

CLAIMS

1. An artificial intelligence-based system for base calling, the system comprising:
 - a neural network-based base caller that processes at least right flanking, center, and left flanking inputs, and produces at least right flanking, center, and left flanking outputs;
 - wherein the right flanking input comprises current image data for a current sequencing cycle of a sequencing run, supplemented with previous image data for one or more previous sequencing cycles preceding the current sequencing cycle, and wherein the right flanking output comprises a right flanking base call prediction for the current sequencing cycle and base call predictions for the previous sequencing cycles;
 - wherein the center input comprises the current image data, supplemented with the previous image data and successive image data for one or more successive sequencing cycles succeeding the current sequencing cycle, and wherein the center output comprises a center base call prediction for the current sequencing cycle and base call predictions for the previous sequencing cycles and the successive sequencing cycles;
 - wherein the left flanking input comprises the current image data, supplemented with the successive image data, and wherein the left flanking output comprises a left flanking base call prediction for the current sequencing cycle and base call predictions for the successive sequencing cycles; and
 - a base call generator coupled to the neural network-based base caller and configured to generate a base call for the current sequencing cycle based on the right flanking, center, and left flanking base call predictions for the current sequencing cycle.
2. The artificial intelligence-based system of claim 1, wherein the current image data for the current sequencing cycle depicts intensity emissions of analytes and their surrounding background captured at the current sequencing cycle.
3. The artificial intelligence-based system of claim 2, wherein the right flanking, center, and left flanking base call predictions for the current sequencing cycle identify likelihoods of a base incorporated in one or more of the analytes at the current sequencing cycle being A, C, T, and G.
4. The artificial intelligence-based system of claim 3, wherein the base call generator is further configured to comprise an averager that
 - base-wise sums the likelihoods across the right flanking, center, and left flanking base call predictions for the current sequencing cycle;
 - determines base-wise averages based on the base-wise sums; and

generates the base call for the current sequencing cycle based on a highest one of the base-wise averages.

5. The artificial intelligence-based system of claim 3, wherein the base call generator is further configured to comprise a consensor that

determines a preliminary base call for each of the right flanking, center, and left flanking base call predictions for the current sequencing cycle based on a highest one of the likelihoods, thereby producing a sequence of preliminary base calls; and

generates the base call for the current sequencing cycle based on a most common base call in the sequence of preliminary base calls.

6. The artificial intelligence-based system of claim 3, wherein the base call generator is further configured to comprise a weighted consensor that

determines a preliminary base call for each of the right flanking, center, and left flanking base call predictions for the current sequencing cycle based on a highest one of the likelihoods, thereby producing a sequence of preliminary base calls;

applies base-wise weights to respective ones of the preliminary base calls in the sequence of preliminary base calls and produces a sequence of weighted preliminary base calls; and

generates the base call for the current sequencing cycle based on a most weighted base call in the sequence of weighted preliminary base calls.

7. The artificial intelligence-based system of claim 3, wherein the likelihoods are exponentially normalized scores produced by a softmax layer.

8. The artificial intelligence-based system of claim 1, further configured to comprise a trainer that, during training,

computes errors between base calls generated by the base call generator for the current sequencing cycle, the previous sequencing cycles, and the successive sequencing cycles based on the right flanking, center, and left flanking outputs of the neural network-based base caller and ground truth base calls;

determines gradients for the current sequencing cycle, the previous sequencing cycles, and the successive sequencing cycles based on the errors; and

updates parameters of the neural network-based base caller by backpropagating the gradients.

9. The artificial intelligence-based system of claim 1, wherein the right flanking base call prediction for the current sequencing cycle accounts for prephasing effect between the current sequencing cycle and the previous sequencing cycles.
10. The artificial intelligence-based system of claim 9, wherein the center base call prediction for the current sequencing cycle accounts for the prephasing effect between the current sequencing cycle and the previous sequencing cycles and phasing effect between the current sequencing cycle and the successive sequencing cycles.
11. The artificial intelligence-based system of claim 10, wherein the left flanking base call prediction for the current sequencing cycle accounts for the phasing effect between the current sequencing cycle and the successive sequencing cycles.
12. An artificial intelligence-based system for base calling, the system comprising:
a host processor;
memory accessible by the host processor storing image data for sequencing cycles of a sequencing run, wherein current image data for a current sequencing cycle of the sequencing run depicts intensity emissions of analytes and their surrounding background captured at the current sequencing cycle; and
a configurable processor having access to the memory, the configurable processor including:
a plurality of execution clusters, the execution clusters in the plurality of execution clusters configured to execute a neural network; and
data flow logic having access to the memory and to the execution clusters in the plurality of execution clusters, configured to provide the current image data, previous image data for one or more previous sequencing cycles preceding the current sequencing cycle, and successive image data for one or more successive sequencing cycles succeeding the current sequencing cycle to available execution clusters in the plurality of execution clusters, cause the execution clusters to apply different groupings of the current image data, the previous image data, and the successive image data to the neural network to produce first, second, and third base call predictions for the current sequencing cycle, and to feedback the first, second, and third base call predictions for the current sequencing cycle to the memory for use in generation of a base call for the current sequencing cycle based on the first, second, and third base call predictions.
13. The artificial intelligence-based system of claim 12, wherein the different groupings include a first grouping comprising the current image data and the previous image data, a second

grouping comprising the current image data, the previous image data, and the successive image data, and a third grouping comprising the current image data and the successive image data.

14. The artificial intelligence-based system of claim 13, wherein the execution clusters apply the first grouping to the neural network to produce the first base call prediction, the second grouping to the neural network to produce the second base call prediction, and the third grouping to the neural network to produce the third base call prediction.

15. The artificial intelligence-based system of claim 12, wherein the first, second, and third base call predictions for the current sequencing cycle identify likelihoods of a base incorporated in one or more of the analytes at the current sequencing cycle being A, C, T, and G.

16. The artificial intelligence-based system of claim 15, wherein the data flow logic is further configured to generate the base call for the current sequencing cycle by

- base-wise summing the likelihoods across the first, second, and third base call predictions for the current sequencing cycle;
- determining base-wise averages based on the base-wise summing; and
- generating the base call for the current sequencing cycle based on a highest one of the base-wise averages.

17. An artificial intelligence-based method of base calling, the method including:

- accessing current image data for a current sequencing cycle of a sequencing run, previous image data for one or more previous sequencing cycles preceding the current sequencing cycle, and successive image data for one or more successive sequencing cycles succeeding the current sequencing cycle;

- processing different groupings of the current image data, the previous image data, and the successive image data through a neural network-based base caller and producing a plurality of call predictions for the current sequencing cycle; and

- generating a base call for the current sequencing cycle based on the plurality of base call predictions.

18. An artificial intelligence-based method of base calling, the method including:

- generating respective base calls for a particular sequencing cycle in response to executing respective iterations of a base caller;

- wherein the respective iterations process respective input sets for respective windows of sequences cycles; and

wherein the respective windows of sequencing cycles have the particular sequencing cycle as at least one overlapping cycle, and one or more non-overlapping cycles.

19. The artificial intelligence-based system of claim 18, wherein the particular sequencing cycle appears at different positions in the respective windows of sequences cycles.

20. The artificial intelligence-based system of claim 18, wherein the particular sequencing cycle appears at a same position between two windows of sequences cycles.