses in imaging speed due to stage shifts are acceptable as traveling distances are short. For axial positioning, i.e., focusing of sample features, the piezo stage insert was used due to its superior accuracy and speed when compared with the microscope body's Z-drive.

Accurate translation of feature coordinates required two steps: At first, the origin of the confocal imaging field of view had to be specified in the widefield coordinate system. These center coordinates, in the following denoted (x_0, y_0), served then as target coordinates for repositioning of centrosomal features in the widefield coordinate system. In the second step, the underlying transformation between the XY stage and camera frame was derived. In doing so, both the rotation angle between camera and stage and deviations between nominal and real stage movements were revealed that are essential for proper positioning.

For matching coordinate systems, initially ten image pairs of different cellular α-tubulin structures were acquired in widefield and confocal mode. In both modes, an identical pixel size of 65 nm was used without intermediate stage movements. Correlating image points that originate from identical sample positions, the
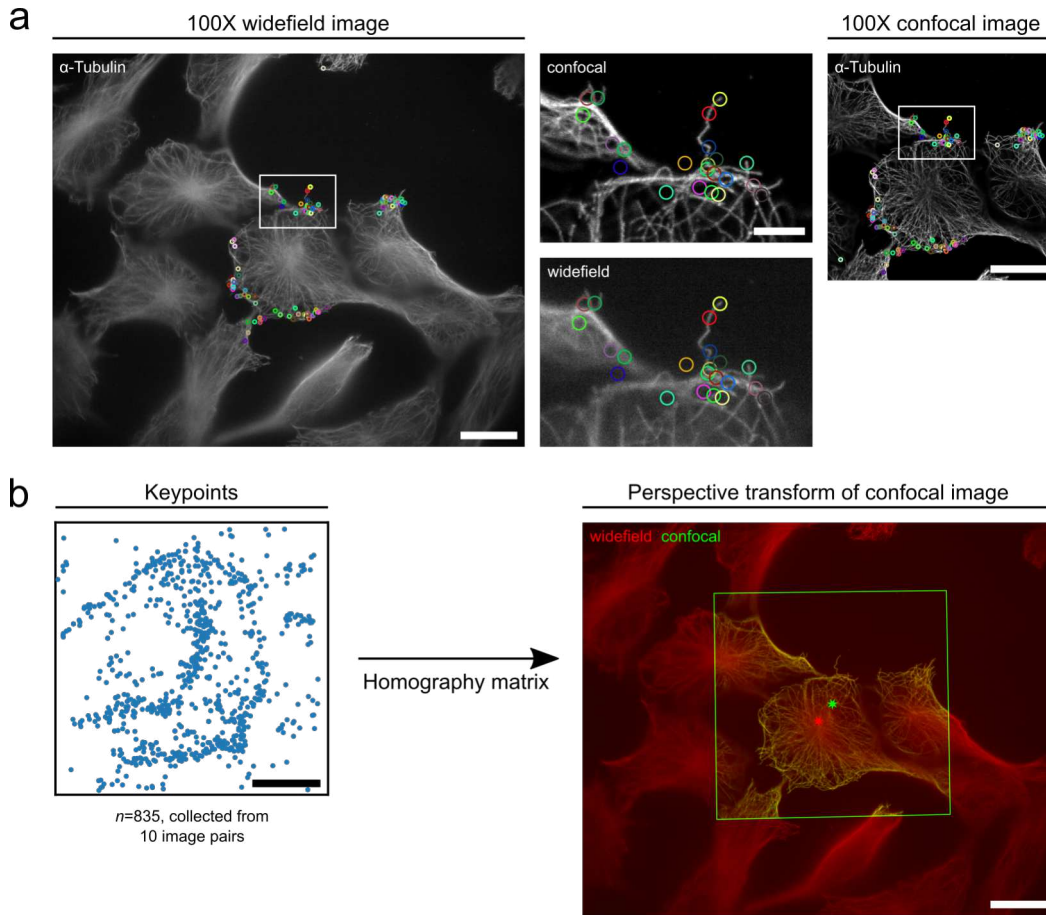
Figure 2.11: Image registration between widefield and confocal imaging systems. (a) Single exemplary image pair of a widefield and confocal α-tubulin fluorescence image. Both images were acquired with a 100X high-NA oil objective lens at the same sample position. Colored circles illustrate matching keypoints as found by a scale-invariant feature transform (SIFT) algorithm (insets). (b) Keypoints from 10 image pairs were collected and used for deriving a homography matrix between both image coordinate systems. With this, a perspective transform of the confocal image can be calculated and overlaid with the widefield. Note that the origin of the confocal field of view (green asterisk) is off-center with respect to the widefield image (red asterisk). Scale bars 20 μm, insets 5 μm.

so-called keypoints, were then identified by applying a scale-invariant feature transform (SIFT, Lowe, 2004) for each image pair (Figure 2.11a). Joint keypoints were routinely found at cell edges, where microtubules appear sharp and with high contrast also in widefield acquisitions. The described scenario corresponds to a situation where the same sample plane is imaged from potentially different viewing angles. Hence, keypoints from all ten image pairs were combined and used to calculate the
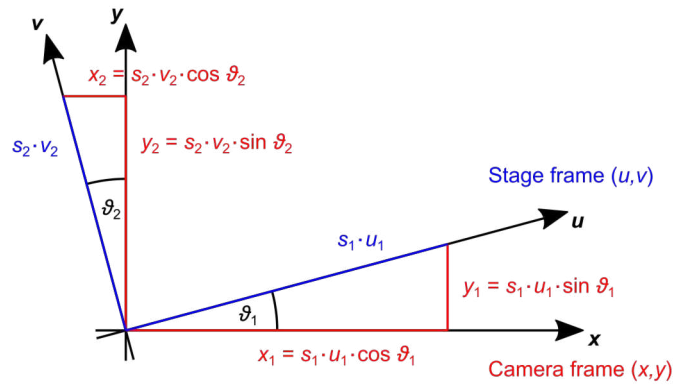
Figure 2.12: Alignment between XY stage and camera imaging frame. Stage movements in u and v directions lead to shifts in x and y in acquired widefield images. By moving the stage only along a single axis, i.e., $(u = u_1, v = 0)$ or $(u = 0, v = v_2)$, the corresponding rotation angles θ_1 and θ_2 and scaling factors s_1 and s_2 could be derived independently.

corresponding homography matrix H (Figure 2.11b). Applying this matrix, the corresponding perspective transform[12]

$$
\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}
$$

was calculated that transforms confocal image points $[x, y]^T$ onto their widefield analogues $[x', y']^T$. Transformation of the center coordinates of the confocal imaging field of view then led to their widefield counterpart, the target coordinates $[x_0, y_0]^T$.

The relationship between identified coordinates in the widefield imaging frame and the needed stage movements for positioning each coordinate at the target coordinate $[x_0, y_0]^T$ was implemented in a similar fashion. Assuming a general scaled rotation transform between the camera frame (x, y) and the stage frame (u, v), the corresponding rotation angles $\theta_{1/2}$ and scaling factors $s_{1/2}$ can be different for the u and v, respectively (Figure 2.12). This assumption reflects the possibility that the stage's directions of travel (u, v) may not be perfectly perpendicular to each other and may also feature differing deviations between nominal and actual traveled distances. When both coordinate systems share the same origin, moving the stage by a

---

[12]Assuming that the widefield and confocal imaging planes are not tilted relative to each other, an affine transform would suffice for valid image registration. However, the more general perspective transform was already implemented in the applied OpenCV library and therefore used for sake of simplicity.

distance (u, v) results in observed movements (x, y) in the widefield image according to

$$
\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = RS \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_1 & -\sin\theta_2 & 0 \\ \sin\theta_1 & \cos\theta_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}
$$

$$
= \begin{bmatrix} s_1 u \cdot \cos\theta_1 - s_2 v \cdot \sin\theta_2 \\ s_1 u \cdot \sin\theta_1 + s_2 v \cdot \cos\theta_2 \\ 1 \end{bmatrix}
$$

where again homogeneous coordinates were used. Herein, R denotes the rotation and S the scaling matrix. By applying stage movements either solely in u or v directions, this set of equations can be decoupled and solved for θ_1, θ_2, s_1 and s_2 by measuring (x_1, y_1) and (x_2, y_2):

$$
s_1 = \frac{\sqrt{x_1^2 + y_1^2}}{u}
$$

$$
s_2 = \frac{\sqrt{x_2^2 + y_2^2}}{v}
$$

$$
\theta_1 = \tan\frac{y_1}{x_1}
$$

$$
\theta_2 = \tan\frac{x_2}{y_2}
$$

To this end, widefield image pairs were acquired for nominal stage movements of $u = 100\,\mu\text{m}$ and $v = 70\,\mu\text{m}$, respectively (Figure 2.13). Related keypoints were again identified by SIFT analysis to calculate the corresponding homography matrices. By applying the respective perspective transforms, the actually traveled distances were measured as $x_1 = 100.84\,\mu\text{m}$, $y_1 = 1.07\,\mu\text{m}$, $x_2 = 0.86\,\mu\text{m}$ and $y_2 = 70.23\,\mu\text{m}$. The corresponding rotation angles amounted to $\theta_1 \approx 0.61°$ and $\theta_2 \approx 0.70°$, thus indeed indicating a small dependence between both directions of motion. By using $\sqrt{1 + x} \approx 1 + \frac{x}{2}$ for $x \ll 1$ which holds true for the observed small angles, the scaling factors can be further simplified to $s_1 \approx \frac{x_1}{u}$ and $s_2 \approx \frac{y_2}{v}$.
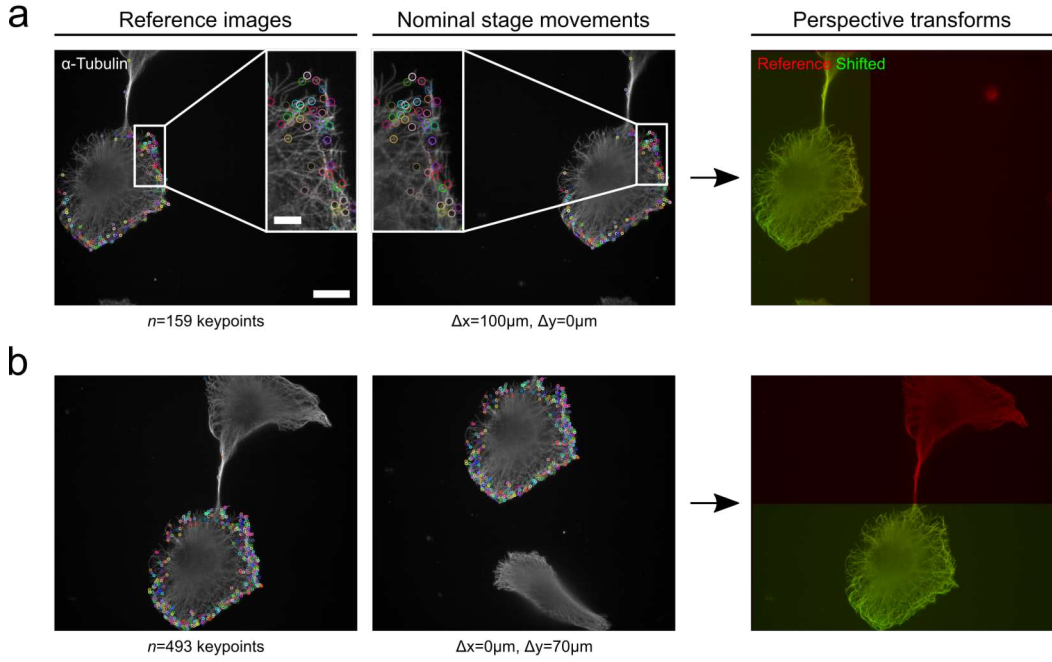
Figure 2.13: Analysis of lateral stage movements for feature positioning. 100X widefield image pairs of α-tubulin were acquired for nominal stage movements of (a) 100 µm in x-direction, and (b) 70 µm in y-direction. Matching keypoints (colored circles) were determined by a scale-invariant feature transform (SIFT) and used for homography calculation. Subsequently, corresponding perspective transforms were calculated that transform the shifted image back to the reference. Note that stage movements in x and y are not decoupled from each other with regard to the widefield image coordinate system. Scale bars 20 µm, insets 5 µm.

In the final application, however, observed feature coordinates must be translated to nominal stage movements which requires the inverse transform, i.e.,

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = (RS)^{-1} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = S^{-1} \cdot R^{-1} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$= \frac{1}{\det R} \begin{bmatrix} \dfrac{x}{s_1} \cos \theta_2 + \dfrac{y}{s_1} \sin \theta_2 \\ -\dfrac{x}{s_2} \sin \theta_1 + \dfrac{y}{s_2} \cos \theta_1 \\ \cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 \end{bmatrix}$$

with the inverse scaling and rotation matrices

$$S^{-1} = \begin{bmatrix} \frac{1}{s_1} & 0 & 0 \\ 0 & \frac{1}{s_2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$R^{-1} = \frac{1}{\det R} \cdot \begin{bmatrix} \cos\theta_2 & \sin\theta_2 & 0 \\ -\sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 \end{bmatrix}.$$

and $\det R = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2$ being the determinant of R. Applying the small-angle approximation $\theta \ll 1 \Rightarrow \sin\theta \approx \tan\theta \approx \theta, \cos\theta \approx 1$, this simplifies for the present small rotation angles further to

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{s_1} \cdot (x + y \cdot \frac{x_2}{y_2}) \\ \frac{1}{s_2} \cdot (y - x \cdot \frac{y_1}{x_1}) \\ 1 \end{bmatrix}.$$

After replacing x for $\Delta x = x - x_0$ and y for $\Delta y = y - y_0$ and inserting the above derived scaling approximations, the stage movements Δu and Δv that are required for positioning a widefield coordinate in the center of the confocal field of view are given by

$$\Delta u = \frac{u}{x_1} \cdot (\Delta x + \Delta y \cdot \frac{x_2}{y_2}),$$

$$\Delta v = \frac{v}{y_2} \cdot (\Delta y - \Delta x \cdot \frac{y_1}{x_1}).$$

When applied for automated measurements, this calibration between the widefield and confocal imaging system worked reliably, such that the FOV size for subsequent STED acquisition could bet routinely set as little as $2 \times 2\,\mu m^2$. Only in few cases, centrosomal features appeared truncated in high-resolution images depending on the exact centriole feature that was found by the blob detection algorithm. Additionally, once calibrated, the transform was robust and stable without the need for recalibration over a period of six month. Moreover, the presented alignment procedure can potentially be fully automatized due to the trustworthy coordinates provided by the SIFT algorithm. The whole imaging cascade finally concludes with confocal and/or STED image acquisitions that were implemented as described in the following section.

2.3.5 Confocal and STED imaging of centrosomal features

For final super-resolution imaging, predefined measurement templates were used that could be specified in the ExpertLine Imspector software platform without restrictions. Depending on the measurement scenario, a suitable template was selected

in the custom-build Python GUI at the beginning of each measurement. If centrosomal features were detected by the blob detection algorithm during widefield imaging, the features' (x, y, z)-coordinates were cached and successively positioned in the center of the confocal field of view. The templates allowed for both single-plane and Z-stack acquisitions and specified already a minimum field of view size for STED imaging. In case that multiple features were combined into a single blob, the field of view size was correspondingly adjusted. All other acquisition parameters were set automatically according to the template.

Unexpectedly, the switch from the widefield to the confocal acquisition mode was quite slow due to hardware restrictions. Once started in the software, the rotations of the filter cube turret and the deflection unit took approximately three seconds and slowed down the successive ExpertLine image acquisition. The exact impact depends heavily on the present sample structure: if a single cell contains only few features of small size, the loss of speed is obviously of greater importance than if a large number of cellular targets can be imaged in STED mode right after a single switch. Importantly, the fan of the widefield camera must be turned off during STED imaging as otherwise the image becomes corrupted by vibration artifacts.

After building and benchmarking the individual feature detection and imaging parts, the entire workflow was finally tested and evaluated by a fully automated proof of concept experiment as detailed in the final results section.

2.4   Automated proof of concept measurements

Ultimately, a first automated proof of concept (POC) STED screen was conducted to evaluate the developed widefield imaging scenario for feature detection. Many potential protein targets to tackle cells with abnormal centrosomal content have already been described in literature (Sabat-Pośpiech et al., 2019). The majority of these proteins play a role in the clustering of supernumerary centrosomes and in many cases small molecule inhibitors have already been successfully developed. These molecular targets can be divided into three main groups, consisting of proteins that are either part of the spindle apparatus or the centrosome itself, those that organize the cortical actin network, and those that are involved in the spindle-assembly checkpoint (SAC) (Kwon et al., 2008). Based on literature research, the following ten active ingredients (Table 2.7) were selected as they previously showed an effect on different target proteins in the context of centrosome amplification (CA): Griseofulvin (Rebacz et al., 2007) and reduced-9-bromonoscapine (Red-Br-Nos) interact directly with microtubules and demonstrated declustering of centrosomes in murine N1E-115 neuroblastoma cells, thereby inducing multipolar mitosis followed by mitotic arrest and apoptosis (Pannu et al., 2014). Aurora-A kinase localizes to centrosomes in the interphase and remains associated with spindle poles during mitosis

Active ingredient	Target protein	Target mechanism
Griseofulvin	Microtubule	Spindle apparatus
Reduced-9-bromonoscapine	Microtubule	Spindle apparatus
Alisertib	Aurora-A	Centrosome
Crenolanib	Cofilin	Cortical actin
CP-673451	Cofilin	Cortical actin
AZ0108	PARP1/2/6	DNA repair
PJ-34	PARP1/2	DNA repair
Niraparib	PARP1/2	DNA repair
Talazoparib	PARP1/2	DNA repair
Centrinone-B	Plk-4	Centrosome inactivation

Table 2.7: Active ingredients applied in the first automated proof-of-concept screen. Substances are described to address different molecular targets and were chosen according to indicated effects on centrosomal clustering in previous studies. Abbreviations: PARP = poly-(ADP-ribose)-polymerase, Plk = polo-like kinase.

