CLAIMS

[Claim 1]. A computer-implemented method of end-to-end sequencing, including integrating neural network-based template generation with neural network-based base calling, comprising:

accessing first image data and second image data that contain pixels in an optical, pixel resolution,

wherein the first image data comprises images of clusters and their surrounding background captured by a sequencing system for initial ones of sequencing cycles of a sequencing run, and

wherein the second image data comprises images of the clusters and their surrounding background captured by the sequencing system for the initial and additional sequencing cycles of the sequencing run;

processing the first image data through a neural network-based template generator, and producing a cluster map that identifies cluster metadata,

wherein the cluster metadata identifies spatial distribution information of the clusters including cluster shapes and

wherein the neural network-based template generator is trained on a task of mapping the images of the clusters to the cluster metadata;

encoding the spatial distribution information of the clusters in a template image in an upsampled, subpixel resolution,

wherein subpixels of the template image and the pixels of the images of the clusters represent a same imaged area;

modifying intensity values of the pixels of the second image data based on the template image, and producing an intensity modified version of the second image data with an intensity distribution that accounts for the spatial distribution information of the clusters; and

processing the intensity modified version of the second image data through a neural network-based base caller, and producing base calls for one or more of the clusters at one or more sequencing cycles of the sequencing run,

wherein the neural network-based base caller is trained on a task of mapping the images of the clusters to the base calls.

[Claim 2]. A computer-implemented method of end-to-end sequencing, including integrating neural network-based template generation with neural network-based base calling, comprising:

accessing first image data and second image data that contain pixels in an optical, pixel resolution,

wherein the first image data comprises images of clusters and their surrounding background captured by a sequencing system for initial ones of sequencing cycles of a sequencing run, and

wherein the second image data comprises images of the clusters and their surrounding background captured by the sequencing system for the initial and additional sequencing cycles of the sequencing run;

processing the first image data through a neural network-based template generator, and producing a cluster map that identifies cluster metadata,

wherein the cluster metadata identifies spatial distribution information of the clusters based on cluster shapes, and

wherein the neural network-based template generator is trained on a task of mapping the images of the clusters to the cluster metadata;

encoding the spatial distribution information of the clusters in a template image in an upsampled, subpixel resolution,

wherein subpixels of the template image and the pixels of the images of the clusters represent a same imaged area;

supplementing the second image data with the template image; and

processing the second image data, supplemented with the template image, through the neural network-based base caller, and producing base calls for one or more of the clusters at one or more sequencing cycles of the sequencing run.

[Claim 3]. The computer-implemented method of claim 1, wherein each subpixel in the template image is identified as either background subpixel, cluster center subpixel, or cluster interior subpixel.

[Claim 4]. The computer-implemented method of any of claims 1 or 3, wherein modifying intensity values of the pixels of the second image data comprises:

calculating an area weighting factor for one or more pixels in the second image data based on how many subpixels in the template image that correspond to a pixel in the images of the second image data contain parts of one or more of the clusters; and

modifying intensities of the pixels based on the area weighting factor.

[Claim 5]. The computer-implemented method of any of claims 1, 3 or 4, wherein modifying intensity values of the pixels of the second image data comprises:

upsampling the images of clusters and their surrounding background to the upsampled, subpixel resolution to produce upsampled images, and assigning a background intensity to those subpixels in the upsampled images that correspond to background subpixels in the template image and assigning cluster intensities to those subpixels in the upsampled images that correspond to cluster center subpixels and cluster interior subpixels in the template image.

[Claim 6]. The computer-implemented method of claim 5, wherein the background intensity has a zero value.

[Claim 7]. The computer-implemented method of any of claims 1 or 3 to 6, wherein the cluster intensities are determined by interpolating intensities of the pixels in the optical, pixel resolution.

[Claim 8]. The computer-implemented method of any of claims 1 or 3 to 7, wherein modifying intensity values of the pixels of the second image data comprises:

upsampling the images of clusters and their surrounding background to the upsampled, subpixel resolution to produce upsampled images, and distributing an entire intensity of a pixel in the optical, pixel domain among only those constituent subpixels of the pixel in the upsampled images that correspond to the cluster center subpixels and the cluster interior subpixels in the template image.

[Claim 9]. A computer-implemented method, comprising:

using a first neural network to determine template image about clusters, wherein the template image identifies spatial distribution of the clusters, including a cluster shape; and

using a second neural network to base call the clusters based on the template image.

[Claim 10]. The computer-implemented method of claim 9, further comprising:

evaluating the template image in an upsampled subpixel domain for at least one particular cluster to identify a pixel that contains part of the at least one particular cluster and adjoining pixels to the pixel that also contain part of the at least one particular cluster;

calculating an area weighting factor for each pixel based on how many subpixels in each of the identified pixels contain parts of the at least one particular cluster; and

modifying a pixel intensity value of the identified pixel and the adjoining pixels for processing based on the area weighting factor for a respective pixel.

[Claim 11]. The computer-implemented method of claim 10, wherein evaluating the template image further comprises:

processing one or more initial image sets respectively generated at one or more initial sequencing cycles of a plurality of sequencing cycles through the first neural network to produce the template image to identify the centers, shapes, and boundaries of the clusters at the upsampled, subpixel resolution; wherein each image set comprises one or more images, each of the images depicting intensity emissions of the clusters and their surrounding background in a respective one of one or more imaging channels captured at the optical, pixel resolution.

