

Improving 2-photon intravital microscopy imaging with Machine Learning

Introduction

Intravital multiphoton microscopy

Intravital multiphoton microscopy has become one of the central tools used in the investigation of dynamic cellular activity and function in living animals under nearly physiological conditions and is particularly important for studying the dynamic immune system. To be able to localize different immune system components to different compartments of the central nervous system (CNS), different CNS barriers have to be fluorescently labeled and imaged simultaneously with the immune cells or mediators. Due to the wide emission spectrum of the fluorophores the efficient identification of different fluorophores is limited to at the greatest number of read-outs photo-multiplier (PMT) channels (in our case 5 channels). Even then, careful setup of the imaging settings is required to allow for successful spectral channel unmixing in postprocessing [1][2][3].

Unsupervised clustering

Unsupervised learning is an approach allowing for search for patterns and structure in the unlabeled data.

K-Means clustering is an unsupervised clustering algorithm, that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype for the cluster.

Aim

The aim of this project is to develop the unmixing pipeline for multichannel image data. This will enable spectral channel unmixing for 2-photon intravital microscopy imaging to facilitate imaging configurations and push the limits of the detection of multiple fluorophores.

Methods and Materials

Dataset

The imaging data used in this project was acquired at the Theodor Kocher Institute using LaVision BioTec TriM Scope II 2-photon Intravital Microscope. This is a laser scanning microscope with 2 independent excitation laser lines. One laser has tunable wavelength in the range 780-1300nm, and the fixed line has wavelength of 1045nm. The microscope has 5 photomultiplier (PMT) detection channels (Figure 1). The

detection range of each channel is shown in Table 1.

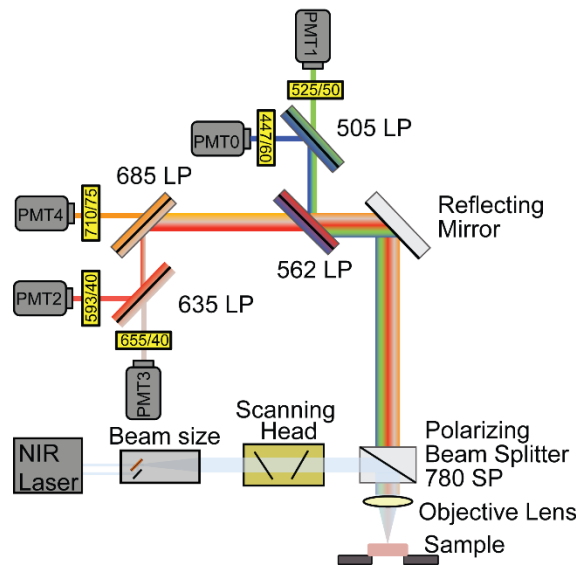


Figure 1. 2-photon microscope setup. Combination of dichroic mirrors and photomultipliers enables simultaneous detection of 5 wavelength bands.

Table 1. Detection bands of the 2-photon microscope.

Channel ID	Channel name	Min wavelength, nm	Max wavelength, nm
0	Blue	417	477
1	Green	505	550
2	Red	573	613
3	Far Red	635	675
4	Far Far Red	685	748

In the T-cell trafficking (TCT) dataset, the VE-Cad-GFP x AQP4-mRuby3 transgenic mouse model was used. It allows for visualization of the vascular adherent junctions [4] and meningeal fibroblasts due to the expression of the C-terminal GFP fusion protein of VE-cadherin in the endogenous VE-cadherin locus. The mRuby3 knock-in in the endogenous aquaporin-4 locus allows for visualization of glia limitans (manuscript in preparation). The imaging data contains 4 imaging channels. Detection of each of the fluorescent signal in the acquisition channels is shown in table 1.

The CD4⁺ T cells were labeled with the DeepRed fluorescent dye. Additionally dew to the second harmonic generation (SHG) in the Collagen-1 [5] the dura matter can be visualized in green and blue when dual excitation (920nm and 1045nm) is used. The imaging data contains 4 imaging channels. Detection of each of the fluorescent signal in the acquisition channels is shown in table 2.

In the cerebrospinal fluid clearance (CSFC) dataset the CLDN5-GFP x Prox1-tdTom transgenic mouse model was used. It allows for visualization of the blood-brain barrier vasculature expressing Clauding-5 tight junctions [6] and the lymphatic vessels expressing Prox-1. In this experiment a cisterna magna injection of the 2.5 um of 200umol of p40D680 fluorescent tracer was performed. Additionally, 2 distinct

autofluorescent signals are present. One is present in all channels except the FarFarRed, while the other is appearing over time in the time-lapse image sequence in Green, Red, and FarRed channels. Imaging was performed with the 860nm excitation. The imaging data contains 5 imaging channels. Detection of each of the fluorescent signal in the acquisition channels is shown in table 3.

Table 2: Visibility of the fluorescent signals of the T-cell trafficking dataset in 4 acquisition channels

	Blue	Green	Red	Far Red
SHG (Collagen 1)	+	+		
GFP (VE-Cadherin)		+	+	
mRuby (AQP4)		+	+	+
Deep Red (T cells)		+	+	+

Table 3: Visibility of the fluorescent signals of the cerebrospinal fluid clearance dataset in 5 acquisition channels

	Blue	Green	Red	Far Red	FarFarRed
SHG (Colagen 1)	+				
GFP (Claudin-5)		+	+		
tdTomato (Prox1)			+	+	
P40D840 (tracer)				+	+
FlashRed (PEG beads)				+	
Autofluorescence 1	+	+	+	+	
Autofluorescence 2		+	+	+	

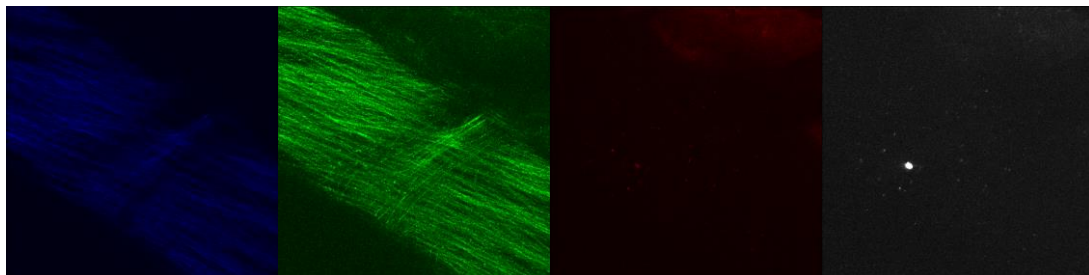


Figure 2. The four channels are Blue, Green, Red and FarRed respectively. They represent different central nervous

system components, SHG (Col1), VE-Cad & SHG, AQP4(Glia limitans), CD4 T cell.