(Crane et al., 2004). As Aurora-A kinase activity promotes the clustering of overamplified centrosomes, Alisertib was included in the set of test compounds as a selective inhibitor that leads to multipolar spindles, prolonged mitosis and subsequent cell death (Navarro-Serer et al., 2019). The cofilin activators Crenolanib and CP-673451 were selected because they trigger spindle multipolarity in U2OS cells with induced CA (Konotop et al., 2016). Furthermore, AZ0108, PJ-34, Niraparib and Talazoparib - four inhibitors of the poly-(ADP-ribose)-polymerase (PARP) enzyme family - were added to the test set. AZ0108 produces a multi-polar spindle phenotype in a high-content screening of HeLa cells (Johannes et al., 2015). PJ-34 leads to similar, but less pronounced centrosomal declustering as compared to Griseofulvin and Red-Br-Nos (Pannu et al., 2014). Niraparib (Sisay and Edessa, 2017) and Talazoparib (Exman, Barroso-Sousa, and Tolaney, 2019) are recently approved anti-cancer drugs that to my knowledge have not yet been systematically studied regarding their effect on centrosomal declustering. Finally, Centrione-B, a polo-like kinase 4 (Plk-4) inhibitor, was added, which holds potential as an alternative way to target cells with CA. In previous studies, it depleted the number of centrosomes in various cancer cell lines by preventing the formation of daughter centrioles,, including BT549 used in this study (Wong et al., 2015).

All ten drug molecules were used in concentrations of 1 µM and 10 µM. The cells were incubated with the respective drug for 48 h prior to fixation and staining[13]. After sample preparation, the entire imaging pipeline was applied for automated image acquisition. As there was no immersion oil dispenser available at the demonstration setup, POC-measurements could not be conducted in multiwell plates. Considering the relevant area of approximately 80 cm^2 over which wells would have to be imaged, the oil film would get lost after a small number of wells without sustained

---

[13]All compounds were provided by the internal Bayer AG compound library, but are also commercially available.
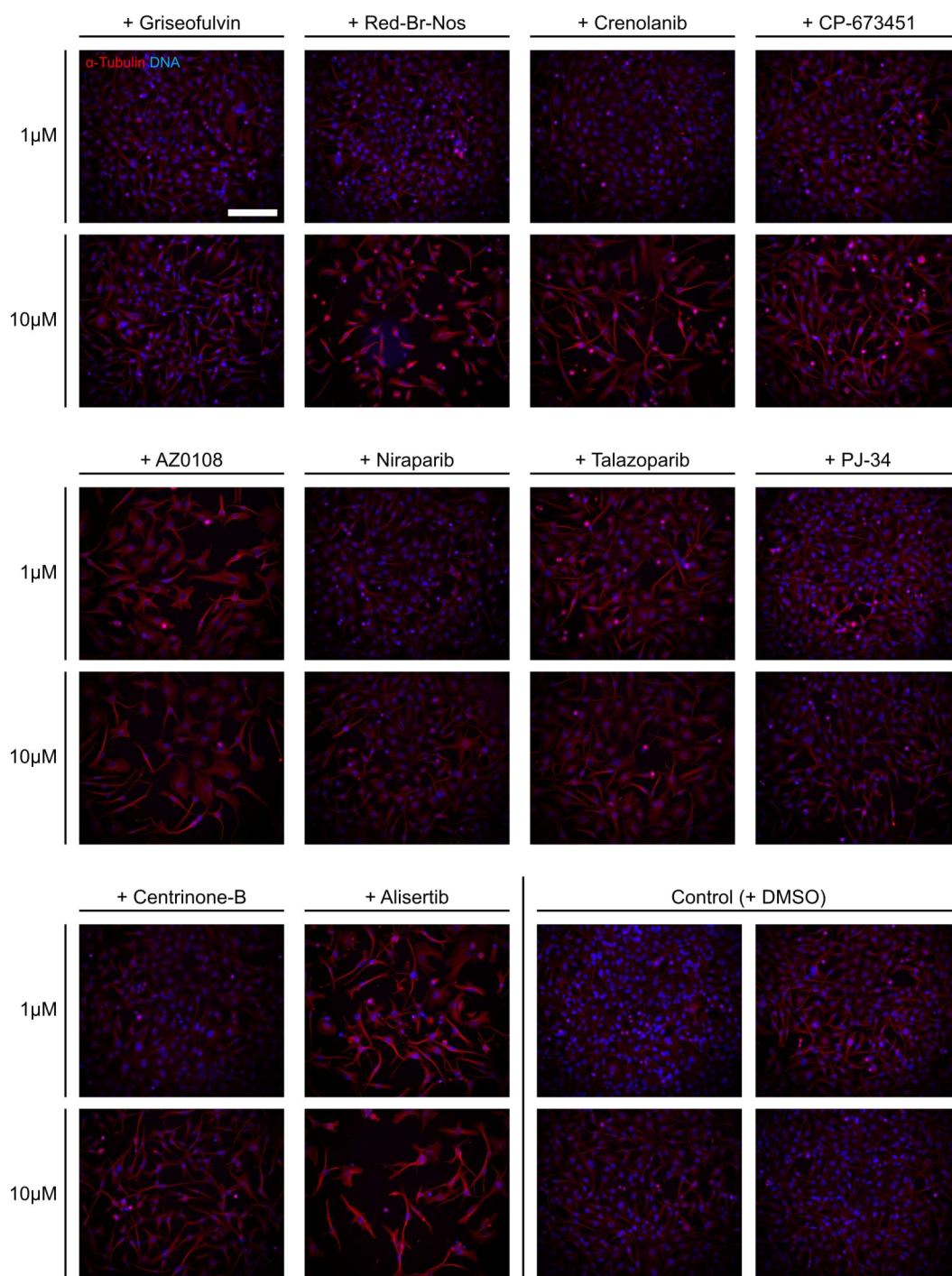
Figure 2.14: Widefield preview images after different drug treatments. Ten published active ingredients were dissolved in dimethyl sulfoxide (DMSO). BT549 cells were incubated for 48 h with fresh medium containing drug concentrations of 1 µM and 10 µM, respectively. Addition of DMSO at an equal concentration was considered as controls. Shown images are illustrative for a total set of 25 preview images that were acquired for each condition. Scale bar 200 µm.

supply. Alternatively, measurements were performed in 8-well ibidi sample slides, for which the application of a single oil droplet is in general sufficient to screen all wells. Additionally, the transfer from the 20X preview air to the 100X high-NA oil objective lens could not be automated. The application of an oil droplet after preview imaging thus remained the only manual operation within the entire imaging cascade. In consequence, some potential problems that might arise in a fully automated instrument, e.g., the occurrence of air bubbles while switching the objective, could not fully be evaluated.

In the first imaging step, 25 single-plane preview images were automatically acquired in a 5×5 pattern using the 20X air objective lens. Drug-mediated effects which influenced a considerable fraction of the cell population could clearly be observed already in this imaging mode (Figure 2.14). For all imaged wells and applied substances, representative images were identified. Whereas Griseofulvin and
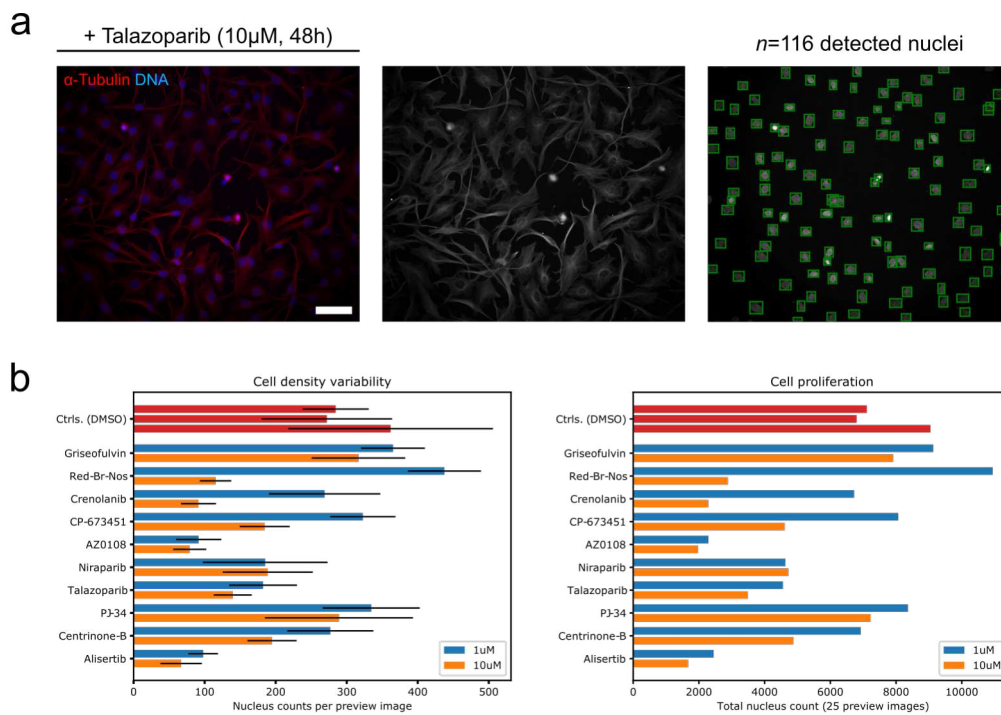

Figure 2.15: Drug treatments have varying influence on cell proliferation. (a) Exemplary widefield preview image of BT549 cells after treatment with $10\,\mu$M talazoparib for 48 h. The number of nuclei was determined using the single-class RetinaNet detection model introduced previously. Identified nuclei are illustrated by green bounding boxes in the DNA image. Scale bar: $100\,\mu$m. (b) A total set of 25 preview images per well was acquired for ten different drug treatments and three control conditions. The number of counted nuclei per preview image varies in part strongly depending on where inside the well the image was acquired (left graph). The black bars indicate the standard deviations of nuclei counts. The accumulated nuclei counts demonstrate variable effects of applied drugs and their concentration on the observed cell proliferation rate (right graph). These results are preliminary due to missing replicates.

PJ-34 showed only little impact on the cell proliferation, other substances, e.g., Red-Br-Nos and Alisertib, showed a clear reduction in proliferation rates and reduced cell densities in comparison to untreated control cells. For some compounds like Centrinone-B and Crenolanib, also the cell morphology was strongly affected, resulting in overly long cell shapes. Whereas some drugs seemed to induce increased cell stress as deduced by clumpy cell appearance ($10\,\mu M$ Red-Br-Nos), none of the applied drugs caused total cell death under the used conditions. Not surprisingly, also the concentration of the applied substances influenced the observed cell densities. In general, however, even within the same well, cell concentrations varied tremendously between neighboring field of views (FOVs). Analyses on the single FOV level are therefore highly questionable. Aiming for a more comprehensive picture on resulting cell proliferation rates, the number of cell nuclei was quantified using all 25 preview images per condition with the single-class RetinaNet cell detector trained previously (Figure 2.15a). The obtained bounding box detections were visually inspected and revealed robust results independent of the actual amount of nuclei per image and differing cell shapes. As a result, the analysis revealed that AZ0108 and Alisertib hampered cell growth the most, already at a concentration of $1\,\mu M$. Among the remaining test compounds, Red-Br-Nos and the cofilin activator Crenolanib showed the strongest decrease in cell proliferation rates when the applied concentration was increased to $10\,\mu M$. In view of the limited data base and lacking replicates, these observations are, however, only preliminary. In general, safe conclusions require that more finely graded concentration ranges are tested at least in duplicates and within independent experiments. On the analysis side, however, the deployed machine learning-based image analysis routine demonstrated its potential with regard to the analysis of extensive datasets without the need for cumbersome parameter tuning.

In the second step of the imaging scheme, identified mitotic cells were imaged as widefield Z-stacks after manually switching to the 100X oil immersion objective. The implemented cell detection and classification approach was further evaluated by manually reviewing all selected nuclei during the POC experiments (Figure 2.16). $90.7\,\%$ out of 1483 singled out nuclei were indeed mitotic and represented different cell division states. With condensed DNA material being the common feature, the number of centrosomal features and the spindle polarity differed strongly between some of the cells. In addition, dense centrosomal clusters were rarely observed, neither in the treated nor in the untreated control cells as revealed by manual inspection of the data. A fraction of $3.8\,\%$ of cells in the mitotic class was actually post-mitotic, which was expected since this class was not separately considered in the training process. Only $2.1\,\%$ and $3.4\,\%$ of nuclei were falsely detected as they belonged to interphase and apoptotic cells, respectively. In consideration of the limited amount of training data, the observed detection precision was satisfactory and posed no bottleneck regarding the remaining imaging steps.
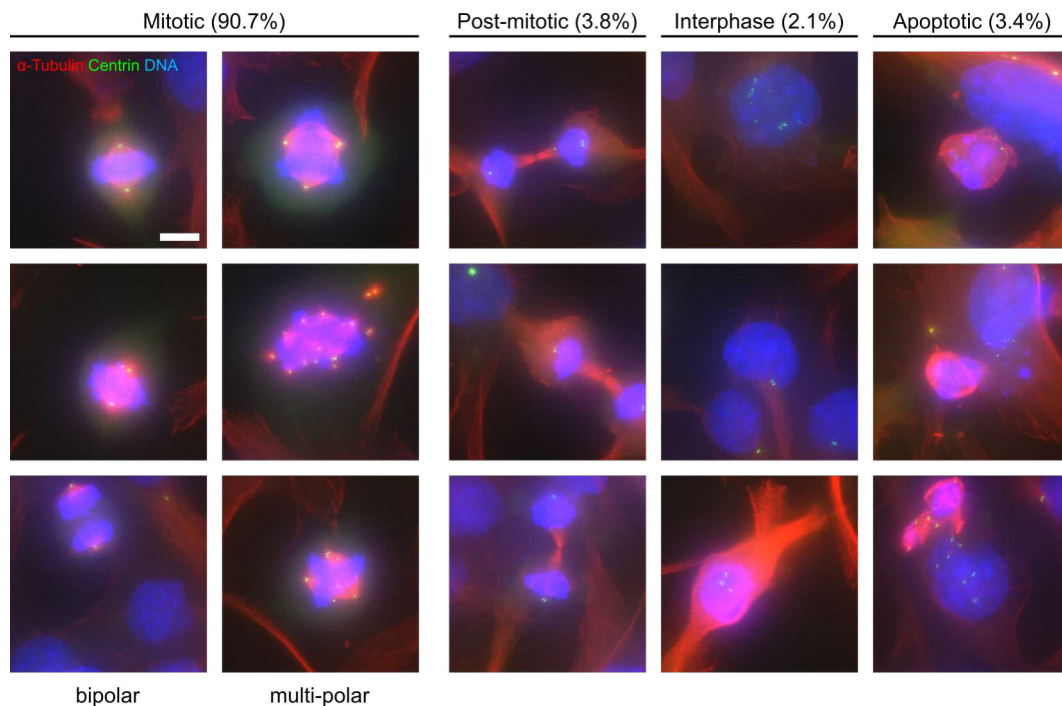
Figure 2.16: Automated mitotic cell detection by machine learning. Exemplary wide-field overview images of mitotic cells identified by the combined approach. The model consisted of a RetinaNet for nucleus detection followed by a ResNet50 for classification. The majority of selected cells was clearly mitotic, featuring either a bipolar, pseudo-bipolar or multi-polar mitotic spindle. Only small subsets of cells were wrongly selected as they belonged to post-mitotic, interphase or apoptotic class. Scale bar 10 µm.

As scheduled, centrosomal features were subsequently detected in maximum intensity projections via a classical blob detection algorithm (Figure 2.17, left panel). Despite the lens switch and differing coordinate centers for both used objectives, virtually all selected cells of interest were located within a central $50 \times 50\,\mu m^2$ area of the total FOV. All centrioles that were found inside this detection window were considered for final confocal and STED imaging. This spatial delimitation was introduced as a simplification step due to time restrictions in the working plan. Depending on the cell density, also centrioles from surrounding interphase cells were detected and afterwards imaged in STED mode. Although these acquisitions can be excluded during post-processing, they substantially prolonged the overall acquisition time. After centriole detection, the feature coordinates were transformed from the widefield overview image to the confocal coordinate system to acquire both confocal and STED Z-stacks of centrosomes only (Figure 2.17, right panel). In the obtained maximum intensity projections, individual centrioles appear either as dot-like or elongated structures depending on their orientation in 3D space. The centriole edges featured a sharper outline in STED images due to the improved resolution. Importantly, the underlying centriole orientation could only be deduced from STED
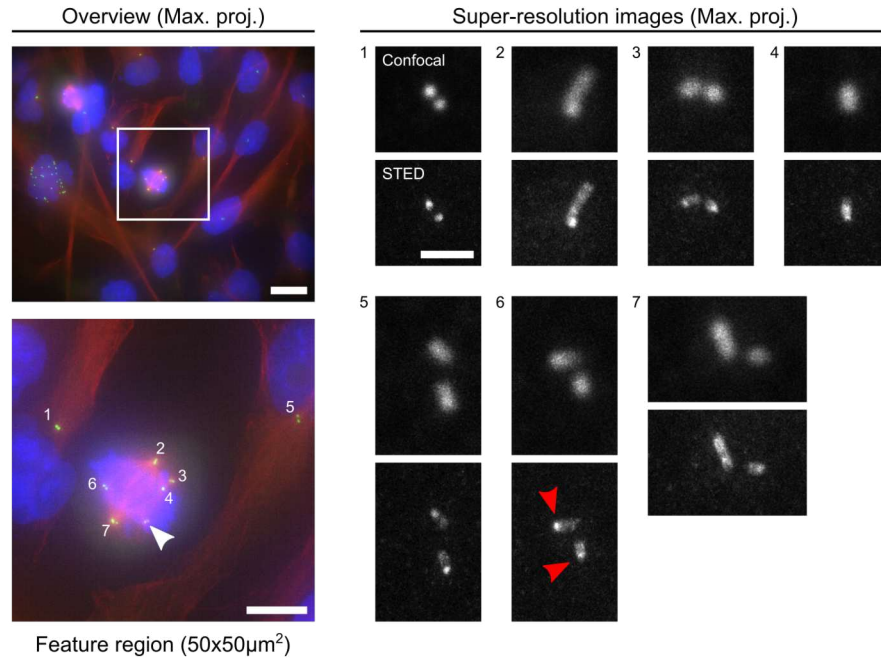
Figure 2.17: Automated confocal and STED imaging of centrosomal features. Exemplified widefield overview image in which individual regions of interest (ROIs) were identified by a difference of Gaussian (DOG) blob detector (numbers; left). Only structures detected inside the central $50 \times 50\,\mu m^2$ square region (white box and inset) were considered for subsequent super-resolution imaging. While two features (1 and 5) belonged to surrounding interphase cells, one feature of the mitotic cell was overlooked (white arrow). Subsequent STED imaging of identified regions enables more precise determination of centriole orientation and protein localization in comparison to confocal images (red arrows). In contrast, STED imaging provided no substantial benefit for counting centrioles due to their sufficiently large spacing. All images are maximum intensity projections of Z-stacks. Scale bars: widefield overview $20\,\mu m$, inset $10\,\mu m$, STED $1\,\mu m$.

acquisitions. Centrin predominantly localizes at the distal ends of individual centrioles where a more intense signal was observed for some imaged features (Figure 2.17, red arrows). Even though a few of the dense structures could only be resolved in STED mode, most of the centrosomal target structures could already be quantified in confocal images. In the given use case, STED microscopy was therefore not absolutely essential to determine centriole copy numbers in cancer cells.