[Claim 12]. The computer-implemented method of claim 10 or 11, wherein evaluating the template image further comprises:

evaluating the cluster shape and boundaries of the at least one particular cluster to identify at least one pixel that contains part of the at least one particular cluster and adjoining pixels to the pixel that also contain part of the at least one particular cluster; and wherein the method further comprises

storing the area weighting factor in the template image; and

generating a modified version of each of the images with pixels having modified pixel intensity values;

processing modified versions of the images through the second neural network to generate an alternative

representation of the modified versions; and

base calling the at least one particular cluster using the alternative representation.

[Claim 13]. The computer-implemented method of claim 12, wherein the base calling further comprises:

accessing one or more images at the optical, pixel resolution in each

of a current image set generated at a current one of the plurality of sequencing cycles,

of a one or more preceding image sets respectively generated at one or more of the plurality of sequencing cycles preceding the current one of the plurality of sequencing cycles, and

of a one or more succeeding image sets respectively generated at one or more of the plurality of sequencing cycles succeeding the current one of the plurality of sequencing cycles;

for pixels in each of the images, modifying a pixel intensity value based on the area weighting factor in the template image for a respective pixel;

generating a modified version of each of the images with pixels having modified pixel intensity values;

for the at least one particular cluster, extracting an image patch from each modified version such that each image patch

has an array of pixels, and

contains in its center pixel a center of the particular cluster identified in the template image;

convolving image patches extracted from modified versions of the images through a convolutional neural network of the second neural network to generate a convolved representation of the image patches;

processing the convolved representation through an output layer to produce, for the center pixel, likelihoods of a base incorporated in the at least one particular cluster at the current one of the plurality of sequencing cycles being A, C, T, and G; and

classifying the base as A, C, T, or G based on the likelihoods.

[Claim 14]. The computer-implemented method any of claims 12 or 13, further comprising:

prior to modifying the pixel intensity values, aligning each of the images captured at the optical, pixel resolution with the template image using cycle-specific and imaging channel-specific transformations.

[Claim 15]. The computer-implemented method of claim 9, further comprising:

evaluating the template image in an upsampled subpixel domain to identify subpixels that contain parts of any cluster; and

assigning a background intensity to subpixels identified in the template image as not contributing to any cluster.

[Claim 16]. The computer-implemented method of claim 15, wherein evaluating the template image in an upsampled subpixel domain further comprises:

calculating how many subpixels in at least one pixel contain parts of any cluster and calculating a per-subpixel area weighting factor for the subpixels in the at least one pixel.

[Claim 17]. The computer-implemented method of claim 15 or 16, wherein the method comprises:

processing one or more initial image sets respectively generated at one or more initial sequencing cycles of a plurality of sequencing cycles through the first neural network to produce the template image at the upsampled, subpixel resolution, wherein each image set comprises one or more images, each of the images depicting intensity emissions of the clusters and their surrounding background in a respective one of one or

more imaging channels captured at the optical, pixel resolution and wherein the template image classifies subpixels into classes including cluster center, background, and cluster interior;

upsampling each of the images captured at the optical, pixel resolution into a subpixel domain and assigning a background intensity to subpixels of each of the images identified in the template image as not contributing to any cluster;

processing the upsampled images through the second neural network to generate an alternative representation of the upsampled images; and

base calling a plurality of the clusters using the alternative representation.

[Claim 18]. The computer-implemented method of claim 17, wherein upsampling each of the images further comprises:

distributing intensity of a particular pixel among first subpixels of the particular pixel identified in the template image as contributing to any cluster by applying the per-subpixel area weighting factor and assigning a background intensity to second subpixels of the particular pixel identified in the template as not contributing to any cluster.

[Claim 19]. The computer-implemented method of claim 18, wherein the prior to upsampling the method comprises:

accessing one or more images at the optical, pixel resolution in each

of a current image set generated at a current one of the plurality of sequencing cycles,

of a one or more preceding image sets respectively generated at one or more of the plurality of sequencing cycles preceding the current one of the plurality of sequencing cycles, and

of a one or more succeeding image sets respectively generated at one or more of the plurality of sequencing cycles succeeding the current one of the plurality of sequencing cycles; and after upsampling the method comprises:

extracting an image patch from each upsampled image such that each image patch has an array of subpixels;

convolving image patches extracted from the upsampled images through the convolutional neural network of the second neural network to generate a convolved representation of the image patches;

processing the convolved representation through an output layer to produce, for each subpixel in the array, likelihoods of a base incorporated at the current one of the plurality of sequencing cycles being A, C, T, and G;

classifying the base as A, C, T, or G based on the likelihoods; and

base calling each one of the plurality of the clusters based on a base classification assigned to a respective subpixel containing a center of a corresponding cluster.

[Claim 20]. The computer-implemented method of claim 18 or 19, further comprising:

prior to the upsampling, aligning each of the images captured at the optical, pixel resolution with the template image using cycle-specific and imaging channel-specific transformations.