Using Neural Network Tools for Fast Point Spread Function Convolutions

University of New Mexico Department of Physics and Astronomy

Joel J. Gates

27 April 2025

Abstract

The field of super-resolution microscopy has made imaging of biological systems at sub-diffraction limit resolutions possible. A prominent technique in super-resolution is single molecule localization microscopy (SMLM). Analysis of the dynamics of live cells is also possible in single molecule imaging using single particle tracking (SPT) techniques. The production of super-resolved images and single particle tracking data using traditional algorithms is computationally expensive. However, it has recently been shown that deep learning methods are particularly useful in SMLM and SPT, expanding imaging capabilities. To train these neural networks, it is necessary to produce large datasets that accurately represent experimental data. Data generation and model training takes significantly more computation time than the applied use of a model. We demonstrate a method, PSFConvolutions, through which neural network architecture is used for the efficient production of deep learning training datasets in SMLM and SPT. This method is implemented in an easy-to-use programming interface which can be applied to produce synthetic data modelling commonly used microscope setups in SMLM/SPT.

Acknowledgements

I would like to specially acknowledge my honors thesis advisor, Dr. Keith Lidke. His advice, contributions and mentorship have been crucial in the completion of this project. Additionally I would like to thank the entire Lidke lab group for their help in completing this project.

Contents

Chapter 1

Introduction

1.1 Diffraction Limited Imaging

The use of optical microscopy for imaging biological systems has an extended history. Many techniques were developed in the 20th century to improve optical microscopy capabilities. Traditional techniques such as phase contrast and confocal microscopy, however, are diffraction limited. This means that the distance between point-like emitters can only be resolved to a certain level defined by the Rayleigh limit defined below

$$R = \frac{0.61\lambda}{\mathrm{NA}} \tag{1.1}$$

where R is the Rayleigh limit, λ is the wavelength of light used in the microscope and NA is the Numerical Aperture of the objective lens. The numerical aperture of a microscope is defined as

$$\mathrm{NA} = n \sin\theta, \tag{1.2}$$

where n is the index of refraction of the objective lens and θ is the angle of the greatest inclined beam from the optical axis that can pass through the optical systems aperture demonstrated in Figure 1.1.

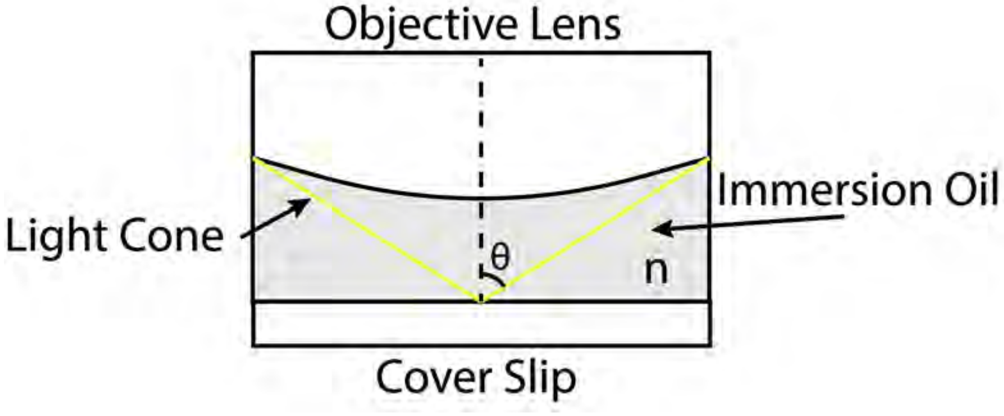


Figure 1.1: Numerical aperture of an objective lens. The maximum angle is denoted as θ and n refers to the index of refraction of the immersion oil used between the cover slip and microscope objective. This corresponds to a numerical aperture of $n \sin\theta$.

This places the best resolution possible for microscopes using light in the visible spectrum at approximately 200 nm [1]. The process of two point sources overlapping until they are indistinguishable is demonstrated in Figure 1.2.

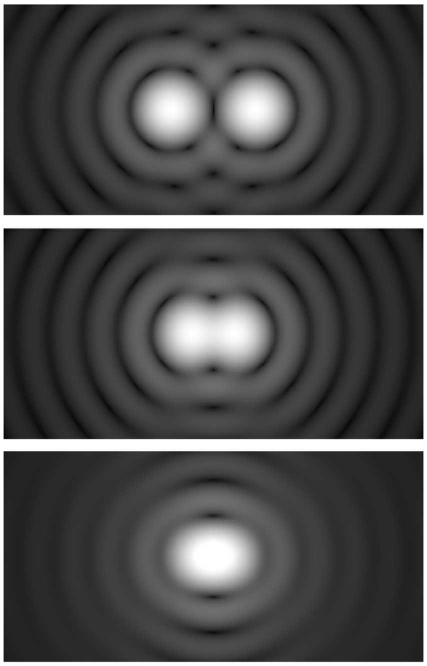


Figure 1.2: Two point sources moving closer together until their point spread functions become indistinguishable from each other and precise localization of each particle becomes impossible [2].

1.1.1 Point Spread Functions

A point source placed at the front focal plane of a microscope is imaged at the image plane as the point spread function (PSF) of the microscope. These point spread functions are obtained by modeling the diffraction of light from the point source through the entire optical setup. They can be considered the impulse response function of the microscope. For a point source passing through a circular aperture then focused to a point, such as in the simplified microscope setup shown in Figure 1.3. The lenses in this figure are circular in shape and act as the aperture of the system.

In the ideal case of an optical system with a circular aperture and no aberrations, the PSF can be calculated simply by a Fourier transform of the aperture [3]. This also makes an assumption of far field or Fraunhofer diffraction limit. The Fourier transform of a circular aperture is known better as the Airy disk. The intensity pattern of an Airy disk is shown as:

$$\begin{aligned} I(r) &\propto \left(\frac{2J_0(V(r))}{V(r)}\right)^2, \\ V(r) &= \frac{2\pi ar}{\lambda R}, \end{aligned} \tag{1.3}$$

where a is the radius of the aperture, R is the distance from the aperture to the image, λ is the wavelength of light used, and J_0 is the spherical Bessel function of the first kind. An example Airy disk is shown in Figure 1.4.

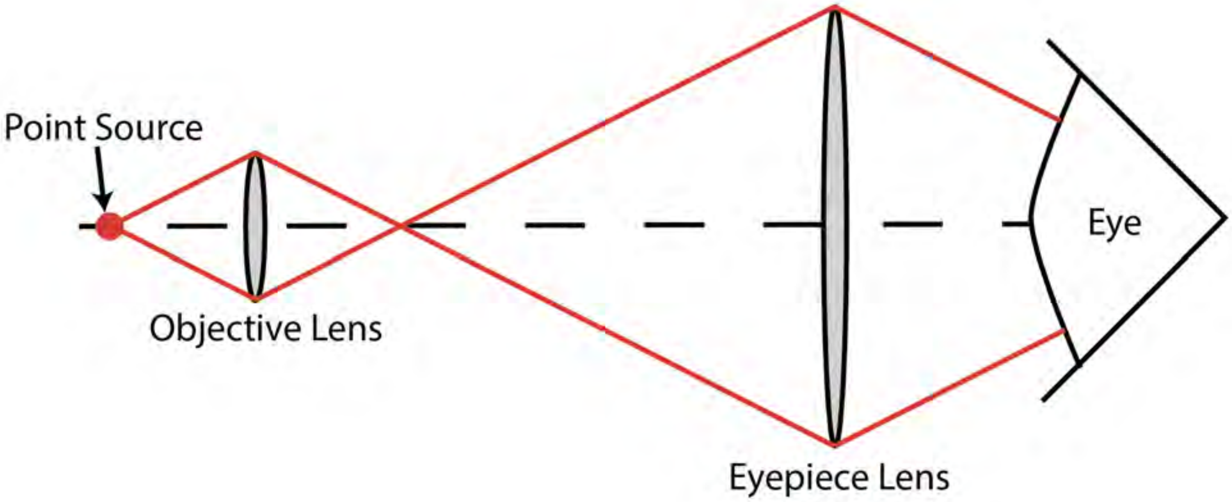


Figure 1.3: A simple two lens microscope system with no additional components. The lenses are assumed to be circularly shaped and an example ray path is shown.