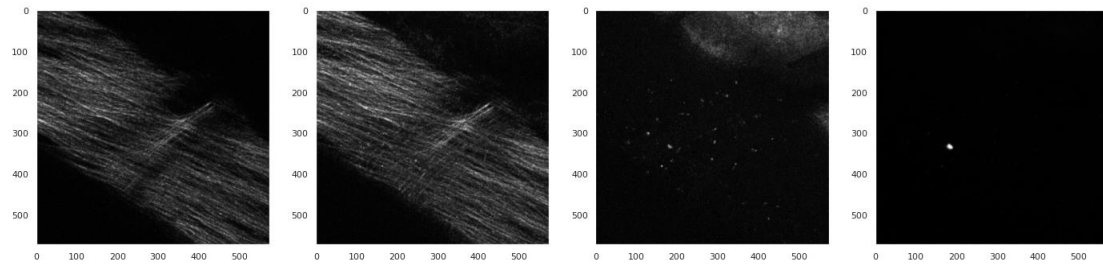


Figure 3. Image of four channels after Gaussian smoothing.

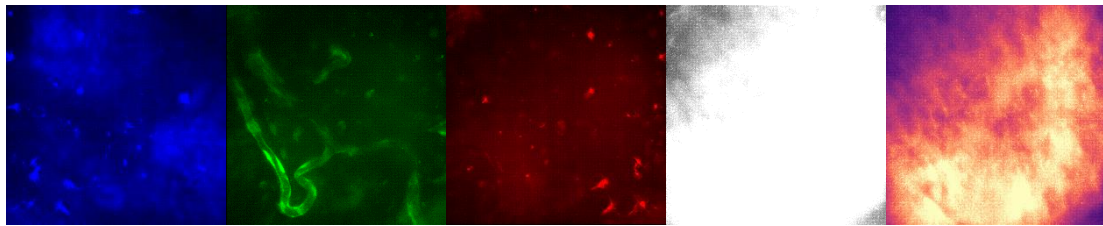


Figure 4. The five channels are Blue, Green, Red, FarRed and FarfarRed respectively. They represent different central nervous system components, SHG (Col1), CLDN-5 GFP, PEG beads, tracer P40D680 + PEG beads, tracer P40D680.

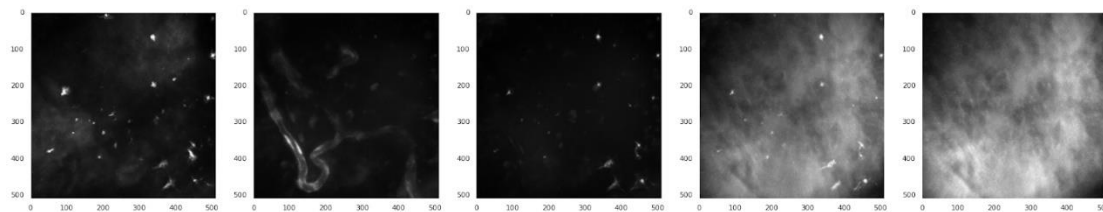


Figure 5. Image of four channels after Gaussian smoothing.

Channel mixing

In a 2-photon intravital microscope, an image is formed by using a laser to excite fluorescent molecules (such as GFP or tdTomato) in a sample, and then detecting the emitted fluorescence using a detector. The position of the emitted fluorescence is used to construct an image of the sample. Emitted fluorescence in bends is proportional total emitted fluorescence, and coefficients of proportionality are constant for a fluorophore and a particular optical system, these coefficients form the elements of a column in the mixing matrix, and that gives us the mixing matrix Q .

The emission spectrum of GFP is typically around 470-500 nm, while the emission spectrum of tdTomato is around 590-610 nm, shown in Figure 6 [7]. These spectra can vary depending on the specific variant of the fluorophore being used.

Detection bands for red and green channels are typically used to split the emission spectrum of the fluorescence. The ratio between signal detection in each of the channel detection bands is independent of the fluorescent signal brightness, with the exception of when the signal is brighter than the dynamic range of the detection

system, leading to saturation in the recorded pixel values.

The green/red pixel value is composed of the corresponding collected bands of the GFP & tdTom Spectrum + pixel noise ϵ . The signal intensity in each channel is proportional to the intensity of fluorescence in that channel's detection band, plus any noise present in the system.

Assuming pixel noise is small, unmixed pixel values will be the measured values without noise. The intensity of the signal in each channel will be proportional to the intensity of fluorescence in that channel's detection band.

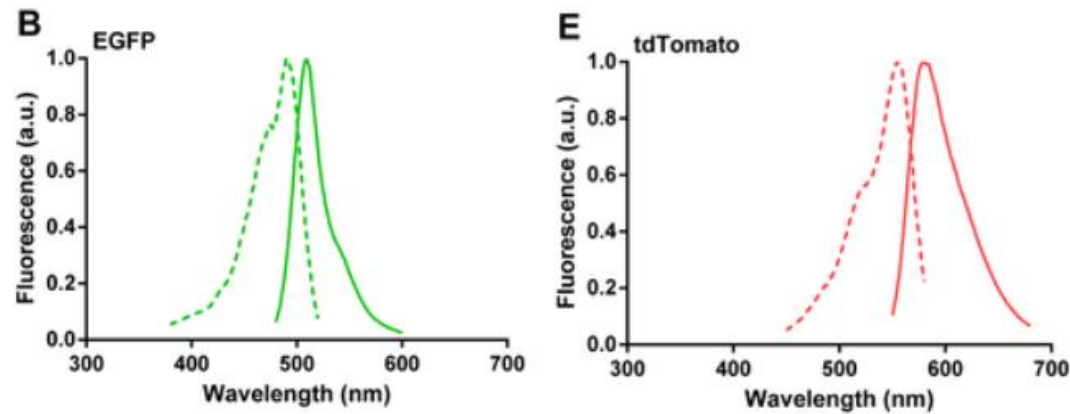


Figure 6. Spectra of fluorescent proteins fused to CNA35.

Excitation (dotted line) and emission (solid line) spectra of CNA35 fusion proteins with EGFP (B), tdTomato. Spectra of the fluorescent proteins were similar for both N- and C-terminal fusion to CNA35. Measurements were performed in 50 mM Tris (pH 8.0), 100 mM NaCl.

Preprocessing

In mixed image channels, the actual fluorescence signal is common across multiple channels, while pixel noise originating from photon counts is independent. As a first step to successfully isolate luciferin, the common part needs to be amplified and pixel noise suppressed.

Gaussian filtering is the process of weighted average of its own and other pixel values in the neighborhood. The width of the filter kernel was set to 1 pixel. Gaussian filtering is used to eliminate noise. After Gaussian smoothing (Figure 3), the data histogram of the green channel is shown, and the peak value appeared near the low pixel value.

By observing the histogram distribution of signal pixel values after Gaussian filtering, it can be seen that the overall signal distribution is positively skewed, the noise pixel value is lower than the pixel value of the actual fluorescence signal, and the pixel values with low values and high counts need to be removed. Therefore, a signal threshold is set to remove signal noise (Figure 7).