During first POC measurements, the proposed acquisition scheme worked in principle and a total number of 4093 confocal and STED Z-stacks of individual centrosomes or small clusters were automatically acquired. According to the small number of identified mitotic cells for some of the applied drugs, and the overall missing replicates, any paradigmatic analysis of the imaged centrosomal features would only lead to preliminary results. However, due to time constraints it was not possible to acquire additional data and establish an detailed analysis to finally assess drug mediated effects in the course of the present thesis.

Chapter 3

Discussion

3.1 Summary

3.1.1 Custom implementation of automated STED microscopy

Aiming at the unsupervised screening of cell biological assays in multiwell plates, the primary goal for any intended high-throughput application is to accelerate the entire image acquisition process. To develop such a process for high-throughput STED microscopy, an existing Abberior Instruments ExpertLine STED microscope was upgraded with a camera system and complemented by online image analysis for automated image acquisitions. By introducing widefield imaging for cell and feature detection, the overall processing time could be significantly reduced in comparison to an entirely confocal acquisition scheme (Results section 2.2.3). The acquisition time estimates for preview and overview imaging could be accelerated by factors of 190 and 5, respectively. Mitotic cell detection was implemented by a machine learning-based image analysis approach, using a RetinaNet convolutional neural network (CNN) architecture. Already a few hundred preview images were sufficient to successfully train an object detector that could be further refined by more training examples in the future. Afterwards, centrosomal features were successfully localized by means of a classical difference of Gaussians blob detector. Found feature coordinates could be reliably transformed from the widefield to the confocal imaging coordinate system without the need for a repeated calibration procedure. The successive confocal and STED imaging of individual centrioles and their clustered arrangements worked stable and without manual user intervention (Results section 2.4). Both the size and number of imaging regions for STED image acquisition hence could be scaled down substantially, from the entire cell population to only micrometer-sized organelles located within a specific sub-population of cells. The upgraded STED microscope was tested in unsupervised overnight imaging sessions and showed no instabilities or drifts that would prevent even longer measurement scenarios in the future.

3.1.2   Evaluation of the centrosomal clustering model use case

The use case selected for evaluation of the intended acquisition scheme addressed centrosomal clusters in cancer cells and could be successfully adapted for STED microscopy. All steps in the sample preparation were initially optimized, ranging from the right choice of fluorophores over superior fixation and labeling conditions up to the identification of a better suited cellular model system (Results section 2.1). The final protocols resulted in specific and consistent staining and intensity levels as observed in multiple repeated experiments. The obtained signal-to-noise ratios were satisfactory, both for the primary target protein centrin and the secondary targets α-tubulin and DNA. Mitotic cells could be clearly identified in most cases, even though additional protein labeling may further simplify the annotation workflow. The extraction of centrosomal features was possible already in widefield images despite their - in comparison to confocal images - intrinsically higher out-of-focus blur. However, the low intensity levels recorded in final STED Z-stacks were challenging, most likely due to the small size of the target structures, i.e., the centrioles, themself. A simple thresholding analysis of STED images to yield accurate number estimates on the centriole count proved to be challenging. The imaging parameters set during the proof of concept (POC) experiments represented a compromise: the accumulated pixel dwell time and excitation laser powers had to be reduced to minimize photobleaching during STED Z-stack acquisition. Due to the fact that predominantly well separated centrioles were observed, a further reduction in the STED laser power, and thereby also of the optical resolution, would be feasible in future experiments to yield increased intensity levels. As this likewise allows for the choice of larger pixel sizes, this would also lead to shorter acquisition periods in upcoming experiments without impeding the necessary STED resolution gain.

Unfortunately, the centrosomal clustering use case prescribed by the Bayer AG proved to be challenging regarding a general evaluation of the potential of STED microscopy for drug screening. Despite searching for a better suited cancer cell model system during the assay development stage, dense centrosomal clusters were only rarely seen in mitotic cells of the finally applied BT549 cell line. Many cells featured multiple centrosomes, but with a normal centriole count, i.e., one parental and one procentriole. Under these conditions, an assay that is specifically designed for the breakup of clustered structures is challenging and would require large sample sizes for significant statistical analyses.

Moreover, individual centrioles could almost always be identified in oversampled confocal images, even in the case of deviant centrosomal structures with higher copy numbers. The processed use case thus turned out to be challenging to reveal the full potential of automated super-resolution imaging. Furthermore, almost all parameters to evaluate the impact of an applied drug were already accessible in standard widefield images. Phenotypic features like the spindle polarity, the existence of additional spindle poles and the number and size of centrosomal structures

could routinely be obtained even without final confocal or STED images. In addition, cellular responses like an unequal chromosome segregation, the occurrence of lagging chromosomes and induced cell stress could be adequately deduced from appropriate widefield data. In conclusion, a use case that exploits the full potential of STED microscopy by resolving a denser cellular target structure would be more suitable for an evaluation of automated high-throughput STED microscopy.

3.1.3 Potential improvements for the custom automated STED microscope

Although the introduced camera-based acquisition scheme was successfully demonstrated by upgrading an existing ExpertLine microscope, the present circumstances and external dependencies led to performance limitations in the current implementation. The overall performance could, however, straightforwardly be improved if sub-optimal components were exchanged or optimized. One bottleneck was the intensity level of the existing LED illumination source in combination with the present quad-band filter set. For an universally bright staining like DAPI-based DNA labeling, the existing LED intensity level was satisfactory. A camera exposure time of only 10 ms was sufficient for full-frame acquisitions of stained nuclei in repeated experiments. Regarding α-tubulin imaging, however, the filter cube's transmission band did not match well with the excitation wavelength needed for exciting the applied Abberior STAR 580 dye. The α-tubulin channel could still be acquired due to the rather broad intensity profile of the 585 nm LED itself, albeit with low illumination levels. Accordingly, the exposure time had to be increased to about 70 ms which was a compromise between a still acceptable intensity level and imaging speed. By mounting a suitable filter set together with a modern laser illumination source for widefield imaging, overall acquisition times could therefore be further decreased. Furthermore, especially during preview imaging using the 20X air lens, a significantly reduced intensity level at the image boundaries was observed. While the implemented machine learning-based cell detector was capable to deal with this inherent variability, a more uniform illumination would be beneficial in future experiments. Although the observed effect was mainly limited by the field number of the widefield observation tube, a flat-top illumination scheme might be favorable in future applications (Stehr et al., 2019).

A second drawback was the present computer hardware that was used to control the ExpertLine microscope and related components. The existing computer featured no potent graphic card which slowed down the entire ML-based mitotic cell detection process. Online analysis of a single preview image took approximately 7 s and used the computer's full capacity. In contrast, the identical analysis lasted less than 1 s when performed on a state-of-the-art working station that was also used for ML training. Moreover, integration of the peripheral component interconnect (PCI) card necessary for reading images from the mounted sCMOS widefield camera was unstable. The shortest feasible exposure times as determined by the samples could

only be used during test measurements. For extended overnight experiments, however, the acquisition rates had to be substantially decreased to prevent the repeated occurrence of spontaneous and erratic hardware overflows. Although this was not a fundamental technical limitation, it limited the achievable high-throughput capabilities under the given circumstances.

Moreover, custom prototype measurements were not entirely autonomous as there was no oil dispenser available. After preview imaging, the proposed acquisition workflow scheduled the exchange of the low magnification air objective for a higher magnifying oil immersion objective. As application of the immersion oil had to be performed manually, this lens switch could not be tested regarding its applicability for automation. The manual intervention was repeatedly tested and ran smoothly, also when conducted by colleagues in the course of remote measurements. The subsequent imaging steps, i.e., widefield overview and confocal or STED acquisitions, then proceeded again completely autonomous.

Not least, due to the lacking oil dispenser, POC measurements could not be performed in 96-well plates as originally intended. Instead, multiple 8-well chamber slides had to be employed which hampered preparation of replicates and a more diversified concentration range in final evaluation experiments. Once these restrictions become solved, an actual evaluation of the attainable acquisition speeds and high-throughput capabilities can be performed.

3.1.4   Estimates of attainable image acquisition speeds

As mentioned above, the employed phenotypic assay was not optimal for the evaluation of the developed high-throughput STED workflow and impedes an exact estimation of the attainable throughput. Due to the fact that only the minority of cells exhibited centrosomal clusters which can be tackled by an active ingredient, large sample sizes are required for significant statistical analyses. This requirement is, however, difficult to meet due to the rare prevalence of mitotic cells and their observed phenotype heterogeneity. Nonetheless, at least a rough estimate based on the concluding POC measurements could be made.

In the case of untreated cells, about four mitotic cells could be identified on average per single widefield preview image. Assuming that 50 cells must be imaged in each well with an estimated average count of three centrosomal features per cell, a total number of 1200 preview images, 4800 widefield overview Z-stacks and 14 400 STED Z-stacks of individual centrosomes have to be acquired for screening an entire 96-well plate. Final Z-stacks had a lateral size of at least $2 \times 2\,\mu m^2$ and comprised 16 slices along the Z-axis with an interslice distance of 250 nm for exact quantification. The pixel size was set to 20 nm with a dwell time of $3 \times 10\,\mu s$ per pixel. Hence, the resulting minimal acquisition time for STED image acquisition sums up to 4.8 s per single STED Z-stack and approximately 19.2 h for the entire well plate. Considering the already derived estimates of 100 ms per widefield preview image and 2 s per

widefield overview Z-stack (Results section 2.2.3), a total acquisition time of approximately 22 h appears theoretically feasible for the given use case and with the widefield feature detection approach. Herein, any time estimates for stage movements, online image analysis and necessary optical switching steps were not yet considered. Assuming a combined travel and settling time of 100 ms per stage movement and a processing time of 1 s per image for ML-based cell detection, both processes together would prolong the overall acquisition time by approximately one hour. The optical switching between widefield and confocal acquisitions, however, would account for an additional time span of about 4 h if a processing time of 1.5 s per single switch could be achieved. In conclusion, a total processing time of 27 h per well plate appears to be in reach, once the above mentioned constraints of the present ExpertLine device become solved.

The obtained image quality and signal intensity levels of the automated STED prototype device developed by Abberior Instruments were overall very convincing. Due to the fact that first test measurements were performed by an external project partner, however, it cannot be confirmed whether these measurements were conducted with a fully optimized parameter set tailored for the highest possible acquisition speed. A direct comparison between the widefield implementation and the commercial prototype can thus only be tentative. Using the identical assumptions for the centrosomal clustering assay, the Abberior Instruments prototype device would require about 6.2 h for preview imaging of the identical sample area and more than 13 h for the same number of confocal overview Z-stacks. The Final STED Z-stack acquisitions apply the identical technological concept and therefore introduce no further difference. Hence, the total measurement time of the prototype device accumulates to almost 39 h per well plate for the given use case. By using shorter pixel dwell times and coarser sampling, i.e., larger slice distances, in overview Z-stack acquisitions, however, the imaging speed could likely be further increased in future experiments.

Notably, the effective overall acquisition speed extraordinarily depends on the specific assay, the therein required imaging modalities (single-plane or Z-stack acquisitions) and the frequency and total amount of features of interest, to name only a few factors. Under the present conditions, the commercial prototype did not fulfill the requirements requested by Bayer AG with regard to obtainable acquisition speed for a widespread use in industrial drug discovery pipelines. In academia, however, where temporal restrictions are less compulsory, automated STED microscopy has great potential to guide super-resolution microscopy into a data-driven quantitative field of science in the future.

3.2 Outlook

In order to exploit the full potential of the proposed method, several further developments remain to be implemented to routinely apply automated STED microscopy

as a versatile technique in various research questions. In terms of sample preparation, all images shown in this thesis originate from specimen that have been prepared by hand. Although manual handling gives the full control especially during pipetting steps, efficient conduct of biological assays on larger scales demands for automated procedures. First very preliminary experiments on 96-well plates with an entry-level class pipetting robot[1] revealed that automated sample preparation could indeed meet the sophisticated labeling requirements for STED microscopy. Whereas test experiments yielded satisfactory results for adherent cells, the preservation of more loosely attached mitotic cells during repetitive washing steps proved to be challenging. Hence, the obtained fractions of mitotic cells were usually lower in comparison to manually prepared specimen. Final conclusions are, however, difficult as the tuneable parameter set was rather limited for the available instrument. Moreover, I had no access to more customizable state-of-the-art equipment regularly used in drug screening campaigns. Nonetheless, early results indicate that an automated, defined and controlled preparation of cellular samples is also feasible for STED microscopy after corresponding optimization.

In addition, the applied feature detection algorithms could be further expanded to allow for both faster acquisitions and more detailed analyses. Referring to mitotic cell detection, the combined model approach consisting of a RetinaNet architecture for nuclei detection followed by a ResNet50 for classification was owed to scarce training data. Besides the limited amount of available training images, there were also considerable time constraints regarding further annotations. By establishing mitotic cell detection solely based on the RetinaNet architecture, calculation overheads issued from the additional classification model could be removed. A more refined training with a larger amount of mitotic cells present in the training dataset would moreover lead to higher recall values. As a result, less images would be required to identify a certain number of mitotic cells in the course of preview imaging.

Furthermore, it would be highly desirable if the automated cell detection algorithm could be able to differentiate between individual stages of mitosis (Eastman and Guo, 2020). By mapping the positions of centrosomal features and their spindle polarity phenotypes directly to mitotic cell cycle stages (like, e.g., prophase, metaphase and anaphase), a more thorough analysis on mediated drug effects could be set up. This would also be a pivotal prerequisite for envisaged live-cell experiments. Accurate annotations of mitotic stages may, however, require the labeling of additional proteins. One potential target could be lamin that localizes into the nuclear membrane (Pannu et al., 2014). In a more advanced scenario, it might be even feasible to detect different mitotic stages in label-free phase-contrast images by applying modern deep learning frameworks (Liu et al., 2017).

Even more importantly, also the process of feature detection in widefield images could be considerably improved. To date, individual centrosomes were selected only based on an imprecise selection window. As a result, also centrosomes

---

[1]Opentrons OT-2 Liquid Handler

from surrounding interphase cells were occasionally selected for STED imaging. In order to avoid unnecessary acquisitions and obstructive post-processing steps, ML-based mitotic cell segmentation should be also implemented for overview images. Thereby, centrosomal features could be confined solely to mitotic cells of interest which would further reduce the overall acquisition time.

Moreover, final STED Z-stack images of centrioles could also be automatically analyzed by means of ML algorithms. In a preliminary approach, I applied the RetinaNet architecture on maximum intensity projections of STED images of centrosomes to determine the present number of centrioles. Although training could not be fully optimized due to temporal constraints, a precision and recall value above 96 % and 92 %, respectively, was already obtained. To exploit the full information content of acquired STED data in the future, three-dimensional instance segmentation algorithms could be applied for a more thorough analysis (Kar et al., 2022). By doing so, both the orientation of individual centrioles in 3D space and the angles between them could be extracted, thereby revealing details that are not resolvable in confocal acquisitions.

Lastly, the most severe bottleneck with regard to image acquisition times is STED imaging itself, independent of the actual use case. Hence, parallelized STED imaging would lead to an immense speed up of any automated STED microscopy platform, irrespective of the acquisition scheme applied for feature detection. Whereas parallel STED microscopy has been successfully established in academic research, either by multiplexing beams (Bingen et al., 2011) or standing wave patterns (Yang et al., 2014; Bergermann et al., 2015), commercial implementations are still lacking. Limiting factors for its widespread use were the need of very intense STED laser, confined camera frame rates and complex optical setups which have to run stable and reliably over extended acquisition periods. In addition, highly parallelized STED imaging relies on camera-based detection which, by virtue of missing pinholes, complicates 3D imaging and acquisitions of thicker biological sample regions due to the increased detection of out-of-focus light. Further technological developments may therefore lead to more advanced and optimized automated STED microscopy approaches that pave the way for its future use in academia as well as, with some limitations, in industrial drug development and target validation pipelines.