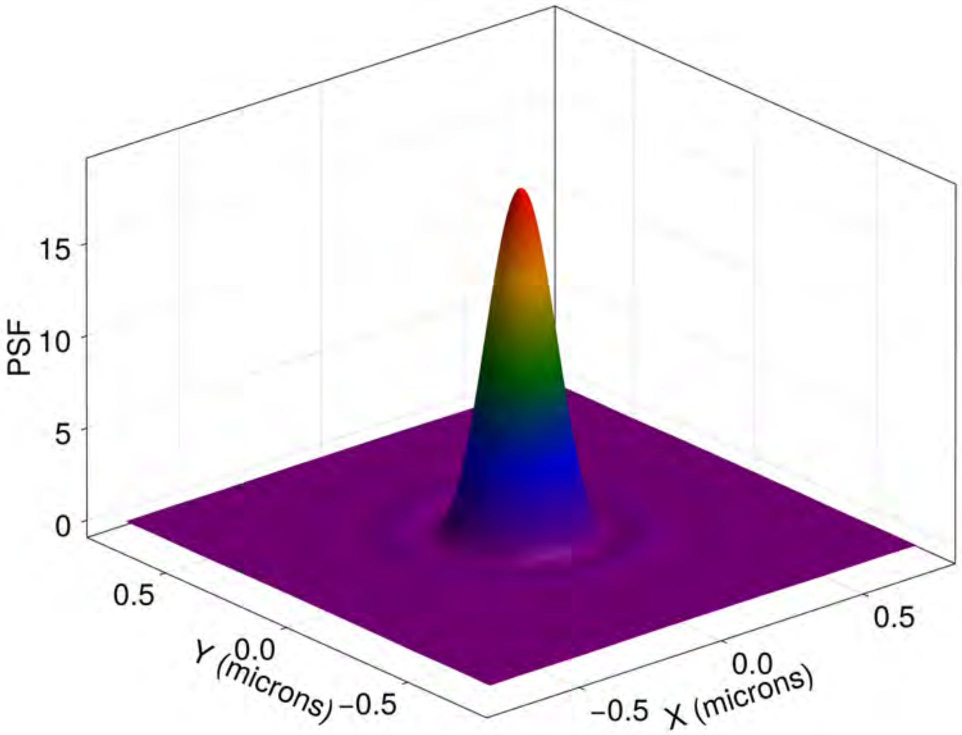


Figure 1.4: An example normalized airy disk pattern from a microscope objective with an NA of 1.3 and 533 nm light source.

In the low photon limit the PSF is normalized and considered a probability distribution of photon distribution. The Rayleigh limit shown in equation 1.1 corresponds to the point at which the maximum of one Airy disk overlaps with the minimum of a neighboring Airy disk. A microscope image will appear as a convolution of the microscope's PSF with the object at the focal plane.

1.2 Single Molecule Localization Microscopy

One may seek to improve resolution in diffracted limited microscopy by the usage of higher frequency electromagnetic radiation outside the visible range. Additionally, electron microscopy can increase resolution by using low wavelength electrons as a light source. In situations where samples may be damaged by high energy photons or electron beams, such as in biological imaging of tissues, cells, and subcellular structures, these techniques cannot often be applied. Instead, changes are made to imaging conditions and processing that allow for the reconstruction of super resolution images, breaking the diffraction limit.

To determine how these sub-diffraction resolutions can be achieved, one must consider the formulation of the Rayleigh limit. The resolution limit is derived in terms of the limit for which two particle's locations can be determined when observed at the image plane of the microscope. However, this is not a measurement of how precisely a single, isolated emitter can be localized.

After making an approximation of an Airy disk by a 2D Gaussian, a separate useful metric is the accuracy to which a singular point source emitter can be localized by fitting a PSF model to the observed light pattern. This maximum precision of this fit can be estimated by analyzing the Cramer Rao lower bound (CRLB) of the fit PSF model [4]. An analytical expression for this minimum uncertainty is described by the equation below [5].

$$\sigma_{\text{loc}} \geq \sqrt{\left(\frac{\sigma_0^2 + \frac{a^2}{12}}{N}\right)\left(\frac{16}{9} + \frac{8\pi(\sigma_0^2 + \frac{a^2}{12}b^2)}{Na^2}\right)}, \tag{1.4}$$

where a is the pixel size and b is the background intensity, σ_0 is the FWHM of the PSF and N is the number of photons collected at the image plane. We see that in the large photon limit, the equation reduces to the following form:

$$\sigma_{loc} \geq \frac{\sigma_0}{\sqrt{N}}, \tag{1.5}$$

where σ_{loc} is the localization precision, σ_0 is the width of the PSF, and N is the number of collected photons. With a stationary point source emitter and a large photon count, high precision localizations far below the Rayleigh limit are possible.

Making use of this principle to improve microscope resolution is the technique of single molecule localization microscopy (SMLM). Samples are carefully prepared with particles that fluoresce under laser illumination known as fluorophores. These fluorophores are stochastically activated such that only a subset of the total number of fluorophores emit light at any given point of time. Dependent on the class of fluorophores, the experimental procedure is altered. Commonly used methods are:

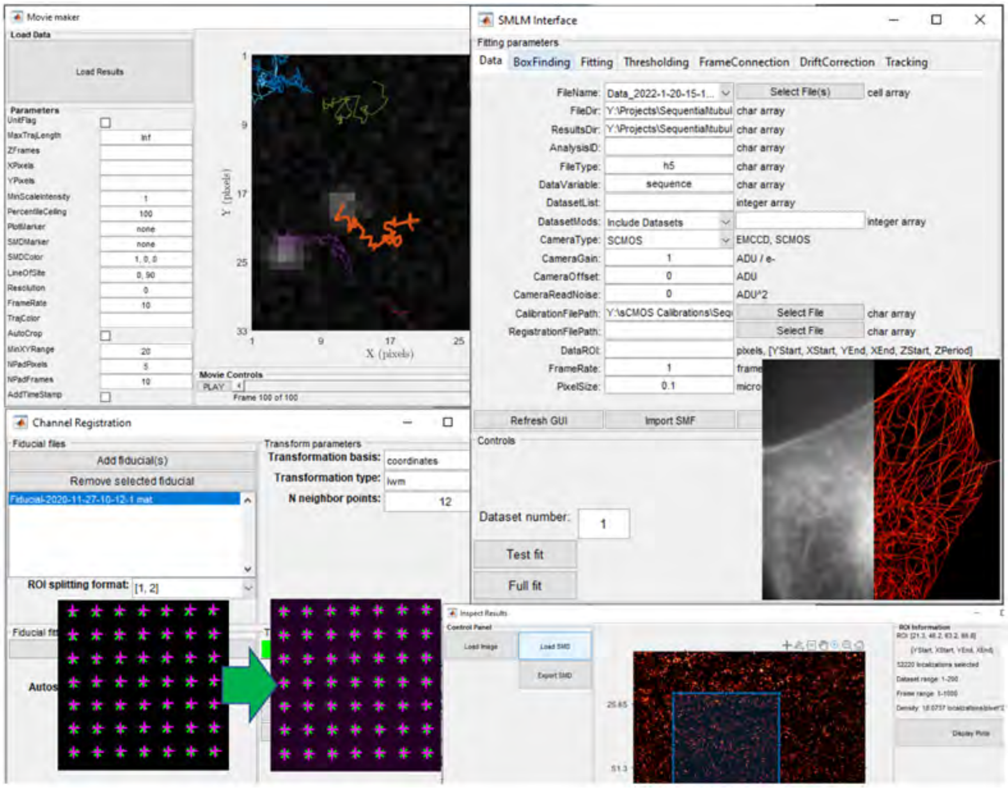
- Photoactivated localization microscopy (PALM) [6]: usage of fluorescent proteins.
- Stochastic optical reconstruction microscopy (STORM) [7]: usage of organic dyes.
- DNA point accumulation for imaging in nanoscale topography (DNA-PAINT) [8]: usage of fluorophores attached to imager DNA strands which bind to docking strands. These act as fluorescent probes.

These methods differ in their choice of fluorescent emitters, but all rely on stochastic emitter activation to create images consisting of non-overlapping PSFs.

1.2.1 Localization Process

Raw images taken in SMLM data acquisition processes appear as sparsely distributed PSFs. To find particle locations in each frame, algorithms are used to fit PSFs to each emitter and extract a centroid location of the emitter. Additional information such as emitter intensity and correlation to other emitters can be extracted from the fit PSFs. Full localization lists of all emitters and their locations are then reconstructed into a final composite image.

This process is computationally complex and software packages have been developed such as Single Molecule Imaging Toolbox Extraordinaire (SMITE) [9] to perform SMLM image processing and image reconstruction. An example graphical user interface is shown in Figure 1.5.