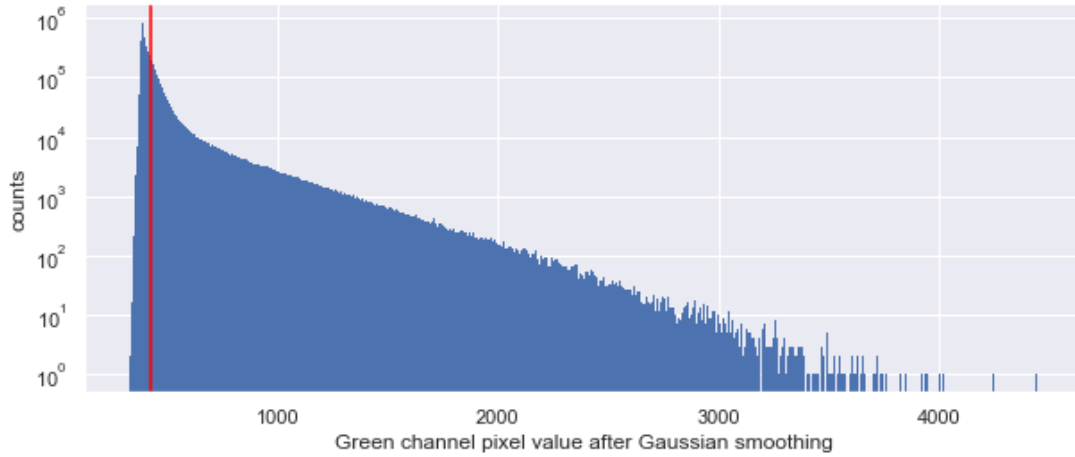


Figure 7. Distribution of the green channel pixel values after Gaussian filtering. The background pixels are the most abundant as seen by the position of the peak on the distribution., Assuming background distribution to be Gaussian the threshold can be defined as a point symmetric to the beginning of the distribution with respect to the distribution peak. The red line shows the position of the signal threshold. Pixel values that are less than the threshold is discarded as background.

After suppressing the pixel noise with Gaussian filtering and eliminating structures which are dim (pixel values are below the selected threshold), the data can be used for channel unmixing. Towards this end, a pixel dataset is prepared, where each sample \mathbf{p}_{ij} is a vector of length N_c representing brightness values of all N_c channels at the image coordinates i, j . Only the pixels where the brightness in at least one channel is higher than the threshold of the corresponding channel are included in the datasets. The sample value distribution (Figure 8) shows that most samples have low brightness values, and thus are concentrated near the origin. The samples with high brightness laying farther from the origin are sparse.

The first step to perform channel unmixing it is to find the directions of each of the fluorophores in the N_c -dimensional channel brightness space. These directions correspond to the ratios between detected signal in each of the channels, and thus are uniquely defined for each of the fluorophores for a given excitation wavelength and optical system.

The imbalance between brighter and dimmer samples, as well as between abundance of different fluorescent signals (compare e.g., samples of the SHG signal as seen on the present in Blue and Green channel against the Deep Red signal in FarRed, Red, and Green channels) can play detrimental role for the machine learning algorithm performance. For example, the traditional classification algorithms with the overall classification accuracy as the learning objective will likely take in to account the most abundant class, thus not being able to achieve sufficient performance on the classes where number of samples is little. This problem can become even more severe in the case of unsupervised learning, such as clustering performed in this work. Therefore, to achieve sufficient clustering performance the subsampling of the pixel dataset in was performed before clustering.

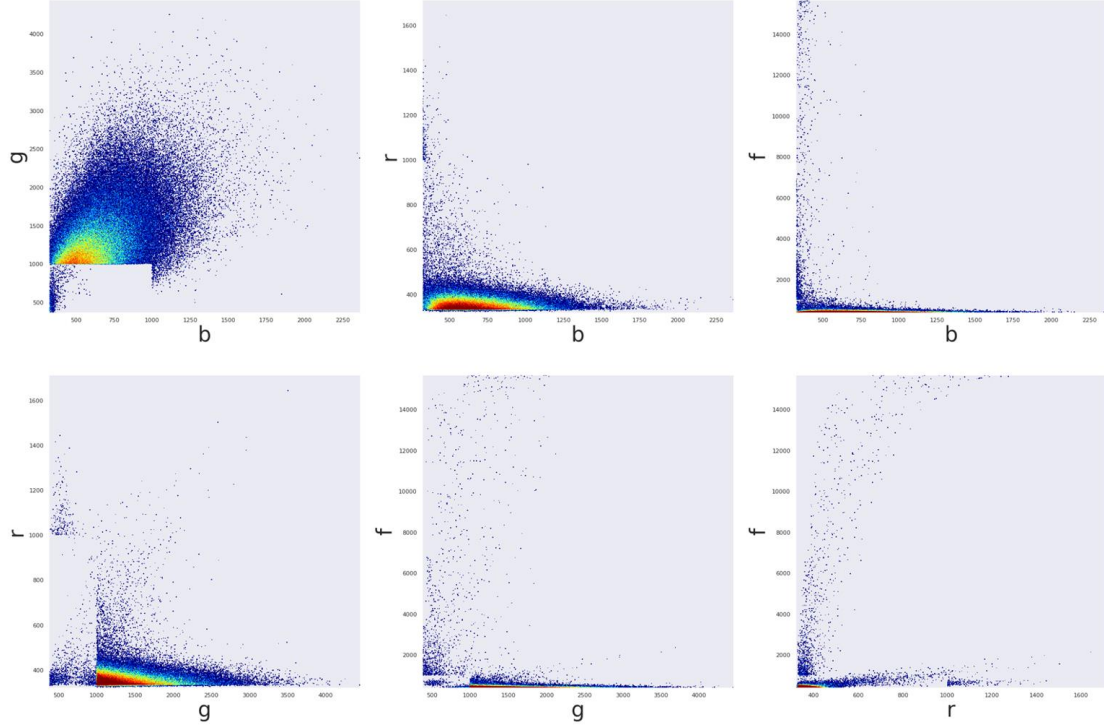


Figure 8. Distribution of the pixel brightness in the 4-dimensional pixel dataset for the TCT dataset. Pairwise distributions are shown for Blue (b), Green (g), Red (r), and FarRed (f) channels. The dots in the figure represent different sample densities with colors. The red area in the center has the highest density, followed by yellow and green, and the lowest density is blue, with the sample density gradually decreasing. Most of samples are concentrated near the origin, and the samples far from the origin are scarce.

Subsampling

The subsampling is performed such that the sample density in the subsampled datasets will be approximately equal in all areas, i.e., closer to the origin as well as farther from it. First the pixel dataset is partitioned into n^{N_c} N_c -dimensional bins. Then subsampling is performed, such that each bin will contain at most m samples after subsampling, where m is a hyperparameter, in the following Figure 8. If number of samples in a given bin is $\leq m$ – all samples are added to the subsampled dataset. Otherwise, the samples in this bin are clustered into m classes by a K-Means algorithm from the scikit-learn package [8] employing Euclidian distance. The coordinates of the obtained m cluster centroids (also vectors of dimension N_c) are then added as samples to the subsampled dataset.