Chapter 4

Materials and Methods

4.1 Materials and instrumentation

4.1.1 Cell culture

All cell lines used in this study were human wild type (*wt*) cancer cells. U2OS bone cancer cells (HTB-96, ATCC) were cultured in DMEM medium containing high glucose levels and GlutaMAX supplement (Gibco 32430-027, ThermoFisher Scientific). BT549 mammary gland breast cancer cells (HTB-122, ATCC) were grown in RPMI-1640 culture medium (Gibco A10491-01, ThermoFisher Scientific). MDA-MB-231 mammary gland breast cancer cells (HTB-26, ATCC) were cultured in Leibovitz's L-15 medium (Gibco 21083-027, ThermoFisher Scientific). All media were supplemented with 10 % fetal bovine serum (FBS; 35-079-CV, Corning), $100\,\text{U}\,\text{mL}^{-1}$ penicillin and $100\,\text{µg}\,\text{mL}^{-1}$ streptomycin (P4333 (100X concentrated mixed solution), Sigma-Aldrich). For BT549 cell line, culture medium was additionally supplemented with $0.023\,\text{U}\,\text{mL}^{-1}$ insulin (Gibco 12585-014, ThermoFisher Scientific). All cell lines were cultured in a humidified 5 % CO_2 incubator at 37 °C. For MDA-MB-231 cells, gas exchange was omitted during incubation as the L-15 medium was formulated for gas exchange with atmospheric air (ATCC handling procedure, 2021). For better adherence to substrates, cell lines were grown in functionalized cell culture flasks (3289 and 3290 with CellBIND® surface, Corning). Cells were passaged twice to thrice a week to maintain confluency levels below 80 %. Before splitting, cells were once washed with prewarmed (37 °C) Dulbeccos's Phosphate-Buffered Saline (DPBS; Gibco 14040-117, ThermoFisher Scientific) and then detached by adding TrypLE Express solution (Gibco 12604-013, ThermoFisher Scientific). Passage numbers were between 3 and 30 for all cell experiments. Cell viability was regularly measured by Trypan blue exclusion assay (Invitrogen T10282, ThermoFisher Scientific) and constantly revealed healthy fractions $\geq$ 98 %. Cells were in addition routinely checked against mycoplasma contamination in an external laboratory (SGS Analytics Germany, Jena, Germany). For investigating cell morphology, confluency levels and potential cell stress, an EVOS XL Core transmitted light microscope (ThermoFisher Scientific) was used.

4.1.2   Immunofluorescence sample preparation

Multi-well chamber slides with #1.5H glass coverslip bottom (8-well μ-Slides 80827/ 80807, Ibidi GmbH) were incubated for 1 h with 20 μg mL^{-1} poly-L-lysin solution (P4707, Sigma-Aldrich) that was diluted in pure distilled water (Invitrogen Ultra-Pure 10977-035, ThermoFisher Scientific). After coating, wells were washed 3 $\times$ 5 min with pure water and sterilized for 1 h under UV-light before use. After seeding into chamber slides, cells were cultured for at least 24 h prior treatment or fixation/labeling.

All fixation, processing and staining solutions were diluted in 1xDPBS (Gibco 14040-117, ThermoFisher Scientific) and were always freshly prepared on the day of the experiment. Required formaldehyde (FA) concentrations were obtained by diluting from a 32 % paraformaldehyde (PFA) aqueous stock solution shipped in individual glass vials (15714, Electron Microscopy Sciences). For every experiment, a fresh glass vial was used. Glutaraldehyde (GA) was diluted from frozen aliquots of a 50 % stock solution (G7651, Sigma-Aldrich) stored at $-20\,^{\circ}$C. If not otherwise stated, the fixation solution was prewarmed to $37\,^{\circ}$C before use. All other incubation steps were carried out at room temperature (RT). Before antibody incubation, cells were permeabilized using 0.2 % Triton X-100 (T8787, Sigma-Aldrich) as detergent. For blocking, 5 % bovine serum albumin (BSA; A9576, Sigma-Aldrich), 300 mM glycine (33226, Sigma-Aldrich) and 0.2 % Tween-20 (Epredia TA-125-TW, ThermoFisher Scientific) were solved in 1xDPBS for quenching remaining aldehyde groups. All primary and secondary antibodies were diluted in 1X-DPBS containing 5 % bovine serum albumin and 0.1 % Tween-20. All washing steps were performed with 1X-DPBS for 5 min. DNA was stained by DAPI (62248, ThermoScientific) diluted 1:1000 (1 mg mL^{-1}) in 1X-DPBS. After preparation, samples were stored in 1X-DPBS at $4\,^{\circ}$C until imaging.

For membrane blebbing experiments (Appendix Figure A.2), U2OS cells were seeded into poly-L-lysine-coated ibidi 8-well μ-slides and cultured for 24 h. For live-cell labeling, the old cell medium was aspirated and fresh, prewarmed ($37\,^{\circ}$C) medium with Vybrant DiD membrane stain (1:200 dilution; Invitrogen V22887, ThermoFisher Scientific) was added for 30 min. After three washing steps (3 $\times$ 2 min) with prewarmed medium, cells were immediately fixed with prewarmed fixative solutions containing 1, 2, 4 or 8 % FA and 0.2 % glutaraldehyde for 10 min at RT. Samples were immediately imaged on a STEDYCON microscope using the 640 nm excitation wavelength.

4.1.3   Abberior Instruments automated STED prototype device

The prototype device was developed by Abberior Instruments GmbH based on the commercially available FacilityLine microscope platform with an Olympus IX83 microscope body. Besides the internal autofocus system provided by Olympus, a second, independent focusing capability by means of a deformable mirror (DM) was newly developed. The focal travel range was limited to 12 μm in the sample when

the DM option was used. The XY stage was fully integrated into the Imspector software platform (Abberior Instruments) used for controlling the microscope. The setup included a liquid oil dispenser (00-74-300-0000, Märzhäuser) for continuous supply of immersion oil during extended measurement periods. It was equipped with four excitation lasers (405 nm, 488 nm, 561 nm and 640 nm) and two STED lasers (595 nm and 775 nm). The detection unit comprised four independent spectral detection channels with configurable detection windows. The microscope was equipped with an auto-alignment routine for maintained calibration of optical beam paths also over extended imaging periods.

4.1.4 Custom STED microscope upgrade for automated camera-based measurements

Custom implementation of an automated STED microscope was performed on an ExpertLine STED microcope (Abberior Instruments) with an Olympus IX83 microscope body as basis. The setup was equipped with four excitation lasers (405 nm, 488 nm, 594 nm and 640 nm) and a 3.0 W high-power 775 nm STED laser (40 MHz, 600 ps pulse duration; Onefive). Fluorescence emission was detected by three APD detection channels with fixed detection windows of 500-550 nm, 605-625 nm and 650-720 nm. A spatial light modulator (SLM) provided optimal overlap of excitation and STED laser beams and was aligned with a 100 nm gold bead calibration sample (Abberior). The employed galvanometer-based scanning system was limited to a maximum line frequency of 250 Hz.

Custom-written software was developed in the Python programming language (Python Software Foundation, version 3.7, https://www.python.org) for full control of all necessary devices. For setting of measurement parameters and viewing of widefield images, a graphical user interface (GUI) was programmed in Python using the wxPython GUI toolkit. All motorized components of the microscope body and the ExpertLine device itself were accessible by means of the SpecPy Python interface provided by the Imspector software platform (Abberior Instruments). This included both the motorized XY stage, the piezo stage insert (P-736 PInano Z, Physik Instrumente) used for focusing and the Olympus autofocus system. A fast widefield camera system (Andor Neo 5.5 sCMOS, Oxford Instruments) was mounted on the camera port of a newly attached trinocular tube. Control and readout of the camera was realized via the Andor SDK3 software development kit and the corresponding pyAndorSDK3 Python wrapper (Oxford Instruments). Image data was transferred by the Camera Link protocol. The already existing widefield illumination source (pE-2, Cool LED) was connected via USB and controlled with the pySerial Python package. Available widefield excitation wavelengths were 400 nm, 490 nm, 585 nm and 635 nm. Separation of excitation and emission light was carried out by a single quadband filter set (F66-866, AHF Analysentechnik) for all widefield excitation wavelengths.

4.2 Experimental work

4.2.1 Optimization of cell fixation and staining protocols

Applied primary antibodies were monoclonal mouse anti-centrin (clone 20H5; 04-1624, Sigma-Aldrich) and polyclonal rabbit anti-α-tubulin antibodies (ab18251, 1 mg mL^{-1}, Abcam). For centrin staining, an anti-mouse nanobody coupled to Abberior STAR RED (FluoTag-X2 IgG, N1202-ABRED, 10 µM, NanoTag-Biotechnology) was used. For α-tubulin labeling, corresponding goat anti-rabbit secondary antibodies were conjugated with Abberior STAR ORANGE (STORANGE-1001-500UG, 1 mg mL^{-1}, Abberior), Abberior STAR 580 (ST580-1002-500UG, 1 mg mL^{-1}, Abberior) and AlexaFluor 594 (ab150080, 2 mg mL^{-1}, Abcam). When comparing signal intensities and cross talk, employed dilutions were adapted according to present antibody concentrations to yield identical amounts of fluorophore in the sample. In the optimization process, various preparation conditions and parameters were continuously adapted. Besides the variations that are indicated in the corresponding results sections, this included also differing washing procedures, fixation conditions and altered concentrations of applied chemicals. Modifications of a specific preparation parameter, e.g., the glutaraldehyde concentration, were always tested within the same cell seeding experiment with all other sample treatments kept unaffected. In final experiments, cells were fixed for 10 min with 1 % formaldehyde and 0.02 % glutaraldehyde. After fixation, samples were three times washed and then permeabilized for 10 min with 0.2 % Triton X-100. After three more washing steps, cells were blocked for 1 h before primary antibodies were incubated for 1 h at RT or overnight at 4 °C. After three times washing, secondary antibodies were incubated for 1 h at RT, followed by three more washing steps. Lastly, DAPI was incubated for 5 min and subsequently removed by three final washing steps. For optimal labeling, both primary antibodies were diluted 1:1000. Secondary antibodies were diluted 1:1000 for centrin (STAR RED) and 1:250 for α-tubulin staining (STAR 580).

During optimization of sample preparation protocols, fluorescence microscopy was performed with a STEDYCON device (Abberior Instruments), a commercial confocal and STED microscope add-on mounted on an Eclipse Ti inverted microscope body (Nikon). It was equipped with four excitation lasers (405 nm, 488 nm, 561 nm and 640 nm) and one STED laser emitting at 775 nm suitable for STED imaging in the red and far-red wavelength regime. Confocal and STED images were acquired with a 100X/NA 1.45 oil immersion objective lens (MRD01905, Nikon). For centrin imaging, the 640 nm excitation laser and the far-red detection channel (650-700 nm) were used. Labeled DNA and α-tubulin structures were excited by the 405 nm and 561 nm lasers, respectively, and detected by the same combined detection channel with two detection windows (500-550 nm and 580-625 nm). Samples were moved by means of an electric XY stage that was controlled by a Tango controller (Märzhäuser). Note that the XY stage was not integrated into the STEDYCON

software and could therefore only be used for manual sample movements. Focusing was realized by a software-integrated objective scanner system (P-721 PIFOC, Physik Instrumente).

During imaging, pixel sizes of 100 nm for pure confocal and 20 nm for STED or combined confocal/STED image acquisitions were routinely applied. For both imaging modes, a pixel dwell time of 10 µs was used. Confocal images were always acquired by a single line accumulation, wheres in STED mode usually two line accumulations were summed up. Gating was always enabled using the default delay settings (1-7 ns). The pinhole had a fixed size of 66.6 µm in the intermediate image plane, corresponding to approximately 1.2 Airy units (AU) for the applied 100X/NA 1.45 objective and an assumed wavelength of 650 nm. When acquiring Z-stacks, the slice distance was usually set to 500 nm for pure confocal and 250 nm for STED image acquisitions.

4.2.2 Abberior Instruments automated STED prototype measurements

Test measurements of the Abberior Instruments prototype device on biological samples were conducted in 96-multiwell plates with glass bottom (ViewPlates 6005430, Perkin Elmer). Well plates were equally prepared, i.e., poly lysin-coated, as described for 8-well μ-Slides. Likewise, also the identical cell seeding, fixation and labeling protocols were applied with the only difference that in the fixation step, besides FA and GA also 0.2 % Triton X-100 were used for pre-permeabilization. All pipetting steps were carried out manually using a multichannel pipette.

Shown prototype measurements were conducted by Abberior Instruments in collaboration with me, either on site or remotely. In test measurements of the automated imaging pipeline, all imaging steps were acquired using a single 60X/NA 1.40 oil immersion objective lens (UPLXAPO60, Olympus). Employed excitation laser wavelengths for the three target structures centrin, α-tubulin and DNA were identical to STEDYCON measurements. For STED imaging, only the far-red 775 nm STED wavelength was applied. Confocal preview images were acquired at a single-plane with a pixel size of 450 nm. Confocal overview Z-stacks of identified cells were acquired using a pixel size of 130 nm and a slice distance of 300 nm. Herein, the DNA (DAPI) signal was split onto two detectors for improved signal intensities. Final confocal and STED imaging of centrosomal features was performed as Z-stacks solely for the centrin channel with a pixel size of 22 nm and the identical slice distance. Embedded cell and feature detection algorithms were contributed by project partners and could not be affected within measurements.

Performance measurements of the prototype device were conducted with a 40 nm fluorescent bead sample (Nanoparticles Fluor Red 40, Abberior) using the 60X/NA 1.40 oil lens. An image size of $10 \times 10 \,\mu\text{m}^2$ was placed at seven positions along the diagonale of the entire field of view, from upper left to lower right side. At every field of view position, the deformable mirror (DM) was set from $-6.0\,\mu\text{m}$ to $6.0\,\mu\text{m}$ in steps of $3.0\,\mu\text{m}$. At the center position, a finer step size of $1.5\,\mu\text{m}$ was used. For every

DM setting, confocal and STED images were acquired with a pixel size of 15 nm and full STED laser power, i.e., maximum attainable resolution. Between consecutive acquisitions, the sample was manually refocused and slightly moved in xy-direction to prevent measurement impairments due to photobleaching.

4.2.3 Widefield imaging for cell and feature detection and image registration

BT549 cells were incubated with 0.1 % dimethyl sulfoxide (DMSO; D2650, Sigma-Aldrich) for 24 h before fixation and labeling of the sample (mock control with identical DMSO concentration as in later drug treatment experiments). The DNA channel (DAPI) was imaged with 400 nm excitation light (10 % LED intensity) and 10 ms exposure time. Abberior STAR 580 (α-tubulin channel) was excited by 585 nm excitation light (100 % LED intensity) and 100 ms exposure time. Abberior STAR RED (centrin channel) was imaged with 635 nm excitation (70 % LED intensity) and an exposure time of 30 ms. The camera readout was full-frame with global shutter, 280 MHz readout rate, 11-bit low noise gain control and no binning in all cases. The sensor was actively cooled to $-30\,°C$ with enabled fan speed.

Preview images for training of cell detection algorithms were acquired as single-plane widefield images using a 20X/NA 0.80 air objective lens (UPLXAPO20X, Olympus). The Olympus autofocus system was always enabled in continuous mode. The related pixel size was 325 nm. For centrosome detection, widefield overview images were acquired using a 100X/NA 1.40 oil objective lens (UPLSAPO100XO, Olympus). Exemplary mitotic cells were imaged as Z-stacks with a slice distance of 1 µm and disabled autofocus. The resulting lateral pixel size was 65 nm.

For image registration, 10 confocal and widefield image pairs of the α-tubulin channel were acquired using the 100X/NA 1.40 oil objective lens. Single-plane images were collected pairwise with enabled autofocus and no intermediate stage movements. Confocal images were acquired with a pixel size of 65 nm and 10 µs dwell time using the 594 nm laser for excitation, while widefield images were acquired as described above. For calibration of XY stage movements, widefield image pairs of the α-tubulin channel were acquired. Image pairs were collected with enabled autofocus and nominal stage movements of either 100 µm solely in x-direction or 70 µm solely in y-direction between acquisitions.

4.2.4 Proof of concept measurements with custom camera-based automated STED microscope

BT549 cells were seeded in 8-well chamber slides and incubated for 24 h before adding drug compounds. All ten applied active ingredients (griseofulvin, reduced-9-bromonoscapine, talazoparib, niraparib, AZ0108, PJ-34, crenolanib, CP-673451, alisertib and centrinone-B) were obtained from the Bayer-internal compound library as 10 mM stock solutions dissolved in DMSO. Stock solutions were further diluted

in DMSO (D2650, Sigma-Aldrich) to a concentration of 1 mM. Both solutions were diluted 1:1000 in fresh prewarmed medium to attain final drug concentrations of 10 µM and 1 µM with identical DMSO concentrations. For mock controls, pure DMSO was diluted 1:1000 in fresh prewarmed medium. When adding drug compounds, the old medium was completely exchanged for the drug-containing media and incubated for 48 h before fixation and labeling. The fixation buffer additionally contained 0.2 % Triton X-100. After labeling, samples were stored at 4 °C in 1X-PBS containing 0.02 % sodium azide (190380050, Acros Organics) before imaging.

25 widefield preview images per well were acquired in a 5x5 snake pattern with an exposure time of 11 ms for the DNA (10 % 400 nm illumination), 70 ms for the α-tubulin (100 % 585 nm illumination) and 30 ms for the centrin (100 % 635 nm illumination) detection channel. Previews were acquired as single-plane images with a 20X air objective and enabled autofocus. When entering a fresh well, the focus search routine was triggered once to retain the substrate in focus. Applied camera settings and used objectives were identically to the cell and feature detection experiments described above.

Mitotic cell detection was performed on the α-tubulin and DNA channel with the combined RetinaNet + ResNet50 model. Appropriate intensity scaling was achieved by averaging the mean intensities of approximately 20 randomly chosen interphase nuclei from test images before starting the measurement. Calibration values were only obtained once and could thereafter be successfully applied for all samples prepared within the same experiment. The parameters for the confidence threshold, the non-maximum suppression (NMS) intersection-over-union (IOU) threshold (both RetinaNet for nuclei detection) and the minimum score (ResNet50 nuclei classification) were all set to a value of 0.50. After acquiring all preview images of a given chamber slide, the objective was manually switched to the 100X oil immersion objective. After applying the immersion oil, the lens was manually approached back to the sample plane before the automated measurement continued.

