THE CLAIMS:

We claim:

1. A computer-implemented method for classifying a digital image, the method comprising:

obtaining, from a host computer, feature data corresponding to the digital image;

determining, by a graphics processing unit, a semi-metric distance based on a Poisson-Binomial distribution between the feature data and one or more reference feature data stored in a memory of the host computer; and

classifying the digital image using the determined semi-metric distance.

2. The method of claim 1, wherein the semi-metric distance is a Poisson-Binomial Radius (PBR).

3. The method of claim 1, wherein classifying the digital image comprises using a support vector machines (SVM) classifier.

4. The method of claim 1, wherein classifying the digital image comprises using a k-nearest neighbor (kNN) classifier.

5. The method of claim 4, wherein the kNN classifier is an adaptive local mean-based k-nearest neighbors (ALMkNN) classifier, wherein the value of the k-nearest neighbors (k) is adaptively chosen.

6. The method of claim 5, wherein the adaptive value of the k-nearest neighbors does not exceed the square root of the number of the one or more reference feature data.

7. The method of claim 1, wherein the obtained feature data and the one or more reference feature data comprise Pairwise Rotation Invariant Co-occurrence Local Binary Pattern (PRICoLBP) data.

8. The method of claim 1, wherein the obtained feature data and the one or more reference feature data comprise Histogram of Oriented Gradients (HOG) data.

9. The method of claim 1, wherein the obtained feature data comprises an N-dimensional feature vector X such that X = (a_1 . . . a_N) and the reference feature data comprises an N-dimensional feature vector Y such that Y = (b_1 . . . b_N), and wherein the determining the semi-metric distance (PBR(X,Y)) comprises calculating:

$$PBR(X,Y) = \frac{\sigma}{N - \mu} = \frac{\sqrt{\sum_{i=1}^{N} p_i(1 - p_i)}}{N - \sum_{i=1}^{N} p_i}$$

,

wherein N is an integer greater than 0,

σ is the standard deviation of vector X,

μ is the mean of vector X, and

p_i is $|a_i - b_i|$.

10. The method of claim 1, wherein the digital image comprises information corresponding to a DNA or RNA sequence, and the obtained feature data comprises a vector X of sequencing quality proximities for a first DNA or RNA sample with a sequencing depth d_x such that $X = (x_1 \ldots x_{dx})$ and the reference feature data comprises a vector Y of sequencing probabilities for a reference DNA or RNA sample with a sequencing depth d_y such that $Y = (y_1 \ldots y_{dy})$, and wherein the determining the semi-metric distance (PBR_{seq}) comprises calculating:

$$\mathrm{PBR}_{seq}(X,Y) = \frac{\sigma_X \sigma_Y}{|\mu_X - \mu_Y|},$$

wherein μ_X is a means for vector X,

μ_Y is a means for vector Y,

σ_X is a standard deviation for vector X, and

σ_Y is a standard deviation for vector Y.

11. The method of claim 10, wherein the classifying the digital image comprises:

determining whether the semi-metric distance (PBR_{seq}) is greater than a threshold value, and

classifying the DNA or RNA sequence as being a tumor or normal based on the determining if the semi-metric distance (PBR_{seq}) is greater than the threshold value.

12. The method of claim 10, wherein classifying the digital image comprises identifying a rare variant in the DNA or RNA sequence.

13. The method of claim 1, further comprising:

determining a closest matching reference feature data of the one or more reference feature data.

14. The method of claim 13, further comprising:

identifying a person based on the determined closest matching reference feature data, wherein the digital image comprises at least one of: an ear, a face, a fingerprint, and an iris.

15. A system for classifying a digital image comprising:

a host computer comprising a processor, wherein the host computer is coupled to a memory comprising one or more reference feature data; and

a graphics processing unit (GPU) comprising a processor,

wherein the GPU is coupled to the host computer and is configured to:

obtain, from the host computer, feature data corresponding to the digital image;

access, from the memory, the one or more reference feature data;

determine a semi-metric distance based on a Poisson-Binomial distribution between the feature data and the one or more reference feature data;

wherein the host computer is configured to:

classify the digital image using the determined semi-metric distance.

16. The system of claim 15, wherein the semi-metric distance is a Poisson-Binomial Radius (PBR).

17. The system of claim 15, wherein the host computer is further configured to classify the digital image using a support vector machines (SVM) classifier.

18. The system of claim 15, wherein the host computer is further configured to classify the digital image using a k-nearest neighbor (kNN) classifier.

19. The system of claim 18, wherein the kNN classifier is an adaptive local mean-based k-nearest neighbors (ALMkNN) classifier, wherein the value of the k-nearest neighbors (k) is adaptively chosen.

20. The system of claim 19, wherein the adaptive value of the k-nearest neighbors (k) does not exceed the square root of the number of the one or more reference feature data.

21. The system of claim 15, wherein the feature data and the one or more reference feature data comprise Pairwise Rotation Invariant Co-occurrence Local Binary Pattern (PRICoLBP) data.

22. The system of claim 15, wherein the obtained feature data and the one or more reference feature data comprise Histogram of Oriented Gradients (HOG) data.

23. The system of claim 15, wherein the feature data comprises an N-dimensional feature vector X such that $X = (a_1 \ldots a_N)$ and the reference feature data comprises an N-dimensional feature vector Y such that $Y = (b_1 \ldots b_N)$, and wherein the GPU is further configured calculate:

$$PBR(X,Y) = \frac{\sigma}{N-\mu} = \frac{\sqrt{\sum_{i=1}^{N} p_i(1-p_i)}}{N - \sum_{i=1}^{N} p_i}$$

,

wherein PBR(X,Y) is a Poisson-Binomial Radius (PBR) distance between the vector X and the vector Y,

N is an integer greater than 0,

σ is the standard deviation of vector X,

μ is the mean of vector X, and

p_i is $|a_i - b_i|$.

24. The system of claim 15, wherein the digital image comprises information corresponding to a DNA or RNA sequence, and the feature data comprises a vector X of sequencing quality proximities for a first DNA or RNA sample with a sequencing depth d_x such that $X = (x_1 \ldots x_{dx})$ and the reference feature data comprises a vector Y of sequencing probabilities for a reference DNA or RNA sample with a sequencing depth d_y such that $Y = (y_1 \ldots y_{dy})$, and wherein the GPU is further configured to determine the semi-metric distance (PBR_{seq}), wherein determining (PBR_{seq}) comprises calculating:

$$\mathrm{PBR}_{seq}(X,Y) = \frac{\sigma_X \sigma_Y}{|\mu_X - \mu_Y|},$$

wherein $PBR_{seq}(X, Y)$ is a Poisson-Binomial Radius (PBR) distance between the vector X and the vector Y,

μ_X is a means for the vector X,

μ_Y is a means for the vector Y,

σ_X is a standard deviation for the vector X, and

σ_Y is a standard deviation for the vector Y.

25. The system of claim 24, wherein the host computer is further configured to:

determine whether the semi-metric distance (PBR_{seq}) is greater than a threshold value, and

classify the DNA or RNA sequence as being a tumor or normal based on the determining if the semi-metric distance (PBR_{seq}) is greater than the threshold value.

26. The system of claim 24, wherein the host computer is further configured to:

identify a rare variant in the DNA or RNA sequence.

27. The system of claim 15, wherein the host computer is further configured to:

determine a closest matching reference feature data of the one or more reference feature data.

28. The system of claim 27, wherein the host computer is further configured to:

identify a person based on the determined closest matching reference feature data, wherein the digital image comprises at least one of: an ear, a face, a fingerprint, and an iris.