Movie maker

Load Data

Parameters
UnitFlag ☐
MaxTrajLength: Inf
ZFrames:
XPixels:
YPixels:
MinScaleIntensity: 1
PercentileCeiling: 100
PlotMarker: none
SMDMarker: none
SMDColor: 1, 0, 0
LineOfSite: 0, 90
Resolution: 0
FrameRate: 10
TrajColor:
AutoCrop ☐
MinXYRange: 20
NPadPixels: 5
NPadFrames: 10
AddTimeStamp ☐

Movie Controls
Frame 100 of 100

SMLM Interface

Fitting parameters

FileName: Data_2022-1-20-15-1... cell array
FileDir: Y:\Projects\SequentialTubul char array
ResultsDir: Y:\Projects\SequentialTubul char array
AnalysisID: char array
FileType: h5 char array
DataVariable: sequence char array
DatasetList: integer array
DatasetMods: Include Datasets integer array
CameraType: SCMOS EMCCD, SCMOS
CameraGain: 1 ADU / e-
CameraOffset: 0 ADU
CameraReadNoise: 0 ADU^2
CalibrationFilePath: Y:\sCMOS Calibrations\Seq char array
RegistrationFilePath: char array
DataROI: pixels, [YStart, XStart, YEnd, XEnd, ZStart, ZPeriod]
FrameRate: 1 frame
PixelSize: 0.1 micro

Controls
Dataset number: 1

Channel Registration

Fiducial files
Fiducial-2020-11-27-10-12-1.mat
ROI splitting format: [1, 2]

Transform parameters
Transformation basis: coordinates
Transformation type: lwm
N neighbor points: 12

Inspect Results

ROI Information
ROI: [21.3, 48.2, 63.2, 88.8]
[YStart, XStart, YEnd, XEnd]
52220 localizations selected
Dataset range: 1-200
Frame range: 1-1000

Figure 1.5: Single Molecule Imaging Toolbox Extraordinaire graphical user interface. This is a MATLAB plugin which allows for easy image reconstruction in SMLM.

1.2.2 3D Information

Imaging with 3D localization information is possible in SMLM. A common method of 3D localization is by the introduction of an astigmatic lens into the microscope setup. The resulting PSF of an astigmatic lens is modeled using linear combinations of Zernike modes at the pupil plane [10]. The modeled PSF is then fit to the observed PSF of each emitter, and the change in shape based on location in lateral space is determined.

1.2.3 Neural Networks applied to SMLM

SMLM data processing is possible by use of neural network frameworks. An implemented software package for SMLM known as the DECODE (DEep COntext DEpendent) method [11] has been

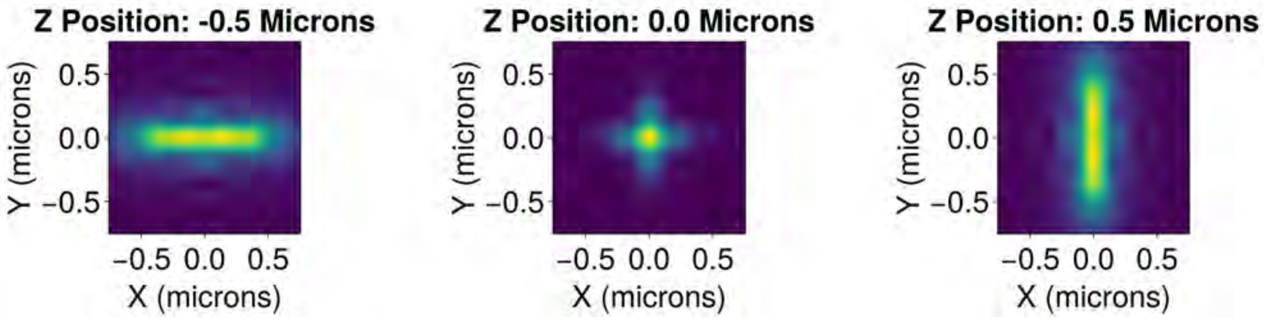


Figure 1.6: Point spread function with primary astigmatism. The change in PSF shape as the emitter moves in and out of focus is demonstrated. The Z Position in microns refers to the distance out of focus with 0.0 microns referring to the emitter being located at the focal plane.

shown to be an accurate and fast way to perform particle localizations and image reconstruction. To perform SMLM with DECODE, one must train a model consisting of parallel U-Nets to produce 10 images for each frame corresponding to the particles location in space, uncertainty in localization and intensity.

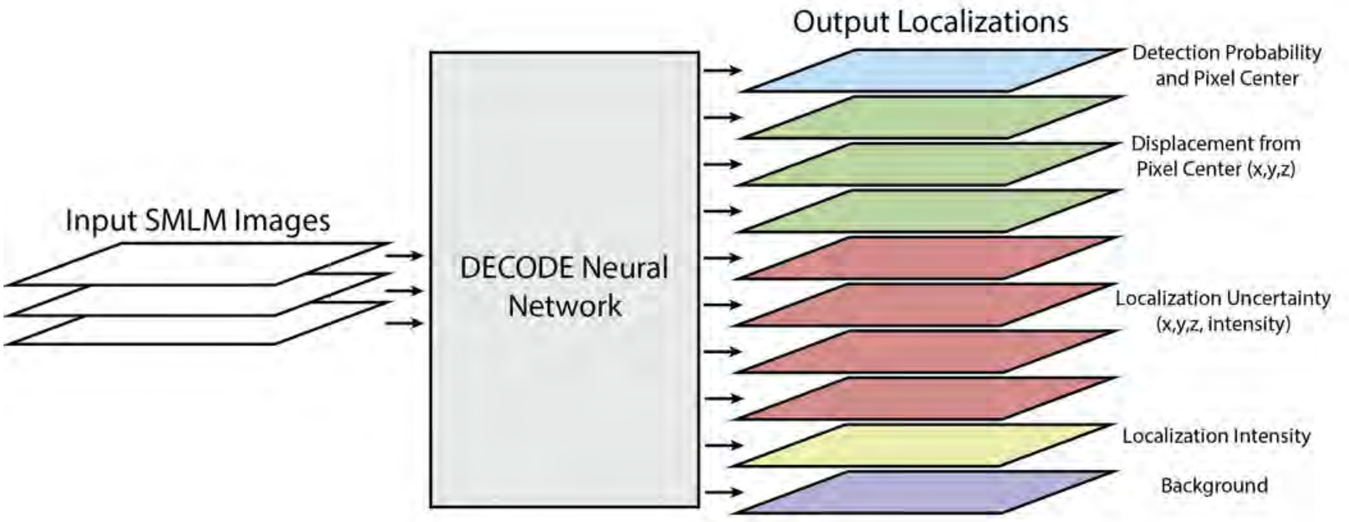


Figure 1.7: Demonstration of the DECODE neural network framework. Input images consisting of three frames in a dataset for temporal context are fed into a CNN based on the UNET architecture [12]. The output images contain information about each detected particles location in space and intensity with uncertainty, and must be processed into reconstructed images.

The usage of neural network models for particle localizations requires training a model with truthful emitter locations and realistic microscope images. Training of a model with synthetic data requires the production of images modeled to microscope specific parameters, as well as correct modeling of the fluorophores acting as emitters and the structures they are attached to.

This process is computationally expensive and reduces the applicability of deep learning frameworks for SMLM as the model must be trained for usage with datasets that accurately reproduce lab dependent imaging conditions.

1.2.4 Single Particle Tracking

Building from the methods of SMLM, single particle tracking (SPT) experiments are achievable. In SPT, structures of interest are labeled with fluorophores and imaged under conditions where they are free to move. At each frame, fluorophores are localized and identified throughout a movie. From this process, information about molecular dynamics of tagged particles can be extracted, such as the diffusion constant of freely diffusing labeled proteins. Software such as SMITE has been developed to produce single particle tracking data, however traditional algorithms are limited significantly by emitter density and the blinking behavior of fluorophores. Additional conditions such as low signal to noise ratio and molecules exhibiting non-diffusive motion patterns also limits the accuracy of traditional single particle tracking algorithms. Recent developments have shown deep learning approaches to be significantly useful in particle tracking applications [13].

1.2.5 Training Data Generation Improvements

To overcome limitations in the application of deep learning baseed SMLM and SPT techniques, we demonstrate an easy to implement algorithm, PSFConvolutions, which produces synthetic image datasets for applications in SMLM and SPT. With the the efficient production of training datasets, larger and more complex networks can be constructed and trained. For example, a neural network such as SPTNet must be trained with numerous multi-frame datasets that contain temporal information of the dynamics of molecules being simulated. The usage of the PSFConvolutions method to produce microscope images from datasets modeling the dynamics of molecules can increase the relevance of neural network models in SPT and SMLM.

Chapter 2

Methods

PSFConvolutions is implemented in the Julia Programming Language, making use of the LUX deep learning framework and ForwardDiff auto differentiation package. It is integrated into the larger JuliaSMLM toolset which contains many Julia packages developed for Single Molecule Localization Microscopy.

Example coordinate lists are generated using the SMLMData structure where for a given single molecule localization dataset (SMLD) the following information is stored:

- PSF
- Emitter Localizations (x,y,z, photon count, frame number, dataset number)
- Camera Information (pixel size, range)

Corresponding to the SMLD, a PSF must be generated. These are stored as distribution functions defined in the MicroscopePSFs package. To produce a PSF, one must specify the imaging conditions of the microscope that is simulated. The MicroscopePSFs package allows for the production of the following PSF models:

- Gaussian: 2D Gaussian approximation of the Airy disk
- Airy: PSF of circular aperture
- Scalar: Scalar diffraction model with Zernike coefficients
- Vector: Vector diffraction model considering the polarization of light
- Spline: Spline fit PSF models

The PSFConvolutions method generates images through a four step process illustrated in Figure 2.1. Prior to the usage of the PSFConvolutions package, one must first provide a list of coordinate data generated to model the fluorophores and structures which they are imaging. The camera and microscope characteristics are stored in the PSF model.