The XY stage consecutively moved to the coordinates of the previously identified mitotic cells and performed a focus search. At every coordinate position, a widefield overview Z-stack was acquired containing 20 images with a slice distance of 1 µm. Any z-movement was performed with the piezo stage insert throughout the Z-stacks. Only the central $50 \times 50\,\mu m^2$ region was considered for identification of centrosomal feature coordinates as detailed below. Final confocal and STED Z-stacks were consecutively acquired immediately after extracting all (x, y, z) feature coordinate triplets in order to preserve z-accuracy due to potential axial drift. Z-stacks were acquired as line interleaved acquisitions and comprised 16 slices with a slice distance of 250 nm. The pixel size was 20 nm with a dwell time of 10 µs. Confocal and STED Z-stacks were acquired with one and three line accumulations, respectively, using a pinhole size of 1.0 airy units (AU). Only the centrin channel was imaged using the 640 nm excitation laser with an applied laser power of 1.5 % for confocal and 2.1 % for STED acquisitions. The 775 nm STED laser was used with a laser power of 30 %.

Typical field-of-view sizes were between $2 \times 2\,\mu m^2$ and $3 \times 3\,\mu m^2$.

4.3   Image and data analysis

All analysis procedures were carried out in custom-written scripts using the Python programming language (Python Software Foundation, version 3.7, `https://www.python.org`). Recurrently used software packages included NumPy (Harris et al., 2020), SciPy (Virtanen et al., 2020), Scikit-image (Walt et al., 2014), Scikit-learn (Pedregosa et al., 2011), OpenCV (Bradski, 2000), TensorFlow (Abadi et al., 2015) and Matplotlib (Hunter, 2007).

4.3.1   Analysis of Abberior Instruments prototype bead measurements

Bead positions were detected in normalized STED images (normalized by maximum image intensity) by a Laplacian of Gaussian (LOG) blob detector (Scikit-image `blob_log()` method with parameters *min_sigma* = 2, *max_sigma* = 5, *num_sigma* = 4, *threshold* = 0.05, *overlap* = 0.5). Found peaks were filtered by their radius ($\leq$ 2 pixels), their intensity ($\geq$ 10 photon counts) and their minimum distance to neighboring beads ($\geq$ 5 pixels; Appendix Figure B.1). Remaining peaks were cropped from the original STED images with a size of 15×15 pixels and then fit to a two-dimensional isotropic Gaussian

$$y = A \cdot \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}\right) + b$$

with amplitude A, Gaussian center coordinates (x_0, y_0), standard deviation σ and background intensity level b. For fitting, the SciPy `curve_fit()` method was used with bounds provided for all fitting parameters. Beads with a fitted standard deviation > 3 pixels, i.e., larger than 45 nm, were disregarded from subsequent analysis as their underlying image data often represented agglomerated beads or background noise (Appendix Figure B.2). Leftover fit results were further analyzed by means of their amplitudes and standard deviations. Fitted σ-values were converted to full width at half maximum (FWHM) values according to FWHM $= 2\sqrt{2\ln 2}\sigma$. Amplitudes and FWHMs were then binned and a histogram was computed by applying the NumPy `hist()` method in density mode with the Freedman Diaconis estimator (*bins = 'fd'*) used for calculating bin boundaries. Binned amplitude distributions were then fit to a normalized single Gaussian distribution

$$y = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(I - I_0)^2}{2\sigma^2}\right)$$

to yield an average intensity value I_0 for each FOV position and DM setting. The distribution of FWHM values was fit to a sum of two normalized Gaussians to distinguish between fitting results of single beads and those originating from small

clusters of two or more beads (Appendix Figure B.3). This discrimination was particularly necessary for lower resolution images to exclude interfering effects from sample imperfections. In final conclusions, only the smaller average FWHM values from the left Gaussian distribution were considered that corresponded in good agreement to fits of individual beads.

4.3.2   Machine learning-based cell detection in widefield images

A total number of 130 20X widefield images was considered for training a RetinaNet Deep Learning network architecture for object detection. A classic feature detection algorithm was implemented in Python to obtain nuclei segmentation estimates for speeding-up the whole annotation workflow (Figure 2.9). At first, the DNA channel image was smoothed by a Gaussian blur (OpenCV `GaussianBlur()` with a kernel size of 11 pixels, i.e., *ksize* = (11,11)). Thereafter, pixels belonging to nuclei were extracted into a binary mask by local thresholding (Scikit-image `threshold_local()` with *block_size* = 501 pixels and *offset* = -1). Potential objects were then identified by a connected component analysis of the obtained mask (Scikit-image `label()` with *background* = 0 and *connectivity* = 2). Closely spaced nuclei were subsequently divided by means of a Watershed algorithm (Scikit-image `watershed()`), for which the thresholded binary image was first transformed into a Euclidean distance map (SciPy `distance_transform_edt()`). The minimum distance between watershed basins was set to 20 pixels. Resulting segments were filtered to range between a minimum and maximum size of 500 and 10 000 pixels, respectively. Finally, contours enclosing individual nuclei were deduced (OpenCV `findContours()` with *mode = RETR_LIST*, *method = CHAIN_APPROX_SIMPLE*) and smoothed afterwards (Scikit-image `approximate_polygon()` with *tolerance* = 3). All polygon segments were subsequently manually checked using the VGG Image Annotator (VIA) software (Dutta and Zisserman, 2019) and assigned to either the interphase, mitotic, apoptotic or post-mitotic nucleus class before training the models.

   All deep learning algorithms were implemented in TensorFlow (Version 2.7.0, Abadi et al., 2015) and trained on a potent workstation PC (ThinkStation P520, Lenovo) with GPU-support (Quadro RTX 6000, 24 GB GDDR6 RAM, Nvidia). Base models were used with weights pretrained on the ImageNet dataset as provided by the Keras API (Chollet et al., 2015). For object detection, the RetinaNet architecture was implemented as exemplified in Keras (Humbarwadi, 2020). Therein, a ResNet50 model with frozen weights from an ImageNet pre-training served as a backbone. 130 fully annotated preview images were randomly split by 60-20-20 ratios into training, validation and test datasets (Scikit-learn `train_test_split()`). Random number generator seed values were always chosen identical to ensure comparability in ablation studies. Polygon annotations were transformed to bounding boxes enclosed individual nuclei. For compatibility with the ImageNet weights, which require a three-channel input image, the α-tubulin and DNA images were used as red and

blue channels while the green channel was filled with zeros. Training results for different RGB mappings showed no significant deviations. Data normalization was performed as discussed in the results section. Raw images (2560×2160 pixels) were resized to 1408×1188 pixels, the maximum image size for which a batch size of 2 still fit into GPU memory. Images were randomly flipped in horizontal and vertical directions for augmentation. Additional rotations or random cropping of image regions led to no improvements in direct comparisons and were therefore omitted. The same was true when image tiles (2×2 split) without resizing were used for training. The RetinaNet loss function was modified by class weights to consider class imbalances as described in the Results section. Training was performed over a period of 80 epochs applying a stochastic gradient descent (SGD) optimizer with a momentum of 0.9. The initial learning rate of 2.5×10^{-6} was repeatedly adjusted to values of 6.25×10^{-4}, 1.25×10^{-3}, 2.5×10^{-3}, 2.5×10^{-4} and 2.5×10^{-5} after 2, 4, 10, 60 and 70 epochs, respectively. All other parameters were set to the default values.

For cell cycle classification, ResNet50, MobileNetV2 and EfficientNetB0 model architectures were used as base models. Training was performed in a transfer learning and fine-tuning scenario as outlined in the Keras API (Chollet, 2020). Annotated nuclei were cropped from raw images according to their bounding boxes. All mitotic (609) and apoptotic (648), but only 1000 randomly selected interphase instances (out of 18 968 available ones) were used for training to deal with class imbalances. Image crops were split 60-20-20 into training, validation and test data. Transfer learning was conducted for 30 epochs followed by 20 epochs of fine-tuning, using a batch size of 32 images. Image normalization, target image sizes for resizing and applied learning rates during transfer learning were chosen as detailed in the results section. During fine-tuning, a learning rate of 1×10^{-5} was used. Image crops were randomly flipped (horizontal and vertical) and rotated during augmentation. Training was performed with the Adam optimizer and sparse categorical cross entropy as loss function.

4.3.3 Centrosomal feature detection in widefield images

The centrin channels of acquired Z-stacks were maximum intensity projected and normalized by its maximum intensity value. Centrosomal features were then detected by a classical difference of Gaussians (DOG) blob detection algorithm (Scikit-image `blob_dog()` with *min_sigma* = 2, *max_sigma* = 5, *sigma_ratio* = 1.3, *threshold* = 0.05, *overlap* = 0.9). Returned σ-values indicate the standard deviation of the Gaussian kernel that detected the blob and were used to specify the width w of each bounding box according to $w = 4\sigma$. Centrosomal features were subsequently combined into a single bounding box if their center positions were closer together as $2\,\mu m$.

4.3.4 Image registration between widefield and confocal image coordinate systems

Acquired 16-bit images were converted to 8-bit by setting their min. and max. intensity values as thresholds for scaling. SIFT keypoints and descriptors were identified in images by the OpenCV SIFT detector (`SIFT_create()` with `detectAndCompute()`) as exemplified in the OpenCV tutorials section (OpenCV feature matching, 2021). In short, related keypoints were specified using the OpenCV brute-force matcher (`knnMatch()` method of `BFMatcher()` class) using the L2 norm as distance measure. Good keypoints were extracted by requesting a descriptor distance < 0.6 for matching keypoints. Homography between confocal and widefield images was then calculated with good keypoints from all 10 image pairs (OpenCV `findHomography()`). Herein, the RANSAC method with a threshold value of 5 pixels was used to exclude remaining outliers. Finally, the resulting homography matrix was used to transform confocal images and coordinates into the widefield image coordinate system (OpenCV `warpPerspective()` and `perspectiveTransform()`).

4.4 Thesis writing

This thesis was written in the TeXstudio editor (`https://www.texstudio.org`) using the LATEXsoftware package (The Latex project, `https://www.latex-project.org`). All images were processed in the ImageJ Fiji distribution (Schindelin et al., 2012). Final figures and illustrations were created in the vector graphics editor Inkscape (The Inkscape Project, `https://inkscape.org`).
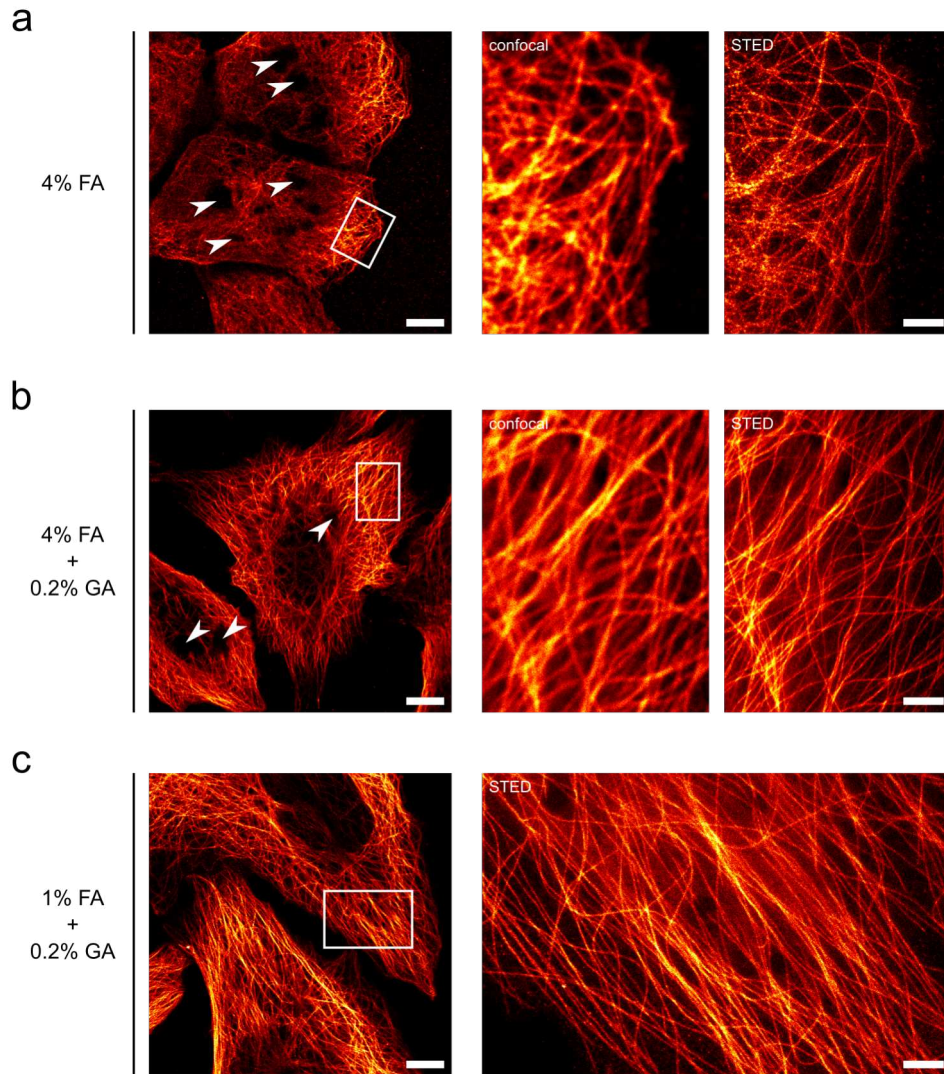
Appendix A

Assay development

Figure A.1: Optimizing cell fixation for microtubule imaging. (a, b, c) Representative immunofluorescence images of the microtubule cytoskeleton in U2OS cancer cells. (a) Standard cell fixation with 4 % formaldehyde (FA) lead frequently to membrane blebbing artefacts (arrows) and insufficient structure preservation of microtubule filaments, predominantly at the center of the cell (left panel). While filaments in the outer regions of the cell appear continuously stained in confocal mode, STED imaging reveals numerous gaps and fractures along individual filaments (insets). (b) Structure preservation improved enormously when 0.2 % glutaraldehyde (GA) were additionally used in the fixation step. Note that dark spots, where the sample has presumably detached from the substrate during fixation, are still present (arrows). (c) Membrane blebbing artefacts were strongly diminished when also the FA concentration was reduced to 1 % (left panel). STED imaging of the dense part of the microtubule network (inset) showed no impairment of labeling density and image quality. Fixative solutions were prewarmed to 37 °C and incubated for 10 min, cell permeabilization was performed with 0.1 % Triton-X 100 for 10 min. After blocking with 5 % bovine serum albumin (BSA) for 1 h, cells were stained with anti-α-tubulin antibodies (T6074, Sigma-Aldrich) for 1 h at room temperature (RT). Secondary antibodies coupled to AlexaFluor 594 (A11032, Invitrogen) were incubated for 1 h at RT. Scale bar 10 μm, insets 2 μm.
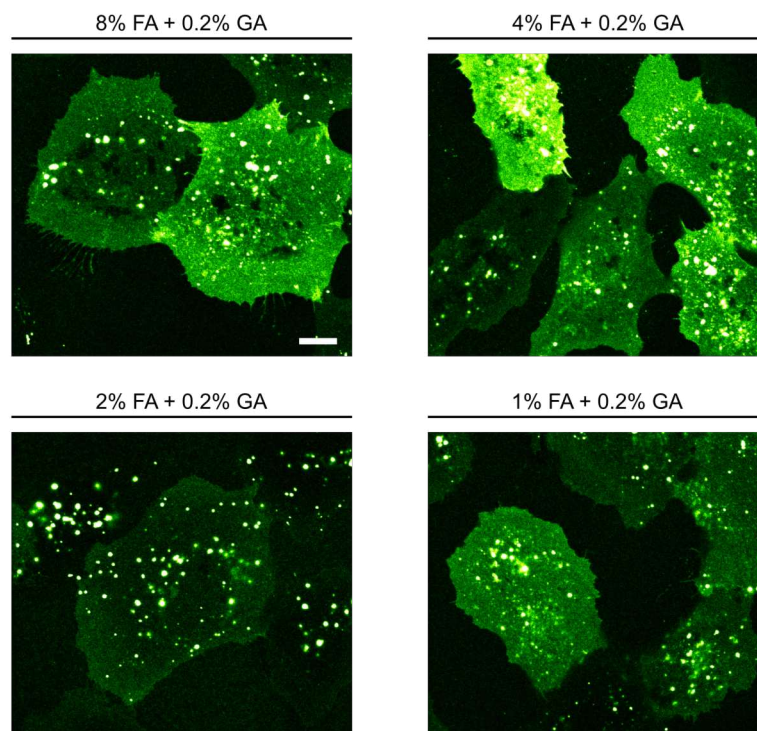
Figure A.2: Membrane blebbing as fixation artifact for high formaldehyde (FA) concentrations. Representative confocal images of fluorescently labeled plasma membrane in U2OS cancer cells. Living cells were stained with Vybrant DiD for approx. 30 min and then fixed with the indicated FA and glutaraldehyde (GA) concentrations. Focus positions were manually adjusted to match a layer close to the surface. Note the more frequent occurrence of unstained spots for high FA concentrations where the membrane has detached from the substrate. Scale bar: 10 µm.
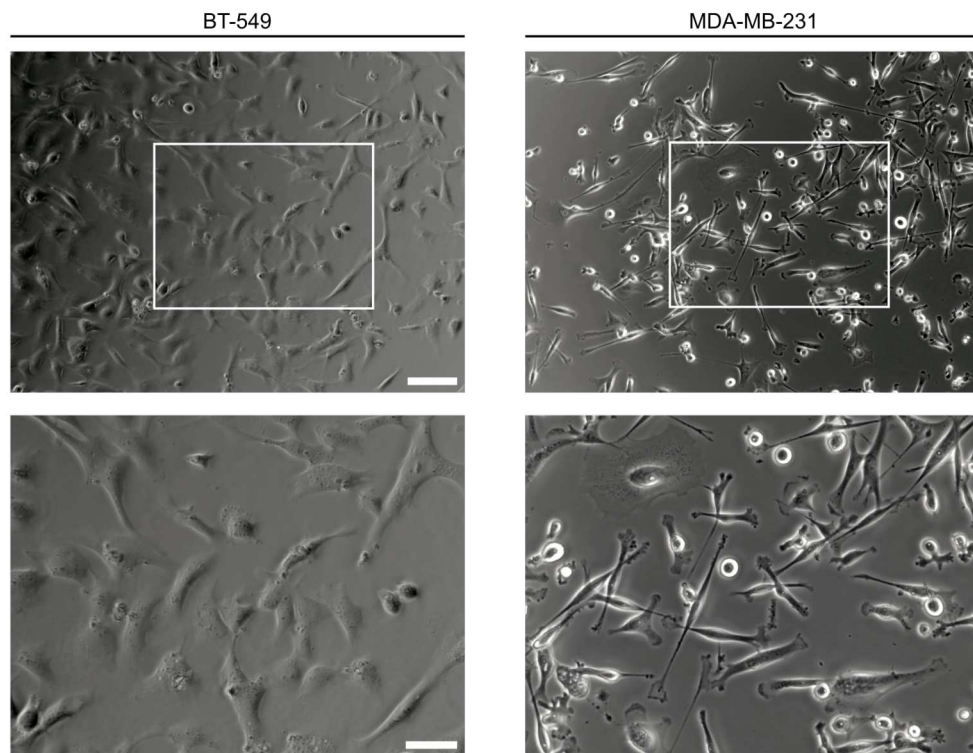
Figure A.3: Cell morphologies matter for advanced imaging workflows. BT-549 *wt* and MDA-MB-231 *wt* cell lines vary with regard to cell morphology and growth behaviour. Cells were imaged in T25 cell culture flasks with an EVOS XL Core phase contrast microscope using a 20X (top) and 40X objective lens (bottom). Scale bar 100 µm, inset 50 µm.