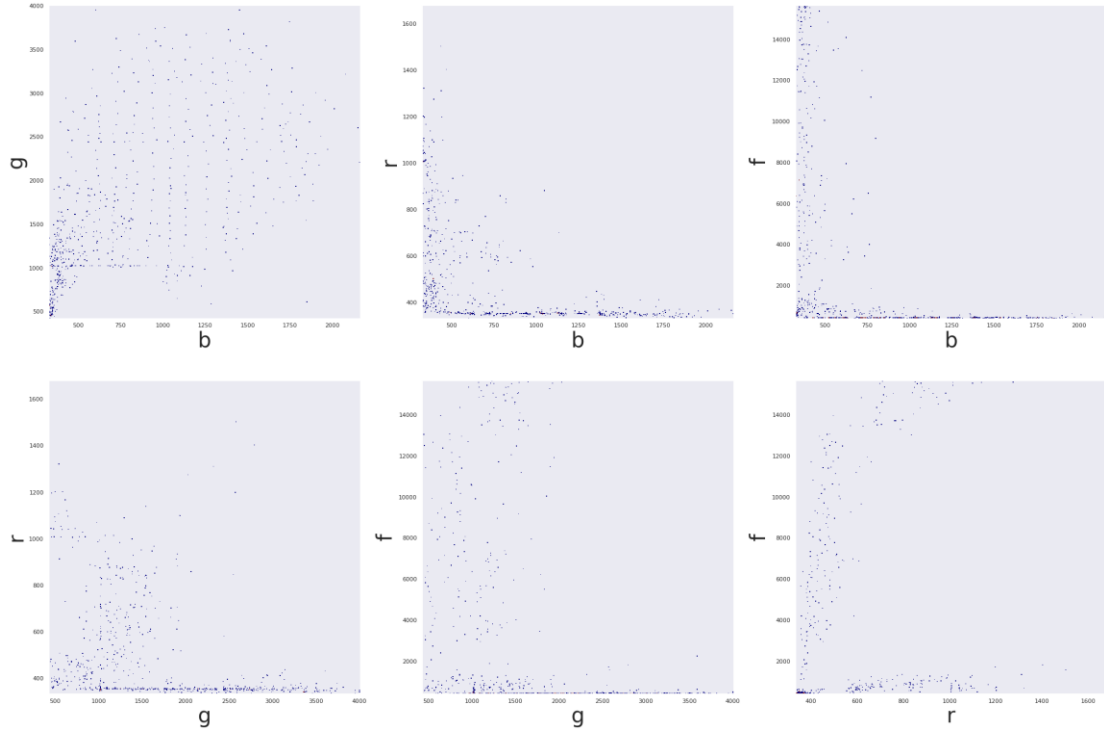


Figure 9. Distribution of the pixel brightness in the 4-dimensional subsampled pixel dataset for the TCT dataset. Samples are distributed more evenly, effectively avoiding points being too concentrated near the origin.

Here number of bins along each dimension $n=6$, number of channels $N_c = 4$ and cluster number in K-Means subsampling $m = 18$.

K-Means

As it can be seen from Figures 8, 9 the signal values of each different fluorophore tend to lay along corresponding direction. At the same time since the absolute brightness for each fluorophore can be arbitrary the points can lay at any distance, i.e., the distribution of the sample distance to the origin is not known. To cluster the samples extending the same direction and thus belonging to same fluorophores into corresponding classes, we employed K-Means clustering with cosine distance as the metric instead of the commonly used Euclidean distance.

Cosine distance is a measure of the difference between two samples using the cosine value of the angle between two vectors from the origin to the samples in a vector space. While employing Euclidian distance groups samples laying in proximity to each other, employing cosine distance groups data that lay in the same direction.

To determine the best K value in K-Means algorithm the elbow method, contour coefficient method, and Silhouette coefficient method were tested in this work.

Silhouette coefficient method can be used to determine the optimal cluster number. The Silhouette plot is shown in Figure 10. However, the K value with the highest score is not necessarily the correct clustering result. In this case since the prongs on which the samples lay do not cross the origin, the best number of clusters (2) does not correspond to the true number of fluorophores (4).

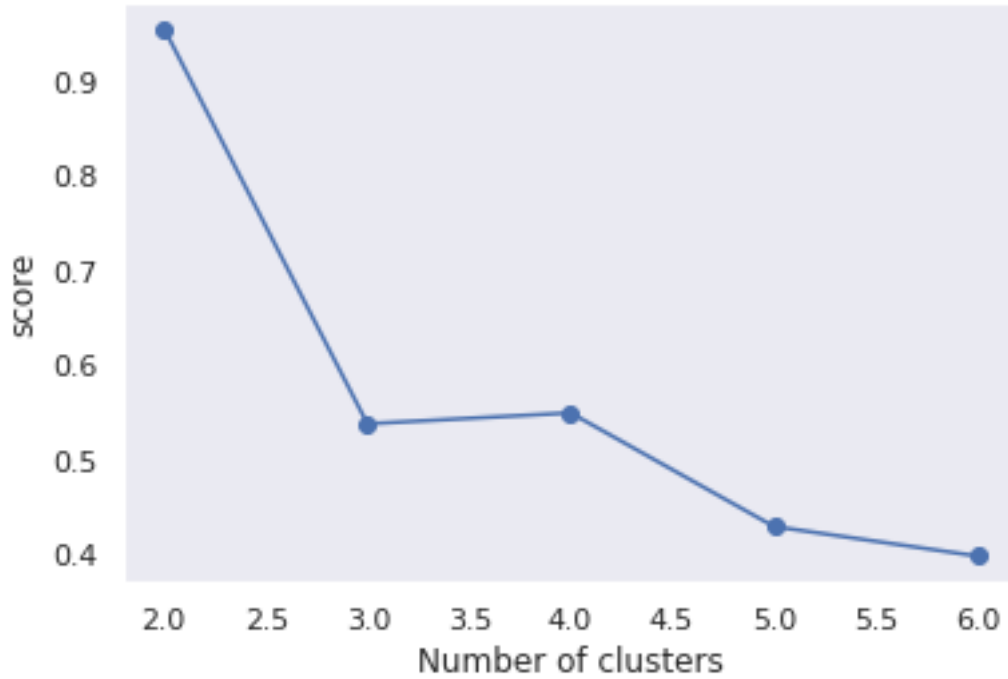


Figure 10. The Silhouette Coefficient combines Cohesion and Separation of clustering and is used to evaluate the effectiveness of clustering. The value ranges from -1 to 1, and the larger the value, the better the clustering effect. The low clustering score at the true cluster number (4) is explained by the position of the origin in the first clustering iteration.

After the first clustering, the samples are assigned to one of the clusters. Obtained cluster centroids (cluster directions) and the sample assignment to the clusters are shown in Figure 11 and Figure 12. To improve the clustering the origin has to be set at the intersection of the prongs corresponding to each fluorophore.