Both the image generation and convolution kernel generation processes differ when the inputted point spread function model is calculated in 2D or 3D space. Systems with PSFs that are engineered to change in shape such as in Figure 1.6 will require modeling using the 3D method. In contrast, systems with limited optical aberrations and imaging only near the focal plane, such as in total internal reflection (TIRF) microscopy, can be sufficiently modeled with an Airy disk or Gaussian PSF in 2D.

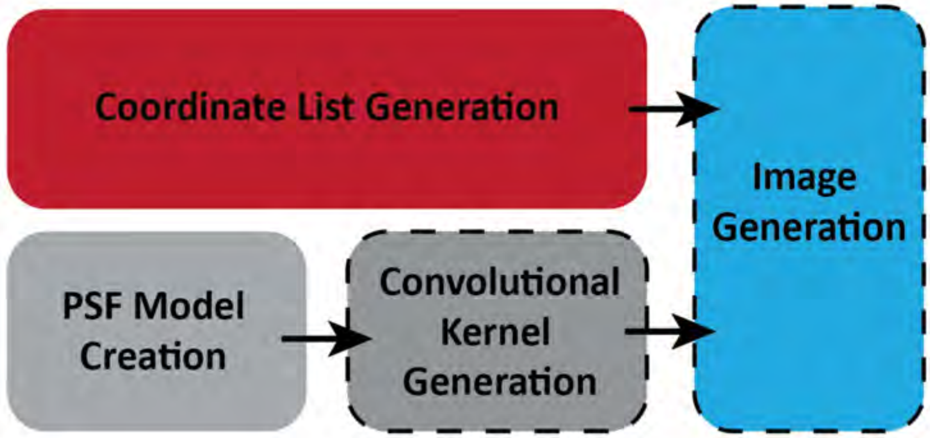


Figure 2.1: Flow chart describing the usage of the PSFConvolutions package for synthetic image generation. The coordinate list and PSF model must be supplied to PSFConvolutions.

From the inputted PSF and emitter coordinates, a series of images corresponding to the emitter pixel location and sub pixel displacement are generated and inputted into a Convolutional Neural Network (CNN) whose convolution kernel weights are determined by the PSFs.

2.1 Convolution Kernel Generation

The first step in image generation with PSFConvolutions is the creation of convolution kernels corresponding to how the point spread function of the microscope appears on the microscope's camera. This corresponds to integrating the analytic point spread function across the pixel size of the camera, creating a pixelated point spread function demonstrated in Figure 2.2. The magnitude of the pixelated PSF in each pixel corresponds to the probability that one photon will be incident on that specific pixel after passing through the microscope system. The size of the pixels in the pixelated PSF is supplied as the size of the pixels of the camera system being modeled. One may also wish to increase the sampling rate/upscale the image for data generation purposes when the analytic form of the PSF is known. This corresponds to sampling the PSF at more locations and decreasing the pixel size by a factor. The upscaled pixel size is described by the following expression:

$$a_u = \frac{a}{F}, \tag{2.1}$$

where a_u is the upscaled pixel size, a is the pixel size of the camera, and F is the upscaling factor applied in PSF generation.

2.1.1 2D Convolution Method

For 2D PSFs, sub-pixel displacement information is obtained by use of a 2D Taylor expansion. A general 2D Taylor expansion around points x and y is described as:

$$f(x+\Delta x, y+\Delta y) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{\delta^{(n+m)} f(x,y)}{\delta x^n \delta y^m} \frac{\Delta x^n \Delta y^m}{n!m!}, \tag{2.2}$$

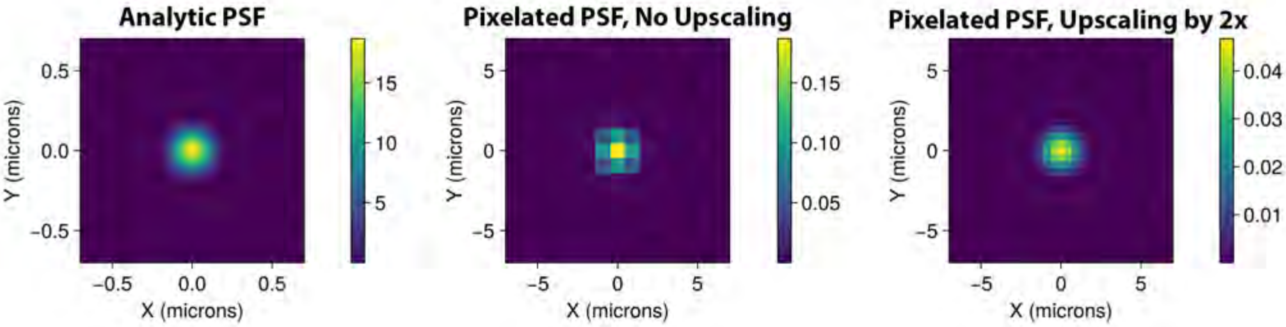


Figure 2.2: An Airy disk and its integrated pixelated versions. In the 2x upscaled version, the PSF is sampled with pixels half the size of the camera pixel.

where $f(x, y)$ corresponds to the function being expanded. Δx and Δy correspond to the distance from the center of the expansion. For calculations we truncate the Taylor expansion at a given order. The partial derivatives of the PSF are calculated using the Julia ForwardDiff auto differentiation package and saved as a list of 2D arrays demonstrated in Figure 2.3.

2.1.2 3D Convolution Method

A 3D Taylor expansion follows the following expression:

$$f(x + \Delta x, y + \Delta y, z + \Delta z) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{p=0}^{\infty} \frac{\delta^{(n+m+p)} f(x, y)}{\delta x^n \delta y^m \delta z^p} \frac{\Delta x^n \Delta y^m \Delta z^p}{n!m!p!}, \tag{2.3}$$

where $f(x, y, z)$ corresponds to the function being expanded. Δx, Δy, and Δz correspond to the distance from the center of the expansion. Similarly to the 2D convolution kernel generation method, 2D convolution kernels are saved which correspond to the partial derivatives of the PSF. To model the behavior of the function in the lateral (z) axis, multiple expansion centers in z for both the zeroth order kernel and its partial derivatives are chosen. An example set of 3D convolution kernels for an astigmatic PSF is shown in Figure 2.4.

2.2 Image Generation

Following the generation of convolution kernels, the image process generation begins. The dataset of emitter locations is sorted into individual frames. Emitters are stored with frame id, photon count, and position in space to sub-pixel accuracy, which is demonstrated in Figure 2.5. To generate images, we must create pixelated images according to the inputted camera range and pixel size. Binned images are created using pixels of size a_u as described in Equation 2.1.

Binned images corresponding to the emitters nearest pixel location and sub-pixel displacement are created and supplied to a neural network that begins with a number of channels specified by the order to which the the PSF is expanded in the kernel generation process. The creation of the convolution layer of the neural network differs based on the use of a 2D or 3D PSF. All activation

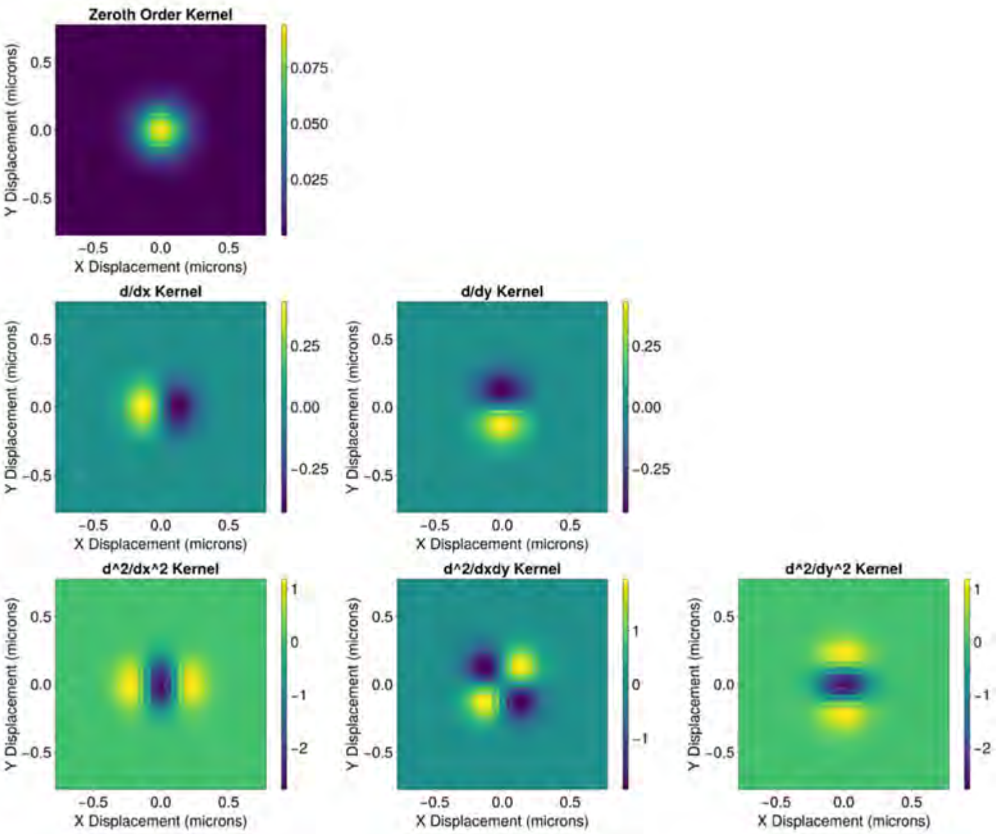


Figure 2.3: Second order expansion of a Gaussian PSF. First and second order partial derivatives of the Gaussian function calculated by forward auto differentiation are shown.

functions in this neural network are chosen to be the identity function and the combination of layers after the convolution layers is a summation. This is demonstrated in Figure 2.6.