Appendix B
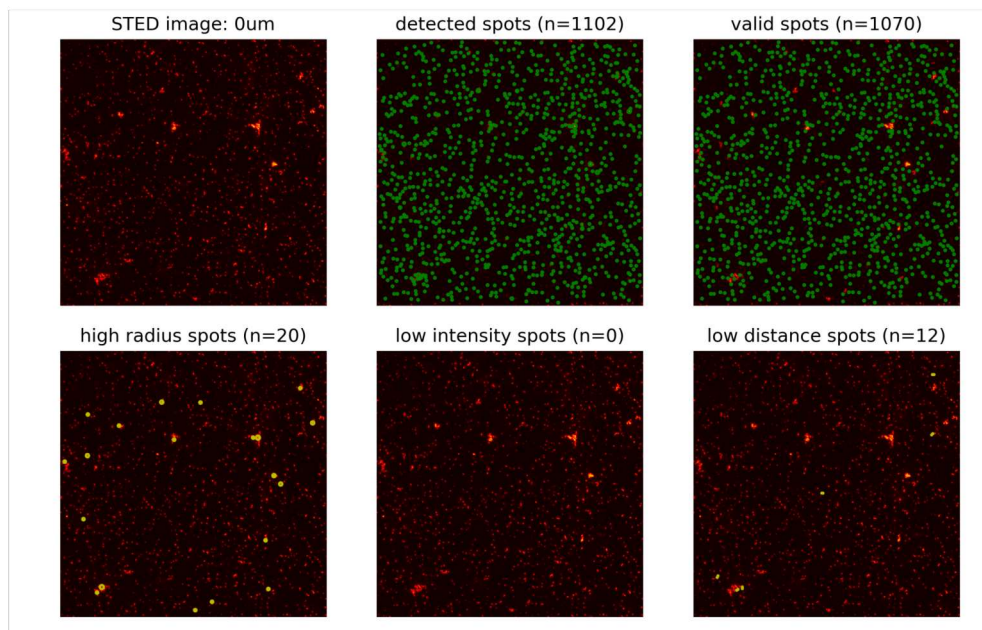
Prototype bead measurements

Figure B.1: Localization of fluorescent beads by classical blob detection. Exemplary STED image of 40 nm fluorescent beads acquired with the Abberior Instruments prototype device using a pixel size of 15 nm. The $10 \times 10 \, \mu m^2$ region of interest was acquired in the center of the field of view at zero focal offset of the deformable mirror. Individual beads were detected by a difference of Gaussians (DOG) blob detector (green circles). Detected clustered bead arrangements and unspecific features were subsequently filtered according to the obtained spot radius, the spot intensity and the distance between spots (yellow circles).
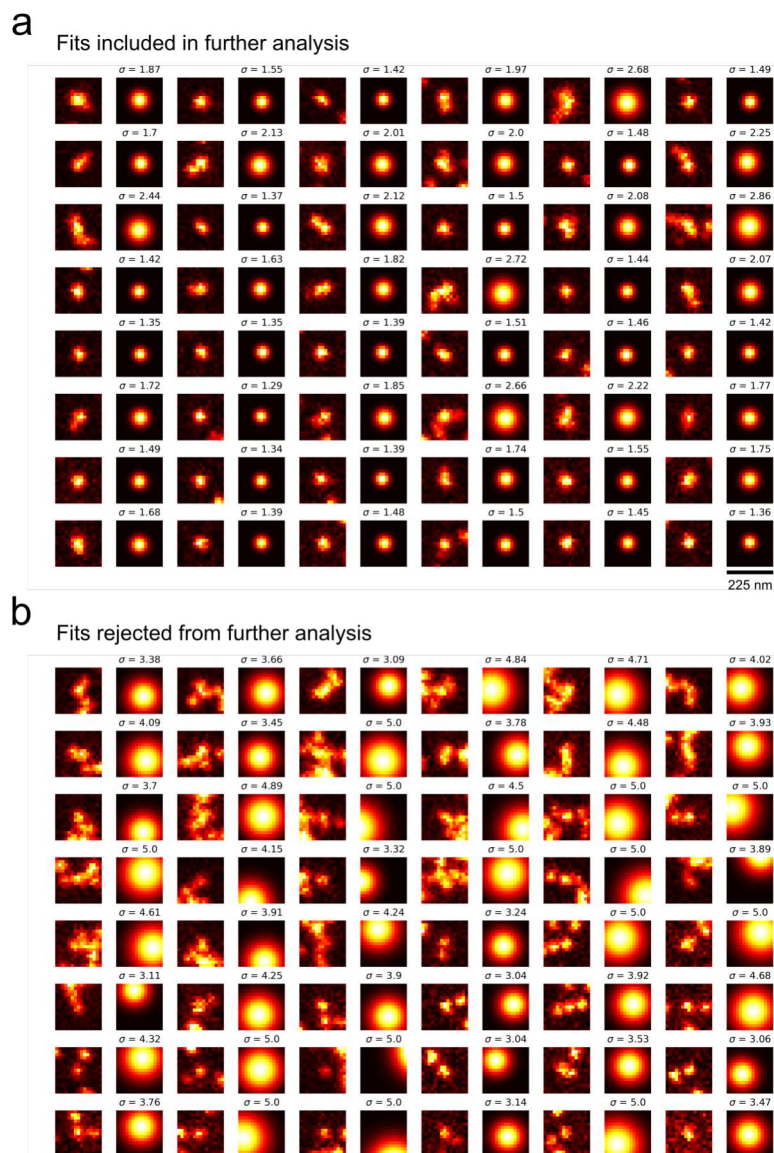
Figure B.2: Determination of 2D STED resolution and intensity estimates obtained by bead measurements. Representative section of exemplary raw data and fit results for a 2D STED image of 40 nm fluorescent beads. Each localized fluorescent bead was cropped by an area of 15×15 pixels before the raw data (respective left columns) was fitted to a two-dimensional isotropic Gaussian function (respective right columns). Obtained standard deviations (σ) are denoted in pixels for every fit. A threshold value of $\sigma = 3$ pixels was applied to filter between fitted σ-values and peak intensities that were (a) included in, or (b) rejected from further analysis. The raw image was acquired with the Abberior Instrument prototype device in the center of the field of view and at zero offset of the deformable mirror, using a pixel size of 15 nm. Scale bar 225 nm.
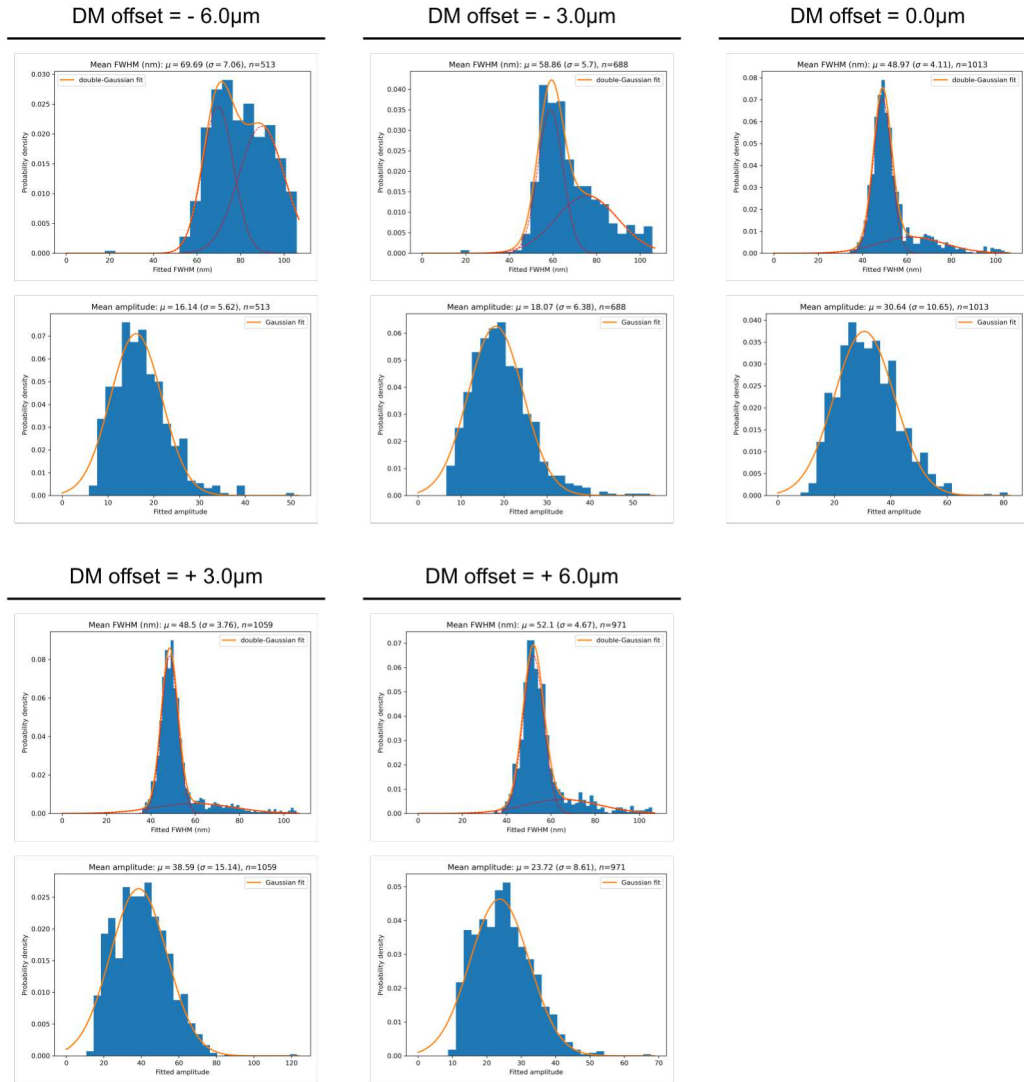
Figure B.3: Determination of the average 2D STED resolution and intensity of the Abberior Instruments prototype device. 2D STED images of 40 nm fluorescent beads were acquired in the center of the accessible field of view, and for different offsets (i.e., focal positions) of the deformable mirror (DM). Images were acquired with a pixel size of 15 nm. Full width at half maximum (FWHM) and peak intensity (amplitude) values were obtained by 2D Gaussian fits. Filtered FWHM distributions (upper graphs) were fitted to a sum of two Gaussian distributions to reduce the influence of small bead clusters still present in the data. Denoted μ and σ values indicate the means and standard deviations of the left fitted Gaussian. Filtered amplitude distributions (lower graphs) were fitted to a single Gaussian distribution. The number of beads considered in each distribution is specified by n.

Appendix C

Machine-learning based cell detection

Measure	Interphase	Apoptotic	Mitotic	Post-mitotic
Instances	18968	648	609	294
μ_{DNA}	142.9	178.4	203.5	169.0
σ_{DNA}	21.2	39.2	44.2	27.3
Med_{DNA}	140.0	170.6	197.0	163.6
$\mu_{\alpha\text{-Tub}}$	119.7	131.0	143.9	129.7
$\sigma_{\alpha\text{-Tub}}$	14.4	23.6	14.8	9.9
$\text{Med}_{\alpha\text{-Tub}}$	118.5	127.7	143.1	129.1