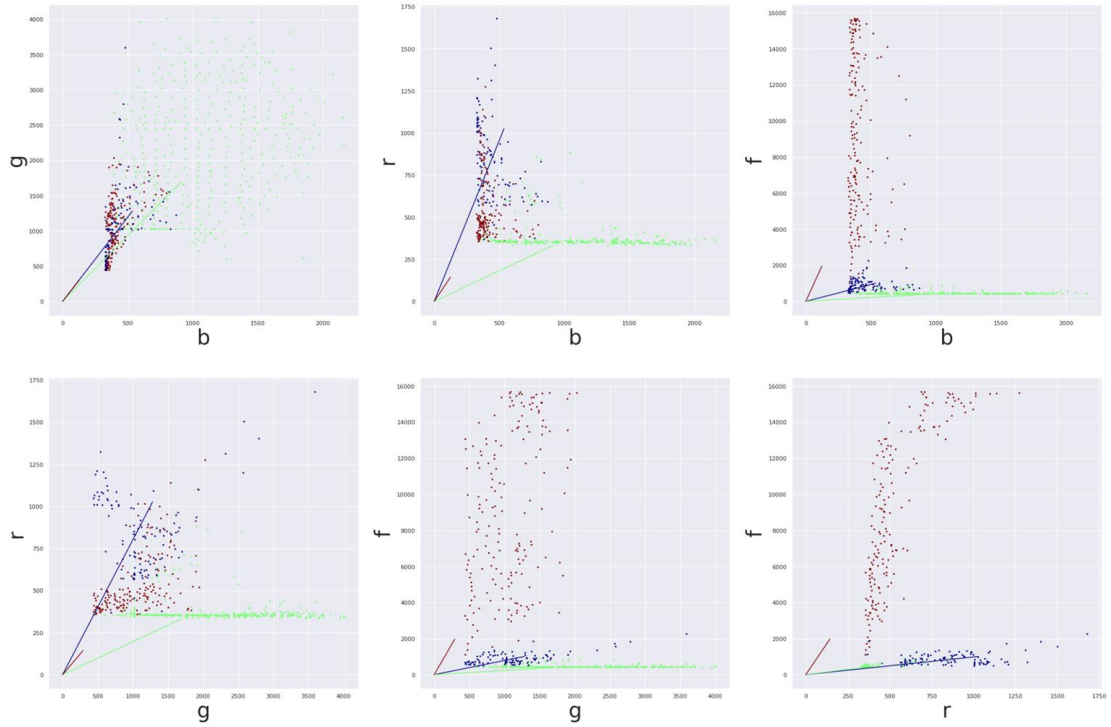


Figure 11. Sample assignment and cluster centroids after the first classification(clusters=3). Different colors represent different clusters, lines show the directions of the cluster's centroids obtained with K-means.

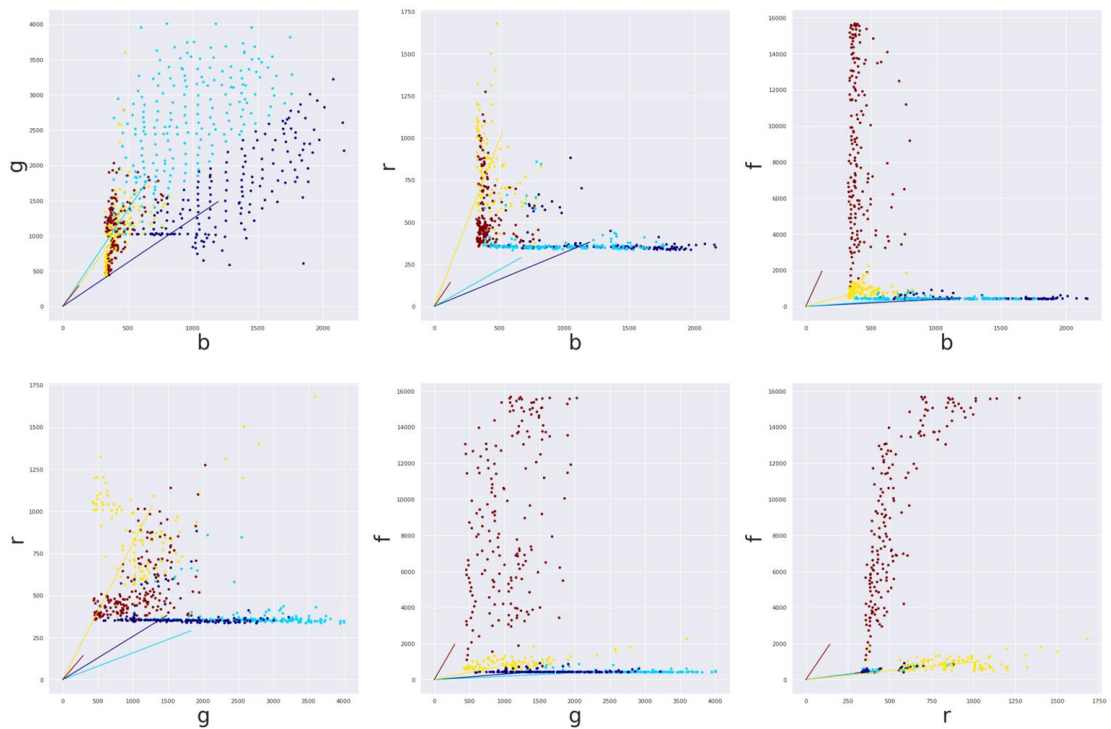


Figure 12. Sample assignment and cluster centroids after the first classification(clusters=4). Different colors represent different clusters, lines show the directions of the clusters centroids obtained with K-means.

PCA

To find the point of cluster intersection we first find the primary direction of each of the clusters. Principal Component Analysis (PCA) is often used to reduce the dimension of a data set while preserving the features in the data set that contribute the most to the variance. The method is performing linear feature decomposition of the covariance matrix [9], so as to obtain the principal orthogonal components of the datapoints (i.e., directions, eigenvectors) and their variances (i.e., eigenvalues [10]). The positions of each cluster's center are calculated as the mean value in the Euclidian space. The primary cluster direction is obtained as the first eigenvector with the PCA for the samples. The cluster axis are then the lines going through cluster center along the primary cluster directions.

Having the axes of each cluster we then search for the point o in the space that is the closest to all of the axes. This point will serve as the origin in the second iteration of K-means clustering.

To find this point we will employ numeric optimization using the Keras interface of the TensorFlow framework [11]. To this end we first define the distance between a point o and a line given by point c and vector of line direction v :

$$d(o, c, v) = \sqrt{\left((\bar{o} - \bar{c}) - \frac{(\bar{o} - \bar{c}) \cdot \bar{v}}{\bar{v}^2} \bar{v} \right)^2}$$

Then the sum of the distances from the point o to all cluster axis then is

$$L = \sum_i d(o, c_i, v_i)$$

where c_i and v_i are the centers and primary directions of i -th cluster. We then employed the Adam optimizer in the TensorFlow framework to perform gradient descent minimization of the loss function L with respect to the coordinates of the point o . Adam method is computationally efficient, widely used for training of neural networks, and is superior to other gradient optimization methods due to the use of first and second moments [12].

After 100 iterations, the sum of the distances tends to be converged (Figure 13), leading to the optimal position of point o .

