

algorithmic technique to determine that the relevant sample has amplified, involves determining whether the associated amplification curve has crossed a threshold value that is either fixed or is calculated based on the characteristics of the amplification curve. If the threshold is crossed, the curve is determined to be amplified; if the threshold is not crossed, the curve is determined to not be amplified.

SUMMARY

[0003] Automated determination of amplification is important to for increasing throughput of sample analysis, which in turn can both advance scientific research and improve provision of time-sensitive, clinically important information. Existing methods of automatically determining amplification have relied on combinations of techniques and parameters to improve accuracy. However, better accuracy is needed. Embodiments of the invention implement deep learning methods using artificial neural networks to provide improved automated predictions regarding whether a particular sample comprises a target molecule.

[0003a] An aspect of the invention provides a computer program product comprising executable code stored in a non-transitory computer readable medium. The executable code are executable on one or more computer processors to perform a method comprising using an artificial neural network to process amplification data. The amplification data comprises an amplification curve generated based on output from a quantitative polymerase chain reaction (qPCR) instrument, the output comprises fluorescence values measured after each cycle of a plurality of thermal cycles of a qPCR assay for a target nucleic acid in a

biological sample conducted using the qPCR instrument, determine, from the processing of the amplification data, an amplification classification of the biological sample, the amplification classification chosen from 1) amplified, indicating presence of the target nucleic acid in the biological sample, 2) not amplified, indicating the target nucleic acid not present in the biological sample, or 3) inconclusive, indicating inconclusive of presence or absence of the target nucleic acid in the biological sample; and output the determined amplification classification, wherein the artificial neural network is trained with training data comprising training amplification curves generated from fluorescence values measured after each cycle of a predetermined number of thermal cycles obtained from conducting a qPCR assay on training samples, wherein the artificial neural network comprises: an input layer comprising a dimension having a size corresponding to the predetermined number of thermal cycles, the input layer configured to receive input data comprising the amplification curve generated from the output of the qPCR assay conducted on the biological sample using the qPCR instrument; a plurality of hidden layers comprising one or more convolutional layers; and an output layer configured to output the determined amplification classification.

[0003b] Another aspect of the present invention provides a method of processing output of a quantitative polymerase chain reaction (qPCR) instrument, the qPCR instrument being configured to generate amplification data from one or more qPCR assays conducted using the qPCR instrument, the method comprising using an artificial neural network, trained with training amplification data comprising fluorescence values measured after each cycle of a plurality of thermal cycles obtained from conducting qPCR on training samples,

to determine an amplification classification corresponding to a particular biological sample subject to a qPCR assay using the qPCR instrument, the amplification classification chosen from 1) amplified, indicating presence of the target nucleic acid in the biological sample, 2) not amplified, indicating the target nucleic acid not present in the biological sample, or 3) inconclusive, indicating inconclusive of presence or absence of the target nucleic acid in the biological sample; output the determined amplification classification, wherein the artificial neural network comprises an input layer configured to receive input data obtained from the amplification data comprising fluorescence values measured after each cycle of the plurality of thermal cycles of the qPCR assay conducted on the particular biological sample using the qPCR instrument, a plurality of hidden layers comprising one or more convolutional layers, and an output layer configured to output the determined amplification classification. Another aspect of the invention provides a quantitative polymerase chain reaction (qPCR) system comprising a qPCR instrument configured to process biological samples and generate amplification data corresponding the samples; and one or more computers configured to receive the amplification data and to receive, store, and execute the executable instruction code of the computer program product as described.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] FIG. 1 illustrates a system in accordance with an embodiment of the present invention.

CLAIMS

What is claimed is:

1. A computer program product comprising executable code stored in a non-transitory computer readable medium, the executable code being executable on one or more computer processors to perform a method comprising:

using an artificial neural network to:

process amplification data comprising an amplification curve generated based on output from a quantitative polymerase chain reaction (qPCR) instrument, the output comprising fluorescence values measured after each cycle of a plurality of thermal cycles of a qPCR assay for a target nucleic acid in a biological sample conducted using the qPCR instrument,

determine, from the processing of the amplification data, an amplification classification of the biological sample, the amplification classification chosen from 1) amplified, indicating presence of the target nucleic acid in the biological sample, 2) not amplified, indicating the target nucleic acid not present in the biological sample, or 3) inconclusive, indicating inconclusive of presence or absence of the target nucleic acid in the biological sample; and

output the determined amplification classification, wherein the artificial neural network is trained with training data comprising training amplification curves generated from fluorescence values measured after each cycle of a predetermined number of thermal cycles obtained from conducting a qPCR assay on training samples, wherein the artificial neural network comprises:

an input layer comprising a dimension having a size corresponding to the predetermined number of thermal cycles, the input layer configured to receive input data comprising the amplification curve generated from the output of the qPCR assay conducted on the biological sample using the qPCR instrument;

a plurality of hidden layers comprising one or more convolutional layers;
and

an output layer configured to output the determined amplification classification.