Table C.1: Intensity measures for different nucleus class annotations. Annotated nucleus instances from 130 preview images were individually cropped to calculate their individual average intensity. Subsequently, for each class the mean intensity μ, standard deviation σ and median Med were calculated over all class instances. DNA and α-Tub indicate the respective acquisition channels of DNA and α-tubulin used for mitotic cell detection.
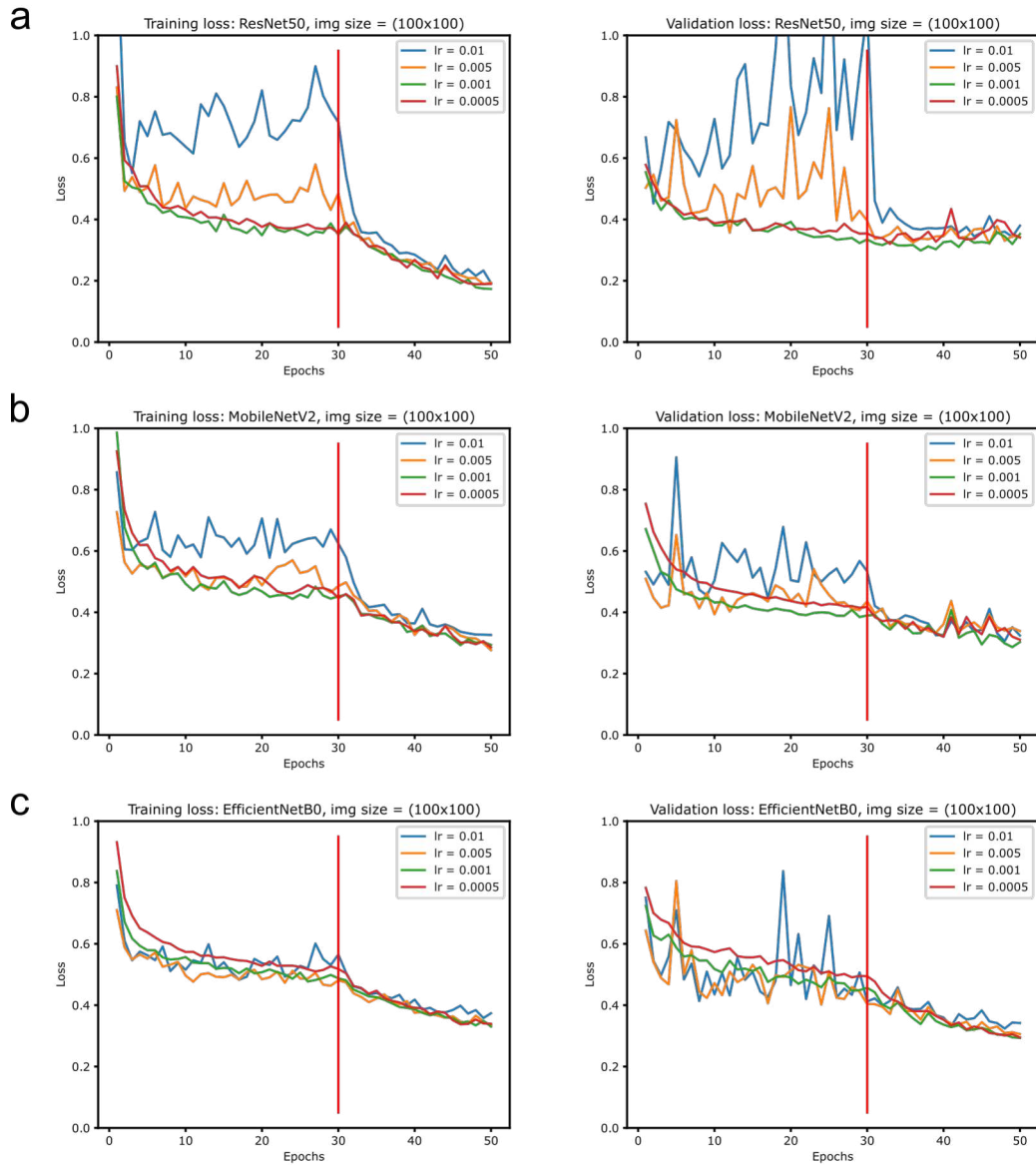
Figure C.1: Evaluation of optimal learning rates for three different deep learning models: (a) ResNet50, (b) MobileNetV2, and (c) EfficientNetB0. Models were trained for classification of manually annotated nuclei crops as interphase, mitotic or apoptotic. All image crops were resized to the same target size of 100×100 pixels before training, and sparse categorical cross entropy was used as loss function. Training was performed by transfer learning of pre-trained weights (trained on the ImageNet dataset) for the first 30 epochs, followed by fine-tuning for 20 epochs. Specified learning rates were solely applied during transfer learning, for fine-tuning a learning rate of 1×10^{-5} was used in all scenarios.
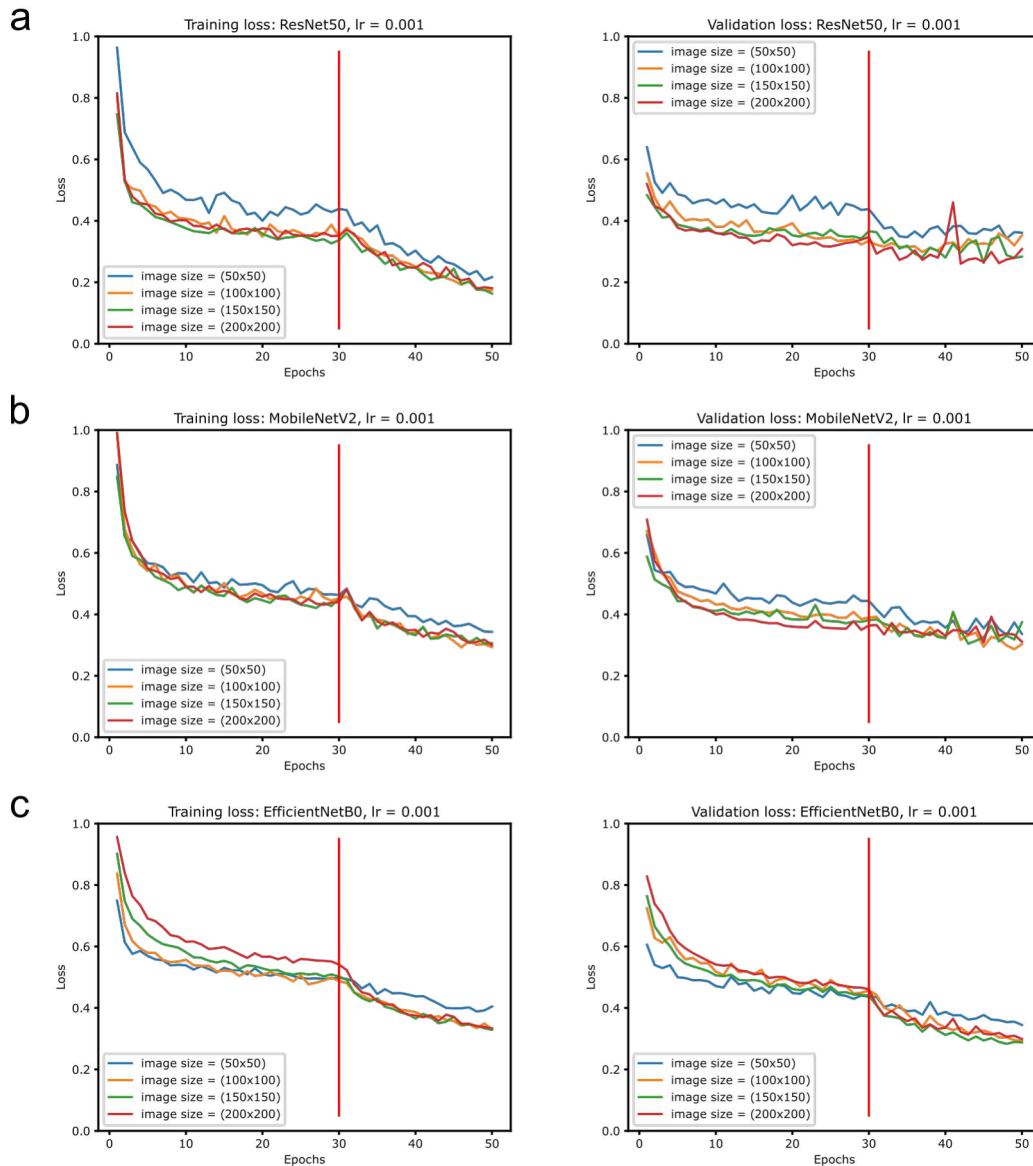
Figure C.2: Evaluation of optimal target image sizes for three different deep learning models: (a) ResNet50, (b) MobileNetV2, and (c) EfficientNetB0. Models were trained for classification of manually annotated nuclei crops as interphase, mitotic or apoptotic. All image crops were resized to the indicated target sizes before training, and sparse categorical cross entropy was used as loss function. Training was performed by transfer learning using pre-trained weights (trained on the ImageNet dataset) with a learning rate of 1×10^{-3} for the first 30 epochs, followed by fine-tuning with a learning rate of 1×10^{-5} for 20 epochs.
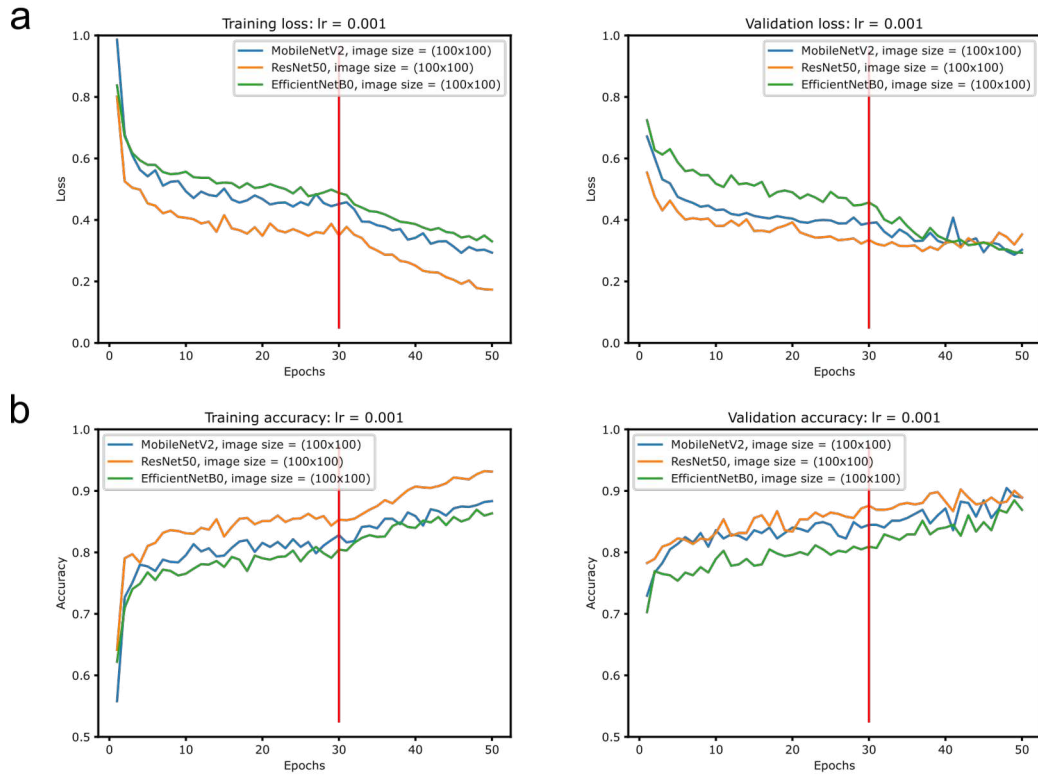
Figure C.3: Optimal model architecture for nuclei classification. Three different deep learning models (ResNet50, MobileNetV2 and EfficientNetB0) were trained for classification of manually annotated nuclei crops as interphase, mitotic or apoptotic. Shown are the training progression of (a) the sparse categorical cross entropy loss, and (b) the sparse categorical accuracy for the training and validation datasets, respectively. All results are for an optimal target image size of 100×100 pixels and learning rate of 1×10^{-3} as previously determined. Training was performed by transfer learning using pre-trained weights (trained on the ImageNet dataset) with a learning rate of 1×10^{-3} for the first 30 epochs, followed by fine-tuning with a learning rate of 1×10^{-5} for 20 epochs.
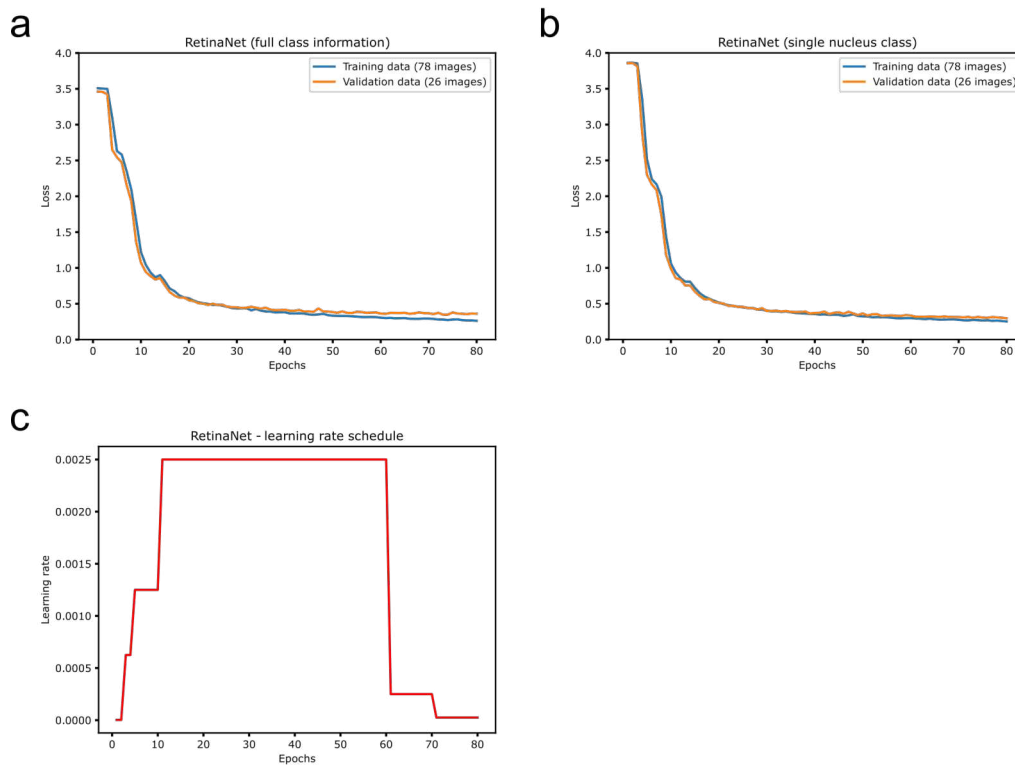
Figure C.4: RetinaNet losses and learning rate schedule. (a) Training and validation loss for RetinaNet training with full class information, i.e., nuclei annotated as interphase, mitotic and apoptotic were considered during training. (b) Training and validation loss for RetinaNet training as nuclei object detector, i.e., all annotated nuclei (interphase, mitotic, apoptotic and post-mitotic) were merged in a single nuclei class. In both cases, training and validation datasets encompassed identical image sets. (c) Applied learning rate schedule for RetinaNet training.

Class	Metric	Training data	Validation data	Test data
Interphase	Instances	11314	3874	3614
	$Prec_{IoU=.50}$	0.955	0.951	0.958
	$Rec_{IoU=.50}$	0.902	0.904	0.908
	$F1\text{-}Score_{IoU=.50}$	0.928	0.927	0.932
	$AP_{IoU=.50-0.95}$	0.696	0.684	0.692
Mitotic	Instances	351	118	140
	$Prec^{IoU=.50}$	0.706	0.664	0.588
	$Rec^{IoU=.50}$	0.815	0.788	0.714
	$F1\text{-}Score^{IoU=.50}$	0.757	0.721	0.645
	$AP^{IoU=.50-0.95}$	0.508	0.429	0.354
Apoptotic	Instances	387	103	158
	$Prec_{IoU=.50}$	0.239	0.146	0.190
	$Rec_{IoU=.50}$	0.618	0.379	0.437
	$F1\text{-}Score_{IoU=.50}$	0.345	0.211	0.264
	$AP_{IoU=.50-0.95}$	0.233	0.075	0.084

Table C.2: RetinaNet results for nucleus detection and classification. RetinaNet was trained with full class information on the interphase, mitotic and apoptotic nucleus classes, post-mitotic annotations were omitted during training. The dataset comprised 130 preview images and was split 60-20-20 into training, validation and test data. Abbreviations: Prec = precision, Rec = recall, AP = average precision, IoU = intersection-over-union.

Class	Metric	Training data	Validation data	Test data
Nucleus	Instances	12213	4168	3972
	$Prec_{IoU=.50}$	0.941	0.947	0.943
	$Rec_{IoU=.50}$	0.945	0.947	0.941
	$F1\text{-}Score_{IoU=.50}$	0.943	0.947	0.942
	$AP_{IoU=.50-0.95}$	0.730	0.715	0.715

Table C.3: RetinaNet results for nucleus detection without further classification. During training, all class annotations, i.e., interphase, mitotic, post-mitotic and apoptotic, were unified in a single *nucleus* class. Training, validation and test datasets were identical to the training scenario with full class information. Abbreviations: Prec = precision, Rec = recall, AP = average precision, IoU = intersection-over-union.

Class	Metric	Training data	Validation data	Test data
Interphase	Instances	601	209	190
	Prec	0.94	0.95	0.93
	Rec	0.97	0.96	0.98
	F1-Score	0.96	0.95	0.95
Mitotic	Instances	362	122	125
	Prec	0.89	0.84	0.90
	Rec	0.90	0.84	0.85
	F1-Score	0.89	0.84	0.87
Apoptotic	Instances	392	120	136
	Prec	0.88	0.79	0.85
	Rec	0.82	0.78	0.82
	F1-Score	0.85	0.78	0.84

Table C.4: Results for nucleus classification with ResNet50. Nucleus crops were extracted from 20X preview images and split 60-20-20 into training, validation and test datasets. Crops contained both the α-tubulin and DNA channel. Only nuclei annotated as interphase, mitotic and apoptotic were considered during training, while post-mitotic annotations were omitted. Abbreviations: Prec = precision, Rec = recall.

Prediction	Post-mitotic
Interphase	0.46
Mitotic	0.28
Apoptotic	0.26

Table C.5: ResNet50 classification of post-mitotic nuclei. Nuclei crops annotated as post-mitotic were evaluated by ResNet50 model solely trained on the interphase, mitotic and apoptotic class. In total, $n = 294$ post-mitotic nuclei were present in 20X preview images.

Class	Metric	Training data	Validation data	Test data
Interphase	Instances	11314	3874	3614
	$\text{Prec}_{\text{IoU}=.50}$	0.944	0.940	0.950
	$\text{Rec}_{\text{IoU}=.50}$	0.909	0.907	0.907
	$\text{F1-Score}_{\text{IoU}=.50}$	0.926	0.923	0.928
	$\text{AP}_{\text{IoU}=.50-0.95}$	0.685	0.668	0.681
Mitotic	Instances	351	118	140
	$\text{Prec}^{\text{IoU}=.50}$	0.864	0.876	0.899
	$\text{Rec}^{\text{IoU}=.50}$	0.652	0.661	0.636
	$\text{F1-Score}^{\text{IoU}=.50}$	0.744	0.754	0.745
	$\text{AP}^{\text{IoU}=.50-0.95}$	0.445	0.419	0.400
Apoptotic	Instances	387	103	158
	$\text{Prec}_{\text{IoU}=.50}$	0.381	0.371	0.418
	$\text{Rec}_{\text{IoU}=.50}$	0.377	0.417	0.405
	$\text{F1-Score}_{\text{IoU}=.50}$	0.379	0.393	0.412
	$\text{AP}_{\text{IoU}=.50-0.95}$	0.101	0.115	0.133

Table C.6: Results for nucleus detection and classification in a combined approach: RetinaNet trained on single nucleus class data was used for object detection, found nuclei were then classified by a ResNet50 trained on instances of the interphase, mitotic and apoptotic class. Training, validation and test datasets were identical to the training scenario with full class information. Abbreviations: Prec = precision, Rec = recall, AP = average precision, IoU = intersection-over-union.

Class	Metric	Cond. 1	Cond. 2	Cond. 3	Cond. 4
Normalization	DNA ch.	[100, 760]	[100, 1160]	[100, 1160]	[100, 760]
	α-Tubulin ch.	[100, 250]	[100, 350]	[100, 250]	[100, 350]
Interphase	Instances	3614	3614	3614	3614
	$\text{Prec}_{\text{IoU}=.50}$	0.952	0.948	0.951	0.955
	$\text{Rec}_{\text{IoU}=.50}$	0.864	0.922	0.863	0.861
	$\text{F1-Score}_{\text{IoU}=.50}$	0.906	0.935	0.905	0.905
	$\text{AP}_{\text{IoU}=.50-0.95}$	0.653	0.688	0.626	0.658
Mitotic	Instances	140	140	140	140
	$\text{Prec}_{\text{IoU}=.50}$	0.687	0.962	0.778	0.923
	$\text{Rec}_{\text{IoU}=.50}$	0.721	0.543	0.650	0.600
	$\text{F1-Score}_{\text{IoU}=.50}$	0.704	0.694	0.708	0.727
	$\text{AP}_{\text{IoU}=.50-0.95}$	0.442	0.356	0.400	0.388
Apoptotic	Instances	158	158	158	158
	$\text{Prec}_{\text{IoU}=.50}$	0.272	0.475	0.275	0.289
	$\text{Rec}_{\text{IoU}=.50}$	0.329	0.424	0.354	0.494
	$\text{F1-Score}_{\text{IoU}=.50}$	0.298	0.448	0.309	0.364
	$\text{AP}_{\text{IoU}=.50-0.95}$	0.092	0.147	0.057	0.155

Table C.7: Influence of image normalization on nuclei detection and classification. Both DNA and α-tubulin detection channels were rescaled as indicated. Results are for the combined model (RetinaNet for nucleus detection and subsequent ResNet50 for classification) that was used during final automated measurements. For training, differing nominal normalization ranges of [100, 960] for DNA and [100, 300] for α-tubulin were used. The combined model was evaluated with test data containing the same 26 images as in the single model cases. Abbreviations: Cond = condition, ch. = channel, Prec = precision, Rec = recall, AP = average precision, IoU = intersection-over-union.

Bibliography

Abadi, Martín et al. (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. URL: `https://www.tensorflow.org/`.

Alberts, B et al. (2008). "Chapter 17: The Cell Cycle". In: *Molecular Biology of the Cell. Fifth edit. Garland Science*. Taylor & Francis Group, Abindong, UK, pp. 1053–1114.

ATCC handling procedure (2021). *MDA-MB-231 handling information*. `https://www.atcc.org/products/htb-26`. Accessed: 2021-02-04.

Balzarotti, Francisco et al. (2017). "Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes". In: *Science* 355.6325, pp. 606–612.

Bergermann, Fabian et al. (2015). "2000-fold parallelized dual-color STED fluorescence nanoscopy". In: *Optics express* 23.1, pp. 211–223.

Bettencourt-Dias, Mónica and David M Glover (2007). "Centrosome biogenesis and function: centrosomics brings new understanding". In: *Nature reviews Molecular cell biology* 8.6, pp. 451–463.

Betzig, Eric et al. (2006). "Imaging intracellular fluorescent proteins at nanometer resolution". In: *science* 313.5793, pp. 1642–1645.