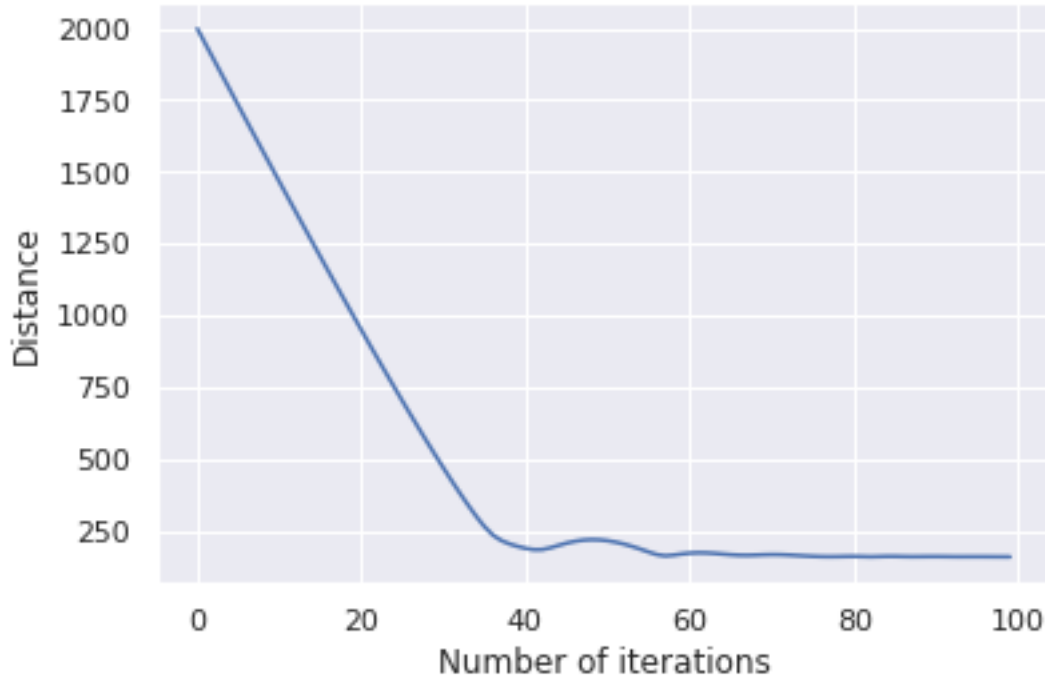


Figure 13. After one hundred iterations, the loss function corresponding to the sum of distances from point p to the cluster axes a_i tends to converge and reaches the minimum value.

K-Means

After the point of intersection of the progs corresponding to the distinct fluorophores is found, the K-means is repeated using the updated point as the new origin for the evaluation of cosine distance. This is implemented by updating the subsampled pixel dataset, subtracting the position of the found point o from each if the pixel value vectors component wise:

$$q_i = p_i - o$$

The pixel dataset q is then used as the input to the second iteration of K-means.

First, we have applied the silhouette coefficient method to find the optimal K value. Figure 14 shows that this score alone again fails to find the true fluorophore number (4). This means that true current method cannot reliably identify the true number of clusters, and it has to be specified by the researcher performing the unmixing for any given experimental dataset. The results of sample assignment after the second K-Means clustering with cluster number of 3 and 4 are shown in Figure 15 and Figure 16 correspondingly.

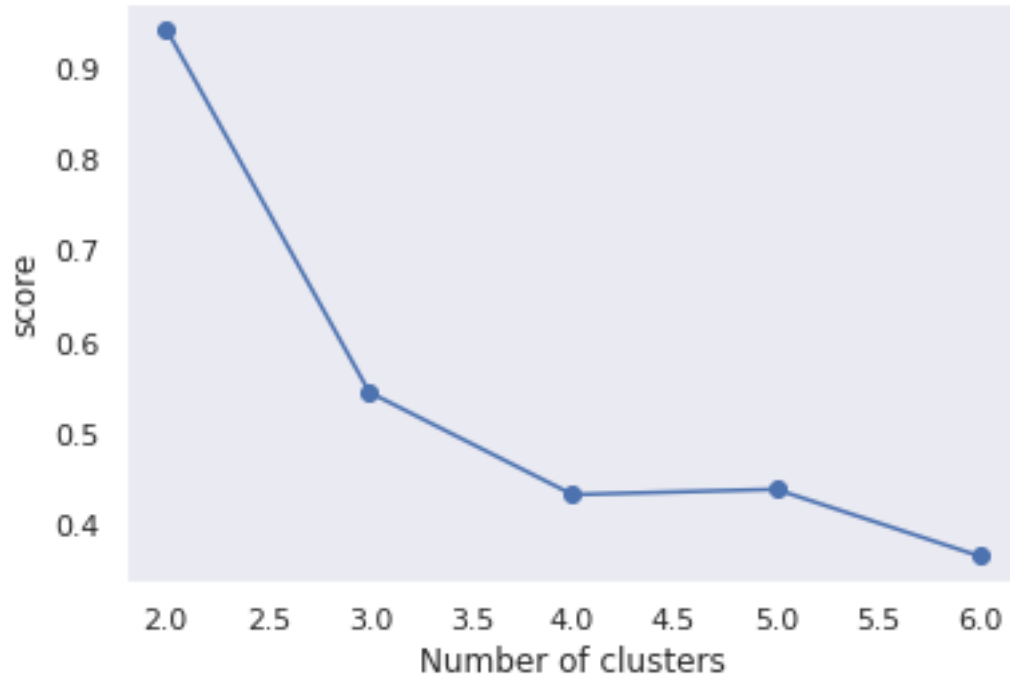


Figure 14. The optimum number of clusters determined by silhouette coefficient method is 2.

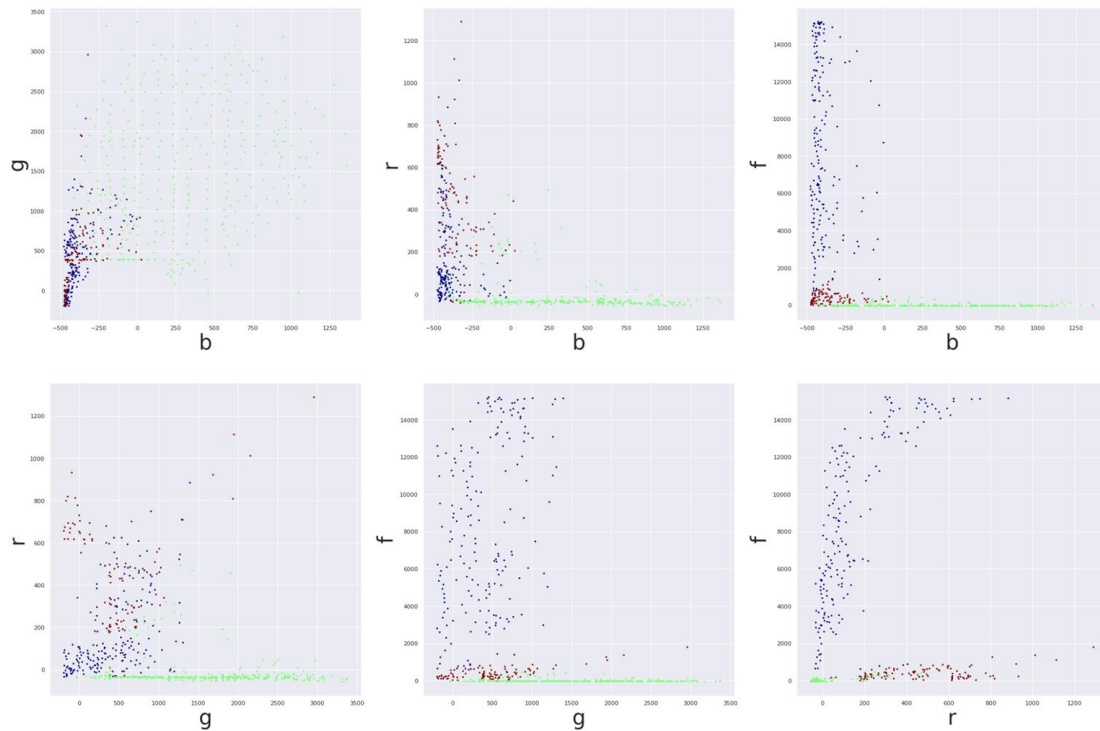


Figure 15. Scatter plot after the second classification (clusters=3). Different colors of the dots represent different clusters, compared with the first classification, the classification effect is significantly improved.