2.2.1 2D Convolution Layer Generation

In the 2D case, The list of emitters from each frame is used to generate a subset of images corresponding to the nearest pixel location and the sub pixel offset. The Δx and Δy values correspond to the distance from the nearest pixel center. For example, if an emitter was specified to be at $x = 12.46, y = 12.53$ microns, and the nearest pixel center was $x = 12.5, y = 12.5$ microns, then the emitter would be placed in the $12.5, 12.5$ pixel with $\Delta x = -0.04$ and $\Delta y = 0.03$. The intensity of each emitter is scaled to equal the number of photons associated with the emitter. The values of the pixel shifts are stored in the nearest pixel center. High order terms such as $\Delta x^2, \Delta x \Delta y$ and Δy^2 are calculated using the sub-pixel shifts as differential values and populated into images placed at the same sub-pixel location. This process is illustrated in Figure 2.7.

Each image is convolved with its respective PSF or derivative of the PSF as specified in Equation 2.2. Correct weighting according to the Taylor series expansion is applied before convolution. For the 2D case, there are a number of input and convolution layers corresponding to the following

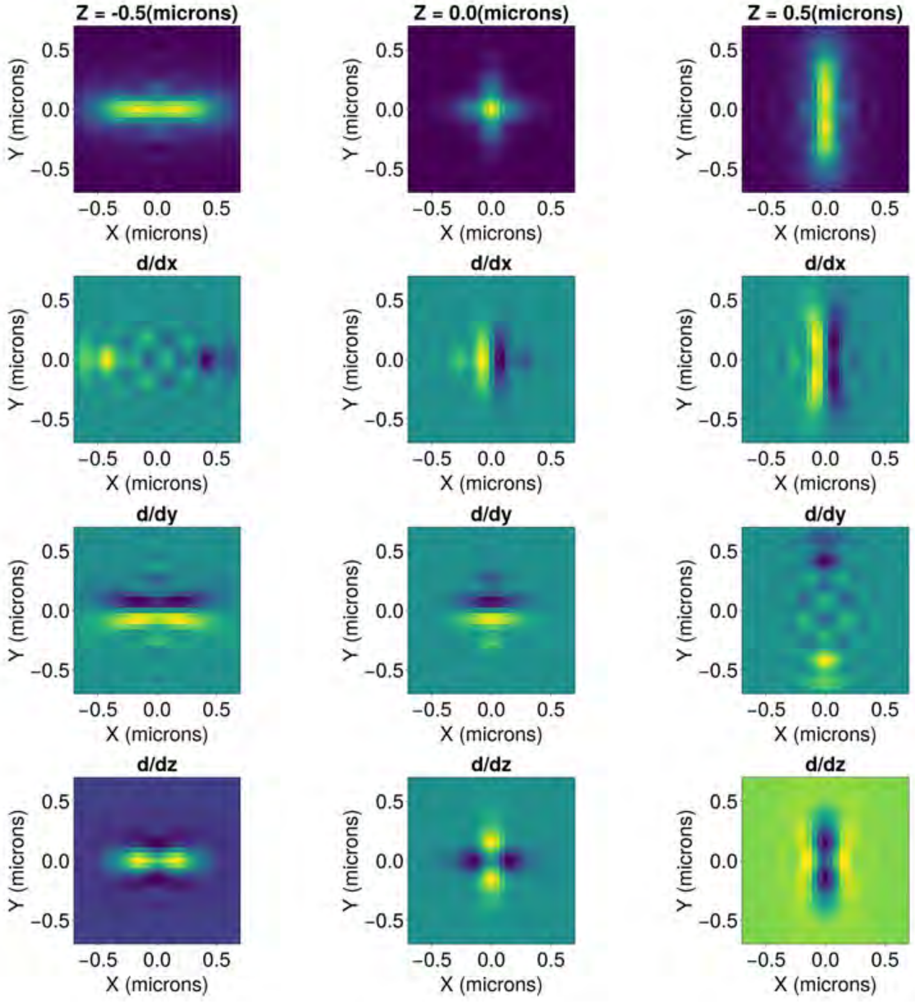


Figure 2.4: First order expansion of a PSF with λ astigmatic phase aberration and expansion centers at $z = -0.5, 0, 0.5$ microns from the focal plane.

expression:

$$\frac{(n+1)(n+2)}{2}, \tag{2.4}$$

where n is the Taylor expansion order.

2.2.2 3D Convolution Layer Generation

In the 3D case, a Δz value must also be calculated. This value is determined to be the distance from the emitter to the nearest expansion center on the lateral (z) axis.

For example, if the emitter location was specified to be located at $x = 12.46, y = 12.53, z = 0.34$ microns and the nearest pixel location in the axial plane was $x = 12.5, y = 12.5$ microns, the Δx and Δy values would still be -0.04 and 0.03 microns respectively. If we used 3 expansion centers at $z = -0.5, 0.0, 0.5$ microns such as in Figure 2.4, the emitter have a $\Delta z = -0.16$ microns relative to the expansion center at $z = 0.5$ microns.

Separate input images are created that correspond to each set of convolution kernels that correspond to each center in the axial direction and are only convolved with their respective convolution

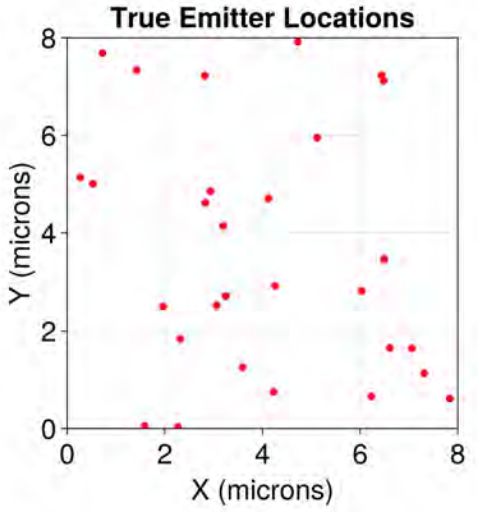


Figure 2.5: Example scatter plot of true emitter locations in one frame. The camera field of view is 8x8 microns. This is for the 2D case, so displacements from the focal plane are not visualized.

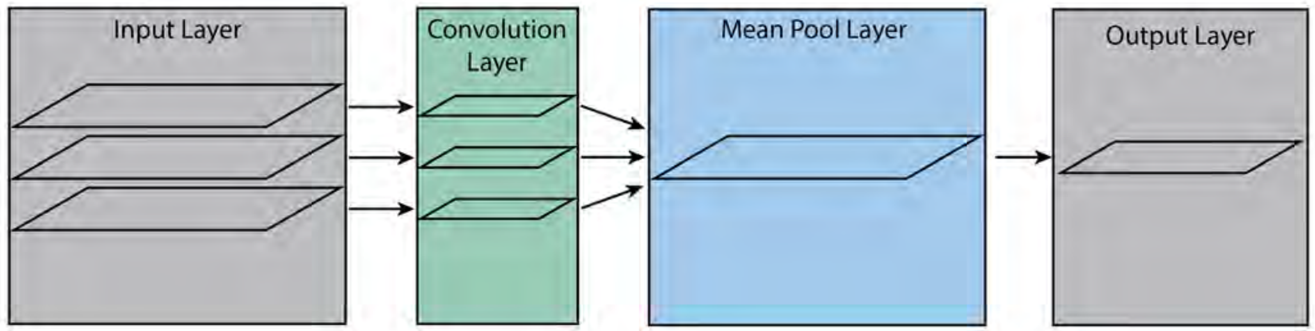


Figure 2.6: Diagram of the PSFConvolutions image generation neural network. A series of images generated from the emitter locations, intensities and sub-pixel displacements are inputted into a convolution layer built from the PSF and its partial derivatives. The mean pool layer averages pixels in a region corresponding to upscaling in PSF and input image generation, and the final image is outputted.

kernels as demonstrated in Figure 2.8.