Bingen, Pit et al. (2011). "Parallelized STED fluorescence nanoscopy". In: *Optics Express* 19.24, pp. 23716–23726.

Bleicher, Konrad H et al. (2003). "Hit and lead generation: beyond high-throughput screening". In: *Nature reviews Drug discovery* 2.5, pp. 369–378.

Booth, Martin et al. (2015). "Aberrations and adaptive optics in super-resolution microscopy". In: *Microscopy* 64.4, pp. 251–261.

Boutros, Michael, Florian Heigwer, and Christina Laufer (2015). "Microscopy-based high-content screening". In: *Cell* 163.6, pp. 1314–1325.

Bradski, G. (2000). "The OpenCV Library". In: *Dr. Dobb's Journal of Software Tools*.

Brito, Daniela A, Susana Montenegro Gouveia, and Mónica Bettencourt-Dias (2012). "Deconstructing the centriole: structure and number control". In: *Current opinion in cell biology* 24.1, pp. 4–13.

Chandrasekaran, Srinivas Niranj et al. (2021). "Image-based profiling for drug discovery: due for a machine-learning upgrade?" In: *Nature Reviews Drug Discovery* 20.2, pp. 145–159.

Chollet, François (2020). *Transfer learning & fine-tuning*. `https://keras.io/guides/transfer_learning`. Accessed: 2022-01-31.

Chollet, François et al. (2015). *Keras*. `https://keras.io`.

Crane, Richard et al. (2004). "Aurora A, meiosis and mitosis". In: *Biology of the Cell* 96.3, pp. 215–229.

Deng, Jia et al. (2009). "Imagenet: A large-scale hierarchical image database". In: *2009 IEEE conference on computer vision and pattern recognition*. Ieee, pp. 248–255.

Descloux, Adrien et al. (2020). "High-speed multiplane structured illumination microscopy of living cells using an image-splitting prism". In: *Nanophotonics* 9.1, pp. 143–148.

Donnert, Gerald et al. (2006). "Macromolecular-scale resolution in biological fluorescence microscopy". In: *Proceedings of the National Academy of Sciences* 103.31, pp. 11440–11445.

Dowden, Helen and Jamie Munro (2019). "Trends in clinical success rates and therapeutic focus". In: *Nat. Rev. Drug Discov* 18.7, pp. 495–6.

Drews, Jurgen (2000). "Drug discovery: a historical perspective". In: *science* 287.5460, pp. 1960–1964.

Durand, Audrey et al. (2018). "A machine learning approach for online automated optimization of super-resolution optical microscopy". In: *Nature communications* 9.1, pp. 1–16.

Dutta, Abhishek and Andrew Zisserman (2019). "The VIA annotation software for images, audio and video". In: *Proceedings of the 27th ACM international conference on multimedia*, pp. 2276–2279.

Eastman, Anna E and Shangqin Guo (2020). "The palette of techniques for cell cycle analysis". In: *FEBS letters* 594.13, pp. 2084–2098.

Exman, Pedro, Romualdo Barroso-Sousa, and Sara M Tolaney (2019). "Evidence to date: talazoparib in the treatment of breast cancer". In: *OncoTargets and therapy* 12, p. 5177.

Fu, Jingyan and Chuanmao Zhang (2019). "Super-resolution microscopy: successful applications in centrosome study and beyond". In: *Biophysics Reports* 5.5, pp. 235–243.

Ganem, Neil J, Susana A Godinho, and David Pellman (2009). "A mechanism linking extra centrosomes to chromosomal instability". In: *Nature* 460.7252, pp. 278–282.

Gould, Travis J et al. (2013). "Auto-aligning stimulated emission depletion microscope using adaptive optics". In: *Optics letters* 38.11, pp. 1860–1862.

Gwosch, Klaus C et al. (2020). "MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells". In: *Nature methods* 17.2, pp. 217–224.

Harris, Charles R. et al. (Sept. 2020). "Array programming with NumPy". In: *Nature* 585.7825, pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2.

Hasin, Yehudit, Marcus Seldin, and Aldons Lusis (2017). "Multi-omics approaches to disease". In: *Genome biology* 18.1, pp. 1–15.

He, Kaiming et al. (2016). "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.

He, Kaiming et al. (2017). "Mask r-cnn". In: *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969.

Heine, Jörn et al. (2017). "Adaptive-illumination STED nanoscopy". In: *Proceedings of the National Academy of Sciences* 114.37, pp. 9797–9802.

Heintzmann, Rainer and Thomas Huser (2017). "Super-resolution structured illumination microscopy". In: *Chemical reviews* 117.23, pp. 13890–13908.

Hell, Stefan W (2007). "Far-field optical nanoscopy". In: *science* 316.5828, pp. 1153–1158.

Hell, Stefan W and Jan Wichmann (1994). "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy". In: *Optics letters* 19.11, pp. 780–782.

Hua, Kiet and Russell J Ferland (2017). "Fixation methods can differentially affect ciliary protein immunolabeling". In: *Cilia* 6.1, pp. 1–17.

Hughes, James P et al. (2011). "Principles of early drug discovery". In: *British journal of pharmacology* 162.6, pp. 1239–1249.

Humbarwadi, Srihari (2020). *Object Detection with RetinaNet*. https://keras.io/examples/vision/retinanet. Accessed: 2022-02-10.

Hunter, J. D. (2007). "Matplotlib: A 2D graphics environment". In: *Computing in Science & Engineering* 9.3, pp. 90–95. DOI: 10.1109/MCSE.2007.55.

Jahr, Wiebke, Philipp Velicky, and Johann Georg Danzl (2020). "Strategies to maximize performance in STimulated Emission Depletion (STED) nanoscopy of biological specimens". In: *Methods* 174, pp. 27–41.

Johannes, Jeffrey W et al. (2015). "Discovery of AZ0108, an orally bioavailable phthalazinone PARP inhibitor that blocks centrosome clustering". In: *Bioorganic & medicinal chemistry letters* 25.24, pp. 5743–5747.

Kar, Anuradha et al. (2022). "Benchmarking of deep learning algorithms for 3D instance segmentation of confocal image datasets". In: *PLoS computational biology* 18.4, e1009879.

Klar, Thomas A et al. (2000). "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission". In: *Proceedings of the National Academy of Sciences* 97.15, pp. 8206–8210.

Konotop, Gleb et al. (2016). "Pharmacological inhibition of centrosome clustering by slingshot-mediated cofilin activation and actin cortex destabilization". In: *Cancer research* 76.22, pp. 6690–6700.

Krämer, Alwin, Bettina Maier, and Jiri Bartek (2011). "Centrosome clustering and chromosomal (in) stability: a matter of life and death". In: *Molecular oncology* 5.4, pp. 324–335.

Kwon, Mijung et al. (2008). "Mechanisms to suppress multipolar divisions in cancer cells with extra centrosomes". In: *Genes & development* 22.16, pp. 2189–2203.

Lang, Paul et al. (2006). "Cellular imaging in drug discovery". In: *Nature Reviews Drug Discovery* 5.4, pp. 343–356.

Lee, Yin Loon et al. (2014). "Cby1 promotes Ahi1 recruitment to a ring-shaped domain at the centriole–cilium interface and facilitates proper cilium formation and function". In: *Molecular Biology of the Cell* 25.19, pp. 2919–2933.

Lelek, Mickaël et al. (2021). "Single-molecule localization microscopy". In: *Nature Reviews Methods Primers* 1.1, pp. 1–27.

Li, Shuaizhang and Menghang Xia (2019). "Review of high-content screening applications in toxicology". In: *Archives of toxicology* 93.12, pp. 3387–3396.

Li, Yue et al. (2017). "The effects of chemical fixation on the cellular nanostructure". In: *Experimental cell research* 358.2, pp. 253–259.

Lin, Sean et al. (2020). "Image-based high-content screening in drug discovery". In: *Drug Discovery Today* 25.8, pp. 1348–1361.

Lin, Tsung-Yi et al. (2017). "Focal loss for dense object detection". In: *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988.

Liu, An-An et al. (2017). "Mitosis detection in phase contrast microscopy image sequences of stem cell populations: a critical review". In: *IEEE Transactions on Big Data* 3.4, pp. 443–457.

Lowe, David G (2004). "Distinctive image features from scale-invariant keypoints". In: *International journal of computer vision* 60.2, pp. 91–110.

Marteil, Gaëlle et al. (2018). "Over-elongation of centrioles in cancer promotes centriole amplification and chromosome missegregation". In: *Nature communications* 9.1, pp. 1–17.

Moen, Erick et al. (2019). "Deep learning for cellular image analysis". In: *Nature methods* 16.12, pp. 1233–1246.

Moffitt, Jeffrey R, Christian Osseforth, and Jens Michaelis (2011). "Time-gating improves the spatial resolution of STED microscopy". In: *Optics express* 19.5, pp. 4242–4254.

Mogilner, Alex (2016). *Quantitative biology on the rise*.

Mullard, Asher (2022). "2021 FDA approvals". In: *Nature Reviews Drug Discovery* 21.2, pp. 83–88.

Navarro-Serer, Bernat et al. (2019). "Aurora A inhibition limits centrosome clustering and promotes mitotic catastrophe in cells with supernumerary centrosomes". In: *Oncotarget* 10.17, p. 1649.

Nigg, Erich A and Andrew J Holland (2018). "Once and only once: mechanisms of centriole duplication and their deregulation in disease". In: *Nature reviews Molecular cell biology* 19.5, pp. 297–312.

Ogden, A, PCG Rida, and R Aneja (2012). "Let's huddle to prevent a muddle: centrosome declustering as an attractive anticancer strategy". In: *Cell Death & Differentiation* 19.8, pp. 1255–1267.

OpenCV feature matching (2021). *OpenCV-Python Tutorial on Feature Matching*. `https://docs.opencv.org/3.4/dc/dc3/tutorial_py_matcher`. Accessed: 2021-07-19.

Pannu, Vaishali et al. (2014). "Centrosome-declustering drugs mediate a two-pronged attack on interphase and mitosis in supercentrosomal cancer cells". In: *Cell death & disease* 5.11, e1538–e1538.

Paul, Steven M et al. (2010). "How to improve R&D productivity: the pharmaceutical industry's grand challenge". In: *Nature reviews Drug discovery* 9.3, pp. 203–214.

Pedregosa, F. et al. (2011). "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Rebacz, Blanka et al. (2007). "Identification of griseofulvin as an inhibitor of centrosomal clustering in a phenotype-based screen". In: *Cancer research* 67.13, pp. 6342–6350.

Rust, Michael J, Mark Bates, and Xiaowei Zhuang (2006). "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)". In: *Nature methods* 3.10, pp. 793–796.

Sabat-Pośpiech, Dorota et al. (2019). "Targeting centrosome amplification, an Achilles' heel of cancer". In: *Biochemical Society Transactions* 47.5, pp. 1209–1222.

Sahl, Steffen J, Stefan W Hell, and Stefan Jakobs (2017). "Fluorescence nanoscopy in cell biology". In: *Nature reviews Molecular cell biology* 18.11, pp. 685–701.

Sandler, Mark et al. (2018). "Mobilenetv2: Inverted residuals and linear bottlenecks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520.

Santos, Rita et al. (2017). "A comprehensive map of molecular drug targets". In: *Nature reviews Drug discovery* 16.1, pp. 19–34.

Schamberger, Jens et al. (2011). "Rendezvous in chemical space? Comparing the small molecule compound libraries of Bayer and Schering". In: *Drug discovery today* 16.13-14, pp. 636–641.

Schindelin, Johannes et al. (2012). "Fiji: an open-source platform for biological-image analysis". In: *Nature methods* 9.7, pp. 676–682.

Schmidt, Roman et al. (2021). "MINFLUX nanometer-scale 3D imaging and microsecond-range tracking on a common fluorescence microscope". In: *Nature communications* 12.1, pp. 1–12.

Schorl, Christoph and John M Sedivy (2007). "Analysis of cell cycle phases and progression in cultured mammalian cells". In: *Methods* 41.2, pp. 143–150.

Shannon, Claude E (1949). "Communication in the presence of noise". In: *Proceedings of the IRE* 37.1, pp. 10–21.

Sisay, Mekonnen and Dumessa Edessa (2017). "PARP inhibitors as potential therapeutic agents for various cancers: focus on niraparib and its first global approval for maintenance therapy of gynecologic cancers". In: *Gynecologic oncology research and practice* 4.1, pp. 1–13.

Staudt, Thorsten et al. (2011). "Far-field optical nanoscopy with reduced number of state transition cycles". In: *Optics express* 19.6, pp. 5644–5657.

Stehr, Florian et al. (2019). "Flat-top TIRF illumination boosts DNA-PAINT imaging and quantification". In: *Nature communications* 10.1, pp. 1–8.

Tan, Mingxing and Quoc Le (2019). "Efficientnet: Rethinking model scaling for convolutional neural networks". In: *International conference on machine learning*. PMLR, pp. 6105–6114.

Tony Yang, T et al. (2015). "Superresolution pattern recognition reveals the architectural map of the ciliary transition zone". In: *Scientific reports* 5.1, pp. 1–13.

Uzbekov, Rustem and Irina Alieva (2018). "Who are you, subdistal appendages of centriole?" In: *Open biology* 8.7, p. 180062.

Vicidomini, Giuseppe, Paolo Bianchini, and Alberto Diaspro (2018). "STED super-resolved microscopy". In: *Nature methods* 15.3, pp. 173–182.

Virtanen, Pauli et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17, pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

Vlijm, Rifka et al. (2018). "STED nanoscopy of the centrosome linker reveals a CEP68-organized, periodic rootletin network anchored to a C-Nap1 ring at centrioles". In: *Proceedings of the National Academy of Sciences* 115.10, E2246–E2253.

Walt, Stefan Van der et al. (2014). "scikit-image: image processing in Python". In: *PeerJ* 2, e453.

Weber, Michael et al. (2021). "MINSTED fluorescence localization and nanoscopy". In: *Nature photonics* 15.5, pp. 361–366.

Weber, Michael et al. (2022). "MINSTED nanoscopy enters the Angstrom localization range". In: *bioRxiv*.

Westrate, Laura M et al. (2014). "Mitochondrial morphological features are associated with fission and fusion events". In: *PloS one* 9.4, e95265.

Whelan, Donna R and Toby DM Bell (2015). "Image artifacts in single molecule localization microscopy: why optimization of sample preparation protocols matters". In: *Scientific reports* 5.1, pp. 1–10.

Wong, Yao Liang et al. (2015). "Reversible centriole depletion with an inhibitor of Polo-like kinase 4". In: *Science* 348.6239, pp. 1155–1160.

Wurm, Christian A et al. (2010). "Sample preparation for STED microscopy". In: *Live cell imaging*. Springer, pp. 185–199.

Yang, Bin et al. (2014). "Large parallelization of STED nanoscopy using optical lattices". In: *Optics express* 22.5, pp. 5581–5589.

Zanella, Fabian, James B Lorens, and Wolfgang Link (2010). "High content screening: seeing is believing". In: *Trends in biotechnology* 28.5, pp. 237–245.

Zheng, Wei, Natasha Thorne, and John C McKew (2013). "Phenotypic screens as a renewed approach for drug discovery". In: *Drug discovery today* 18.21-22, pp. 1067–1073.

Acknowledgements

First and foremost, I want to thank Prof. Dr. Stefan Jakobs for supervising my PhD thesis and for giving me the opportunity to implement my concept of automated STED microscopy in his laboratory in Göttingen. In addition, my gratitude goes to the further members of my thesis advisory committee, Prof. Dr. Andreas Janshoff and Prof. Dr. Stefan W. Hell, for providing me with invaluable feedback over the past three years. I also cordially thank the members of my examination board, Prof. Dr. Stefan Jakobs, Prof. Dr. Andreas Janshoff, Prof. Dr. Stefan W. Hell, Prof. Dr. Sarah Köster, apl. Prof. Dr. Alexander Egner and Dr. Helmut Berg.

Furthermore, I am grateful to all the members of the Jakobs research group for many fruitful discussions and their kind and candid help whenever needed. In particular, I want to thank Dr. Daniel Jans and Dr. Christian Brüser for their support during my measurement visits in Göttingen, their great help in all situations regarding the microscope and their valuable assistance during remote measurements.

Likewise, my gratitude goes to my supervisors at Bayer, Dr. Helmut Berg and Dr. David Schneider, for the open and trusting athmosphere as well as their continuous support in the course of the *Screening-STED* project. I want to especially thank Dr. Helmut Berg for giving me the freedom to leave the beaten track and pursue my ideas in Göttingen. I am also grateful to Dr. Marc Osterland and Dr. Paula Zapata from Bayer Machine Learning Research as well as to my colleagues from Bayer Computer Vision Innovation, especially Dr. Anne Gehre and Dr. Wei Liao, for the useful exchange and advice regarding all topics of machine learning and classical image analysis.

Moreover, I would like to thank my colleagues from the Bayer Screening Technology group for their support in the past two years, especially Robin Gsell for many hints regarding object-oriented programming and Dr. Pascal Lambertz for illuminating insights into new modalities for disease treatment. Not least, I also want to thank my colleagues Jens Sicking, Dr. Theresa Vogt and Nedal Darif from the former Bayer Applied Physics group for the great time together in the lab.

I am grateful for the peer cooperation with Abberior Instruments within the jointly executed *Screening-STED* project. I especially want to thank Dr. Christian Wurm for his helpful advice regarding sample preparation and Dr. Jörn Heine for his great support and valuable assistance concerning all aspects of microscope automation. I also appreciate the enlighting technological discussions with Dr. Joachim Fischer and the helpful support regarding antibodies and labeling strategies provided by Dr. Jan-Gero Schlötel, Jasmin Rehman and Dr. Florian Grimm.

I am highly indebted to Jana Marquet, Dr. Pascal Lambertz, Dr. David Schneider, Dr. Helmut Berg, Dr. Anne Gehre, Dr. Marc Osterland, Dr. Katja Grossmann and Dr. Daniel Jans for proofreading my thesis manuscript. Thank you very much!

My deepest gratitude is due to my family, Jana and Luca, for their extensive and continuous encouragement and their motivating spirit within the past three years.