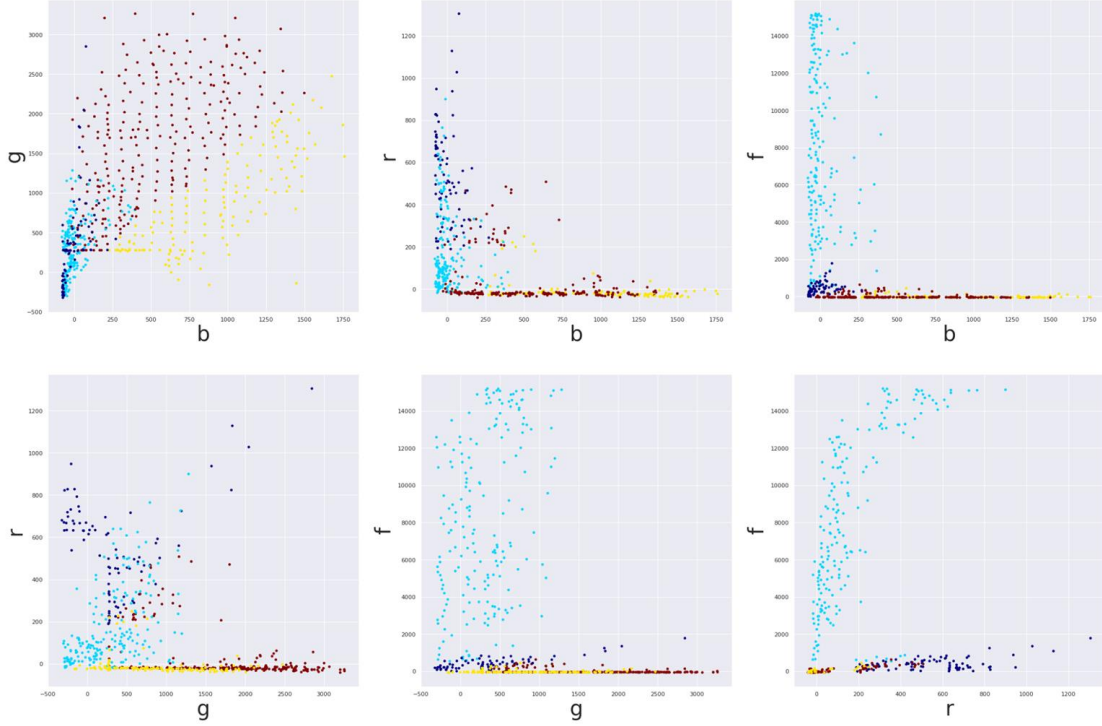


Figure 16. Scatter plot after the second classification($clusters=4$). Different colors of the dots represent different clusters, compared with the first classification, the classification effect is significantly improved.

Unmixing procedure

After the corrected directions of each cluster were found by the means of second K-Means clustering, the unmixing can be performed.

There are three possible situations to consider when unmixing images: number of fluorophores and correspondingly number of clusters is smaller, equal, or more than the number of image channels. In this study the last scenario is not considered as it does not allow for linear unmixing procedure.

When the number of newly unmixed channels is equal to the number of original channels, the centroids of each cluster after the second classification is orthogonalized to form new coordinate system, which will form the basis for unmixed pixel brightness space. Given the mixing matrix Q , each row of which corresponds to the unit vector v_i of the i -th cluster centroid, the unmixing matrix U can be found as the inverse of Q :

$$U = Q^{-1},$$

Then unmixing can be performed pixelwise:

$$u_{ij} = U(p_{ij} - o),$$

where p_{ij} is the pixel brightness vector at location i, j of length N_c after noise suppression, u_{ij} is the vector of unmixed pixel brightness of length n at the same location, and o is the position of the found cluster intersection.

When the number of newly unmixed images is less than the number of original channels, the matrix Q is not square and the inverse of it cannot be calculated. In this case, the left pseudoinverse can be evaluated as:

$$U = (Q^T * Q)^{-1} Q^T,$$

Where $(Q^T * Q)^{-1} Q^T$ Afterwards unmixing is performed exactly as in the previous case.

Results

Figure 17 indicating CD4⁺ cells in FarRed channel. CD4⁺ cells can be observed in the mixed Green channel and Red channel respectively in the original figure. After the unmixing, the CD4⁺ cells can be seen only on single channel (unmixed channel 4 in Figure 18), indicating successful unmixing. Similarly, the GFP⁺ blood vessel junctions visualized by the endogenous VE-Cadherin are unmixed and are present only in the unmixed channel 2. The SHG was not fully unmixed into a separate channel. This is caused by the fact that the SGH is formed independently in the blue channel by the 920 nm excitation, while the green signal is formed by the 1045 nm excitation. Due to the presence of chromatic aberrations in the excitation optical path, the two formed images are not perfectly aligned. This is one of the limitations of the current method.

In the case of three clusters (Figure 19), AQP4-mRuby3, Deep Red CD4⁺ cells, and the SHG+VE-Cad-GFP are successfully unmixed into 3 channels.

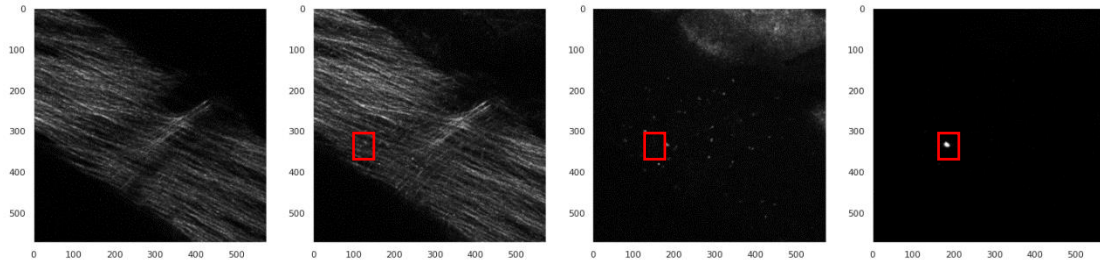


Figure 17. Raw image data for the TCT dataset. Each channel is shown as grayscale to simplify comparison.