This results in a total number of input and convolution layers described by the following expression:

$$\frac{m(n+1)(n+2)(n+3)}{6}, \tag{2.5}$$

where m is the number of z expansion centers and n is the specified expansion order.

2.2.3 Mean Pool Layer

Following the convolution layer, a mean pool layer is supplied. This averages neighboring pixels in a region of dimensions (F, F) where F is the same upscaling factor in Equation 2.1. A clever reader will note that if the PSF was properly normalized in the convolution kernel generation process according to the upscaled pixel size, the image after applying a mean pooling layer must be

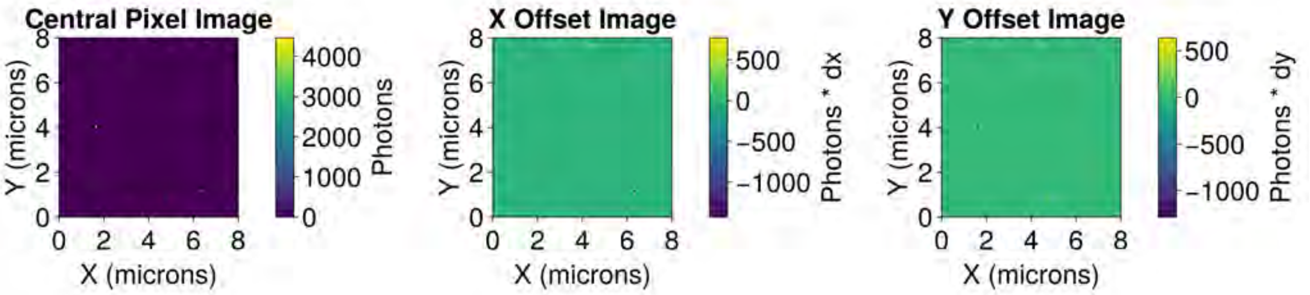


Figure 2.7: Zeroth order image and corresponding first order differential binned images. The intensity of each emitter is scaled to the number of photons emitted by each emitter.

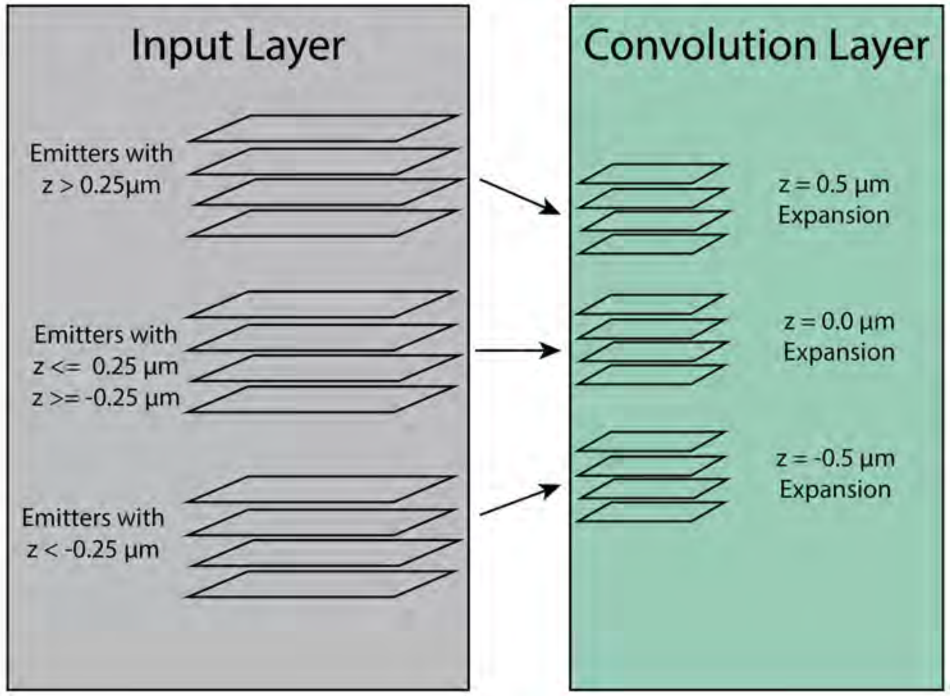


Figure 2.8: 3D input and convolution layers for a Taylor expansion of order 1 and 3 expansion centers. There are a total of 12 input and convolution layers.

multiplied by F^2. We address this by creating our upscaled convolution kernels such that they are normalized to F^2 instead of 1.

2.2.4 Output Layer

After mean pooling, the image is outputted from the neural network. This outputted image models the effects of diffraction and additional noise sources can be modeled after the use of PSFConvolutions.

2.3 Programming Tools Used

The following tools were used in the development of this module:

- Visual Studio Code (VS Code)
- GitHub Copilot VS Code Extension
- Julia Language Support VS Code Extension

Chapter 3

Results

Usage of the PSFConvolutions method for generating data is compared in accuracy and computation speed to the current most computationally efficient method of generating PSFs using spline fit models [14].

3.1 Accuracy

To benchmark the accuracy of the PSFConvolutions method, separate procedures were applied to the 2D and 3D methods. For 3D PSFs modeled using combinations of Zernike polynomials error relative to the spline fit models of the PSF is computed, providing an estimate of accuracy as the spline fit model will contain error.

3.1.1 2D Accuracy Results

To examine the error in sub-pixel shifts using the PSFConvolutions, we created a series of images of a single emitter of 1 photon intensity with a 2D Gaussian PSF. In separate frames, the emitter was stored with shifts up to 0.5 pixels from the central pixel. This allows for a calculation of error relative to an image generated by integrating a Gaussian PSF with the same sub-pixel displacement across each pixel. The analytical form of a 2D isotropic Gaussian distribution integrated across a square pixel region is given by the following expression:

$$I = \frac{1}{4}\left(\operatorname{erf}\left(\frac{x_1-\mu_x}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{x_0-\mu_x}{\sqrt{2}\sigma}\right)\right)\left(\operatorname{erf}\left(\frac{y_1-\mu_y}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{y_0-\mu_y}{\sqrt{2}\sigma}\right)\right), \tag{3.1}$$

where I is the intensity of the region, x_1 and x_0 are the pixel boundaries along the x axis, y_1 and y_0 are the pixel boundaries on the y axis, μ_x is the center of the point spread function in x, μ_y is the center of the point spread function in y, and σ is the standard deviation of the PSF.

Mean squared error (MSE) is calculated between the analytically computed image and PSFConvolutions image according to the following relation,

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2, \tag{3.2}$$

where N is the number of pixels, y_i are pixel values in the PSFConvolutions image, and x_i are pixel values in analytically computed image.