2. The computer program product of claim 1 wherein the input data comprises, at each of a plurality of cycles, normalized fluorescence values and one or more derivative values of a curve fit to the amplification data.
3. The computer program product of claim 2 wherein the curve fit to the amplification data is a curve fit to the normalized fluorescence values.
4. The computer program product of claim 2 wherein the one or more derivative values comprises at least three derivative values.
5. The computer program product of claim 2 wherein the one or more derivative values comprises at least four derivative values.
6. The computer program product of any one of claims 1-5 wherein the plurality of hidden layers comprises one or more separable convolutional layers.
7. The computer program product of any one of claims 1-5 wherein the plurality of hidden layers comprises one or more separable convolutional layers and one or more normal convolutional layers.

8. The computer program product of claim 1 wherein the plurality of hidden layers further comprises a plurality of fully connected layers.
9. The computer program product of any one of claims 1-7 wherein the plurality of hidden layers comprises pooling layers but do not comprise any fully connected layers.
10. The computer program product of any one of claims 1-9 wherein the determined amplification classification is expressed as a probability.
11. A method of processing output of a quantitative polymerase chain reaction (qPCR) instrument, the qPCR instrument being configured to generate amplification data from one or more qPCR assays conducted using the qPCR instrument, the method comprising:
 - using an artificial neural network, trained with training amplification data comprising fluorescence values measured after each cycle of a plurality of thermal cycles obtained from conducting qPCR on training samples, to:
 - determine an amplification classification corresponding to a particular biological sample subject to a qPCR assay using the qPCR instrument, the amplification classification chosen from 1) amplified, indicating presence of the target nucleic acid in the biological sample, 2) not amplified, indicating the target nucleic acid not present in the biological sample, or 3) inconclusive, indicating inconclusive of presence or absence of the target nucleic acid in the biological sample;
 - output the determined amplification classification,

wherein the artificial neural network comprises:

an input layer configured to receive input data obtained from the amplification data comprising fluorescence values measured after each cycle of the plurality of thermal cycles of the qPCR assay conducted on the particular biological sample using the qPCR instrument,

a plurality of hidden layers comprising one or more convolutional layers,
and

an output layer configured to output the determined amplification classification.

12. The method of claim 11 wherein the artificial neural network is configured to use a plurality of respective class-weighted error determinations corresponding to the respective amplification classifications, and wherein the respective amplification classifications that is less represented in the training data is more heavily weighted than another of the respective amplification classifications that is more represented in the training data.

13. The method of any one of claims 11-12 wherein the determined amplification classification is expressed as a probability.

14. The method of any one of claims 11-13 wherein a cross-entropy loss function is used in combination with a class weight to obtain the class-weighted error determination.

15. The method of claim 11, further comprising pre-processing the amplification data generated from the qPCR assay to obtain pre-processed data

to provide as input data to the input layer of the artificial neural network, the pre-processing comprising:

determining if the amplification data corresponds to a number of cycles of the qPCR instrument that is different than a size of a corresponding dimension of the input layer of the artificial neural network; and

if the number of cycles is different than the size, then fitting the amplification data to the size.

16. The method of claim 15 wherein fitting the pre-processed data to the size comprises truncating the amplification data such that truncated data fits the size.

17. The method of claim 15 wherein fitting the pre-processed data to the size comprises interpolating the amplification data such that interpolated data fits the size.

18. The method of claim 16 wherein the pre-processed data further comprises one or more derivative values for each cycle represented in the truncated data, wherein the one or more derivative values are determined based on a curve fit to non-truncated data.

19. The method of claim 15, wherein the pre-processed data further comprises one or more derivative values for each cycle represented in the interpolated data, wherein the one or more derivative values are determined based on a curve fit to non-interpolated data.

20. A quantitative polymerase chain reaction (qPCR) system comprising:

a qPCR instrument configured to process biological samples and generate amplification data corresponding the samples; and

one or more computers configured to receive the amplification data and comprising the computer program product of any one of claims 1-9.