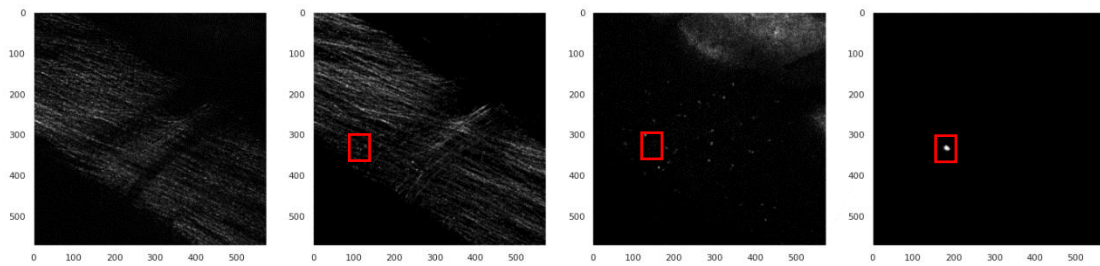


Figure 18. Unmixed Image with number of clusters $n=4$.

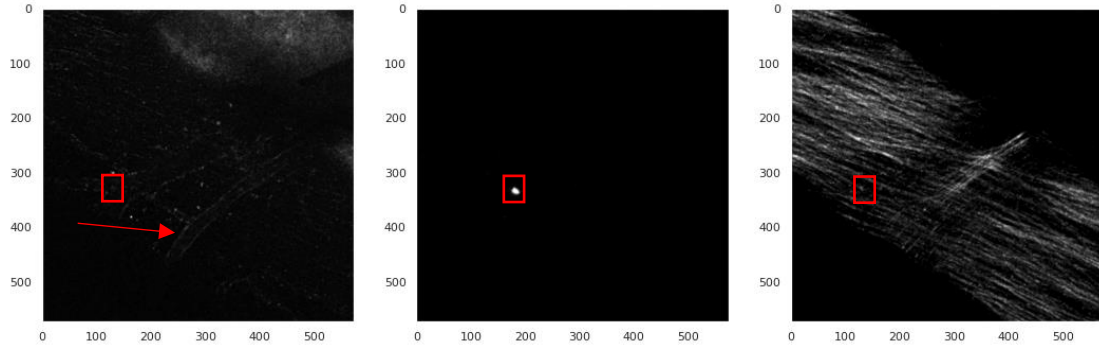


Figure 19. Unmixed Image with number of clusters $n=3$.

After unmixing, the background noise in Figure 21 is obviously eliminated. now the four unmixed channels correspond well to the P40D640, CLDN5-GFP, Prox1-tdTom + autofluorescence, and autofluorescence. The auto-fluorescent spots are saturated, limiting the possibility to properly unmix it.

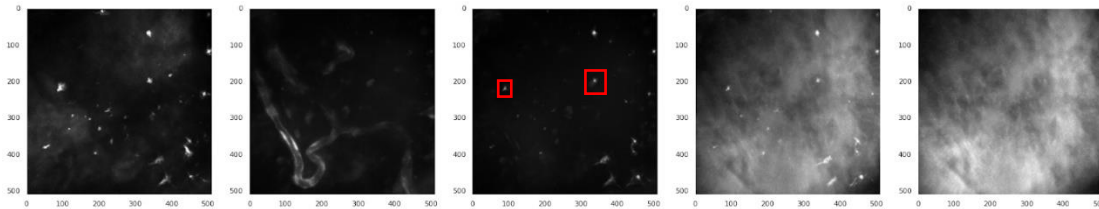


Figure 20. Mixed Image, which is labeled Figure 4

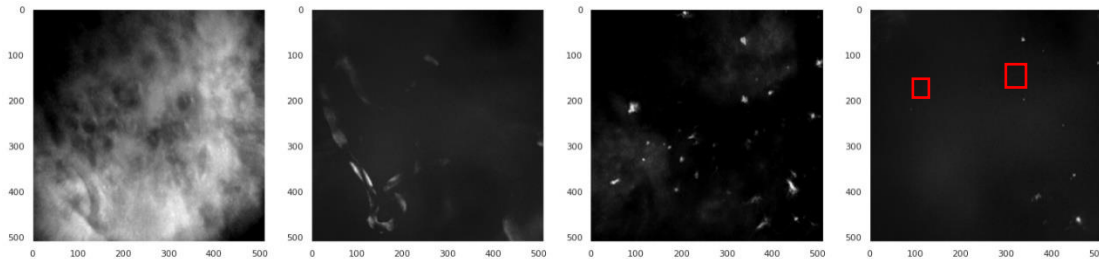


Figure 21. Unmixed Image while clusters=4.

Conclusion

In this research, the pipeline for unmixing multichannel 2-photon fluorescent images was developed. It includes image pre-processing, data subsampling used for unmixing matrix generation, two iterations of K-Means based on cosine distance into n clusters, calculation of the unmixing matrix based on the obtained mixing matrix, and finally the Orthogonalization knowledge unmixing of the image data to unmixed image data of n channels. The pipeline was tested on two datasets revealing meaningful unmixed images.

This pipeline has several limitations. Moreover, the unmixing quality has not been

evaluated quantitatively, and was only evaluated visually. Current method does not allow to unmix more fluorophores than available readout channels.

In the future work the pipeline will be optimized for user-friendly operation and extended to support more unmixed channels than imaging channels with deep learning-based approaches.

References

1. Vladymyrov, M, Abe, J, Moalli, F, Stein, JV, Ariga, A. 2016. "Real-Time Tissue Offset Correction System for Intravital Multiphoton Microscopy." *Journal of Immunological Methods* 438 (November):35–41.
2. Vladymyrov M, Haghayegh Jahromi N., Kaba, E, Engelhardt B. and Ariga A. "VivoFollow 2: Distortion-Free Multiphoton Intravital Imaging." *Front. Phys.* 2020.
3. Pizzagalli D, Bordini J, Morone D, Pulfer A, Carrillo-Barberà P, Thelen B, Ceni K, Thelen M, Krause R, Gonzalez SF, CANCOL, a Computer-Assisted Annotation Tool to Facilitate Colocalization and Tracking of Immune Cells in Intravital Microscopy, *The Journal of Immunology* 2022, 208 (6) 1493-1499
4. Winderlich, M.; Keller, L.; Cagna, G.; Broermann, A.; Kamenyeva, O.; Kiefer, F.; Deutsch, U.; Nottebaum, A.F.; Vestweber, D. Ve-ptp controls blood vessel development by balancing tie-2 activity. *J. Cell Biol.* **2009**.
5. Zipfel, W.R.; Williams, R.M.; Christie, R.; Nikitin, A.Y.; Hyman, B.T.; Webb, W.W. Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 7075–7080.
6. Vanlandewijck, M.; He, L.; Mae, M.A.; Andrae, J.; Ando, K.; Del Gaudio, F.; Nahar, K.; Lebouvier, T.; Lavina, B.; Gouveia, L.; et al. A molecular atlas of cell types and zonation in the brain vasculature. *Nature* **2018**.
7. <https://journals.plos.org/plosone/article/figures?id=10.1371/journal.pone.0114983>
8. <https://scikit-learn.org/stable/modules/clustering.html#clustering>
9. Abdi. H., & Williams, L.J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2010, 2: 433–459.
10. Shaw P.J.A. (2003) *Multivariate statistics for the Environmental Sciences*, Hodder-Arnold. ISBN 978-0-340-80763-7.
11. https://www.tensorflow.org/api_docs/python/tf/keras
12. Kingma, D.P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *CoRR*, abs/1412.6980.