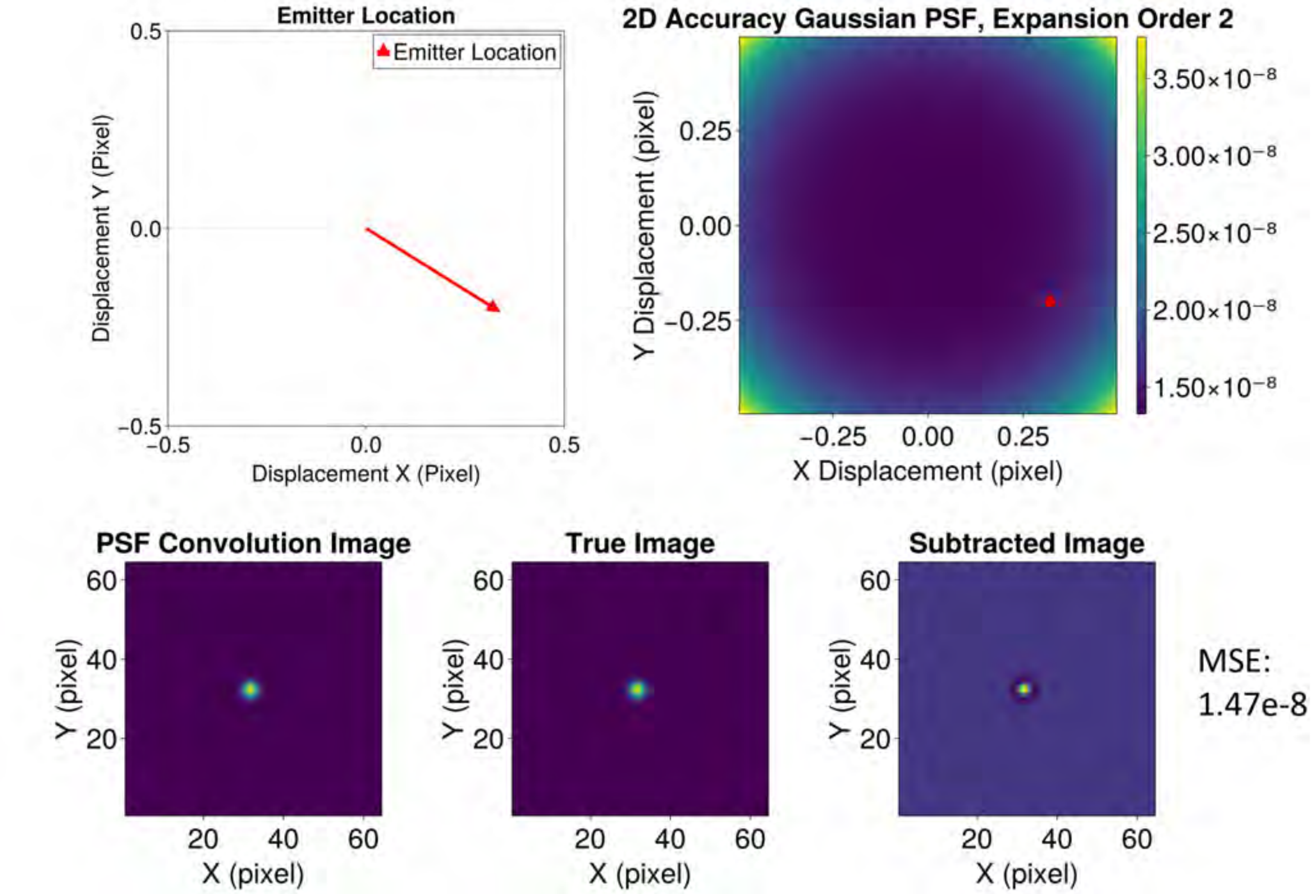


Figure 3.1: Mean squared error of the PSFConvolutions method plotted as a function of radial displacement with no upscaling and second order Taylor expansion. The mean squared error calculation of a sample point is demonstrated.

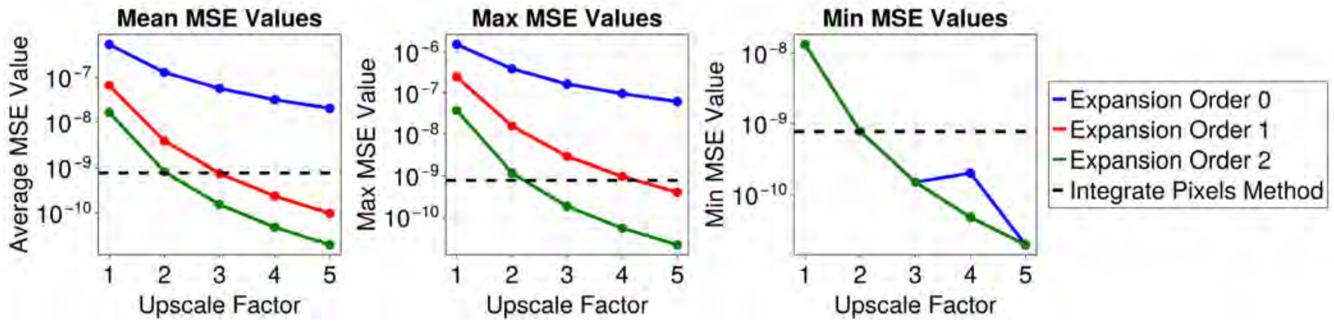


Figure 3.2: Mean squared error results for the PSFConvolutions method with varying upscale factor and expansion order. The inputted PSF is calculated on a 15x15 pixel region with a pixel size of 0.1 microns. The PSF modeled is a 2D Gaussian PSF with a standard deviation of 0.13 microns. For an upscale factor of 3 with a second order Taylor expansion, higher precision than the integrate pixels method is achieved.

Additionally we compare the MSE of the PSFConvolutions image with the MSE of images calculated using the integrate _ pixels method in MicroscopePSFs Julia package within the JuliaSMLM ecosystem. This method calculates an estimate of the integrated pixel image using Riemann sums with $\frac{1}{F}$ pixel size.

Examining MSE as a function of radial displacement from the central pixel demonstrates the increasing error of the PSFConvolutions method as the PSF is displaced from pixel centers demonstrated in Figure 3.1. However, increasing upscaling can provide a reduction in this error, resulting in accurate image generation demonstrated in Figure 3.2.

3.1.2 3D Accuracy Estimations

In the 3D case, accuracy of the PSFConvolutions method was compared to PSFs generated using a spline fit model. All PSFs generated for demonstrating the accuracy of the 3D convolutions method were made using a 3D scaler model of a point spread function with phase aberration of λ in the primary vertical/horizontal astigmatism mode illustrated in Figure 1.6. The wavelength of light modeled is 533 nm, with a 1.4 Numerical Aperture objective, and a refractive index of 1.5

MSE in this case is a metric of deviation between values calculated by the PSFConvolutions method and an integrated spline fit of the point spread function. Our spline model was calculated by sampling the PSF in a 2x2x1 micron range in 0.05 micron intervals. For PSFConvolutions, we create convolution kernels with an upscale factor of 3 and second order taylor expansion.

Our emitter dataset consists of emitters displaced from the $z = 0$ micron focal plane in a region of $(-0.5, 0.5)$ microns. The deviation from the Spline fit model increases as a function of distance from each z expansion center specified in creation of the input layers. This difference is shown in Figures 3.3 and 3.4.

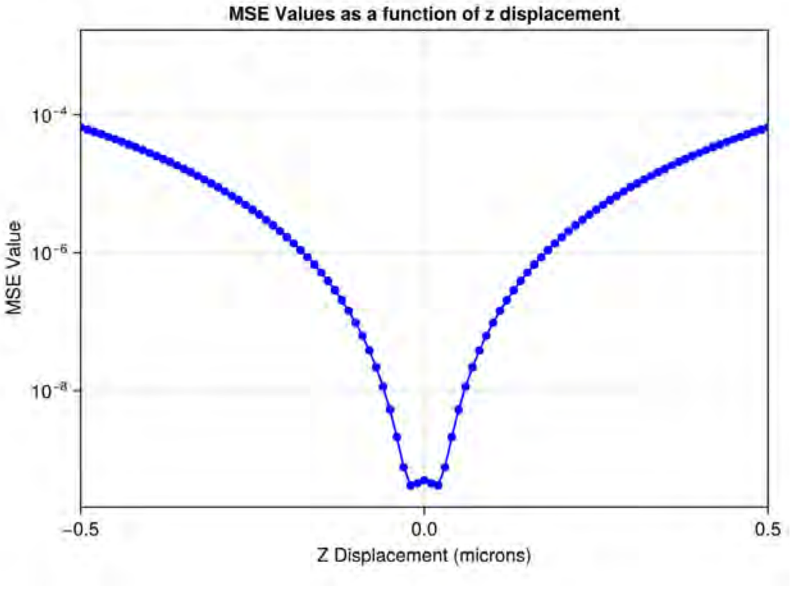


Figure 3.3: MSE values of as a function of displacement in emitter location from the focal plane. This is using 1 expansion center at the focal plane.

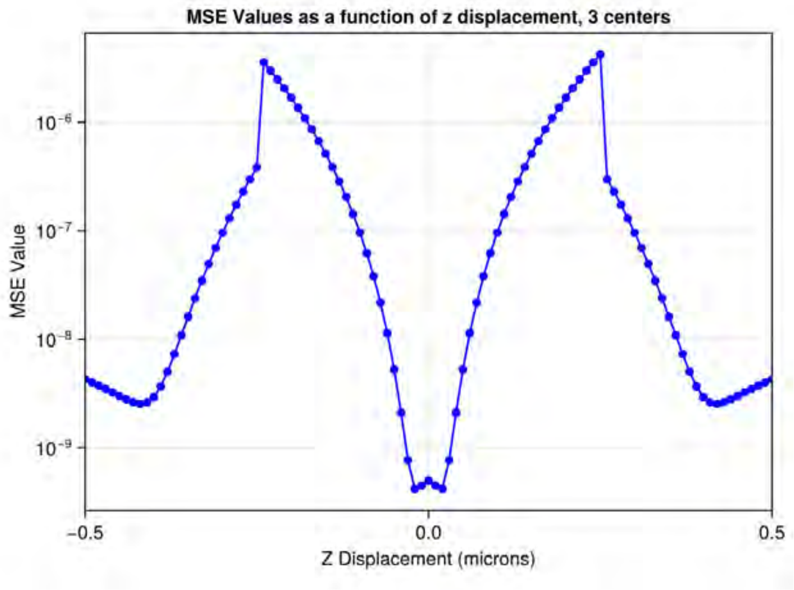


Figure 3.4: MSE values of as a function of displacement in emitter location from the focal plane. This is using 3 expansion centers at $z = (-0.5, 0.0, 0.5)$. We see a decrease in deviation from the spline fit model near the expansion centers and overall greater accuracy than the single expansion center figure.

3.2 Performance

To benchmark the performance of the PSFConvolutions method, we compare image generation using a test dataset of varying emitter density over a 6.4x6.4 micron camera region with 0.1 micron pixels. Test datasets contain 50 frames of emitters randomly placed at the specified densities ranging from $(\frac{1}{10}, 10)$ emitters per square micron. Spline PSF calculations are performed using a 16 core CPU whereas PSFConvolutions calculations are performed on a Nvidia RTX 4090 GPU. This is pictured in Figure 3.5.

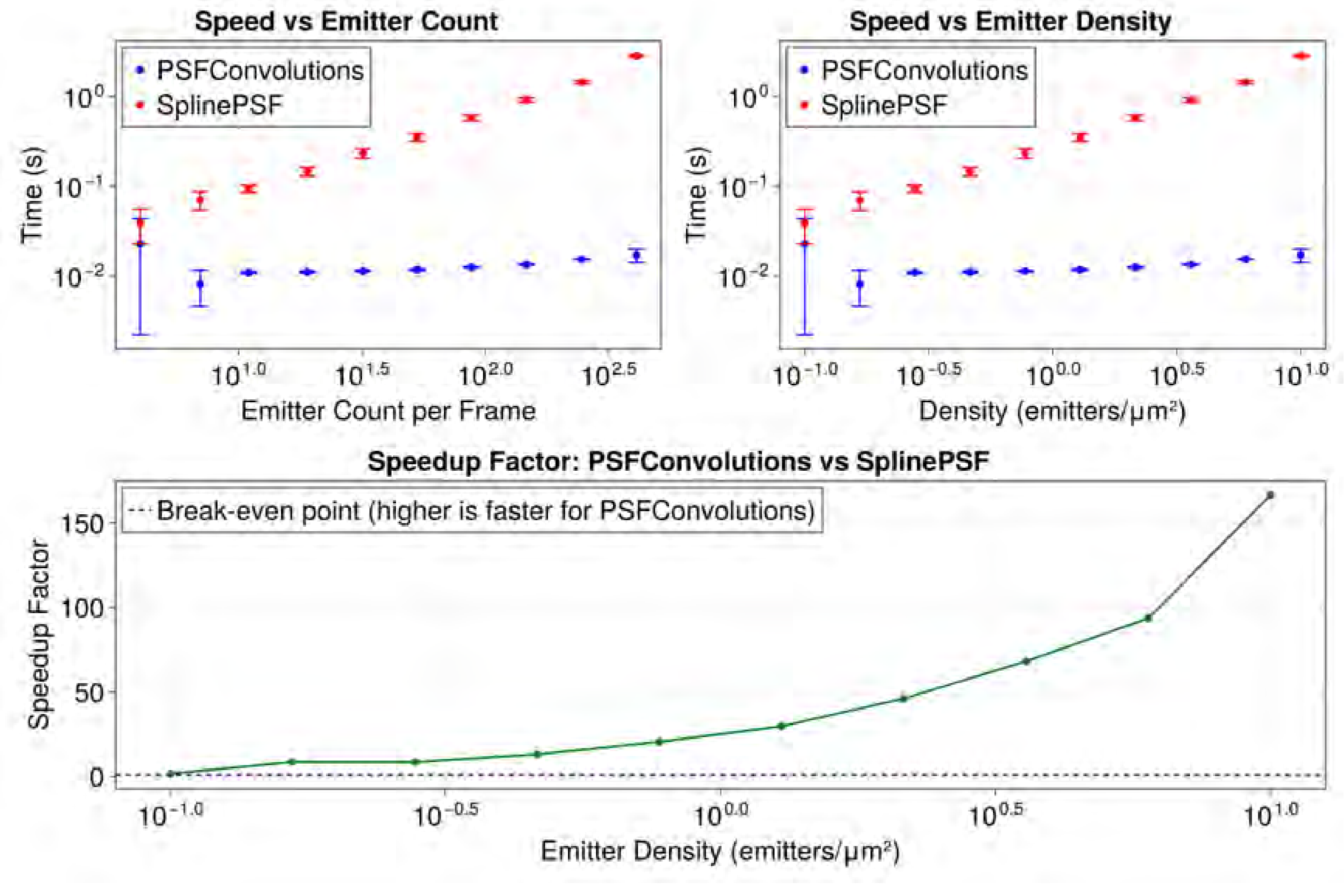


Figure 3.5: Speed of PSFConvolutions method vs image generation using an integrated spline PSF. For high emitter densities, the PSFConvolutions method performs image generation two orders of magnitude quicker than integrating a spline fit PSF.

Chapter 4

Conclusions

The PSFConvolutions image generation method is shown to be an accurate and efficient method of producing diffraction modeled images for SMLM and SPT purposes. This method produces images using a CNN with convolution kernels weighted by the Taylor expansion of an inputted PSF. This method is up to 150 times faster than existing image generation for high emitter density image generation, and performs with increased accuracy in comparison to existing image generation in the 2D case. Calculations are performed using GPUs for increased performance over commonly used CPU image generation methods. This method is highly applicable for use in training complex neural networks for SMLM and SPT such as DECODE and SPTNet. Improvements to computation speed of training data allows for the generation of larger and more complex training datasets which improves the relevance of deep learning methods for super resolution microscopy applications.

Bibliography

[1] David Mayerich and Ruijiao Sun. "Chapter Twelve - Superresolution Image Processing". In: *Microscope Image Processing (Second Edition)*. Ed. by Fatima A. Merchant and Kenneth R. Castleman. Second Edition. Academic Press, 2023, pp. 319–334. ISBN: 978-0-12-821049-9. DOI: `https://doi.org/10.1016/B978-0-12-821049-9.00011-3`. URL: `https://www.sciencedirect.com/science/article/pii/B9780128210499000113`.

[2] Wikimedia Commons. *File:Airy disk spacing near Rayleigh criterion.png — Wikimedia Commons, the free media repository.* [Online; accessed 21-April-2025]. 2023. URL: `https://commons.wikimedia.org/w/index.php?title=File:Airy_disk_spacing_near_Rayleigh_criterion.png&oldid=787147764`.

[3] B.D. Guenther. "DIFFRACTION — Fraunhofer Diffraction". In: *Encyclopedia of Modern Optics*. Ed. by Robert D. Guenther. Oxford: Elsevier, 2005, pp. 239–257. ISBN: 978-0-12-369395-2. DOI: `https://doi.org/10.1016/B0-12-369395-0/00695-3`. URL: `https://www.sciencedirect.com/science/article/pii/B0123693950006953`.

[4] Carlas S. Smith et al. "Fast, single-molecule localization that achieves theoretically minimum uncertainty". In: *Nature Methods* 7.5 (May 2010), pp. 373–375. ISSN: 1548-7105. DOI: `10.1038/nmeth.1449`. URL: `https://doi.org/10.1038/nmeth.1449`.

[5] Russell E Thompson, Daniel R Larson, and Watt W Webb. "Precise nanometer localization analysis for individual fluorescent probes". en. In: *Biophys J* 82.5 (May 2002), pp. 2775–2783.

[6] Eric Betzig et al. "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution". In: *Science* 313.5793 (2006), pp. 1642–1645. DOI: `10.1126/science.1127344`. eprint: `https://www.science.org/doi/pdf/10.1126/science.1127344`. URL: `https://www.science.org/doi/abs/10.1126/science.1127344`.

[7] Michael J Rust, Mark Bates, and Xiaowei Zhuang. "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)". en. In: *Nat Methods* 3.10 (Aug. 2006), pp. 793–795.

[8] Joerg Schnitzbauer et al. "Super-resolution microscopy with DNA-PAINT". In: *Nature Protocols* 12.6 (June 2017), pp. 1198–1228.

[9] David J Schodt et al. "SMITE: Single Molecule Imaging Toolbox Extraordinaire (MATLAB)". en. In: *J Open Source Softw* 8.90 (Oct. 2023).

[10] Colm McAlinden, Mark McCartney, and Jonathan Moore. "Mathematics of Zernike polynomials: a review". In: *Clinical & Experimental Ophthalmology* 39.8 (2011), pp. 820–827. DOI: `https://doi.org/10.1111/j.1442-9071.2011.02562.x`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1442-9071.2011.02562.x`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1442-9071.2011.02562.x`.

[11] Artur Speiser et al. "Deep learning enables fast and dense single-molecule localization with high accuracy". en. In: *Nat Methods* 18.9 (Sept. 2021), pp. 1082–1090.

[12] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *CoRR* abs/1505.04597 (2015). arXiv: `1505.04597`. URL: `http://arxiv.org/abs/1505.04597`.

[13] Cheng Bi et al. "SPTnet: a deep learning framework for end-to-end single-particle tracking and motion dynamics analysis". In: *bioRxiv* (2025). DOI: `10.1101/2025.02.04.636521`. eprint: `https://www.biorxiv.org/content/early/2025/02/08/2025.02.04.636521.full.pdf`. URL: `https://www.biorxiv.org/content/early/2025/02/08/2025.02.04.636521`.

[14] Serafim Korovin et al. "Using splines for point spread function calibration at non-uniform depths in localization microscopy". In: *bioRxiv* (2024). DOI: `10.1101/2024.01.24.577007`. eprint: `https://www.biorxiv.org/content/early/2024/03/28/2024.01.24.577007.full.pdf`. URL: `https://www.biorxiv.org/content/early/2024/03/28/2024.01.24.577007`.